



Ordem dos Médicos Veterinários

# **Introdução à Biotecnologia Animal Moderna**

## **Elementos de Biologia Molecular Animal**

José Henrique Rocha Dias Correia  
António Agostinho Dias Correia

•  
2007

**ORDEM DOS MÉDICOS VETERINÁRIOS**

**ELEMENTOS DE BIOLOGIA  
MOLECULAR ANIMAL**

2007

José H. R. Dias Correia e A. A. Dias Correia

CIISA, Departamento de Morfologia e Função,  
Faculdade de Medicina Veterinária de Lisboa,  
Avenida da Universidade Técnica, Pólo Universitário da Ajuda,  
1300-477 Lisboa, Portugal  
e-mail [jhrdcorreia@fmv.utl.pt](mailto:jhrdcorreia@fmv.utl.pt)

Prefácio

Apresentação e justificação de todos os textos de “Elementos de Biologia Molecular Animal”

Índice geral

Índices temáticos

Conteúdos e objectivos mais relevantes de cada um dos temas tratados

Nota

Sumários, textos científicos, figuras , quadros e bibliografia de cada tema tratado

## PREFÁCIO

### APRESENTAÇÃO E JUSTIFICAÇÃO DE TODOS OS TEXTOS DE “ELEMENTOS DE BIOLOGIA MOLECULAR ANIMAL”

A sequenciação praticamente completa do genoma humano e a sua divulgação em 2001 constituiu um marco extraordinariamente importante com reflexos em todas as ciências biológicas.

A sequenciação completa do DNA mitocondrial humano tinha sido feita algum tempo antes, o mesmo sucedendo com os genomas de algumas dezenas de outros seres vivos, sobretudo microrganismos.

Também pouco depois da elaboração destes textos “Elementos de Biologia Molecular Animal” e da divulgação da sequenciação completa do genoma humano foi possível completar e divulgar a sequenciação do genoma do ratinho de laboratório (*Mus musculus*) e da ratazana (*Rattus norvegicus*), do peixe japonês (*Fugu rubripes*) que possui o mais curto genoma conhecido de qualquer espécie de vertebrado, e já em 2005 o genoma de frango (*Gallus gallus*).

Em 2003 foi anunciado nos U.S.A. que tinha sido constituído um consórcio com fundos dos U.S.A. (NHGR, NIH, USDA ARS; CSREES e estado do Texas), Canadá, Austrália (CSIRO) e Nova Zelândia, que se propõe fazer a sequenciação completa do genoma de bovinos (*Bos taurus*).

Toda esta informação está contida em bases de dados por esse mundo fora, dados esses de relativamente fácil acesso através de meios bioinformáticos convenientes.

O mesmo se passa com a muita informação correspondente a genes isolados e produtos da sua expressão, que tinham sido identificados e caracterizados, sobretudo a partir da data em que foram disponibilizadas metodologias para sequenciar moléculas biológicas de vários tipos, sua automação e meios informáticos e estatísticos para a sua análise.

Como se compreende em todo este contexto surgiu a necessidade de articular e integrar conhecimentos de diversas áreas tais como bioquímica e biologia molecular, genética molecular e das populações, bioinformática, estatística, etc.

Nos cursos de biologia aplicada como é o caso da Medicina Veterinária, os estudos do genoma, transcriptoma e proteoma, no futuro próximo, adquirirão o maior relevo há medida que as tecnologias para a sua identificação e caracterização, como é o caso dos “genes chips” e “proteínas microarray”, se forem aperfeiçoando e sobretudo ficarem mais acessíveis para todos aqueles envolvidos nestas matérias.

Uma evolução marcada destas matérias tem decorrido nos últimos anos, sobretudo na viragem do século, o que torna necessária a sua introdução, estudo e desenvolvimento nos cursos de graduação, tal como implica o indispensável refrescamento e apresentação aos já licenciados dentro da sua previsível vida profissional activa.

A ultraestrutura cromossomal e respectivos genes, as características específicas e funcionais do seu DNA, as tecnologias que podem permitir a sua manipulação aliadas com as estratégias desenvolvidas para a sequenciação dos genomas, são fundamentais para que se possam conhecer as características gerais destes genomas e a forma como nos animais esses genomas podem expressar-se, permitindo caracterizar determinados tipos de animais e/ou dos seus produtos.

São desenvolvidos nestes textos “Elementos de Biologia Molecular Animal” numa parte inicial, “Biologia Molecular e Genoma”, seis temas que abordam estas matérias.

A forma como se processam e operam sinais extracelulares e intracelulares em ordem a desencadear, inclusive ao nível da expressão dos vários genes, as respostas correspondentes, é desenvolvida sob a designação de “Biologia Molecular, Transdução de Sinais e sua Expressão” em três temas distintos. Nestes se procura aprofundar o conhecimento do funcionamento molecular entre diversos sectores dos meios intracelular, intercelular e extracelular, a transdução para o meio intracelular de vários sinais e o seu tráfego intracelular através de interacções entre diversas populações moleculares bem como a regulação da transcrição e expressão de genes correspondentes em certas circunstâncias.

Surge depois a necessidade, dentro deste contexto de vislumbrar perspectivas fisiológicas e patológicas desta interacções entre populações moleculares distintas, quer através dos ritmos circadianos, quer de sinais ópticos e olfactivos, quer ainda através da determinação, multiplicação, diferenciação, necrose ou apoptose celulares, com os inerentes stress, agressões e até formações malignas. Este conjunto de matérias constituem sete temas englobados na designação geral de “Biologia Molecular Fisiológica e Patológica”.

Também aquilo que vai sendo conhecido nos domínios da biologia molecular de tecidos muito importantes em produção animal (muscular e mamário) é abordado na “Biologia Molecular da Produção Animal” com dois temas.

# ELEMENTOS DE BIOLOGIA MOLECULAR ANIMAL

## **ÍNDICE GERAL**

I - BIOLOGIA MOLECULAR E GENOMA

II - BIOLOGIA MOLECULAR, TRANSDUÇÃO DE SINAIS E SUA EXPRESSÃO

III - BIOLOGIA MOLECULAR FISIOLÓGICA E PATOLÓGICA

IV - BIOLOGIA MOLECULAR NA PRODUÇÃO ANIMAL

## INDICES TEMÁTICOS

<b>I - <u>BIOLOGIA MOLECULAR E GENOMA</u></b>	26
1 - CROMOSOMAS E SUA ESTRUTURA	28
2 - -SEQUÊNCIAS NUCLEOTIDICAS, ESTRUTURAS DO DNA E SEU FUNCIONAMENTO	67
3 - TECNOLOGIA DO DNA RECOMBINANTE	99
4 - CONSIDERAÇÕES INTRODUTÓRIAS AO ESTUDO DOS GENOMAS	129
5 - ALGUMAS CARACTERÍSTICAS GERAIS DOS GENOMAS	145
6 - PROTEOMAS E GENOMAS. CLASSES DE FUNÇÕES MOLECULARES DE GENES E SEUS POLIMORFISMOS. PROTEÍNAS E FAMÍLIAS E SEUS DOMÍNÍOS. CARACTERIZAÇÃO DE ANIMAIS, TECIDOS E PRODUÇÃO ANIMAL	172
<b>II - <u>BIOLOGIA MOLECULAR, TRANSDUÇÃO DE SINAIS E SUA EXPRESSÃO</u></b>	240
1 - COMUNICAÇÃO INTRA E INTERCELULAR, MATRIZ EXTRACELULAR E INTEGRINAS, MEMBRANAS, VESÍCULAS E CITOESQUELETO	242
2 - SINAIS, RECEPTORES, CASCATAS E TRANSDUÇÃO DE SINAIS.	269
3 - TRANSCRIÇÃO, SEUS FACTORES E REGULAÇÃO	294
<b>III - <u>BIOLOGIA MOLECULAR FISIOLÓGICA E PATOLÓGICA</u></b>	442
1- RITMOS CIRCADIANOS NOS MAMÍFEROS	444
2 - TRANSDUÇÃO DE SINAIS VISUAIS	454
3 - FAMÍLIAS DE RECEPTORES DA OLFAÇÃO E VIAS DE SINALIZAÇÃO	461
4 - STRESS E RESPOSTA INFLAMATÓRIA. QUIMOCINAS. CITOCINAS	468
5 - NECROSE E APOPTOSE	495
6 - CICLO CELULAR	513
7 – ONCOGÉNESE	523

<b>IV - <u>BIOLOGIA MOLECULAR NA PRODUÇÃO ANIMAL</u></b>	539
1 - . GLÂNDULA MAMÁRIA E SEGREGAÇÃO LÁCTEA	541
2 - MIOGÉNESE, MÚSCULOS DO ESQUELETO E PRODUÇÃO ANIMAL	618



## CONTEÚDOS E OBJECTIVOS MAIS RELEVANTES DE CADA UM DOS TEMAS TRATADOS

### I - BIOLOGIA MOLECULAR E GENOMA

#### **1 - CROMOSSOMAS E SUA ESTRUTURA**

##### Alguns assuntos envolvidos e seus objectivos e conteúdos mais relevantes

###### ...Objectivos

O objectivo de fundo é a caracterização dos cromossomas, relacionando a ultraestrutura cromatínica com a sua remodelação no contexto da regulação da expressão dos genes, para depois se considerar a localização de genes nos cromossomas de diversas espécies animais, assim como a replicação dos cromossomas e o papel especial das telomerasas sobre as extremidades cromossomais.

###### ...Conteúdos

Faz-se uma revisão integrada do DNA, histonas e outras proteínas constituintes dos nucleossomas e cromossomas numa perspectiva de presença ou ausência de transcrição ao nível dos genes e correlativo grau de condensação da cromatina.

Também a localização de alguns genes nos cromossomas é revista tal como o esquema geral dos cromossomas e suas bandas.

Apesar de existirem *on-line* acessos fáceis da localização e características de alguns genes relacionados com doenças nos seres humanos e animais, faz-se um apanhado da localização de alguns destes genes responsáveis por doenças monogénicas e desordens poligénicas.

Dado o interesse em endocrinologia e reprodução dos suínos de um catalogo com genes porcinos (64 loci), também disponíveis *on-line*, refere-se a este propósito, a localização de alguns destes genes, o seu símbolo, designação e cromossoma onde residem.

Analogamente é feita uma compilação de alguns dos 1.563 genes de bovinos, disponíveis *on-line*, referindo o seu símbolo e localização.

A replicação celular e cromossomal revelam diversas características muito importantes, tal como o papel dos telómeros e telomerasas envolvidas neste contexto, com características muito próprias na semi-vida dessas populações celulares.

Admite-se que o leitor está de posse dos mecanismos gerais da replicação celular e da replicação cromossomal com a respectiva replicação e/ou reparação do DNA.

## 2 - SEQUÊNCIAS NUCLEOTIDICAS, ESTRUTURAS DO DNA E SEU FUNCIONAMENTO

### Alguns assuntos envolvidos e seus objetivos e conteúdos mais relevantes

#### ...Objectivos

Visa-se, a caracterização dos ácidos nucleicos. O corte em sequências precisas dos ácidos nucleicos assume o maior interesse nos estudos a partir dos genomas, bem como os diversos tipos e características das sequências que se podem encontrar nos genomas, pois constituem a base para a caracterização dos tipos de estruturas ao longo dos ácidos nucleicos e das suas respectivas conformações, bem como da estrutura dos diversos genes.

#### ...Conteúdos

Aborda-se como é possível “cortar” em sítios específicos qualquer molécula de DNA com o recurso a enzimas de restrição (endonucleases) bem como as características dos palindromas cortados ao nível do DNA.

A partir daqui verifica-se como as sequências nucleotídicas ao longo da molécula de DNA tendem a ser agrupadas (DNA de cópia única e sequências repetidas de vários tipos).

Estes diversos tipos de estruturas primárias do DNA originam depois outros tipos de estruturas que podem assumir diversas conformações (B, A e Z) existindo troços de DNA que se curvam ou formam dobragens angulares o que se repercute na replicação, transcrição, recombinação sítio-específica, etc. As estruturas peculiares de DNA cruciforme e as do dos DNA têm funções biológicas ainda mal conhecidas.

O DNA de hélices triplas, quádruplas, DNA deslizante, têm muito a ver com a relevância da sua dinâmica o que leva à introdução destas matérias, tal como as características dos diferentes tipos de recombinação genética.

Entendeu-se como pertinente desenvolver a estrutura dos genes e a sua transcrição em geral nas células eucariotas procurando refrescar e actualizar conhecimentos já possuídos em ordem a sustentar conteúdos a abordar em temas posteriores.

### 3 - TECNOLOGIA DO DNA RECOMBINANTE

#### Alguns assuntos envolvidos e seus objetivos e conteúdos mais relevantes.

##### ...Objectivos

São essencialmente o de introduzir e refrescar conhecimentos sobre a “manipulação” do DNA através das tecnologias do DNA recombinante.

?Como poderão ser obtidos fragmentos de DNA de interesse, a partir de um ser vivo e como introduzi-los noutra ser vivo com diversas finalidades?

?E como construir um reportório de todos os genes de um ser vivo ou como caracterizar e identificar cada um desses genes em ordem a melhorá-los, suprimi-los ou corrigi-los, com finalidades produtivas ou terapêuticas evidentes?

Procura-se a este propósito introduzir algumas noções que complementarão outras já adquiridas pelos Médicos Veterinários recentemente licenciados e introduzir bases porventura novas destas matérias para os licenciados já a médio prazo.

##### ...Conteúdos

As tecnologias do DNA recombinante, embora sejam metodologias apenas utilizadas por um número muito restrito de investigadores, contêm em si fundamentos básicos para a sua realização que se afiguram de grande interesse em todos os sectores da biologia aplicada, mormente no sector Médico Veterinário.

Em primeiro lugar aborda-se a fragmentação do DNA com enzimas de restrição que produzem para cada DNA, um mapa de restrição específico, bem como a forma como podem depois ser ligados entre si, se isso interessar, e como podem esses fragmentos ser inseridos em diversos tipos de vectores que os incorporem depois em células hospedeiro adequadas, produzindo a transformação destas com a produção de clones, clones estes que são facilmente identificáveis.

É possível esta clonagem de todo o DNA genómico de organismos multicelulares complexos, constituindo um repertório ou “biblioteca” genómica.

Estes repertórios podem inclusivamente ser feitos a partir de mRNA ou de cDNA, podendo o DNA clonado ser separado do DNA do seu vector, visualizado e sequenciado.

Há hoje uma estratégia para estudo dos genomas, inclusive de animais superiores, através de tecnologia de DNA recombinante, sendo possível purificar qualquer gene bem como determinar a sua sequência e as suas regiões funcionais, assim como preparar grande número de moléculas de DNA idênticas.

A formação e utilização de DNA sintético obtido por síntese química a partir de oligonucleótidos, com uma única hélice e com a sequência nucleotídica desejada, pode ser utilizada com finalidades muito diversas (sondas para identificação, porções para substituição de sequências naturais, para estudo de mutações específicas, etc).

Dada a importância, na construção de repertórios genômicos, da utilização, como vector de clonagem, do bacteriófago lambda ( $\lambda$ ) desenvolve-se esta matéria, assim como algumas características de outros vectores de clonagem.

A construção de repertórios genômicos com a utilização de sondas e a identificação de clones específicos assumem em conjunto com a bioinformática papel de particular relevo para o que se indicam uma série de fontes de bases de dados de sequências genômicas.

Em todo este contexto os locais de início da transcrição assumem particular relevância o que nos leva a abordar a forma como esse locais podem ser mapeados nos ácidos nucleicos.

A produção de animais transgênicos com a substituição dos genes endógenos por genes mutados é hoje assunto candente e de investigação de ponta.

Genes estranhos podem ser introduzidos em organismos pluricelulares (ovinos, caprinos, aves, bovinos suínos, etc) e passados estavelmente de uma geração a outra numa linha transgênica.

A terapia por genes a ensaiar em humanos e animais assim como alguns dos produtos biológicos obtidos pela tecnologia do DNA recombinante constituem a última parte deste tema.

## 4 - CONSIDERAÇÕES INTRODUTÓRIAS AO ESTUDO DOS GENOMAS

### Alguns assuntos envolvidos e seus objectivos e conteúdos mais relevantes

#### ...Objectivos

Pretende-se transmitir uma panorâmica das estratégias utilizadas para sequenciação dos genomas.

Aquilo que se pode extrair a partir do conhecimento do genoma completo de um dado ser vivo é hoje a pedra de toque de toda a biologia molecular moderna e quando se torna possível comparar, neste contexto, diversas espécies animais com características bem diversas, o realce ainda é maior.

Foi há poucos anos possível ter informação a este propósito, informação esta armazenada em diversas bases de dados acessíveis.

Como foi isto possível, e como foram sendo dados sucessivos passos, consoante iam surgindo novas metodologias utilizáveis, parece-nos de iniludível para todos os Médicos Veterinários, pois no futuro próximo estes assuntos serão intensamente explorados para as diversas espécies animais (já estão a sê-lo para os bovinos, frangos e peixes) com impactos económicos e em saúde pública por demais evidentes. É pois objectivo desta matéria dar uma panorâmica sobre as estratégias a desenvolver para caracterização dos genomas, sobretudo dos animais superiores.

#### ...Conteúdos

Pode dizer-se que os estudos dos genomas tiveram diversas fases ao longo dos tempos consoante as metodologias que iam entretanto surgindo.

Em 1977, com a introdução por F. Sanger de métodos para a sequenciação do DNA, começou a ser possível pensar em conhecer a sequencia nucleotídica completa do DNA de diversas espécies animais.

Contudo, o conhecimento para alguns seres vivos dos mapas genéticos e físicos, agora melhor definidos e esclarecidos pelas novas metodologias disponíveis, permitiu ir mais além.

Assim entre 1977 e 1982 foi possível alcançar a sequenciação de alguns vírus bacterianos e de alguns vírus animais (SV 40) e do DNA mitocondrial humano completada em 1994, concluindo-se que era possível a “assemblagem” de pequenos fragmentos sequenciados em genomas completos.

Também em 1987 surgiu o primeiro sequenciador automático de DNA que veio depois a ser sucessivamente melhorado.

Em 1989 foi criado o *Human Genome Organization (HUGO)* que se propunha promover uma colaboração internacional entre os cientistas envolvidos no *Human Genome Project (HGP)*.

Em 1990 foi oficialmente iniciado nos USA o *HGP* sob a égide do Instituto Nacional de Saúde tendo como objectivo a realização da sequenciação do genoma humano.

Tudo isto constituía uma tarefa ciclópica pois embora se conhecesse a sequência de muitos genes e não genes individuais, tornava-se necessário ordená-los em ordem a conhecer na totalidade do DNA como eles se ligavam entre si. Apenas em 2001 foi concluída esta tarefa e para um número muito restrito de seres vivos.

Para se chegar aqui desenvolveram-se trabalhos que levaram à construção de mapas genéticos e físicos dos genomas humanos e de ratinho, obtendo-se vários dados chave e pontos de “ancoragem” na sequência genómica.

Também as sequências de cDNA (obtidas pela transcrição reversa a partir de RNA) foram essenciais para anotar e validar previsões de genes baseadas na expressão de sequências TAG (EST ou sequências expressas tag eram obtidas por leitura de uma simples sequência nucleotídica não “assemblada” de um clone de DNA completamente aleatório).

Os números aumentados de sequências EST necessitavam de novos algoritmos nos computadores, capazes de analisar grandes quantidades de dados destas estruturas, e em 1993 foi criado um algoritmo que permitiu “assemblar” e analisar centenas de milhares de EST o que permitiu produzir “assemblagens” de cromossomas com elevada ordem de precisão e orientação necessitando de menos de dez sobreposições.

Foi também possível elaborar um dispositivo que permitia fazer a sequenciação de todo o genoma.

A sequenciação do genoma humano envolveu pois a sequenciação directa de fragmentos de DNA e a “assemblagem” das sequências destes fragmentos em unidades maiores na base das suas sobreposições.

Este trabalho fabuloso da sequenciação do genoma humano foi desenvolvido através de duas estratégias, uma no projecto público (HGP) e outra num projecto privado (CELERA) que acabaram por produzir resultados similares.

No projecto público estiveram envolvidos mais de 240 investigadores de diversíssimos países, sedeados em vinte centros de sequenciação do genoma (os seus resultados finais bem como uma lista complementar de outros investigadores envolvidos neste projecto podem ser consultados na revista *Nature* de 15 de Fevereiro de 2001).

No projecto privado encabeçado pela CELERA estiveram envolvidos 273 investigadores sedeados em 14 locais de actividade e os seus resultados podem ser consultados na revista *Science* de 15 de Fevereiro de 2001).

Ambos os projectos fizeram a apresentação simultânea da sequenciação de praticamente todo o genoma humano, nos USA (Casa Branca) em Junho de 2000.

Como se compreende é fácil perspectivar que a sequenciação e comparação de genoma s mamíferos completos será no futuro não muito distante uma operação de rotina nos laboratórios de biologia.

Estas razões levaram-nos a pensar ser fundamental dar á profissão veterinária uma perspectiva de como se passaram as coisas, dentro deste contexto. E o facto de já hoje serem utilizados em clínica veterinária e mesmo no âmbito da produção animal e tecnologias dos produtos animais, ferramentas baseadas no conhecimento parcelar de algumas sequências dos genomas animais, reforçam o interesse do refrescamento e actualização de conhecimentos desta índole.

## 5 - ALGUMAS CARACTERÍSTICAS GERAIS DOS GENOMAS

### Alguns assuntos envolvidos e seus objectivos e conteúdos mais relevantes.

#### ...Objectivos

Assinalar e sopesar as características gerais dos genomas após a sequenciação completa destes.

Após o conhecimento da sequenciação completa de alguns genomas foi possível através dos milhões de dados contidos em diversíssimas bases de dados, analisar as características dessas sequências com fins muito variados, não só de identificação e caracterização, mas também para aquilatar do seu funcionamento fisiológico e patológico.

É objectivo deste estudo encontrar uma linha de rumo que permita evolução de fundo neste contexto.

#### ...Conteúdos

A publicação atempada das sequências do genoma humano levaram a uma disponibilização de informação extraordinária que permitiu a análise global da organização do DNA dos vertebrados.

No contexto de um refrescamento e actualização para a profissão veterinária, dentro deste âmbito, começa-se por abordar a estrutura organizativa dos genes bem como as suas características operativas gerais, tal como a sua organização geral no contexto cromossomal e as suas localizações conjuntas ou não consoante os seres vivos.

Os algoritmos computacionais permitiram encontrar a localização de genes na imensidão de todo o genoma.

Nos mapas citogenéticos os diversos tipos de bandas heterocromáticas ou eucromáticas, a sua menor ou maior riqueza em G+C, a presença de ilhas CpG bem como outras características foram sendo assinaladas, contribuindo para uma definição muito mais fina de características gerais do genoma.

Os polimorfismos de nucleótidos simples (SNP) dada a sua importância na identificação e caracterização de situações fisiológicas e patológicas em cada ser vivo (mapa de SNP) e a sua distribuição contribuem de uma forma fundamental e decisiva para o grande impacto do conhecimento da sequenciação genómica.

Por outro lado, tornou-se possível verificar a organização dos genes nos cromossomas e a sua possível classificação genérica em famílias consoante o seu carácter único ou repetitivo, ou a semelhança ou dissemelhança com outros genes, realçando-se também a importância de alguns genes transcritos, mas não expressando estruturas proteicas.

## **6 - PROTEOMAS E GENOMAS. CLASSE DE FUNÇÕES MOLECULARES DE GENES E SEUS POLIMORFISMOS. PROTEÍNAS E FAMÍLIAS E SEUS DOMÍNIOS. CARACTERIZAÇÃO DE ANIMAIS, TECIDOS E PRODUÇÕES ANIMAIS**

### Alguns assuntos envolvidos e seus objetivos e conteúdos mais relevantes

#### ...Objectivos

Compulsar aspectos dominantes apurados nos proteomas após a sequenciação dos genomas para caracterização e identificação de animais e seus produtos.

Dispondo-se hoje de algoritmos e outros meios bioinformáticos que nos permitem identificar e caracterizar com um certo grau de fiabilidade milhares ou dezenas de milhares de sequências nucleotídicas num dado genoma ou de as contrastar com genomas conhecidos, compreende-se o interesse que isto pode ter para um Médico Veterinário, pois permite-lhe vislumbrar quais são os genes, as famílias de genes mais abundantes, bem como outras características inerentes desejáveis (melhores “performances” produtivas, maior resistência a certas doenças) ou indesejáveis (anomalias de diversa natureza) quando se comparam uns com os outros.

Por outro lado antes do conhecimento de alguns genomas completos, já as actividades ligadas com a Medicina Veterinária procuravam identificar estruturas nucleotídicas ou mesmo genes, que em alguns animais tivessem importância, sobretudo económica, evidente e que em muitos casos deram origem até a patentes.

#### ...Conteúdos

Quando se dispõe de um “apanhado” geral dos genes contidos num dado genoma é tentador arregimentá-los em função das suas características dominantes (funções moleculares) e saber da sua importância relativa dentro de cada ser vivo, e depois comparar espécies com espécies, relacionamento provável, evolução possível, etc.

E a sequenciação praticamente total, sobretudo do genoma humano, permite ao sector Médico Veterinário encarar como fortemente provável que ocorra uma estreita analogia com os outros seres vivos que o interessam e para os quais não se dispõe ainda de dados para todo o genoma.

A repartição em categorias das funções moleculares dos genes humanos que exprimem proteínas (e desta próprias) permite admitir a importância, certamente partilhada por muitos outros animais, das actividades relacionadas com a biologia dos ácidos nucleicos e com a transmissão de sinais, que são as funções moleculares que aparecem em maior número, parecendo que a maquinaria de transcrição e tradução tem sido muito conservada através da evolução das espécies, desde as células bacterianas até aos eucariotas mais complexos.

No entanto, ao longo dos anos, vinham sendo feitos trabalhos de muito interesse e que visavam aplicações práticas no sector pecuário e veterinário, utilizando para esse efeito “ferramentas” mais rudimentares que aquelas utilizadas na sequenciação do genoma total.



Assim já era feito o isolamento de alguns genes e seu mapeamento no respectivo genoma, através da fragmentação dos cromossomas com a subsequente sobreposição das porções dos fragmentos obtidos.

Foi assim possível localizar diversos genes e marcadores em diversas espécies animais, bem com identificar a localização de algumas características produtivas.

A identificação e caracterização dos animais perspectivando a sua selecção, também foi sendo realizada recorrendo a métodos que detectavam variações nos genes de *locii* de características quantitativas (QTL) ou então através de métodos que detectavam variações genéticas em genes funcionais.

A importância dos polimorfismos nucleotídicos para identificação de espécies e animais, análises de descendência e pedigree, etc, era utilizada, havendo inclusive uma série de testes comercializados e patentes para detecção de doenças nos animais, baseados na detecção de mutações ocorridas.

Neste capítulo da resistência ou sensibilidade às doenças dos animais, patentes foram registadas baseadas na identificação de características dos genes, havendo inclusive, nas desordens e doenças identificadas nos animais com hereditariedade Mendeliana, uma base de dados *on-line Mendelian Inheritance in Animals-MIA*.

As características produtivas dos animais, quer no que se refere a características reprodutivas melhoradas, quer no que se refere a produção de carne e leite têm suscitado uma profusão de testes baseados nas características de determinados tipos de genes.

Também a identificação dos sexos em algumas espécies animais recorre à análise das características do material genético.

Perspectiva-se num futuro que já começou que as análise aos genomas, transcriptomas e proteomas por técnicas de “microarray” que permitirão a apresentação global e individual de cada uma destas importantíssimas fracções biológicas, constitua um novo meio de intervenção crucial em toda a vida biológica.

## **II - BIOLOGIA MOLECULAR, TRANSDUÇÃO DE SINAIS E SUA EXPRESSÃO**

### **1 -COMUNICAÇÃO INTRA E INTERCELULAR,MATRIZ EXTRACELULAR E INTEGRINAS, MEMBRANAS, VESÍCULAS E CITOESQUELETO**

Alguns assuntos envolvidos e seus objetivos e conteúdos mais relevantes.

...Objectivos

Sopesar como comunicam genericamente as células entre si e com o seu meio envolvente.

Os complexos seres que são os animais superiores constituídos por milhões e milhões de células integradas em maior ou menor extensão em matrizes extracelulares, necessitam de transmitir milhões de mensagens, sinais ou metabolitos em ordem a estabelecer a homeostasia característica de cada animal e em cada etapa do seu desenvolvimento.

É essencial conhecer a nível ultraestrutural como se desenrola esta troca de informação, nas interfaces entre o meio extracelular e/ou intracelular e vice-versa, até para compreender como pode este fluxo informativo condicionar a expressão do material genético porventura responsável pelas respostas obtidas.

...Conteúdos

Em todos os seres pluricelulares as comunicações entre as células ou entre as células e as matrizes extracelulares assumem um papel fundamental pois permitem que essas comunicações se processem nos dois sentidos, de fora para dentro das células, e do interior da célula para o seu meio envolvente, através do concurso de uma série de populações moleculares que residem ou actuam nas respectivas interfaces.

Por outro lado, a compartimentação intracelular nas células eucariotas, necessita por seu turno de meios para operar essa comunicação, inclusive entre os vários organelos contidos no interior das células.

O papel de diversas proteínas é essencial para a formação de vesículas, segregação, fusão e endocitose, em diversos compartimentos subcelulares, havendo proteínas específicas que alvejam organelos subcelulares específicos.

O transporte de vesículas através do aparelho de Golgi também assume particular relevância.

A dinâmica dos microfilamentos e microtúbulos, os motores proteicos. Moleculares (inclusive miosinas especiais e proteínas relacionadas com as cinesinas) está bem envolvida na motilidade das células e na sua divisão celular.

O conhecimento detalhado dos diversos componentes que estabelecem as comunicações intracelulares e intercelulares é essencial assim como os mecanismos sistémicos envolvidos.

## **2 - SINAIS, RECEPTORES, CASCATAS E TRANSDUÇÃO DE SINAIS.**

### Alguns assuntos envolvidos e seus objectivos e conteúdos mais relevantes.

#### ...Objectivos

Abordar aspectos específicos da comunicação extracelular e intracelular e a expressão dessa comunicação.

No contexto pluricelular dos animais superiores a forma como os diversos sinais extracelulares, intercelulares e intracelulares operam e desencadeiam os seus efeitos inclusive ao nível de variadíssimos genes, pressupõe vias para a transdução de sinais em que estão envolvidas populações moleculares muito distintas.

O conhecimento destes mecanismos implica não só a justificação de um determinado tipo de resposta mas também pode sugerir em que condições será possível intervir com consequências clínicas ou produtivas notórias.

#### ...Conteúdos

Nos animais as vias de transdução de sinais desencadeadas ao nível dos receptores das superfícies celulares ou no interior das células, controlam a proliferação celular, o crescimento e mobilidade celular, além de outros acontecimentos biológicos.

No conjunto de sinais anteriormente referidos, as hormonas e factores de crescimento assumem particular relevância, o que nos leva a um breve resumo das funções e sinais das principais hormonas polipeptídicas e esterólicas dos organismos animais.

A transdução desses sinais pode ser desencadeada através de receptores membranários à superfície das células ou através de receptores intracelulares.

Por outro lado existem na transdução de alguns sinais uma série de proteínas cinases de tipos um tanto diversificados.

Em todo este contexto são dados exemplos de algumas proteínas cinases envolvidas na transdução de sinais.

Na transdução de sinais envolvendo receptores membranários à superfície das células, são revistos os seus tipos principais. Assim são estudados os receptores de sete hélices, as adenilato-ciclases e as proteínas G e a sua regulação. Também a transdução de sinais através de receptores com actividade Tirosina-cinases intrínseca é abordada, as famílias que os constituem, as moléculas adaptadoras e cascatas envolvendo CREB ou desencadeadas pela insulina, EGF e SRF.

De igual modo é abordada a transdução de sinais através de Tirosina-cinases citosólicas, famílias que as integram e cascatas desencadeadas pela IFN e GH.

Na transdução de sinais utilizando receptores intracelulares é abordada a síntese e libertação de hormonas esterólicas a sua acção, receptores e seus elementos de resposta no DNA.

Termina-se com a abordagem da transcrição nuclear a ser modelada pela transdução dos respectivos sinais.

### 3 - TRANSCRIÇÃO, SEUS FACTORES E REGULAÇÃO

#### Alguns assuntos envolvidos e seus objectivos e conteúdos mais relevantes

##### ...Objectivos

Aquilar acerca de diversos tipos de expressões celulares e seus factores intervenientes.

As interacções funcionais entre diversas populações moleculares que constituem os seres vivos são decisivas em toda a vida biológica.

Conhecer os objectivos das interacções entre os ácidos nucleicos e as proteínas assume particular relevância, pois permite compreender como pode por exemplo o DNA ser condicionado e regulado pelas interacções com moléculas proteicas activadas por uma série de factores e sinais de distintas proveniências.

A replicação, reparação e transcrição do DNA carecem destas interacções.

No funcionamento de todas as células, a forma regulada ou não, como a partir do seu material genético se exprimem as diversas populações moleculares responsáveis pelo fenótipo funcional de cada tipo de célula, intervêm uma série de mecanismos operativos comuns aos vários tipos de células, mas um tanto diversificados consoante os produtos finais resultantes.

Nesta transcrição, a ultraestrutura molecular envolvida e os seus detalhes morfológicos são essenciais para se alcançar os efeitos fisiológicos finais.

##### ...Conteúdos

A transcrição nas células eucariotas dos animais superiores é regulada por uma série de factores.

Os diversos tipos de RNA polimerases envolvidos na transcrição são controlados pela associação combinada de diversas proteínas.

A estrutura destas proteínas, que são factores de transcrição que interactuam com o DNA, é extraordinariamente importante e podem essas estruturas, consoante os seus tipos, ser agrupadas em famílias estruturais a que correspondem certos princípios para interacção com o DNA.

As estruturas secundárias e os domínios que integram cada uma destas famílias são decisivas nesta regulação da transcrição.

É feita uma revisão actualização dos factores gerais de transcrição e da actuação das RNA polimerases dos tipos II, I e III, sendo depois posta a tónica na transcrição dos genes codificadores das proteínas com a respectiva remodelação da cromatina.

Dada a importância do factor de transcrição TF II D é visto em mais detalhe o seu funcionamento, assim como se aprofunda o conhecimento dos complexos coactivadores da transcrição.

### **III - INTRODUÇÃO À BIOLOGIA MOLECULAR FISIOLÓGICA E PATOLÓGICA**

#### Alguns assuntos envolvidos e seus objectivos e conteúdos mais relevantes

##### ...Objectivos

Conhecer factores intervenientes nos relógios circadianos dos animais, nos fenómenos visuais e olfactivos, e ainda os factores moleculares condicionadores da determinação celular, diferenciação, necrose, apoptose, ciclo celular e oncogénese.

As biologies moleculares envolvidas em todas estas situações são essenciais para se compreenderem as patogenias envolvidas e possíveis diagnóstico, prognóstico e tratamentos a operar.

##### ...Conteúdos

Começa-se por abordar aspectos da biologia molecular dos ritmos circadianos nos mamíferos, focando a trama molecular desse relógio, assim como as vias de input-fótico, o output dos osciladores circadianos e aspectos dos genomas correspondentes.

A transdução de sinais visuais é revista não só no que se refere às proteínas envolvidas no ciclo visual e cascatas operadas, mas também a biologia conhecida da visão a cores nas células em cone.

As famílias de receptores da olfacção e vias de sinalização são abordadas através do estudo dos receptores de sete hélices e das guanilil-ciclases.

A complexidade da resposta inflamatória e o envolvimento do stress, a biologia das quimocinas e das citocinas são abordadas, fazendo uma revisão e actualização dos seus aspectos estruturais e funcionais mais importantes.

A consequente necrose e apoptose são estudadas através das sinalizações desencadeadas por diversas vias e proteínas intervenientes.

O ciclo celular e a sua dependência das ciclinas e CdK é revisto tal como a sua regulação e desregulação (oncogénese).

É aprofundada depois a oncogénese atendendo a factores de transcrição envolvidos e a outras proteínas, as suas características, e numa pequena parte e dado o seu particular interesse a transdução de sinais mediada por oncogenes em células epiteliais mamárias de ratos transgénicos.

## **IV - BIOLOGIA MOLECULAR NA PRODUÇÃO ANIMAL**

### **1 . GLÂNDULA MAMÁRIA E SEGREGAÇÃO LÁCTEA**

#### Alguns assuntos envolvidos e seus objetivos e conteúdos mais relevantes

##### ...Objetivos

A biologia da glândula mamária e das suas segregações, numa perspectiva produtiva ou mesmo fisiopatológica, fundamenta-se na intervenção de diversíssimos factores moleculares que inclusivamente condicionam as características dos componentes do leite, o que implica a necessidade do seu conhecimento operacional.

O desenvolvimento cíclico dos tecidos mamários em vários períodos de lactação ao longo da vida dos animais pressupõem mecanismos alternativos e interacções muito complexas entre o seu meio envolvente e as diversas populações celulares constituintes dos tecidos, bem como com os produtos da sua segregação subsequente.

Compreende-se que numa perspectiva Médico-Veterinária os novos dados que vêm sendo obtidos dentro de todo este contexto permitirão uma intervenção clínica mais esclarecida e eficiente assim como uma orientação para uma produção de leite mais vantajosa.

##### ...Conteúdos

As glândulas mamárias dos animais domésticos são órgãos especiais com uma evolução muito própria ao longo dos diversos períodos de lactação decorrentes durante a vida dos animais.

As primeiras etapas da formação dos tecidos mamários, bem como os factores para o seu crescimento e diferenciação, são considerados após uma introdução genérica à biologia da glândula mamária.

As características próprias da puberdade bem como da involução mamária estão condicionadas por uma série de factores intervenientes que interactuam entre si.

No desenvolvimento da glândula mamária o papel especial da somatostatina, opióides, progesterona e progestinas, bem como o desempenho fundamental da prolactina e dos vários factores de transcrição e cascatas de transdução de sinais, são elementos em que sinergias e antagonismos interactivos entre essas populações moleculares condicionam e regulam a expressão dos genes responsáveis pelas proteínas que constituem a segregação láctea.

Nos últimos anos grande soma de conhecimentos têm sido alcançados, inclusive ao nível da organização dos nucleossomas e remodelação da cromatina, e das interacções matriz celular células da glândula mamária em diversas fase da sua actividade.

A estrutura e modo de funcionamento dos genes implicados na expressão das proteínas do leite tem também sentido avanços notáveis que têm permitido conhecer variações na constituição dessas proteínas lácteas (polimorfismos) om consequências fisiológicas e industriais evidentes.

Todos estes conhecimentos têm inclusivamente sugerido intervenções quer ao nível da fracção proteica quer lipídica do leite, com notável impacto nas características do produto final obtido e com iniludíveis reflexos na actividade Médico Veterinária neste sector



## 2 - MIOGÉNESE, MÚSCULOS DO ESQUELETO E PRODUÇÃO ANIMAL

### Alguns assuntos envolvidos e seus objectivos e conteúdos mais relevantes.

#### ...Objectivos

A base das produções animais cárneas passa pelo conhecimento da miogénese e dos diversos factores que a afectam nas diversas espécies animais.

Aquilo que leva a uma génese acrescida de tecidos miogénicos, ou à sua regeneração ou hipertrofia funcional, passa por diversas populações moleculares actuando fora ou dentro dos tecidos miogénicos, desencadeando neste tecidos diversos tipos de sinais que podem ou não ser transduzidos até ao núcleo celular originando aqui profundas implicações na regulação de diversos tipos de genes.

Os conhecimentos adquiridos nos últimos anos neste âmbito, dotam a Medicina Veterinária de um melhor conhecimento para intervir no âmbito dos tecidos miogénicos quer com finalidades funcionais, terapêuticas ou de obtenção de produções acrescidas.

#### ...Conteúdos

O conhecimento básico dos factores e mecanismos implicados na miogénese ao longo da vida dos animais assume em Medicina Veterinária enorme acuidade, quer numa perspectiva clínica, quer numa perspectiva de produção animal.

Para esse efeito é feita uma abordagem da miogénese desde fase iniciais da vida dos animais (embriogénese) com a intervenção de factores de transformação do crescimento, interacções células com células e células-matrizes extracelulares, como condicionadoras da determinação e diferenciação miogénica.

O papel dos genes envolvidos na regulação da miogénese é considerado em diversas situações de regulação positiva e negativa.

É focada a hipertrofia muscular em espécies animais com reguladores genéticos negativos actuantes, aprofundando-se o papel de um destes reguladores, a miostatina e de outros reguladores envolvidos na regulação global da miogénese.

Numa segunda parte são abordados diversos factores que afectam o crescimento e desenvolvimento dos músculos do esqueleto, tais como hormonas, exercício e nutrição, terminando-se com um curto capítulo sobre alguns aspectos da endocrinologia dos peixes em produção animal.

## **NOTA**

A elaboração destes textos, que corresponde em alguma parte á consulta, recolha selectiva e sinopse das obras fontes que lhe deram suporte, foi feita ao longo dos últimos anos.

Salienta-se que em alguns textos em alguns sub-capitulos e rubricas, logo a seguir ao titulo respectivo aparece indicação da fonte bibliográfica de suporte com o que pretendemos indicar que o texto desenvolvido de seguida foi fortemente influenciado pela fonte indicada.

Entendeu-se que como se tratava de textos para divulgação e actualização de generalistas (a maioria dos textos) que não interessava adoptar sempre o figurino habitual para publicação dos trabalhos de investigação científica e desenvolvimento experimental.

Assim á semelhança de muitos textos recentes e actuais por esse mundo fora, os assuntos, na maioria dos textos, á medida que são desenvolvidos não contém expressamente indicações de cada fonte bibliográfica, salvo algumas excepções, aparecendo esta na bibliografia final sem ser por ordem alfabética do nome dos autores, mas sim pela ordem em que intervêm no texto, o que está implícito na denominação de cada artigo referido e o assunto em desenvolvimento.

Foi o que sucedeu com os quatro primeiros temas englobados na designação de “Biologia Molecular e Genoma”, embora nos dois últimos temas desta mesma rubrica, ou sejam as “Algumas Características Gerais dos Genomas” e “Preoteomas e Genomas...” tenhamos adoptado um dos sistemas clássicos de referências numeradas no texto e contidas na bibliografia final por ordem de citação por entendermos que assim poderiam ser mais úteis para os leitores.

Nos temas incluídos na área da “Biologia Molecular, Transdução de Sinais e sua Expressão”, o primeiro tema (“Comunicação Intra e Intercelular, Matriz Extracelular e Integrinas. Membranas.Vesículas e Citoesqueleto”, adopta o figurino dos quatro primeiros temas da “Biologia Molecular e Genoma”, mas os dois temas seguintes, (“Sinais, Receptores, Cascatas e Transdução de Sinais” e “Transcrição.Seus Factores e Regulação”) dada a profusão de figuras e esquemas importantes incluídos nos textos levaram-nos após a bibliografia, a detalhar na iconografia, as respectivas fontes bibliográficas donde foram adaptadas.

No que se refere aos temas de “Biologia Molecular Fisiológica e Patológica”, e de “Biologia Molecular da Produção Animal”, assuntos mais especializados e com os textos mais compartimentados e independentes entendeu-se pertinente distribuir a bibliografia final por cada assunto tratado nos textos em ordem a facilitar outras consultas eventualmente desejadas.

**SUMARIOS, TEXTOS CIENTIFICOS, FIGURAS, QUADROS E BIBLIOGRAFIA DE CADA TEMA**

# **BIOLOGIA MOLECOLAR E GENOMA**

## **I – BIOLOGIA MOLECULAR E GENOMA**

<b>1 - <u>Cromossomas e sua estrutura</u></b>	28
1.1 - DNA, histonas e outras proteínas, nucleossomas e cromossomas.	28
1.2 - Cromossomas das células eucariotas e expressão dos seus genes. Genes, <i>loci</i> e alelos. Localização nos cromossomas de alguns genes.	33
<u>Genes, <i>Loci</i> e Alelos</u>	34
<u>Localização nos cromossomas de alguns genes</u>	36
1.3 - Localização nos cromossomas humanos de alguns genes relacionados com doenças.	37
1.4 -Catálogo de genes porcinos (64 <i>loci</i> ) de interesse em endocrinologia e reprodução.	38
1.5 - Lista de alguns genes de bovinos contidos no <i>Bovmap Database (world wide web version 2.00)</i> .	40
1.6 - Replicação celular e cromossomal.	54
<u>Replicação dos cromossomas</u>	55
1.7 - Telomeros e telomerasas.	59
1.8 - Bibliografia.	64

## **1 - Cromossomas e sua estrutura**

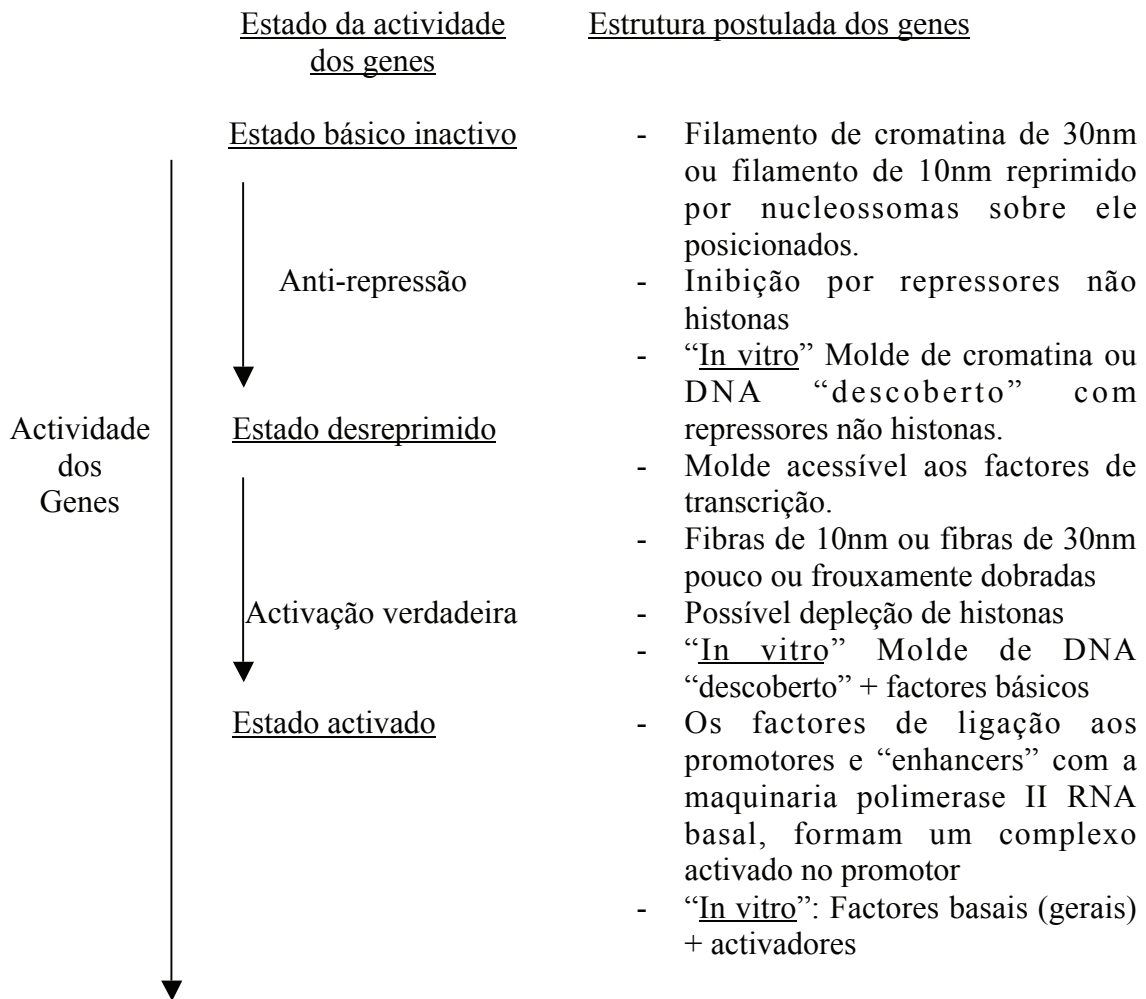
### **1.1 – DNA, histonas e outras proteínas, nucleossomas e cromossomas.**

(Paranjape,S.M. et al.,1994)

Os cromossomas são estruturas celulares para armazenagem e transmissão da informação genética. Os cromossomas das células eucariotas são complexos de DNA (10-30%), RNA (3-15%) e proteínas (40-75%). As proteínas cromossomais podem ser divididas em duas classes: as histonas, básicas, e as proteínas não-histonas, mais acídicas. Há seis classes de histonas, enquanto que as proteínas não-histonas dão centenas de bandas após separação umas das outras por electroforese em gel. As histonas combinam-se com o DNA formando complexos fibrilares. As proteínas não-histonas estão provavelmente implicadas na regulação da transcrição (englobam também as polimerases). Os cromossomas das células eucariotas formam estruturas complexas antes da divisão celular (mitose ou meiose) e, nesta altura, cada cromossoma consiste de duas metades longitudinalmente idênticas: os cromatídeos, unidos num ponto denominado centrómero. Também sequências especializadas de DNA, os telómeros, encontram-se nas extremidades dos cromossomas lineares. No núcleo o DNA está pois “empacotado” num complexo nucleoproteico conhecido por cromatina, a qual é grosseiramente constituída por duas partes de proteínas para uma de DNA, sendo ainda a proporção histonas / DNA de um para um. Os centrómeros (que são responsáveis pelo movimento dos cromossomas para as células filhas) são o local onde os cromossomas se constroem durante a metafase e em que se separam os braços curtos dos compridos do cromossoma. Tanto os centrómeros como os telómeros contêm conjuntos curtos de DNA repetidos em tandem (*vide* adiante). Como também veremos adiante, os telómeros são compostos por um grande número de curtas sequências repetidas e que são mantidas, independentemente da replicação, pela enzima telomerase. Por sua vez, a cromatina é um material corável do núcleo em interfase, consistindo de DNA, RNA e proteínas, dispersas um tanto ao acaso no núcleo das células eucariotas. A maior parte da cromatina num núcleo em interfase encontra-se presente como filamentos de 30 nm de diâmetro, mais do que na forma de fibras distendidas de 10nm de diâmetro (colar de pérolas) (*vide* figura adiante). O DNA cromossomal nesta fase está intacto, ou seja, desenrolado ou relaxado. Antes da divisão celular a cromatina condensa-se em corpos densos ( cromossomas ) que podem ser corados intensamente. A estrutura da cromatina regula a expressão dos genes tal como se esquematiza seguidamente.

Um gene inactivo existe, pois, como uma estrutura de cromatina reprimida (quanto à transcrição) e que pode encontrar-se “dobrada”, ou compactada, em ordem a torná-la inacessível aos factores de transcrição. No entanto, no estado activo, ou desreprimido, a estrutura da cromatina deve apresentar-se descondensada, ou “desdobrada”, e a estrutura dos nucleossomas pode, por sua vez, ser alterada (*vide* adiante) sendo que provavelmente as histonas H1, mais do que quaisquer outras histonas, podem ter sido removidas. Uma importante função dos factores que se ligam ao promotor e ao “enhancer” (*vide* adiante) consiste em remover a repressão mediada pela cromatina, o que por vezes se designa por anti-repressão. Por outro lado, os factores de transcrição específicos destas sequências possuem capacidades para facilitar a transcrição do DNA (activação verdadeira).

Paradigma do papel da estrutura da cromatina  
na regulação da expressão dos genes (Adaptado de Paranjape, S.M. et al., 1994)



Um aspecto característico do núcleo das células eucariotas é o da associação do DNA com pequenas proteínas básicas, as histonas, formando assim os nucleossomas; estes são grupos de histonas, aos pares, constituindo uma partícula à volta da qual o DNA enrola-se.

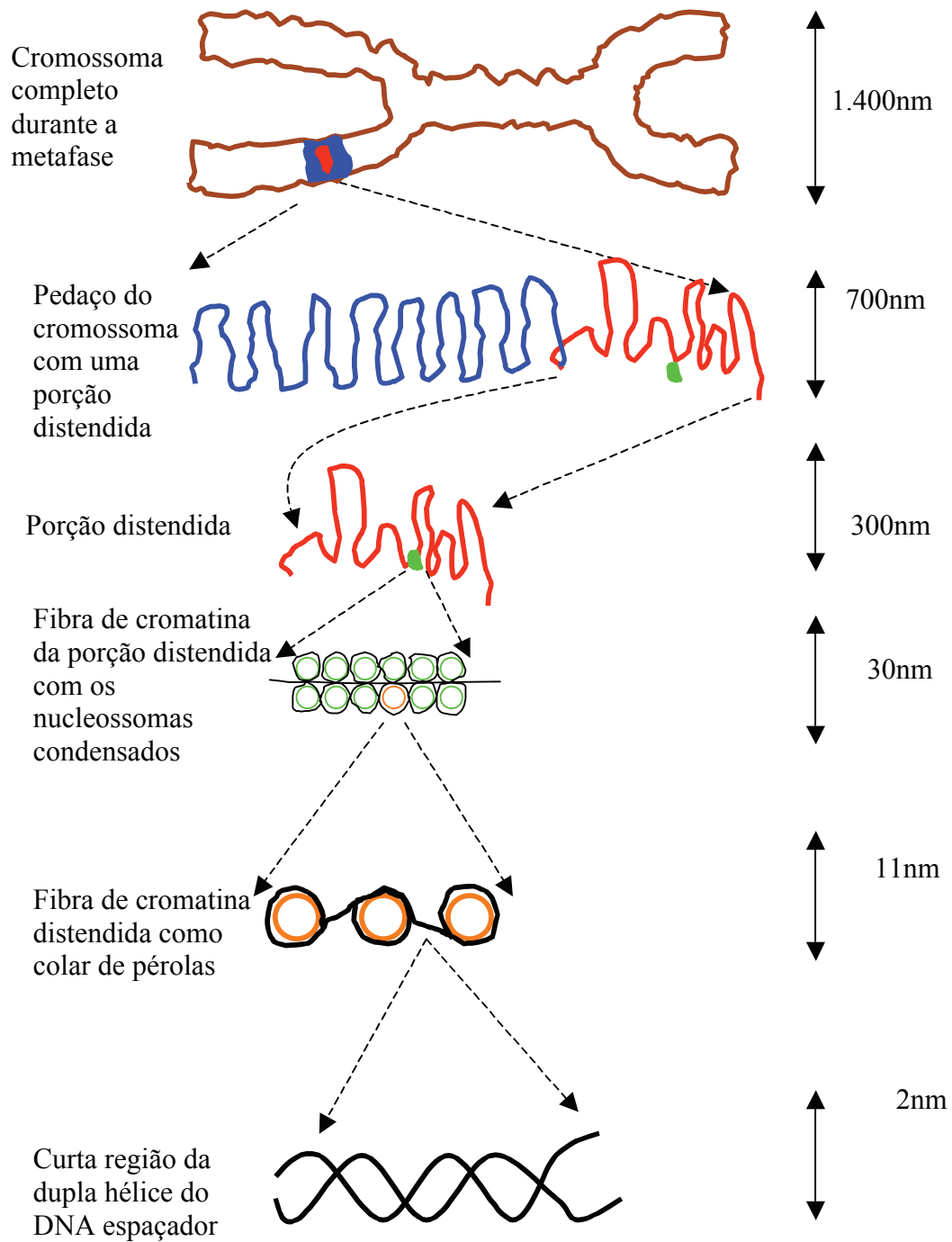
<u>Tipos de histonas</u>			
<u>Tipo</u>	<u>Relação Lys/Arg</u>	<u>Nº de resíduos de aminoácidos</u>	<u>Mr x 10<sup>-3</sup></u>
H1	20,0	215	21,0
H2A	1,25	129	14,5
H2B	2,5	125	13,8
H3	0,72	135	15,3
H4	0,79	102	11,3

} Componentes dos Nucleossomas

(Adaptado de Campbell, P.N. & Smith, A.D., 1993)

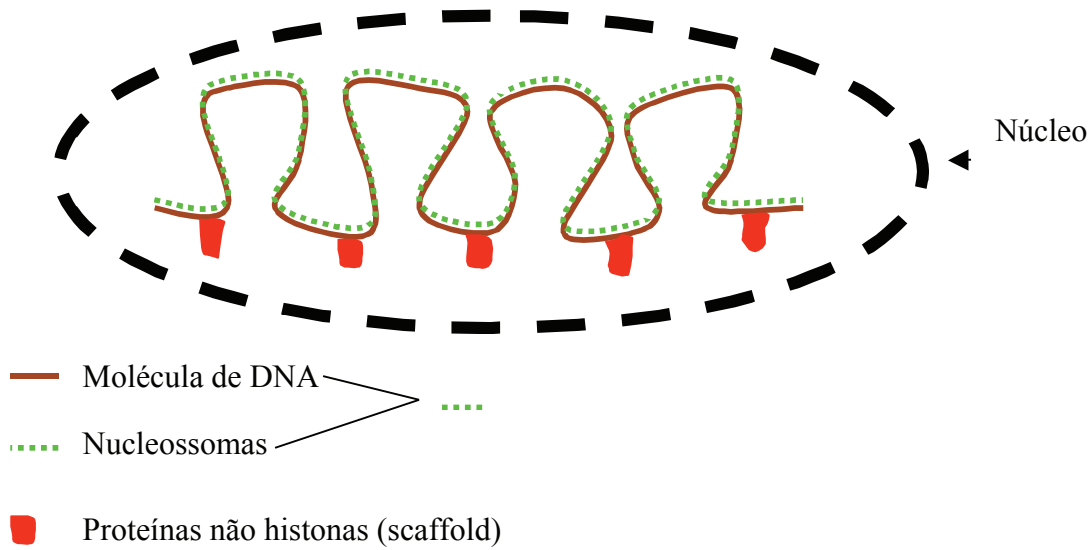
O DNA na cromatina está pois organizado em unidades repetidas regularmente, os nucleossomas, que contêm sucessivos segmentos da hélice dupla de DNA com  $\pm 200$  pares de bases de comprimento associados com um octamero de quatro proteínas histonas, H3, H4, H2A e H2B, estando o DNA enrolado cerca de duas vezes á volta deste octamero. No ponto de entrada e de saída o DNA está unido a outra histona: a H1.(Campbell e Smith, 1993). Esta unidade estrutural repete-se ao longo de todo o DNA formando um filamento que pode enrolar-se numa estrutura helicoidal, ou “solenóide”, para formar um segundo nível de enrolamento (*vide* figuras adiante). Estes complexos nucleoproteicos são também designados por cromatina, a qual quando condensada forma os cromossomas. Durante a interfase celular a maioria da cromatina não se encontra condensada, sendo neste estado feita uma distinção morfológica entre a heterocromatina firmemente “packed” (empacotada) e a eucromatina, muito menos “packed” (empacotada), sendo esta última transcrita activamente. As proteínas histonas, ricas em ácidos aminados básicos com cadeias laterais ionizadas positivamente, contactam com os grupos fosfatos ionizados negativamente do DNA. Alguns destes ácidos aminados básicos das histonas têm as cadeias laterais modificadas pela adição, após a tradução, de grupos metilo, acetilo ou fosfato, neutralizando a carga positiva da cadeia lateral, ou convertendo-a numa carga negativa. A sequência em ácidos aminados das quatro histonas (H2A, H2B, H3 e H4) são muito semelhantes entre as diversas espécies animais. A sequência em ácidos aminados da H1 varia mais de organismo para organismo, sendo até em certos tecidos substituída esta H1 por histonas especiais, como por exemplo a H5. A cromatina existe em formas condensadas e em formas não condensadas ou distendidas, como se fosse um colar de pérolas (proteínas histonas separadas umas das outras mas unidas por um filamento único (DNA)) (*vide* figura seguinte). Os nucleossomas têm cerca de 10nm de diâmetro, podendo assumir também formas mais condensadas de vários nucleossomas, constituindo neste caso fibras de 30nm de diâmetro. O DNA que envolve as histonas é muito menos sensível à digestão por enzimas do que o DNA que se encontra ligando os sucessivos nucleossomas entre si. Os nucleossomas contêm cerca de 146 pares de bases de DNA enquanto o DNA que liga dois nucleossomas sucessivos (DNA espaçador) parece apresentar cerca de 14 a 54 pares de bases com o DNA de 146 pares de bases enrolado duas vezes à volta da superfície do octamero. Há proteínas acídicas, sem serem histonas, associadas, no entanto, com as histonas e que não estão “assembladas” nos nucleossomas, embora estejam envolvidas na organização e estrutura dos cromossomas constituindo a armação (*scaffold*) dos cromossomas (*vide* esquemas seguintes). Estas proteínas não-histonas formam pois um *scaffold* (esqueleto ou armação) estrutural para a formação de longas ansas de DNA nos cromossomas constituindo a armação (*scaffold*) dos cromossomas (*vide* esquemas seguintes).

Esquema das diversas ordens de condensação da cromatina a partir de um cromossoma altamente condensado da metáfase(Adaptado de Campbell e Smith,1993)

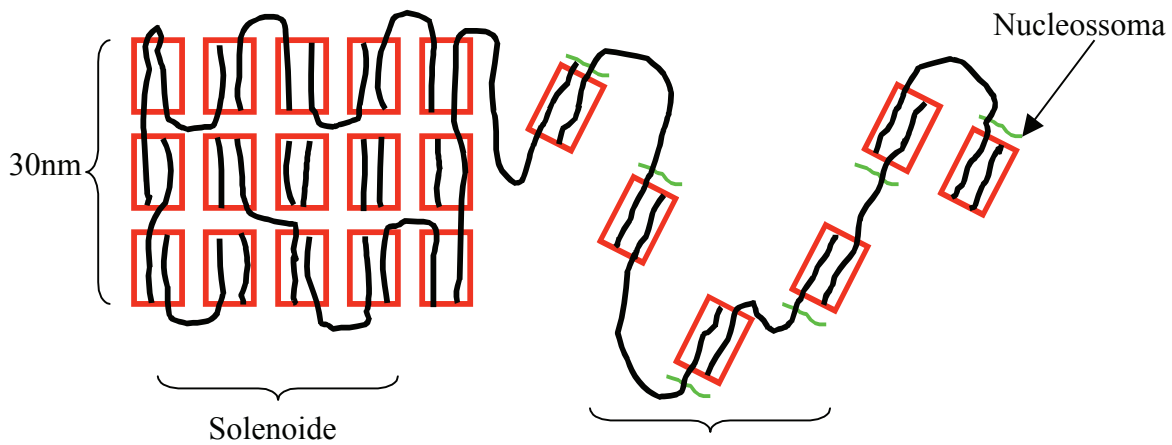




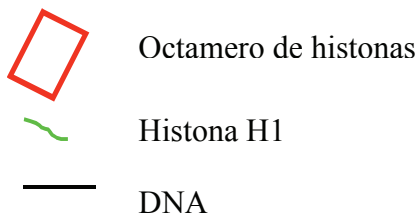
Ansas de cromatina ligadas a regiões “scaffold”



A cromatina contém pequenas quantidades de outras proteínas que se ligam ao DNA sem serem histonas ou proteínas do *scaffold* como, por exemplo, factores de transcrição e outros tipos de proteínas. Na sua forma condensada a cromatina apresenta uma estrutura de tipo solenóide com fibras de cerca de 30nm de diâmetro e com uma histona H1 associada a cada nucleossoma.



Forma não condensada espiralada com seis nucleossomas por espira



Portanto duas formas de cromatina podem existir: a forma condensada e a não condensada ou distendida, dentro do mesmo cromossoma. As regiões cromossomais que não estão sendo transcritas encontram-se predominantemente na forma condensada, enquanto as distendidas estão provavelmente sendo transcritas (*vide* adiante parte II. 9. transcrição). A cromatina, organizada em grandes unidades de centenas a milhares de kilobases de comprimento, constituindo os cromossomas, encontra-se em número e forma que são específicos das espécies de seres vivos (*vide* quadro seguinte).

Número de cromossomas de diversas espécies animais:

<u>Nome comum</u>	<u>Espécies</u>	<u>Número diplóide</u>
Humano	<i>Homo sapiens</i>	46
Macaco resus	<i>Macaca mulatta</i>	42
Bovinos	<i>Bos taurus</i>	60
Suínos	<i>Sus scrofa</i>	40
Cães	<i>Canis familiaris</i>	78
Gatos	<i>Felis domesticus</i>	38
Equídeos	<i>Equus caballus</i>	64
Rato	<i>Mus musculus</i>	40
Ratazana	<i>Rattus norvegicus</i>	42
Hamster dourado	<i>Mesocricetus auratus</i>	44
Cobaio	<i>Cavia cobaya</i>	64
Coelho	<i>Oryctolagus cuniculus</i>	44
Frango	<i>Galus domesticus</i>	78 x
Carpa	<i>Cyprinus carpio</i>	104

x – Varia consoante as estirpes

Na maioria dos organismos o número de cromossomas não varia e cada cromossoma contem uma molécula de DNA linear.

**1.2 -Cromossomas das células eucariotas e expressão do seus genes.** (Stryer,L.,1995)

**Genes, loci e alelos** (Bradley et ali.,2001)

**Localização nos cromossomas de alguns genes** (Bradley et ali.,2001)

Os cromossomas das células eucariotas são maiores e têm uma maior complexidade estrutural do que os cromossomas das células procariotas. Como se disse anteriormente um cromossoma numa célula eucariota contem uma molécula simples e linear de DNA em hélice dupla, encontrando-se profundamente ligada às histonas que constituem cerca de metade da massa dos cromossomas. Os cinco tipos de histonas, (não consideramos o 6º tipo: as H5), H1, H2A, H2B, H3 e H4, todas elas contêm um alto teor de ácidos aminados básicos com as cadeias laterais dos ácidos aminados a estarem carregadas electricamente com sinal positivo. Cada tipo de histona pode existir em diversas formas em virtude das modificações (após a tradução) de certas cadeias laterais dos seus resíduos de ácidos aminados. As sequências de ácidos aminados das histonas H3 e H4 são quase as mesmas em todas as plantas e animais.

A maioria do DNA está enrolado por fora do ângulo das histonas, sendo o DNA que se encontra entre os nucleossomas (o *linker* ou espaçador) aquele que contribui para a flexibilidade da fibra de cromatina. Uma fibra de cromatina é pois uma cadeia de nucleossomas unidos de uma forma flexível.

Os genomas dos eucariotas superiores contêm diversas sequências de DNA repetidas, calculando-se que nos humanos mais de 30% do genoma consiste em sequências repetidas pelo menos 20 vezes (Stryer, L., 1995). Famílias de sequências de DNA repetidas de 100 a 500 bp (*base pairs* ou pares de bases) de comprimento são dispersas no genoma e denominadas SINES (*short interspersed repeats*)(Stryer,L.,1995). Outras sequências repetidas mais longas (varios Kb ou mais) são denominadas LINES (*long interspersed repeats*). Alu são sequências especialmente abundantes nos SINES, sequências com 300bp que se repetem cerca de um milhão de vezes no genoma humano. A maior parte de cada fragmento humano de 20Kb (*kilo base pairs* ou milhares de pares de bases) contem uma sequência Alu (a designação provem da endonuclease de restrição Alu I que corta várias destas sequências). O grau de identidade das sequências entre dois membros desta família é de cerca de 85%. As sequências Alu são similares ao 7SL RNA, a pequena molécula de RNA citoplásmico da partícula reconhecadora de sinal (Stryer, L., 1995). Os bovinos, que têm 29 pares de autossomas e dois cromossomas do sexo, possuem no seu genoma 50% de sequências repetitivas (SINES e LINES) e com microssatélites (*vide* adiante) altamente polimórficos. As funções dos SINES e LINES são um mistério. Algum DNA altamente repetitivo é agrupado mais do que disperso no cromossoma, podendo muito deste DNA ser isolado por centrifugação em gradiente de densidades porque os fragmentos que os contêm tem densidades *buoyant* distintas (bandas satélites) Exemplo (Stryer, L., 1995)

5'-ACAAACT-3'	5'-ATAAACT-3'	5'-ACAAATT-3'
3'-TGTTTGA-5'	3 - TATTTGA-5'	5'-TGTTTAA-3'
Satélite I	Satélite II	Satélite III

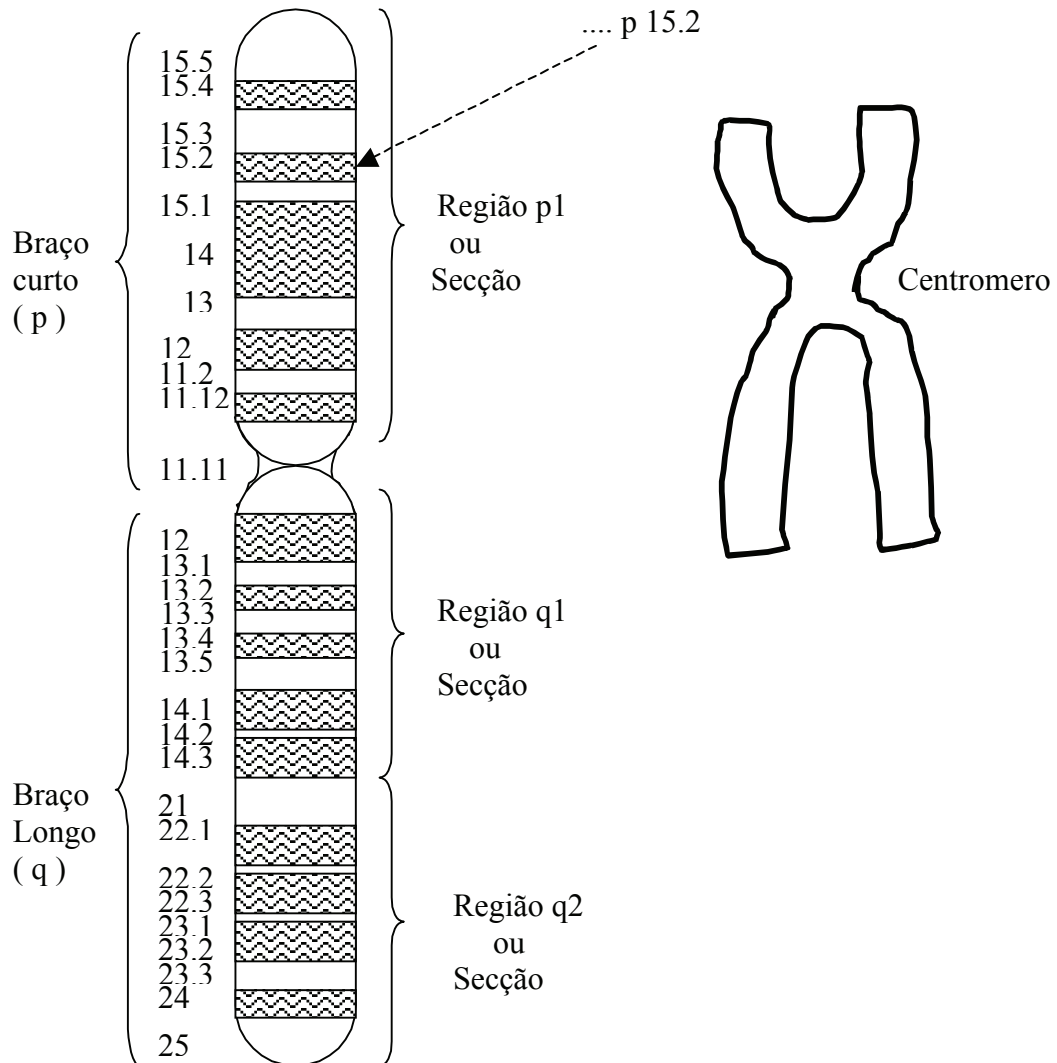
O DNA satélite pode ser determinado por hibridação *in situ* e quase todo o DNA satélite está localizado nos centromeros. Também os genes para os RNAs ribossomais estão repetidos em tandem várias centenas de vezes. Os genes das histonas são agrupados e repetidos em tandem várias vezes. Os genes das histonas são desprovidos de intrões e não tem os mRNAs dessas histonas caudas poly(A) (extensões 3' polidesoxiadenilatadas). Várias proteínas importantes são codificadas por um único gene, portanto com uma cópia só, podendo um simples gene ser suficiente para a síntese de 10<sup>9</sup> moléculas da respectiva proteína. No entanto genes de que apenas existe uma cópia no genoma podem ser altamente amplificados por pressão selectiva, tornando-se assim genes altamente repetidos, e chegando em certos casos a contarem-se centenas de cópias dum único gene, podendo uma região do cromossoma adquirir um conjunto em tandem desse gene. Esta descoberta levou a encarar a natureza dinâmica do genoma. Apenas uma pequena parte do genoma dos mamíferos codifica proteínas, calculando-se que apenas 2% do DNA tem esta função.

Genes, Loci e Alelos  
(Bradley, J. et ali., 2001)

“Uma forma particular de um gene é denominada um alelo e a posição do gene no cromossoma é o seu *locus*”(Bradley et ali.,2001). O alelo é pois uma de duas ou mais formas alternativas de um gene localizado no correspondente sítio ou *locus* sobre cromossomas homólogos. Geralmente, existe apenas um ou poucos alelos diferentes para um dado gene, mas alguns genes são altamente polimórficos (podendo existir pois diferentes alelos). Quando um indivíduo tem dois alelos idênticos, ele é homocigoto nesse *locus*. Quando tem um só alelo, com tais genes no cromossoma X e Y, o indivíduo é hemizigoto nesse *locus*. Quando um indivíduo tem dois alelos diferentes, ele é heterocigoto nesse *locus*. Apesar de cada cromossoma de um par ser morfologicamente similar, o seu material genético exacto varia em consequência da recombinação dos cromossomas e da junção ao acaso durante a meiose, e das mutações do DNA. A prevalência das mutações varia ao longo do genoma. Onde ocorre maior variação é ao nível das regiões não codificadoras como os intrões, as regiões flaqueadoras, e as sequências de DNA

repetitivas, mais do que nos exões. As mais altas taxas de mutações ocorrem nas sequências repetitivas não codificadoras como é o caso das regiões minisatélites. Cada cromossoma contem como se disse uma simples molécula de DNA e certos corantes coram selectivamente certas regiões dos cromossomas em metafase, mais intensamente do que outras regiões, produzindo um perfil de bandas específicas e características para cada cromossoma. Embora não se conheça bem a base molecular deste comportamento ele é no entanto muito útil até para distinguir não só os cromossomas uns dos outros, mas também para distinguir cromossomas de tamanho idêntico e forma, e identificar partes específicas de um dado cromossoma. Por outro lado, sondas de DNA clonado que hibridam com sequências específicas nos cromossomas podem ser localizadas em bandas particulares. Os cromossomas condensados durante a metafase podem ser tratados por vários processos, como por exemplo corados pelo Giemsa ou pela quinacrina, o que promove neles a formação de bandas transversais coradas mais claras e mais escuras, alternadamente, que dividem assim os cromossomas em diversas áreas facilitando a localização por marcadores adequados de variadíssimos genes e outros detalhes estruturais, no braço mais curto (p) ou no mais longo (q) e a sua posição relativamente ao centromero (*vide* figura seguinte). As regiões de um e do outro lado do centromero (uma parte compacta do cromossoma) são chamadas braços e como o centromero não se encontra no centro do cromossoma um braço é mais comprido do que o outro.

Esquema de um Cromossoma com bandas G (Adaptado de Bradley, J., 2001)



O sistema de numeração indica a posição sobre os cromossomas como se ilustra.  
p = braço curto      q = braço longo

Cada braço é dividido em regiões ou secções, bandas e sub-bandas ou sub-regiões ou sub-secções.

No exemplo da figura anterior exemplifica-se com a seta a banda cuja localização refere primeiro o número do cromossoma (...) depois p, a indicar o braço curto, 1 a indicar a região 1, depois 5 a indicar a sub-região ou banda dentro da região e finalmente 2 a indicar a sub-banda.

Localização nos cromossomas de alguns genes

Dado o interesse comparativo da identificação de alguns genes, quer para efeitos de caracterização de algumas situações patológicas, quer de interesse na endocrinologia e na reprodução, quer ainda dos genes em geral já identificados em algumas espécies de animais, passamos a transcrever, uma selecção por nós efectuada daquilo que já hoje se conhece a este propósito.

**1.3 - Localização nos cromossomas humanos de alguns genes relacionados com doenças**  
(Bradley, J. et ali., 2001)

<b>Doenças monogénicas</b>	<b>Localização</b>
Deficiência em adenosina de aminase ( ADA )	20q12-13
Adrenoleucodistrofia	Xq28
Doença do rim poliquístico adulto	16p13.3 (PKD1) 4q13-23 (PKD2)
Deficiência alfa 1-antitripsina	14q
Fibrose quística	7q31.2
Síndrome George	22
Distrofia muscular Duchenne	Xp21.2
Hipercolesterolemia familiar	19p
Síndrome frágil X	Xq27
Ataxia de Friedreich	9q13-21
Febre mediterrânica familiar	19p13.3
Haemocromatose	6p21.3
Haemoglobinopatias	cromossoma 11 _globina 16 ( _ globina) 11p15.5 ( _ talassemia)
Hemofilia	Xq28 (factor coagulante VIII) xq27 (factor IX)
Doença de Huntington	4p16.3
Síndrome de Marfan	15q21.1
Mucopolissacaridoses	4p16.3 Xq28
Distrofia miotónica	19q133
Fenil cetonúria	12q22-24
Doença de Wilson	13q14.2-21
Síndrome Wiskott-Aldrich	Xp11
Hipertensão forma monogénica	16

Desordens poligénicas

Diabetes mellitus	6p21 (tipo 1) 11p15 (tipo 2)
Doença de Alzheimer	21 14 1

Cancro

Limfoma de Burkitt	14 q 32 8 q 34
Cromossoma Philadelphia	22 q 11 9 q 34
Ataxia telangiectasia	11 q 23.1
Cancro da mama	17 q 21 (BRCA 1) 13 q 12.3 (BRCA 2)
Cancro do colon	3 p 21.3
Síndrome Li-Fraument	17 p 13.1
Neoplasia múltipla endócrina 1	11 q 13

Neurofibromatose 1	17 q 11
Retinoblastoma	13 q 14.2
Esclerose tuberosa	9 q 34 (TSC 1) 16 p 13 (TSC 2)
Síndrome Von Hippel-Lindau	3 p 26
Tumor de Wilm	11 p 13

Refere-se seguidamente o catalogo de genes porcinos (64 *loci*) de interesse em endocrinologia e reprodução, transcritos do *site*:

<http://www.genome.iastate.edu/maps/reprogenes.html>.

Esta página foi preparada por A. G. Amador & R. King e espelha o seu *site* original.

Referem-se em cada cromossoma dos suínos o número de *loci* e a localização, símbolo, e nome de cada gene.

#### 1.4 - Catálogo de genes porcinos de interesse em endocrinologia e reprodução (64 *loci*)

Localização	Símbolo	Designação	Cromossoma
1 p 25 – p 24	<i>ESR</i>	Receptor estrogénio	<b>1</b>  <b>5 loci</b>
1 q 15 – q 16	<i>CGA</i>	Hormona glicoproteína da pituitária, polipeptideo alfa	
1 q 25	<i>IFN A</i>	Interferão, alfa, leucócito	
1 q 25 – q 26	<i>IFN B1</i>	Interferão, beta 1	
1 q 28 – q 29	<i>RLN</i>	Relaxina	

Localização	Símbolo	Designação	Cromossoma
2 p 16 – p 12	<i>FSHB</i>	Folitrofina, beta polipeptideo	<b>2</b>  <b>7loci</b>
2 q 11 – q 21	<i>INSR</i>	Receptor da insulina	
2 q 12 – q 13	<i>LCIP</i>	Proteína insulina like da célula de Leyding	
2	<i>PDGFRB</i>	Receptor do factor de crescimento derivado das plaquetas beta polipeptideo	

Localização	Símbolo	Designação	Cromossoma
3p13-p11	<i>PRM1</i>	Protamin 1	<b>3</b>  <b>7 loci</b>
3p12-p11	<i>TNP2</i>	Transition protein 2	
3q11-q13	<i>IL1A</i>	Interleukin 1 alpha	
3q11-q14	<i>IL1B</i>	Interleukin 1 beta	
3q22-q23	<i>FSHR</i>	Follitropin receptor	
3q22-q23	<i>LHCGR</i>	Lutropin/choriogonadotropin receptor	
3	<i>PRM2</i>	Protamine 2	

Localização	Símbolo	Designação	Cromossoma
4q15-q24	<i>NGFB</i>	Nerve growth factor, beta polypeptide	4 5 Loci
4	<i>AFD</i>	Abdominal fat deposition	
4	<i>HSP</i>	Heat shock protein	
4	<i>TGR</i>	Total growth rate (birth 70 Kg)	
4	<i>TSHB</i>	Thyrotropin (thyroid stimulating hormone), beta polypeptide	

Localização	Símbolo	Designação	Cromossoma
5p15	<i>ACR</i>	Prepro-acrosin	5 3 Loci
5p12-q11	<i>IFNG</i>	Interferon, gamma	
5q23-q24	<i>IGF1</i>	Insulin-like growth factor 1	

Localização	Símbolo	Designação	Cromossoma
7p11	<i>HSPA1</i>	Heatshock protein 70 kDa protein 1	7 5 Loci
7p11-q11	<i>TNFA</i>	Tumor necrosis factor, alpha	
7p11-q11	<i>TNFB</i>	Tumor necrosis factor, beta	
7	<i>CYP21</i>	Cytochrome P450, subfamily XXI (steroid 21-hydroxylase)	
7	<i>PRL</i>	Prolactin	

Localização	Símbolo	Designação	Cromossoma
13q31	<i>TF</i>	Transferrin	
13	<i>IGR</i>	Initial growth rate (birth to 30 Kg)	13
13	<i>PIT1</i>	Pituitary-specific transcription factor (growth hormone factor 1)	3 Loci

Localização	Símbolo	Designação	Cromossoma
16q13-q14	<i>FSA</i>	Follistatin	16
16q13-q14	<i>GHR</i>	Somatotropin receptor	2 Loci

Localização	Símbolo	Designação	Cromossoma
17	<i>ENDO</i>	dynorphin	17 :1 Locus

Localização	Símbolo	Designação	Cromossoma
18q24	<i>IGFBP 3</i>	Insulin-like growth factor binding protein 3	
18q24	<i>INHBA</i>	Inhibin beta A	18
18q	<i>CFTR</i>	Cystic fibrosis transmembrane conductance regulator	3 Loci



Localização	Símbolo	Designação	Cromossoma
X	ZFX	Zinc finger protein, X-linked	X : 1 Locus

### 1.5 - Lista de genes de bovinos

*Bovmap Database*

*World Wide Web version 2.00*

*Institute Nacional de Recherche Agronomique*

*Laboratoire de Génétique Biochimique de Jouy-en-Josas*

Lista parcial entre 1553 genes

Transcreve-se seguidamente a lista de alguns dos 1.563 genes contidos no *Bovmap Database* do Laboratório de genética bioquímica-Jouy-en-Josas do Instituto Nacional de Investigação Agronómica na versão *World Wide Web 2.00* e que se encontra no *site*:

<http://locus.jouy.inra.fr/cgi-bin/lgbc/mapping/common/main.pl?BASE=>

Nesta lista são referidas numa coluna as designações abreviadas dos genes, na coluna seguinte a sua localização cromossómica nos bovinos, e na última coluna pormenorização das suas características que podem ser aprofundadas “clicando” no símbolo do gene, o que fornece mais detalhes.

A lista de genes OMIM com milhares e milhares de genes humanos e suas características, pode ser consultada no *site*: <http://www.ncbi.nlm.nih.gov/omim/>

Para criar uma link do documento www para um registo específico OMIM, utilizar o seguinte URL:

<http://www.ncbi.nlm.nih.gov/htbin-post/OMIM/dispim?MIMnumero>

O *guideline* para as nomenclaturas de genes humanos foi actualizado em Abril de 2002 após a recente publicação da sequências do genoma humano completo e que calculou um número total de 26.000 – 40.000 genes podendo ser consultado no *site*: <http://www.gene.UCL.ac.UK/nomenclature/guidelines.html>

Todos os símbolos dos genes humanos aprovados podem ser encontrados no *Genew database* referido por:

Wath, H. M., Lush, M., Dduzeau, F., Pouy, S. (2002). *Genew: The human nomenclature database. Nucleic Acids Research. 30: 169-171.*

Genes homólogos em diferentes espécies vertebradas (ortólogos) devem quando possível ter a mesma nomenclatura nos genes.

A concordância entre a nomenclatura dos genes humanos e do rato para diversos genes homólogos deve ser continuada e estendida a outras espécies de vertebrados tanto quanto possível.

<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
ACACA	19q13-q14	Acetyl-coenzyme A carboxylase alpha [acetyl-coA carboxylase] [acetyl coenzyme A carboxylase A]
ACADM	3	Acyl-coenzyme A dehydrogenase, C-4 to C-12 straight chain
ACADVL	19	Acyl-coenzyme A dehydrogenase, very long chain
ACF7	3	Actin-crosslinking protein
ACHE		Acetylcholinesterase
ACK1	6	Activated p21cdc42Hs kinase
ACLY	19q17	ATP citratelase
ACO1	8	Aconitase 1, soluble
ACO2	5q35	Aconitase 2, mitochondrial
ACP1		Acid phosphatase 1, soluble [low molecular weight phosphotyrosine protein phosphatase] [ptpase]
ACP2	15	Acid phosphatase 2, lysosomal
ACP5	7	Tartrate-resistant acid phosphatase type 5 precursor (human homolog)
ACR	5q35	Acrosin
ACTA1	28	Actin, alpha 1, skeletal muscle
ACTA2	26	Actin, alpha 2, smooth muscle, aorta
ACTB		Actin, beta
ACTBP		Actin, beta, pseudogene
ACTG2	11	Actin, gamma 2, smooth muscle, enteric [actin]
ACTL6	25	Actin-like 6 (human homolog)
ACTR3	2	ARP3 (actin-related protein 3, yeast) homolog
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
ACVR2	2q23-q24	Activin A receptor, type II [activin receptor, type IIA] [activin receptor 2]
ACVR2B	22	Activin A receptor, type IIB
ACY1	22	Aminoacylase 1
AD-020	3	Protein x 013 (human homolog) [Bovine STS AF213862]
ADA	13	Adenosine deaminase
ADAM10	10	A disintegrin and metalloproteinase domain 10
ADAM11	19	A disintegrin and metalloproteinase domain 11 (human homolog)
ADCY1	4	Adenylate cyclase 1 (brain)
ADCY2	15q14-q21	Adenylate cyclase 2
ADCY7	18	Adenylate cyclase 7
ADCYAP1	24q23	Adenylate cyclase activating polypeptide 1 (pituitary) [pituitary adenyl cyclase activating peptide]
AGTR1	1	Angiotensin receptor 1 [angiotensin II receptor 1]
AGTR1B	1	Angiotensin II receptor type 1B
ALB	6	Albumin
ALCAM	1	Activated leucocyte cell adhesion molecule
ALDH1	8	Aldehyde dehydrogenase 1, soluble
ALDH2	17	Aldehyde dehydrogenase 2, mitochondrial
ALDH3		Aldehyde dehydrogenase 3, dimeric NADP- preferring

		[corneal protein 54] [transparentin]
ALDH5		Aldehyde dehydrogenase 5, mitochondrial X
ALDOB	8	Aldolase B, fructose bisphosphate
ALOX12	19	Arachidonate 12-lipoxygenase
ALPI	2q45	Alkaline phosphatase, intestinal
ALPL	2	Alkaline phosphatase, liver / bone / kidney
AMY1A	3	Amylase, alpha 1A; salivary
AMY2A	3	Amylase, alpha 2A; pancreatic
AQP4	24	Aquaporin 4
AR	Xq26	Adrogeenn receptor
ARCNI		Archain – coatomer protein delta-COP
ARPC2	2	Actin related protein 2/3 complex, subunit 2 (34 kD)
ARPP-19	10	Cyclic AMP phosphoprotein, 19 kDa (human homolog)
ARRB1	15q25	Arrestin, beta 1 [arrestin B]
ARRB2	19	Arrestin, beta 2
ASL	25	Argininosuccinate lyase
ASMT	Y	Acetylserotonin O-methyltransferase
ASPA	19	Aspartoacylase [aminoacylase 2]
ASS	11	Argininosuccinate synthetase
AVP	13q21-q22	Arginine vasopressin [neurophysin II]
AW289261	3	EST AW289261
B2M	10	Microglobulin, beta-2
B4GALT1	8	Beta-galactosyltransferase 2 (glycoprotein) [N-acetylglucosamine galactosyltransferase] [beta-1. 4-galactosyltransferase]
B4GALT5	13	UDP-Gal :betaGlcNAc beta 1, 4- galactosyltransferase, polypeptide 5
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
B9	19	Homeo box-B9 (human homolog)
BAK1	23	BCL2-antagonist / killer 1
BLVRB	18	Biliverdin reductase B [flavin reductase (NADPH)]
BM045	25	Uncharacterized bone marrow protein BM045 (human homolog)
BMP1	8q21	Bone morphogenetic protein 1
BMP3	6	Bone morphogenetic protein 3 (osteogenic)
BOLA	23	Major histocompatibility complex
BOLA-A	23q22	Major histocompatibility complex, class I, A
BZRP	5	Benzodiazepine receptor (peripheral)
C12ORF22	5	Chromosome 12 open reading frame 22 (human homolog)
C1ORF9	16	Membrane protein CH1 (human Homolog) [chromosome 1 open reading frame 9 (human homolog)] [Bovine STS AF213847]
C20ORF19	13	Chromosome 20 open reading frame 19 (human homolog)
C20ORF33	13	Chromosome 20 open reading frame 33 (human homolog)
C2ORF2	11	Chromosome 2 open reading frame 2 (human homolog)
C2ORF	11	Transcription factor 9 (binds GC-rich sequences)

CA4		
CACNB4	19	Carbonic cnhydrase IV
CAD	2	Calcium channel, voltage-dependent, beta 4 subunit
CAPN3	11	Carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase (human homolog)
CAPNS1	10	Skeletal muscle-specific calpain [calcium activated neutral protease] [calpain 3]
CAPZA1	18	Calpain 4, small subunit (30K)
CASR	3	Capping protein (actin filament) muscle Z-line, alpha 1
CAST	1	Calcium-sensing receptor
CAT	7	Calpastatin
CCNA	15	Catalase
CCNB2	6	Cyclin A
CCNB2	10	Cyclin B2
CCND1	29qter	Cyclin D1 [parathyroid adenomatosis 1]
CCND2	5	Cyclin D2
CCNG1	19	CyclinG1
CCNH	7	Cyclin H
CDH2	5	Cyclin-dependent kinase 2
CDK5	4	Cyclin-dependent kinase 5
CDK5R1	19	Cyclin-dependent kinase 5, regulatory subunit 1 (p35)
CFTR	4q23-q25	Cystic fibrosis transmembrane conductance regulator
CHS1	28q13-q14	Lysosomal trafficking regulator
CINF	1	Bovine chronic interstitial nephritis with diffuse zonal fibrosis
CITED1	X	Melanocyte specific gene 1 [cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 1]
CKB	21	Creatine kinase, brain
CLCN1	4	Chloride channel 1, skeletal muscle
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
CNGA1	6	Cyclic nucleotide gated channel (photoreceptor), cGMP gated
CNP	19	2', 3'-cyclic nucleotide 3' phosphodiesterase
CNR1	9q22	Cannabinoid receptor (brain)
CNTRF	8q21	Ciliary neurotrophic factor receptor
COL10A1	9	Collagen, type X, alpha 1 [Schmid metaphyseal chondrodysplasia]
COL1A1	19	Collagen, type I, alpha 1
COL1A2	4	Collagen, type I, alpha 2
COL2A1	5	Collagen, type II, alpha 1 [collagen 2 alpha 1]
COL3A1	2q12	Collagen, type III, alpha 1 chain, fibril forming [colagenn 3 alpha 1]
COL4A1	12	Collagen, type IV, alpha 1 chain, fibril forming [colagenn 4 alpha 1]
COL5A2	2q14-q15	Collagen, type V, alpha-2 chain
COL6A1	1q12-q14	Collagen, type VI, alpha-1
COL6A2	1	Collagen, type VI, alpha-2
COL9A2	3	Collagen, type IX, alpha-1 (human homolog)

COMP	7	Cartilage oligomeric matrix protein
COX4		Cytocrome c oxydase subunit 4
COX6A1	17	Cytocrome c oxydase subunit Via polypeptide 1
COX6A2	25	Cytocrome c oxydase subunit Via polypeptide 2
COX8	29	Cytocrome c oxydase subunit VIII, heart
COXP1	5	Cytocrome c oxydase, pseudogene
CP	1	Ceruloplasmin [ferroxidase]
CRI1	10	CREBBP/EP300 inhibitory protein 1 (human homolog)
CROT	4	Carnitine O-octanoyltransferase (human homolog)
CSF1	3	Colony stimulating factor 1 (macrophage)
CSF1R	7	Colony stimulating factor 1 receptor
CSF2	7	Colony stimulating factor 2 [granulocyte-macrophage]
CSF2RA	Xq42-q43	Colony stimulating factor 2 receptor, alpha, low-affinity, (granulocyte-macrophage)
CSF3R	3	Colony stimulating factor 3 receptor (granulocyte)
CSH1	23	Placental lactogen [chorionic somatomammotropin hormone 1]
CSN10	6q32	Casein kappa
CSN1S1	6q31	Casein alpha-S1
CSN1S2	6q31	Casein alpha-S2
CSN2	6q31	Casein beta
CSN@	6q31	Casein ( cluster )
CSNK1G2	7	Casein kinase 1, gamma 2
CSNK2A1	13	Casein kinase 2, alpha 1 polypeptide
CSNK2A2	18	Casein kinase 2, alpha prime polypeptide
CSNK2B	23	Casein kinase 2, beta polypeptide
CYP11B1	14	Cytochrome P450, subfamily XI B [steroid 11-beta-hydroxylase], polypeptide 1
CYP17	26q22-31	Cytochrome P450, subfamily XVII [steroid 17-alpha-hydroxylase]
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
CYP19	10q31-q32.1	Cytochrome P450, family XIX (conversion of androgen to estrogen), aromatase
CYP19P1	10q26	Cytochrome P450, family XIX (conversion of androgen to estrogen), aromatase, pseudogene
CYP19p1.2	10q26	CYP19 promoter 1.2
CYP21A2	23q12d-q13p	Cytochrome P450, subfamily XXI [steroid 21-hydroxylase] [cytochrome P450 21 hydroxylase]
CYP2C19	26	Cytochrome P450, subfamily IIC, polypeptide 19 [mephenytoin 4-hydroxylase]
CYP2D@	5	Cytochrome P450, subfamily II D [debrisoquine, sparteine, etc., metabolizing] cluster [cytochrome P2D]
CYP2E	26	Cytochrome P450, subfamily IIE (ethanol-inducible)
CYP51	4	Cytochrome P450, 51 (human homolog) [lanosterol 14-alpha-demethylase (human homolog)]
DMD	Xq34	Dystrophin [muscular dystrophy]
DNAJC5	13	Dnaj (Hsp40) homolog, subfamily C, member 5 (human homolog)

DNASE2	7	Deoxyribonuclease II, lysosomal
DNMT1	7q15	DNA methyltransferase [(cytosine 5) methyltransferase]
DNMTA3	11	DNA (cytosine-5-)-methyltransferase 3 alpha (human homolog)
DNTT	26	Deoxynucleotidyltransferase, terminal [deoxyribonucleotidyl transferase]
DOCK2	20	Dedicator of cyto-kinesis 2
DPH2L2	3	Diphtheria toxin resistance protein required for diphthamide biosynthesis-like 2 (S. Cerevisiae) (human homolog)
DPP4	2	Dipeptidylpeptidase IV (human homolog) [CD26, adenosine deaminase complexing protein 2 (human homolog)]
DPYD	3	Dihydropyrimidine dehydrogenase
DRD2	15	Dopamine receptor D2
EGF	6	Epidermal growth factor [beta-urogastrone]
EPHX1	16	Epoxide hydrolase 1, microsomal (xenobiotic)
EPO	25	Erythropoietin
EPOR	7	Erythropoietin receptor
ESD	12	Esterase D [formylglutathione hydrolase]
ESR1	9	Estrogen receptor 1 [estrogen receptor alpha] [estrogen receptor]
FDFT1	8	Farnesyl- diphosphate farnesyltransferase 1
FDX1	15	Ferredoxin 1
FECH	24	Ferrochelatase [protoporphyrin]
FEN1	29	Flap structure-specific endonuclease 1
FES	21	Feline sarcoma viral (v-fes) oncogene; v-fps oncogene homolog
FGB	17	Fibrinogen, B beta polypeptide
FGF1	7	Fibroblast growth factor, acidic [endothelial growth factor] [fibroblast growth factor alpha]
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
FGF2	17	Fibroblast growth factor 2 (basic) [fibroblast growth factor B]
FGF3	29	Fibroblast growth factor 3
FGF5	6	Fibroblast growth factor 5
FGFR1	27q18-q19	Fibroblast growth factor receptor 1
FGFR2	26	Fibroblast growth factor receptor 2 [bacteria-expressed kinase] [keratinocyte growth factor receptor]
FGFR3	6	Fibroblast growth factor receptor 3
FGG	17q13	Fibrinogen, gamma polypeptide [fibrinogen gamma]
FN1	2	Fibronectin 1
FNTA	27q18-q19	Farnesyltransferase, CAAX box, alpha
FOS	10	Murine FBJ osteosarcoma viral (v-fos) oncogene homolog
FOXO1A	1	Forkhead box L2 (human homolog)
FOXO1A	12	Forkhead box o1a [ human: rhabdomyosarcoma]
FRZB	2	Frizzled-related protein
FSHB	15q25dist-q26	Follicle stimulating hormone, beta polypeptide
FSHR	11	Follicle stimulating hormone receptor

FUCA1	2	Follicle, alpha-L-1, tissue
G6PD	X	Glucose-6-phosphate dehydrogenase
G6PT1	15	Glucose-6-phosphate, transport (glucose-6-phosphate) protein 1
GAA	19	Glucosidase, alpha; acid
GABRA1	7	Gamma-aminobutyric acid (GABA) A receptor, alpha 1
GABRA2	6	Gamma-aminobutyric acid (GABA) A receptor, alpha 2
GAD1	2q21	Glutamic acid decarboxylase 1
GADD45B	7	Myeloid differentiation primary response gene
GAL	29	Galanin
GALK2	10	Galactokinase 2
GAP43	1	Growth associated protein 43
GAPD	5	Glyceraldehyde-3-phosphate dehydrogenase
GCG	2q21-q22	Glucagon
GCNT1	8	Glucosaminyl (N-acetyl) transferase 1, core 2 [beta-1, 6-N-]
GDF8	2q14-q15	Growth differentiation factor 8 [myostatin][muscular hypertrophy]
GDH	5	Glucose dehydrogenase
GDI1	X	GDP dissociation inhibitor 1
GFAP	19	Glial fibrillary acidic protein
GFI1B	11	Growth factor independent 1B (human homolog) [potential regulator of CDKN1A, translocated in CML (human homolog)]
GGCX	11	Gamma-glutamyl carboxylase
GGTA1	11q26	Alpha-galactosyltransferase 1 (glycoprotein) [alpha-1-3 galactosyl transferase]
GH1	19q22prox	Growth hormone 1 [growth hormone beta]
GHR	20q17	Growth hormone receptor
GHRH	13	Growth hormone releasing hormone
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
GHRHR	4	Growth hormone releasing hormone receptor
GPI	18q24prox	Glucosephosphate isomerase
GPRK5	26	G protein-coupled receptor kinase 5
GPX1	22	Glutathione peroxidase 1
HBA1	25	Hemoglobin, alpha 1
HBB	15q22-q27	Hemoglobin, beta [beta globin] [haemoglobin beta]
HBP1	4	High blood pressure type 1 (human homolog)
HCK	13q22	Hemopoietic cell kinase
HK1	28	Hexokinase 1
HOXA		Homeo box
HOXA @	4	Homeobox region 1
HOXB @	19q17-qter	Homeobox region 2
HOXC @	5q13-q22	Homeobox region 3
HOXD8	2	Homeo box D8 (human homolog) [Bovine STS AF213873]
HP	18	Haptoglobin
HPRT	X	Hypoxanthine-guanine phosphoribosyltransferase [hypoxanthine phosphoribosyltransferase]

HPRT 1	X	Hypoxanthine phosphoribosyltransferase [human Lesch-Nyhan syndrome]
HPS	1	Horned / polled syndrome
HRH1	22	Histamine H1 receptor
HSD11B2	18	Hydroxysteroid (11-beta) dehydrogenase 2
HSD3B1	3q21dist	Hydroxy-delta-MAY-steroid dehydrogenase, 3 beta-and steroid delta-isomerase
HSPA1A	23q13	Heat shock 70 kD protein 1
HSPA1B	23q22	Heat shock 70 kD protein 2
HSPA3	10q34	Heat shock 70 kD protein 3
HSPA5	11	Heat shock 70 kD protein 5 [glucose-regulated protein, 78 kD]
HSPA6	3q13	Heat shock 70 kD protein 4
HSPA8	15	Heat shock cognate 71 kD protein [heat shock 70 kD protein 10 (HSC71)]
IDH1	2	Isocitrate dehydrogenase 1 (NADP+), soluble
IDH3B	13	Isocitrate dehydrogenase 3 (NADP+) beta
IF	6	Complement component 1 [I factor (complement)]
IFN1@	8q15prox	Interferon, alpha, leukocyte
IFNAR1	1	Interferon, alpha receptor
IFNAR2	1	Interferon, beta ; receptor
IFNB1	8q15	Interferon, beta 1, fibroblast
IFNG	5q23	Interferon, gamma or immune type [interferon gamma type 2]
IFNT	8q15	Inteferon, trophoblast
IFNW1	8q15	Interferon, omega
IGF1	5	Insulin like growth factor 1 [insulin-like growth factor I] [somatomedin C]
IGF1R	21	Insulin-like growth factor 1 receptor
IGF2	29q24dist	Insulin-like growth factor 2
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
IGF2R	9q26dist	Insulin-like growth factor 2 receptor
IGFBP3	4	Insulin-like growth factor binding protein 3
IGFBP4	19	Insulin-like growth factor binding protein 4
IL10	16q12	Interleukin 10
IL12A	1q34-q36	Interleukin 12A [natural killer cell stimulatory factor 1, cytotoxic lymphocyte maturation factor 1, p35]
IL12B	7q23-q24	Interleukin 12A [natural killer cell stimulatory factor 2, cytotoxic lymphocyte maturation factor 2, p40]
IL15	17q13-q14	Interleukin 15
IL18	15q13-q14	Interleukin 18 [interferon-gamma-inducing factor]
IL1A	11	Interleukin 1, alpha
IL1B	11q23-q24	Interleukin 1, beta
IL1R1	11q12	Interleukin 1 receptor, alpha ; type I receptor [interleukin receptor 1]
IL1R2	11	Interleukin 1 receptor, beta ; type II receptor [interleukin receptor 2]
IL1RN	11	Interleukin 1 receptor antagonist
IL2	17q22-q23	Interleukin 2



IL2RA	13q13dist	Interleukin 2 receptor
IL2RG	Xq22-q23	Interleukin 2 receptor, gamma
IL3	7q15-q21	Interleukin 3 [colony-stimulating factor, multiple]
IL4	7q15-q21	Interleukin 4
IL5	7q15-q21	Interleukin 5 [colony-stimulating factor, eosinophil]
IL6	4	Interleukin 6 [interferon, beta 2]
IL6R	19	Interleukin 6 receptor
IL7	14q24-q26	Interleukin 7
IL7R	20q15-q17	Interleukin 7 receptor
IL8	6	Interleukin 8 [neutrophil activating peptide 1]
IL8R	2	Interleukin 8 receptor
IMPA1	14	Inositol (myo)-1 (or 4)-monophosphatase 1
INHA	2q36-q42	Inhibin, alpha
INSR	7	Insulin receptor
ISG15	16	Interferon-stimulated protein, 15 kDa
ITGA2	20	Integrin, alpha 2 [CD49B, alpha 2 subunit of VLA-2 receptor]
ITGA2B	19	Integrin, alpha 2b (human homolog) [platelet glycoprotein IIb of Iib/IIIa complex, antigen CD41B (human homolog)]
ITGA5	5	Integrin, alpha 5 [fibronectin receptor, alpha polypeptide]
ITGAD	25	Integrin, alpha D (human homolog)
ITGAL	25	Integrin, alpha L [antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide]
ITGB1	13	Integrin, beta 1 subunit [fibronectin receptor] [fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12]
ITGB2	1	Integrin, beta 2 [antigen CD18 subunit (p95), lymphocyte function-associated antigen 1] [integrin B2]
ITPA	13	Inosine triphosphatase A
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
ITSN1	1	Intersectin 1 (SH3 domain protein) (human homolog) [intersectin (SH3 domain protein 1A) (human homolog)] [Bovine STS AF213841]
IVL	1q41-q46	Involucrin
JUN	3	v-jun avian sarcoma virus 17 oncogene homolog
JUNB	7	Jun B proto-oncogene
JUND	7	Jun D proto-oncogene
LALBA	5q21	Lactalbumin, alpha
LALBAP	5	Bovine alpha-lactalbumin pseudogene
LDHA	29q22	Lactate dehydrogenase A
LDHB	5	Lactate dehydrogenase B
LDLR	7	Low density lipoprotein receptor
LECT2		Leukocyte cell-derived chemotaxin 2
LEP	4q32	Leptin [obesity] [murine obesity homolog]
LEPR	3q33	Leptin receptor
LFC	3q21	Murine lfc oncogene homolog
LGB	11q28	Lactoglobulin, beta
LHB	18	Luteinizing hormone beta polypeptide

LHCGR	11	Luteinizing hormone/choriogonadotropin receptor
MAP1B	20q13	Microtubule associated protein 1B
MAP2	19	Microtubule associated protein 2C
MAP4	22	Microtubule associated protein 4
MAPK13	23	Gene MPK13
MAPRE1	13	Microtubule-associated protein, RP/EB family, member 1 [human: adenomatous polyposis coli-binding protein EB1]
MAPT	19	Microtubule associated protein tau
MAS1	9	MAS1 oncogene (human homolog) [Bovine STS AF213854]
MB	5	Myoglobin
MFGE8	21	Milk fat globule-EGF factor 8 protein
MTATP6	MT	ATP synthase F0 subunit 6
MTND1	MT	NADH dehydrogenase subunit 1
MTTA	MT	Transfer RNA Alanine
MYF5	5q13	Myogenic factor 5, transcription activator
MYF6	5	Myogenic factor 6, transcription activator [herculin]
MYL4	19	Myosin, light polypeptide 4, alkali, atrial, embryonic
MYO1D	19	KIAA0727 protein
MYOD1	15	Myogenic differentiation 1, transcription activator [myogenic factor 3]
MYOG	16	Myogenin [myogenic factor 4]
NGFB	3q23	Nerve growth factor, beta polypeptide
NPPA	16	Natriuretic peptide precursor A
NPR2	8	Natriuretic peptide receptor B/guanylate cyclase B [atriuretic peptide receptor B]
NPR2L	22	Homologous to yeast nitrogen permease (human homolog)
NPR3	20	Natriuretic peptide receptor C [atriuretic peptide receptor C]
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
NR2F2	21	Transcription factor COUP 2 [chicken ovalbumin upstream promoter 2] [apolipoprotein regulatory protein] [apolipoprotein A1 regulatory protein 1]
NR3C1	7	Glucocorticoid receptor alpha [nuclear receptor subfamily 3, group C, member 1]
NR3C2	17	Mineralocorticoid receptor [aldosterone receptor] [nuclear receptor subfamily 3, group C, member 2]
ODC1	11	Ornithine decarboxylase
OXTR	22	Oxytocin receptor
PDE1B	5	Phosphodiesterase 1B, calmodulin-dependent
PDE2A	15	c-GMP stimulated phosphodiesterase
PDE5A	6	Phosphodiesterase 5A, cGMP-specific
PDE6A	7	Phosphodiesterase, cyclic GMP (rod receptor), alpha polypeptide [cGMP phosphodiesterase alpha subunit]
PDE6B	6	Phosphodiesterase, cyclic GMP (rod receptor), beta polypeptide [phosphodiesterase B]
PDE6C	26	Phosphodiesterase, cyclic GMP (cone receptor)

PDE6G	19	Phosphodiesterase, cyclic GMP (rod receptor), gamma polypeptide
PGRMC2	3	Progesterone receptor membrane component 2 (human homolog) [progesterone membrane binding protein]
PKM2	10	Pyruvate kinase, muscle
POLR2A	19	Polymerase (RNA) II (DNA directed) polypeptide A (220 kD)
POLR2E	7	Polymerase (RNA) II (DNA directed) polypeptide E (25 kD) (human homolog)
POLR3F	13	Polymerase (RNA) III (DNA directed) polypeptide F (39 kD) (human homolog)
PRKACA	7	Protein kinase, cAMP-dependent, catalytic, beta
PRKACB	3	Protein kinase, cAMP-dependent, catalytic, beta
PRKAG1	5q21-q22	AMP-activated protein kinase subunit gamma 1 (non-catalytic subunit)
PRKAR2A	22	Protein kinase, cAMP-dependent, regulatory, type II, alpha (human homolog)
PRKAR2B	4	Protein kinase, cAMP-dependent, regulatory, type II, beta
PRKCA	19	Protein kinase, C alpha
PRKCB1	25	Protein kinase C, beta 1 polypeptide
PRKCG	18	Protein kinase C, gamma
PRKCI	1q34-q36	Protein kinase C, iota
PRKDC		DNA protein kinase catalytic subunit [DNA-dependent protein kinase]
PRL	23q23med	Prolactin
PRLR	20q17	Prolactin receptor
PRP1	23	Prolactin-related protein 1
PRP10	23	Prolactin-related protein 10
PRP3	23	Prolactin-related protein 3
PRP6	23	Prolactin-related protein 6
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
PRP@	23	Prolactin-related protein, gene cluster (PRP1, PRP3, PRP6, PRP10)
PRPS1	9	Phosphoribosyl pyrophosphate synthetase 1
PTH	15q22-q27	Parathyroid hormone
PTK2	14	PTK2 protein tyrosine kinase 2
PTN	4	Pleiotrophin [heparin binding growth factor 8, neurite growth-promoting factor 1]
PTP4A2	2	Protein tyrosine phosphatase type IVA, member 2
PTPN13	6	Protein tyrosine phosphatase, non-receptor type 13 [APO-1/CD95 (Fas)-associated phosphatase]
RAB18	9	RAB18 small GTPase
RAB2	14	RAB 2, member RAS oncogene family
RAB2L	23	RAB 2, member RAS oncogene family-like (human homolog)
RAB31	24	RAB 31, member RAS oncogene family (human homolog)
RAB3A	7	RAB 3A, member RAS oncogene family
RAB7	22	RAB 7, member RAS oncogene family
RAC1	25	Ras-related C3 botulinum toxin substrate 1 [rho family,

		small GTP binding protein Rac1]
RAD54L	3	RAD54-like (human homolog)
RAF1	22	Murine leukemia viral (v-raf-1) oncogene homolog 1
RASA1	7q25	RAS p21 protein activator [GTPase activating protein RAS p21] [GAP RAS activator]
RASGRP1	10	RAS guanyl releasing protein 1 (calcium and DAG-regulated)
RASSF2	13q21	Rasfadin [Ras association (Ra1GDS/AF-6) domain family 2]
REN	16	Renin
RENT1	7	Regulator of nonsense transcripts 1 (human homolog)
RGR	28	Retinal G protein coupled receptor
RGS16	16	Regulator of G-protein signalling 16
RGS20	14	Regulator of G-protein signalling 20 (human homolog) [small membrane protein 1] [ cathepsin E]
RGS7	16	Regulator of G-protein signalling 7
RPE65	3	Retinal pigment epithelium-specific protein (65kD)
RPL18A	7	Ribosomal protein L18a
RPL19	19	Ribosomal protein L19
RPL28	18	Ribosomal protein L28
RPL3	5	Ribosomal protein L3
RPL4	10	Ribosomal protein L4
RPL7A	11	Ribosomal protein L7a
RPLP0	17	Ribosomal protein, large, P0 (human homolog) [Bovine STS AF275603]
RPN2	13	Ribophorin II
RPS13	15	Ribophorin protein S13
RPS14	7	Ribophorin protein S14
RPS16	18	Ribophorin protein S16
RPS17	21	Ribophorin protein S17
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
RPS24	28	Ribophorin protein S24
RPS27A	11	Ribophorin protein S27a (human homolog)
RPS3A	17	Ribophorin protein S3a
RPS4	X	Ribophorin protein S4
RPS8	3	Ribophorin protein S8
SFRS12	20	Splicing factor, arginine/serine-rich 12 (human homolog)
SFRS3	23	Splicing factor, arginine/serine-rich 3
SLC25A4	27	Adenine nucleotide translocator 1 (skeletal muscle) [solute carrier family 25 member 4]
SLC25A5	X	Adenine nucleotide translocator 2 (fibroblast) [solute carrier family 25 member 5]
SLC25A6	X	Adenine nucleotide translocator 3 (liver) [solute carrier family 25 member 6]
SLC2A3RS	18	Glucose transporter type 3 [solute carrier family 2 member 3] related sequence
SLC2A4	19	Solute carrier family 2 [facilitated glucose transporter], member 4

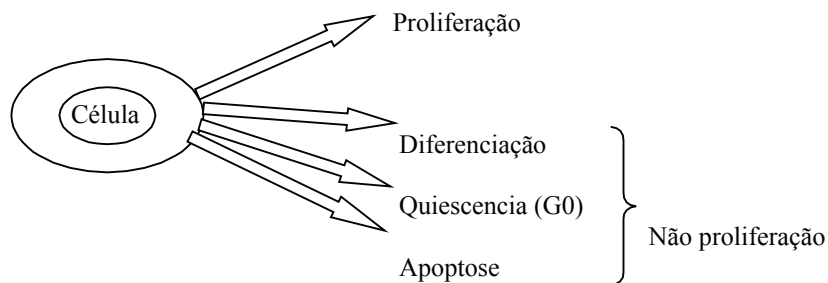
SLC30A4	10	Zinc transporter 4 [solute carrier family 30 (zinc transporter), member 4]
SLC34A1	7	Solute carrier family 34 [sodium phosphate], member 1
SLC5A3	1	Osmoregulatory Na <sup>+</sup> /myo-inositol cotransporter [solute carrier family 5 (inositol transporters), member 3] [sodium-myoinositol-cotransporter]
SOD1	1q14	Superoxide dismutase 1, soluble
SOD1L1	13q21-q22	Superoxide dismutase 1, soluble-like
SOD2	9	Superoxide dismutase 2, mitochondrial
SOD3	6	Superoxide dismutase 3, extracellular
SOX2	1q33	SRY-related HMG-box 2 [sex determining region Y-box 2] [SRY-box 2]
SP1	5	Sp1 transcription factor
SRY	Y	Sex determining region Y
SSBP1	2	Single-stranded DNA-binding protein [single-strand binding protein]
SSR1	23	Signal sequence receptor, alpha (human homolog) [translocon-associated protein alpha (human homolog)]
SSRP1	15	Structure-specific recognition protein 1
SST	1q23-q25	Somatostatin
ST13		Suppression of tumorigenicity 13 – colon carcinoma [Hsp70-interacting protein]
STAG3	25	Stromal antigen 3 (human homolog)
STAR	27q18-q19	Steroidogenic acute regulatory protein
STAT1	2	Signal transducer and activator of transcription 1, 91kD
STAT3	19	Signal transducer and activator of transcription 3 (acute-phase response factor)
STAT4	2q34	Signal transducer and activator of transcription 4
STAT5A	19q17	Signal transducer and activator of transcription 5A
STAT5B	19	Signal transducer and activator of transcription 5B
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
TBP	9	TATA boxbinding protein (human homolog) [Bovine STS AF213838]
TCAP	19	Titin-cap (human homolog) [telethonin (human homolog)]
TCP1	23	T-complex 1
TCTEL1	9	T-complex-associated-testis-expressed 1-like 1
TEGT	5	Testis enhanced gene transcript [BAX inhibitor 1]
TEK	8	TEK tyrosine kinase, endothelial [human: venous malformations, multiple cutaneous and mucosal ]
TF	1q41-q46	Transferrin
TFAP2B	23	Transcription factor AP-2 beta [activating enhancer-binding protein 2 beta]
TFF2	1	Spasmolytic polypeptide
TFPI	2q14-q15	Tissue pathway factor inhibitor [lipoprotein-associated coagulation inhibitor]
TG	14q13dist	Thyroglobulin [hereditary goitre]
TG737	12	Transgene insert site 737, insertional mutation, polycystic kidney disease (human homolog) [bovine EST U83031]

TGFB1	18q24-q25	Transforming growth factor, beta 1
TIE	3	Tyrosine kinase with immunoglobulin and epidermal growth factor homology domains
TNF	23q14dq15p	Tumor necrosis factor, alpha [cachectin] [TNF superfamily, member 2]
TNFRSF12	16	Tumor necrosis factor, receptor superfamily, member 12 [translocating chain-association membrane protein]
TNFRSF1A	5	Tumor necrosis factor, receptor superfamily, member 1A
TNFRSF6	26q13	Apoptosis (APO-1) antigen 1 (FAS) [tumor necrosis factor receptor superfamily, member 6] [Fas/APO-1 oncogene]
TNFSF5	X	Tumor necrosis factor, (ligand) superfamily, member 5 [human: hyper-IgM syndrome]
U3-55K	22	U3 snoRNP-associated 55-kDa protein (human homolog)
U5-116KD	19	U5 snRNP-specific protein, 116 kD (human homolog)
U83006	28	Bovine EST U83006
UBE1C	22	Ubiquitin-activating enzyme E1C (homologous to yeast UBA3)
UBE2D3	6	Ubiquitin-conjugating enzyme E2D 3 (homologous to yeast UBC4/5)
UBE2H	4	Ubiquitin-conjugating enzyme E2H 3 (homologous to yeast UBC8)
UBE2L3	17	Ubiquitin-conjugating enzyme E2L 3 (human homolog) [Bovine STS AF275591]
UBE2N	5	Ubiquitin-conjugating enzyme E2N (homologous to yeast UBC13)
UBL3	12	Ubiquitin-like 3
ZFHX1B	2	Zinc finger homeobox 1b (human homolog) [Bovine STS AF213836]
<b>Símbolo</b>	<b>Localização</b>	<b>Gene</b>
ZFP36	18	Zinc finger protein homologous to Zfp-36 in mouse
ZFX	Xq34	Zinc finger protein X-linked
ZFY	Yp12dist	Zinc finger protein Y-linked
ZNF144	19	Zinc finger protein 144 (Mel-18) (human homolog)
ZNF146	18q24	Zinc finger protein 146
ZNF164	17q24	Zinc finger protein 164
ZNF189	8	Zinc finger protein 189
ZNF192	23	Zinc finger protein 192
ZNF236	24	Zinc finger protein 236 (human homolog)
ZNF302	18	Zinc finger protein 302 (human homolog)
ZNF325	25	Zinc finger protein 325 (human homolog) [gonadotrophin inducible transcription repressor-3 (human homolog)]
ZNF33A	28	Zinc finger protein 33a (KOK31)

## 1.6 – Replicação celular e cromossomal

(Diffley & Labib, 2002).

O ciclo celular origina a duplicação e transmissão da informação genética, sendo o DNA replicado e duas cópias idênticas de cada cromossoma distribuídos pelas duas células filhas. O ciclo celular divide-se nas quatro fases, G1, S, G2, M. A fase M designa-se por mitose, enquanto as fases G1, S e G2 fazem parte da interfase. Durante a fase G1 a célula é estimulada por mitogénios extracelulares e por factores de crescimento. Na fase S ocorre a síntese e duplicação do DNA celular. A fase G2 é o intervalo entre a biossíntese completa do DNA (S) e a fase da mitose (M). Nesta fase M ocorre a produção dos feixes ou fusos mitóticos bipolares, a segregação dos cromátídeos irmãos e a divisão celular. A a regulação do ciclo celular deve assegurar-se de que cada fase está completa antes de se passar à fase seguinte. Existem pois pontos de controlo (*check points*) para averiguar da integridade do DNA, um na parte final da fase G1, e outro na interface entre G2/M. Com isto procura-se evitar que células mutantes ou alteradas progridam e se propaguem. A fase G0 corresponde a células quiescentes que se encontram temporariamente ou permanentemente fora do ciclo celular.



A diferenciação é um processo que envolve em geral mudanças na expressão dos genes e através das quais uma célula precursora se transforma num tipo de célula especializada distinta. Afirma-se desde já que a proliferação e diferenciação das células dos músculos do esqueleto são processos que se excluem mutuamente. Os efeitos epigenéticos passam de uma célula específica para a sua descendência sem qualquer alteração. Na epistasia um gene suprime a acção de outro gene.

As células normais para saírem daquela fase G0 e entrarem na G1 precisam de estímulos externos (mitogénicos ou factores de crescimento). Estes estímulos externos desencadeiam cascatas de transdução de sinal que levam a fosfatações intracelulares reguladas pela expressão de ciclinas e cínases dependentes destas ciclinas (CDKs). As ciclinas e as CDKs funcionam durante fases distintas do ciclo celular. Os complexos ciclinas/CDK fosfatam substratos proteicos específicos com activação da síntese do DNA (na fase tardia da G1 e na S) e a formação de componentes estruturais ligados á mitose (na fase tardia do G2 e na M). A célula na fase G0 através de estímulos externos avança para a fase G1, mas na ausência destes estímulos pode sofrer diferenciação, apoptose ou entrar num estado quiescente (G0) como referimos anteriormente.

Na fase G1 aumenta a expressão de ciclinas D ( D1, D2, D3) que se associam com CDK4 e CDK6 originando estes complexos a fosfatação e activação das CDKs. Estes complexos fosfatados vão, por sua vez, fosfatar a proteína retinoblastoma (RB) que é crítica para a regulação de progressão da G1. Esta família de proteínas RB sequestram as proteínas de transcrição E2F. RB não fosfatado ou hipofosfatado liga firmemente E2F e inibe a transcrição. Após a fosfatação da RB pelas CDK4/6, o RB dissocia-se da

E2F permitindo que esta interactue com o DNA e transcreva uma série de genes que respondem inclusivé pela ciclina E. Este RB é um guardião do ciclo celular pois se estiver hipofosfatado interactiva com o ponto de restrição e impede a progressão do ciclo celular; mas se estiver o RB hiperfosfatado liga-se com o E2F e o ponto de restrição é ultrapassado continuando-se o ciclo celular, e continuando no resto do ciclo na forma hiperfosfatada até a mitose estar completada. Na fase tardia da G1 há um aumento da expressão da ciclina E. O complexo ciclina E/CDK2 é necessário para a passagem da G1 para S. Nesta transição ocorre também um aumento das expressões da ciclina A que persiste através da fase S. Com a ligação da ciclina A a CDK2 a síntese da DNA processa-se. Na parte final da fase S, a ciclina A associa-se com CDK1. Aumentadas concentrações de ciclinas A, B complexadas com CDK1 favorecem a mitose. A integridade do genoma da célula é controlada pelo factor de transcrição p53. Alterações genómicas levam o p53 a interromper o ciclo celular (dando tempo para a reparação do DNA) o que resulta da inibição da fosfatação do RB.

### Replicação dos cromossomas (Diffley & Labib., 2002)

Nas células eucariotas a síntese do DNA nos diversos cromossomas encontra-se coordenada, começando a cópia desse DNA no momento próprio do ciclo da divisão celular, em várias centenas de diversas origens, algumas das quais são desencadeadas no princípio da fase S, enquanto outras são desencadeadas mais tarde. O iniciador da replicação, ou seja, o complexo ORC actua em conjunto (para controlar esta iniciação da replicação do DNA) com outras proteínas adicionais que se ligam a estas origens de iniciação e permanecerem aqui ligadas até a fase S começar. Estas outras proteínas são as cínases dependentes de ciclinas (CDKs) que operam associadas com ciclinas. Estas CDKs e ciclinas controlam o ciclo celular “empurrando” as células para a fase S e iniciação da síntese de DNA, evitando que nova síntese de DNA seja iniciada numa segunda etapa, de forma que só ocorra uma fase S em cada ciclo. A degradação das CDKs remove o sinal que inibe a divisão celular e o ciclo celular desloca-se então para a mitose. Ocorre pois a formação de um complexo de pré-replicação com adição ou remoção da ORC de ciclinas e CDKs de uma maneira cíclica. Nas células eucariotas a replicação do DNA inicia-se em múltiplas origens da replicação existentes ao longo de cada cromossoma. Estas origens de replicação são curtas e as respectivas sequências desse DNA são bem definidas. Na maioria dos tipos celulares dos eucariotas, estas origens para replicação são continuamente activadas através das fases S do ciclo celular, através de um programa regulado no tempo.

Para efeitos expositivos consideraremos seguidamente as onze etapas no ciclo da replicação dos cromossomas, as primeiras cinco na fase G1 do ciclo celular, as 6, 7, 8 e 9 na fase S, a 10 na fase G2 e a 11 na fase M.

#### ETAPA 1

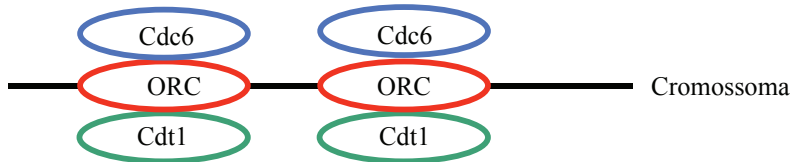


Nesta etapa ocorre a interacção com as origens de replicação situadas ao longo do DNA de cada cromossoma de um complexo de reconhecimento (ORC) com seis sub-unidades. Para esta ligação dos ORC ao DNA é necessária ATP (adenosina trifosfatada) que sofre uma reacção de hidrólise. Nos



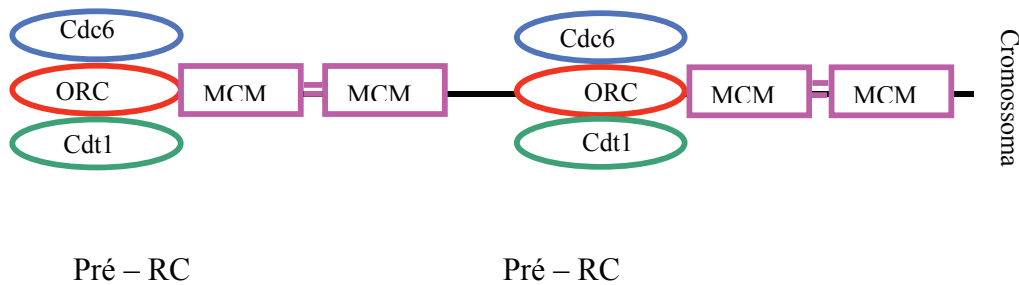
eucariotas pluricelulares algumas das sub-unidades ORC continuam ligadas ao DNA através da fase G1, da S e da fase G2, não se sabendo bem como se comportam durante a mitose.

### ETAPA 2



Nesta etapa duas outras proteínas a Cdc6 e Cdt1 são recrutadas separadamente para as ORC, já no fim da mitose. A proteína Cdc6 (602627 Omim) é um regulador crítico da replicação do DNA, expresso selectivamente nas células em replicação e não nas células quiescentes de mamíferos. Esta proteína Cdc6 é regulada pelas proteínas E2F que são factores de transcrição.

### ETAPA 3



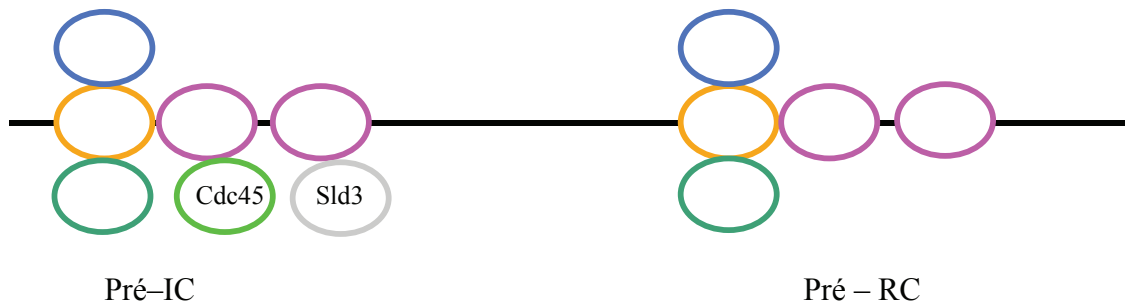
Nesta etapa, seis proteínas relacionadas, constituindo o complexo MCM 2 – 7 (proteínas para iniciação da replicação do DNA – 600592 Omim) são recrutadas pelo complexo formado na etapa 2 constituindo o complexo pré-replicativo (pré – RC). Este complexo MCM 2 – 7 é necessário através da fase S para que ocorra iniciação e alongamento do DNA. Parte deste complexo, ou seja o sub-complexo 4, 6 e 7, tem actividade de helicase do DNA parecendo desenrolar o DNA nas forquilhas de replicação. Sabe-se que as origens de replicação próximas dos telómeros tendem a ser activadas tardiamente na fase S e que as proteínas ORC que actuam mais cedo se deslocam depois para regiões sub-teloméricas onde actuam mais tarde.

### ETAPA 4

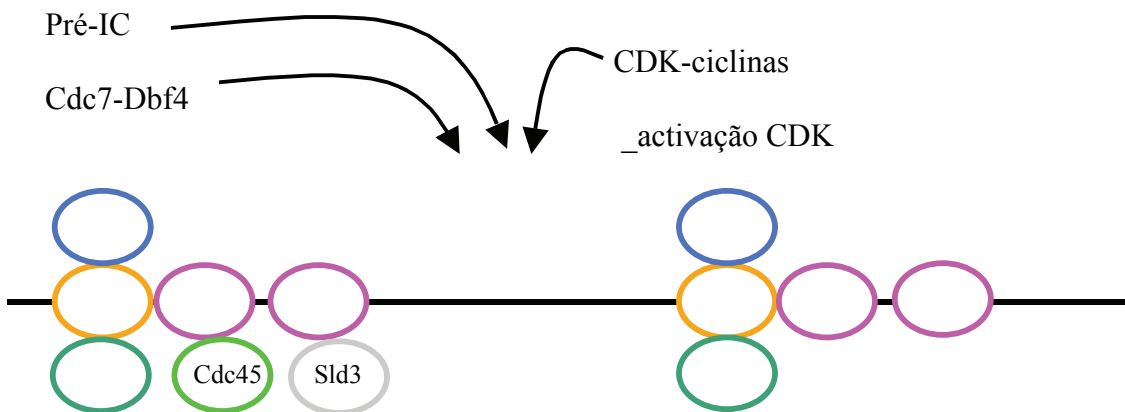
Diversas células eucariotas a partir da fase G1 entram em quiescência, parecendo que tais células perderam os seus pré – RC, ou pelo menos não os assemblaram previamente.



### ETAPA 5



Ocorre a maturação do complexo pré-RC em complexo de pré-iniciação (pré-IC) pela interacção com factores adicionais inclusive o Cdc45 e Sld3 que são importantes não só para a iniciação mas também para o alongamento do DNA.



### ETAPA 6

Esta etapa marca a transição da fase G1 (até aqui) para a fase S com a iniciação da replicação do DNA. São essenciais nesse acontecimento duas proteínas cínases, uma delas uma cínase dependente da ciclina (CDK). Nos organismos eucariotas multicelulares o ciclo celular é controlado por várias CDKs. A outra proteína cínase é a Cdc7. Os pré-Is apenas são assemblados depois da activação dos CDK e Cdc7. Após a acção destas duas cínases, as origens de replicação do DNA são desenroladas sendo agora recrutada uma proteína heterotrimerica, a RPA, que liga-se a uma das hélices desenrolada do DNA. A RPA (179835 Omim) é pois uma proteína de replicação. As CDKs além de desencadear a iniciação da replicação regulam ainda a replicação do DNA pois inibem o *assembly* de novos pré-RCs, o que evita a re-iniciação a partir das origens de replicações.

### ETAPA 7

Após o desenrolamento, é recrutada a primeira DNA polimerase, mas esta é incapaz de iniciar a síntese de novo de uma cadeia de DNA, para o que é necessário uma “primase” que sintetiza um curto segmento de ribonucleótidos (6 a 10) e estes vão actuar como *primer* na síntese do DNA. Esta primase é constituída por quatro sub-unidades que integram este complexo DNA polimerase -primase. As helicases desenrolam o dúplice de DNA, criando duas hélices simples de DNA antiparalelas. As duas cadeias do DNA inicial são antiparalelas, sendo cada uma delas replicada por um mecanismo diferente

durante a progressão da forquilha de replicação. Uma das hélices iniciais do DNA, a que é 3'–5' em relação à direcção do desenrolamento, é replicada continuamente por uma DNA polimerase sintetizando de 5' para 3', progredindo na mesma direcção que a helicase, e produzindo a chamada cadeia continua ou condutora. A outra hélice inicial do DNA, como as DNA polimerases não podem sintetizar DNA na direcção de 3' para 5', é replicada numa hélice descontínua ou vagarosa começando com a interacção de uma DNA polimerase  $\alpha$ -primase que sintetiza um *primer* híbrido RNA-DNA que origina os fragmentos de Okazaki (*vide* a replicação do DNA nos textos convencionais de Bioquímica).

## ETAPA 8

A replicação processa-se portanto em duas direcções a partir das origens de replicação nos cromossomas, ou melhor, as forquilhas de replicação movem-se em direcções opostas a partir das origens.

A síntese de DNA na hélice contínua requer DNA polimerase que sintetiza longos trechos de DNA de uma forma progressiva, sem interrupções, enquanto a DNA polimerase  $\alpha$ -primase tem uma baixa processividade, admitindo-se que a maioria dos fragmentos de Okasaki do DNA sejam sintetizados por outras polimerases além da DNA polimerase  $\alpha$ -primase.

## ETAPA 9

A hélice descontínua sintetizada como híbrido que é de moléculas de RNA-DNA necessita que estes fragmentos de Okasaki removam o RNA *primer*, o que é feito por endonucleases Fen1 e helicases/endonucleases DNA2 seguido pela ligação dos vários segmentos descontínuos pela ligase I. Acoplados com a replicação do DNA, nos cromossomas em replicação encontram-se outras situações a considerar tais como:

- A coesão dos dois cromossomas obtidos.
- A assemblagem da cromatina.
- A possível reparação após a replicação.
- O alongamento dos telómeros e a possível homeostasia.
- “Imprinting”/hereditariedade e epigenética

As duas moléculas filhas de DNA obtidas depois da replicação permanecem firmemente ligadas uma à outra naquilo que se chama a coesão dos cromátídeos irmãos. Esta coesão dos cromátídeos necessita de uma série de factores tais como Sco 1-4, Eco-1, Smc1,3, DNA polimerase- $\alpha$ , Ctf8,18, Dcc1, Pds5 e Rfc 2-5 e podem ocorrer apenas durante a replicação do DNA. Também a assemblagem da cromatina associada com a replicação do DNA precisa do factor CAF1. Este factor (601245 Omim) é uma proteína constituída por três sub-unidades que assemblam o octamero de histonas sobre o DNA. O alongamento e manutenção dos telómeros acoplados com a replicação do DNA será considerado mais adiante em capítulo próprio. A terminação ocorre quando duas forquilhas de replicação opostas se encontram, e o DNA que resulta das duas forquilhas está todo ligado.

## ETAPA 10

Os dois cromátídeos irmãos resultantes são mantidos unidos pela coesão. Esta coesão não parece desenvolver-se ao longo de todos os cromátídeos, parecendo ocorrer antes em numerosos e específicos locais de coesão, inclusivé nos centómeros.

## ETAPA 11

Por fim, após a assemblagem do feixe ou fuso mitótico bipolar na metafase, ocorre uma proteólise controlada das coesinas (que são complexos de sub-unidades proteicas específicas que mantêm a coesão dos cromátídeos irmãos na mitose e na meiose) o que permite que os dois irmãos sejam deslocados para pólos celulares opostos. Após a segregação dos cromossomas estar completa as CDKs são inactivadas numa reacção complexa acoplada com a citocinese pelo desaparecimento do redanho mitótico (*mitotic exit network* ou MEN) e pela trama de iniciação da septação (*septation initiation network* ou SIN).

### **1.7 – Telómeros e telomerasas.**

(Lodish et ali.,1995; Bonney,2000).

“A divisão celular é um processo limitado, podendo as células normais dividir-se 60 a 100 vezes, tudo dependendo do tipo de tecido, após o que elas envelhecem e possivelmente morrem. Um mecanismo que leva á morte das células normais é a erosão dos telómeros. Em cada divisão celular, os telómeros encurtam-se e eventualmente desaparecem originando a morte das células. Cada replicação celular, desencadeia a perda de cerca de ~100bp nas extremidades dos cromossomas eucariotas. As extremidades dos cromossomas, os telómeros, contêm repetições TTAGGG. O comprimento destes telómeros reflecte o número de replicações de uma dada célula e a sua capacidade replicativa. Contudo, as células cancerosas ajudadas pelas telomerasas podem reparar os telómeros desgastados tornando-se a célula imortal. As extremidades 5’ e 3’ expostas no DNA dos cromossomas lineares são muito vulneráveis à acção degradativa das nucleases. Os telómeros defendem-se desta situação criando complexos de DNA – proteínas compostos por grandes extensões de sequências nucleotídicas repetidas (por ex: 5’-TTAGGG-3’) juntamente com proteínas que se ligam a sequências específicas do DNA, estabilizando e protegendo assim as extremidades dos cromossomas. Nas proteínas que se ligam a sequências específicas do DNA para este efeito, as TRF1 e TRF2 que reconhecem especificamente as sequências repetidas TTAGGG, desempenham um papel essencial, a primeira (TRF1) regulando o comprimento do telómero ao inibir o alongamento do telómero quando este atinge um tamanho crítico, e a segunda (TRF2) suprimindo a fusão, terminação com terminação entre os cromossomas, e estabilizando as terminações cromossomais” (Bonney,2000). A telomerase encontra-se nas células germinativas (*germcells*) e nas primeiras células embrionárias, mas é normalmente desactivada nos tecidos somáticos adultos normais. Uma célula normal é incapaz de replicar as últimas de algumas bases dos telómeros de uma das hélices de DNA, o que leva a que os telómeros se encurtem em cada divisão celular que ocorra. Depois de várias divisões celulares os telómeros são eventualmente perdidos e daí o cromossoma tornar-se instável e as células morrerem.(Bonney,2000)

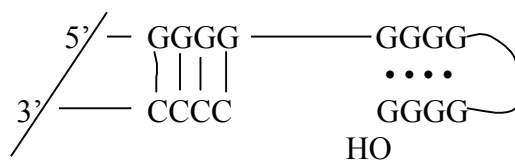
As células normais tem capacidade para proliferar podendo tornar-se malignas se mutações ao acaso ou herdadas activam oncogenes ou desactivam genes supressores dos tumores (Bonney,2000). Cerca de 90% das células cancerosas humanas tem capacidade para reactivar as suas telomerasas e reparar assim os seus telómeros o que facilita a replicação imortal destas células. A actividade telomerasica é observável em quase todos os tumores avançados (mais de 30 tipos de tumores estudados) sendo assim a telomerase um alvo de eleição para a terapia cancerígena, enquanto as células somáticas

normais que perderam a sua actividade telomerásica não são afectadas por esta possível terapia (Bonney,2000).

Os telómeros são pois as regiões terminais de cada cromossoma eucariota que contêm sequências teloméricas características (TEL) e que são replicadas por um processo especial através da telomerase, o que contraria a tendência de cada cromossoma para ser encurtado durante cada ciclo de replicação do DNA. O DNA telomero tem uma sequência característica que lhe é adicionada (às extremidades das moléculas de DNA) pela telomerase, ou seja, uma enzima telomero-terminal-transferase.

As sequências dos telómeros são oligómeros repetidos com um elevado conteúdo em G na hélice que corre no sentido 5' → 3' para o telomero (este oligómero em algumas espécies é G<sub>4</sub>T<sub>2</sub>, noutras G<sub>4</sub>T<sub>4</sub> ou G<sub>1</sub>-3T e AGGGTT nos humanos e noutras vertebrados). Estas sequências simples são repetidas nas terminações dos cromossomas, num total de algumas milhares de pares de bases nos vertebrados. (Lodish et al., 1995)

A extremidade 3' da hélice rica em G estende-se 12-16 nucleótidos para além da extremidade 5' da hélice complementar rica em C (ver figura abaixo). Contudo as terminações dos telómeros não são um simples dúplex seguido por uma região de hélice simples. Com efeito a porção de hélice simples dobra-se para trás sobre si mesma formando pares de bases G-G especiais (não de Watson Crick) como se esquematiza na figura seguinte. (Adaptado de Lodish et al., 1995, fig.9.61).



A extremidade 3' da região da hélice rica em G dobra-se para trás sobre si mesma e é estabilizada por invulgares emparelhamentos de bases G-G na conformação Syn mais do que na conformação anti.

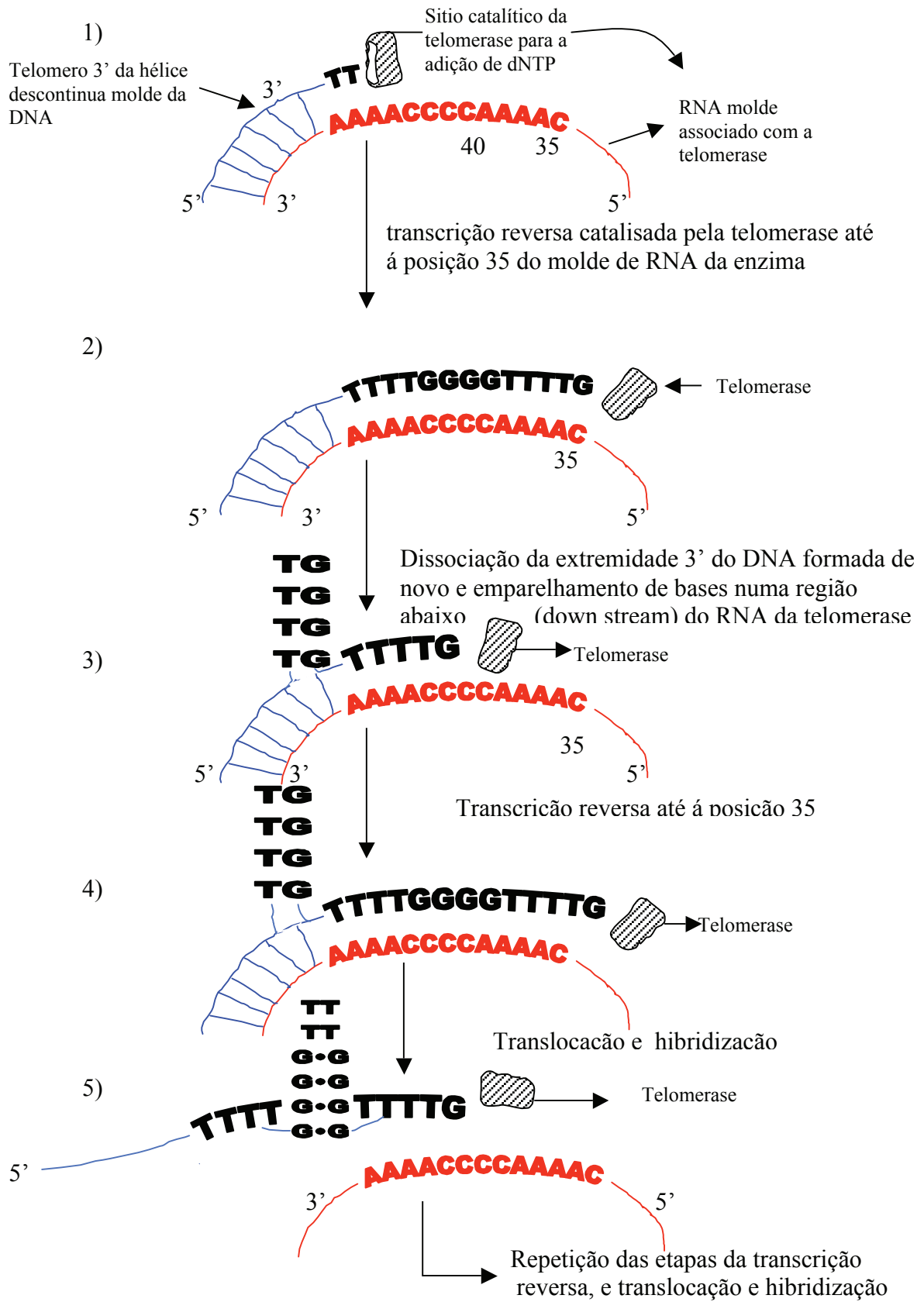
A fonte da telomerase, e não a sequência do DNA do telomero, é que determina a sequência dos nucleótidos de DNA a adicionar ao *primer*. Entende-se por *primer* uma curta sequência de ácido nucleico contendo um grupo hidroxilo 3' livre e que funciona como ponto de partida para a adição de nucleótidos para copiar a cadeia emparelhada com a cadeia singular que serve de molde. O sítio catalítico da telomerase é composto por uma subunidade de RNA (o molde para a adição ao telomero) conhecida com hTR (*human telomerase RNA*) e uma subunidade proteica que catalisa a síntese do telomero a HTER ( *human telomerase reverse transcriptase*). Todas as DNA polimerases conhecidas biossintetizam as cadeias de DNA no sentido 5' → 3' necessitando de um *primer*.

Quando a forquilha de replicação se aproxima do fim de um cromossoma linear, a síntese da hélice de DNA continua sem problemas até à extremidade da hélice de DNA molde, sendo o cromatídeo filho resultante libertado. Mas como a hélice descontínua ou vagarosa é biossintetizada, como a sua denominação significa (copiada de uma forma descontínua), ela não pode ser replicada na totalidade, havendo um mecanismo especial que impede que o cromatídeo filho resultante da biossíntese da hélice vagarosa não seja encurtado em cada divisão celular que ocorre. A enzima implicada neste mecanismo especial é uma transcriptase-reversa modificada chamada telomerase que tem capacidade para alongar a hélice vagarosa molde a partir da sua extremidade hidroxílica 3' (parte-se do princípio que a replicação do DNA e os seus mecanismos, bem como a sua eventual reparação, já são conhecidas do leitor). Com efeito os cromossomas lineares não se podem replicar completamente sem a ajuda de etapas adicionais que operem a replicação das suas extremidades ou telómeros.

A síntese da extremidade do DNA da hélice descontínua não fica completa. Este vazio (*gap*) resulta da remoção do *primer* que foi utilizado para iniciar a replicação, e surge na hélice descontínua. O tamanho exacto deste *gap* depende da localização do último fragmento de Okasaki sintetizado. Com efeito, a DNA-polimerase não pode preencher este espaço pois precisa de um *primer*. Os produtos da replicação do DNA ficam assim mais curtos em relação ao DNA parental, originando uma perda gradual de DNA nas extremidades dos cromossomas dos animais, tornando-se necessário um mecanismo distinto para a replicação destas extremidades das moléculas de DNA. Na maioria dos eucariotas, a replicação dos telómeros utiliza para este efeito uma transcriptase reversa, chamada telomerase. Esta telomerase contém um local catalítico que polimeriza desoxiribonucleótidos trifosfatados (dNTP) sobre um molde de RNA. A sequência repetida adicionada pela telomerase é determinada pelo RNA associado com a enzima.

Na figura seguinte (adaptada de Lodish et ali.,1995, fig.10.17) esboça-se um modelo da acção da telomerase embora o mecanismo preciso implicado na síntese do telómero 3' da hélice descontínua seja desconhecido.

Certas proteínas ligam-se firmemente a estas estruturas na extremidade dos telómeros talvez para as protegerem da acção de exonucleases.



A telomerase é um complexo ribonucleoproteico que funciona como uma transcriptase-reversa modificada. A enzima na figura anterior junta unidades T4G4 repetidas à extremidade 3' da hélice descontínua do DNA. Outras telomerasas adicionam sequências ligeiramente diferentes.

Na figura anterior em 1) a extremidade 3' da hélice descontínua (a azul) emparelha com uma região única do RNA (a vermelho) associada com a enzima telomerase. A hibridação é facilitada pela sequência no fim do telomero e pela sequência do lado da extremidade 3' do RNA. Em 2) o sítio catalítico da telomerase junta desoxirribonucleótidos utilizando como molde a molécula de RNA, e esta transcrição reversa decorre até atingir a posição do nucleótido 35 do molde de RNA. Em 3) as duas hélices do híbrido DNA-RNA deslizam seguidamente uma em relação à outra, originando a deslocação da região de hélice única da hélice de DNA telomera, e deixando a descoberto parte da sequência molde do RNA. Em 4) e 5) a sequência telomera da hélice descontínua é depois novamente alongada (biossintetizada) até à posição 35 pela telomerase e as duas hélices DNA – RNA sofrem nova translocação e hibridização como sucedeu anteriormente. A deslocação é facilitada pelo emparelhamento não do tipo Watson Crick entre os pares de bases deslocados G•G na extremidade 3' do telomero do DNA. As telomerasas podem adicionar longas extensões destas repetições pelo processo antes referido.



## 1.8 - Bibliografia e outros textos

- Wallace, M, R., et al. 1991. A de novo Alu insertion results in neurofibromatosis type 1 Nature 353 :864-868.
- Ments, G., and E. N. Moudrianakis, 1993 Topography of the histone octamer surface : repeating structural motifs utilized in the docking of nucleosomal DNA. Proc. Nat'l. Acad. Sci. USA 90:10489-10493.
- Graziano, V., S. E. Gerchman, D. K. Schneider, and V. Ramakrishnan, 1994. Histone H1 is located in the interior of the chromatin 30-nm filament. Nature 368:351-354.
- Hebbes, T. R., A. W. Thorne, and C. Crane-Robinson. 1988. A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J. 7:1395-1402.
- Phi-Van, L., and W. H. Strating. 1988. The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. EMBO J. 7:655-664.
- Schmid, M. B. 1990. More than just "histone-like" proteins. Cell 63:451-453.
- Travers, A. A., and A. Klug. 1987. The bending of DNA in nucleosomes and its wider implications. Phil. Trans. Roy. Soc. (Lond.) B317:537-561.
- Sawyer, J. R., and J. C. Hozier. 1986. High resolution of mouse chromosomes: banding conservation between man and mouse. Science 232:1632-1635.
- Blackburn, E. H. 1991. Structure and function of telomeres. Nature 350:569-773.
- Blackburn, E. H., and J. W. Szostak. 1984. The molecular structure of centromeres and telomeres. Ann. Rev. Biochem. 53:163-194.
- Eaenshaw, W. C., and J. E. Tomkiel. 1992. Centomere and Kinetochore structure. Curr. Opin. Cell Biol. 4:86-93.
- Morin, G. B. 1991. The human telomere terminal transferase enzymes is a ribonucleoprotein that synthetizes TTAGGG repeats. Cell 59:521-529.
- Zakian, V. A. 1989. Structure and function of telomeres. Ann. Rev. Genet. 23:579-604.
- Deininger, P. L. 1989. SINES: Short interspersed repeated DNA elements in higher eukaryotes, in Berg, D. E. and M. M. Howe, eds. Mobile DNA, American Society for Microbiology pp. 593-618.
- Kornberg, J. R., and M. C. Rykowski. 1988. Human genome organization: Alu, LINES, and the molecular structure of metaphase chromosome bands. Cell 53:391-400.
- Miklos, G. L. G., et al. 1988. Microcloning reveals a high frequency of repetitive sequences characteristic of chromosome 4 in the -heterochromatin of Drosophila. Proc. Nat'l. Acad. Sci. USA 85:2051-2055.
- Moyziz, R. K., et al. 1987. Human Chromosome-Specific Repetitive DNA Sequences : Novel Markers for Genetic Analysis. Chromosoma 95:375-386.
- Singer, M. F. 1982. SINES and LINES: highly repeated short and long interspersed sequences in mammalian genomes. Cell 28:433-434.
- Benjamin, H. W., and N. Kleckner. 1989. Intramolecular transposition by Tn 10. Cell 59:373-383.
- Federoff, N. V. 1989. About maize transposable elements and development. Cell 56:181-191.
- Evans, J. P., and R. D. Palmiter. 1991. Retrotransposition of a mouse L1 element. Proc. Nat'l. Acad. Sci. USA 88:8792-8796.
- Fanning, T. G., and M. F. Singer. 1987. LINE-1 mammalian transposable element. Biochem. Biophys. Acta 910:203-212.
- Jurka, J., and T. Smith. 1988. A fundamental division in the Alu family of repeated sequences. Proc. Nat'l. Acad. Sci. USA 85:4775-4778.

- Narita, N., et al. 1993. insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. J. Clin. Invest. 91:1862-1867.
- Israels, E.D. and L.G. Israels. 2001. The cell cycle. Stem Cells 19, 1, 88-91.
- Sherr, C.J. and J.M. Roberts. 1999. CDK inhibitors: positive and negative regulators of G1-phase progression. Genes Dev 13, 1501-1512.

### **Outros textos básicos**

- Diffley, J.F.X. and K. Labb.2002. The chromosome replication cycle. Journal of Cell Science 115, 869-872.
- Bonney, R. C., ( 2000 ) Telomerase : the immortality enzyme and its role in cancer. European Clinical Laboratory. November 6.
- Paranjape,S.M. et ali.,1994.Role of chromatin structure in the regulation of transcription by RNA polymerase II.Ann. Rev. Biochem.63:265-97.
- Campbell,P.N. & Smith,A.D.,1993.Biochemical Illustrated-Third Edition, Churchill Livingstone.
- Bradley,J. et ali.,2001.Molecular Medicine. Sec. Ed.,Blackwell Science.
- Stryer,L.,1995.Biochemistry- Fourth Edition. W.H. Freeman and Company. New York.
- Lodish,H. et ali.,1995-Molecular Cell Biology. Third Edition. W. H: Freeman and Company. New York.

<b>2 - Sequências nucleotídicas, estruturas do DNA e seu funcionamento</b>	<b>67</b>
2.1 - Endonucleases de restrição e palindromas.	67
2.2 - Sequências nucleotídicas	69
2.2.1 - Sequências nucleotídicas solitárias ou de cópia única	69
2.2.2 - Sequências repetidas	70
Sequências moderadamente reiteradas	70
Sequências altamente repetidas	72
Sequências repetidas invertidas	72
2.2.3 – DNA mitocondrial (mtDNA)	73
2.3 - Tipos de estruturas do DNA e suas conformações alternativas	74
Conformações do DNA linear e circular e do DNA	74
Dobragens angulares ou curvaturas do DNA	75
DNA cruciforme	75
DNA de hélices triplas	77
DNA de quatro hélices (quadruplex)	79
DNA “slipped” (deslizante)	80
Relevância da dinâmica estrutural do DNA	82
2.4 - Recombinação do DNA	83
2.5 - Estruturas dos genes, e sua transcrição e translação nas células eucariotas	85
2.6 – Bibliografia	95

## **2 - Sequências nucleotídicas, estruturas do DNA e seu funcionamento**

A composição em bases púricas e pirimídicas do DNA, e mesmo a proporção entre as bases que o constituem, apenas caracterizam o DNA de uma forma parcial. Para caracterizá-lo melhor é muito mais importante conhecer a sequência das bases dos desoxirribonucleótidos que o integram. '

A determinação destas sequências nos DNA de diversas espécies constitui hoje um dos objectivos mais importantes da Biologia Molecular.

### **2.1 – Endonucleases de restrição e palindromas**

As endonucleases de restrição são enzimas provenientes de bactérias que tem a propriedade de serem altamente específicas para certas sequências palindromas contidas nas moléculas de DNA que são seus substratos. Por sequência palindroma, ou palindromo, entende-se uma série de bases que apresenta uma certa ordem numa das cadeias dum *duplex* de DNA (lida da esquerda para a direita) e exactamente a mesma ordem em sentido contrário (lida da direita para a esquerda) na outra cadeia de DNA. As moléculas de DNA são cindidas por hidrólise por meio destas enzimas dentro duma sequência palindroma, dando origem por cada palindromo a dois fragmentos da molécula de DNA original que têm cada uma deles uma terminação 3'-OH e outra 5'-P livres.

Estas endonucleases de restrição são designadas consoante a sua proveniência, indicando habitualmente as três primeiras letras da respectiva designação a fonte ou espécie bacteriana donde foram extraídas, a letra ou letras seguintes a estirpe respectiva, e o número romano que se lhe segue a ordem temporal da sua caracterização na estirpe de onde provêm.

Estão identificadas e caracterizadas hoje em dia centenas e centenas destas enzimas de restrição (mais de 2000) sendo perfeitamente conhecidas as sequências nucleotídicas sobre as quais elas actuam e que mais habitualmente são sequências palindromas com quatro, seis ou oito pares de bases de extensão. Estas sequências são chamadas palindromas e são caracterizadas por serem constituídas completamente por repetições invertidas e simétricas dos respectivos desoxirribonucleótidos que as integram (esta simetria no DNA interactua com a enzima de restrição que é um homodímero, ou seja, a enzima é composta por duas subunidades proteicas idênticas) sendo a ordem das bases destas sequências a mesma quando as hélices do palindromo são lidas em direcções opostas.

No quadro seguinte dão-se diversos exemplos de locais da cisão do DNA por endonucleases de restrição com diversas especificidades.

Dispõem-se hoje de enzimas de restrição que reconhecem mais de 200 sequências específicas diferentes o que permite pois a produção de diversíssimos fragmentos de DNA.

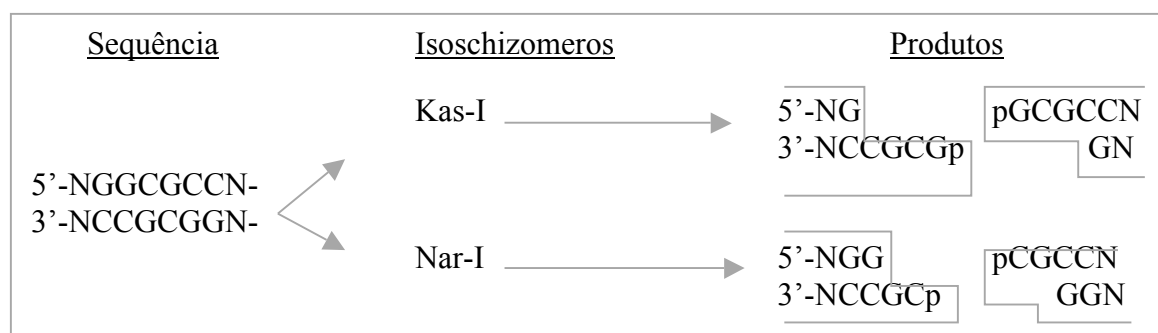
Exemplos de locais de cisão do DNA por enzimas de restrição com diversas especificidades (Adaptado de Devlin, T.M., 1997, Table 14.8).

( as setas _ indicam os locais de cisão ) <u>Enzima</u>	<u>Microorganismo</u>	<u>Sequências específicas</u>	<u>Nº de locais</u>	
			<u>cindidos em</u> Øx174	<u>2</u> pBR322
Eco RI	<i>E. Coli</i>	5'- G _AATT-C3' 3'-C-TTAA _G-5'	25	9
Hae III	<i>Haemophilus aegyptus</i>	5'-GG _CC-3' 3'-CC _GG-5'	11	22
Hpa II	<i>Haemophilus parainfluenza</i>	5'-C _CG-G-3' 3'-G-GC _C-5'	5	26
Hind III	<i>Haemophilus influenza Rd</i>	5'-A _AGCT-T-3' 3'-T-TCGA _A-5'	0	1

Conforme se pode verificar pela leitura do quadro anterior é extraordinariamente importante a orientação na sequência nucleotídica 5' – bases – 3'.

Assim no caso da Eco R1 esta enzima de restrição cinde a sequência 5' – GAATTC – 3' mas já não cinde a mesma sequência com orientação oposta (ou seja 3' – GAATTC – 5').

Estão assinaladas diversas enzimas de restrição que cindem a mesma sequência desoxirribonucleotídica, sendo designadas por isoschizmeros. Contudo, tal como se exemplifica seguidamente, essas enzimas nem sempre cindem o mesmo local de uma dada sequência palindroma.



(Adaptado de Bradley et ali., 2001-Fig.2.11)

Da diversa selectividade das enzimas de restrição resultam vantagens de vária índole, pois uma dada enzima de restrição produz uma única família de fragmentos de DNA a partir de uma certa molécula de DNA, o que constitui o chamado digesto, mapa ou perfil de restrição.

Refira-se ainda que os palindromos no DNA permitem que enzimas de outra natureza, as metilases, possam identificá-los como locais para metilação de duas bases dentro de palidromo. Após a metilação acontecer ela impede que esses palindromos sejam cindidos pelas endonucleases de restrição da própria bactéria, donde resulta a protecção do DNA bacteriano da hidrólise.

## 2.2 – Sequências nucleotídicas

As sequências nucleotídicas ao longo de uma molécula de DNA tendem a ser agrupadas consoante as suas funções conhecidas ou desconhecidas, podendo algumas dessas sequências denominadas genes estruturais codificar a expressão de proteínas específicas, enquanto outras não codificam para a expressão de proteínas, havendo sequências que são apenas transcritas em RNA, enquanto outras parecem apenas servir como espaçadores ou estruturas intervalentes entre outras sequências.

### 2.2.1 - Sequências nucleotídicas solitárias (DNA de cópia única)

As sequências desoxirribonucleotídicas ao longo das moléculas de DNA podem ser únicas (aparecem uma só vez) ou repetidas. Sequências desoxirribonucleotídicas únicas (genes solitários com uma única cópia) ou sejam genes estruturais codificam para proteínas específicas. Sequências repetidas podem ser repetidas em número de vezes muito variável e com características muito diversas.

Muitas proteínas importantes são codificadas por genes solitários ou seja por um só gene por genoma haplóide.

Contudo um só gene é suficiente para produzir milhares ou milhões das moléculas proteicas por ele expressas.

No entanto genes de cópia única podem ser consideravelmente amplificados por pressão selectiva.

O DNA de cópia única corresponde a cerca de metade do genoma humano (uma única sequência desoxirribonucleotídica). No entanto, só uma pequena parte destas sequências é que corresponde a genes estruturais codificadores de proteínas específicas, pois a maior parte dessas sequências não codifica proteínas, constituindo os pseudogenes que são parte do DNA e que tem homologia nucleotídica significativa com um gene funcional mas que contêm mutações que evitam a sua expressão (Devlin, 1997).

Nas células eucariotas apenas uma pequena parte do DNA codifica para genes estruturais, ou sejam, aqueles que exprimem proteínas específicas. No entanto, nas células procariotas a maioria do DNA codifica para proteínas específicas. Os genes estruturais e as sequências desoxirribonucleotídicas que os controlam correspondem pois a uma pequena parte do conteúdo total em DNA nas células eucariotas.

Os genes eucariotas não se sobrepõem havendo entre eles intervalos de cerca de 40Kbp, embora alguns desses genes se encontrem juntos em regiões que são expressas de uma forma altamente coordenada (famílias de genes).

Também em regra os genes eucariotas contêm intrões e exões conhecendo-se no entanto exemplos, como o do gene humano que exprime o interferão  $\gamma$ , que não contêm intrões.

Boa parte do DNA, assim como o que se encontra nos centomeros e telomeros e ainda muito outro DNA não codificador, não revela qualquer função específica.

A cinética de reassociação do DNA dos eucariotas, após segmentação desse DNA em determinadas condições, revela a existência de quatro classes de sequências, a saber:

- A – as sequências únicas que se reassociam lentamente.
- B – as repetições invertidas que se reassociam muito rapidamente.
- C – as sequências altamente repetitivas que se reassociam muito rapidamente também.
- D – as sequências moderadamente reiteradas que se reassociam a velocidades intermédias.

### 2.2.2 – Sequências repetidas

As sequências desoxirribonucleotídicas repetidas no DNA das células procariotas são muito poucas, ao contrário do que sucede nas eucariotas onde podem representar (consoante as espécies) 3 a 80% da totalidade, referindo-se para os mamíferos valores de 25 a 35%, ou mesmo 50%.

Por vezes é feita uma distinção entre sequências reiteradas e repetitivas, aceitando-se que sequência reiterada é uma sequência singular de DNA, geralmente com diversas centenas de nucleótidos de comprimento, e que se repete em múltiplas cópias no genoma. Na sequência repetitiva a sequência desoxirribonucleotídica é habitualmente curta e repetida muitas vezes ao longo do genoma.

As sequências altamente repetitivas têm uma composição em bases diferente da maioria do restante DNA, podendo estas sequências ser isoladas por fragmentação do DNA em segmentos de algumas centenas de desoxirribonucleótidos cada, e a sua separação feita por centrifugação em gradiente de densidades, onde dão origem ao denominado DNA satélite, uma vez que aparecem como satélites da banda principal correspondendo à maioria do DNA.

Refira-se que algumas sequências altamente repetidas não podem ser isoladas por centrifugação, embora possam ser identificadas pela sua rápida reanelagem (*reannealing*).

#### Sequências moderadamente reiteradas

São portanto sequências relativamente compridas (contendo de centenas a vários milhares de desoxirribonucleótidos), idênticas ou muito parecidas, e que ocorrem até milhares de vezes.

Normalmente, e curiosamente, as sequências de cópia única e estas sequências moderadamente reiteradas assumem nos cromossomas um perfil de distribuição ordenado em que ocorrem blocos alternados de DNA de cópia única e DNA moderadamente reiterado.

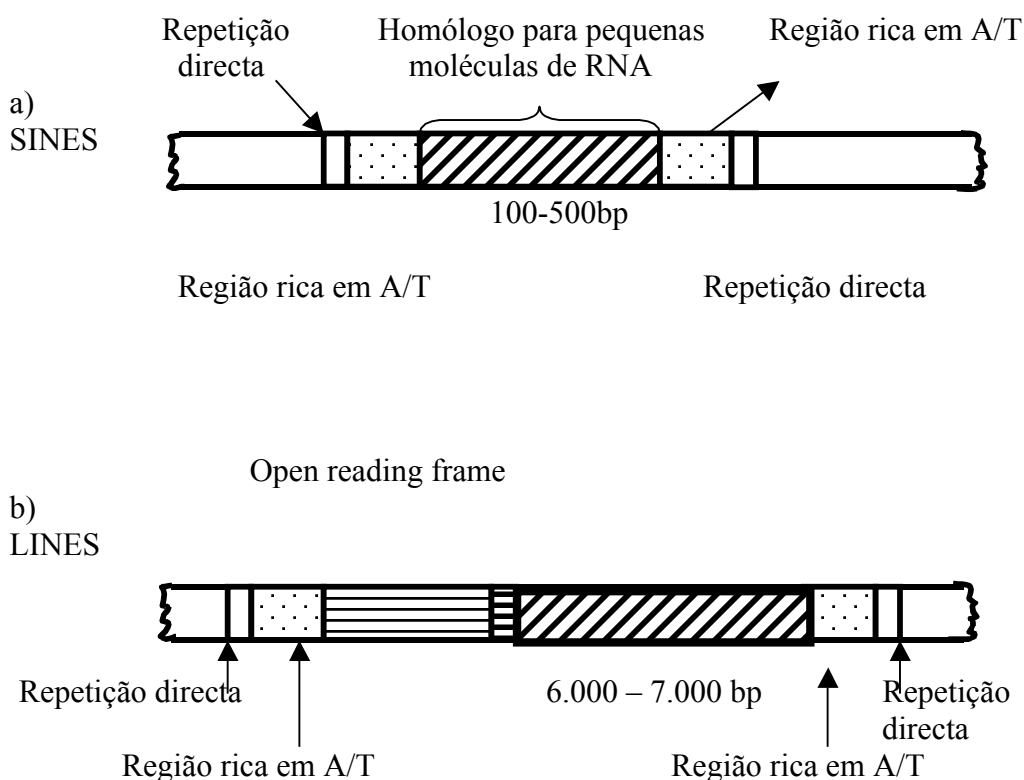
As sequências moderadamente reiteradas podem ser classificadas (ver esquema seguinte) em SINES, da designação “*short interspersed repeats*”, que são famílias relacionadas, embora distintas, contendo sequências com 10-500bp de comprimento, e LINES, da designação “*Long interspersed repeats*”, que são quaisquer sequências de 100bp até vários milhares de desoxirribonucleótidos de comprimento. Tanto das SINES como das LINES encontram-se até 100 000 cópias ou mais por genoma.

#### A família “Alu” no DNA

É a mais abundante família de sequências repetitivas nos mamíferos, presentes ao longo do genoma, contendo nos humanos estas sequências 300bp (*short interspersed repeat*) e aparecendo cerca de 300 000 vezes. Esta sequência é libertada pela acção das enzimas de restrição Alu o que dá o nome a

estas sequências. Estas estruturas Alu são reminiscências de elementos móveis do DNA ( transposões ) de função desconhecida e que representam cerca de 5% do genoma humano (Devlin,1997).

Curtas (SINES) e longas (LINES) repetições disseminadas no DNA das células eucariotas (Adaptado de Devlin,1997-Fig.14.51)



- a) As repetições curtas disseminadas são sequências com cerca de 100 – 500 bp de comprimento que são homólogas para pequenas moléculas de RNA tais como os tRNA, 5S RNA ou 7SL RNA ( partícula reconhedora de sinal ). A versão humana do 7SL RNA é referida como a sequência AU e representa cerca de 10% do DNA humano.
- b) As repetições longas disseminadas são homólogas para genes tRNA e contêm longos trechos de tripletos de codões no DNA que não são interrompidos por codões de paragem da translação “*open reading frames*”, com sequências adicionais codificando proteínas que se assemelham a genes retrovíricos.
- Ambos os tipos de repetições contêm curtas sequências ricas em A/T nas terminações 3’ que são flanqueadas por curtas *direct repeat* DNA ( ou conformações alternativas do DNA ).

O perfil da disseminação das sequências moderadamente reiteradas faz com que estas estejam implicadas no controlo da transcrição de genes estruturais, pois a maioria destes tem adjacentes sequências reiteradas.



As sequências moderadamente reiteradas podem ainda assumir-se formando conjuntos em tandem que são sobretudo utilizados para dar expressão a produtos que são necessários em grande quantidade, como por exemplo o RNA ribossomal e as histonas.

É conhecido que os genes das cinco histonas formam um conjunto em tandem, sendo os intervalos entre os genes de cada histona separados uns dos outros por DNA espaçador com de 400 a 900 desoxirribonucleótidos, com a característica destes espaçadores serem ricos em AT e serem facilmente separáveis como DNA satélite.

Nos genomas das células eucariotas calcula-se que as sequências de cópia única e as sequências moderadamente reiteradas representam cerca de 80% da totalidade do DNA.

### Sequências altamente repetidas (Devlin,1997)

Evocando a afirmação do parágrafo anterior pode referir-se que o restante DNA consiste de sequências nucleotídicas inferiores a vinte desoxirribonucleótidos que se repetem no entanto milhares ou milhões de vezes.

No rato estão assinaladas repetições de um milhão de vezes em cada célula de curtas sequências de 10bp. Estas curtas sequências de desoxirribonucleótidos são também designadas como sequências simples de DNA, não devendo ser confundidas com as sequências de cópia única.

Por vezes ocorre um tipo principal desta sequência simples como sucede no rato, em que o tipo de sequência simples 5' – GCACAC – 3' se repete de seis em seis bases, havendo no entanto outras muito mais compridas.

Estes DNA de sequência simples também podem ser isolados como DNA satélite.

Nos centromeros das células eucariotas superiores o DNA satélite aí residente consiste de milhares de cópias em tandem de uma ou de algumas destas sequências simples, estando assinaladas algumas apenas com 5-10bp de comprimento.

### Sequências repetidas invertidas (Devlin,1997)

Estas sequências repetidas invertidas são um motivo estrutural do DNA (*vide* adiante tipos de estruturas do DNA e sua conformação alternativa).

Estas sequências repetidas invertidas podem ser curtas, com até seis desoxirribonucleótidos, estando identificada nos humanos uma sequência palíndroma deste tipo (GAATTC) de 3000 em 3000 desoxirribonucleótidos.

Estas estruturas com sequências repetidas invertidas e curtas não podem formar estruturas estáveis em gancho de cabelo, o que já é possível com sequências palíndromas mais compridas.

### 2.2.3 - DNA mitocondrial ( mtDNA ) (Devlin,1997).

É uma estrutura pequena, circular, em hélice dupla de cerca de 16500 bp representando nos mamíferos cerca de 1% de todo o DNA celular. A mitocôndria possui múltiplas cópias de DNA habitualmente organizadas dentro de vários “clusters”(Devlin,1997).

A sequência de mtDNA humano contém 37 genes, 13 codificando para proteínas necessárias para a manutenção da síntese do ATP mitocondrial e os restantes 24 codificando RNA específicos para a mitocôndria, 2 para ribossomas e para 22 RNA de transferência.

As mutações são em maior número no mtDNA do que no DNA nuclear. Os genes mitocondriais são herdados da mãe uma vez que as mitocôndrias do espermatozóide não penetram no óvulo a ser fertilizado.

Mutações somáticas no DNA mitocondrial, com deleção de bases ou de segmentos de oligodesoxirribonucleótidos, são originadas por alterações pelo oxigênio, estando associado com alterações da fosforilação oxidativa podendo também estar envolvidas no envelhecimento e no desenvolvimento de doenças degenerativas.

O genoma mitocondrial completo dos bovinos (*Bos taurus*) com 16337 bp está completo e nele estão incluídos 13 genes codificadores de proteínas tal como se refere no quadro seguinte.

	Localização	Hélice	Comprimento	Gene	Produto
	3101-4057	+	319	ND1	NADH desidrogenase
	4266-5309	+	348	ND2	NADH desidrogenase
	5687-7231	+	515	Cox1	Citocromo C oxidase
	7374-8057	+	228	Cox2	Citocromo C oxidase
Região de Sobreposição	{ 8129-8329	+	67	ATP8	ATP sintase Fo
		+	227	ATP6	ATP sintase Fc
	8970-9750	+	260	Cox3	Citocromo C oxidase
	9823-10.168	+	115	ND3	NADH desid
	10.239-10.535	+	99	ND4L	NADH desid
	10.529-11.906	+	459	ND4	NADH desid
	12.109-13929	+	607	ND5	NADH desid
	13.913-14.440		176	ND6	NADH desid
	14.514-15.653	+	380	CYTB	Citocromo b

<http://www.ncbi.nlm.nih.gov>

Estão também assinalados neste genoma mitocondrial bovino 24 genes de RNA que se referem no quadro seguinte.

<u>Localização</u>	<u>Hélice</u>	<u>Produto</u>
<u>RNA ribossomal</u>		
431-1385	+	12 S ribossomal RNA
1453-3023	+	16 S ribossomal RNA
<u>RNA de transferencia</u>		
364-430	+	tRNA – Phe
1386-1452	+	tRNA – Val

3024-3098	+	tRNA – Leu
4057-4125	+	tRNA – Ile
4123-4194	-	tRNA – Gln
4197-4265	+	tRNA – Met
5308-5374	+	tRNA – Trp
5376-5444	-	tRNA – Ala
5446-5518	-	tRNA – Asn
5551-5617	-	tRNA – Cys
5618-5685	-	tRNA – Tyr
7229-7299	-	tRNA – Ser
7304-7372	+	tRNA – Asp
8061-8127	+	tRNA – Lys
9754-9822	+	tRNA – Gly
10170-10238	+	tRNA – Arg
11907-11976	+	tRNA – His
11977-12036	+	tRNA – Ser
12039-12108	+	tRNA – Leu
14441-14509	-	tRNA – Glu
15658-15726	+	tRNA – Thr
15726-15791	-	tRNA - Pro

### **2.3 – Tipos de estruturas do DNA e suas conformações alternativas**

#### Conformações do DNA linear e circular e dosDNA

Recorda-se que as hélices de DNA podem assumir diversas conformações. A mais comum, a forma B-DNA, consiste em hélices enroladas para a direita com dez pares de bases em cada espira possuindo pois duas depressões ou entalhes (*groves*) distintas, uma maior e outra menor.

A forma A-DNA tem hélices também enroladas para a direita mas é mais curta e mais larga que a forma B-DNA.

A forma Z-DNA tem hélices enroladas para a esquerda e nelas a alternância entre purinas e pirimidinas origina um aspecto de zig zag.

Também se recorda que o DNA pode ser uma molécula linear (por exemplo nos núcleos das células eucariotas) ou circular (por exemplo nas mitocôndrias, bactérias, etc.), podendo, em princípio, uma molécula linear de DNA converter-se numa molécula circular.

O DNA circular em hélice dupla (pois algum pode existir em hélice simples em certos fagos) tem uma topologia curiosa, dado que essa estrutura circular torce-se e contem assim superenrolamentos.

O DNA circular na forma relaxada (ou seja o formado pela ligação fósfodiéster das duas extremidades de um DNA linear) sem qualquer outra modificação tem uma estrutura na sua hélice dupla na forma de B-DNA. Contudo, antes de fechado o círculo, podem formar-se estruturas de DNA superenroladas quando uma das extremidades do DNA for enrolada em relação a outra e podem resultar assim estruturas torcidas que no caso de terem menos voltas são designadas como DNA superhelicoidal

negativo, havendo no entanto vários tipos de superhélices do DNA, pois podem ocorrer superenrolamentos positivos ou negativos.

As variações que podem ocorrer na conformação dos constituintes dos desoxirribonucleótidos do DNA estão associadas com as variantes conformacionais observáveis no DNA, ou sejam, as formas A, B e Z.

Por outro lado, podem surgir no DNA estruturas invulgares, pois o DNA não é uma estrutura rectilínea e uniforme em toda a sua extensão, podendo surgir arranjos cruciformes, hélices triplas ou quadruplas, dobragem angular, tudo isto condicionando as interacções do DNA com diversíssimos tipos de proteínas factores de transcrição ou não.

Estas estruturas invulgares do DNA são o corolário da ocorrência de motivos específicos na sequência do DNA, que se designam abreviadamente por dos DNA (sequências ordenadas definidas de DNA) como é o caso das repetições invertidas, das repetições em espelho, das repetições directas, das sequências homopurinas-homopirimidinas, dos tractos faseados A e das regiões ricas em G, etc. que passamos a abordar seguidamente (Devlin,1997).

É conhecido que as sequências ricas em AT facilitam a separação das duas hélices do DNA, na zona próxima do início da replicação do DNA, e também estão identificadas no genoma humano muitas sequências homopurinas-homopirimidinas e troços alternados de purinas-pirimidinas.

#### Dobragens angulares ou curvaturas do DNA

É conhecido que a existência de troços no DNA contendo 4 a 6 bases de adenina intervaladas por 10bp originam nesse DNA conformações curvas, o que facilita interacções com diversas moléculas proteicas implicadas na replicação, transcrição, recombinação sítio-específica, e assim por diante.

No entanto esta sequência nucleotídica não necessita de ser tão exacta para serem desencadeadas curvaturas no DNA quando este interactua com enzimas e com histonas.

Mas também a ausência de espaçadores entre estas sequências com poliadeninas, ou mesmo a presença de certos troços de purinas-pirimidinas, podem originar menos "flexibilidade" do DNA, que passa a então a designar-se por DNA anisomórfico.

#### DNA cruciforme (Devlin,1997)

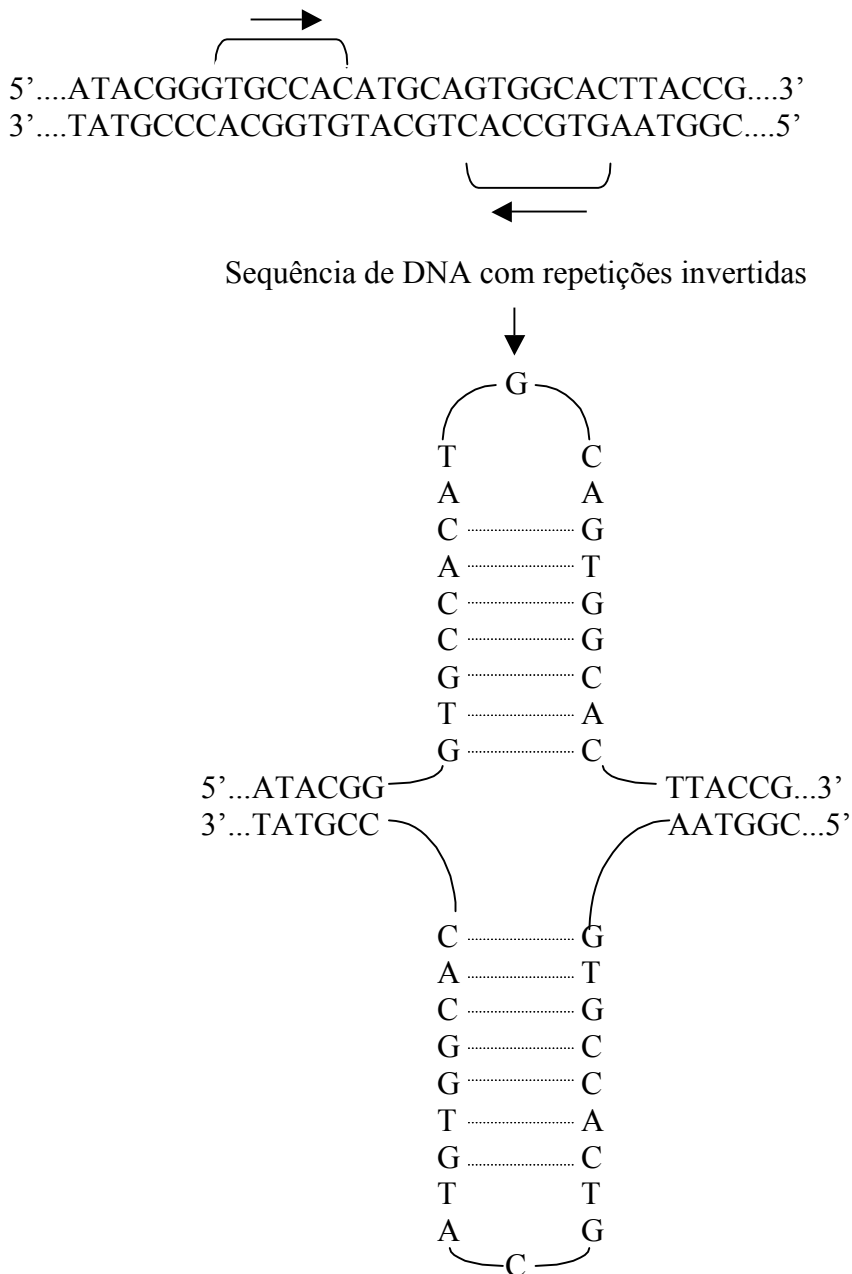
São as regiões de DNA não codificantes que contêm geralmente os dos DNA. Estes dos DNA consistem de diversos elementos de simetria (ver figuras seguintes) tais como as repetições invertidas, as repetições invertidas completamente simétricas ou sejam os palíndromos, as repetições ao espelho, e as repetidas directas.

Dentro deste dos DNA o emparelhamento clássico das bases de Watson e Crick pode ser rompido e formarem-se conformações variadas do DNA dentro dos dos DNA, como é o caso de junções e outras estruturas invulgares como as cruciformes e arranjos em hélices triplas e quadruplex ou ainda DNA deslizante não emparelhado (*slipped mispaired*).

As funções biológicas destas estruturas invulgares são mal conhecidas.

Há no entanto quem admita que as conformações cruciformes do DNA podem estar implicadas no controlo da replicação e transcrição, com a rotura das pontes de H entre hélices complementares e a formação de pontes de H dentro da mesma hélice dentro da região repetida a produzir uma estrutura cruciforme (ver figura seguinte).

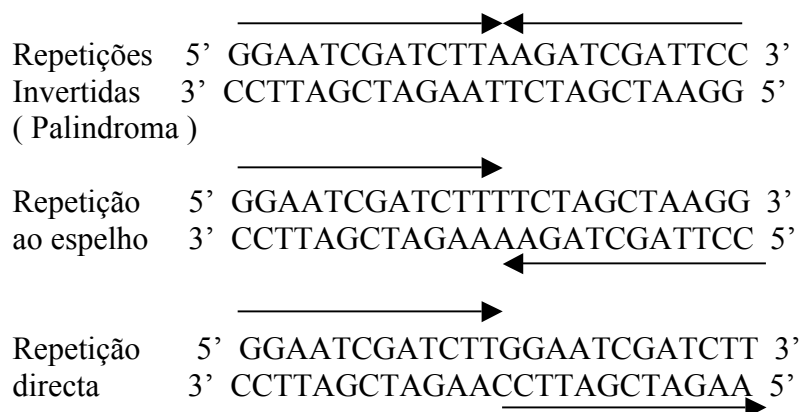
Formação de estruturas cruciformes no DNA



Emparelhamento de bases entre segmentos complementares da mesma hélice de DNA (Adaptado de Devlin, 1997 Fig.14.33).

A existência de repetições invertidas na dupla hélice de DNA é necessária mas não é suficiente das estruturas cruciformes.

### Elementos de simetria das sequências de DNA (Adaptado de Devlin, Fi.14.32)



Nos três tipos de simetria as setas indicam as relações específicas destes elementos em cada uma destas sequências.

Nos palindromos cada hélice simples do DNA é complementar dentro da região invertida que contem os elementos de simetria.

Na repetição em espelho há a presença de bases idênticas á mesma distância de um centro de simetria dentro do segmento de DNA.

Na repetição directa uma sequência particular é repetida. As repetições não necessitam ser adjacentes umas das outras.

No DNA relaxado não é viável a formação da conformação cruciforme uma vez que o DNA linear acomoda mais bases emparelhadas por pontes de H. O desenrolamento seguido pela formação de pontes de H dentro da mesma hélice de DNA entre duas partes simétricas da repetição leva à formação da conformação cruciforme. As ansas (*loops*) formadas nesta conformação requerem o não emparelhamento de 3-4 pares de bases na extremidade das estruturas em gancho de cabelo o que exige gasto de energia.

Em regiões que consistam de repetição ao espelho não se formam conformações cruciformes. Estas repetições ao espelho tendem mais a formar hélices triplas.

### DNA de hélices triplas (Devlin,1997)

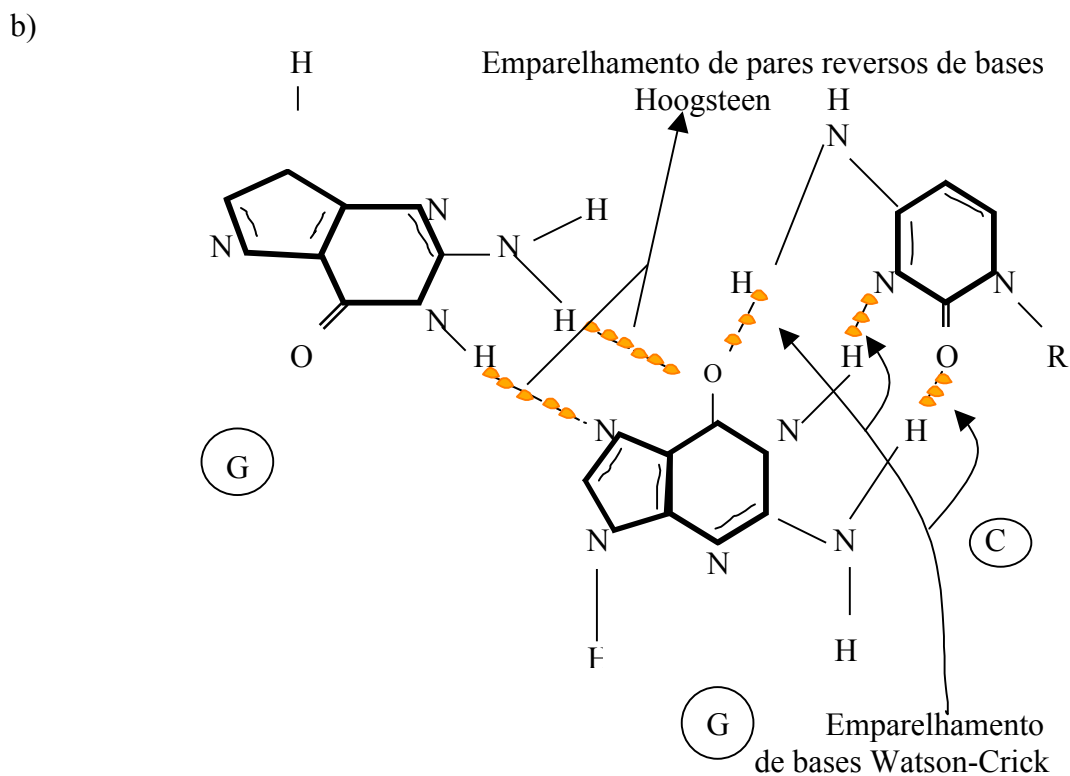
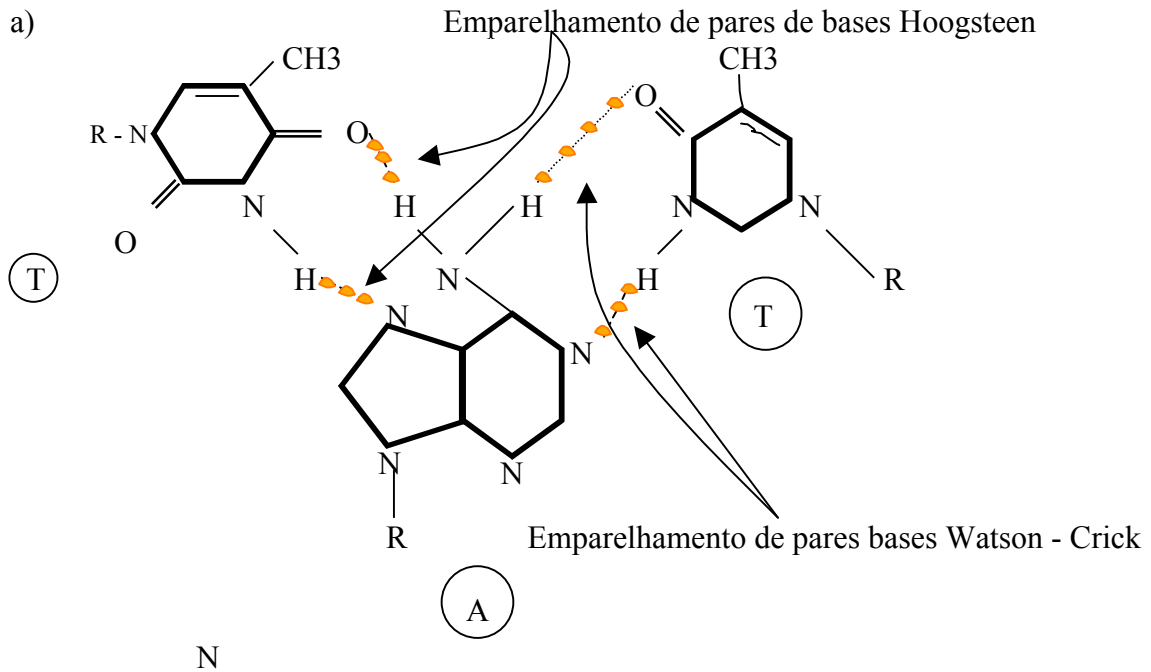
No genoma humano, sobretudo em regiões implicadas na regulação dos genes, pode ocorrer a formação de estruturas invulgares em hélice tripla. Esta hélice tripla pode resultar da interacção da hélice dupla de DNA com um outro polidesoxirribonucleótido (intermolecular) ou então forma-se dentro da mesma estrutura de DNA (intramolecular).

São as regiões de DNA possuidoras de séries contínuas de homopurinas-homopiridinas (troços de 25 desoxirribonucleótidos de polipurinas constituem cerca de 0,5% de alguns genomas eucariotas) que tendem a originar estas hélices triplas. Parecem estas estruturas estar implicadas na regulação da transcrição, iniciação da replicação, terminação desta, ou ainda funcionarem como *enhancers* da estabilidade dos telomeros, bem como iniciadores de recombinação genética.

A hélice tripla é produzida pela formação de pontes de hidrogénio de uma 3ª hélice com o entalhe “groove” principal do B-DNA.

Este B-DNA forma pares de bases, e a 3ª hélice forma pontes de hidrogénio com a outra superfície da dupla hélice através de pares de Hoogsteen, podendo formar-se a estrutura em hélice tripla com quatro tipos de bases triplete-TAT, CGC, GGC e AAT. Na figura seguinte mostra-se a estrutura de dois destes tripletos.

Emparelhamento de bases em DNA triplex(Adapt. Devlin Fig.14.35)



Em a) mostram-se as três hélices, uma poly T, outra poly A e outra poly T. Neste triplex T – A – T, a purina A participa num emparelhamento de Watson e Crick com T, e num emparelhamento de Hoogsteen com outro T.

Em b) mostram-se as três hélices, uma poly G, outra poly G e a terceira poly C. A purina G forma com C um par de Watson-Crick, e com a outra G um emparelhamento de Hoogsteen, estando os grupos ribose das duas purinas numa orientação *trans* gerando um emparelhamento chamado reverso de Hoogsteen com outro T.

Na prática hélices triplas intermoleculares de DNA só podem formar-se dentro de regiões homopurina – homopirimidina do DNA.

Hélices triplas intramoleculares podem formar-se por rotura de pontes de hidrogénio em regiões de DNA com hélices de polipurinas.

Para se formar a estrutura do DNA em triplex a região contendo polipurinas – polipirimidinas deve conter uma simetria repetida em espelho.

Um tipo distinto de hélice tripla intermolecular é formado por acção de enzimas durante a fase intermédia da recombinação geral, sendo estas hélices triplas atípicas pois não são limitadas a regiões polipurina – polipirimidina, mas envolvem antes hélices de DNA de sequências desoxirribonucleotídicas idênticas ou quase idênticas.

Estas hélices são estruturas desenvolvidas onde depois se liga uma terceira hélice no lado da “*groove*” principal da dupla hélice de uma forma paralela.

Sequências longas polipurina – polipirimidina podem formar outra variante estrutural do DNA, o módulo DNA que consiste de um par de duas regiões triplex intermoleculares.

O papel biológico da formação do DNA triplex está bem evidenciado no caso da persistência da hemoglobina fetal nos adultos e ainda no seu possível potencial terapêutico noutras situações.

#### DNA de quatro hélices ( quadruplex ) (Devlin, 1997)

Estas estruturas de quatro hélices de DNA podem formar-se como estruturas paralelas e anti-paralelas, as primeiras podendo formar-se durante a recombinação, podendo encontrar-se nos genes das cadeias pesadas das imunoglobulinas.

Estruturas de DNA paralelas e antiparalelas de quatro hélices de DNA formam-se nos telomeros que contêm sequências oligodesoxirribonucleotídicas simples repetitivas (tais como G4 T2) que são geralmente ricas em purinas numa das hélices. As sequências repetitivas tornam possível a formação de DNA quadruplex.



DNA “*slipped*” (deslizante)  
(Devlin, 1997)

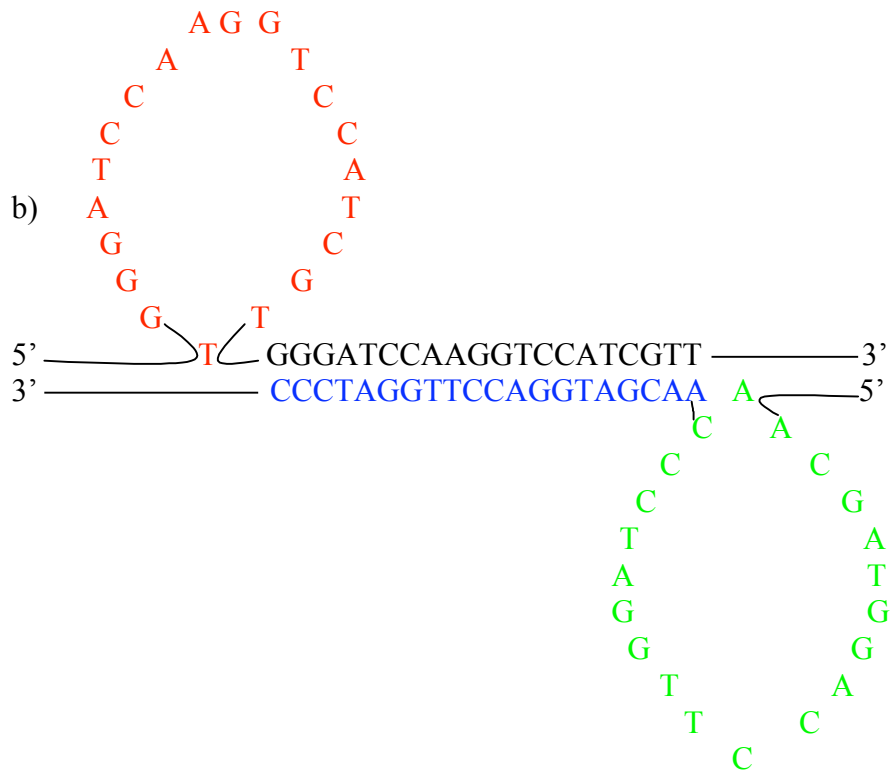
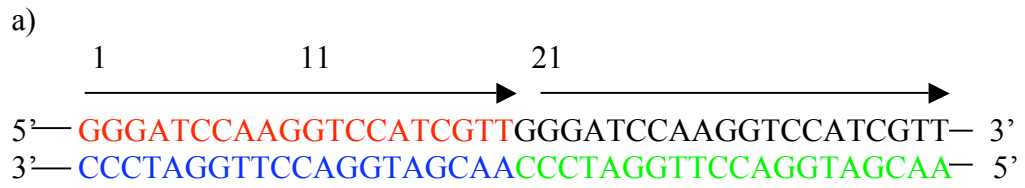
As regiões do DNA com simetria repetida directa (ver figura atrás) podem formar “*Slipped mispaired DNA* ( SMP – DNA )” ou seja DNA deslizante mal emparelhado.

A sua formação implica o desenrolamento da dupla hélice e o realinhamento e emparelhamento de uma cópia da repetição directa de uma hélice com outra cópia adjacente na outra hélice, produzindo uma ansa de cadeia simples (*vide* figura seguinte) sendo possível a formação de dois isómeros SMP – DNA.

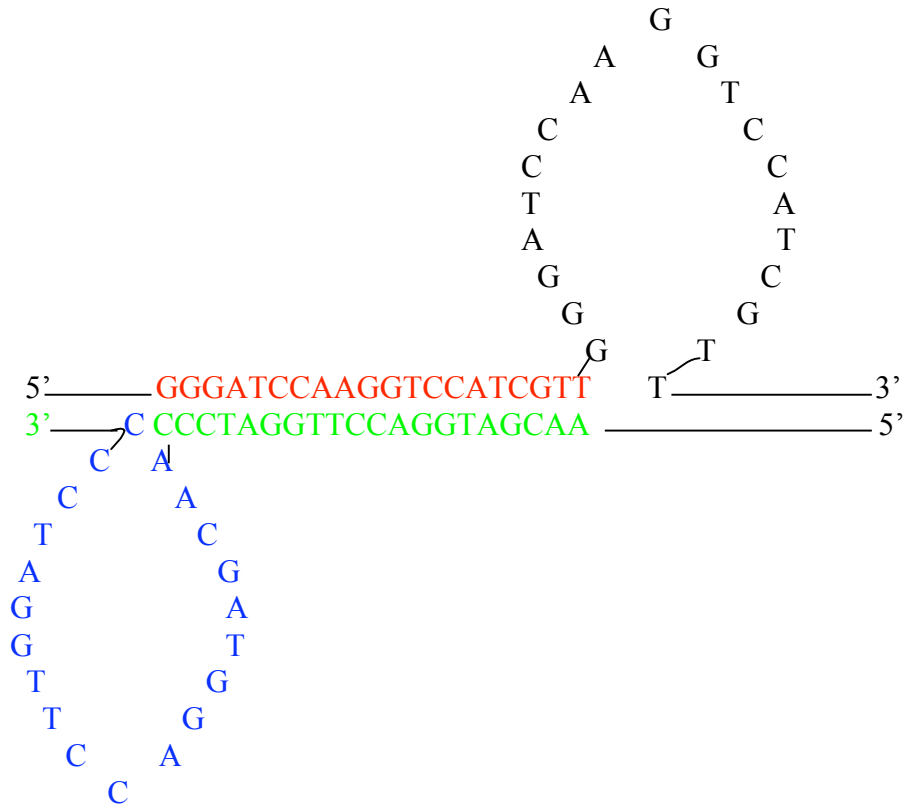
Este tipo de DNA está indiscutivelmente implicado na mutagénese espontânea “*frameshift*” que se manifesta como a adição de bases ou sua deleção.

Mudanças num simples par de bases podem produzir uma de três tipos de mutações: a “*missense*”, que origina uma proteína que tem um ácido aminado substituído por outro; a “*nonsense mutation*”, na qual um codão de *stop* substitui um codão de ácido aminado, levando a uma terminação prematura da translação; uma “*frameshift mutation*”, que leva a introdução de um ácido aminado não relacionado e geralmente codões de *stop*. Pequenas deleções têm efeitos similares às mutações “*frameships*”.

DNA deslizante não emparelhado  
( slipped mispaired DNA )(Adapt. Devlin Fig.14.40)



c)



A presença de duas repetições adjacentes em tandem a) pode dar origem a dois isómeros b) e c) nos quais em b) a segunda cópia da repetição directa da hélice superior emparelha com a primeira cópia da repetição da hélice inferior produzindo um par de ansas (*loop*) com uma simples cadeia.

No isómero c) as coisas passam-se com o emparelhamento da primeira cópia da hélice superior com a segunda cópia da hélice inferior gerando também duas ansas de cadeia simples.

### Relevância da dinâmica estrutural do DNA

A molécula de DNA é dinâmica e polimorfa por natureza, apesar da forma B enrolada para a direita ser a conformação do DNA predominante. Conformações não de B-DNA, como a Z-DNA, triplex, tetraplexes, gancho de cabelo, cruciforme, etc., ocorrem nos sistemas biológicos em diversas condições fisiológicas.

Pouco é conhecido acerca da topologia genómica do DNA, com formas B-DNA enroladas para a direita e sem emparelhamentos de Watson Crick, a sua capacidade funcional e as suas possíveis implicações em patologias diversas como por exemplo a doença de Alzheimer.

Também muito pouca informação existe disponível sobre o papel da apoptose na modulação da tipologia do DNA.

## 2.4 - Recombinação do DNA

(Devlin, 1997)

A recombinação do DNA ocorre por processos distintos que levam a que o material genético seja rearranjado por cisão e junção de porções dentro da mesma molécula de DNA ou de porções de moléculas de DNA diferentes, havendo a possibilidade de ocorrer no DNA das células eucariotas ou das células procariotas, podendo mesmo acontecer entre DNAs de diferentes organismos para gerar novo “*composite*” DNA.

Os três tipos bem conhecidos de recombinação genética geral referem-se no quadro seguinte.

A recombinação origina novas combinações dos genes nos cromossomas aumentando assim a diversidade genética.

Na recombinação genética homóloga ocorre uma permuta entre um par de moléculas distintas de DNA, por vezes duas cópias com diferenças notáveis, no mesmo cromossoma, ou entre dois segmentos de DNA produzido a partir da mesma molécula de DNA.

Para que possa ocorrer este processo torna-se necessário que os DNA recombinantes sejam homólogos, isto é que os dois DNA tenham uma sequência de bases muito semelhante numa região extensa que pode conter vários milhares de pares de bases. E um exemplo deste processo de recombinação homóloga nas células eucariotas é a permuta de secções de cromossomas homólogos durante o desenvolvimento inicial dos gametas (óvulo e espermatozóide), desta forma produzindo-se durante a meiose versões bastante diferentes do mesmo gene (alelos).

Nas bactérias o processo de recombinação geral é muito frequente.

Esta recombinação homóloga envolve um mecanismo em várias etapas catalisadas por diversas proteínas, sendo de salientar entre estas a Rec A proteína, que nas recombinações homólogas da *E. coli* são codificadas pelos genes Rec.

A recombinação homóloga é acompanhada da formação de uma região de DNA heteroduplex e necessita de quebra e junção do DNA cromossomal.

Os genes das imunoglobulinas são “*assembled*” por recombinação.

A recombinação homóloga na *E. coli* necessita de cerca de vinte e cinco enzimas ( Rec A, Rec BCD, Ruv AB, Ruv C, etc., etc. ).

Proteínas homólogas ao Rec A tem sido isoladas em leveduras e células humanas.

Na recombinação sítio-específica, ou recombinação conservadora sítio-específica, é necessária a presença de apenas curtas sequências homólogas de DNA. Contudo, estas recombinações ocorrem apenas em sequências específicas do DNA presentes em ambas as moléculas de DNA interveniente, sendo o processo catalisado por recombinases.

Esta recombinação sítio-específica é limitada a regiões próprias do genoma sendo catalisada pelas recombinases que reconhecem sequências específicas curtas ( 20-200bp ) em ambos os locais de recombinação.

A recombinação sítio-específica transposicional ou transposição difere da recombinação sítio-específica em não necessitar de uma sequência específica no DNA no cromossoma alvo, sendo catalisada por transposases.

As transposases e as recombinases reconhecem e actuam sobre sequências específicas do DNA.

Esta transposição é melhor conhecida nas bactérias, mas o DNA de todas as células, inclusive eucariotas como a *Drosophila*, milho e levedura, contem segmentos transposões que podem mover-se mas com baixas frequências ( $10^{-4} - 10^{-7}$  por geração celular) de um local dador para outro local alvo dentro de um cromossoma.

A transposição difere da recombinação homóloga não necessitando de sequências homólogas entre o dador e o sítio alvo, só o sítio dador, ou seja o transposão, tem sequências nucleotídicas específicas localizadas de um lado e do outro do transposão que servem de local de ligação para as transposases.

Características dos diferentes tipos  
de recombinações genéticas (Devlin, Table 15.8)

Tipo	Homologia da sequência	Sequências do heteroduplex	Proteínas implicadas	Síntese de DNA
Homóloga	Extensa, mas a homologia é independente da sequência de DNA	Longa	Rec A, Rec B,C,D, Ruv ABC e enzimas de reparação do DNA	Alguma
Sítio-específica	Curta, mas sequências específicas do DNA são necessárias em ambos os DNA	Curta	Recombinases	Alguma
Transposicional	Homologia não é necessária: Sequências específicas necessárias apenas num dos DNA	Nenhuma	Transposases	Muito pouca ( apenas no “ <u>fil gaps</u> ” )

A recombinação de qualquer tipo é responsável pela inserção de vírus, plasmidos e elementos transposáveis (transposões) para o DNA cromossomal.

Os transposões são pois elementos de DNA que se podem deslocar de local para local dentro do genoma bacteriano ou dos eucariotas.

A recombinação mais frequente é a do tipo homólogo, sendo os outros dois tipos relativamente raros, gerando a recombinação homóloga novas combinações de genes levando à diversidade genética.

## **2.5– Estrutura dos genes e sua transcrição e translação nas células eucariotas**

Calcula-se que as células eucariotas dos animais superiores contenham cerca de 35000 genes, sendo alguns deles expressos ao mesmo tempo em todas as células (é o caso dos genes chamados de “*house-keeping*” e que são responsáveis, por exemplo, pelas funções metabólicas comuns a todas as células). Contudo, outros genes só são expressos quando a célula se diferencia para dar um tipo especial de célula. Outros ainda são expressos sempre em células que se diferenciaram para dar um tipo especial de célula (por exemplo genes implicados na síntese de anticorpos) havendo ainda aqueles genes que apenas são expressos em condições de mudança no ambiente celular (regulação positiva ou negativa desencadeada por exemplo por uma hormona ou por um factor de crescimento, etc.).

A expressão dos genes pode ser regulada nas células eucariotas de diversas formas tais como:

- Alteração de ritmo de transcrição do gene em causa.
- Alteração do ritmo de processamento do RNA transcrito ainda ao nível do núcleo.
- Alteração da estabilidade e degradação das moléculas de RNA.
- Alteração da eficiência da tradução nos ribossomas do mRNA.

É conhecido para o genoma humano que apenas uma pequena parte desse genoma codifica para a biossíntese de proteínas e que mais de 95% não codifica, não se sabendo qual o papel biológico de toda esta fracção.

Este DNA não codificante pode encontrar-se com diversas localizações quer dentro dos próprios genes (por exemplo no caso dos intrões) ou entre diversos genes diferentes.

Repete-se, tal como já se afirmou anteriormente, que algum deste DNA aparentemente sem funções pode encontrar-se na forma de sequências desoxirribonucleotídicas repetidas e que podem ter significado biológico como sucede por exemplo nos centómeros e telómeros onde podem encontrar-se na forma em tandem.

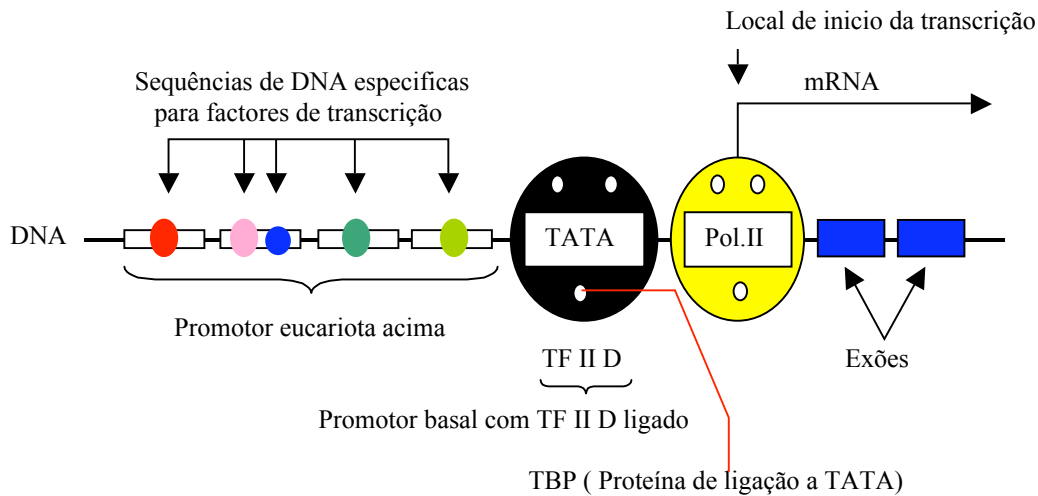
Os genes codificadores de proteínas além dos exões e intrões, contem ainda (*vide* figura seguinte) o local de início para a transcrição, o local promotor-compreendendo o promotor âmago ou basal situado cerca de 40bp do local de início para a transcrição e o promotor acima do qual se pode estender cerca de 200bp acima, e ainda os locais favorecedores (*enhancers*) e os locais silenciadores.

“O promotor basal contem uma sequência chamada a caixa TATA onde se liga o factor de transcrição IID (TFIID) que é um complexo de cerca de 10 ou mais proteínas diferentes, inclusive a TBP (*TATA-binding protein*) que reconhece e interaccua com a caixa TATA, e outros factores proteicos que se ligam ao TBP mas não ao DNA.

O promotor basal encontra-se em todos os genes codificadores de proteínas enquanto o promotor acima difere de gene para gene.

O que torna um dado gene activo num determinado tipo de célula é provavelmente a combinação única dos locais promotores e dos factores de transcrição que com eles interaccuam”(Gene regulation in eukariotes).

(Adaptado de Gene Regulation in Eukariotes)



Os silenciadores e os favorecedores são regiões de controlo do DNA podendo localizar-se a milhares de pares de bases de distância do local de início da transcrição do gene que controlam. No entanto, quando os factores de transcrição se ligam a eles a expressão do gene é condicionada.

Os favorecedores quando interactuam com determinados factores de transcrição (“*enhancer binding proteins*”) aumentam a velocidade de transcrição do gene correspondente.

Os favorecedores podem situar-se acima, abaixo, ou mesmo no interior do gene que controlam.

Esta regulação desencadeada pelos favorecedores sobre os genes situados a milhares de pares bases de distância desencadeiam a formação de ansas (*loops*) no DNA talvez devido ao facto das “*enhancer-binding proteins*” se ligarem ao DNA e poderem também interactuar com outros factores de transcrição assemblados no promotor do gene.

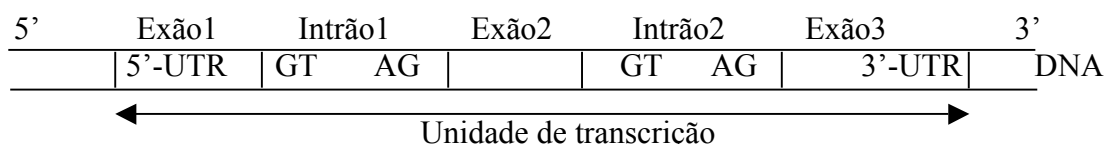
Genes adjacentes, quer se trate de codificadores de RNA, quer se trate de codificadores de proteínas, são por vezes separados uns dos outros por isoladores que evitam interacções cruzadas entre os promotores e favorecedores e as silenciadoras de vários genes, evitando que por exemplo um favorecedor de um gene situado a grande distância deste possa promover a formação de ansas no DNA, e levar à activação do promotor de um outro gene na mesma região do cromossoma.

Os isoladores são pois extensões de algumas dezenas de pares de bases situados entre os favorecedores e o promotor, ou entre o silenciador e o promotor de genes adjacentes, ou de conjuntos destes, procurando evitar que um gene seja influenciado pela activação ou repressão de genes vizinhos.

Estrutura dos genes e transcrição nas células eucariotas  
(Bradley et ali., 2001)

A unidade de transcrição de um gene contém a região transcrita num RNA primário (precursor do mRNA) constituído por exões e por intrões (os intrões na sua sequência desoxirribonucleotídica começam sempre com GT e terminam sempre com AG).

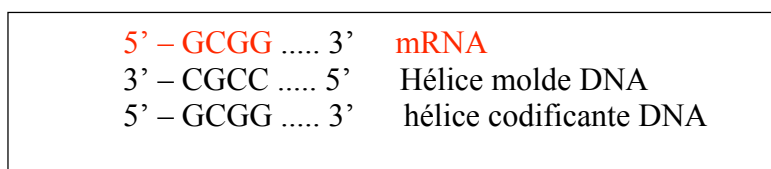
Unidade de transcrição (Adaptado de Bradley et ali., 2001, Fig.1.11).



As regiões codificantes no primeiro exão e no último exão são flanqueadas por regiões não transladas (UTR) que fazem parte dos exões e são transcritas em mRNA mas não transladas em proteínas.

As sequências que se encontram do lado 5' da região do DNA (as que se encontram à esquerda numa sequência 5' – 3') são chamadas sequências acima ou a montante (“*upstream*”) enquanto as que se encontram abaixo ou a jusante do lado 3', são (“*downstream*”).

“O primeiro nucleótido (o ponto de partida) da sequência de DNA transcrita é assinalado com +1, o segundo +2, etc. e o nucleótido precedente do ponto de partida é –1. Estas designações referem-se à hélice (*strand*) codificante (*coding*) do DNA. Recorde-se que a sequência da cadeia molde (*template strand*) do DNA é complementar à do RNA transcrito.” (Stryer, 1995).



A sequência de bases do mRNA (a vermelho) é complementar da hélice do DNA molde *template* (a do meio) (*anti-sense*). A outra hélice de DNA (a de baixo) é chamada codificante *coding* hélice (*sense*) porque tem a mesma sequência que o RNA transcrito excepto T em lugar de U.

As sequências de bases dos locais promotores (sequência de DNA que determina o local de início da transcrição para uma RNA polimerase) não são todas idênticas, contudo elas possuem aspectos comuns e muitas vezes diferem apenas em duas ou poucas bases chamando-se-lhes sequências consenso.

O ritmo a que os genes são transcritos é regulado e controlado por sequências de DNA normalmente fora das unidades de transcrição, embora na mesma hélice de DNA, encontrando-se próximas do gene em causa ou a milhares de pares de bases de distância.

“Estas sequências de DNA reguladoras da transcrição denominadas como elementos de controlo *cis*-actuantes (*cis-acting, control elements*) que não codificam proteínas mas actuam como locais de

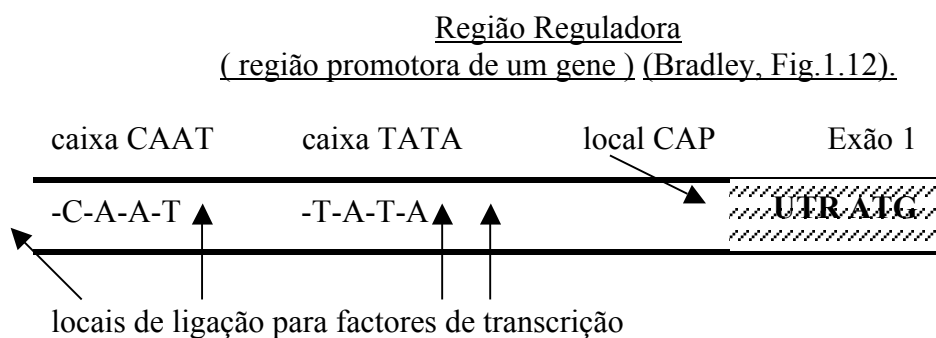


ligação para proteínas produzidas por outros genes, chamados factores de transcrição trans-actantes (*trans-acting transcription factors*) ou proteínas de ligação ao DNA.

Não se conhece bem o mecanismo de actuação destas proteínas de ligação ao DNA mas admite-se que induzam o desenrolamento ou dobragens ou enrolamentos do DNA em ordem a expor a transcrição de certas sequências.

As regiões de controlo cis-actantes são muitas vezes organizadas em “*clusters*” localizados nas regiões promotoras ou facilitadoras (*enhancer*) do gene.

Algumas sequências de DNA são encontradas em regiões promotoras da maioria dos genes (é o caso das caixas TATA e CAAT)” (Bradley et ali.,2001).

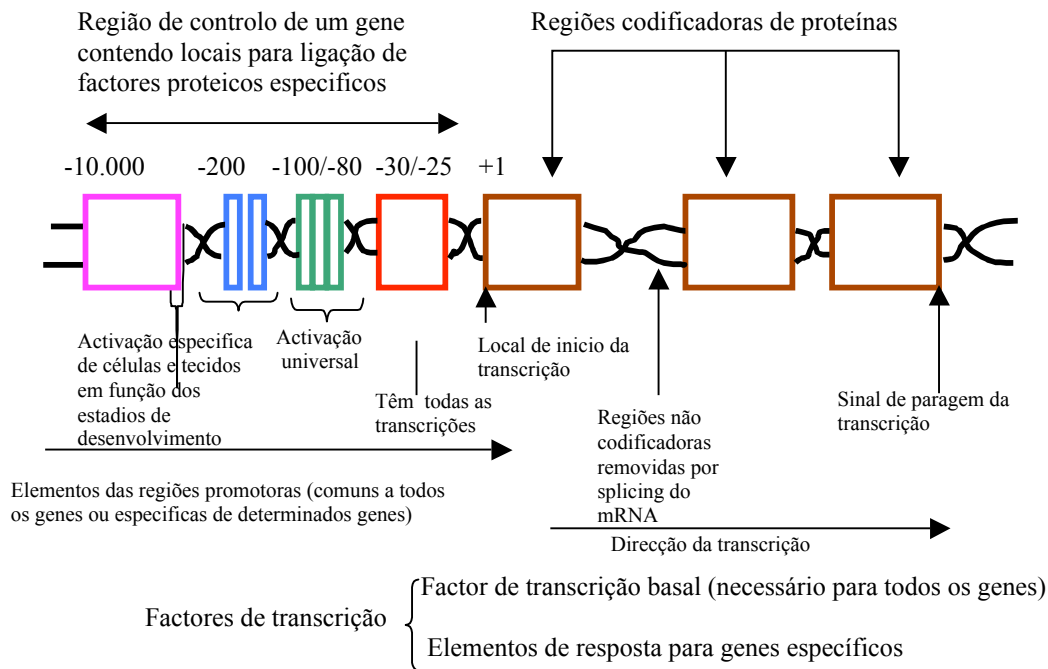


A caixa TATA é constituída por uma sequência rica em AT que se encontra cerca 30bp acima do local de começo da transcrição (muitas vezes é assinalada como - 30 indicando que se encontra 30 desoxirribonucleótidos acima do local de início da transcrição indicado como +1).

As caixas TATA não se encontram muitas vezes nos promotores dos genes “house Keeping” que são genes que codificam por exemplo as proteínas estruturais actina e que se encontram a ser expressas continuamente embora em baixas concentrações, e por vezes tem sequências ricas em GC tais como GGGCGG nas suas regiões promotoras.

As caixas CAAT contêm esta sequência cerca de 80bp acima ( -80 ) do local de partida.

a) Expressão dos genes pode ser esquematizada da forma seguinte



Adaptado de <http://www.copewithcytokines.de/cope.cgi?003305>

Referimos no quadro seguinte as características de alguns factores de transcrição.

Factores de transcrição seleccionados e seus locais de ligação no DNA

<u>Locais de ligação</u>	<u>Proteína</u>	<u>Actividades</u>
TATAA	TF II D	Proteína chave para o “assembly” do complexo de transcrição
GGGCGG6	Sp-1	Activador universal o-glicosilado
ATTTGCAT	Oct-1	Activador universal
TGGNNNNNGCCAA	NF1	Factor de replicação e transcrição
AGGGACTTTC	NF-KB	Factor específico das células B
TGACGTCA	CREB	Factor dependente do C-AMP
TGACTCA	C-Jun	Factor dependente do TPA
GGTACANNNTGTTCT	GR	Factor glucocorticoide dependente
TGGTTAATAATCTACA	LF-B1	Factor específico da célula hepática
NTTCNNGAAN	HSTF	Factor de choque térmico
CGGAGGACTGTCCTCC G	GAL4	Transactivador de genes da levedura governando o metabolismo da glucose

(Adaptado de Gene expression: <http://www.copewithcytokines.de/cope.cgi?003305>)

A transcrição inicia-se no local CAP (*vide* figura adiante) ou seja a extremidade 5’ do mRNA que é capeada neste local por um desoxirribonucleótido especial metilado chamado 7 – metilguanosina.

Após este ponto de início CAP segue-se uma sequência indicadora do codão de início (ATG) que indica o começo da translação (ver figura anterior ou a seguinte). Depois do sinal de paragem da translação (TAA, TAG ou TGA) segue-se uma UTR que inclui um sinal poli(A) (AATAAA) que assinala a cisão do RNA formado de novo, numa posição ligeiramente abaixo (*down stream*) e pela adição de uma cauda poliadenilatada (poli A), não sendo pois esta codificada pelo gene.

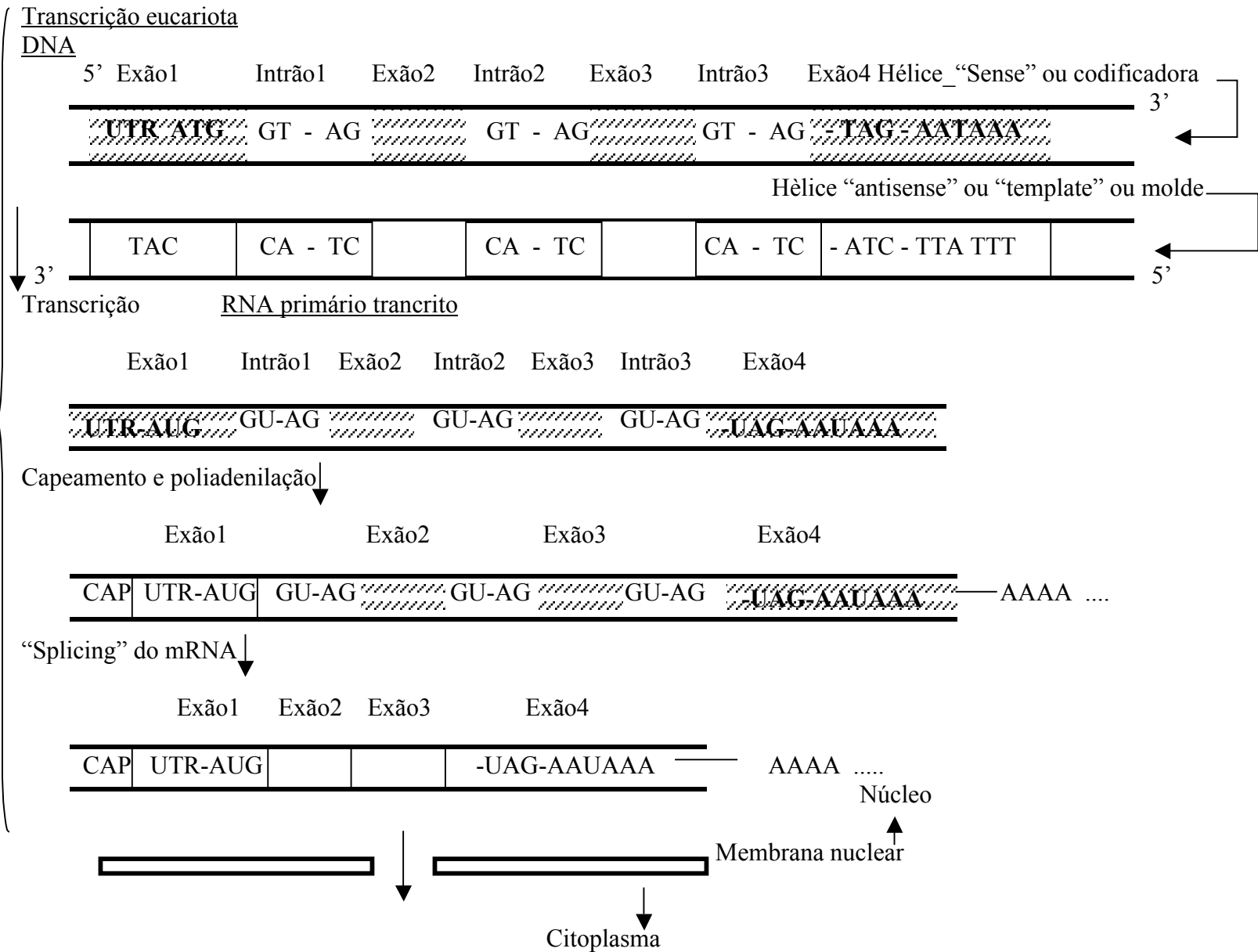
Reafirme-se pois que cada terminação de um mRNA contém uma UTR não translacionada em proteína, parecendo que a região 5' (*upstream*) influencia a translação enquanto a 3' parece ser importante na estabilização do mRNA.

Um mecanismo adicional de controle da translação, chamado "*antisense control*" ocorre nas células bacterianas, mediado por um RNA *anti-sense* que contém sequências complementares da região de um mRNA contendo um codão de iniciação. A hibridação do RNA *antisense* complementar bloqueia o reconhecimento do codão de iniciação e a ligação da sub-unidade ribossomal 30S à sequência de Shine-Delgarno, impedindo assim o início da translação.

O RNA *anti-sense* regula a translação da transposase mRNA nas bactérias.

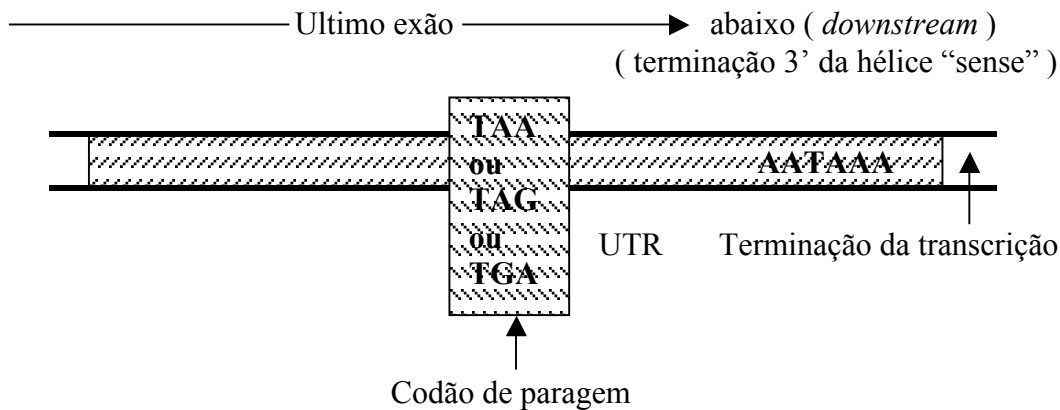
Em certas bactérias a síntese de um mRNA que seja complementar do mRNA da translação é chamado *antisense* mRNA.

N  
U  
C  
L  
E  
O



(Adaptado de Bradley et ali.,2001, Fig 1.13)

## Terminação da Transcrição Eucariota



(Adaptado de Bradley et ali.,2001, Fig.1.14)

A extremidade ( CAP ) 5' liga-se á pequena sub-unidade do ribossoma e é a primeira etapa da translação.

Os transcritos primários de RNA produzidos pela RNA polimerase do tipo II são também denominados RNA heterogénios nucleares ( hrRNA ) porque revelam notável variedade ( heterogeneidade ) do seu tamanho ao contrário dos tRNA e rRNA.

A maioria do rRNA é transcrito pela RNA polimerase do tipo I no nucléolo, enquanto o pequeno ( 5 S ) rRNA e os tRNA são transcritos pela RNA polimerase tipo III na região extranucleolar.

Antes da polimerase II poder reconhecer e ligar-se ao promotor de um dado gene e começar a transcrição, é necessário a presença de proteínas ou sejam factores de transcrição todos eles com o prefixo TF II ( por exemplo o A, D, E, F e H pela cronologia da sua descoberta ).

O TF II D liga-se à caixa TATA e também é conhecido como o factor TATA.

“Quando a RNA polimerase II se liga ao promotor origina a separação e desenrolamento das duas hélices de DNA numa curta porção com a extensão de cerca de 20bp e quando se forma a molécula de mRNA esta hibrida com o DNA em cerca de 10bp” (Bradley et ali., 2001).

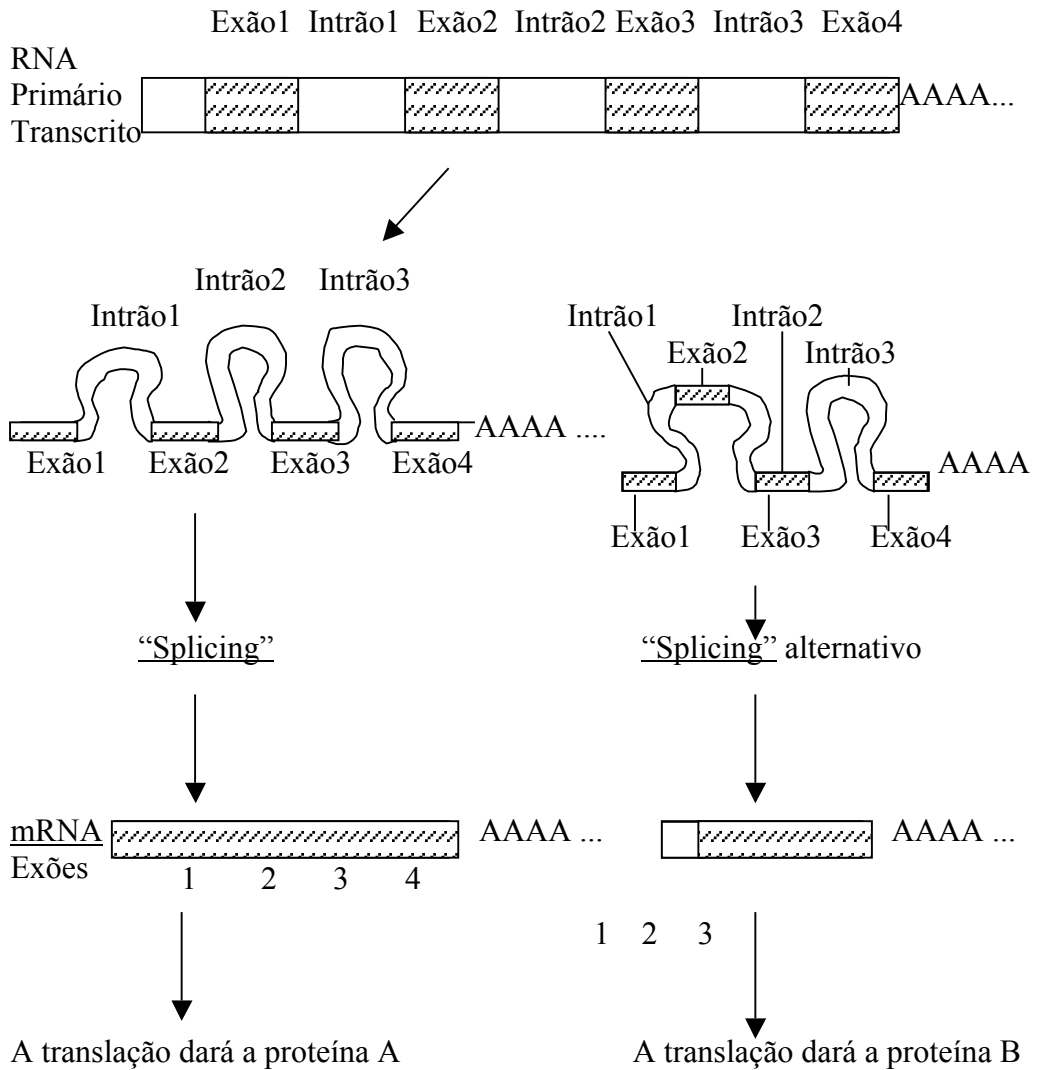
Apenas a hélice molde “*antisense*” é utilizada como molde para a síntese do RNA primário, sendo este uma cópia exacta da hélice de DNA codificante “*sense*” excepto no facto do U substituir a T.

Processamento após a transcrição:

Após a síntese no núcleo do RNA primário, este, ainda no núcleo, é modificado através das seguintes etapas:

- Capeamento, pela adição do nucleótido 7 – metilguanósina.
- Poliadenilatação por adição de uma série de A.
- “*Splicing*” em que as sequências correspondentes aos intrões são excisadas e os restantes exões ligados entre si (*vide* figura).

“Splicing” diferente alternativo sobre um mesmo RNA primário transcrito originando mRNA diferente



(Adaptado de Bradley et ali., 2001, Fig. 1.15)

Após estas etapas o RNA está então pronto para ser transportado para o citoplasma. A translação leva à síntese das proteínas e ocorre nos ribossomas que nos eucariotas são constituídos por uma sub-unidade de 40S ( pequena ) e por uma sub-unidade mais de 60S, que juntas formam partículas 80S. Nas duas subunidades ribossomais 60S as proteínas respectivas estão complexadas com Rna 28S, 5.8 e 5S, ao passo que na sub-unidade 40S as proteínas estão complexas com RNA 18S. Para início da biossíntese proteica ocorre a formação a um complexo macromolecular no qual o codão de iniciação AUG do respectivo mRNA interactua com a mais pequena sub-unidade ribossomal a qual transporta um tRNA acoplado com a metionina, para início da cadeia polipeptídica. Intervém depois uma série de factores proteicos para alongamento (eIFs) da cadeia polipeptídica em formação que interactuam com a grande

sub-unidade ribossomal e levam esta a interactuar com o complexo anteriormente referido, desencadeando-se a movimentação do ribossoma sobre o mRNA.

Recorda-se que os diversos tRNA transportando os vários ácidos aminados que compõem as proteínas, cada um desses tRNA através do seu anti-codão identificador, reconhece e interacciona com o correspondente codão, ou seja, o trio de bases que lhe são complementares (específicas) no mRNA.

Intervem depois a actividade enzimática de uma peptidil-transferase que vai efectuar uma ligação peptídica entre os dois ácidos aminados contíguos na cadeia polipeptídica em formação. São moléculas de RNA situadas à superfície do ribossoma que tem esta actividade de peptidil-transferase, daí serem também denominadas como ribozimas.

A princípio verificou-se que estas moléculas de RNA, ou ribozimas, tinham actividades hidrolíticas sobre o DNA, mas actualmente sabe-se que são as responsáveis pela síntese das ligações peptídicas.

Recorda-se ainda que a biossíntese das cadeias peptídicas é terminada quando o ribossoma encontra sobre o respectivo mRNA um codão indicador para esta terminação como é o caso de qualquer dos codões UAA, UAG e UGA.

Quando isto sucede a cadeia polipeptídica é libertada do último tRNA que lhe acoplou o último ácido aminado dessa cadeia, e as sub-unidades ribossomais dissociam-se também do mRNA.

Há depois todo um processamento da cadeia polipeptídica, após a translação, podendo formar-se simplesmente a sua estrutura complexa, ou as proteínas a segregar passarem através do retículo endoplásmico rugoso para o aparelho de Golgi, para serem processadas, havendo pois translocação de proteínas sendo o seu destino marcado pelos sinais na sua estrutura (*Protein Targeting and sorting*).

Após a sua síntese diversos polipéptidos são modificados de diversas formas como por exemplo por hidroxilação, ou fosforilação de certos ácidos aminados ou ainda pela adição de monossacáridos, polissacáridos, componentes lípidicos, etc.

#### Endereçamento das proteínas targeting (Stryer, 1995)

As cadeias polipeptídicas em vias de serem biossintetizadas já contem na sua sequência sinais que determinam qual o seu último destino, se se destinam por exemplo a ser segregadas para fora da célula ou se são dirigidas para o retículo endoplásmico, ou para os lisosomas, mitocôndrias, núcleo, peroxisomas ou se ficam acantonadas em biomembranas, ou são contidas em vesículas ou se controlam a exocitose.

Nas células eucariotas, logo após se iniciar a síntese proteica, um ribossoma permanece livre no citosol a não ser que ele seja dirigido para o retículo endoplásmico por uma determinada sequência de ácidos aminados no polipéptido em biossíntese que funciona como sinal. Havendo este sinal os ribossomas ligados á membrana permitem que a cadeia polipeptídica a ser biossintetizada seja translocada através da membrana do retículo endoplásmico (ER) e atingido o lumen do ER, podem depois ser glicosiladas e modificadas, passando depois para o aparelho de Golgi onde continuam a ser modificadas, acabando depois por sofrer o seu destino final ou destinadas aos lisosomas, vesículas secretórias, ou plasmamembrana. O transporte das proteínas é feito por vesículas que emanam dos compartimentos dadores e se fundem com os compartimentos alvo.

## 2.6 - Bibliografia

- Bradley, J., Johnson, D., e Rubenstein, D. ( 2001 ) Lecture notes on Molecular Medicine Second Edition. Blackwell Science.
- Bates, A. D. 1993 DNA Topology. Oxford, England: IRL Press,
- Bindem, R. R. 1994 DNA Structure and Function. San Diego, CA: Academic Press.
- Bouffler, S., Silver, A., and Cox, R. 1993 The role of DNA repeats and associated secondary structures in genomic instability and neoplasia. Bioessays 15:409.
- Bray, A. B., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. 1994 Molecular Biology of the Cell, New York: Garland.
- Calladine, C., and Drew, H. Understanding DNA 1992 – The Molecule and How It Works, San Diego, CA: Academic Press.
- Clark, D. J., and Felsenfeld, G. 1991 Formation of nucleosomes on positively supercoiled DNA. EMBO J. 10:387.
- Clayton, D. A. 1991 Replication and transcription of vertebrate mitochondrial DNA. Annu.-Rev. Cell Biol. 7:453.
- Cortopassi, G., and Liu, Y. 1995 Genotypic selection of mitochondrial and oncogenic mutations in human tissue suggests mechanisms of age-related pathophysiology. Mutat. Res. 338:151.
- Getzenberg, R. H., Pienta, K. J., Ward, W. S., and Coffey, D. S. 1991 Nuclear structure and the three-dimensional organization of DNA. J. Cell Biochem. 47:289.
- Haran, T. E., Kahn, J. D., and Crothers, D. M. 1994 Sequence elements responsible for DNA curvature. J. Mol. Biol. 244:135.
- Kim, M. G., Zhurkin, V. B., Jernigan, R. L. and Camerini-Otero, R. D. 1995 Probing the structure of a putative intermediate in homologous recombination: the third strand in the parallel DNA triplex is in contact with the major groove of the duplex. J. Mol. Biol. 247:874.
- Leach, D. R. 1994 Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. Bioessays 16:893.
- Lebowitz, J. 1990 Through the looking glass : the discovery of supercoiled DNA Trends Biochem. Sci. 15:202.
- Lee, M. S., and Garrard, W. T. 1991 Positive DNA supercoiling generates a chromatin conformation characteristic of highly active genes. Proc. Natl. Acad. Sci. USA 88:9675.
- Lewin, B. Genes V. New York: Oxford University Press, 1994
- Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., Haworth, I. S. 1995 Haipin properties of single-stranded DNA containing a GC-rich triplet repeat (CTG)<sub>15</sub> Nucleic Acids Res. 23:1050.
- Neidle, S. DNA Structure and Recognition, 1<sup>st</sup> ed. Oxford, England: IRL Press, 1994.
- Saluz, H. P., and Wiebauer, K. DNA and Nucleoprotein Structure in Viro New York: Demos Vermande, 1995.
- Soyfer, V. N., and Potaman, V. N. Triple-Helical Nucleic Acids. New York Springer-Verlag, 1995.
- Trinh, T. Q., and Sinden, R. R. 1991 Preferential DNA secondary structure mutagenesis in the lagging strand of replication in E. Coli. Nature 352:544.
- Turner, B. M. 1991 Histone acetylation and control of gene expression J. Cell Sci. 99:13.
- Vologodskii, A. V., and Cozzarelli, N. R. 1994 Conformational and thermodynamic properties of supercoiled DNA. Annu. Rev. Biophys. Biomol. Struct. 23:609.
- Wolffe, A. Chromatin Structure and Function. San Diego, CA: Academic Press, 1992.
- Lee, H. S., et al. 1992. Dna sequence requirements for generating paused polymerase at the start of *hsp70*. Genes and Dev. 6:284-295.



- Spencer, C. A., and M. GROUDINE. 1990. Transcription elongation and eukaryotic gene regulation. *Oncogene* 5:777-785.
- Dreyfuss, G., et al. 1993. hnRNP proteins and the biogenesis of mRNA, *Ann. Rev. Biochem.* 62:289-321.
- Nenroth, S., W. KELLER, and E. WAHLE. 1993. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J.* 12:585-594.
- Nenroth, S., et al. 1991. Purification of the cleavage and polyadenylation factor involved in 3' processing of messenger RNA precursors *J. Biol. Chem.* 266:19768-19769.
- Gunderson, S. I., et al. 1994. The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly (A) polymerase. *Cell* 76:531-541.
- Murthy, K. G. K., and J. L. Manley. 1992. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem.* 267:14804-14811.
- Raabe, T., F. J. BOLLUM, and J. L. MANLEY. 1991. Primary structure and expression of bovine poly (A) polymerase. *Nature* 353:229-234.
- Wahle, E. 1991. Purification and characterization of a mammalian polyadenylate polymerase involved in the 3' end processing of messenger RNA precursors. *J. Biol. Chem.* 266:3131-3139.
- WAHLE, E., and W. KELLER. 1992. The biochemistry of 3' -end cleavage and polyadenylation of messenger RNA precursors. *Ann. Rev. Biochem.* 61:419-440.
- AGABIAN, N. 1990. Trans-splicing of nuclear pre-mRNAs. *Cell* 61:1157-1160.
- GREEN, M. R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Ann. Rev. Cell Biol.* 7:559-559.
- JOROWITZ, D. S., and A. R., KRAINER, 1994, Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing. *Trends Genet.* 10:100-105.
- KANDELS-LEWIS, S., and B. SERAPHIN. 1993. Role of U6 snRNA in 5' splice site selection. *Science* 262:2035-2039.
- MANIATIS, T. 1991. Mechanisms of alternative pre-mRNA splicing, *Science* 251:33-34.
- McKEOWN, M. 1992. Alternative mRNA splicing, *Ann. Rev. Cell Biol.* 8:113-155.
- McKEOWN, M. 1993. The role of small nuclear RNAs in RNA splicing. *Curr. Opin. Cell biol.* 5:448-454.
- SALDANHA, R., et al. 1993. Group I and group II introns. *FASEB J.* 7:15-24.
- CARTER, K. C., et al. 1993. A three-dimensional view of precursor messenger RNA metabolism within the mammalian nucleus. *Science* 259:1330-1335.
- KRUG, R. M. 1993. The regulation of export of mRNA from nucleus to cytoplasm. *Curr. Opin. Cell Biol.* 5:944-949.
- LEGRAIN, P., and M. ROSBASH. 1989. Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell* 57:573-583.
- ROSBASH, M., and R. H. SINGER. 1993. RNA travel: tracks from DNA to cytoplasm. *Cell* 75:399-401.
- XING, Y., et al. 1993. Higher level organization of individual gene transcription and RNA splicing. *Science* 259:1326-1330.
- SACHS, A. B. 1993. Messenger RNA degradation in eukaryotes. *Cell* 74:413-421.

## Outras obras básicas

DEVLIN ,T.M. 1997.Textbook of biochemistry with clinical correlations.Fourth Edition. Wiley-Liss. A John Wiley & Sons ,Inc., Publication.New York/ Chichester/ Weinheim/Brisbane/Singapore /Toronto. Cha. 14.

BRADLEY, J. et ali. 2001. Molecular Medicine.Blackwell Science.

Gene regulation in eukariotes .

<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/P/Promoter.html>

STRYER,L.,1995.Biochemistry. Fourth Edition

Gene expression. <http://www.copewithcytokines.de/cope.cgi?003305>

<b>3. - <u>Tecnologia do DNA recombinante</u></b>	99
3.1 - Fragmentação de DNA, vectores, clones e bibliotecas genómicas	99
3.2 - Bibliotecas genómicas a partir de mRNA e Cdna	100
3.3 - Separação de DNA clonado	100
3.4 - Estratégia para estudo de um genoma. Mapa de restrição. Fragmentos de restrição e suas características.	101
3.5 - Formação e utilização de DNA sintético. Síntese química de oligonucleótidos	108
3.6 - Vectores de clonagem do bacteriófago $\lambda$ e construção de um repertório ou biblioteca genómica	110
3.7 - Sondas e identificação de clones específicos	115
3.8 - Informática e análises de dados sobre a sequência de ácidos nucleicos e proteínas	115
3.9 - Locais de início da transcrição podem ser mapeados pela chamada protecção S1 e pela extensão com um primer	117
3.10 - Substituição de genes, animais transgénicos e terapia por genes	119
3.11 – Bibliografia	126

### **3. - Tecnologia do DNA recombinante**

#### **3.1 – Fragmentação de DNA, vectores, clones, e bibliotecas genómicas**

(Lodish et ali., 1995)

É um lugar comum referir que as tecnologias do DNA recombinante (rDNA) redesenharam toda a investigação biológica que se faz nos dias de hoje, e que estas tecnologias se baseiam na utilização de enzimas de restrição que cortam o DNA em sequências específicas de 4 a 8 bp (pares de bases) gerando um conjunto reprodutível de fragmentos de restrição a partir de qualquer genoma.

É conhecido que em média as enzimas de restrição que reconhecem e cortam segmentos constituídos por quatro pares de bases originam fragmentos com cerca de 256 bases de comprimento, e as que reconhecem e cortam segmentos com seis pares de bases produzem fragmentos com cerca de 4.000 bases, enquanto as que têm oito pares de bases no seu local de reconhecimento produzem fragmentos com cerca de 69.000 bases de comprimento.

Os fragmentos produzidos, qualquer que seja a sua proveniência, podem, em condições adequadas, serem ligados entre si através de ligações covalentes desencadeadas pela acção catalítica de DNA ligases. Produz-se assim um DNA recombinado o qual pode ser inserido em diversos tipos de vectores como por exemplo plasmidos e estes depois incorporados em células hospedeiras em cultura, como por exemplo *Escherichia coli*, transformando-as (as raras células bacterianas que tomam estes plasmidos são depois cultivadas e recorda-se que cada colónia provem de uma única célula bacteriana transformada constituindo um clone de células, cada clone transportando um fragmento de restrição particular).

Os vectores para clonagem são pois elementos genéticos que se replicam de forma autónoma e que transportam um cDNA (DNA complementar) ou um fragmento do DNA genómico para uma célula hospedeira para clonagem de genes, sendo os vectores mais utilizados os plasmidos bacterianos, os genomas de bacteriófagos modificados, e os cromossomas artificiais de leveduras (YACs).

Recorda-se ainda que os plasmidos são moléculas de DNA extracromossomal, pequenas, circulares e capazes de se replicar de forma autónoma numa célula.

Os “clones” são uma população idêntica de células ou moléculas de DNA descendentes de um único progenitor.

Os fragmentos de restrição, noutras metodologias que utilizam outros tipos de vectores, podem ser ligados por exemplo a “braços” do genoma do bacteriófago *in vitro*, produzindo viriões infecciosos. Estes viriões infecciosos infectando bactérias *E. coli* cultivadas em placas de Petri com meios de cultura adequados originam nessas bactérias placas de lise, cada placa de lise correspondendo a um dado clone de fago recombinante contendo o fragmento específico de restrição inserido.

É conhecida a eficiência desta clonagem em vectores que permite que praticamente todo o DNA genómico de organismos multicelulares complexos possa ser representado num reportório genómico de alguns milhões de clones independentes.

Noutro tipo de vectores plasmidos, os cósmidos, que contêm as sequências específicas COS do fago *in vitro* (vide adiante) permitem utilizar nos cósmidos fragmentos de restrição de  $\approx 45$  kbp (kbp ou milhares de pares de bases) enquanto a utilização, como vector, do bacteriófago P1 permite utilizar fragmentos de restrição com até  $\approx 100$  kbp de comprimento.

O repertório ou biblioteca genómica (DNA Library) é pois uma colecção de moléculas de DNA clonado, consistindo de fragmentos da totalidade do genoma, ou cópias do DNA de todos os mRNA produzidos por um dado tipo de células (cDNA Library ou biblioteca de DNA complementar) e inseridos num vector de clonagem apropriado.

### **3.2 – Bibliotecas genómicas a partir de mRNA e cDNA**

(Lodish et ali.,1995)

Outro aspecto de muita utilidade na tecnologia do DNA recombinante é o facto dos mRNA poderem ser convertidos em cDNAs com a enzima transcriptase reversa e como quase todos os mRNA eucariotas contem uma extremidade 3' com uma cauda poliadenilatada (poliA), uma mistura de cDNAs correspondendo a praticamente todos os mRNA extraídos de um tipo particular de tecido ou células pode ser sintetizado pela acção catalisada pela transcriptase reversa principiando com oligo-dT hibridando com a cauda poliA.

Esta estrutura de cDNAs pode depois facilmente ser convertida em DNA de hélice dupla e este inserido em vectores apropriados e estes clonados para produzir uma cDNA library repertório ou biblioteca de DNA complementar (cDNA).

Utilizam-se dois métodos para identificar clones codificando um dado tipo de proteínas dentro de um genoma ou numa cDNA library. Um envolve hibridação em membrana com sondas marcadas radioactivamente com cDNA de cadeia simples ou com oligonucleótidos sintéticos. No outro método o clone específico é identificado através da interacção da proteína produzida com um anticorpo monoclonal.

### **3.3 – Separação de DNA clonado**

(Lodish et ali.,1995)

Um DNA clonado pode ser separado do DNA do seu vector pela cisão com uma enzima de restrição apropriada, que é a mesma que foi utilizada para gerar os fragmentos de DNA a inserir no vector. O vector e o DNA clonado podem depois ser separados por electroforese em gel, com poderes de resolução tal que chegam a separar nucleótidos unidos numa cadeia com até 500 nucleótidos de comprimento e diferindo no comprimento apenas em um nucleótido.

As técnicas de gel electroforese podem resolver fragmentos de DNA contendo até cerca de  $\cong$  20kbp. DNAs maiores como por exemplo  $2 \times 10^4$  a  $10^7$  pares de bases (20kbp a 10 megabases-Mbp) de comprimento, podem ser separados pelo tamanho através da *pulsed-field gel electrophoresis* (Lodish et ali., 1995).

A sua visualização após electroforese é fácil com o corante fluorescente brometo de etídeo ou através de marcação radioactiva feita oportunamente.

A sequenciação dos fragmentos de restrição e do cDNA pode ser feita pelos métodos de Maxam-Gilbert ou de Sanger (dideoxi), ambos os métodos utilizando a electroforese em gel de poliacrilamida.

Para detectar um DNA específico ou RNA específico numa mistura complexa, o Southern blotting (DNA) e o Northern blotting (RNA) são utilizados, ambos os métodos envolvendo a electroforese em gel

e a transferência depois (*blotting*) das moléculas separadas para um filtro e exposição deste a uma sonda adequada.

A PCR (*polimerase chain reaction*) é utilizada para ampliar uma dada sequência de DNA quando as sequências das extremidades são conhecidas o que tem inúmeras aplicações em investigação básica e médica. Com esta técnica o DNA de uma simples célula humana (embrionária) pode ser suficientemente ampliado em ordem a permitir sequenciar a uma dada região, para detectar por exemplo mutações associadas com doenças hereditárias.

### **3.4 – Estratégia para estudo de um genoma. Mapa de restrição. Fragmentos de restrição e suas características**

A quebra do DNA de um animal superior (por exemplo com cerca de  $3 \times 10^9$  pares de bases) com enzimas de restrição que hidrolizem este genoma de 3.000 em 3.000 pares de bases, pode produzir cerca de 1 milhão de fragmentos sendo praticamente impossível separá-los directamente uns dos outros.

No entanto é hoje em dia possível obter amostras puras de DNA de grandes genomas através da tecnologia do DNA recombinante. É hoje também possível purificar qualquer gene bem como determinar a sua sequência e as suas regiões funcionais, assim como preparar grande número de moléculas de DNA idênticas. Estas moléculas podem inclusive possuir um fragmento de DNA de interesse ligado a uma molécula vectora de DNA (ligados através de ligações fosfodiéster 3'–5') podendo estas moléculas de DNA recombinado replicar-se quando introduzidos numa célula hospedeira adequada.

Nestas últimas circunstâncias são reproduzidos em grande número os dois tipos de DNA, o DNA do fragmento de DNA adicionado e o DNA do vector.

Como referimos anteriormente os fragmentos de DNA podem ser clonados por diversos tipos de vectores, como por exemplo o fago lambda ( $\lambda$ ), ou plasmídeos de *E. coli*, etc.

Os plasmídeos mais frequentemente utilizados nesta tecnologia recombinante do DNA replicam-se em bactérias da espécie *E. coli*, tendo estas sido optimizadas para este efeito. Contudo, o número de clones individuais que podem ser obtidos por este processo é limitado dada a pouca eficiência da transformação da *E. coli* e o pequeno número de colónias individuais transformadas (apenas algumas centenas) que podem ser detectadas numa simples placa de Petri semeada contendo o meio de cultura adequado.

Isto torna impraticável a clonagem em plasmídeos de todo o DNA genómico dos organismos superiores (por exemplo  $\approx 1,5 \times 10^5$  clones transportando fragmentos de 20kbp de DNA, eram necessários para representar a totalidade de um genoma haplóide contendo  $\approx 3 \times 10^9$  pares de bases) o que é bem mais fácil de alcançar utilizando como vectores de clonagem derivados do bacteriófago lambda ( $\lambda$ ) (Lodish et al., 1995).

Com efeito, esta assinalado que apenas 20 a 30 placas de Petri, cada uma delas contendo cerca de  $5 \times 10^4$  placas lambda (placas de lise do fago  $\lambda$ ) são suficientes para representar a totalidade do genoma com  $3 \times 10^9$  pares de bases, enquanto para rastrear  $10^6$  plasmídeos recombinantes transportando o mesmo genoma seriam necessárias cerca de 5.000 placas de Petri, porque apenas cerca de 200 colónias de *E. coli* transformada podem ser detectadas em cada placa de Petri.

Os plasmídeos variam em tamanho desde alguns milhares de pares de bases até mais de 100 kilobases (kbp).

Um fragmento de DNA contendo de alguns pares de bases até cerca de  $\approx 20$  kilobases (kbp) pode ser inserido num plasmídeo vector.

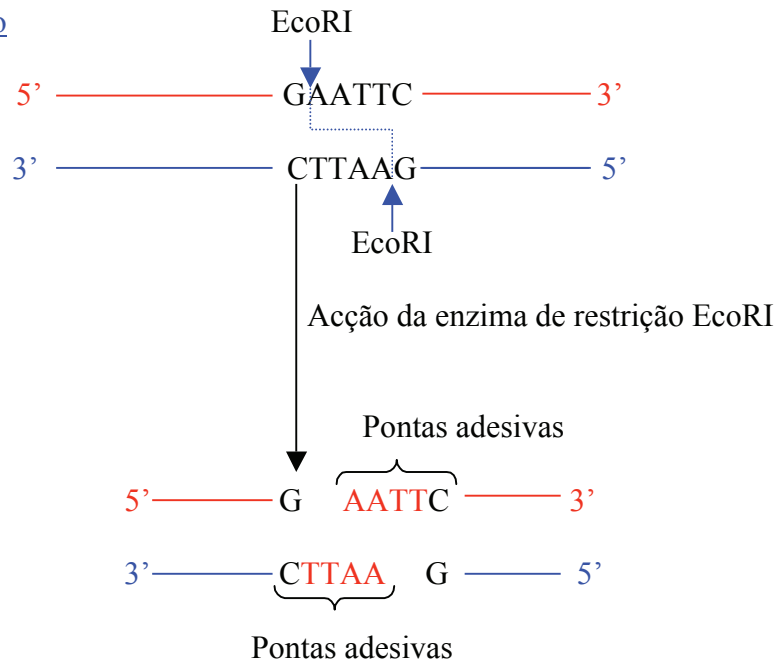
Como já referimos, para clonar fragmentos específicos de DNA num plasmídeo vector, ou noutros vectores, os fragmentos devem ser produzidos e depois inseridos em DNA vector entrando em jogo as enzimas de restrição (endonucleases de restrição) e as DNA ligases, as primeiras cortando as duas hélices das moléculas de DNA em sequências específicas (4 a 8 pares de bases) nos locais de restrição, enquanto as DNA ligases voltam a unir aquilo que as enzimas de restrição separaram.

Como já referimos também diversos locais de restrição, como o local para a Eco R1 (ver na figura seguinte) são curtas sequências repetidas invertidas (palíndromas).

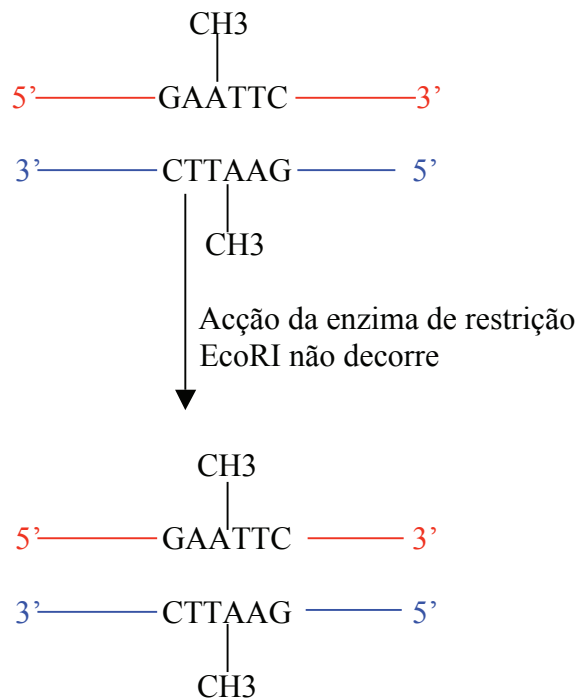
Figura 1

Exemplo de locais de restrição e sua inativação por metilação

DNA não metilado



DNA metilado



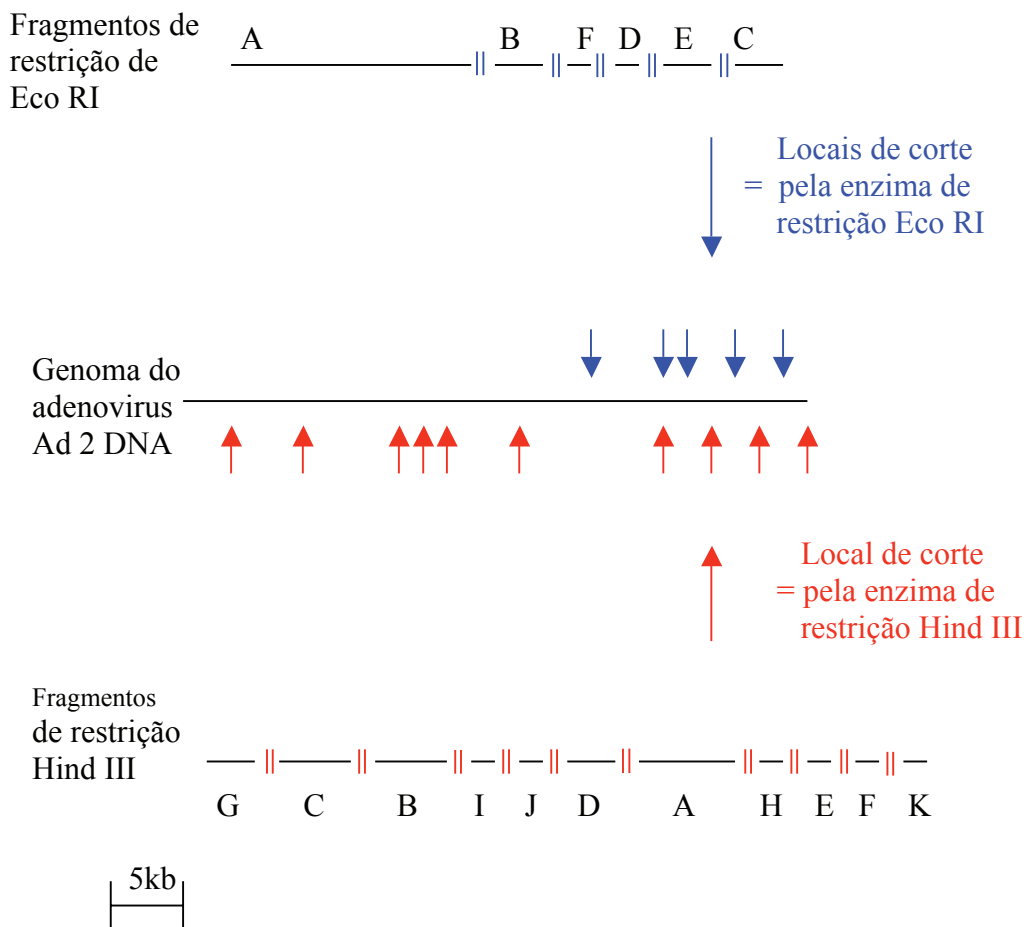
(Adaptado de Lodish et ali., 1995, Fig.7.5)



Repete-se portanto que a sequência do local de restrição é a mesma em cada uma das hélices do DNA quando lidas no sentido 5' \_ 3'.

Como o DNA de uma dada origem tem uma sequência específica, as enzimas de restrição cortam o DNA num conjunto de fragmentos chamados perfil dos fragmentos de restrição ou mapas de restrição de que se dá um exemplo seguidamente (Fig.2).

Figura 2  
Mapa de restrição



(Adaptado de Lodish et ali.,1995,Fig.7.6)

Referem-se na figura 2 anterior os fragmentos de restrição produzidos por quebra do DNA do genoma de  $\approx 36\text{kbp}$  do adenovirus 2 (Ad2) (linha do meio que representa a dupla hélice de DNA) pelas enzimas de restrição Eco RI (em cima) e Hind III (em baixo). O Eco RI produz seis fragmentos em consequência dos locais de restrição (GAATTC) na sequência do Ad2, enquanto o Hind III com o local de restrição AAGCTT produz onze fragmentos específicos. Os fragmentos específicos obtidos convencionou-se designá-los de A a Z por ordem de tamanho decrescente (Lodish et ali., 1995).

Por técnicas adequadas é possível determinar a ordem de encadeamento dos fragmentos de restrição no DNA original.

O mapa com os locais de restrição obtido com várias enzimas de restrição é uma característica única de cada DNA.

Diversas enzimas de restrição produzem fragmentos de DNA com extremidades coesivas (“Sticky ends”), (ver Fig.1 em DNA não metilado).

As extremidades dos fragmentos de restrição produzidos num dado local de restrição são complementares daquelas produzidas noutros fragmentos de restrição de um DNA diferente produzidos pelo mesmo tipo de enzima de restrição, independentemente da origem do DNA. Este emparelhamento de bases das extremidades coesivas permite que DNA de espécies muito diferentes possa ligar-se entre si formando moléculas quiméricas (Lodish et ali., 1995).

Quadro 1

Exemplos de acções de endonucleases de restrição (Adap. Lodish, Table 7.1)

Enzima	Local de reconhecimento (restrição) (só se indica uma hélice dentro de cada DNA)	Número de cortes efectuados sobre o genoma (kb) de			
		Fago_ genoma 50kb	Adenovirus Ad2 36kb	Vírus SV40 5.2kb	Plasmideo pBR322 4.3kb
Alu I	AG _CT	143	158	34	14
Taq I	T _CGA	121	50	1	13
Hph I	GGTGA+5	168	99	4	18
Eco RI	G _AATTC	5	5	1	1
Hind III	A _AGCTT	6	12	6	1
Not I	GC _GGCCGC	0	7	0	0
Sfi I	GGCN <sub>4</sub> _NGGCC	0	3	1	0

Na tecnologia do DNA recombinante ligases purificadas do DNA podem utilizar-se (no sentido 3’\_5’) para unir covalentemente *in vivo* as extremidades dos fragmentos de restrição (Lodish et ali., 1995).

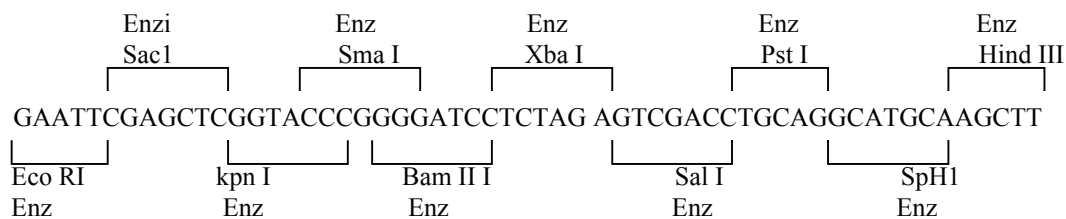
Quando a ligase do DNA e ATP são adicionados aos fragmentos de restrição com extremidades coesivas estas são covalentemente ligadas formando-se uma ligação fosfodiester entre 3’ e 5’.

Nas extremidades não coesivas (como no quadro 1 anterior na Alu I) ambas as hélices de DNA são cortadas exactamente a meio não se gerando pontas coesivas mas sim pontas rombas “blunt” (“flush”), mas também a ligase do DNA extraída do bacteriófago T4 pode ligar estas extremidades, embora exigindo mais elevada concentração do DNA para isso ocorrer.

Os fragmentos de restrição de DNA estranho podem ser facilmente inseridos em plasmidos vectores *in vitro*.

Plasmideos vectores de *E. coli* podem ser construídos com um polilinker, ou seja uma sequência com múltiplos locais de clonagem ou melhor uma sequência sintética que contem uma cópia de vários locais de restrição diferentes como se refere seguidamente na figura 3 (*vide* também a seguir a formação e utilização de DNA sintético, síntese química de oligonucleótidos).

Figura 3 - Sequência de um polilinker (Adap. Lodish, Fig.7.8a).



Plasmídeos vectores contendo um “polilinker” como exemplificado em cima (sequência com múltiplos locais para clonagem) são utilizados frequentemente para produzir plasmídeos recombinantes transportando fragmentos de DNA exógeno.

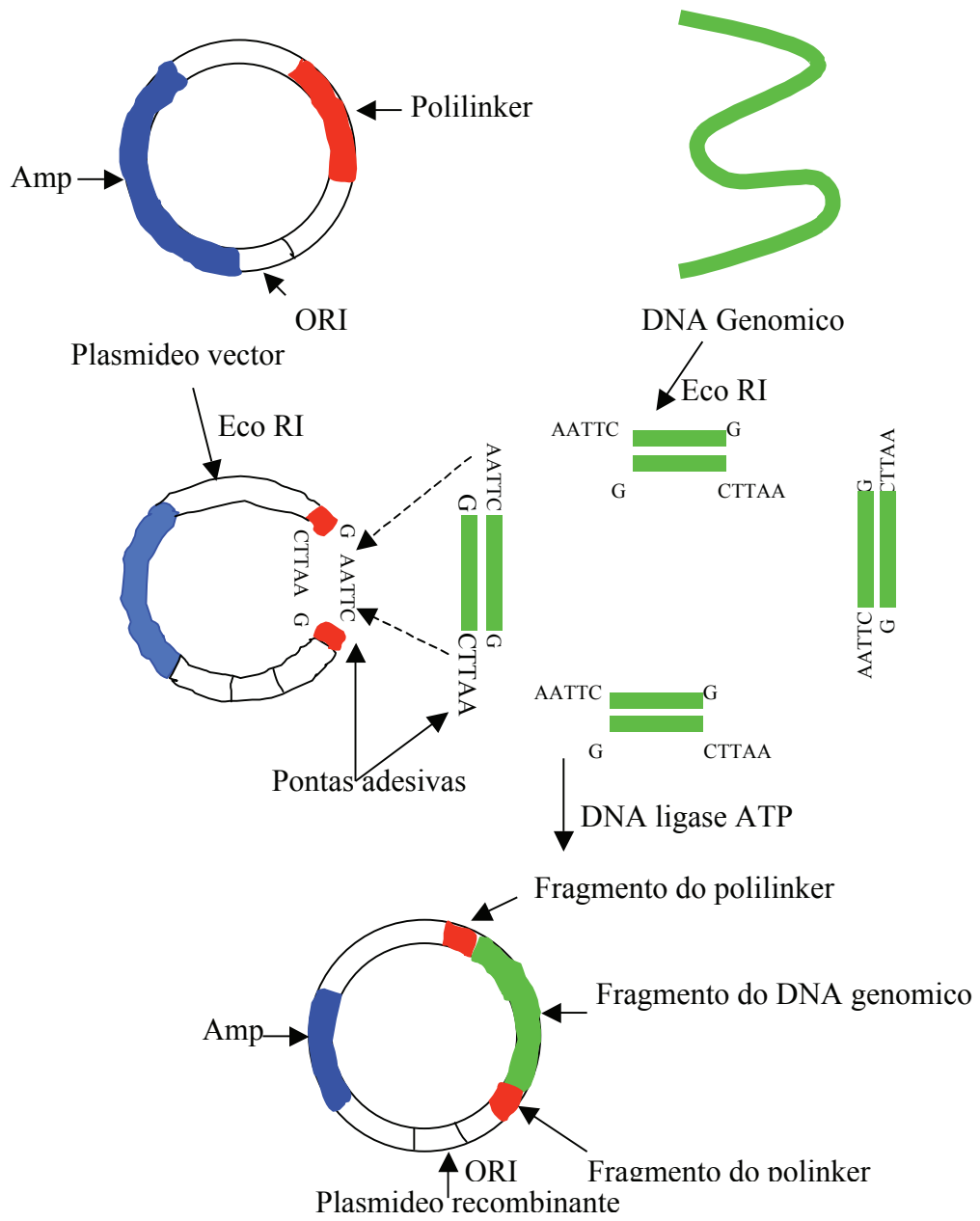
Na sequência de polilinker referido acima mostra-se uma cópia de local de reconhecimento (indicado pelos colchetes ) para cada uma das dez enzimas de restrição indicadas Figura 3).

Os polilinkers são sintetizados por via química e depois inseridos num plasmídeo vector.

Uma das enzimas de restrição cujo local de reconhecimento se encontra no polilinker é utilizada para cortar moléculas do plasmídeo que contem o polilinker e do DNA genómico (estranho) que se pretende clonar produzindo plasmídeos com um único corte e fragmentos de restrição com pontas coesivas complementares (ver na figura 4 seguinte).

Figura 4

Inserção de fragmentos de restrição de um DNA genômico do Eco RI num plasmídeo vector



Adap.Lodish, fig.7.8b)

### 3.5 - Formação e utilização de DNA sintético Síntese química de oligonucleótidos

(Lodish et ali., 1995)

A possibilidade para manipular DNA recombinante foi consideravelmente ampliada com o desenvolvimento de técnicas químicas que permitem sintetizar moléculas de DNA com uma única hélice, com a sequência nucleotídica desejada, contendo até cerca de 100 nucleótidos e que podem ser utilizados para diversas finalidades.

É pois possível a síntese química de moléculas de DNA de hélice simples (ss DNA) com a sequência desejada até cerca de 100 nucleótidos de comprimento, o que tem muito interesse na tecnologia do DNA recombinante.

ssDNA (com uma só cadeia) complementar pode ser também sintetizado e hibridado com o outro inicial, formando DNA com pontas coesivas (*sticky ends*). Tal DNA de cadeia dupla obtido completamente por síntese química pode ser clonado em plasmídeos vectores tal como fragmentos de restrição do DNA preparados a partir de organismos vivos.

DNAs sintéticos podem ser utilizados na sequenciação de DNA e como sondas para identificar clones de interesse, podendo também ser utilizados para substituir sequências de DNA naturais em DNA clonado para estudar os efeitos de mutações específicas.

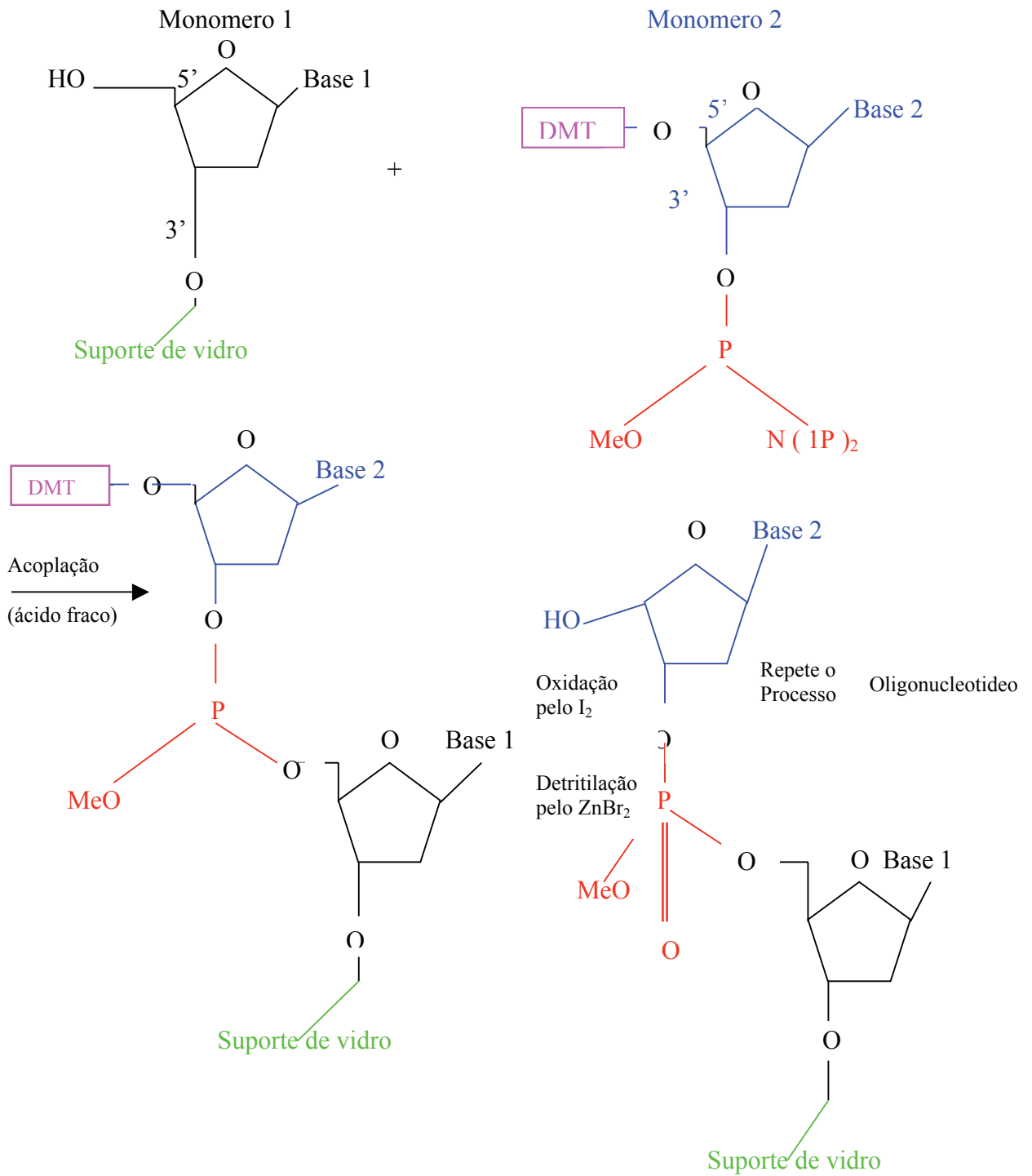
A técnica para a síntese química de oligonucleótidos refere-se seguidamente na figura.

É de salientar que a cadeia cresce no sentido 3' \_ 5' ao contrário do que fazem as DNA polimerases, estando hoje estes processos automatizados.

Na figura 5 o 1º nucleótido (monómero 1) está ligado a um suporte de vidro através do hidroxilo 3' estando o hidroxilo 5' disponível. O nucleótido seguinte da sequência (monómero 2) é derivado pela adição do 4',4-dimeoxitritilo (DMT) ao seu hidroxilo 5', inibindo este grupo para reagir. A adição do grupo fosforamidite diisopropil metilado (a vermelho na figura) altamente reactivo é ligado ao hidroxilo 3' deste monómero 2, e quando os dois monómeros são misturados na presença de um ácido fraco formam a ligação fosfodiester 5' \_ 3' com o fósforo no estado trivalente. A oxidação com o iodo (I<sub>2</sub>) aumenta a valência do fósforo para cinco. A remoção do grupo DMT por detritilação com brometo de zinco (Zn Br<sub>2</sub>) liberta o grupo hidroxilo 5' sendo depois o processo sucessivamente repetido (Lodish et ali., 1995).

Quando a síntese está completa todos os grupos metilo dos fosfatos são libertados ao mesmo tempo em ambiente de pH alcalino tal como é cindida a ligação do monómero 1 ao vidro de suporte.

Figura 5



Síntese química do Oligonucleotidos

(Adaptado de Lodish, Fig.7.9)

Os genes codificantes de proteínas necessárias para o “assemble” da cabeça e da cauda do fago encontram-se referidas tais como aqueles necessários ao ciclo lítico. Algumas regiões ( a vermelho ) do genoma ( actividades lisogénicas ) são substituíveis ou até suprimidas ( a pontado preto ) sem afectar a capacidade do fago para infectar células hospedeiras e montar, os novos viriões, permitindo ainda que no seu local entre os genes J e N se faça a inserção de DNA estranho com até ≈ 25kb, formando-se pois um DNA recombinante que pode formar viriões capazes de se replicar e formar placas liticas nas células hospedeiras de E. coli.

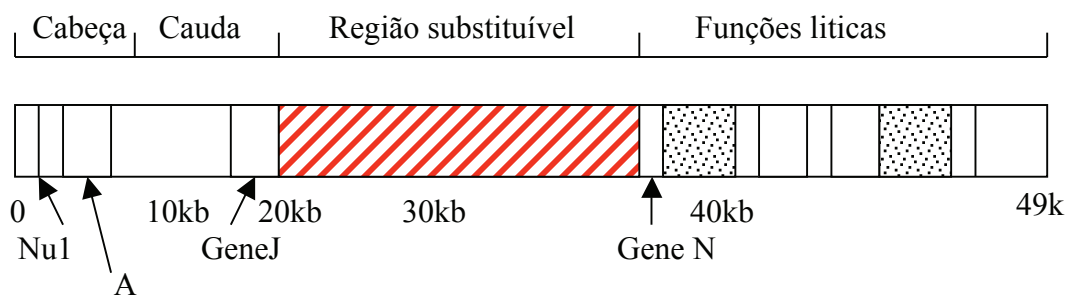
A chave da alta eficiência da clonagem do fago \_ reside na sua capacidade para “montar” os viriões \_ “in vitro”. Com efeito à semelhança do que se passa in vivo, é possível in vitro desenvolver o processo de “montagem” do fago. Para isso utilizamos células de E. coli infectadas com uma mutante do fago \_ que consegue sintetizar as sub-unidades proteicas correspondentes á cabeça do fago (vazia sem o respectivo DNA) e a cauda. Estas células de E. coli são depois lisadas experimentalmente, pois a mutante do fago \_ não possui essa capacidade, sendo preparado um extracto contendo altas concentrações das cabeças vazias e caudas do fago. Quando este extracto é misturado com a proteína A (produzida pelo fago \_ normal não mutado) e DNA \_ recombinante (vide figura seguinte) que contem um local COS (vide em baixo), o DNA é introduzido dentro das cabeças vazias do fago, acoplado-se depois as caudas contidas no extracto produzindo-se assim viriões completos que possuem um DNA \_ recombinante. [Em condições normais o fago \_ não mutado e ao contrário do mutado, clonado tem capacidades para montar viriões \_ in vitro e in vivo. Para isto in vivo monta separadamente dentro das células hospedeiras infectadas, cabeças e caudas do fago a partir das proteínas suas precursoras, enquanto o DNA \_ é produzido na célula

### 3.6 - Vectorios de clonagem do bacteriofago lambda e construção de um repertório ou biblioteca genómica (Library)

Uma colecção de clones \_ que inclua todas as sequências de DNA de uma dada espécie é chamada, a essa colecção, o repertório genómico, ou biblioteca genómica ou “genomic library” e uma vez esta preparada, pode ser rastreada no sentido de encontrar os clones \_ que contem uma dada sequência nucleotídica de interesse.

O bacteriofago \_ pode ser modificado e “montado” in vitro em ordem a ser utilizado como vector de clonagem. Com efeito um \_ virião (uma partícula viral individual) contem genes codificantes das proteínas implicadas nas suas vias lítica e lisogénica (vide figura 6 seguinte).

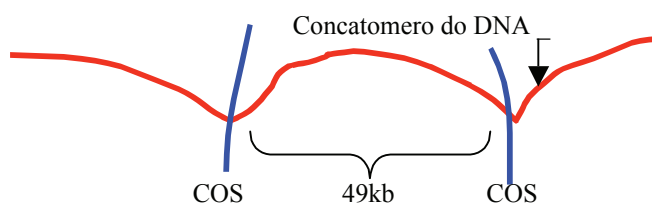
Figura 6  
Mapa simplificado do genoma (tem cerca de 60 genes) do fago  
(Adaptado de Lodish et ali., 1995, Fig.7.10b)



Os genes codificantes de proteínas necessárias para o “assemble” da cabeça e da cauda do fago encontram-se referidas tais como aqueles necessários ao ciclo lítico. Algumas regiões ( a vermelho ) do genoma ( actividades lisogénicas ) são substituíveis ou até suprimidas ( a pontado preto ) sem afectar a capacidade do fago para infectar células hospedeiras e assemblar, os novos viriões, permitindo ainda que no seu local entre os genes J e N se faça a inserção de DNA estranho com até  $\approx 25\text{kb}$ , formando-se pois um DNA recombinante que pode formar viriões capazes de se replicar e formar placas líticas nas células hospedeiras de E. coli.

A chave da alta eficiência da clonagem do fago  $\lambda$  reside na sua capacidade para “assemblar” os viriões  $\lambda$  “*in vitro*”. Com efeito à semelhança do que se passa *in vivo*, é possível *in vitro* desenvolver o processo de “assembly” do fago. Para isso utilizamos células de E. Coli infectadas com uma mutante do fago  $\lambda$  que consegue sintetizar as sub-unidades proteicas correspondentes á cabeça do fago ( vazia sem o respectivo DNA ) e a cauda. Estas células de E. coli são depois lisadas experimentalmente, pois a mutante do fago  $\lambda$  não possui essa capacidade, sendo preparado um extracto contendo altas concentrações das cabeças vazias e caudas do fago. Quando este extracto é misturado com a proteína A ( produzida pelo fago  $\lambda$  normal não mutado ) e DNA  $\lambda$  recombinante que contem um local COS (*vide* figura 7 em baixo) o DNA é introduzido dentro das cabeças vazias do fago, acoplando-se depois as caudas contidas no extracto e produzindo-se assim viriões completos que possuem um DNA  $\lambda$  recombinante. [Em condições normais o fago  $\lambda$  não mutado e ao contrário do mutado, clonado tem capacidades para assemblar viriões  $\lambda$  *in vitro* e *in vivo*, para isto *in vivo* assembla separadamente dentro das células hospedeiras infectadas, cabeças e caudas do fago a partir das proteínas suas precursoras, enquanto o DNA  $\lambda$  é produzido na célula hospedeira na forma de longas moléculas multiméricas de DNA chamadas concatómeros que resultam de várias cópias múltiplas do genoma vírico ligadas extremidade com extremidade e separadas por sequências nucleótídicas específicas chamadas locais COS (Fig. 7). Duas proteínas do fago  $\lambda$  a Nu1 e a A (*vide* figura 6 anterior) ligam-se aos locais COS e dirigem a inserção do DNA ( $\approx 50\text{kbp}$ ) situado entre dois locais COS consecutivos para uma cabeça do fago preassemblada, acabando depois a cabeça por assemblar, tal como a cauda que se agregam originando um virião completo infeccioso.

Figura 7

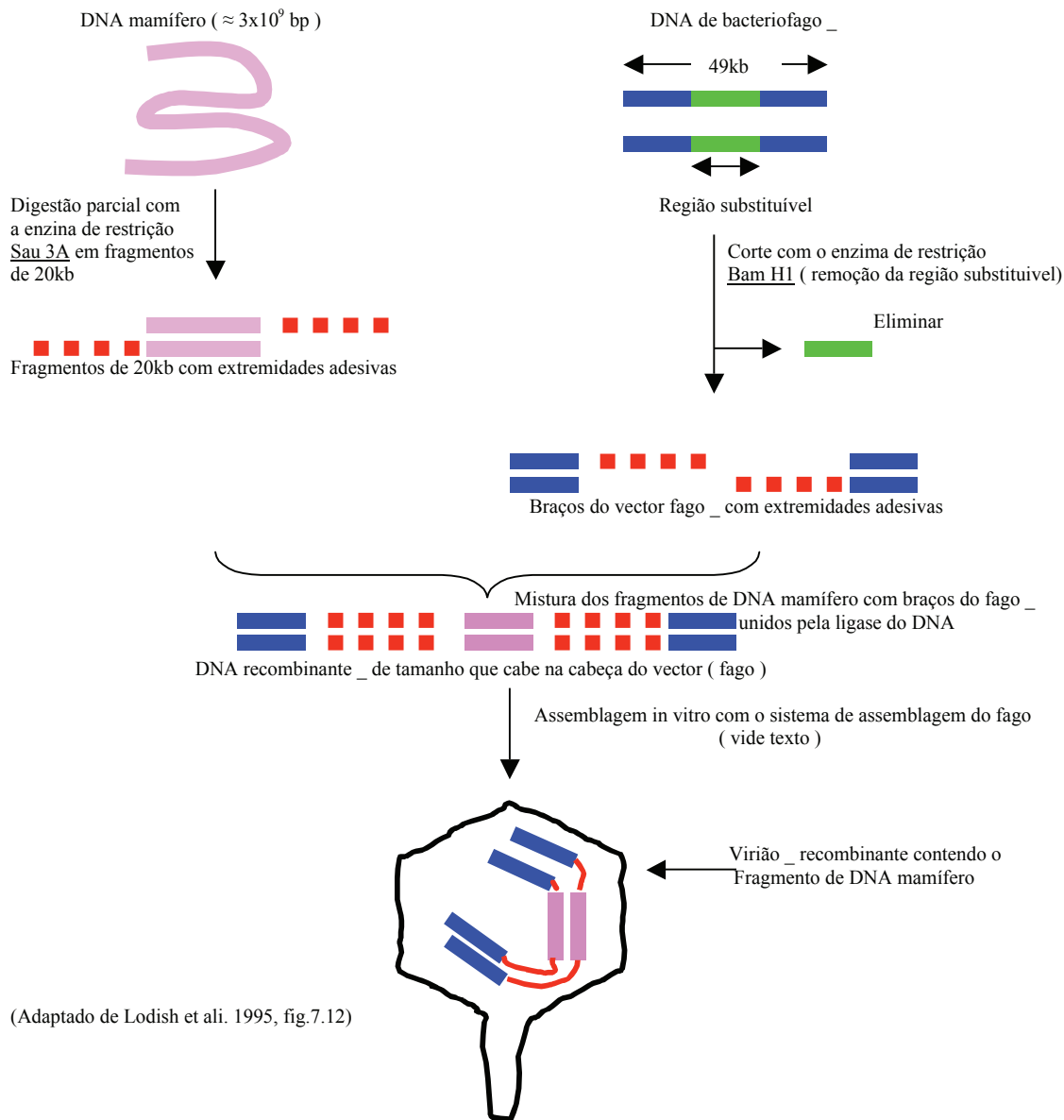


Através da utilização pela  $\lambda$  clonagem é possível preparar repertórios genómicos quase completos de organismos superiores como se esquematiza seguidamente (Figura 8)



Figura 8.

Construção de um repertório ou biblioteca ( library )  
Genômica de DNA mamífero com um vector bacteriofago



Como se esquematizou na figura anterior o DNA do fago \_ tem primeiro que ser removido da sua região substituível (a verde na figura) o que é feito com o concurso da enzima de restrição Bam HI, produzindo os braços do vector e uma parte a eliminar.

O DNA mamífero (a rosa na figura) de que se pretende construir o repertório genómico é parcialmente digerido com a enzima de restrição Sau 3A.

Os dois enzimas de restrição utilizados Sau 3A e Bam HI produzem fragmentos de restrição com extremidades adesivas complementares entre si (linhas verticais vermelhas na figura).

Os braços do vector  $\lambda$  e os fragmentos de restrição do DNA mamífero de cerca de 20kbp são misturados, unidos pela acção de ligase e “empacotados” *in vitro* para produzir viriões de fago  $\lambda$  recombinante, viriões estes depois semeados em bactérias *E. coli*.

Podem ser produzidos diversos fragmentos de restrição do DNA mamífero (neste exemplo) por uma variedade de enzimas de restrição e clonados em vectores  $\lambda$ , sendo assim construídos múltiplos  $\lambda$  vectores.

Para nos apercebermos da complexidade da elaboração e construção de um repertório genómico de um animal superior transcrevemos um exemplo referido por Lodish e col in Molecular Cell Biology (1995, 3ª Ed) em relação a esta matéria.

O tamanho de um repertório genómico de um dado organismo depende da quantidade de DNA do genoma haplóide deste organismo. Se, por exemplo, o genoma contiver  $3 \times 10^9$  pares de bases e se for cindido em fragmentos de 20kbp para serem inseridos num vector  $\lambda$ , podem ser necessários á roda de  $1,5 \times 10^5$  viriões do fago  $\lambda$  recombinante diferentes para termos constituído um repertório completo.

Como os fragmentos de restrição são incorporados nos fagos ao acaso, são necessários cerca de  $10^6$  fagos recombinantes para assegurar que cada região do DNA do organismo em questão tem 90 a 95% de hipóteses de ser incluído no repertório.

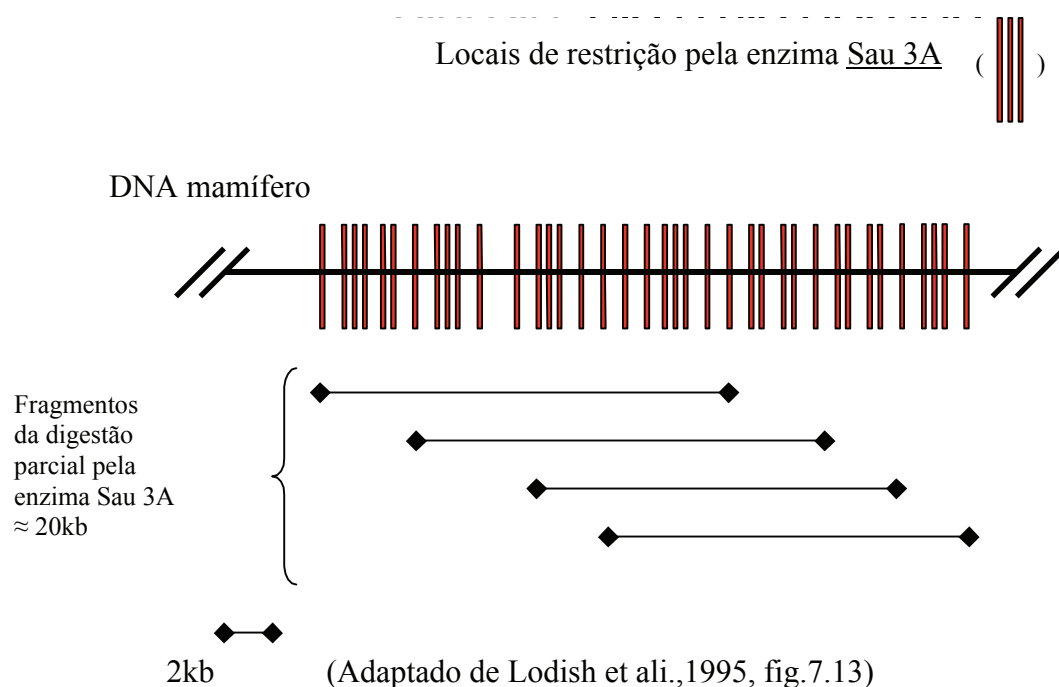
Cada placa de lise originada pelo fago  $\lambda$  recombinante sobre a cultura bacteriana (*E. coli*) em placa de Petri contem grande número de viriões recombinantes, ou seja, dos fragmentos de restrição do DNA clonados que podem ser identificados por exemplo por métodos de hibridação.

É possível detectar especificamente placas de lise muito pequenas podendo em cada placa de Petri ser restreadas  $5 \times 10^4$  placas de lise. Portanto, cerca de 20 a 30 placas de Petri nestas condições, cada uma contendo cerca de  $5 \times 10^4$  placas  $\lambda$ , são suficientes para representar um genoma de  $3 \times 10^9$ bp, enquanto que o rastreio de  $10^6$  plasmideos recombinantes tendo inserida a totalidade do mesmo genoma ( $3 \times 10^9$ bp) necessitaria de cerca de 5.000 placas de Petri, uma vez que apenas cerca de 200 colónias de *E. coli* são detectáveis em cada placa.

Uma região particular de 20kbp, num genoma contendo  $3 \times 10^9$ bp, pode ocorrer cerca de uma vez em cada  $1,5 \times 10^5$  viriões recombinantes. Contudo são necessários cerca de  $10^6$  fagos recombinantes para assegurar que fragmentos que representem todo o genoma mamífero ( $\approx 3 \times 10^9$ bp) tenham 90 – 95% de hipóteses de ser incluídos no repertório genómico.

Em ordem a aumentar as probabilidades de que todas as regiões de um dado genoma sejam clonadas com sucesso e representadas no repertório genómico  $\lambda$ , o DNA desse dado genoma deve ser parcialmente digerido em ordem a produzir fragmentos de restrição sobreponíveis de  $\approx 20$ kbp (*vide* figura 9 seguinte). Num grande repertório genómico  $\lambda$  construído a partir de tais fragmentos de restrição, uma sequência específica do DNA genómico pode estar contida em vários clones que se sobrepõem, tal como se refere no esquema seguinte.

Figura 9  
Produção de fragmentos de restrição  
sobreponíveis por digestão parcial do DNA



A enzima de restrição Sau 3A reconhece sequências GATC e produz fragmentos com pontas coesivas com a mesma sequência do lado 5'.

A linha de cima representa uma região de DNA hipotética com os locais de restrição pela Sau 3A. A digestão parcial desta região do DNA originará uma variedade de fragmentos ( $\approx$  20kbp de comprimento) que se sobrepõem em parte aumentando a probabilidade de todas as sequências do DNA estarem representadas no repertório genómico \_.

Fragmentos de DNA de comprimento maiores do que 20kbp podem ser clonados em cósmidos e outros vectores como o P1 ou o YAC (cromossoma artificial de levedura) (Quadro 2)

Quadro 2.

Tamanho máximo do DNA que pode ser clonado em vectores (Adpt. Lodish ,Table 7.2)

Tipo de vectores	Comprimento do DNA clonado ( kbp)
Plasmido	20
Bacteriofago _	25
Cósmido	45
P1 vector	100
YAC	1.000

Os BAC (*bacterial artificial chromosome*) são uma estrutura cromossoma – “like” construída por engenharia genética e que transporta DNA genómico para ser clonado.

Os BAC “clone” ou vector de um cromossoma bacteriano artificial transporta um DNA inserido com cerca de 100-200kbp. A maioria dos clones com grandes inseridos ou insertos sequenciados no projecto do HGP (human genome project) que referimos em outro trabalho foram BAC clones.

### **3.7 - Construção de um repertório, sondas e identificação de clones específicos**

(Lodish et ali., 1995)

Nas células eucariotas de organismos superiores muitos genes são transcritos em mRNA apenas em certos tipos de células, sendo clássicos os mRNA para as globinas que apenas são encontrados nos reticulocitos (células precursoras dos eritrocitos) e os mRNA para a albumina que apenas são produzidos pelas células hepáticas.

Como vimos as cópias de DNA obtidas de mRNA são chamadas DNAs complementares (cDNAs).

Alguns cDNAs e oligonucleótidos sintéticos são utilizados como sondas, pois a identificação de clones específicos pela técnica da hibridação em membrana depende da disponibilidade em sondas marcadas. O método para preparar uma sonda depende de quando o gene que nos interessa codifica uma proteína que é expressa em altas taxas ou em baixas taxas.

Em certos casos a taxa de um dado mRNA particular é suficientemente abundante num dado tecido o que permite facilmente isolá-lo e utilizá-lo para preparar uma sonda, no entanto em muitos casos o gene que nos interessa codifica uma proteína que se exprime com baixas taxas de concentração, sucedendo que é o próprio mRNA que é muito escasso para ser isolado e marcado como sonda. Nestes casos sondas oligonucleotídicas para estes genes e para estas proteínas que se encontram pouco abundantemente podem ser sintetizadas quimicamente baseadas na sequência dos ácidos aminados que integram essa proteína.

Uma sonda contendo cerca de  $\approx 20$  nucleótidos é suficiente e apenas uma pequena porção (6 ou 7 ácidos aminados) da proteína total necessitam ser sequenciados. Para este efeito, para preparar uma sonda nestas condições a proteína que interessa é purificada e depois digerida em péptidos específicos, sendo depois as sequências N-terminais de alguns destes péptidos determinadas. Depois, baseado no código genético, sondas oligonucleotídicas codificando a sequência do péptido desejado são sintetizadas e marcadas considerando a degenerescência do código genético. Assim são de evitar as escolhas das sequências que contenham Arg, Leu ou Ser pois seis codões diferentes codificam cada um destes ácidos aminados. Os melhores ácidos aminados para fazer tais sondas são os que têm apenas um codão.

Na prática muitas vezes utilizam-se uma mistura de sondas de uma só vez.

### **3.8 - Informática: a recolha, distribuição e análise dos dados sobre a sequência do DNA**

Hoje são conhecidas grande número de sequências do DNA de várias espécies e no futuro espera-se que muitas mais venham a ser conhecidas.

Para guardar e distribuir este enorme volume de dados da análise e sequência de DNA e de proteínas são necessários meios computacionais existindo já imensos centros por esse mundo fora que recolhem essa informação e a disponibilizam gratuitamente ou não, consoante as instituições, a todos os interessados (*vide* quadro abaixo).

É possível assim comparar as sequências obtidas para estabelecer semelhanças, ou sequências homólogas.

Regiões do DNA codificando proteínas podem ser convertidas em sequências de ácidos aminados e em virtude da degenerescência do código genético, proteínas relacionadas entre si exibem por vezes maior homologia do que os próprios genes que as codificam.

Proteínas com funções semelhantes por vezes contêm sequências homólogas o que pode facilitar a caracterização de genes clonados.

Fontes disponíveis de bases de dados de sequências e outras informações genômicas relevantes referidas na Nature vol.409 13 february 2001, pag. 876

<http://genome.ucsc.edu/>

University of California at Santa Cruz

Contains the assembly of the draft genome sequence used in this paper and updates

<http://genome.wustl.edu/gsc/human/Mapping>

Washington University

Contains links to clone and accession maps of the human genome

<http://www.ncbi.org>

EBI/Sanger Centre

Allows access to DNA and protein sequences with automatic baseline annotation

<http://www.ncbi.nlm.nih.gov/genome/guide/>

NCBI

Views of Chromosomes and maps and loci with links to other NCBI resources

<http://www.ncbi.nlm.nih.gov/genemap99/>

gene map 99: contains data and viewers for radiation hybrid maps of EST-based STSs

<http://compbio.ornl.gov/channel/index.html>

Oak Ridge National Laboratory

Jawa viewers for human genome data

<http://hgrep.ims.u-tokyo.ac.jp/>

RIKEN and the University of Tokyo

Gives an overview of the entire human genome structure

<http://snp.cshl.org/>

The SNP Consortium

includes a variety of ways to query for SNPs in the human genome

<http://www.ncbi.nlm.nih.gov/Omin/>

Online Mendelian inheritance in Man

Contain information about human genes and disease

<http://www.nhgri.nih.gov/ELSI/> and <http://www.ornl.gov/hgmis/elsi/elsi.html>

NHGRI and DOE

Contains information, links and articles on a wide range of social, ethical and legal issues

### **3.9 - Locais do início da transcrição podem ser mapeados pela chamada protecção S1 e pela extensão com um primer**

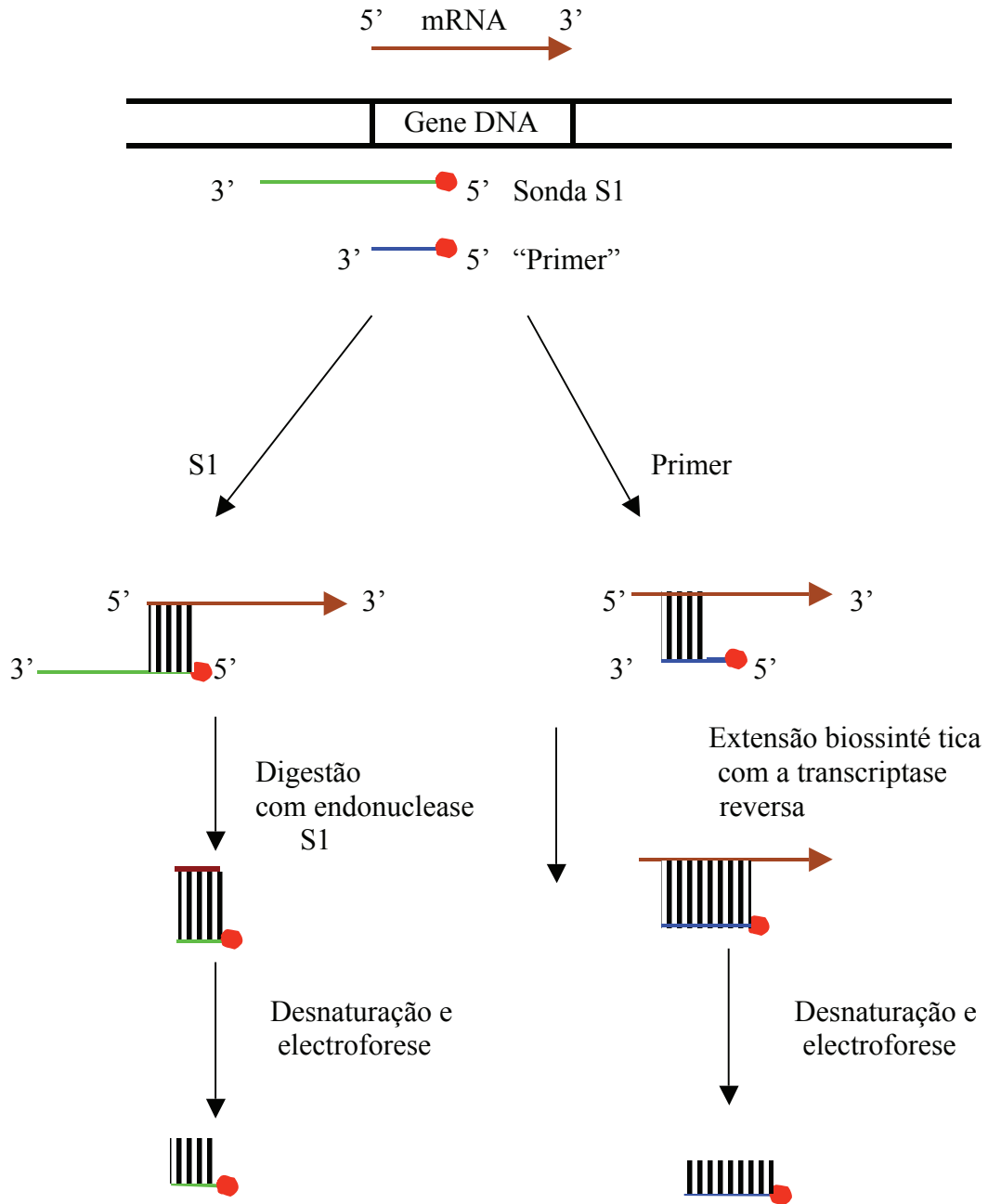
(Lodish et ali., 1995).

É possível analisar ácidos nucleicos específicos contidos em misturas complexas. Por exemplo: obtido um DNA clonado este pode ser utilizado como sonda para detectar ácidos nucleicos complementares em misturas complexas contendo todo o DNA e RNA celular, permitindo-se também localizar as regiões do DNA que codificam mRNA específico, assim como os locais de início da transcrição.


Os elementos reguladores que controlam a transcrição de um gene estão habitualmente situados próximo do local do início da transcrição ( +1 ). Assim o mapeamento deste local de início de transcrição de um dado gene pode ajudar muito a identificar as sequências de DNA que regulam essa transcrição.


Dois métodos são utilizados para mapear a extremidade 5' de um dado mRNA sobre o DNA que lhe é complementar, a chamada protecção S1 ( endonuclease de *Asp. Oryzae*) e a extensão com *primer*, sendo necessário em qualquer um deles identificar primeiro a região geral de um DNA que inclui o local do início da transcrição que pretendemos conhecer o que pode ser feito pela análise pelo *Northern blot* ou por protecção pela nuclease utilizando vários fragmentos de restrição clonados (Figura 10).

Figura 10  
Métodos para mapear o local de início  
para a transcrição de um dado gene  
( numa região de DNA de sequência conhecida )(Adapt. Lodish ,fig. 7.34)



Na pagina anterior o fragmento de DNA que contem o gene cujo ponto de inicio de transcriçãõ queremos conhecer encontra-se em cima e o correspondente mRNA está representado a castanho \_ .

Um fragmento de DNA de cadeia simples representado a verde  (marcado com bola vermelha) na extremidade é uma sonda na técnica de mapeamento S1. Nesta técnica (á esquerda na figura) a sonda é hibridada com o mRNA e os ácidos nucleicos não emparelhados são digeridos com endonucleases S. A desnaturaçãõ subsequente deixa um fragmento da sonda marcada cujo comprimento marca exactamente a distância do nucleótido de inicio de transcriçãõ do mRNA para o nucleótido que hibrida com a terminaçãõ 5' da sonda de DNA marcada.

Na outra técnica (lado direito da figura) é utilizado um “primer” oligonucleotídico (azul marcado (bola  vermelha) na extremidade que vai ser utilizada na técnica de extensãõ com primer. Nesta técnica o curto oligodesoxinucleótido (de cerca de 20 nucleótidos) sintetizado e marcado na extremidade é hibridado com o mRNA e estendido pela acçãõ de transcriptase reversa até atingir o primeiro nucleótido do mRNA.

### **3.10 - Substituiçãõ de genes e animais transgênicos**

#### **Terapia por genes**

(Lodish et ali., 1995; Picciotto e Wickmam, 1998)

Existem uma sêrie de técnicas capazes de induzir in vitro mutações em diversos genes, podendo esses genes mutados ser introduzidos em diversos seres vivos (por exemplo leveduras, ratos, etc.) substituindo os genes endógenos selvagens ou nativos.

Tal “knockout” do gene alvo constitui um poderoso meio para avaliar da funçãõ biológica do gene clonado pois a mutaçãõ de um gene específico e a substituiçãõ de um gene normal pelo mutado pode permitir conhecer a sua funçãõ *in vivo*.

Genes estranhos podem também ser introduzidos em organismos unicelulares ou pluricelulares e passados estavelmente de uma geraçãõ a outra numa linha transgênica. O DNA tem sido injectado em um de dois pronúcleos de um óvulo fertilizado e múltiplas cópias de genes em tandem incorporados no genoma. Em outros casos o gene que interessa é inserido num elemento transposável de um plasmido sendo este injectado num embrião adequado onde o DNA é transposto do plasmido para o genoma de uma forma ao acaso. Nestas técnicas pode não ocorrer a substituiçãõ do gene endógeno, mas sim a integraçãõ de cópias adicionais sendo esse genes chamados transgenes e os seres que os receberem transgênicos.

A expressãõ transgênico é utilizada para designar organismos que contêm nos seus cromossomas cópias de genes integrados estavelmente, ou construções de genes de outras espécies que não se encontram habitualmente nos organismos hospedeiros.

Um animal transgênico transporta pois um gene estranho que foi deliberadamente introduzido no seu genoma.

O gene estranho foi constituído utilizando metodologias do DNA recombinante.

Além de gene estrutural, o DNA recombinante habitualmente inclui outras sequências que facilitam a sua incorporaçãõ no DNA hospedeiro e que facilitam ainda que sejam expressos correctamente pelas células do hospedeiro.



Ovinos e caprinos transgênicos têm sido produzidos exprimindo proteínas estranhas no seu leite, tal como frangos transgênicos capazes de sintetizar proteínas humanas nas brancas (albumen) dos seus ovos.

Dois métodos são utilizados por exemplo para a produção de ratos transgênicos sendo a primeira mais versátil que a segunda;

Metodo 1 – a transformação de células estaminais ou “Stem Cells” embrionicas (células ES) crescendo em cultura com o DNA desejado.

Metodo 2 – a injeção do gene desejado no pró-núcleo de um óvulo fertilizado de rata.

No metodo 1 as “Stem Cells” embrionicas são colhidas da massa celular interior de blastocitos de ratos os quais podem ser desenvolvidos em cultura e estes o gene pleno potencial para produzir células do animal adulto inclusive gâmetas.

Os passos a desenvolver são os seguintes:

Construir o DNA desejado utilizando as metodologias do DNA recombinante, contendo este o gene estrutural desejado mais o vector do DNA que facilite a entrada de moléculas a serem inseridas nas moléculas de DNA hospedeiras e ainda as sequências promotoras e facilitadoras (enhancer) que facilitem a expressão do gene pelas células hospedeiras.

Depois é necessário transformar as “Stem Cells” embrionicas em cultura, em ordem a que estas incorporem algum do DNA recombinado a que são expostas.

Seguidamente é necessário seleccionar as células que foram realmente transformadas pela incorporação desse DNA.

Estas células seleccionadas são injectadas na massa celular interna de blastocitos de rato, procedendo depois à transferência de embriões. Para isto é preparada uma rata pseudo-grávida (acopulando a fêmea com um macho vasectomizado) transferindo-se depois os embriões para o interior do seu útero e aguardando que estes implantes tenham sucesso originando crias saudáveis (costuma ser de 1/3 este êxito).

Por fim é necessário testar a descendência obtida retirando por exemplo um pequeno fragmento de tecido da cauda e verificando se o DNA contem o gene incorporado desejado. Cerca de 10 – 20% da descendência estará nesta condições produzindo genes heterozigotos.

Por último é necessário cruzar dois destes animais heterozigotos e rastrear a descendência destes para que 1:4 seja homozigoto para o transgene seleccionado e cruzando-se depois estes animais para obter um estirpe transgénica.

No método 2 (metodo do pró-núcleo) o DNA é preparado como no metodo 1 anterior e depois é necessário transformar os óvulos fertilizados. Para isso são preparadas ratas com superovulação (as ratas em estro são injectadas com hormonas inclusive hCG para aumentar o número de células fertilizadas disponíveis podendo assim obter-se até 35 óvulos.

Estas fêmeas em superovulação são depois cruzadas com machos inteiros e os ovos fertilizados (zigotos pré-implantados) são arrastados a partir dos ovidutos e recolhidos.

Pouco tempo após a fertilização, tanto o pró-núcleo macho como o pró-núcleo fêmea existem separados um do outro, sendo o primeiro maior que o segundo sendo facilmente identificáveis ao microscópio.

O DNA transgene preparado é então microinjectado no pró-núcleo macho.

Após a fusão dos dois pró-núcleos o zigoto divide-se por mitose e forma duas células, quatro células, oito células.

São depois transplantados 20 a 40 zigotos para o oviduto de uma mãe pseudo-grávida (que foi cruzada na noite anterior com um macho vasectomizado. Três semanas depois a mãe pseudográvida dará uma ninhada de ratos, estes são rastreados a partir de um fragmento de tecido da cauda e a partir das linhas heterozigotas obtém-se linhas homozigóticas.

Resumindo são colhidos óvulos fertilizados recentemente antes da cabeça do espermatozóide se tornar num pró-núcleo. São nestes óvulos injectados os pró-núcleos macho com o DNA preparado e quando o pró-núcleo se funde para formar o núcleo diplóide do zigoto, deixa-se o zigoto dividir-se por mitose até se formar um embrião de duas células. Segue-se a implantação destes embriões em fêmeas pseudográvidas e proceder como no método 1.

Estão também descritas metodologias para obtenção de ovinos, frangos, cabras, vacas e porcos transgênicos.

Os transgênicos podem ser expressos ou não, animais transgenicos de diferentes especies tem sido produzidos através da introdução de DNA com genes estranhos clonados, em oócitos fertilizados de animais dadores ou por micro-injecção em pró-núcleos.

Também aqui os oócitos são depois transferidos para o útero de uma fêmea recipiente pseudogravida e desenvolvidos até ao seu termo, ou então utilizam-se em vez de oocitos, “Stem Cells” embrionicas, podendo estas ser manipuladas *in vitro* antes de utilizadas para gerar transgênicos. Estas “Stem Cells” manipuladas podem depois ser agregadas com células no estado de morulas ou injectadas em blastocistos, qualquer um deles sendo depois implantados em fêmeas pseudográvidas onde se desenvolvem até ao seu termo.

Na descendência resultante pode ser testada a presença do transgene através de “Southern blotting” ou pela Polimerase Chain Reaction (PCR). Os animais obtidos que são hemizigotos para o transgene, podem ser cruzados para obter linhas homozigóticas puras, e o sucesso da manipulação pode ser confirmado *in vitro* antes das células manipuladas serem utilizadas para gerar novos animais transgênicos vivos.

Os animais gerados a partir de células manipuladas contêm o gene estranho em todas as células somáticas e também nas células germinativas se o gene estranho for integrado no genoma da célula recipiente antes da primeira divisão celular.

A integração numa fase mais tardia origina mosaicos geneticos de celulas normais com um genoma normal e celular com o genoma modificado.

Referem-se seguidamente as características de alguns animais transgênicos obtidos em diversas espécies animais (Quadro 3).

Quadro 3-Construção de animais transgenicos de diferentes especies animais  
Animais transgenicos de diversas espécies

Especie animal	Tempo de gestação	Tamanho da ninhada	Características da fertilidade	Metodo	Frequência da integração do DNA	aplicação
Ratinho	~20 dias	5-10	6-8 semanas	EI, ES	~20%	Investigação base
Rato	~22 dias	5-10	9-13 “	EI, ES?	~20%	Modelos p/doença
Coelho	~30 dias	5-10	~30 “	EI	~15%	“
Suino	4 meses	9-10	7-8 meses	EI, ES?	~10%	Xenotransplante Estudo de farmacos
Ovino	5 meses	1-2	7-8 “	EI	~1%	“
Cabra	5 meses	1-2	7-8 “	E1	~1%	“
Vaca	9 meses	1-2	~12 “	E1	~5%	“

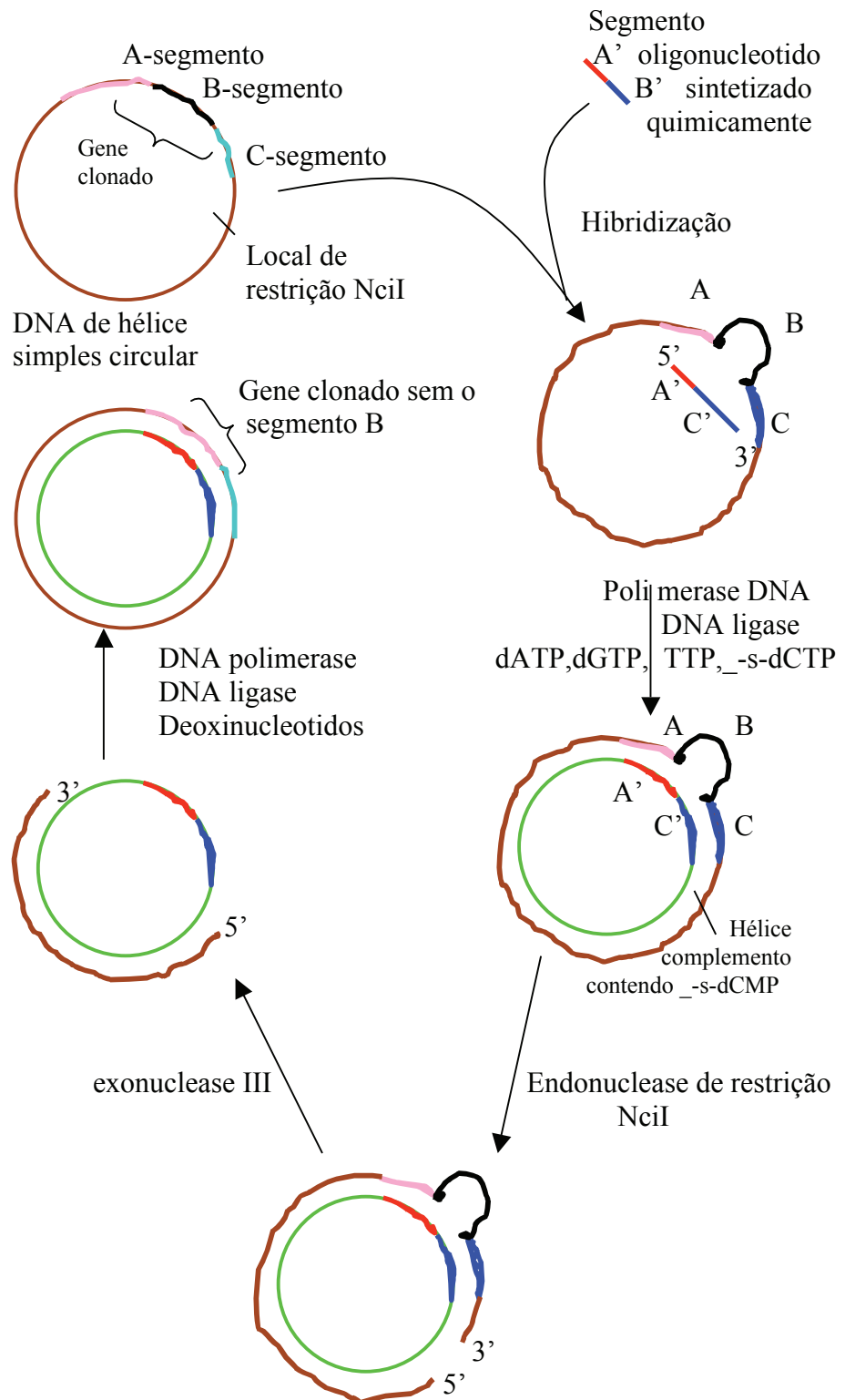
EI= injeção de embrião ES=manipulação de “Strem Cells” embrionicas

(Adaptado de Transgenic:<http://www.copewithcytokines.de/cope.cgi?008161>)

Sequências específicas de genes clonados podem ser alteradas in vitro e depois introduzidas em organismos experimentais ou em células adequadas. Nas técnicas mais frequentes para conseguir este efeito utilizam-se oligonucleótidos com a sequência desejada como mutagénios, gerando neles durante a sua síntese química as deleções, inserções ou mutações pontuais pretendidas mostrando-se na figura 11 seguinte como pode ser alcançada por exemplo uma mutação por deleção.

Figura 11

Mutagenese in vitro de genes clonados com  
Oligonucleotidos obtidos por síntese química



(Adaptado de Lodish et al., 1995, fig.8.32)

Na figura anterior o gene clonado constituído por três segmentos A, B e C é mutagenizado pela deleção do segmento B. Para isso foi sintetizado quimicamente um oligonucleótido com os segmentos A' e C' híbridos de A e C que depois era hibridado com o DNA de cadeia simples que continha o gene que se queria suprimir ou seja (B). Após hibridação o gene a suprimir não hibridado forma uma ansa, enquanto o oligonucleótido sintetizado quimicamente forma um *primer* para a síntese da cadeia complementar (a verde na figura) tendo no entanto além dos nucleótidos normais um nucleótido substituído com enxofre (\_-s-dCMP). O heteroduplex gerado é tratado com a enzima de restrição NciI que não consegue cindir os nucleótidos com enxofre da hélice respectiva; mas origina um simples corte apenas na outra hélice que depois é hidrolizada pela exonuclease III que remove todo o gene clonado. Seguidamente a actuação da DNA polimerase, DNA ligase e deoxinucleótidos incorporados permite a reconstrução do gene clonado (sem o segmento B-knockout) que pode depois ser posteriormente introduzido em células eucariotas ou procaríotas.

O DNA pode ser transferido para células eucariotas de diversas formas.

Muitas células podem tomar o DNA a partir do meio em que se encontram, ou então o DNA pode ser injectado no núcleo de células em cultura ou em embriões em desenvolvimento. Na electroporação as células submetidas a um breve choque eléctrico de vários milhares de volts são transitoriamente permeáveis pelo DNA.

Após internamento do DNA no interior das células provavelmente as enzimas de reparação do DNA e de recombinação juntam os fragmentos de DNA com os cromossomas do hospedeiro em nenhum sítio específico. Como apenas uma pequena quantidade de células é receptiva ao DNA estranho é necessário identificar as células transgénicas sendo pois necessário introduzir um marcador no DNA exógeno.

Nos ratos Knockout a integração específica do DNA exógeno no genoma do rato ocorre por recombinação homóloga mas com uma frequência muito baixa. Pelo contrário a frequência da integração ao acaso em locais não homólogos é muito alta.

Na terapia pelo genes os obstáculos são de várias naturezas. Assim como muitos dos genes associados com doenças genéticas são apenas expressos em certos tipos de tecidos ou células, são necessários métodos que selectivamente introduzam os genes onde se deseja (células sanguíneas, tecido pulmonar, muscular, etc.).

Por outro lado há genes muito grandes de difícil introdução quer nas células somáticas quer nas células germinativas.

De qualquer forma há experiências de terapias pelos genes com êxito como por exemplo se refere no quadro seguinte.

#### Produtos obtidos pela tecnologia do DNA recombinante e que são utilizados em terapias nos humanos e animais

- Insulina (tratamento da diabetes)
- Factor VIII, factor IX (tratamento de hemofilias)
- Hormona do crescimento humano, bovino (produção de leite) e suíno (produção de carne magra)
- Eritropoietina (EPC) (tratamento de anemias)
- Interferões (3 tipos)

- Diversas interleucinas
- GM-CSF (factor estimulador de colónias granulocito-macrofago)
- TPA (actividade do plasminogénio dos tecidos para dissolução de coágulos sanguíneos)
- ADA (admosina desamina) (tratamento de imunodeficiências)
- Angiostatina e Endostatina (tratamento anti-canceroso)
- Hormonas paratiróide

Hoje as atenções centram-se mais na terapia de genes somáticos nos quais genes defectivos são corrigidos em células específicas ou órgãos mais do que nos óvulos, nos espermatozóides ou embriões precoces.

Por exemplo, *in vivo* são recolhidas células de um doente, inserido nelas o novo gene replicando-se seguidamente as células. O direccionamento dos genes para os tecido alvo é feito habitualmente infectando estes com vírus que contem o novo gene, o que implica que o gene novo tem primeiramente que ser clonado e sequênciado em ordem a caracterizar se a sua forma era correcta ou não, depois tem que se verificar se a introdução do gene para o alvo era eficiente e inócua e por ultimo se ele se exprima no seu novo local.

Além da utilização dos vírus como vectores dos novos genes para terapias também se podem utilizar a injeção directa, a microprecipitação com cálcio e a electroporação, sendo no entanto os vírus de melhor utilização, sobretudo os vírus derivados de retrovirus murinos embora decorram sempre riscos potenciais. Nos humanos têm sido realizados ensaios de terapia com genes em situação de deficiência em deaminase da adenosina, na fibrose quística e na hipercolesterolemia familiar, não referidos no quadro anterior.

Também se tem tentado modificar as características do leite produzido por certas espécies animais, inclusive bovinos, introduzindo determinado DNA recombinante veiculado por fagos no tecido mamário, durante o periodo de seca, vindo a desencadear acções na lactação seguinte. Se por um lado a tecnologia do DNA recombinante permite a clonagem, identificação e caracterização de genes e a sua expressão com a produção de proteínas revelantes sobretudo com finalidades terapêuticas, por outro lado a engenharia de proteínas pode permitir a redução da sua antigenicidade, prolongar a sua semi-vida, aumentar a sua afinidade.

Para ultrapassar problemas de imunogenicidade desencadeados pelos anti-corpos, estão hoje em franco desenvolvimento tecnologias para a selecção de anti-corpos recombinantes que permitem por exemplo tornar mais parecidos com os animais a que se destinam, anticorpos monoclonais (mAb) de determinadas espécies dadoras nos quais foram introduzidas sequências peptídicas dos anticorpos das especies animais receptoras diferentes a que se destinam com a ajuda de fagos filamentosos adequados.

### 3.11 - Bibliografia

Ausubel, F. M. et al. 1987. Current Protocols in Molecular Biology. Wiley  
Sambrook, J., E. F. Fritsch, and T. Maniatis, 1989. Molecular Cloning. Cold Spring Harbor Laboratory.

Watson, J. D., et al., 1992, Recombinant DNA, 2d ed. Scientific American Books.  
Caruthers, M. H. 1985 Gene synthesis machine: DNA chemistry and its uses. Science 230:281-285.

Jendrisak, J., R. <sup>a</sup> Young, and J. D. Engel, 1987. Cloning cDNA into Agt10 and Agt11. Methods Enzymol. 152:359-385

Singh, H., et al. 1988. Molecular cloning of an enhancer binding protein : isolation by screening of an expression library with a recognition site DNA. Cell 52:415-423

Davies, K. E., ed. 1988. Genome Analysis: A Practical Approach. Oxford: IRL Press

Nelson, M., and M. McClelland. 1992. The effect of site-specific methylation on restriction-modification enzymes, Nucl. Acids Res. 20 (Suppl.):2145-2157.

Roberts, R. J. 1992, Restriction enzymes and their isoschizomers. Nucl. Acids Res. 20 (Suppl.):2167-2180.

Carleg. F., M. Frank, and M. V. Olson. 1986. Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. Science 232:65-70.

Sanger, F. 1981. Determination of nucleotide sequences in DNA. Science 214:1205-1210.

Elroy-Stein, O., T. R. And B. Moss. 1989. Cap-independent translation of mRNA conferred by encephelomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. Proc. Nat'l. Acad. Sci. USA 86:6126-6130

Luckow, V. A., and M. D. Miller. 1989. High level expression of nonfused foreign proteins with Autographa californica nuclear polyhedrosis virus expression vectors. Virology 170:31-39

Studier, F. W., and B. A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189:113-134.

Erlich, H., ed. 1992. PCR Technology: Principle and Applications for DNA Amplification. W. H. Freeman and Company.

Davis, I. G., Kuehl. W. M., and Battey. J. F. 1994 Basic in Molecular Biology, 2<sup>nd</sup> ed. Norwalk. CT: Appleton & Lange.

Jaenisch. R. 1988 Transgenic animals. Science 240:1468.

Marshall. E. 1995 Gene therapy's growing pains. Science 269:1050.

McPherson. M. J., Quirke. P., and Taylor. G. R. (Eds.) 1994 PCR. A Practical Approach, Vol. 1. Oxford, England: Oxford University Press.

Mulligan. R. C. 1993 The basic science of gene therapy. Science 260:926.

Palmiter. R. D., Brinster. R. L., Hammer, R. E., et al. 1982 Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. Nature 300:611.

Weintraub. H. M. 1990 Antisense RNA and DNA. Sci. Am. 262:40.

Williams. N. 1995 Closing in on the complete yeast genome sequence. Science 268:1560.

#### Outras leituras

Lodish, H. et al., 1995. Molecular Cell Biology. Third Edition. Scientific American Books. W. H. Freeman and Company, New York.

Picciotto, M. R. & Wickman K., 1998. Using knockout and transgenic mice to study neurophysiological and behaviour. Physiological Reviews, 78:1131-1163.

Transgenic-<http://www.copewithcytokines.de/cope.cgi?008161>

Gene expression- <http://www.copewithcytokines.de/cope.cgi?0033305>

Transgenic Animals- <Http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/T/transgenicAnimals.html>

Making transgenic and knockout animals.-[http://www.biology.iupui.edu/biocourses/Bio\\_540/16\\_transgenic\\_2K2.html](http://www.biology.iupui.edu/biocourses/Bio_540/16_transgenic_2K2.html)

Recombinant DNA and gene cloning- <http://users.rcn.com/jkimbal.ma.ultranet/BiologyPages/R/RecombinantDNA.html>

Graddis, T.J. et al., 2002. Designing proteins that work using recombinant technology. Curr. Pharm. Biotechnol., December 1; 3 (4):285-297.

Gene therapy. - [http://www.amsci.org/amsci/articles/99\\_articles/kmiec.html](http://www.amsci.org/amsci/articles/99_articles/kmiec.html)



<b>4. - <u>Considerações introdutórias ao estudo dos genomas</u></b>	129
4.1 - Breve história da sequenciação do DNA	129
4.2 - Estratégias para a sequenciação dos genomas e metodologias genéricas utilizadas sobretudo no estudo do genoma humano	130
No projecto público (HGP)	131
No projecto privado (Celera)	135
4.3 – Bibliografia	139

## **4. - Considerações introdutórias ao estudo dos genomas**

### **4.1 – Breve história da sequenciação do DNA**

Por genoma dum organismo entende-se a sequência nucleotídica completa do DNA do organismo em causa.

Tal como é comumente referido a história moderna da sequenciação do DNA começou em 1977 quando F. Sanger descreveu a sua metodologia para determinar a ordem de encadeamento dos nucleótidos no DNA, tendo sido nesse mesmo ano isolado e sequenciado o primeiro gene humano.

Posteriormente, com a melhoria das metodologias descritas por Sanger e com a utilização de corantes fluorescentes ligados aos nucleótidos, foi possível passar a fazer a leitura da sequenciação em computador.

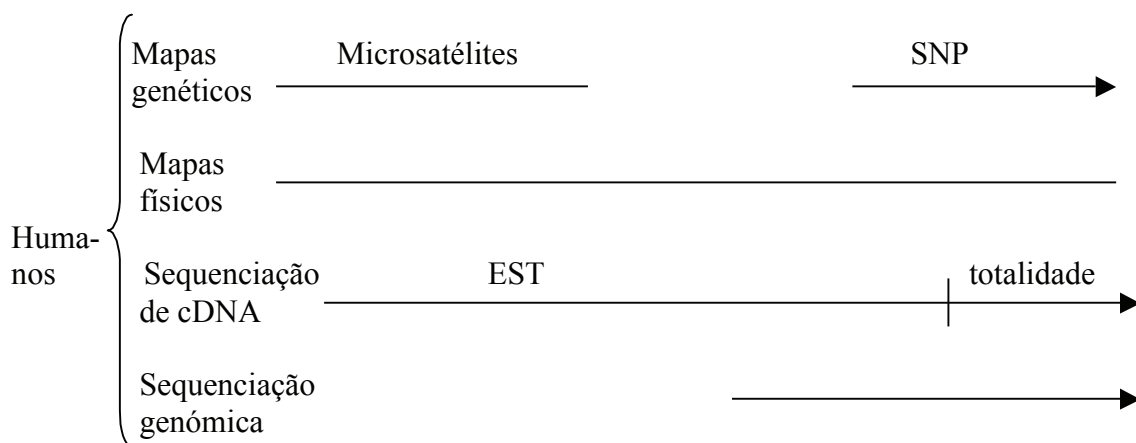
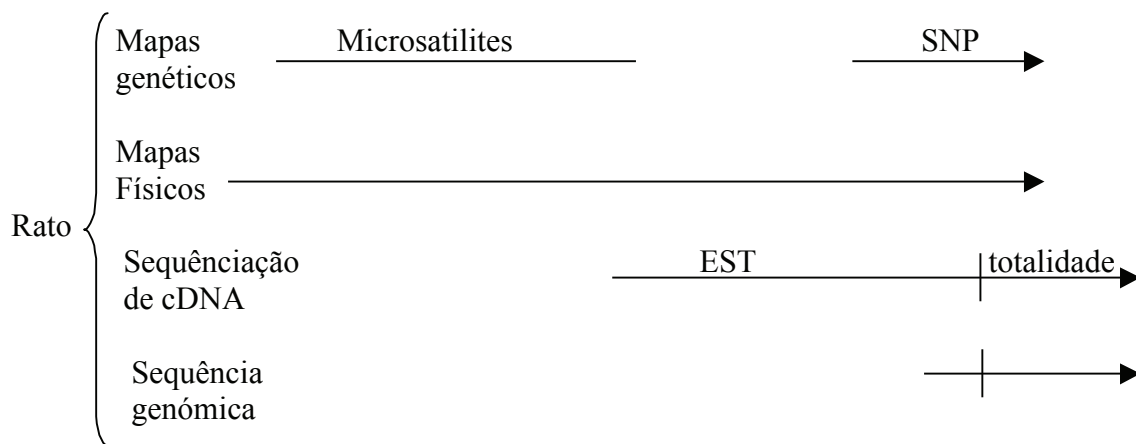
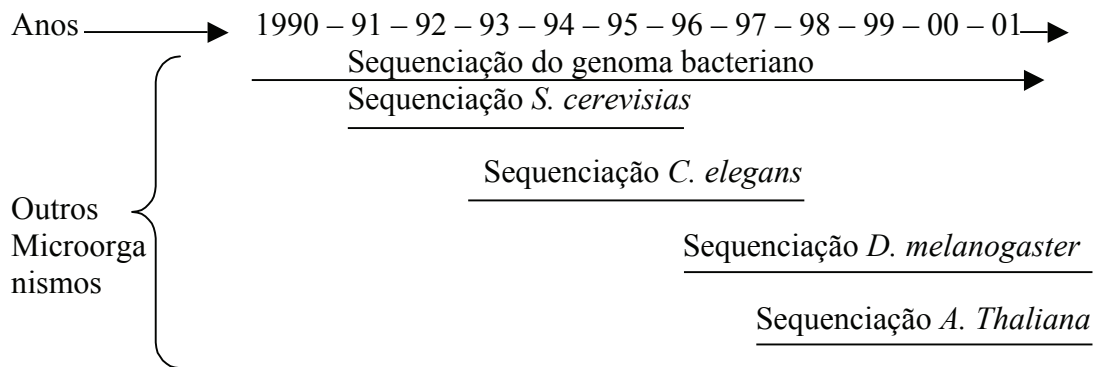
Contudo a sequenciação dos vírus bacterianos  $\phi$ x 174 e lambda (1), do vírus animal SV40(2) e do DNA mitocondrial humano entre 1977 e 1982, através de vários projectos, vieram mostrar que era possível a “assemblage” de pequenos fragmentos sequenciados, em genomas completos, tal como revelaram o valor, para este efeito, dos catálogos completos disponíveis de genes e de outros elementos funcionais.

Também em meados da década de 1980 os programas para criar mapas físicos de clones cobrindo os genomas de leveduras (3) e vermes (4) permitiram o isolamento de genes e regiões baseados apenas na sua posição cromossomal.

Por outro lado, o desenvolvimento da sequenciação por “Random Shotgun” (peças ou fragmentos não seleccionados de DNA que são sequenciados e “assemblados” a partir do genoma total) de fragmentos de DNA complementares, após armar sequências Tags expressas ( EST = expressed sequence tags, são curtas peças de sequências de DNA correspondentes a um fragmento de cDNA- feito a partir de um mRNA celular; os EST tem sido utilizados para “caçar” genes e assim centenas de milhares deles encontram-se nas bases de dados de sequências ) e depois a sequenciação automatizada permitiram ir muito mais além.

Na cronologia em larga escala da análise genómica podem assinalar-se diversas fases e etapas tanto em modelos não vertebrados, bem como no rato e nos humanos, e os primeiros projectos envolvidos no estudo de SNP ( polimorfismos de um simples nucleótido ) e sequências Tags expressas ( EST ).

Cronologia da análise genômica que permitiu chegar à sequenciação de “todo” o genoma humano



Por aqui se podem ver os passos graduais que foram sendo dados a pouco e pouco à medida que se disponibilizavam metodologias adaptáveis e eficientes.

## **4.2 – Estratégias para a sequenciação dos genomas e metodologias genéricas utilizadas sobretudo no estudo do genoma humano**

### No Projecto Público (HGP)

A Human Genome Organization (HUGO) foi criada em 1989 para promover uma colaboração internacional entre os cientistas envolvidos no Human Genome Projects (HGP).

Em 1990, o HGP foi oficialmente iniciado nos USA, sob a égide do Instituto Nacional de Saúde, tendo como objectivo a realização da sequenciação completa do genoma humano.

A sequência de genoma mitocondrial foi completada em 1994 e revelou maior diversidade nos humanos africanos do que noutras populações.

Tudo isto constitui por ora uma tarefa ciclópica pois mesmo conhecendo a sequência de muitos genes e não genes, individuais, torna-se necessário ordená-los em ordem a que se saiba na totalidade do DNA, como eles se interligam entre si, tarefa que apenas em 2000-2001 foi concluída para um número muito restrito de seres vivos.

Em 1995 os trabalhos decorriam essencialmente em duas frentes, na primeira construíram-se mapas genéticos e físicos dos genomas humanos e do rato (5-9) obtendo-se dados chave para identificação de genes responsáveis por doenças e pontos de ancoragem na sequência genómica. Na segunda frente sequenciaram-se os genomas de leveduras (10) e vermes (11) assim como as regiões alvo de genomas mamíferos (12-13-14-15).

Em 26 de Junho de 2000 o HGP anunciou que tinha “assembled” a sequência do DNA do genoma humano, estando disponível na Internet o respectivo “draft” e no mesmo dia uma entidade privada, a Celera publicitou atitude idêntica.

Estes “drafts” cobrem 97% do genoma humano com uma certeza média de 97%. A etapa seguinte consistirá em “acabar” a sequência preenchendo os intervalos (gaps) não conhecidos ainda e aumentando a certeza para 99,99% (1 erro em 10.000).

A sequência acabada do cromossoma humano 22 referida em Dezembro de 1999 assinalava uma sequência de 33.5 milhões de bp, e esta sequência tem um erro médio de menos de 1 em 50.000bp e contem só 11 gaps que não puderam ser preenchidos utilizando a tecnologia corrente, estando identificados nesse cromossoma 545 genes, 134 pseudogenes (genes sem função conhecida) prevendo-se ainda 200-300 genes. Os genes têm de 1.000 a 583.000bp com uma média de 190.000bp. Cerca de 39% deste cromossoma codifica exões e intrões e apenas 3% codifica proteínas. Muitas regiões estão duplicadas em tandem.

A sequência terminada do cromossoma humano 21 foi anunciada em Maio de 2000.

O primeiro sequenciador automático de DNA foi criado em 1987 pela Applied Biosystems na Califórnia que em 1998 o melhorou consideravelmente (ABI PRISM 3.700 DNA Analyser) baseado em protocolos com os métodos de sequenciação com os análogos didesoxi o qual produzia tipicamente apenas 500 a 750pb da sequência por reacção, com uma velocidade de 175.000 leituras por dia sendo a sequenciação do DNA suportada por dispositivos computacionais de alta *performance*.

A sequência completa do genoma do bacteriófago lambda de 40-kbp foi determinada por um método de digestão por restrição “shotgun” e foi conseguida em 1982 (16).

Em 1991 (17) quando foram avaliados métodos para sequenciação do genoma do vírus da varíola, o método de sequenciação por “shotgun” de todo o genoma foi recusado devido á falta de meios adequados de software para fazer o “assembly” do genoma. Mas em 1994 quando foi contemplado um projecto de sequenciação de um genoma bacteriano já foi possível utilizar o método de “shotgun” para o genoma total uma vez que se dispunha já do algoritmo de “assembly” TIGR EST (Institute for Genomic Research).

Em 1995 foi completada a sequenciação do genoma de 1,8Mbp do *Haemophilus influenza* com este método.

A sequência nucleotídica da porção de 120Mbp eucromática do genoma da *Drosophila* foi determinada também em um ano pouco depois (18-20).

Por outro lado também ficou claro que sequências de cDNA (que são obtidas por transcrição reversa a partir de RNA) são essenciais para anotar e validar previsões de genes. Estes estudos foram a base em parte para um método de identificação de genes baseado na expressão de sequências tag (EST ou sequências expressas tag que são obtidas por leitura de uma simples sequência não assemblada de um clone de DNA complementar aleatório) que permitia sequenciar e caracterizar as “bibliotecas” do cDNA.

Foi assim possível a descoberta rápida e o mapeamento dos genes.

Mas os números aumentados de sequências EST necessitavam do desenvolvimento de novos algoritmos nos computadores, capazes de analisar grandes quantidades de dados das sequências tendo sido criado em 1993 um algoritmo que permitiu “assemblar” e analisar centenas de milhares de EST (na base de 30.000 “assemblies EST”).

O “assembly” dos algoritmos podia produzir “assemblies” de cromossomas com uma elevada ordem de precisão e orientação, com sobreposições menores que dez ou seja em menos de dez sobreposições realizadas.

Em conjunto com isto foi possível elaborar um dispositivo que permitia fazer a sequenciação do “shotgun” melhorado da totalidade do genoma.

Um aspecto chave no sistema de sequenciação para estes genomas de megabases de tamanho e maiores, foi a utilização de sequências finais das extremidades emparelhadas (mate pairs) derivadas de repertórios ou “bibliotecas” de sub-clones com distintos tamanhos de inseridos e características de clonagem. Estas sequências finais das extremidades emparelhadas (matepairs) eram sequências de 500pb de comprimento de ambas as extremidades do clone de DNA de cadeia dupla.

Com estes meios foi possível mapear simultaneamente e sequenciar o genoma humano através das sequências finais de 150kpb de cromossomas bacterianos artificiais ou BAC (os BAC clones eram cromossomas bacterianos artificiais vectores que transportavam inseridos de DNA genómico de 100-200kb).

A construção de repertórios ou “bibliotecas” de plasmidos é a primeira etapa crítica na sequenciação por “shotgun”.

Se as “bibliotecas” de DNA não são uniformes no tamanho, não quiméricas e não representam de uma forma aleatória o genoma, então as etapas subsequentes não permitem reconstruir correctamente a sequência do genoma.

Foram utilizados processos de sequenciação do DNA automatizados e de elevada rapidez e infra-estruturas computacionais que articulavam eficientemente as enormes quantidades de informação das sequências (27,3 milhões de leituras de sequências e 14.9 biliões de sequências de pb).

A sequenciação e articulação de ambas as extremidades dos clones dos plasmidos, com “bibliotecas” de 2,10 e 50kpb foi essencial para a reconstrução pelos computadores, do genoma, tendo o emparelhamento das sequências finais atingido uma precisão de 98%.

Para a elaboração do genoma humano foram colhidas amostras de sangue de cinco dadores (mulheres e homens de várias raças) bem como esperma de dois homens.

Linhas celulares limfoblastóides permanentes foram criadas pela imortalização com o vírus de Esptein-Barr.

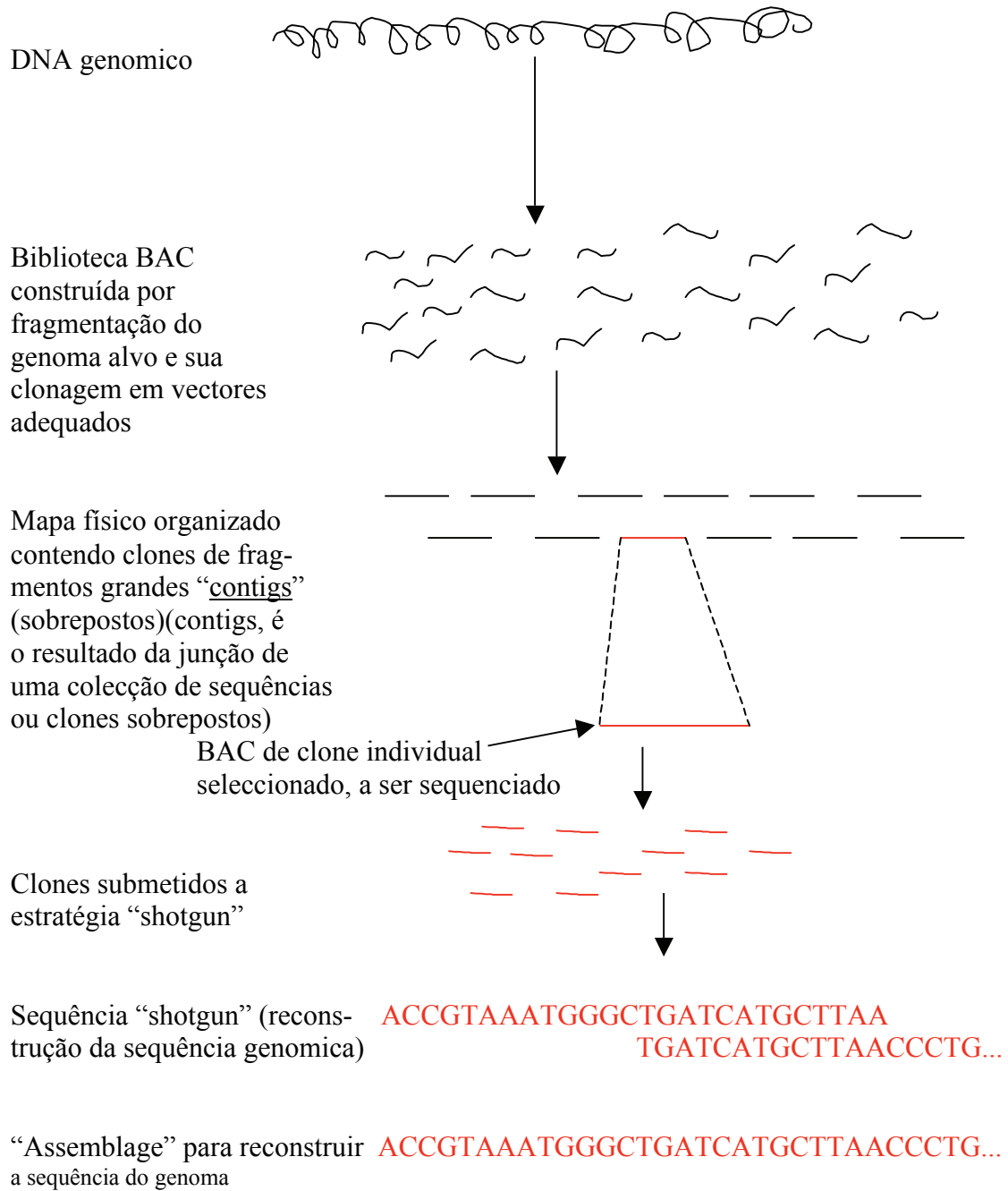
Foi possível verificar que 99,9% do genoma humano é idêntico de indivíduo para indivíduo e que apenas a pequeníssima parte da sequência que varia é que nos torna todos diferentes e pode predispor-nos para a doença ou para diversos tipos de comportamentos e de respostas a meios terapêuticos idênticos.

A sequenciação do genoma humano envolveu pois a sequenciação directa de fragmentos de DNA e a assemblagem das sequências destes fragmentos em unidades maiores na base das suas sobreposições (shotgun assembly).

O HGP (Human Genome Project) utilizou um mapeamento hierárquico e sequenciação, envolvendo a produção de uma série de clones sobreponíveis que cobriam a totalidade do genoma e faziam a sequenciação “shotgun” de cada clone. A sequência do genoma era reconstituída por assemblagem dos fragmentos na base das sequências sobreponíveis e da informação contida nos mapas e nas posições cromossomais de cada clone.

No HGP, grandes fragmentos de DNA cuja posição no genoma era conhecida, eram convertidos em pequenos fragmentos e estes sequenciados e depois reassemblados na base das sobreposições das sequências.

Representação idealizada da estratégia para sequenciação por “shotgun”  
hierarquizada (21)



## No Projecto Privado (CELERA)

A CELERA Genomics utilizou um método de sequenciação “shotgun” do genoma total, sem produzir uma série de clones sobreponíveis, mas incorporando também a informação do HGP quando disponível.

A Celera utilizava a estratégia “whole-genome shotgun” na qual o genoma era quebrado ao acaso (“randomly”) em pequenos fragmentos, e estes sequenciados e depois colocados novamente juntos por ordem correcta o que se baseava em máquinas sequenciadoras de altíssima capacidade e análise computacional sofisticadíssima que identificava e alinhava as sequências sobrepostas.

Claro que a criação de mapas com pontos de referência, através de todo o genoma foi essencial, como os mapas genéticos e físicos.

Os mapas genéticos descreviam a posição dos genes ou outras sequências do DNA identificáveis tais como os RFLP (restriction length polymorphisms) e os marcadores de microsatélites.

Os mapas físicos relacionavam variações características químicas do DNA (bandas G, etc.).

Os “assemblies” da Celera consistiam num conjunto de “contigs” (o resultado da junção de uma colecção de sequências de clones sobrepostos) que eram ordenados e orientados em armações “Scaffolds” (que eram o resultado da conexão de “contigs” por inter-relacionamento da informação da leitura de paired-ends de plasmídeos, leituras de paired-ends de BACs, mRNAs conhecidos ou outras fontes: os “contigs” num scaffold são ordenados e orientados uns em relação aos outros) que eram depois mapeados em localizações cromossomais utilizando para isso marcadores conhecidos.

Os “contigs” consistiam numa colecção de leituras de sequências sobreponíveis que permitiam a reconstituição de um consenso para um dado espaço contíguo no genoma. Os “mate pairs” eram um componente essencial nesta estratégia de “assemblage”. Estes eram utilizados para produzir “scaffold” nas quais o tamanho dos intervalos entre “contigs” consecutivos é conhecido com alguma exactidão.

O diagrama da sequenciação efectuada pela Celera era o seguinte:



Extracção DNA / RNA

Construção da Biblioteca

Pré-sequenciação

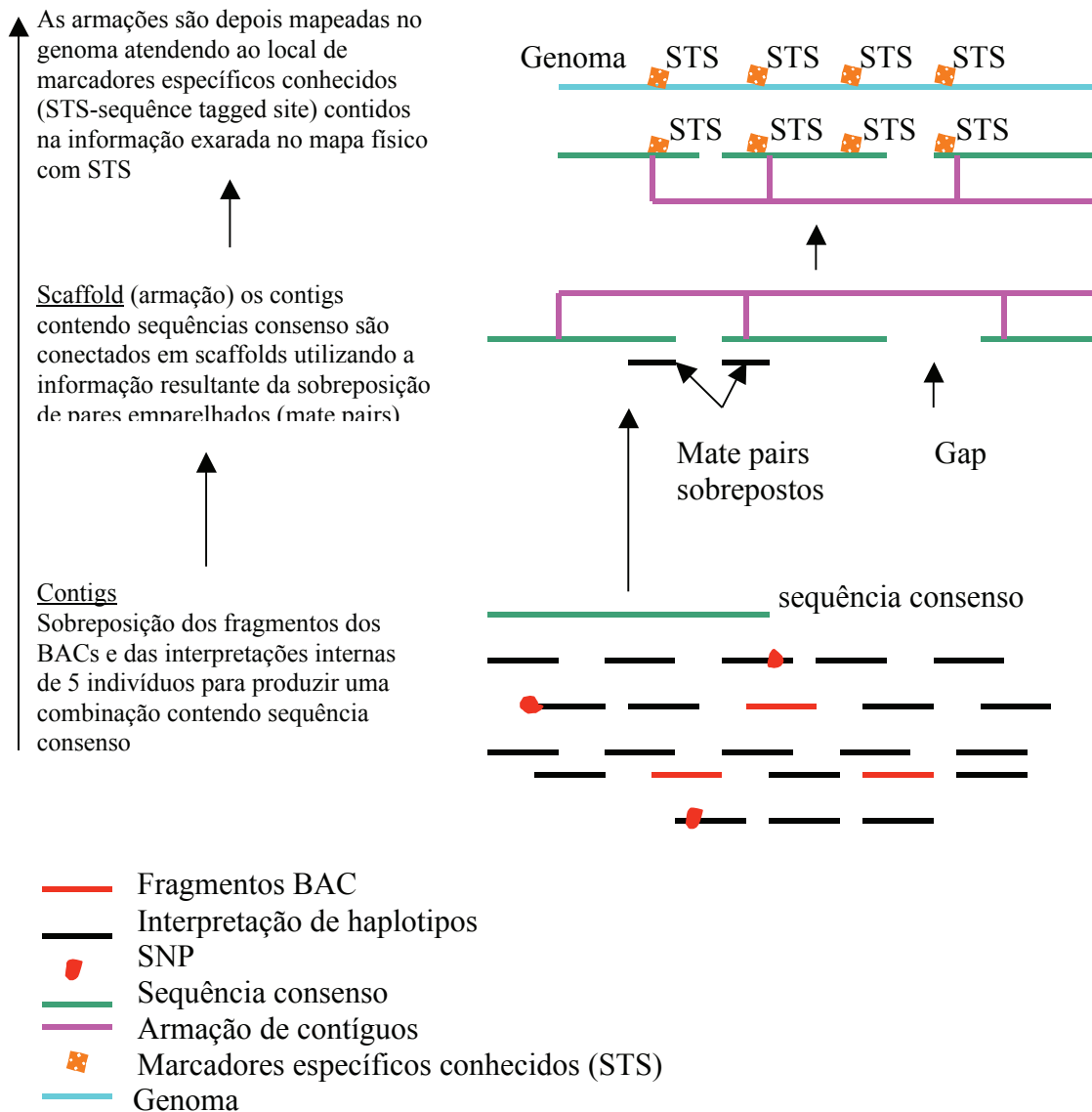
Sequenciação

Post-sequenciação

Pré-asmblagem

Assemblagem

E a anatomia da assemblagem de todo o genoma podia esquematizar-se da seguinte forma (22) (ler de baixo para cima)



Dentro de poucos anos admite-se que haverá capacidade para ler seis biliões de pares de bases do DNA numa célula humana ou animal com a mesma facilidade e despesa com que hoje se lêem um número similar de bits num CD ou num DVD.

Perspectiva-se um futuro em que a ressequenciação e comparação de genomas mamíferos completos será uma operação de rotina nos laboratórios de biologia.

Resumindo e concluindo no projecto público HGP para a realização da análise do genoma humano adaptou-se uma estratégia que consistiu no “clone-a clone” ou seja o genoma total foi cortado em fragmentos contendo várias centenas de milhares de pares de bases de comprimento (100-200kb) e estes fragmentos eram depois inseridos em cromossomas sintéticos os chamados cromossomas bacterianos artificiais ou BACs (estes fragmentos derivavam de sequências cuja localização cromossomal era conhecida e existiam como clones de BACs).

Seguia-se a etapa de mapeamento ou cartografia de cada BAC sendo cada um destes BACs posicionado sobre os genomas dos cromossomas baseado em sequências marcadoras distintas ou específicas os STS, cuja localização tinha sido minuciosamente assinalada, obtendo-se desta forma um mapa de alta resolução do genoma total ( os STSs de local de sequência marcada correspondem a um curto e único locus genómico, tipicamente com menos de 500bp e com o qual pode ser desenvolvida uma reacção de PCR ).

Os clones de BACs eram então seccionados em pequenos fragmentos por um processo designado por “shotgunning”. Cada um destes pequenos fragmentos era sequenciado e os algoritmos computacionais por contrastação reconheciam a informação da sequência dos fragmentos sobrepostos, sendo depois utilizados para reconstruir a sequência completa inserida em cada BAC.

Graças a métodos automatizados (novos sequenciadores multicapilares) e a programas informáticos sofisticados, milhares de sequências nucleotídicas foram diariamente trabalhadas pelos mais potentes ordenadores do mundo.

Entretanto Craig Venter, nos USA ligado á iniciativa privada tinha desenvolvido um método de sequenciação similar ao do HGP mas mais rápido embora reputado como menos fiável.

Com efeito em 1997, Gene Myers, vice-presidente da área de informática da Celera (sector privado) proclamou que a etapa de mapeamento era desnecessária e que os algoritmos utilizados para o reassemble dos fragmentos de DNA “shot-gunned” poderiam ser aplicados aos fragmentos clonados aleatoriamente tomados a partir da totalidade do genoma (Genome R. 7, 401-409, 1997).

Nesta estratégia do “shotgun” do genoma total, os fragmentos eram primeiro “assembled” por algoritmos, em armações (scaffold) muito maiores. A posição correcta destes “scaffolds” sobre o genoma era então trabalhada utilizando as sequências marcadas conhecidas (STSs).

Nesta sequenciação aleatória escolhida por Craig Venter da Celera, em lugar de sequenciar o DNA em grandes fragmentos que servem de pontos de referência sobre esta imensa cadeia de DNA, o DNA era cindido em milhões de pequenos fragmentos e depois estes eram submetidos aos supra-ordenadores para os colocar por ordem.

Estas duas estratégias de “assembly”-assembly do genoma como um todo e assembly cromossomal regional, foram utilizadas cada uma combinando os dados das sequências obtidas pela Celera e pela organização subsidiada pelos fundos públicos a HGP.

As duas estratégias de “assembly” produziram resultados similares que concordaram amplamente com os dados de mapeamento independente, e cobriram efectivamente as regiões eucromáticas dos cromossomas humanos, dando essencialmente a mesma sequência de DNA.

No método desenvolvido pela Celera foi apurada uma sequência consenso da porção eucromática do genoma humano de 2,91 biliões de pares de bases (bp) através do método de sequenciação do “shotgun” do genoma total. Para isso foram produzidas ao longo de nove meses, 14,8 biliões de pares de bases da sequência do DNA a partir de 27. 271. 853 leituras de sequências de alta qualidade obtidas a partir de ambas as extremidades de plasmídeos clones feitos de DNA retirado a cinco indivíduos. A bioinformática foi fundamental em todo este contexto.

A bioinformática e a consequente utilização de bases de dados desta índole parece assumir também grande relevância em medicina veterinária e no melhoramento animal.

Com efeito a grande soma de dados genéticos e sobre os genomas animais, despoletada com o estudo sobretudo dos genomas humanos, levou a uma grande proliferação das bases de dados destas matérias na Internet esperando-se que esta riqueza permita ir mais além na medicina veterinária e na produção animal.

É necessário saber como utilizar toda esta informação acumulada, nas estratégias a delinear para a melhoria da medicina veterinária e da produção animal.

Assim é importante tomar conhecimento do que são estas bases de dados acima referidas e como podem elas ser utilizadas. Como extrapolar informações úteis nelas contidas e como comparar sequências de DNA contidas nestas bases de dados.

Para isto além de técnicos de produção animal com conhecimentos de biologia molecular é necessário o envolvimento de geneticistas e cientistas de meios computacionais.

A contribuição da sequenciação do genoma humano para a melhoria da saúde animal e da produção animal parece poder identificar-se nas características do genoma assinaladas nos humanos e que são extrapoláveis para outras espécies animais sobretudo mamíferas.

Com efeito, a predominância dos genes das famílias implicadas no metabolismo nucleico e sua regulação, factores de transcrição e transmissores de sinais, bem como das proteínas componentes do citoesqueleto, e ainda da família das hidrolases, levam-nos a admitir que o melhoramento da produção animal no que toca à produção de leite e carne terá a implicação de todas estas famílias que numericamente e percentualmente assumem a maior importância na totalidade da maquinaria biológica animal. (vide adiante)

### 4.3 - Bibliografia

1. Sanger, F. et al. (1978) The nucleotide sequence of bacteriophage  $\phi$ x174, J. Mol. Biol. **125**, 225-246
2. Fiers, W. et al. (1978) Complete nucleotide sequence of SV40DNA. Nature, **273**, 113-120
3. Olson, M. V. et al. (1986) Random-clone strategy for genomic restriction mapping in yeast. Proc. Natl. Acad. Sci. USA **83**, 7826-7830
4. Coulson, A., Sulston, J., Boenner, S. & Karn, J. (1986) Toward a physical map of the genome of the nematode Caenorhabditis Elegans. Proc. Natl. Acad. Sci. Usa **83**, 7821-7825
5. Donis-Keller, H. et al. (1987) A genetic linkage map of the human genome, Cell **51**, 319-337
6. Gyapay, G et al. (1994) The 1993-94 Genethon human genetic linkage map Nature Genet **7** 246-339
7. Hudson, T. J., et al. (1995) An STS-based map of the human genome Science **270** 1945-1954
8. Dietrich, W. F. et al. (1996) A comprehensive genetic map of the mouse genome Nature Genet **380** 149-152
9. Nusbaum, C. et al. (1999) A YAC-based physical map of the mouse genome Nature Genet **22** 388-393
10. Oliver, S.G. et al. (1992) The complete DNA sequence of yeast chromosome III Nature **357** 38-46
11. Wilson, R. et al. (1994) 2.2 Mb of *Confluens* nucleotide sequence from chromosome III of *C. Elegans* Nature **368** 32-38
12. Chen, E. Y. et al. (1989) The human growth hormone locus : nucleotide sequences biology and evolution Genomics **4** 479-497
13. McCombie, W.R. et al. (1992) Expressed genes, Alu repeats and polymorphisms in cosmids sequenced from chromosome 4p 16.3. Nature Genet **1** 348-353
14. Martin-Gallardo, A. et al. (1992) Automated DNA sequencing and analysis of 106 kilobases from human chromosome 19p13.3. Nature Genet **1** 34-39
15. Edwards, A. et al. (1990) Automated DNA sequencing of human HPRT locus. Genomics **6** 593-608
16. Sanger, F. et al. (1982) J.Mol Biol. **162** 729
17. Mahy, B. W. Y. et al. (1991) Am. Soc. Microbiol : News **57** 577
18. Adams, M. D. et al., (2000) Science **287** 2185
19. Rubin, G. M. et al., (2000) Science **287** 2204
20. Myers, E. W., et al., (2000) Science **287** 2196
21. International Human Genome Sequencing Consortium. Initial Sequencing and analysis of the human genome (2001) Nature **409** 860-921.
22. Venter, J. C., et al The sequence of the human genome (2001) Science **29** 1304-1351

## Outra Bibliografia

- Olson, M. V. (2001) Clone by Clone by Clone Nature 409 816-818
- Bark, D. and Coply, R. (2001) Filling in the gaps Nature 409 818-820
- Rubin, G. (2001) Comparing Species Nature 409 820-821
- Wolfsberg, T. G., McEntyres, J. And Schuler, G. (2001) Guide to the draft human genome Nature 409 824-826
- Birney, E., Bateman, A., Clamp, M. and Hubbard, T. J. (2001) Mining the draft human genome Nature 409 827-828
- International Human Genome sequencing Consortium (2001) Initial sequencing and analysis of the human genome Nature 409 860-921
- Venter, J. Craig et al. (2001) The sequence of the human genome Science 291 1304-1351
- Sanger, E., et al. (1977) Nucleotide sequence of bacteriophage  $\lambda$  DNA, Nature 265 687-695
- Sanger, E., et al. (1978) The nucleotide sequence of bacteriophage  $\lambda$  I. Mol. Biol. 125 225-246
- Sanger, E., Coulson, A. R., Hong, G. F., Hill, D. F. & Petersen, G. B. (1982) Nucleotide-sequence of bacteriophage Lambda DNA, I. Mol. Biol. 162 729-773
- Anderson, S., et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290 457-465
- Botstein, D., White, R. L., Skolnick, M. & Davis R. W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms Am. J. Hum. Genet. 32 314-331
- Milner R. J. & Sutcliffe, I. G. (1983) Gene expression in rat brain . Nucleic Acids Res. 11 5497-5520
- Houlgatte, r., et al (1995) The Genexpress Index: a resource for gene discovery and the genicmap of the human genome. Genome Res. 5 272-304
- Dietrich, W. F., et al (1996) A comprehensive genetic map of the mouse genome. Nature 380 149-152
- Nusbaum, C., et al A Yac-based physical map of the mouse genome. (1999) Nature Genet 22 388-393
- Oliver, S. G., et al (1992) The complete DNA sequence of yeast chromosome III Nature 357 38-46
- Wilson, R., et al (1994) 2,2Mb of contiguous nucleotide sequence from chromosome III of *C. Elegans* Nature 368 32-38
- Chen E. Y., et al (1989) The human growth hormone locus nucleotide sequence biology and evolution Genomics 4 479-497
- McCombie W. R., et al (1992) Expressed genes. Alu repeats and polymorphisms in cosmids sequenced from chromosome 4p16.3 Nature Genet 1 348-353
- Martin Gallardo, A., et al (1992) Automated DNA sequencing and analysis of 106 kilobases from human chromosome 19p13.3 Nature Genet 1 34-39
- Edwards, E., et al (1990) Automated DNA sequencing of the human HPRT locus Genomics 6 593-608
- Sanger, F. & Coulson, A. R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase I. Mol. Biol. 94 441-448
- Maxam, A. M. & Gilbert W. (1977) A. new method for sequencing DNA Proc. Nat. Acad. Sci. USA 74 560-564

- Anderson, S. (1981) Shotgun DNA sequencing using cloned Dnase I generated fragments. Nucleic Acids Res. 9 3015-3027
- Gardner, R. C., et al (1981) The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. Nucleic Acids Res. 9 2871-2888
- Deininger, P. L. (1983) Random subcloning of sonicated DNA application shotgun DNA sequence analysis. Anal. Biochem. 129 216-223
- Smith, L. M., et al (1968) Fluorescence detection in automated DNA sequence analysis. Nature 321 674-679
- Ju J. Y., Ruan, C. C. Fuller, C. W., Glazer, A. N. & Mathies, R. A. (1995) Fluorescence energy-transfer dyelabeled primers for DNA sequencing and analysis Proc. Natl. Acad. Sci. USA 92 4347-43.51
- Lee L. G., et al (1997) New energy transfer dyes for DNA sequencing. Nucleic Acids Res. 25 2816-2822
- Roseblum, B. B., et al (1997) New dye-labeled terminators for improved DNA sequencing patterns. Nucleic Acids Res. 25 4500-4504
- Metzker, M. L. Lu I. & Gibbs R. A. (1996) Electrophoretically uniform fluorescent dyes for automated DNA sequencing Science 271 1420-1422
- Dietrich, W., et al. (1992) A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131 423-447
- Deloukas, P., et al (1998) A physical map of 30.000 human genes Science 282 744-746
- Purit, K. D. & Maglon, D. R. (2001) RetSeq and LocusLink : NCBI gene-centered resources. Nucleic Acids Res. 29 137-140
- Duret, L. Mouchiroud, D. & Gautier, C. (1993) Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC rich isochores I. Mol. Evol. 40 308-317
- Ewing, B. & Green, P. (2000) Analysis of expressed sequence tags indicates 35.000 human genes. Nature genet 23 232-234
- Gregory, T. R. & Hebert, P. D. (1999) The modulation of DNA content: proximate causes and ultimate consequences Genome Res. 9 317-324
- Hartl, D. L. (2000) Molecular melodies in high and low C. Nature Rev. Genet 1 143-149
- Smit, A. F. (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr. Opin. Genet Dev. 9 657-663
- Prak, E. L. & Haig, H. K. (2000) Ir mobile elements and the human genome. Nature Rev. Genet 1 134-144
- Esnault, C. Maestre, I. & Heidmann, T. (2000) Human LINE retrotransposons generate processed pseudogenes. Nature Genet 24 363-367
- Wei, W., et al (2001) Human L1 retrotransposition: eis-preference vs, trans-complementation. Mol. Cell. Biol. 21 1429-1439
- Smit, A. E. (1996) The origin of interspersed repeats in the human genome. Curr. Opin. Genet Dev. 6 743-748
- The international SNP Map Working Group. An SNP map of the human genome generated by reduced representation shotgun sequencing (2000) Nature 407 513-516
- Kruglyak, S., Durrett, R. T., Schug, M. D. & Aquadro. C. F. (1998) Equilibrium distribution of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc. Natl. Acad. Sci. Usa 95 10774-10778
- Toth, G., Gaspari, Z. & Jurak, J. (2000) Microsatellite in different eukaryotic genomes: survey and analysis. Genome Res. 10 967-981

- O'Keefe, C. & Eichler, E. in Comparative Genomics: Empirical and Analytical Approaches to Gene Order Dynamics, Map alignment and the Evolution of Gene Families (eds Sankoff, D. & Nadeau, J.)29-46 (Kluwer Academic, Dordrecht. 2000)
- Lander, E. S. (1996) The new genomics: Global views of biology. Science 274 336-339
- Eddy, S. R. (1999) Noncoding RNA genes, Curr. Op. Genet. Dev. 9 695-699
- Ban, N. Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000) the complete atomic structure of the large ribosomal subunit at 2,4 angstrom resolution. Science 289 905-920
- Nissen. P., Hansen, J. Ban. N., Moore. P. B. & Steitz. T. A. (2000) the structural basis of ribosome activity in peptide bond synthesis. Science 289 920-930
- Weinstein. L. B. & Steitz, J. A. (1999) Guided tours: from precursor snoRNA to functional snoRNP. Curr. Opin. Cell Biol. 11 378-384
- Bachellerie, J. P. & Cavaille, J. (1997) in Modification and Editing of RNA (ed. Benne, H. G. a. R.) 255-272 (ASM, Washington DC, 1998)
- Burge, C. & Sharp P. A. Classification of introns: U2-type or U12-type. Cell 91 875-879
- Hatlen, L. & Attardi, G. (1971) Proportion of the HeLa cell genome complementary to the transfer RNA and 5S RNA, J. Mol. Biol. 56 335-553
- Long, E. O. & David. I. B. (1980) Repeated genes in eukaryotes. Annu. Rev. Biochem. 49 727-764
- Sylvester, J. E., et al. (1986) The human ribosomal RNA genes : structure and organization of the complete repeating unit. Hum. Genet. 73 193-198
- Hawkins, J. D. (1988) A survey on intron and exon lengths. Nucleic Acids Res. 16 9893-9908
- Burge, C. & Karlin, S. (1997) Prediction of complete gene structures in human genomic DNA I. Mol. Biol. 268 78-94 Labeit, S. & Kolmerer, B. Titins: giant proteins in charge of muscle ultrastructure and elasticity Science 270 293-296 (1995)
- Lewin, B. Gene Expression: Wiley New York 1980
- The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM consortium functional annotation of a full-length mouse cDNA collection. (2001) Nature 409 685-690
- Barbazuk, W. B., et al. (2000) The syntenic relationship of the zebrafish and human genomes. Genome Res. 10 1351-1358
- McLysaght, A., Enright, A. I. Skrabanek, L. & Wolfe, K. H. (2000) Estimation of synteny conservation and genome compaction Between pufferfish (Fugu) and human. Yeast 17 22-36
- Trachtulec, Z., et al. (1997) Linkage of TATA- binding protein and proteasome subunit C3 genes in mice and humans reveals synteny conserved between mammals and invertebrates. Genomics 44 1-7
- Nadeau, J. H. (1989) Maps of Linkage and synteny homologies between mouse and man. Trends Genet 5 82-86 DeBry, R. W. & Seldin, M. F. Human/mouse homology relationships. Genomics 33 337-351 (1996)
- Novacek, M. J. (1992) Mammalian phylogeny: shaking the tree. Nature 356 121-125
- O'Brien, S. J., et al. (1999) Genome maps 10. Comparative genomics. Mammalian radiations Wall chart. Science 286 463-478
- Collins, F. S. (1995) Positional cloning moves from perditional to traditional. Nature Genet 9 347-350
- Maeshall, E. (2000) Public private project to deliver mouse genome in 6 months. Science 290 242-243

- J. Gocayne et al. (1987) Proc. Natl. Acad. Sci. Usa 84 8296
- A. Martin-gallardo et al., DNA Sequence 3, 237 (1992); W. R. McCombie et al., Nature Genet 1 348 (1992) M. A . Jensen et al., DNA sequense 1 233 (1991)
- F. Sanger, A.Q. R. Coutson, G. F. Hong. D. F. Hill, G. B. Petersen, (1982) J. Mol. Biol. 162 729
- B. W. J. Mahy, J. J. Esposito, J. C. Venter, (1991) Am Soc. Microbiol. News 57 577
- C. M. Fraser et al., (1995) Science 270 397
- C. J. Bult et al., Science 273 1058 (1996) ; J. f. Tomb et al ., Nature 388 539 (1997) ; H. P. Klent et al., Nature 390 364 (1997)
- S. Zhao et al., (2000) Genomics 63 321
- J. C. Venter et al., (1998) Science 280 1540
- M. D. Adams et al., (2000) Science 287 2185
- G. M. Rubin et al (2000) Science 287 2204
- E . W. Myers et al (2000) Science 287 2196
- G. L. Miklos, (1979) B. John Am. J. Hum. Genet 31 264; U. Fracke Cytogenet. Cell Genet. 65 206 (1994)
- J. E. Horvath, S. Schwartz, E. E. Eichler, (2000) Genome Res. 10 839
- G. P. Holmquist, (1992) Am. J. Hum. Genet. 51 17
- G. Bernardi, (2000) Gene 241 3
- S. Zoubak , O. Clay, G. Bernardi (1996) Gene 174 ,95
- M. j. McEachem, A. krauskopf, E. H. Blackburn, (2000) Annu. Rev. Genet. 34 331
- A. Bird, (1987) Trends Genet 3 342
- M. Gardiner-Garden, M. Frommer, (1987) J. Mol. Biol. 196 261
- S. H. Cross, A. Bird, (1995) Curr. Opin. Genet. Dev. 5 309
- D. brett et al., (2000) FEBS Lett. 474 83
- H. J. Muller, H. Kern, (1967) Z. Naturforsch B 22 1330lable at



<b>5. - <u>Algumas características gerais dos genomas</u></b>	145
5.1 - Introdução	145
5.2 - Sintenia, genoma humano e genoma de outros animais, Comparando espécies	147
5.3 - Eucromatina, heterocromatina, genoma e mapa citogenético e mapa de “Linkage”	149
5.4 - Diversidade genómica e SNP	151
5.5 - Organização dos genes nos cromossomas. Famílias de genes e sua classificação	155
5.6 - Conclusão	167
5.7 – Bibliografia	169

## 5. - Algumas características gerais dos genomas

### 5.1 - Introdução

Um gene é um *locus* (sítio específico de um gene dentro de um cromossoma) de exões cotranscritos. Um único gene pode originar diversos transcritos, produzindo assim proteínas distintas múltiplas, com múltiplas funções, devido a diversos tipos de “splicing” e a diversos locais de iniciação e de terminação.

“Splicing” é o processo que remove intrões (porções não codificantes de proteínas) dos RNA transcritos. Os exões (porções codificadoras de proteínas) podem também ser removidos. Consoante os exões que são removidos, proteínas diferentes podem ser produzidas a partir do mesmo RNA inicial ou do mesmo gene. Proteínas diferentes produzidas desta forma são “variantes de corte” ou “cortadas alternativamente”.

Estes sinais para início de transcrição e para corte (*splicing*) dos exões, separados por alguns ou por centenas de milhares de pares de bases, no contexto de biliões de pares de bases que constituem o genoma de DNA, são no entanto perfeitamente identificáveis pelas células.

Em cada gene há pois que definir a sua estrutura e a unidade de transcrição respectiva.

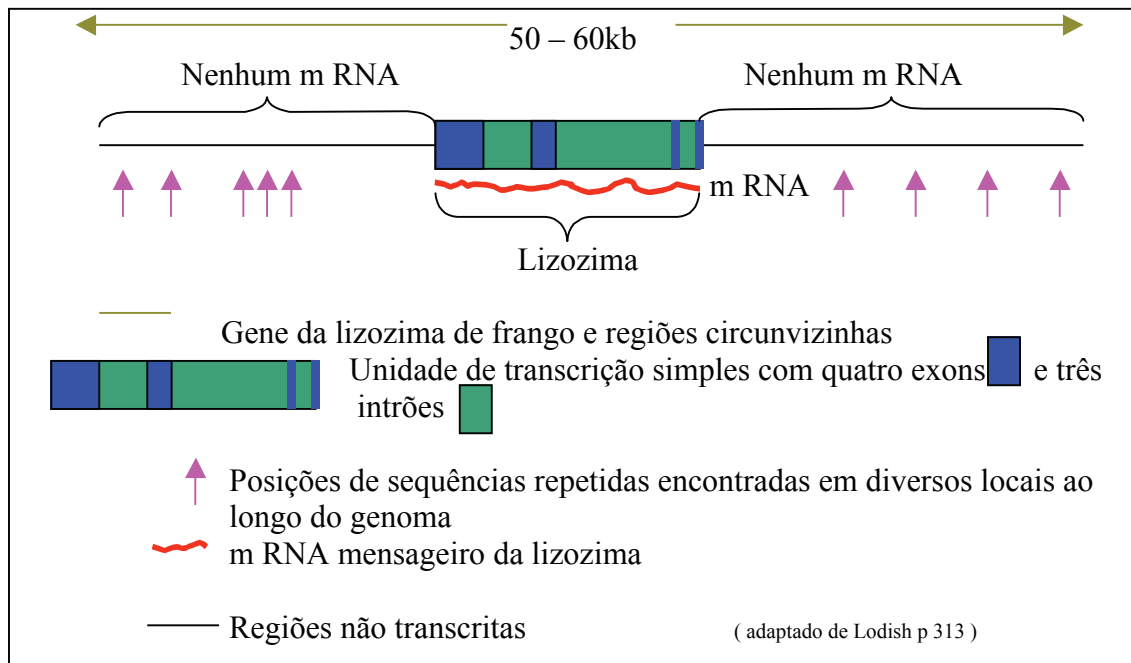
Os genomas dos eucariotas superiores contem muito DNA não funcional.

O conteúdo em DNA celular nos diferentes animais não varia em consonância com a aparente complexidade dos organismos.

Muito do DNA em certos organismos não codifica RNA nem tem qualquer função reguladora ou estrutural.

Cerca de um quarto a metade dos genes codificadores de proteínas em todos os organismos multicelulares são únicos ou solitários (são representados no genoma apenas uma vez) e os restantes genes codificadores de proteínas pertencem a famílias de dois ou mais genes similares.

A este propósito servimo-nos do exemplo de um gene único ou solitário que codifica a lisozima nos frangos. Uma sequência de 15kbp de DNA codifica a lisozima que contem quatro exões e três intrões, mas as regiões que flanqueiam estes intrões e exões estendem-se 20kb “acima” e outros 20kb “abaixo” da unidade de transcrição não codificando qualquer mRNA detectável.



A ORF (*open reading frame* ou sejam longos trechos de trios de códons no DNA que não são interrompidos por códons de paragem da translação) é a parte de um gene codificando a sequência de ácidos aminados do seu produto proteína.

O número de genes que codificam proteínas nos mamíferos calcula-se em cerca de vinte e sete a trinta e cinco mil. Calcula-se também que um gene “típico” dos humanos alcance em média um tamanho de 27.894 bases e também se estima que cada gene tenha em média 7,8 exões embora também existam estimativas que referem 3,7 exões. O maior número de exões identificados num transcrito é de 234 para o mRNA da titina.

Para estudo do genoma humano mais de 90% deste genoma foi “assemblado” em armações (*scaffolds*) de 100.000 pares de bases ou mais e 25% do genoma foi assemblado em armações de 10 milhões de pares de bases ou mais.

Há fortes provas de que existam transcrições codificantes para 26.588 proteínas e talvez cerca de mais 12.000 genes adicionais, isto derivado do trabalho de computação efectuado e outras provas de suporte embora fracas e com semelhanças no rato.

Conjuntos com grande densidade de genes são evidentes, e quase metade dos genes encontram-se dispersos em sequências de baixo conteúdo em G+C, separados por grandes extensões de sequências aparentemente não codificadoras.

Apenas 1,1% do genoma humano abarca exões, enquanto 24% corresponde a intrões, com 75% do genoma correspondendo a DNA entre genes.

Duplicação de blocos contendo vários segmentos (segmentais), variando no seu tamanho são abundantes.

Uma análise comparativa do genoma humano com os de outras espécies, indica que nos vertebrados ocorreram expansões dos genes associados com a função neuronal, com a regulação do desenvolvimento específico dos tecidos, com a hemostase e com os sistemas imunitários.

Estão também referidos 2,1 milhões de polimorfismos com sede em nucleótidos simples (SNPs). Um par aleatório de genomas humanos haplóides difere em cerca de um par de bases por cada 1.250 nucleótidos, em média, mas há uma marcada heterogeneidade ao nível dos polimorfismos ao longo do genoma. Menos de 1% de todos os SNPs originam uma variação nas proteínas, mas o objectivo de determinar qual o SNP que tem consequências funcionais permanece por esclarecer.

Avaliações recentes permitiram chegar à conclusão de que existem 25.000 a 35.000 genes humanos (1,2).

Mais de 10.000 genes humanos estão catalogados no Online Mendelian Inheritance in Man (OMIM) (3), que documenta todas as doenças humanas hereditárias e os genes mutados que as causam, o que foi facilitado pelo conhecimento das sequências dos RNA mensageiros (4,5).

Neste caso as posições dos diversos genes conhecidos foi determinada pelo alinhamento dos mRNA com as sequências genómicas. Para os restantes genes utilizaram-se os métodos computacionais para encontrar genes (6,7).

Quando espécies de mRNA se alinham diferentemente sobre sequências genómicas, isto indica que ocorreram processos alternativos de “splicing”.

No conjunto de mRNAs completo, de referência, 11.174 transcritos foram sequenciados a partir de 10.742 genes distintos, 2,4% dos genes tinham variantes múltiplas de “splicing” (8). Contudo há sugestões de que cerca de 60% dos genes humanos têm variantes múltiplas de “splicing” (9).

O mapa de homologia rato – humano revela, apesar das duas espécies estarem separadas por uma evolução de 100 – 200 milhões de anos, que os genes das duas espécies, rato e homem, têm regiões cromossomais homólogas similares que possuem uma ordem de genes conservada (sintenia).

O genoma humano, sendo a primeira sequência genómica de um vertebrado a ser determinada, parece ser muito representativa daquilo que se poderá encontrar noutros genomas de vertebrados (10).

As regiões codificantes de genes parecem representativas apenas de 3% do DNA humano, enquanto as sequências repetidas formam uma grande parte do restante DNA, cerca de 46%. Estas repetições podem ter ou não uma função, mas são certamente características dos genomas dos grandes vertebrados. O resto da sequência contém promotores, sequências reguladoras da transcrição e outros aspectos ainda não conhecidos.

## **5.2 - Sintenia, genoma humano e genoma de outros animais. Comparando espécies** (9,18)

Nas bactérias genes localizados próximos uns dos outros, por vezes codificam proteínas de uma mesma via metabólica e são regulados pelo mesmo operão (operão no DNA bacteriano é um conjunto de genes contíguos transcritos de um promotor que dão lugar a um mRNA policistrónico).

Nos mamíferos genes que se encontram próximos uns dos outros só raramente têm funções comuns embora tenham uma história comum.

No que se refere a segmentos do DNA conservados nos humanos e nos ratos é possível afirmar que eles tiveram um ancestral comum há cerca de 100-200 Myr (milhões de anos). Apesar de uma

distância evolutiva de 200 milhões de anos entre estas espécies, uma significativa fracção de genes apresenta sintenia entre os dois, sendo preservada dentro de segmentos conservados. Genes firmemente ligados numa espécie mamífera tendem a encontrar-se ligados noutras espécies mamíferas.

De facto segmentos conservados têm sido observados entre humanos e peixes (11,12) e mesmo com invertebrados como insectos e vermes (13).

Sintenia conservada indica que pelo menos dois genes que residem num mesmo cromossoma numa espécie, são também localizados num mesmo cromossoma noutras espécies (14).

Os *loci* sinténicos diz-se que estão ligados num “segmento conservado” quando não é apenas a posição cromossomal mas também a ordem linear dos *loci* que foi preservada, sem interrupção por outros rearranjos cromossomais.

Numa observação inicial dos *loci* homólogos similares nos humanos e no rato (15) parece que o número total de segmentos conservados nas duas espécies será de cerca de 180.

Quase todos os genes humanos do cromossoma 17 se encontram no cromossoma 11 do rato, com dois membros da família lactogénica placentária inseridos no cromossoma 13 do rato.

O cromossoma 20 parece inteiramente ortólogo (têm a mesma função) do cromossoma 2 do rato, aparentemente num segmento simples.

O segmento conservado aparentemente contíguo, maior, no genoma humano tem sede no cromossoma 4 incluindo cerca de 90,5Mbp do DNA humano, que é ortólogo do cromossoma 5 do rato.

Com cerca de 200 segmentos conservados entre o rato e o humano e cerca de 100 milhões de anos de evolução a partir de um ancestral comum (16) podemos calcular uma média de cerca de 1,0 rearranjo cromossomal por milhão de anos.

Contudo há boas provas de que a velocidade de rearranjo cromossomal (tal como a velocidade de substituição de nucleótidos) difere entre as duas espécies.

Entre os mamíferos, os roedores podem apresentar uma invulgarmente rápida alteração dos cromossomas.

Estudos em diversos tipos de mamíferos tais como gatos, vacas, ovinos e suínos, sugerem uma velocidade média de alteração dos cromossomas de apenas 0,2 rearranjos por milhão de anos nestas linhagens (17).

As duplicações segmentais podem ocorrer devido a um desigual “crossing over” criando-se assim famílias de genes em regiões específicas cromossomais. Este mecanismo no genoma humano pode criar pequenas famílias como por exemplo a família de cinco genes relacionados com a  $\alpha$ -globina no cromossoma 11, ou grandes famílias como o conjunto de genes dos receptores olfatórios que atingem cerca de 1.000 genes e pseudo-genes.

Por outro lado sabe-se que há um mais elevado grau de “splicing” alternativo nos humanos e nos vertebrados do que noutras espécies vivas (18). Assim mais proteínas podem ser codificadas por cada gene naquelas espécies.

Também as proteínas individuais por vezes apresentam unidades estruturais discretas, chamadas domínios que são conservados durante a evolução. Mais de 90% dos domínios identificados nas proteínas humanas, encontram-se também presentes nas proteínas dos frutos e vermes.

Das proteínas humanas que se prevê existirem, 60% tem alguma similaridade de sequências com proteínas de outras espécies cujos genomas foram sequenciados.

O genoma do rato (*Mus musculus*) é cerca de 14% mais pequeno que o genoma humano (2,5Gb comparado com 2,9Gb).

Mais de 90% dos genomas de rato e humano podem ser partilhados em regiões correspondentes de sintenia conservada.

Ao nível nucleotídico, cerca de 40% do genoma humano pode ser alinhado com o genoma de rato.

Os genomas de rato e humanos parecem conter cada um cerca de 30.000 genes codificadores de proteínas.

A proporção de genes de rato com um único ortólogo identificável no genoma humano parece ser de cerca de 80%.

A proporção de genes de rato sem qualquer homólogo correntemente detectável no genoma humano (e vice-versa) parece ser menor do que 1%.

### **5.3 - Eucromatina, heterocromatina, genoma e mapas citogenético e de “linkage”**

No genoma humano as sequências de cerca de 90% das regiões eucromáticas (regiões fracamente coradas e ricas em genes ao contrário das regiões heterocromáticas que são profundamente coradas e pobres em genes) estão completadas, calculando-se que o tamanho total do genoma humano seja de 3.2Gb (gigabases ou bilião de bases). Destes cerca de 2,95Gb é eucromatina e apenas 1,1% a 1,4% destes corresponde a genes codificadores de proteínas ou seja 5% dos 28% de sequências que são transcritas em RNA.

Mais de metade do DNA consiste de sequências repetidas de diversos tipos: 45% em quatro classes de elementos DNA parasitas, 3% em repetições contendo apenas algumas bases e cerca de 5% são duplicação recente de grandes segmentos do DNA.

A maioria do DNA parasita provem da acção da transcriptase reversa sobre RNA.

A maioria dos genes estão pois localizados fora das regiões heterocromáticas.

Nas regiões do genoma ricas em bases GC, a densidade de genes é maior e a média do tamanho de intrões, menor. Estes intrões constituídos por grandes sequências nucleotídicas sem significado que interrompem as sequências codificadoras de proteínas, os exões, são muito mais longos nos DNA humanos do que noutros genomas sequenciados.

As sequências reguladoras dos genes tem sido procuradas mas sem grande êxito.

Como já se referiu a heterocromatina são áreas genómicas condensadas altamente repetitivas e pobres em genes.

Para o genoma humano completo contendo 3.200 milhões de pares de bases o consórcio público (HGP) refere cerca de 32.000 genes, cerca de 15.000 deles sendo genes conhecidos e 17.000 sendo previsões.

Mas destes 32.000 genes admitem 24.500 genes actuais, alguns dos genes previstos podendo ser pseudogenes.

Todos os genes são expressos ou seja copiados (transcritos) em RNA mensageiro e a maioria deste traduzida em proteína. Mas há provas de que uma extensão de DNA transcrita não significa que essa extensão de DNA seja um gene.

Não conhecemos como as células controlam eficientemente a transcrição: parece que algumas sequências de DNA sem serem genes são transcritas com relativa frequência, mas não se sabe como a célula identifica transcritos que não são translidos numa proteína funcional.

Por outro lado, proteínas que não podem servir para qualquer função útil (por exemplo porque não se podem “dobrar” correctamente) podem ser feitas, mas são rapidamente removidas.

Para se chegar a definir genes codificadores de proteínas apenas os estudos computacionais não chegam sendo necessário caracterizar as proteínas e as suas funções.

Por outro lado algoritmos computacionais para encontrar genes não podem facilmente prever a existência de formas alternativas de um mRNA sem informação experimental, mas esta informação é difícil de obter no caso de mRNAs raros. Por exemplo um exão que é utilizado apenas em algumas células do cérebro humano, pode nunca ter sido detectado experimentalmente num mRNA.

Espera-se no futuro saber mais sobre genes codificando RNA que não codifica proteínas, pontos de início das replicações de DNA e elementos genéticos que controlam a estrutura dos cromossomas.

Nos vertebrados foi possível verificar a elaboração e o aparecimento de novo de dois tipos de genes, os responsáveis por capacidades específicas dos vertebrados como é o caso da complexibilidade neuronal, da coagulação do sangue e da resposta imunitária e aqueles que aumentam capacidades genéricas como é o caso dos genes para sinalização intra e extracelular, desenvolvimento, morte programada da célula, e controle da transcrição dos genes.

### Mapas citogenéticos

Um dos aspectos mais característicos da estrutura do genoma é o aspecto em bandas produzido pela coloração pelo Giemsa.

Estudos destas bandas dos cromossomas revelaram que cerca de 17% a 20% nos humanos, consiste de bandas C ou heterocromatina constitutiva (19). Muita desta heterocromatina (ou sejam regiões de cromatina que permanecem altamente condensadas e inactivas no que se refere a transcrição, durante a interfase) é altamente polimórfica e consiste de diferentes famílias de DNA alfa satélite (centroméricas) com várias estruturas repetidas.

Diversos cromossomas revelam duplicações complexas inter e intracromossomais nas regiões pericentroméricas. Cerca de 5% das sequências interpretadas foram identificadas como sequências alfa satélites centroméricas (20).

Os restantes 80% do genoma correspondentes ao componente eucromático são divisíveis em bandas G, R- e T. (21) que diferem na sua composição nucleotídica e densidade de genes embora seja muito difícil determinar com precisão os limites de cada banda a nível molecular.

As bandas T. são mais ricas em G+C e em genes enquanto as bandas G. são pobres em G+C.

Bernardi denominou de isocores, na porção eucromatina, longas extensões de DNA com diferentes composições em bases (designando as L, H1, H2 e H3) com mais de 300kbp de comprimento.

As isocores L (leves) eram pobres em G+C (menos que 43%) enquanto as isocores H (pesadas) eram repartidas em três classes ricas em G+C, representando 24, 8 e 5% do genoma.

Os genes eram muito poucos nos isocores L e vinte vezes mais abundantes nos isocores H2 e H3 (23).

A densidade de genes é maior nas regiões ricas em G+C do que nas regiões de baixo conteúdo em G+C.

Nos humanos os cromossomas 17, 19 e 22 que têm um elevado número de isocores H3 têm também a mais alta densidade de genes. Por outro lado os cromossomas X, 4, 18, 13 e Y têm a mais baixa densidade de genes e têm também as mais baixas concentrações de bandas H3.

Os mapas de encadeamento (linkage) são uma base para a análise genética e para estudo de características hereditárias e da posição de genes de clonagem. A distância métrica, centimorgans (cM) é baseada na proporção de recombinação entre cromossomas homólogos durante a meiose.

Em geral a proporção de recombinação nas fêmeas é maior do que nos machos, bem como nas regiões teloméricas dos cromossomas (24).

Ilhas de CpG são extensões de DNA não metiladas e com uma elevada frequência de dinucleótidos CpG quando comparados com o restante genoma total (25).

As ilhas de CpG ocorrem preferencialmente no início da transcrição dos genes. Também tem sido verificado que a maioria dos genes “House Keeping” têm ilhas CpG na extremidade 5’ do transcrito (26).

A metilação das ilhas CpG está correlacionada com a inativação dos genes (27).

No genoma humano calculam-se existirem 30.000 a 45.000 ilhas de CpG, parecendo existir uma forte correlação entre as ilhas de CpG e os primeiros exões codificantes.

#### **5.4 - Diversidade genómica e SNP**

Uma sequência de DNA é uma combinação linear de quatro nucleótidos. Se compararmos duas espécies, ou dois seres distintos posição a posição de cada nucleótido, quando encontramos diferentes nucleótidos na mesma posição temos um SNP que reflecte mutações no passado.

A base das diferenças de humano para humano e de animal para animal reside no polimorfismo de simples nucleótidos.

Os seres humanos diferem entre si em cerca de um par de bases por cada mil bases do genoma.



Estes polimorfismos num simples nucleótido (SNP) são marcadores que podem permitir aos epidemiologistas identificar a base genética de diversas doenças.

A diversidade tem pois sede na identificação e mapeamento de 1,4 milhões de polimorfismos em simples nucleótidos (SNP) no genoma humano.

Calcula-se que se comparar-mos duas sequências de DNA humanos, em cada 1.000 – 2.000 nucleótidos surge um SNP. Como no genoma humano há cerca de 3,2 biliões de nucleótidos, existirão cerca de 1,6 a 3,2 milhões de SNP, mas apenas 0,1% de 1,6 milhões de SNP está verificado e testado.

A diversidade nucleotídica em regiões contendo genes tem sido calculada em oito diferenças por 10 kilobases (28,29).

Nos humanos as diferenças são menores nos cromossomas do sexo. No cromossoma X as diferenças são de cerca de 4,69 diferenças por 10 kilobases e muito menos no cromossoma Y (1, 5' diferenças por 10 kilobases). Isto é devido a que os cromossomas do sexo têm perfis de mutação e recombinação (o “swapping” de segmentos similares de DNA durante a geração de óvulos e esperma) que diferem um do outro e dos cromossomas não ligados ao sexo.

Os locus HLA que codificam proteínas que apresentam antígenos ao sistema imunitário, revelam a maior diversidade.

A principal utilização do mapa de SNP pode ser o de averiguar as contribuições dos genes individuais para doenças que tenham uma base complexa de multigenes.

Também as variantes dos genes por tecidos e órgãos podem indicar incompatibilidades reflectindo-se no sucesso dos transplantes.

E ainda o conhecimento de variantes de genes raros que podem estar implicados em doenças hereditárias.

Variações nas sequências dos genes podem indicar diferenças na susceptibilidade ou na protecção para todas as espécies de doenças, na idade em que pode surgir a doença e a severidade desta e na forma como os organismos reagem aos tratamentos.

A título de exemplo sabe-se, nos humanos, que a simples diferença numa base, do gene APOE está associada com a doença de Alzheimer e que a simples deleção dentro do gene CCR5 receptor – “chemokine” origina resistência ao HIV e SIDA.

Comparando perfis e frequências dos SNP em doentes e controlos é possível identificar quais os SNP associados com as doenças (30, 31, 32).

Tudo isto será muito mais eficiente quando alguns problemas relacionados com os SNP estejam resolvidos.

Identificando as variações através de todo o genoma, o mapa de SNP pode ser a melhor via para compreender os papéis da natureza e da alimentação sobre os seres vivos.

Os SNPs podem ser utilizados para traçar a hereditariedade de qualquer gene contribuindo para as características que fazem de cada indivíduo um ser único, e estão subjacentes às susceptibilidades para as doenças comuns como o cancro, a diabetes ou as doenças cardíacas.

Também se admite que os SNP, ajudam a explicar porque razão certos indivíduos respondem de forma diferente aos mesmos agentes terapêuticos.

Pode afirmar-se que o mapa SNP representa uma companhia inseparável do “ Book of Life” o *blueprint* genético de 3 biliões de letras anunciado no verão de 2000.

Como quaisquer dois seres humanos são 99,9% similares na sua sequência genómica, torna-se necessário mapear 0,1% desse DNA, que é ao fim e ao cabo aquele que determina as características individuais únicas e que pode ajudar a explicar as razões porque alguns indivíduos são mais susceptíveis a certas doenças do que outros.

Sabe-se que as doenças causadas por simples genes e herdadas de acordo com as leis Mendelianas são actualmente raras.

As doenças mais frequentes como a diabetes e as doenças cardíacas são causadas pela interacção de diversos genes, mas encontrar estes genes é uma tarefa difícil, e, ultimamente, tem avançado a ideia de que o risco de contrair tais doenças pode ser traçado utilizando pequenas diferenças no código genético ou substituições de um simples nucleótido no DNA.

Por comparação dos perfis de SNP em indivíduos afectados e não afectados com diabetes, e com controlos adequados saudáveis, por exemplo, os cientistas podem catalogar variações específicas no DNA que estejam subjacentes ao risco para uma dada doença e respostas diferentes de cada indivíduo aos tratamentos.

Nos humanos calcula-se que 60.000 SNP caiem dentro dos exões e que 85% dos exões estão situados, dentro de 5kb, do SNP mais próximo.

Centenas de genes de doenças humanas (33) estão identificados (<http://ncbi.nlm.nih.gov/OMIM>) mas quase todos eles são condições raras nas quais a mutação de um simples gene é necessária e suficiente para causar a doença.

Também nos animais o mesmo acontece tal como pode ser verificado na lista OMIM (<http://www.angis.org.au/databases/BIRBX/omia/species-list.html>) e que teremos oportunidade de apresentar oportunamente.

Na maioria das doenças a arquitectura genética é mais complexa, aceitando-se que cada *locus* contribui modestamente para a etiologia da doença.

A maioria das variações na sequência do genoma humano são devidas a SNPs, sendo o resto atribuível a inserções ou deleções de uma ou mais bases, polimorfismos repetidos e rearranjos.

Em contraste com marcadores mais mutáveis, tais como os microsátélites (34), os SNPs têm uma baixa taxa de mutação recorrente, tornando-os assim indicadores estáveis de história humana.

Refere-se seguidamente a distribuição de SNP nos cromossomas humanos (35).

### Distribuição de SNP nos cromossomas humanos

<u>Nº. do cromossoma</u>	<u>Comprimento em bp</u>	<u>Nº. de SNP</u>
1	214.066.000	129.931
2	222.889.000	103.664
3	186.938.000	93.140
4	169.035.000	84.426
5	170.954.000	117.882
6	166.022.000	96.317
7	149.414.000	71.757
8	125.148.000	57.834
9	107.440.000	62.013
10	127.894.000	61.298
11	129.193.000	84.663
12	125.198.000	59.245
13	93.711.000	53.093
14	89.344.000	44.112
15	73.467.000	37.814
16	74.037.000	38.735
17	73.367.000	34.825
18	73.078.000	45.135
19	56.044.000	25.676
20	63.317.000	29.478
21	33.824.000	20.916
22	33.786.000	28.410
X	131.245.000	34.842
Y	27.753.000	4.193

Dependendo da localização genómica, a consequência fenotípica de uma SNP pode variar. Assim se a SNP tem sede numa região codificante dos genes, a sequência em ácidos aminados da proteína codificada pode ser alterada, afectando assim a sua estrutura e função.

Mas se a SNP tem lugar em regiões reguladoras de um gene pode ser afectada a ligação de factores de transcrição, sendo assim influenciada a taxa de expressão do gene.

A maioria das SNP ocorre em regiões não codificantes do genoma não tendo assim impacto conhecido no fenotipo respectivo.

Os SNP podem ser úteis como marcadores genéticos em identificação forense, na tipagem de tecidos, em estudo da evolução de populações genéticas assim como nas doenças genéticas.

Nas desordens monogenicas uma simples mutação pontual posicionada habitualmente em regiões codificantes ou em regiões reguladoras, é uma causa suficiente para originar doenças.

Nos humanos são conhecidas 1.600 doenças monogenicas.

O efeito combinado de SNP em genes codificadores de enzimas metabolizadoras de medicamentos, transportadores destes, ou seus receptores, é responsável pela diversidade das respostas de

cada indivíduo ao tratamento com idênticos agentes terapêuticos, quer no que se refere à eficácia do tratamento, quer aos efeitos secundários desencadeados.

Hoje análises farmacogenéticas rotineiras nos humanos, incidem na tipagem dos genes responsáveis pela metabolização de agentes terapêuticos, como é o caso da família de genes pertencentes ao citocromo P450.

Contudo as doenças complexas devidas à variação genética de diversos genes, actuando só por si ou em conjunto com factores ambientais, são mais frequentes do que as doenças monogénicas.

Os mais importantes princípios em que se baseiam as reacções utilizadas na genotipagem, são a hibridização com sondas com sequências oligonucleotídicas específicas, e a utilização de enzimas modificadoras do DNA como as DNA polimerase e ligases.

A utilização de SNP como marcadores de doenças, tem um largo futuro no diagnóstico clínico.

Através da genotipagem do DNA, a presença de variações genéticas ligadas a doenças pode ser facilmente averiguada, tal como a farmacogenética irá permitir um mais eficiente “desenho” do agente terapêutico e da forma de tratamento.

## **5.5 - Organização dos genes nos cromossomas**

### **Famílias de genes e sua classificação**

A organização dos genes nos cromossomas revelou nos eucariotas superiores que os genomas contêm muito DNA não funcional (que por exemplo não codifica proteínas).

Por outro lado também se verificou que a quantidade total de DNA cromossomal nos diferentes animais, não varia de forma consistente com a aparente complexidade dos organismos.

Muito do DNA em certos organismos não codifica RNA nem tem qualquer função reguladora ou estrutural.

Muitos cromossomas eucariotas têm quantidades variáveis de DNA sem qualquer função demonstrável, incluindo-se nestas fracções extensões repetidas de DNA algumas das quais nunca são transcritas e parecendo a maioria delas perfeitamente dispensáveis.

Referem-se no quadro adiante as diferentes classes de sequências de DNA eucariotas.

Cerca de um quarto a metade de todos os genes eucariotas codificadores de proteínas são solitários ou seja representados uma só vez no genoma haplóide.

As famílias de genes são formadas pela duplicação de genes e codificam proteínas homólogas. Os pseudogenes são genes duplicados que se tornaram não funcionais.

As famílias de genes formadas pela duplicação de genes codificando proteínas homólogas, provêm provavelmente de um gene ancestral comum e são designados genes codificadores de proteínas duplicados.

Admite-se que nos genomas dos vertebrados estes genes duplicados constituem metade do DNA que codifica proteínas.

O conjunto de genes duplicados que codificam proteínas com sequências de ácidos aminados similares, mas não idênticas é chamado família de genes e as proteínas afins produzidas ou parecidas entre si ou proteínas homólogas constituem uma família de proteínas.

Há famílias de proteínas com vinte membros ou até centenas tal como se refere no quadro seguinte – (*vide* ainda adiante genes codificadores de proteínas e proteoma).

Famílias de proteínas nos vertebrados

Proteínas comuns a vertebrados e invertebrados	{	<b><u>Famílias</u></b>	<b><u>Nº. de proteínas na família</u></b>
		Actinas	5 – 30
		Proteínas de choque térmico 70k	3
		Ceratinas	20
		Miosina, cadeia pesada	5 – 10
		Proteína cinases	10 – 100
		Factores de transcrição	10 – 100 ( ? )
		Tubulinas, _ e _	3 – 15
		Globinas (diversas espécies)	
		_ – globina	1 – 3
		_ – “like” globinas	5
		Imunoglobulinas, regiões variáveis (várias espécies)	500
		Ovalbumina (galinha)	3
		Antigénios de transplantação (rato e humano)	50 – 100
		Proteína do pigmento-visual (humano)	4
		Vitelogenina (galinha)	5

Pseudogenes são genes duplicados que se tornaram não funcionais, mantendo a mesma aparente estrutura exões-intrões que os genes funcionais correspondentes, resultando provavelmente de um gene ancestral comum, mas durante a evolução a acumulação de sequências que terminam a translação ou que bloqueiam o processamento de mRNA tornaram tais regiões não funcionais. Tais genes estão assinalados na família das globinas, tubulinas, actinas, etc.

Genes que se repetem em tandem são responsáveis pela formação de rRNA, tRNA e das histonas. Estes genes distinguem-se dos genes duplicados de famílias de genes por terem múltiplas repetições em tandem que codificam idênticas ou quase idênticas proteínas ou RNAs funcionais.

A maioria destas cópias da sequência aparecem umas a seguir às outras.

Estes genes repetidos em tandem para o rRNA, tRNA e histonas servem para responder a uma grande solicitação pelas células, destes transcritos. Genes codificando diversos RNA funcionais, outros

que os mRNA, existem em múltiplas cópias nas células eucariotas. Nos humanos existem cerca de 250 cópias de genes de pré-rRNA, 2.000 cópias de genes 5S rRNA e 1.300 cópias de genes tRNA.

Nas frações de DNA repetidas “repetitious” verificou-se que existiam sequências, deste DNA, muito curtas e que ocorriam em repetições em tandem, enquanto outras eram muito mais compridas e encontravam-se interdispersas em diversos locais do genoma.

#### Classificação de DNA eucariota

- DNA codificando proteínas
  - Genes solitários
  - Genes duplicados e genes que “divergiram” (famílias de genes funcionais e pseudogenes não funcionais)
- Genes repetidos em Tandem
  - Codificadores de rRNA, 5srRNA, tRNA e alguns genes de histonas
- Genes repetidos (similar mas não identico, “repetitious”)
  - DNA de sequência simples
  - DNA de repetição intermediada (elementos moveis de DNA)
    - Curtos elementos interdispersos (retrotransposões não víricos)
    - Longos elementos interdispersos (retrotransposões víricos e não víricos)
- DNA espaçador não classificado

Se fragmentar-mos o DNA total de um organismo em fragmentos com um tamanho médio de cerca de 1.000 pares de bases e se depois desnaturarmos esses fragmentos em hélices simples e os colocarmos seguidamente em condições de renaturação (reassociação) todos os fragmentos de DNA retomarão a forma de duplexes sensivelmente com a mesma rapidez, se não houver sequências repetidas no genoma.

Contudo um segmento contendo uma sequência repetida várias vezes no genoma poderá encontrar uma hélice “partner” mais rapidamente do que um segmento com uma sequência que ocorra uma só vez por genoma haplóide, uma vez que a sequência repetida se encontra presente em muito mais elevada quantidade. Logo sequências repetidas reassociam-se muito mais rapidamente do que um fragmento de sequência única ou não repetida. Assim o DNA codificando pré-rRNA e 5SrRNA reassocia-se mais depressa do que o DNA não repetido.

Experiências deste tipo de reassociação revelaram três tipos principais de DNA nos mamíferos estudados, podendo no entanto as proporções em que se encontram várias um tanto; a saber:

1. DNA de reassociação lenta
2. DNA de sequência simples de reassociação rápida
3. DNA com velocidade de reassociação intermediada

1) DNA de reassociação lenta (DNA de cópia única)

Representam cerca de 50-60% do DNA dos mamíferos. Cada cópia única de sequência do DNA não realiza necessariamente uma função genética só uma pequena fracção da totalidade do DNA humano, cerca de 5% codifica para proteínas ou moléculas de RNA funcionais. O resto do DNA de cópia única que aparentemente não tem função conhecida, além de separar as sequências de DNA funcionais, designa-se por DNA espaçador (“spacer”).

2) DNA de sequência simples de reassociação rápida

Cerca de 10-15% do DNA dos mamíferos reassocia-se rapidamente sendo designado esta fracção como DNA de sequência simples. Esta sequência simples de DNA é composta grandemente de diversos e diferentes conjuntos de curtos oligonucleótidos, 5 a 10 pares de bases, embora nos vertebrados possam conter 20-200 nucleótidos, repetidos em conjuntos de longos tandens.

Os eucariotas superiores contêm vários tipos de DNA de sequências simples que geralmente ocorrem em extensões muito longas de DNA até  $10^5$  bp de comprimento. Estas sequências longas de DNA de sequências simples que é facilmente distinguindo e separado do DNA celular de cópia única, é por vezes designado como DNA satélite em virtude das distintas classes de densidades do DNA.

DNA satélite:

<b>Designação das repetições</b>	<b>Tamanho</b>	<b>Tamanho total</b>	<b>Outras características</b>
<b>1 – Satellite x</b> ( satellite)	5-200pb	Até vários milhões de pb	Encontra-se na heterocromatina e nos centromeros. Não é transcrito
<b>2 – Minisatélites</b> a) Família hipervariável <b>xx</b>	10-60pb	1.000 a 20.000 pb	Tem uma sequência (motivo) comum GGGGAGGANG (N/= qualquer base dispersa, repetição em tandem
b) Família Telomérica <b>xxx</b>	6pb	1.000 a 20.000 pb	Habitualmente TTAGGG é repetida cerca de mil vezes. Protegeas extremidades cromossomais
<b>3 – Microsatélite</b> <b>xxxx</b>	1-4pb	Menos de 1.000 pb	Repetições de A e CA são as mais frequentes. Dispersas através do genoma

x – O DNA satélite normalmente não é transcrito e forma a maioria da heterocromatina.

xx – A família minisatélite hipervariável tem uma sequência motivo comum e encontra-se dispersa. O número destas repetições num dado *locus* varia constituindo um método para identificação.

xxx – A família telomérica de minisatélites protegem as extremidades dos cromossomas.

xxxx – Os microsatélites consistem de sequências repetidas de dois nucleótidos. As variações no número de repetições dos minisatélites (regiões) varia de animal para animal com excepção dos gémeos idênticos, e constitui a base do respectivo “fingerprint” (*vide* adiante).

Cada animal tem vários tipos de DNA de sequência simples caracterizados por diferentes unidades de repetição. Em algumas espécies as unidades oligoméricas repetidas de DNA de sequência simples têm sequências muito similares, o que sugere um ancestral comum.

Nos humanos existem pelo menos dez tipos de DNA de sequência simples.

A maioria do DNA de sequência simples está localizado em regiões específicas dos cromossomas. No rato e na maioria dos outros mamíferos, predominam embora não na totalidade próximo dos centromeros.

Diferenças nos comprimentos das sequências simples de conjuntos em tandem permitem-nos o “fingerprint” do DNA.

Longas extensões de DNA de sequência simples dentro de uma espécie exibem uma notável conservação na sua sequência nucleotídica. Pelo contrário, diferenças no número de repetições, e assim no comprimento de conjuntos de sequências simples em tandem contendo a mesma unidade repetitiva, são muito comuns. Estas diferenças no comprimento resultam de “crossing over” desiguais dentro de regiões de DNA de sequências simples durante o desenvolvimento de precursores dos espermatozóides e do oócito durante a meiose. Em consequência deste diferente “crossing over” conjuntos de sequência simples em tandem em indivíduos diferentes diferem no comprimento, uma característica que constitui a base para o “fingerprint” do DNA, uma técnica que permite identificar amostras de DNA individual.

Nos mamíferos, algum do DNA de sequência simples existe em regiões relativamente curtas de 1 a 5kb constituídas por 20 a 50 unidades repetidas cada uma contendo 15 a 100 pares de bases.

Estas regiões são chamadas minisatélites para as distinguir de outras regiões mais comuns também de DNA de sequência simples referidas em tandem as quais têm  $10^5$  a  $10^6$  pares de bases de comprimento.

Ligeiras diferenças nos comprimentos totais de minisatélites diferentes de indivíduos distintos podem ser detectados pela análise convencional de Southern-blot.

Para se obter o “fingerprint” do DNA de um indivíduo, uma amostra do seu DNA é submetida à acção de enzimas de restrição que não cortem o interior da sequência dos minisatélites. Os fragmentos obtidos são sujeitos a análise por Southern-blot com sondas radioactivas minisatélites (*vide* figura seguinte). Análises em duplicado com duas ou quatro sondas minisatélite diferentes são suficientes para identificar o “finger print” de um DNA que é único para cada indivíduo o que pode ter profundas implicações em medicina forense.

### 3) DNA com velocidade de reassociação intermediada(9)

Este DNA com repetição intermediada ou moderada representa 25-40% do DNA dos mamíferos e é composto por um grande número de cópias de famílias com algumas sequências e são interdispersos através do genoma dos mamíferos e são classificados em dois tipos:

- Elementos de curtas interdispersões ou SINES contendo 150-300 pares de bases.
- Elementos de longas interdispersões ou LINES com 5.000-7.000pb de comprimento.

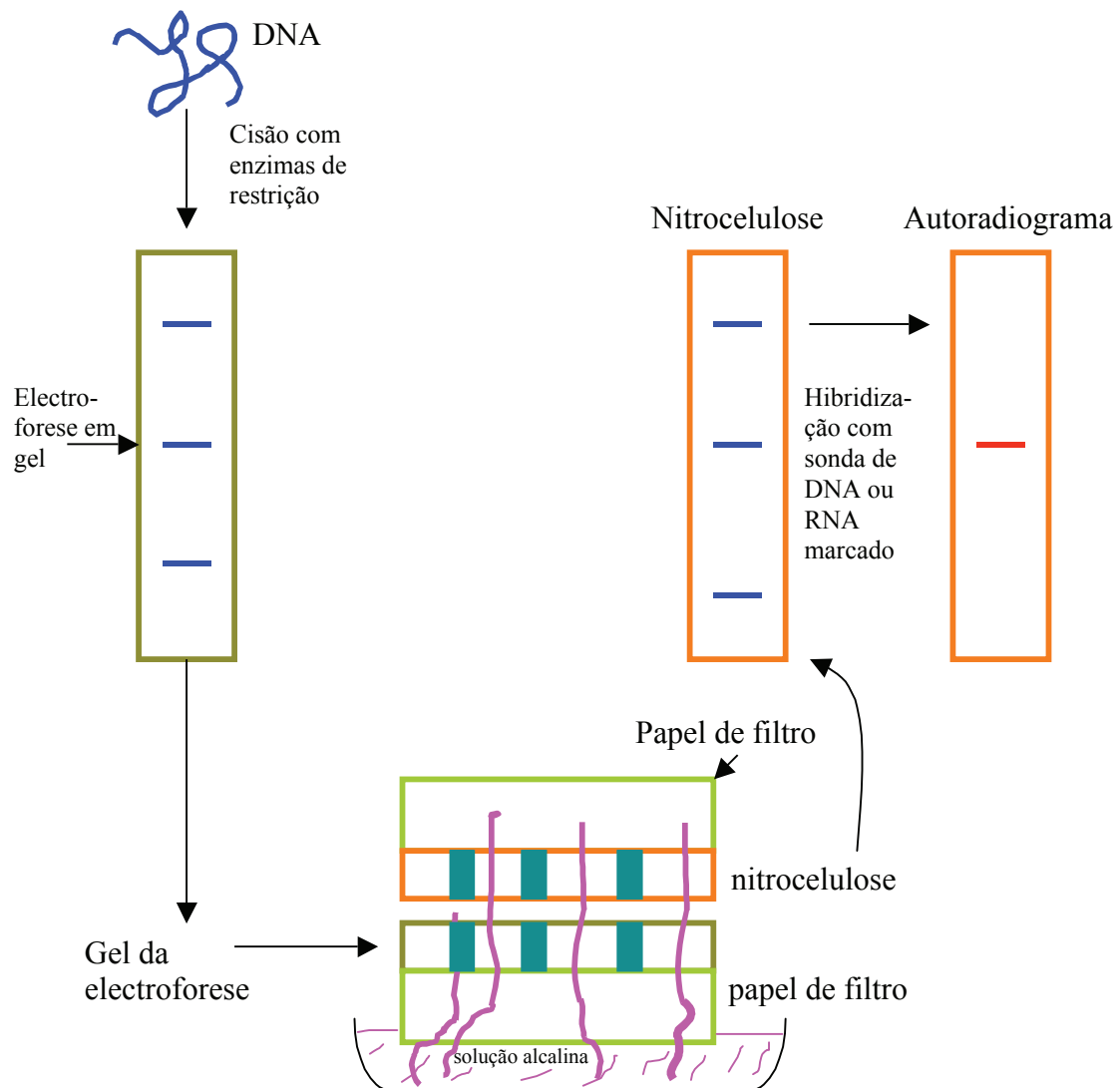
As cópias múltiplas de SINES e LINES podem não ser repetições exactas e nem os SINES nem a vasta maioria dos LINES codificam proteínas.

Nos genomas mamíferos foi encontrada uma classe principal de SINES e talvez cerca de 10 classes de LINES encontrando-se de cada classe milhares de cópias.

Vário deste DNA com sequências repetidas intermediadas pode movimentar-se ou transpor (“transpose”) para novos locais no DNA celular, daí serem chamados elementos transposáveis ou elementos de DNA móvel.



Técnica de Southern – blot para detectar seqüências específicas de DNA após separação electroforetica de uma mistura de fragmentos de restrição



Como já referimos 25 a 40% do DNA dos mamíferos renaturam-se a uma velocidade intermédia, constituindo, uma pequena porção desta fracção, os genes repetidos em tandem codificantes de rRNA, tRNA e das histonas e genes relacionados com estes, formando várias famílias de genes.

Nesta fracção estão também incluídas sequenciais de DNA moderadamente repetitivas e interdispersas através do genoma e capazes de se moverem, ou transporem para novos locais no genoma.

Estes elementos móveis de DNA são essencialmente parasitas moleculares que não parecem ter qualquer função específica e que existem apenas para se manter a si próprios. Eles são eliminados nos eucariotas a uma velocidade muito baixa, por deleção dos segmentos de DNA que os contêm, podendo assim acumular-se durante a evolução. Nos seres humanos os elementos móveis de DNA representam cerca de 30% do DNA genómico total.

No entanto apesar destes elementos móveis de DNA não terem funções directas, a sua presença pode ter impacto significativo na evolução, pois podem por exemplo inserir-se dentro ou próximo de uma unidade de transcrição e originar uma mutação espontânea, tal como recombinações homólogas entre elementos móveis dispersos através dos genomas podem originar rearranjos durante a evolução.

Há provas de que durante a evolução dos eucariotas superiores ocorreram recombinações entre intrões de genes distintos, gerando novos genes feitos de novas combinações dos exões pré-existentes.

Os termos “exon shuffling” refere-se a este tipo de processo evolutivo.

Os elementos móveis podem ser transpostos por um de dois mecanismos:

- a) directamente como DNA (transposões)
- b) através de um RNA intermediário transcrito (retrotransposição) de um elemento móvel por uma RNA polimerase e depois convertido por uma transcriptase reversa numa dupla hélice de DNA (*vide* figura seguinte).

Há pois duas categorias principais de retrotransposões nas células eucariotas.

Nos humanos estão assinalados retrotransposões em centenas de milhares de locais dispersos no genoma.

As duas categorias de retrotransposões são as retrotransposões víricas e não víricas. Os primeiros exibem semelhanças com os genomas dos retrovírus e contêm repetições directas terminais de cerca de 250 a 600 bp, também chamados “long terminal repeats” ou LTR. Os segundos não têm LTR e daí serem designados retrotransposões não víricas.

Nos mamíferos os mais abundantes destes assemelham-se a moléculas de RNA celular que são sintetizadas pela RNA polimerase III.

Nos mamíferos LINES e SINES são os elementos móveis mais abundantes sendo retrotransposões não víricas.

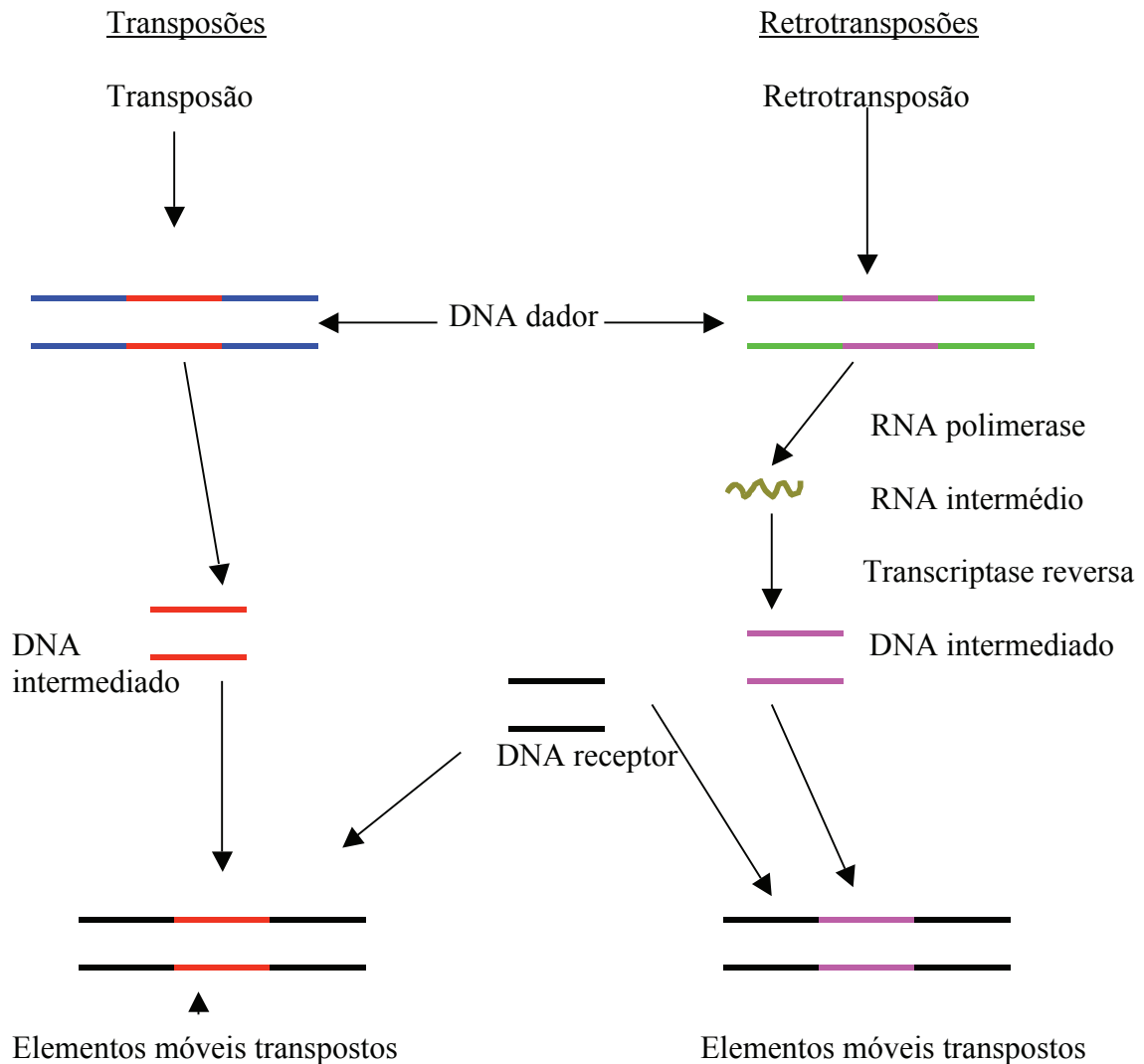
Nos humanos o comprimento dos LINES é de 6-7kb enquanto nos SINES é de cerca de 300bp e apesar destas sequências de DNA não conterem LTRs diversas provas indicam que elas são transpostas através da transcrição reversa de um RNA intermediário como veremos adiante.

Sequências repetidas com características de LINES são particularmente abundantes nos genomas mamíferos e os SINES também.

A família L1 LINES é muito frequente no genoma humano onde existem cerca de 50.000 cópias representando cerca de 5% do DNA humano total.

As curtas seqüências designadas por SINES encontram-se em 500.000 ou mais locais no genoma humano, representando cerca de 5% deste, encontrando-se SINES com seqüências relacionadas também com alta freqüência nos genomas de outros mamíferos.

Classificação do mecanismo de transposição de elementos móveis



Diversas centenas de exemplos de SINES de vários mamíferos tem sido seqüenciadas e clonadas e apesar de duas cópias destes “intermediate repeats” não serem idênticas, a sua similaridade na seqüência geral revela um ancestral comum dentro da espécie e entre as espécies. A conservação da seqüência é de cerca de 80% dentro de uma espécie, mas cai para 50-60% de espécie para espécie dentro dos mamíferos. Como muitas destas seqüências repetitivas contêm um local de reconhecimento para o enzima de restrição Alu I, elas são chamadas colectivamente família Alu.

Contudo como estes muito curtos elementos interdispersos não são precisamente idênticos a muitos falta o local Alu.

Tal como noutros elementos móveis as sequências Alu habitualmente são flanqueadas por “direct repeats” (repetições directas), não codificam proteínas e são transcritas pela RNA polimerase III não parecendo as sequências Alu ter qualquer função.

Cópias retrotransposáveis de RNAs celulares encontram-se presentes nos cromossomas eucariotas. Cópias mutadas de DNA de uma larga variedade de mRNAs também têm sido descobertas integradas no DNA cromossomal.

Calcula-se que cerca de 43% de genoma humano é ocupado por quatro classes de elementos repetitivos interdispersos:

- 1) Curtos elementos interdispersos (SINEs)
- 2) Longos elementos interdispersos (LINEs)
- 3) Elementos com repetições terminais longas ( elementos LTR)
- 4) Transposões de DNA

No rato e nos humanos existe um forte aumento da densidade de SINES e uma diminuição na densidade de L1 á medida que o conteúdo (G+C) aumenta.

Há mais de 4,3 milhões de elementos repetidos no genoma humano, com Alu e LINE 1 ( L1 ) sendo os mais frequentes.

A inserção de elementos repetidos num gene não é deletéria, e vários elementos repetidos translidos são encontrados em diversas proteínas.

No quadro seguinte referem-se o número encontrado de elementos repetitivos no genoma humano.

Elementos repetitivos no genoma humano(Li e al 2001.Nature,409 ,847).

Tipo	Número encontrado	% do genoma
SINE (todos)	1.404.300	12,5%
Alu	1.010.400	10,7%
LINE (todos)	1.045.800	18,9%
L 1	661.000	15,4%
DNA transposões	308.800	2,7%
LTR	531.900	7,9%
Outros	7.300	0,1%
Total	3.959.200	42,5%

Conteúdo repetido no genoma (9)

Tal como já referimos o tamanho do genoma não está correlacionado com a complexidade dos seres vivos o que parece ser devido ao facto dos genomas conterem grandes quantidades de sequências repetidas além daquelas que correspondem aos genes codificadores de proteínas.

Com efeito estes últimos correspondem a menos de 5% do genoma total enquanto as sequências repetidas representam 50% ou mais deste, sendo estas sequências repetidas distribuídas em cinco classes, a saber:

- 1) Repetições derivadas de transposões também por vezes designadas com repetições interdispersas.
- 2) Cópias “retroposed” inactivas (parcialmente) de genes celulares inclusivé genes codificadores de proteínas e pequenos RNAs estruturais, habitualmente designados por pseudogenes processados.
- 3) Repetições de sequências simples (SSR) constituídas por repetições directas de relativamente curtas k-mers (sequências nucleotídicas) tais como (A) n, (CA) n ou (CGG) n.
- 4) Duplicações segmentais contendo blocos de cerca de 10-300kb copiados de uma região do genoma para outra.
- 5) Blocos de sequências repetidas em tandem, como nos centromeros e telomeros, nos braços curtos de cromossomas acrocentricos, (cromossomas cujo centromero está próximo de uma extremidade) e em conjuntos de genes ribossomais.

Nos mamíferos quase todos os elementos transposáveis, são enquadráveis em quatro tipos (LINES, SINES, Elementos *retrovirus like* e DNA transposões) dos quais três são transpostos através de RNA intermediários e um directamente como DNA havendo pois elementos longos interdispersos (LINES), elementos curtos interdispersos (SINES), LTR retrotransposões e DNA transposões.

Os transposões LINES com cerca de 6kbp de comprimento albergam um promotor polimerase II interno e codificam duas ORF (*open reading frame*). Após translação um RNA LINE reúne-se com as próprias proteínas por si codificadas e movimenta-se para o núcleo onde uma endonuclease faz um corte simples numa hélice do DNA e uma transcriptase reversa utilizando esta hélice cortada de DNA faz a transcrição reversa a partir da extremidade 3' do RNA LINE. Esta transcrição reversa muitas vezes não consegue fazer o processamento até á extremidade 5' resultando assim diversas inserções truncadas não funcionais. Contudo a maioria das repetições LINE derivadas são curtas com um tamanho médio de 900bp.

No genoma humano estão assinaladas três famílias LINE distintas umas das outras, a LINE 1, LINE 2 e a LINE 3, só a primeiro sendo activa.

As SINES são curtas (cerca de 100-400bp) e carregam um promotor interno da polimerase III derivado de sequências de tRNA e não codificam proteínas, utilizando a maquinaria LINES para efectuar a transposição.

No genoma humano estão assinaladas três famílias SINES, a activa Alu, e a inactiva MIR e Ther2/MIR3.

Os retrotransposões LTR são flanqueados por longas repetições terminais directas que possuem todos os elementos reguladores de transcrição necessários. Os elementos autónomos (retrotransposões) contêm genes gag e pol que codificam uma protease, uma transcriptase reversa, RNase H e integrase.

Os retrovírus exógenos parecem ter resultado de retrotransposões endógenos por mecanismos retrovíricos com transcrição reversa ocorrendo numa partícula citoplásmica *like-virus* derivada de um tRNA (contrariamente à localização nuclear e preparação cromossomal das LINES).

Existe uma diversidade de LTR retrotransposões, embora apenas retrovírus endógenos específicos dos vertebrados (ERVs) pareçam ser activos no genoma dos mamíferos.

Os retrovírus mamíferos podem ser agrupados em três classes (I-III). Como já referimos um pseudogene é uma cópia não funcional muito similar a um gene normal mas que foi alterado ligeiramente não se exprimindo em consequência disso.

As características estruturais gerais destes pseudogenes processados incluem a perda completa das sequências intervenientes encontradas nos genes funcionais, um tracto poli (A) na extremidade 3' e repetições directas flanqueando a sequência do pseudogene.

Pseudogenes processados ocorrem em consequência da retrotransposição, enquanto pseudogenes não processados provêm de uma duplicação segmental do genoma.

Estas duplicações segmentais de porções da sequência genómica envolvem a transferência de blocos de 1-200kbp para um ou mais locais do genoma, sendo os locais das regiões dadora e recipiente não dispostos em tandem sugerindo que os mecanismos implicados sejam outros que não um “crossing-over” desigual.

Há duas categorias destas duplicações segmentais. A primeira categoria são duplicações intercromossomais em que segmentos são duplicados entre cromossomas não homólogos, sendo a segunda categoria para as duplicações intracromossomais, dentro de um dado cromossoma ou braço de cromossoma.

As regiões pericentroméricas dos cromossomas próximas dos centromeros consistem quase inteiramente de segmentos duplicados intercromossomais, com poucas ou nenhuma sequências únicas.

Regiões mais pequenas de duplicação intercromossomais são também observadas próximo dos telomeros (regiões subteloméricas).

As regiões pericentroméricas são estruturalmente muito complexas.

As repetições de sequências simples (SSR) são estruturas repetitivas perfeitas ou levemente imperfeitas de repetições em tandem de um particular K-mer (determinada sequência nucleotídica). Os SSR com uma unidade de repetição curta ( $n \leq 1-13$  bases) são por vezes chamados microsátélites, enquanto se a unidade repetitiva é mais comprida ( $n \geq 14-500$  bases) são chamadas minisátélites.

Com excepção das caudas poliA das mensagens transcritas reversamente, os SSR provêm da “slippage” (deslize entre duas hélices do DNA) (ver estrutura do DNA e seu funcionamento) durante a replicação do DNA (36,37).

Os transposões DNA assemelham-se aos transposões bacterianos, possuindo repetições terminais invertidas e codificando uma transposase, que se liga próximo das repetições invertidas.

O genoma humano contém pelo menos sete classes principais de transposões DNA que podem ser subdivididas em diversas famílias tendendo a ter vidas curtas dentro das espécies.

Resumindo quanto ao conteúdo em genes do genoma humano, os genes (ou pelo menos as suas regiões codificantes) são apenas uma pequena fracção do DNA humano.

Será importantíssimo compilar uma lista completa de todos os genes humanos bem como das proteínas por eles codificados para serem utilizados como uma “tabela periódica” pela investigação biomédica (38) o que constitui uma tarefa muito difícil.

Os genes humanos tendem a possuir pequenos exões separados por longos intrões (alguns excedendo 10kbp) o que leva a que os programas de computadores para a predição directa de genes tenham apenas uma certeza limitada ou relativa. Por outro lado milhares de genes humanos produzem RNAs não codificadores (ncRNA) (39).

### RNAs não codificadores (nc RNA) (9)

Existem diversas categorias de nc RNA, assim:

- a) tRNA adaptadores na translação.
- b) rRNA importantes na maquinaria de translação e recentes provas de cristalografia assinalam que a formação da ligação peptídica é catalisada por rRNA e não por proteínas (40,41).
- c) RNA pequeno nucleolar (s ncRNA) necessário para o processamento do rRNA e modificação de bases no nucléolo (42-43).
- d) RNA pequeno nuclear (snRNA) crítico como componente dos spliceossomas, os grandes complexos ribonucleoproteicos (RNP) que cortam os intrões do pré-mRNA no núcleo. Estão assinalados nos humanos um spliceossoma mais importante o U2 sn RNA que corta a maioria dos intrões, e um spliceossoma menor o U12 snRNA que corta um tipo raro de intrão que por vezes tem AT/AC dinucleótidos no lugar de corte em vez dos canónicos GT/AG.

Incluem-se ainda nos ncRNA outros RNA de funções conhecidas como a RNA telomerase e a partícula RNA reconhecadora de sinal 7SL e ncRNA de funções enigmáticas.

Os nc RNA não têm ORF traduzidas, são pequenos e não são poliadenilados. Referem-se seguidamente as características de alguns genes implicados na formação de ncRNA.

#### a) Genes RNAt

As experiências clássicas calculavam cerca de 1.310 genes tRNA nos humanos apesar de no “draft” apenas terem sido encontrados 497.

Estes genes tRNA encontram-se dispersos através do genoma mas não é uma dispersão ao acaso, havendo alguns tRNA agrupados em “clusters”.

Nos humanos mais de 25% dos genes tRNA encontram-se numa região de apenas cerca de 4Mb no cromossoma 6. Esta pequena região correspondendo apenas a 0.1% do genoma, contem um conjunto de genes tRNA suficientes para todo o genoma.

Nos humanos mais de metade de todos os genes tRNA (280 em 497) tem sede no cromossoma 1 ou no cromossoma 6, enquanto os cromossomas 3, 4, 8, 9, 10, 12, 18, 20, 21 e X parecem ter menos de 10 genes tRNA cada um, e os cromossomas 22 e Y não tem nenhum.

#### b) Genes RNA ribossomais

As duas sub-unidades ribossomais contêm quatro espécies de rRNA e diversas proteínas.

A grande sub-unidade ribossomal (LSV) contem rRNA 28S e 5.8S e também 5s rRNA.

A pequena sub-unidade (SSV) contem 18S rRNA. Os genes para os rRNA da LSV e SSV ocorrem no genoma humano como uma unidade repetitiva em tandem de 44kb, pensando-se que existam 150-200 cópias desta unidade repetida nos braços curtos dos cromossomos acrocentricos 13, 14, 15, 21 e 22 (44,45). Os genes 5s rRNA ocorrem em tandem, o maior dos quais encontra-se no cromossoma 1 humano entre 1q41.11 e 1q42.13 próximo do telomero (46,47) havendo 200-300 genes verdadeiros 5s nestes conjuntos. O número de sequências relacionadas com 5s no genoma inclusive numerosos pseudogenes dispersos são de 2.000 (48).

c) Pequenos genes RNA nucleolares

Os rRNA eucariotas são profundamente processados e modificados no nucléolo (s nc RNA) sendo distribuídos em duas famílias: a caixa C/D snc RNA envolvida em metilações 2'-O-ribose de outros RNA e a H/ACA snc RNA envolvida em pseudouridilação.

d) Genes RNA de spliceossomas e outros genes ncRNA

Encontram-se 44 genes dispersos para o U6sn RNA e 16 para o U1 sn RNA.

## 5.6 - Conclusão

### Algumas características do Genoma humano global

( science, 291, 16-2-01 p 1327 )

Tamanho do genoma (incluindo "gaps")	2,91Gbp
Tamanho do genoma (excluindo "gaps")	2,66Gbp
"contig" mais comprido	1,99Mbp
"scaffold" mais comprido	14,4Mbp
Percentagem de A+T no genoma	54
Percentagem de G+C no genoma	38
Percentagem de bases não determinadas no genoma	9
Mais rico em GC em 50kb	Cromoss. 2(66%)
Menos rico em GC em 50kb	Cromoss. X(25%)
Percentagem do genoma classificado como repetido	35
Número de genes anotados	26.383
Número de genes anotados com funções desconhecidas	42
Número de genes (hipotéticos e anotados)	39.114
Percentagem de genes hipotéticos e anotados com função desconhecida	5 <sup>9</sup>
Gene com mais exões	Titina (234 exões)
Tamanho médio dos genes	27kb
Cromossoma mais rico em genes	Cromoss. 19(23genes/Mb)
Cromossoma menos rico em genes	Cromoss. 13(5genes/Mb) Cromoss. Y(5genes/Mb)
Dimensão total não ocupada por genes (>500kb sem genes anotados)	605Mbp
Percentagem de bases abarcadas pelos genes	25,5 a 37,8*
Percentagem de bases abarcadas pelos exões	1,1 a 1,4*
Percentagem de bases abarcadas pelos intrões	24,4 a 36,4*
Percentagem de bases do DNA entre genes	74,5 a 63,6*



Cromossoma com mais alta proporção de DNA em exões anotados	Crom.19(9.33)
Cromossoma com mais baixa proporção de DNA em exões anotados	Crom.Y(0.36)
Região entre genes mais longa (entre genes anotados+genes hipotéticos)	Crom13(3.038.416bp)
Proporção de variação SNP	1/1250bp

\* Nestes casos as percentagens correspondem ao conjunto de genes anotados (26.383 genes) e ao conjunto de genes hipotéticos + genes anotados (39.114 genes) respectivamente.

Pontos claros referentes ao genoma humano e de alguns vertebrados  
(Nature, 409,15-2-01,p860)

- Há notáveis diferenças no genoma no que se refere a uma série de aspectos tais como genes, elementos transferíveis, conteúdo em GC, ilhotas CpG e grau de recombinação.
- Calcula-se a existência de 30.000 a 40.000 genes codificadores de proteínas, sendo os genes mais complexos do que em alguns vertebrados pois possuem mais locais alternativos para corte o que origina maior número de proteínas.
- O proteoma também parece ser mais complexo nos vertebrados do que nos invertebrados.
- Cerca de metade do genoma deriva de elementos transferíveis. Os transposões de DNA parece terem-se tornado completamente inactivos tal como os LTR (*long terminal repeat*).
- As regiões pericentroméricas e subteloméricas dos cromossomas parecem ter sido preenchidas com grandes duplicações de segmentos de sequências provenientes de qualquer lado do genoma, sendo muito mais frequente nos vertebrados do que nos invertebrados.
- As grandes regiões pobres em GC estão correlacionadas nos cariótipos com bandas G escuras.
- As recombinações tendem a ser mais elevadas nas regiões distais dos cromossomas (à volta de 20 megabases – Mb) e nos braços mais curtos dos cromossomas em geral.
- Foram identificados mais de 1,4 milhões de polimorfismos de simples nucleótidos (SNP).

## 5.7 - Bibliografia

1. Ewing, B. & Green, P. Analysis of expressed sequence tags indicates 35,000 human genes, (2000). Nature Genet 25, 232-234.
2. Roest Collins, H. et al. Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence (2000). Nature Genet 25, 235-238.
3. McKusick, V.A. Mendelian Inheritance in Man: Catalog of Human Genes and Genetic Disorders. Johns Hopkins Univ. Press Baltimore, 1998.
4. Maglott, D.R., Katz, K.S., Sicolte, H. & Pruitt, K.D. NCBI's Locuslink and RefSeq (2000). Nucleic Acids Res. 28, 126-128.
5. Pruitt, K.D., Katz, K.S., Sicolte, H & Maglott, D.R. Introducing Ref. Seq. And Locuslink: curated human genome resources at the NCBI (2000). Trends Genet. 16, 44-77.
6. Guigo, R., et al. An assessment of gene prediction accuracy in large DNA sequences (2000). Genome Res. 10, 1631-1642.
7. Stormo, G.D. Gene-finding approaches for eukaryotes (2000). Genome Res. 10, 394-397.
8. Wolfsberg Tyra G. Et al. Guide to the draft human genome (2001). Nature 409, 824-826.
9. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome (2001). Nature 409, 860-921.
10. Birney E., et al. Mining the draft human genome (2001). Nature 409, 827-828.
11. Barbozuk, W.B. et al. The syntenic relationship of the zebra fish and human genomes (2000). Genome Res. 10, 1351-1358.
12. McLysaght, A., et al Estimation of synteny conservation and genome compactation between pufferfish ( *Fugu* ) and human(2000). Yeast 17, 22-36.
13. Trachtulec, Z., et al. Linkage of TATA-binding protein and proteasome subunit C5 genes in mice and humans reveals synteny conserved between humans and invertebrates (1997). Genomics 44, 1-7.
14. Nadeau, J.H. Maps of linkage and synteny homologies between mouse and man (1989). Trends Genet 5, 82-86.
15. Nadeau, J.H. & Taylor, B.A. Lengths of chromosomal segments conserved since divergence of man and mouse (1984). Proc. Nat. Acad. Sci. USA 81, 814-818.
16. Novacek, M.J., Mammalian phylogeny shaking the tree (1992) Nature 365,121-125.
17. O'Brien, S.J. et al Genome maps to comparative genomics Mammalian radiations. Wall Chart, (1999). Science 286, 463-478.
18. Rubin, G.M. Comparing species (2001). Nature 409, 820-821.
19. Miklos, G.L. and John, B (1979) Am. J. Hum. Genet 31, 264.
20. Horvath, J.E., Schwartz, J. And Elohler, E.E. (2000) Genome Res. 10, 839.
21. Holmquist, G.P. (1992) Am. J. Hum. Genet 51, 15.
22. Bernardi, G. (2000) Gene 241, 3.
23. Zoubak, S. et al. (1996) Gene 174, 95.
24. McEachern, M.J. et al (2000) Annu. Rev. Genet 34, 331.
25. Bird, A. (1987) Trends Genet 3, 342.
26. Gardiner-Garden, M. and Prommer, M. (1987) J. Mol. Biol. 196, 261.
27. Cross, S.H. and Bird, A. (1995) Curr. Opin. Genet Dev 5, 304.
28. Halushka, M.K. et al (1999) Nature Genet 22, 231-238.
29. Cargill, M. Et al. (1999) Nature Genet 22, 231-238.
30. Riskip, N.S. Merikangas, K. (1999) Science 273, 1516-1517.
31. Lander, E.S. (1999) Science 274, 536-539.
32. Collins, F.S. et al (1997) Science 278, 1580-1581 .

33. Collins, F.S. of needles and haystacks finding human disease genes by positional cloning (1991) Clin. Res 39, 615-623.
34. Jorde, L.B. Linkage disequilibrium and the search for complex disease genes (2000) Genome Res 10, 1435-1444.
35. The International SNP map working group – a map of human genome sequence variation containing – 42 million single nucleotide polymorphisms (2000) Nature 409, 928-933.
36. Kruglyak, S. Et al. Equilibrium distribution of microsatellite repeat length results from a balance between slippage events and point mutation. (1998) Proc. Natl. Acad. Sci USA 95, 10770-10778.
37. Tolh, G. Et al. Microsatellites in different eukaryotic genomes ; survey and analysis. (2000) Genome Res 10, 967-981.
38. Lander, E.S. the new genomics; Global view of biology (1996) Science 274, 516-539.
39. Eddy, S.R. Noncoding RNA genes (1999) Curr. Of Genet. Dev 9, 595-699.
40. Ban, N et al the complete atomic structure of the large ribosomal subunit at 24 angstrom resolution (2000) Science 289, 905-920.
41. Nissen, P. Et al the structural basis of ribosome activity in peptide bond synthesis (2000) Science 289, 920-930.
42. Weinstein L.B. & Steitz J.A. guided tours from precursor sno RNA to functional snoRNA (1999) Curr. Opin. Cell Biol. 11, 378-384.
43. Bachellerie J.P. & Cavaille J. In Modification and Editing a RNA (ed. Renne, H.G. a R.) 255-272 (As M. Washington D.C. 1998).
44. Long E.O. & Dawid L.R. Repeated genes in eukaryotes (1980) Annu Rev. Biochem 49, 727-764.
45. Sykes J.E. et al the human ribosomal RNA genes: structure and organization of the complete repeating unit (1986) Human Genet 73, 193-198.
46. Sorensen P.D. & Frederiksen S. Characterization of human 5s ribosomal RNA genes. (1991) Nucleic Acids Res. 19, 4147-4151.
47. Timofeero M. et al organization of a 5s ribosomal RNA gene cluster in the human genome (1993) Mod. Biol. (Mosk.) 27, 861-8687.
48. Hatlen L. & Altaroli G. Proportion of the HeLa cell genome complementary to the transfer RNA and 5s RNA (1971) J. Mol. Biol. 56, 535-553.

<b><u>6. - Proteomas e genomas. Classes de funções moleculares de genes e seus polimorfismos. Proteínas e famílias e seus domínios. Caracterização de animais, tecidos e produções animais</u></b>	172
6.1 - Genes codificadores de proteínas e proteoma	172
Classes das funções moleculares dos genes	174
Proteínas, famílias, domínios e proteoma	177
6.2 - Isolamento de genes e seu mapeamento no respectivo genoma	182
6.3 - Genoma, identificação e caracterização dos animais	185
6.3.1 - Genoma, doenças dos animais e testes comercializados para a sua detecção	189
6.3.2 - Genoma, resistência e/ou sensibilidade às doenças nos animais	192
6.3.3 - Desordens e doenças identificadas nos animais com hereditariedade Mendeliana	194
6.3.4 - Genoma e características produtivas quantitativas dos animais	220
6.3.4.1 - Genoma e características reprodutivas melhoradas	220
6.3.4.2 - Genoma e produção de carne	224
6.3.4.3 - Genoma e produção de leite	228
6.3.5- Genoma e triagem de sexos	233
6.3.6 - Genoma, sua expressão e “ <i>microarrays</i> ” do DNA. Possibilidades futuras.	234
6.4 – Bibliografia	237

## **6. - Proteomas e genomas. Classes de funções moleculares de genes e seus polimorfismos. Proteínas e famílias e seus domínios. Caracterização de animais, tecidos e produções animais**

### **6.1 - Genes codificadores de proteínas e proteomas**

#### **Classes das funções moleculares dos genes**

#### **Proteínas, famílias, domínios e proteoma**

Por proteoma entende-se o conjunto de proteínas codificadas pelo genoma.

Como se sabe este é o grande desafio que se coloca ao mundo científico actual que desenvolve neste momento um projecto (HUPO) muito similar àquele que foi feito para o genoma humano (HUGO).

A literatura da especialidade refere que são conhecidos já alguns proteomas relativamente simples.

Numa análise computacional, face ao “draft” do genoma humano, do conjunto de proteínas previstas foi possível verificar (1) que a cerca de 40% do conjunto destas proteínas não pode ser atribuída uma função molecular pelos métodos que habitualmente são utilizados para integrar proteínas em famílias de proteínas conhecidas.

A enumeração fina das famílias de proteínas, bem como os detalhes da sua estrutura proteica necessitam de mais trabalho experimental bem como de interpretação manual compreensível.

Na análise e classificação preliminar dos 26.383 genes codificadores de proteínas previstos nos humanos, foram utilizados dois métodos. O primeiro método baseou-se nos elementos disponíveis nos Pfam database (2,3) e Celera Panther classification (CPC). O segundo método baseou-se na análise ao nível dos domínios das proteínas, com os Pfam e SMART databases (4, 5).

Foram sequenciadas e assembladas ~95% da sequência euromática do genoma humano e utilizado um novo método automatizado para predizer, produzindo-se um catálogo preliminar dos genes humanos.

Encontraram-se 26.000 a 38.000 genes, bastante menos que as previsões iniciais que apontavam para 50.000 até 140.000.

Certamente a análise ainda se encontra incompleta e no futuro chegarão muito mais achegas.

Cerca de 40% dos genes humanos são alternativamente “spliced”.

Os genomas mamíferos devem conter no máximo não mais de 30.000 genes, previa já Muller em 1967 (6).

A identificação deste tipo de genes, genes codificadores de proteínas, é uma das mais importantes aplicações dos dados das sequências estudadas, mas é também um dos desafios mais difíceis de alcançar.

Referem-se seguidamente os esforços feitos para criar um índice inicial dos genes e das proteínas humanas. Contudo antes de tentar identificar novos genes, deve-se explorar o que se pode aprender alinhando as sequências cDNA de genes conhecidos com a sequência genómica apurada no *draft* (7).

Esta sequência genómica do *draft* foi a sequência produzida por combinação da informação obtida de sequências de clones individuais (criando *merged sequence contigs*\* e depois empregando a informação correlativa para criar *Scaffolds*\*\*\*) e posicionando esta sequência ao longo do mapa físico dos cromossomas.

\**Merged sequence contig* são contíguos produzidos tomando as sequências contíguas iniciais contidas em clones com sobreposição e amalgamando aqueles que se sobrepunham.

*Contigs* é o resultado da junção de uma colecção de sequências de clones que se sobrepõem por alinhamento.

\*\**Scaffolds* é o resultado da conexão entre *contigs* utilizando a informação das leituras das pontas emparelhadas dos plasmídeos, ou das leituras das pontas emparelhadas dos BACs, ou dos mRNAs conhecidos ou de outras fontes.

Os *contigs* num *scaffold* são ordenados e orientados uns em relação aos outros.

Estes alinhamentos até agora (2001) só foram permitidos para \_ dos genes conhecidos. Os alinhamentos genómicos permitiram o estudo da estrutura exão – intrão e do conteúdo local em GC e são valiosos para estudos biomédicos porque relacionam os genes com o mapa genético e citogénico ligando-os com sequências reguladoras e facilitando o desenvolvimento de PCR *primers* (*polimerase chain reaction*) para ampliar exões.

Os genes conhecidos estudados até agora são os contidos no Ref Seq. *Database* (8) que contem 10.272 mRNAs, ou seja uma colecção da maioria de sequências mRNA humanas em todo o seu comprimento contidos no GenBank (Ref Seq intencionalmente contem algumas formas alternativas de “Splice” dos mesmos genes).

Os genes Ref Seq foram alinhados com a sequência do *draft* utilizando dois programas de computadores (Spidey e Acembly). Como esta sequência está incompleta e contem erros nem todos os genes puderam ser plenamente alinhados e alguns puderam ter sido incorrectamente alinhados, sendo citadas percentagens destas situações.

Alinhamentos completos foram feitos de 5.364 genes e alinhamentos imperfeitos ou parciais de 9.212 genes conhecidos distintos.

Na tabela seguinte referem-se as características básicas das estruturas dos genes humanos.

Há considerável variação no tamanho dos genes e no tamanho dos intrões. Muitos genes têm mais de 100kbp de comprimento e o maior conhecido, o gene da distrofina (DMD) tem 2,4Mbp.

#### Características dos genes humanos (9)

	Media
Exões / internos	145 bp
Nº de exões	8,8
Intrões	3.665 bp
3' UTR	770 bp
5' UTR	300 bp
Sequência codificante	1.340 bp
Tamanho médio da proteína	447 ácidos aminados
Extensão genómica	27kb

As sequências ricas em purinas podem facilitar o *splicing* (10, 11). As regiões ricas em GC tendem a ser densas em genes, enquanto as regiões ricas em AT tendem a ser pobres em genes.

No estudo dos intrões foi possível verificar que de 53.295 intrões estudados, 98,12% tem os canônicos dinucleótidos GT na extremidade de corte 5' e AG na extremidade 3' (GT-AG perfil), 0,76% utiliza o perfil GC-AG e cerca de 0,1% utiliza AT-AC.

*Splicing* alternativo pode permitir que diversas proteínas produzidas por um único gene possam ser utilizadas para regulação complexa dos genes.

Calcula-se que 35% dos genes humanos sejam sujeitos a *splicing* alternativo (12, 13).

70% das formas alternativas de *splicing* encontradas nos genes dos cromossomas 19 e 22 humanos afectam a sequência codificante, mais do que a simples alteração da região 3' ou 5' UTR (*untranslated region*).

Nos genes humanos as sequências codificantes representam apenas alguns por cento do genoma com uma média de 5% em cada gene.

Calcula-se que existam 31.000 genes nos humanos e se calcula também que cada proteína tenha em média 447 aminoácidos.

Estamos muito longe de ter um conjunto completo dos genes humanos.

Havendo 30.000 a 35.000 genes, cada um com um comprimento codificante de cerca de 1.400 bp e uma extensão genómica de 30kb então cerca de 1,5% do genoma humano consiste de sequências codificantes e 1/3 do genoma será transcrito em genes.

Algumas classes de genes podem ter sido esquecidas nos métodos utilizados para encontrar genes como pode suceder com os genes expressos com uma taxa muito baixa ou em tecidos raros.

**Classes identificadas das funções moleculares dos genes**  
**Categorias ou classes das funções moleculares**  
**dos 26.383 genes humanos ( 14 )**

Classe	<u>Função molecular</u>	número	percentagem
1	Enzimas dos ácidos nucleicos	2.308	7,5%
2	Factores de transcrição	1.850	6%
3	Receptores	1.543	5%
4	Miscelânea	1.318	4,3%
5	Hidrólases	1.227	4,0%
6	Moléculas reguladoras selectivas	988	3,2%
7	Protooncogenes	902	2,9%
8	Proteínas estruturais do citoesqueleto	876	2,8%
9	Cinases	868	2,8%
10	Oxido reductases	656	2,1%
11	Transferases	610	2,0%
12	Adesão celular	577	1,9%
13	Transportadores	533	1,7%
14	Matrix extracelular	437	1,4%

15	Canais iónicos	406	1,3%
16	Motores	376	1,2%
17	Moléculas sinais	376	1,2%
18	Transportadores intracelulares	350	1,1%
19	Proteínas estruturais dos músculos	296	1,0%
20	Sintase e sintetases	313	1,0%
21	Imunoglobulinas	264	0,9%
22	Proteínas transferidora/portadoras	203	0,7%
23	Isomerasas	163	0,5%
24	Chaperones	159	0,5%
25	Liasas	117	0,4%
26	Proteínas virais	100	0,3%
27	Ligases	56	0,2%
28	Proteínas ligadoras de cálcio selectivas	34	0,1%
	Função molecular desconhecida	12.809	41,75%

Os métodos automáticos utilizados para classificação das proteínas nesta análise inicial, referem-se apenas a famílias de proteínas relativamente grandes. Para o conjunto de 60% das proteínas cujas previsões foram determinadas automaticamente quanto aos seus papéis funcionais, as funções específicas foram englobadas em classes muito amplas.

A tónica foi posta na função molecular (mais do que nos processos celulares mais elevados) em ordem a classificar tantas proteínas quanto possível.

Estas previsões funcionais são baseadas na sua semelhança com sequências de funções conhecidas. Dos genes previstos com baixa taxa de segurança ou confiança (apenas com um método dando possível confirmação) apenas 5% poderão ter funções moleculares atribuíveis pelos métodos automatizados. Um terço destes 5% de genes hipotéticos possíveis, são representados por proteínas retrovirais endógenas o que sugere que a maioria desses genes de funções desconhecidas não são realmente genes.

As funções moleculares mais frequentes são as correspondentes aos factores de transcrição, bem como aquelas enzimas implicadas no metabolismo dos ácidos nucleicos.

Nesta categoria das enzimas dos ácidos nucleicos inclui-se sobretudo a maquinaria da transcrição (essencialmente DNA/RNA metiltransferases, DNA/RNA polimerases, helicases, DNA ligases, DNA e RNA factores de processamento, nucleases e proteínas ribossomais).

A maquinaria de transcrição e translação parece ter sido muito conservada através da evolução das espécies, desde as células bacterianas até aos eucariotas mais complexos.

Genes que se encontram presentes em dois organismos distintos dizem-se conservados. A conservação pode ser detectada medindo a semelhança de duas sequências das bases ao nível do DNA ou RNA ou dos ácidos aminados ao nível das proteínas. Quanto mais semelhantes as sequências mais altamente conservadas estão.

Também diversas ribonucleoproteínas implicadas no *splicing* do RNA parecem ter sido conservadas entre os animais.



Outras funções também muito frequentes são os receptores, hidrolases (a maioria das hidrolases são proteases) e cinases.

Há também muitas proteínas que são membros das famílias proto-oncogenes, assim como famílias de moléculas reguladoras selectivas ou sejam:

- a) proteínas implicadas nas etapas específicas da transdução de sinais tais como as heterotrimericas proteínas de ligação ao GTP (G proteínas) e reguladores do ciclo celular.
- b) proteínas que medeiam a actividade de cinases, G proteínas, e fosfatases.

Numa perspectiva comparativa da evolução de diferentes espécies foi possível verificar factos muito interessantes comparando o genoma humano com genomas de levedura (*S. cerevisiae*) e de dois invertebrados, um nematodo (*C. elegans*) e um insecto (*D. melanogaster*) e de uma planta (*A. thaliana*).

Assim há categorias de genes que estão super representados em todos estes genomas estudados como é o caso das enzimas implicadas no metabolismo dos ácidos nucleicos, sobretudo da maquinaria de transcrição, outros tipos de enzimas (transferases, oxido reductase, ligases, liases e isomerases) e ainda transportadores proteicos e *chaperones*.

Salienta-se que nestes últimos dois grupos são sobretudo as proteínas implicadas no transporte mediado por vesículas revestidas e os *chaperones* implicados na dobragem das proteínas e de resposta a choques térmicos (sobretudo família DNAJ) e famílias HSP60, HSP70, HSP90.

Nas diferenças entre o genoma humano e os outros genomas eucariotas sequenciados vistos numa perspectiva famílias proteína/domínios [definida pela similitude da sequência, por ex: proteína cinases serina treonina e super-famílias definidas pela função molecular análoga o que pode incluir diversas famílias de sequências relacionadas, como por exemplo as citocinas] verifica-se que no genoma humano e afirmamos nós provavelmente nos genomas animais as expansões de genes mais notória são as correspondentes às proteínas implicadas,

- a) nas funções de imunidade adquirida.
- b) no desenvolvimento neural, sua estrutura e funções.
- c) nas vias de sinalização intercelular e intracelular do desenvolvimento e da homeostase.
- d) da hemostase.
- e) da apoptose.

O maior desenvolvimento dos genes responsáveis pela imunidade adquirida parece justificar-se pelo facto desse sistema apenas existir nos vertebrados.

Nos humanos estão assinalados 22 genes da classe I do complexo principal de histocompatibilidade (MHC) e 22 da classe II e 114 outros genes de imunoglobulinas, além de 59 genes na família dos receptores de imunoglobulina conhecidos.

Também o número acrescido de membros de famílias de proteínas implicadas no desenvolvimento neural em relação aos outros genomas estudados inclui factores neurotrópicos, factores de crescimento

dos nervos, moléculas sinal como as semaforinas assim como proteínas envolvidas directamente na estrutura neural e função como as proteínas da mielina, canais de iões franqueados por voltagens e por proteínas sinápticas.

No que se refere aos genes implicados nas vias de sinalização intercelular e intracelular no desenvolvimento e na homeostase, eles estão mais expandidos nos vertebrados do que nos invertebrados nelas se incluindo hormonas segregadas, factores de crescimento, receptores.

Também a expansão dos genes responsáveis por moléculas de adesão extracelular e sinalização parece dever-se ao papel crítico que elas assumem na defesa do organismo, morfogénese e reparação dos tecidos.

A hemostase é regulada sobretudo pelas proteases do plasma sanguíneo da via da coagulação e pelas interações que ocorrem entre o endotélio vascular e plaquetas.

A proteólise da matrix extracelular é crítica para o desenvolvimento dos tecidos e para a degradação dos tecidos em doenças tais como o cancro, artrites, doença de Alzheimer e uma série de situações inflamatórias (15, 16).

### **Proteínas, famílias, domínios e proteoma**

#### Número de algumas proteínas assinaladas em famílias e sub-famílias e identificadas nos humanos (H. sapiens) (17)

##### I – Da estrutura neuronal

Canais iónicos – 267

##### II – De resposta imunológica

Citocinas – 86

Interferon – 8

Interleucinas – 26

TNF – 9

Receptor de citocinas – 62

Receptor de interferon – 3

Receptor de interleucinas – 32

Receptor de TNF – 3

Receptores de imunoglobulinas – 59

##### III – Reguladores do desenvolvimento e homeostase

Moléculas assinalantes

Calcitonina – 3

Glucagina – 4

Insulina – 1

Insulin-like hormona – 3

Factor de crescimento dos nervos – 3

Neurogulina – 6

PDGF – 1

## Receptores

### IV – Cinases e fosfatases

Fosfatases com dupla especificidade – 29

Proteínas cinases serina/treonina e dupla especificidade – 395

Fosfatase serina/treonina – 15

Proteína cinase Y – 106

Fosfatase Y – 56

### V – Transdução de sinais

Fosfodiesterases de nucleótidos cíclicos – 25

Receptores acoplados a G proteínas – 616

### VI – Factores de transcrição / organização da cromatina

Proteínas contendo zinc finger  $C_2H_2$  – 607

CREB – 7

Relacionada com ETS – 25

Relacionada com cabeça de forquilha – 34

Histona H1 – 5

Histona H2 A – 24

Histona H2 B – 21

Histona H3 – 28

Histona H4 – 9

Leucina Zipper – 6

Receptor nuclear de hormonas – 59

### VII – Adesão da matrix extra celular

Caderina – 113

Conexina – 14

ICAM – 6

Integrinas \_ – 24

Integrinas \_ – 9

Proteoglicanas – 22

### VIII – Apoptose

Bcl – 2-12

Calpains – 22

Caspases – 13

### IX – Outras enzimas

Citocromo P450 – 60

### X – Splicing e translação

EF1 \_ – 56

Ribonucleoproteínas – 269

Proteínas ribossomais – 812

As proteínas são compostas por unidades mais pequenas chamadas domínios, que desempenham funções particulares. A actividade de uma proteína é a soma destas funções.

Os domínios são normalmente compostos por sequências discretas de ácidos aminados contíguos não sobrepostas.

Um domínio facilmente detectado é por vezes denominado “Tag”.

**Domínios e número de algumas proteínas contendo o domínio em causa (nos humanos) (18)**

- a) Número de acesso Pfam    b) Designação do domínio  
c) Descrição do domínio    d) Número de proteínas

<b>a</b>	<b>b</b>	<b>C</b>	<b>d</b>
		<b>Reguladores do desenvolvimento e homeostase</b>	
PF00212	ANP	Atrial natriuretic péptido	2
PF00976	Domínio ACTH	Domínio ACTH corticotrofina	1
PF00167	FGF	Factor de crescimento fibroblástico	23
PF00049	Insulina	Insulina / IGF / família relaxina	7
PF00219	IGFBP	Insulin-like growth factor binding proteins	10
PF02024	Leptina	Leptina	1
PF00243	NGF	Família do factor de crescimento dos nervos	3
PF00341	PDGF	Factor de crescimento derivado das plaquetas	5
PF01033	Somatomedina B	Domínio somatomedina B	5
PF00103	Hormona	Somatotrofina	1
PF00019	TGF -	Factor de crescimento transformante domínio - like	27
		<b>Respostas imunitária</b>	
PF00047	Ig	Domínio imunoglobulina	381
PF00143	Interferon	Interferão / domínio	7
PF00714	IFN -	Gama interferão	1
PF00726	IL - 10-15-2-etc.	Interleucina	1 cada
PF00229	TNF	Família do factor de necrose tumoral	12
		<b>PI - PY - rho GTPase assinalante</b>	
PF00503	G -	Subunidades da G-proteína	27
PF00631	G -	Subunidade - like	16
PF0071	Ras	Família Rab	126
PF0017	SH <sub>2</sub>	Domínio Src homólogo 2 (SH <sub>2</sub> )	87
PF0018	SH <sub>3</sub>	Domínio Src homólogo 3 (SH <sub>3</sub> )	143
PF01017	STAT	Proteína STAT	7
		<b>Domínios envolvidos na apoptose</b>	
PF00452	Bcl-2	Bcl-2	9
PF00619	CARD	Domínio de recrutamento de caspases	16
PF00531	Death	Domínio de morte	16
PF01335	DED	Domínio efector de morte	4
PF00656	ICEp20	Domínio p20 ICE-like protease (caspases)	11
PF00653	BiR	Inibidor do domínio de apoptose	8
		<b>Citoesqueleto</b>	

PF00022	Actina	Actina	61
PF00435	Spectrina	Repetição espectrina	31
PF00418	Ligação de tubulina	Proteínas Tour r MAP	4
PF00992	troponina	troponina	4

		<b>Adesão à matrix extracelular</b>	
PF01391	Colagénio	Repetição em hélice tripla de colagénio (20copias)	65
PF00008	EGF	EGF-like domínio	108
PF00357	Integrina A	Integrina região citoplasmica	3
PF00362	Integrina B	Integrina cadeia	8
		<b>Domínios de interacção nuclear</b>	
PF00250	Cabeça de forquilha	Domínio cabeça de forquilha	35
PF00010	HLH	Domínio hélice-ansa-hélice de ligação ao DNA	60
PF00104	Hormona receptor	Domínio receptor por ligação ao ligando	47
PF00352	TBP	TBP-TATA binding proteína	2
PF00642	Zf-CCCH	Zinc finger tipo c-xb-cx5-cx3-H é similar	17
PF00096	Zf C <sub>2</sub> H <sub>2</sub>	Zinc finger tipo C <sub>2</sub> H <sub>2</sub>	564
PF00097	Zf C <sub>3</sub> HC <sub>4</sub>	Zinc finger tipo C <sub>3</sub> HC <sub>4</sub>	35

O velho conceito paradigmático de que um gene faz uma proteína foi hoje substituído (depois da informação publicada na Science e Nature de 2001 sobre a sequenciação do genoma humano), por “um gene faz diversas proteínas”, sendo a base para esta explicação o processo de alternativos “splicing” do mesmo mRNA (19).

Nesta era *post*-genómica, o centro das atenções recai agora no estudo da proteómica ou seja a compreensão da função de cada proteína codificada dentro de uma célula durante a sua vida. É sabido que as proteínas são entidades multifuncionais. A proteómica é pois a análise em larga escala da função dos genes.

O estudo da proteómica envolve a comparação das amostras de tecidos de indivíduos saudáveis com as de indivíduos doentes ou diferentes.

De início as tecnologias para este efeito, recorriam á 2-D gel electroforese, para desenhar o perfil em proteínas específicas dos tecidos ou das células e depois procedia-se á análise das proteínas separadas, pela espectrometria de massa. A partir daqui obtinham-se bases de dados e bioinformáticos das proteínas, assim como mapas das interacções entre as proteínas.

Na análise das proteínas, uma ferramenta chave na tecnologia da proteómica foi a descoberta há cerca de uma década do *electrospray* (ES) e do *matrix-assisted laser desorption ionization* (MALDI) métodos que permitiam a franca ionização de grandes biomoléculas que com a espectrometria de massa permitiram ir muito mais além (20).

A ES e a MALDI são técnicas de ionização que permitiram o maior sucesso á espectrometria de massa aplicada ás ciências da vida.

Também em 1993 surgia o primeiro “software” de algoritmos que permitia correlacionar os dados da espectrometria de massa obtidos para uma proteína com a base de dados das respectivas sequências.

O European Bioinformatics Institute – EBI – (U.K.) dispõe de uma série de informação chave para a análise “in silico”\* de proteomas completados, que em Outubro de 2001, cobrem mais de quarenta proteomas, fornecendo uma colecção de páginas de análise do proteoma completado de cada organismo.

Esse mesmo instituto refere o fácil acesso (Dr. Rolf Apweiler – Swiss – PROT coordinators, EMBL/EBI, U.K.) ao conjunto completo de sequências proteicas, e estatísticas sobre os dados das estruturas primárias, secundárias e terciárias, famílias baseadas na Inter Pro e informação dos domínios, e ainda meios para comparar cada organismo com o proteoma acabado de outro organismo, etc.

No quadro seguinte referem-se algumas bases de dados disponíveis dentro deste contexto.

#### Base de dados sobre proteínas (21)

- Swiss PROT – Swiss institute of Bioinformatics (Genova. Switzerland) [www.expasy.ch](http://www.expasy.ch)  
(Primary Sequence Information of proteins)  
([www.expasy.ch/www/tool's.html](http://www.expasy.ch/www/tool's.html))  
(Swiss – 2D PAGE 2 – D – gel electrophoresis and MELANIE)
- University of California, San Francisco  
(MS database to identify proteins)  
([www.prospector.UCSF.edu](http://www.prospector.UCSF.edu))
- Rockefeller University – PROWL – (New York NY)  
(Peptide sequence ID from MS data)  
([prowl1.rockefeller.edu/prowl/prowldescription.html](http://prowl1.rockefeller.edu/prowl/prowldescription.html))
- EMBLs (European Molecular Biology Laboratory, Heidelberg, Germany).  
([www.mann.emblheidelberg.De/services/peptidesearch/peptidesearchIntro.html](http://www.mann.emblheidelberg.De/services/peptidesearch/peptidesearchIntro.html))
- Biomolecular Interaction Network Database (BIND) at the Samuel Lunenfeld Research Institute of Toronto's Mount Sinai Hospital  
([bioinfo.mshri.on.ca/cgi-bin/bind/bindl/dataman](http://bioinfo.mshri.on.ca/cgi-bin/bind/bindl/dataman))
- Hybrigenics SA (Paris, France)  
([www.pim.hybrigenics.com](http://www.pim.hybrigenics.com))
- Proteome Inc. (Beverly, MA)  
([www.proteome.com/YPDhome.html](http://www.proteome.com/YPDhome.html))

\* *In silico* – método utilizando sofisticados modelos de computadores, mais do que experiências em laboratório (*in vitro*) e em animais (*in vivo*).

## **6.2 - Isolamento de genes e seu mapeamento no respectivo genoma**

Desde meados dos anos 70 que se tem procurado isolar e interpretar os genes ou as sequências de DNA, tal como surgiu a ideia de os mapear nos respectivos genomas.

Os genomas dos bovinos, ovinos, ratos e suínos têm sido profundamente estudados nesta perspectiva, havendo hoje já construídos em maior ou menor extensão mapas para diversas espécies bem como para equídeos, frangos, perus, gato, salmonídeos, tilápias, etc.

A sequenciação de todos os genes bovinos segundo notícias de 2003 será efectuada nos U.S.A, a partir deste ano pelas Faculdade de Medicina de Baylor em Houston e pela Universidade A&M do Texas.

O mapeamento dos genes permite-nos localizar os genes que controlam características hereditárias inclusivé de determinadas doenças.

Assim e no caso dos suínos foi possível descobrir um gene responsável pela PSE (suínos susceptíveis ao *stress*), tal como são correntemente identificados genes correlacionados com o tamanho da ninhada de leitões, o ritmo de crescimento, a espessura da gordura, a eficiência dos alimentos e área do “longo dorsal”.

Está localizado já e pode ser determinado um gene (denominado ES R) que influencia grandemente o “tamanho” da ninhada das marrãs com as vantagens facilmente percebíveis pois pode-se determinar quais as fêmeas que produzirão maior número de crias.

Também nas vacas leiteiras associadas com mais altas produções de leite, gordura e proteína parecem estar os alelos *\_SI - CNB* e *\_ - CNA*, bem como provavelmente a *K - CN BB* e *\_ - LGAA* sobre a produção de proteínas (*vide* adiante biologia molecular na produção animal – glândula mamária e excreção láctea).

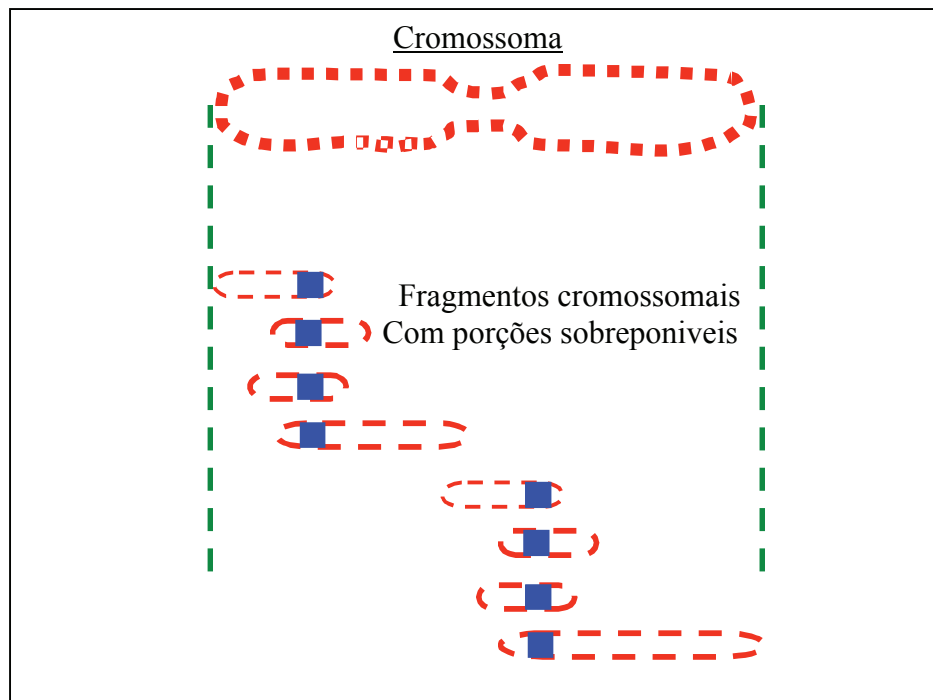
Também são bem conhecidos em bovinos de garupa dupla as mutações ao nível do gene da miostatina GDF8 ou MSTN (Mh locus) no cromossoma 2 responsáveis por esta hipertrofia muscular, tal como em ovinos está também assinalada uma hipertrofia muscular, simultaneamente com a produção de carne menos gorda e obtida com melhores índices de conversão, nos animais com um gene autossomal dominante mutado (*clpg*) no locus callipyge, situado no cromossoma 18. (*vide* adiante biologia molecular na produção animal).

Características complexas (como por exemplo o crescimento dos animais) são controladas por diversos genes cada um deles com pequeno efeito sobre as características em causa.

Para localizar os genes e seu mapeamento podem ser utilizados o mapa citogenético ou o mapa de “linkage”. Este último é baseado na análise da cosegregação de marcadores genéticos em populações descendentes.

No mapa citogénico obtém-se um conjunto de bandas nos cromossomas, após coloração adequada (cariótipo) sendo o perfil em bandas característico e único para cada cromossoma.

O mapa citogenético ou mapa físico pode ser criado pelo corte, quimicamente, dos cromossomas em pequenas peças e depois reordenando estas em ordem a reconstituir o cromossoma. Para isto utilizam-se dois tipos de marcadores os genes e os microssatélites (marcadores de DNA anónimos).



Um marcador genético pode ser ou pode não ser uma parte de um gene.

São comumente referidos os seguintes mapas genéticos calculados para diversas espécies animais (22).

<u>Espécies animais</u>	<u>Genes e marcadores</u>
Suínos	4.000
Bovinos	4.000
Frangos	2.000
Ovinos	1.000
Humanos	60.000
Rato	50.000

O mapa de “linkage” dá uma localização específica para um dado gene baseado na distância entre genes. Este tipo de mapa é utilizado quando se deseja obter uma visão global de uma grande quantidade de DNA num cromossoma e para seleccionar os genes implicados em certas características. Quanto menor a distância (em unidades chamadas centimorgans ou cM) mais próximos estarão os genes uns dos outros.

Podemos resumir nos seguintes quadros o mapeamento em diversas espécies animais de genes ou de características produtivas, ou responsáveis por situações de doenças em diversos cromossomas.



### Mapeamento de características produtivas em suínos ( 23 )

Crescimento	Cromossoma 2, 3, 4, 6, 7, 8, 9, 11, 13, 14,
Corpulência	Cromossoma 1, 2, 4, 5, 6, 7, 13, 14, 18
Qualidade da carne	Cromossoma 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 17
Tamanho das ninhadas	Cromossoma 6, 7, 8
Resposta imunitária/stress	Cromossoma 1, 4, 6, 8

### Mapeamentos no genoma de frango (24)

Doença de Marek	Cromossoma 1, 2, 4, 7
Resistência as salmoneloses	Cromossoma 5, 7
Resposta imunitária geral	Cromossoma 2
Características produtivas	Diversos cromossomas

### Características produtivas em bovinos leiteiros ( 25 )

Cromossoma	Características
1	Leite e proteína
3	% de gordura no leite
4	Doenças várias
5	Doenças várias
6	% de gordura e proteína e produção de leite
9	Produção de gordura e proteína
10	Produção de gordura
14	% de gordura e % de proteína do leite
15	Sindactilia
16	Vida produtiva
17	Produção de leite
28	% de proteínas no leite

Nos últimos anos os mapas de “linkage” e físico, do genoma dos suínos tem tido avanços notáveis, estando assinalados no mapa de “linkage” mais de 1800 *loci* incluindo cerca de 250 genes.

Estão mesmo identificadas importantes regiões cromossómicas e genes principais associados com características de interesse económico nos suínos, inclusivé QTL (*quantitative trait loci\**) para o crescimento e gordura de depósito (nos cromossomas 1, 2, 3, 4, 5, 6, 7, 8, 13, 14) qualidade da carne (cromossoma 2, 3, 4, 6, 7, 12,15) e reprodução ( cromossomas 4, 6 ,7, 8).

Também as mutações responsáveis pelo síndrome de *stress* suíno (HAL ou CRC1 ou gene do halotano) e da cor da pele estão identificados. Também há resultados muito interessantes no que se refere ao número de animais nascidos em cada gestação (nos genes ESR- receptor dos estrogénios e PRLR – receptor da prolactina) na qualidade da carne (FABP – *heart fatty acid binding protein*) e na resistência a doenças (FUT1 ou gene fucosil transferase).

-----\* É uma medição feita numa escala quantitativa (linear). Estas características são tipicamente afectadas por mais de um gene e pelo ambiente exterior ao genoma.

### **6.3 - Genoma, identificação e caracterização dos animais**

Os animais podem ser seleccionados através do seu valor procriativo calculado na base de medições fenotípicas das suas características produtivas. Tradicionalmente os programas de melhoramento animal baseavam-se na selecção pelas características fenotípicas dos animais e hoje a selecção é mais feita baseada nas características genotípicas dos animais que estão associados com melhores características produtivas.

A selecção pelas características fenotípicas era mais feita pela descendência ou produtos animais obtidos enquanto se utilizarmos as características genotípicas podemos detectar mais cedo e mais especificamente os animais de maior interesse.

Dentro dos métodos que seleccionam características genotípicas específicas temos a considerar.

- a) métodos que detectam variações nos genes de *loci* de características quantitativas (QTL) que estão meramente associadas com características produtivas dos animais.
- b) Métodos que detectam variações genéticas nos genes funcionais que influenciam directamente aquelas características produtivas.

Um dos primeiros métodos é a selecção baseada num marcador na medida em que polimorfismos nos marcadores identificados de forma aleatória podem estar associados com características da produção. São pois métodos de biologia molecular aplicados no campo da procriação animal de espécies pecuárias, procurando seleccionar animais cujas características produtivas podem melhorar o seu valor nos programas de melhoramento animal.

Como se disse um marcador genético pode ser ou pode não ser uma parte de um gene.

Para este efeito podem-se considerar (26):

#### Genes de tipo I

Enzimas marcadores

RFLP, PCR-RFLP (comprimento polimórfico de fragmentos de restrição)

SNPs (polimorfismo de nucleótidos simples)

#### Marcadores ao acaso ( tipo II )

Microssatélites (  $\leq 13$  nucleótidos repetidos *single locus* SSR )

Minissatélites (  $\geq 14$  nucleótidos repetidos multi-locus, elementos repetido em tandem)

RAPDs (polimorfismo de DNA ampliado ao acaso)

AFLPs (comprimento polimórfico do fragmento ampliado)

### Polimorfismos e identificação de espécies

O estudo do polimorfismo dos ácidos nucleicos constitui um meio para identificar espécies, serotipos, estirpes, variedades, raças ou indivíduos, baseado nas diferenças da sua constituição genética.

Os polimorfismos nos ácidos nucleicos podem ser devidos à substituição de nucleótidos, à sua inserção ou deleção.

O interesse em determinar os polimorfismos genéticos reside nas suas aplicações para por exemplo o mapeamento genético, diagnóstico médico, estudos epidemiológicos, forenses ou agrícolas, etc.

Existem diversos métodos para determinar se existem ou não polimorfismos quando se comparam segmentos homogêneos de DNA (*vide* por exemplo patente US6,287,564 de 2001) (27).

Contudo todos estes métodos têm as suas limitações o que prejudica a sua aplicação prática.

O método mais directo para detectar polimorfismos é a comparação directa das sequências dos diferentes genomas. E apesar do desenvolvimento de métodos de sequenciação automatizados terem diminuído em muito o tempo e esforço necessários para sequenciar o DNA, o tamanho deste genoma mesmo nos seres procariotas mais simples é tão grande que a sua sequenciação directa é impraticável como meio de rotina para a detecção de polimorfismos.

Um método para determinação do polimorfismo genético do DNA genómico é a electroforese em gel em campo pulsátil ou PFGE que é utilizada para detectar diferenças em grandes fragmentos de DNA genómico.

Nesta PFGE são aplicadas ao gel campos eléctricos ortogonais, pulsáteis, alternados.

As grandes moléculas de DNA migram através do gel, progressivamente até que o DNA tenha tempo suficiente para se reorientar ele próprio ao longo do novo eixo do campo eléctrico. O grau de resolução do PFGE depende de vários factores, mas a técnica é útil apenas para a detecção de relativamente grandes diferenças, geralmente de 5 – 10kbp entre as moléculas de DNA.

Outro método para a detecção de polimorfismo são o polimorfismo do comprimento de fragmentos de restrição ou RFLP. Este RFLP é baseado na especificidade com que as enzimas de restrição cortam o DNA. Uma enzima de restrição quebrará o DNA apenas quando se encontrar presente uma sequência particular de DNA específica desse enzima de restrição.

Qualquer alteração nesta sequência do DNA impedirá que essa enzima de restrição seja capaz de cindir esse DNA nesse local, o que originará uma alteração do comprimento do fragmento produzido. Alterações que ocorram em qualquer local da sequência podem criar novos locais de reconhecimento dentro da sequência, alterando assim a distribuição do tamanho dos fragmentos de restrição.

Assim alterações na sequência do DNA devidas a inserção, deleção e inversão podem originar variações no tamanho dos fragmentos produzidos que podem assim ser detectados por RFLP.

Diferenças no tamanho dos fragmentos devidos à acção das enzimas de restrição, são tipicamente determinados pelos “Southern blots” em que após a digestão pela enzima de restrição escolhida, os fragmentos resultantes são separados pelo seu tamanho em gel electroforese, seguindo-se a sua transferência para uma membrana e hibridação com uma sonda marcada correspondente à área particular do genoma.

Este método é limitado a sequências que possam ser detectadas pelas sondas de DNA utilizadas. Por outro lado, mutações pontuais podem ser detectadas apenas se elas ocorrerem num local ou sítio de restrição.

Outro método utilizado para identificar polimorfismos utiliza “primers” de uma sequência arbitrária para amplificar DNA pela reacção de PCR (*polymerase chain reaction*).

Como os “primers” não são desenhados para amplificar uma sequência específica a técnica é chamada RAPD (*randomly amplified polymorphic DNA*). Os “primers” utilizados têm pelo menos sete nucleótidos de comprimento e tipicamente têm entre 9 e 13 bases com um conteúdo G+C de cerca de 50-80% e nenhuma sequência palíndroma. Em condições adequadas, diferenças num simples nucleótido podem afectar a ligação do “primer” ao DNA molde, originando assim diferenças na distribuição dos produtos de amplificação produzidos entre os genomas.

Este método é mais rápido e mais susceptível de automação do que o RFLP embora tenha uma limitação que é a natureza ao acaso dos “primers” utilizados, não permitirem alcançar a máxima cobertura do genoma. Pelo contrário, a natureza ao acaso da amplificação torna possível que grandes porções do genoma não sejam amplificadas, abrindo a porta a que áreas úteis dos polimorfismos não sejam detectadas.

Existem processos patenteados (Patente USA 4.683.195 de julho de 1987) (28) para amplificar, detectar e/ou clonar sequências de ácidos nucleicos alvo, contidas em ácidos nucleicos ou em misturas destes.

Nestes processos as hélices complementares separadas do ácido nucleico são tratadas com um excesso molar de dois oligonucleótidos “primers” estendendo estes *primers* de forma a formarem-se produtos ou sejam extensões desse “primer” complementares, as quais actuam como moldes para se sintetizar a sequência desejada de ácido nucleico e detectando-se a sequência desta forma amplificada.

As etapas da reacção podem ser repetidas à medida que se deseje. Além disso uma sequência específica de ácido nucleico pode ser clonada num vector utilizando “primers” para amplificar a sequência, a qual contém locais de restrição nas suas extremidades não complementares, podendo ser preparado um fragmento de ácido nucleico a partir de um fragmento mais curto existente, utilizando o processo de amplificação.

Um outro método para identificar e mapear polimorfismos genéticos é o AFLP (*amplified fragment length polymorphism*) que combina a utilização das enzimas de restrição como na RFLP com a utilização da PCR como na RAPD. O AFLP é no entanto um processo com muitas etapas e a detecção dos polimorfismos é limitada a áreas no local ou entre os sítios de restrição.

Outro método para detectar polimorfismos baseia-se no alto grau de variação do comprimento de certas sequências nucleotídicas repetidas em tandem na maioria, se não em todos os eucariotas, chamadas simples sequências repetidas (SSR), simples sequências de comprimento polimórfico (SSLP), repetições de nucleótidos, trinucleótidos, tetranucleótidos, ou pentanucleótidos e microssatélites.

Os microssatélites podem ser utilizados na detecção de polimorfismos. Para isso são desenhados “primers” baseados na sequência das sequências dos microssatélites. Os “primers” contem 2 a 4 nucleótidos adicionais que flanqueiam a sequência microssatélite em ordem a ancorar os “primers” a um sítio particular em cada *locus* microssatélite. Os “primers” dirigidos pelos microssatélites são altamente eficientes para detectar polimorfismos mas não podem ser utilizados nos procariotas aos quais falta DNA microssatélite.

Dentro dos genomas existem sequências oligonucleotídicas super representadas ou seja que se encontram presentes com uma frequência superior àquela que seria expectável estatisticamente. Novas

sequências nucleotídicas úteis para a identificação genética de bovinos encontram-se referidas na patente FR2779153 de Dezembro de 1999 (29).

### Identificação de animais e análise da sua paternidade em bovinos (30, 31, 32)

Qualquer programa de melhoramento animal pode utilizar a identificação do DNA como ferramenta muito valiosa, uma vez que a identificação dos animais através do seu DNA é muito mais precisa do que os métodos tradicionais como por exemplo o “tipo de sangue”.

A identificação pelo DNA pode ser feita através de marcadores (como por exemplo os microssatélites) que originam um “fingerprint” único para cada animal.

No caso dos bovinos, os marcadores universalmente aceites pela Sociedade Internacional de Genética Animal (ISAG) com três outros marcadores adicionais são utilizados para produzir o “fingerprint”.

Este “fingerprint” do DNA pode ser utilizado para registar animais e estabelecer *pedigree* ou para verificar a respectiva paternidade.

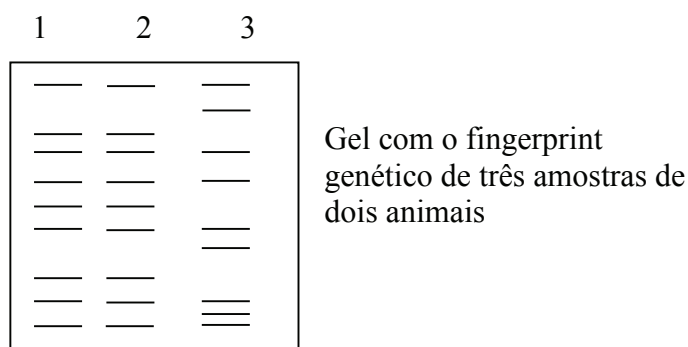
Os benefícios de utilizar para identificação dos animais o DNA são vários.

Esta genotipagem pela análise do DNA envolve a criação de um perfil do DNA de cada animal (*fingerprint* genético) o qual em teoria facilita ou permite a identificação de um animal em seis biliões de animais.

Este “fingerprint” genético pode ser obtido a partir de diversos tecidos tais como sangue, pele, pêlos com os folículos pilosos, penas com os respectivos canudos no caso das aves e após o abate, carnes ou produtos cárneos. A comparação do perfil do DNA de um dado animal ou produto animal, com o perfil do DNA obtido ou criado antes e depositado numa base de dados, permite a prova inequívoca da sua identidade. Quando não for possível a comparação entre amostras, as origens de um animal podem ser averiguadas através das raças dos animais.

É também assim possível certificar a origem, de uma forma credível, da carne.

O “fingerprint” genético é produzido utilizando o diagnóstico pelos microssatélites. Estes microssatélites são sectores do genoma não codificantes, e nos quais existem ou são constituídos por elementos ou sequências simples de DNA, como por exemplo GA, GAG, CAA, que se repetem por diferentes vezes, o que faz com que estes sectores tenham comprimentos variados. Cada animal possui uma combinação única destes diversos sectores de microssatélites com variados comprimentos. Uma análise destes comprimentos feita em gel revela um perfil de bandas único, semelhante a um código de barras. Este perfil de bandas é também chamado perfil do DNA ou “fingerprint” genético.



Fingerprint genético

Para os bovinos, são utilizados para esta análise do *fingerprint* genético 11 marcadores microssatélites, 9 das quais são recomendados pela Sociedade Internacional do Genética Animal (ISAG), o que permite uma identificação de cada animal com eficiência.

Este *fingerprint* genético de um dado animal resultante da análise do comprimento dos marcadores microssatélites, é uma combinação dos *fingerprint* genético do pai e da mãe, uma vez que o genoma de cada animal representa uma combinação do genoma dos seus pais, o que permite determinar com segurança a origem (paternidade) de cada animal.

São hoje oferecidos em diversos países serviços de diagnóstico veterinário baseados na análise do genoma. Assim a genotipagem, para prova de descendência, paternidade ou identificação é feita baseada na análise de microssatélites (em equídeos, asininos, bovinos, suínos, ovinos, caprinos, veados, cães, gatos e macacos) e na análise AFLP *fingerprint* (em animais exóticos, aves, peixes e insectos).

A certificação da origem da carne e de produtos cárneos também é feita por análise do DNA, tal como a determinação do sexo em bovinos e aves.

Também a análise dos *loci* com características quantitativas em melhoramento animal (QTL) é feito recorrendo á análise do DNA, tal como uma série de diagnósticos são hoje feitos baseados em pesquisa de genes como por exemplo:

- Em bovinos a deficiência de adesão dos leucócitos (BLAD), a deficiência em uridina monofosfato sintetase (DUMPS), o Red Factor, o tipo de Kappa-caseína.
- Em cavalos a paralisia periódica hipercalcémica (HYPP) e a imunodeficiência severa combinada (SCID).
- Em suínos o síndrome de hipertermia maligna (MHS) e ainda o diagnóstico de uma série de microorganismos patogénicos para os animais.

### 6.3.1 – Genoma, doenças do animais e testes comercializados para a sua detecção

Referem-se seguidamente alguns testes comercializados para efeitos de diagnóstico de algumas doenças e mais adiante desordens hereditárias nos animais bem como as suas características (outros testes para esta mesma finalidade são referidos em diversos locais destes textos).

### Deficiência de uridina monofosfato sintase em bovinos (33)

Esta situação patológica, também designada oroticoacidúria I orótico acidúria, deficiência em UMP sintase, UMPS deficiência e DUMPS é devida a uma mutação *nonsense* no codão 405 do gene UMPS o que origina uma completa deficiência de UMPS funcional.

A UMPS ou uridina monofosfato sintase é uma enzima necessária para a conversão do ácido orótico em uridina mono fosfato (UMP) que é fundamental para a formação dos nucleótidos pirimídicos.

Esta UMPS actua a dois níveis, como fosforibosil transferase orotica (OPRTase) e como descarboxilase da orotidina monofosfatato (OMPDcase), as duas últimas etapas na via de biossíntese das pirimidinas de constituição dos ácidos nucleicos.

Nos bovinos esta desordem é uma das causas de morte dos embriões. As mortes dos embriões são difíceis de identificar e a sua única manifestação é o “regresso ao serviço” das mães.

A única causa desta desordem nos bovinos é uma mutação “nonsense”\* no codão 405 do gene UMPS que origina uma deficiência completa da enzima UMPS funcional.

Esta mutação origina a morte dos embriões á volta dos 40 dias no útero, originando um atraso no “retorno ao serviço” uma vez que algumas das gestações terminam num aborto natural precoce.

Como estes atrasos no “ retorno ao serviço” podem ter diversas causas, a detecção inicial desta mutação letal embriónica pode ser muito importante.

As portadoras são fenotipicamente normais, mas possuem apenas metade da actividade normal da enzima uridina monofosfato sintase. Durante a lactação as portadoras excretam uma elevada concentração de ácido orótico no leite e na urina.

O gene UMPS está mapeado no meio do cromossoma 1 (q31-36).

O teste genotípico baseia-se na transição C \_ T identificada por Schwenger et al (1993) como a mutação causal, a qual remove a sequência de reconhecimento para a enzima de restrição Ava I.

-----\*mutação *nonsense* é aquela em que um codão de *stop* substitui um codão de ácido aminado, levando a uma prematura terminação da translação.

### Deficiência de adesão leucocitária nos bovinos (34)

Esta deficiência LAD (*Leucocyte Adhesion deficiency*) tem vários sinónimos tais como deficiência em beta-2-integrina, CD18 deficiência, CD11b, BLAD (*bovine leucocyte adhesion deficiency*) etc.

Os animais afectados morrem em virtude da extrema susceptibilidade às infecções devido à incapacidade dos leucócitos para passarem da corrente sanguínea para os tecidos infectados, o que é devido á falta da glicoproteína da membrana, chamada subunidade de integrina beta-2 do leucócito ou CD18.

Em 1992, Shuster et al provaram que esta desordem é devida a uma mutação “mis-sense”\* no gene CD18.

-----\* Mutaç o *missense* origina que uma prote na tenha um amino cido substituído por outro amino cido.

### Doena da imunodefici ncia severa combinada nos cavalos (35)

A CID ou SCID ou doenas de imunodefici ncia combinada, agamaglobulin mia, su ssa de tipo alimfocit tica   devida nos cavalos a uma “frameshift”\*\* muta o no gene para a subunidade catal tica da prote na cinase dependente do DNA (DNA-PK), originando uma falta de toda a extens o da cinase, e aus ncia de actividade cin sica.

-----\*\* A muta o *frameshift* leva   introdu o de um amino cido n o relacionado e geralmente cod o de *stop*.

### Paralisia peri dica II no cavalo (36)

Esta HYPP (*hyperkalaemic periodic paralysis*)   devido a uma muta o “mis-sense” no gene que codifica a cadeia alfa do canal de s dio do m sculo do esqueleto adulto, organizando uma permeabilidade aumentada para o s dio atrav s da membrana das c lulas dos m sculos do esqueleto.

Os aspectos cl nicos assinalam fascicula o muscular e espasmo; debilidade e recorr ncia.

### Detec o de outras situa es patol gicas

A patente U. S. 5.582.987 de Dezembro de 1996 (37) contem m todos para testar a resist ncia ou susceptibilidade dos bovinos   limfocitose persistente, detectando polimorfismos no ex o 2 do gene BoLA-DR3 (BoLA= *bovine lymphocyte antigen*).

A metodologia inclui as etapas seguintes:

- 1) obten o de uma amostra de DNA do animal, que contem o ex o 2 do gene DRB3
- 2) identifica o dos nucle tidos deste ex o 2 do gene DRB3
- 3) determina o de quando os nucle tidos do gene DRB3 contem cod es que codificam na posi o 70 um res duo de  cido glut mico, na posi o 71 uma arginina, na posi o 75 uma valina, na posi o 76 um  cido asp rtico, na posi o 77 uma treonina, e uma tirosina na posi o 78.

A presena destes cod es indica susceptibilidade   limfocitose persistente.

Os *kits* inclu dos na patente cont m “primers” que hibridam com as posi es nucleot dicas 70 e 71 ou 75-78 do ex o 2 do gene DRB3, e envolvem a tipagem designada PCR – RFLP (*polymerase chain reaction-restriction fragment length polymorphism*).

O sistema BoLA (*bovine lymphocyte antigen*)   o complexo principal de histocompatibilidade (“MHC”) dos bovinos e na maioria dos animais contem diversos genes da classe I e II altamente polim rficos. Estes genes codificam mol culas da superf cie celular cuja fun o   a de apresentarem



péptidos antigénicos às células T desempenhando assim um papel muito importante na resposta imunológica a agentes estranhos na medida em que governam as interacções entre as células do sistema imunitário.

O gene DRB3 é o gene mais activamente transcrito da sub-região BoLA DRB.

Diversos métodos têm sido utilizados para caracterizar o polimorfismo da classe II de genes do sistema BoLA. Nos bovinos este complexo principal de histocompatibilidade BoLA e o PRL têm sido mapeados no cromossoma 21.

Outra patente U. S. 5.358.649 de outubro de 1994, (38) refere o diagnóstico da hipertermia maligna suína, utilizando uma molécula purificada de DNA que engloba uma sequência de DNA com aproximadamente 15,1kbp e codificadora de uma proteína RYR1 normal ou mutante, com um peso molecular aproximado de 564.740 daltons.

Esta situação maligna origina uma subida de temperatura progressiva dos animais, com rigidez muscular e acidose metabólica levando à morte rápida dos animais.

### **6.3.2 – Genoma, resistência e/ou sensibilidade às doenças nos animais**

As doenças zoonóticas microbianas originam grandes prejuízos nas explorações animais a despeito da larga utilização de esquemas de vacinação, isolamento e quarentena, diagnóstico e abate, bem como da aplicação de antimicrobianos etc.

A falta de sucessos na erradicação das doenças infecciosas nos animais com a utilização dos meios anteriormente referidos, aponta para outra estratégia que passa pela utilização de meios que permitam identificar as sequências genéticas associadas com a resistência e/ou susceptibilidade desses animais à doença, o que permitiria a eliminação dos animais susceptíveis e a utilização em programas de melhoramento animal, dos animais resistentes.

A necessidade de arranjar métodos para identificação de animais ungulados ruminantes resistentes ou susceptíveis a doenças tais como a brucelose, tuberculose, paratuberculose e salmoneloses levou inclusivamente ao aparecimento da patente U. K. 6.114.118 de Setembro de 2000 (39). Nesta patente relata-se a identificação de um gene NRAMP1 e da sua sequência em bovinos que está associado com a susceptibilidade ou resistência de um animal, a doenças tais como brucelose, tuberculose, paratuberculose e salmonelose, tal como se refere o método para identificar aquelas sequências em ordem a identificar os animais que são susceptíveis ou resistentes.

Em mais pormenor a invenção contida nesta patente relata a identificação de sequências específicas da região 3' não translada (3'UTR) do gene bovino NRAMP1 que está associada com a resistência ou susceptibilidade antes referida, e a utilização de perfis da sequenciação geral para identificar os animais contendo aquelas sequências “in situ” permitindo assim a identificação dos animais previstos como sendo resistentes ou susceptíveis às doenças referidas.

O mecanismo através dos quais variações nas sequências da 3'UTR do NRAMP1 contribuem para a susceptibilidade ou resistência às doenças não é precisamente conhecido, embora se admita que estas variações afectem a translocação da mensagem do NRAMP1 bovino, com uma sequência sendo transcrita mais ou menos do que outras.

Utilizando análises SSCA (*single stranded conformational analysis*) e SSCP (*single stranded conformational polymorphism*), foi descoberto um polimorfismo genético na 3'UTR do gene. Este polimorfismo tem duas formas diferentes que se associam significativamente com os fenótipos de resistência natural e susceptibilidade natural às doenças referidas nos animais.

O polimorfismo bovino NRAMP, resulta da transversão na posição 1782 do NRAMPcDNA bovino encontrando-se timina na sequência resistente e guanina na sequência susceptível. Adicionalmente ocorre uma diferença na sequência microssatélite DNA polimórfica, entre os bovinos resistentes e os susceptíveis, envolvendo o número de repetição dinucleotídica (GT) e o "spacing" na 3'UTR do Nramp1 bovino. Esta sequência nos animais resistentes começa na posição 1779 e é SEQID N°15:GGGTGT(GT).sub.10AT(GT).sub.3(N).sub.61(GT).sub.5(N).sub.24 (GT). Sub13 onde "N" simboliza qualquer uma das quatro bases nucleotídicas A, C, G ou T.

Contrariamente as sequências de DNA associadas com bovinos susceptíveis têm as seguintes formas.

SEQ ID N°32: (GT). Sub < 10 AT(GT). Sub. 3 (N). Sub > 61 (GT) . sub 5 (N) . sub < 24 (GT) . sub > 13 onde "N" representa qualquer das quatro bases nucleotídicas A, C, G ou T.

Há uma ligação entre o *stress* e a ocorrência de doenças nos animais e tem sido mesmo sugerido que o *stress* pode levar a um comprometimento do sistema imunitário.

Em populações animais, inclusive de bovinos, suínos, aves, equídeos, e peixes, está provado que o *stress* pode estar relacionado com a inibição do crescimento, infertilidade, e diminuição da produção de leite e ovos (quando aplicável).

O período periparto nos animais é um período de *stress*.

Dificuldades nas defesas dos bovinos, durante este período podem estar associadas, com a elevada ocorrência de diversas situações patológicas.

Resistências perturbadas nos animais podem ser devidas a factores endócrinos associados com perturbações metabólicas e físicas que decorram durante a gestação, parto e lactação. Neste período *peripartum* as doenças infecciosas mais frequentes são as mastites, metrites e pneumonias, retenção de placenta, febre do leite, cetose e deslocação do abomaso.

A resistência às doenças pode ser pesquisada por selecção directa, quer seleccionando os animais mais resistentes nas condições ambientais normais, ou então por contrastação dos animais com agentes patogénicos específicos.

A selecção indirecta para pesquisa dos animais mais resistentes é baseada na identificação de marcadores indirectos fiáveis da resistência à doença.

Indicadores fenotípicos incluem marcadores morfológicos (como por exemplo: pigmentação margin no olho na queratoconjuntivite infecciosa dos bovinos) marcadores fisiológicos (por exemplo: tipo de hemoglobina na malária) e características de resposta imunitária inata (como por exemplo: PMN função, resposta de anticorpos e CMI). Os indicadores genotípicos incluem genes candidatos (como por exemplo: genes MHC, genes Ig, genes TcR) e marcadores genéticos moleculares anónimos (por exemplo: RFLPs, *loci* repetidos em tandem, *loci* microssatélites).

Em Setembro de 2001 a patente U. S. 6.287.564 (40) refere um método para identificação de animais com elevadas respostas imunológicas em situação de *stress*. Os animais são identificados por um processo de “ranking” que classifica a resposta imunológica dos animais a um antígeno ao longo de um período de tempo que abarca a situação de *stress*.

O método envolve a avaliação da resposta em anticorpos do animal a um antígeno durante um período de tempo que abarca o *stress* que ocorre por exemplo: no período de “periparturition”, de pré-e post-parto.

Com base na resposta aos antígenos, os animais são classificados como de resposta imunológica alta, média ou baixa.

Foi possível verificar com a utilização destas metodologias que vacas leiteiras de alta resposta imunológica têm uma baixa incidência de mastites relativamente aos animais de resposta imunológica baixa. Os animais seleccionados como vantajosos podem depois ser utilizados em programas de melhoramento animal inclusive bovinos, suínos, frangos e outros animais de interesse comercial.

No rastreio de defeito molecular originador de condroplastia hereditária (“Spider Lamb Syndrome” ou SLS) nos ovinos, a patente 6.306.591 de Outubro de 2001 (41) assinala marcadores genéticos para este síndrome e métodos para a sua triagem. Estes marcadores são baseados na presença ou ausência de certos polimorfismos no gene receptor 3 do factor de crescimento fibroblástico (FGFR3) nos ovinos, sendo este polimorfismo de preferência detectado como PCR-RFLP (*polymerase chain reaction-restriction fragment length polymorphisms*) ou SSCP (*single strand conformational polymorphisms*).

A “condrodisplasia hereditária” ou SLS é uma desordem congénita nos ovinos causando severas anormalidades do esqueleto (pernas anormalmente longas e parecidas com as das aranhas coluna vertebral corcunda e torcida costelas e externo deformados, deformidades faciais, perda de gordura corporal e subdesenvolvimento muscular).

### **6.3.3 – Desordens e doenças identificadas nos animais com hereditariedade Mendeliana**

Transcrevem-se da Internet – Mendelian Inheritance in Animals – MIA – estas situações mantendo as designações em Inglês das desordens e doenças em ordem a facilitar em cada caso o seu conhecimento mais profundo o que pode ser conseguido nos seguintes sites:

- <http://www.angis.org.au/databases/BIRBX/omia/species-list.html>
- <http://morgan.angis.su.oz.au>
- <http://probe.nalusola.gov>

#### **Desordens e características com uma base molecular conhecida, da MIA ( 6 Setembro 2000 )**

##### **GATO**

- GANGLIOSIDOSIS, GM2 in CAT
- HYPERLIPOPROTEINAEMIA in CAT
- MANNOSIDOSIS, ALPHA in CAT
- MUCOPOLYSACCHARIDOSIS I in CAT
- MUCOPOLYSACCHARIDOSIS VI in CAT
- MUCOPOLYSACCHARIDOSIS VII in CAT

- MUSCULAR DYSTROPHY, DUCHENNE AND BECKER TYPES in CAT

## **BOVINO**

- CHEDIAK-HIGASHI SYNDROME in CATTLE
- CHRONIC INTERSTITIAL NEPHRITIS WITH DIFFUSE ZONAL FIBROSIS in CATTLE
- CITRULLINAEMIA in CATTLE
- COAT COLOUR, EXTENSION in CATTLE
- COAT COLOUR, ROAN in CATTLE
- DEFICIENCY OF URIDINE MONOPHOSPHATE SYNTHASE in CATTLE
- EHLERS-DANLOS SYNDROME in CATTLE
- EHLERS-DANLOS SYNDROME, TYPE VII in CATTLE
- GLYCOGEN STORAGE DISEASE II in CATTLE
- GLYCOGEN STORAGE DISEASE V in CATTLE
- GOITRE, FAMILIAL in CATTLE
- LEUKOCYTE ADHESION DEFICIENCY in CATTLE
- MANNOSIDOSIS, ALPHA in CATTLE
- MANNOSIDOSIS, BETA in CATTLE
- MAPLE SYRUP URINE DISEASE IN CATTLE
- MUSCULAR HYPERTROPHY in CATTLE
- PROTOPORPHYRIA in CATTLE
- SEX REVERSAL: XY FEMALE in CATTLE
- SPHEROCYTOSIS in CATTLE

## **FRANGO**

- DWARFISM, SEX-LINKED in CHICKEN
- FEATHER COLOUR, ALBINISM in CHICKEN
- FEATHERING, Z-LINKED in CHICKEN
- HENNY FEATHERING in CHICKEN
- NANOMELIA in CHICKEN
- RESISTANCE TO AVIAN SARCOMA AND LEUKOSIS VIRUSES, SUBGROUP B in CHICKEN
- RIBOFLAVINURIA in CHICKEN
- RIBOSOMAL DNA DEFICIENCY in CHICKEN

## **CÃO**

- C3 DEFICIENCY in DOG
- COAT COLOUR, EXTENSION in DOG
- FUCOSIDOSIS, ALPHA in DOG
- GLYCOGEN STORAGE DISEASE I in DOG
- GLYCOGEN STORAGE DISEASE VII in DOG
- HAEMOPHILIA B in DOG
- KRABBE DISEASE in DOG
- LEUKOCYTE ADHESION DEFICIENCY in DOG
- MUCOPOLYSACCHARIDOSIS I in DOG
- MUCOPOLYSACCHARIDOSIS VII in DOG
- MUSCULAR DYSTROPHY, DUCHENNE AND BECKER TYPES in DOG
- MYOTONIA in DOG

- NARCOLEPSY in DOG
- NEPHRITIS, X-LINKED in DOG
- PYRUVATE KINASE DEFICIENCY OF ERYTHROCYTE in DOG
- RETINAL PIGMENT EPITHELIAL DYSTROPHY in DOG
- ROD-CONE DYSPLASIA-1 in DOG
- ROD-CONE DYSPLASIA-3 in DOG
- SEVERE COMBINED IMMUNODEFICIENCY DISEASE, X-LINKED in DOG
- TREMOR, X-LINKED in DOG
- VON WILLEBRAND DISEASE III in DOG

### **RAPOSA**

- COAT COLOUR, AGOUTI in FOX
- COAT COLOUR, EXTENSION in FOX

### **CABRA**

- GOITRE FAMILIAL in GOAT
- MANNOSIDOSIS, BETA in GOAT
- MUCOPOLYSACCHARIDOSIS IIID in GOAT
- REDUCED CASEIN CONCENTRATION in GOAT

### **HAMSTER**

- CARDIOMYOPATHY, DILATED in HAMSTER
- CARDIOMYOPATHY, HYPERTROPHIC in HAMSTER

### **CAVALO**

- COAT COLOUR, EXTENSION in HORSE
- MEGACOLON in HORSE
- PERIODIC PARALYSIS II in HORSE
- SEVERE COMBINED IMMUNODEFICIENCY DISEASE, AUTOSOMAL in HORSE
- SEX REVERSAL: XY FEMALE in HORSE

### **MEDAKA**

- COAT COLOUR, ALBINISM in MEDAKA

### **MACACO**

- KRABBE DISEASE in MONKEY, RHESUS

### **SUINO**

- COAT COLOUR, DOMINANT WHITE in PIG
- HYPERCHOLESTEROLAEMIA in PIG
- MALIGNANT HYPERTHERMIA in PIG
- MEAT QUALITY in PIG

## **CODORNIZ**

- HYPOTROPHIC AXONOPATHY in QUAIL

## **COELHO**

- ADRENAL HYPERPLASIA, CONGENITAL in RABBIT
- C8 DEFICIENCY in RABBIT
- TREMOR, X-LINKED in RABBIT

## **OVINO**

- CEROID LIPOFUSCINOSIS in SHEEP
- CHONDRODYSPLASIA in SHEEP
- COAT COLOUR, EXTENSION in SHEEP
- GLYCOGEN STORAGE DISEASE V in SHEEP

### **Desordens e características com um marcador ligado, da MIA ( 6 Setembro 2000 )**

- ATAXIA, PROGRESSIVE in PIG
- CEROID LIPOFUSCINOSIS in DOG
- COAT COLOUR, DOMINANT WHITE in CATTLE
- COAT COLOUR, SPOTTED in CATTLE
- DERMATOFIBROSIS in DOG
- FECUNDITY, BOORoola in SHEEP
- HORNS in CATTLE
- HORNS in GOAT
- HORNS in SHEEP
- HYPOTRICHOSIS WITH INCISOR ANODONTIA in CATTLE
- NEONATAL DIARRHOEA in PIG
- PROGRESSIVE RETINAL ATROPHY in DOG
- RENAL CYSTADENOCARCINOMA AND NODULAR DERMATOFIBROSIS in DOG
- RENAL DYSPLASIA in CATTLE
- RESISTANCE TO AVIAN SARCOMA AND LEUKOSIS VIRUSES, SUBGROUP A in CHICKEN
- ROD-CONE DEGENERATION, PROGRESSIVE in DOG
- SPINAL DYSMYELINATION in CATTLE
- SPONGIFORM ENCEPHALOPATHY in SHEEP
- SYNDACTYL in CATTLE
- TREMOR, HIGH-FREQUENCY in PIG
- VON WILLEBRAND DISEASE in DOG
- WILSON DISEASE in DOG

### **Desordens e características com um gene / péptido conhecido, da MIA ( 6 Setembro 2000 )**

- ADRENAL HYPERPLASMA, CONGENITAL
- AFIBRINOGENAEMIA
- ANALPHALIPOPROTEINAEMIA
- ARGININAEMIA

- C3 DEFICIENCY
- C4 DEFICIENCY
- C8 DEFICIENCY
- CARDIOMYOPATHY, DILATED
- CARDIOMYOPATHY, HYPERTROPHIC
- CHRONIC INTERSTITIAL NEPHRITIS WITH DIFFUSE ZONAL FIBROSIS
- CITRULLINAEMIA
- COAT COLOUR, AGOUTI
- COAT COLOUR, ALBINISM
- COAT COLOUR, DOMINANT WHITE
- COAT COLOUR, EXTENSION
- DEFICIENCY OF URIDINE MONOPHOSPHATE SYNTHASE
- DIPLOPODIA-5
- DWARFISM, CROOKED NECK
- DWARFISM, GROWTH-HORMONE-RECEPTOR DEFICIENCY
- DWARFISM, SEX-LINKED
- EHLERS-DANLOS SYNDROME
- EHLERS-DANLOS SYNDROME, TYPE VII
- EPIDERMOLYSIS BULLOSA, DYSTROPHIC
- FACTOR VII DEFICIENCY
- FACTOR X DEFICIENCY
- FACTOR XI DEFICIENCY
- FACTOR XII DEFICIENCY
- FEATHER COLOUR, ALBINISM
- FUCOSIDOSIS, ALPHA
- GANGLIOSIDOSIS, GM1
- GANGLIOSIDOSIS, GM2
- GAUCHER DISEASE, TYPE I
- GLUCOSE-6-PHOSPHATE DEHYDROGENASE DEFICIENCY
- GLYCOGEN STORAGE DISEASE I
- GLYCOGEN STORAGE DISEASE II
- GLYCOGEN STORAGE DISEASE IV
- GLYCOGEN STORAGE DISEASE V
- GLYCOGEN STORAGE DISEASE VII
- GOITRE, FAMILIAL
- HAEMOPHILIA A
- HAEMOPHILIA B
- HENNY FEATHERING
- HYPERBILIRUBINAEMIA I
- HYPERCHOLESTEROLAEMIA
- HYPERLIPOPROTEINAEMIA
- HYPOPROTHROMBINAEMIA
- HYPOTROPHIC AXONOPATHY
- INTESTINAL COBALAMIN MALABSORPTION
- KRABBE DISEASE
- LEUKOCYTE ADHESION DEFICIENCY
- LIMBLESS
- MALIGNANT HYPERTHERMIA
- MANNOSIDOSIS, ALPHA
- MANNOSIDOSIS, BETA

- MAPLESYRUP URINE DISEASE
- MEMBRANOPROLIFERATIVE GLOMERULONEPHRITIS TYPE II
- MENKES SYNDROME
- MUCOLIPIDOSIS II
- MUCOPOLYSACCHARIDOSIS I
- MUCOPOLYSACCHARIDOSIS II
- MUCOPOLYSACCHARIDOSIS IIIA
- MUCOPOLYSACCHARIDOSIS IIID
- MUCOPOLYSACCHARIDOSIS VI
- MUCOPOLYSACCHARIDOSIS VII
- MUSCULAR DYSTROPHY, DUCHENNE AND BECKER TYPES
- MUSCULAR HYPERTROPHY
- MYOCLONUS
- NANOMALIA
- NARCOLEPSY
- NEPHRITIS, X-LINKED
- NIEMANN-PICK DISEASE
- OBESITY
- OXALOSIS II
- PERIODIC PARALYSIS II
- PORPHYRIA CUTANEA TARDA
- PORPHYRIAL, CONGENITAL ERYTHROPOIETIC
- PROTAMINE-2 DEFICIENCY
- PROTEIN C DEFICIENCY
- PROTOPORPHYRIA
- PSEUDO-VITAMIN D DEFICIENCY RICKETS
- PYROVATE KINASE DEFICIENCY OF ERYTHROCYTE
- REDUCED CASEIN CONCENTRATION
- REDUCED GLUTATHIONE DEFICIENCY DUE TO GCS DEFICIENCY
- RESISTANCE TO OEDEMA DISEASE
- RESTRICTED OVULATOR
- RETINAL DEGENERATION I
- RETINAL DEGENERATION II
- RIBOFLAVINURIA
- RIBOSOMAL DNA DEFICIENCY
- ROD-CONE DYSPLASIA-1
- ROD-CONE DYSPLASIA-3
- SEVERE COMBINED IMMUNODEFICIENCY DISEASE, X-LINKED
- SEVERE COMBINED IMMUNODEFICIENCY DISEASE, AUTOSOMAL
- SEX REVERSAL: XY FEMALE
- SHAKER
- SPHEROCYTOSIS
- SPINAL MUSCULAR ATROPHY
- SPONGIFORM ENCEPHALOPATHY
- TALPID-2
- TESTICULAR FEMINIZATION
- TREMOR, X-LINKED
- TYROSINE TRANSAMINASE DEFICIENCY
- VITAMIN-K-DEPENDENT BLOOD COAGULATION FACTORS DEFICIENCY
- VON WILLEBRAND DISEASE



- XANTHINURIA

**Modelos animais potenciais de desordens humanas, da MIA ( 6 Setembro 2000)**

- ADDISON DISEASE
- ADRENAL HYPERPLASIA, CONGENITAL
- AFIBRINOGENAEMIA
- AGNATHIA
- ALEXANDER DISEASE ALKAPTONURIA
- ALOPECIA
- ALPHA-1-ANTITRYPSIN DEFICIENCY
- ALZHEIMER DISEASE
- AMYLOIDOSIS
- AMYLOIDOSIS, AA
- ANALPHALIPOPROTEINAEMIA
- ANENCEPHALY
- ANISOCORIA
- ANKYLOSING SPONDYLITIS
- ANODONTIA
- ANTERIOR SEGMENT DYSGENESIS SYNDROME
- ARGININAEMIA
- ARNOLD-CHIARI MALFORMATION
- ARTHRITIS
- ARTHRITIS, RHEUMATOID
- ARTHROGRYPOSIS
- ASCITES
- ATAXIA
- ATHEROSCLEROSIS
- ATOPY
- ATRESIA COLI
- ATRIAL FIBRILLATION
- ATRIAL SEPTAL DEFECT
- AUTOIMMUNE THYROIDITIS, SPONTANEOUS
- BINUCLEATED RED CELL
- BUDD-CHIARI SYNDROME
- BULLOUS PEMPHIGOID ANTIGEN
- C3 DEFICIENCY
- C4 DEFICIENCY
- C8 DEFICIENCY
- CARDIOMYOPATHY, DILATED
- CARDIOMYOPATHY, HYPERTROPHIC
- CATARACT
- CEREBELLAR HYPOPLASIA
- CEROID LIPOFUSCINOSIS
- CEROID LIPOFUSCINOSIS, JUVENILE ONSET
- CHEDIAK-HIGASHI SYNDROME
- CHRONIC INTERSTITIAL NEPHRITIS WITH DIFFUSE ZONAL FIBROSIS
- CITRULLINAEMIA
- CLEFT LIP
- CLEFT PALATE

- COAT COLOUR, AGOUTI
- COAT COLOUR, ALBINISM
- COAT COLOUR, DOMINANT WHITE
- COAT COLOUR, EXTENSION
- COELIAC SPRUE
- COLOBOMA
- COMEDO SYNDROME
- CONOTRUNCAL HEART MALFORMATIONS
- CORONAL SUTURE SYNOSTOSIS
- CRANIOSYNOSTOSIS
- CRYPTORCHIDISM
- CISHING DISEASE
- CYANOSIS
- CYCLOPIA
- CYSTIC OVARY
- DANDY-WALKER SYNDROME
- DEFICIENCY OF URIDINE MONOPHOSPHATE SYNTHASE
- DERMATOFIBROSIS
- DERMATOSIS
- DETACHED RETINA
- DIABETES INSIPIDUS
- DIABETES MELLITUS
- DIABETES MELLITUS, TYPE I
- DIABETES MELLITUS, TYPE II
- DWARFISM, CROOKED NECK
- DWARFISM, GROWTH-HORMONE-RECEPTOR DEFICIENCY
- DWARFISM, PITUITARY
- DWARFISM, SEX-LINKED
- DYSAUTONOMIA
- ECLAMPSIA
- ECTODERMAL DYSPLASIA, X-LINKED
- ECTRODACTYLY
- ECTROPION
- EHLERS-DANLOS SYDROME
- EHLERS-DANLOS SYDROME,TYPE VII
- ENCEPHALOMYELOPATHY
- EPIDERMOLYSIS BULLOSA
- EPIDERMOLYSIS BULLOSA,DYSTROPHIC
- EPIDERMOLYSIS BULLOSA,DYSTROPHIC
- EPIDERMOLYSIS BULLOSA, JUNCTIONALIS
- EPILEPSY
- EPISTAXIS
- EPITHELIOGENESIS IMPERFECTA
- ERYTHROCYTOSIS
- EXOPHTHALMOS WITH STRABISMUS
- EXOSTOSIS, MULTIPLE
- FACTOR VII DEFICIENCY
- FACTOR X DEFICIENCY
- FACTOR XI DEFICIENCY
- FACTOR XII DEFICIENCY

- FANCONI SYNDROME
- FATTY METAMORPHOSIS OF VISCERA
- FEATHER COLOUR, ALBINISM
- FECUNDITY, BOORoola
- FIBRODYSPLASIA OSSIFICANS
- FRAGILE X
- FRAGILE SITE
- FUCOSIDOSIS, ALPHA
- GALACTOKINASE DEFICIENCY
- GALACTOSIALIDOSIS
- GANGLIOSIDOSIS, GM1
- GANGLIOSIDOSIS, GM1
- GANGLIOSIDOSIS, GM2
- GAUCHER DISEASE, TYPE I
- GENU VALGUM
- GLAUCOMA
- GLOMERULONEPHRITIS
- GLUCOSE-6-PHOSPHATE DEHYDROGENASE DEFICIENCY
- GLUTEN-SENSITIVE ENTEROPATHY
- GLYCOGEN STORAGE DISEASE I
- GLYCOGEN STORAGE DISEASE II
- GLYCOGEN STORAGE DISEASE IV
- GLYCOGEN STORAGE DISEASE V
- GLYCOGEN STORAGE DISEASE VII
- GLYOXALASE II DEFICIENCY
- GOITRE, FAMILIAL
- GONIODYSPLASIA, MESODERMAL
- HAEMOCHROMATOSIS
- HAEMOLYTIC ANAEMIA, AUTOIMMUNE
- HAEMOPHAGOCYTTIC SYNDROME
- HAEMOPHILIA A
- HAEMOPHILIA B
- HEMERALOPIA
- HEMIVERTEBRAE
- HENNY FEATHERING
- HEPATIC FIBROSIS, IDIOPATHIC
- HERNIA, DIAPHRAGMATIC
- HERNIA, HIATAL
- HERNIA, INGUINAL
- HETEROCHROMIA IRIDIS
- HIP DYSPLASIA
- HODGKIN DISEASE
- HORNER SYNDROME
- HORNS
- HUNTINGTON DISEASE
- HYDROCEPHALUS
- HYDROMELIA
- HYDROPS FOETALIS
- HYMEN, IMPERFORATE
- HYPERADRENOCORTICISM

- HYPERALDOSTERONISM
- HYPERBILIRUBINAEMIA I
- HYPERBILIRUBINAEMIA II
- HYPERBILIRUBINAEMIA, UNCLASSIFIED
- HYPERCHOLESTEROLAEMIA
- HYPERGAMMAGLOBULINAEMIA
- HYPERHIDROSIS
- HYPERKERATOSIS, PALMOPLANTAR
- HYPERLIPIDAEMIA
- HYPERLIPOPROTEINAEMIA
- HYPEROSTOSIS
- HYPERPARATHYROIDISM
- HYPERPHOSPHATAEMIA
- HYPERTHYROIDISM
- HYPOADRENOCORTICISM
- HYPOCATALASIA
- HYPOLIPOPROTEINAEMIA
- HYPOPARATHYROIDISM
- HYPOPHOSPHATAEMIA
- HYPOPROTHROMBINAEMIA
- HYPOSPADIAS
- HYPOTHYROIDISM
- HYPOTRICHOSIS
- HYPOTRICHOSIS WITH INCISOR ANODONTIA
- HYPOTRICHOSIS, DOMINANT
- HYPOTRICHOSIS, RECESSIVE
- HYPOTROPHIC AXONOPATHY
- INTESTINAL COBALAMIN MALABSORPTION
- INTUSSUSCEPTION
- KARTAGENER SYNDROME
- KERATITIS
- KRABBE DISEASE
- LEGG-CALVE-PERTHES DISEASE
- LETHAL TRAIT A46
- LEUKOCYTE ADHESION DEFICIENCY
- LEUKOCYTE ADHESION DEFICIENCY
- LINEAR IGA DISEASE
- LYMPHANGIECTASIA
- LYSOSOMAL STORAGE DISEASE
- MACROSTOMUS
- LEUKOCYTE ADHESION DEFICIENCY
- MALIGNANT HYPERTHERMIA
- MANNOSIDOSIS, ALPHA
- MANNOSIDOSIS, BETA
- MAPLESYRUP URINE DISEASE
- MARFAN SYNDROME
- MASTICATORY MUSCLE MYOSITIS
- MECKEL DIVERTICULUM
- MEGACOLON
- MELANOMA, SINCLAIR SWINE CUTANEOUS MALIGNANT

- MEMBRANOPROLIFERATIVE GLOMERULONEPHRITIS TYPE II
- MENKES SYNDROME
- MESANGIOCAPILLARY GLOMERULONEPHRITIS, TYPE I
- METHAEMOGLOBINAEMIA
- MICROMELIC SYNDROME
- MITOCHONDRIAL MYOPATHY
- MOTOR NEURON DISEASE
- MUCOLIPIDOSIS II
- MUCOPOLYSACCHARIDOSIS I
- MUCOPOLYSACCHARIDOSIS II
- MUCOPOLYSACCHARIDOSIS IIIA
- MUCOPOLYSACCHARIDOSIS IIID
- MUCOPOLYSACCHARIDOSIS VI
- MUCOPOLYSACCHARIDOSIS VII
- MUSCULAR DYSTROPHY, DUCHENNE AND BECKER TYPES
- MYASTHENIA GRAVIS
- MYCOSIS FUNGOIDES
- MYOCLONUS
- MYOCLONUS EPILEPSY OF LAFORA
- MYOTONIC DYSTROPHY
- NARCOLEPSY
- NECROTISING ENCEPHALOPATHY, SUBACUTE, OF LEIGH
- NEPHRITIS, X-LINKED
- NEPHRITIS, AUTOSOMAL
- NEPHROTIC SYNDROME
- NEUROFIBROMATOSIS
- NEURONAL ABIOTROPHY
- NEUROPATHY, GIANT AXONAL
- NEUROPATHY, PERIPHERAL
- NEUTROPENIA, CYCLIC
- NIEMANN-PICK DISEASE
- NIEMANN-PICK DISEASE, TYPE C
- NIPPLES, INVERTED
- OBESITY
- OMPHALOCELE
- ONYCHODYSTROPHY
- OSTEOCHONDROMATOSIS
- OSTEOCHONDROSIS
- OSTEOCHONDROSIS DISSECANS
- OSTEODYSTROPHY
- OSTEOGENESIS IMPERFECTA
- OSTEOPETROSIS
- OTITIS MEDIA, SUSCEPTIBILITY TO
- OXALOSIS II
- PANCREATIC INSUFFICIENCY, EXOCRINE
- PAPILLOMATOSIS, CUTANEOUS
- PATENT DUCTUS ARTERIOSUS
- PELGER-HUET ANOMALY
- PEMPHIGUS
- PERIODIC PARALYSIS II

- PEROMELIA
- PERSISTENT MULLERIAN DUCT SYNDROME
- PERSISTENT TRUNCUS ARTERIOSUS WITH VENTRICULAR SEPTAL DEFECT AND
- PATENT FORAMEN OVALE
- PHEOCHROMOCYTOMA
- PLATELET DELTA-STORAGE POOL DISEASE
- PNEUMOTHORAX
- POLYCYSTIC KIDNEY DISEASE
- POLYCYTHEMIA
- POLYGLANDULAR AUTOIMMUNE SYNDROME, TYPE II
- POLYURIA
- PORPHYRIA CUTANEA TARDA
- PORPHYRIA, CONGENITAL ERYTHROPOIETIC
- PORPHYRIA, UNCLASSIFIED
- PORTOSYSTEMIC SHUNT
- PREKALLIKREIN DEFICIENCY
- PRIAPISM
- PROGNATHISM
- PROTAMINE-2 DEFICIENCY
- PROTEIN C DEFICIENCY
- PROTOPORPHYRIA
- PROTOPORPHYRIA
- PSEUDO-VITAMIN D DEFICIENCY RICKETS
- PSEUDOCHOLINESTERASE DEFICIENCY
- PYRUVATE KINASE DEFICIENCY OF ERYTHROCYTE
- REDUCED CASEIN CONCENTRATION
- REDUCED GLUTATHIONE DEFICIENCY DUE TO GCS DEFICIENCY
- RENAL DYSPLASIA
- RESPIRATORY DISTRESS SYNDROME
- RESTRICTED OVULATOR
- RETINAL AND SKELETAL DYSPLASIA
- RETINAL DEGENERATION I
- RETINAL DEGENERATION II
- RETINAL PIGMENT EPITHELIAL DYSTROPHY
- RETINOSCHISIS
- RHABDOMYOLYSIS
- RIBOSOMAL DNA DEFICIENCY
- RIGHT VENTRICULAR CARDIOMYOPATHY
- ROD-CONE DEGENERATION, PROGRESSIVE
- ROD-CONE DYSPLASIA-1
- ROD-CONE DYSPLASIA-3
- SARCOID
- SCLERODERMA
- SCOLIOSIS
- SELF-MUTILATION SYNDROME
- SEVERE COMBINED IMMUNODEFICIENCY DISEASE, X-LINKED
- SEVERE COMBINED IMMUNODEFICIENCY DISEASE, AUTOSOMAL
- SEX REVERSAL: XX MALE
- SEX REVERSAL: XY FEMALE
- SHAKER

- SITUS INVERSUS
- SPHEROCYTOSIS
- SPHEROCYTOSIS
- SPINA BIFIDA
- SPINAL MUSCULAR ATROPHY
- SPONDYLOSIS DEFORMANS
- SPONGIFORM ENCEPHALOPATHY
- SPONGIFORM MYELOPATHY
- STOMATOCYTOSIS
- STRABISMUS
- SUBAORTIC STENOSIS
- SYNDACTYLY
- SYNOVIAL CHONDROMATOSIS
- SYRINGOMYELIA
- SYSTEMIC LUPUS ERYTHEMATOSUS
- TACHYCARDIA
- TAIL, SHORT
- TEAT NUMBER
- TESTICULAR FEMINIZATION
- TETRALOGY OF FALLOT
- THROMBASTHENIA
- THROMBOCYTOPAENIA
- THROMBOCYTOPATHY
- THROMBOCYTOPENIC PURPURA, AUTOIMMUNE
- THROMBOPATHIA
- TIBIAL HEMIMELIA
- TORTICOLLIS
- TREMOR
- TREMOR, X-LINKED
- TREMOR, X-LINKED
- TREMOR, HIGH-FREQUENCY
- TYROSINE TRANSAMINASE DEFICIENCY
- UROLITHIASIS
- URTICARIA PIGMENTOSA
- VASCULOPATHY
- VERTEBRAL ANOMALIES
- VITAMIN-K-DEPENDENT BLOOD COAGULATION FACTORS DEFICIENCY
- VITAMIN-K-DEPENDENT BLOOD COAGULATION FACTORS DEFICIENCY
- VITILIGO
- VON WILLEBRAND DISEASE
- VON WILLEBRAND DISEASE I
- VON WILLEBRAND DISEASE III
- WAARDENBURG SYNDROME
- WILMS TUMOUR
- WILSON DISEASE
- WOLFF-PARKINSON-WHITE SYNDROME
- XANTHINURIA

**Erros inatos do metabolismo da MIA ( 6 Setembro 2000 )**

- ADRENAL HYPERPLASIA, CONGENITAL

- ARGININAEMIA
- CEROID LIPOFUSCINOSIS
- CEROID LIPOFUSCINOSIS WITH BROWN BOWEL SYNDROME
- CEROID LIPOFUSCINOSIS JUVENILE ONSET
- CITRULLINAEMIA
- COAT COLOUR, ALBINISM
- DEFICIENCY OF URIDINE MONOPHOSPHATE SYNTHASE
- FUCOSIDOSIS, ALPHA
- GALACTOKINASE DEFICIENCY
- GALACTOSIALIDOSIS
- GANGLIOSIDOSIS
- GANGLIOSIDOSIS, GM1
- GANGLIOSIDOSIS, GM2
- GAUCHER DISEASE, TYPE 1
- GLUCOSE-6-PHOSPHATE DEHYDROGENASE DEFICIENCY
- GLYCOGEN STORAGE DISEASE
- GLYCOGEN STORAGE DISEASE I
- GLYCOGEN STORAGE DISEASE II
- GLYCOGEN STORAGE DISEASE IV
- GLYCOGEN STORAGE DISEASE V
- GLYCOGEN STORAGE DISEASE VII
- GLYOXALASE II DEFICIENCY
- HAEMOLYTIC ANAEMIA
- HENNY FEATHERING
- HYPERLIPOPROTEINAEMIA > HYPERTRIGLYCERIDAEMIA
- HYPOCATALASIA
- KRABBE DISEASE
- LYSOSOMAL STORAGE DISEASE
- MANNOSIDOSIS, ALPHA
- MANNOSIDOSIS, BETA
- MAPLE SYRUP URINE DISEASE
- MUCOLIPIDOSIS II
- MUCOPOLYSACCHARIDOSIS
- MUCOPOLYSACCHARIDOSIS I
- MUCOPOLYSACCHARIDOSIS II
- MUCOPOLYSACCHARIDOSIS IIIA
- MUCOPOLYSACCHARIDOSIS IIID
- MUCOPOLYSACCHARIDOSIS VI
- MUCOPOLYSACCHARIDOSIS VII
- NIEMANN-PICK DISEASE
- NIEMANN-PICK DISEASE, TYPE C
- PORPHYRIA CUTANEA TARDA
- PORPHYRIA, CONGENITAL ERYTHROPOIETIC
- PORPHYRIA, UNCLASSIFIED
- PROTOPORPHYRIA
- PSEUDO-VITAMIN D DEFICIENCY RICKETS
- REDUCED GLUTATHIONE DEFICIENCY DUE TO GCS DEFICIENCY

**Doenças de depósito lisossomal da MIA ( 6 Setembro 2000 )**



- CEROID LIPOFUSCINOSIS
- CEROID LIPOFUSCINOSIS WITH BROWN BOWEL SYNDROME
- CEROID LIPOFUSCINOSIS, JUVENILE ONSET
- FUCOSIDOSIS, ALPHA
- GALACTOSIALIDOSIS
- GANGLIOSIDOSIS
- GANGLIOSIDOSIS, GM1
- GANGLIOSIDOSIS, GM2
- GAUCHER DISEASE, TYPE 1
- GLYCOGEN STORAGE DISEASE
- GLYCOGEN STORAGE DISEASE I
- GLYCOGEN STORAGE DISEASE II
- GLYCOGEN STORAGE DISEASE IV
- GLYCOGEN STORAGE DISEASE V
- GLYCOGEN STORAGE DISEASE VII
- KRABBE DISEASE
- LYSOSOMAL, STORAGE DISEASE
- MANNOSIDOSIS, ALPHA
- MANNOSIDOSIS, BETA
- MUCOLIPIDOSIS II
- MUCOPOLYSACCHARIDOSIS
- MUCOPOLYSACCHARIDOSIS I
- MUCOPOLYSACCHARIDOSIS II
- MUCOPOLYSACCHARIDOSIS IIIA
- MUCOPOLYSACCHARIDOSIS IIID
- MUCOPOLYSACCHARIDOSIS VI
- MUCOPOLYSACCHARIDOSIS VII
- NIEMANN-PICK DISEASE
- NIEMANN-PICK DISEASE, TYPE C

**Desordens hereditárias de hemorragia ( 6 Setembro 2000 )**

- AFIBRINOGENOGENAEMIA
- BLEEDING DIATHESIS
- BLEEDING DISORDER
- FACTOR VII DEFICIENCY
- FACTOR X DEFICIENCY
- FACTOR XI DEFICIENCY
- FACTOR XII DEFICIENCY
- HAEMOPHILIA A
- HAEMOPHILIA B
- HYPOPROTHROMBINAEMIA
- THROMBOPATHIA

**Doenças congénitas do coração, da MIA (6 Setembro 2000)**

- AORTIC STENOSIS, SUBVALVULAR
- AORTICOPULMONARY SEPTAL DEFECT
- ATRIAL FIBRILLATION
- ATRIAL SEPTAL DEFECT

- ATRIAL SEPTAL DEFECT AND ATRIAL FIBRILLATION
- CARDIAC ANOMALY
- CARDIOMYOPATHY
- CARDIOMYOPATHY AND WOOLLY HAIRCOAT SYNDROME
- CARDIOMYOPATHY, DILATED
- CARDIOMYOPATHY, HYPERTROPHIC
- CARDIOMYOPATHY, SPONTANEOUS
- CARDIOVASCULAR MALFORMATIONS
- CHRONIC VALVULAR DISEASE
- CONOTRUNCAL HEART MALFORMATIONS
- HEART DEFECT, CONGENITAL
- HEART DISEASE, HYPERTROPHIC
- MITRAL STENOSIS
- MITRAL VALVE DISEASE
- MITRAL VALVE STENOSIS
- PATENT DUCTUS ARTERIOSUS
- PATENT DUCTUS ARTERIOSUS AND PULMONARY HYPERTENSION
- PATENT DUCTUS VENOSUS
- PERSISTENT RIGHT AORTIC ARCH
- PERSISTENT TRUNCUS ARTERIOSUS
- PERSISTENT TRUNCUS ARTERIOSUS WITH VENTRICULAR SEPTAL DEFECT AND PATENT FORAMEN OVALE
- SUBAORTIC STENOSIS
- SUBVALVULAR PULMONARY STENOSIS
- VENTRICULAR ARRHYTHMIAS AND SUDDEN DEATH
- VENTRICULAR SEPTAL DEFECT
- VENTRICULAR SEPTAL DEFECT AND ATRIAL FIBRILLATION
- VENTRICULAR SEPTAL DEFECT WITH ATRIOVENTRICULAR VALVULAR ANOMALY

**Nanismo da MIA (6 Setembro 2000)**

- ACHONDROPLASIA
- ACHONDROPLASIA FOETALIS
- ACHONDROPLASIA, CREEPER
- CHONDRODYSPLASIA
- CHONDRODYSTROPHY
- DWARFISM
- DWARFISM WITH ANAEMIA
- DWARFISM WITH RETINAL DYSPLASIA
- DWARFISM, ANCON
- DWARFISM, DEXTER
- DWARFISM, AUTOSOMAL
- DWARFISM, CHONDRODYSTROPHY
- DWARFISM, CHONDROPLASTIC
- DWARFISM, CROOKED NECK
- DWARFISM, PITUITARY
- DWARFISM, PROPORTIONATE
- DWARFISM, SEX-LINKED
- DWARFISM, SNORTER

- DWARFISM, STUMPY
- JOINT LAXITY AND DWARFISM, CONGENITAL
- MINIATURE
- NANOMELIA
- PSEUDOACHONDROPLASTIC DYSPLASIA

**Desordens e características com hereditariedade num único locus da MIA (6 Setembro 2000)**

- ACHONDROPLASIA, CREEPER in CHICKEN
- ACHROMATOSIS in CHICKEN
- ADRENAL HYPERPLASIA, CONGENITAL in RABBIT
- AFIBRINOGENAEMIA in DOG
- AFIBRINOGENAEMIA in GOAT
- AMETAPODIA in CHICKEN
- ANALPHALIPOPROTEINAEMIA in CHICKEN
- ANTERIOR SEGMENT DYSGENESIS SYNDROME in HORSE
- ARGININAEMIA in SHEEP
- ARTHROGRYPOSIS in PIG
- ATAXIA, PROGRESSIVE in PIG
- ATYPICAL MITOTIC METAPHASE in QUAIL
- BALDNESS, CONGENITAL in CHICKEN
- BARRING in CHICKEN
- BINUCLEATED RED CELL in TURKEY
- BLACK HAIR FOLLICLE DYSPLASIA in DOG
- BLASTODERM DEGENERATION in CHICKEN
- BLOOD GROUP SYSTEM in SHEEP
- BLOOD GROUP SYSTEM A in SHEEP
- BLOOD GROUP SYSTEM B in SHEEP
- BLOOD GROUP SYSTEM C in SHEEP
- BLOOD GROUP SYSTEM D in HORSE
- BLOOD GROUP SYSTEM D in SHEEP
- BLOOD GROUP SYSTEM F30 in SHEEP
- BLOOD GROUP SYSTEM F41 in SHEEP
- BLOOD GROUP SYSTEM I in SHEEP
- BLOOD GROUP SYSTEM M in SHEEP
- BLOOD GROUP SYSTEM O in PIG
- BLOOD GROUP SYSTEM R in SHEEP
- BLOOD GROUP SYSTEM RH in DOG
- BLOOD GROUP SYSTEM X in SHEEP
- BLOOD RING in TURKEY
- BLUE EGGSHELL in CHICKEN
- BOBBER in TURKEY
- C3 DEFICIENCY in DOG
- C3 DEFICIENCY in RABBIT
- C8 DEFICIENCY in RABBIT
- CARDIOMYOPATHY in CATTLE
- CARDIOMYOPATHY in HAMSTER
- CARDIOMYOPATHY, DILATED in CATTLE
- CARDIOMYOPATHY, DILATED in HAMSTER

- CARDIOMYOPATHY, HYPERTROPHIC in CAT
- CARDIOMYOPATHY, HYPERTROPHIC in HAMSTER
- CARPAL SUBLUXATION in DOG
- CATARACT in DOG
- CEREBELLAR DEGENERATION in CAT
- CEROID LIPOFUSCINOSIS in CAT
- CEROID LIPOFUSCINOSIS in CATTLE
- CEROID LIPOFUSCINOSIS in DOG
- CEROID LIPOFUSCINOSIS in HUMAN
- CEROID LIPOFUSCINOSIS in SHEEP
- CEROID LIPOFUSCINOSIS, JUVENILE ONSET in DOG
- CEROID LIPOFUSCINOSIS, JUVENILE ONSET in SHEEP
- CHEDIAK-HIGASHI SYNDROME in CAT
- CHEDIAK-HIGASHI SYNDROME in CATTLE
- CHEDIAK-HIGASHI SYNDROME in FOX
- CHEDIAK-HIGASHI SYNDROME in HUMAN
- CHEDIAK-HIGASHI SYNDROME in MINK
- CHEDIAK-HIGASHI SYNDROME in MOUSE
- CHEDIAK-HIGASHI SYNDROME in RAT
- CHONDRODYSPLASIA in CATTLE
- CHONDRODYSPLASIA in SHEEP
- CHONDRODYSTROPHY in QUAIL
- CHRONIC INTERSTITIAL NEPHRITIS WITH DIFFUSE ZONAL FIBROSIS in CATTLE
- CIRCADIAN TIMING in HAMSTER
- CITRULLINAEMIA in CATTLE
- CLENCH in QUAIL
- COAT COLOUR in DOG
- COAT COLOUR in SALAMANDER
- COAT COLOUR, AGOUTI in CATTLE
- COAT COLOUR, AGOUTI in FOX
- COAT COLOUR, AGOUTI in SHEEP
- COAT COLOUR, ALBINISM in CAT
- COAT COLOUR, ALBINISM in CATTLE
- COAT COLOUR, ALBINISM in CHICKEN
- COAT COLOUR, ALBINISM in DOG
- COAT COLOUR, ALBINISM in HAMSTER
- COAT COLOUR, ALBINISM in LOCUST
- COAT COLOUR, ALBINISM in MEDAKA
- COAT COLOUR, ALBINISM in RABBIT
- COAT COLOUR, ALBINISM in SHEEP
- COAT COLOUR, BLACK CRYSTAL in MINK
- COAT COLOUR, BROWN in CATTLE
- COAT COLOUR, CHAMPAGNE in HORSE
- COAT COLOUR, DOMINANT WHITE in CATTLE
- COAT COLOUR, DOMINANT WHITE in PIG
- COAT COLOUR, EXTENSION in CATTLE
- COAT COLOUR, EXTENSION in DOG
- COAT COLOUR, EXTENSION in FOX
- COAT COLOUR, EXTENSION in HORSE
- COAT COLOUR, EXTENSION in SHEEP

- COAT COLOUR, LETHAL DOMINANT ROAN in HORSE
- COAT COLOUR, MERLE in DOG
- COAT COLOUR, ORANGE in CAT
- COAT COLOUR, PINK-EYED DILUTION in HAMSTER
- COAT COLOUR, ROAN in CATTLE
- COAT COLOUR, ROAN in HORSE
- COAT COLOUR, SILVER-BLUE in MINK
- COAT COLOUR, SPOTTED in CATTLE
- COAT COLOUR, WHITE, LETHAL, DOMINANT in HORSE
- COAT COLOUR, WHITE, LETHAL, RECESSIVE in HORSE
- COLOBOMA in CHICKEN
- CRANIOSYNOSTOSIS in RABBIT
- CREST in CHICKEN
- CREST in GOOSE
- CYSTINURIA in DOG
- DEAFNESS in DOG
- DEFICIENCY OF URIDINE MONOPHOSPHATE SYNTHASE in CATTLE
- DERMATOFIBROSIS in DOG
- DIPLOPODIA in CHICKEN
- DIPLOPODIA-2 in CHICKEN
- DIPLOPODIA-3 in CHICKEN
- DIPLOPODIA-4 in CHICKEN
- DIPLOPODIA-5 in CHICKEN
- DWARFISM in PIG
- DWARFISM, ANCON in SHEEP
- DWARFISM, AUTOSOMAL in CHICKEN
- DWARFISM, CROOKED NECK in CHICKEN
- DWARFISM, DEXTER in CATTLE
- DWARFISM, GROWTH-HORMONE-RECEPTOR DEFICIENCY in CATTLE
- DWARFISM, HYPOCHONDROPLASTIC in DOG
- DWARFISM, PITUITARY in DOG
- DWARFISM, SEX-LINKED in CHICKEN
- DWARFISM, SNORTER in CATTLE
- DYSERYTHROPOIESIS in CATTLE
- ECTODERMAL DYSPLASIA, X-LINKED in DOG
- EHLERS-DANLOS SYNDROME in CAT
- EHLERS-DANLOS SYNDROME in CATTLE
- EHLERS-DANLOS SYNDROME in DOG
- EHLERS-DANLOS SYNDROME in MINK
- EHLERS-DANLOS SYNDROME in RABBIT
- EHLERS-DANLOS SYNDROME in SHEEP
- EHLERS-DANLOS SYNDROME, TYPE VII in CAT
- EHLERS-DANLOS SYNDROME, TYPE VII in CATTLE
- EHLERS-DANLOS SYNDROME, TYPE VII in SHEEP
- ENDOGENOUS VIRAL ELEMENT 1 in CHICKEN
- EPIDERMOL, YSIS BULLOSA, DYSTROPHIC in DOG
- EPIDERMOL, YSIS BULLOSA, DYSTROPHIC in SHEEP
- EPILEPSY in CHICKEN
- EPILEPSY in DOG
- EPITHELIOGENESIS IMPERFECTA in CAT

- EPITHELIOGENESIS IMPERFECTA in CATTLE
- EPITHELIOGENESIS IMPERFECTA in SHEEP
- EUMELANIN DILUTION BLUE in CHICKEN
- EUMELANIN EXTENSION in CHICKEN
- EUMELANIN RESTRICTOR in CHICKEN
- EXOSTOSIS, MULTIPLE in HORSE
- EXTENDED BLACK in CHICKEN
- FACTOR VII DEFICIENCY in DOG
- FACTOR X DEFICIENCY in DOG
- FACTOR XI DEFICIENCY in CATTLE
- FACTOR XI DEFICIENCY in DOG
- FACTOR XII DEFICIENCY in CAT
- FACTOR XII DEFICIENCY in DOG
- FADED SHAKER in CHICKEN
- FEATHER COLOUR in GUINEA-FOWL
- FEATHER COLOUR, ALBINISM in CHICKEN
- FEATHER COLOUR, ALBINISM, SEX-LINKED, IMPERFECT in CHICKEN
- FEATHER COLOUR, BLACK AT HATCH in QUAIL
- FEATHER COLOUR, CHOCOLATE in CHICKEN
- FEATHER COLOUR, DILUTE DOWN LETHAL in QUAIL
- FEATHER COLOUR, LIGHT DOWN LETHAL in QUAIL
- FEATHER COLOUR, ORANGE in QUAIL
- FEATHER COLOUR, RECESSIVE WHITE in CHICKEN
- FEATHER COLOUR, ROUX in QUAIL
- FEATHER COLOUR, SILVER in CHICKEN
- FEATHER COLOUR, SILVER in QUAIL
- FEATHER CURLING in CHICKEN
- FEATHERING, INHIBITED in TURKEY
- FEATHERING, Z-LINKED in CHICKEN
- FEATHERING, Z-LINKED in GUINEA-FOWL
- FEATHERING, Z-LINKED in TURKEY
- FECUNDITY, BOORoola in SHEEP
- FECUNDITY, CAMBRIDGE in SHEEP
- FECUNDITY, ICELAND in SHEEP
- FECUNDITY, INVERDALE in SHEEP
- FECUNDITY, JAVA in SHEEP
- FRAGILE SITE in SHEEP
- FRIZZLE in CHICKEN
- FUCOSIDOSIS, ALPHA in CAT
- FUCOSIDOSIS, ALPHA in DOG
- GANGLIOSIDOSIS in EMU
- GANGLIOSIDOSIS, GM1 in CAT
- GANGLIOSIDOSIS, GM1 in CATTLE
- GANGLIOSIDOSIS, GM1 in DOG
- GANGLIOSIDOSIS, GM1 in SHEEP
- GANGLIOSIDOSIS, GM2 in CAT
- GANGLIOSIDOSIS, GM2 in DOG
- GANGLIOSIDOSIS, GM2 in PIG
- GAUCHER DISEASE, TYPE I in DOG
- GAUCHER DISEASE, TYPE I in SHEEP

- GLAUCOMA in QUAIL
- GLAUCOMA in RABBIT
- GLUCOSE-6-PHOSPHATEDEHYDROGENASE DEFICIENCY in CAT
- GLUCOSE-6-PHOSPHATEDEHYDROGENASE DEFICIENCY in HORSE
- GLUTEN-SENSITIVE ENTEROPATHY in DOG
- GLYCOGEN STORAGE DISEASE I in DOG
- GLYCOGEN STORAGE DISEASE II in CAT
- GLYCOGEN STORAGE DISEASE II in CATTLE
- GLYCOGEN STORAGE DISEASE II in DOG
- GLYCOGEN STORAGE DISEASE II in QUAIL
- GLYCOGEN STORAGE DISEASE II in SHEEP
- GLYCOGEN STORAGE DISEASE IV in CAT
- GLYCOGEN STORAGE DISEASE V in CATTLE
- GLYCOGEN STORAGE DISEASE V in SHEEP
- GLYCOGEN STORAGE DISEASE VII in DOG
- GOITRE FAMILIAL in CATTLE
- GOITRE FAMILIAL in GOAT
- GOITRE FAMILIAL in SHEEP
- GONIODYPLASIA, MESODERMAL in DOG
- HAEMOPHILIA A in CAT
- HAEMOPHILIA A in DOG
- HAEMOPHILIA A in HORSE
- HAEMOPHILIA A in PIG
- HAEMOPHILIA A in SHEEP
- HAEMOPHILIA B in CAT
- HAEMOPHILIA B in DOG
- HAIR LENGTH in CAT
- HAIR LENGTH in HAMSTER
- HALO HAIR 1 in SHEEP
- HENNY FEATHERING in CHICKEN
- HEREDITARY MULTIPLEMALFORMATION in QUAIL
- HIND LIMB PARALYSIS in HAMSTER
- HIND LIMB PARALYSIS in PIG
- HORNS in CATTLE
- HORNS in GOAT
- HORNS in SHEEP
- HYPERBILIRUBINAEMIA I in HORSE
- HYPERBILIRUBINAEMIA I in SHEEP
- HYPERBILIRUBINAEMIA II in SHEEP
- HYPERBILIRUBINAEMIA, UNCLASSIFIED in CATTLE
- HYPERCHOLESTEROLAEMIA in PIG
- HYPERCHOLESTEROLAEMIA in RABBIT
- HYPERKERATOSIS, PALMOPLANTAR in DOG
- HIPERLIPOPROTEINAEMIA in CAT
- HYPOPROTHROMBINAEMIA in DOG
- HYPOTRICHOSIS in CATTLE
- HYPOTRICHOSIS in DOG
- HYPOTRICHOSIS WITH INCISOR ANODONITIA in CATTLE
- HYPOTRICHOSIS DOMINANT in PIG
- HYPOTRICHOSIS RECESSIVE in PIG

- HYPOTROPHIC AXONOPATHY in QUAIL
- INTESTINAL COBALAMIN MALABSORPTION in DOG
- KARTAGENER SYNDROME in DOG
- KRABBE DISEASE in CAT
- KRABBE DISEASE in DOG
- KRABBE DISEASE in MONKEY, RHESUS
- KRABBE DISEASE in SHEEP
- LADYKILLER in CHICKEN
- LEGG-CALVE-PERTHES DISEASE in DOG
- LEGLESS in PIG
- LETHAL TRAIT A 46 in CATTLE
- LEUKOCYTE ADHESION DEFICIENCY in CATTLE
- LEUKOCYTE ADHESION DEFICIENCY in DOG
- LIMBLESS in CHICKEN
- LUSTROUS WOOL in SHEEP
- LYMPHOSARCOMA in PIG
- LYMPHOSARCOMA in RABBIT
- MACROSTOMUS in RABBIT
- MALIGNANT HYPERTHERMIA in PIG
- MANNOSIDOSIS, ALPHA in CAT
- MANNOSIDOSIS, ALPHA in CATTLE
- MANNOSIDOSIS, ALPHA in SHEEP
- MANNOSIDOSIS, BETA in CATTLE
- MANNOSIDOSIS, BETA in GOAT
- MAPLE SYRUP URINE DISEASE in CATTLE
- MEAT QUALITY in PIG
- MEGACOLON in HORSE
- MEMBRANOPROLIFERATIVE GLOMERULONEPHRITIS TYPE II in PIG
- MENKESA SYNDROME in DOG
- MENKESA SYNDROME in SHEEP
- MICROMELIC SYNDROME in CHICKEN
- MICROMELIC SYNDROME in QUAIL
- MICROPTHALMIA in HAMSTER
- MUCOLIPIDOSIS II in CAT
- NUCOPOLYSACCHARIDOSIS I in CAT
- NUCOPOLYSACCHARIDOSIS I in CATTLE
- NUCOPOLYSACCHARIDOSIS I in DOG
- NUCOPOLYSACCHARIDOSIS II in DOG
- NUCOPOLYSACCHARIDOSIS IIIA in DOG
- NUCOPOLYSACCHARIDOSIS IIID in GOAT
- NUCOPOLYSACCHARIDOSIS VI in CAT
- NUCOPOLYSACCHARIDOSIS VI in DOG
- NUCOPOLYSACCHARIDOSIS VII in CAT
- NUCOPOLYSACCHARIDOSIS VII in DOG
- MUSCULAR DYSTROPHY in CHICKEN ( ALGUMA COISA ) DOENÇA
- MUSCULAR DYSTROPHY in DOG
- MUSCULAR DYSTROPHY in CHICKEN ( ALGUMA COISA ) DOENÇA
- MUSCULAR DYSTROPHY, DUCHENNE AND BECKER TYPES in CAT
- MUSCULAR DYSTROPHY, DUCHENNE AND BECKER TYPES in DOG
- MUSCULAR HYPERTROPHY in CATTLE



- MUSCULAR HYPERTROPHY in SHEEP
- MYASTHENIA GRAVIS in DOG
- MYOCLONUS in CATTLE
- MYOCLONUS in HORSE
- MYOCLONUS EPILEPSY OF LAFORA in DOG
- MYOTONIA in DOG
- MYOTONIA in SHEEP
- MYOTONIC DYSTROPHY in QUAIL
- NAKED NECK in CHICKEN
- NANOMELIA in CHICKEN
- NARCOLEPSY in DOG
- NEONATAL DIARRHOEA in PIG
- NEPHRITIS, AUTOSOMAL in DOG
- NEPHRITIS, X-LINKED in DOG
- NEPHRITIS, X-LINKED DOMINANT in DOG
- NEUTROPENIA, CYCLIC in DOG
- NIEMANN-PICK DISEASE in DOG
- NIEMANN-PICK DISEASE, TYPE C in CAT
- NUCLEOSIDE TRANSPORT DEFECT in PIG
- NUCLEOSIDE TRANSPORT DEFECT in SHEEP
- OPEN EYELIDS AT BIRTH in SHREW
- OPTIC CHIASM, ABSENCE OF in DOG
- OSTEOCHONDRODYSPLASIA in CAT
- OSTEODYSTROPHY in CAT
- OXALOSIS in CAT
- PAROXYSM in CHICKEN
- PEA COMB in CHICKEN
- PEMPHIGUS in DOG
- PERIODIC PARALYSIS II in HORSE
- PEROMELIA in GOAT
- PEROSIS in CHICKEN
- PERSISTENT MULLERIAN DUCT SYNDROME in DOG
- PHOTORECEPTOR DYSPLASIA in DOG
- PLUMAGE PATTERN in CHICKEN
- POLYCYSTIC KIDNEY DISEASE in CAT
- POLYCYSTIC KIDNEY DISEASE in DOG
- POLYCYSTIC KIDNEY DISEASE in RAT
- POLYDACTYLY in CHICKEN
- POP-EYE in CHICKEN
- PORPHYRIA, CONGENITAL ERYTHROPOIETIC in CATTLE
- PORPHYRIA, CONGENITAL ERYTHROPOIETIC in PIG
- PORPHYRIA, UNCLASSIFIED in CAT
- PORPHYRIA, UNCLASSIFIED in PIG
- POTASSIUM TRANSPORT in SHEEP
- PRENATAL in CHICKEN
- PROGRESSIVE DEGENERATIVE ENDOENCEPHALOPATHY in CATTLE
- PROGRESSIVE MYOPATHY in PIG
- PROGRESSIVE RETINAL ATROPHY in DOG
- PROGRESSIVE RETINAL ATROPHY, X-LINKED in DOG
- PROTEMINE-2 DEFICIENCY in CATTLE

- PROTEMINE-2 DEFICIENCY in PIG
- PROTEIN C DEFICIENCY in HORSE
- PROTOPORPHYRIA in CATTLE
- PROTOPORPHYRIA in CHICKEN
- PSEUDO-VITAMIN D DEFICIENCY RICKETS in PIG
- PYRUVATE KINASE DEFICIENCY OF ERYTHROCYTE in DOG
- REDUCED CASEIN CONCENTRATION in GOAT
- REDUCED GLUTATHIONE DEFICIENCY DUE TO AMINO-ACID TRANSPORT DEFECT in SHEEP
- REDUCED GLUTATHIONE DEFICIENCY DUE TO GCS DEFICIENCY in SHEEP
- REDUCED GLUTATHIONE DEFICIENCY, UNCLASSIFIED in SHEEP
- RENAL CYSTADENOCARCINOMA AND NODULAR DERMATOFIBROSIS in DOG
- RENAL CYSTS in PIG
- RENAL DYSPLASIA in CATTLE
- RESISTENCE TO AVIAN SARCOMA AND LEUKOSIS VIRUSES, SUBGROUP A in CHICKEN
- RESISTENCE TO AVIAN SARCOMA AND LEUKOSIS VIRUSES, SUBGROUP B in CHICKEN
- RESISTENCE TO OEDEMA DISEASE in PIG
- RESPONSE TO NORADRENALINE in SHEEP
- RESTRICTED OVULATOR in CHICKEN
- RETINAL DEGENERATION in ZEBRAFISH
- RETINAL DEGENERATION I in CHICKEN
- RETINAL DEGENERATION II in CAT
- RETINAL DEGENERATION, EARLY in DOG
- RETINAL DYSTROPHY, EARLY-ONSET in CAT
- RETINAL PIGMENT EPITHELIAL DYSTROPHY in DOG
- BIBOFLAVINURIA in CHICKEN
- RIBOSOMAL DNA DEFICIENCY in CHICKEN
- ROD-CONE DEGENERATION, PROGRESSIVE in DOG
- ROD-CONE DYSPLASIA-1 in DOG
- ROD-CONE DYSPLASIA-2 in DOG
- ROD-CONE DYSPLASIA-3 in DOG
- ROSE COMB in CHICKEN
- SCALELESS in CHICKEN
- SEVERE COMBINED IMMUNODEFICIENCY DISEASE, AUTOSOMAL in HORSE
- SEVERE COMBINED IMMUNODEFICIENCY DISEASE, X-LINKED in DOG
- SEX REVERSAL: XX MALE in CARP
- SEX REVERSAL: XX MALE in DOG
- SEX REVERSAL: XX MALE in LLAMA
- SEX REVERSAL: XX MALE in PIG
- SEX REVERSAL: XX MALE in CATTLE
- SHORT BEAK in QUAIL
- SPASTIC LETHAL in CATTLE
- SPERM DEGENERATION in CHICKEN
- SPHEROCYTOSIS in CATTLE
- SPHEROCYTOSIS in MOUSE, DEER
- SPINAL DYSMYELINATION in CATTLE
- SPINAL MUSCULAR ATROPHY in CATTLE
- SPINAL MUSCULAR ATROPHY in DOG

- SPONGIFORM ENCEPHALOPATHY in SHEEP
- STUMPY LIMB in QUAIL
- SYNDACTYLY in CATTLE
- TAILLESSNESS in CAT
- TALPID-2 in CHICKEN
- TALPID-3 in CHICKEN
- TESTICULAR FEMINIZATION in CAT
- TESTICULAR FEMINIZATION in CATTLE
- TESTICULAR FEMINIZATION in DOG
- TESTICULAR FEMINIZATION in HORSE
- TESTICULAR HYPOPLASIA in CATTLE
- TESTICULAR HYPOPLASIA in GOAT
- TETANIC TORTICOLLAR SPASMS in TURKEY
- THROAT TUFT in QUAIL
- THROMBOXANE RESPONSIVENESS OF PLATELETS in DOG
- TIBIAL HEMIMELIA in CATTLE
- TREMOR, HIGH-FREQUENCY in PIG
- TREMOR, X-LINKED in DOG
- TREMOR, X-LINKED in PIG
- TREMOR, X-LINKED in RABBIT
- TYROSINE TRANSAMINASE DEFICIENCY in MINK
- VERTICAL FIBRE HIDE DEFECT in CATTLE
- VIBRATOR in TURKEY
- VITAMIN-K-DEPENDENT BLOOD COAGULATION FACTORS DEFICIENCY in CAT
- VITAMIN-K-DEPENDENT BLOOD COAGULATION FACTORS DEFICIENCY in SHEEP
- VON WILLEBRAND DISEASE in DOG
- VON WILLEBRAND DISEASE in HORSE
- VON WILLEBRAND DISEASE in PIG
- VON WILLEBRAND DISEASE in RABBIT
- VON WILLEBRAND DISEASE I in DOG
- VON WILLEBRAND DISEASE III in DOG
- WAARDENBURG SYNDROME in HAMSTER
- WHITE SKIN in CHICKEN
- WILSON DISEASE in DOG
- WINGLESS in CHICKEN
- WOOLLY HAIR in PIG
- XANTHINURIA in DOG
- YELLOW FAT in SHEEP

### **6.3.4 – Genoma e características produtivas quantitativas dos animais**

#### **6.3.4.1 – Genoma e características reprodutivas melhoradas**

Quanto aos suínos estão patenteados uma série de métodos para identificar genes e marcadores genéticos de características reprodutivas melhoradas nos animais, como é o caso do “tamanho das ninhadas” e do peso dos animais à desmama.

A genética quantitativa sugere que características complexas dos animais, como é o caso do tamanho das ninhadas, são controladas por um grande número de genes cada um deles tendo um pequeno efeito sobre as características em causa. Contudo alguns genes, genes principais (major) tem efeitos mais amplos neste contexto, sendo hoje possível identificar os genes que afectam características quantitativas dos animais (*quantitative trait loci*, QTL).

Assim em 1994 a patente 5,374.526, (42) continha métodos para identificar tais marcadores e métodos para rastrear suínos para determinar quais os que produzem maior número de animais à nascença. Estes marcadores eram baseados na presença ou ausência de certos polimorfismos no gene para o receptor de estrogénio nos suínos, preferivelmente o polimorfismo e o comprimento do polimorfismo de um fragmento de restrição (RFLP). Estão associados com um maior tamanho da ninhada, um fragmento de 3,7 kilobases e um fragmento de 4,3 kilobases, obtidos pela digestão do DNA genómico de suíno, com a endonuclease de restrição Pvu II e detecção dos fragmentos com uma sonda marcada que contém o gene receptor para o estrogénio humano.

A reprodução nos mamíferos corresponde a uma série de acontecimentos que ocorrem entre o cérebro e os órgãos reprodutivos, sendo as hormonas esteróides como o estrogénio fundamentais. As hormonas esteróides interagem com células e tecidos. Nos suínos os estrogénios que são sobretudo produzidos nos ovários, actuam sobre o útero, cérebro e hipófise, modelando o aparecimento da puberdade, os comportamentos reprodutivos, a libertação cíclica das gonadotrofinas e o comportamento alimentar. Os estrogénios ligam-se aos seus receptores específicos que se encontram no núcleo das células sensíveis e que respondem aos estrogénios.

O gene responsável pela codificação do receptor dos estrogénios nos humanos foi isolado inclusivamente noutras espécies animais embora não para os suínos. Contudo a patente a que fizemos referência anteriormente, baseia-se na descoberta de polimorfismos no gene para o receptor de estrogénios nos suínos, os quais estão correlacionados com um tamanho maior das ninhadas.

Nesta patente é descrita a tipagem genética dos suínos nos seus genes receptores de estrogénios e a determinação das inter-relações dos RFLP específicos com o tamanho aumentado da ninhada, permitindo a identificação dos machos e fêmeas dos quais se pode esperar que produzam ninhadas maiores do que o que sucede com a média dessa raça.

Para isso uma amostra de DNA genómico é obtida de uma porca e analisada para determinar a presença ou ausência de um polimorfismo no gene do receptor do estrogénio que esteja correlacionado com um maior tamanho da ninhada. Preferivelmente o polimorfismo é uma RFLP (*restriction fragment length polymorphism*).

Como se compreende a presença ou ausência deste RFLP específico é feito através das etapas convencionais: o DNA genómico é digerido com enzimas de restrição que cindem o gene do receptor de estrogénio dos suínos pelo menos num local.

Os fragmentos obtidos pela digestão são separados por exemplo por gel electroforese, e depois os fragmentos detectados com uma sonda capaz de hibridar com eles, o que gera um perfil de RFLP. Este perfil é depois comparado com perfis conhecidos deste gene que estão correlacionados com ninhadas aumentadas. Utiliza-se também para comparação um segundo perfil de controlo feito com a mesma enzima de restrição e a mesma sonda.

Outra patente do mesmo inventor, Max F. Rothschild, a 5.939.264 (43) de 17 de Agosto de 1999 foi patenteada para a mesma finalidade da proposta anterior: genes e marcadores genéticos de

características reprodutivas melhoradas tais como maior número de animais nascidos por ninhada e maior peso á desmama em suínos. Estes marcadores são baseados na presença ou na ausência de certos polimorfismos em genes reprodutivos dos suínos, tais como a proteína 4 de ligação do retinol, o receptor gama do ácido retinoico, o receptor Ia da melatonina e a molécula 1 de adesão vascular celular.

Um gene reprodutivo identificado como uma proteína tendo um *loci* polimórfico é a molécula 1 de adesão da célula vascular (VCAM1) que era induzida no endotélio dos vasos sanguíneos nos locais de inflamação, constituindo um marcador valioso para aquilatar da resistência às doenças e da resposta imunológica inflamatória.

Contudo a VCAM1 é necessária também para a conexão vascular entre o embrião e as membranas não embrionárias que mais tarde formam a placenta pois sem VCAM1 não se formará nem a estrutura pré-umbilical chamada “allantois”(ou seja a membrana fetal vascular que na placenta dos mamíferos está intimamente associada com o córion na formação da placenta) nem quaisquer vasos placentários, tal como é essencial para a formação do miocárdio, sendo um elemento crítico que controla a sobrevivência do embrião e a velocidade do crescimento fetal através do fluxo sanguíneo que atravessa a placenta.

Nesta patente descreve-se a identificação de polimorfismos dentro do gene reprodutivo VCAM1 que está associado com um maior tamanho das ninhadas obtidas.

Outros genes reprodutivos polimórficos identificados nesta patente incluem o receptor gama do ácido retinoico (RARG) e a proteína 4 de ligação ao retinol (RBP 4).

Os receptores ácido retinóicos (RARs) são factores de transcrição induzidos por ligandos, tendo diversos subtipos consoante as espécies, e que se exprimem diferentemente durante o alongamento do blastocisto, altura em que é elevada a taxa de mortalidade dos embriões.

O RBP4 é também expresso durante o alongamento do blastocito e transporta e regula a quantidade de retinol recebida pelos fetos.

Os ácidos retinóicos têm potentes efeitos morfogénicos sobre o desenvolvimento de embriões podendo aumentar o tamanho das ninhadas de porcos.

O RBP4 é expresso durante o período crítico do alongamento dos embriões. O RBP4 pode regular a expressão dos genes durante o alongamento e é um candidato a ser um bom indicador do tamanho das ninhadas.

No quadro seguinte referem-se uma série de exemplos de genes candidatos a indicadores de características vantajosas nos suínos.

- Reprodução	{	ESR - Receptor do estrogénio
		PRLR - Receptor da prolactina
		RBP4 - Proteína 4 de ligação do retinol
- “Performance”	{	MC4R
		Miostatina

Foram identificados segundo esta invenção da patente polimorfismos nos RARG e RBP4.

Outro gene reprodutivo com formas alélicas alternativas identificado por esta invenção engloba o gene Ia para o receptor da melatonina (MTNR1A).

A melatonina é uma hormona segregada pela glândula pineal que ajuda a regular os ritmos circadianos e as variações na reprodução que ocorre na reprodução sazonal dos mamíferos. A melatonina liga-se com alta afinidade para actuar, a receptores acoplados a proteínas G.

Nesta patente descrevem-se pois meios para a detecção de diferenças genéticas quanto á eficiência reprodutiva dos animais, identificando marcadores genéticos em diversos genes, indicativos de fenótipos hereditários associados com características reprodutivas melhoradas, sobretudo tamanho das ninhadas. Com efeito a média desta entre várias raças ocidentais de suínos é de 4-16 leitões por ninhada, sendo a idade média da puberdade aos 3-7 meses de idade.

Há raças chinesas que atingem a puberdade mais cedo e com maior número de animais por ninhada (mais cerca de quatro leitões por ninhada).

Nos machos as melhores características reprodutivas englobam o tamanho testicular, volume, concentração e qualidade do esperma, líbido e cópula. Nas fêmeas englobam o tamanho da ninhada, o número de nado-vivos, peso da ninhada á nascença, desmama, idade da puberdade, reaparecimento do estro, intervalo entre as ninhadas, velocidade de ovulação, capacidade uterina e sobrevivência de embriões por porca.

Através do cômputo destes objectivos são identificáveis os animais machos e fêmeas que possuem genótipos mais vantajosos.

Como é de norma o método envolve as etapas clássicas 1) obtenção da amostra de DNA genómico do animal 2) análise deste DNA genómico para determinar qual o alelo ou alelos presentes.

Portanto a amostra do material genético obtido é analisada a fim de determinar a presença ou ausência de um polimorfismo num gene que está correlacionado com a característica reprodutiva desejável.

Numa perspectiva o polimorfismo corresponde a um RFLP (*restriction fragment length polymorphism*) e a análise implica identificar o gene reprodutivo (neste caso) do material genético isolado, para o que se expõe o gene a uma enzima de restrição que produz fragmentos de vários tamanhos os quais são depois separados por exemplo por electroforese ou HPLC, comparando-se depois este “perfil” obtido dos fragmentos de restrição, com testemunhas adequadas (contendo ou não o marcador desejado).

Algumas vezes o gene ou o fragmento que contem o polimorfismo é isolado com a utilização de “primers” e DNA polimerase para amplificar esta região, seguindo-se a sua separação, sequenciação, ou a sua digestão com uma enzima de restrição, e os fragmentos obtidos separados de novo. A visualização dos fragmentos separados pode ser feita por simples coloração dos fragmentos ou por marcação dos “primers” ou dos nucleótidos trifosfatos utilizados na amplificação.

Os *kits* para realizar estas análises estão disponíveis na respectiva patente.

Numa outra patente, a 6.291.174 de Setembro de 2001, investigadores chineses (Li, Ning et al) (44) ainda no que se refere a marcadores de DNA implicados no tamanho das ninhadas dos suínos, referem a determinação de qual do ou dos alelos da subunidade beta da FSH se encontra presente no genoma dos suínos, e os *kits* para a execução destas metodologias.

A hormona foliculo estimulina dos mamíferos (FSH) é uma glicoproteína composta de duas subunidades, uma alfa (que é comum a outras hormonas glicoproteicas como a LH e TSH) e uma subunidade beta, única. A FSH é segregada pela hipófise anterior por acção da GnRH e através da circulação sanguínea atinge os tecidos alvo nas gónadas, onde interaccua com os seus receptores promovendo a maturação e diferenciação dos folículos do ovário. A FSH e LH desempenham um papel importante no desenvolvimento do oócito antes da fertilização.

Em certas raças de porcos há uma mutação no gene da subunidade beta da FSH o que origina a inserção de um retrotransposição, que contem um promotor completo para a RNA polimerase II assim como outras possíveis regiões de transcrição, tendo sido verificado pelos investigadores chineses desta patente que o tamanho das ninhadas dos suínos que possuem esta mutação, são significativamente diferentes daqueles suínos que não têm esta mutação.

A determinação dos alelos da subunidade beta da FSH é feita através de marcadores particulares do DNA ligados directa ou indirectamente ao gene da subunidade beta FSH.

A mutação antes referida consiste como referimos na inserção de um retroposição e a presença desta está associada com um mais pequeno tamanho da ninhada. Este retrotransposição está situado no limite entre o intrão I e o exão II do gene da subunidade beta da FSH, no 809bp (com o sítio de início da transcrição a ser o O) e tem 292bp de comprimento com uma sequência perfeitamente conhecida, tal como as suas sequências flanqueadoras em 5' e em 3'.

As etapas para esta determinação feitas com *kit* adequado consistem no seguinte:

- a) Obtenção de uma amostra do DNA genómico do suíno
- b) Sua hibridação com um ou mais "primers" adequados
- c) Execução de um ou mais ciclos de PCR utilizando o ácido nucleico hibridado de b)
- d) Análise do comprimento do produto obtido pelo PCR em c)

Também a cor da pele dos suínos é importante na indústria suinícola havendo métodos patenteados (6.183.955 de Fevereiro de 2001) (45) para determinar esta por genotipagem, baseado no gene KIT dos suínos e analisando a diferença entre os alelos I, I. Sup. P., e i.

Em Julho de 1997, a patente USA 5.646.040 (46) apresentava métodos de diagnóstico e avaliação de desordens do peso corpóreo dos mamíferos, incluindo obesidade, caquexia, anorexia e predisposição para estas situações e terapêuticas possíveis. Estas moléculas de ácido nucleico representam os genes correspondentes aos genes *tub* dos mamíferos, um gene que está envolvido na regulação do peso corporal.

No mesmo sentido foi desenvolvida a patente USA 6.309.853 de Outubro de 2001 (47).

A patente US 6.383.751 de Maio de 2002 (48) da autoria de Barendse William John (Au) relata métodos e sondas de ácidos nucleicos para avaliar características do metabolismo lipídico nos animais, e sobretudo métodos para predizer os níveis de gorduras na carne, leite ou outros depósitos de gordura nos animais.

O método consiste na testagem dos animais para averiguar a presença de ausência de um ou mais marcadores seleccionados do grupo que compreende:

- a) um alelo da região 5' não translida da tiroglobulina.
- b) um alelo do polimorfismo do DNA do CSSM34, associado com o gene codificante do receptor gama do ácido retinóico (RARG).

c) um alelo do polimorfismo do DNA do ETH10, associado com o 11-cis,9-cis retinol desidrogenase (RDH5).

Esta metodologia é sobretudo aplicável para prever a disposição muscular da gordura, responsável pelo marmoreado das carnes e para prever o conteúdo butírico do leite, podendo servir para seleccionar animais sobretudo bovinos produtores de altas ou baixas concentrações de gordura no leite.

#### 6.3.4.2 – Genoma e produção de carne

A produção de carne nos animais está estreitamente ligada com a formação de tecido muscular embrionário estando esta produção concentrada em tecidos animais definidos como é o caso do tecido muscular.

Os músculos são tecidos complexos compostos de vários tipos de células; miócitos compostos de miofibras e células satélites, adipócitos intramusculares, fibroblastos, células endoteliais, neurocitos, etc.

É o número de miofibras presentes no momento do nascimento dos animais que determina a máxima capacidade de produção de carne magra nos suínos (49).

Também os bovinos de garupa dupla revelam um maior número de miofibras desenvolvidas antes do nascimento dos animais do que os outros bovinos sem garupa dupla (50) o que indica que a capacidade de produção de carne magra é determinada pelo desenvolvimento embrionário do número de miócitos.

A miogénese envolve uma via com as seguintes etapas:

- 1) Determinação das células progenitoras em linhagem miogénica.
- 2) Migração das “stem cells” miogénicas (mioblastos) para locais apropriados no embrião inicial.
- 3) Proliferação dos mioblastos e células não miogénicas dos tecidos musculares.
- 4) Diferenciação terminal dos miócitos (como por exemplo a fusão dos mioblastos) e expressão e organização dos produtos específicos dos genes apenas nas células musculares diferenciadas terminalmente.
- 5) Manutenção do estado diferencial terminal e modelação das miofibras em vários tipos de miofibras em função da idade e dos desempenhos fisiológicos.

#### Modelo de mecanismo da regulação genética da miogénese baseada na acção da família de genes das proteínas Myo D

Esta família Myo D nos vertebrados possui quatro membros a Myo D (ou Myf-3), a miogenina (Myf-4), a Myf-5 e a MRF4 (Myf-6, ou herculina).

As proteínas Myo D são expressas especificamente no tecido muscular onde actuam como factores de transcrição específicos. *In vitro* e *in vivo* elas são activas após formação de dímeros complexos com proteínas do gene E2A expresso ubiquamente. A expressão destes genes E2A depende de diversos motivos das sequências das suas regiões “enhancer” e das suas regiões promotoras. Uma classe destas sequências que são chamadas “E boxes” (caixas box) são locais onde se ligam proteínas e encontram-se na maioria dos “enhancer” específicos dos músculos.



Heterodímeros formados entre proteínas específicas básico hélice-ansa-hélice (bHLH) dos tecidos e os produtos dos genes E2A desempenham papéis muito importantes no destino das células específicas dos tecidos.

Estes complexos ligam-se a sequências reguladoras específicas da transcrição de genes específicos dos músculos, chamadas regiões favorecedoras (enhancer) dos promotores, activando a expressão de uma série de genes tais como os da actina, tropomiosina, e titina.

Uma vez activados, cada membro da família de genes Myo D pode autoregular positivamente a sua própria expressão e regular a expressão de outros genes Myo D continuando a via da diferenciação. Uma vez activada uma via, a miogénese continua até se estabelecer a diferenciação terminal.

Determinadas células (mioblastos) são capazes de migrar (etapa 2) da via miogénica antes referida e proliferar (etapa 3). A diferenciação terminal irreversível (etapa 4) é conduzida pela fusão dos mioblastos em miofibras multinucleadas, sendo esta fusão induzida pela activação do gene da miogenina (Myf-4) nos mioblastos (etapa 4).

O gene miogenina de suínos tal como o humano e o do rato tem três exões: o primeiro exão codifica o domínio bHLH, que tem sequências de ácidos aminados idêntica nas três espécies animais, o exão 2 codifica vinte e sete ácidos aminados do domínio de transactivação e o exão 3 o segmento C terminal conservado, é comum ás quatro Myf proteínas.

A sequência em ácidos aminados da miogenina suína é 97% e 96% idêntica á humana e á do rato. O primeiro intrão tem 758bp e é mais comprido nos suínos do que nos humanos e rato, e o segundo intrão 639bp é mais longo nos suínos do que no rato mas mais curto do que nos humanos.

O promotor miogenina e sequências 3' não codificantes estão também caracterizadas.

O promotor miogenina nos suínos tem uma sequência consenso para ligação dos factores de transcrição, em três sequências E-box (CANNTG) localizadas nas posições 508 a 513 (E1), 210 a 215 (E2) e 22 a 27 (E1).

Estão também identificados, um factor 2 enhancer específico do miocito (MEF-2) no local 456 a 464, um factor nuclear 1 (NF1) localizado no local 469 a 481 e uma sequência TATA box (494 a 498).

Todos estes factores estão envolvidos na regulação da expressão da miogenina. Uma sequência repetida CA de comprimento variável (19 a 22 repetições) está identificada entre as posições 603 e 560 do promotor de miogenina do rato.

A caracterização de um local polimórfico Msp1 do lado 3' do gene miogenina porcino está também descrita, bem como a sua detecção, tal como detecção e caracterização de dois locais adicionais polimórficos Msp1 no locus do gene da miogenina porcina.

Recentemente (51) foi assinalado o primeiro polimorfismo no gene Myf-4 de miogenina de suínos.

A miogenina é o único gene Myf expresso em todas as linhas celulares de músculos do esqueleto. Esta proteína tem um papel essencial na diferenciação muscular regulando a fusão dos mioblastos e a formação de fibras musculares funcionais.

RFLP (polimorfismos no comprimento de fragmentos de restrição) estão assinalados no gene *myo D* de rato, no gene *myf-5* de bovino, na miogenina de suíno e nos genes *myf-6*.

A frequência dos alelos difere consoante as diferentes raças de suínos.

A miogénese pode ser uma determinante importante do número de fibras musculares à nascença dos animais.

Como a miogenina regula o momento em que decorre a diferenciação e a fusão dos mioblastos existentes no fim da proliferação dos mioblastos, níveis de expressão tardios ou alterados podem influenciar o número de células musculares formadas.

Existem metodologias que tornam possível averiguar da variação genética dentro dos genes da miogenina e da variação da regulação da sua expressão e relacionar isto com características da produção animal (patente US6,143.880 de 2000) (52).

Estão pois disponíveis métodos para localizar, identificar ou marcar genes de alelos ou *loci* de características quantitativas (QTL) nas amostras (células de tecidos, como por exemplo: cabelo, peles ou sangue) a partir das quais se faz a amplificação específica de fragmentos genómicos daqueles genes, alelos ou *loci* de características quantitativas. Como o marcador utilizado para a selecção dos animais é frequentemente caracterizado pela variação genética que existe dentro de genes funcionais que influenciam directamente uma característica da produção, pode utilizar-se um método que identifique ou marque tais genes.

Também é possível utilizar métodos em que locais de restrição polimórficos dentro de genes funcionais (e assim diferentes alelos daqueles genes), são identificados ao realizar a amplificação específica de fragmentos genómicos daqueles genes.

Os métodos de amplificação são bem conhecidos e o mais divulgado é o PCR.

Após amplificação, um método adequado para a identificação dos alelos desejados é o tratamento com uma endonuclease de restrição, como por exemplo a *MspI* ou outras.

Por estes métodos é possível genotipar um grande número de animais para averiguar da variação genotípica que está associada com as características de crescimento dos animais e outras características produtivas.

Existem outros métodos para identificar polimorfismos nos alelos, quer ao nível nucleico (DNA / RNA) quer ao nível do produto proteína.

Ao nível das proteínas é possível utilizar imunoensaios enquanto ao nível nucleico se pode utilizar diversos métodos tais como diversos tipos de hibridação com “primers” ou ácidos nucleicos marcados.

Os fragmentos do gene miogenina podem ser isolados por amplificação PCR de DNA genómico.

Os fragmentos são subclonados e sequenciados (pelo método de Sanger dideoxy).

DNA genómico porcino isolado a partir de sangue tratado é digerido com *MspI*. Após precipitação, DNA digerido foi electroforesado em agarose e transferido para membrana de nylon positivamente carregadas por Southern blotting.

Após fixação do DNA as membranas são hibridadas em condições definidas na patente antes referida e os fragmentos de hibridação visualizados por autorradiografia.

### Regulação da miogénese e miostatina

O crescimento dos músculos, quer se trate de músculos do esqueleto quer se trate do coração é regulado por controlo positivo (por exemplo pela expressão dos genes MRF da família Myo D) e por controlo negativo (por exemplo pela expressão de genes da miostatina, MSTN).

A miostatina MSTN é um poderoso regulador negativo da massa de músculos do esqueleto, parecendo que a função da miostatina se encontra conservada entre as diversas espécies animais.

Investigação de diversa natureza tem demonstrado que o “switching off” do gene da miostatina em diversos mamíferos, desencadeia o aparecimento nesses animais de hiperplasia muscular ou seja aquilo que na gíria pecuária se designa de “garupa dupla”. Também em animais experimentais em que foram desencadeados enfartes de miocárdio, verificou-se que nas células musculares lesadas destes corações, existem altas concentrações de miostatina, ao contrário do que sucede em células do miocárdio normais.

A identificação de diversos inibidores da miostatina pode assim ter grande interesse quer na promoção do crescimento das massas musculares, quer no tratamento de situações patológicas que afectem estas.

A miostatina das células mamíferas consiste num complexo de um propéptido N-terminal ligado não covalentemente com fragmentos de um dímero C-terminal ligado por pontes dissulfureto.

Este dímero C-terminal da miostatina tem capacidade para interactuar com receptores de activina do tipo II, Act RIIB e numa extensão menor também com o Act RIIA.

A interacção da miostatina com o Act RIIB pode ser inibida pela folistatina que é uma proteína que se liga á activina, e em mais elevadas concentrações pelo propéptido da miostatina.

Parece que este propéptido, a folistatina e outras moléculas podem bloquear a via de transdução do sinal desencadeado pela miostatina e favorecer assim o crescimento muscular.

A proteína “folistatina” produzida em quantidades excessivas (por engenharia adequada) no rato provoca nestes o aparecimento de massa muscular muito mais desenvolvida, que nos ratos normais, como sucede se os ratos produzirem um excesso de mutantes de receptores para a activina II, o mesmo sucede com o propéptido da miostatina.

No entanto a administração directa das proteínas referidas em cima (folistatina, receptor para a activina II, propéptido da miostatina e outros bloqueadores da miostatina) quer para produção de animais mais corpulentos, quer para o tratamento de doenças que desgastem o tecido muscular, ou ainda a alteração do genoma desencadeando mutação ao nível da miostatina que vão neste sentido necessitam ser encaradas com cuidado.

Não surpreende portanto que tenham também surgido recentemente patentes que referimos a seguir relacionadas com a miostatina.

A US patent nº 6, 103, 466 intitulada “Double-Muscling in Mammals” de 15 de Agosto de 2000 e a Australian Patent nº 199884571 designada “Mutation in the myostatin gene cause double-muscling in mammals” de 16 de Janeiro de 2003, abordam a temática da miostatina como condicionante da produção de carne nos animais.

#### 6.3.4.3 – Genoma e produção de leite

##### Marcadores genéticos indicativos de produção leiteira acrescida ou melhorada em bovinos

Estão patenteados nos USA diversos tipos de ensaios para identificar marcadores genéticos de produção leiteira aumentada em bovinos.

Assim em 1991, uma patente utilizando a tecnologia do DNA recombinante, determinava em células bovinas por hibridação em fragmentos de restrição, quando o polimorfismo de genes estava associado com produção leiteira aumentada (Patente US 5.041.371 de 1991) (53).

Com efeito as técnicas para avaliar o “progeny testing” de cada touro levavam anos a realizar e eram muito dispendiosas e vários genes bovinos têm sido identificados como responsáveis pela expressão de proteínas que são importantes no controlo do crescimento mamário, lactogenese e ou lactação. Um destes genes, o da prolactina bovina tem cerca de 10kilobases de comprimento. Polimorfismos (que não afectam a composição em ácidos aminados) adjacentes ao gene da prolactina bovina têm sido assinalados.

Contudo foi patenteado na patente antes referida, um ensaio para assinalar a presença na sequência de um gene bovino de um marcador genético localizado dentro de 1,5kb do exão codificador na sequência da prolactina bovina. Este marcador é indicativo de uma característica hereditária na descendência de produção leiteira aumentada.

Neste ensaio é exposta a sequência do gene a uma enzima de restrição (por exemplo Ava II) de forma que se produzam fragmentos do gene de diversos comprimentos, sendo depois separados estes uns dos outros por electroforese, seguindo-se a hibridação com uma série de sondas de DNA radio-marcadas) que contêm uma porção da sequência do gene da prolactina bovina. Depois comparavam-se os resultados da hibridação com os resultados de ensaios com uma sequência de genes bovinos conhecidos como contendo um marcador e com uma sequência do gene bovino conhecida ou identificada como não possuindo o referido marcador.

São mesmo fornecidos *kits* para ensaiar a presença do marcador genético (na sequência do gene bovino) localizado dentro de 1,5kb no exão codificador da prolactina bovina. O *kit* contem uma sonda que possui uma porção da sequência do gene da prolactina bovina e também uma sequência do gene bovino que se sabe conter o respectivo marcador.

As etapas gerais deste ensaio são pois a extracção do DNA por exemplo do sémen, a sua digestão e separação dos fragmentos obtidos, a hibridação e finalmente a correlação com os dados sobre a produção leiteira do efectivo descendente.

Em 1994 outra patente (U. S. 5.374.523) (54) refere um ensaio para determinar a presença no material genético bovino de um marcador genético localizado no gene da somatotrofina bovina indicativo de uma produção leiteira aumentada transmissível á descendência. O marcador é um polimorfismo no gene codificador da somatotrofina no ácido aminado na posição 126, existindo pois duas formas de somatotrofina nos bovinos. Este ensaio realiza pois a análise do bovino para determinar o seu genótipo no que respeita ao gene da somatotrofina. O bovino que tem o marcador desejado pode depois ser

seleccionado para ser incluído no programa de selecção para o melhoramento da produção leiteira. O marcador desejado indicativo da produção aumentada de leite depende da raça do bovino. Assim bovinos da raça Holstein que são homocigotos para as formas de leucina na somatotrofina são desejáveis, e bovinos Jersey que são homocigotos para a forma valina da somatotrofina são também desejáveis.

Já tínhamos referido anteriormente que a patente U. S. Pat. Nº 5.041.371 ensaiava e identificava um marcador genético envolvendo a presença de um polimorfismo adjacente á sequência do gene da prolactina bovina.

Parece vantajoso que outros marcadores genéticos no material genético bovino possam ser identificados fornecendo indicações adicionais do potencial genético em ordem a incluir esse animal num programa de melhoramento leiteiro, É pois desejável uma pluralidade de marcadores indicativos das características desejadas para o efectivo leiteiro como por exemplo uma produção leiteira aumentada.

São conhecidas quatro variantes do gene da somatotrofina nos bovinos (55). Estas variantes provêm da combinação de dois possíveis ácidos aminados N terminais (alanina ou fenilalanina). Na posição 126 da somatotrofina podem existir leucina ou valina, o que resulta da mudança de um simples nucleótido no codão para o ácido aminado 126 no gene da somatotrofina.

É conhecido que formas recombinantes da proteína somatotrofina podem ser injectadas em vacas leiteiras aumentando a produção leiteira destas. Também é conhecido que vacas leiteiras Holstein injectadas com formas recombinantes de somatotrofina para aumentar a sua produção láctea, têm uma resposta mais eficiente se injectadas com a variante contendo a valina na posição 126 relativamente á variante contendo a leucina na posição 126.

O polimorfismo do gene da somatotrofina no ácido aminado na posição 126 fornece assim um marcador que pode estar correlacionado com a característica de uma maior produção láctea.

Nas populações bovinas existem pois duas formas de somatotrofina, uma exprimindo leucina na posição 126 e outra exprimindo valina nesta mesma posição.

Este ensaio compreende a identificação do gene da somatotrofina bovina, a partir de material genético isolado de bovinos, e depois a exposição do gene a enzima de restrição que produz fragmentos de restrição de variados tamanhos os quais após separação (por electroforese ou HPLC) formam um perfil dos fragmentos, comparando-se depois este “perfil de fragmentos” resultante, com um perfil de fragmentos de um gene de somatotrofina bovina conhecida por ter ou não ter este marcador.

Este ensaio é vantajoso utilizado quer com machos quer com fêmeas bovinas, os primeiros sendo testados quanto á sua capacidade progenitora de produção leiteira acrescida nas fêmeas geradas. É preferível seleccionar touros que exibam um genótipo que seja homocigoto para o marcador desejado nesta raça. As vitelas e as vacas são seleccionadas baseados no potencial dos seus “óvulos” para produzir descendência com produções leiteiras aumentadas, sendo também preferível seleccionar os genótipos homocigotos para o marcador desejado nesta raça. Os embriões também podem ser testados na mesma perspectiva.

Preferivelmente o gene da somatotrofina é isolado por amplificação do gene utilizando a tecnologia PCR com os “primers” oligonucleotídicos adequados. Certamente que o gene da somatotrofina pode ser identificado a partir do DNA genómico bovino utilizando técnicas de hibridação com sondas radiomarcadas, ou técnicas de reacção em cadeia com ligases.

É sabido que existem dois genes da somatotrofina no genoma diplóide bovino e que ambos os genes são transcritos e translidos. O mapa genómico e a sequência da somatotrofina bovina são

conhecidos desde 1982. “Primers” oligonucleótidos foram desenvolvidos capazes de amplificar e isolar um fragmento do gene da somatotrofina bovina de 428 pares de bases (bp) que contem o polimorfismo que serve como marcador genético.

Como se disse o polimorfismo que origina duas variantes da somatotrofina na posição do ácido aminado 126 é uma simples alteração de um nucleótido no codão para o ácido aminado nesta posição. Na variante leucina, o codão é CTG, e na valina GTG. Estas variantes originam fragmentos de restrição de diversos comprimentos quando sujeitos a endonucleases de restrição particulares, como a Alu I que corta a dupla hélice de DNA na sequência 5'-AGCT-3'. O codão da sequência do gene da somatotrofina da variante com leucina nas posições 125 e 126 é 5'-GAGCTG-3' e a sequência da variante com valina, nestas mesmas posições, é 5'-GAGGTG-3'. A mudança de C para G origina uma mudança no ácido aminado (leucina para valina) e uma perda do local de cisão pela Alu I).

Quando o gene da somatotrofina, variante leucina, é exposto à acção da enzima de restrição Alu I, formam-se fragmentos de restrição de 265,96,51 e 16 pares de bases de comprimento e na variante com valina formam-se fragmentos de restrição de 265, 147 e 16 pares de bases de comprimento. O polimorfismo está presente no fragmento de 147bp.

Vacas homozigóticas para o gene com a variante leucina originam assim fragmentos de DNA 265,96 e 51bp (o fragmento de 16bp não é habitualmente visível com técnicas de coloração com o brometo de etídio), vacas heterozigóticas exibem fragmentos de 265,147,96 e 51bp, e vacas homozigóticas para o gene com a variante valina, exibem fragmentos de DNA 265 e 147bp.

Os fragmentos resultantes do DNA são separados por gel electroforese e os perfis dos fragmentos de restrição são comparados com padrões que possuem ou não possuem o marcador em questão.

Uma outra patente US 5,614,364 de 1997 (56) refere um marcador genético para produção acrescida de leite nos bovinos baseado no gene PIT-1 bovino que está associado com uma produção láctea aumentada bem como o respectivo conteúdo em proteínas e gordura. Novas sequências nas regiões intrão do gene são assinaladas as quais podem ser utilizadas na amplificação pela PCR para rastrear a presença do marcador.

O PIT-1 é um factor de transcrição específico da pituitária responsável pelo desenvolvimento da pituitária e pela expressão hormonal nos mamíferos e é um membro da família POU de factores de transcrição que regulam o desenvolvimento dos mamíferos.

O PIT-1 ou factor de transcrição específico de pituitária tem outras designações alternativas tais como GHF1 (growth hormone factor 1) e POU1 F1.

A proteína PIT-1 funciona ligando-se e transactivando os promotores dos genes da GH e da PRL e do próprio gene PIT-1.

Portanto o gene PIT-1 regula a expressão do gene da GH e da PRL na hipófise.

O PIT-1 também medeia a estimulação dos genes da TRH (hormona libertadora da tirotrófina) e estimulação da PRL pelo cAMP e da beta TSH.

O gene PIT-1 e a proteína estão assinaladas nos ratos, ratazanas, bovinos, humanos e suínos.

Transcritos de PIT-1 mRNA têm sido encontrados nos cinco tipos celulares das pituitárias maduras (galactotrofos, gonadotrofos, corticotrofos, tirotrofos e somatotrofos) mas a proteína PIT-1 apenas tem sido detectada nos lactotrofos, somatotrofos e tirotrofos do lobo anterior da hipófise.

A PIT-1 contém dois domínios proteicos, chamados POU- específico e POU- homeo, sendo ambos necessários para se ligarem ao DNA com alta afinidade, nos genes que codificam a hormona do crescimento (GH) e prolactina (PRL), sendo também importante na regulação dos genes que codificam proteína e subunidade beta da hormona estimuladora da tiróide (TSHB) pela hormona libertadora tirotrofina (TRH) e AMP cíclico.

Mutações no rato no gene PIT-1 que se situa no cromossoma 16 do rato são responsáveis por deficiências pleiotróficas da hormona de crescimento, prolactina e hormona estimuladora da tiróide, enquanto a produção da hormona adrenocorticotrófica, hormona luteinizante (LH) e hormona estimuladora do folículo (FSH) são preservadas.

Mais uma vez nas metodologias utilizadas, para rastrear este marcador PIT-1 se utilizam metodologias gerais incluindo nelas, a extração do DNA, a sua digestão com enzimas de restrição a separação dos fragmentos resultantes, e a hibridação com sondas rádiomarcadas e finalmente a correlação dos polimorfismos identificados com as características da produção leiteira em ordem a identificar o marcador que ajude nesta selecção. Também a PCR para amplificar o fragmento do gene relevante para posterior análise pode também ser utilizada originando esta nova tecnologia maior rapidez na análise, pois necessita menos material genético. A fase de hibridação para detectar fragmentos específicos do gene pode ser evitada e é mais barata.

Têm sido referidos marcadores genéticos que indicam a presença de deficiente adesão dos leucócitos bovinos, bem como marcadores localizados dentro do gene da prolactina bovina e na Kappa caseína associados com produção leiteira aumentada, tal como tem sido utilizados marcadores para detectar variantes de DNA mitocondrial bovino que estão associados com as melhores performances das vacas leiteiras.

Tem sido tentados outros meios para melhorar as performances dos bovinos tal como o aumento o “boosting” da concentração de hormonas do crescimento através da introdução de hormonas adicionais *per os* ou injectados. É o caso da GH de bovino recentemente aprovada pela FDA.

Contudo estes outros meios, nada oferecem no que se refere á selecção dos animais geneticamente superiores, sendo preferível nesta perspectiva a utilização de marcadores genéticos como é o caso do gene PIT-1 bovino indicativo de uma produção leiteira aumentada e de uma melhor composição do leite obtido.

Nesta patente referente ao gene PIT-1 de bovinos averigua-se dos polimorfismos deste gene que estão associados com produção leiteira melhorada. Nesta patente o perfil de fragmentos de restrição produzidos pela enzima *HinfI* identifica pelo menos um polimorfismo que pode ser utilizado para assinalar a presença ou ausência de um marcador associado com características leiteiras desejáveis.

O genótipo não marcador consiste de pelo menos um polimorfismo dentro do exão 6 do gene bovino PIT-1 o qual cria um local adicional para hidrólise pela *HinfI*. O genótipo marcador, que não possui qualquer sítio adicional *HinfI* para hidrólise, está associado com produções aumentadas de leite em média de 13,2% e com um aumento médio de 17,2% do conteúdo proteico total.

A patente inclui o ensaio para a detecção do marcador assim como a caracterização da sequência de um dos polimorfismos e inclui novas sequências do intrão no gene PIT-1 que podem ser utilizadas para desenhar os primers para amplificação para ensaios.

Outra patente USA 5,292,639 de 1994 (57) associa o DNA mitocondrial bovino com características de importância económica, criando um método para geneticamente avaliar estes animais. Os polimorfismos no DNA mitocondrial são detectados por isolamento, fragmentação e sequenciação do DNA. Os perfis de restrição das sequências nucleotídicas do DNA mitocondrial de diferentes animais são correlacionados com as características expressas pelos animais.

As mitocôndrias transportam múltiplas cópias de um genoma circular que é replicado e expresso dentro do organelo, sendo o único elemento genético herdado citoplasmicamente nos mamíferos. O DNA mitocondrial nos animais codifica 13 polipéptidos componentes do sistema de síntese do ATP e codifica também todos os RNA de transferência utilizados na síntese proteica mitocondrial. A única região não codificante deste genoma mitocondrial é uma região de cerca de 900 pares de bases correspondente á “D-Loop”. Esta D-Loop está implicada no controlo da transcrição e replicação do DNA mitocondrial.

A avaliação genética dos animais habitualmente centra-se nos genes contidos no núcleo, contudo também as mitocôndrias contêm o seu próprio DNA (hereditariedade citoplasmica) tendo sido recentemente verificado que existe variação ou seja polimorfismos dentro deste DNA mitocondrial dos animais e que estes polimorfismos podem ser utilizados para avaliar características hereditárias. Os genes mitocôndriais são herdados da mãe pois os machos não transmitem genes mitocôndriais à sua descendência. Os efeitos da hereditariedade citoplasmica (DNA mitocondrial ou mtDNA) tem sido conotados com o número de dias que decorre entre o parto e a concepção seguinte, com os dias que decorrem entre o parto e o primeiro estro depois detectado, etc.

Estão assinalados métodos para isolar e preparar fragmentos de DNA mitocondrial e associa-los com a expressão fenotípica de características particulares dos animais.

Assim a presença de substituições de certos nucleótidos na região da ansa D (D Loop) do DNA mitocondrial está associada com um volume aumentado da produção de leite e com um mais elevado teor butiroso deste.

A variação molecular do mtDNA bovino tem sido assinalada através da análise dos RFLP e comparação das sequências nucleotídicas.

O DNA mitocondrial dos animais é isolado, fragmentado e sequenciado e os polimorfismos detectados, sendo depois estes correlacionados com a presença ou ausência das características desejadas. O isolamento do DNA mitocondrial pode ser feito por clonagem em *E. coli* ou por amplificação pela PCR. A fragmentação é depois feita com enzimas de restrição seleccionadas, e a sua resolução por gel electroforese.

### **6.3.5 – Genoma e triagem de sexos**

#### Escolha de sexo nos animais mamíferos (58)

A capacidade de separar os espermatozóides portadores do cromossoma X e Y permite aos criadores predeterminar o sexo dos animais quando uma amostra de sémen seleccionado é utilizada para inseminação artificial.



Uma nova técnica para este efeito foi patenteada em 1989 após os estudos efectuados por Lawrena Johson *et alii*. Este método baseia-se na separação do esperma consoante o seu conteúdo em DNA. Na técnica original que podia ser utilizada em coelhos, suínos e bovinos havia a limitação de que apenas 1,5 – 2 milhões de espermatozóides podia ser escolhido por dia e com uma certeza que variava entre 75-90%. Hoje a técnica foi modificada e é também utilizável com ovinos, animais de laboratório e outras espécies.

A técnica baseia-se na utilização de um corante fluorescente que adere “sticks” ao DNA. O corante liga-se ao DNA de espermatozóides consoante quanto DNA existe na cromossoma X e Y. Os espermatozóides X produtores de fêmeas contêm 2,8 a 7,5% mais de DNA do que os espermatozóides Y produtores de machos, consoante as espécies. As células espermatozóides coradas são analisadas e escolhidas utilizando um repartidor (sortex) contador de células citométrico do fluxo de células. Quando uma fonte laser ilumina o corante, cada espermatozóide emite uma luz proporcional ao seu conteúdo em DNA e é separado para diferentes tubos, consoante a quantidade de luz emitida.

Hoje 35 a 40 milhões de espermatozóides X e Y podem ser repartidos em oito horas de trabalho. A precisão da técnica depende da diferença de conteúdo em DNA dos espermatozóides X e Y o que varia consoante as espécies animais. Os suínos por exemplo têm mais 3,6% de DNA nos X do que nos Y, os bovinos mais 3,8%. Quanto maior a diferença maior a precisão.

#### Sondas de identificação dos sexos nas aves

Enquanto a genética e bioquímica da determinação do sexo nos mamíferos tem sido muito investigada, o mesmo não tem sucedido no que se refere às aves, apesar da enorme importância que nelas pode revestir sobretudo nos frangos e perus.

Marcadores genéticos específicos do sexo são de grande importância, pois tais marcadores poderão ser utilizados para identificar aves ainda imaturas e antes do desenvolvimento das suas diferenças morfológicas específicas (características sexuais externas).

Os cromossomas do sexo, diferem um do outro no tamanho e na composição genética. Assim algumas regiões de um cromossoma do sexo contêm genes que não têm alelo correspondente no outro cromossoma do sexo. A ave fêmea é sexo “heterogamético”, tendo cromossomas do sexo, Z e W. Nos mamíferos, o macho é sexo “heterogamético” tendo ambos os cromossomas X e Y enquanto a fêmea é “homogamética”.

Sondas marcadoras consistem de sequências polinucleotídicas que hibridam especificamente com uma região do cromossoma. Estas regiões cromossomais de hibridação são reveladas como sendo polimórficas entre indivíduos da mesma espécie, quando o DNA cromossomal é digerido por endonucleases da restrição e analisado por análise de hibridação.

Alelos diferentes distinguem-se um do outro na base dos perfis de bandas de hibridação produzidos após a separação pelos seus tamanhos.

Na patente U: S: nº 5.707.809 de janeiro de 1998 (59) são fornecidos marcadores genéticos adequados para a identificação de regiões do DNA que permitem o diagnóstico do sexo nas aves, dado que uma região de DNA pode ser única num único dos cromossomas do sexo, de forma que a quantidade ou a presença do marcador determinará o sexo da ave.

Tais marcadores genéticos permitem a identificação do sexo especificado cromossomalmente numa ave baseado na análise de uma preparação de DNA obtido a partir da ave.

Na patente e no *kit* sobre ela desenvolvido, sequências nucleotídicas aviárias são utilizadas como fontes de sondas e “primers”, sendo indicados os processos a utilizar para identificar o sexo das aves. Há sondas que hibridam com ambos os cromossomas ZeW, permitindo a diferenciação entre os dois cromossomas na base de diferentes comprimentos dos polimorfismos. Outras sondas podem hibridar exclusivamente com um dos dois cromossomas do sexo.

### 6.3.6 – Genoma, sua expressão e “*microarrays*” do DNA. Possibilidades futuras

#### Monitoragem do genoma sobre um chip

A tecnologia do DNA *microarray* (*genome chip* ou *biochip*) (60. 61) é uma tecnologia que vai redesenhar e reorientar a biologia molecular.

Com efeito milhares de genes e seus produtos como é o caso dos RNA e proteínas em cada ser vivo funcionam de uma forma organizada criando o mistério da vida, e nos métodos tradicionais de biologia molecular geralmente trabalha-se na base de um gene em cada experiência. Contudo há relativamente poucos anos surgiu uma nova tecnologia, a do DNA *microarray* que pretende fazer o rastreio da totalidade do genoma num simples chip de forma a dar uma melhor panorâmica das interações entre milhares de genes simultaneamente.

Portanto esta tecnologia do DNA *microarray* permite uma monitoragem da totalidade do genoma num simples chip.

O emparelhamento (por exemplo A-T e G-C no DNA, e A-U e G-C no RNA) ou hibridação é o principio de fundo em que se baseia o DNA *microarray*.

Este DNA *microarray* consiste num arranjo ordenado de amostras, em que é criado um ambiente em que são postas em jogo amostras conhecidas e desconhecidas de DNA baseado nas regras de emparelhamento das bases, e automatizando o processo de identificação das amostras desconhecidas.

Um ensaio de DNA *microarray* pode ser realizado utilizando microplacas de membranas standard de “blotting” e pode ser feito á mão ou utilizando meios roboticos para depositar as amostras. Há dois tipos de “arrays” os *macroarrays* quando a largura das manchas (*spots*) de depósito das amostras é de cerca de 300 microns ou mais de diâmetro, e as *microarrays* quando o tamanho das manchas (*spots*) de depósito é menor que 299 microns de diâmetro.

As imagens dos *microarrays* são facilmente visualizadas no gel existente e “blot scanners”. As manchas dos *microarrays* são muito pequenas e estes arrays contêm milhares de manchas necessitando ser lidos por meios roboticos especializados e exigindo ainda equipamento para obtenção de imagens que não se encontra disponível no comércio na forma de um sistema completo pronto a utilizar.

Os DNA chips ou DNA *microarrays* são fabricados por meios robóticos de alta velocidade sobre placas de vidro ou nylon, sobre os quais sondas convenientemente identificadas são utilizadas para determinar a ligação complementar que se pode estabelecer, permitindo assim estudos de identificação de genes ou dos produtos da sua expressão. Uma experiência com um único DNA *chip* pode dar informação sobre milhares de genes simultaneamente.

Há dois tipos principais de aplicação deste DNA chip, que são:

- 1) A identificação de genes (gene / gene com mutação).
- 2) Determinação da taxa de expressão (abundância dos genes).
- 3)

Há duas variantes desta biotecnologia do DNA *microarray* em termos das propriedades das sequências de DNA dispostas no *chip (arrayed)* com identificação conhecida.

- a) Formato I: sondas de cDNA (de 500-5.000bases de comprimento) são imobilizadas na superfície sólida do chip como por exemplo: vidro, utilizando “spotting” (depósito) robótico, e expostas a um conjunto de alvos (ou *targets* ou seja amostras de ácidos nucleicos livres cuja identidade / quantidade se quer conhecer) separadamente ou em mistura.
- b) Formato II: um conjunto (array) de sondas de oligonucleótidos (20 –80-mer oligos) ou péptidos ácido nucleico (PNA) é sintetizado quer *in situ* (sobre o chip) ou por síntese convencional, seguido pela sua imobilização sobre o chip. O conjunto (array) é exposto á amostra de DNA marcado hibridada, e determinadas a identidade / quantidade das sequências complementares.

Existem diversas firmas que comercializam produtos desta natureza fabricados fotolitograficamente, ou chips baseados em oligonucleotidos manufacturados utilizando síntese *in situ* ou tecnologia depositantes.

Diversas organizações universitárias e académicas têm posto de pé facilidades sobre estas tecnologias do DNA microarray em ordem a tornar esta tecnologia acessível ao seu corpo de investigadores.

A tecnologia do DNA *microarrays* são perfeitamente adequadas para comparar a expressão dos genes em diferentes populações celulares, e as principais etapas numa experiência comparativa de hibridação de cDNA são as seguintes:

- 1) escolha da população de células a estudar
- 2) extracção do mRNA e sua transcrição reversa
- 3) marcação fluorescente dos cDNA
- 4) hibridação com um DNA microarray
- 5) scanning do array hibridado
- 6) interpretação da imagem “scanned”

Os caminhos do futuro na área da bioquímica e da biologia molecular passam obrigatoriamente pelo aprofundamento de conhecimentos dos genomas, transcriptomas e proteomas dos seres vivos, e as suas interacções com todo o seu meio envolvente e ambiente, e as metodologias que acabamos de referir parecem extraordinariamente promissoras em todo este contexto.

#### 6.4 - Bibliografia

1. Craig Venter, J. et al (2001) The sequence of the human genome. Science 291. 1304-1351
2. Sonnhammer, E. L., S. B. Eddy, R. Durbin (1997) Proteins 28, 403
3. Bateman, A. et al (2000) Nucleic Acids Res. 28, 263
4. Referencia 116 vide em 1
5. Pointing, C. P. et al (1999) Nucleic Acids Res. 27, 229
6. Muller, H. J. and Kern, H. Z. (1967) Naturforsch B.22, 1330
7. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome Nature 409, 860-921
8. Pruitt, K. D. & Magdon, D. R. ( 2001 ) Ref Seq and locus link NCB gene centered resources Nucleic acids Res. 29, 137-140
9. Idem 7 pag. 896
10. Sun, Q, Mayeda, A. Hampson, R. K. et al (1993) General splicing factor SF/2 ASF promotes alternative splicing by binding to an exonic splicing enhancer Genes Dev 7, 2598-2608
11. Tanaka, K, Watakabe, A. S. Shimura, Y. ( 1994 ) Polypurine sequences within a downstream exon function as a splicing enhancer, Mol. Cell Biol. 14, 1347-1354
12. Mironov, A. A., Fickert, J. W. & Gelfond, M. S. (1999) Frequent alternative splicing of human genes Genome Res 9, 1288-1293
13. Brett, D. et al (2000) EST comparison indicator 38% of human mRNA contain possible alternative splice forms FEBS lett 474, 83-83

14. Idem 1 pag. 1335-1336
15. Hurskainen, T. L. et al (1999) J. Biol. Chem 274, 25555
16. Black, R. A. & White, J. M. (1998) Curr. Opin. Cell Biol 10, 654
17. Idem 1 pag. 1343-1345
18. Idem 1 pag. 1338-1342
19. Rathin Das (2001) Proteomics The new vista in drug discovery International Biotechnology Laboratory – August, 6
20. Mann, M., Hendrikson, R. C. & Pandya A. ( 2001 ) Analysis of proteins and proteomes by mass spectrometry Annu. Rev Biochem 70, 437-473
21. Idem referencia 19
22. Rothschild, Max F. (2000) New Technologies in animal breeding Lecture 2 – curso VILA Real – 16e 17 de Outubro de 2000
23. Idem referencia 22, Lecture 4
24. Idem referencia 22, Lecture 4
25. Idem referencia 22, Lecture 4
26. Idem referencia 22
27. Wagter – Lesperence, et al September 11, 2001 – Method of Identifying High Immune Response Animals – United Sates Patent 6,287,564
28. Mullis, et al – July 18, 1987 – Process for amplifying, detecting, and por cloning nucleic acid sequences – USA Patent 4,683,195
29. Francis, Grosclande, et al – 1999 – 12 – 03 – New nucleotide sequences, useful for genetic identification of cattle – France 2779153
30. Dr. Van Haeringen Laboratorium b. V. (Netherlands) <http://easynet.webmind.nl/vhl/>
31. Medigenomix(Germany) <http://www.medigenomix.de/services/schotyfing/genotyfing.htm>  
<http://www.medigenomix.de/services/food/foodscheme.htm>
32. Geneseek ( USA ) <http://www.geneseek.com/productsserv/diagnostic>
33. Idem 31
34. <http://www.medigenomix.de/services/vetmed/vetmed/htm>
35. <http://www.medigenomix.de/services/vetmed/vetmed/htm>
36. <http://www.medigenomix.de/services/vetmed/vetmed/htm>
37. Lewin, et al December 10, 1996 Methods for testing bovine for resistance or susceptibility to persistent lymphocutosis by detecting polymorphism in BOLA-DR3 exon. USA patent 5,582,487
38. Mac Lemman, et al – October 25, 1994 – Diagnosis for porcine malignant hyperthemia. USA patent 5,358,649
39. Templeton, et al – September 5, 2000 – Method of identification of animals resistant or susceptible to disease such as ruminant brucellosis, tuberculosis, paratuberculosis and salmonellosis – USA patent 6,114,118
40. Wagter-Lesperence, et al, September 11, 2001 – Method of identifying high immune response animals – USA patent 6,287,564
41. Cockett, et al. October 23, 2001 – Screening for the molecular defect causing spider lamb syndrome in sheep. USA patent 6,306,591
42. Rothschild, et al – December 20, 1994 – Method for determining genetic marker for increased pig litter size – USA patent 5,374,526
43. Rothschild, et al – August 17, 1999 – Genes and genetic markers for improved reproductive traits in animals – USA patent 5,939,264
44. Li, et al – September 18, 2001 – DNA markers for pig litter size – USA patent 6,291,174
45. Anderson, et al – February 6, 2001 – Methods for determining the coat color genotype of a pig – USA patent 6,183,955
46. Kleyn, et al – July 8, 1997 – Mammalian tub gene. USA patent 5,646,040

47. Friedman, et al – October 30, 2001 – Modulators of body weight, corresponding nucleic acids and proteins, and diagnostic and therapeutic uses there of – USA patent 6,309,853
48. Barendse William John (AU) (2002) 05-07. Assessing lipid metabolism – patent number US 6383751
49. Handes, S. E. & Stickland, N. C. (1998). Catch-up growth in pigs relationship with muscle cellularity. *Anim. Prod.* 47, 291-293
50. Swatland, H. J. & Kiefer, N. M. (1974) Fetal development of double muscled condition in cattle *J. Anim. Sci.* 38: 752-757
51. Ernst, C. W. Vaske, D. A., Larson, R. G. & Rothschild, M. T. (1993) – Rapid communication Msp1 restriction fragment length polymorphism at the swine myogenin locus, *J. Anim. Sci.* 71: 3479
52. Soumillion, et al. November 7, 2000 – Pig myogenin gene and method to identify polymorphisms related to muscle growth – USA patent 6,143,880
53. Cowan, et al – August 20, 1991 – Genetic marker for superior milk products in dairy cattle – USA patent 5,041,371
54. Collier, et al – December 20, 1994 – Allelic variants of bovine somatotrophin gene; genetic marker for superior milk production in bovine – USA patent 5,374,523
55. Wood, et al ( 1989 ) Purification and characterization of pituitary bovine somatotrophin *J. Biol. Chem* 264: 14741-7
56. Tuglye, et al – March 25, 1997 – Genetic marker for improved milk production traits in cattle \_ USA patent 5,614,364
57. Beitz, et al March 8, 1994 – Association of bovine mitochondrial DNA with traits of economic importance – USA patent 5,292,639
58. <http://www.genomic.iastate.edu/edu/newtech/sexing.html>
59. Halverson, et al. January 13, 1998 – Avian sex identification probes – USA patent 5,707,809
60. Lemming, Shi – DNA Microarray ( Genome Chip ) Monitoring the Genome on a chip <http://www.gene.chips.com/>
61. Jeremy Buhler – Anatomy of a comparative gene expression study <http://www.cs.wvstl.edu/~jbuhler//research/array/>

Todas as patentes anteriormente referidas são links do site <http://www.genome.iastate.edu/pig>

**BIOLOGIA MOLECULAR, TRANSDUÇÃO DE SINAIS E SUA  
EXPRESSÃO**

## ÍNDICE

<b>1. <u>Comunicação intra e intercelular, matriz extracelular e integrinas, membranas, vesículas e citoesqueleto</u></b>	242
1.1 - Matriz extra celular moléculas de adesão e integrinas	242
1.2 - Membranas e vesículas	253
1.3- Genoma citoesqueleto e motilidade	262
1.4 - Proteínas RHO/RAC/CDC42 como reguladores da reorganização do citoesqueleto e expressão de genes	265
1.5 – Bibliografia	267



# **1. Comunicação intra e intercelular, matriz extracelular e integrinas, membranas, vesículas e citoesqueleto**

## **1.1 - Matriz extracelular, moléculas de adesão e integrinas.**

### Adesão celular, junções celulares e a matriz extracelular

As células nos tecidos encontram-se habitualmente em contacto com um redanho de macromoléculas extracelulares segregadas constituindo estas a matriz extra celular.

Em alguns casos as células estão ligadas a esta matriz, em regiões especializadas das suas plasma membranas, chamadas junções célula / matriz. Célula em contacto directo com as células vizinhas são por vezes unidas umas às outras em junções intercelulares especializadas.

Nos tecidos conectivos dos animais a matriz extracelular é extensa e as células encontram-se escassamente dispersas nela. A matrix é rica em polímeros fibrosos, sobretudo colagénio e é a matriz mais do que as células que sofre a maioria do stress a que são sujeitos estes tecidos. As células estão ligadas aos componentes da matriz sendo as ligações mecânicas entre as células pouco importantes neste tipo de tecido.

Nos tecidos epiteliais, pelo contrário as células encontram-se firmemente ligadas entre si formando lamínas (os epitélios) sendo a matriz extracelular escassa, e a maioria do respectivo volume ocupada por células, sendo neste caso as células os principais suportes do stress, através dos fortes filamentos proteicos intracelulares (componentes do citoesqueleto) que atravessam o citoplasma de cada célula epitelial ligando-se no interior da plasma membrana da célula onde se forma uma junção especializada com a superfície de outras células ou com a matriz extracelular.

A matriz extracelular compreende uma diversidade de polissacáridos e proteínas que são segregadas localmente e se “assemblam” num redenho organizado.

Como se disse esta matrix extracelular é sobretudo abundante nos tecidos conectivos.

Variações nas concentrações relativas dos diferentes tipos de macromoléculas da matriz e na forma com elas se organizam, originam uma grande diversidade de formas adaptadas às características funcionais de cada tecido. Assim a matriz pode ser calcificada (ossos e dentes) ou formar uma matriz transparente como no caso da córnea, etc.

Na interface entre o epitélio e o tecido conectivo a matriz forma a lâmina basal muito importante no controlo do comportamento celular.

A matriz extracelular consiste sobretudo de proteínas fibrosas embebidas num gel polissacarídico hidratado, sendo estas macromoléculas segregadas pelas células vizinhas, sobretudo pelos fibroblastos no tecido conectivo, e em alguns tecidos conectivos especializados, como na cartilagem e no osso, são segregadas por células da família fibroblasto como por exemplo condroblastos e osteoblastos.

As duas principais classes das macromoléculas que formam a matriz são as glicosaminoglicanas polissacarídicas ligadas normalmente por ligação covalente á proteína formando proteoglicanas e as proteínas fibrosas colagénio e elastina, (proteínas estruturais) fibronectina e laminina (proteínas adesivas).

Nas glicosaminoglicanas são de considerar o ácido hialurónico, a condroitina sulfato, a dermatana sulfato, heparana sulfato, heparina e queratana sulfato.

As cadeias de glicosaminoglicanas podem encontrar-se altamente organizadas na matriz extracelular.

O colagénio é a principal proteína da matriz extracelular, sendo a sua organização em fibrilas adaptada às necessidades dos respectivos tecidos.

A fase aquosa do gel polissacarídico da matriz permite a difusão dos nutrientes, metabolitos e hormonas entre o sangue e as células dos tecidos. As fibras de colagénio enrijam e ajudam a organização da matriz e as fibras elásticas dão-lhe resistência. As proteínas adesivas ajudam as células a ligar-se à matriz extracelular.

A lâmina basal é uma matriz extracelular especializada composta sobretudo de colagénio do tipo IV, proteoglicanas, laminina, e realiza diversas e complexas funções. Esquemáticamente a lâmina reticularis mais a lâmina basal constituem a membrana “basement” enquanto a lâmina basal é constituída pela lâmina rara mais a lâmina densa.

As fibronectinas e lamininas são grandes glicoproteínas adesivas da matriz, a fibronectina mais distribuída nos tecidos conectivos e a laminina sobretudo na lâmina basal. Estas proteínas através dos seus múltiplos domínios de ligação, ajudam as células a aderir à matriz e muitas destas proteínas adesivas contêm uma sequência tripeptídica comum (RGD) que forma parte da estrutura reconhecida por uma superfamília de receptores homólogos transmembranários para a matriz chamados integrinas. As integrinas ajudam a ligação das células à matriz extracelular.

Algumas proteoglicanas são componentes integrais da plasma membrana, podendo ter a sua parte proteica inserida através da bicamada lipídica da membrana ou covalentemente ligada a ela, ligando-se à maioria dos tipos de componentes da matriz extracelular. Estas proteoglicanas ajudam assim a ligar as células à matriz. Contudo as componentes da matrix extracelular também se ligam à superfície celular através de receptores específicos glicoproteicos.

Os receptores para a matriz diferem dos receptores da superfície celular para as hormonas e para outras moléculas assinalantes visto unirem-se aos seus ligandos com relativamente baixa afinidade, e se encontrarem presentes a uma mais elevada concentração na superfície das células.

O receptor fibronectina dos fibroblastos mamíferos (com duas cadeias  $\alpha$  e  $\beta$ ) é um dos receptores para a matriz melhor conhecido (é uma glicoproteína da plasma membrana) funcionando como um elo de ligação transmembranar mediando interações entre a actina do citoesqueleto dentro da célula e a fibronectina na matriz extracelular.

O receptor da fibronectina pertence a uma grande superfamília de receptores homólogos para a matriz chamados integrinas a maioria dos quais reconhecem sequências RGD nas proteínas extracelulares a que se ligam.

Vários outros receptores para a matriz inclusivé alguns que ligam colagénio e laminina são também chamados integrinas sendo todos heterodímeros com cadeias  $\alpha$  e  $\beta$  homólogas daquelas do receptor de fibronectina e a maioria deles reconhece, como já se disse, sequências RGD nos componentes da matriz a que se ligam.

Contudo nem todos os receptores para a matriz pertencem a esta superfamília.

O citoesqueleto e a matrix extracelular comunicam pois através da plasma membrana, podendo a matriz influenciar a organização do citoesqueleto através de uma interação recíproca. Diversos componentes da matrix extracelular estão ligados a proteínas do citoesqueleto das células através de proteínas transmembranárias.

A matriz como referimos anteriormente é um complexo de diferentes colagénios, proteoglicanas (PG), ácido hialurónico, laminina, fibronectina e diversas glicoproteínas inclusive enzimas proteolíticas implicadas na remodelação da matriz extracelular.

As metaloproteinases da matriz ou matrixinas (abreviadamente MMP) são endopeptidases dependentes do zinco que funcionam fora das células a pH neutro, requerendo iões cálcio e contendo um centro activo constituído por um domínio catalítico conservado que contém três histidinas complexando um átomo de zinco.

As MMP são enzimas centrais no metabolismo das proteínas da matriz extracelular, inclusive colagénios fibrilares e não fibrilares, fibronectina, laminina e membrana “basement” ou glicoproteínas intersticiais do estroma, estando implicadas na sua degradação extracelular e quebra durante os processos de remodelação normal dos tecidos, como por ex: a cicatrização, gravidez e angiogénese.

Pensa-se que também facilitam a migração celular através das membranas “basement”.

A destruição dos tecidos durante uma série de processos patológicos está muitas vezes correlacionada com um desequilíbrio entre as MMP e os seus inibidores proteicos os TIMPs.

A expressão das MMPs e dos TIMPs é regulada por citocinas, factores de crescimento e hormonas, sendo controlada ao nível da transcrição e/ou post-translação. Algumas MMP existem também na forma de proteínas ligadas a membranas, parecendo desempenhar um papel na modulação das interacções célula-matriz.

Historicamente as MMP individualmente eram designadas em função do seu substrato principal ou da sua fonte celular (colagenases, gelatinases, stromelisinases, etc.) mas hoje são designadas por MMP seguida por um número (MMP-1;MMP-14;MMP-26,etc).

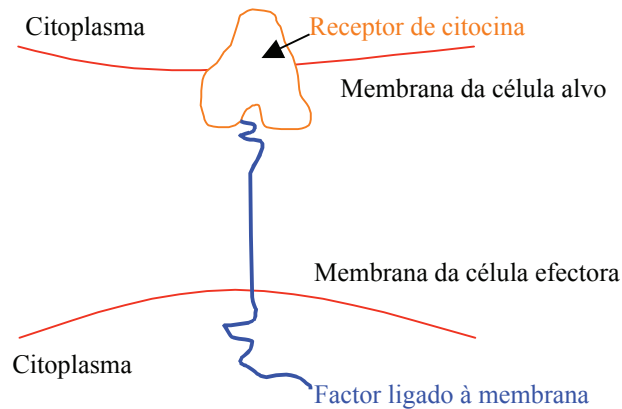
As MMP não são habitualmente expressas constitutivamente, sendo a sua expressão modificada por uma série de sinais fisiológicos e patológicos, sendo sintetizadas na forma de preproenzimas e segregadas pelas células como proenzimas (promatrixinas) que requerem activação por proteólise regulada que origina a perda do pró-peptido.

A activação destes proenzimas é uma das etapas críticas que leva à quebra da matriz extracelular.

Os inibidores das metaloproteinases ou TIMPs são proteínas relacionadas que podem formar complexos com todas as MMP conhecidas.

Diversos componentes da matriz estão envolvidos em processos mediadores de interacções célula-célula (justacrina), e em muitas circunstâncias a capacidade celular para proliferar, diferenciar-se ou exprimir funções especializadas depende da matriz.

O mecanismo justacrinico controla o crescimento e as comunicações intercelulares implicando contactos célula-célula, através de formas de factores de crescimento ligadas a membranas que são normalmente segregadas e de receptores de células adjacentes (vide figura seguinte).



As formas de citocinas ligadas a membranas, são frequentemente incompletamente processadas a partir de precursores activos de formas segregadas do factor, podendo ser geradas por splicing alternativo do mRNA correspondente.

As proteoglicanas da matriz extracelular funcionam como moduladores de actividades dos factores de crescimento.

O ambiente extracelular através da sua organização, composição e propriedades físicas, modela as funções das células endoteliais tal como a angiogénese.

As proteoglicanas são proteínas modificadas por glicosaminoglicanas (GAG) que são cadeias longas de unidades dissacarídicas repetidas, havendo quatro tipos principais como já referimos: as sulfato heparana sulfatadas/heparina, as condroitina sulfato/dermatana, as queratanas sulfato e o ácido hialurónico ou seja uma glicosaminoglucana não sulfatada.

Diversas proteoglicanas contem um âmago proteico que as liga à membrana celular e o ácido hialurónico é o único oligossacárido extracelular que não está ligado covalentemente a uma proteína.

Como as glicosaminoglicanas estão carregadas negativamente elas podem ligar-se inespecificamente a outras substâncias tais como factores de crescimento, sendo as suas interacções com as citocinas um dos mecanismos mais importantes de comunicação entre as células, que são assim medeados por factores segregados e que actuam localmente.

O metabolismo da matriz extracelular é regulado pelas citocinas, que inibem a síntese de proteoglicanas e/ou promovem a degradação das proteoglicanas, ou então podem estimular a síntese de proteoglicanas e/ou inibindo a degradação das proteoglicanas.

Alguns exemplos de citocinas que se ligam a proteoglicanas são:

- 1) Factor de crescimento dos fibroblastos que se liga a betaglicanas da matrix extracelular e HSPG (heparana sulfato proteoglicana) da plasma membrana e a sindicanas
- 2) GM-CSF ou seja o factor estimulador de colónias granulocito-macrófago que se liga a HSPG da matrix extracelular
- 3) IFN gama interferão que se liga a HSPG da matrix extracelular
- 4) IL-3 que se liga a HSPG da matrix extracelular

- 5) TGF-beta ou factor de crescimento transformante que se liga a betaglicanas, biglicanas e decorinas
- 6) Outras citocinas que influenciam a matrix extracelular.

As heparanas sulfato proteoglicanas (HSPG) sobretudo as sindecanas, ligam-se a bFGF (factor de crescimento fibroblástico) com baixa afinidade protegendo assim este factor da degradação, criando ao mesmo tempo um seu reservatório na matrix podendo esse factor ser libertado quer pela heparina, quer por enzimas específicas.

Outros factores como os GM-CSF, IL-3, inibidor das “stem cells”, LIF (factor inibidor da leucemia), IFN-gama, PF4, anfiregulina, etc, podem ligar-se de uma forma similar na matrix formando complexos com heparana sulfato proteoglicanas ou heparina.

O TGF-beta liga-se a proteoglicanas betaglicanas e a decorinas.

### Moléculas de adesão

As células imunitárias comunicam entre si e comunicam com outras células alvo para além da barreira sangue-tecidos, através de citocinas (factores solúveis) e através de interacções resultantes do íntimo contacto célula-célula, sendo neste último caso os mediadores as moléculas da superfície celular também chamadas moléculas de adesão que actuam através da união do ligando com o receptor.

Estas moléculas de adesão quanto à sua estrutura, podem ser classificadas em cinco famílias a saber:

- 1) Superfamília imunoglobulinas
- 2) Selectinas
- 3) Integrinas
- 4) Caderinas
- 5) Outras moléculas

1) A superfamília imunoglobulinas com mais de setenta membros, engloba os receptores das células T, imunoglobulinas, antígenos MHC, CD2, CD3, CD4, CD8, NCAM, ICAM1-5, VCAM-1 e PECAM-1, sendo todas caracterizadas pela presença de um ou mais domínios Ig.

A dobra imunoglobulina é composta por 70-100 ácidos aminados em duas folhas paralelas  $\beta$  estabilizadas por pontes dissulfureto.

Os membros desta superfamília ligam-se a outros membros da mesma família (por ex. os antígenos MHC ligam-se a receptores das células T), a integrinas (por ex: LFA-1 e Mac1 ligam-se a ICAM-1) e a diversos outros receptores opostos.

2) As selectinas com os três membros E-P e L selectinas ligam-se a carboidratos, ao contrário da maioria de outras moléculas de adesão que se ligam a proteínas. As selectinas têm três tipos de domínios na sua estrutura. Uma porção N-terminal dependente do cálcio e que interacciona com o ligando. Segue-se depois um domínio relacionado com o EGF (factor de crescimento epidérmico) e depois uma série de domínios repetidos também presentes na CRP (proteínas reguladoras do complemento). A região transmembranaria é seguida por uma curta cauda citoplásmica.

3) As integrinas são os principais mediadores para adesão das células à matriz extracelular, embora estejam também implicadas na adesão célula a célula. Integram portanto o citoesqueleto intracelular com a matriz extra celular.

São glicoproteínas de membrana compostas por duas subunidades,  $\alpha$  e  $\beta$ . As subunidades  $\beta$  revelam entre si uma homologia de 40-48%, e as subunidades  $\alpha$  uma homologia muito menor.

A porção N-terminal das subunidades  $\alpha$  contem três a quatro locais de ligação para catiões bivalentes (Mg).

Nas integrinas estão até hoje descritas 12 subunidades  $\alpha$  diferentes e 8 subunidades  $\beta$ , podendo as subunidades  $\alpha$  individuais estar associadas com mais de um tipo de subunidade  $\beta$ . O local de união parece residir nas duas subunidades e os seus domínios citoplásmicos formam conexões com o citoesqueleto estabelecendo a ligação entre este e a matriz extracelular (vide adiante mais desenvolvimentos sobre integrinas).

4) As caderinas são moléculas de adesão que dependem do cálcio e que têm um elevado grau de homologia (43-58%).

Elas actuam como receptores e como ligandos, sendo responsáveis pela adesão selectiva célula a célula ou pela distribuição celular para levar diferentes tipos de células ao seu destino adequado durante o desenvolvimento.

Elas também mantêm a integridade de diversas estruturas multicelulares. Têm um precursor polipeptídico que sofre uma série de modificações post-translaçionais (glicosilação, fosfatação e cisão proteolítica), tendo as proteínas maduras 723 a 748 ácidos aminados, contendo o domínio extracelular três a cinco repetições integrais de cerca de 110 ácidos aminados.

5) Nas outras moléculas de adesão inclui-se o CD44 uma glicoproteína muito ubíqua expressa nas células hematopoiéticas (células B e T, monocitos e neutrófilos), células epiteliais, fibroblastos e células gliais, estando envolvida nas interacções célula a célula e na adesão célula - matriz.

Apesar de derivado de um só gene o seu mRNA sofre diversos processos de splicing. A sua forma padrão liga-se a hialuronato, mas as suas variantes isoformas ligam-se à fibronectina, laminina e colagénio.

Na interacção dos leucócitos com as células endoteliais os leucócitos têm capacidade para migrar através de todos os organismos animais, mediando a observação imunitária e desencadeando as respostas inflamatórias a antigénios estranhos. Inicialmente, na inflamação, os leucócitos e as células endoteliais exprimem selectinas que reconhecem sequencias específicas dos carboidratos residentes nos leucócitos ou nas células endoteliais desencadeando um retardamento do fluxo das células em circulação e o seu encarquilhamento ao longo da parede do vaso, acabando por ocorrer uma forte adesão aos endotélios resultante das integrinas dos leucócitos que se ligam aos seus receptores ICAM (intercellular adhesion molecules) com a ajuda de citocinas. Após isto à uma migração dos leucócitos para o interior dos tecidos.

Por outro lado os linfocitos derivados do timo (células T) são fundamentais na resposta imunitária o que é mediado através do complexo receptor das células T (TCR)-CD3 que interactuam com os antigénios associados com o MHC (complexo maior de hiscompatibilidade).

O reconhecimento específico do antigénio origina a activação das células T e a sua proliferação contribuindo ainda para esta activação outras moléculas de superfície, como é o caso das moléculas acessórias ICAM-1 e do seu receptor LFA-1.

No quadro seguinte referem-se algumas moléculas de adesão importantes, locais da sua expressão e receptores com que interactuam.

Moléculas de adesão, locais de expressão e receptores com que interactuam

<u>Moléculas de adesão</u>	<u>Expressão</u>	<u>Receptores</u>
<b>1) Superfamília Ig</b>		
ICAM 1	Células endoteliais, monocitos. Varias células induzíveis com os linfocitos T e B, células dendríticas, fibroblastos, células endoteliais.	LFA 1 Mac 1
ICAM 2	Células endoteliais, linfocitos, monocitos, células dendríticas	LFA 1
ICAM 3	Leucócitos, células dendríticas	LFA 1
PECAM-1	Células vasculares, plaquetas, monocitos, neutrófilos, etc.	
VCAM-1	Células endoteliais activadas, macrófagos, células dendríticas, fibroblastos, mioblastos, miotubos	VLA-4
EpCAM 1	Células epiteliais	
<b>2) Família selectina</b>		
E-selectina	Endotélio activado	
L-selectina	Linfocitos, monocitos, granulocitos	
P-selectina	Plaquetas, endotélio activado	PSGL-1
PSGL-1	Linhas mielóides, linfóides e dendríticas	P-selectina
<b>3) Família _ 2 integrinas</b>		
LFA-1	Linfocitos, neutrófilos, monocitos, macrófagos	ICAM-1,2
Mac-1	Monocitos, macrófagos, granulocitos	ICAM-1 C3b, LPS
p150,95	Macrófagos, monocitos, células NK	Fibrinogenio, iC3b
<b>4) Caderinas</b>		
VE-caderina	Células endoteliais	VE-caderina
<b>5)Outras</b>		
Cd44	Linfocitos, monocitos, neutrófilos, células epiteliais, fibroblastos, monocitos	Hialuronato

## Integrinas

Todos os acontecimentos que decorrem ao nível das zonas de adesão celular participam directamente na diferenciação celular, proliferação e migração celular, tendo assim impacto na embriogénese, cicatrização de feridas, inflamação e cancro.

Muitas das propriedades adesivas das células são devidas a um ou mais membros da família das “integrinas” dos receptores de adesão celular.

As integrinas são receptores heterodiméricos da superfície celular que na presença de catiões divalentes promovem interacções célula-matriz e célula a célula.

As integrinas mediando o contacto das células com a matriz extracelular, regulam uma série de funções celulares, podendo regular segundos mensageiros intracelulares.

Mudanças na expressão dos genes, na fosfatação da tirosina, na concentração do cálcio intracelular e no pH intracelular podem ser desencadeadas pela aglomeração “clustering” das integrinas que podem também influenciar a proteína cínase C e o metabolismo do inositol, desencadeando assim transdução de sinais.

As integrinas podem inclusivé alterar a expressão de proteinases remodeladoras da matriz extracelular.

As proteínas da matriz extracelular ligam-se a múltiplas integrinas e estas ligam-se a múltiplos ligandos.

As integrinas ligam-se por outro lado ao citoesqueleto, sobretudo aos filamentos de actina, na face citoplásmica da plasma membrana.

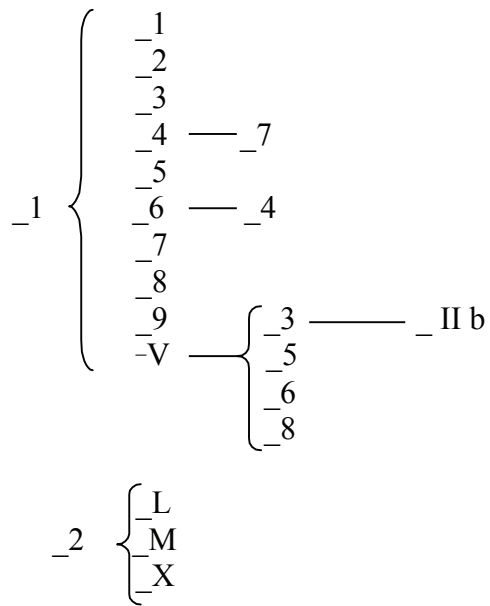
As integrinas podem pois funcionar como moléculas transdutoras de sinal.

A transdução de sinal pode ser iniciada pela ligação de proteínas intracelulares á integrina, induzindo esta ligação a alteração importante da sua conformação, passando esta alteração da conformação para o domínio extracelular da integrina através da membrana aumentando a afinidade do domínio extracelular para certos ligandos, podendo existir pois dois estados, um de baixa afinidade para os ligandos e outro de alta afinidade para os ligandos.

Uma integrina individual é composta de uma sub-unidade  $\alpha$  e de uma sub-unidade  $\beta$  cada uma das quais atravessa a plasma membrana, estabelecendo assim conexão entre o ambiente extracelular e o intracelular.

Há pelo menos vinte e um membros na família das integrinas definidas pelo emparelhamento de uma sub-unidade  $\alpha$  com uma sub-unidade  $\beta$  como se mostra na figura seguinte.





Família das integrinas

Uma dada célula pode exprimir na sua superfície integrinas distintas funcionalmente e estruturalmente, e em geral integrinas de composição distinta têm diferentes capacidades de ligação como se indica na figura seguinte.

## Integrinas

## Ligandos

_1	{	_1	—	Laminina, Colagenio
		_2	—	Laminina, Colagenio _3_1
		_3	—	Laminina, Colagenio, Fibronectina, Epiligrina, Entactinas _2_1
		_4	—	Fibronectina ( CS-1 ) VCAM
		_5	—	Fibronectina ( RGD ) invasina
		_6	—	Laminina, Merosina, Kalinina, Invasina
		_7	—	Laminina
		_8	—	?
		_9	—	?
		-V	—	Fibronectina, Vitronectina
_2	{	_L	—	ICAM-1, ICAM-2, ICAM-3
		_M	—	IC3b, Fibrinogenio, Factor X, ICAM X
		_X	—	IC3b, Fibrinogenio
-V	{		— II b	Fibrinogenio, Fibronectina, Factor de Von Williebrand, Vitronectinas
		_3	—	Vitronectina, Fibrinogenio, Factor de Von Williebrand, Fibronectina, Colagénio desnaturado, Laminina, Tenascina, Osteopontina, Trombospondina
		_1	—	Fibronectina, Vitronectina
		_5	—	Vitronectina
		_6	—	fibronectina
	—	_8	—	?
_4	{	_6	—	Laminina, Kalaninas
_7	{	_4	—	Fibronectina ( CS-1 ) VCAM, MADCAM
		_IEL	—	?

Está no entanto provado que duas integrinas com as mesmas ou similares propriedades de ligação podem veicular sinais biológicos distintos para a célula. Contudo certas integrinas são específicas para certas células.

As \_1 integrinas tendem a mediar interações celulares com os componentes da membrana basement tais como o colagénio, laminina e fibronectina, enquanto as integrinas \_v promovem a adesão a proteínas provisórias da matrix tais como o fibrinogénio, vitronectina, fibronectina, trombospondina e osteopontina durante acontecimentos relacionados com a reparação e organização dos tecidos.

A ligação ás integrinas transmite sinais para o interior das células os quais desencadeiam mudanças no comportamento biológico dessas células.

A glândula mamária constitui o mais claro exemplo dos efeitos e das intersecções entre a célula e a matriz extracelular (ECM) e da forma como a remodelação da ECM condiciona o funcionamento do tecido mamário. Dado o interesse desta matéria desenvolvemos um pouco mais sobre este assunto.

Pelo menos no tecido epitelial mamário murino a expressão dos genes específicos do tecido parece extraordinariamente sensível à estrutura e composição da ECM. Nestas células a expressão das proteínas do leite em resposta às hormonas lactogénicas, parece necessitar da participação de integrinas e ser regulada “in vivo” pela remodelação da ECM, resultando esta remodelação de mudanças na expressão de proteinases da ECM e dos seus inibidores.

A expressão do gene da  $\kappa$ -caseína é controlada por interacção célula-ECM, provavelmente em parte pela fosfatação de proteínas e muito deste controlo é exercido ao nível da transcrição.

Outras proteínas do leite são também reguladas pela ECM, mas existem diferenças significativas na sua regulação.

A expressão dos componentes da ECM das células epiteliais mamárias é sensível a regulação por contactos ECM.

O fenótipo segregatório das células epiteliais mamárias durante a lactação e a involução desta é regulada pela remodelação da ECM.

Durante a gravidez as células epiteliais mamárias proliferam e por partição a glândula enche-se de alvéolos excretores, consistindo de células epiteliais polarizadas ligadas por tight junctions, contactando uma membrana basement uniforme.

Esta estrutura mantém-se durante a lactação. Após a desmama os alvéolos regressam e os genes das proteínas do leite são reprimidos e a glândula assume uma estrutura similar à dos animais virgens.

Uma série de proteinases segregadas, e inibidores de proteinases são expressas durante este ciclo, como o caso da estromelina<sup>1</sup> e da gelatinase 72kDa.

Durante a lactação, o mRNA da  $\kappa$ -caseína é expresso abundantemente, ao contrário dos mRNA das proteinases e inibidores.

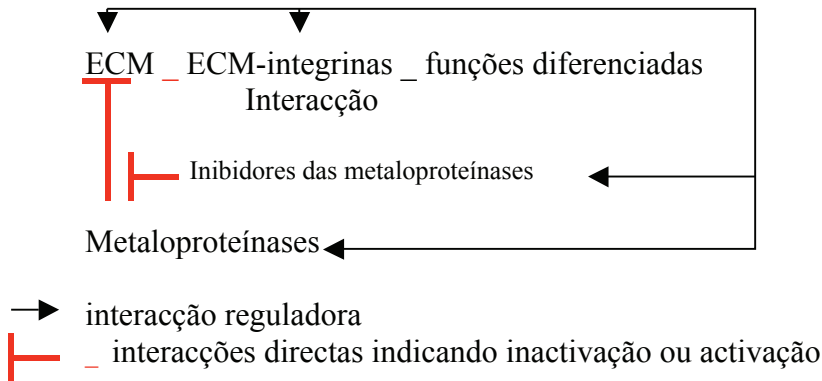
Durante a involução, a membrana basement é degradada e as células epiteliais mamárias perdem a capacidade de exprimir os genes das proteínas do leite e são depositadas dentro da lamina, onde acabam por morrer.

A degradação da ECM parece envolver a cooperação entre as células epiteliais e as mesenquimatosas.

Quando a involução começa a primeira alteração é a indução dos inibidores das proteinases, depois aumenta a expressão das proteinases e declina a expressão dos seus inibidores.

Parece que o equilíbrio entre proteinases e seus inibidores é que modula o estado diferenciado das células mamárias mais do que a sua expressão, regulando os contactos integrina – ECM através do controlo do “turnover” da membrana basement. Os sinais através das integrinas parecem ser necessários para o controlo da expressão dos genes específicos do tecido e para a viabilidade celular.

## Modelo de interações ECM-integrinas



O equilíbrio entre as metaloproteinases e os inibidores das proteinases determina a estabilidade dos componentes da matriz extracelular e assim afecta a natureza dos ligandos que têm acesso às integrinas e a outros receptores da superfície celular.

A sinalização mediada pelas integrinas regula o fenótipo diferenciado das células e modula a expressão ou actividade das metaloproteinases, dos seus inibidores, das integrinas e dos componentes do matriz extracelular.

### **1.2-Membranas e vesículas**

Um aspecto marcante das células eucariotas é a capacidade para segregar reacções bioquímicas dentro de organelos revestidos por membranas, como sucede no caso do retículo endoplásmico, do complexo de Golgi, endosomas e lisossomas.

Nestes compartimentos sub-celulares existem proteínas residentes características de cada um destes compartimentos e responsáveis pelas suas funções bioquímicas próprias e proteínas em trânsito que passam através do seu percurso até ao seu destino final.

Nas vias transportadoras intracelulares a carga é transportada através de uma vesícula formada de novo a qual se destaca de um organelo dador e é transportada para o organelo receptor onde se fixa e funde para lhe transferir os seus conteúdos (ver figura anexa seguinte).

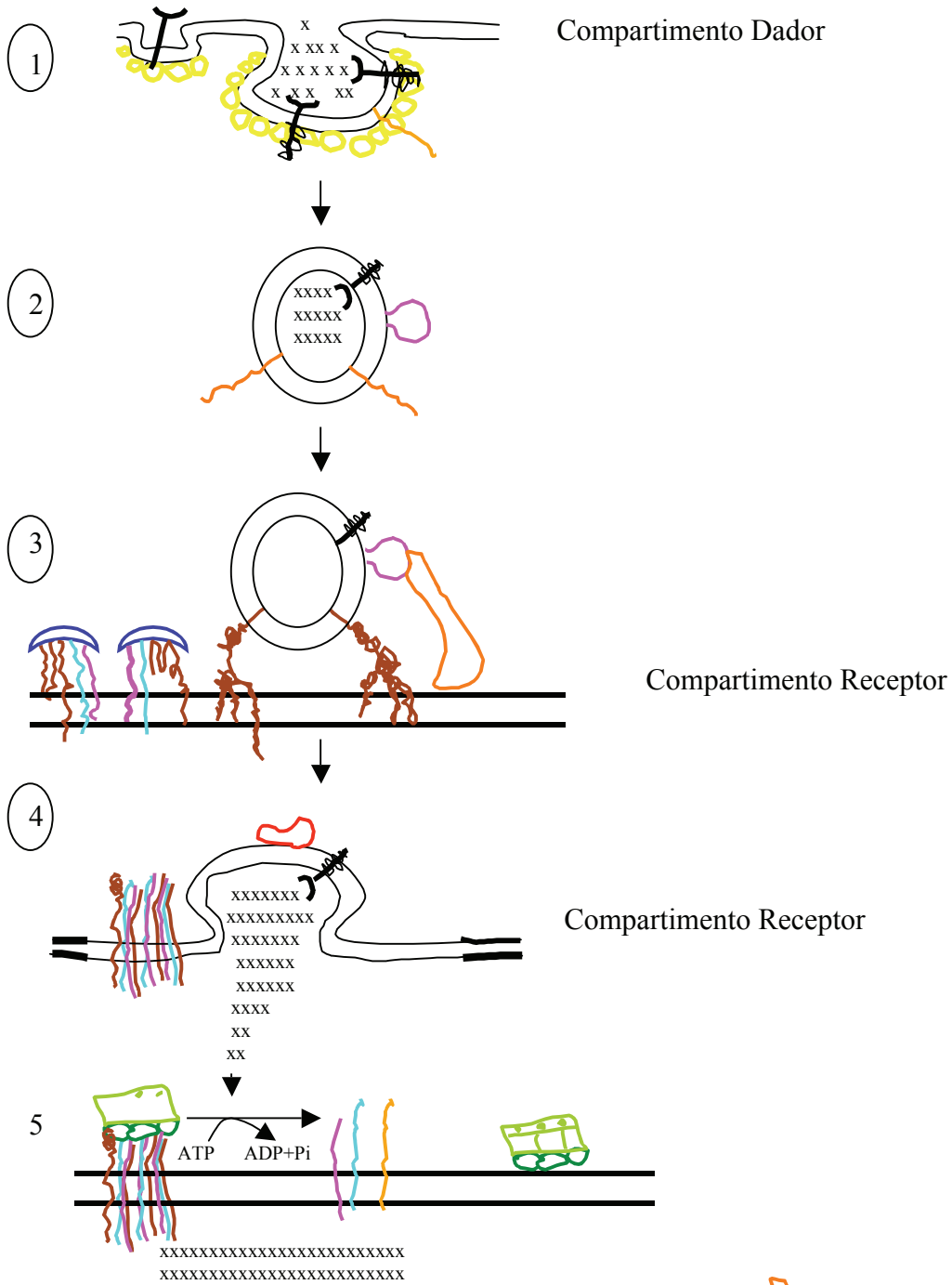
Estas etapas distintas ocorrem em todas as vias transportadoras através de vesículas como por exemplo no transporte de uma proteína do retículo endoplásmico para o aparelho de Golgi para aí ser glicosilada, ou do endossoma para o lisossoma para aí ser catabolizada, embora cada um destes transportes seja específico.

Na organização dos compartimentos membranários das células eucariotas dos animais superiores e no respectivo tráfego de vesículas nos mecanismos responsáveis por esta organização, estão identificadas famílias membros dos chamados SNARE, dos complexos de revestimento e das famílias de proteínas Rab e Sec1.

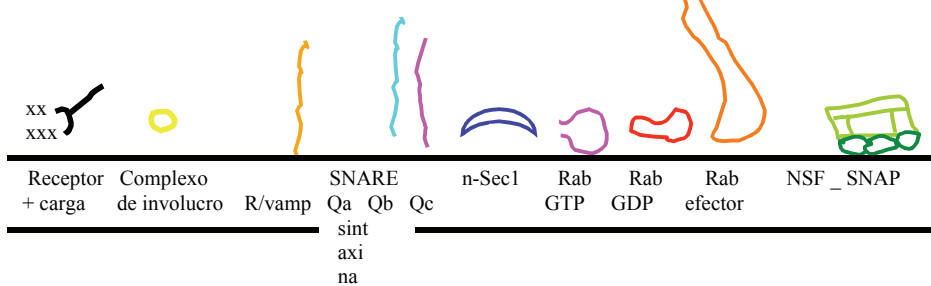
Nos humanos através do estudo do respectivo genoma foi possível identificar 35 famílias SNAREs com 4 sub-famílias, 60 famílias Rabs e 53 sub-famílias do complexo de revestimento ou invólucro.

As famílias de proteínas implicadas neste tráfego por vesículas encontra-se altamente conservadas entre diversas espécies. Cada etapa do transporte por vesículas envolve um conjunto de proteínas similares contidas dentro destas famílias antes referidas, sendo os diferentes membros de cada família contidos nos distintos compartimentos das biomembranas, os responsáveis pela especificidade de cada transporte.

Ciclo de uma vesícula de transporte de um compartimento dador para um compartimento receptor



Bicamada lipídica das biomembranas



### Legenda da figura

**1 e 2** - Gemulação (formação de uma vesícula). A membrana dadora recruta de início alguns factores que não estão referidos na figura e a seguir recruta complexos de invólucro (ver legenda de símbolos na figura) ao mesmo tempo que as proteínas a transportar ou de carga se difundem para dentro da gémula em formação onde interactuam com os complexos de invólucro e com os receptores de cargas.

A membrana esboça a formação de uma gémula revestida dando origem a uma vesícula (2).

Uma proteína GAP activadora de GTPase que faz parte do invólucro estimula a actividade GTPase, a qual leva a perdas dos complexos de envólucro, e das proteínas GTPase que se libertam para o citosol onde podem reiniciar um novo ciclo de transporte.

As proteínas são transportadas do local onde se formaram na membrana, o compartimento dador e com intervenção de elementos do citoesqueleto e motores de transporte atingem o compartimento receptor.

**3** – Na acostagem as vesículas são levadas aos locais onde ocorrerá a fusão o que envolve a intervenção da Rab-proteína que liga GTP, proteínas encabrestantes alongadas, em serpentina (coiled coil) com a ajuda ainda de outros factores. Após a acostagem uma R-SNARE (VAMP) junta-se com Qa-SNARE (syntaxina), com Qb-SNARE (SNAP-N) e um Qc-SNARE (SNAP-C) para formar um conjunto de quatro hélices paralelas.

Admite-se que as interacções da Sec1 com as Qa-SNARE seja decisiva para formar o complexo de fusão.

**4** – A fusão ocorre após as hélices paralelas levarem a vesícula e a membrana receptora a juntarem-se.

**5** – A desassemblagem, após fusão, leva a que o Rab-GDP é libertado da membrana.

O complexo SNAREs é também desassemblado por um factor ATPase N-sensível a etilmaleimida (NSF) e  $\gamma$ -SNAP de forma que o SNARE possa ser reciclado para novos ciclos de transporte.

### Papel das pequenas G proteínas sobretudo no transporte de vesículas

As pequenas proteínas que se ligam ao GTP ou G proteínas existem nos eucariotas e constituem uma superfamília com mais de cem membros, classificada em pelo menos cinco famílias as Ras, Rho, Rab, Sar1/Arf e Ran.

Estas pequenas proteínas que ligam GTP são G proteínas monoméricas, com massas moleculares de 20-40 kDa estando algumas delas (as proteínas Ras) relacionadas com as G proteínas heterotriméricas tais como as GS, Gi e G proteínas envolvidas na síntese proteica como é o caso do factor de alongamento Tu(EF-Tu).

As pequenas proteínas que se ligam ao GTP funcionam como temporizadores biológicos e também determinam no espaço certas funções específicas das células. Assim a família Ras regula a expressão dos genes, a família Rho regula a reorganização do citoesqueleto e expressão de genes, as famílias Rab e Sar1/Arf regulam o tráfego de vesículas e a família Ran regula o transporte

nucleocitoplásmico durante as fases G1,S e G2 do ciclo celular e a organização dos microtúbulos durante a fase M.

No quadro seguinte referem-se as pequenas G proteínas identificadas nas famílias Rab, Sar1/Arf e Ran dos mamíferos

<i>Reguladoras do tráfego de vesículas</i>		<i>Reguladoras do transporte Nucleocitoplasmico</i>
<b>Família Rab</b>	<b>Família Sar1/Arf</b>	<b>Família Ran</b>
Rab1A	Arf1	Ran
Rab1B	Arf2	
Rab2	Arf3	
Rab3A,B,C,D	Arf4,5 e 6	
Rab4	Sar1ae1b	
Rab5A,B,D	Arf1,2,3,4,5,6,7	
Rab6,7,8,9 e 10	Arf1	
Rab11A e 11B		
Rab12,13,14,15,16,17,18,19,20,21,22,23,24,25 e 26		
Rab27A e 27B		
Rab28,29,30,31 e 32		
Rab33A e 33B		

Todas as pequenas G proteínas têm sequências consenso em ácidos aminados responsáveis pela interacção com o GDP e GTP e para actividade GTPase. Têm também uma região que interactua com efectores. As pequenas G proteínas dos tipos Ras, Rho/Rac/Cdc 42 e Rab têm sequências na porção C-terminal que sofrem modificações postranslação com lípidos como farnesilo, geranylgeranilo, palmitoilo e metilo, e proteólise. As proteínas Arf têm na extremidade N um resíduo de glicina que é modificado com ácido merístico. Aquelas modificações lipídicas são necessárias para a sua ligação por efectores.

A maioria das pequenas G proteínas são localizadas quer no citosol quer nas membranas.

As Rab proteínas são reguladoras do tráfego de vesículas intracelularmente, sendo necessárias para o tráfego do aparelho de Golgi para a plasma membrana e do retículo endoplásmico para o aparelho de Golgi.

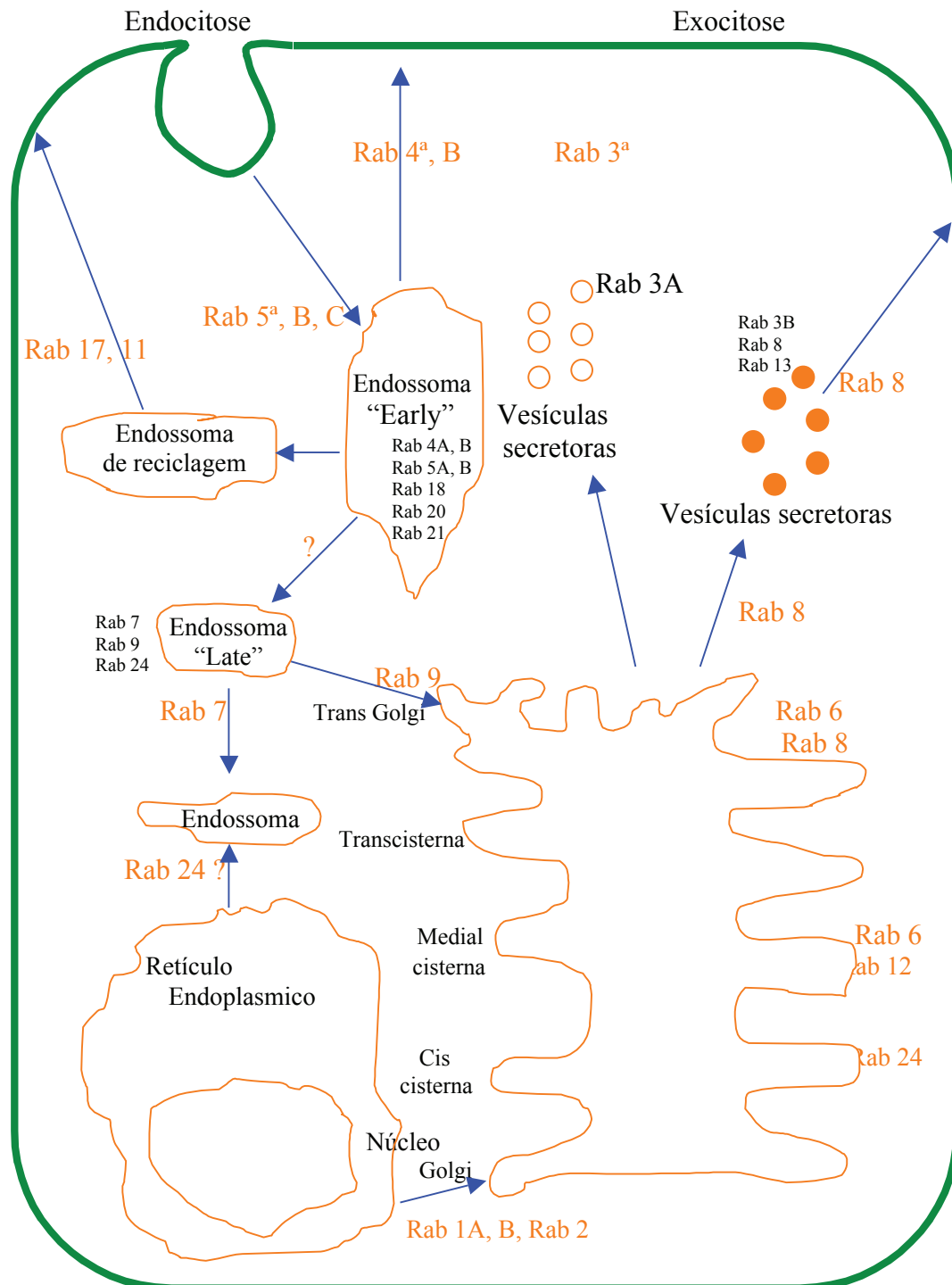
Proteínas transmembranárias e proteínas solúveis segregadas são transportadas a partir de um dado compartimento por vesículas sendo as proteínas segregadas translocadas para o retículo endoplásmico e depois transportadas para a plasma membrana através do aparelho de Golgi por vesículas.



Por outro lado macromoléculas tomadas a partir da plasma membrana são transportadas para os endossomas e lisossomas por vesículas, inclusivé certos receptores para ligandos extracelulares.

Assim a exocitose, endocitose e reciclagem são realizadas por tráfego intra celular em vesículas (vide figura seguinte).

Localização sub-celular das proteínas Rab



A exocitose pode se feita por duas vias uma regulada e outra constitutiva.

Há quatro etapas em cada transporte intracelular de vesículas, ou seja:

- 1) Rebento ou gemulação de uma vesícula a partir de uma membrana dadora
- 2) Destaque da vesícula e seu direcionamento para a membrana receptora
- 3) Docagem da vesícula a esta membrana receptora
- 4) Fusão da vesícula com esta membrana receptora

As proteínas Rab regulam estes processos, embora a etapa da formação do rebento seja sobretudo regulada pelas proteínas Sar1/Arf.

Um tipo de proteínas Rab regula a segregação e outro tipo está envolvido no transporte em forma de vesículas.

10 a 50% de uma dada Rab proteína encontra-se no citoplasma na forma ligada a GDP por interacção com Rab GDIs (inibidores da dissociação do GDP).

Cada tipo das diversas proteínas Rab está localizado nos organelos implicados.

Estas proteínas Rab ciclam entre a forma inactiva ligada a GDP e a forma activa ligada a GTP e entre o citosol e as membranas e estes acontecimentos, activação, inactivação e translocação são regulados por GEPs (proteínas permutadoras de nucleótidos de guanina), GDIs e GAPs (proteínas activadoras deGTPases).

As vesículas e as SNAREs da membrana alvo (v- e t- SNAREs) interactuam entre si originando a docagem das vesículas às membranas alvo, apesar de estar assinalado que vesículas derivadas do retículo endoplásmico com membranas do aparelho de Golgi podem ocorrer na ausência de SNAREs.

Também proteínas encabrestantes (p115,US01,EEA1) são fundamentais na marcação e docagem das vesículas. Estas proteínas encabrestantes ligam-se às membranas antes da formação dos complexos SNARE.

Há diversas evidências que assinalam que as Rab proteínas também são importantes no processo da formação de rebentos que irão dar origem a vesículas.

As proteínas Arf são homólogas das Sar1 em cerca de 35%.A Arf 1 está localizada no aparelho de Golgi de algumas linhas celulares.

A formação de rebentos de vesículas requiere a colaboração de proteínas específicas de revestimento da face citoplásmica da membrana dadora, havendo três classes de vesículas revestidas, as clatrina, as COP (proteína de revestimento) e as vesículas COPII revestidas.

As clatrininas de revestimento contêm clatrina e complexos de proteínas adaptadoras heterotetraméricas (AP 1 a 4).

As proteínas Arf também ciclam entre uma forma com GDP ligado (inactiva) e uma forma activa com GTP ligado, sendo esta ciclagem regulada por GEP e GAPs específicos.

No funcionamento das proteínas Ran no transporte nucleocitoplásmico, e na organização dos microtúbulos, durante a fase M do ciclo celular, elas desempenham um papel central, sendo as macromoléculas transportadas nos dois sentidos entre o núcleo e o citoplasma através de um complexo NPC (transportador nucleocitoplásmico) que contém mais de 50 proteínas diferentes. Este movimento ocorre por dois mecanismos, um por difusão passiva e outro por transporte activo.

As pequenas moléculas difundem-se rapidamente através do NPC em qualquer direcção enquanto se forem maiores que 50-60 kDa, inclusive proteínas e RNAs, são transportados activamente e selectivamente requerendo três tipos de factores solúveis. As moléculas receptoras de transporte, as moléculas adaptadoras e as proteínas Ran e proteínas que se ligam a esta. As moléculas receptoras para importação para o núcleo são chamadas importinas e as de exportação exportinas.

Estas moléculas receptoras de transporte constituem uma superfamília que se liga a um domínio de GTP-Ran.

Tal como outras pequenas G-proteínas também as proteínas Ran ciclam entre uma forma activa e uma forma inactiva.

A melhor conhecida exportina Ran é a CRM 1 também chamada exportina 1.

No papel das proteínas Ran na organização dos microtúbulos sabe-se que durante as fases G1, S e G2 do ciclo celular, o transporte nucleocitoplásmico é activo mas durante a fase M, o invólucro nuclear é rompido e este transporte pára.

A formação inicial das vesículas de transporte começa com o recrutamento de um grande complexo de diferentes sub-unidades proteicas do invólucro que são fundamentais não só para a selecção da carga a transportar mas também para induzir a deformação da bicamada lipídica que leve á formação da vesícula. Cada complexo destas subunidades proteicas é específico para um dado transporte.

Por exemplo o complexo adaptador AP-1 implicado no transporte do trans-Golgi, para os endossomas consiste de quatro subunidades (\_ , \_ , \_1 e \_1) cada sub-unidade com funções discretas, desde a localização até à regulação e afinidade do complexo para os componentes da carga a transportar.

#### Numero de membros das famílias de proteínas transportadoras de vesículas

	Humanos	S. cerevisiae	Células eucario., C. elegans	D. melanogaster
Complexos do invólucro (Subunidades)	7	6	6	6
Rabs	53	31	29	24
SNAREs	60	11	29	26
Qa – SNARE/Syntaxina	35	21	25	20
Qb – SNARE/SNAP Ns	12	7	9	7
Qc – SNARE/SNAP Cs	9	6	7	5
R – SNARE – VAMPs	8	6	4	5
Sec1s	9	5	6	5
	7	4	6	3

O número de complexos do invólucro aumenta apenas ligeiramente da levedura, mosca e vermes para os humanos, mas o número de sub-unidades de invólucro, individuais, aumenta significativamente nos humanos.

Isto faz-nos supor que também nos mamíferos superiores, os produtos a transportar necessitem de grande e diverso número de sub-unidades de proteínas de invólucro, (embora o número de complexos se mantenha o mesmo) para que possam ser transportados intracelularmente os variados tipos de carga a transportar o que pode ser um tanto variado consoante as espécies animais, sendo provavelmente características a atender numa perspectiva de produção animal.

As etapas seguintes no ciclo de vida do transporte intracelular de vesículas implicam proteínas da família Rab que são uma sub-família da superfamília Ras GTPase de baixa massa molecular e que ciclizam entre uma forma GTP ligada a membrana, estado activo e um estado GDP citosólico e inactivo. Quando ligada á GTP as Rabs interactivam com um hospedeiro de diferentes proteínas chamados efectores Rab.

Estes efectores Rab desempenham diversas funções desde a segregação de vesículas até ao seu transporte pelo citoesqueleto e fixação na membrana alvo.

Diferentes Rabs estão localizados em vesículas distintas e organelos.

O genoma humano contém 60 Rabs, que parecem ter também um papel regulado.

A etapa final no tráfego das vesículas é a sua fusão com as membranas alvo, supondo-se que isto é mediado por uma família de proteínas as SNAREs (1) que são proteínas integrais das membranas que se encontram quer nas vesículas quer nas membranas alvo.

A formação dos complexos SNARE empurra a vesícula para a membrana alvo, podendo fornecer a energia para a fusão de duas bicamadas lipídicas, parecendo que se forma um complexo SNARE com quatro hélices em cada local de fusão.

Diferentes subclasses existem de SNAREs, consoante o seu perfil em proteínas, sendo os diferentes membros das famílias SNAREs localizadas em diferentes compartimentos membranosos através das células eucariotas, o que constitui a base da formação de complexos específicos bem como das cargas específicas a transportar.

As famílias SNARE têm 35 membros nos humanos, mais do que noutras células eucariotas estudadas, parecendo que os organismos multicelulares funcionam com um conjunto âmagado de SNAREs utilizando SNAREs adicionais para ulterior especialização específica pelos tecidos no tráfego membranário.

As Sec-1 são proteínas citosólicas que estão perifericamente associadas com membranas em parte por interacções com syntaxinas, sendo provavelmente chaperones que colocam as syntaxinas na conformação necessária para interagirem com outros SNAREs.

O aumento de genes para Sec-1 tal como para as SNAREs indica a especialização dos acontecimentos relacionados com o transporte.

É pois possível verificar que as quatro famílias de proteínas envolvidas no tráfego de vesículas tem mecanismos básicos comuns dos organismos simples para os multicelulares (envolucros SNAREs,

Sec-1), mas a família Rab expande-se notavelmente nos organismos multicelulares indicando que estas exercem uma maior regulação sobre as vias de transporte das vesículas.

Nos mamíferos parece haver um significativo aumento de todas as famílias estudadas o que talvez signifique uma regulação muito mais fina, mas também uma especialização no tráfego consoante os tecidos. Isto parece até ser mais claramente ilustrado no cérebro.

### **1.3 - Genoma, citoesqueleto e motilidade**

A maioria das proteínas do citoesqueleto e da motilidade foram identificadas por processos de isolamento bioquímico, métodos de clonagem tradicionais ou sequências de DNA complementar aleatórias.

Nos animais superiores existem três sistemas no citoesqueleto; os filamentos de actina, os microtúbulos e os filamentos intermediários.

A família miosina de motores moleculares desloca-se sobre filamentos de actina ou movem cargas ao longo destes, enquanto os motores de Dineina e Cinesina movem microtúbulos ou movem cargas ao longo dos microtúbulos.

Estão assinaladas por processos bioquímicos tradicionais e métodos genéticos seis actinas nos mamíferos e mais de setenta famílias de proteínas ligantes de actina nos vertebrados, 6  $\alpha$ -tubulinas e  $\beta$ -tubulinas (formando as subunidades diméricas dos microtúbulos) 12 famílias de proteínas dos microtúbulos, cerca de 31 proteínas de filamentos intermediários (nos humanos) e 5 famílias de proteínas associadas.

A maioria das famílias de proteínas que ligam um dos polímeros do citoesqueleto consistem de várias isoformas, resultando de múltiplos genes de splicing alternativos.

Actinas divergentes denominadas proteínas relacionadas com a actina ou Arp em funções especiais originam diversidade do sistema.

Foi possível verificar através do draft do genoma humano que a família actina incluía seis genes actina funcionais, mais de vinte pseudogenes e seis famílias de Arp codificadas por nove genes (quadro seguinte). O draft localiza vários dos pseudogenes e revela pelo menos 14 genes novos (7 genes de actina altamente divergentes e 7 genes Arp novos), sendo necessário averiguar destes genes quais são funcionais.

Foram encontrados apesar das dificuldades notórias, 40 genes para a miosina e 40 genes para a cinesina.





A miosina e a cinesina tem os seus próprios domínios catalíticos definidos. A maioria das cinesinas e miosinas têm caudas carboxi-terminais grandes contendo domínios divergentes (ver figura seguinte). Algumas têm domínios aminoterminais e diversos fragmentos de intrões nos genes.

Um investigador experiente com profundos conhecimentos das famílias de genes pode colocar e “obter” genes completos começando com um domínio catalítico homólogo e procurando cDNA e sequências genómicas, na base de dados publicados para as sequências flanqueadoras.

Famílias de genes humanos de actina e Arp (proteínas relacionadas com actina)

Tipo de gene	Nº de genes e pseudogenes	Localização nos cromossomas
_actina músculo esquelito	1	1
_actina músculo cardíaco	1	Desconhecido
_actina músculo liso	1	10
_actina citoplásmica	1 (22-23)	7(1,1,22,33,5,5,5,5,5,5,6,16,16,16,17)
_actina citoplásmica	1 (6)	17(1,3,11,19,X)
_actina músculo liso	1	2

Mapa de domínios nos quatro tipos de proteínas do citoesqueleto

Família de genes (número)	Domínios terminal variável	N- Domínio variável (core)	Domínio âmo	Domínios variável	C-terminal
Profilina (2)					
Actina (6)					
Miosina (40)		Domínio catalítico 750 resíduos	Domínio 1-7 IQ 750 resíduos	Domínio 1-7 IQ 1.100 resíduos (de 15 coiled-Coil, resíduos SH <sub>3</sub> , ligante de membrana, PH, talin-like, FERM, TH4)	Domínio 30-300 resíduos (coiled-coil, cadeia leve ligante, ligante de cargas)
Cinesina (40)		Domínio de O- 380 resíduos (coiled-coil) ligação de carga	Domínio motor 340 resíduos	Domínios de 0-2. 200 resíduos (coiled-coil, cadeia leve ligante, ligante de cargas)	

Nem a profilina nem a actina tem domínios acessórios.

Os genes da miosina e da cinesina têm uma gama de domínios variável nas extremidades N e ou C e dos domínios âmo catalítico e do domínio motor.

As miosinas tem 1 a 7 domínios IQ que ligam cadeias leves de miosina ou calmodulina.

Os motivos IQ referidos no quadro anterior tem cerca de 20 ácidos aminados contendo um domínio de ligação CaM (calmodulina). A interação entre CaM e o domínio IQ nas neuromodulina e na neurogranina é mais elevada na ausência do que na presença de Ca<sup>2+</sup> e a ligação de CaM é inibida pela fosfatação PKC (proteína cínase K) dentro do domínio IQ, numa serina específica.

Admite-se que o domínio IQ sirva como um regulador geral das proteínas para ligação de CaM e fosfatação PKC.

Os motivos IQ repetidos nas miosinas convencionais ligam cadeias leves de miosina.

Os domínios PDZ são módulos de reconhecimento proteína-proteína que desempenham um papel muito importante na organização de diversas cascatas de transdução de sinais, reconhecendo estes domínios motivos peptídicos curtos C-terminais, podendo no entanto reconhecer outras sequências internas que imitem uma terminação.

Diversas proteínas contendo domínios PDZ desempenham um papel importante no transporte, localização e assembly de diversos complexos assinalantes supramoleculares.

Os PH domínios cooperam na associação da plasma membrana e na função celular. Estes domínio PH (de pleckstrin homologia) é um elemento estruturalmente conservado encontrado em cerca de 100 proteínas assinalantes e está implicado na associação da plasma membrana. Mais, diversos domínios PH ligam-se *in vitro*, embora com fraca afinidade a fosfoinositídeos fosfatidilinositol-(4,5) bisfosfato e trifosfato.

Os domínios “talin-like” ligam-se a actinas.

O FERM domínio é um módulo envolvido na ligação de proteínas citoplásmicas a membranas parecendo regular as interacções citoesqueleto/plasmamembranas.

A troponina C (TnC) proteolisada pela trombina produz quatro fragmentos com capacidade para se ligarem ao  $Ca^{2+}$  sendo o domínio TH4 o sítio III com 101-120 resíduos de ácidos aminados.

O inventário genético completo fará avançar a nossa compreensão sobre as doenças. Suspeita-se que a maioria dos genes humanos possam contribuir para a doença, por exemplo variantes genéticas em proteínas contrácteis importantes podem originar disfuncionamento do músculo cardíaco. E o mesmo se passa noutros órgãos.

Há 30 anos os resultados destas mutações eram chamadas cardiopatias “idiopáticas” e ainda em 1986 a ideia de que as substituições de ácidos aminados nas proteínas contrácteis como causa de cardiomiopatias era considerado especulativa.

O inventário genético também pode ser útil no desenvolvimento de novas terapias.

Durante as últimas dezenas de anos para se compreender a biologia estudava-se um modelo ou um sistema em profundidade e depois procuravam-se extrapolar os seus princípios para os sistemas fisiológicos em que estavam implicados.

Para se compreender qualquer proteína em profundidade (biologia molecular) era necessário um década ou mais de intensos esforços laboratoriais que conduzissem ao conhecimento da sua estrutura, interacções com moléculas afins e papéis e desempenhos na fisiologia celular.

Hoje a tentação para analisar as sequências da totalidade dos genomas sem cair na genomania, não pode nunca retardar o trabalho de laboratório conducente ao conhecimento completo dos mecanismos intervenientes nos fenómenos fisiológicos.

#### 1.4 - Proteínas RHO/RAC/CDC42 como reguladores da reorganização do citoesqueleto e expressão de genes.

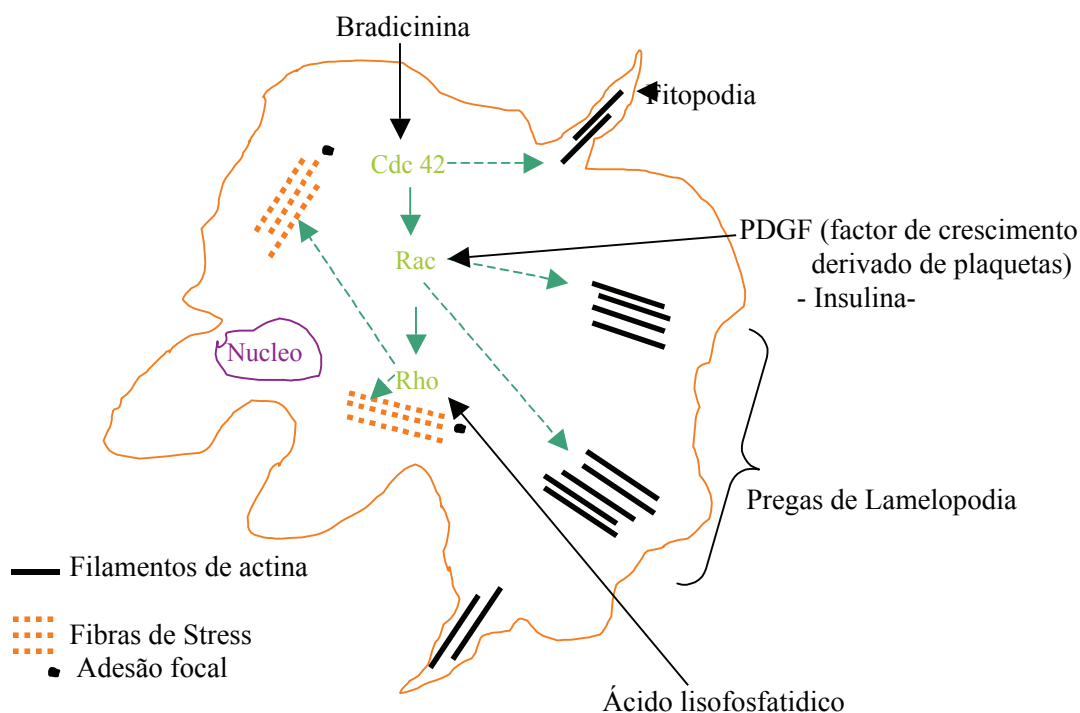
##### Pequenas G proteínas da família Rho

Rho A	Rac 1
Rho B	Rac 2
Rho C	Rac 3
Rho D	Cdc 42
Rho E/	Rnd 1/
Rho 8	Rho 6
Rho G	Rnd 21
TTF	Tc 10
Rho H/	Rho 7

As proteínas Rho regulam a formação de fibras do stress e a adesão focal nos fibroblastos.

As proteínas Rac e Cdc 42 regulam a formação de lamelopodia (folhas de actina finas em protuberância que se encontram nas extremidades das células) e filopodia (protuberancias em dedo que contêm um firme feixe de longos filamentos de actina na direção da protuberância) respectivamente (vide figura seguinte).

##### Reorganização do citoesqueleto em fibroblasto Swiss 3T3



Adaptado de *Physiol. Rev.* 81: 153-208,2001



As proteínas Rho/Rac/Cdc 42 regulam a reorganização do citoesqueleto em resposta a sinais extracelulares (Bradicinina, PDGF, insulina e ácido lisofosfatídico).

A reorganização do citoesqueleto de actina desempenha um papel muito importante na mudança de forma das células, na sua mobilidade, adesão e citocinese. O citoesqueleto de actina é composto de filamentos de actina e diversas outras proteínas que se ligam à actina.

A actina filamentosa organiza-se em estruturas de três tipos, as fibras de stress de actina (feixes de filamentos que atravessam a célula e se ligam à matriz extracelular, lamelopodia e filopodia.)

As proteínas Rho regulam a formação de fibras de stress, as proteínas Rac regulam o enrugamento e formação de lamelopodia e as proteínas Cdc 42 regulam a formação de filopodium (ver figura anexa).

A activação e inactivação das proteínas Rho/Rac/Cdc 42 é feita pelo mesmo mecanismo das proteínas Ras ou seja pelas GEPs e GAPs.

As proteínas Rho/Ras/Cdc 42 são modificadas após a translação, com lípidos.

As proteínas Rho/Rac/Cdc 42 activam o factor de resposta sérica (SRF) e também activam o factor de transcrição NFkB.

Os rearranjos do citoesqueleto estão muito relacionados com o desencadear da fagocitose. Na fagocitose de tipo I as protuberâncias da plasma membrana englobam a partícula e levam-na para o interior da célula o que é feito com as acções coordenadas das proteínas Rac/Cdc 42. Na fagocitose de tipo II as partículas afundam-se em zonas da plasma membrana revestidas de actina e aqui o internamento depende das proteínas Rho A.

Também na endocitose as proteínas Rho/Rac estão implicadas, tal como na regulação do transporte de vesículas secretórias.

## 1.5 - Bibliografia

- Kay E. Gottschalk, Adamag, D.D., Bruger, A.T., e Kessler, H. Transmembrane signal transduction of the  $\alpha$  II b  $\beta$  3 integrin Protein Science 11: 1800-1812( 2002 ).
- Bock, J.B., Matern, H.T., Peden, A.A. e Scheller, R.H. A genome perspective on membrane compartment organization Nature 409: 839-841. ( 2001 )
- Pollard, T.D .Genome: the cytoskeleton and motility Nature 409: 842-843. . ( 2001 )
- Lin, R.C., & Scheller, R.H., Mechanisms of synaptic vesicle exocytosis, -Annu. Rev. Cell Dev. Biol. 16 19-49 ( 2000 ).
- Rubin, G.M., et al Comparative genomics the eukariotes Science 287: 2204-2215 ( 2000 ).
- The C. elegans Sequencing Consortium, Genome sequence of the nematode C. Elegans: a platform for investigating biology [ published errata appear in Science283, 35 ( 1999 ), Science283, 2103 ( 1999 ) and Science285, 1493 ( 1999 ) ]. Science282, 2012-2018 ( 1998 ).
- Hirst, J. & Robinson, M.S. Clathrin and adaptors. Biochim. Biophys. Acta 1404, 173-193 ( 1998 ).
- Novick, K.P., & Zerial, M. The diversity of Rab proteins in vesicle transport. Curr. Opin. Cell Biol.9, 496-504 ( 1997 ).
- Skolnick, J., Fetrow, J.S. & Kolinski, A. Structural genomics and its importance for gene function analysis. Nature Biotechno 118, 283-287 ( 2000 ).
- Kreis, T. & Vale R. (eds) Guidebook to the Cytoskeletal and Motor Protein 2nd edn ( Oxford Univ. Press. Oxford, New York, 1999 ).
- Schroer, T.A. et al Actin-related protein nomenclature and classification, J. Cell Biol.127, 1777-1778 ( 1994 ).
- Higgs, H.N. & Pollard, T.D. Regulation of actin polymerization by Arp2/3 complex and WASp/Scar proteins, J. Biol. Chem.274, 32531-32534 ( 1999 ).
- Towbin, J.A. The role of cytoskeletal proteins in cardiomyopathies, Curr. Opin. Cell Biol.10, 131-139 ( 1998 ).
- Gunning, P. et al Isolation and characterization of full length cDNA clones for human  $\alpha$ -,  $\beta$ - and  $\gamma$ -actin mRNA: skeletal but not cytoplasmic actins have an amino-cystine that is subsequently removed. Mol. Cell Biol.3, 787-795 ( 1983 ).
- Schafer, D.A. & Schroer, T.A. Actin-related proteins, Annu. Rev. Cell Dev. Biol.15, 341-363 ( 1999 ).
- Takai, Y, Sasaki, T, & Matozaki, T. -Small GTP-binding proteins. Physiol. Rev. 81:153-208(2001)

<b>2. <u>Sinais, receptores, cascatas e transdução de sinais</u></b>	<b>269</b>
2.1 - Hormonas e factores de crescimento	269
2.2 - Principais hormonas polipeptídicas e esterólicas dos organismos animais, suas funções e sinais	271
2.3 - Proteínas cinases na transdução de sinais	277
PTK, RTK e NRTK	277
Alguns exemplos da actuação de proteínas cinases dos tipos A, C e G.	284
2.4 - Transdução de sinais e receptores membranários à superfície das células	293
2.4.1 - Tipos de receptores membranários	293
2.4.1.1 - Transdução de sinais e receptores de sete hélices	298
Receptores de Sete Hélices	298
Cascatas de Adenilato Ciclase	300
Proteínas G e sua regulação	301
2.4.1.2 - Transdução de sinais e receptores com tirosina cinases intrínsecas	311
Famílias de receptores tirosina cinase	311
Moléculas adaptadoras e sinalização	316
Moléculas adaptadoras e factores de transcrição respondendo a cascatas de transdução de múltiplos sinais (CREB) inclusivé a partir de receptores de sete hélices	323
Cascatas de transdução de sinal e insulina	332
Cascatas de transdução de sinal e IGFs	338
Cascatas de transdução de sinal e EGF	341
Cascatas de transdução de sinal e SRF	343
2.4.1.3 - Transdução de sinais e tirosinas cinases citosólicas	346
Famílias de receptores para outras citocinas	346
Moléculas adaptadoras e sinalizantes	352
Cascatas de transdução de sinais e IFN	357
Cascatas de transdução de sinais e GH	364
2.5 - Transdução de sinais e receptores intracelulares	375
Síntese e libertação de hormonas esteroides	375
Acção das hormonas esterólicas	378
Receptores esteróides e elementos de resposta no DNA	380
2.6 - Transdução de sinais e Transcrição	387
2.7 – Bibliografia	389
Iconografia	391

## 2. Sinais, receptores, cascatas e transdução de sinais

### 2.1 - Hormonas e factores de crescimento

Os diversos estímulos podem desencadear alterações na expressão dos genes das células a que têm acesso (cascatas de transdução de sinais e sua expressão proteica) e a forma como isso acontece é extraordinariamente importante para se conhecer a forma como essas células se adaptam ao seu meio envolvente.

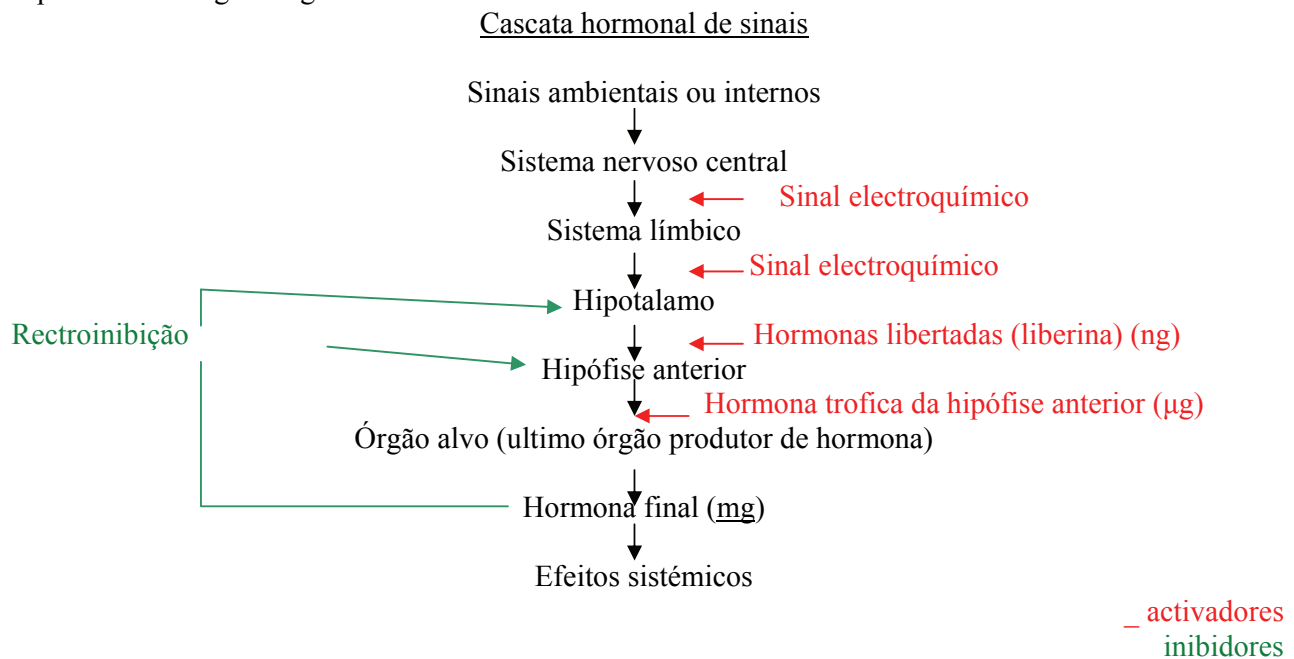
É conhecido que nos animais superiores as cascatas de transdução de sinais desencadeadas por muitos sistemas hormonais têm o seu início no sistema nervoso central e terminam na célula ou órgão alvo. Estes estímulos desencadeadores da transdução de sinais podem ter a sua origem no meio exterior dos animais ou no próprio organismo destes e esses sinais transmitidos podem corresponder a pulsos eléctricos (potencial de acção) ou a sinais químicos.

Estes sinais podem a partir do sistema nervoso central ir através do sistema límbico, hipotálamo, hipófise chegar até ao órgão alvo onde promovem a libertação da hormona correspondente a qual por seu turno atinge diversas células alvo em função do número de receptores específicos que essas células disponham para essa hormona.

Estas vias de transdução de sinais ou cascatas hormonais de sinais podem ser reguladas através de sistemas de retroinibição que podem actuar quando quantidades suficientemente elevadas da hormona final são segregadas para a circulação. Esta retroinibição pode assumir três formas, a forma longa, a forma curta e a forma ultra curta.

Na forma longa a hormona final interagindo com um receptor específico na hipófise anterior, hipotálamo e sistema nervoso central bloqueia a formação de mais hormonas por estes órgãos. Na forma curta é a própria hormona produzida na hipófise que vai interagir com um receptor específico do hipotálamo e bloquear a acção deste. Na forma ultracurta por exemplo o factor de libertação do hipotálamo bloqueia a sua própria produção no hipotálamo.

Estes são mecanismos gerais de regulação e controlo da cascata hormonal tal como se esquematiza na figura seguinte.



Refere-se seguidamente uma classificação dos tipos de sinais químicos extracelulares bem como as etapas principais da intercomunicação celular por sinais químicos.

### Tipos de sinais químicos extracelulares

Autócrinos – Actuam na própria célula que os produz e segrega  
(por exemplo factores de crescimento)

Parácrinos – Actuam em células adjacentes  
(por exemplo neurotransmissores e neurohormonas)

Endócrinos – Produzidos e segregados por órgãos endócrinos e pela circulação sanguínea atingem órgãos alvo (por exemplo: insulina)

Mistos – Mais de um mecanismo

### Etapas principais da intercomunicação celular por sinais químicos

- 1 – Síntese do sinal
- 2 – Segregação
- 3 – Transporte para as células alvo
- 4 – Interação com o receptor específico
- 5 – Modificação do comportamento celular
- 6 – Remoção do sinal e interrupção da resposta desencadeada

Nos organismos vivos as diversas células que os constituem para realizarem o seu crescimento e diferenciação permutam entre si centenas e centenas de diversos sinais que na maioria das vezes são estruturas polipeptídicas difusíveis denominadas factores de crescimento.

Estes factores de crescimento são normalmente produzidos numa forma solúvel que após proteólise e processamento adequados realizados no interior das próprias células, originam a forma madura que é finalmente segregada.

Há situações ou melhor ocorrem certos tipos de factores de crescimento, que interactivam com receptores tirosina cinases que existem à superfície das células e que após isso são internalizados nas células por endocitose envolvendo uma via clatrina-“coated pits”, sendo depois estes complexos factor de crescimento-receptor dissociados em vesículas endocíticas ácidas. Posteriormente o factor de crescimento ou ligando é degradado nos lisossomas e o receptor respectivo devolvido novamente para a superfície celular ou então degradado também.

Há outro tipo de factores de crescimento que intervêm na estimulação intercelular justacrínica e que resulta do facto desses factores de crescimento terem a sua origem em precursores que estão ligados ou ancorados às membranas. Após processamento adequado atingem a superfície das células podendo então interactivar com receptores específicos em células vizinhas sem ser necessária a formação do factor de crescimento na forma difusível.

A estimulação intercelular justacrínica ao contrário da estimulação autocrínica, paracrínica e endócrina não utiliza um factor de crescimento na forma difusível.

Pode no entanto ocorrer a conversão dos factores de crescimento ligados a membranas de uma forma activa ligada á membrana e não difusível, numa outra forma activa, solúvel e difusível. É o que sucede quando ocorre a cisão do domínio extracelular destes factores ligados a membranas, libertando formas solúveis e difusíveis. No entanto as formas ligadas á membrana podem ser muito mais importantes que as formas solúveis.

Existem numerosos factores de crescimento ligados a membranas, como é o caso da família de factores de crescimento epidérmico (EGF) com pelo menos dez membros e outros que não revelam semelhanças na sua estrutura primaria com estes EGF, como sucede com os precursores do factor estimulador de colónias (CSF-1) do ligando receptor C-Kit(KL) e do factor  $\alpha$  de necrose tumoral(TNF- $\alpha$ ), do boss, etc.

Ocorrem características estruturais comuns a todos estes factores de crescimento, tendo todos eles, com excepção do boss, uma pequena região transmembranaria e interactuando todos com excepção do TNF- $\alpha$  com receptores da família tirosina cinases (vide adiante).

Podem ocorrer alterações no processamento de factores de crescimento ancorados a membranas o que sucede por exemplo numa proteólise deficiente e nestas circunstancias estes factores acumulados á superfície das células interactuam com receptores de células vizinhas provocando interacções e adesões intercelulares e consequente estimulação intercelular. Contudo a maior parte dos factores de crescimento ancorados a membranas são internados nas células após interactuarem com os respectivos receptores.

A proteólise para conversão de factores de crescimento ligados a membrana em formas solúveis é realizado próximo da superfície celular e regulada por sinais

## **2.2 - Principais hormonas polipeptídicas e esterolicas dos organismos animais. Suas funções e sinais**

Nos animais superiores e numa perspectiva bioquímica há três grupos de moléculas que medeiam os sinais de umas células para as outras.

1º - As moléculas lipo-solúveis como as hormonas esteroides e tiroxina que se difundem através da plasma membrana celular e interactuam com receptores proteicos no citosol ou no núcleo formando-se um complexo que induz ou reprime a transcrição de determinados genes.

2º - As hormonas peptídicas como é o caso da insulina, glucagina e vários factores de crescimento e pequenas moléculas hidrófilicas como as catecolaminas que se ligam a receptores existentes á superfície das células desencadeando efeitos a curto e a longo prazo.

3º - As prostaglandinas que embora sendo lipídeo solúveis ligam-se a receptores da superfície das células.

Os sinais a considerar são muitos e variados, e para muitos deles não se conhece em toda a extensão cada uma das cascatas de transdução desses sinais, cascatas simples ou múltiplas que podem desencadear diversos tipos de acções desde a activação ou inactivação de sistemas enzimáticos pré-formados com diversas ligações sub-celulares até regulação de factores de transcrição gerais ou específicos ou mesmo, ao nível do genoma, qual ou quais os elementos de resposta a esses sinais implicados e a correspondente expressão desses genes simples ou múltiplos (dezenas ou centenas de genes como já é sabido em alguns casos particulares).

Parece-nos no entanto que embora sem se poder descrever em todo o pormenor a biologia molecular de cada uma das vias ou cascatas de transdução dos diversos sinais que podem actuar, que terá interesse vislumbrar para alguns desses sinais mais importantes quais são as acções globais que estão associadas a cada um deles. É o que pretendemos fazer no quadro seguinte em que referimos as principais hormonas peptídicas e esterólicas e as funções (acções globais) atribuídas a cada uma delas.

Principais hormonas polipeptídicas dos organismos animais, e suas funções

<u>Fonte</u>	<u>Hormona</u>	<u>Função</u>
<u>Hipotálamo</u>	Hormona da libertação da tirotrófina (TRH)	Actua sobre a tirotrófina (hipófise) libertando TSH
	Hormona da libertação das gonadotrofinas (GnRH)	Actua sobre os gonadotrofos (hipófise) libertando LH e FSH da mesma célula
	Hormona para libertação da hormona do crescimento ou somatocrinina (GRH)	Actua sobre somatotrofos (hipófise) libertando GH
	Hormona inibindo a libertação da hormona do crescimento ou somatostatina (GIH)	Actua sobre o somatotrofos (hipófise) inibindo a libertação de GH
	Hormona de libertação da corticotrofina (CRH) a vasopressina é uma hormona auxiliar da CRH na libertação de ACTH; a angiotensina II também estimula a acção da CRH libertando ACTH	Actua sobre o corticotrofo (hipófise) libertando ACTH e $\beta$ - lipotrofina
	Factor de libertação da prolactina (PRF) (não bem estabelecido)	Actua sobre o lactotrofo (hipófise) libertando PRL
	Factor inibidor da libertação da prolactina (PIF) (não bem estabelecido; pode ser uma hormona peptídica sob controlo da dopamina ou pode ser a própria dopamina )	Actua sobre o lactotrofo inibindo a libertação da PRL
<u>Hipófise anterior</u>	Tirotrófina (TSH)	Actua sobre as células do folículo da tiróide induzindo a libertação da T4 (T3)
	Hormona luteinizante (LH) (a hCG gonadotrofina corionica humana é uma hormona similar da placenta)	Actua sobre as células de Leyding nos testículos aumentando a síntese e libertação de testosterona; actua sobre o corpo lúteo do ovário aumentando a produção e libertação de progesterona
<u>Fonte</u>	<u>Hormona</u>	<u>Função</u>
	Hormona estimuladora do folículo (FSH)	Actua sobre as células de Sertoli do túbulo seminífero

		aumentando as proteínas no esperma; actua sobre os folículos do ovário estimulando a maturação do óvulo e a produção de estradiol
	Hormona do crescimento (GH)	Actua sobre diversos tipos de células produzindo IGF (ou somatomedinas), crescimento celular, e sulfatação óssea
	Hormona adrenocorticotrófica (ACTH)	Actua sobre células das cápsulas supra-renais aumentando a produção e segregação de cortisol
	$\beta$ -endorfina	Actua sobre células e neurónios produzindo efeitos analgésicos e outros
	Prolactina (PRL)	Actua sobre a glândula mamária produzindo diferenciação das células segregadoras (com outras hormonas) e estimulam a síntese de componentes do leite
	Hormona estimuladora dos melanocitos (MSH)	Actua sobre as células da pele originando a dispersão da melanina (escurecimento da pele)
<u>Hormonas de glândulas endócrinas alvo derradeiro</u>	Insulin-like growth factors (IGF)	Respondem á GH e produzem efeitos de crescimento estimulando a mitose das células
	Hormonas da tiróide (T4/T3)	Respondem a TSH e estimulam a oxidação em diversas células
	Peptídeos opióides	Podem resultar da cisão da lipotrofina ou $\beta$ -endorfina ou serem produtos de genes específicos; podem responder a CRH ou dopamina e podem produzir analgesia e outros efeitos
<u>Fonte</u>	<u>Hormona</u>	<u>Função</u>
	Inibina	Respondem a FSH no ovário e nas células de



		Sertoli regulam a secreção de FSH na hipófise; uma segunda forma da inibina (activina) pode estimular a secreção de FSH
	Peptído intermediário corticotrofina-like (CLIP)	Deriva da hipófise intermédia por degradação da ACTH; contem actividade $\beta$ -cell trofina que estimula a libertação de insulina das células $\beta$ na presença de glucose
<u>Hormonas peptídicas respondendo a outros sinais não provenientes de hormonas da hipófise anterior</u>	Arginina vasopressina (AVP hormona antidiurética ADH)	Responde a um aumento nos osmoreceptores, os quais assinalam o $[Na^+]$ extracelular; aumentam a reabsorção de água a partir do tubulo renal distal
	Oxitocina	Respondem ao reflexo de segregação e estradiol; originam a ejeção de leite nas fêmeas lactentes; envolvida nas contracções uterina no trabalho de parto; factor luteolítico produzido pelo corpo lúteo; diminui a síntese de esteróides nos testículos
Células $\beta$ do pâncreas respondem á glucose e a outros constituintes do sangue para libertar insulina	Insulina	Aumentam a utilização de glucose pelos tecidos
Células $\alpha$ do pâncreas respondem a baixas concentrações de glucose e á falta de cálcio no soro sanguíneo	Glucagina	Diminui a utilização de glucose elevando a glicémia
Derivado das proteínas sanguíneas em circulação por acção da renina e enzima conversora	Angiotensina II e III (A II e A III)	A renina inicialmente responde a uma diminuição do volume de sangue ou à diminuição de $[Na^+]$ na <u>macula densa</u> do rim a A II e A III estimulam as corticais das supra-renais para sintetizar e libertar aldo-esterona
Libertado do <u>atria</u> cardíaco em resposta à hipovolémia; regulado por outras	Factor atrial natriurético (ANF) ou atriopeptina	Actua sobre as células supra-renais mais externas diminuindo a libertação de

hormonas		aldoesterona e outros efeitos
Produzida a partir do plasma sanguíneo, intestino e outros tecidos	Bradicinina	Modula uma vasodilatação extensa originando hipotensão
Hipotálamo e mucosa intestinal	Neurotensina	Efeitos sobre o intestino, pode ter acção neurotransmissoras
Hipotálamo, sistema nervoso central e intestino	Substância P	Transmissor de dor, aumenta a contracção de músculos lisos no tubo gastrointestinal

<u>Fonte</u>	<u>Hormona</u>	<u>Função</u>
Nervos e células endócrinas do intestino; hormona hipotérmica	Bombesina colecistocinina (CCK)	Aumenta a segregação ácida gástrica; estimula a contracção da vesícula biliar e o fluxo da bilis; aumenta a segregação de enzimas pâncreaticos
Antrum estômago	Gastrina	Aumenta a segregação de ácido gástrico e pepsina
Duodeno a valores de pH inferiores a 4.5	Secretina	Estimula as células acinases pancreáticas para libertarem bicarbonato e água para elevar o pH duodenal
Hipotálamo e tracto gastrointestinal	Péptido vaso-intestinal (VIP)	Actua como um neurotransmissor no sistema nervoso autónomo periférico, relaxa a musculatura lisa da circulação; aumenta a segregação de água e electrólitos do pâncreas e intestino
Rim	Eritropoietina	Actua sobre a medula óssea para a diferenciação e inicia a síntese de hemoglobina
Corpo luteo ovárico	Relaxina	Inibe a contracção miométriais, a sua segregação aumenta durante a gestação
	Lactogénio placentário humano	Actua como a PRL e GH em virtude da grande quantidade de hPL produzida
Glândulas salivares	Factor de crescimento epidérmico	Estimula a proliferação de células derivadas da ectoderme e mesoderme juntas com soro; inibem a segregação gástrica
Timo	Timopoiatina ( $\alpha$ -II-Timosina)	Estimula os fagocitos; estimula a diferenciação de precursores em células T imunologicamente competentes
Células parafoliculares C da tiroide	Calcitonina (CT)	Baixa o cálcio no soro sanguíneo

Paratiróides	Hormona paratiróide (PTH)	Estimula a reabsorção óssea; estimula a excreção de fosfato pelos rins; sobe a calcémia
Células endoteliais dos vasos sanguíneos	Endotelina	Vasoconstrição

Principais hormonas esteroides e suas funções

<u>Hormona</u>	<u>Sinal para secreção</u>	<u>Fonte</u>	<u>Função</u>
Cortisol	ACTH	Células da fasciculata	Adaptação ao stress; leve elevação do glicogénio hepático eleva a pressão sanguínea; tomada de sódio no epitélio luminal
Aldosterona	Angiotensina	Células glomerulosas supra-renais (córtex)	Origina tomada do ião sódio; valores elevados durante o stress; sobe a pressão sanguínea
Testosterona	LH	Células de Leyding dos testículos; (supra-renal); ovário	No macho após conversão em dihidrotestosterona, produção de proteínas nas células de Sertoli; características sexuais secundárias
17 $\beta$ -estradiol	FSH	Folículo ovárico; corpo lúteo; (células Sertoli)	Fêmea; regula a secreção de gonadotrofina no ciclo ovárico, mantém o endométrio uterino; diferenciação da mama. No macho-inibição nas células de Leyding da síntese de testosterona
Progesterona	LH	Corpo lúteo	Mantém (com o estradiol) o endométrio uterino para a implantação; factor de diferenciação para a glândula mamária
1,25 (OH) <sup>2</sup> -vitamina D <sub>3</sub>	PTH	Provem das células da pele após irradiação e sucessivas hidroxilações no fígado e rim	Facilita a absorção do cálcio e fosfato no intestino; induz a ligação intracelular de cálcio as proteína
Dihidroepiandrosterona	ACTH	Células <u>reticularis</u>	Diversos efeitos protectores; fraco androgénio; não tem receptores identificados, inibe a G6-PDH, regula os coenzimas NAD <sup>+</sup>

ACTH = Hormona adrenocorticotropica  
LH = Hormona luteinizante  
FSH = Hormona foliculo estimulante  
PTH = Hormona paratiróide

Diversas hormonas, factores de crescimento, neurotransmissores e mesmo algumas toxinas regulam a actividade das células através destes diversos ligandos que se ligam a receptores adequados situados na superfície das células (hormonas polipeptídicas e derivadas de ácidos aminados) embora existam outros importantes, mecanismos para atravessar a membrana das células como sucede com as hormonas esteróides e hormonas da tiróide que interactuam com receptores no interior das células.

### **2.3 - Proteínas cinases na transdução de sinais**

#### **PTK, RTK E NRTK**

As proteínas cinases são sistemas enzimáticos caracterizados por realizarem fosfatações pertencendo á classe 2 de enzimas ou sejam transferases. Esta transferência é feita a partir do grupo  $\gamma$  fosfato do ATP ou de qualquer outro nucleótido trifosfatado, sendo o substrato um grupo alcool ou aminado formando-se uma ligação covalente.

Muitas proteínas podem ser reguladas na sua actividade por diversas proteínas cinases que portanto fosfatam resíduos de ácidos aminados dessas proteínas contendo grupos hidroxilo como é o caso da serina, treonina e tirosina. Há no entanto especificidade nestas fosfatações, havendo uma classe de proteínas cinases que fosfatam especificamente serinas e treoninas, e outra classe de proteínas cinases que fosforilam tirosinas.

Salienta-se contudo que estas fosfatações podem ser reversíveis por acção de enzimas proteínas fosfatases.

Por outro lado as proteínas que são inteiramente extracelulares não são reguladas por fosfatação reversível. Ainda pode suceder que uma só proteína cinase activada fosfate centenas de proteínas substrato em breves instantes.

Ainda no que se refere a especificidade das proteínas cinases, refere-se que algumas destas enzimas apenas fosfatam um tipo de proteína ou proteínas muito relacionadas, enquanto outras são multifuncionais fosfatando substratos um tanto diferentes embora reconheçam sequências de ácidos aminados no substrato relacionadas, como é o caso da proteína cinase. A (PKA) que reconhece a sequencia Arg-Arg-X-Ser-Z ou Arg-Arg-X-Thr-Z correspondendo X a um pequeno acido aminado e Z a um grande resíduo hidrofobo, sendo a Ser ou Thr através dos seus grupos hidroxilo o local da fosfatação.

Logo esta especificidade destas proteínas cinases é comandada pela sequência linear de ácidos aminados que se encontram á volta da Ser ou Thr, apesar de também poderem influenciar esta actividade outros ácidos aminados do substrato situados a maior distância.

Nas células eucariotas muitos dos efeitos desencadeados por AMP cíclico (AMPC) resultam da intervenção de uma simples proteína cinase que intervem nesta cascata de transdução de sinal.

Mas já o mesmo não sucede com a actuação de por exemplo da proteína cinase C (PKC) que é multifuncional e é activada pelo diacilglicerol e pela presença de  $\text{Ca}^{2+}$  e fosfatidilserina indo a PKC activada fosfatar resíduos de serina e treonina existentes em muitas proteínas alvo.

Outra situação na comunicação intercelular nos organismos superiores, na embriogénese, na manutenção dos tecidos adultos, etc, é a fosfatação dos resíduos de tirosina que integram muitas proteínas pela acção de enzimas proteínas tirosina cinases globalmente designadas por PTK.

Todas estas fosfatações que temos vindo a referir e que podem ocorrer sobre variadíssimas proteínas substrato originam locais onde podem ligar-se outras proteínas assinalantes que ocorram por exemplo em vias de transdução de sinais.

Abordaremos mais adiante em capítulo próprio exemplos de proteínas cinases dos tipos A, C e G.

Salienta-se que a actividade de muitas proteínas enzimáticas, de canais de membranas e muitas outras proteínas substrato é regulada por fosfatação que é o meio mais disseminado nos seres vivos para modificação covalente reversível de todos esses substratos e consequentemente condicionando as suas actividades e funções biológicas.

Desenvolvemos seguidamente o estudo das proteínas cinases PTK de que existem duas classes de famílias nas células; os receptores membranários proteína tirosina cinases (RTK) e as proteínas tirosina cinases sem serem receptores (NRTK).

A família das RTK corresponde a glicoproteínas transmembranárias receptoras que interagindo com determinado tipo de ligandos ou sinais fazem a transdução desse sinal extracelular para o interior da célula ou seja para o citoplasma através da autofosfatação dos resíduos de tirosina contidos no próprio receptor, desencadeando depois uma cascata para transdução desse sinal, através do concurso de uma serie de proteínas assinalantes envolvidas em diversas vias intracelulares e que induzem a proliferação celular, a diferenciação, migração ou alterações metabólicas.

Pertencem a esta família RTK os receptores que interactivam com a insulina e diversos outros factores de crescimento como é o caso do EGF (factor de crescimento epidérmico), FGF (factor de crescimento fibroblástico), PDGF (factor de crescimento derivado das plaquetas, VEGF (factor de crescimento vascular endotelial) e NGF (factor de crescimento dos nervos).

Estruturalmente estes RTK apresentam nas suas porções extracelulares domínios globulares diversos discretos, do tipo imunoglobulina, ou do tipo fibronectina (tipo III), domínios ricos em cisteína, ou do tipo EGF. As porções intracelulares destes RTK têm o domínio catalítico tirosina cinase, após a hélice transmembranária e uma região carboxílica terminal.

Conforme se pode ver na figura seguinte a maioria destes RTK são constituídos apenas por uma cadeia polipeptídica, na ausência de ligandos.

Para activação das RTK por interacção com ligandos adequados é necessário que a actividade catalítica intrínseca do próprio receptor seja promovida, assim como é necessário que se estabeleçam na estrutura do receptor regiões que possam interactuar com proteínas assinalantes.

A activação da actividade catalítica ocorre quando da autofosfatação do resíduo de tirosina dentro do domínio cinásico do RTK e por outro lado a autofosfatação de outros resíduos de tirosina com outras localizações desencadeia outros efeitos. Assim a autofosfatação de tirosinas na justamembrana e na região

carboxílica terminal leva á formação de locais (por exemplo domínio SH2 e domínio de ligação a fosfotirosinas ou PTB) para interacções com domínios modulares de proteínas assinalantes.

No entanto curiosamente no receptor para o EGF a substituição da tirosina por fenilalanina no domínio cinasico do RTK não parece afectar as propriedades assinalantes deste RTK:

Na família NRTK ou seja a das proteínas tirosina cinases que não são receptores estão englobadas as proteínas Src, as Janus cinases (Jak) e as Abl entre outras sendo estas proteínas NRTK componentes integrais de cascatas de transdução de sinais desencadeadas pelas RTK e por outros receptores da superfície das células como é o caso dos receptores acoplados com G-proteínas e receptores do sistema imunitário.

Nesta família NRTK há que considerar diversas subfamílias, sendo a maior a Src que possui nove membros e que está envolvida na mitogénese, na activação das células T e B e na reestruturação do citoesqueleto.

Outra subfamília das NRTK é a Jak que reside no citoplasma e aí interactua não covalentemente com as porções citoplásmicas de receptores do tipo citocina como é o caso do interferão  $\gamma$ . Essas proteínas Jak são activadas quando se oligomerizam devido á interacção receptor-ligando.

Após activação destas proteínas Jak elas podem depois fosfatar os receptores do tipo citocina com que estão associadas, criando assim locais para interacção com factores de transcrição do tipo STAT que são por sua vez fosfatados.

Fosfatados os STATs pelas Jak ocorre a dimerização daqueles, a sua translocação para o núcleo onde vão desencadear a transcrição específica de genes.

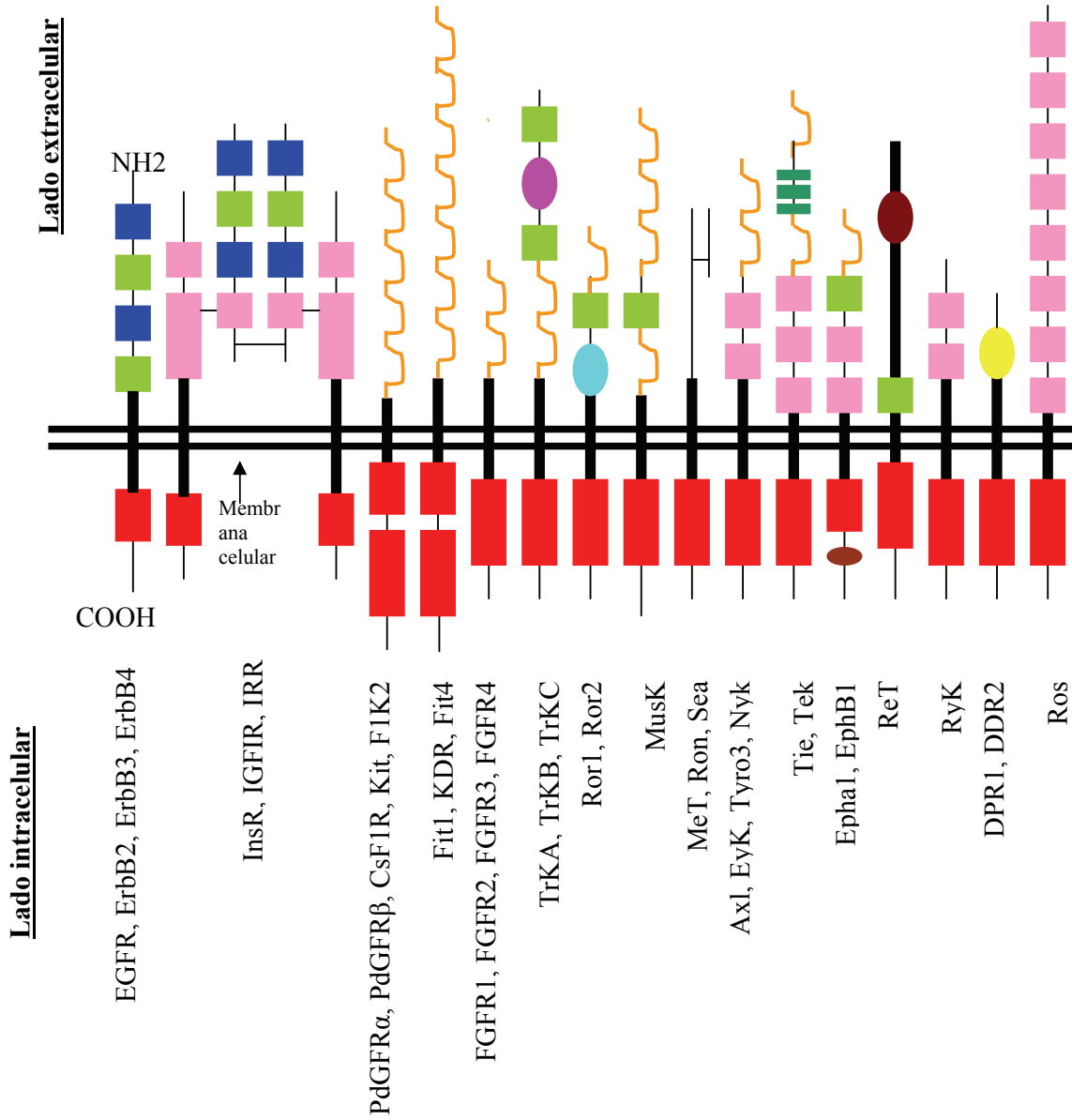
Portanto estas proteínas NRTK não dispõem de regiões extracelulares tendo a sua franca maioria residência no citoplasma. Existem no entanto algumas proteínas NRTK ancoradas á membrana através de uma modificação lipídica da porção amino terminal por miristoilação ou palmitoilação.

Como se compreende na estrutura das NRTK existem domínios tirosina cinases, domínios para interacção com outras proteínas, com lípidos, com DNA, desempenhando os domínios SH2 e SH3 papéis muito importantes nas interacções proteína-proteína.

Refira-se que o domínio SH2 tem cerca de 100 resíduos de ácidos aminados, constituindo um domínio compacto que interactua com resíduos de fosfotirosina, ou com uma forma específica da sequência proteica, enquanto o domínio SH3 é mais pequeno com cerca de 60 resíduos e interactua com sequências proteicas contendo prolina podendo formar uma hélice do tipo poliprolina.

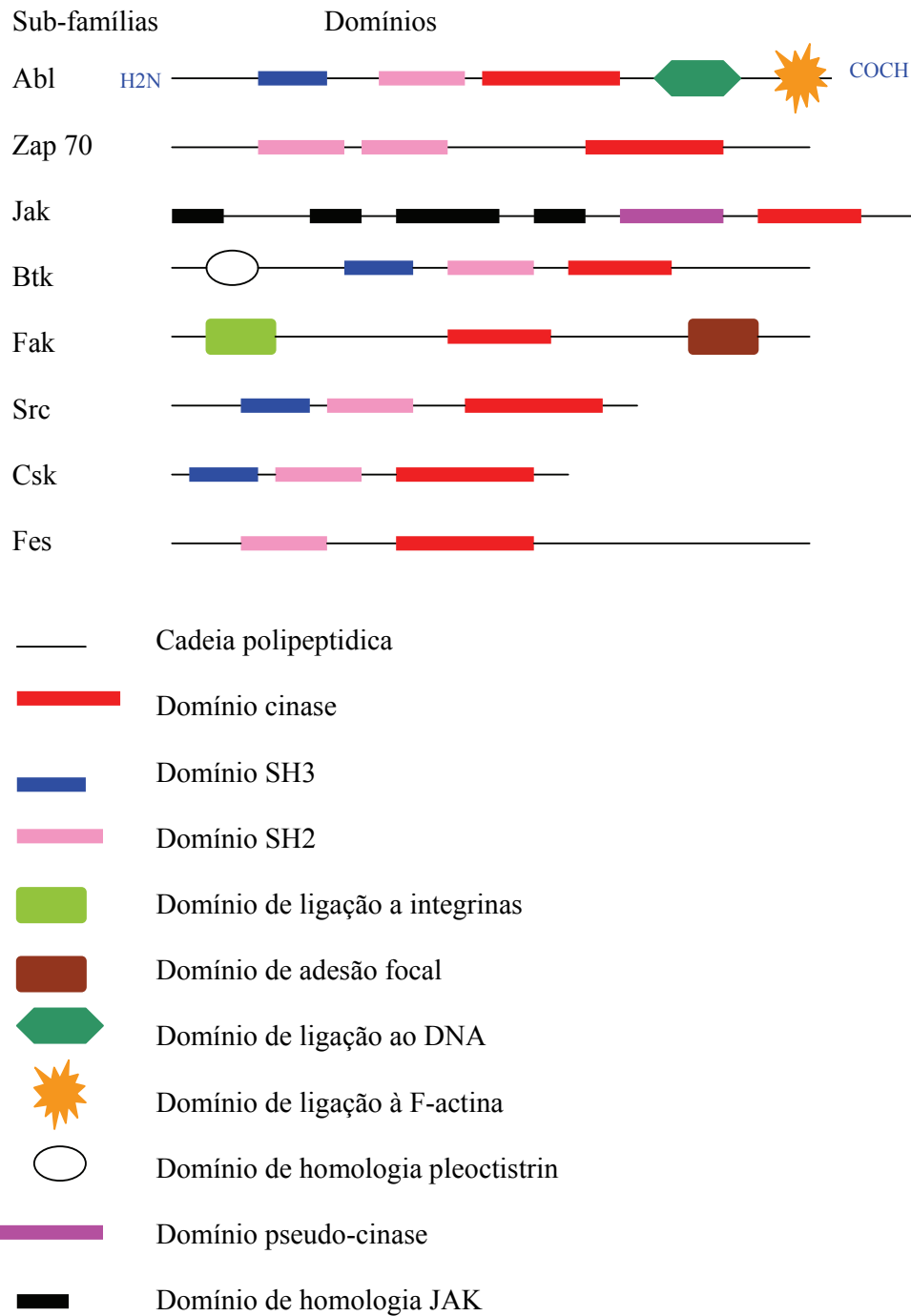
Em algumas NRTK não existem domínios SH2 e SH3 como é o caso da família Jak mas possuindo nessas circunstâncias domínios específicos para interacção com outras proteínas.

Organização em domínios de uma série de famílias de RTK



- Domínio L
  Domínio rico em cisteínas
  Domínio tirosina cinase
- Domínio tipo Ig
  Domínio fibronectina tipo III
  Domínio rico em leucinas
- Domínio tipo “Kringle”
  Domínio EGF
  Domínio caderina
- Domínio SAM
  Domínio discoina

Organização em domínios das principais subfamílias de NRTK





### Acção intracelular e proteínas cinases

Muitos sinais como por exemplo diversas hormonas polipeptídicas ou derivadas de ácidos aminados desencadeiam uma cascata para transdução do sinal ao interagirem com receptores da membrana celular e podem fazê-lo de formas variadas por diversas vias tais como:

1) por uma via que implica elevação do AMP cíclico e transmissão do sinal através da via da proteína cinase A.

2) por hidrólise do fosfatidilinositol 4,5 bifosfato e estimulação das vias proteína cinase C e IP3-Ca<sup>2+</sup>.

A concentração de cálcio no interior da célula pode aumentar devido á activação de canais da membrana celular que são sensibilizados pela ligação da hormona com o receptor permitindo a entrada de calcio extracelular havendo outra possibilidade em que por acção da hormona-receptor a fosfolipase C é activada e vai hidrolizar o PIP<sub>2</sub> originando diacilglicerol fosfato (DAG) e IP<sub>3</sub> que são dois segundos mensageiros transducidos por uma G-proteína,activando o DAG a via da proteína cinase C.Estas G-proteínas servem nas células como transdutores de sinais hormonais.

3) por uma via em que o sistema proteína cinase G pela elevação do GMP cíclico citosólico é estimulado podendo também ser directamente activado por exemplo pelo ANF(atrial natriuretic factor) sem necessitar de que o sinal da hormona seja transducido por uma G proteína.

4) por outras vias embora menos frequentes.

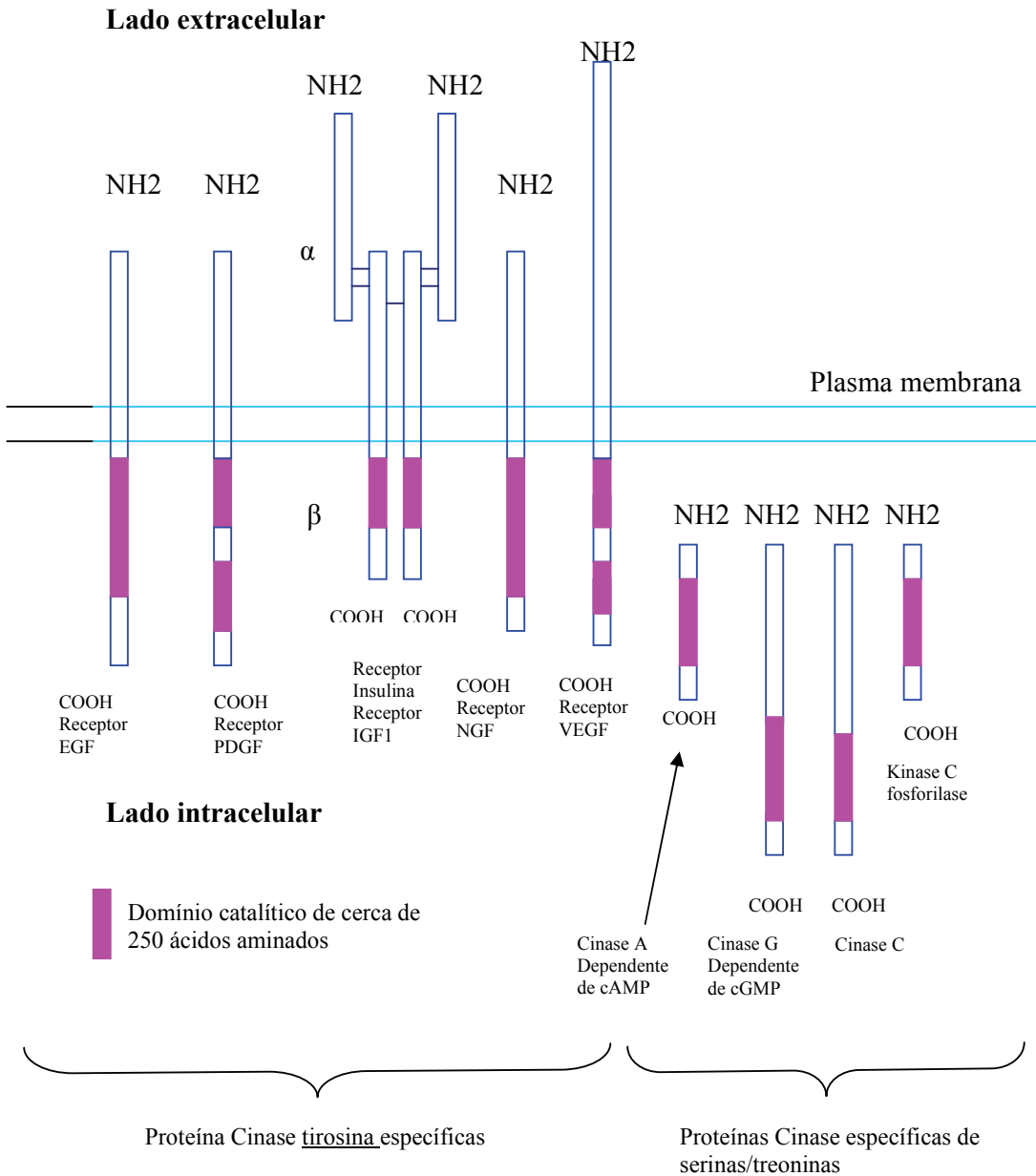
É possível que mais do que uma via seja activada pela transdução do mesmo sinal.

Ambas as proteínas cinases A e C fosfatam diversas proteínas que possuam resíduos de serinas e treoninas.

Como já referimos anteriormente um outro sistema de proteínas cinases fosfatam tirosinas situadas nos domínios citoplásmicos de alguns receptores de membrana, é o caso de receptores para factores de crescimento (no receptor para a insulina, no receptor da IGF e em certos oncogenes intervem este sistema).

Na figura seguinte mostra-se a localização celular destas proteínas cinases realçando que os seus domínios catalíticos têm estruturas primárias similares.

Proteínas cinases e tamanho e localização dos seus domínios catalíticos



As sub-unidades reguladoras normalmente associadas com a Cinase A e com a fosfatase cinase não são apresentadas na figura anterior.

- EGF – factor de crescimento epidérmico
- NGF – factor de crescimento nervoso
- VEGF – factor de crescimento endotélio vascular
- PDGF – factor de crescimento das plaquetas

Na figura anterior as proteínas cinase específicas da tirosina são receptores transmembranarios que quando activados por união a ligandos específicos extracelulares fosfatam outras proteínas e fosfatando

também as suas próprias proteínas de constituição (auto fosfatam-se), em resíduos de tirosina no interior das células.

As cadeias  $\alpha$  e  $\beta$  dos receptores da insulina são codificados por um único gene.

O domínio extra celular do receptor PDGF admite-se que é constituído por cinco domínios imunoglobulina-like sugerindo que esta proteína pertence à superfamília das imunoglobulinas.

As proteínas que são reguladas através de fosfatações e desfosfatações podem conter diversos locais para fosfatação e podem ser fosfatadas por mais de um tipo de proteínas cinases.

### **Alguns exemplos de proteínas cinases dos tipos A, C e G**

#### **Proteínas cinases A**

##### Um exemplo da acção intracelular de Proteínas Cinase A

A actividade da hormona arginina vasopressina (AVP) é um exemplo de activação de uma via por proteína cinase A ao nível das células renais distais.

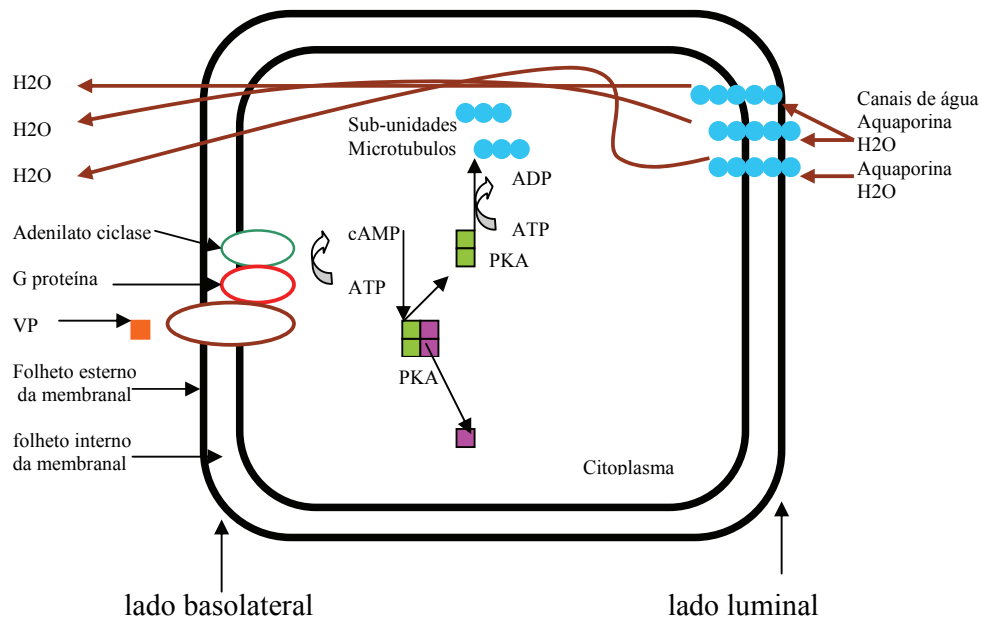
A vasopressina (VP) também chamada hormona antidiurética origina uma reabsorção de água aumentada a partir da urina, no rim distal (vide figura seguinte).

A segregação e acção da arginina vasopressina no rim distal pode delinear-se da forma seguinte. A libertação da AVP ou VP do lobo posterior da hipófise começa através de um sinal vindo de um osmoreceptor (que responde a um aumento da concentração extracelular de sal) ou baroreceptor (que responde a uma queda da pressão sanguínea). O sinal pode ser portanto um aumento da concentração extracelular de cloreto de sódio que origina que o neurónio osmoreceptor se contraia e mande uma mensagem eléctrica ao longo do axónio que faz interface com o corpo de uma célula vasopresinergica onde origina uma despolarização das terminações nervosas originando a libertação por exocitose, do complexo VP-neurofisina aí armazenado, que entra na circulação local e depois na circulação geral. Depois a neurofisina (proteína estabilizadora) dissocia-se do VP e este liga-se a um seu receptor na membrana celular das células do tubulo distal do rim (outros receptores situam-se no corticotrofo da hipófise anterior e nos hepatocitos embora com mecanismos intervenientes diferentes). Através de uma adenilato ciclase proteína-G estimulada do lado citoplásmico da membrana, gera-se um aumento de cAMP o qual vai estimular proteínas cinases A dependentes de cAMP que por sua vez vão fosfatar diversas proteínas inclusivé sub-unidades dos microtúbulos, as quais agregando-se formam canais aquíferos (aquaporinas) na plasma membrana luminal aumentando assim a reabsorção de água.

A água é transportada através da célula renal para o seu lado basolateral e depois para a grande circulação originando uma diluição da concentração de sal (o sinal) e um aumento da pressão arterial.

As aquaporinas são uma família de proteínas integrais de membrana com seis domínios transmembranários  $\alpha$  helicoidais e quando ocorrem mutações específicas das sequências dos seus ácidos aminados podem surgir aquaporinas não funcionais e desenvolver-se a diabetes insípida caracterizada pelo aumento de sede e produção de grande volume de urina.

## Célula do tubulo renal distal



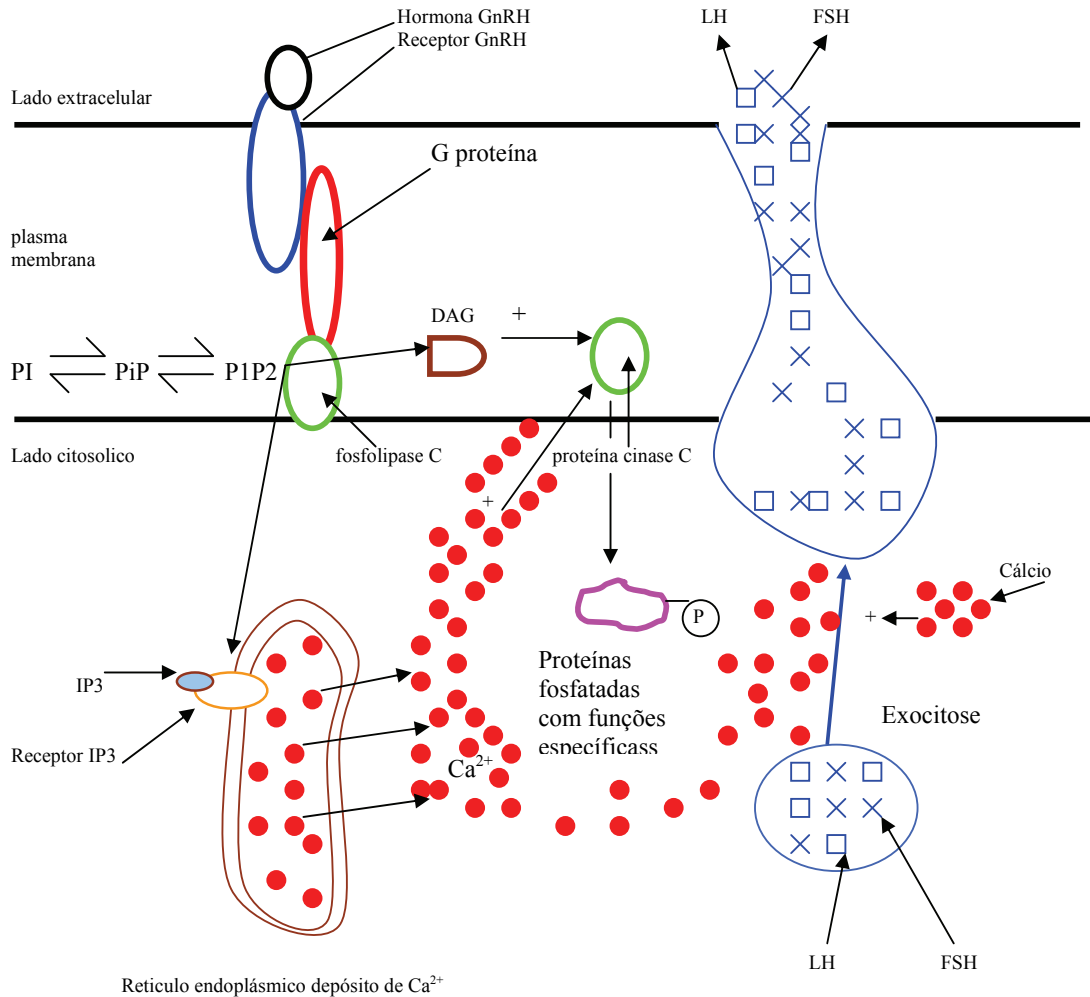
### Hormonas utilizando o c AMP como segundo mensageiro

Calcitonina  
Gonadotrofina coriónica  
Corticotrofina  
Adrenalina  
Hormona estimuladora do folículo  
Glucagina  
Lipotrofina  
Hormona luteinizante  
Hormona estimuladora dos melanocitos  
Nor-adrenalina  
Hormona da paratiróide  
Hormona estimuladora da tiróide  
Vasopressina

### Proteínas cinases C

Acção intracelular da proteína cinase C: exemplo da hormona libertadora Gonadotrofina (GnRH).

Este sistema hormonal actua através da estimulação da via do fosfatidilinositol e com a subsequente activação do sistema proteína cinase C.



A hormona libertada a partir das terminações nervosas (no nosso exemplo a GnRH) liga-se a receptores na membrana da célula gonadotrofa e o sinal do complexo hormona-receptor é transduzido através de uma G-proteína sendo activada uma fosfolipase C a qual catalisa a hidrólise do PIP2 para DAG e IP3- O DAG (diacilglicerol) activa a proteína cinase C a qual fosforila proteínas específicas algumas das quais podem intervir no processo segregador para transportar LH e FSH para fora da célula.

O IP3 por seu turno liga-se a receptores da membrana do retículo endoplasmático de depósito de cálcio e estimula a libertação deste.

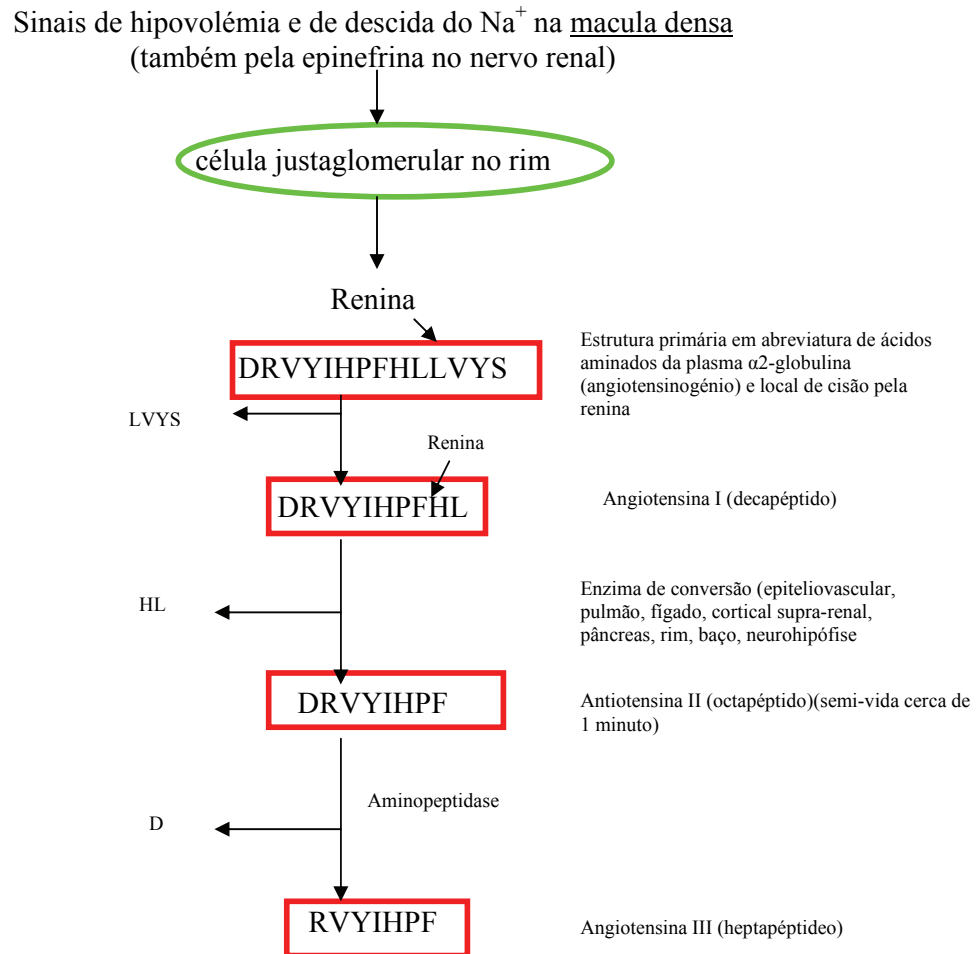
A elevada concentração de  $Ca^{2+}$  no citosol estimula a proteína cinase C e ajuda a exocitose da LH e FSH.

Também a síntese e segregação da aldosterona é bastante complexa e estimulada também pela proteína cinase C (vide figuras adiante).

A sua principal força impulsionadora é a angiotensina II produzida pelo sistema renina-angiotensina esquematizado seguidamente.

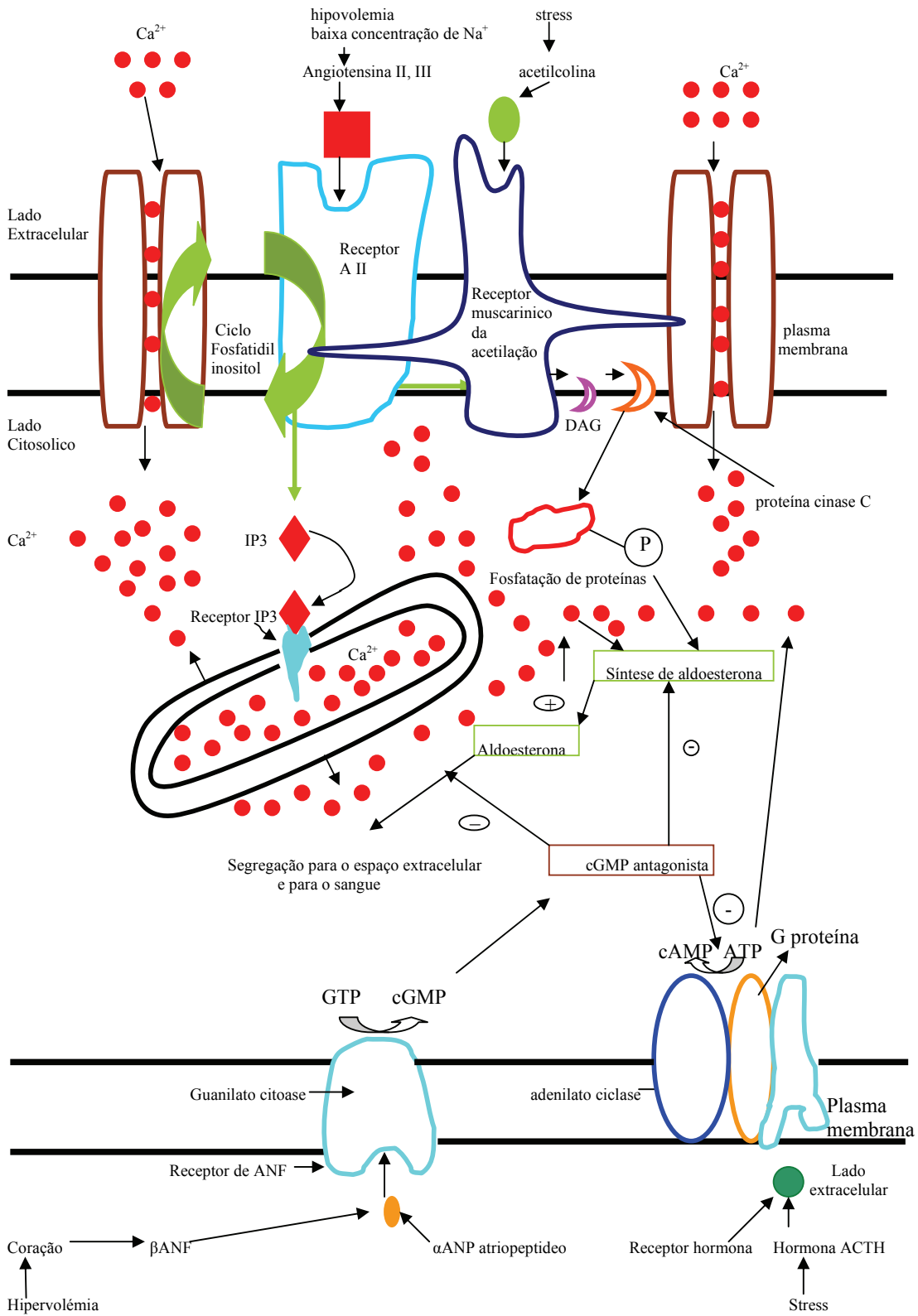
A angiotensina II e III (A II e A III) derivam de proteínas em circulação no sangue por acção da renina e da enzima conversora.

A renina inicialmente responde à diminuição do volume de sangue ou diminuição do  $[Na^+]$  na macula densa do rim. A A II / A III estimulam as células mais externas das supra-renais sintetizando e libertando a aldosterona.



O sinal (angiotensina III ou AIII) para a síntese e segregação da aldosterona é gerado quando é necessário aumentar a pressão sanguínea (volume de sangue) e concentração de  $Na^+$  no sangue. A  $\alpha$ 2-globulina (angiotensinogénio) em circulação no sangue é cindido pela renina e por uma aminopeptidase. Tanto a angiotensina II como a III se podem ligar ao receptor da angiotensina indicado na figura seguinte e quando isto sucede é activado o ciclo fosfatidilinositol produzindo IP<sub>3</sub> e DAG (vide figura). O IP<sub>3</sub> estimula a libertação de iões cálcio a partir dos seus depósitos intracelulares, tal como é estimulada a actividade dos canais de  $Ca^{2+}$  o que aumenta consideravelmente a concentração citoplásmica de  $Ca^{2+}$ .

Reacções levando à segregação da aldosterona nas células da zona glomerulosa da supra-renal

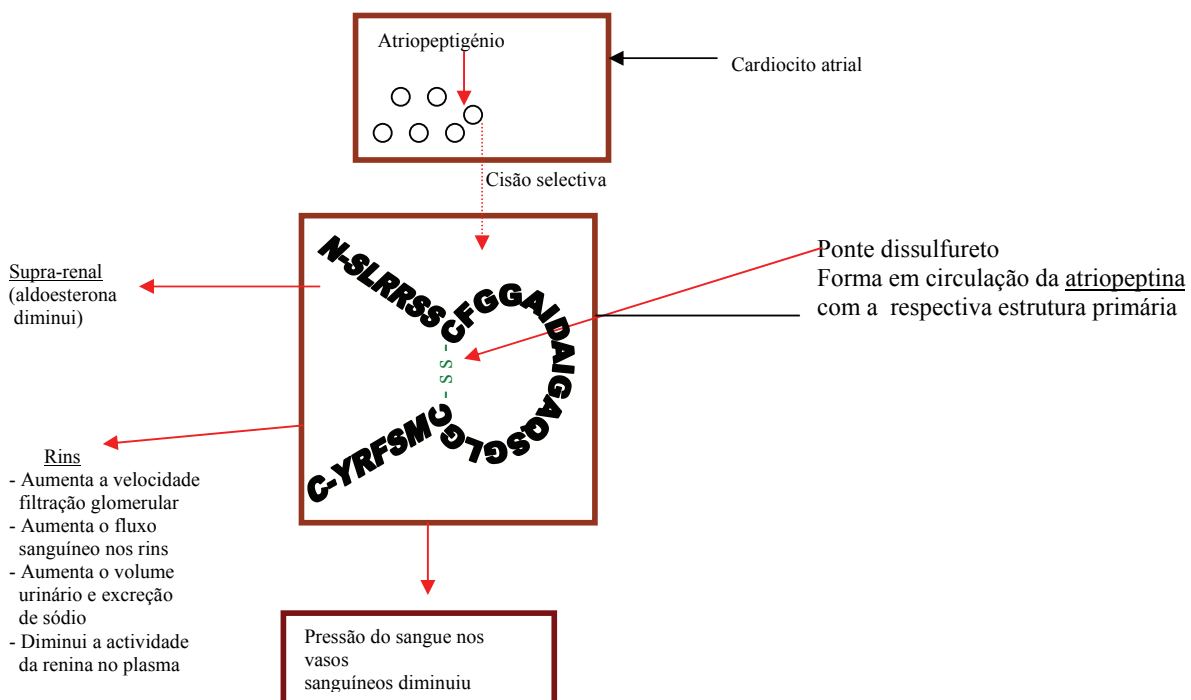


Esta maior concentração de cálcio em conjunto com o diacilglicerol (DAG) estimula a proteína cinase C.

Também a acetilcolina libertada pelos neurónios sob stress, têm um efeito similar.

A proteína cinase C activada leva a fosfatação de proteínas que induzem um aumento de concentração de aldosterona que é depois segregada para o sangue e chegando ao rim (células distais) liga-se ao seu receptor citoplásmico e estimula a expressão de proteínas que aumentam o transporte de  $\text{Na}^+$  a partir do filtrado glomerular para o sangue.

Sinais opostos aos desencadeados pela angiotensina são os produzidos pelo ANF (factor atrial natriurético) ou atriopeptina formada no átrio cardíaco como se esquematiza seguidamente.



O ANF ao ligar-se ao seu receptor específico na membrana celular da zona glomerulosa da supra-renal activa a guanilato ciclase (proteína cinase G - vide adiante) que faz parte da mesma cadeia polipeptídica do receptor e aumenta a concentração de cGMP que antagoniza a síntese e segregação de aldosterona, o mesmo sucedendo com a formação de cAMP pela adenilato cidase, induzida pela ACTH.

A aldosterona è pois também uma hormona de stress.

Péptidos natriuréticos são uma família de hormonas peptídicas que têm propriedades excretoras de sódio (natriureticas) e hipotensoras.

Nesta família incluem-se o ANF ou ANP, o BNP isolado inicialmente de um cérebro suíno e o CNP ou péptido natriurético do tipo C.

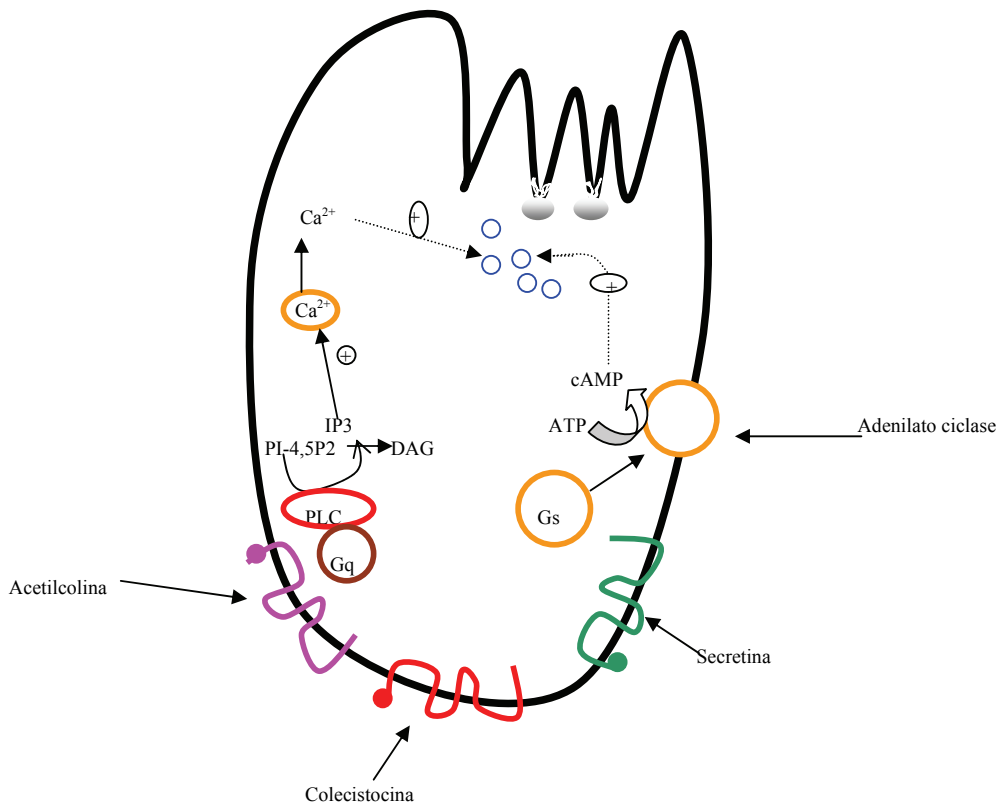
O ANP tem 28 resíduos de ácidos aminados, o BNP 32 e o CNP22.



O gene codificador de cada péptido natriurético produz uma pro-hormona, ou precursor que é subsequentemente cindido para produzir um fragmento activo. Assim o ANP é inicialmente sintetizado com 126 ácidos aminados sendo este precursor cindido produzindo um fragmento N-terminal de 96 ácidos aminados e um fragmento C-terminal de 28 ácidos aminados, sendo este último fragmento biologicamente activo. O BNP precursor tem 108 ácidos aminados sendo depois cindido produzindo um fragmento C-terminal de 32 ácidos aminados biologicamente activo.

Uma propriedade estrutural comum destes péptidos activos é a existência de uma ponte dissulfureto que origina uma estrutura de tipo cíclico envolvida na ligação destes péptidos aos seus receptores específicos (NPRs). Estes receptores transmembranários têm um domínio extracelular onde se liga a hormona e um domínio intracelular “cinase-like” seguido por um domínio guanilil ciclase.

Outro exemplo de activação pelo aumento da concentração de cálcio e DAG é a regulação celular da segregação exócrina no pâncreas



Diferentes células exócrinas regulam as suas segregações exócrinas em diferentes glândulas e para isso possuem diferentes conjuntos de receptores e após a ligação de efectores adequados (acetilcolina, amins biogénicas, hormonas neurotransmissoras peptídicas), desencadeiam uma cadeia de acontecimentos assinalantes que terminam com a fusão de grânulos de zimogénios com a plasma membrana, estando assinaladas duas vias principais para este efeito (vida figura acima).

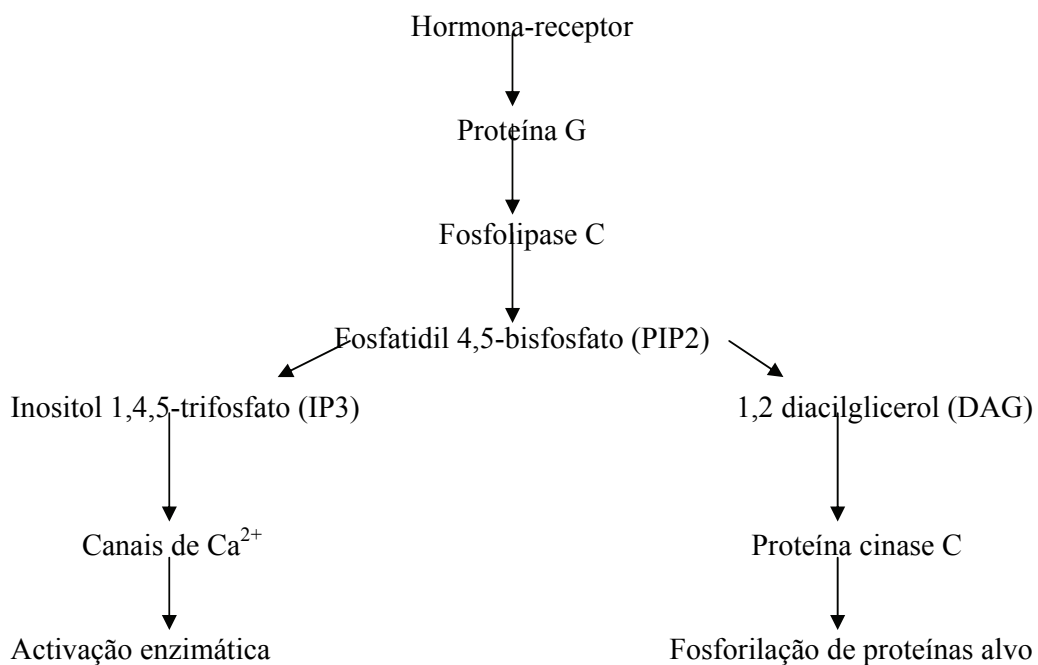
A activação da fosfolipase C (PLC) específica do fosfatidilinositol (PI-4,5 P2) liberta diacilglicerol (DAG) e IP3 que desencadeiam a libertação de Ca<sup>2+</sup> no citosol e a activação da proteína cinase C.

Por outro lado a activação da adenilato ou guanilato ciclase, eleva as concentrações de cAMP ou cGMP, podendo as segregações serem estimuladas por esta via ou pela anterior.

Os efeitos mediados pela cascata fosfoinositideo podem ser:

- Glicogenólise nas células hepáticas
- Segregação de histamina pelas “mast cells”
- Libertação de serotonina pelas plaquetas sanguíneas
- Agregação das plaquetas sanguíneas
- Segregação de insulina pelas células das ilhotas pancreáticas
- Segregação da adrenalina pelas células cromafínicas da supra-renal
- Contração da musculatura lisa
- Transdução visual nos fotoreceptores dos invertebrados

Cascata de fosfoinositois



As intervenções metabólicas do cálcio ionizado podem ser as seguintes:

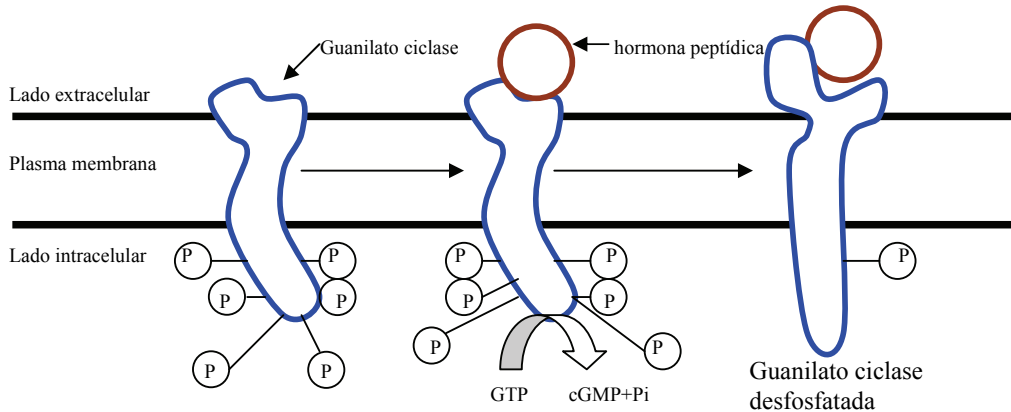
- Transporte transmembranário
- Mediador intracelular de acção hormonal
- Actividade enzimática
- Excitabilidade neuromuscular
- Coagulação do sangue
- Mecanismos de segregação celular
- Descarga de hormonas e neurotransmissores
- Mineralização óssea

## Proteínas cinases G

### Acção intracelular

#### Proteína cinase G: exemplo da actividade do factor atrial natriurético (ANF).

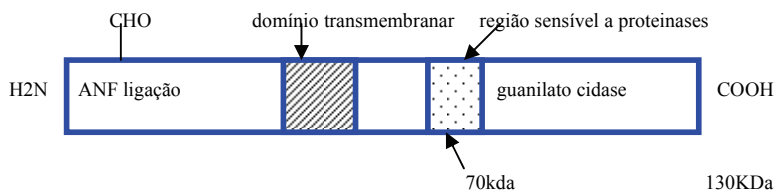
Nestes sistemas a proteína cinase G é estimulada pela elevação citosólica do cGMP pela guanilato ciclase.



A guanilato ciclase está ligada à biomembrana plasmática e pode receber sinais biológicos específicos através da porção receptora que é ao fim e ao cabo o domínio extracelular da ciclase que está directamente acoplado ao domínio citosólico.

O receptor liga-se ao ligando ANF, e temos assim na mesma cadeia polipeptídica o local de ligação a hormona, o domínio transmembranar e o domínio com actividade de guanilato ciclase.

Os domínios funcionais do receptor do ANF referem-se seguidamente:



O ANF é libertado por uma série de sinais e é segregado como um dímero que é inactivo, tendo que ser convertido no plasma sanguíneo num monómero capaz de interaccionar com o receptor.

O ANF aumenta a velocidade de filtração glomerular sem aumentar o fluxo sanguíneo renal.

O ANF opera através dos seus receptores membranários e o cGMP produzido activa proteínas cinases G que depois fosfatam diversas proteínas celulares com diversos tipos de acções que são no entanto muito mal conhecidas.

## **2.4 - Transdução de sinais e receptores membranários à superfície das células**

### **2.4.1 - Tipos de receptores membranários**

#### Sinais de célula para célula; hormonas e receptores

Na sinalização desencadeada a partir do meio extracelular, os sinais responsáveis operam a diversas distâncias nos animais, indo esse sinais interactuar com proteínas receptoras com as quais interactuam de forma muito específica, desencadeando cada tipo e ligando uma especificidade correspondente.

De entre estes sinais avultamos hormonas que consoante a sua solubilidade e a localização dos seus respectivos receptores assim são classificadas. As hormonas lipossolúveis interactuam especificamente com receptores de localização intracelular enquanto as hormonas hidrossolúveis interactuam com receptores localizados á superfície das células.

As diversas hormonas desencadeiam os seus respectivos efeitos através do concurso ou melhor com a intervenção de uma série de segundos mensageiros tais como o cAMP (3',5'AMP cíclico) o cGMP (3',5'GMP cíclico) o 1,2diacilglicerol (DAG) o inositol 1,4,5 trifosfato(IP3) e o cálcio(Ca<sup>2+</sup>).

No que se refere aos receptores de sinais com sede á superfície das células eles podem ser distribuídos em quatro classe principais:

- 1) os receptores ligados a G-proteínas
- 2) os receptores que são canais de iões
- 3) os receptores que não possuem actividade catalítica intrínseca mas que estão associados com proteínas citosólicas essas sim dotadas de actividade tirosina cinasica.
- 4) os receptores que possuem actividade catalítica intrínseca.

Mais tarde abordaremos a classificação dos receptores de sinais com localização intracelular.

Na classe 1) de receptores com sede à superfície das células o ligando sinal (como por exemplo a adenalina, serotonina e glucagina) interactuando com o receptor respectivo vai promover a activação de uma G-proteína que por seu turno pode activar ou inibir, na cascata de transdução de sinal induzida, uma proteína enzimática responsável pela formação de um segundo mensageiro específico, ou então condicionando a actividade de um canal iónico e o respectivo potencial de membrana.

Na classe 2) destes receptores de que é exemplo o receptor que interactua com a acetilcolina na junção neuromuscular, a interacção sinal-receptor canal iónico modifica a configuração deste induzindo a passagem de iões específicos através destes canais com alteração do potencial eléctrico.

Na classe 3) de que são exemplo os receptores da superfamilia das citocinas, receptores estes sem actividade catalítica, mas associados com proteínas citosólicas com actividades de tirosina cinases, a interacção desta classe de receptores com o respectivo ligando sinal (por exemplo citocinas, interferons, factor humano de crescimento, etc)promove a dimerização dos receptores o que desencadeia por seu turno a activação de uma ou mais proteínas tirosina cinases citosólicas.

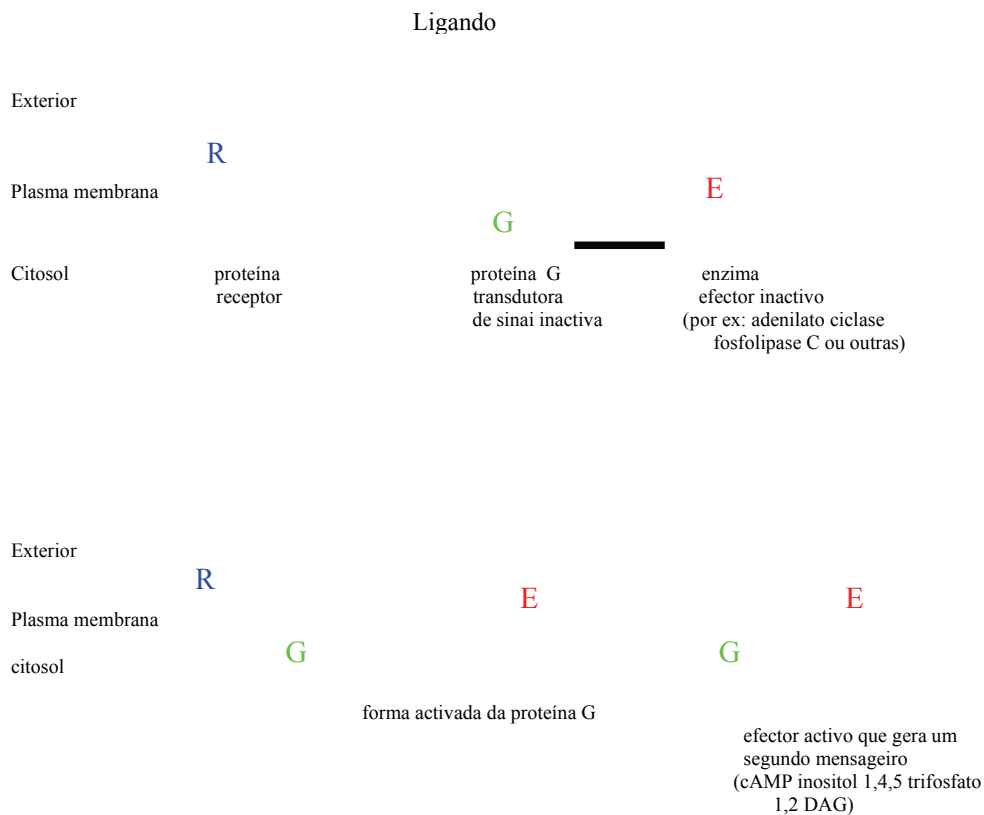
Na classe 4) dos receptores que possuem actividade catalítica intrínseca, existem vários tipos. Um primeiro tipo é aquele em que alguns ligandos activam uma actividade de guanilato ciclase e num outro

tipo é activada uma proteína fosfatase com sede no domínio citosólico do receptor. Ainda nesta classe 4) se englobam os receptores específicos para interacção com a insulina e com diversos factores de crescimento, qualquer um deles dimerizando pela interacção ligando-receptor o que activa a serina/treonina cinase ou tirosina cinase que têm sede no próprio receptor, sendo estes receptores por vezes designados por RTK.

Ocorre portanto neste caso a fosfatação de resíduos de ácidos aminados do próprio receptor situados no domínio citosólico deste ou seja a autofosfatação dos RTK. Quando isto sucede são criados locais para docagem de diversas proteínas citosólicas as quais modificam as suas actividades podendo depois gerar segundos mensageiros nas cascatas de transdução de sinais.

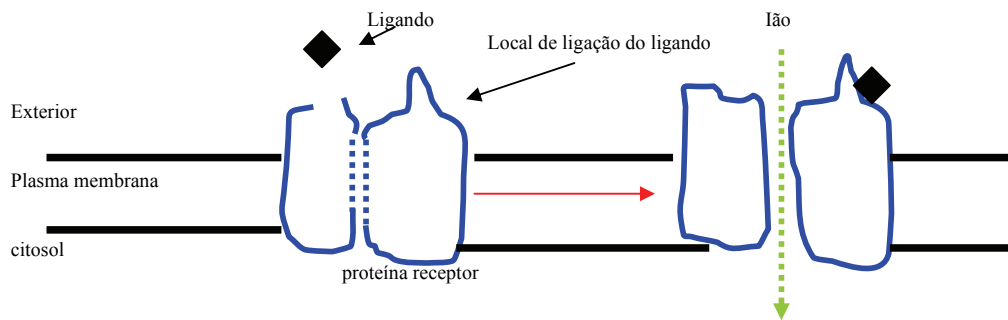
Nas figuras seguintes mostra-se a interacção de vários tipos de sinais ligandos com diversos tipos de receptores da superfície celular.

1. Receptores ligados a proteínas G (ligandos comuns accionando este tipo de receptores, epinefrina, glucagina, serotonina)



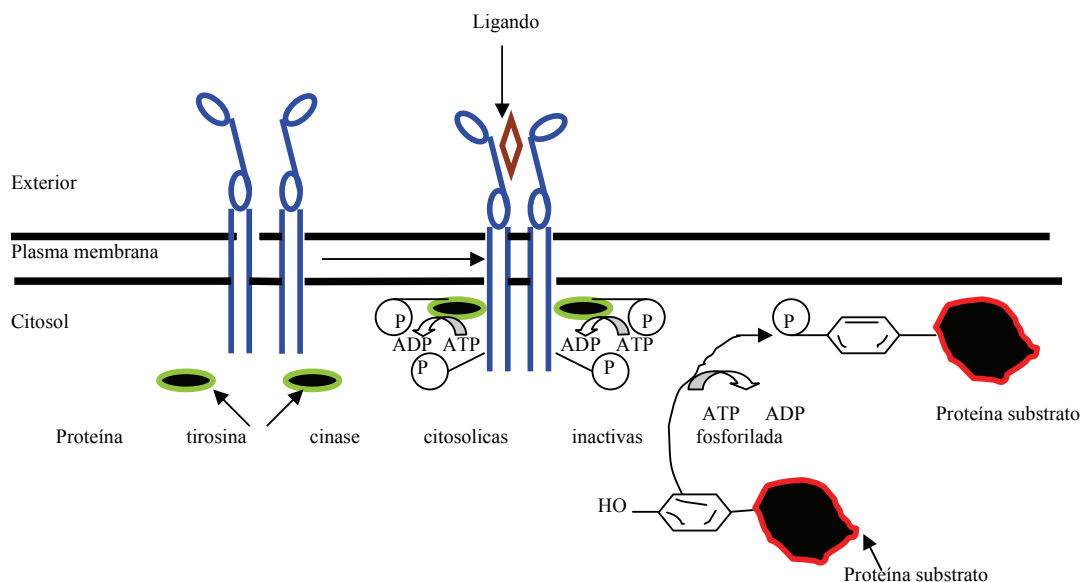
A união do ligando ao receptor induz a activação da proteína G a qual tende depois a ligar-se e activar uma enzima que cataliza a síntese de um segundo mensageiro específico.

2. Receptores canais iónicos (accionados por ligando do tipo acetilcolina)



Neste tipo de receptor uma alteração da sua conformação induzida pela união com o ligando abre o canal permitindo a passagem de um fluxo de iões.

3. Receptores ligados a tirosina cinases que existem independentes no citossol (accionados por ligandos do tipo eritropoietina, interferons)



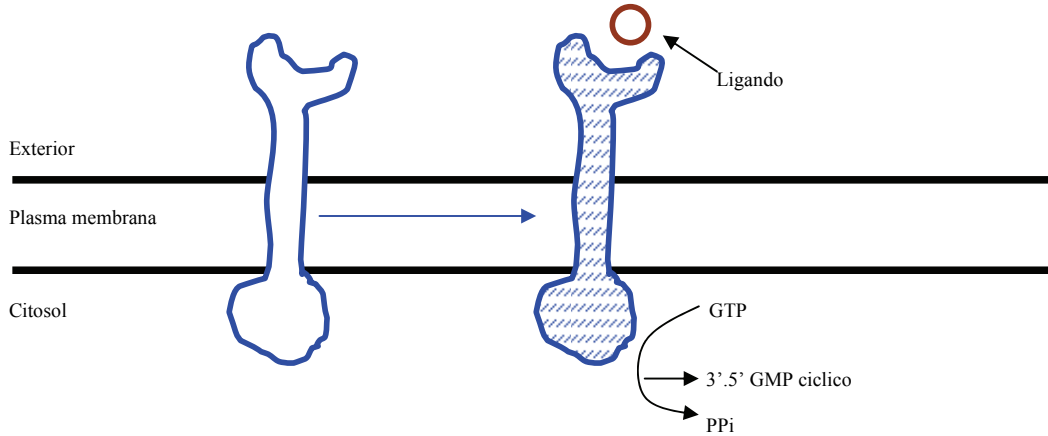
Neste tipo de receptor o ligando origina a formação de um homodímero ou de um heterodímero induzindo a ligação e activação de uma proteína tirosina cinase citosólica.

Esta enzima activada por seu lado fosfata resíduos de tirosina no receptor. Proteínas substrato ligam-se depois a estes resíduos de fosfotirosinas e são assim fosfatadas.

A eritropoietina do rim actua sobre a medula óssea para a diferenciação terminal e inicia a síntese da hemoglobina. Os interferons são um grupo de pequenas proteínas libertadas pelos macrófagos após estimulação ou por várias células após infecção por vírus e que se ligam a receptores ligados a tirosina cinases em células alvo induzindo alterações na expressão dos genes levando a um estado antiviral (IFN $\alpha$ , IFN $\beta$ , e IFN $\gamma$ ) ou a outras respostas imunitárias (IFN- $\gamma$ ).

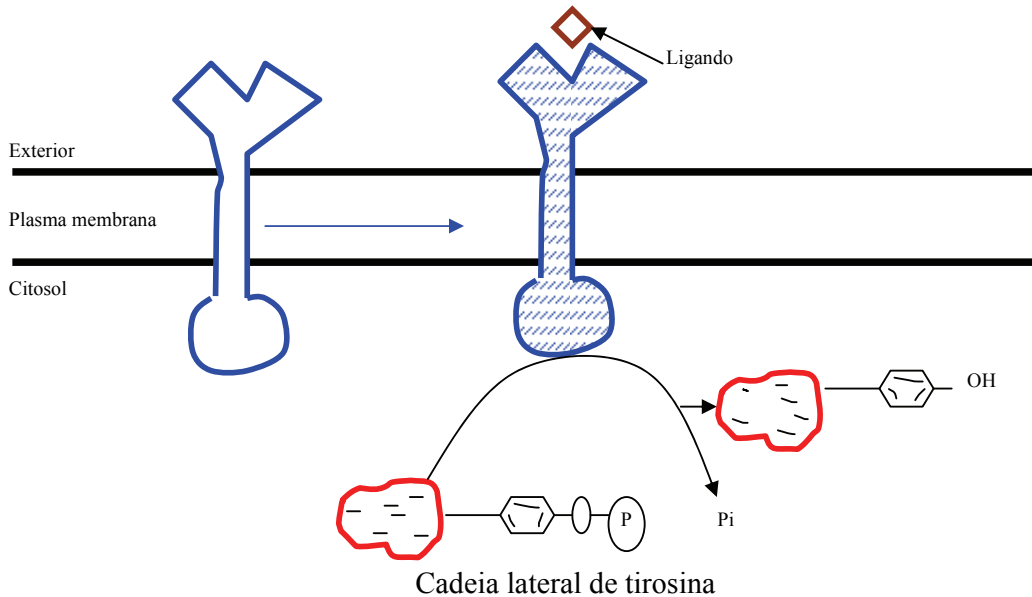
#### 4. Receptores com actividade enzimática no seu domínio citosólico desencadeada pelo ligando

4.1 Exemplifica-se com o receptor guanilato ciclase (activado pelo ligando factor atrial natriurético (ANF) ou atriopeptina que são uma família de péptidos) que relaxam a musculatura lisa vascular e produzem uma vasodilatação e natriurese bem como outros efeitos. Actua sobre as células adrenais exteriores diminuindo a libertação da aldosterona e tem outros efeitos. O ANF é libertado do atria do coração em resposta a hipovolémia; é regulada por outras hormonas.



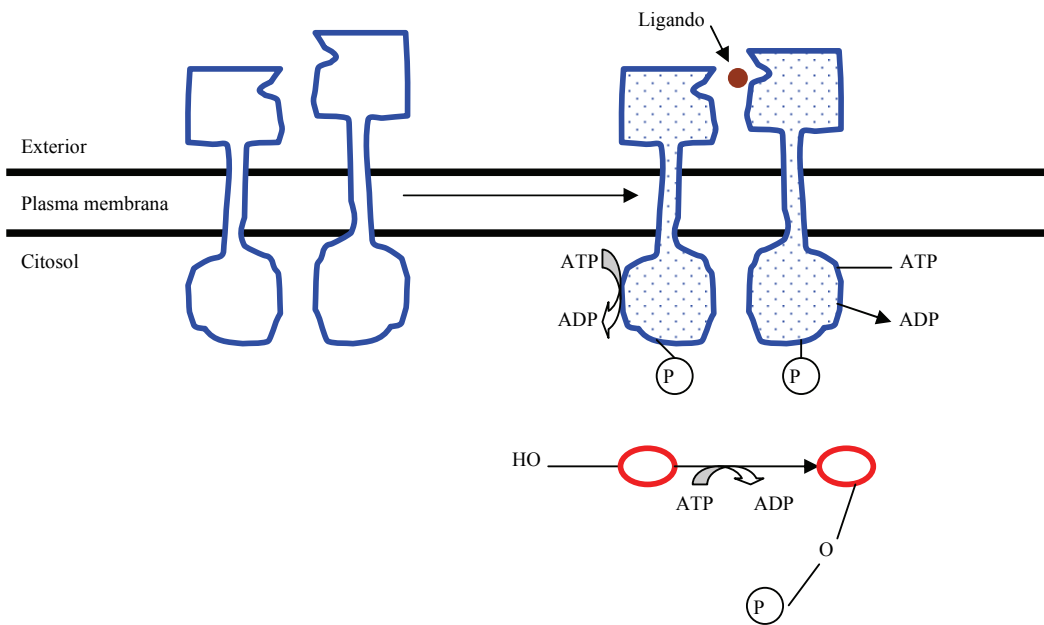
Neste caso o receptor é um monómero que depois de interaccionar com o ligando (hormona neste caso) é activado e funciona como guanilato ciclase gerando um segundo mensageiro ou seja o c GMP.

#### 4.2 Receptor tirosina fosfatase (accionado pela proteína CD45 dos leucocitos).



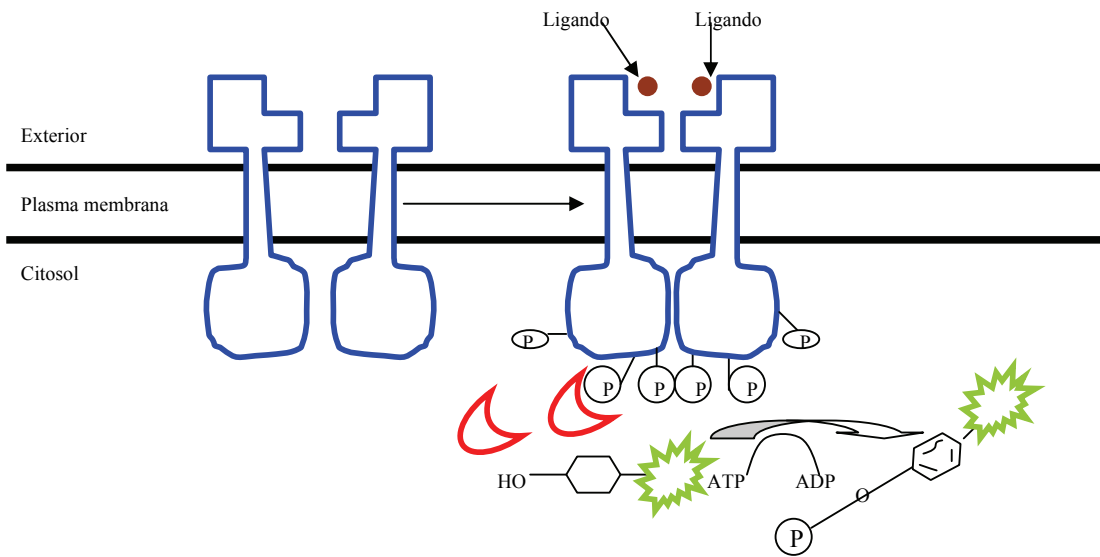
A interacção do receptor com o ligando activa a actividade de Tirosina fosfatase, podendo estes receptores removerem grupos fosfato de resíduos de fosfotirosina em proteínas substrato, modificando assim a actividade destes.

4.3 – Receptores heterodímero serina / treonina cinases accionados pelo factor de crescimento transformante  $\beta$ .



Os receptores para vários factores de crescimento têm proteínas com actividade cinasica intrínseca. A união com o ligando origina dimerização entre monómeros receptores idênticos ou não idênticos activando assim a sua actividade enzimática. No caso da figura acima os receptores activados com actividade cinasica serina treonina são heterodímeros.

4.4 – Receptores homodímeros Tirosina cinase accionados pela insulina e factor de crescimento epidérmico.



Estes receptores com actividade Tirosina cinasica são homodímeros.



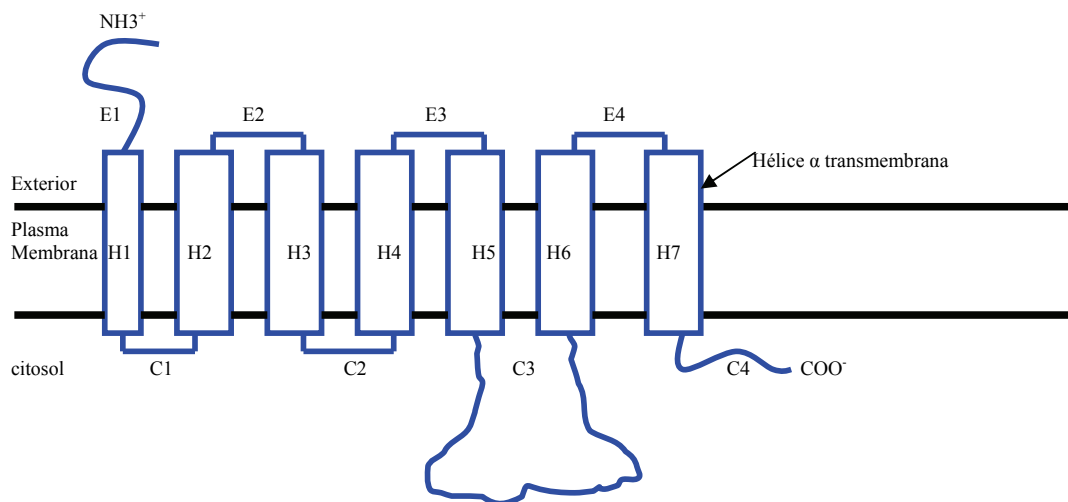
Nos dois tipos de receptores referidos em 4.3 e 4.4 o receptor dimérico activado fosfata diversos resíduos no seu próprio domínio no lado citosólico mas enquanto os receptores com actividades Tirocína cinase podem também fosfatar certas proteínas substracto alterando a actividade destas, em relação aos receptores com serina / treonina cinases, não se sabe quando estas fosfatam algumas proteínas substracto específicas.

#### 2.4.1.1 - Transdução de sinais e receptores de sete hélices

### Receptor de sete hélices

Este tipo de receptores de sete hélices constituem uma classe importante de receptores da superfície celular que interactuam com G-proteínas.

No esquema seguinte representa-se a estrutura geral destes receptores de sete hélices. Todos os receptores deste tipo contêm sete cadeias transmembranárias daí a sua designação de receptores de sete hélices



E1 a E4 = ansas extracelulares  
H1 a H7 = hélice transmembranárias  
C1 a C4 = ansas citosólicas

A sequência de ácidos aminados deste polipéptido receptor está contida em sete hélices, cada uma delas contendo cerca de 22-24 resíduos de ácidos aminados hidrófobos, que se constituem em hélices  $\alpha$  transmembranárias.

A unir estas diversas hélices  $\alpha$  transmembranárias ou nas extremidades N e C terminais do polipéptido existem no entanto estruturas em ansa importantes. Assim as hélices numeradas de H1 a H7 na figura anterior apresentam entre as hélices 5(H5) e 6(H6) uma ansa (C3) que tal como o segmento C-terminal (portanto do lado citosólico) são importantes para o receptor interactuar com as proteínas G, proteínas G que são as transdutoras de sinal que interactua com o receptor funcionando pois como um interruptor biológico molecular ligando e desligando a via transdutora de sinal, estando no estado desligado quando se ligam ao GDP.

Os receptores  $\beta$  adrenérgicos que são receptores de sete hélices que podem interagir com as catecolaminas, sobretudo a norepinefrina, possuem dois subtipos, os  $\beta_1$  e os  $\beta_2$  que diferem na sua afinidade para a norepinefrina e para antagonistas sintéticos. Assim os  $\beta_1$  têm maior afinidade para a norepinefrina do que para a epinefrina sucedendo o inverso com os receptores  $\beta_2$  adrenérgicos.

No entanto estes dois subtipos de receptores de sete hélices  $\beta_1$  e  $\beta_2$  revelam uma profunda analogia.

Explorando ainda as características estruturais dos receptores de sete hélices na figura anterior verifica-se que a partir da hélice 1(H1) se destaca para o espaço extracelular a porção N-terminal do polipéptido, enquanto do lado do citosol se formam ansas entre as hélices 1 e 2(H1 e H2) hélices 3 e 4(H3 e H4) e hélices 5 e 6(H5 e H6) havendo depois uma extensa porção C-terminal também citosólica que se destaca da hélice H7(H7). Esta extensa porção C-terminal contém resíduos de serina e treonina que são locais que podem ser fosfatados o que se revela muito importante na regulação da actividade destes receptores envolvendo a sua desensibilização.

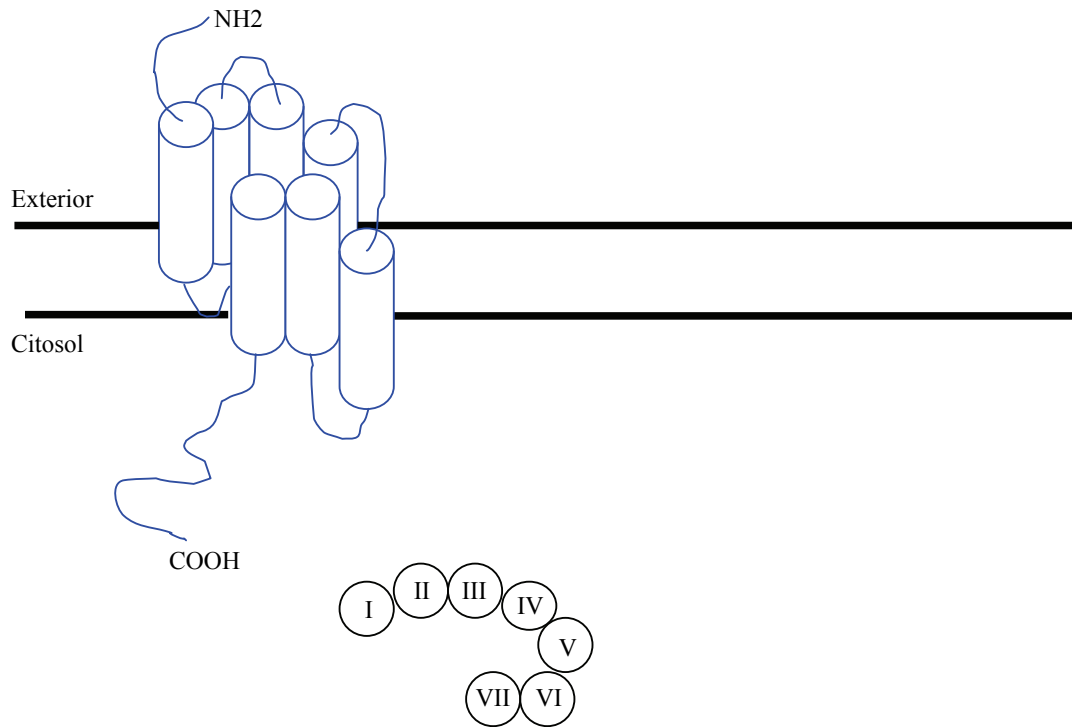
Com efeito estas fosfatações na cauda citoplasmática do receptor induzem a interacção com uma proteína inibidora a  $\beta$  arrestina que impedem portanto a activação pelos receptores da subunidade Gs das proteínas G.

Ainda na figura anterior, bem como na figura seguinte verificam-se ansas unindo do lado extracelular algumas hélices transmembranárias do receptor. Assim entre as hélices 2 e 3, 4 e 5, e 6 e 7 ocorrem ansas peptídicas, que segundo estudos realizados com mutantes não parecem interagir com o ligando sinal, parecendo que este interage com uma espécie de bolsa formada pela disposição dos cilindros correspondentes às hélices de 1 a 7 (ver figura seguinte).

No entanto parece que o domínio transmembranário 6 do receptor intervém na activação da adenilato ciclase um elemento da cadeia de transdução de sinal como veremos adiante.

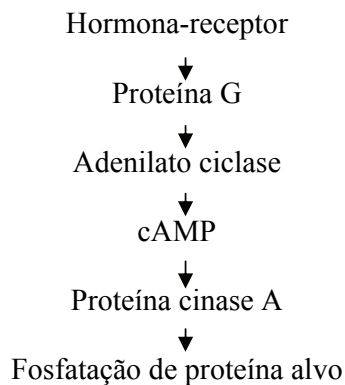
Como é evidente alterações na estrutura primária destes receptores, na medida em que estes possuem características específicas, pode alterar as suas funções. Contudo é conhecido que a substituição de resíduos de cisteína podem não perturbar a capacidade de interacção do receptor com o ligando embora diminua a activação da adenilato ciclase.

Arranjo proposto para as sete hélices transmembranarias dos receptores  $\beta$ -adrenérgicos das membranas

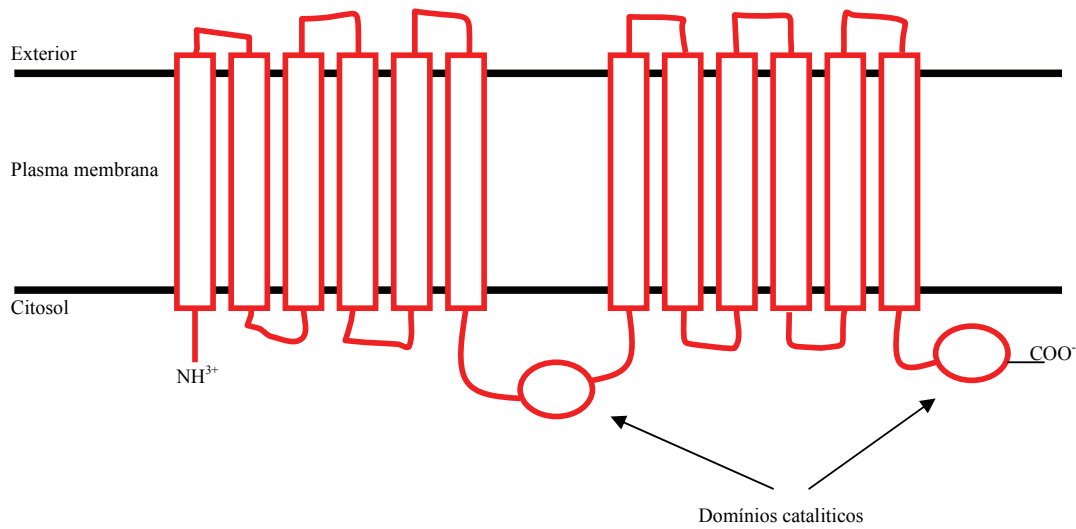


A enzima adenilato ciclase dos mamíferos é regulada pelas G proteínas heterotriméricas, mas não pelas Ras proteínas (vide adiante) que são pequenas G proteínas monoméricas constituindo estas uma superfamília com cinco famílias (Ras, Rho, Rab, Sar 1 – Arf e Ran), a família Ras regulando a expressão de genes, a família Rho regulando a reorganização do citoesqueleto e expressão de genes, as famílias Rab e Sar 1 – Arf regulando o tráfego de vesículas nas células e a família Ran regulando o transporte nucleocitoplasma e organização de microtúbulos.

**Cascata da adenilato-ciclase**



### Esquema da adenilato ciclase dos mamíferos



A proteína enzimática adenilato ciclase que também se encontra ligada á membrana celular(vide figura anterior) revela no polipéptido dois domínios dotados de actividade catalítica,situados no lado citosólico e dois outros domínios cada um constituído por seis hélices  $\alpha$  transmembranárias,embutidos na membrana celular.

Estão identificados nesta adenilato ciclase nove isoformas ou sejam múltiplas formas de uma mesma proteína diferindo ligeiramente na sua estrutura primária, mas com a mesma actividade geral.

Esta enzima adenilato ciclase pode ser activada ou inibida pelas proteínas G que ao fim e ao cabo são as transdutoras do sinal desencadeado por este no receptor.

Sabe-se que no cérebro se encontra uma isoforma da adenilato ciclase que é activada pelos iões cálcio complexados com a proteína transportadora de cálcio, a calmodulina.

A enzima adenilato ciclase dos mamíferos é regulada pelas G-proteínas heterotriméricas, mas não pelas Ras proteínas (vide adiante) que são pequenas G-proteínas monoméricas, constituindo estas uma superfamilia com cinco famílias(Ras,Rho,Rab,Sar/Arf Ran).

### **Proteínas G e sua regulação**

Como referimos as proteínas G são transdutores de sinal que interactuam com o receptor funcionando como um interruptor biológico molecular ligando e desligando a via transdutora de sinal, estando no estado desligado quando se liga ao GDP.

A interacção do receptor com um ligando origina a libertação do GDP que estava ligado a proteína G levando esta proteína G a interactuar com o GTP o que constutue a activação desta proteína G: Após esta activação o complexo proteína G-GTP interactua e activa ou inibe certas entidades enzimáticas e estas se activadas, por sua vez originam a formação de segundos mensageiros (é o caso do cAMP, inositol 1,4,5-trifosfato,1,2-diacilglicerol) através da acção respectivamente da adenil ciclase,fosfolipases, fosfodiesterases e canais iónicos para o cálcio e o potássio).

Por seu turno a hidrólise do GTP ligado á proteína G activada, origina a inactivação deste complexo.

A título de exemplo refere-se que os sinais epinefrina/norepinefrina interactuando com receptores de sete hélices activam a proteína G transdutora do sinal a qual vai activar a enzima adenilato ciclase responsável pela formação de cAMP. Contudo podiam ser outros receptores deste tipo, outros sinais e outras proteínas G a modificar outros sistemas enzimáticos responsáveis por diferentes processos metabólicos.

Este tipo de sinal epinefrina e nor-epinefrina pode ser produzido pelas células da medular das cápsulas supra-renais e pelas células nervosas derivadas do mesmo tecido embrionário que as células da medular da supra-renal, sendo ainda segregadas pelas células nervosas diferenciadas.

Genericamente designam-se por proteínas G um numeroso grupo de proteínas heterotriméricas com capacidade para interactuar com nucleótidos de guanina, proteínas essas que habitualmente se encontram á superfície das células interactuando com receptores de sete hélices.

Os sinais hormonais servem-se pois destas G-proteínas para a transdução desses sinais.

As G-proteínas são heterotriméricas constituídas por três tipos de subunidades, as  $\alpha$ ,  $\beta$  e  $\gamma$ , cada uma delas com as suas características próprias. Assim a subunidade  $\alpha$  é aquela que interactua com o nucleótido de guanina, interactuando por outro lado indirectamente e por intermédio das subunidades  $\beta$  e  $\gamma$  com o receptor do sinal.

Esta subunidade  $\alpha$  pode existir em duas formas, uma designada  $\alpha_s$  porque é estimuladora da actividade da adenilato ciclase que se segue na cascata de transdução de sinal em que intervem e outra designada  $\alpha_i$  porque desempenha um papel inibidor neste contexto. Portanto existem dois complexos o G  $\alpha_s$  e o G  $\alpha_i$ .

No quadro seguinte referem-se uma série de subfamílias de proteínas G, sobretudo das subunidades G  $\alpha$ , bem como a expressão da sua actividade e através de que intermediários desencadeiam as suas acções efectoras e como o fazem.

São conhecidas uma série de sub-famílias de proteínas G que fazem a transdução de uma série de actividades

<u>Sub-unidade <math>\alpha</math></u>	<u>Expressão</u>	<u>Efeito efector</u>
Gs	Ubíqua	↑ Adenilato ciclase, canal $Ca^{2+}$
Golf	Olfatória	↑ Adenilato ciclase
G+1 (transducina)	Fotoreceptores bastonete	↑ cGMP fosfodiesterase
G+2 (transducina)	Fotoreceptores do cone	↑ cGMP fosfodiesterase
Gi1	Neural > outros tecidos	↓ Adenilato ciclase
Gi2	Ubíqua	
Gi3	Outros tecidos > neural	
Go	Neural, endócrino	↓ Canal $Ca^{2+}$
Gq	Ubíqua	↑ Fosfolipase C
G11	Ubíqua	
G14	Fígado, pulmão, rim	
G15/16	Células do sangue	

Vejamos como alguns ligandos sinais desencadeiam determinados efeitos metabólicos.

A glucagina da circulação sanguínea ao interactuar com os receptores hepáticos que interactuam com proteínas G, promovem a glicogenólise hepática. A sua interacção com os receptores da plasma membrana das células hepáticas activa a adenilato ciclase e esta por acção mediada pelo cAMP activa a enzima glicogénio fosfatase e inactiva a enzima glicogénio sintase, o que origina a hidrólise do glicogénio, inibindo por outro lado a glicólise ao nível de dois sistemas enzimáticos, da 6-fosfofructo-1-cinase e da piruvato cinase.

Da mesma forma a epinefrina promove a glicogenólise hepática. Os receptores  $\beta$  adrenérgicos ao interactuarem com este sinal, activam a adenilato ciclase e consequentemente o teor em cAMP, do que resulta a inibição da síntese de glicogénio e da glicólise.

No entanto ao nível da célula hepática se forem outros os tipos de receptores a interactuarem com a epinefrina, por exemplo os receptores  $\alpha$  adrenérgicos, a cascata de transdução de sinal opera de forma diferente com a produção de inositol 1,4,5 trifosfato (IP3) e diacilglicerol (DAG) pela catálise pela fosfolipase C sobre o fosfatidilinositol 4,5 bifosfato contido na plasma membrana hepática.

Aquele segundo mensageiro o IP3 aumenta a concentração do cálcio ( $Ca^{2+}$ ) citosólico devido á sua libertação do retículo endoplásmico e daí activação da fosfatase cinase e depois da glicogénio fosfatase resultando na glicogenólise.

Esta interacção da epinefrina com receptores  $\beta$  adrenérgicos pode ocorrer em variadíssimos tipos de tecidos com a cascata de transdução de sinal antes referida, com aumento do cAMP, com diversas respostas metabólicas.

Respostas metabólicas variadas de diversos tecidos ao aumento de cAMP induzido por diversos tipos de hormonas

<u>Tecido</u>	<u>Hormonas indutora da subida de cAMP</u>	<u>Resposta metabólica</u>
Adiposo	Epineprina, ACTH glucagina	Aumento da hidrólise de triacilglicerois; diminuição da absorção de ácidos aminados
Fígado	Epinefrina, nor- epinefrina, glucagina	Aumento da glicogenólise; inibição da síntese do glicogenio; aumento da tomada de A. A. aumento da gluconeogénese
Foliculo dos ovários	FSH; LH	Aumento da síntese de estrogénio, progesterona
Córtex supra-renal	ACTH	Aumento da síntese de aldosterona, cortisol
Células músculo cardíaco	Epinefrina	Aumento da velocidade de contracção
Tiróide	TSH	Segregação de tiroxina
Células ósseas	Horm. Paratiroide	Aumento da reabsorção do cálcio do osso
Músculo esquelto	Epinefrina	Glicogenolise
Intestino	Epinefrina	Segregação de fluidos
Rim	Vasopressina	Reabsorção de água
Plaquetas sanguíneas	Prostaglandina I	Inibição da agregação e segregação

Todos os receptores de sete hélices têm estruturas tridimensionais muito semelhantes, possuindo regiões análogas para interagirem com as proteínas G e com os vários tipos de ligandos. Contudo a estrutura primária destes receptores proteicos são muito desiguais o que é extraordinariamente importante. Cita-se o caso dos receptores  $\beta_1$  e  $\beta_2$  adrenérgicos anteriormente referidos serem apenas iguais em 50% da suas sequências e as diferenças entre os receptores  $\beta$  e  $\alpha$  adrenérgicos ainda são maiores.

Como se compreende são estas sequências específicas em ácidos aminados de cada receptor que determinam qual o tipo de ligando ou de proteína G com que pode interagir cada receptor.

Detenhamo-nos agora sobre a regulação das concentrações do AMP cíclico (cAMP).

Já vimos anteriormente como a interacção de diversos sinais com receptores de sete hélices pode activar a adenilato ciclase e conseqüentemente subir a concentração de cAMP. Mas pode ocorrer uma situação inversa ou seja a diminuição de cAMP através da sua hidrólise pelas enzimas fosfodiesterases. Estas fosfodiesterases podem por seu turno ser activadas pelo aumento de  $Ca^{2+}$  citosólico que já referimos anteriormente.

Sucedo assim que algumas células podem regular a concentração de cAMP através da segregação deste metabolito para o espaço extracelular.

Pode portanto afirmar-se que o cAMP é um sinal transitório sendo produzido pela adenilato ciclase e reduzido pela fosfodiesterase, por inibidores da proteína cinase A (PKA) ou por fosfatases.

Como vimos no quadro anterior os efeitos metabólicos da subida do cAMP intracelular depende do tipo de tecidos e células.

Com efeito o cAMP bem como outros segundos mensageiros influenciam as actividades metabólicas ao interactuarem com proteínas cinase específicas as quais por seu turno modificam especificamente a actividade de diversas enzimas de vários tipos de células. A transferência de fosfatos provenientes do ATP por acção da proteína cinase específica para resíduos de ácidos aminados das proteínas enzimáticas substrato pode activar ou inibir a actividade destas.

É bem conhecido o caso da proteína cinase A (PKA) constituída por quatro subunidades sendo neste tetramero duas subunidades catalíticas (C) e duas outras subunidades reguladoras (R). Quando o cAMP se liga á PKA altera-se a estrutura deste complexo, dissociando-se as duas subunidades R da duas subunidades C e estas ultimas, livres, podem agora fosfatar especificamente determinadas proteínas substrato condicionando a actividade destas.

Outros segundos mensageiros que não o cAMP podem activar proteínas cinases por um mecanismo idêntico.

A justificação para este comportamento reside no facto deste tipo de cinases activadas por segundos mensageiros, serem inibidas na sua actividade devido a interacção de uma sequência peptídica com o seu centro activo catalítico, podendo aquela sequência peptídica fazer parte de uma subunidade reguladora como sucede no caso das cinases dependentes do cAMP) ou então fazer parte de um domínio regulador dentro da própria cadeia polipeptídica que constitui a cinase.

Pode dizer-se de uma forma geral que os diversos efeitos desencadeados pelas proteínas cinases que são activados pelo cAMP dependem dos tipos de proteínas substrato que são fosfatadas pelas cinase activadoras.

Por outro lado há que considerar o efeito de cascata com a ampliação sucessiva a partir do sinal ligando das acções desencadeadas nas várias etapas.

Todas as fosfatações e desfosfatações anteriormente referidas podem constituir ou estar integradas numa série de reacções em cascata ou sejam reacções em sequência em que o produto de uma reacção anterior vai activar (ou inibir) a reacção seguinte.

Uma cascata de reacção é um local óptimo para regular as reacções que a integram, bem como os seus efeitos metabólicos. Esta regulação pode ser feita por um simples tipo de molécula e a ampliação pode ser desencadeada por um pequeno sinal inicial.

A propósito desta ampliação de sinal inicial cita-se o exemplo clássico da epinefrina no sangue que em baixas concentrações (dez elevado a menos dez M), pode desencadear este sinal a glicogenólise levando a aumentos da glicémia de 50%. Este estímulo pela epinefrina nestas concentrações produz uma concentração intracelular de cAMP de dez elevado a menos 6 M ou seja uma ampliação de dez elevado a 4. Como no nesta cascata podem depois intervir outras etapas tais como a activação da proteína cinase, e depois a interacção desta com proteínas substrato varias, o efeito da ampliação pode ser consideravelmente aumentado de etapa para etapa pois as moléculas produzidas numa reacção activam varias moléculas da reacção seguinte.

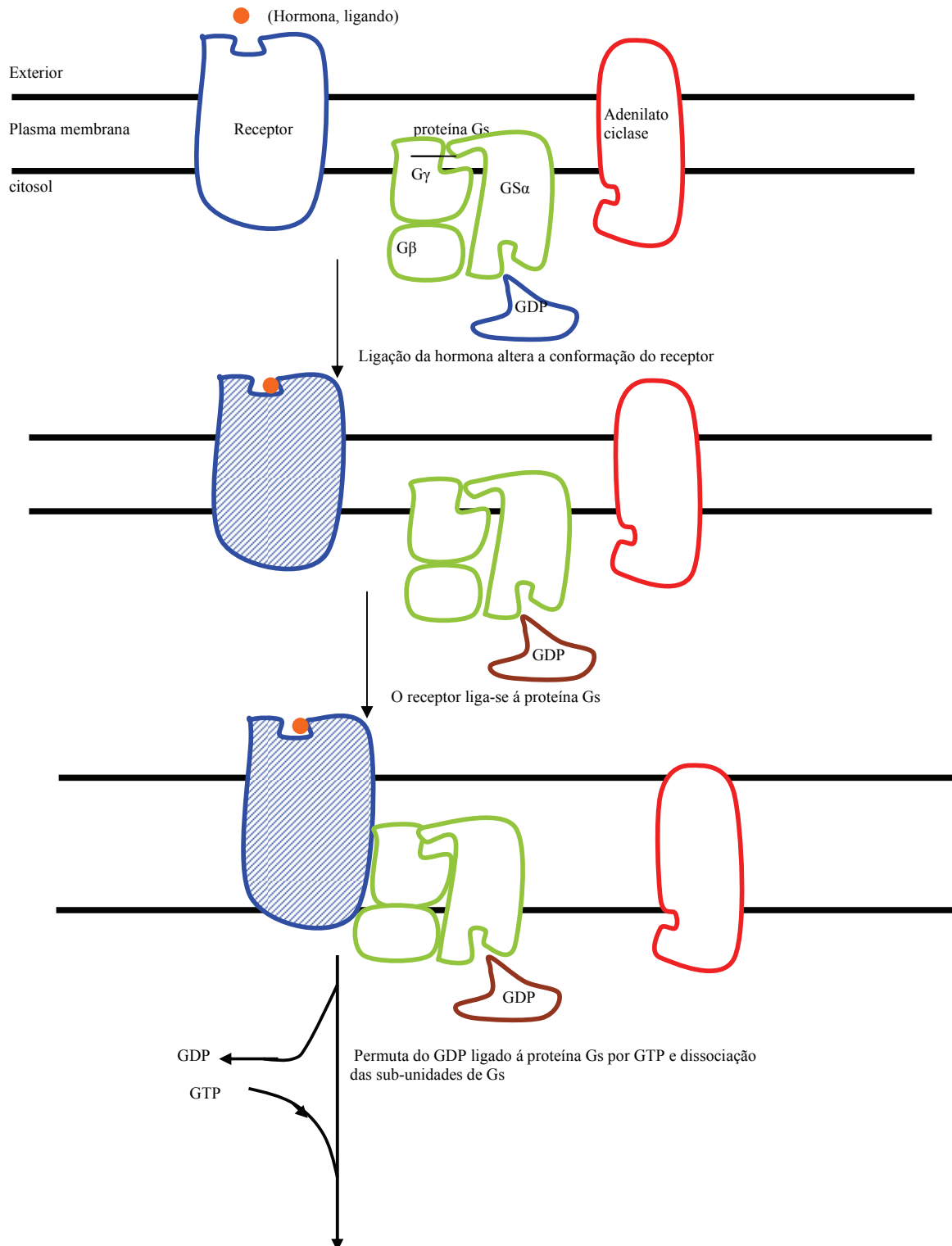


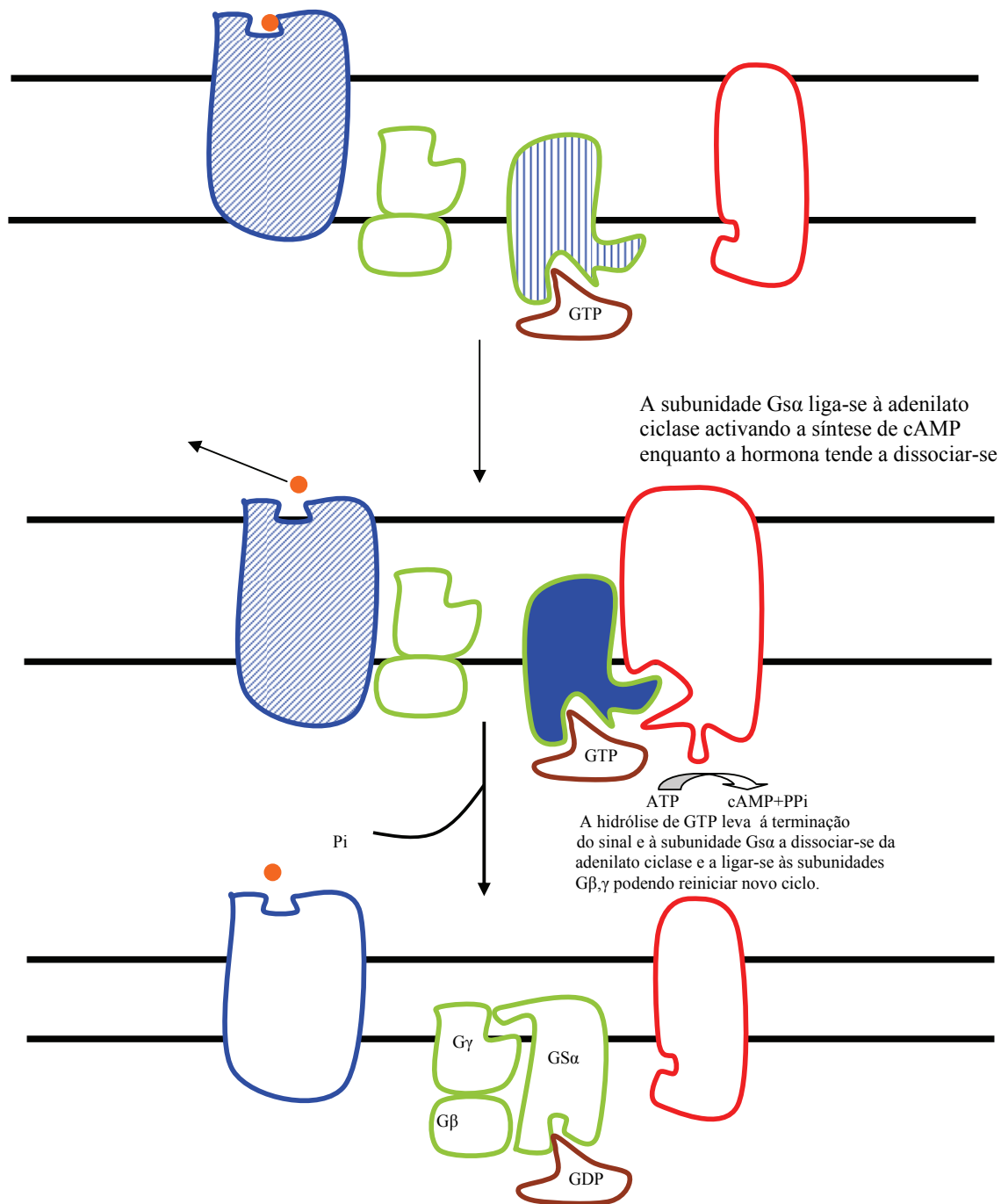
Portanto na transdução intracelular de um sinal extracelular através de uma “cascata” quando maior o número de etapas da cascata maior será a possível ampliação do sinal de partida.

No caso da síntese e catabolismo do glicogénio os efeitos do cAMP fazem-se sobretudo sentir no fígado e nos músculos, contudo o cAMP medeia as respostas intracelulares em muitas outras células.

Na figura seguinte mostra-se a activação da adenilato ciclase após interação de uma hormona com o receptor de sete hélices

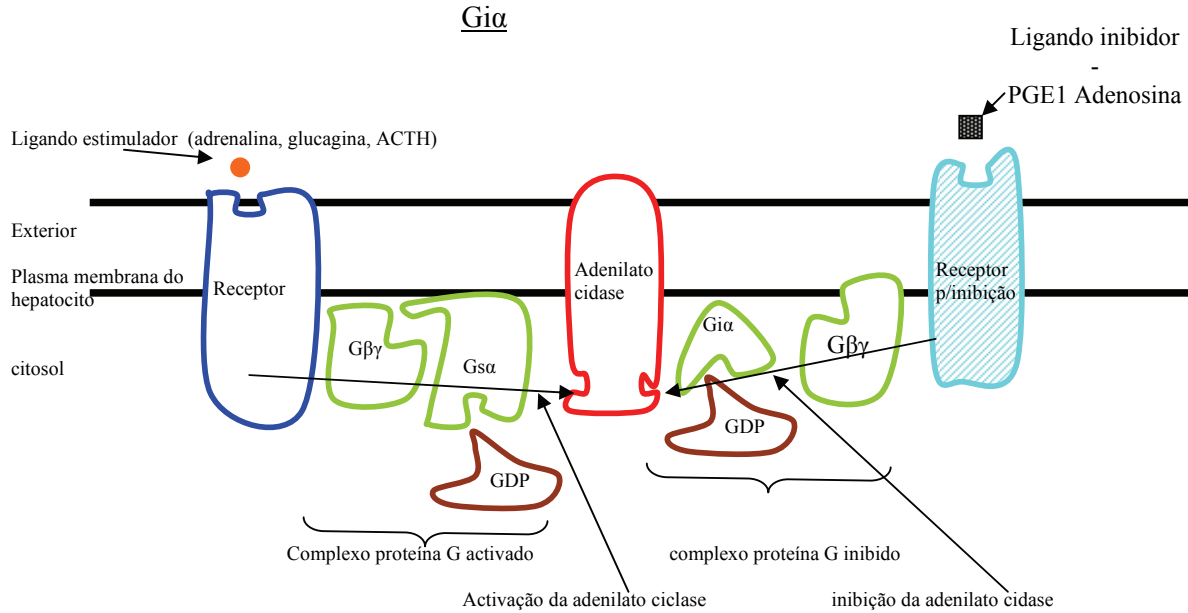
Activação da adenilato ciclase após ligação de uma hormona (por exemplo: epinefrina ou glucagina) ao receptor de sete hélices





Na figura seguinte mostra-se como a activação ou inibição da adenilato ciclase, promovida por diversos tipos de sinais hormonais indicados é mediada pela proteína  $G_{\alpha s}$  ou pela  $G_{\alpha i}$

A activação e inibição da adenilato ciclase induzida por hormonas é mediada pela proteína  $G\alpha$  ou pela



A ligação da sub-unidade  $G\alpha$ -GTP á adenilato ciclase activa esta enzima enquanto a ligação a ela de  $G\alpha$  a inibe, sucedendo que as subunidades  $G\beta\gamma$  nas proteínas G activadoras ou inibidoras são idênticas, mas as subunidades  $G\alpha$  e os receptores são diferentes na activação e na inibição.

Há múltiplas G proteínas acopladas a diferentes proteínas efectoras. Nos mamíferos estão assinaladas 16 subunidades  $G\alpha$  distintas, 4  $G\beta$  distintas e 5  $G\gamma$ .

Várias proteínas triméricas G ligam receptores de 7 hélices a variadíssimas proteínas efectoras tais como canais iónicos, adenilato ciclase, fosfolipase C, podendo assim ser modeladas diversas funções celulares.

A mesma célula pode exprimir diversas G proteínas. É possível também que um simples ligando possa iniciar a transdução de sinal através de mais de uma proteína efectora.

Também parece que em algumas células mamíferas complexos de subunidades  $G\beta\gamma$  podem regular directamente certas proteínas efectoras.

### Processos fisiológicos mediados por proteínas G

<u>Estímulo</u>	<u>Receptor</u>	<u>G proteína</u>	<u>Efector</u>	<u>Resposta fisiológica</u>
Epinefrina	Receptor adrenérgico	$\beta$ - Gs	Adenilato ciclase	Quebra de glicogénio
Serotonina	Receptor serotonina	de Gs	Adenilato ciclase	Sensibilização comportamental e aprendizagem na <u>aplysia</u>
Luz	Rodopsina	Transducina	cGMP fosfodiesterase	Excitação visual
Cheiros	Receptores olfativos	Golf	Adenilato ciclase	Olfacção
fMetPéptido	Receptor quimotáctico	Gq	Fosfolifase C	Quimotaxia
Acetilcolina	Receptor muscarínico	Gi	Canal de potássio	Actividade de retardamento do pacemaker

Como comentário resumo das últimas figuras repete-se que nestas cascatas de transdução de sinal a hormona ou sinal interacciona especificamente com o domínio extracelular do receptor de sete hélices (como os receptores  $\beta$  adrenérgicos) levando este a alterar a sua conformação e a activar as adenilato ciclases com o concurso de G-proteínas trimericas que também são activadas e designadas nesse estado como Gs.

Estas proteínas triméricas Gs são constituídas por três subunidades  $\alpha$ ,  $\beta$  e  $\gamma$ , podendo encontrar-se numa forma activa ou inactiva.

A proteína trimérica Gs, através da sua subunidade  $\alpha$  ou seja  $G\alpha$  pode estar complexada com o GDP e com as subunidades  $\beta$  e  $\gamma$  como sucede quando não se encontra nenhum sinal a interaccionar com o receptor. Mas se porventura o sinal interacciona com o receptor a alteração de conformação do receptor induz que o GDP seja substituído pelo GTP na interacção com a subunidade  $\alpha$  da G-proteína Gs o que leva á dissociação das subunidades  $\beta$  e  $\gamma$  da proteína G e á interacção e activação da adenilato ciclase.

As proteínas G são portanto intermediários entre as proteínas membranárias do receptor de sete hélices e da adenilato ciclase, podendo encontrar-se em dois estados ou complexada com GTP, o que equivale á forma activa, ou complexada com GDP ou seja a forma inactiva.

Existe um numeroso grupo de proteínas intracelulares que tal como as proteínas  $G\alpha$  pertencem a uma superfamília chamada GTPases. Nesta superfamília integra-se uma proteína a Ras que é uma proteína interruptora de cascatas e transdução de sinal e que pode também existir numa forma activa, a que tem GTP ligado, ou numa forma inactiva a que tem GDP ligado.

Quando activadas as proteínas Ras medeiam a interacção com outras proteínas efectoras especificas, implicada na regulação do crescimento e diferenciação celular.

Outras proteínas desta superfamília GTP ases são as Rab que regulam o tráfego de vesículas transportadoras dentro das células.

Quanto á Ras sabe-se que a sua estrutura tridimensional é semelhante á da G $\alpha$  embora mais pequena, sendo a activação da Ras induzida pela interação de varias hormonas e crescimento com o respectivo receptor.

Quando esta última situação ocorre forma-se um complexo activo Ras-GTP, havendo depois um outro factor específico o GEF (factor de troca de nucleótido de guanina) que permuta o GDP por GTP produzindo o complexo activo Ras-GTP e libertando-se o GEF.

É a hidrólise deste GTP do complexo Ras-GTP que leva á formação da forma inactiva Ras-GDP o que é acelerado por uma GTPase a GAP.

Estão identificadas mutações das proteínas Ras humanas, que podem complexar-se com o GTP mas não conseguindo hidrolizar este e daí a permanência do estado activo deste complexo e a formação de cancros.

Em condições humanas normais as proteínas Ras e GAP estão envolvidas na regulação do crescimento e diferenciação celular, por interação com sinais provenientes de células vizinhas.

Também é conhecido que certas toxinas bacterianas modificam irreversivelmente as proteínas G.

Os receptores acoplados a G-proteínas são substrato de cinases, cinases estas que constituem uma família de seis proteínas cinases serina/treonina eucariotas (GRKs).Estas enzimas cinases são muito importantes na dessensibilização da sinalização desencadeada pelos receptores acoplados a G-proteínas (GPCR), que constituem a maior família conhecida de proteínas de transdução de sinal.

A fosfatação do receptor desencadeada pelas GRK rapidamente dificultam a transmissão de sinal pelo receptor levando à dessensibilização.

Estas cinases GRK podem ser reguladas na sua actividade por diversos factores alostéricos bem como por outros factores tais como as subunidades  $\beta$   $\gamma$  das G-proteínas triméricas na forma GTP, cofactores fosfolipídicos, proteínas que interactuam com o cálcio como a calmodulina e recoverina, modificações postranslação como a isoprenilação e palmitoilação, autofosfatação, etc

#### 2.4.1.2 - Transdução de sinais e receptores com tirosina cinases intrínsecas

##### **Família de receptores tirosina cinase**

Em cada momento o estado fenotípico das células está a ser condicionado por diversos sinais extracelulares sinais estes que podem induzir na população celular em causa respostas muito variadas como por exemplo a proliferação celular e sua diferenciação, a sobrevivência celular ou a apoptose celular.

Nestes sinais extracelulares e intracelulares englobam-se muitos de diversa natureza havendo no entanto um grupo deles, o dos factores de crescimento polipeptídicos que para desencadear os seus efeitos biológicos, interactuam com receptores á superfície das células, receptores do tipo tirosina cinases (RTK) que têm alta afinidade para esse tipo de sinais, existindo nos vertebrados diversas subfamílias com características próprias. Assim os membros de uma determinada subfamília têm características estruturais idênticas, mas diferentes daquelas de outras subfamílias.

### Sub-famílias de receptores tirosina cinases (RTK) nos vertebrados

- receptor do factor de crescimento epidérmico (EGFR)
- receptor da insulina (IR)
- receptor PDGF/MCSF-1/steel (factor de crescimento derivado das plaquetas)
- receptor do factor de crescimento das células endoteliais vasculares (VEGF)
- receptor do factor de crescimento do hepatocito (HGF e SF)
- receptor neurotrofina (NGF, BDNF, NT-3, NT-4 e NI5)
- receptor do factor de crescimento dos fibroblasto (FGF)
- receptor Eph- like
- receptor Axl

Todas as RTK dos mamíferos apresentam um domínio extracelular que é o local que interaccua com o sinal activador (factor de crescimento ou outro) e um domínio intracelular com actividade tirosina cinase quando activado por indução do ligando.

Esta actividade tirosina cinase promove a activação de uma série de moléculas assinalantes comuns que (muito importante) acabam por desencadear mudanças na expressão de certos genes.

Esta série de moléculas assinalantes comuns são por exemplo a fosfolipase C, a fosfatidilinositol 3-cinase(PI3-cinase),a proteína activadora GTPase,a pp60c-src,a p21 ras,a Raf-1 cinase,a ERK 1 e ERK-2 cinases(também denominadas MAP cinases) e as cinases SG ribossomal (SG).

Ocorre ainda que para um único tipo de RTK são as respostas biológicas promovidas em função do tipo de células em que intervêm.

### Família EGFR

Estes receptores são expressos em diversos tipos de células, sendo o receptor para o factor de crescimento epidérmico(EGF) uma proteína tirosina cinase transmembranaria de 170 kDa.Nesta família EGFR englobam-se também outros membros tais como a p185 neu(também chamada erb B-2 ou HER-2) e a erb B-3.

Como já referimos da interacção do receptor com um ligando resulta a dimerização do receptor e a sua autofosfatação.

O EGF é um factor polipeptídico mitogénico que pertence a uma grande família de proteínas nos vertebrados que engloba o factor transformante de crescimento  $\alpha$  (TGF- $\alpha$ ), a amfíregulina (AR), o factor de crescimento derivado do Schawanoma (SDGF), o designado HB-EGF(heparin-binding EGF-like factors), vários ligandos que se ligam ao receptor HER-2/erb2/neu,etc.

Como características desta família EGF avulta o facto de apresentarem uma ou mais unidades estruturais EGF (domínios) no seu domínio extracelular.

### Família do receptor da insulina

Este receptor para a insulina (IR) apresenta uma característica distintiva dentro dos RTKs pois funciona como um heterotetramero com duas cadeias polipeptídicas  $\alpha$  e duas cadeias polipeptídicas

$\beta$ . Cada subunidade extracelular que interacciona com o ligando, está ligada por uma ponte dissulfureto à subunidade transmembranária  $\beta$ .

Nesta família incluem-se o receptor para a insulina (IR) e o receptor para o factor-1 "insulin-like growth" (IGF-1R) e o receptor "insulin-related" (IRR) cujos ligandos ainda são muito mal conhecidos.

A insulina, o IGF-1 e IGF-2 possuem uma homologia estrutural significativa, mas quanto aos respectivos receptores não se observa qualquer homologia entre o que interacciona com a IGF-2 e os que interaccionam com a insulina e o IGF-1.

Como veremos mais adiante a interacção da insulina com o receptor promove efeitos metabólicos e efeitos promotores do crescimento.

Apesar dos receptores para a insulina e para o IGF-1 conservarem os mesmos elementos estruturais nos seus domínios extracelulares, estas duas hormonas quando interaccionam com os domínios extracelulares dos seus receptores fazem-no em regiões distintas das subunidades  $\alpha$ .

Refira-se desde já que a proteína IRS-1 (pp185) é um importante substrato que interacciona com o receptor da insulina, e é dotado de mais de dez resíduos de tirosina fosfatáveis.

#### Família de receptores PDGF/MCSF-1/Steel

Esta família de receptores é caracterizada pela existência de cinco domínios (parecidos com os domínios das imunoglobulinas) na sua região extracelular pertencendo a esta família os receptores  $\alpha$  e  $\beta$  para o factor de crescimento derivado das plaquetas (PDGF), o receptor para o factor-1 estimulador de colónias de macrófagos (MCSF-1) e a proteína c-kit receptora para o ligando "steel".

Qualquer um destes três factores desempenha funções muito distintas. Assim o PDGF é um potente agente mitogénico de diversos tipos de células tais como as dos músculos lisos, das células gliais, dos fibroblastos e de certas células endoteliais, interaccionando com os respectivos receptores de tipo  $\alpha$  e de tipo  $\beta$ , embora só este último pareça mediar as interacções promovidas pelo ligando ao nível do citoesqueleto das células. Também o PDGF interaccionando com os respectivos receptores promove a dimerização destes por ligações não covalentes, parecendo intervir na cascata de transdução de sinal desencadeada, P13 cinase, GAP, PLC- $\gamma$ , raf e pp60c-src tirosina cinase.

O factor MCSF-1 é um regulador da proliferação, diferenciação e sobrevivência dos macrófagos não parecendo estar associado, na cascata de transdução de sinal por ele desencadeada com o PLC- $\gamma$  ou o GAP.

O factor ou ligando "steel" é importante no desenvolvimento dos melanocitos, das células germinativas e na hematopoiese.

#### Família de receptores (VEGF)

Este tipo de receptores é expresso apenas nas células endoteliais em resposta ao sinal VEGF que é caracterizado por possuir sete domínios do tipo imunoglobulina e um domínio tirosina cinase citosólico interrompido por uma inserção de um domínio de uma grande cinase.

Este factor VEGF aumenta a permeabilidade vascular provavelmente afectando as "light junctions" entre as células endoteliais.



### Família de receptores para o factor de crescimento do hepatocito

Os dois factores de crescimento do hepatocito o HGF e o SF (scatter factor) desencadeiam nas células epiteliais diversos efeitos tais como mitogénese e estimulação da mobilidade, sendo estes dois factores proteínas idênticas, que após interagirem com o respectivo receptor desencadeiam uma cascata de transdução de sinal que in vitro e in vivo está associada com as moléculas assinalantes, PLC- $\gamma$ , GAP, cinases relacionadas com src e com cinase P-13.

### Família de receptores neurotrofina

Para o crescimento, diferenciação e sobrevivência dos neurónios os ligandos neurotrofina são muito importantes, além de estimularem a proliferação celular em células não neuronais.

Nesta família há que considerar cinco membros; o factor de crescimento dos nervos (NGF), o factor neurotrófico derivado do cérebro (BDNF), a neurotrofina-5 (NT-5), variando nas suas afinidades, altas e baixas, para os respectivos receptores.

Há duas classes distintas de receptores para as neurotrofinas, as p75 LNGFR (ou gp80 LNGFR ou LNGFR) e as trk (trk A, trk B e trk C).

Nas cascatas de transdução de sinal de vias assinalantes desencadeadas há dúvidas sobre a utilização de cAMP e do Ca<sup>2+</sup>.

Parece que a fosfatação aumentada promovida pelo ligando NGF activaria a ERK 1 e ERK 2, a SG cinase ribossomal e a RAF-1, embora também sejam referidas como activadas a MAP cinase ou MAPKK e a ras.

### Família dos receptores do factor de crescimento dos fibroblastos

Estes factores (FGFs) desencadeiam diversas respostas tais como proliferação, diferenciação e sobrevivência das células tais como angiogenese.

Estão identificados quatro genes distintos codificantes destes receptores (FGFR1, FGFR2, FGFR3, FGFR4).

Estão identificados sete membros na família FGF incluindo a FGF acídica (aFGF), básica (bFGF), o produto do oncogene int-2, o produto do oncogene hst (também chamado sarcoma de kaposi FGF), o FGF5, o FGF6, e o factor de crescimento do ceratinocito (KGF).

Os FGFs ligam-se também às moléculas de heparina sulfato proteoglicana localizadas na superfície celular ou na matriz extra-celular.

A região extra-celular do receptor contém três domínios consenso análogos aos das imunoglobulinas.

No caso das genes FGFR1 e FGFR2, múltiplas formas do receptor FGF podem ser produzidas devido a um diferente “splicing”.

A fosfatação dos receptores dimerizados parece ocorrer através de um mecanismo de transfosfatação intermoléculas.

A estimulação desencadeada pela FGFs aumenta o pH intracelular e as concentrações de  $\text{Ca}^{2+}$ , o turnover PI, assim como a fosfatação de uma série de proteínas intracelulares, tais como PLC- $\gamma$ , Raf-1, ERK1 e ERK2 cinases e SG cinase ribossomal.

### Receptor Eph-Like Tirosina cinase

Estas proteínas são a maior sub-família das RTK. Os genes que as codificam incluem Eph, Eck, Elk, Feek, Erk, Cek4 e Cek5.

Por ora não são conhecidos ligandos que se liguem a qualquer das Eph-like proteínas.

### Receptor Axl Tirosina cinase

É uma proteína 140KDa relacionada com as sub-famílias de receptores tirosina cinase da insulina.

Não estão identificados ligandos para este tipo de receptor.

### Regulação dos receptores da superfície das células

Constitue um enigma o facto de um simples factor ou ligando poder actuar sobre diferentes tipos de células induzindo a activação de uma série de moléculas assinalantes idênticas como por exemplo a ras, na cascata de transdução de sinal e com fenótipos completamente diferentes.

Outro enigma consiste em diferentes factores de crescimento actuarem sobre o mesmo tipo de células induzindo diferentes fenótipos.

È bem conhecido que o número e actividade dos receptores à superfície das células é variável havendo diversas formas de operar esta regulação.

De um lado pode ocorrer o internamento nas células dos receptores, por endocitose. Após isto podem ser destruídos ou armazenados em vesículas intracelulares afectando tudo isto a quantidade desses receptores.

Assim receptores por exemplo para a insulina, glucagina, EGF e PDGF sofrem em parte endocitose sendo depois degradados ao nível dos lisossomas, enquanto outra parte pode ser reciclada e devolvida à superfície das células por exocitose.

Contudo a maioria dos receptores para hormonas peptídicas não é reciclada eficientemente.

Quanto à sua actividade também esses receptores podem ser regulados. É o que sucede por exemplo com a fosfatação ou desfosfatação desses receptores da superfície das células.

Estas proteínas que são reguladas através de fosfatações e desfosfatações podem conter ao longo da sua estrutura diversos locais para estas ocorrências podendo as fosforilações ser efectuadas por mais de um tipo de cinases e as desfosforilações também podendo ser desencadeadas por mais e um tipo de sistema enzimático.

A título de curiosidade refira-se desde já que as duas cadeias  $\alpha$  e  $\beta$  dos receptores para a insulina são codificados por um único gene.

## **Moléculas adaptadoras e sinalização**

O numeroso grupo dos receptores tirosina cinases pode ser activado pela interacção com variadíssimos tipos de ligandos, que podem ser hormonas solúveis, ou péptido/proteínas ligadas a membranas. Referimos anteriormente uma série de famílias destes ligandos sinais e teremos oportunidade mais adiante de estudar em maior detalhe as cascatas e transdução e sinal desencadeadas pela insulina, EGF, etc.

Nas características estruturais de todos estes receptores tirosina cinase avulta uma porção extracelular, a que interacciona com o ligando, e uma porção intracelular ou citoplasmica dotada de actividade tirosina cinase, que interagindo com moléculas adaptadoras desencadeia diversas respostas metabólicas.

É na porção citoplásmica destes receptores que se situa a actividade tirosina cinase que quando activada transduz o sinal originando alterações na fisiologia da célula e/ou diferentes comportamentos da expressão dos genes. Pode desta forma regular-se a proliferação e diferenciação celular, a apoptose ou sobrevivência celular e ainda regulações do metabolismo das células.

Estas cascatas de transdução de sinais funcionam ora ligadas ora desligadas. Pode suceder que mutações ao nível destes receptores RTK, não permitam que a cascata de transdução de sinal seja desligada e daí, mesmo na ausência de factores de crescimento, essas células exibirem uma proliferação desordenada e a formação de situações oncológicas.

Conforme já tivemos oportunidade de referir a interacção ligando-receptor promove a dimerização dos receptores com a consequente modificação da estrutura conformacional deste complexo, o que leva a que cada monómero do receptor seja fosfatado, ou seja autofosfatado no conjunto distinto de resíduos de tirosina contidos neste domínio citosólico do receptor.

Após isto é possível a interacção deste complexo activado com duas classes diferentes de proteínas ou sejam:

1) proteínas ditas adaptadoras como por exemplo as GRB 2 que acoplam os receptores activados com outras moléculas sem actividade assinalante intrínseca.

2) com proteínas enzimáticas envolvidas em vias assinalantes, como por exemplo a GAP (uma proteína associada com a Ras), a Syp (uma proteína fosfatase) e enzimas implicadas na formação de derivados do fosfatidilinositol.

As proteínas destas duas classes antes referidas têm a propriedade de poder interaccionar com os resíduos de fosfotirosina formados no lado citosólico do receptor devido a possuírem na sua estrutura um domínio polipeptídico conservado, o domínio Src de homologia 2 ou SH2 codificado pelo gene src. Há também nestas proteínas um outro domínio SH3 homólogo com outra região do Src que em certas vias de transdução de sinal está implicado nas interacções entre diversas proteínas.

Estes domínios SH2 nas diferentes proteínas em que se encontram revelam estruturas tridimensionais muito parecidas mas diferentes estruturas primárias, as que são específicas para cada proteína, o que permite a cada uma destas reconhecer especificamente os resíduos específicos de fosfotirosina do receptor com que irão interaccionar.

Como referimos atrás as proteínas Ras intervêm em várias cascatas de transdução e sinal eucariotas interconectando os receptores RTK com proteínas serina/treonina cinases que intervêm na cascata, podendo controlar o crescimento e a diferenciação celular.

É o caso dos factores PDGF e EGF que realizam cascatas deste tipo com a intervenção da Ras (vide figura adiante).

Na via de transdução de sinal esboçada nesta figura intervêm duas proteínas citosólicas a GRB2 e a Sos, com a GRB2 a interagir através de um domínio SH2 com um resíduo específico de fosfotirosina da porção citosólica do receptor RTK activado pelo factor de crescimento. Salienta-se que ainda no GRB2 dois domínios SH3 ligam-se e activam a proteína Sos.

Esta proteína Sos activada funciona com uma proteína permutadora de nucleótido (GEF) e a sua ligação à proteína Ras vai activar esta quando ocorre a substituição de GDP por GTP, sendo depois induzida uma cascata proteína cinase que altera a expressão de genes específicos.

Refira-se a talhe de foice que no olho na Drosophila estão identificadas três proteínas que ligam os receptores RTK com uma cascata cinase e que são a Ras proteína, uma GEF chamada Sos e uma proteína adaptadora contendo SH2, semelhante à proteína GRB2, como sucede nos mamíferos.

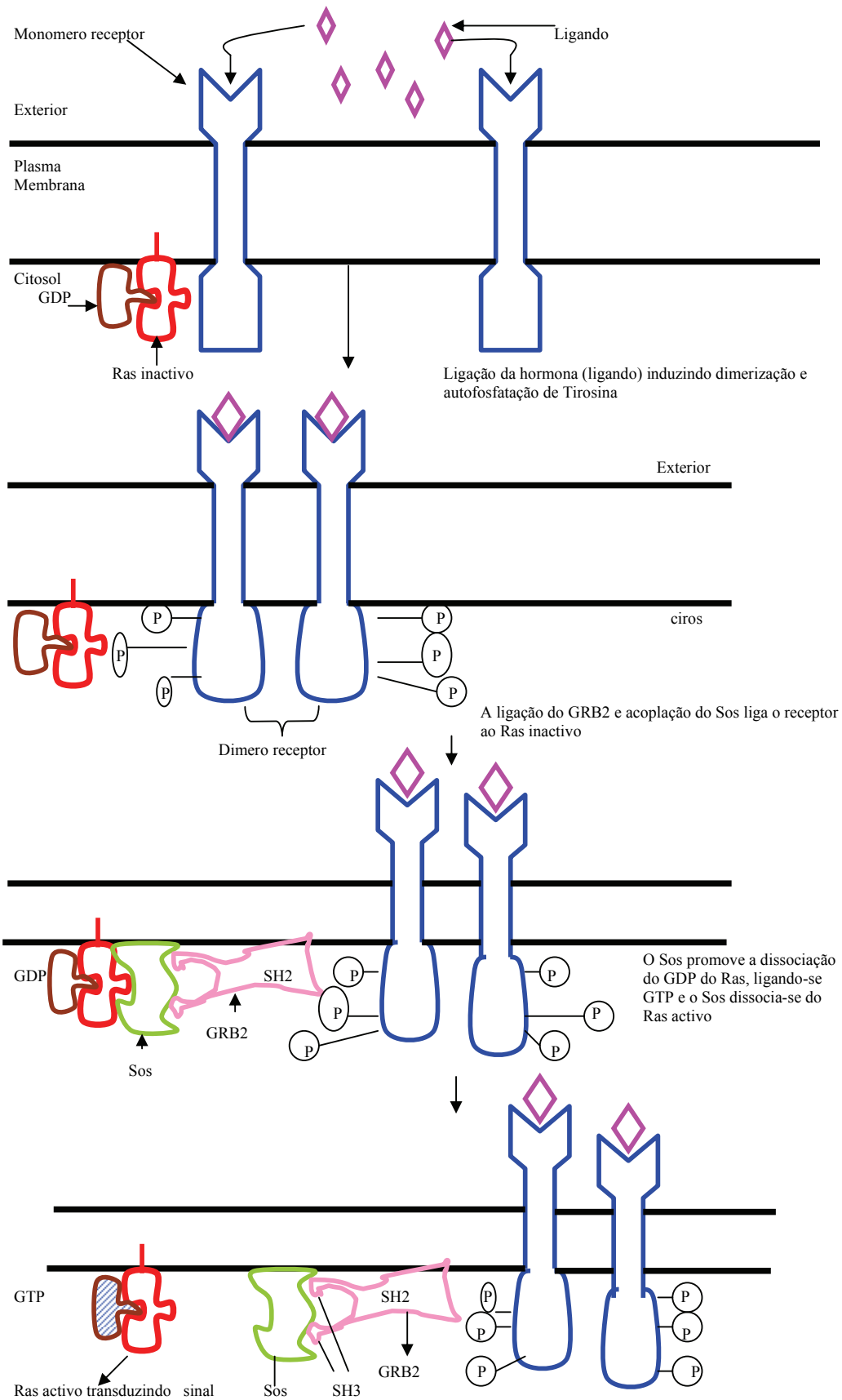
Resumindo a GRB2 é uma proteína adaptadora que interage por um lado com os receptores RTK activados e por outro com a Sos. A Sos está localizada na membrana celular e interage com a GRB2 através de dois domínios SH3 (estes domínios SH3 contêm cerca de 55-70 ácidos aminados e encontram-se em grande número de moléculas assinalantes, possuindo estruturas tridimensionais semelhantes, mas diferentes e específicas estruturas primárias). Por outro lado a GRB2 contém um domínio SH2 que interage com a fosfotirosina dos receptores RTK:

Após o sinal desencadeado pela hormona sobre o RTK o complexo activado contém GRB2 e Sos do lado citosólico da plasma membrana, sendo este último recolocado próximo da proteína Ras ligada à membrana celular que é activada por acção da Sos (vide figura seguinte).

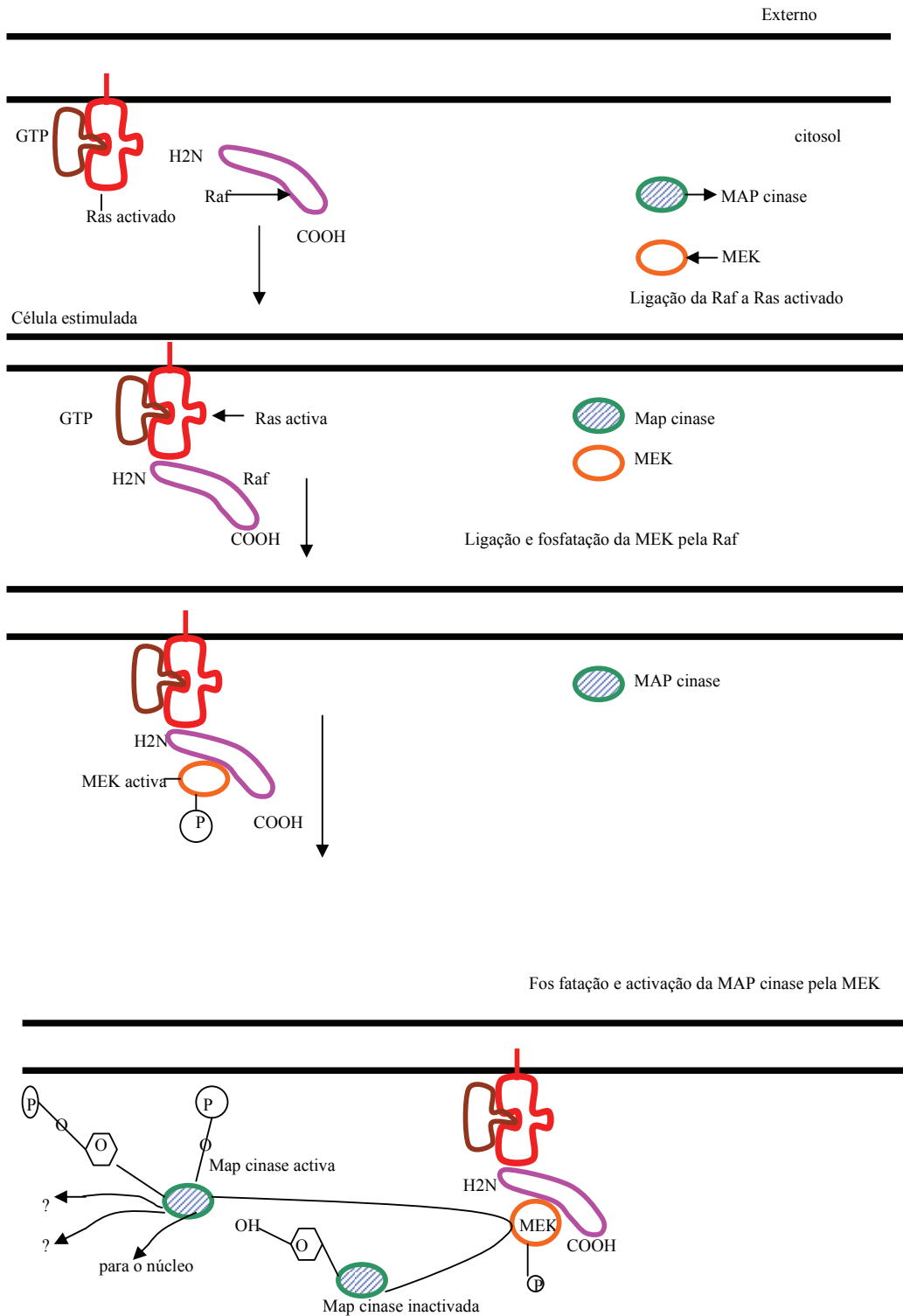
Esta activação da Ras induz a diferenciação celular ou a proliferação celular consoante o tipo e células.

Em condições normais o Ras é activado pela ligação do ligando ao RTK mas quando ocorrem mutantes em células, pode formar-se um Ras constitutivo activo que permanece activo na ausência de um ligando indutor no RTK.

Activação da Ras após ligação de uma hormona (por ex: EGF) ao receptor RTK



Vejamos agora como funciona o Ras ativado na transdução do sinal através de uma série de proteínas de transdução de sinal.



### Função e regulação das proteínas Ras

As proteínas Ras são portanto transdutoras de sinais, sendo estas proteínas Ras codificadas por genes ras, podendo mutações variadas, alterar estas proteínas e desencadear fenótipos neoplásicos.

Em todos os eucariotas se encontram normalmente estes genes ras que são reguladores críticos do crescimento e da diferenciação celular funcionando como intermediários chave em vias de transdução de sinais.

As proteínas Ras interactivam com alta afinidade para nucleótidos de guanina (GTP e GDP) e possuem intrinsecamente actividade GTPasica sendo portanto GTPases, sendo a superfamília de proteínas Ras constituída por proteínas que têm uma certa homologia com a Rãs e cerca de 20-25 kDa.

Dentro da família Ras há que referir grupos de pequenas G-proteínas que regulam nos mamíferos a reorganização do citoesqueleto e expressão de genes (vide quadro seguinte)

Família Ras	Família Rho
Há-Ras	RhoA
Ki-Ras	RhoB
N-Ras	RhoC
R-Ras	RhoD
M-Ras	RhoE/
Ra1A	Rnd3/
Ra1B	Rho8
Rap1A	RhoG
Rap1B	RhoH/
Rap2A	TTF
Rap2B	Rac1
Tc21	Rac2
Rit	Rac3
Rin	Cdc42
Rad	Rnd1/
Kir/Gem	<u>Rho6</u>
Rheb	
KB-Ras1	
KB-Ras2	

Takai, Y et al – Physiological Reviews 81; 153-208, 2001

Nas células mamíferas as proteínas Ras ligam-se directamente e activam a proteína cinase Raf a qual induz a expressão de genes através de uma cascata com intervenção da MAP (mitogen-activated protein kinase) em resposta a diversas moléculas assinalantes extra - celulares.

As proteínas Ras regulam a proliferação celular, a sua diferenciação, morfologia e apoptose.

A actividade das Ras proteínas é regulada por GEPs (proteínas de troca de nucleotídeos de guanina) e GAPs (proteínas activadoras de GTPases) e a sua activação é induzida por muitos sinais extra - celulares, sobretudo sinais que activam receptores com actividade intrínseca ou associada de tirosina cinases.

Estas fosfotirosinas servem como locais de docagem para proteínas adaptadoras como é o caso das GRB2 e do complexo SHC/GRB2, que depois recrutam SOS do citosol para a plasma membrana estimulando a Ras proteína localizada na face citoplásmica da plasma membrana e convertendo-a da forma com GDP ligado na forma activa com GTP ligado.

Outros receptores não directamente associados com tirosina cinases, como é o caso dos receptores das células T, podem activar directamente as Ras.

Estão assinaladas três GEPs da Ras os SOS, Cdc25 e Ras GRF.

Nos GAPs o melhor caracterizado é o 120 Ras GAP.

As Ras proteínas medeiam a proliferação celular através da activação de uma cascata de proteínas cinases como as Raf proteínas cinases (cRaf1, A-Raf, e B-Raf), MEK (MAPcinases cinases 1 e 2) e MAPcinase. A MAPcinase activada transloca-se para o núcleo onde fosfata e activa vários factores de transcrição inclusive o E1K-1.

Versões da Ras proteínas, com mutações nos três genes humanos Ras (Há-Ras, Ki-Ras e N-Ras) foram detectadas em cerca de trinta por cento de todos os cancros humanos.

Apesar destas notáveis homologias estruturais estão assinalados três genes ras funcionais, nos humanos e nos roedores, os H-ras, K-ras e N-ras situados em cromossomas diferentes e possuindo promotores com elevado conteúdo GC, faltando-lhes a caixa TATA. O gene K-ras codifica duas proteínas p21.

A Ras proteína pode ser regulada positivamente ou negativamente por uma série de factores, encontrando-se em todos os tecidos fetais e adultos.

A actividade de GTPase da Ras pode por exemplo ser activada por uma proteína GAP que se liga de preferência á GTP-Ras, parecendo que todas as proteínas Ras normais são sensíveis á acção desta GAP.

Outros genes codificam proteínas que activam a actividade GTPase da Ras nos mamíferos, como é o caso do gene NF1 que exprime neurofibromina para este efeito.

As proteínas Ras são pois importantes na proliferação das células normais estando assinalado que no Sac.cerevisiae apenas a adenil ciclase é um effector da proteína Ras.

Diversas serina/treonina cinases mitogénicas são activadas através da participação de proteínas Ras, embora possa ser necessário para aquele efeito outros effectores diferentes das Ras. Estão neste caso as proteínas Raf, as cinases de mitoses activadas por proteínas (MAP), as cinases que activam as MAP cinases (MAP cinase cinase), a cinase SG da proteína ribossomal (R1sk cinase) e a PKC.

Os efeitos desencadeados pela activação da Ras são muito parecidos com os provocados pelos factores de crescimento do soro conforme veremos adiante.

Na figura anterior a Ras activada liga-se ao domínio N-terminal da Raf que é uma serina/treonina cinase e esta fosfata a MEK que é uma cinase que fosforila resíduos de serina e de tirosina.



A MEK fosfata uma serina/treonina cinase a MAP cinase que se encontra associada com microtúbulos na célula e que por seu turno fosfata diversas proteínas inclusivé factores de transcrição envolvidos na regulação da expressão de genes e na diferenciação celular.

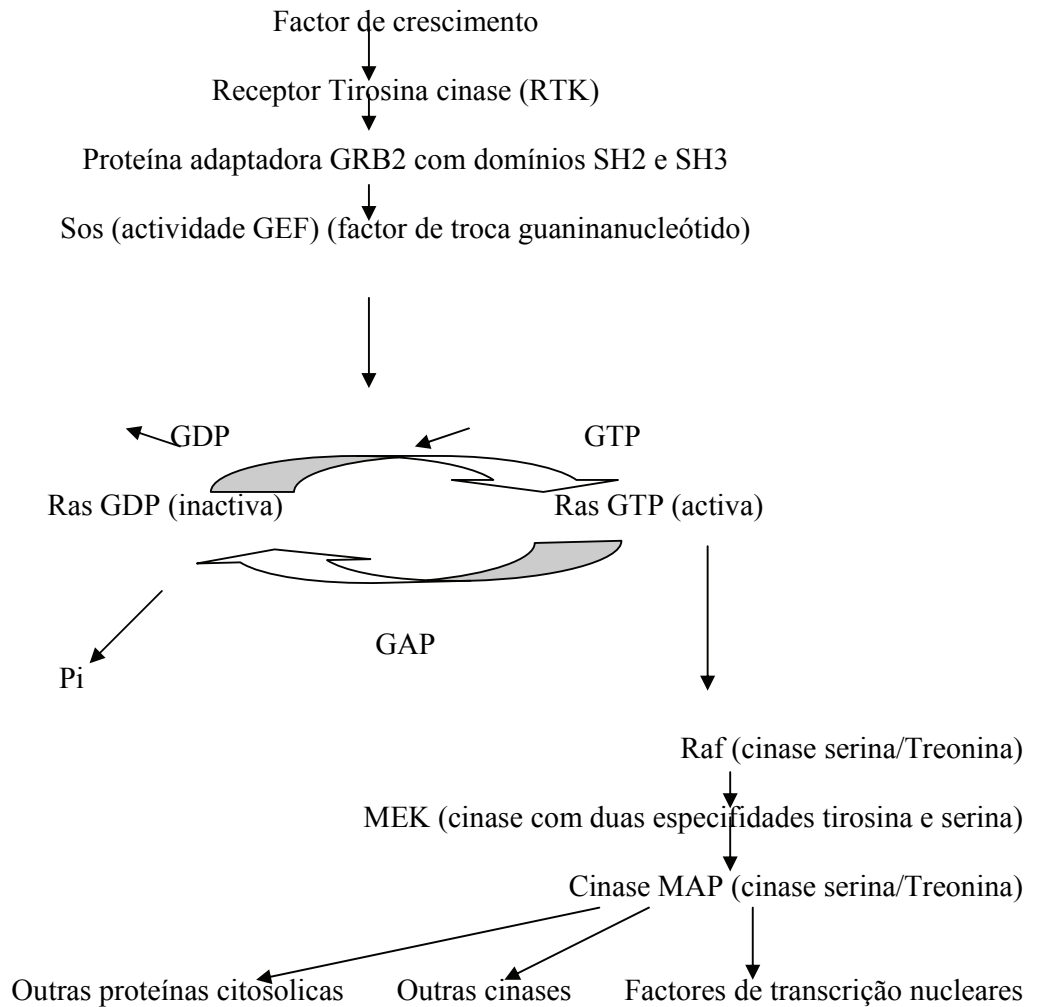
RTK »»»»Ras»»»»Raf»»»»(?)»»»»MAP cinase

Refira-se a talhe de foice que em células quiescentes activadas para entrarem em ciclo celular, podem ser necessárias induções e genes precoces imediatos que exprimam factores de transcrição, como por exemplo o fos e o jun, que são indispensáveis para que o ciclo celular prossiga.

A maioria dos receptores RTK desencadeia a via de transdução de sinais anteriormente referida mas podem ocorrer também outras vias diferentes.

A activação dos receptores RTK está pois acoplado com a proteína Ras levando à activação de uma MAP cinase como se esquematiza seguidamente

Via de transdução de sinal RTK-Ras



### Outros importantes segundos mensageiros

Parece que os RTK ao contrário dos receptores de sete hélices não têm capacidade para regular as concentrações do AMP cíclico, não o utilizando pois como segundo mensageiro.

Já no que se refere ao  $Ca^{2+}$ , como ao inositol 1,4,5 trifosfato e ao 1,2 diacilglicerol, tanto os RTK como os receptores de sete hélices utilizam-nos como segundos mensageiros em cascatas de transdução de sinais.

O cálcio desempenha efeitos celulares variados por vezes mediado por calmodulina consoante a sua concentração citosólica.

O inositol 1,4,5 trifosfato promove a libertação do cálcio a partir do retículo endoplásmico, enquanto o 1,2-diacilglicerol activa a proteína cinase C.

Esta proteína cinase C conforme já referimos anteriormente origina diversas respostas consoante os tipos celulares em que actua, sendo importantes no crescimento celular e no metabolismo. Assim no caso do fígado fosfata a glicogénio sintase que nestas circunstâncias passa ao estado inactivo e fosfata ainda uma série de factores de transcrição, consoante as linhas celulares, regulando a síntese de certos RNA mensageiros. O factor de transcrição NF-KB ao nível do citoplasma das células não estimuladas é inibido por uma proteína que quando fosfatada pela proteína cinase C é libertada do factor NF-KB permitindo que este se transloque para o núcleo e aí induza a transcrição de certos genes.

### **Moléculas adaptadoras e factores de transcrição respondendo a cascatas de transdução de múltiplos sinais inclusivé a partir de receptores de sete hélices**

É possível a modificação da expressão dos genes dentro das células, através da interacção de factores de crescimento, hormonas hidrossolúveis e neurotransmissores com receptores específicos da superfície celular desencadeando cascatas de transdução desse sinais.

Esta interacção do ligando ou sinal com os receptores específicos activa proteínas cinases que actuam sobre factores de transcrição fosfatando serinas, treoninas ou tirosinas que se encontram nas estruturas desse factores de transcrição.

Esta activação dos factores de transcrição pode resultar da estimulação promovida por ligandos que interactuam com receptores ligados a proteínas G, ou de ligandos que interactuam com receptores tirosina cinases (RTK).

Teremos oportunidade adiante de vislumbrar a activação de diversos factores de transcrição por diversas vias de transdução de sinais desencadeadas por diversos tipos de ligandos. Assim veremos vias mediadas através de proteínas G ligadas a receptores de sete hélices, vias mediadas por receptores tirosina cinases (RTK) e ainda a estimulação da expressão de genes por sinais, interferão  $\gamma$  por exemplo (IFN  $\gamma$ ), em que a transdução do sinal se faz através de um receptor não dotado de actividade tirosina cinase (NRTK) mas com o concurso de uma proteína citosólica cinasica.

Veremos também como dois factores de transcrição melhor conhecidos, a proteína de ligação ao elemento CRE do DNA ou seja a CREB e o factor de resposta ao soro (SRF) que estimulam a expressão de diversos genes são desencadeados por diversas hormonas hidrossolúveis e por diversos factores de crescimento.

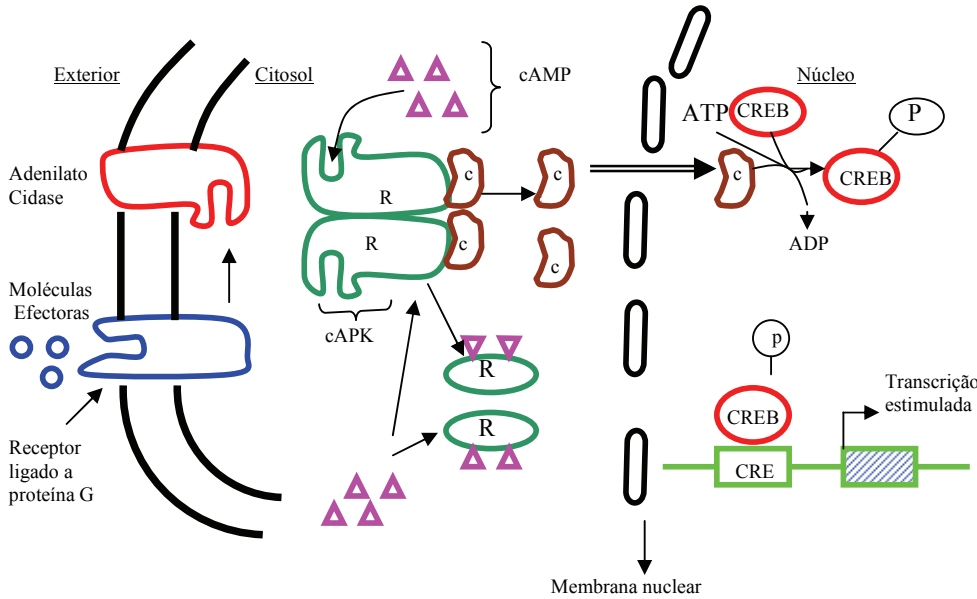
A CREB é uma proteína nuclear que modula a transcrição de genes com elementos nos seus promotores e que responde a diversos tipos de estímulos tais como já referimos (hormonas peptídicas, factores de crescimento e actividade neuronal) estímulos estes que activam uma série de proteínas cinases como por exemplo a proteína cinase A (PKA), proteínas cinase que activam a mitogénese(MAPKs),proteínas cinase dependentes do Ca<sup>2+</sup>/calmodulina(CAMKs),etc,todas elas activando proteínas CREB quando fosfatam esta no seu ácido aminado serina na posição 133.

O tipo de estímulo é que vai condicionar o mecanismo e a via pela qual a transcrição desencadeada pela CREB ocorre, sendo desta forma regulados os tipos de genes que vão ser expressos.Certamente que apesar da estrutura tridimensional parecida dos vários tipos das várias proteínas CREB, haverá depois particularidades específicas na estrutura primária dos diversos tipos de proteínas CREB.

O factor de transcrição CREB é pois um componente de acontecimentos intracelulares assinalantes e que regula uma larga gama de funções biológicas, desde a espermatogénese, até aos ritmos circardianos e memória.

Na figura seguinte esquematizamos a activação do factor de transcrição CREB numa via em que intervem cAMP e que vai modelar a expressão de certos genes.

Via de transdução de sinal levando à activação do factor de transcrição (CREB proteína) e modelação da expressão dos genes (via proteína G-cAMP)



Nas células mamíferas a interacção de efectores adequados com os seus receptores específicos de sete hélices localizados na membrana celular, desencadeiam através da proteína G não indicada na figura a activação da adenilato ciclase que eleva a concentração de cAMP citosólico.

Como já referimos este vai interactuar com as subunidades reguladoras ( R ) da proteína cinase A ( PKA ou APK) que passa assim da forma inactiva para a forma activa em que as duas subunidades catalíticas( C) se translocam para o núcleo onde vão induzir a expressão de vários genes,como por

exemplo o gene da somatostatina em certas células endócrinas ou genes de células hepáticas responsáveis pela formação de enzimas da gliconeogénese.

Estes genes têm uma característica estrutural que corresponde a sequências do respectivo DNA cis-actuante ou sejam sequências que controlam genes no mesmo cromossoma e que se chamam neste caso elementos de resposta ao cAMP ou seja CRE.

Como se verifica na figura anterior em apreço ao nível do núcleo as subunidades C da PKA activadas vão fosforilar o factor de transcrição nuclear CRE (porque se liga ao CRE depois de activado) convertendo-o em CREB.

É esta forma activa CREB que se liga a genes alvo contendo CRE, estimulando a sua transcrição.

Esta proteína CREB é um alvo para uma série de cascatas de transdução de sinais mediando respostas a estímulos extracelulares.

O elemento CRE do DNA de resposta ao cAMP está identificado, tem 8-bp e tem a sequência 5'-TGACGTCA-3'.

As cinases referidas anteriormente fosfatam a CREB no seu resíduo serina 133. Mas podem ainda ocorrer vias assinalantes que marquem locais adicionais na CREB ou em proteínas associadas com a CREB levando a que esta regule as expressões distintas de genes em diferentes condições de estimulação.

A fosfatação da serina 133 da CREB, ao nível do núcleo pode afectar a sua capacidade para se dimerizar com diferentes bZIP "partners" (vide adiante) ou então promover a interacção CREB-CRE, levando à activação da transcrição por interacção também com componentes do mecanismo básico da transcrição como o TFIID, a RNA polimerase II, etc.

Diferentes estímulos podem levar àquela interacção da CREB com bZIP "partners".

Como é natural a CREB possui diversas isoformas e é conhecido que a sua Kd para a interacção CREB-CRE varia entre 1 e 180 o que é devido sobretudo à estrutura primária das proteínas, no local da interacção.

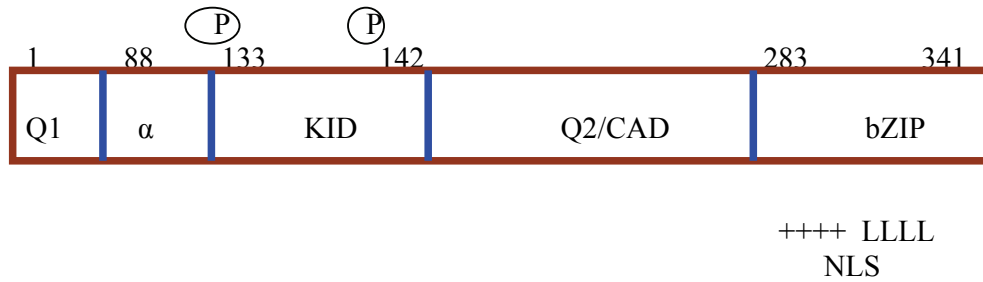
A proteína CREB é fundamental como já referimos para a proliferação e diferenciação celular, para variadíssimas respostas adaptativas, inclusive a aprendizagem e memorização, para o controlo de processos metabólicos tais como a regulação da gluconeogénese pela glucagina e pela insulina.

Como referimos atrás a CREB forma um dímero para interactuar com o CRE alvo no DNA. Este dímero resulta da interacção de um motivo estrutural conservado na porção C-terminal da CREB constituído por uma heptade repetida de leucinas, chamada leucina "zipper".

Por outro lado a ligação da CREB ao DNA é feita por um seu domínio básico, uma sequência de ácidos aminados rica em lisina e arginina aminoterminais e pela leucina "zipper" (bZIP).

Estas características estruturais levam a colocar a CREB dentro de uma família maior de proteínas factores de transcrição bZIP que engloba nos mamíferos a c-Fos, c-Jun, cMYc e cEBP.

Na figura seguinte esquematizam-se os principais domínios da CREB.



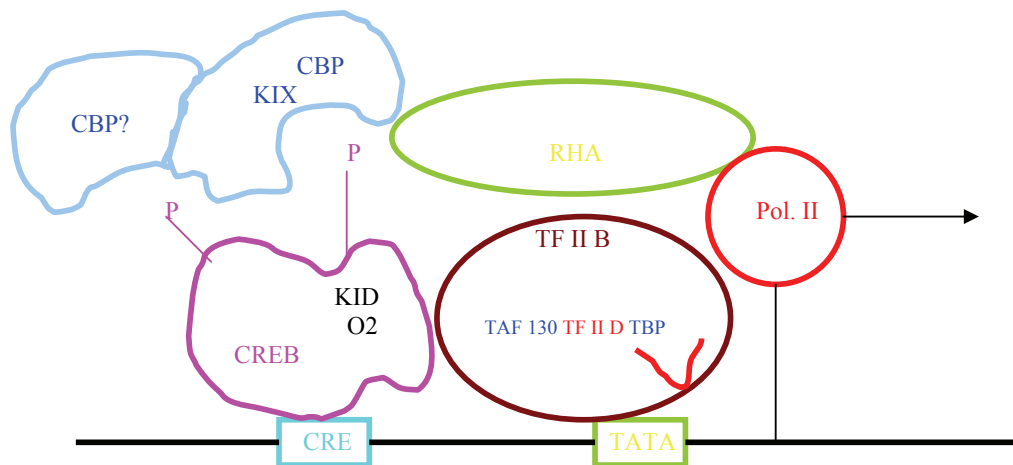
### Domínios do CREB

Os locais de fosfatação na serina 133 e na serina 142, são referidos com os números de cima que indicam a posição dos resíduos de ácidos aminados. Os sinais +++++ referem-se ao domínio básico carregado positivamente e LLLL refere o domínio leucina zipper.

A localização NLS corresponde ao sinal da posição nuclear.

Vejamos agora como os múltiplos domínios da CREB estão envolvidos na activação da transcrição.

### Activação da transcrição pelos múltiplos domínios da CREB



O CBP é um factor nuclear de 265 kDa que se liga á proteína CREB fosfatada servindo para “recrutar” um complexo de transcrição incluindo a RNA polimerase II (Pol.II) para acaixa TATA com a ajuda de uma proteína RNA helicase (RHA).

A CREB fosforilada utiliza estas duas proteínas de interacção a CBP e a RHA para estabilizar a sua interacção com a PolII.

Portanto diferentes domínios da CREB interactuam com coactivadores distintos e factores básicos da transcrição para activar esta.

O dímero da CREB interactua com o elemento CRE no promotor do gene alvo da CREB, encontrando-se mais abaixo a caixa TATA com a qual interactua o factor de transcrição básico TF II D através da proteína TBP.

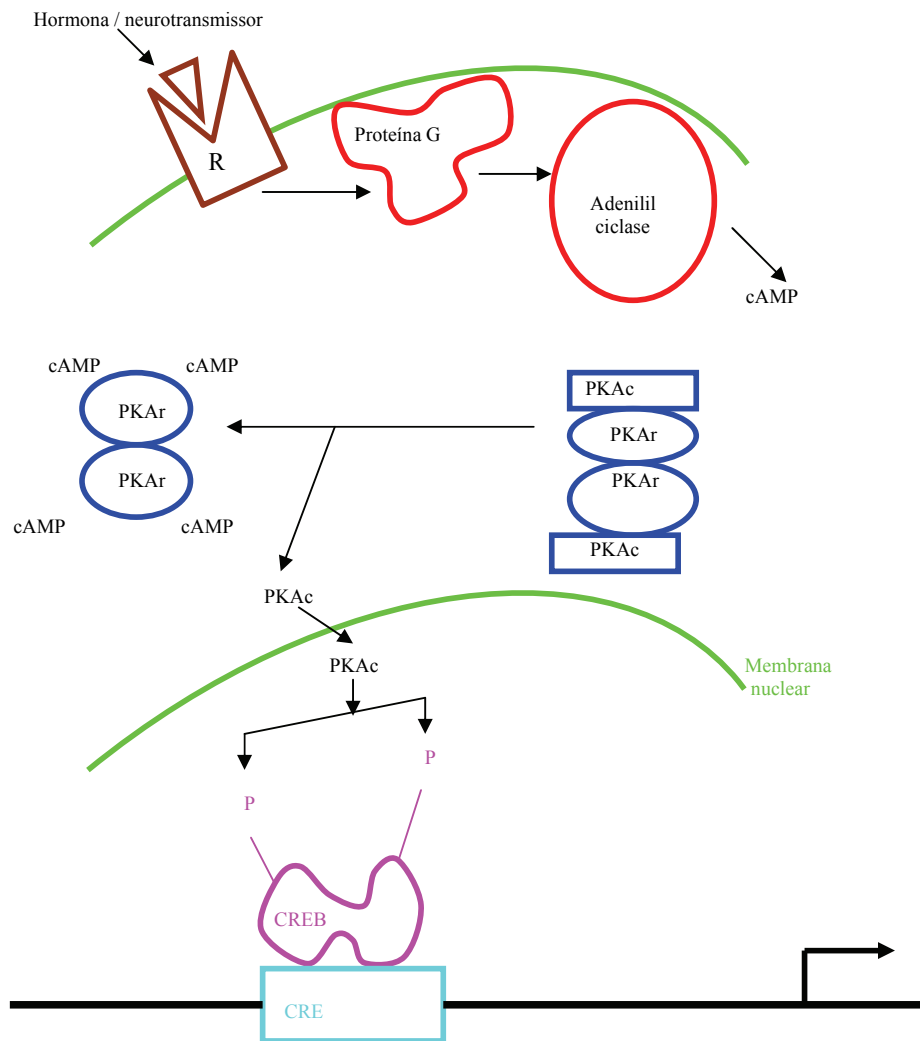
Esquematiza-se ainda que outro factor pertencente ao TF II D, oTAF130 interactua com o domínio Q2 (vide figura anterior com os domínios da CREB), e este domínio Q2 da CREB ainda interactua com o factor básico TF II B.

Um outro domínio da CREB fosfatada na serina 133,o domínio KID pode interactuar com o domínio KIX da CBP.

De todas estas interacções resultantes da activação da CREB origina-se a expressão dos genes, a partir dos sinais estimuladores extracelulares e com o concurso das cascatas de transdução de sinais, citoplásmicas e nucleares que desembocam na expressão dos genes correspondentes.

Na figura seguinte esboçamos novamente uma cascata de transdução de sinal em que o sinal, uma hormona ou um neurotransmissor, interactua especificamente com um receptor de sete hélices transmembranárias e através da G-proteína, adenilato ciclase, cAMP, PKAc, CREB fosfatado e CRE,activam a expressão dos genes correspondentes.

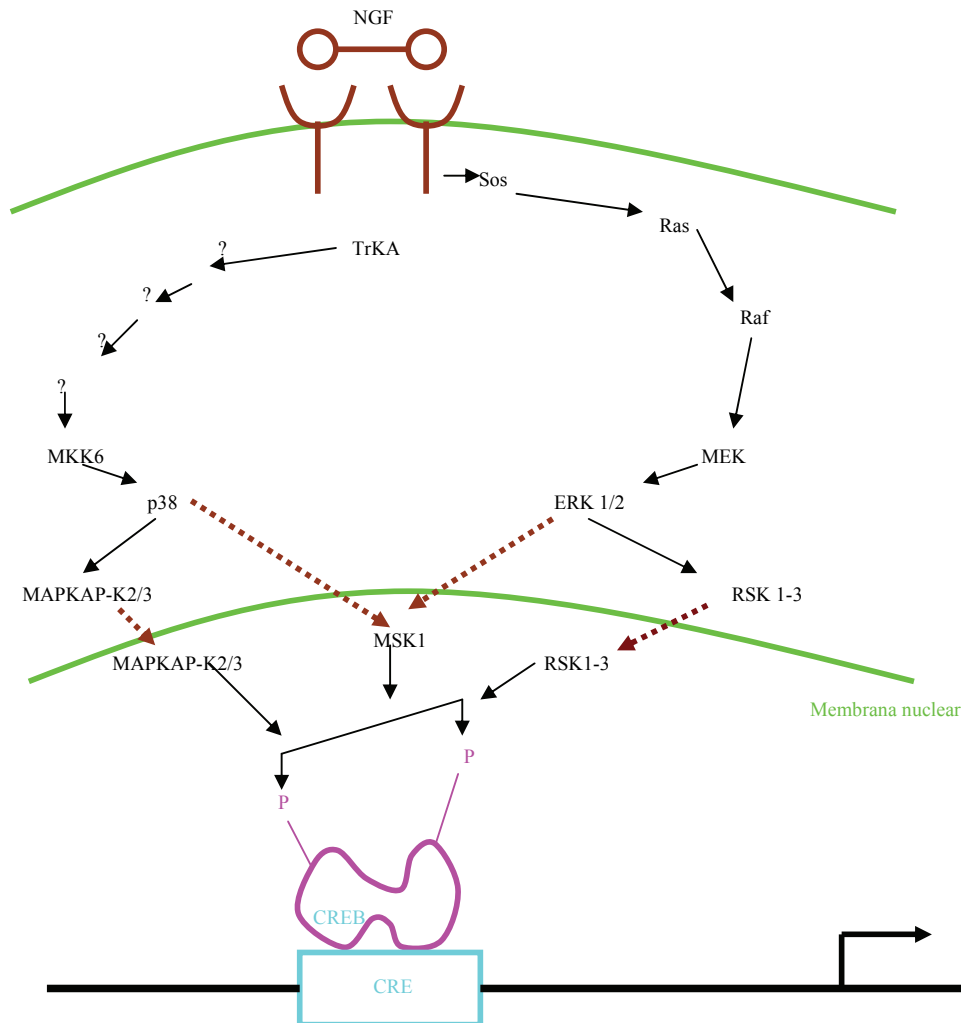
## Activação da CREB pela via assinalante cAMP/PKA



As hormonas e neurotransmissores ligam-se a receptores R e activam as proteínas G que directamente activam a adenilato ciclase que catalisa a produção de cAMP. Este cAMP activa a proteína cinase A (PKA). Esta PKA consiste de duas sub-unidades catalíticas (PKAc) e duas sub-unidades reguladoras (PKAr). A elevação do cAMP, liberta as subunidades PKAc activas que se translocam para o núcleo da célula onde fosfatam a proteína CREB na serina 133.

Esboçamos noutras cascatas de transdução de sinal como o ligando factor de crescimento dos nervos (NGF) pode induzir a fosfatação da serina 133 da CREB

Fosforilação da serina 133 da CREB por vias assinalantes activadas pelo NGF (factores de crescimento dos nervos)



Como se esquematiza na figura o sinal NGF interaccua na forma de dímero com o seu receptor específico, TrKA, activando este o qual por seu turno com o concurso de um complexo de adaptadores Grb 2/Shc não desenhado na figura vai por seu turno interaccuar com o factor Sos. Este factor Sos que interaccua com a Grb 2 promove a permuta do GDP ligado à proteína Ras por GTP, e esta Ras activada activa subsequentemente as cinases Raf, MEK, e MAPK, ERK ½.

São as ERK activadas que vão activar as RSKs (RSK 1-3) sendo estas proteínas translocadas para o núcleo onde também fosfatam a serina 133 da CREB.

Contudo pode haver uma variante na via, pois que as ERKs podem translocar-se para o núcleo e activar da mesma forma a CREB mas através da cinase MSK 1.

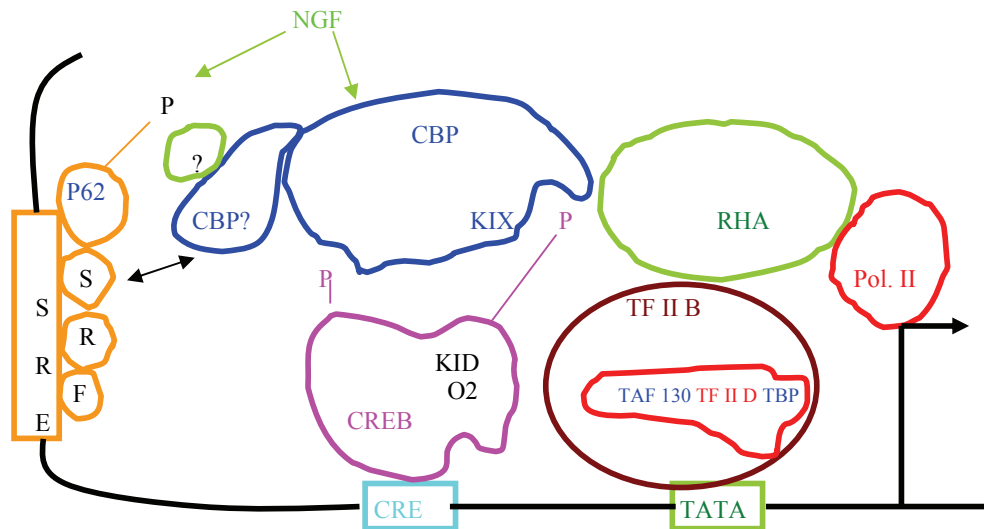
Noutra variante de transdução do mesmo sinal NGF, o receptor activado (lado esquerdo da figura) activa a MKK 6 uma cinase MEK “like” que activa por seu lado a proteína p38 MAPK e esta pode activar também a MSK 1 ou então outras cinases como as MAPKAPK 2/3, ambas activando a proteína CREB da mesma forma anteriormente referida.



Na figura as linhas a castanho indicam a translocação do citoplasma para o núcleo.

Vejamos agora como a proteína CREB activada em resposta ao NGF coopera com outros factores de activação da transcrição de certos genes como por exemplo o da expressão da proteína c-Fos.

A CREB coopera com outros factores na activação da transcrição na resposta ao NGF (factor de crescimento dos nervos)



A proteína c-Fos é uma fosfoproteína nuclear que interacciona não covalentemente formando um complexo, com o factor de transcrição c-Jun/AP-1, tendo a c-Fos um papel crítico na regulação do desenvolvimento das células pensando-se que contribuam para a proliferação celular e diferenciação.

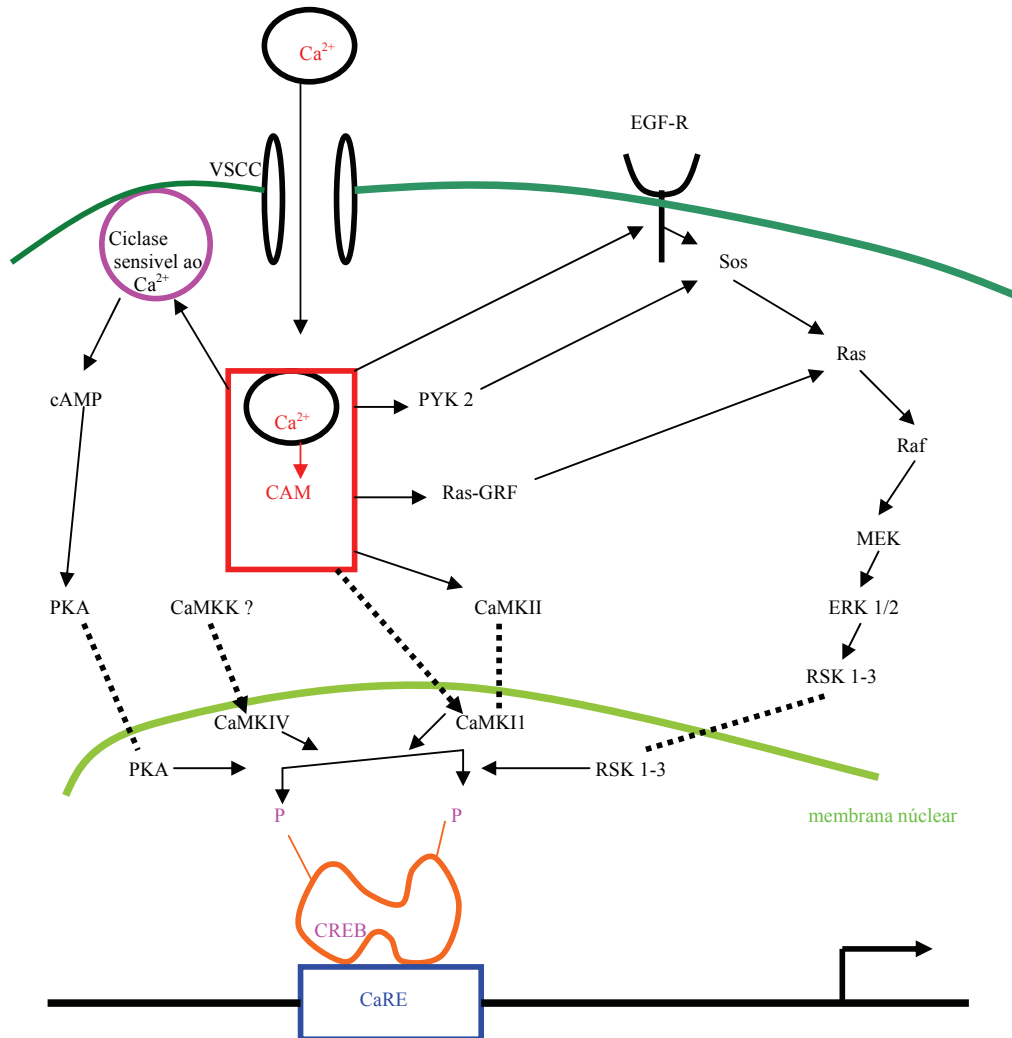
A expressão da c-Fos pode ser aumentada por diversos tipos de estímulos, tais como factores de crescimento, citocinas, neurotransmissores, hormonas neuropeptídicas, stress, agressões celulares, etc, pertencendo esta proteína c-Fos à família bZIP, subfamília Fos.

No caso da figura anterior a activação da transcrição do gene c-fos pelo NGF necessita de outros factores que interaccionem com outros elementos promotores além do CRE. Concretamente factores ligados acima dos elementos SRE são necessários para este efeito.

Na figura anterior esquematiza-se a cooperação entre o CREB e um complexo factor de transcrição, composto por um dímero de SRF associado com um factor P62 que é fosfatado pelo sinal NGF. Este complexo SRF-P62 fosfatado liga-se ao SER associando-se depois com proteínas tais como a CBP e outras para estabilizar estes coactivadores sobre o promotor.

O cálcio é um importante segundo mensageiro pleotrófico que activa uma diversidade de cascatas de transdução de sinal, tal como esquematizamos na figura seguinte, todas elas induzindo a fosfatação da CREB com as consequências inerentes.

Vias assinalantes múltiplas contribuem para a fosfatação da CREB em resposta ao influxo de  $Ca^{2+}$



A plasmamembrana dos neurónios por estimulação eléctrica sofre uma despolarização o que vai permitir a entrada de cálcio extracelular através dos canais de  $Ca^{2+}$  sensíveis a voltagem (na figura VSCC), aumento de cálcio este que vai activar diversas cinases de diversas cascatas de transdução de sinal, algumas delas fosfatando a serina 133 da proteína CREB com activação desta.

O cálcio entrado na célula neuronal complexa-se com a proteína calmodulina (CAM) e este complexo pode activar as diversas vias indicadas na figura em apreço.

Do lado esquerdo da figura esquematiza-s que o complexo  $Ca^{2+}/CAM$  vai activar a via da proteína cinase A (PKA) através da interacção com adenilato sensível ao cálcio. Já vimos anteriormente que a PKA pode translocar-se para o núcleo e aí fosfatar a serina 133 da CREB.

No entanto o complexo  $Ca^{2+}/CAM$  também activa outras proteínas da família das cinases dependentes do  $Ca^{2+}/CAM$  como é o caso da CAMK e todas estas cinases podem também fosfatar a CREB.

Na figura indica-se que o próprio complexo Ca<sup>2+</sup>/CAM pode translocar-se para o núcleo e aqui activar as CAMKIV e CaMKII, podendo mesmo algumas isoformas da CaMKII translocar-se do citoplasma para o núcleo.

No lado direito da figura evidencia-se que o complexo Ca<sup>2+</sup>/CAM pode também activar a via Ras/MAPK podendo contudo a activação da Ras ocorrer por diversos mecanismos. Assim o influxo de cálcio com a formação do complexo Ca<sup>2+</sup>/CAM pode levar à activação do receptor EGF-R independentemente do ligando EGF o que induz a activação do Sos e da Rãs.

A activação da Ras estimula a cascata cinasica Raf, MEK e ERK ½ e estas últimas activam directamente membros da família pp90 RSK, as proteínas cinases RSK-1 que activadas se translocam para o núcleo onde activam a CREB por fosfatação da serina 133.

Mas a Ca<sup>2+</sup>/CAM também pode activar a Ras através da Ras-GRF que é um factor de troca de nucleótidos de guanina.

Ou então a PyK2 tirosina cinase activada pelo Ca<sup>2+</sup>/CAM pode também activar a Sos e estimular a via Ras.

As linhas a tracejado na figura indicam translocação das moléculas em causa do citoplasma para o núcleo.

Refira-se no entanto que a actividade da CREB fosfatada pode ser regulada não só pelas referidas anteriormente PKA, CaMKs e RSKs mas também por outras cinases como é o caso da PKC, da GSK-3 (cinase da glicogénio sintase 3) e da caseína cinase II (CKII).

### **Cascatas de transdução de sinal e insulina**

#### Vias múltiplas transdutoras de sinal

A insulina ligando induz a sua cascata de transdução de cinases através de um receptor dotado de actividade de tirosina cinase (RTK) com características específicas mas a via desencadeada difere significativamente daquela desencadeada por outros RTK.

Salienta-se desde já que a insulina além de poder interactivar com o seu próprio receptor RTK, ou IR, pode também interactivar com outros receptores como é o caso do receptor para o IGF-1 (factor de crescimento insulina-like 1) que pode nessa interacção induzir inclusivé diversas acções promotoras do crescimento (por exemplo síntese do DNA).

Refira-se que este factor IGF-1 tem uma estrutura e sequência em ácidos aminados muito semelhante à da insulina, e a sua síntese é induzida por acção da hormona do crescimento (GH) sobretudo ao nível do fígado.

Estão descritos para a insulina efeitos desencadeados a curto preazo por exemplo sobre o metabolismo da glucose e ainda efeitos a longo prazo indutores do crescimento, tal como se refere no quadro seguinte.

## Acção da insulina

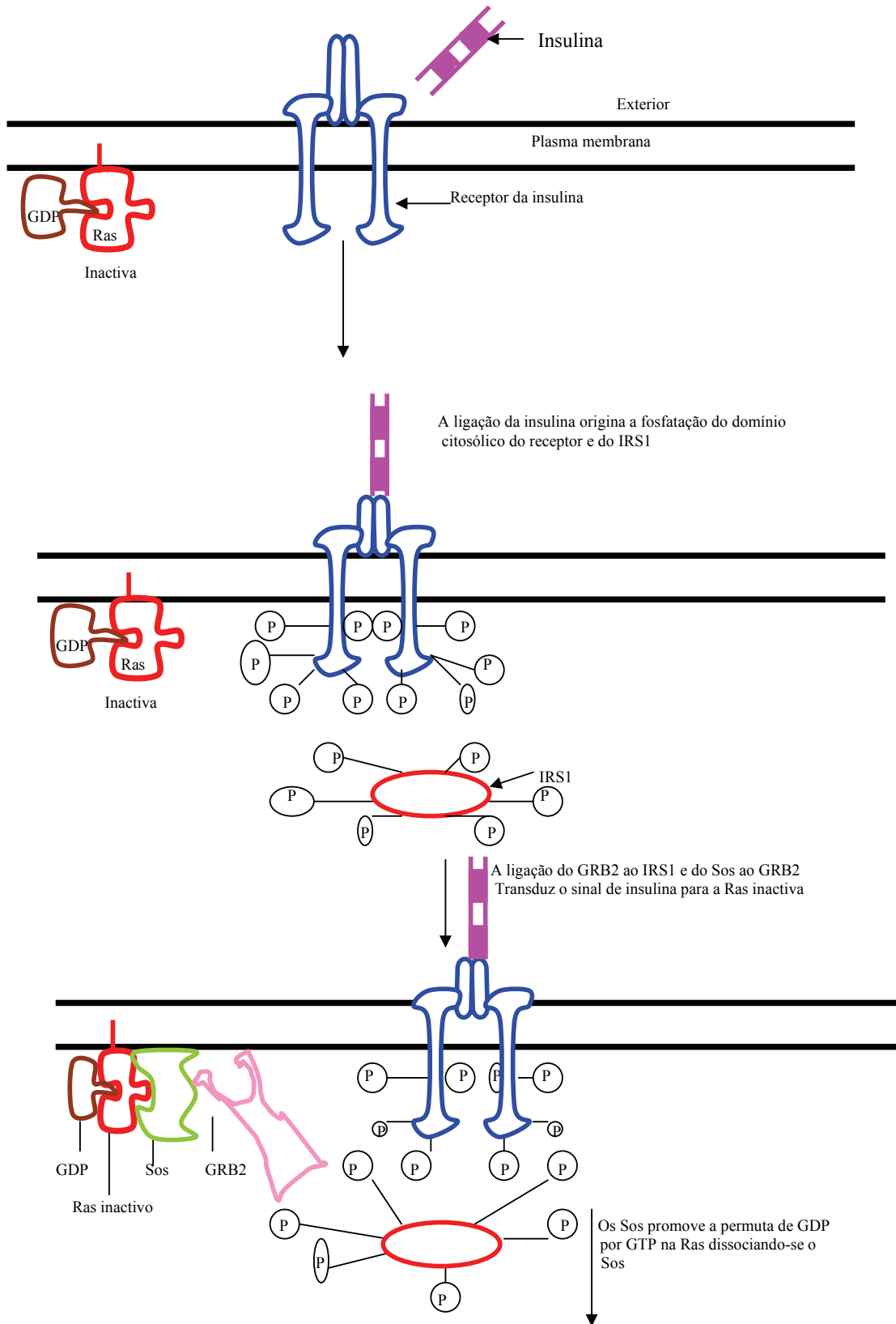
<u>Em segundos:</u>	Ligação ao receptor da insulina Autofosfatação do receptor Activação da proteína receptora tirosina cinase
<u>Em minutos:</u>	Activação do transporte de hexoses Alteração de actividades enzimáticas intracelulares Mudanças na regulação dos genes Internamento do receptor a que ligou a insulina e sua regulação Fosfatação do receptor da insulina por outras proteínas cinases
<u>Em horas:</u>	Indução de sínteses do DNA, RNA, da proteína e lipidica Crescimento celular Máxima regulação do receptor da insulina

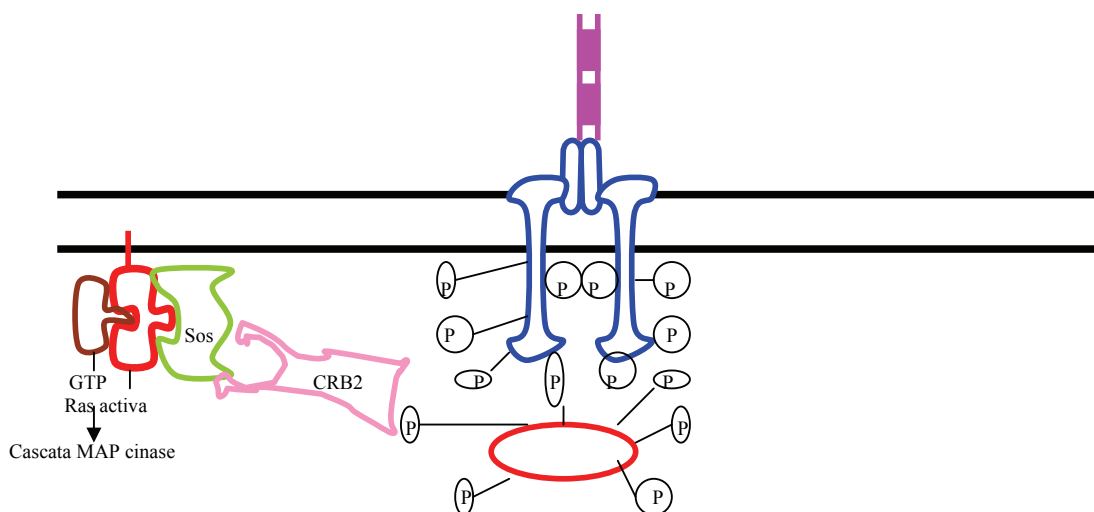
A cascata de transdução de sinal desencadeada pela insulina, envolve uma proteína solúvel substrato 1 receptor da insulina, a IRS1, que não se liga ao receptor (vide figura adiante) e que contém na sua estrutura resíduos de tirosina que são rapidamente fosfatados após indução pela insulina, podendo também sê-lo pelo receptor relacionado com o IGF1, ao contrário da maioria de outros RTK que não conseguem fazê-lo.

Após fosfatação da IRS1 esta proteína interacciona com domínios SH2 de diversas proteínas tais como as GRB2, PI-3 cinase e a Syp uma tirosina fosfatase.

Na figura seguinte mostra-se a activação da proteína Ras induzida pela ligação da insulina ao seu receptor que é pois um receptor tirosina cinase atípico.

Ativação da proteína Ras induzida pela ligação da insulina ao seu receptor Tirosina cinase atípico





A Ras activada desencadeia a mesma cascata com a proteína cinase que referimos na figura anterior.

A ligação da proteína Syp ao IRS1 aumenta a actividade fosfatase cujo papel não está bem demonstrado.

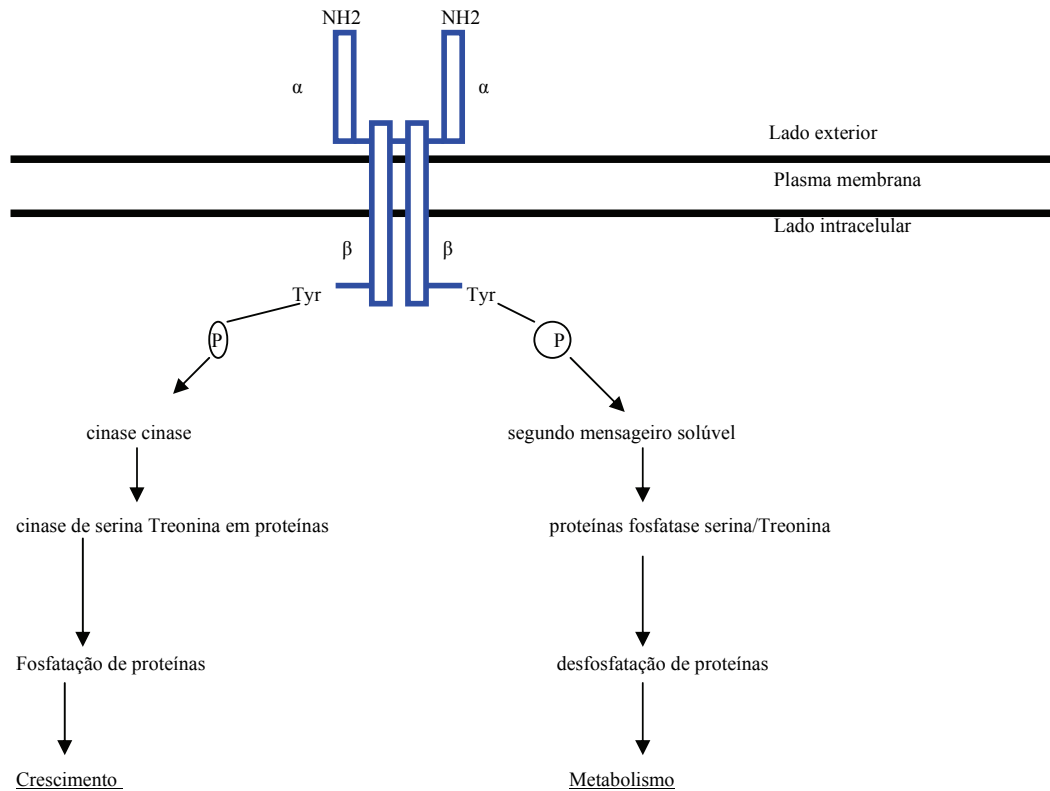
### Receptores de insulina e transdução de sinal através de tirosina cinases

Nos tetrameros receptores da insulina as duas subunidades  $\alpha$  têm uma localização inteiramente extracelular, sendo a este nível que ocorre a interacção com a hormona sinal. Por outro lado estas subunidades  $\alpha$  estão interligadas através de pontes dissulfureto com as subunidades  $\beta$  de localização predominantemente transcelular e intracelular. O ligando ao interactuar com as subunidades  $\alpha$  induz a alteração da conformação de todo o complexo formado com o receptor o que propicia a autofosfatação dos resíduos de tirosina existentes do lado citoplásmico das subunidades  $\beta$  (vide figura adiante), indo esta autofosfatação induzir a fosfatação de outras proteínas (por exemplo a tomada aumentada de glucose pela célula) ou efeitos a longo prazo (por exemplo sobre a diferenciação e crescimento celular).

Salienta-se que os locais de autofosfatação nos receptores da insulina, são resíduos de tirosina, mas já nas proteínas citosolicas onde também é induzida a fosfatação, são sobretudo resíduos de serina e treonina os fosfatados (vide figura seguinte) o que corresponde a activação de cinases, podendo também ocorrer a activação de proteínas fosfatases.

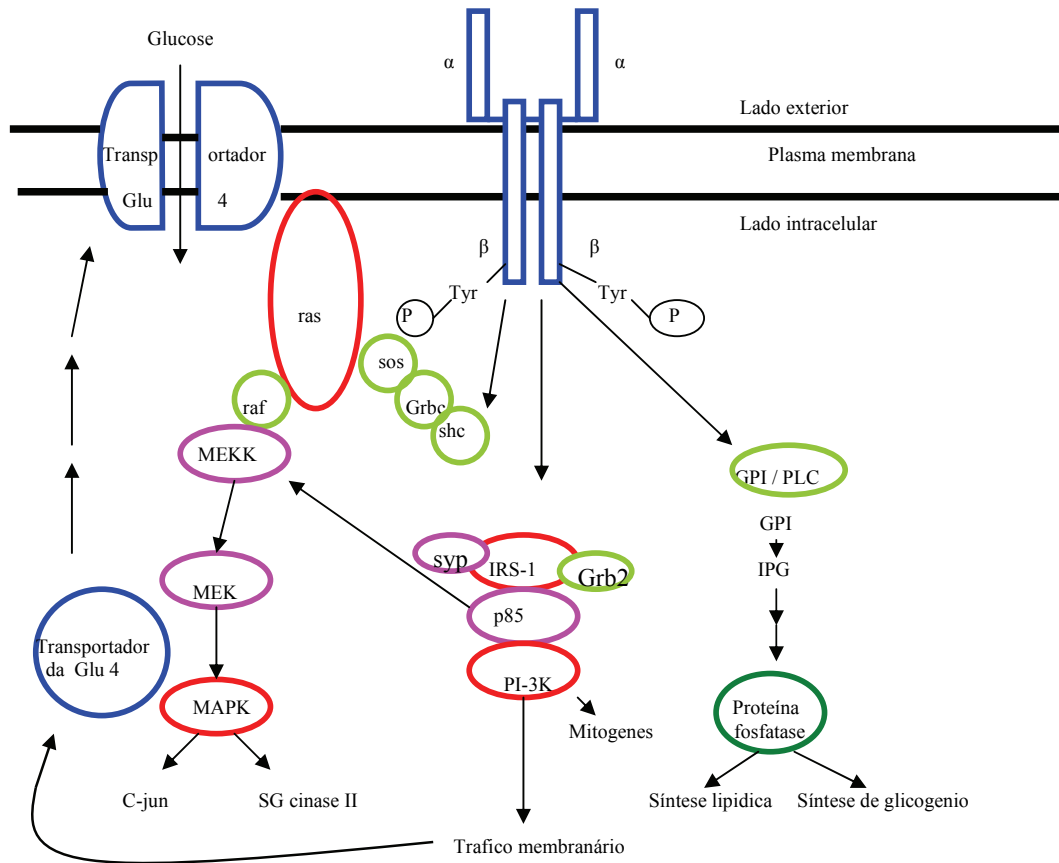
Na figura seguinte verifica-se que o receptor activado da insulina produz um segundo mensageiro solúvel que activa directa ou indirectamente serina e treonina de proteínas fosfatases (à direita na figura), tal como estimula uma cascata de proteínas cinases (à esquerda na figura) que induz por seu turno a fosforilação de uma série de proteínas.

Possíveis efeitos da insulina sobre a fosfatação de proteínas em duas vias metabólicas distintas



Esta série de acontecimentos pode levar à activação ou inibição de enzimas específicas mediando assim a acção da insulina que desencadeia uma série de acções pleiótricas como se admite hipoteticamente na figura seguinte.

## Transdução de sinais induzida pela insulina



IRS = substrato 1 do receptor da insulina

p85  
Syp  
Grb2 } Proteínas contendo domínios SH2

MAPK = proteína cinase ativadora da mitogénese

GPI = glicosil fosfatidilinositol

PLC = fosfolipase C

IPG = inositol fosfato glicana

Nesta última figura são delineadas diversas vias para transdução de sinais induzidas pela insulina.

Assim a autofosforilação das tirosinas do domínio citosólico do receptor induzem a fosforilação de substratos intracelulares como é o caso do IRS-1 e de proteínas SHC que podem depois interactuar com outras proteínas que contenham domínios SH2 como é o caso, na figura, da p85, da Syp e da Grb2.

O complexo formado pela IRS-1 e p85 activa depois a PI-3cinase (PI-3K) e o complexo IRS-1 e Syp activam a Syp activando a MEK.



Noutra via a formação do complexo SHC e Grb2 estimula a P21 Ras ligando GTP induzindo uma cascata de fosfatação provavelmente com intervenção da raf-.proto-oncogenes, da MEK, de MAP cinases e SG cinase II.

Noutra via ainda (à direita na figura) o receptor da insulina induz provavelmente a activação dde uma fosfolipase C específica que leva à hidrólise de moléculas de glicosil-PI na plasmamembrana, e o produto resultante desta hidrolise, o inositol fosfato glicana é um segundo mensageiro que vai activar serinas/treoninas proteínas fosfatases que regulam o metabolismo dos lípidos e da glucose. Esta activação aumenta o metabolismo da glucose e a sua inibição diminui a hidrólise do depósito de glucose ou de ácidos gordos.

### **Cascatas de transdução de sinal e IGFs**

Os factores de crescimento "insulina-like" (IGF) intervêm na modelação do crescimento e no metabolismo sendo regulados pela hormona do crescimento (GH) e por outros modeladores endócrinos, existindo duas formas de IGFs, a IGF 1, e a IGF 2. São ambas simples cadeias polipeptídicas com uma estrutura similar á da proinsulina contendo um c-peptido intacto e três pontes dissulfureto.

Dado o seu interesse para aliviar situações patológicas ou para favorecer a produção animal, têm sido produzidas por métodos de DNA recombinante e são comercializadas as hormonas do crescimento (GH) e as IGF.

A estrutura primária das IGF dos diversos mamíferos encontra-se altamente conservada. Assim a IGF-1 dos bovinos e suínos é idêntica à IGF-1 humana, com 70 ácidos aminados, e um pI de 8.5 embora sem conterem histidina nem triptofano.

A IGF-1 de ovino apresenta na sua estrutura primária e na posição 66 uma prolina, em vez de alanina.

Quanto a IGF-1 aviária ela já difere em oito ácidos aminados, da IGF-1 humana, enquanto a IGF-2 só difere em dois ácidos aminados.

A IGF-2 humana contem 67 ácidos aminados, tem um pI < 6.5 e difere da IGF-2 porcina em um ácido aminado e da IGF-2 bovina em três e da congénere ovina em quatro.

### **Os receptores IGF**

Os receptores para a IGF-1 e para a IGF-2 diferem na sua estrutura pois enquanto os primeiros têm muita analogia com os receptores para a insulina (IR) sendo glicoproteidos heterotetrameros com duas subunidades  $\alpha$  e duas subunidades  $\beta$ , já os segundos são constituídos apenas por uma cadeia glicoproteica.

No receptor IGF-1 as subunidades extracelulares e intracelulares são unidas por ligações covalentes pontes dissulfureto e da mesma forma a interacção do efector com este receptor promove a autofosfatação das subunidades  $\beta$  nos seus resíduos de tirosina o que leva posteriormente à fosfatação de outras proteínas citoplásmicas contendo tirosinas, sendo a proteína IRS-1 a principal destes substratos.

Recorda-se que a proteína IRS-1 é induzida para fosfatação pelos dois tipos de receptores, os da IGF-1 e da insulina.

Este IRS-1 fosfatado analogamente pode activar outros mensageiros intracelulares, como é o caso da fosfoinositol-3'-cinase e outras proteínas contendo domínios SH2.

O receptor IGF-2 é um receptor de cerca de 250kDa idêntico ao receptor manose-6-fosfato.

Nos frangos este receptor-6-fosfato não interacciona com o IGF-2 e as acções do IGF-1 e IGF-2 são mediadas aparentemente pelos receptores do tipo IGF-1 e IR.

### IGF como mediador da acção da GH

Nos animais domésticos, ovinos, bovinos e suínos a síntese e segregação da IGF-1 é activada pela GH, parecendo as aves ser uma excepção.

Contudo há algumas características curiosas pois a secreção de GH é pulsátil, mas os níveis de IGF-1 no soro sanguíneo são relativamente constantes.

Também se pensou a princípio que os IGF tinham apenas um papel endócrino, mediando os efeitos da GH, mas hoje sabe-se que eles têm efeitos locais independentes do sistema circulatório e dos efeitos endócrinos, aceitando-se contudo que os IGF circulantes regulam o crescimento e metabolismo dos animais.

Assim os níveis de IGF-1 estão correlacionados com o peso do corpo e velocidade de crescimento de suínos, ovinos, bovinos e canídeos.

Os IGFs em circulação são sintetizados e segregados sobretudo pelo fígado, apesar de outros tecidos fetais e adultos o fazerem também e quando o fazem é para actuação no local de biossíntese ou em células adjacentes (acção paracrina) ou nas próprias células que os produzem (acção autocrina).

No caso do tecido muscular esquelético, sabe-se que ele exprime IGF-1 e que o seu m-RNA nestes músculos do esqueleto aumenta em resposta a injecções de GH.

Estas IGF-1 regulam aqui o crescimento muscular e desempenham portanto um papel no crescimento dos animais, o que está bem demonstrado.

In vitro e in vivo a IGF-1 estimula o anabolismo proteico e inibe a proteólise ao nível do tecido muscular esquelético.

No caso das aves (frangos) os IGF 1 e 2 parecem estar mais envolvidos no metabolismo lipídico do que na regulação do crescimento.

Para o IGF-2 está assinalada uma acção de crescimento durante o desenvolvimento fetal.

Resumindo pode afirmar-se que os IGFs e a insulina têm estruturas e funções biológicas semelhantes e se a insulina interagindo com o receptor da IGF-1 promove o crescimento e desencadeia efeitos metabólicos ao nível da glucose, lípidos e prótidos, também concentrações elevadas de IGF-1 podem imitar os efeitos metabólicos da insulina quando interacciona com o receptor IR da insulina.

A família de proteínas IGF inclui três conhecidos ligandos, os IGF-1, IGF-2 e a insulina, seis proteínas de ligação (IGFBP-1 a 6) que veremos a seguir e receptores à superfície das células.

A homologia das estruturas da IGF-1 e IGF-2 com a insulina é de cerca de 50% e os receptores para a IGF-1 e para a insulina (IR) revelam 60% de identidade dos seus ácidos aminados.

Ao contrário da insulina as IGFs em circulação estão ligadas a proteínas com alta afinidade (IGFBPs) das quais são conhecidas seis o que parece indicar que o sistema IGF é complexo.

### Proteínas que ligam IGFs (IGFBP)

Os IGFs actuam na modelação do crescimento e metabolismo englobados num sistema mais complexo, sendo regulados pela GH e por outros modeladores endócrinos que elevam ou baixam as concentrações locais e sistémicas dos IGF, e também pela presença de proteínas que ligam especificamente as IGF e que se chamam as IGFBP.

Estas IGFBP no sangue sequestram as IGFs prolongando a sua semi-vida na circulação.

As IGFs ligadas às IGFBP são menos activas do que as formas livres da IGF, embora as IGFBP também possam favorecer a actividade das IGFs.

Estão caracterizadas seis, IGFBP de 1 até 6. Todas elas são similares na sua estrutura global (200 a 300 ácidos aminados com sequências altamente conservadas em variadas espécies animais nas porções N e C terminal ocorrendo maior variabilidade no meio da estrutura).

Todas as seis IGFBPs formam complexos binários de 30 a 50 KDa com as IGFs. Por outro lado a IGFBP-3 pode ligar-se à IGF para formar um complexo binário ou combinar-se com uma proteína maior, a ALS (sub-unidade ácido lábil) e IGF formando um complexo ternário de 150 KDa.

As IGFBPs têm vários modos de acção (endocrina, parácrina e autócrina) e dois mecanismos de acção distintas, estimulador ou inibidor, sendo expressas em quase todos os tecidos, excepto a IGFBP5 cuja expressão é muito escassa ou até nula no fígado.

As IGFBPs solúveis atenuam a actividade da IGF e as IGFBPs que estão associadas com células podem atenuar ou favorecer a actividade IGF.

A maioria das IGFs sistémicas está ligada à IGFBP-3 na forma de complexo ternário e este complexo não atravessa o endotélio capilar o que aumenta a sua semi-vida em circulação.

Os complexos binários IGF-IGFBP atravessam o endotélio capilar e transportam IGF ligada para tecidos alvos.

As IGFBPs solúveis ligam IGFs com alta afinidade evitando que estes se liguem aos receptores na superfície das células inibindo desta forma a actividade do IGF.

As IGFBP-1,2,3 e 5 favorecem a actividade mitogénica das IGFs in vitro devido à sua capacidade (das IGFBP) para aderirem à superfície das células e ligarem IGF-1.

As formas e proporções relativas de IGFBP presentes nos fluidos extracelulares podem determinar o efeito das IGFs.

Por exemplo a IGFBP-5 associada com o colagénio ou ECM tem uma menor afinidade para a IGF-1 do que a IGFBP-5 na forma solúvel e daí potenciar naquele caso a actividade mitogénica da IGF-1.

Também a actividade da IGFBP pode ser regulada por fosfatação em resíduos serina ou glicosilação. As formas fosfatadas têm uma afinidade muito mais elevada para os IGF do que as formas não fosfatadas.

Parece que as IGF controlam a degradação das IGFBP.

As IGFBP-5 solúveis são mais rapidamente degradadas na ausência de IGF, enquanto as IGFBP-5' dentro da ECM permanecem intactas.

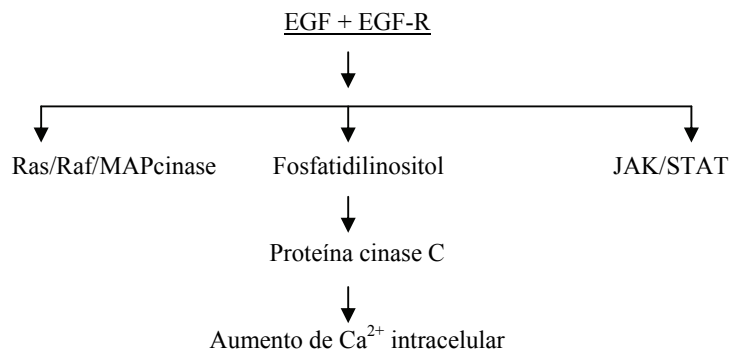
Anote-se que a designação IGF-1 corresponde a somatomedina C.

### Cascatas de transdução de sinal e EGF

O factor de crescimento epidérmico (EGF) é uma cadeia polipeptídica com 53 ácidos aminados, que regula a proliferação das células ligando-se nestas ao receptor EGF (EGF-R) sendo este receptor uma proteína transmembranária com actividade intrínseca de tirosina cinase, actividade esta que é activada quando o EGF se liga ao respectivo receptor o qual sofre autofosfatação. A união do EGF ao seu receptor promove a dimerização deste. Após esta autofosfatação de um receptor ocorrem interacções com diversos substratos através de domínios SH2, como é o caso da proteína adaptadora Grb2 e do complexo Grb2-SOS.

Estes acontecimentos originam a fosfatação da Ras (uma proteína que se liga ao GTP) e a activação da via Ras/Raf/MAPcinase, que por seu turno origina a fosfatação de proteínas reguladoras e de factores de transcrição que levam a proliferação celular.

As outras vias activadas pelo EGF envolvem a via do fosfatidilinositol (levando a activação da proteína cinase C e a um aumento do cálcio citosólico) e a via assinalante JAK/STAT.



As cascatas de transdução de sinal desencadeadas pelo EGF podem ser portanto: a) a via através do fosfatidilinositol que origina activação da proteína cinase C induzindo o aumento do Ca<sup>2+</sup> intracelular; b) a via através do Ras originando uma activação da MAP cinase;c)a via JAK-STAT.

O receptor EGF é uma proteína que se liga a actina, originando uma rápida despolimerização desta e a formação de “pregas” membránarias que parecem constituir estruturas para transdução do sinal.

Por outro lado o EGF e o seu receptor têm um papel a desempenhar ao nível do núcleo.

A expressão do c-fos e do c-jun são induzidas pelo EGF e esta expressão é poderosamente inibida pela microgravidade, parecendo que o citoesqueleto constitui o componente celular sensível a gravidade.

O receptor da EGF é uma glicoproteína com 130kDa (1186 ácidos aminados) e 40kDa de cadeias hidrocarbonadas, sendo este receptor constituído por quatro domínios; um domínio extracelular glicosilado onde se liga o EGF, um domínio transmembranar, um domínio tirosina cinase e um domínio de autofosfatação.

A ligação do EGF ao seu receptor origina a replicação de DNA e a proliferação celular, ocorrendo durante esta etapa o pregeamento da membrana celular, a fosfatação do EGF-R, o internamento, mudança de pH, activação enzimática, reorganização do filamento de actina e indução de proteína oncogene.

A semi-vida do EGF-R é de cerca de vinte horas mas quando interactua com o EGF a degradação passa a ser mais rápida em cinco horas.

#### Estrutura do receptor EGF e suas posições de fosfatação

-241							619	644	694			944					1186
			N					T					Y	S			Y
	N			N	N	N		T					S		Y		
	N		N		N	N		S						S		Y	Y
Péptido sinal	Região de ligação do EGF região de admissão de carboidratos						Domínio transmembranário			Domínio tirosina cinase			Domínio de autofosfatação				
N = Asparagina T = Treonina S = Serina Y = Tirosina																	

## Cascatas de transdução de sinal e SRF

### Activação do factor de transcrição SRF

As células transduzem sinais extracelulares em programas específicos da expressão dos genes.

É o caso da interacção cooperativa entre o factor de resposta ao soro (SRF) e algumas proteínas com o motivo estrutural homeodomínios. Está provado que certas proteínas com homeodomínios recrutam SRF para se ligar a sequências do DNA que não são reconhecidas apenas pelo SRF.

A maioria das cascatas ou vias de transdução de sinais citoplásmicas são compostas por proteínas expressas ubiquamente e ligadas a diversos tipos de receptores.

Distintos receptores podem desencadear sinais intracelulares idênticos.

O factor de resposta ao soro (SRF) é um factor de transcrição que se liga ao elemento de resposta ao soro (SRE) que se encontra associado com uma série de genes tais como o c-fos, fosB, junB, erg-1 e 2, genes neuronais e genes dos músculos tais como os das actinas e miosinas, podendo a expressão destes genes ser regulada pelo SRF controlando assim este factor o crescimento e diferenciação celular, a transcrição pelos neurónios e o desenvolvimento dos músculos.

O SRF pode ser activado por uma série de factores (lipopolissacáridos, citocinas, TNF $\alpha$ , etc.) e também por estímulos extracelulares como antioxidantes e luz U.V.

O SRF é regulado por vias de transdução de sinais celulares e por interacções com outros factores de transcrição ( por exemplo Sp1 e factores reguladores miogénicos).

A activação do SRF depende de variações na dinâmica da actina controladas pelo Rho.

Parecem existir duas classes de genes alvo SRF, uma sensível à via Rho-actina e outra sensível à via MEK-ERK. O factor de transcrição SRF controla diversos genes de resposta imediata, cuja transcrição é induzida por sinais extracelulares, havendo diversos genes constitutivos expressos no tecido muscular.

A actividade do SRF é pois regulada por vias de transdução de sinal celular e por interacção com outros factores de transcrição como por exemplo Sp1, ATF6 e factores reguladores miogénicos.

O SRF humano e do rato (*Mus musculus*) já conhecidos são glicoproteínas de 508 e 504 ácidos aminados respectivamente com uma estrutura primária muito parecida, cuja função é como se disse a de um factor de transcrição que se liga ao elemento de resposta ao soro (SRE) uma curta sequência de simetria diada localizada a 300 pares de bases do início da transcrição de alguns genes (como por exemplo o c-fos).

Liga-se ao DNA como um multímero provavelmente como um dímero e tem uma localização nuclear pertencendo à família de domínios MADS dos factores de transcrição. Os domínios MADS (MCM1, AGAMOUS, DEFICIENS e SRF) são uma região conservada de ligação/dimerização ao DNA, presente numa série de factores de transcrição. Na estrutura do SRF os resíduos de ácidos aminados 168 a 222 estão implicados na dimerização o domínio compreendendo os resíduos 143 a 197 são denominados MADS e nos 277, 307, 309, 316, 383 ligam-se carbohidratos.

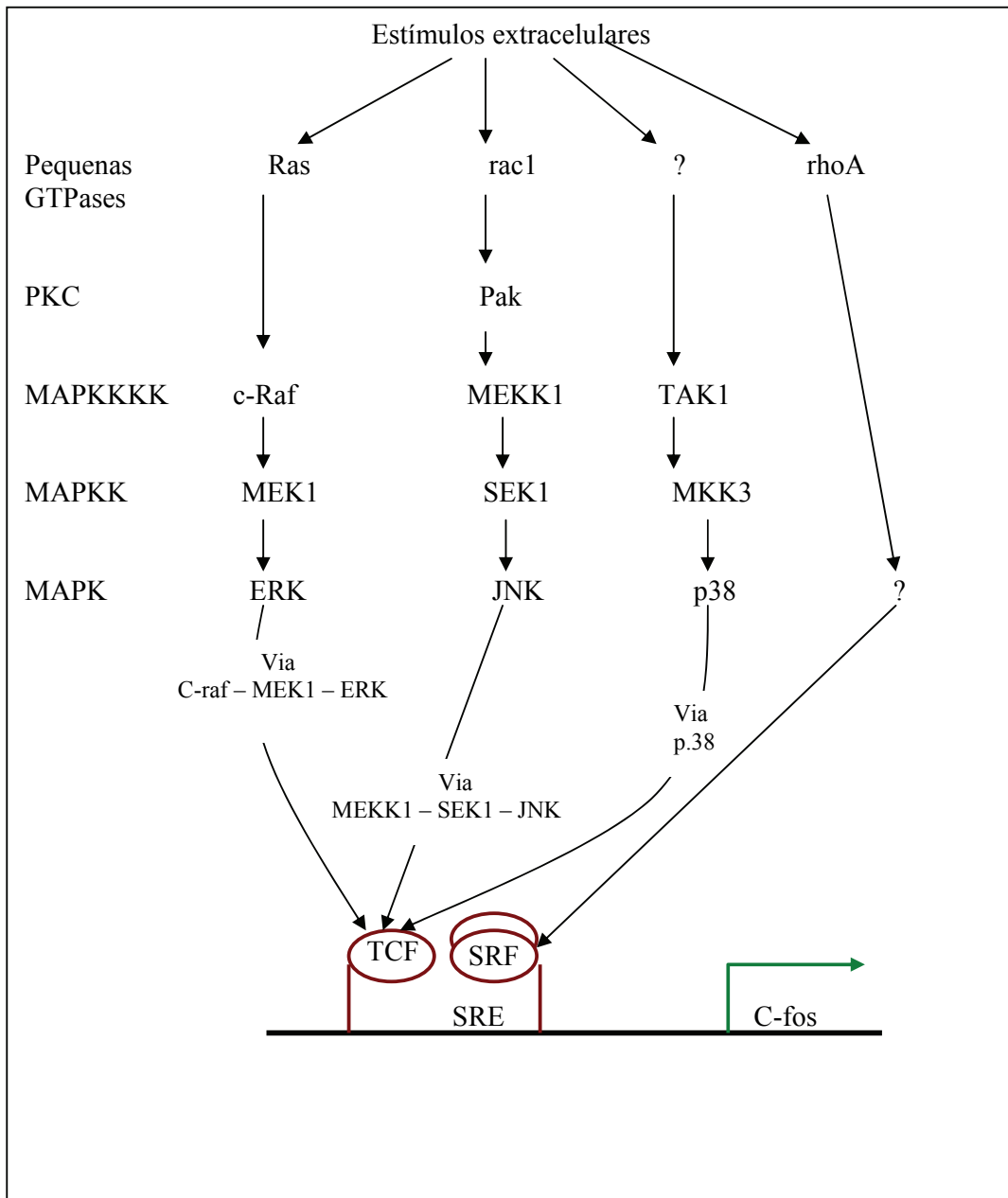
O SRF de frangos (*Gallus gallus*) embora com as mesmas características dos SRF humano e do rato têm uma estrutura primária muito diferente como é natural.

Resumindo o SRF é um factor de transcrição que controla diversos genes de resposta imediata, cuja transcrição é induzida por sinais extracelulares e que controla também diversos genes constitutivamente expressos nos músculos, sendo a actividade do SRF regulada por vias de transdução de sinal celulares e pela interacção com outros factores de transcrição.

Por exemplo nos promotores de resposta imediata fos e erg1, o SRF forma complexos ternários com membros da família TCF. O SRF coopera com uma série de outros factores de transcrição como o Sp1, ATF6, GATA4, NKx2.5 e factores mitogénios.

As Rho GTPases activam o SRF através da sua capacidade para induzir a deplecção do “pool” de G-actina.

Na figura seguinte esquematizam-se as vias de transdução de sinal envolvendo proteínas cinases C (PKC) que levam a activação de SRF ou da via através do rho A levando à activação do SRF.



As três vias da esquerda na figura levam a activação dos factores de transcrição TCF e SRF que formam um complexo ternário com o elemento promotor SRE do gene C-fos.

A via rho-A leva à activação do factor de transcrição SRF.

O factor de transcrição TF II.1 favorece a activação do promotor c-fos através da interacção com elementos acima (upstream) dos activadores SRF, STAT1 e STAT3.

Está provado que culturas quiescentes de células mamíferas na fase Go por adição de factores de crescimento (por exemplo EGF ou PDGF) apresentam um aumento rápido da expressão de mais de 100 genes diferentes, os genes de resposta precoce induzida antes das células entrarem na fase S e replicarem o seu DNA.

Um destes genes de resposta precoce codifica um factor de transcrição c-Fos que com outras proteínas leva as células a avançar no seu ciclo celular.

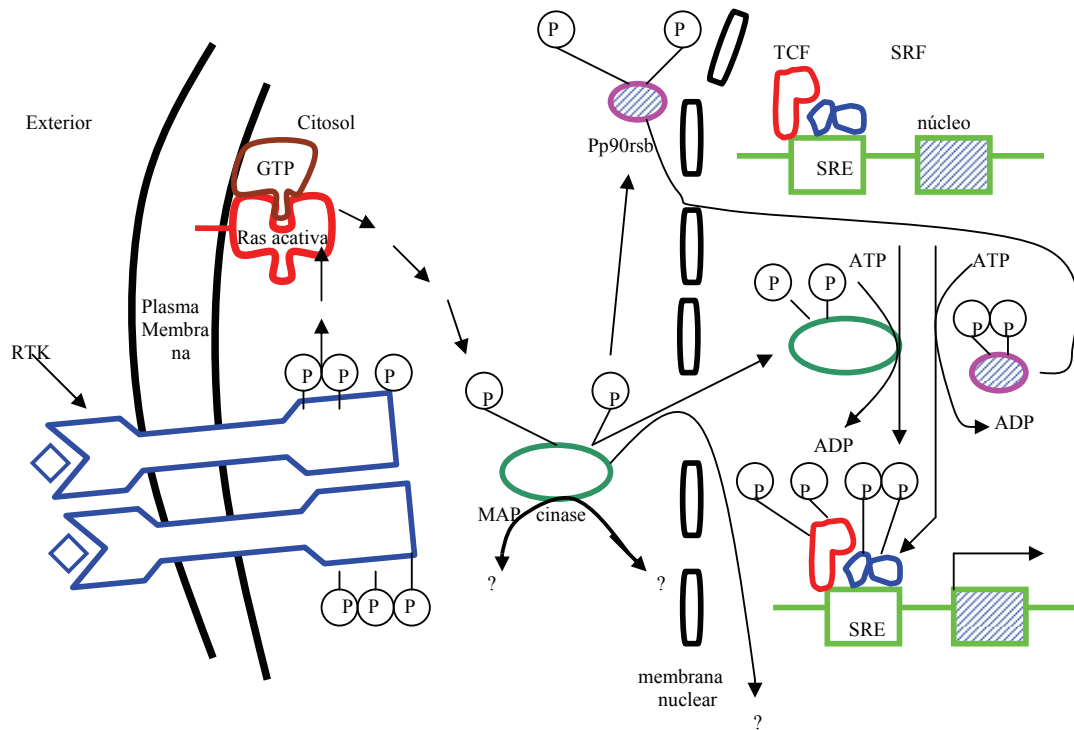
A indução do gene c-fos por vários factores de crescimento é mediado conforme vimos em figura anterior por várias vias intracelulares de transdução de sinal levando a activação de diferentes proteínas cinases, aquelas envolvendo cAMP (proteínas cinases dependentes de cAMP), Ca<sup>2+</sup>/DAG (proteína cinase C) e proteína cinase Ras (MAP cinase).

A sequência reguladora do gene c-fos contém como vimos um elemento de soro resposta (SRE) assim chamado porque ele é activado por diversos factores de crescimento no soro.

A ligação do SRF e proteínas associadas aos SREs activam a expressão de genes. A via RTK-Ras activa uma cascata MAP cinase e esta translocada para o núcleo fosforila locais específicos no domínio C-terminal de uma proteína o factor complexo ternário (TCF) que depois se associa com duas moléculas de SRF formando um complexo factor activo trimérico para ligação ao DNA (vide figura seguinte). O SRF e TCF podem permanecer ligados aos promotores in vivo.



Via de transdução de sinais levando à activação do factor de transcrição SRF e modulação da expressão de genes (via RTK – Ras)



2.4.1.3 - Transdução de sinais e tirosina cinases citosolicas

**Famílias de receptores para outras citocinas**

As citocinas são pequenas proteínas (por exemplo interferões, interleucinas) e muito numerosas que se ligam aos receptores da superfície celular de certas células para desencadear a sua diferenciação e proliferação. Algumas citocinas são chamadas linfocinas porque regulam a intensidade e duração da resposta imunitária como por exemplo IL-2, o interferão  $\gamma$  ( $IFN\gamma$ ), IL-4, IL-3, GM-CSF, IL-6, IL-7, segregadas por células  $T_H$  (timo) em resposta ao antígeno.

<u>Citocinas</u>
Efeitos autócrinos e / ou parácrinos
<ul style="list-style-type: none"> <li>- Influenciam o metabolismo e as actividades celulares</li> <li>- Controlam a proliferação e diferenciação celular</li> <li>- Controlam a hematopoiese e a resposta imunológica, inflamatória, fagocitária e citotóxica</li> <li>- Remodelação óssea</li> <li>- Reparação de feridas</li> </ul>

### Classes de citocinas

Factores de crescimento (exemplos EGF, PDGF, FGF, IGF, NGF).

Linfocinas (exemplos IL 1 a 8).

Factor indutores de colónias (exemplos GM-CSF, G-CSF, M-CSF, EPO, LIF).

Factores de transformação e crescimento (exemplos TGF $\alpha$  e  $\beta$ ).

Factores de necrose tumoral (exemplos TNF  $\alpha$  e  $\beta$ ).

Interferão (exemplos IFN  $\alpha$ ,  $\beta$  e  $\gamma$ ).

### Transdução de sinais das citocinas (Principais mecanismos de acção intracelular)

- Proteína G-cAMP-proteínas cinase A
- Tirosina cinases
- Fosfolipases A2, C e D
- Inositol fosfato-cálcio intracelular
- Diacilglicerol (DAG)-proteínas cinase C
- Araquidonato-eicosanoides

A família de receptores que se ligam a citocinas é uma família muito grande e heterogénea que pode ser dividida em dois grupos ou classes. A classe I que engloba os receptores e ligandos hematopoiéticos IL-2, 3, 4, 5, 6,7, 9, 11, 12-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTP e trombopoietina.

A classe II que é mais divergente e inclui os receptores para o IFN  $\alpha$ ,  $\gamma$  e IL-10.

Os receptores hormonais como vimos permitem que sinais exteriores sejam traduzidos em acontecimentos tais como a proliferação celular, a sua diferenciação ou até a morte.

A superfamília das citocinas da classe I hematopoiética consiste de mais de vinte citocinas e factores de crescimento conhecidos, e respectivos receptores, possuindo vários aspectos estruturais comuns e funcionais.

Estas hormonas extracelulares activam os seus receptores transmembranários levando estes a oligomerizar-se. Estes receptores oligomerizados por sua vez activam moléculas de tirosina cinases intracelulares as quais por ser turno activam factores de transcrição (ou sejam vias JAK-STAT).

Vejamos na figura seguinte, consoante o tipo de ligando interveniente, as características dessa oligomerização e da transdução de sinal desencadeada.

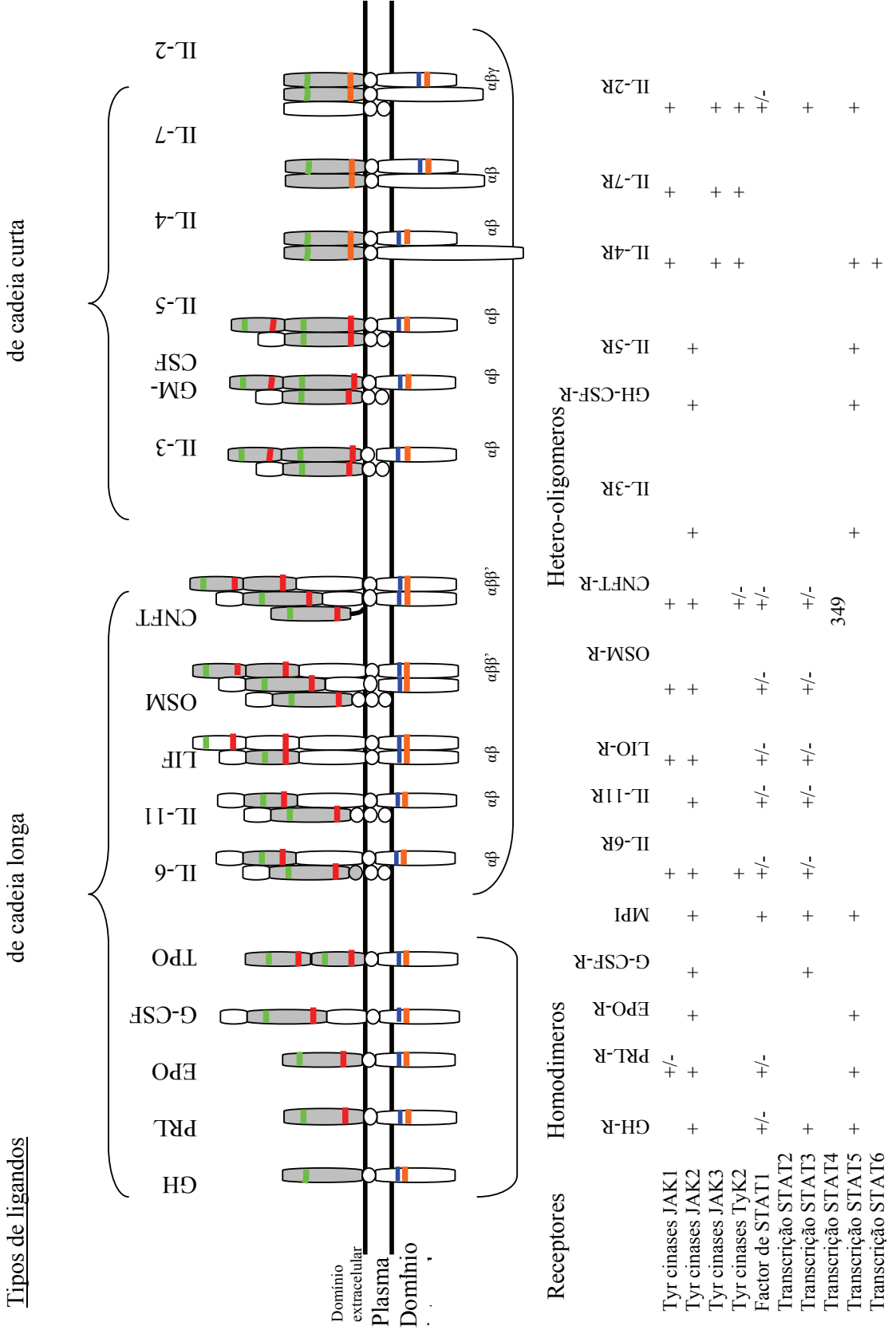
Todos os ligandos têm pelo menos dois locais para ligar duas moléculas receptoras e estes locais são geralmente localizados nas mesmas regiões destas moléculas e a sua união ocorre sequencialmente. A ordem pela qual os ligandos se ligam aos seus receptores é normalmente semelhante à ordem seguida pela hGH, apesar da IL-4 se ligar aos receptores de forma inversa.

A simples dimerização é tudo aquilo que é necessário para a activação de diversos receptores.





As reacções de oligomerização levam como dissemos à activação intracelular de tirosina cinases associadas desencadeando a fosfatação de proteínas de uma ou mais proteínas STAT. A cinase exacta e a proteína ou proteínas STAT utilizadas dependem do domínio intracelular do receptor  $\alpha$  ou receptor  $\beta$  para os receptores homo ou heterooligoméricos respectivamente.

Conforme se pode ver na figura seguinte as estruturas dos domínios intracelular do receptor, tem duas regiões conservadas as chamadas caixa 1 e caixa 2.

Ligandos hematopoiéticos, receptores e Transdução de sinais



## Legenda da figura anterior

- Cada receptor contém pelo menos uma região homóloga citocina receptora (CRH) a sombreado que pode unir-se a um ligando.
  - Nas regiões CRH bandas a verde  representam cisteínas conservadas e bandas a vermelho  motivos WSXWS.
  - Alguns destes receptores têm domínios adicionais, na sua extremidade amino terminal tais como os parecidos com os domínios IgG ou três repetições parecidas com as fibronectinas, próximo da membrana.
  - Os domínios intracelulares dos receptores variam bastante na sequência e no comprimento mas os receptores assinalantes contêm duas regiões conservadas a caixa 1 a azul  e a caixa 2 a laranja .
- Os receptores são distribuídos em dois grupos, o homodímero e o hetero-oligomérico consoante o número de subunidades não idênticas necessárias para a ligação e sinalização.
- O grupo hetero-oligomérico é dividido em três subgrupos baseado nas subunidades  $\beta$  utilizadas; gp130, gh140 ou IL-2 $\gamma$ .
  - A sinalização destes complexos é conduzida pela ligação não covalente a várias tirosinas cinases (JAK 1, 2, 3 ou TyK 2) e factores de transcrição (STAT 1 a 6); + e  $\pm$  indicam o papel mais ou menos importante desencadeado na sinalização.

### Abreviaturas:

- GH-R – receptor da hormona do crescimento
- PRLR – receptor da prolactina
- EPO-R – receptor da eritropoietina
- G-cSFR – receptor do factor estimulador de colónias do granulócito
- MpI – receptor de leucemia mieloproliferativa de TPO receptor
- IL-6R – receptor interleucina 6
- LIF-R – receptor do factor inibidor da leucemia
- OSM-R – receptor da oncoestatina M

Nos aspectos estruturais gerais dos ligandos há que considerar que as estruturas básicas para os ligandos de cadeia longa (160 a 200 ácidos aminados) como a hGH e de cadeias curtas (104 a 145 ácidos aminados) como o GM-CSF (factor de estimulação dos granulócitos macrófagos) têm um ângulo constituído por um feixe de quatro hélices antiparalelas comum mas diferem no comprimento das suas hélices, e no tamanho das ansas (loops) de conexão e nas posições das suas pontes dissulfureto.

A marca de membros da família de receptores hematopoiéticos é a presença de uma região conservada citocina receptor (CRH) na porção extracelular, habitualmente comprometida na ligação da hormona (vide figura anexa). Cada CRH consiste de 200 a 250 resíduos e supõe-se ser constituída por dois domínios parecidos com os das imunoglobulinas e conectados por uma curta ligação. O domínio N-terminal contém quatro resíduos de cisteínas estritamente conservadas enquanto próximo da extremidade C do segundo domínio se encontra a sequência Trp-Ser-X-Trp-Ser altamente conservada, a chamada caixa WSXWS (X representando qualquer resíduo).

Nas estruturas das porções extracelulares dos receptores da hGH e hPRL os dois domínios parecidos com as imunoglobulinas são melhor classificados como módulo do tipo fibronectina III (FNIII).

A região de ligação à hormona está localizada entre os dois domínios do tipo fibronectina III e a região implicada nas interações receptor-receptor está localizada num lado do domínio C-terminal e o segmento WSXWS está localizado no lado oposto

#### Complexos receptores hematopoiéticos assinalantes.

A estrutura destes complexos formados pelos domínios dos receptores ligados com a hormona solúvel, é uma estrutura oligomérica, podendo estes oligómeros ser distribuídos em duas categorias ou grupos, o grupo homodimérico e o grupo heterodimérico (vide fig. Anexa).

No grupo homodimérico do qual a hGH é o paradigma, a subunidade receptora  $\alpha$  que liga a hormona é responsável pelo sinal.

O grupo hetero-oligomérico, do qual a interleucina IL-6 é um bom exemplo tem pelo menos dois receptores ( $\alpha$  e  $\beta$ ), sendo o  $\alpha$  o responsável pela ligação e o  $\beta$  o assinalante.

No grupo homodimérico está provado que uma molécula de hGH pode ligar-se a duas moléculas do domínio extracelular do receptor hGH (também chamado hGHbp), existindo naquela hormona dois locais distintos para se ligar a dois locais equivalentes do hGHbp ocorrendo portanto a dimerização com a formação de um homodímero. O mecanismo desta dimerização também é razoavelmente conhecido.

O receptor da prolactina m (PRL) é também um membro do grupo de receptores homodiméricos bem como outras hormonas hematopoiéticas tais como a EPO (eritropoietina), G-CSF (factor estimulador do granulócito) e TPO (trombopoietina).

No grupo hetero-oligomérico de complexos receptores, uma série de citocinas tem mais do que uma subunidade receptora, sendo a cadeia  $\alpha$  geralmente responsável pela ligação inicial da hormona e uma ou mais cadeias  $\beta$  responsáveis pela sinalização. Em alguns casos apenas o domínio extracelular do receptor  $\alpha$  com o receptor  $\beta$  na totalidade em toda a sua extensão é competente para assinalar a presença de hormonas.

O melhor caracterizado receptor hetero-oligomérico é a citocina IL-6 (interleucina) que parece ter dois locais, o local 1 para o receptor  $\alpha$  e o local 2 para o receptor  $\beta$ .

GM-CSF, IL-3 e IL-5 são citocinas afins, cada uma delas tendo um receptor  $\alpha$  específico e ligando uma subunidade  $\beta$  receptora comum, chamada Sp140.

A oligomerização da IL-4 parece decorrer de forma inversa à de outras citocinas. O  $\alpha$  receptor é ainda importante para a ligação inicial e o  $\beta$  receptor para a sinalização; o  $\beta$  receptor parece ser o  $\gamma$  receptor do complexo IL-2. O IL-4 tem dois locais para interactuar com os seus receptores e são topologicamente similares aqueles utilizados pela hGH.

A IL-2 tem três subunidades receptoras, a  $\alpha$ ,  $\beta$  e  $\gamma$ , a hormona pode formar complexos de afinidade aparentemente crescente:  $\beta < \alpha < \beta/\gamma < \alpha/\beta < \alpha/\beta/\gamma$ .

Dois outros membros da família da classe I que utilizam gp130 (OSM e CNTF) também têm três receptores e nestes casos o receptor pode funcionar da mesma forma que o receptor  $\alpha$  funciona para a IL-2.

A classe de I de hematopoiéticas tem muitas semelhanças nas suas estruturas e mecanismos de oligomerização dos receptores.

Apesar da pouca identidade de sequências entre a hGH e a prolactina (cerca de 23%) bem como dos seus receptores (cerca de 28%) a hGH liga e activa o receptor da prolactina.

### **Moléculas adaptadoras e sinalizantes**

As cinases Janus (JAK) revelam ausência de domínios SH2 e SH3 e a presença de dois domínios em tandem tirosina-cinases.

Em geral estas cinases são cataliticamente inactivas nas células em repouso mas estão associadas com domínios citoplásmicos das cadeias dos receptores sendo rapidamente activadas quando um ligando estimula a fosfatação de tirosinas dentro do domínio cinase.

Todas estas cinases, excepto a JAK3 são expressas em diversos tecidos. A JAK3 aparece sobretudo nas linhagens mielocíticas e linfocíticas. As JAK cinases medeiam a activação das proteínas STAT.

O receptor da GH assim como outros receptores da família citocinas não possuem actividades tirosina cinase intrínseca e em vez disso os componentes receptores multimerizados rapidamente activam tirosina cinases associadas (por exemplo JAK cinases).

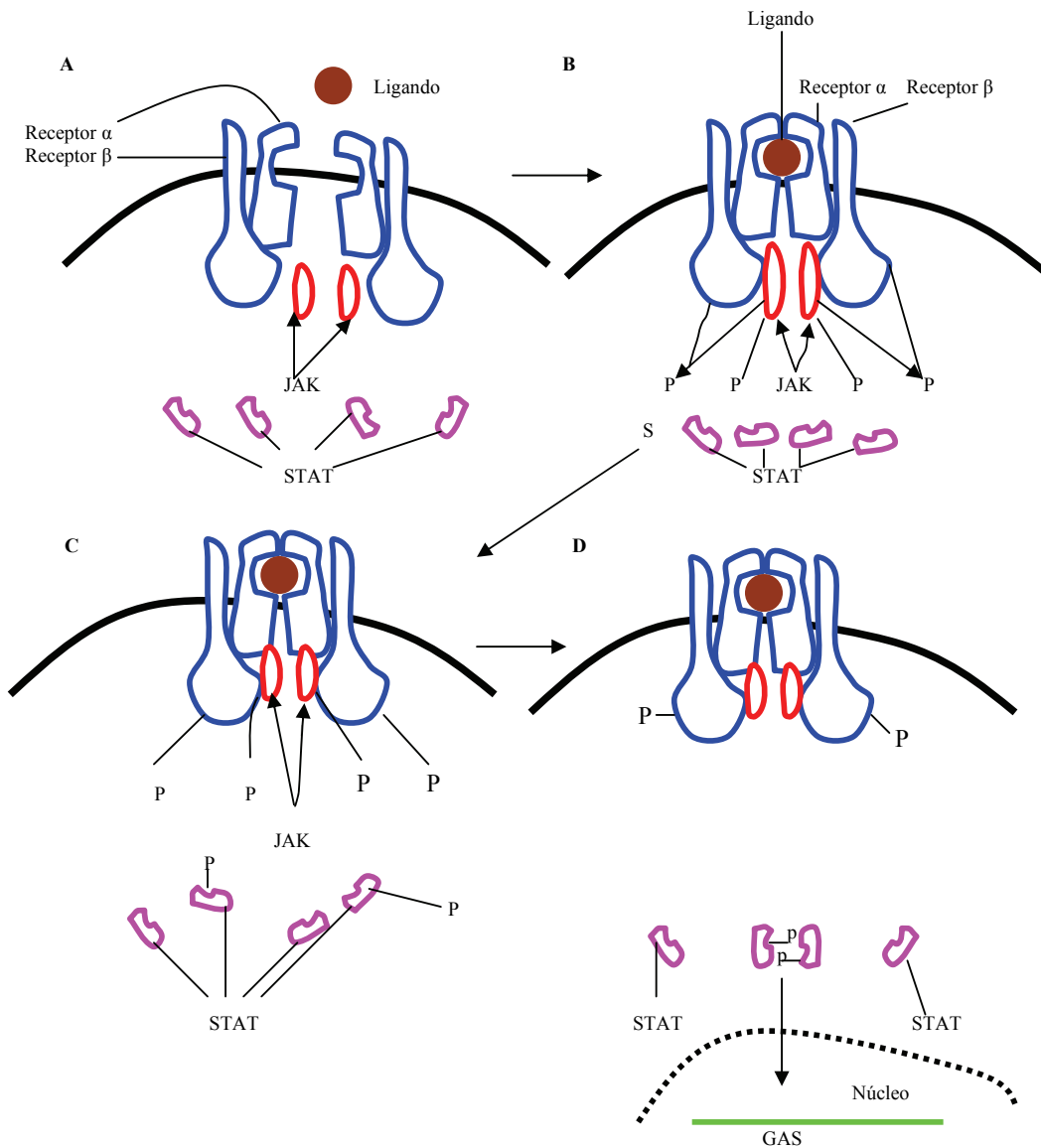
Esta capacidade é essencial para transmitir a maioria de sinais intracelulares para fosfatação, indução de genes e activação de moléculas de sinalização como a cinase fosfatidilinositol 3 (Pi3), a MAP cinase e a ras.

Na figura seguinte em que se mostra uma cascata de transdução de sinal tudo começa com um ligando que ao unir-se com um receptor específico, promove a dimerização do componente do receptor que traduz o sinal (cadeias  $\beta$ ). As cinases JAK associam-se com domínios específicos da região proximal da membrana das cadeias  $\beta$ . As cinases depois fosfatam uma tirosina na região distal das cadeias  $\beta$  do receptor que activado é capaz de interactuar com um membro específico da família STAT através dos domínios SH2 desta família.

As proteínas STAT são subsequentemente activadas por fosfatação da tirosina tornando-se então possível a sua dimerização (das proteínas STAT por heterodimerização e / ou homodimerização) sendo depois translocadas para o núcleo ligando-se ao DNA, sendo assim transducidos sinais para os seus alvos no núcleo.

As várias STAT proteínas com cerca de 700 ácidos aminados tem sequências semelhantes, e tem duas funções na transdução de sinais no citoplasma e na activação da transcrição no núcleo. Todas elas têm um simples resíduo de tirosina na região do resíduo 700 que é fosfatado durante a activação no citoplasma, e após activação todas elas se podem ligar ao DNA numa determinada sequência.

## Transdução de sinal das citocinas



**A** – a sinalização começa quando o ligando induz a dimerização do receptor

**B** – a dimerização do receptor leva á aposição de JAK cinases associadas que se transfosfatam uma á outra. As cinases, assim activadas fosfatam uma tirosina distal no receptor.

**C** – este resíduo receptor fosfatado é depois reconhecido por um domínio SH2 de uma proteína STAT que se adapta ao complexo receptor onde é activada por fosfatação.

**D** – as proteínas STAT activadas e tornadas competentes pela hetero ou homodimerização, são translocadas para o núcleo e ligam-se ao GAS (elemento do DNA contendo TTNNNNNAA).



Activação de JAK cinases por diversos ligandos

<u>Ligando</u>	<u>JAKs</u>			
	<b>TYK2</b>	<b>JAK1</b>	<b>JAK2</b>	<b>JAK3</b>
<b>Família IFN</b>				
IFN $\alpha/\beta$	+	+	-	-
IFN $\gamma$		+	+	-
IL-10	+	?	?	-
<b>Família gp130</b>				
IL-6	+	+	+	?
IL-11	?	+	?	?
onM	?	+	+	?
Lif	?	+	+	?
CNTF	-/+	+	+	?
G-CSF	?	+	?	?
IL-12	+	-	+	-
<b>Família <math>\gamma</math>-c</b>				
IL-2	-	+	-	+
IL-4	-	+	-	+
IL-7	-	+	-	+
IL-9	-	+	-	+
IL-13	-	+	?	?
IL-15	?	+	?	+
<b>Família gp140</b>				
IL-3	-	-	+	-
IL-5	-	-	+	-
GM-CST	-	-	+	-
<b>Família da hormona do crescimento</b>				
Epo	?	-	+	-
GH	?	-	+	-
PRL	?	+/-	+	-
<b>Receptor tirosina cinases</b>				
EGF	?	+	+	?
PDGF	?	+	+	?
CSF	?	+	+	?

Estão assinalados nos mamíferos seis genes que codificam proteínas STAT. Quatro destes genes que codificam as proteínas STAT 1 a 4 estão já localizados nos mapas cromossómicos. As STATs 1 e 4 estão próximas e juntas na região proximal do cromossoma 1 do rato enquanto o STAT2 está no cromossoma 10 do rato e o STAT3 no 11.

As STAT 1 e 3 encontram-se em muitos se não em todos os tipos de células, embora a sua concentração seja muito variável.

A STAT2 também aparece em muitos tipos de tecidos. As STATs 4 e 5 têm uma distribuição mais limitada, a STAT4 encontra-se em alta concentração nos testículos, e moderada no timo.

A STAT5 foi inicialmente chamada factor de crescimento mamário.

A STAT1 $\alpha$  ou  $\beta$  e a STAT2 contactam o cDNA. A p48 é a única proteína não STAT que funcionalmente interactiva com STATs.

Locais de ligação ao DNA nos promotores de uma série de vários genes que se ligam a proteínas activadas por ligandos, são similares ao consenso GAS (TTNNNNNAA).

Diferentes locais de ligação para diferentes genes não ligam igualmente a proteínas activadas por diferentes ligandos.

Muitos dos ligandos que operam através da via JAK-STAT circulam no sangue e as células podem possuir receptores simultaneamente para muitos ligantes diferentes.

O complexo receptor cinase tem efeitos qualitativos e quantitativos sobre a transcrição. O número de cinases e o número de STATs é menor que o número de receptores. É o próprio receptor que transmite a especificidade da interacção ligando receptor dentro das células.

O potencial de activação de um dado ligando depende do número de receptores e do tempo de duração do sinal, inclusivé da velocidade de internamento se isso sucede. Assim dois ligandos que activem o mesmo STAT, podem obter resultados de nível diferente durante o mesmo período de actuação. A afinidade de cada STAT também pode variar consoante o complexo receptor-cinase.

Há ainda que atender ás interacções entre as STATs, a STAT1 e 3 podem homodimerizar, enquanto a 1 e 2 e a 1 e 3 podem heterodimerizar.

Activação de proteínas STAT por diversos ligandos

<u>Ligandos</u>	<u>STATs</u>						STF Factor de transdução de sinal
	1	2	3	4	5	6	
<b>Família IFN</b>							
IFN $\alpha$	+	+	+	-			
IFN $\gamma$	+	-	-	-			
IL-10	+/-	-	+	-			
<b>Família gp130</b>							
IL-6	+/-	-	+	-			
IL-11	+/-	-	+	-			
onM	+/-	-	+	-			
LiF	+/-	-	+	-			
CNTF	+/-	-	+	-			
<u>Ligandos</u>	1	2	3	4	5	6	STF Factores de transdução de sinal
<b>Família gp130 (continuação)</b>							
G-CSF	+/-	-	+	-			
IL-12			+	+			
<b>Família <math>\gamma</math>-c</b>							
IL-2	+/-		+		+		+
IL-4						+	+
IL-7					+		+
IL-9					+		+
IL-13						?	+
IL-15					+		+
<b>Família gp140</b>							
IL-3					+		+
IL-5					+		+
GM-CSF					+		+
<b>Família GH</b>							
EPO	-	-	-	-	?		+
GH	+/-	-	+	-	+		
PRL	+/-	-	-	-	+		
<b>Receptores tirosina cinases</b>							
EGF	+	-	+	-			
PDGF	+	-	+	-			
CSF-1	+		+				

Referem-se neste quadro à esquerda os ligandos e depois à direita as proteínas STAT activadas por fosfatação (tirosina) e que se ligam ao DNA.

O factor STF indica uma actividade induzível de ligação ao DNA por um elemento do DNA contendo a GAS consenso (TTNNNNNAA) e na maioria dos casos também contem tirosina fosfato.

O ponto final da cascata de transdução de sinais reside na indução de genes específicos embora pouco se conheça acerca disto.

Diferenças na natureza dos elementos GAS nos promotores dos genes podem explicar algumas diferenças.

Após ligação aos seus receptores, diversos genes promovem uma alteração no fenótipo das células.

### **Cascatas de transdução de sinal e IFN**

#### **Como vimos os STAT são factores de transcrição activados por proteínas tirosina cinase associadas com receptores da superfície celular**

Em culturas de células mamíferas tratadas com diversos tipos de interferons (IFNs) foram identificados e caracterizados novos tipos de proteínas que estabeleciam a conexão entre os receptores activados por estes polipeptídeos e por outros e a expressão de diversos genes. Estas novas classes de proteínas eram proteínas cinases e factores de transcrição denominados STAT por serem sinais de transdução e activação de transcrição.

Foi ainda caracterizado que após adição do IFN $\alpha$  a curto prazo (15 a 30 minutos) alguns genes aumentavam a sua expressão mais de vinte vezes, e em muitos destes genes os seus elementos promotores continham uma sequência nucleotídica semelhante chamada ISRE ou seja elemento de resposta desencadeada pelo interferon.

Verificou-se depois que este ISRE interactuava com quatro proteínas que se ligam a ele e que são as STAT1 $\alpha$ , STAT1 $\beta$ , STAT2 e uma proteína p48.

Nas duas primeiras, embora idênticas a STAT1 $\alpha$  tem mais 38 resíduos de ácidos aminados extra na região C-terminal, resultando de um mesmo gene, mas com processamentos distintos (splicing).

A STAT3 tem 59% de homologia com as duas STATs anteriores, e curiosamente estas três STATs (1 $\alpha$ , 1 $\beta$  e 3) citosólicas, contêm na sua estrutura um domínio SH2.

A proteína p48 que existe no citoplasma e no núcleo interactua normalmente, sem estimulação hormonal nem a presença de STATs, com o DNA e é caracterizada ainda por não conter domínios SH2 e a sua estrutura não estar relacionada com a sequência primária das STATs.

As três proteínas STAT, antes referidas após activação desencadeada pela interacção do interferon  $\alpha$  com o receptor respectivo, translocam-se para o núcleo.

Os três tipos de interferons (IFN $\alpha$ ,  $\beta$  e  $\gamma$ ) interactuam com receptores que pertencem á superfamilia dos receptores-citocina que não são dotados de actividade intrínseca de proteína cinase.

No delineamento geral das cascatas de transdução de sinal desencadeadas pelos IFNs temos que após interacção destes com o receptor respectivo o domínio citosólico deste é activado o que vai permitir

a interação com outras proteínas independentes existentes no citosol e que são como tal activadas como proteínas cinases.

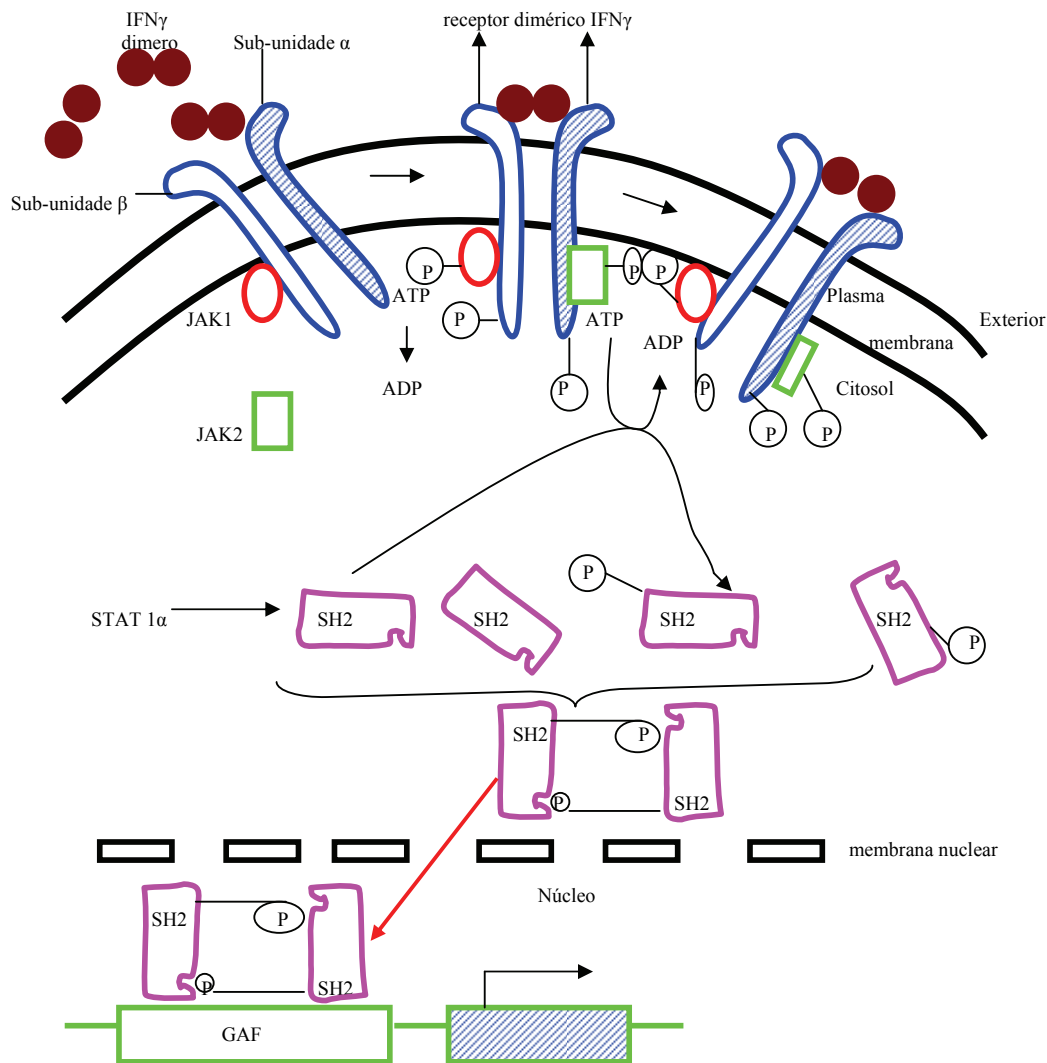
No caso do receptor activado pelo IFN $\alpha$ , duas proteínas cinases citosólicas a TYK2 e a JAK1 são assim activadas indo por seu turno fosfatar depois resíduos de tirosina que fazem parte da estrutura dos STATs e são estes que depois se dimerizam formando heterodímeros STAT1-STAT2 dímeros que se deslocam para o núcleo onde vão interagir com a proteína p48 originando a interação no DNA com as sequências ISRE e activando a expressão dos respectivos genes (ver figura seguinte).

Cada tipo de interferon é responsável pela especificidade da cascata sinal-interferon-STAT desencadeada e como tal é promovida a transcrição de um subconjunto único de genes, tudo dependendo das cinases associadas com o receptor das STATs intervenientes e dos elementos de resposta cis-activos sobre o DNA.

#### Via de transdução de sinal com ligação do IFN $\gamma$ a um receptor da superfamília dos receptores citocinas, via ( JAK – STAT)

O receptor activo é um heterodímero, a subunidade  $\beta$  é diferente consoante o tipo de célula.

Após estimulação do receptor e sua dimerização, os resíduos de Tirosina no domínio citosólico do receptor e das proteínas JAK 1 e JAK 2 são fosfatados, o que induz a que se forme um complexo receptor / JAK 1 / JAK 2. Os monómeros de STAT 1 admitem-se que se liguem através dos seus domínios SH 2 ao complexo activado, o que leva á fosfatação dos STAT 1 $\alpha$ , sendo depois libertados no citosol onde se dimerizam, translocando-se este homodímero de STAT 1 $\alpha$  para o núcleo activando a transcrição de genes controlados pelo elemento regulador GAF.



Vias de transdução de sinais induzidas por outros membros da superfamília receptores-citocina são semelhantes à via antes referida para o interferon-JAK-STAT, com dimerização dos receptores, e com os domínios citosólicos ativados destes a interagirem e ativarem cinases citosólicas da família JAK.

Outros ligandos ligando-se a outro tipo de receptores podem também activar STATs. Por exemplo o EGF, liga-se a uma RTK que estimula fosfatando um STAT citosólico o qual depois pode ligar-se uma sequência no promotor-c-fos chamado elemento de indução do soro (SRE).

Os interferões desempenham papéis muito importantes mediando respostas antivirais e anticrescimento celular, e modelando as respostas imunitárias.

Os humanos têm pelo menos doze IFN $\alpha$  funcionais, e um simples IFN $\beta$  antigenicamente distinto e um IFN $\gamma$  relacionado, enquanto outras espécies têm múltiplos sub-tipos IFN $\beta$ .

Os IFN inibem o crescimento celular e controlam a apoptose, actividades que afectam a supressão de cancro e da infecção.

As principais vias assinalantes dos IFN são rápidas e directas e envolvem fosfatações de tirosina e activação de transdutores de sinal e de activadores de factores de transcrição por tirosinas cinases Janus na membrana celular, seguidas pela libertação de transdutores de sinal e activadores da transcrição e sua migração para o núcleo onde induzem a expressão de genes alvo.

Os interferões (IFN) do tipo I (sobretudo  $\alpha$  e  $\beta$ ) e do tipo II ( $\gamma$ ) desencadeiam sinais através de vias distintas embora relacionadas. Estas vias implicam os receptores específicos do tipo I e do tipo II que se ligam a cinases Janus (JAK), sendo depois estes sinais induzidos através de transdutores de sinal e activadores de transcrição (STATs) que propagam esses sinais.

Como vimos as JAKs podem autofosfatar-se e transfosfatar fosfatando receptores e proteínas que possam interactuar com elas, como é o caso das STATs.

As JAK 1, JAK 2 encontram-se no núcleo, no citoplasma e nas membranas.

As STATs induzem factores de transcrição.

Também se sabe que os JAKs e STATs estão envolvidos em diferentes vias mediadas por citocinas, e factores de crescimento.

Conhecem-se hoje quatro JAKs mamíferas e sete STATs.

Após activação pelas JAK 1 através de fosfatação de um resíduo específico de tirosina as STATs formam homo ou heterodímeros através de interacções mútuas fosfotirosina-Src homologia região 2 (SH2).

Os dímeros STAT ligam-se a elementos sequência gama activados (GAS) os quais conduzem a expressão de genes alvo vizinhos.

Diferentes elementos GAS preferem diferentes dímeros STAT consoante as especificidades.

Ambos os heterodímeros STAT 1 e 2 e STAT 1 homodímeros ligam p48, um membro da família de factores reguladores do interferon. Os trimeros resultantes (ISGF3) ligam-se a elementos reguladores estimulados pelo IFN (ISREs) que são distintos dos elementos GAS; conduzindo os ISRE a expressão da maioria dos genes regulados pelo IFN  $\alpha/\beta$  e alguns genes regulados pelo IFN $\gamma$ .(vide figura adiante)

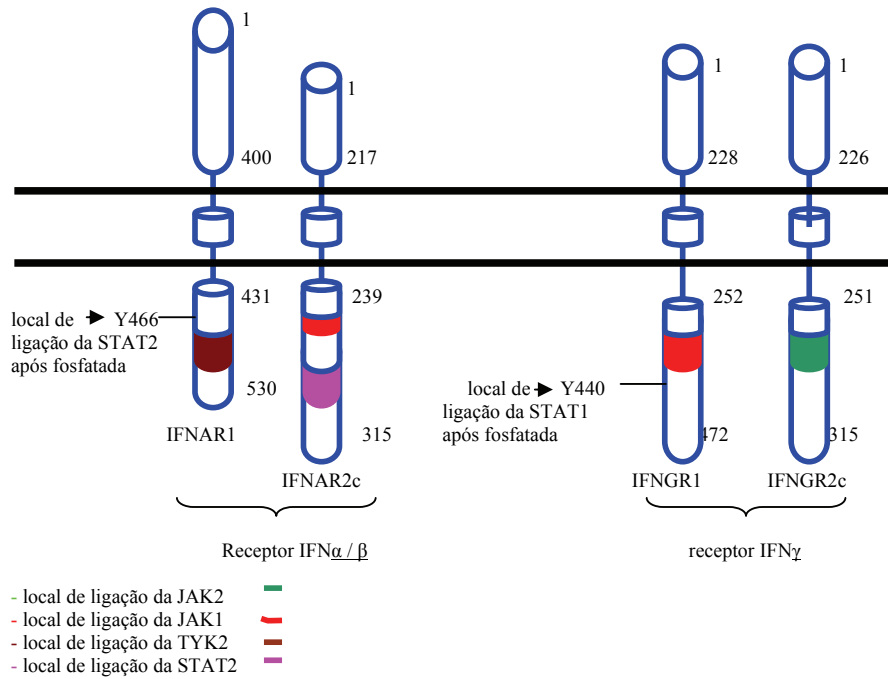
### Interferão $\gamma$ (IFN $\gamma$ )

Os primeiros acontecimentos da actuação do IFN $\gamma$  requerem a participação de cinco proteínas distintas: as proteínas integrais de membrana do tipo I, IFNGR1 e IFNGR 2 (as subunidades do receptor IFN $\gamma$ ), a JAK 1, a JAK 2 e a STAT 1.

Os receptores IFN  $\gamma$  são expressos em aproximadamente todos os tipos de células, possivelmente com excepção dos eritrocitos.

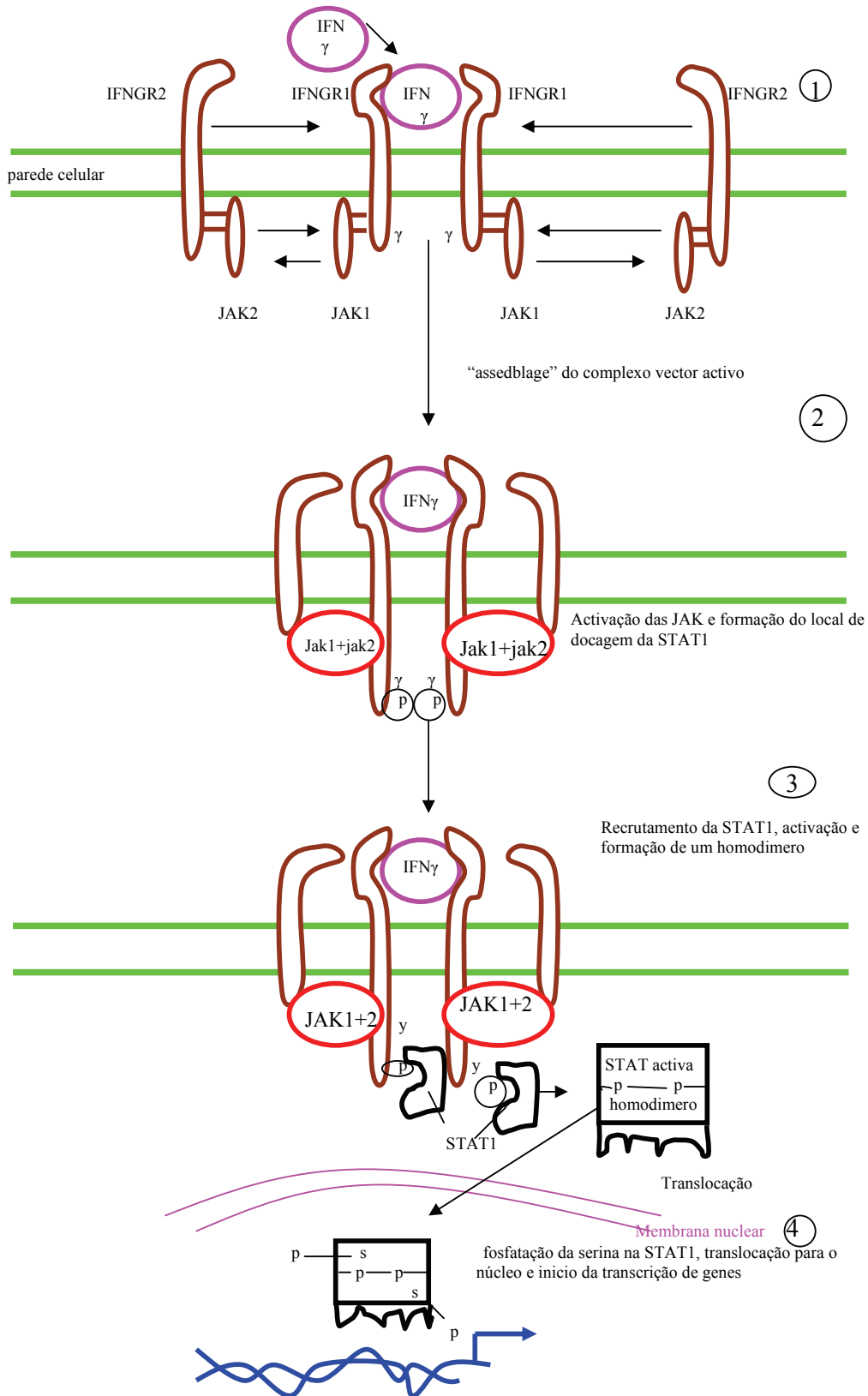
Os receptores IFN  $\gamma$  funcionalmente activos são constituídos por duas espécies de cadeias polipeptídicas, a IFN GR1, dantes chamada cadeia  $\alpha$ , que medeia a ligação ao ligando, e a IFN GR2, previamente chamada cadeia  $\beta$  que tem um papel menor na união ao ligando mas é necessária para a sinalização.

Esquema dos receptores dos interferões (IFN) humanos  $\alpha / \beta$  (à esquerda) e  $\gamma$  (à direita)



No esquema seguinte mostra-se o início de transduções de sinal pelo IFN $\gamma$  a partir do respectivo receptor.





Nas células não estimuladas, ① na figura anterior, a sub-unidade IFNGR1 associa-se com a cinase Janus (JAK1) e a sub-unidade IFNGR2 associa-se com a JAK2. A estimulação pelo IFN $\gamma$  ② induz a oligomerização das duas sub-unidades do receptor o que leva à transfosfatação e activação da JAK1 e da JAK2 (ver ② na figura anterior). As JAKs activadas então fosfatam o  $\gamma$  (tirosina) 440 da subunidade IFNGR1, criando um local para “docagem” do STAT1 (sinal transdutor e activador de transcrição) – vide ③ na figura.

Ainda ligada ao receptor a STAT1 é fosfatada no  $\gamma$  (tirosina) 701 e depois libertada do receptor, formando um homodímero que se transloca para o núcleo ④ e promove o início da transcrição de genes.

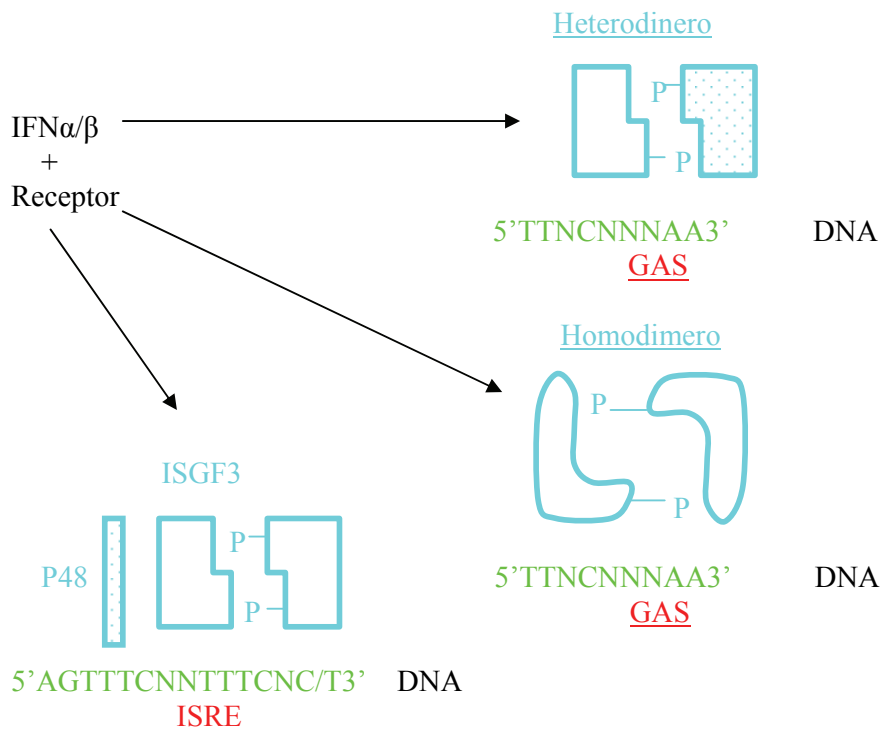
O IFN $\gamma$  utiliza pois a via JAK-STAT para a transdução de sinal.

A principal via de resposta ao IFN  $\alpha / \beta$  requer duas subunidades receptoras (IFNAR1e2c), duas JAKs, duas STATs e o factor de transcrição p48 da família IRF. A via geral de sinalização pelos IFN  $\alpha/\beta$  envolve cinco etapas principais; primeiro a dimerização do lado extracelular dos receptores induzida pelo IFN, depois a iniciação da cascata de fosfatação tirosinica dentro da célula, levando á dimerização das STAT fosforiladas e sua activação, com o conseqüente transporte para dentro do núcleo onde se liga a sequências de DNA específicas estimulando a transcrição.

O IFN $\alpha/\beta$  pode activar diversas vias de transdução de sinal desencadeando respostas distintas.

Os diversos factores de transcrição desencadeados pela interacção dos IFN $\alpha/\beta$  com os respectivos receptores tem diversos elementos de reconhecimento ao nível do DNA tal como se representa na figura seguinte.

Factores de transcrição formados pela interacção interferões  $\alpha/\beta$  – receptores e seus elementos de reconhecimento no DNA

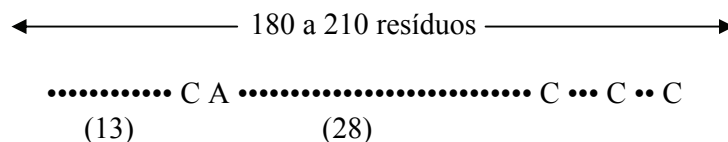


### Cascatas de transdução de sinal e GH

A hormona somatotropina (hormona do crescimento, GH), que controla o crescimento, a coriomamotropina (lactogénio), um análogo placentário da GH e a prolactina que actua sobretudo sobre as glândulas mamárias promovendo a lactação constituem uma família de proteínas homólogas, estando ainda incluídas nesta família outras hormonas a seguir referidas:

- Lactogénicos placentários I e II dos roedores
- Lactogénio dos bovinos e ovinos
- Proliferina I, II e III do rato e proteína proliferina relacionada
- Proteínas I, II e III relacionadas com a prolactina placentária bovina
- Proteínas A e B prolactina like placentária do rato
- Variantes 1 (GH-V1) e 2 (GH-V2) da hormona do crescimento humano
- Somatolactina (SL) de diversos peixes
- 

A estrutura esquemática das proteínas desta família é a seguinte:



- C - corresponde a resíduos de cisteína conservados envolvidos na formação de pontes dissulfureto
- ... - resíduos de ácidos aminados (número entre parentesis)

A hormona do crescimento dos frangos (c GH) uma hormona polipeptídica sintetizada e segregada pela glândula pituitária, está implicada no crescimento, composição do corpo dos animais, produção de ovos, desenvolvimento e reprodução.

Os polimorfismos destas GH dos frangos tem sido referidos como estando associados com certos fenótipos.

Nos humanos a regulação da síntese e libertação da GH é modelada por uma família de genes que incluem factores de transcrição tais como o PROP1 e PIT1 que regulam a diferenciação das células da pituitária em somatotrofos que sintetizam e libertam a GH.

Genes importantes na libertação da GH envolvem genes GHRH e a GHRHR, o primeiro sintetizado e libertado no hipotálamo, vai depois para a pituitaria anterior onde se liga a GHRHR desencadeando a transdução do sinal e promovendo a libertação da GH pré-sintetizada que é armazenada em granulosa segregatários.

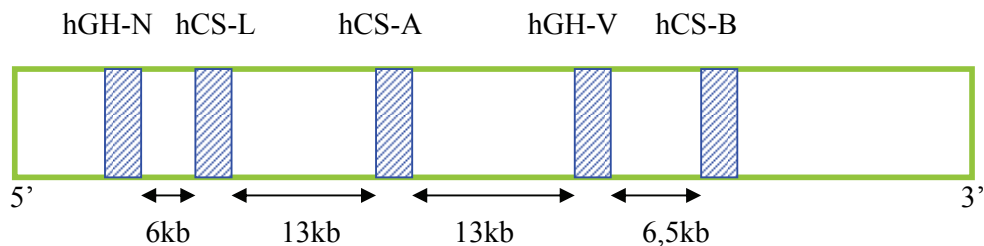
Outros produtos de genes importantes na síntese e libertação da GH são os GHR e as proteínas de ligação à GH ou sejam as GHBP que derivam da GHR e que permanecem ligadas à GH em circulação. Após ligação da GH a duas moléculas de GHR é transduzido um sinal para produzir IGF1.

A somatomamotrofina coriónica (hCS) ou lactogénio placentário sintetizado na placenta e contribuindo na preparação da mama para a lactação durante a gravidez é uma proteína também com 191 ácidos aminados como a GH e com uma homologia entre as duas de cerca de 85%.

As duas hormonas hCS e GH são formadas em tecidos diferentes e têm funções distintas.

Nos seres humanos está identificada uma família de genes da hormona do crescimento (hGH), com cinco genes estruturais situados no cromossoma 17 no seu braço longo e ocupando uma extensão nucleotídica de 55 kb. Cada um destes genes possui cinco exões e cinco intrões e ocorrem entre algumas sequências Alu.

#### Organização da família de genes da hormona do crescimento humano (hGH)



A ordem dos genes é a indicada na figura anterior sendo o primeiro gene expresso na forma de um precursor maior do que os 191 ácidos aminados que possui a forma madura, pelos somatotrofos da hipófise anterior e os restantes na placenta.

É também conhecido que existem diferentes sequências nos genes hGH e hCS cerca de 100 bp à frente do local de poliadenilação.

A expressão do gene hGH-N promove a formação de uma GH normal com 22 kDa, podendo no entanto um corte alternativo do intrão 3 promover um GH de menor peso molecular e função desconhecida.

Os genes hCS-A e hCS-B embora expressem a mesma hormona madura, fazem-no ao nível da placenta em diversos níveis.

O pseudogene  $\psi$ hCS-L não permite a maturação normal do seu transcrito primário em mRNA dada uma diferente sequência num local de corte exão-intrão.

A regulação da expressão dos genes GH e CS pode ser feita de diversos modos. Estão identificados dois locais promotores acima do hGH-N, havendo ainda um factor específico de transcrição, o GHF-1 ou Pit.

Este último factor GHF-1 tem um homeodomínio na sua estrutura.

A tiroxina e o cortisol estimulam a transcrição dos genes GH e CS depois de transportadas para o núcleo através dos respectivos receptores, ou então fazem-no associadas com o GHF-1.

São bem conhecidas diversas deleções nesta família de genes da GH, que em casos de severa deficiência em GH podem ser responsáveis pelas baixas estaturas e pelas baixas concentrações de GH detectáveis no sangue.

Recorda-se que a GH ao nível do fígado induz a síntese de outras hormonas ou somatomedinas que são factores “insulina-like”(IGFs) que promovem a proliferação de tecidos mesodermicos como os ossos ,cartilagens e músculos.

### Controlo do crescimento pelo eixo somatotrófico

A teoria de que a maioria dos efeitos biológicos da GH eram mediados pela circulante IGF-1 (factor endócrino insulina-like growth) tem sido modificada ultimamente desde que se descobriu que a maioria dos tecidos exprimem IGF-1 que pode actuar de uma forma autócrina/parácrina, tendo por outro lado a GH e a IGF-I efeitos independentes em diversos órgãos alvo.

A IGF-1 é sobretudo derivada do fígado mas não é essencial para o crescimento normal após o nascimento.

A regulação da segregação da hormona do crescimento (GH) é a determinante mais importante do tamanho do corpo.

A GH é regulada pela nutrição, e pelo meio hormonal e genético que controlam o “timing” e a velocidade do crescimento.

Os principais efeitos da GH são após o nascimento e não têm praticamente efeito sobre a regulação do crescimento fetal.

Muitos dos efeitos da GH sobre o crescimento e metabolismo são mediados indirectamente através do controlo e síntese de outros factores do crescimento dos quais o IGF-1 é o mais importante.

A GH e o seu receptor (GHR) pertencem à grande família dos péptidos citocina e seus receptores. Parece haver na segregação da GH, nos machos, discretos pulsos com baixos níveis interpulsos, enquanto nas fêmeas, estes pulsos são mais baixos e os intervalos entre estes, mais elevados.

Em todas as espécies estudadas os níveis de IGF-1 no soro e velocidades de crescimento correlacionam-se com a amplitude dos picos das GH e não com a concentração desta entre os picos.

A GH é produzida na hipófise, cérebro, linfocitos, placenta, mama e glândula pineal.

Factores hipotalâmicos e outros periféricos controlam a libertação da GH e a sua acção.

A GH é uma hormona anabólica que induz um balanço azotado positivo, favorecendo a tomada de ácidos aminados pelos músculos do esqueleto parecendo que este tecido é o alvo principal dos efeitos fisiológicos da GH.

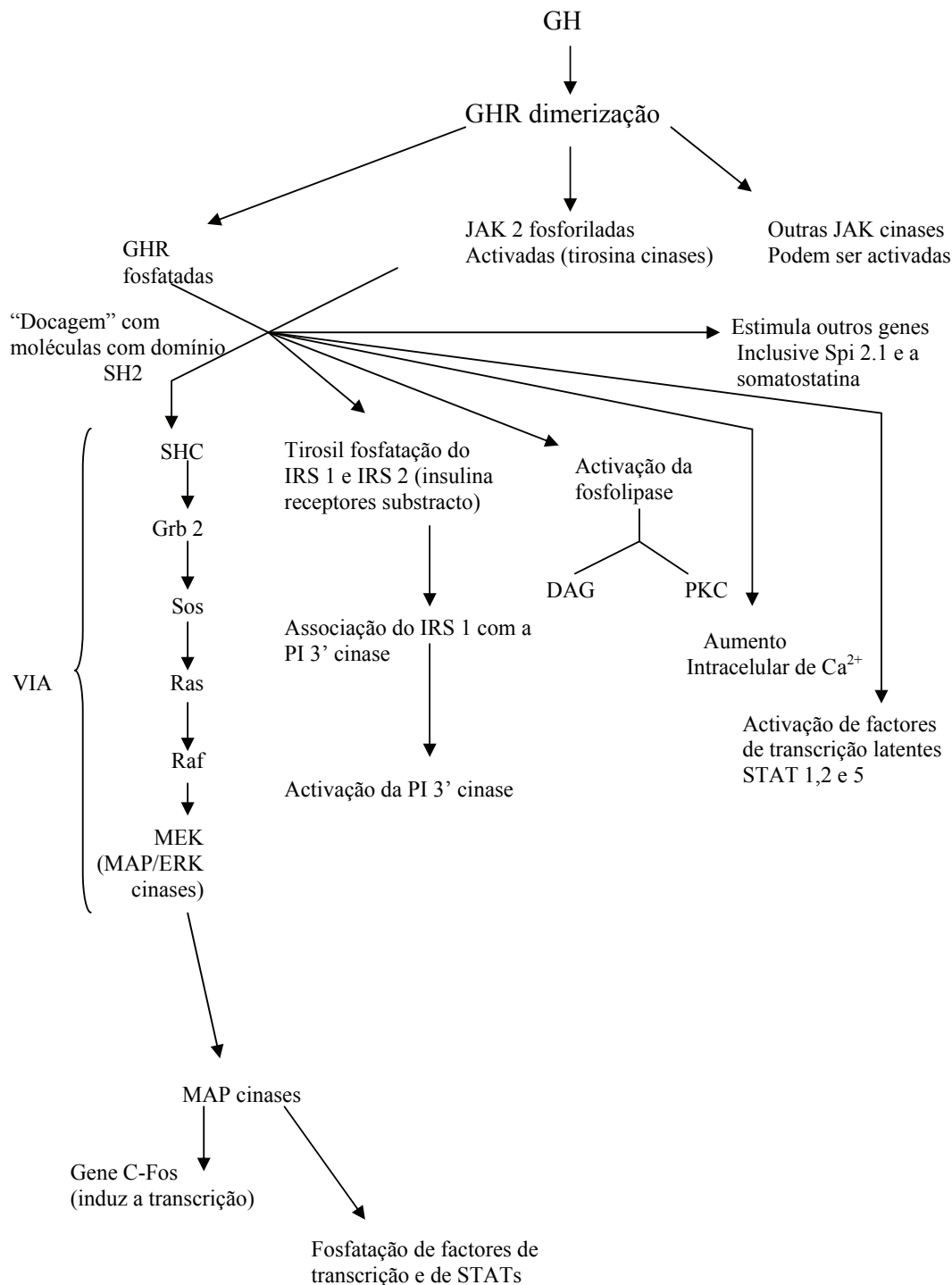
A longo termo a GH diminui a deposição de gordura e aumenta a mobilização desta.

### Mecanismo molecular da acção da hormona do crescimento

A forma como a GH se liga ao seu receptor (GHR) e desencadeia diversos efeitos começa pela ligação da hormona ao receptor desencadeando a dimerização deste e a activação da JAK2 tirosina cinase associada aos receptores e pela tirosil fosfatação da JAK2 e dos GHR.

Estes acontecimentos activam uma diversidade de moléculas sinais, inclusive MAP cinases (MAPK), substâncias receptoras da insulina (IRS), fosfatidilinositol 3' fosfato cinase (PI3'K), diacilglicerol (DAG), proteína cinase C (PKC), cálcio intracelular e factores de transcrição STAT.

## Vias sinalizadas pela GHR



Está confirmado em vários tipos de culturas celulares, que a GHR activada aumenta a biossíntese da insulina, a proliferação celular tal como estimula a fosfatação da tiosina numa série de proteínas celulares e a actividade cinase de proteínas associadas com microtúbulos (MAP), aumentando a concentração de Ca<sup>2+</sup> intracelular, e activando ainda, a STAT proteína, a expressão da C-fos, a síntese proteica e o transporte de glucose.

Todas estas moléculas sinais contribuem para as mudanças induzidas, pela GH, na actividade enzimática, transportes e expressão de genes que acabam por dar origem a mudanças no crescimento e no metabolismo.

Para o conhecimento da acção da GH foram determinantes os seguintes factos:

1. A clonagem do cDNA codificante da GHR numa série de animais domésticos rato, vaca, suíno, ovino, frango.
2. A classificação do GHR como pertencentes à superfamília dos receptores citocina / hematopoietina.
3. A verificação de que uma molécula de GH induz a dimerização do GHR.
4. A identificação das moléculas JAK2 e GHR associada a tirosina cinase que são activadas pela ligação da GH ao respectivo receptor.

Foram depois identificadas uma série de moléculas que interactuaram com a GHR, JAK2 e outras proteínas associadas.

É possível hoje caracterizar as estruturas da hormona do crescimento e do respectivo receptor, bem como as cascatas de sinais, desencadeadas, após a sua interacção ao nível do citoplasma e membrana celular bem como os seus reflexos ao nível do núcleo celular.

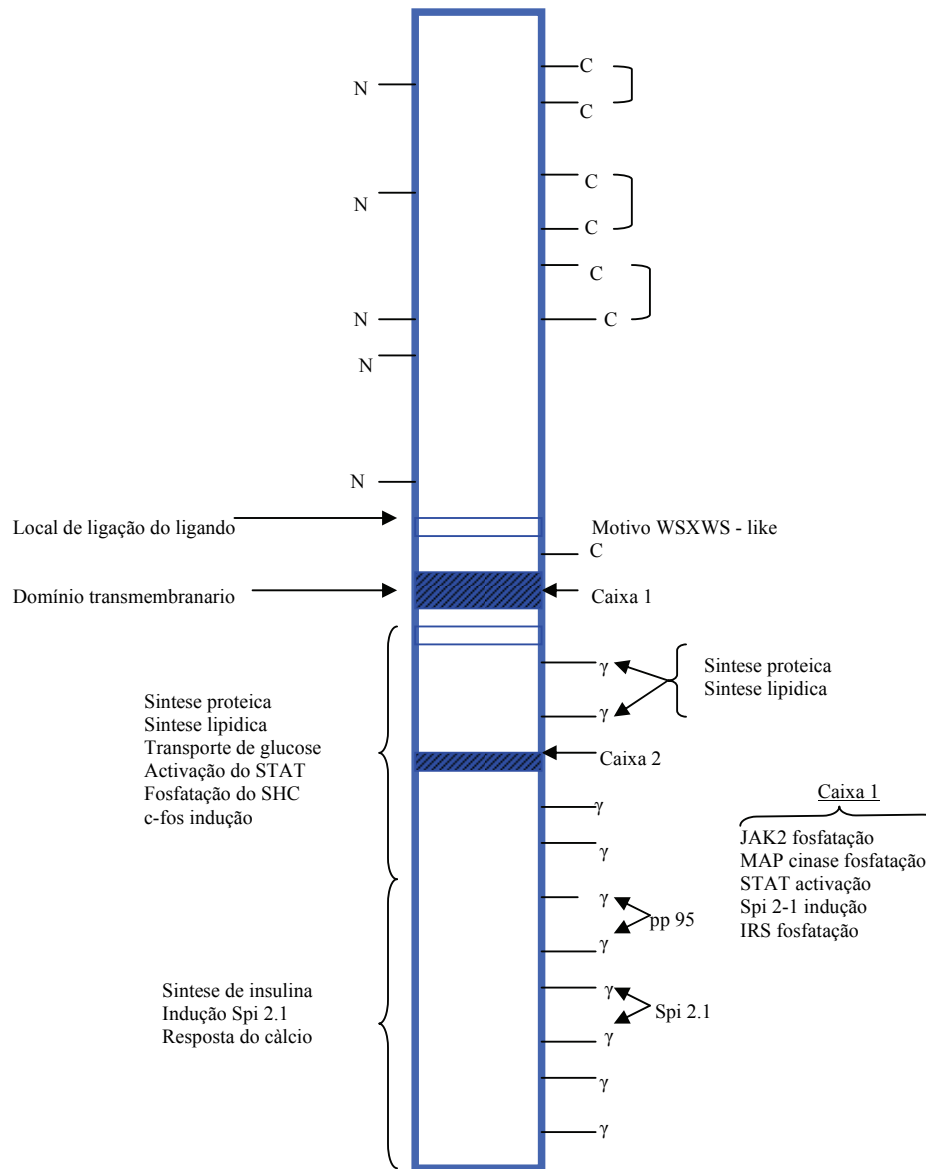
A superfamília dos receptores citocina / hematopoietina engloba a GHR, tal como as subunidades receptoras para a prolactina (PRL) e uma diversidade de citocinas e factores de crescimento, inclusive eritropoietina (EPO), interleucinas (ILs), trombopoietina, factor inibidor da leucemia (LIF).

Ainda nesta superfamília são considerados os receptores para os interferões  $\alpha$  /  $\beta$  e  $\gamma$  e IL-10 embora as suas sequências os afastem um tanto levando-se a integrá-los na classe II desta superfamília de receptores.

Apesar de já termos visto anteriormente algumas características estruturais desta família de receptores que interactuam com a GH, desenvolvem-se seguidamente alguns aspectos dos GHR.



## Estrutura do receptor da hormona do crescimento e funções desempenhadas GHR



O domínio extracelular dos GHR relativamente a outros receptores daquela superfamília, tem uma fraca homologia, numa extensão de cerca de 210 ácidos aminados apenas 15 a 35% são conservados, assinalando-se nele a presença de cisteínas e asparaginas e próximo da membrana existe o motivo WSXWS que parece ser crítico para a interação com o ligando.

Nas caixas 1 e 2 situadas no domínio citoplásmico do dGHR e que são comuns a outros membros desta superfamília de receptores citocina, a caixa 1 situada próximo da membrana nuclear é rica em prolinas, possuindo oito ácidos aminados ILPPVPVP nos mamíferos.

A caixa 2 começa com uma série de ácidos aminados hidrofobos e termina com ácidos aminados carregados positivamente, sendo separada a caixa 2 da caixa 1 por cerca de 30 ácidos aminados.

Como se compreende mutações ou deleções ao nível de qualquer uma destas estruturas desencadeia interações deficientes com os ligandos respectivos pelo menos em determinados tipos celulares.

### Cascatas de transdução de sinal desencadeadas pela GH

Uma molécula de GH complexa-se com dois receptores GHR originando a sua dimerização.

Após esta ocorrem uma série de respostas celulares rapidamente como é o caso da ligação da tirosina cinase JAK2 ao GHR e activação da JAK2 que é um membro da família Janus de tirosina cinases citoplásmicas que incluem a JAK1, JAK2, JAK3 e Tyk2, parecendo que a JAK1 também sofre alguma fosfatação nos seus resíduos tirosilo, tal como a JAK3, embora a mais importante neste contexto seja a JAK2. Em relação a Tyk2 existem dúvidas sobre se são ou não activadas nesta cascata.

As diversas JAK cinases desta superfamília de receptores citocina, são activadas também por diferentes tipos de ligandos.

Admite-se que a ligação da GH a duas moléculas de GHR aumenta a afinidade da JAK2 para cada molécula de GHR possibilitando a transfosfatação de uma ou mais tirosinas no domínio cinase da JAK2 emparelhada.

A activação da JAK2 aumenta a fosfatação tirosilo das proteínas SHC, IRS1 e 2 e ERKs 1 e 2, aumentando ainda a síntese de insulina, induzindo também o gene da Spi 2.1, e a mitogénese de algumas células.

Em resposta a GH, a JAK2 e o GHR são fosfatados nas tirosinas, e como na JAK2 do rato existem 48 tirosinas admite-se que múltiplas moléculas sinal possam interactuar com a JAK2.

Parece que o GHR é fosfatado nas tirosinas das metades C-terminal e N-terminal do domínio citoplásmico.

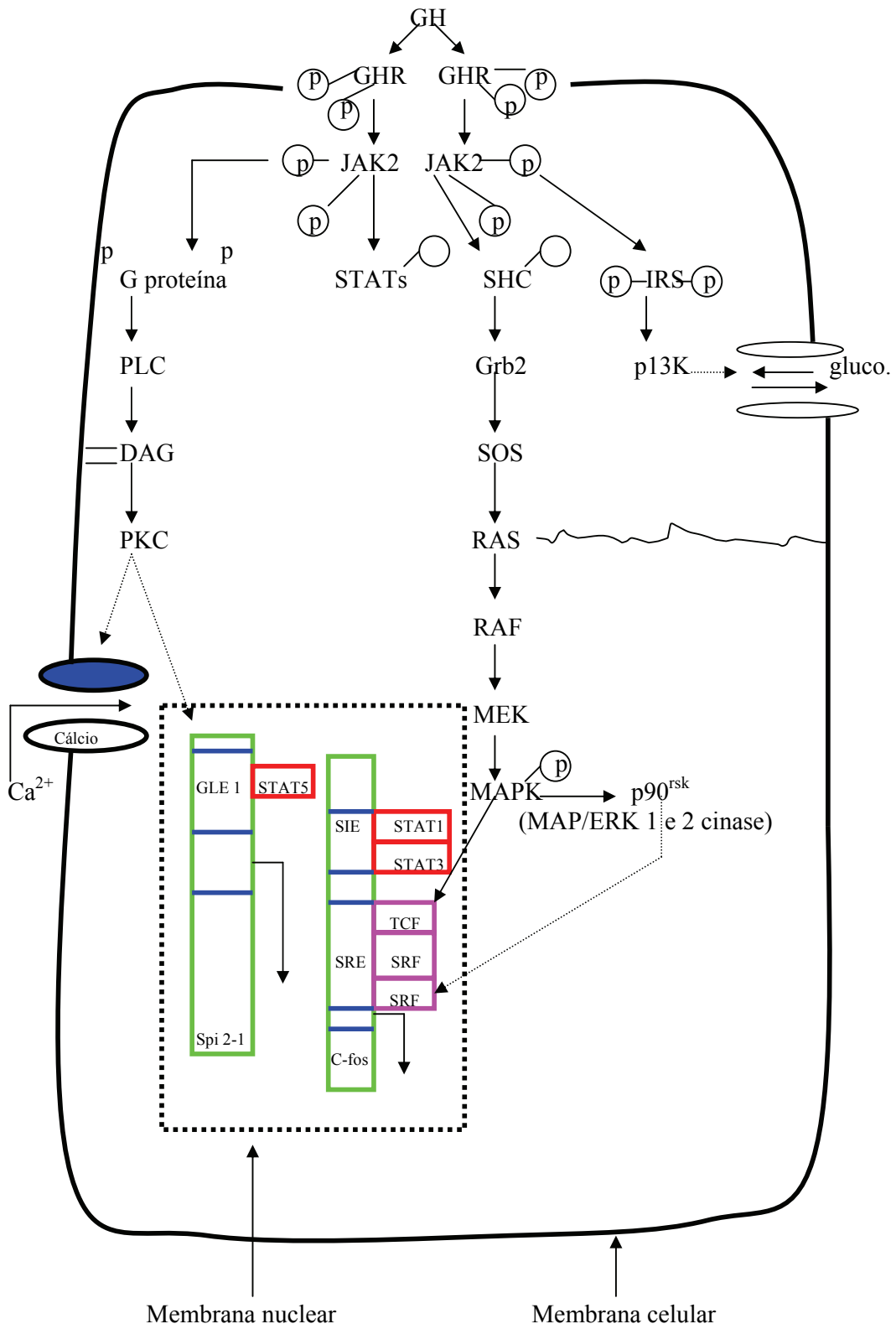
Na via de activação dos MAP cinases pela GH verifica-se além do GHR e da JAK2, um largo número de outras proteínas podem ser fosfatadas nos seus tirosilos em resposta á GH. Parece que a JAK2 fosfata além da JAK2 e da GHR outras proteínas que se ligam á JAK2 activada e á GHR iniciando uma via de transdução de sinal envolvendo outras tirosina cinases ou fosfatases. É o caso das MAP cinases designadas ERK1 e ERK2. que activadas desempenham um papel crítico na regulação do crescimento e da diferenciação celular.

Uma via que parte do receptor de membrana tirosina cinase para as MAP cinases, implica SHC, Grb2; Sos, ras, raf e MAP/ERK cinase (MEK), (vide esquema anexo na página seguinte). Parece que a GH estimula a associação da Grb2 com a SHC e destas proteínas com o complexo JAK2-GHR através de domínios SH2 das proteínas SHC as quais são então fosfatadas pela JAK2 permitindo a ligação da Grb2 á SHC.

Parece que a GH não activa as restantes moléculas desta cascata de transdução de sinais (Sos, ras, raf e MEK).

Nos substratos da MAP cinase estão outras proteínas cinases (por ex: C-Raf-1, p70<sup>rsk</sup> e p90<sup>rsk</sup>), fosfolipase A2, proteínas do citoesqueleto e factores de transcrição (por ex: C-Jun e complexos ternário TCFs) inclusivé p62<sup>tef</sup>/ELK-1).

Vias sinalizadas pelo GHR



Parece que a GH e a insulina podem ter as mesmas moléculas para transdução de sinal, uma vez que ambas estas hormonas podem desencadear efeitos rápidos inclusivamente o aumentado número na membrana celular de transportadores da glucose como o Glu 1 e Glu 4 o que aumenta o transporte da glucose.

A GH desencadeia através do IGF-1 a fosfatação de tirosilos no principal substrato do receptor da insulina o IRS-1 assim como do IRS-2.

Estes dois tipos de moléculas IRS-1 e IRS-2 fosfatados nos seus tirosilos pela insulina. ou pela IGF-1 podem depois interactuar com outras moléculas dotadas de domínios SH2, como pode ser o caso da PI3' cinase e da p70 rsk implicada esta na regulação celular e aquela no transporte da glucose e síntese do DNA.

A proteína cinase C (PKC) desempenha um papel na transdução de sinal a partir da GH, estando assinalado um pequeno aumento rápido de DAG, um conhecido activador da PKC.

Parece que a GH induz a produção de DAG através da quebra de fosfatidilcolina pela fosfolipase C acoplada ao GHR através de uma G proteína.

Não se sabe quando a activação da PKC dependente da GH necessita da via JAK2 ou é uma via independente da JAK, havendo ainda a possibilidade da GH activar a PKC através de uma via que envolva proteínas IRS e P13' cinase.

O aumento do cálcio livre intracelular é induzido pela GH parecendo a GH activar os canais de cálcio tipo L por um mecanismo que implica a hidrólise fosfolipídica e activação pela PKC.

### A GH regula acontecimentos ao nível do núcleo

Muitos dos sinais das cascatas fosfatantes de transdução de sinal antes referidas e reguladas pela GH, acabam por se reflectir nas expressões dos genes contidos nas células alvo, parecendo que esta regulação se processa através dos factores de transcrição e dos genes específicos dos tecidos.

A activação da família dos transdutores de sinal e activadores de transcrição (Signal Transducers and Activators of Transcription – STAT) está bem assinalada como interveniente na acção da GH entre o receptor e o núcleo.

Como já referimos as proteínas STAT são factores citoplásmicos contendo domínios SH2 e SH3.

As proteínas STAT após fosfatadas levam á formação de complexos proteicos com outras proteínas STAT ou não STAT que se translocam para o núcleo onde interactuam com o DNA. Esta interacção activa a transcrição dos genes respectivos.

Nas proteínas STAT ou relacionadas, fosfatadas por indução pela GH encontram-se identificadas as STAT1, a STAT3/APRF (factor de resposta de fase aguda), STAT5/MGF (factor da glândula mamária).

Encontram-se STAT proteínas em diversos complexos ligados ao DNA, induzidos pela GH para vários genes.

No promotor C-fos, a GH induz a ligação de três complexos SIS-induzíveis (SIE) pensando-se que contem STAT1 e STAT3.

A activação dos STATs pela GH parece depender do tipo de célula em questão, podendo os STATs ser activados pela GH independentemente uns dos outros. Sugere-se mesmo uma via em que a GH pode eleger acontecimentos no núcleo directamente.

GHR → JAK – STATs → Gene alvo

Parece que a activação das proteínas STAT é um mecanismo comum através do qual as citocinas e os factores de crescimento regulam a transcrição.

A GH induz rapidamente a expressão de diversos genes que codificam factores de transcrição implicados no crescimento e diferenciação celular, como é o caso do c-fos e c-jun do c-myc os dois primeiros podendo regular genes de resposta tardia á GH

Parece que uma via envolvendo GHR → MAP kinase → SRE proteína → Fos permitiria que a GH desencadeasse acontecimentos ao nível do núcleo promovendo o crescimento celular e a diferenciação.

As vias da GHR para o SRE e mediadas pela via STAT para o SIE não são conhecidas.

Nalgumas situações torna-se necessário a exposição a longo termo à GH (horas a dias) para que surjam efeitos ao nível da expressão dos genes. É o caso da diferenciação dos adipocitos e respostas metabólicas destes.

Parece ser aceitável a via GH → Fos → genes específicos do adipocito → diferenciação dos adipocitos.

Nos genes regulados pela GH no fígado está o gene responsável pela formação do insulina-like growth factor1 (IGF1) um mediador do soro de algumas acções da GH.

O gene spi 2-1, codificando um inibidor de uma serina-protease específica do fígado, é estimulado pela GH in vivo e in vitro.

Também a GH regula o gene da somatoestatina.

Parecem pois haver múltiplos mecanismos para a regulação da expressão dos genes pela GHR.

A JAK2 é necessária para a activação da SHC, Grb2, MAP cinase, IRS-1 e 2 e PI-3 cinase, e STATs 1 e 3.

Interacções directas do SHC, IRS1 e 2 e STATs1 e 3 com o GHR não estão demonstradas o que sugere que estas moléculas se ligam a resíduos tirosilo fosfatados na própria JAK2 ou em moléculas acessórias ainda não identificadas.

## 2.5 - Transdução de sinais e receptores intracelulares

### Síntese e libertação das hormonas esteróides

#### Comunicação célula a célula e controlo da síntese e libertação de hormonas esteróides

A segregação das hormonas esteróides a partir das células em que são biossintetizadas é desencadeada por uma série de hormonas tal como se refere no quadro seguinte.

#### Hormonas implicadas na síntese e libertação de hormonas esteróides

<u>Hormona esteróide</u>	<u>Célula ou Sinal</u> <u>estrutura</u> <u>responsável</u>	<u>Segundo</u> <u>mensageiro</u>	<u>Sistema sinal</u>
Cortisol	Cortical da ACTH supra-renal (fasciculata)	cAMP, ciclo PI, Ca <sup>2+</sup>	Cascata hipotálamo- hipófise
Aldosterona	Glomerulosa supra-renal	Angiotensina II / III	Sistema renina- angiotensina
Testosterona	Células Leyding	LH	Cascata hipotálamo- hipófise
17β Estradiol	Folículo ovário	FSH	Hipotálamo hipófise-ciclo ovário
Progesterona	Corpo amarelo	LH	Hipotálamo hipófise-ciclo ovário
1,25 (OH) <sub>2</sub> - vitamina D <sub>3</sub>	Rim	PTH	Luz diurna glândulas paratiróides taxa de Ca <sup>2+</sup> no plasma sanguíneo

ACTH = hormona adrenocorticotropa

LH = hormona luteinizante

FSH = hormona estimuladora dos folículos

PI = fosfatidilinositol

PTH = hormona paratiróide

cAMP = AMP cíclico

Os sinais para a estimulação da biossíntese e segregação das hormonas esteróides são hormonas peptídicas que interactuam com receptores específicos das membranas tal como se esquematiza na figura seguinte.

Todas as hormonas esteróides são sintetizadas a partir do colesterol de uma forma regulada mediada pela concentração intracelular de cAMP e de cálcio, podendo também o inositol trifosfato estar implicado.

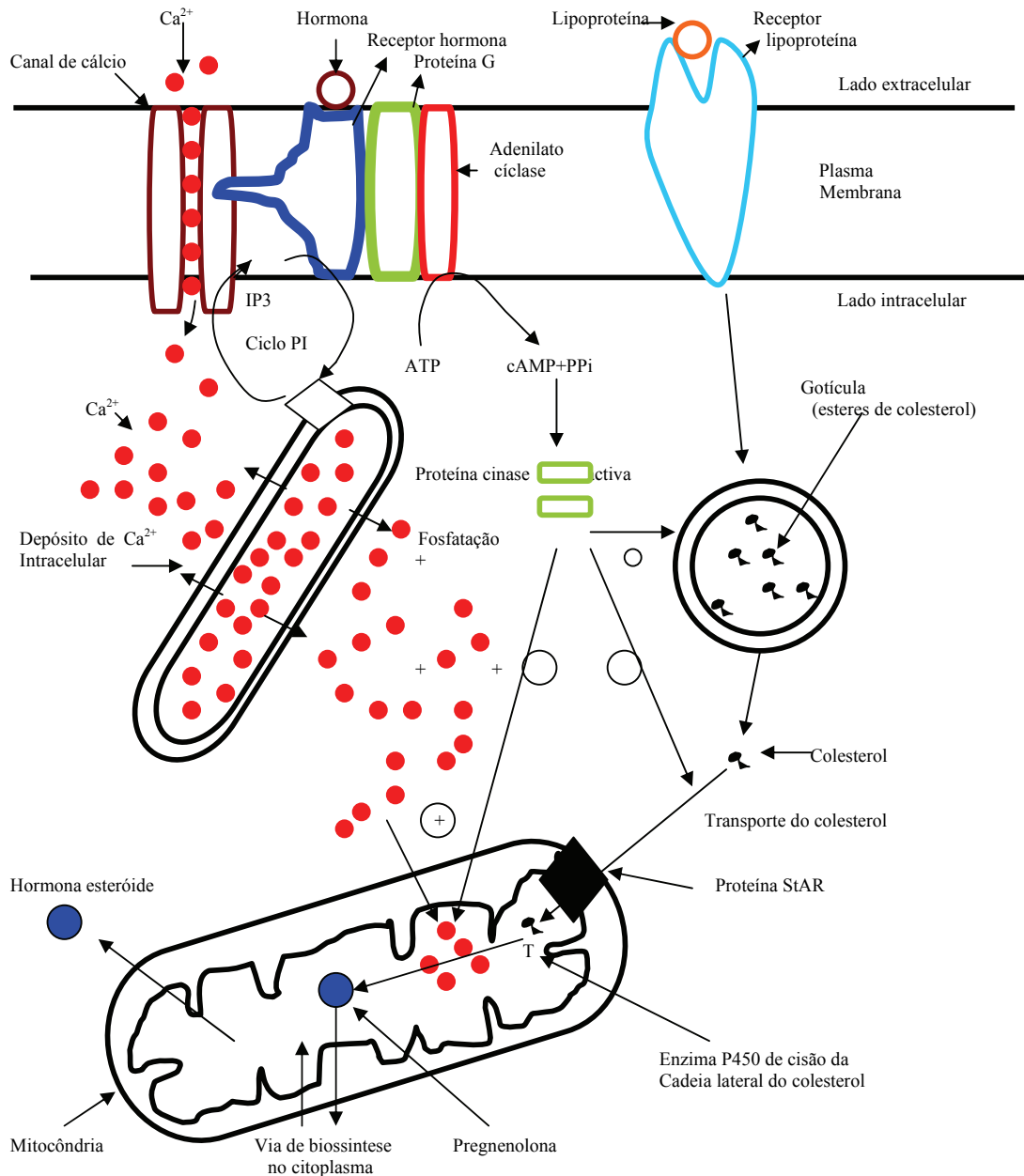
A resposta ao estímulo pelo cAMP pode assumir um carácter agudo (em segundos ou minutos) ou crónico (necessitando de várias horas).

No carácter agudo o colesterol é mobilizado para o folheto interno das mitocôndrias onde é convertido em pregnenolona um precursor de todas as hormonas esteróides pela enzima citocromo p-450 que quebra a cadeia lateral do colesterol.

Os efeitos crónicos traduzem-se no aumento da transcrição dos genes que codificam as enzimas implicadas no biossintese das hormonas esteróides.

Nos efeitos agudos a translocação do colesterol para o folheto interno das mitocôndrias é facilitado por uma fosfoproteína de 30-kda chamada proteína esteroideogénica aguda reguladora (StAR) cuja biossintese é induzida pela hormona.

## Estimulação hormonal da biossíntese das hormonas esteróides em geral



Consoante o tipo de célula e o tipo de receptor por elas contido assim a natureza da hormona que com eles vai interagir e assim a natureza da hormona produzida.

O ACTH induzirá a síntese de cortisol, a FSH a do estradiol, a LH a da testosterona, etc.

Tal como se esquematiza na figura em cima, a hormona interagindo com o seu receptor específico, activa uma adenilato ciclase através de uma G-proteína o que provoca o aumento de cAMP e consequente activação da proteína cinase A que por sua vez promove fosforilações de outras proteínas



enzimáticas que aumentam a hidrólise dos ésteres de colesterol depositados nas gotículas (à direita na figura) do que resulta a libertação do colesterol e seu transporte para dentro das mitocôndrias, o que é facilitado pela indução pela hormona da proteína StAR.

No entanto a hormona ao activar o receptor respectivo pode também condicionar um canal para o cálcio existente na plasma membrana celular, o que pode acontecer directamente ou indirectamente através de um ciclo fosfatidilinositol (ciclo PI) indicado na figura entre o receptor da membrana e um depósito intracelular de  $Ca^{2+}$ .

Se o ciclo PI for activado, a concentração citosólica de IP3 e de cálcio aumenta uma vez que este é libertado dos seus depósitos intracelulares.

As concentrações aumentadas de  $Ca^{2+}$ , a indução da biossíntese da proteína StAR, e as fosfatações desencadeadas pelo PKA em conjunto promovem o aumento da cisão da cadeia lateral do colesterol aumentando por outro lado a biossíntese das hormonas esteróides, a partir do precursor pregnenolona.

As hormonas esteróides libertadas para o sangue são aqui transportadas em formas ligadas com proteínas específicas e com albumina, em ordem a proteje-las do seu metabolismo e consequente inactivação

Estas proteínas de ligação são as globulinas para ligação aos corticoesteróides (CBG) ou transcortinas, as globulinas para ligação às hormonas sexuais (SHBG), as de ligação aos androgénios (ABP), além das albuminas e de outras proteínas de transporte e que acabam por fornecer as células de esteróides livres após dissociação e que podem interaccionar então com as células alvo.

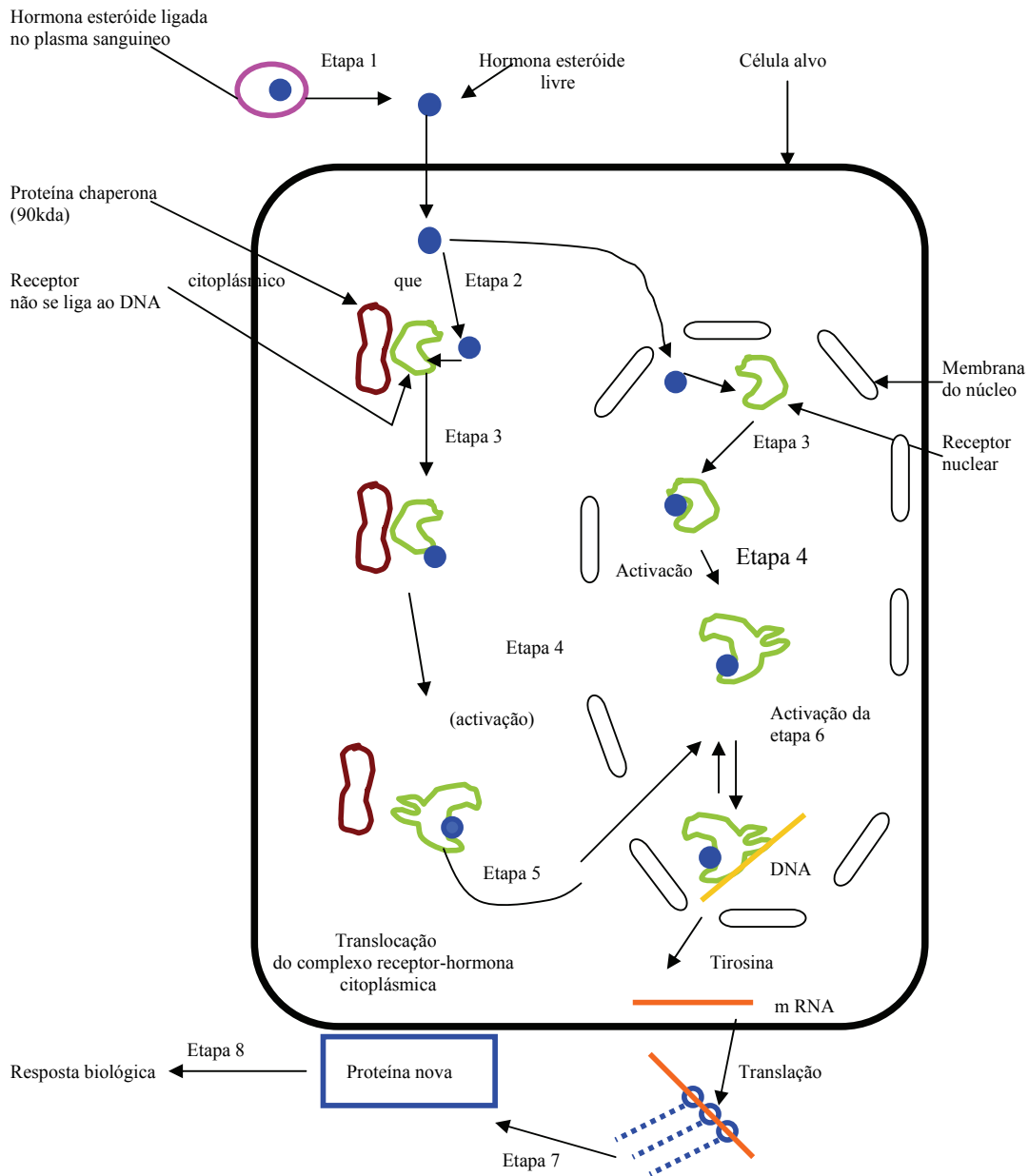
Como já referimos as hormonas esteróides livres ligam-se com receptores proteicos intracelulares específicos citoplásmicos ou nucleares tal como se pode verificar na figura seguinte.

Realça-se mais uma vez que ao contrário do que sucede com os receptores para as hormonas peptídicas que geralmente se encontram situados nas superfícies celulares, os receptores das hormonas esteróides e outros receptores de hormonas não esteróides, como por exemplo hormonas da tiróide, ácido retinóico, Vit. D3, localizam-se no interior das células.

### **Esquema da acção das hormonas esteróides**

O receptor glucocorticóide e possivelmente também o da aldo esterona parecem residir no citoplasma, enquanto os receptores para o estradiol, progesterona, androgénios e vit D3 parecem residir no núcleo.

## Esquema da acção das hormonas esteróides



Na etapa 1 a hormona esteróide ligada a uma proteína transportadora adequada, no sangue, dissocia-se desta proteína o que lhe permite interagir na etapa 2 com um receptor específico inactivo que pode consoante os casos ter localização citoplásmica ou nuclear.

A interacção da hormona converte o receptor de uma forma inactiva numa forma activa (etapa 3 e 4) o que decorre das modificações da configuração do complexo formado salientando-se que no receptor citoplásmico esta interacção é ajudada por uma proteína chaperona o que subsequentemente permite que o complexo receptor hormona de transloque para o núcleo.

Na etapa 6 estes complexos activados, no interior do núcleo, reconhecem e interactivam com elementos de resposta específica no DNA o que induz a expressão de proteínas novas codificadas pelos genes que respondem às hormonas (etapa 7) e consequente resposta biológica(etapa 8) ao sinal desencadeado pela hormona e que são alterações fenotípicas ou metabólicas mediadas nas células alvo pelas proteínas induzidas especificamente.

Salienta-se que após a ligação do complexo activado hormona-receptor aos locais alvo no DNA podem ocorrer activações ou pelo contrário inibições da transcrição.

### **Receptores esteróides e elementos de resposta no DNA**

A identificação de sequências de DNA, habitualmente localizadas nas regiões flangeadoras 5' dos genes regulados que medeiam os efeitos da regulação hormonal sobre a transcrição dos genes é extraordinariamente importante e tais sequências de DNA são chamadas elementos de resposta hormonal (HREs).

Estão definidas para diversas hormonas estes HRE, bem como são conhecidas algumas proteínas que após activação pelas hormonas, podem ligar-se directamente ou indirectamente às HRE e medeiam efeitos transcripcionais.

Glucocorticóides, androgénios, mineralocorticóides e progestina, todos podem regular a transcrição através de uma sequência consenso HRE idêntica.

No quadro seguinte referem-se HREs consenso para uma série de hormonas, referidas por Lucas em 1992.

Elemento de resposta às hormonas (HRE)

GRE (glucocorticóides, androgénios, mineralocorticóides, progestinas)	←—— ———→ AGAACA nnn TGTTCT
ERE (estrogenios)	←—— ———→ AGGTCA nnn TGACCT
TRE (hormonas das tiróides, ácido retinoico)	←—— ———→ AGGTCA - - - TGACCT
CRE (cAMP)	←—— ———→ 5'-TGACGTCA-3'
AP-1-Site (ésteres de forbol)	←—— ———→ TGA (b/c) TCA

É sabido que os HRE para as hormonas de família esteróide são palíndromas na estrutura, por vezes até consistindo em sequências repetidas invertidas separadas por um intervalo de três nucleótidos. É o caso da sequência AGAACA nnn TGTTCT do quadro anterior que liga o complexo hormona glucocorticóide-receptor mas que liga também vários outros complexos ligando-receptor, com a transactivação dos genes correspondentes.

O HRE para os estrogénios já difere do anterior mas também consiste de uma sequência repetida invertida separada por três nucleótido enquanto o HRE para as hormonas de tiróide é idêntico ao da ERE mas sem ter os três nucleótidos a separar a sequência repetida invertida.

Leves alterações na sequência primária de um HRE pode ter profundas consequências sobre a sua actividade.

A identificação de um HRE pressupõe a demonstração da ligação com o complexo com o ligando-receptor e na transactivação da expressão dos genes o que pode implicar a necessidade de outros intervenientes.

Os receptores intracelulares, específicos para cada hormona têm grande semelhança entre si na sua organização estrutural, mas por outro lado tem diversos outros domínios funcionais dentro de molécula receptora. Cada um destes receptores possui um domínio de cerca de 70 ácidos aminados que se liga ao DNA, domínio este que se dobra em dois motivos “zinco finger”.

Parece que a discriminação na ligação ao DNA deste domínio, para os diversos membros da superfamília de receptores dos esteróides / tiróide / ácido retinóico, se baseia nos três ácidos aminados na base do primeiro “finger”. Um segundo domínio funcional comum a todos estes receptores é o domínio C-terminal que se liga à hormona e que tem uma sequência de ácidos aminados que varia de receptor para receptor reconhecendo cada uma das hormonas de per se.

Um domínio para dimerização existe ainda nesta família de receptores consistindo este domínio de nove héptades repetidas de ácidos aminados hidrófobos que podem dimerizar-se na interface ao longo de um lado da “coiled-coil  $\alpha$  helix” numa estrutura semelhante ao motivo de dimerização “leucine zipper” presente noutros factores de transcrição.

Esta dimerização dos complexos hormona receptor ocorre antes da ligação ao DNA.

As sequências consenso no DNA correspondentes a elementos específicos de resposta às hormonas (HRE) para ligação de diversos complexos receptor-hormona\_\_activados referem-se seguidamente, com mais promenor do que anteriormente.

Elementos do DNA que respondem a receptores-hormonas esteróides  
(loais aceitadores consenso, referidas por Beato em 1989)

<u>Elemento</u>	Sequência de DNA
<u>Positivo</u> (aumentam o ritmo de transcrição do gene associado)	
- Elemento de resposta a glucocorticóides (GRE)	} 5'-GGTACAnnnTGTTCT-3' vide que esta sequência esquema parece ser AGAACA segundo Lucas em 1992
- Elemento de resposta a mineralocorticóides (MRE)	
- Elemento de resposta à progesterona (PRE)	
- Elemento de resposta aos androgénios (ARE)	
- Elemento de resposta aos estrogénios (ERE)	5'- <u>AGGTCAnnnTCACT</u> -3'
- <u>Negativo</u> (reprimem a transcrição de genes específicos)	
- Elemento de resposta aos glucocorticóides	5'- <u>ATYACnnnTGATCW</u> -3'

n = qualquer nucleotido  
Y = purina  
W = pirimidina

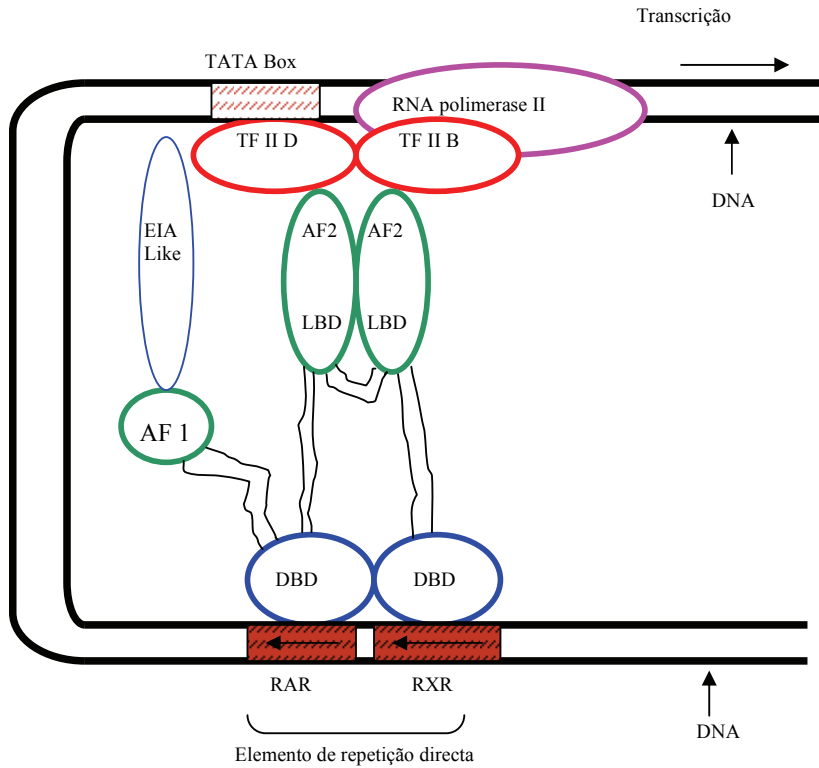
No DNA o mesmo HRE liga-se com os complexos receptores para os glicocorticóides, mineralocorticóides, progesterona e androgénios logo é o tipo de células e o tipo de receptor expresso que determinará a resposta hormonal. Neste contexto é conhecido que os receptores para as hormonas do sexo apenas são expressas em alguns tipos celulares, tal como o receptor para a progesterona só é expresso em certos tipos de células, enquanto outros tipos de receptores como os para os glucocorticóides já são muito ubíquos. Ainda no caso de serem coexpressos os receptores para a aldosterona e para o cortisol apenas uma forma predomina consoante o tipo de célula.

Para a eficiente ligação ao DNA dos complexos receptores das hormonas esterólicas, é necessária a sua dimerização, homodimerização ou heterodimerização, só depois ocorrendo a correspondente transcrição.


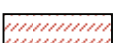






Nestas estruturas destes receptores ocorrem pois domínios implicados nessa dimerização. Estes domínios apresentam uma estrutura helicoidal que pode conter uma série de sequências hidrófobas do que pode resultar a formação de estruturas "leucina-zipper like" ou um motivo hélice-volta-zipper que já se sabe que são necessários para a dimerização de outros factores de transcrição.

Embora a maioria destes receptores esterólicos formem homodímeros, também existem heterodímeros como no caso do exemplo da figura seguinte. Nesta figura apresentam-se receptores do ácido retinóico (uma classe distinta de receptores) classificados como retinóides X receptores (RXR) que regulam a expressão dos genes através de uma heterodimerização com outra forma distinta do receptor ácido retinóico a RAR, o receptor da hormona tiróide e outros membros desta superfamília de receptores.

Complexo de pré-iniciação de transcrição e sua estabilização pelo heterodímero RXR / RAR

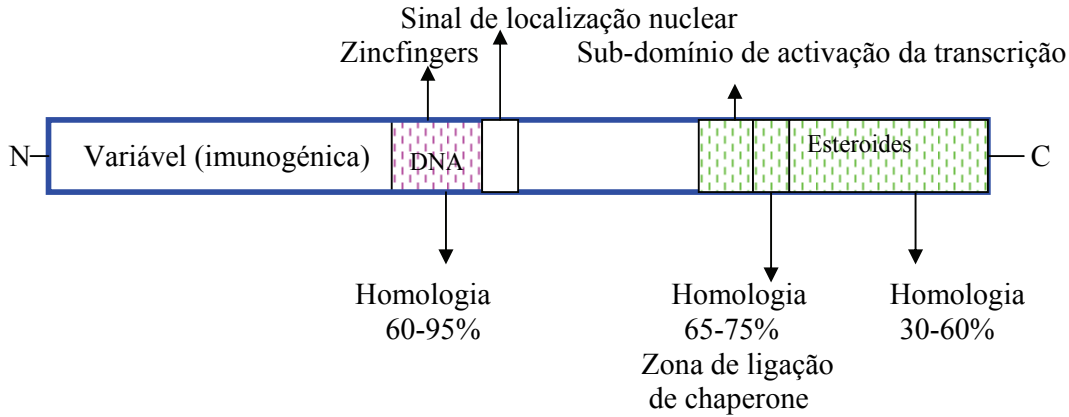


Legenda da figura

-  DNA
-  Caixa TATA
-  Local do DNA para interacção com os receptores heterodimericos
-  RNA polimerase II
-  Factores de transcrição gerais (TF II D e TF II B)
- Receptor {
  -  -LB D (domínio de ligação ao ligando). AF2 função de activação dentro do domínio de ligação ao ligando que interactua directamente com a maquinaria da transcrição.
  -  AF1 função de activação dentro das regiões amino terminal do receptor que interactua com proteínas específicas da célula.
  -  -DBD (Domínio de ligação ao DNA)

O receptor glucocorticóide pode ser dividido em três domínios funcionais como se esquematiza na figura seguinte, mas estes aspectos a seguir referidos são comuns a todos os receptores esteróides.

Módulo de um receptor típico glucocorticóide para hormona esteróide



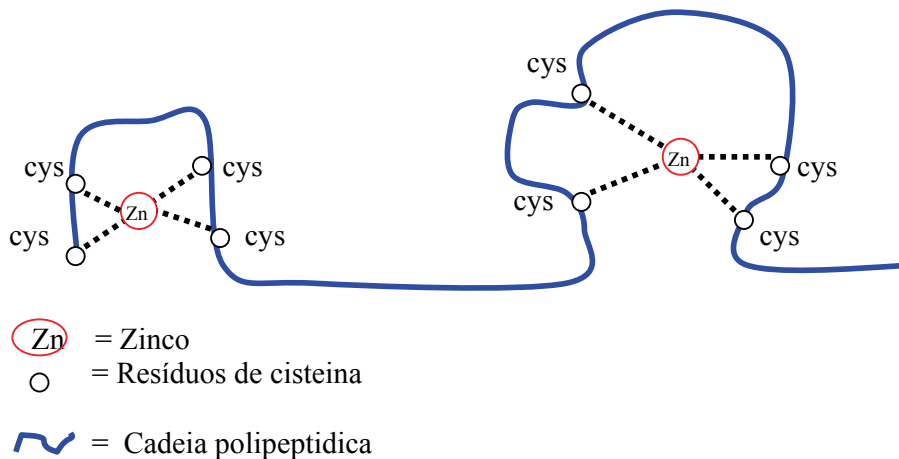
Na extremidade C o receptor tem um domínio que tem uma homologia de 30 a 60% com os domínios de ligação a ligandos de outros receptores na família de receptores esteróides.

Ainda neste domínio existe uma sequência que se liga a proteínas chaperones e á esquerda desta uma região que modifica a transcrição.

No centro existe um domínio que se liga ao DNA sendo este domínio homólogo de 60 a 95% com todos os outros receptores esteróides, havendo dois “zinc fingers” que interactuam com o DNA e cuja estrutura se mostra na figura seguinte.

A extremidade N do receptor contem domínios antigénicos e um local que modela a activação da transcrição variando muito estas sequências de ácidos aminados consoante os receptores.

Estrutura do “zinc finger” localizado no domínio de ligação ao DNA do receptor glucocorticóide



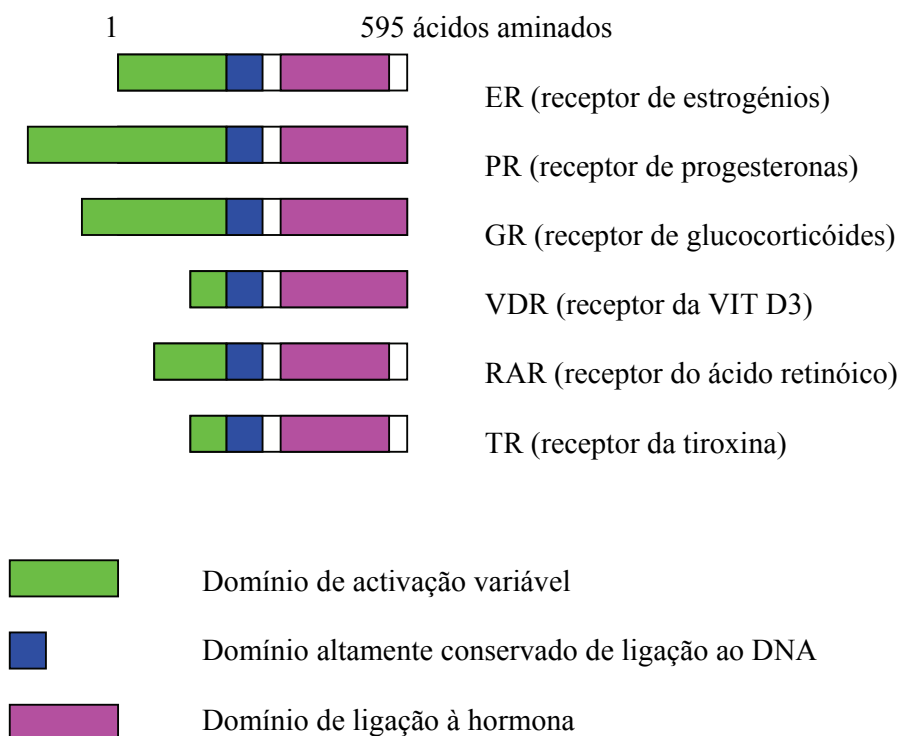
Como vimos diversas hormonas e factores de crescimento desempenham os seus papeis biológicos através da activação de receptores nas superfícies das células.

Como vimos também as hormonas esteróides penetram nas células alvo onde se ligam a receptores específicos, constituindo-se complexos hormona-receptor que migram para o núcleo onde se ligam a locais específicos do DNA, cada um deles induzindo ou reprimindo a transcrição de um conjunto particular de 50 a 100 genes.

As hormonas esteróides alteram sobretudo o perfil da expressão dos genes mais do que a actividade de um enzima em particular ou de um transportador membranário.

Os receptores para as hormonas esteróides tais como estrogénios, progesterona e glucocorticóides têm motivos comuns (vide figura seguinte), contendo um domínio conservado de 66 resíduos que se liga ao DNA e um domínio conservado de 240 resíduos que se liga à hormona, enquanto o domínio amino-terminal, que não é conservado, facilita a interacção do receptor com outros reguladores da transcrição.

#### Estrutura dos Domínios da superfamília dos receptores nucleares



Os receptores activados ligam-se a sequências específicas do DNA chamadas como vimos elementos de resposta às hormonas (HRE) que regulam a expressão dos genes vizinhos.

Como dissemos anteriormente HREs para os receptores esteróides são palíndromas constituídos por sequências de 6-bp separados por um espaçador de 3-bp (vide figura seguinte).



### Elemento de resposta glucocorticóide (GRE)

N = qualquer nucleótido

5'- NAG**AA**CANNNTG**TT**CTN-3'



3'-NTC**TT**GTNNNACA**AG**AN-5'

Elemento de resposta aos estrogénios (ERE)

5'-NAG**GT**CANNNTG**AC**CTN-3'



3'-NTC**CAG**TNNNACT**GG**AN-5'

● = Eixo de simetria

Na figura anterior o elemento de resposta glucocorticóide difere do elemento de resposta estrogénico em ter TT em vez de AC nas posições +5 e +6 do palindroma.

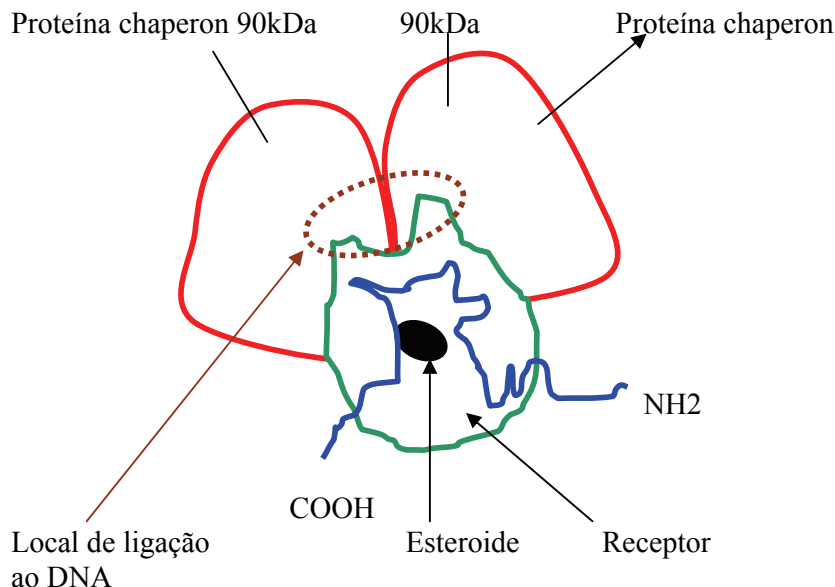
Os domínios de ligação ao DNA dos receptores glucocorticóides e estrogénios são capazes de distinguir entre estes elementos em virtude de diferirem em 3 resíduos chave.

Cada cadeia polipeptídica do receptor é estabilizada como já vimos por um par de “clusters” zinco, ou seja um ião zinco está coordenado tetrahedricamente com quatro cadeias laterais de cisteínas.

Alguns receptores esteróides fazem parte da família de proto-oncogenes cErbA ou vErbA. O V-ErbA é um oncogene que se liga ao DNA mas que não tem nenhum domínio para união a um ligando.

De uma maneira geral a activação dos receptores esteróides converte uma forma do receptor que não se liga ao DNA numa forma activada capaz de se ligar não especificamente ao DNA ou especificamente ao DNA no HRE (elemento que responde à hormona).

A forma dos receptores citoplásmicos que não se liga ao DNA é um trimeroheteromérico constituído por uma molécula do receptor e um dímero de proteínas chaperon de 90kDa (vide figura seguinte).



O local de ligação ao DNA do receptor é bloqueado pelas proteínas heteroméricas, e por activação este complexo desagrega-se levando o local de ligação ao DNA a ficar plenamente exposto para interactuar com o DNA, o que pode ser iniciado pela ligação do esteróide que produz uma alteração da conformação da proteína receptora.

## **2.6 - Transdução de sinais e transcrição**

Tal como vimos anteriormente em mais detalhe, na transdução de sinais com receptores de sete hélices, na transdução de sinais com receptores tirosina cinases intrínsecas, na transdução de sinais com tirosina cinases citosólicas e na transdução de sinais com receptores intracelulares, há muitos destes sinais com potencialidades para condicionar a transcrição de genes.

Nas hormonas actuando através de receptores à superfície das células, há uma série delas que influenciam a transcrição de genes alvo através de mecanismos intracelulares variados e complexos, com cascatas de segundos mensageiros como sucede no caso das hormonas peptídicas, activando por esta via um ou mais factores de transcrição.

Um dos melhores conhecidos exemplos da regulação da transcrição por hormonas peptídicas é aquele em que intervem a adenilatociclase, com elevação do cAMP e activação da proteína cinase A. Numa série de genes regulados pelo c-AMP uma sequência curta palindroma CRE (vide quadro anterior) é capaz de mediar a transcrição de genes induzidos pelo c-AMP. Este elemento de resposta CRE provavelmente forma um dímero que se liga ao genes promotores com a ajuda de uma proteína de ligação CREB que provavelmente se dimeriza através de um distinto motivo leucina-zipper ligando-se ao DNA através de uma região básica. Parece haver numerosas isoformas CREB que são todas membros da família dos factores de transcrição.

A transcrição dos genes é controlada através de outros meios, inclusivé por alguns factores de crescimento e pela via da proteína cinase C dependente do diacilglicerol.

Também os produtos dos proto-oncogenes nucleares C-fos e C-jun medeiam a estimulação da transcrição iniciada pela proteína cinase C e são factores de transcrição, tal como CREB.

O complexo heterodimérico da Fos e Jun, denominado factor de transcrição AP-1 liga-se a curtas sequências de DNA (vide quadro anterior) presentes numa série de genes regulados por ésteres de forbol.

Contudo a regulação da transcrição por uma série de hormonas polipeptídicas é muito mal conhecida.

Na insulina que regula a transcrição de numerosos genes apenas são conhecidas algumas sequências de DNA responsáveis por esses efeitos.

Também a activação de genes pelos interferões é feito através de elementos de DNA muito discretos.

Enquanto a maioria das HRE medeiam efeitos positivos das hormonas sobre a transcrição, há numerosos exemplos nos quais as hormonas inibem a transcrição.

A insulina por exemplo induz a expressão de vários genes, mas também é capaz de inibir outros.

Elementos de resposta hormonal negativa (nHREs) têm sido descritos para varias outras hormonas, como por exemplo a T3 que inibe a transcrição do promotor do gene do receptor do factor de crescimento epidérmico (EGF).

## 2.7 - Bibliografia

- Daniel, P.B., Walker, W.H. e Habener, J.F. (1998). Cyclic AMP signaling and gene regulation. Annu. Rev. Nutr. 18: 353 - 83
- Carter, Su, C., Schwartz, J. E Smit, L.S. (1996). Molecular mechanisms of growth hormone action. Annu. Rev. Physiol 58: 187 - 207
- Argetsinger, L.S., Campbell, G.S., Yang, X., et al. (1993). Identification of JAK2 as a growth hormone receptor-associated tyrosine kinase. Cell 74: 237 - 44
- Cobb, M.H., Robbins, D.J., Boulton, T.G. (1991). ERKs extracelular signal-regulated MAP-2 kinases. Curr. Opin. Cell Biol 3: 1025 - 32
- Hubbard, S.R., e Till, J.H. (2000). Protein tyrosine kinase structure and function. Ann. Rev. Biochem 69: 373 - 98
- Sato, K., Sato, A., Aoto, M., Fukami, Y. (1995). Biochem Biophys Res. Commun 215: 1078 - 87
- Biscondi, J.S., Maa, M.C., Tice, Da, Lox, M.E., Leu, T.H. et al (1999) J. Biol. Chem 274: 8335 - 43
- Fantl, W.J., Johson, D.E. e Williams, L.T., (1993) Signaling by receptor tyrosine kinase. Annu. Rev. Biochem 62: 453 - 81
- Davis, S., Lu, M.L., Lo, S.H., Lin, S., Butler, J.A., et al (1992). Science 256 : 712- 15
- Schindler, C., e Darnell Jr, J.E., (1995). Transcriptional responses to polypeptide Ligands : The JAK – STAT pathway. Ann. Rev. Biochem 64: 621 - 51
- Brockman, R.P. e La arveld, B., (1986). Hormonal regulation of metabolism in ruminants: a review. Livestock Production Science 14: 313 - 334
- Hannonne, J. e Defer, N. (2001). Regulation and nole of adenylyl cyclase isoforms. Ann. Rev. Pharmacol. Toxicolo 91 :145-174
- Daniel, P.B.,Walker W.H. e Habener, J.F. (1998). Cyclic AMP sinaling and gene regulation. Ann. Rev. Nutr. 18: 353 - 383
- Baggiolini, M., Dewald, B. E Moser B. (1997). Human chemokines: an update. Annu. Rev. Immunol 15 : 675 - 705
- Fernandez, E.J. e Lolis, E. (2002). Structure, function, and inibition of chemokines. Annu. Rev. Pharmacol Toxicol 42: 469 - 499
- Silva, A.J., Kogan, J.H., Frankland, P.W. e Kida, S. (1998). Creb and memory. Annu.Rev. Neurosci 21: 127 - 148
- Hossner, K.L., McCusker, R.H., e Dodson, M.V., (1997) Insulin-like growth factors and their binding proteins in domestic animals. Animal Science 64: 1 - 15
- Butler, A.A., e Roith, D.L., (2001). Contos of growth by the somatropic axis: growth hormone and the insulin like growth factors have related and independent roles. Annu. Rev. Physiol. 63: 141 - 64
- Klapper, D.G., Svoboda, M.E., Van Wyk, J.J., (1983). Seqeunce analysis of somatomedin-C: confirmation of identity with insulin-like growth factor I. Endocrinology 112: 2215 - 17
- Pitcher, J.A., Freedman, N.J., e Lefkowitz, R. (1998). G-protein-coupled receptor kinase. Ann. Rev.Biochem 67: 653 - 92
- Shaywitz, A.J., e Greenberg, M.E., (1999) – CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. Annu.Rev. Biochem 68: 821 - 861
- Montminy, M. (1997). Annu. Rev. Biochem.66: 807 - 22
- Stark, G.B., Kerr, I.M., Williams, B.R.G., Silverman, R.H., e Schreiber, R.D. (1998). How cells respond to interferons. Annu. Rev. Biochem 67: 227 - 64
- Kaibuchi, K., Kuroda, S., e Amano, M. (1999) – Regulation of the cytoskeleton and cell adhesion by the RHO family GTPases in mammalian cells. Annu. Rev. Biochem 68: 459 – 86
- Stock A.M., Robinson, V.L., e Goudreau, P.N., (2000) Two-component signal transduction. Ann. Rev. Biochem 69: 183 - 215

- Lucas, P.C., e Granner, D.K. (1992). Hormone response domains in gene transcription. Annu. Rev. Biochem 61: 1131 - 73
- Massagué, J., e Pandieilla, A., (1993) – Membrane – anchored growth factors. Annu. Rev. Biochem 62: 515 - 41
- Lowy, D.B., e Willumsen, B.M. (1993) – Function and regulation of RAS. Annu. Rev. Biochem 62: 851 - 91
- Walker, R.A., e Sheetz, M.P. (1993) – Cytoplasmic microtubule-associated motors. Annu. Rev. Biochem 62: 429-51
- Nuoffer, C. E Balch, W.E. (1994). GTPases: multifunctional molecular Switches regulating Vesicular traffic. Annu. Rev. Biochem 63: 949-96
- Pryer, N.C., Wvestihube, L.J., e Schekman, R., (1992) Vesicle-mediated protein sorting. Annu. Rev. Biochem 61: 471-516
- Wells, J.A., e Vos, A.M., (1996)-Hematopoietic receptor complexes. Annu. Rev. Biochem 65: 609-34
- Nicola, N.A., (1994). In Guidebook to cytokines and their receptors, pp1-194. Oxford: Oxford Univ. Press
- Lodish, H., Baltimore, D., Berk, A., Zipursky, S.L., Matsudaira, P., e Darnell, J. (1995) Molecular Cell Biology 3<sup>m</sup>Ed. – Scientific American Books
- Devlin, T.N. (1997) – Textbook of Biochemistry with clinical correlations -4<sup>th</sup> ed Wiley-Liss/New York

## **Iconografia**

As figuras e quadros(bem como os textos que lhe dão apoio) indicados na respectiva iconografia foram adaptados e/ou redesenhados a partir das seguintes fontes:

Pagina 4-Cascata hormonal de sinais.In Devlin,T.M. Textbook of biochemistry with clinical correlations,Wiley-Liss,1997-pag.841.

Pagina 7 a 12-Principais hormonas peptidicas dos organismos animais e suas funções.In idem anterior pag. 847 a 849.

Pagina 18-Organização em domínios de uma série de famílias de RTK.In Hubbard,S.R.&Jeffrey,H.Till-Protein tyrosine kinase struture and function.Ann.Rev.Biochem. 2000;69,pag 377

Pagina 19-Organização em domínios das principais sub-famílias de NTRK.In idem anterior pag.378.

Pagina 21-Proteínas cinases e tamanho e localização dos seus domínios catalíticos.In Devlin,T.M. Textbook of biochemistry with clinical correlations,Wiley-Liss,1997-pag.879.

Pagina 24-Celula do tubulo renal distal. In idem anterior pag.882.

Pagina 24-Hormonas utilizando o cAMP como segundo mensageiro.In Stryer,L.Biochemistry,W.H. Freeman and Company,New York,1995,pag.341.

Pagina 25-Acção intracelular da proteína cinase C.In Devlin,T.M. anterior,pag.884.

Pagina 26-Angiotensina. In idem anterior pag.905.

Pagina 27-Reacções levando á secreção da aldosterona nas células da zona glomerulosa da supra renal.In idem anterior pag.904.

Pagina 28-Atriopeptidina.In idem anterior pag.887.

Pagina 29-Outro exemplo de activação pelo aumento da concentração de cálcio e DAG é a regulação celular da secreção exócrina do pâncreas.In idem anterior pag.1061.

Pagina 31-Proteínas cinases G.In idem anterior pag.886 e 887.

Pagina 35 a 38-Receptores.InLodish, H.,Baltimore, D., Berk, A., Zipursky, S.L., Matsudaira, P., & Darnel, J.Molecular cell biology.Scientific American Books.W.H. Freeman and Company,New York,1995,pag.862.

Pagina 41-Receptor de sete helices.In idem anterior pag.869.

Pagina 43 -Receptor de sete hélices.In Devlin, T.M. anterior pag.876.

Pagina 44-Adenilato ciclase.In Lodish, H. E all.anterior pag. 875.

Pagina 49 e 50.-Adenilato ciclase. In idem anterior pag. 876 a 879.

Pagina 51-Processos fisiológicos mediados por proteínas G. In Stryer L. anterior pag. 346.

Paginas 60 e 61-Activação da Ras após ligação de uma hormona( por exemplo EGF) ao receptor RTK. In Lodish, H. e all. Anterior pag. 890 e 895.

Pagina 64-Via de transmissão de sinal RTK-Ras. In idem anterior pag.897.

Pagina 66-Vias de transdução de sinal.In idem anterior pag. 915.

Paginas 67 a 72-CREB. In Shaywitz,A. J. & Greenbery,M.E..CREB.A stimulus-induced transcription factor activated by a diverse array of extracellular signals.Ann. Rev. Biochem. 1999;65,pag. 832,837,841,842 e 843.

Pagina 74- Acção da insulina. In Lodish, H. et all. anterior pag. 908.

Paginas 75 e 76-Activação da proteina Ras pela ligação da insulina ao seu receptor tirosina cinase atípico. In Lodish, H. et all. Pag 910.

Pagina 77-Possíveis efeitos da insulina sobre a fosforilação de proteínas em duas vias metabólicas distintas. In Devlin, T. M., anterior pag. 880.

Pagina 78- Transdução de sinais induzida pela insulina. In Devlin,T. M., anterior pag. 881.

Pagina 86-Vias de transdução de sinal levando á activação do factor de transcrição SRF. In Lodish,H. et all. anterior pag. 915.

Página 88-Citocinas. Martins Silva,A.-Apontamentos de Química Fisiológica. Faculdade de Medicina de Lisboa.

Página 90-Ligandos hematopoiéticos, receptores e transdução de sinais. In Wels,J.A. & Vos.A. M.Hematopoietic receptor complexes.Ann. Rev. Biochem. 1996;65, pag.611.

Página 94-Transdução de sinal das citocinas. In Schindler,C & Darnell,Jr,J.E. Transcriptional responses to polypeptide ligands.The JAK-STAT pathways.Ann. Rev. Biochem.,1995;64,pag. 627.

Páginas 95 a 97-Activação de JAK cinases por diversos ligandos.Activação de proteínas STAT por diversos ligandos. In idem anterior pag. 625 e 633.

Página 99- IFN $\gamma$  . In Lodish, H. et all. anterior pag. 917.

Página 101-Esquema de receptor de interferões. In Stark, G. B., Kerr, I. M. et all.How cells respond to interferons.Ann. Rev. Biochem. 1998;67, pag. 229.

Página 102-IFN  $\gamma$  . In idem anterior pag. 231.

Página 103-IFN  $\alpha/\beta$  . In idem anterior pag.237.

Página 105-Organização h GH. In Devlin, T. M. idem anterior pag.829.

Página 109-GHR. In Carter-Su, C.,Schwartz,J. & Smit, L. S.Molecular mechanism of growth hormone action.Ann. Rev. Physiol. 1996;58, pag. 188.

Página 111-Vias sinalizadas pela GHR. In idem anterior pag.191.

Página 115-Hormonas esteroides. In Devlin, T.M. idem anterior pag. 903.

Página 116-Estimulação hormonal da biossíntese das hormonas esteroides. In Devlin, T. M. idem anterior pag 898.

Página 118-Esquema de acção das hormonas esteroides. In Devlin, T. M. idem anterior pag 909.

Página 120-Elementos do DNA que respondem a receptores-hormonas esteroides. In Devlin, T. M. idem anterior pag. 911.

Página 122- Complexo de pré-iniciação da transcrição. In Devlin,T. M. idem anterior pag. 912.

Página 123-Modulo de um receptor para hormona esteroide.Estrutura “zinc finger”. In Devlin, T. M. idem anterior pag. 913.

Página 124-Estrutura dos domínios da superfamilia dos receptores nucleares. In Styer,L. idem anterior pag.1001.

Página 125-Modelo hipotético de uma forma não se ligando ao DNA de um receptor esteroide. In Devlin, T. M. idem anterior pag. 915.

<b>3. - <u>Transcrição, seus factores e regulação</u></b>	294
3.1 – Introdução	294
Regulação da iniciação da transcrição	294
RNA polimerase II	295
RNA polimerase I e III	297
A transcrição nos eucariotas é controlada pela associação combinada de diversas proteínas	297
3.2 - Importância da estrutura das proteínas factores de transcrição nas interacções com o DNA	298
Factores de transcrição. Famílias estruturais e princípios no reconhecimento do DNA	298
As estruturas secundárias são elementos importantíssimos da arquitectura das proteínas	402
Factores de transcrição eucariotas	405
3.3.- Factores gerais de transcrição e RNA polimerase II, I e III	411
Factores de transcrição	412
Transcrição pela RNA polimerase II	413
RNA polimerase II e o complexo de início da transcrição	414
Regulação da actividade dos factores de transcrição eucariotas	415
Transcrição dos genes RNA ribossomais	416
Transcrição pela RNA polimerase III	418
3.4 - Transcrição de genes codificadores de proteína	420
Cromatina e sua remodelação	420
RNA polimerase II e cofactores de iniciação	426
RNA polimerase II e cofactores de alongamento	427
3.5 - Biologia do factor de transcrição II D (TF II D)	428
3.6 - Complexos coactivadores da transcrição	433
Bibliografia	440
Iconografia	441



### **3. - Transcrição, seus factores e regulação**

#### **3.1 – Introdução**

##### **Regulação da iniciação da transcrição**

Em todos os seres vivos cada um dos diversos genes integrados no respectivo genoma pode ser regulado de várias formas no seu funcionamento, o que equivale a dizer que o mesmo se passa com os produtos da expressão desses genes.

Encontrando-se nos animais superiores esses genes integrados na cromatina, a forma como esta é aberta em ordem a permitir a transcrição da mensagem, estará dependente de uma série de complexos proteicos que actuam sobre os componentes dessa cromatina ou sejam o DNA e as proteínas histonas ou outras.

Como se compreende estes acontecimentos podem ocorrer num determinado momento e num determinado ponto ou parte da totalidade do genoma, ocorrendo variações na quantidade dessa expressão, tudo dependendo dos sinais desencadeadores dessas actividades que chegam até ao genoma e ainda das características próprias dos órgãos ou tecidos implicados.

O próprio DNA é regulado na sua transcrição através da sua própria estrutura, de diversas formas, por exemplo através de isoladores, repetições e vírus endógenos, metilação de ilhas CpG, ou de regiões favorecedoras dos chamados promotores de transcrição ou de regiões intrónicas que regulam a transcrição. Como se compreende no genoma em toda a sua extensão a distribuição de algumas destas características é muito irregular.

Como se compreende também a tradução ou translação dos genes em proteínas é um processo muito complexo, em que ocorre primeiramente a transcrição do mRNA, havendo nesta transcrição uma fase de iniciação, outra fase de alongamento e uma fase de terminação, sendo a primeira destas fases aquela que é mais finamente regulada.

Esta regulação permite por exemplo que células procariotas se adaptem ao seu ambiente nutritivo e que células eucariotas se diferenciem originando os diversos tipos de tecidos e órgãos.

Nas bactérias um conjunto de genes contidos num único operão (uma simples unidade de transcrição) e o complexo composto pela RNA polimerase e por diversas outras moléculas proteicas inclusivé um factor de iniciação (o factor sigma), liga-se a um local específico de bases no sítio promotor, podendo esta ligação efectuar-se com maior ou menor intensidade (força do promotor).

Este início da transcrição pode pois ser regulado por repressores ou activadores, estes últimos induzindo a formação de complexos abertos da cromatina levando à transcrição e os repressores inibindo a formação deste complexo aberto.

Como se compreende alterações na sequência nucleotídica ao nível do operador podem perturbar as interacções com os reguladores de tipo activador ou repressor, podendo inclusivamente não ocorrer a ligação com os repressores o que leva a uma transcrição continua, também chamada constitutiva.

Estas alterações das sequências operadoras ou promotoras podem controlar um gene que exista no mesmo cromossoma onde ocorre a alteração e daí dizer-se que são cis-actantes, podendo no entanto

esses elementos cis-activos (operador ou promotor) estarem próximos dos genes controlados como sucede nas bactérias, ou um tanto afastados como sucede nos eucariotas.

Sequências do DNA trans-ativadoras, são por outro lado sequências responsáveis pela síntese de proteínas difusíveis que se difundem nas células e podem controlar genes em diversos outros cromossomas ou no mesmo cromossoma.

Uma característica estrutural de muitas moléculas proteicas repressoras e activadoras nas células bacterianas é possuírem nos dímeros e tetrameros entretanto formados, hélices \_ que se unem no entalhe (groove) principal da molécula de DNA interagindo entre si.

Nas células eucariotas e comparativamente com as procariotas, a iniciação da transcrição é mais complexa, embora existam aspectos idênticos, sendo as RNA polimerases de três tipos, a I, II e III.

As funções destes três tipos de polimerases são um tanto diferentes pois a I é responsável pela formação dos pré-r RNA (que é depois processado em 28 S, 18 S e 5.8 S r RNAs).

A RNA polimerase II é responsável pela transcrição de todos os genes codificadores de proteínas e ainda de quatro outros genes codificando pequenos RNAs que intervêm no “splicing” dos precursores dos m RNA, e a polimerase III que biossintetiza 5sr RNA, t RNAs e ainda pequenos outros RNAs necessários ao splicing dos m RNAs.

### **RNA polimerase II**

Esta polimerase para os genes que exprimem proteínas, para iniciar a transcrição, liga-se indirectamente a promotores fortes que podem ser uma simples TATA box ou um outro local mais extenso situado no local de partida.

É esta polimerase adjuvada por vários factores de iniciação que podem actuar ligando-se primeiro às sequências promotoras e orientando a enzima RNA polimerase II para o sítio do início da transcrição.

Nesta TATA box ligam-se uma série de proteínas factores de iniciação a saber: primeiramente liga-se a proteína TF II D (composta por várias subunidades) através da subunidade TBP, seguindo-se depois o TF II A, TF II B e um complexo formado pelo TH II F, RNA polimerase II, TF II E, TF II H e TF II J formando-se um complexo com mais de 2.000kDa.

Este início da transcrição pode pois ser regulado, como já referimos anteriormente por proteínas activadoras ou repressoras que se ligam ao DNA num local distante do complexo de iniciação e hidrolizando ATP para formar o complexo de iniciação onde se possa iniciar a transcrição.

Estão assinaladas proteínas activadoras que se ligam ao DNA a dezenas e dezenas de milhares de pares de bases do complexo de iniciação, havendo no entanto outras proteínas activadoras que se ligam ao complexo de iniciação directamente.

Locais de ligação para os promotores activados dentro dos cerca de 200 pares de bases de um promotor, são elementos proximais do promotor, e se situados a maior distância, ou mesmo dentro de intrões ou do lado 3' da unidade de transcrição são chamados favorecedores (enhancers) e podem conter centenas de pares de bases.

Na maioria dos promotores nos eucariotas a transcrição pode ser regulada ao nível dos elementos proximais ou de favorecedores.

Na mesma molécula da proteína reguladora podem existir diversos módulos e domínios flexíveis que interactivam com o DNA ou com outros domínios de outras moléculas.

Nos eucariotas, os domínios dos reguladores activadores da transcrição revelam diversos tipos de estruturas proteicas como por exemplo uma hélice  $\pi$  que interactiva e encaixa no entalhe principal (groove) do DNA.

Estas proteínas reguladoras activadoras podem interactivar umas com as outras formando estruturas heterodiméricas o que aumenta o repertório de sequências do DNA que pode interactivar com estes domínios de activação destas proteínas reguladoras.

A primeira proteína a ligar-se ao conjunto do complexo de iniciação é a TF II D o que resulta da interacção de proteínas activadoras com factores associados TBP (TAFs) com o TBP, podendo também contactar o TBP e o TF II B directamente.

As proteínas reguladoras inibidoras ou repressoras, são também constituídas por módulos nos eucariotas e parecem desencadear a sua acção a partir de locais de ligação distantes. Aqueles módulos que as constituem contêm domínios distintos para ligação ao DNA e para a repressão.

Os diversos tipos de células num organismo multicelular exprimem os seus genes consoante o tipo de proteínas reguladoras activadoras e repressoras que contêm.

Assim há células que podem ter activadores apenas expressos nesse tipo de células enquanto outros podem ser expressos em variadíssimos tipos de células. É a história do desenvolvimento de cada célula que determina os factores de regulação da transcrição que são expressos.

Não pode também esquecer-se que no que se refere a reguladores da expressão dos genes, existem uma série de sinais extracelulares tais como as hormonas e factores de crescimento que também contribuem para esta regulação dos genes.

Como dissemos anteriormente nos eucariotas, para que possa ocorrer a transcrição do DNA, uma vez que este se encontra complexado na forma de cromatina, é necessária a intervenção de activadores, e não sendo assim a transcrição encontra-se reprimida.

Muitos genes que não são transcritos nos vertebrados, possuem grupos 5-metilo no resíduo de C (citosina) nos dinucleótidos CG na região de controlo da transcrição. Mas se estas regiões se encontram submetiladas e hipersensíveis à acção da Dnase a transcrição já é activa.

As alterações na estrutura e acessibilidade do DNA são essenciais para a transcrição e tudo aquilo que possa modificar as histonas, por exemplo por fosfatação e ubiquitinação, modificam a afinidade destas proteínas para o DNA e conseqüentemente a acessibilidade deste para a transcrição.

O C-5 da citosina pode ser metilado por metil-transferases específicas e condicionar assim, inactivando, a transcrição, ao interferir na ligação de factores activadores da transcrição com o DNA através de ligação ao entalhe (groove) principal deste. Cerca de 70% das sequências de C-G no DNA dos mamíferos estão metiladas.

Esta metilação do DNA não é contudo um regulador universal para todos os eucariotas superiores, sabendo-se no entanto que os genes mamíferos activos estão menos metilados que os genes activos.

Morfologicamente nos cromossomas os “puffs” correspondem a regiões de transcrição muito activas.

### **RNA polimerase I e III**

Também a transcrição desencadeada por estas polimerases necessita, antes da polimerase se ligar ao DNA, de factores de iniciação a interagirem com regiões de controlo do DNA:

Estas proteínas de iniciação interactivam próximo do local de início da transcrição com a subunidade TBP (que é comum para as polimerases I, II e III) que se liga à caixa TATA, caixa esta que não existe nos promotores do rRNA e tRNA.

Nas mitocôndrias o respectivo genoma é transcrito por uma RNA polimerase composta por duas cadeias polipeptídicas, curiosamente codificadas pelo núcleo da célula.

### **Proteínas que em combinação associada controlam a transcrição nos eucariotas.**

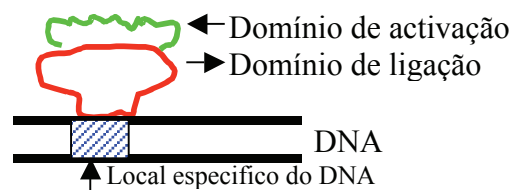
Como tivemos oportunidade de verificar anteriormente a maioria dos genes dos eucariotas não são expressos, salvo se forem especificamente activados por proteínas reguladoras adequadas. Para isso e como referimos anteriormente é necessário que a cromatina e sobretudo que os nucleossomas sejam removidos do sítio de partida para que a transcrição possa ocorrer.

Nos procariotas as RNA polimerases só por si transcrevem o DNA ao contrário do que sucede nos eucariotas (como vimos anteriormente) que necessitam de manter o complexo de transcrição, onde intervêm activadores e repressores.

A maioria dos genes eucariotas são regulados por diversas proteínas e não apenas por uma ou duas como sucede nos procariotas.

Cita-se o exemplo do gene da albumina do sangue do rato que contem seis locais para interacção com proteínas reguladoras próximo da TATA box, havendo ainda outros locais reguladores mais distantes.

Na estrutura dos reguladores de transcrição do tipo activadores, existem dois domínios muito importantes, um que interactiva com o DNA numa sequência específica deste, e outro domínio, o de activação que contribue para a formação do complexo de transcrição na TATA box.



Estas interacções podem estender-se da proteína que se liga ao DNA para outras proteínas, desempenhando as cargas dos grupos químicos de uma superfície o suporte responsável pela interacção com cargas de sinal contrário na superfície de outra proteína.

É mesmo possível que estas proteínas reguladoras que interactivam entre si e com o DNA, se encontrem as distâncias razoáveis, simplesmente a flexibilidade de que é dotada a molécula de DNA permite a sua aproximação espacial.

Existem diversas classes de proteínas reguladoras da transcrição nos eucariotas. Assim referem-se os “dedos “ de zinco (fingers) que se podem encontrar associados em tandem e que ligando-se a sequências estendidas de DNA podem regular assim a expressão dos genes aí contidos.

Por outro lado as hormonas esteróides também activam proteínas que se podem ligar ao DNA regulando desta forma a correspondente transcrição.

Também as proteínas leucina “zipper” controlam a transcrição nos eucariotas através de características estruturais peculiares como é o caso de possuírem uma \_ “coiled coil” helicoidal e um par de domínios que se ligam ao DNA. Estas proteínas leucina “zipper” apresentam um resíduo de leucina em cada sétima posição de uma extensão de cerca de 35 resíduos de ácidos aminados, formando estas proteínas, entre si, dímeros ligados através do domínio \_helicoidal”coiled coil”, dímeros que por seu turno se ligam a duas sequencias de DNA contíguas.

Está assinalada uma proteína leucina zipper a GCN4 que coordena a transcrição de mais de quarenta genes que exprimem enzimas para biossínteses.

Também é conhecido que as proteínas leucina “zipper” regulam o efeito do cAMP (AMP cíclico) sobre a transcrição.

Também outro tipo de proteínas reguladoras da transcrição do DNA nos eucariotas, são as codificadas pelas caixas homeo (uma sequência de 480-bp), caixas estas que se repetem nos genes que controlam o desenvolvimento dos insectos e dos invertebrados. Estas proteínas expressas pelas caixas homeo possuem 60 resíduos de ácidos aminados, são chamados homeodomínios, ligam-se ao DNA actuando também como reguladores da transcrição em todos os eucariotas.

### **3.2 - Importância da estrutura das proteínas, factores de transcrição nas interacções com o DNA**

#### **Factores de transcrição: Famílias estruturais e princípios no reconhecimento do DNA.**

Como é sabido ao DNA podem ligar-se variadíssimos tipos de proteínas, umas responsáveis pela replicação ou reparação do DNA enquanto outras intervêm na transcrição de genes activos.

Podemos afirmar que estas ultimas constituem um grande grupo que corresponde aos factores de transcrição que regulam a expressão dos genes, controlando o desenvolvimento celular, tal como a diferenciação e crescimento celular, interactivando especificamente com um ou mais locais do DNA do respectivo genoma.

Nos factores de transcrição há que considerar as grandes famílias onde se incluem:

- a) as proteínas HTH (hélice, ”turn”, hélice)
- b) proteínas homeodomínios

- c) proteínas "zinc finger"
- d) proteínas receptoras de esteróides
- e) proteínas leucina "zipper"
- f) proteínas HLH (hélice, "loop", hélice)
- g) outras

Nas proteínas HTH factores de transcrição, os domínios que se ligam ao DNA são uma pequena parte da proteína total.

As proteínas homeodomínios com 60 resíduos formam uma estrutura estável, dobrada que se pode ligar ao DNA, embora as proteínas homeodomínios possam conter um motivo HTH.

As proteínas "zinc finger" onde se inclui o factor de transcrição TF III, contêm repetições em tandem com motivos de 30 resíduos "zinc finger" cada motivo com a sequência em ácidos aminados Cys-X2 a 4-Cys-X12-His-X3-5-His. Estas proteínas ligam-se ao B-DNA no maior entalhe ("major groove") sendo de realçar as posições conservadas das cisteínas e histidinas.

As proteínas receptores de esteróides e reguladoras da transcrição são uma família muito ampla onde se incluem também os receptores para os retinóides, vitamina D, hormonas da tiróide e ainda outros compostos importantes, possuindo cada uma destas proteínas domínios distintos para ligação à hormona, para ligação ao DNA (com cerca de 70 resíduos com 8 resíduos conservados de cisteína) e para activação da transcrição.

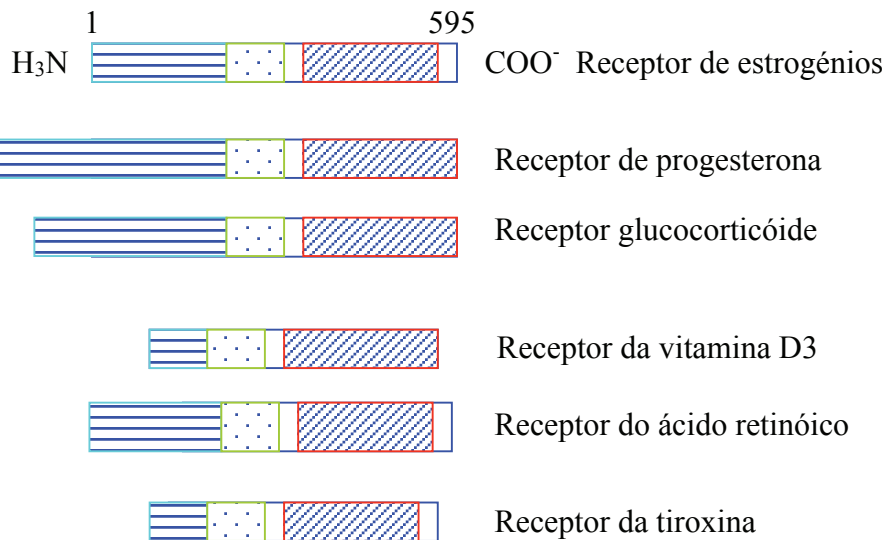
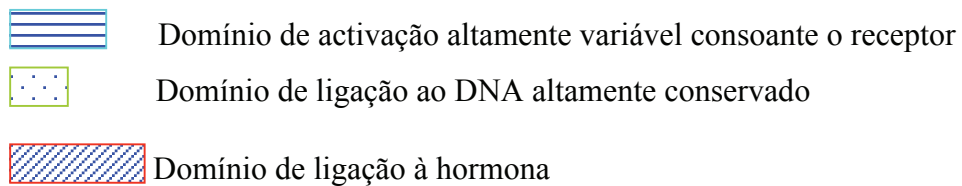
Estas hormonas esteróides lipossolúveis têm características próprias e dada a sua lipossolubilidade disseminam-se através da célula para atingirem o núcleo, interagindo consoante o tipo de hormonas, ou com receptores citoplásmicos ou com receptores nucleares, mas atingindo sempre o núcleo onde activam determinados factores de transcrição como é o caso da superfamília de receptores esteróides. Sem a presença destas hormonas os factores de transcrição estão ligados a inibidores, inibidores estes que se dissociam dos factores de transcrição quando estes interactivam com hormonas. É bem conhecido o caso do receptor glucocorticóide que é impedido de se ligar ao DNA pela presença de um inibidor, e também é referido o receptor para a hormona da tiróide que está ligado ao DNA em determinadas zonas, funcionando portanto como um inibidor, mas que quando interactiva com a hormona da tiróide passa a funcionar como um activador.

As hormonas esteróides e os chamados morfogénios interactivam pois com factores de transcrição regulando a transcrição do DNA, o estradiol, progesterona, testosterona e cortisol interagindo com receptores específicos no citosol indo depois este complexo regular ao nível do núcleo o funcionamento de determinados locais do DNA que podem abarcar a transcrição de 50 a 100 genes.

Como se compreende este modo de interacção hormona esteróide - receptor, ligação ao DNA, leva horas a decorrer desde o inicio da transdução do sinal até à sua expressão em síntese de proteínas.

Na figura seguinte esquematizam-se os diversos domínios de diversos tipos de receptores, sublinhando-se que todos eles possuem um domínio conservado de 66 resíduos de ácidos aminados para se ligarem ao DNA, um outro domínio também altamente conservado e com 240 resíduos de ácidos aminados onde se liga a respectiva hormona, e um ultimo domínio de activação, altamente variável consoante o tipo de receptor.

### Estrutura esquemática dos domínios da superfamília de receptores nucleares



O domínio do receptor que interacciona especificamente com a hormona só consegue interaccionar com o DNA se estiver ligado à hormona respectiva, o que é impedido de acontecer se o receptor não estiver complexado com a respectiva hormona.

O domínio amino terminal destes receptores, que é bastante variável de receptor para receptor, está vocacionado para interaccionar com outros reguladores da transcrição.

A título comparativo refere-se que quando os sinais desencadeadores são hormonas hidrossolúveis e como estas não parece que se difundam através da plasma membrana celular, elas interaccionam com receptores específicos superficiais, ocorrendo depois a transdução do sinal para o citoplasma, podendo mesmo esses receptores possuir actividades do tipo cinase.

Alguns exemplos deste tipo de hormonas são a epinefrina, a glucagina, insulina, factor epidérmico de crescimento, interferão, etc.

Assim a epinefrina interaccionando com um receptor de sete hélices contido na plasma membrana activa a adenilato ciclase, desencadeando várias respostas intracelulares, tal como sucede com a glucagina, a insulina e o factor epidérmico do crescimento interaccionando com receptores com actividade tirosina cinase intrínseca e o interferão  $\gamma$  com receptores com actividade tirosina cinase citosólica, todos eles desencadeando respostas intracelulares.

No caso do interferão  $\alpha$  há uma via de transdução de sinal em que o receptor da superfície celular activa uma proteína cinase que por fosfatação activa depois um factor de transcrição no citoplasma o qual se dissemina para o núcleo onde activa a transcrição de diversos genes.

No caso das hormonas esterólicas e outros morfogénios (como a tiroxina e o ácido retinóico induzindo diferenciação celular e controlo desta) os receptores activados vão ligar-se a determinadas sequências do DNA, sequências estas específicas denominadas elementos de resposta às hormonas (HRE) que controlam por sua vez a transcrição de uma série de genes próximos. Todos estes elementos (HREs) de resposta às hormonas esteróides possuem características estruturais peculiares pois são palíndromas com um par de sequências de seis pares de bases separadas por um espaçador de três pares de bases (vide esquema seguinte)

“Elemento” do DNA de resposta à hormona glucocorticóide



“Elemento” do DNA de resposta à hormona estrogénio

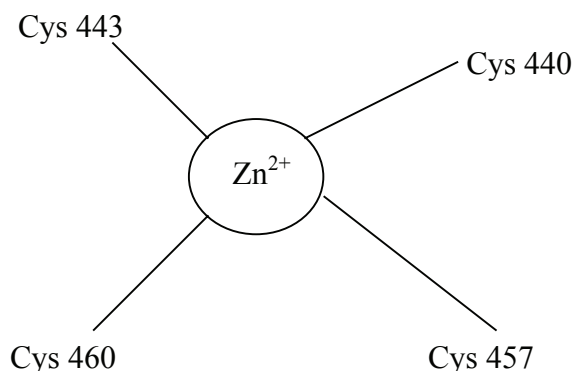


São os eixos de simetria (bola azul) que relacionam as sequências da 6-bp.

Como se verifica o elemento de resposta (HRE) no DNA para os glucocorticóides é diferente do que corresponde aos estrogénios. Com efeito nas posições +5 e +6 do palíndromo existe TT no HRE para os glucocorticóides e AC no HRE para os esteróides.

Estas diferenças nos HREs no DNA são distinguidas pelos receptores dos glucocorticóides e dos esteróides com as consequências biológicas respectivas.

O receptor glucocorticóide para se ligar ao HRE no DNA, encontra-se na forma de um dímero, sendo a cadeia polipeptídica deste domínio de ligação estabilizada por um par de “clusters” ou dedos de zinco (zinc clusters) em que um ião de zinco ( $Zn^{2+}$ ) está coordenado tetrahedricamente com cadeias laterais de cisteínas.





O domínio dimeriza-se quando a cadeia se liga ao elemento de resposta palíndromo.

Cada sub-unidade proteica contém dois “Zinc clusters”. Contudo as estruturas secundárias e terciárias destes complexos receptor-hormona que se ligam ao DNA são diferentes das do dedo de zinco pois aquele domínio é mais globular do que alongado. sucedendo ainda que as duas subunidades do receptor esteróide reconhecem a mesma sequência de DNA, e os dedos de zinco adjuvantes que actuam independentemente reconhecem diferentes sequências do DNA.

Ainda dentro das proteínas factores de transcrição há que referir as proteínas leucina zipper e as proteínas HLH (helice, loop, hélice) condicionadoras da diferenciação e desenvolvimento celular.

As proteínas com motivos leucina zipper existem numa série de factores de transcrição animais, tendo uma sequência conservada cujos domínios de ligação ao DNA contêm 60-80 resíduos de ácidos aminados, com dois subdomínios distintos sendo uma região responsável pela dimerização e outra uma região básica, a qual interacciona com o DNA.

São caracterizados estes domínios leucina zipper por possuírem uma sequência repetida de sete leucinas numa região compreendendo 30-40 resíduos, região esta que forma duas hélices \_ paralelas. Estas proteínas podem formar homo ou heterodímeros.

Os factores de transcrição do tipo proteínas HLH são parecidos com as proteínas leucina zipper visto terem também uma região básica a contactar com o DNA e depois uma outra região vizinha que é responsável pela formação de dímeros, sendo também estas proteínas muito importantes para a diferenciação e desenvolvimento celular.

Por exemplo a proteína Myo D fundamental na diferenciação das células musculares, forma um heterodímero com a proteína E2A ligando-se ao núcleo. Esta última proteína contém uma região HLH.

As interacções destas proteínas factores de transcrição com as moléculas de DNA fazem-se na maioria destas famílias, sobretudo através de hélices \_ que interaccionam com o entalhe principal do DNA (major groove), e também através de folhas \_ ou regiões de cadeias polipeptídicas distendidas. As interacções entre as proteínas e o DNA formam-se através de pontes de hidrogénio e entre as cadeias laterais dos ácidos aminados da proteína e as bases do ácido nucleico, havendo também contactos hidrofóbos, pontes de hidrogénio ocasionais mediadas por moléculas de água, etc, sendo também os oxigénios das ligações fosfodiéster do DNA, e as estruturas cíclicas da desoxirribose envolvidas em muitos destes contactos proteínas factores de transcrição molecular de DNA.

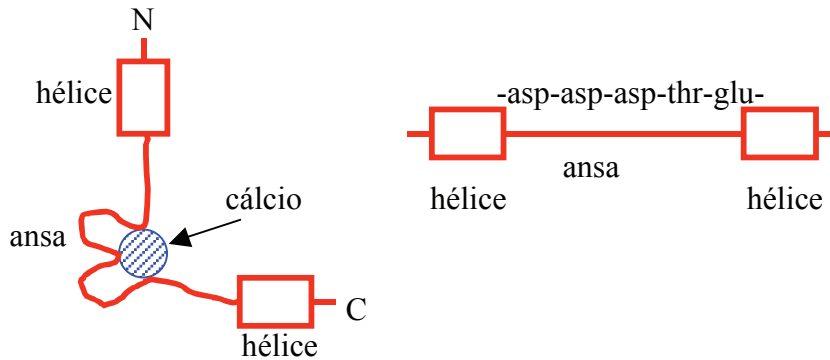
### **As estruturas secundárias são elementos importantíssimos da arquitectura das proteínas**

A maior parte das cadeias polipeptídicas que constituem as proteínas existem na forma de hélices \_ ou de folhas \_ (cerca de 60%). A porção restante da cadeia polipeptídica assume dobras ao acaso ou ansas (random coils) e voltas ou inversões (turns).

Parte-se do princípio que o leitor está identificado com as características das hélices \_ e das folhas \_ realçando-se no entanto que as voltas ou inversões (turns) são constituídas por um pequeno número de ácidos aminados, três ou quatro, adoptando formas em U compactas, que se encontram à superfície das proteínas e que estabelecem contactos entre o primeiro e o último ácido aminado da volta, através de pontes de hidrogénio. Estas voltas formam um ângulo na cadeia polipeptídica condicionando assim a sua reorientação.

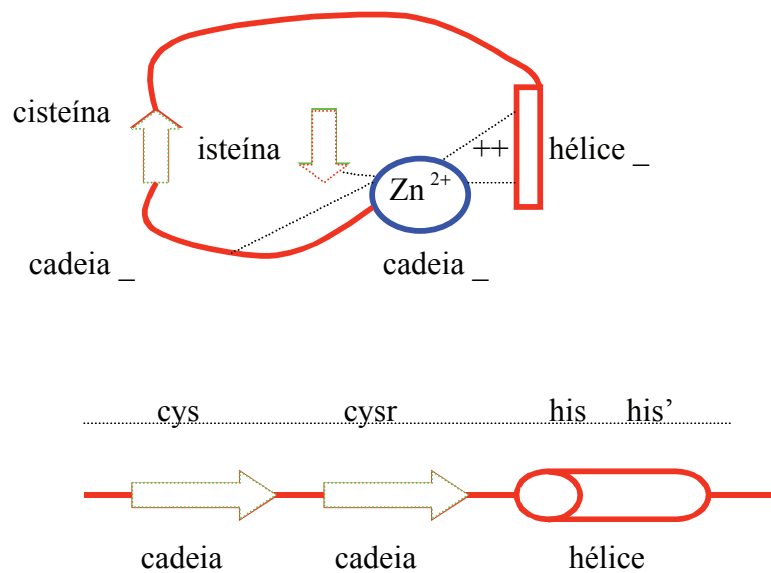
Outras partes das cadeias polipeptídicas que constituem as proteínas podem assumir curvas mais ou menos largas ou ansas (loops) dando variadíssimas geometrias às moléculas em que se integram.

Ainda do conhecimento dos estudos de bioquímica, recorda-se que motivos são combinações regulares de diversas estruturas secundárias, tendo um dado motivo uma sequência característica e uma actividade biológica característica. Servimo-nos a seguir de um exemplo com o motivo hélice-ansa-hélice que existe em bastantes proteínas podendo complexar iões cálcio



As cadeias laterais dos ácidos aminados que constituem a ansa e que contêm oxigénio estabelecem interações com o ião cálcio. Com efeito esta ansa contém cerca de 14 ácidos aminados sobretudo hidrofílos que têm alta afinidade para o cálcio.

Ainda no que se refere a motivos nas proteínas voltamos ao motivo em dedo de zinco (zinc fingers) dada a sua forma peculiar que complexa zinco e que faz parte de diversas proteínas com capacidade para interagir com os ácidos nucleicos. Estas proteínas contêm um par de cadeias \_ antiparalelas e uma hélice \_ simples que complexam um ião zinco por interacção com um par de resíduos de cisteínas e com duas histidinas tal como se esquematiza na figura seguinte:



Como se vê em baixo na figura o dedo é formado de três estruturas secundárias, duas cadeias \_ antiparalelas e uma hélice \_ e as cadeias laterais de dois resíduos de histidina e de dois resíduos de cisteína situados nas duas pontas do “dedo, formando uma caixa dentro da ansa que liga os cordões \_ e a base da hélice \_ (ver na parte de cima da figura) complexando o ião zinco.

Este complexo proteico dede com zinco interactua depois com o DNA no entalhe maior (groove) podendo também motivos desta natureza encontrarem-se em proteínas que interactuam com o RNA:

Ainda no que se refere à estrutura das grandes proteínas e quanto aquilo que se denomina como estruturas terciárias estas são por vezes subdivididas em domínios. Domínios são por exemplo as regiões globulares ou as regiões fibrosas e com actividades características, de uma dada proteína. Um domínio é pois uma unidade estrutural ou funcional de uma proteína contendo 100 a 300 residuos de ácidos aminados em diversas combinações de hélices \_\_, folhas \_\_, voltas ou ansas.

Diversas proteínas podem conter o mesmo tipo de domínio, podendo algumas proteínas conterem o mesmo tipo de domínio repetido várias vezes, dentro de uma mesma molécula proteica, ou então existirem diversos tipos de domínios dentro da mesma molécula proteica. Nos seres humanos o maior número de domínios identificados numa proteína é de nove, mas o número total de domínios numa proteína pode atingir 130.

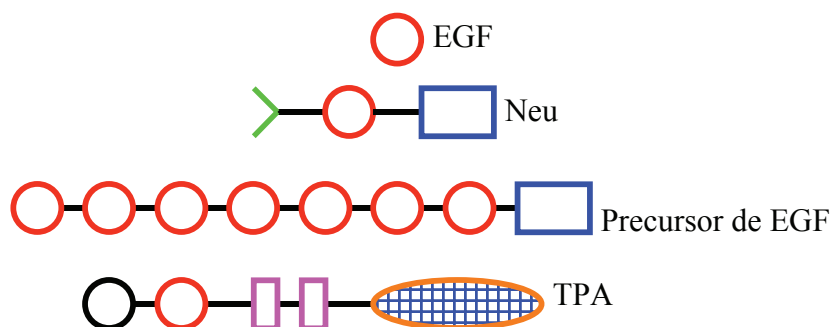
Os domínios muitas vezes são caracterizados por aspectos estruturais peculiares, como por exemplo um número muito elevado de um mesmo tipo de ácido aminado (por exemplo domínio rico em prolina, ou glicina, etc.) ou noutro caso ocorrerem sequências idênticas de ácidos aminados em diversos tipos de proteínas (como por exemplo SH3 ou região de homologia Src3), ou ainda assumir dentro do domínio um motivo com estrutura secundária particular (por exemplo domínio “kringle”, motivo dede de zinco, etc.).

Os domínios são melhor definidos através das suas características estruturais do que através da suas características funcionais.

Muitas vezes um domínio funcional é constituído por diversos domínios estruturais.

Os domínios estruturais e funcionais são módulos da estrutura terciária das proteínas em que estão integrados.

Diferentes tipos de proteínas podem assim incorporar diversos tipos e módulos. Pode afirmar-se que o desenho em módulos de uma dada proteína é que lhe imprime as diversas funções que pode desempenhar ao mesmo tempo, consoante portanto o tipo de domínio contido na sua arquitectura.



Exemplifica-se seguidamente, com um domínio que faz parte de diversos tipos de proteínas. Este domínio é o factor do crescimento epidérmico (EGF).

O módulo EGF **O** é uma hormona peptídica solúvel que se encontra pois integrando diversas proteínas, podendo coexistir com outro tipo de módulos, mas quando é libertado da proteína que o contém desempenha a sua actividade biológica.

Assim esse módulo EGF constitui uma hormona peptídica solúvel com capacidade para interagir com as células da epiderme e tecido conectivo promovendo a respectiva divisão celular, e essa funcionalidade é operante desde que seja libertado de qualquer das proteínas que o integram como módulo.

Na figura anterior a proteína precursora de EGF que se encontra ligada à membrana celular através de um domínio hidrófobo que penetra nesta, por proteólise liberta a hormona péptido solúvel EGF.

No entanto as duas outras proteínas referidas na figura em apreço também contêm o módulo EGF, juntamente com outros módulos. É o caso do activador tissular do plasminogénio (TPA) da proteína Neu implicada na diferenciação embrionária e da proteína Notch que facilita a adesão intercelular.

As sequências homólogas (homologia) entre diversas proteínas sugerem que há inter-relações funcionais e evolutivas entre essas proteínas.

Dois genes que derivam da duplicação de um gene são designados parálogos e dois genes (em duas espécies) são ortólogos se eles derivam do mesmo gene através de “speciation”.

A previsão de quando duas proteínas são parálogos é relativamente simples quando a identidade da sequência é alta (>40% para sequências longas) mas torna-se difícil quando essa identidade da sequência é média (20-35%) ou menor, sobretudo para sequências curtas.

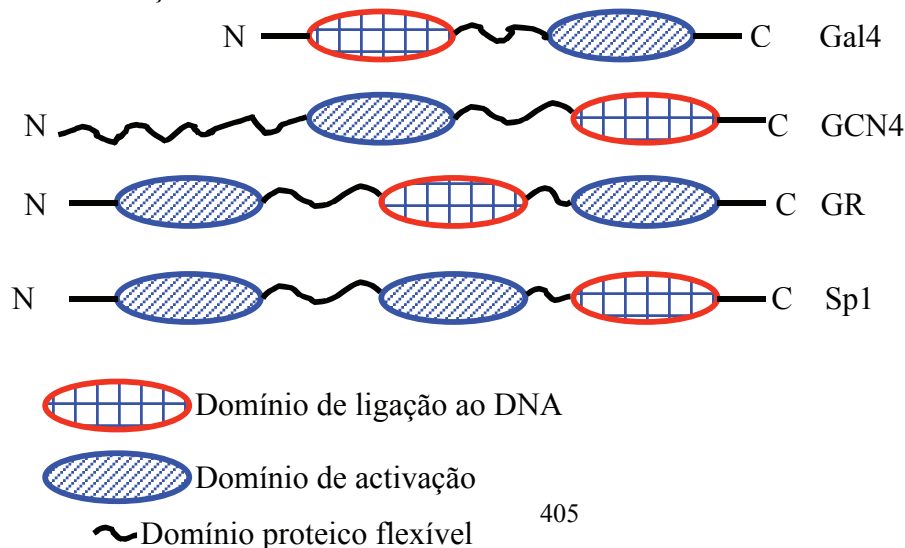
É necessário definir melhores métodos para decidir quando duas proteínas são homólogas sobretudo para proteínas curtas.

### Factores de transcrição eucariotas

Dado o interesse deste tema em Medicina Veterinária justifica-se que o desenvolvamos um pouco mais relativamente aquilo que referimos anteriormente.

Vários dos factores de transcrição dos eucariotas são compostos por módulos, módulos estes compostos por domínios funcionais, com localizações separadas. Assim haverá um domínio de ligação ao DNA que interacciona especificamente com este e outro domínio de activação que interaccionará com outra proteína para estimular a transcrição de um promotor próximo.

Na figura seguinte representa-se a estrutura em módulos (modular) de diversos activadores eucariotas de transcrição .



Curiosamente os vários factores de transcrição desta figura contêm apenas um domínio para ligação ao DNA mas podem conter diversos domínios de activação, podendo mesmo suceder que porções de um domínio de ligação ao DNA participem também na activação da transcrição.

Há que considerar de entre os factores de transcrição os chamados factores gerais de transcrição, ao lado dos factores de transcrição específicos para determinados tipos de genes. Nos primeiros englobam-se os TF II D, TF II A, TF II B, TF IIF; TF II E, TF II H, TF II J que são indispensáveis para o início da transcrição de todos os genes.

Salienta-se ainda que os domínios de activação estabelecem a interacção de proteínas com outras proteínas inclusivé com factores de transcrição ligados ao promotor.

Por outro lado os domínios que se ligam ao DNA, dos factores de transcrição têm uma grande diversidade de motivos ou estruturas proteicas.

De qualquer forma está identificado que as hélices \_ destes domínios se dispõem de forma que lhes permite interactuar com o entalhe maior (major groove) do DNA através de ligações químicas fracas (pontes de hidrogénio e forças de van der Waals).

Encontra-se estabelecida uma classificação dos factores de transcrição segundo o tipo de domínio de ligação ao DNA que evidenciam, tendo a maioria destes domínios sequencias em ácidos aminados consensuais e característicos.

As classes mais comuns de domínios de ligação ao DNA encontram-se seguidamente referidas:

- a) Homeodomínio
- b) Domínio C2H2 zinc-finger
- c) Domínio C4 zinc-finger
- d) Domínio C6 zinc-finger
- e) Dominio winged-helix (forkhead)
- f) Domínio leucina-zipper
- g) Basic zipper
- h) Domínio hélice – ansa – hélice (HLH)

Está calculado que os genomas dos eucariotas superiores codifiquem dezenas de classes de domínios para ligação ao DNA e centenas de factores de transcrição.

Desenvolvemos seguidamente algumas características das classes mais comuns de domínios de ligação ao DNA.

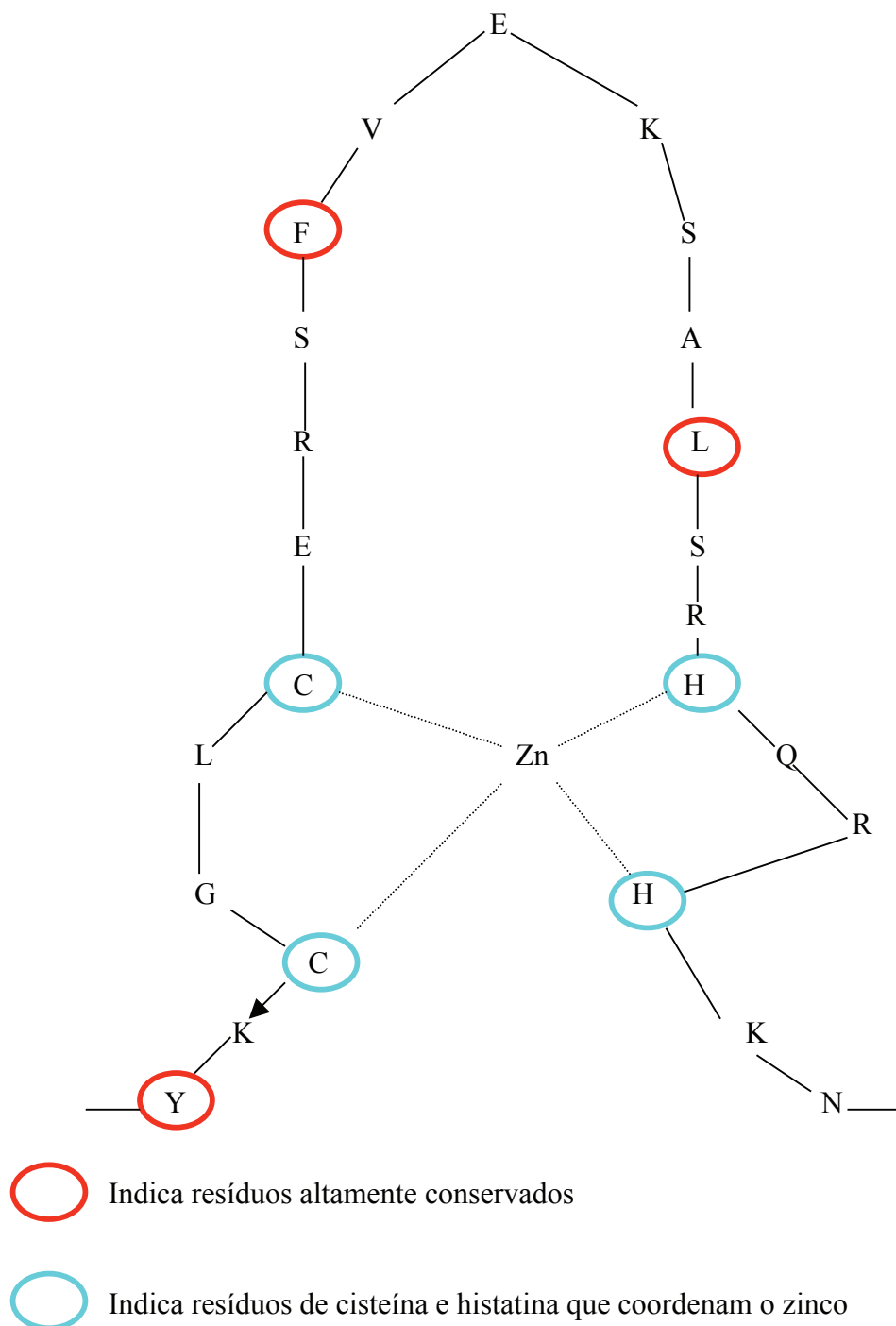
a) As caixas homeo (uma sequência de 180 bp) repete-se nos genes que controlam o desenvolvimento dos insectos e dos invertebrados, codificando proteínas denominadas homeodomínios, com sessenta ácidos aminados que parecem ligar-se ao DNA, servindo como reguladores da expressão dos genes em todos os eucariotas.

b), c) e d) Estas três classes de proteínas em dedo de zinco encontram-se em proteínas que se ligam ao DNA e também em proteínas que não se ligam ao DNA:

Todas elas possuem regiões com um domínio compacto que contem um ião zinco central coordenado.

Estas proteínas em dedo de zinco que se ligam ao DNA contêm motivos repetidos constituídos pelos complexos formados pelo átomo de zinco com as cadeias laterais de duas cisteínas e duas histidinas como se refere na figura seguinte

Estrutura primária de proteína “dedo” de zinco



Salienta-se que por vezes as histidinas são substituídas por cisteínas e o motivo que contém as duas cisteínas próximas que se ligam ao átomo de zinco, estão afastadas cerca de 12 ácidos aminados do segundo par de ácidos aminados próximos (histidinas ou cisteínas) que se ligam também ao átomo de zinco, contendo este motivo da proteína um segmento em hélice \_ que interacciona com o entalhe maior do DNA alvo, e enrola-se à volta da dupla hélice do DNA.

A classe I (motivo C2H2) destas proteínas em dedo de zinco que existe no factor de transcrição III A, implicado na transcrição pela polimerase III dos genes 5sr RNA, possui nos motivos repetitivos a estrutura primária consenso seguinte (Tyr/Phe) X1Cys X2-4 Cys3(Phe/Tyr) X5 Leu X2 H15 X3-4 His representando X qualquer ácido aminado.

Na maior parte dos factores de transcrição eucariotas é este tipo de domínio CH2 ou classe I que estabelece ligações com o DNA:

A classe II ou C4 destas proteínas em dedo de zinco encontra-se numa superfamília com mais de 100 factores de transcrição de receptores de hormonas esteróides. Os domínios de ligação ao DNA desta classe II das proteínas em dedo de zinco revela sequências consenso, nesses domínios de ligação ao DNA, com as seguintes características Cys X2 Cys X13 Cys X2 Cys X14-15 Cys X3 Cys X9 Cys X2 Cys. Como têm dois grupos de quatro cisteínas a ligar-se ao ião zinco são designadas proteínas C4 (ou classe 2) em dedo de zinco.

As proteínas de dedo de zinco da classe I ou C2 H2 ligam-se ao DNA na forma de monómeros enquanto as da classe II (domínio C4) ligam-se ao DNA como homodímeros de diferentes proteínas em dedo de zinco C4.

A classe III (domínio C6) de proteínas de dedo de zinco têm a seguinte sequência consenso Cys X2 Cys X6 Cys X5-6 Cys X2 Cys X6 Cys e denomina-se C6 porque tem seis resíduos de cisteína ligados ao ião zinco.

e) Outra classe de factores de transcrição eucariotas são as proteínas em hélice alada (cabeça de forquilha, "winged-helix ou forkhead protein) que existem em diversos factores de transcrição e também nas histonas H5, ligando-se esse domínio ao DNA como monómeros.

f e g) Nova classe de factores de transcrição são as proteínas leucina zipper que apresentam na sua posição C terminal que se liga ao DNA, a leucina em cada sétima posição, sendo esta leucina que está na base da formação de dímeros que se ligam ao DNA, contendo estas proteínas duas hélices \_ estendidas que interaccionam com a molécula de DNA colocada transversalmente entre as duas hélices.

Conforme se pode ver na figura seguinte estas proteínas leucina zipper são heterodímeros, cada uma das cadeias peptídicas contendo uma região zipper básica e depois sequências com a leucina na sétima posição.

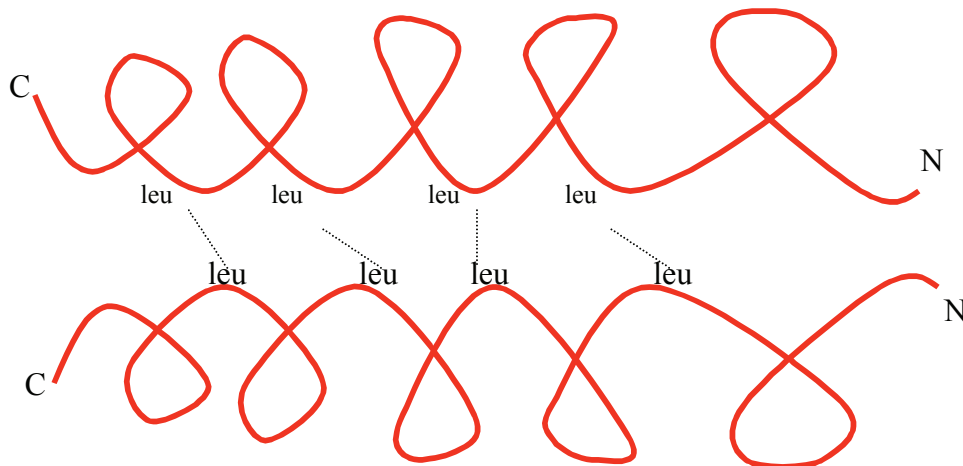
Estes motivos leucina zipper (b ZIP) têm sede em regiões da proteína em hélice \_, regiões estas que contêm pelo menos quatro resíduos de leucina, cada leucina separada da leucina seguinte por seis ácidos aminados (Leu-X6- Leu-X6-Leu- correspondendo X a qualquer tipo de ácido aminado).

Na figura seguinte cada uma das hélices \_ contendo 3,6 ácidos aminados em cada volta apresenta a leucina na margem da hélice com uma leucina aparecendo em cada segunda volta de cada hélice. As duas cadeias polipeptídicas distintas com estas características permitem-lhe estabelecer contactos hidrofóbicos entre si, formando aquilo que se chama um domínio ou "zipper" das duas subunidades sendo

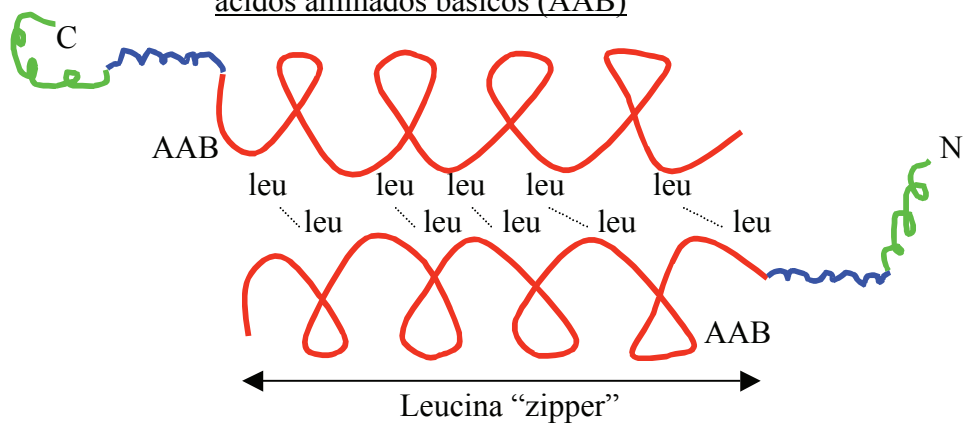
os contactos a fazer com o DNA estabelecidos com zonas adjacentes das hélices \_ dos motivos zipper, zonas estas que são muito ricas em ácidos aminados básicos (AAB) sobretudo arginina e lisina que interactuam electrostaticamente com as cargas negativas dos fosfatos das ligações fosfodiéster do DNA.

Estes motivos leucina zipper encontram-se também em diversas proteínas reguladoras oncogénicas (como é o caso das Myo, Jun e Fos) na forma de heterodímeros. Como elas se ligam em locais de regulação de genes, se sofrem mutações ou são produzidas de forma irregular podem desencadear a formação de células cancerosas.

Duas cadeias polipeptídicas distintas com hélices \_ contendo Leu em cada segunda volta da hélice (leucina “zipper” ou bZIP)

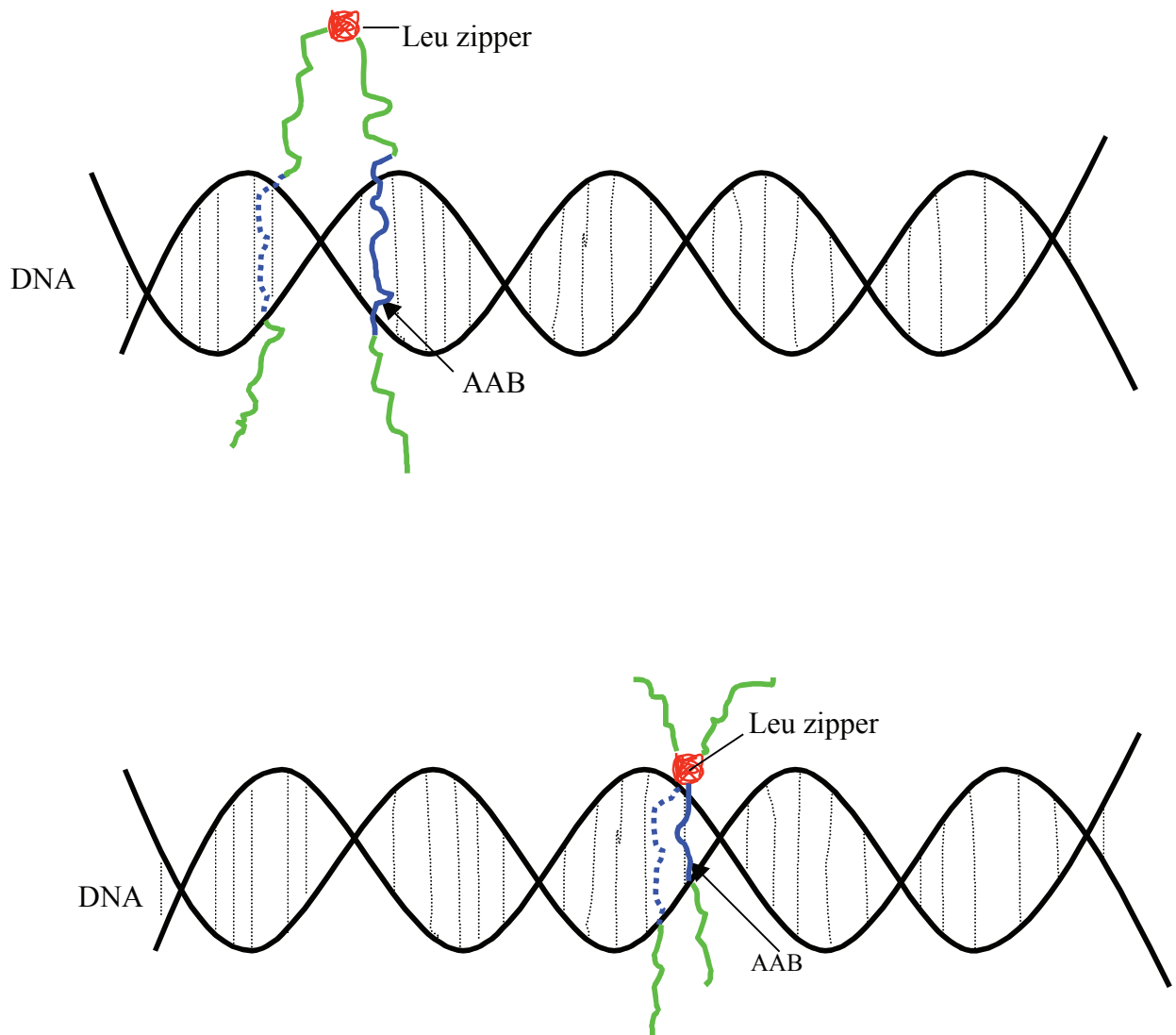


Interacção da leucina “zipper” num dímero e zonas adjacentes contendo ácidos aminados básicos (AAB)





### Complexo de dois bZIP diferentes interagindo com DNA



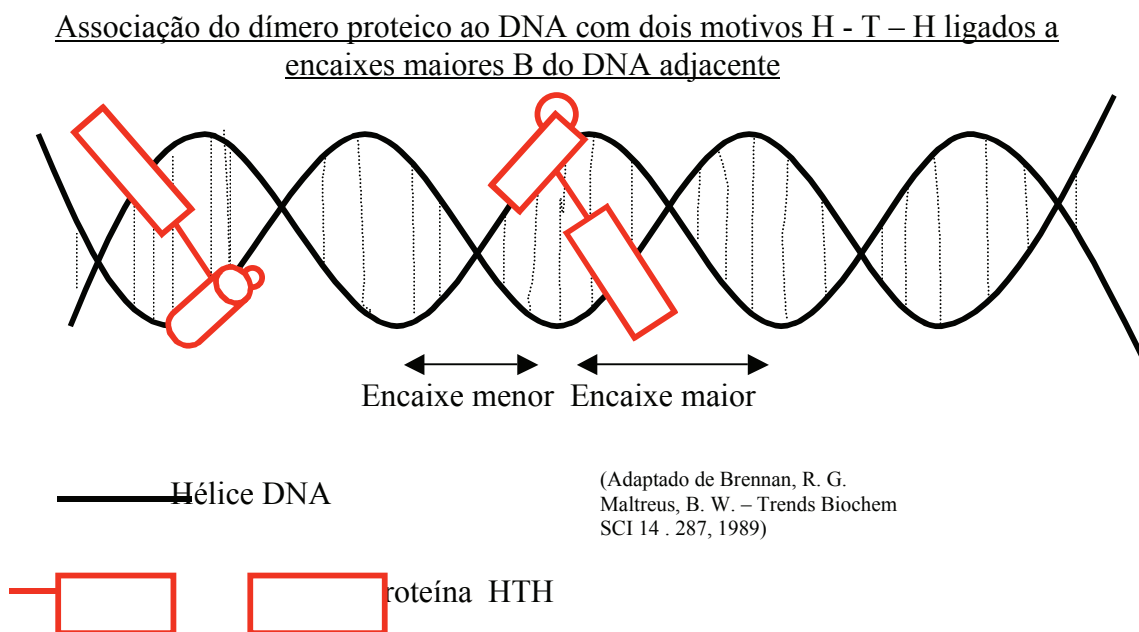
i) Esta classe de domínios HLH encontra-se em factores de transcrição na forma de dímeros que se ligam ao DNA de uma forma muito análoga aquela que o domínio básico dos b-ZIP faziam. Contudo as características estruturais destes factores de transcrição são diferentes pois cada monómero do dímero é constituído por um motivo que tem duas regiões \_ helicoidais ou seja o motivo HLH (hélice-ansa-hélice), motivo este que tem portanto uma hélice \_ N-terminal com resíduos de ácidos aminados básicos, no meio uma porção em ansa, e depois uma região C-terminal com ácidos aminados hidrofóbicos dispostos com intervalos certos (vide figura da pagina 13).

Todas estas proteínas factores de transcrição como as C4 em dedo de zinco, proteínas básicas zipper e proteínas HLH, assim como outras, formam heterodímeros para se ligarem ao DNA o que aumenta consideravelmente o repertório de interacções possíveis com o DNA.

Recorda-se que as moléculas de DNA na forma B possuem ao longo da sua estrutura helicoidal duas depressões, entalhes, ou encaixes um maior (major groove) e outro menor (minor groove) que

constituem locais que podem interactuar com as proteínas anteriormente referidas e a que de resto já fizemos algumas alusões a este propósito.

Outro motivo que existe em outras proteínas que podem interactuar com o DNA, são os motivos HTH (hélice, volta, hélice), proteínas que também se oligomerizam na forma de dímeros. Quando o motivo HTH interactua com o DNA fá-lo através de uma das suas hélices \_ (hélice de reconhecimento) que interactua com o entalhe maior do DNA através de cadeias laterais dos ácidos aminados desta hélice que se ligam com os fosfatos do DNA alvo provocando a modificação da conformação do DNA-B nesta região.



Estes dois motivos ligam-se a duas voltas adjacentes de “encaixe” maior no DNA, formando uma forte interacção entre o DNA e a proteína.

### **3.3 - Factores gerais de transcrição e RNA polimerase II, I e III**

Como já tivemos oportunidade de referir a complexa regulação da expressão dos genes ocorre sobretudo no início da transcrição.

Nos eucariotas para a biossíntese do RNA mensageiro se iniciar torna-se necessária a participação de diversos elementos ou classes funcionais da sequência de DNA a transcrever.

Assim há que considerar o elemento promotor âmago (core promoter element onde irá interactuar a RNA polimerase II) e que controla o local onde se inicia a transcrição.

Alem deste elemento promotor âmago, existem os elementos promotores acima (upstream) e os favorecedores (enhancers) que influenciam o ritmo a que a RNA polimerase II recomeça novas rondas de transcrição.

Este conjunto de elementos com sede no DNA dirigem duas classes de factores de transcrição ou sejam os factores gerais para iniciação e os factores reguladores. Os primeiros, os factores gerais, são

suficientes para comandar a iniciação e um nível básico de transcrição a partir de uma grande parte de todos os promotores âmagos, os segundos não são precisos para a iniciação mas são intermediários para a interacção dos elementos promotores acima com os favorecedores.

### **Factores de transcrição**

A expressão dos genes é um mecanismo complexo que ocorre primeiramente ao nível do núcleo com a transcrição do DNA e depois com a translação do RNA mensageiro ao nível do citoplasma.

Ao nível do núcleo torna-se necessário a biossíntese ou transcrição do RNA pré-mensageiro, seguindo o seu processamento (splicing) e a formação da extremidade 3'.

Nos eucariotas os genes codificadores de proteínas, com a intervenção da RNA polimerase II necessitam de dois tipos de factores para operarem a transcrição: o primeiro tipo são os factores gerais ou básicos da transcrição (GTF) e o segundo tipo são os factores activadores da transcrição. Os GTF englobam a própria RNA polimerase II e pelo menos seis GTFs que são os TFIID, TFIIA, TFIIB, TFIIE, TFIIF e TFIIH.

A interacção do conjunto de GTFs sobre o promotor do DNA leva à formação do chamado complexo de pré-iniciação (PIC), começando com o TFIID a ligar-se à TATA box. Neste factor TFIID encontra-se uma proteína que se liga à TATA box chamada TBP e ainda vários outros factores TBP associados designados como TAFs.

Este dispositivo básico para a transcrição encontra-se numa grande diversidade de seres vivos desde leveduras até aos seres humanos, e nestes últimos parece que o respectivo genoma apenas contém uma cópia para os componentes da RNA polimerase II, e das TFIIB, TFIIE, TFIIF, e TFIIH, admitindo-se que talvez no gene para este último factor TFIIH possam existir mais dois outros genes relacionados com uma ciclina dependente de cinase (Cdk 7).

Contudo estão assinaladas sequências relacionadas nos genes que exprimem várias subunidades TFIID, inclusive TBP, TFIIA e vários TAFs.

Os factores de transcrição activadores (o segundo tipo antes referido) são proteínas específicas para determinados locais do DNA ou sejam os promotores alvo, onde é fortemente activada a transcrição.

Existe um grande número de factores de transcrição activadores, sendo estes classificados como vimos anteriormente em função da natureza dos seus domínios que se ligam ao DNA, estando assinalados mais de 2.000 genes seus codificadores.

Já referimos anteriormente dentro destas proteínas as C2H2 em dedo de zinco que parece ser a maior de todas estas famílias proteínas activadoras, com cerca de 900 membros.

Também já nos referimos dentro destas proteínas activadoras da transcrição às proteínas leucina “zipper” (bZIP), às proteínas H-L-H, aos receptores nucleares para hormonas esterólicas e outros morfogénios, etc.

O processamento do pré RNA mensageiro opera-se no spliceosoma, resultante da assemblagem de quatro pequenas ribonucleoproteínas nucleares (snRNP) as U1, U2, U5 e U4/U6 com outras proteínas não snRNP. Estas interacções decorrem segundo uma ordem definida, assim nos metazoários tudo começa

pela identificação do local de cisão 5' pelo U1 snRNP e do tracto polipirimidina (Py) reconhecido pelo U2 snRNP, U2AF.

A terminação dos mRNA eucariotas consta de uma cauda 3' poliadenilada (A) com cerca de 200 nucleótidos, sendo esta cauda acoplada, depois de uma cisão endonucleotídica do pré-mRNA e após a transcrição, sendo esta poliadenilatação conduzida por uma sequência AAUAAA situada 10-30 bases acima do local da poliadenilatação. Para este efeito são necessárias várias proteínas inclusive um factor específico (CPSF) para a cisão antes referida, um factor estimulador desta cisão (CstF) bem como outros factores auxiliares desta cisão, além de uma polimerase de adenilatos (PAP). A cauda 3' poli (A) madura liga-se depois com diversas proteínas chamadas PABs, sendo conhecidos diversos genes implicados na formação dos PAPs e dos PABs.

### **Transcrição pela RNA polimerase II**

Na transcrição de todos os genes pela RNA polimerase II ocorrem três facetas idênticas. Assim:

1) A regulação da transcrição pode ser feita por uma série de elementos ou sejam sequências do DNA distintas, além do respectivo promotor, intervindo portanto diversas sequências do DNA neste contexto que funcionam de uma forma articulada permitindo regular de forma muito eficiente a transcrição em causa.

2) São diversas moléculas proteicas ou sejam factores de transcrição, que interagindo com determinadas sequências nucleotídicas do DNA pelas quais são reconhecidas, que condicionam esta regulação da transcrição.

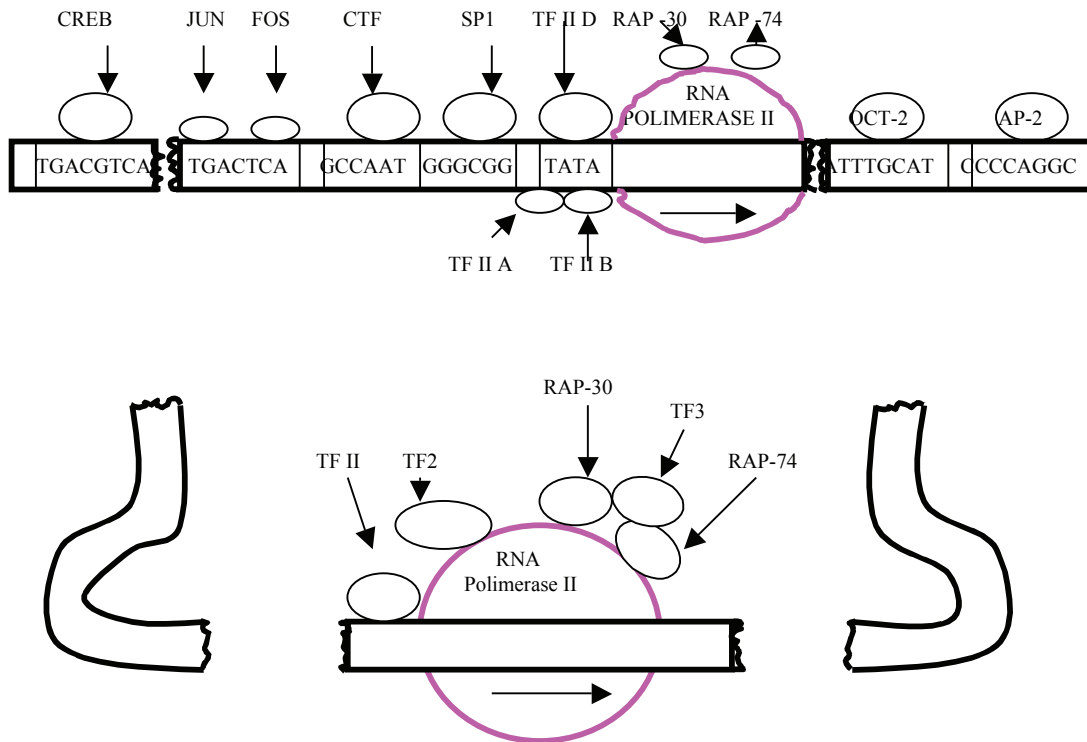
3) Por outro lado estas moléculas proteicas factores de transcrição interagem entre si e com a RNA polimerase II em ordem a activar a transcrição, o que implica que estas moléculas proteicas estejam dotadas de domínios para interagir com sequências determinadas do DNA e possuam também outros domínios que lhes permitem interagir com outros factores de transcrição.

Nos procariotas a RNA polimerase actua de forma diferente pois apenas interage uma simples sequência promotora.

Nos eucariotas a RNA polimerase II reconhece indirectamente, como veremos na figura seguinte, uma série de sequências nucleotídicas consenso acima do sítio em que se inicia a síntese do mRNA. Assim a primeira destas sequências é a caixa TATA seguindo-se a caixa CAAT (esta sequência CAAT podendo em alguns promotores não existir, ou possuir uma sequência pouco conservada).

Estas duas caixas como se indica na figura não contactam directamente com a RNA polimerase II, mas sim através das diversas proteínas factores de transcrição específicos indicados na figura, que interagem com os promotores.

## Interacção de factores de transcrição com promotores



Grande número de factores de transcrição interactuam com os promotores.

Em cima um conjunto hipotético de factores interactuam com sequências de DNA específicas próximas do promotor. Este induz um factor TF II D que se liga a caixa TATA e as proteínas Jun e Fos que são proto-oncogenes.

Em baixo faz-se um arranjo hipotético no qual vários factores de ligação ao DNA se ligam uns aos outros e à RNA polimerase II.

### RNA polimerase II e o complexo de início da transcrição

Salienta-se que a maior parte destes factores gerais de transcrição utilizados para a transcrição de todos os genes são compostos por várias subunidades proteicas. Por exemplo o factor de transcrição TFIID tem uma subunidade TBP que se liga à caixa TATA e mais oito subunidades as TAFs.

O domínio C-terminal destas subunidades TBP nos eucariotas é muito análogo, contendo cerca de 180 ácidos aminados.

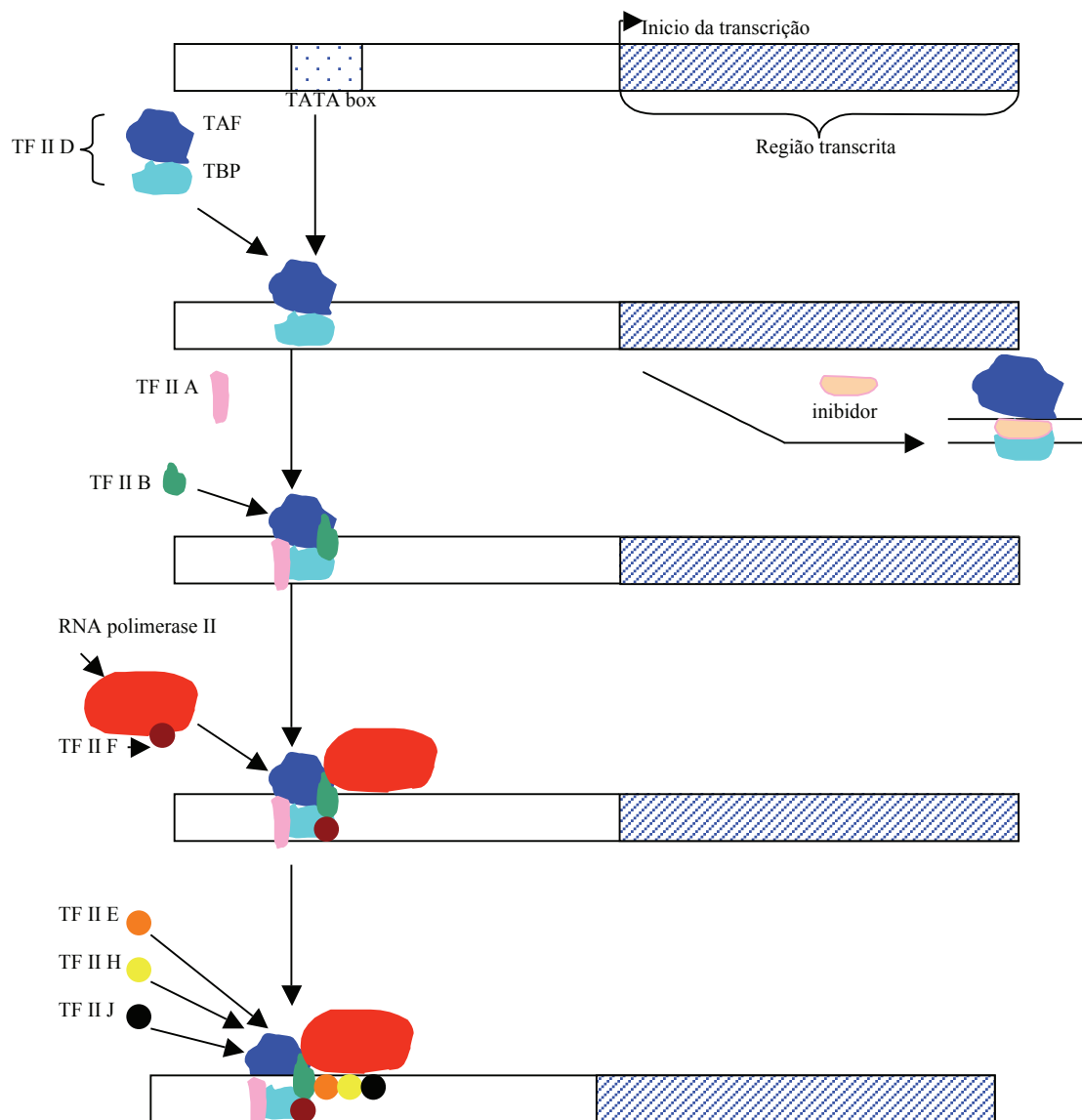
Tanto as RNA polimerase II, I e III nos eucariotas necessitam da subunidade TBP para começar a síntese do respectivo RNA, mas os promotores para a RNA polimerase II possuem uma caixa TATA, enquanto os promotores para as RNA polimerase I e III não têm essa caixa TATA.

## Regulação da actividade dos factores de transcrição eucariotas

Embora o genoma de cada célula em cada ser pluricelular seja idêntico, consoante o tipo de célula e o seu perfil fisiológico, assim a expressão de genes responsáveis pelos variados factores de transcrição.

A biossíntese de alguns factores de transcrição é regulada por hormonas lipido-solúveis e no caso das hormonas polipeptídicas estas podem regular estes factores de transcrição através de fosfatações de diversos tipos.

### Esquema para o “assemble” do complexo de iniciação da RNA polimerase II sobre o promotor TATA



No esquema da figura anterior o “assembly” inicia-se com a ligação do factor geral de transcrição TF II D (composto de várias subunidades as TBP e as TAFs) à caixa TATA.

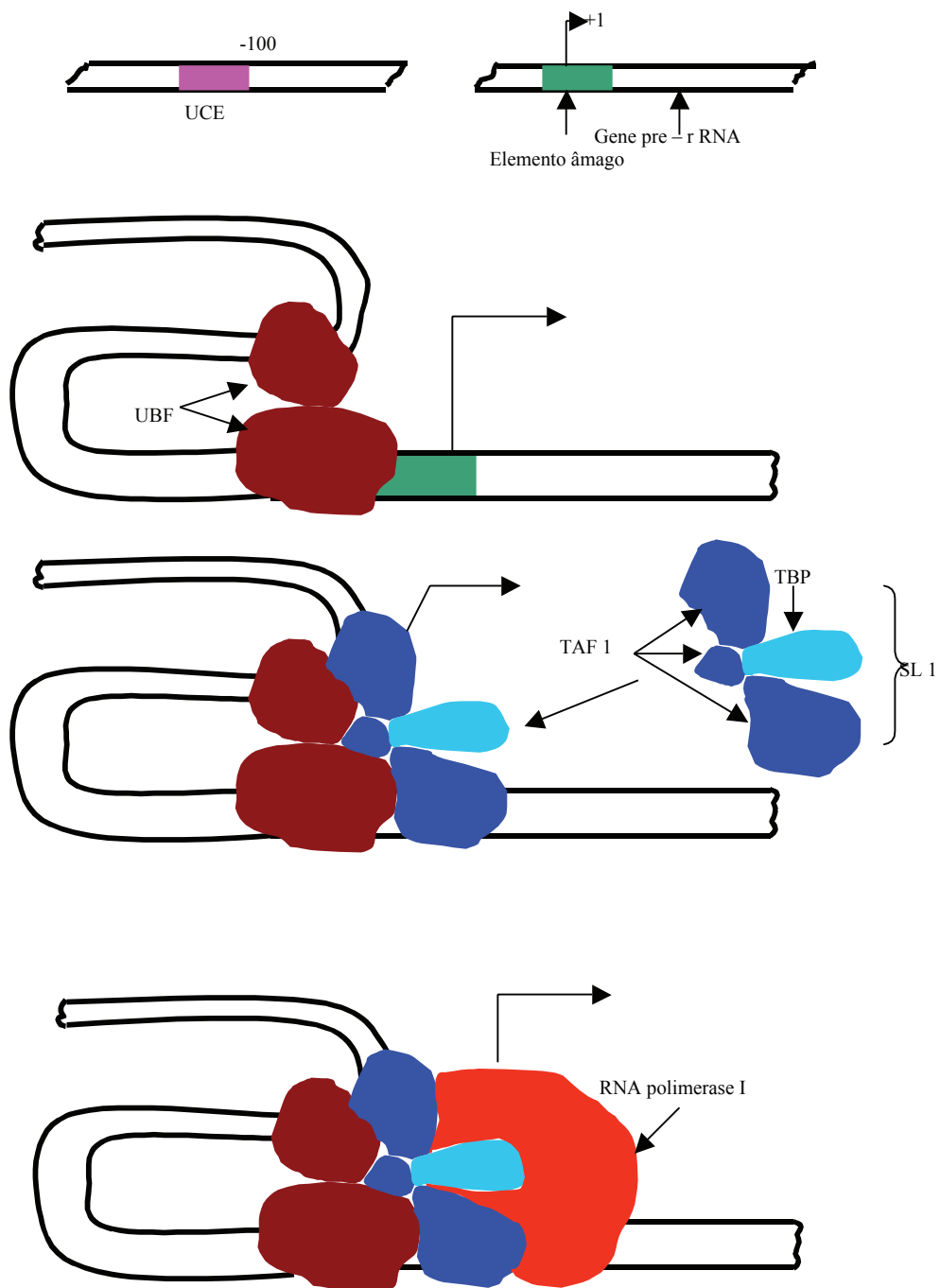
Contudo a presença de um inibidor que interactue com o complexo constituído pelo promotor (caixa TATA) e pelo TFIID pode impedir a interacção com outros factores de transcrição. Se o factor de

transcrição TFIIA se unir ao complexo constituído pelo promotor (caixa TATA) e pelo TFIID, já o inibidor não pode interactuar com este complexo, sendo depois possível a interacção do complexo promotor-TFIID-TFIIA com o factor de transição TFIIB e depois ainda a ligação com o produto da interacção da RNA polimerase II com outro factor TFIIF.

Apesar de tudo só depois da interacção com outros factores de transcrição como os TFIIE, TFIIH e TFIIJ é que se inicia a transcrição.

### Transcrição dos genes RNA ribossomais

#### “Assembly” do complexo de iniciação da RNA polimerase I



A formação do complexo de iniciação da RNA polimerase I, a região para controlo das unidades de transcrição do pré-rRNA, possui dois elementos, um que é o elemento promotor ângulo e que se sobrepõe ao local de início da transcrição (+1) e um outro elemento de regulação localizado cerca de ~100 pares de bases acima (upstream) e que é o UCE (upstream control element).

Um factor de ligação acima o UBF (upstream binding factor) interacciona com os dois elementos antes referidos o que leva o DNA a curvar-se em ansa, levando a que um factor de selectividade 1 (SL1) constituído por diversas sub-unidades proteicas associadas (TAF) interaccione com o complexo anteriormente formado, estabilizando-o permitindo depois a interacção com a RNA polimerase I.

Esta RNA polimerase I apenas é responsável pela formação de um tipo de RNA o pré-rRNA que origina depois os 18S, 5.8S e 28S rRNA.

Realça-se que o DNA que é transcrito em pré-rRNA é muito similar em todos os eucariotas, contendo os genomas de todos de todos os animais, múltiplas cópias em tandem da unidade de transcrição pré-rRNA, sendo no entanto as regiões espaçadoras não transcritas, existentes entre estes genes em tandem, muito variadas no seu tamanho consoante as espécies.

Na transcrição dos genes ribossomais as unidades repetidas contêm uma cópia de cada sequência de RNA (28S, 5.8S e 18S).

Cada unidade é transcrita produzindo transcritos em quantidades iguais dos três tipos de ácidos ribonucleicos (28S, 5.8S e 18S) após maturação dos transcritos primários.

A terminação da transcrição ocorre na região espaçadora antes da RNA polimerase I atingir o promotor da unidade de repetição seguinte.

Esta transcrição dos rRNA é muito rápida e sabe-se que a fosfatação da RNA polimerase I pode inibir essa transcrição.

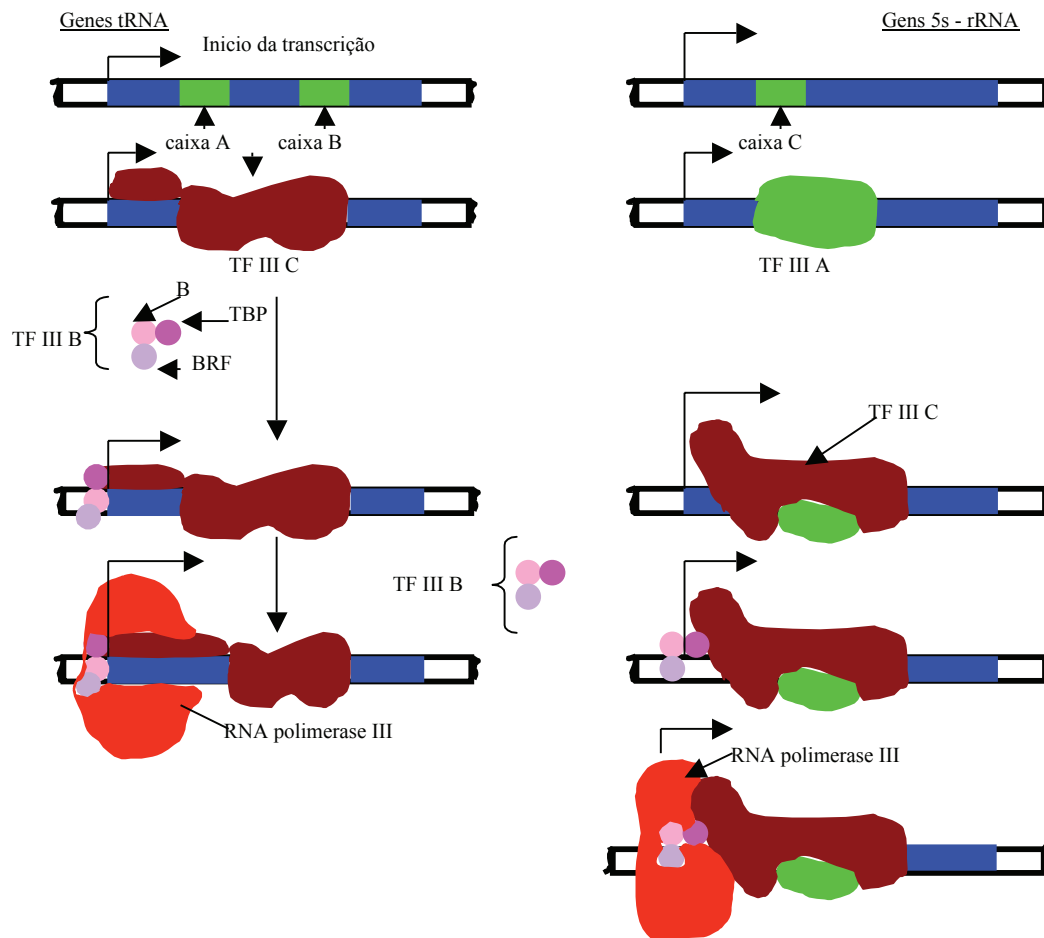
Por outro lado a RNA polimerase I utiliza factores de iniciação específicos da espécie.

Recorda-se que a iniciação da transcrição pela RNA polimerase I não hidroliza ATP ao contrário do que sucede com operação análoga realizada pela RNA polimerase II.



## Transcrição pela RNA polimerase III

### Assembly da RNA polimerase III

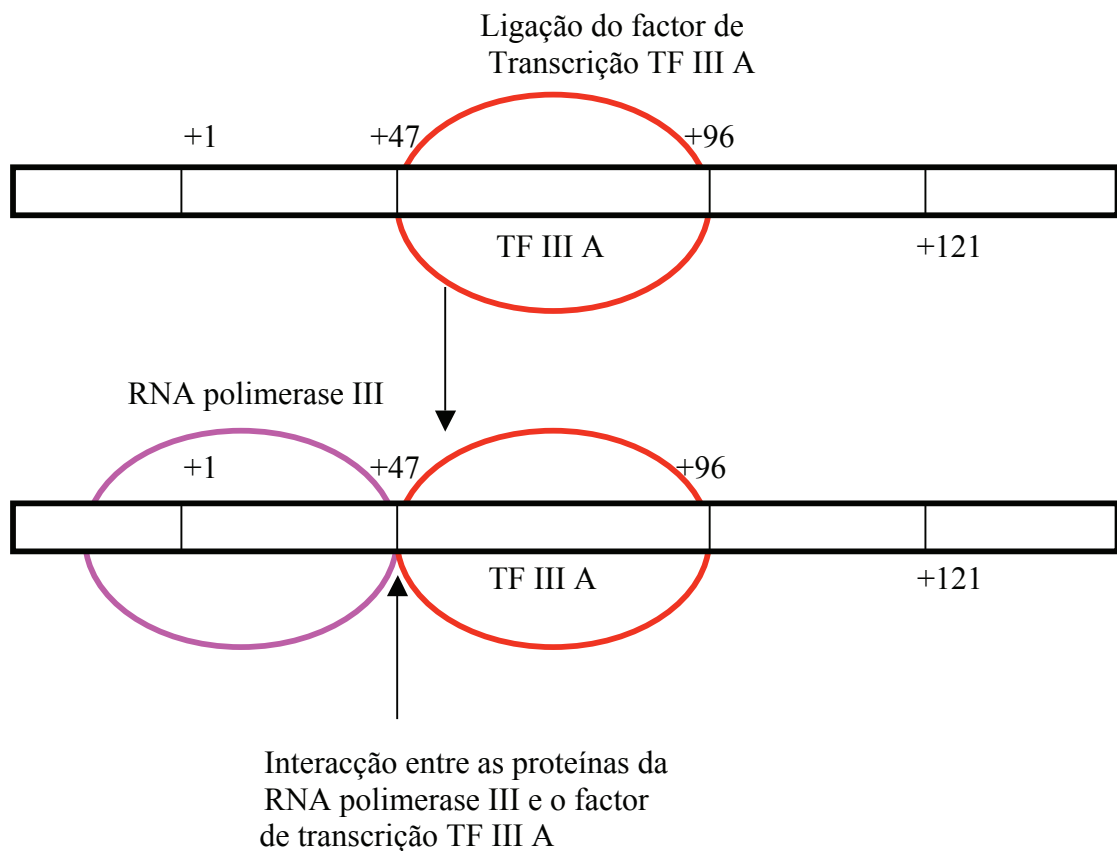


O “assembly” para formação do complexo de iniciação para a RNA polimerase III vai permitir a transcrição dos genes tRNA para o que se torna necessário que uma proteína a TFIIC composta por varias subunidades, interactue com maior intensidade com o promotor da caixa B do que com o promotor da caixa A. Depois desta interacção é a vez do trimerio TFIIB referido na figura anterior interactuar com qualquer sequência do DNA acima do gene do tRNA o que viabiliza a interacção deste complexo com a RNA polimerase III começando a transcrição.

Os genes tRNA são pois controlados por duas regiões que se encontram no interior da sequência transcrita sendo regiões de controlo interno que interactuam com dois factores de transcrição formados por varias subunidades, o TFIIC constituído por seis subunidades que interactuam com a caixa A e a caixa B no promotor, e o TFIIB constituído por três subunidades que interactua acima da caixa A.

Na transcrição pela RNA polimerase III do 5sRNA e dos tRNA, os factores de transcrição interactuando com o DNA regulam aquela enzima com a particularidade de na transcrição do 5sRNA a sequência que interactua com o factor de ligação encontra-se no interior da sequência que codifica o RNA.

## Factor de transcrição para a classe III de genes eucariotas



Como é evidente nesta figura o factor de transcrição liga-se a uma sequência no interior do gene para o 5srRNA, ligando-se depois a RNA polimerase III ao factor e iniciando a transcrição, não sendo necessário qualquer sequência específica além da sequência de ligação ao factor.

Quanto aos genes implicados na transcrição do 5s-rRNA o factor de transcrição TF III A (ver lado direito da penúltima figura) interacciona com a caixa C (controlo interno) permitindo depois a sua interacção com o trímico TFIIC após o que é viável a interacção com o TFIIB e depois a RNA polimerase III começando a transcrição.

Também a iniciação da transcrição pela RNA polimerase III não hidroliza ATP.

A transcrição do DNA mitocondrial revela características idênticas aquelas dos núcleos eucariotas, consistindo o DNA mitocondrial nos vertebrados de um círculo de ~16Kb responsáveis pela síntese de 13 proteínas de membrana, assim como pelos rRNA e tRNAs necessários para a translação dos mRNA mitocondriais.

### **3.4 - Transcrição de genes codificadores de proteínas**

Nos eucariotas na transcrição dos genes codificadores de proteínas são extraordinariamente importantes os seguintes aspectos:

- a) A complexidade do “molde” de cromatina
- b) Os complexos formados pela interacção dos activadores de transcrição e que vão regular a estrutura da cromatina e o aparelho de transcrição
- c) As formas holoenzimáticas da RNA polimerase II que vão desencadear a iniciação da transcrição e depois o alongamento dos mRNAs.
- d) Os mecanismos que interligam o processamento do mRNA e a sua biossíntese.

#### **Cromatina e sua remodelação**

In vitro as RNA polimerases sem factores de transcrição basais ou gerais (GTFs) não conseguem iniciar uma transcrição sobre promotores. In vivo as coisas ainda mais se complicam pois torna-se necessário reguladores a actuarem sobre os genes.

Como vimos anteriormente para estes genes codificadores de proteínas os promotores possuem locais para interactuarem com activadores de transcrição e outros factores que modificam a cromatina e a transcrição. A cromatina pode ser modificada por diversas enzimas que fazem parte de complexos proteicos que intervêm na iniciação da transcrição e no alongamento dos mRNA:

Recorda-se que a estrutura da cromatina foi abordada anteriormente com maior desenvolvimento na parte referente a cromossomas e sua estrutura.

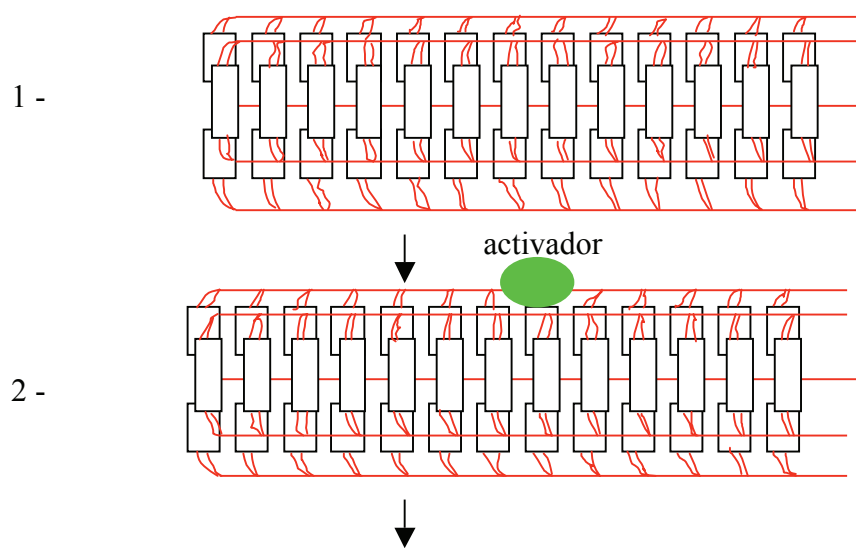
Como nucleoproteína (uma parte de DNA para duas de proteínas) que é, há que considerar nela uma unidade fundamental, repetida, o nucleossoma. Neste temos a porção proteica com um octamero de moléculas de histonas, e o DNA, que se enrola à volta de cada octamero de histonas numa extensão de 146 bp, que se enrolam em redor de cada octamero 1,65 voltas, orgaizando-se grosseiramente na forma de cilindro constituído por dois heterodímeros de histonas H3 e H4 flanqueados por dois outros heterodímeros de histonas H2A e H2B.

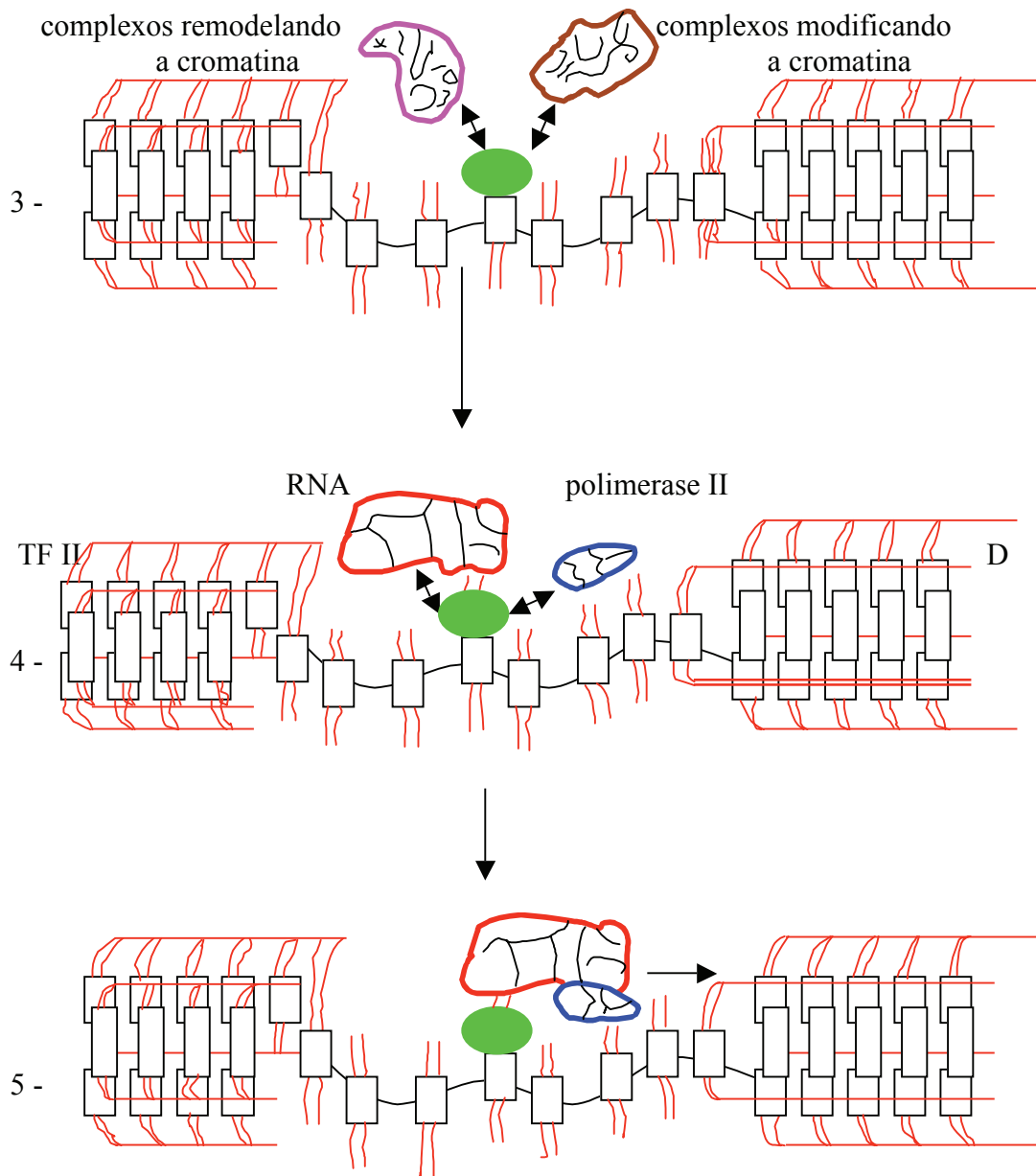
Estas moléculas complexas que formam o nucleossoma resultam das interacções químicas entre as proteínas das histonas e os fosfatos ou desoxiriboses do DNA, havendo ainda interacções entre as caudas amino terminais dos polipéptidos das histonas e os nucleossomas adjacentes.

Contudo apesar destas interacções existe uma certa flexibilidade ou instabilidade dentro deste complexo inclusivé nas interacções anteriormente referidas, o que constitui um acontecimento basilar para que os genes aí residentes possam ser expressos.

Salienta-se que as quatro subunidades das histonas contidas no octamero, são proteínas que nos eucariotas se revelam altamente conservadas de espécie para espécie, embora nas aves (frangos) estejam referidas variantes.

Modelo proposto para activação dos genes





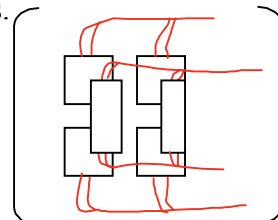
No diagrama anterior apresenta-se o início da transcrição com um simples activador● ligado a um promotor, mas os promotores habitualmente contêm múltiplos locais para ligação de múltiplos activadores.

Os activadores ligam-se a sequências específicas do DNA e induzem a remodelação da cromatina com complexos adequados ( — ) e com outros complexos modificantes ( — ) influenciam a estrutura local da cromatina.

Os activadores também recrutam e ligam o aparelho de iniciação da transcrição ( — ) e ( — ) para os promotores. A maioria destes processos são reversíveis e são regulados por repressores.

No diagrama referem-se as seguintes etapas:

1. O DNA está contido nos nucleossomas ( ) e estas estruturas de cromatina organizadas em formas mais condensadas.



2. Um activador liga-se a um elemento regulador.
3. O activador actuando sobre a estrutura de cromatina remodela-a e modifica-a através de complexos que este activador recruta.
4. A activação recruta outros componentes do aparelho de transcrição.
5. O activador pode estimular a actividade do aparelho de transcrição “assemblado”.

Salienta-se contudo que as interacções entre a porção proteica e nucleica dentro dos nucleossomas assim como a organização dos nucleossomas em contactos entre si podem condicionar enormemente a acessibilidade do DNA para que possa ocorrer expressão dos genes correspondentes.

Tudo aquilo que contribua para diminuir esta complexidade interactuante inclusivé ao nível do nucleossoma vai facilitar a expressão de alguns genes, podendo no entanto simultaneamente ocorrer a diminuição da expressão de outros genes (vide figura seguinte).

De qualquer forma é também possível que os sucessivos nucleossomas se organizem em super-estruturas enroladas ou dobradas em fibras cromatínicas contribuindo para isto as histonas H1 que são histonas de ligação (“linkers”), assim como as porções amino terminais flexíveis das histonas irradiando dos nucleossomas podem também favorecer a formação de fibras.

Recorda-se novamente que no DNA têm sede os chamados elementos promotores para a transcrição, assim como outros elementos, havendo três aspectos idênticos na maior parte dos genes codificadores de proteínas e que são:

- a) o local onde se inicia a transcrição
- b) a caixa TATA
- c) outras sequências do DNA a que se ligam reguladores da transcrição.

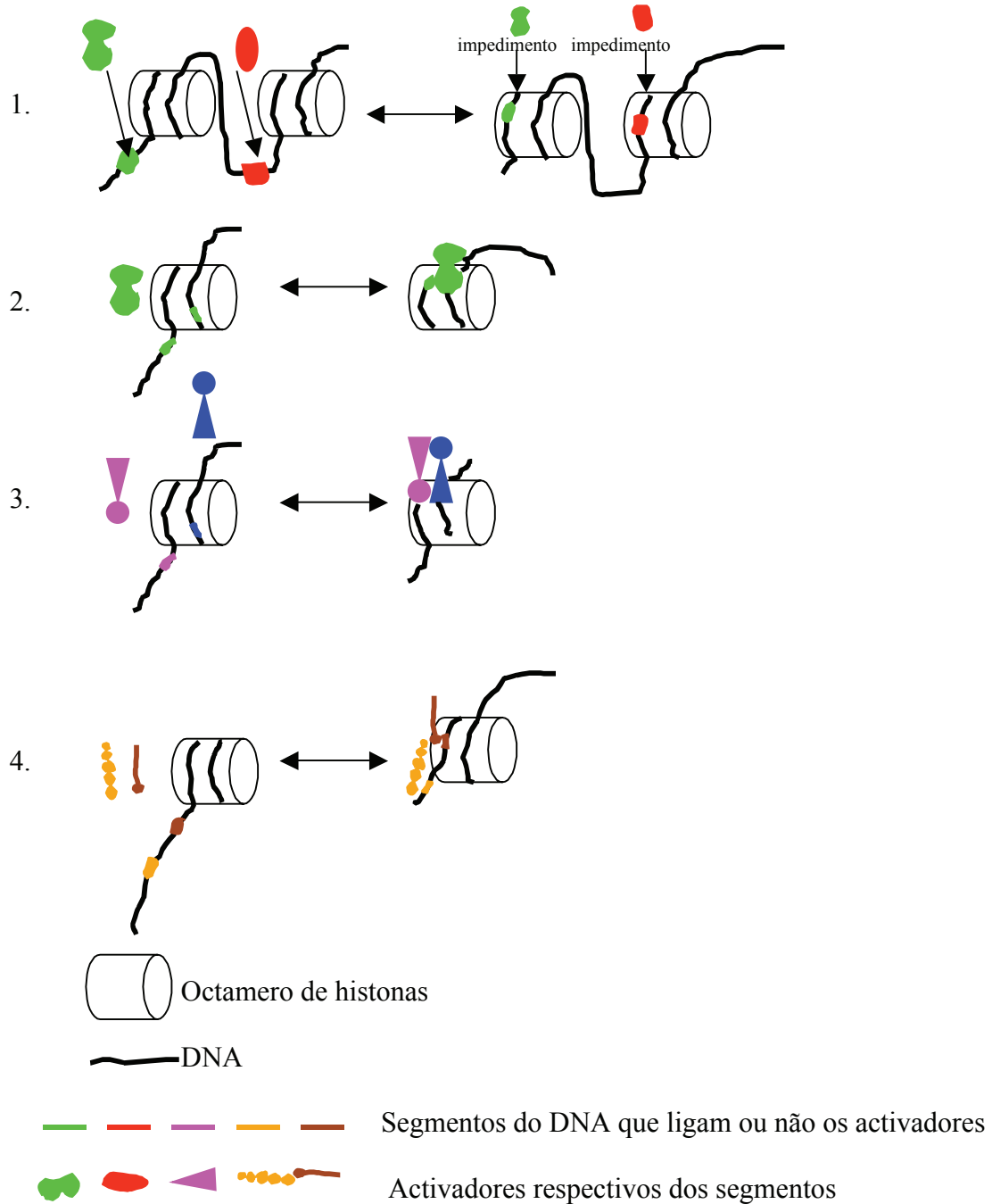
Chama-se elemento promotor âmago a um conjunto de cerca de 100bp que engloba o local onde se inicia a transcrição e a caixa TATA, e este elemento é suficiente para que a transcrição se inicie.

As sequências do DNA a que se ligam reguladores da transcrição, englobam as sequências activadoras acima (UASs), as dos favorecedores (“enhancers”), as sequências repressoras acima (URSs) e as sequências silenciadoras.

A activação promovida pelos activadores pode ser esquematizada como se refere na figura anexa, recordando-se mais uma vez que os activadores possuem um domínio para interactuar especificamente com o DNA e outro para interactuar com outros factores de transcrição, e repete-se novamente que a caixa TATA serve para interacção com a proteína TBP.

Diversos genes podem ser activados por um único activador e pelo contrário um único gene pode ser activado por diversos activadores, sendo as enzimas acetilases actuantes sobre as histonas componentes de diversas actividades de transcrição.

Efeitos dos nucleossomas sobre a ligação dos activadores e “maquinaria” de transcrição



A posição relativa do DNA em relação às histonas pode condicionar fortemente a transcrição.

- 1) A “embalagem” do DNA nos nucleossomas em princípio é um bloqueio á transcrição uma vez que o nucleossoma interfere com a ligação de activadores ou de elementos da maquinaria de transcrição.

- 2) Os nucleossomas podem favorecer a transcrição ao posicionar dois segmentos distintos do DNA de forma a que eles constituam um local de ligação adequado para o activador.
- 3) Os nucleossomas podem ainda proporcionar a constituição de um local para ligação de dois activadores que se entre-ajudam.
- 4) Os nucleossomas podem alterar as distâncias ou as posições entre diversos factores de transcrição, estimulando a interacção entre estes factores.

No que se refere aos repressores da transcrição há que referir duas classes, uma classe dita de factores gerais e outra classe que corresponde a factores específicos apenas para certos genes. Alguns destes repressores actuam por interacção com a proteína TBP que se liga á caixa TATA.

Desenvolvemos seguidamente algumas modificações que podem ocorrer na cromatina condicionando a transcrição de diversas formas.

Como já referimos anteriormente o empacotamento do DNA nos nucleossomas tal como a formação de mais complexas estruturas dificulta a transcrição, embora o empacotamento do DNA dentro dos nucleossomas possa ter efeitos antagónicos, ou de repressão ou de activação da transcrição que vai depender das “exigências” do promotor.

Os complexos de cromatina formados por interacção entre os nucleossomas resultam em parte da interacção com as porções N-terminais dos polipéptidos histonas. Mais, alterações químicas desta caudas N-terminais das histonas como sucede por exemplo quando são acetiladas estão correlacionadas com a actividade de transcrição, assim como com a rotura das estruturas de cromatina mais complexas.

Por outro lado a desacetilação das histonas está correlacionada com a indução da formação de estruturas cromatínicas mais complexas e daí repressão da transcrição.

Mas além da acetilação ou desacetilação as histonas podem ser modificadas por fosfatação, desfosfatação, metilação, desmetilação e ubiquitinação conhecendo-se em muitas situações os locais ou sejam os pontos da estrutura proteica onde ocorrem essas modificações. É o caso da histona H3 que se sabe já ser acetilada nas lisinas nas posições 9, 14, 18 e 23, fosfatada na serina na posição 10 ou metilada nas lisinas nas posições 9 e 27.

A combinação de modificações químicas operadas nas caudas (porções N-terminais) das histonas pode ser a causa ou melhor constituir um código que induz o DNA molde (“template”) a operar por exemplo para assembly, mitose, transcrição ou replicação.

Muitas destas modificações químicas são reversíveis sendo a melhor conhecida a acetilação reversível das histonas. Como já se afirmou histonas hiperacetiladas permitem uma transcrição mais activa e se hipoacetiladas sucede o inverso. A acetilação vai promover a rotura das estruturas dos nucleossomas, deixando o DNA disponível ou livre para ser transcrito, podendo também alterar a interacção com nucleossomas vizinhos. Existem bastantes acetiltransferases assinaladas para as histonas, podendo essa acetilação estender-se a outras proteínas como por exemplo factores gerais de transcrição TFIIE e TFIIIF, tendo a acetilação destes factores efeitos diversos (positivos ou negativos) sobre a transcrição e mesmo ao nível dos promotores.



Também as fosfatações desencadeadas por cinases sobre as histonas (H3) favorece a transcrição. A RSK.2 H3 cinase e a MSK1 cinase que é activada por um factor de crescimento fosfatam as histonas H3.

As modificações por metilação são pois pior conhecidas quanto aos seus efeitos sobre a transcrição.

Alem destas modificações químicas referidas que são modificações covalentes, a cromatina pode ser modificada sem ser por modificações covalentes que originem maior fluidez remodelante dos nucleossomas com alteração da sua conformação.

Concluindo a cromatina pode ser modificada pela acção de vários tipos de enzimas modificação esta regulada por activadores e repressores da transcrição, visando tornar esse substrato mais ou menos passível de ser transcrito.

### **RNA polimerase II e cofactores de iniciação**

Para que o aparelho para iniciação da transcrição possa funcionar torna-se necessário que os factores activadores de transcrição recrutem a RNA polimerase, contendo aquele aparelho para os promotores dos genes codificadores de proteínas.

Este complexo do aparelho para iniciação da transcrição contem:

- a) a enzima RNA polimerase II com 12 sub-unidades
- b) os factores de transcrição geral
- c) um ou mais complexos de varias sub-unidades denominadas coactivadores ou mediadores

<u>RNA polimerase II e factores gerais de transcrição</u>
Polimerase II - 12 sub-unidades
TF II D
TF II A
TF II B
TF II F
TF II E
TF II H

In vitro para que a RNA polimerase II se possa ligar especificamente com o promotor do respectivo gene são necessários o conjunto básico de factores gerais de transcrição (GTFs) referidos no quadro anterior.

In vivo é possível que as coisas ainda sejam mais complexas talvez com a necessidade de mais factores.

È todo este complexo ligado ao promotor do DNA que “funde” este formando uma estrutura complexa aberta em que 12-15bp da sequência promotora estão na forma de uma bolha de cadeia simples (formação do complexo aberto). Seguidamente forma-se a primeira ligação fosfodiéster entre o primeiro e segundo ribonucleótidos incorporados na biossíntese.

Existe na RNA polimerase II, na sua maior subunidade, um detalhe ultraestrutural de grande importância. Esta subunidade maior da RNA polimerase II possui um domínio carboxílico-terminal único e repetido (CTD). Este domínio contém um tandem de repetições de uma sequência consenso heptapeptídica (Tyr-Ser-Pro-Thr-Ser-Pro-Ser). (Recordar-se que sequências consenso são os nucleótidos ou ácidos aminados mais normalmente encontrados em cada posição nas sequências de DNA, RNA ou proteínas-ver homologia também).

Aquela sequência heptapeptídica altamente conservada nos eucariotas pode ter 26 ou 27 repetições do heptapeptídeo no *S. cerevisiae* e até 52 nos humanos.

O papel destas CTD é importantíssimo na viabilização das células dos metazoários, pois são locais que se podem fosforilar, encontrando-se vários graus dessa fosforilação. Assim as RNA polimerase II não fosforiladas neste domínio CTD são as que se encontram nos complexos de iniciação da transcrição, mas no estado posterior de alongamento da transcrição os CTD encontram-se fortemente fosforilados.

Logo esta fosforilação do CTD ocorre na transição para o alongamento, sendo conhecidas duas cinases para este efeito. É o caso da subunidade Cdk7 do factor de transcrição TFIIF e da srb10/cdk8 cinase que faz parte do complexo srb mediador.

Além das fosforilações podem ocorrer as remoções dos fosfatos dos CTD e essas fosfatases existem nos factores de transcrição TFIIF e TFIIB.

### **RNA polimerase II e cofactores de alongamento**

A transição da fase de iniciação da transcrição para a fase de alongamento implica algumas modificações, a saber:

- a) A fosforilação do domínio CTD da RNA polimerase II como acabamos de referir
- b) A permuta dos factores associados com a RNA polimerase II por outros já identificados.

Como já é conhecido ocorrem uma série de modificações sobre estes transcritos primários mRNA produzidos pela RNA polimerase II destacando-se:

- 1) Modificações nas extremidades 5' e 3'.
- 2) Mecanismos de "splicing"
- 3) A extremidade 5' é "capeada" com uma guanósina trifosfato 7-metilada
- 4) A extremidade 3' após cisão é acrescida com poliadenilatos

Parece que o domínio CTD intacto da RNA polimerase é indispensável para muitos destes acontecimentos.

Podemos pois concluir que a iniciação da transcrição, o alongamento e o processamento dos mRNAs estão ligados através de variadas funções, com a RNA polimerase II CTD

### **3.5 - Biologia do factor de transcrição IID (TF IID)**

Para poderem reconhecer o seu promotor alvo e dar início ao respectivo processo de transcrição as diversas RNA polimerases I, II e III necessitam da colaboração de uma série de factores acessórios dessa transcrição ou sejam os factores gerais para iniciação da transcrição, pois não conseguem reconhecer o seu promotor alvo directamente. E isto é necessário para a transcrição dos promotores correspondentes aos genes nucleares da classe I, II e III.

Para regular a transcrição dos promotores da classe II são precisas diversas entidades enzimáticas que colaborem com os factores gerais de transcrição com activadores da transcrição e com co-activadores.

É fulcral em todo este contexto o papel de um complexo de 700KD, o TFIID.

Neste complexo TFIID como já referimos existe uma proteína que se liga á caixa TATA a TBP e uma série de outros factores associados com ela ou sejam os TAFII, que em conjunto conduzem a assemblagem da maquinaria de transcrição, contribuindo ainda para regular a expressão dos genes.

Como dissemos em cima qualquer das RNA polimerases I, II e III não consegue reconhecer o seu promotor alvo directamente e para o fazerem socorrem-se da subunidade proteica TBP que é pois um factor universal.

Repete-se que os genes nucleares da classe II contêm nos promotores do seu DNA três classes de alvos e que são os elementos âmagos ou promotores básicos, os elementos promotores proximais e os elementos favorecedores (enhancers) distais, ocorrendo portanto interacções de diversos factores de transcrição com estas três classes de alvos, além de interacções entre os diversos factores.

Os melhor conhecidos destes alvos são os elementos promotores âmagos ou básicos, sobretudo o elemento TATA situado 25 bp acima do local onde ocorre o início da transcrição, e um elemento iniciador localizado no sitio do início da transcrição e rico em pirimidinas (Inr).

É possível que existam promotores âmagos ou básicos individuais, possuindo ambos estes elementos TATA e Inr, ou nenhum deles ou alguns deles, sendo identificados como TATA+Inr+, TATA+Inr-, TATA-Inr+ ou TATA-Inr-.

Os alvos dos elementos promotores proximais situados variavelmente entre as 50 e 200 bp acima do local “cap” são interactuados com activadores transcriptionais, que desta forma também regulam a transcrição.

Os terceiros alvos ou elementos favorecedores (enhancers) podem situar-se em qualquer direcção e orientação bastante afastado do local de início da transcrição, interactuando com diversos factores regulando também a actividade da RNA polimerase II.

Cada uma destas três sequências nucleotídicas alvo de um gene promotor pode interactuar com diversos tipos de proteínas.

Assim os factores gerais iniciais TFIIA, B, D, E e F interactuam sobre o promotor âmagos com a RNA polimerase II antes de começar a transcrição, tal como se esquematiza na figura seguinte.

De todos estes factores apenas o TFIID pode interactuar com um local específico do DNA, a caixa TATA e só depois de formado este complexo para início da transcrição é que ocorre a interacção com

outros factores gerais da iniciação e a RNA polimerase II originando-se o complexo de pré-iniciação denominado PIC.

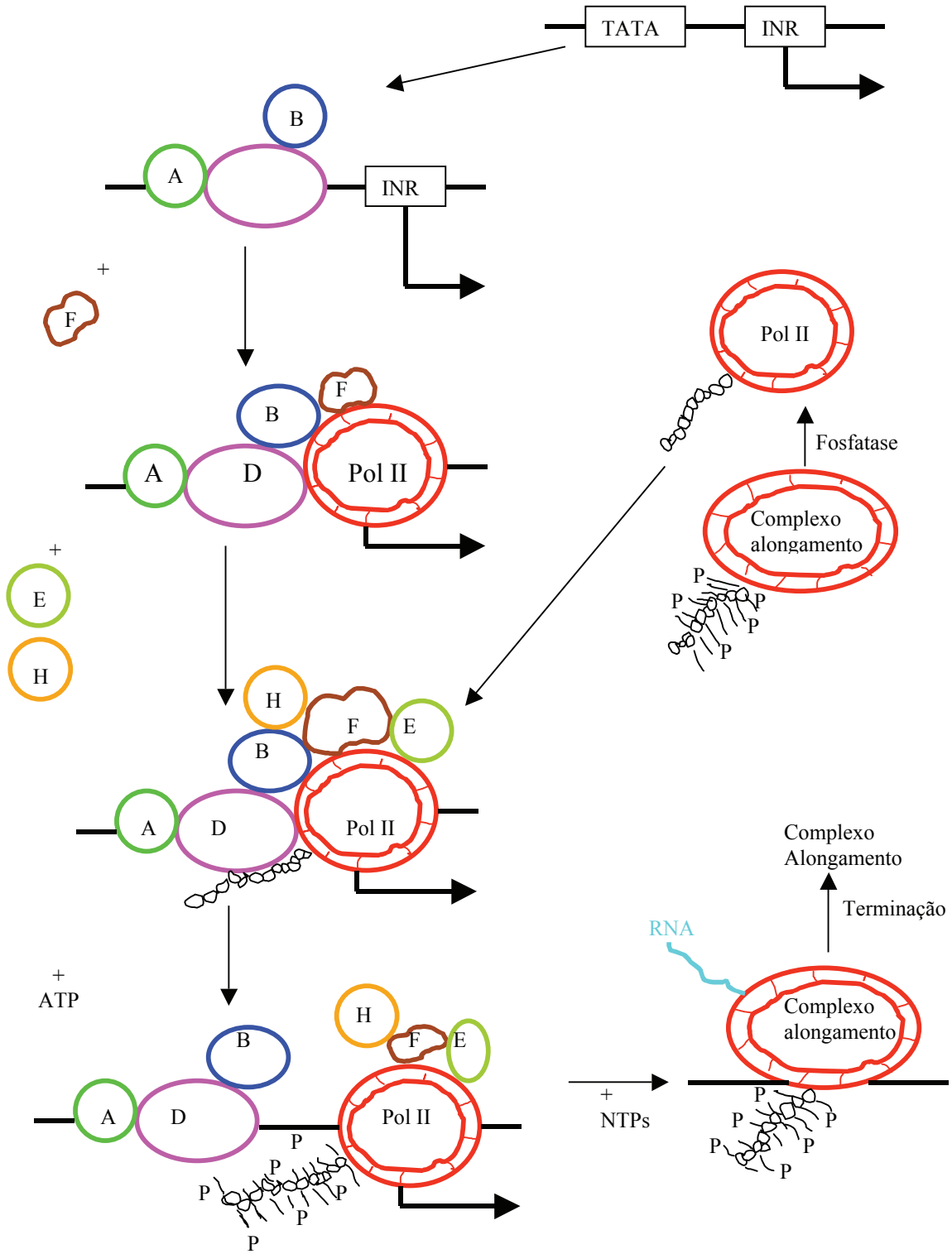
No TFIID a proteína TBP interaccua com uma série de factores associados os TAFs. Mas a TBP também interaccua com os factores de iniciação TFIIA e TFIIB bem como com a porção N-terminal da grande subunidade da RNA polimerase II.

In vivo a iniciação da transcrição pela RNA polimerase II necessita de TFIIE, TFIIIF, TFIIH e talvez TFIIA.

Formado o complexo de pré-iniciação PIC e com o concurso de trinucleótidos trifosfatados (NTP) processa-se a separação ou abertura das duas hélices do DNA no local de início da transcrição (complexo aberto), sendo por outro lado o domínio C-terminal da grande subunidade da RNA polimerase II fosfatado dando-se início á transcrição.

Os mecanismos de acção da TBP na transcrição pelas RNA polimerase I e III é muito mal conhecido.

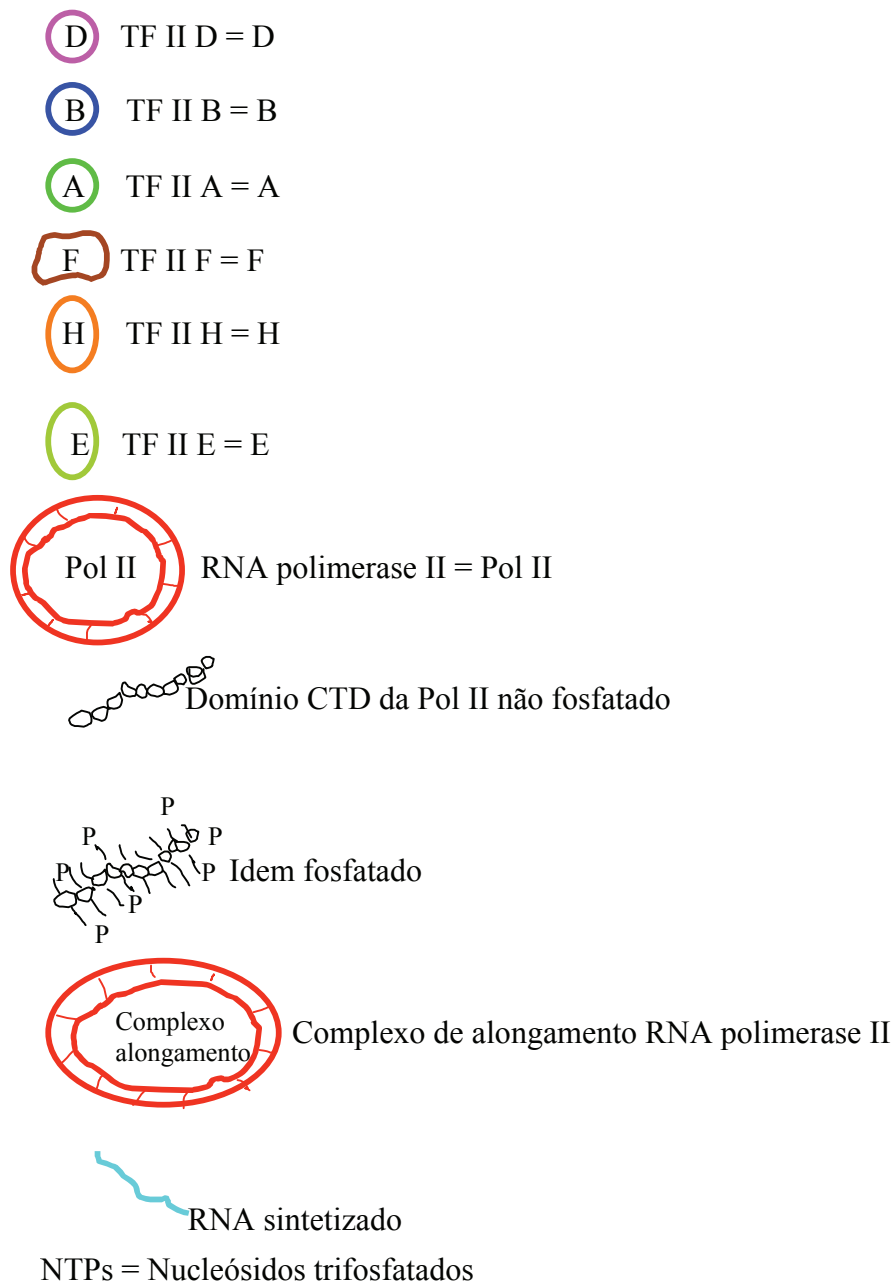
### Formação do complexo de pré-iniciação para transcrição



A formação do complexo de pre-iniciação da transcrição começa quando o TF II D reconhece o elemento TATA e se liga a ele unindo-se depois ao TF II A e TF II B, á forma não fosfatada da Pol II e ao TF II F, TF II E e TF II H.

Antes do alongamento da cadeia polipeptídica, a Pol II é fosfatada pela TF II H.

A seguir a terminação da cadeia polipeptídica, uma fosfatase desfosfata a Pol II permitindo que esta enzima reinicie novo ciclo de transcrição.



Sinais provenientes do meio ambiente podem condicionar a iniciação da transcrição dos genes da classe II na medida em que podem regular a formação deste grande complexo formado por diversas proteínas sobre o gene promotor, tal como se representa na figura seguinte onde se esquematizam uma série de acontecimentos moleculares que representam distintos pontos de controle para regular a actividade da RNA polimerase II.

A assemblagem do grande complexo multiproteico sobre o gene promotor implica interacções entre os diversos componentes da maquinaria de transcrição (TBP, TAFs, TFIIA, TFIIB, RNA polimerase II, TFIIF, TFIIE, e TFIIH) e interacções também com os activadores de transcrição ligados ao elemento promotor proximal ou ao elemento favorecedor (“enhancer”) distal.

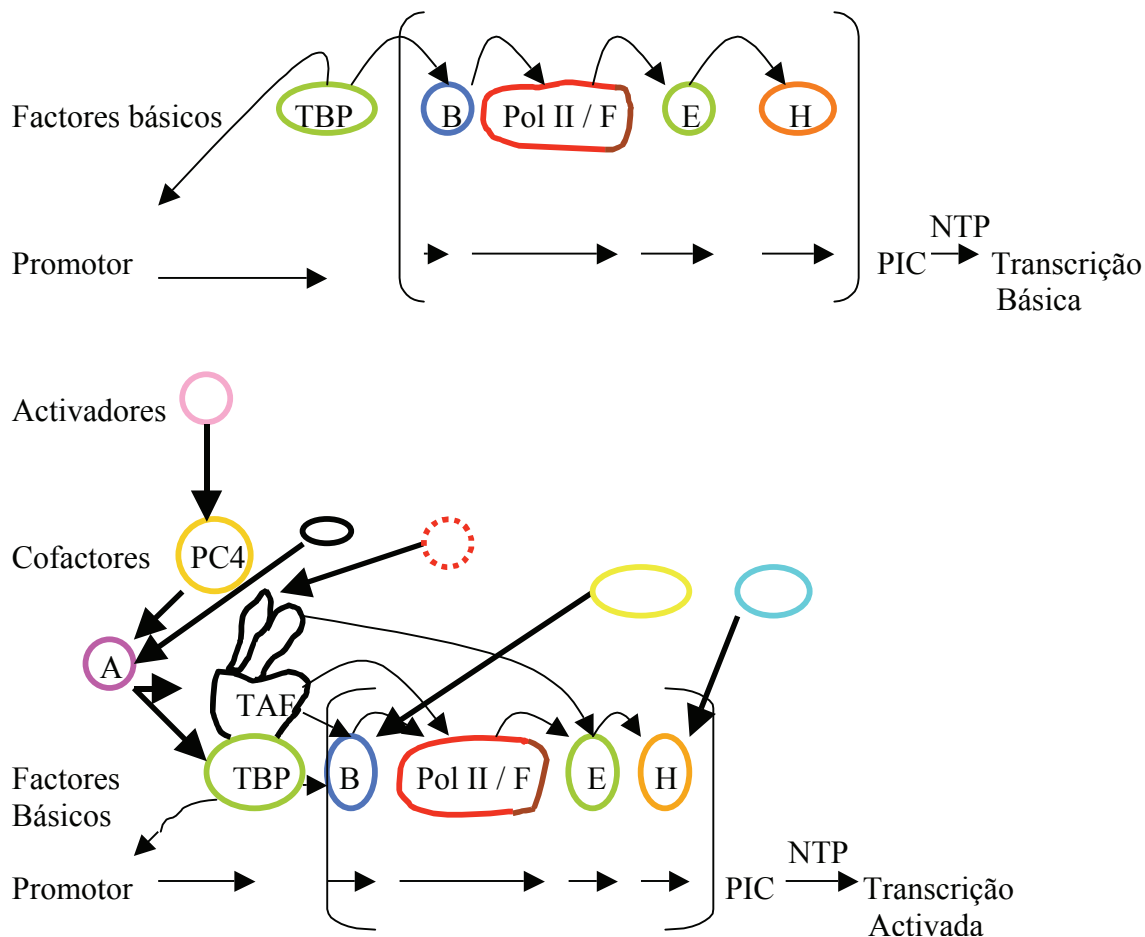
Também é já reconhecido que os activadores de transcrição não necessitam de estar ligados directamente aos factores gerais de transcrição no PIC.

Coactivadores podem actuar como adaptadores entre os activadores e os factores básicos.

Na figura em apreço, na parte de baixo mostram-se os activadores da transcrição e os coactivadores interagindo com os seus alvos no PIC e estabilizando este.

Nos promotores TATA<sup>-</sup> Inr<sup>-</sup>, activadores podem ligar-se a qualquer promotor proximal ou elementos enhancer distais e recrutar TF II D para o promotor, suportando o início da transcrição de uma série de sítios para inicio de transcrição.

Interacções que modelam a transcrição básica (em cima) e a dependente de activação (em baixo)



### **3.6. – Complexos coactivadores de transcrição**

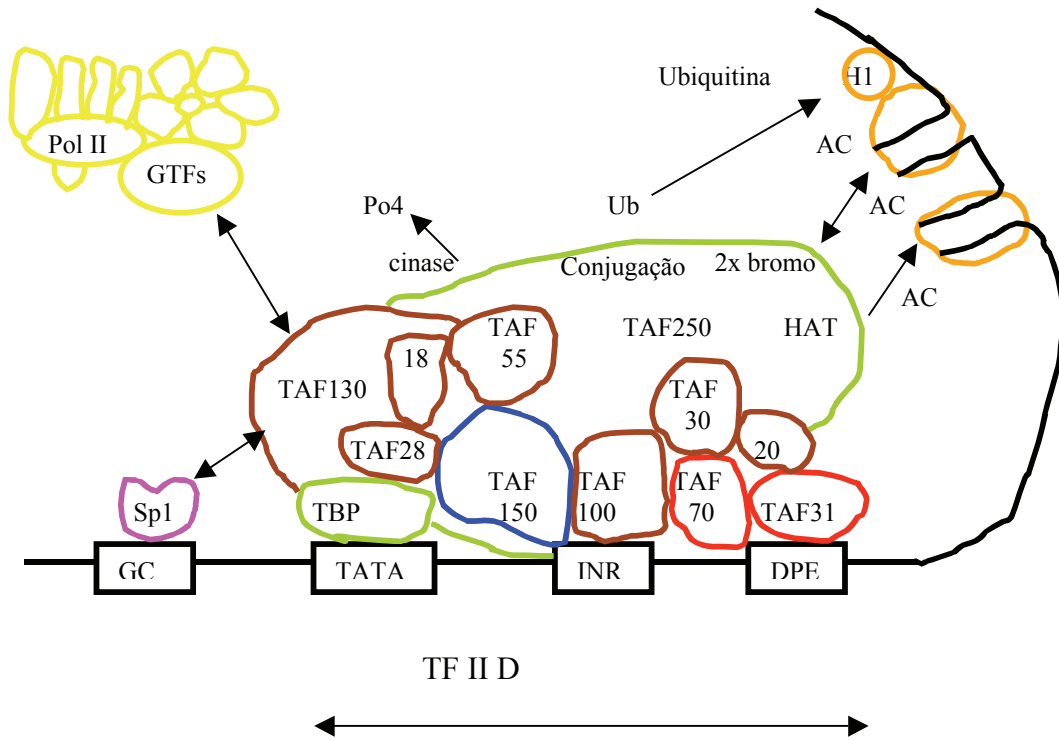
Definem-se como coactivadores de transcrição uma série de factores proteicos que interactivam, por um lado com os activadores ligados a sequências específicas do DNA, e por outro lado com o aparelho geral de transcrição e que ajudam todos estes complexos a “navegar” através da cromatina. Logo há duas classes de coactivadores ou sejam a classe que através das suas interacções orientam a instalação do aparelho de transcrição, e a classe que corresponde a proteínas enzimáticas que modificam remodelando a cromatina.

A anteceder o início da transcrição ocorre a interacção do factor TFIID com a caixa TATA situada nos vários promotores do DNA, seguindo-se a interacção com o complexo de pré-iniciação dos factores gerais de transcrição (GTF). No TFIID é a subunidade TBP, portanto a proteína que se liga á caixa TATA, e uma série de factores proteicos os TAFs que funcionam como coactivadores ao interagirem e interligarem as sequências específicas activadoras do DNA e a adaptarem estas á maquinaria básica da transcrição.

Na figura seguinte podemos verificar que o TFIID integra alem da subunidade TBP, dez ou mais outros factores TAFs (tudo dependendo dos organismos).



Funções associadas com a TF II D e suas sub-unidades




—  = DNA

INR = Elemento iniciador

DPE = Elemento promotor abaixo

Ub = Ubiquitina

 = Nucleossomas

AC = Acetilação

H1 = Histona H1

Pol II = RNA polimerase II

GTFs = Outros factores de transcrição gerais

Na TFIID a maior subunidade proteica é a TAF 250 que é dotada de diversas actividades catalíticas de proteínas cinases e histonas acetiltransferases (HAT), possuindo ainda dois bromodomínios, que são módulos estruturais que existem em diversos coactivadores e em outros factores implicados nas modificações da cromatina.

Na figura anteriormente referida, resumem-se as várias funções da TFIID na activação da transcrição bem como a participação de varias subunidades TAFs neste contexto. Nesta figura o activador alvo é indicado pela seta de interacção entre o TAF 130 e o transactivador Sp1, estando diversas subunidades proteicas componentes do TFIID a interactuar com as sequências do DNA, com a subunidade TBP ligada á caixa TATA, enquanto a subunidade TAF 150 interactua com o elemento iniciador no DNA(INR) e as subunidades TAF 70 e TAF 31 a interactuarem com o elemento promotor abaixo no DNA(DPE).

Ainda explorando a mesma figura verifica-se que a maior subunidade do TFIID, ou seja a TAF 250 revela variadas actividades tais como a da proteína cinase, a de ubiquitina ligase, a de histona acetiltransferase (HAT) possuindo ainda uma porção estrutural com bromodomínios em tandem (2 vezes bromo) que pode interactuar com as histonas assinaladas.

Este complexo de factores, o TFIID constitue o cerne para que se forme o complexo de pré-iniciação com a interacção com a RNA polimerase II bem como outros factores gerais.

Surgem ainda mediadores (“mediators”) que são complexos activadores implicados na fosfatação do domínio carboxilterminal (CTD) da maior subunidade da RNA polimerase II regulando desta forma esta actividade.

Recorda-se nesta altura que a cromatina pode assumir-se com regiões genómicas descondensadas e activamente transcritas ou então condensada na forma de heterocromatina (estrutura complexa e praticamente silenciada ou não transcrita). Em todo este contexto da cromatina as histonas de ligação H1, assim como outras proteínas associadas com a cromatina como por exemplo proteínas de alta mobilidade (HMG) e proteínas reguladoras do silenciamento da transcrição (SIR) envolvem e empacotam o DNA em estruturas cromatínicas complexas.

Torna-se necessário pois criar em toda esta estrutura locais de acesso para os activadores da transcrição e respectiva maquinaria, sendo para isso indispensável operar a remodelação da cromatina, com complexos de remodelação nos nucleossomas dependentes de ATP, bem como com a actividade enzimática das acetiltransferases.

Tudo isto desestabiliza a estrutura da cromatina, na medida em que ocorrem deslocamentos nos contactos entre as histonas e o DNA:

É conhecido que genes activamente transcritos estão correlacionados com acetilações aumentadas das histonas e vice-versa, histonas hipoacetiladas estão correlacionadas com genes silenciados como sucede nos que se encontram na heterocromatina.

Mas paradoxalmente existem diversos coactivadores de transcrição e correpressores que têm sede em subunidades proteicas, umas com actividade de histonas acetilases (HAT) enquanto outras têm sede em subunidades com actividade de histonas desacetilases (HDAC).

Locais que se encontrem diacetilados são facilmente reconhecidos e interactivam com complexos co-activadores como por exemplo o TFIID. Estes locais diacetilados encontram-se nas regiões terminais das histonas, sendo activamente transcritos.

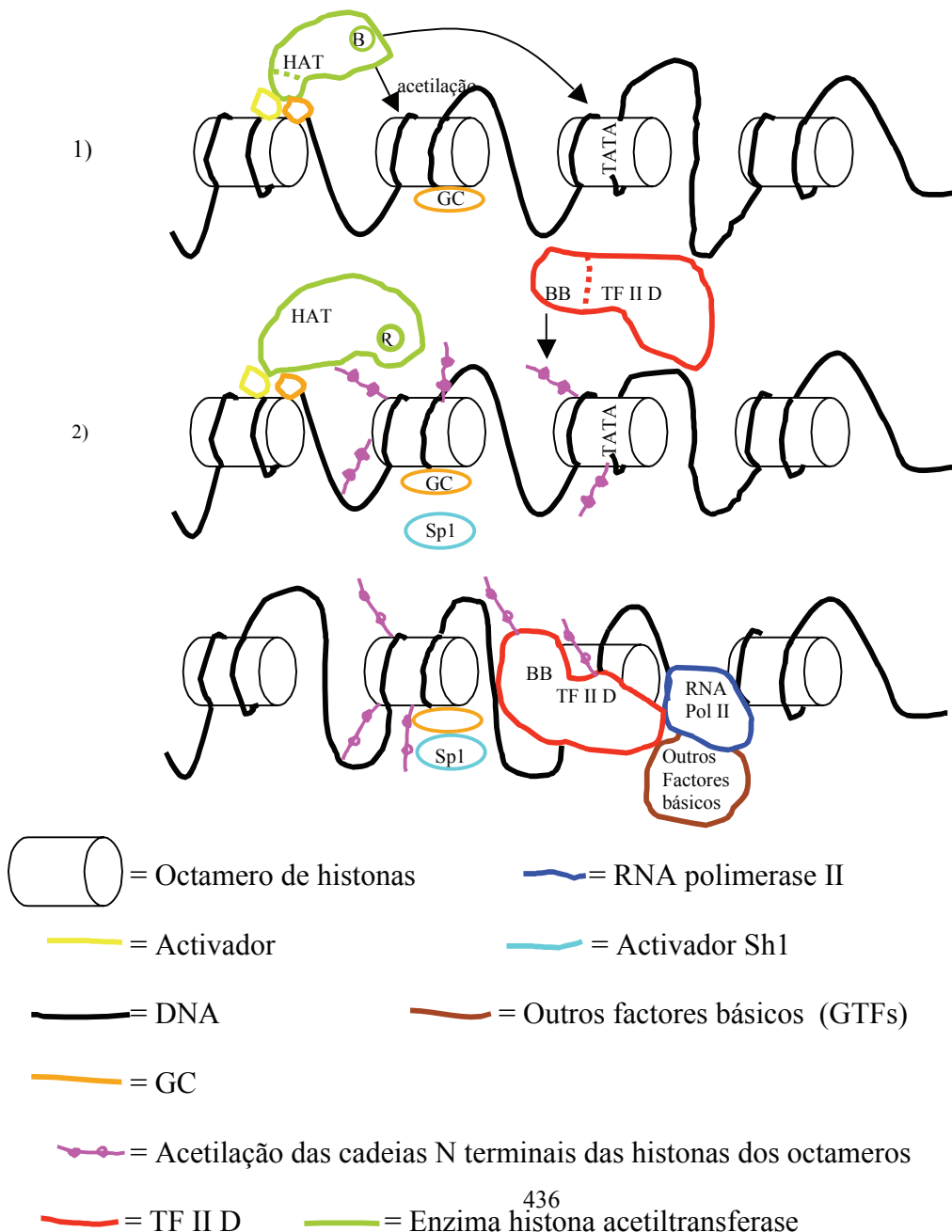
No esquema da figura seguinte procura-se mostrar o papel potencial de acetilação das histonas e dos bromodomínios da TAF250 no recrutamento e estabilização do coactivador TF II D complexado com um promotor ângulo.

Assim em 1) os activadores recrutam HAT, enzima que acetila as histonas em redor do ângulo do promotor.

Em 2) os bromodomínios (BB) do TAF250 (TF II D) reconhecem as histonas diacetiladas.

Em 3) os bromodomínios do TAF250 estabilizam a ligação do complexo TF II D ao ângulo do promotor.

Acetilação das histonas e estabilização da interacção do TF II D com regiões activas da cromatina através dos bromodomínios ( BB ) em tandem da TAF250



Na figura seguinte esquematiza-se um modelo para algumas das diversas actividades desenvolvidas pelos diferentes tipos de complexos co-activadores da transcrição na regulação da actividade dos genes.

1) Complexos de remodelação do nucleossoma e dependentes de ATP ajudam os activadores e a transcrição ao poderem interactuar com sequências promotoras favorecedoras (“enhancer”), mobilizando assim os nucleossomas.

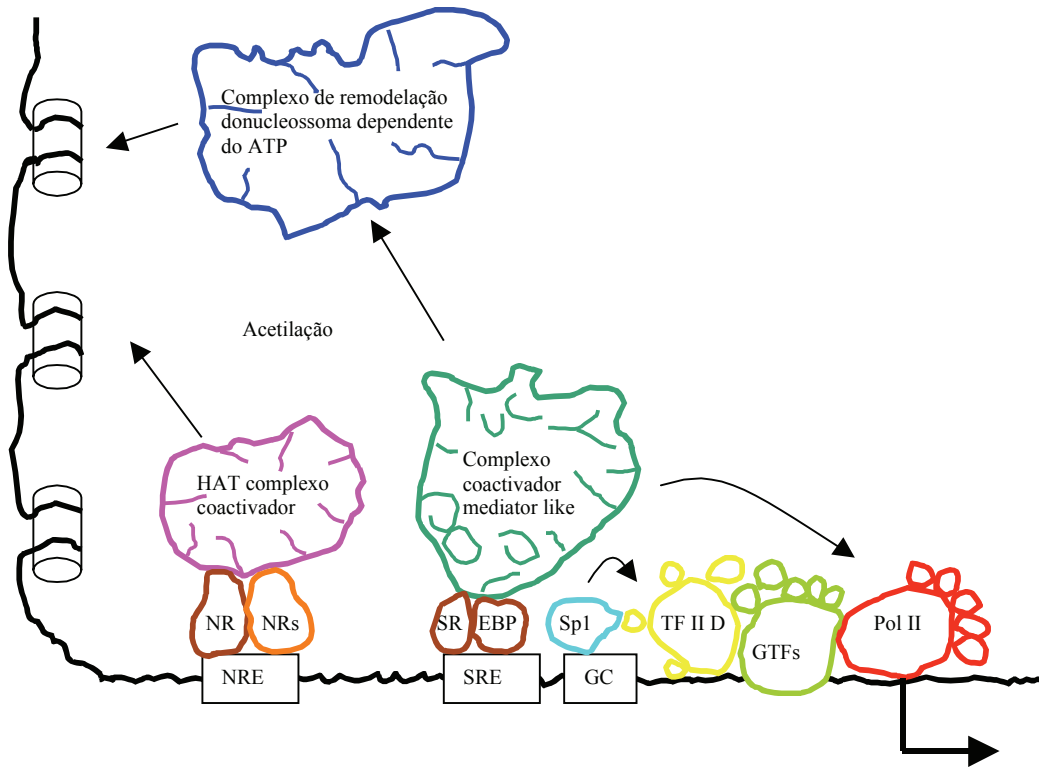
2) As subunidades proteicas com actividade HAT (acetiltransferases das histonas) contendo complexos co-activadores são solicitadas para genes alvo específicos através de activadores do tipo NRs nucleares.

3) A acetilação dos nucleossomas pelas HAT facilitam na estrutura da cromatina a formação de alvos que interactuam com os domínio de bromo. Outros polipéptidos podem também ser acetilados.

4) Outros complexos co-activadores como por exemplo o “Mediator-like” por interacção com activadores como é o caso dos SREPs podem ser orientados para determinados genes mediando o recrutamento da RNA polimerase II.

5) Outros activadores como é o caso do Sp1 podem ao interactuar o TFIID através da TAF formar um complexo de pré-iniciação com os factores gerais de transcrição (GTFs) e com a RNA polimerase II ,iniciando a transcrição.

Diferentes ativadores de distintos complexos coativadores de transcrição na regulação da actividade dos genes



- = Octâmero de histona
- = DNA
- = Complexo de remodelação do nucleossoma dependente do ATP
- = Complexo coativador “ Mediator – Like”
- = Complexo coativador contendo HAT ( Histonas acetil transferase )
- = Receptor nuclear NRs
- = Activador SREBPs
- = RNA polimerase II
- = Activador Sp1
- = Componentes TAF do TF II D
- = Factores gerais de transcrição ( GTF )

Os complexos co-activadores de transcrição regulam a transcrição ao facilitar a acção dos activadores sobre determinadas sequências específicas do DNA dos genes, sendo a presença ou ausência de certas subunidades proteicas nestes complexos co-activadores sujeita a controlo.

## **Bibliografia**

- Myers, L.C. e Kornberg, R.D. (2000) – Mediator of transcriptional regulation. Annu. Rev. Biochem 69: 729 – 49
- Lee, T.I e Young, R.A. (2000) – Transcription of eukariotic protein-coding genes. Annu. Rev. Genete 34: 77 – 137
- Näär, A.M., Lemon, B.D. e Tjian, R. (2001) – Transcriptional coativator complexes. Annu. Rev. Biochem 70: 475 – 501
- Pabo, C.O. e Sauer, R.T. (1992) – Transcription factors structural families and principles of DNA recognition. Annu. Rev. Biochem 61: 1053 – 95
- Parajape, S.M., Kamakaka, R.T. e Kadonaga, J.T. (1994) – Role of chromatin structure in the regulation of transcription by RNA polymerase II. Annu. Rev. Biochem 63: 265 – 97
- Conaway R.C. e Conaway, J.W. (1993) – General initiation factors for RNA polymerase II. Annu. Rev. Biochem 62: 161 – 90
- Lewin, B. (1990) – Continent and activation at Pol II promoters: A tail of protein-Protein interactions: Cell 61: 1161 – 1164
- Silva, A.J., Smith, A.M. e Giese, K.P. (1997) – Gene targeting and the biology of learning and memory. Annu. Rev. Gene 31: 527 – 46
- Burley, S.K. & Roedes, R.O. (1996) – Biochemistry and structural biology of transcription factor II D (TF II D). Annu. Rev. Biochem 65: 769 – 799
- Kobayashi, K. et al (1998) – Nature394 : 388
- Feinberg, A.P. (2000) – Curr Top Microbiol Immunol 249: 87
- Holmquist, G.P. (1992) – Am. J. J. Hum. Gen. 51: 17
- Orphanides, G., Lagrange, T. & Reinberg. (1996) – The general transcription factors of RNA polymerase II. Genes Dev 10: 2657 – 83
- Coin, F. & Egly, J.M. (1998) – Ten years of TF II H. Colds primg Harb. Symp. Quant. Biol 63: 105 – 110
- Burge, C.B., Tushi, T.H. & Sharp, P.A. (1999) – In The RNA World 2<sup>nd</sup> ed
- Staley, J.P. & Guthrie. (1998) – Mechanical devices of the spliceosome motors, clocks, springs and things. Cell 92: 115 – 126
- Reed, R. (1998) – Initial splice-site recognition and parring during pre-m RNA splicing. Curr. Opini. Genet. Dev 6: 215 – 220
- Stryer,L.-Biochemistry, W. H. Freeman and Company,New York,1995,4ed.
- Lodish,H et al.-Molecular Cell Biology,Scientific American Books,W.H.Freeman and Company,New York,1995

## **Iconografia**

As figuras e quadros (bem como os textos que lhes dão apoio) indicados na respectiva iconografia, foram adaptados e/ ou redesenhados a partir das seguintes fontes:

Página 6-In Stryer,L.Biochemistry. W.H.Freeman and Company,New York, 1995, pagina 998.

Página 10-Estrutura esquemática dos domínios da superfamilia de receptores nucleares .In Lodish, H., et all.-Molecular cell biology.Scientific American Books.W.H. Freeman and Company,1995, pagina 470.

Página 11 –Elementos do DNA de resposta. In Stryer,L.,referido anteriormente,pagina 1001.

Página 13 – In Lodish,H., et all.,referido anteriormente,pagina 67.

Página 14 – Idem anterior,pagina 68

Página 15 – Idem anterior,pagina 447.

Página 17-Estrutura primária da proteína “dedo” de zinco. In Devlin,T.M.Textbook of biochemistry and clinical correlations. Wiley-Liss, 1997,pagina 109.

Página 20 -Complexo de dois b ZIP diferentes interagundo com DNA. Idem anterior pagina 111.

Página 21 – Associação do dímero proteico ao DNA com dois motivos H-T-H.,Idem anterior pagina 109.

Página 25 – Interacção de factores de transcrição com promotores , Idem anterior pagina 698.

Página 26 –Esquema para o assembly do complexo de iniciação da RNA polimerase II . In Lodish,H. et all. referido anteriormente,pagina 454.

Página 27 – Assembly do complexo de iniciação da RNA polimerase I,In Idem anterior pagina 470.

Página 29 – Assembly da RNA polimerase III, In Idem anterior pagina 472

Página 30 – Factores de transcrição para a classe III de genes eucariotas In Devlin, T. M., referido anteriormente pagina 700.

Páginas 33 e 34 –Modelo proposto para activação de genes. In Lee, T.I. & Young, R. A..Transcription of eucaryotic protein-coding genes.Annu.Rev.Genet. 2000,34:pag117.

Página 36 –Idem anterior figura 2.

Página 42 – Formação do complexo de pré-iniciação para transcrição. In Burley, S. K. & Roeder, R. G.,Biochemistry and structural biology of transcription factor II D (TFIID).Annu.Rev. Biochem.,1996,63, pagina 772.

Página 44 – Interacções que modelam a transcrição básica (em cima) e a dependente de activação (em baixo). In idem anterior pagina 790.

Página 47 – Funções associadas com a TF II D e suas subunidades. In Naar,A., et all..Transcriptional coactivator complexes. Ann.Rev. Biochem.,2001,70, pagina 480.

Página 49 – Acetilação das histonas e estabilização da interacção do TF II D com regiões activas da cromatina através dos bromodomínios( BR) em tandem da TAF-250. In idem anterior pagina 492.

Página 51 – Diferentes activadores de distintos complexos coactivadores de transcrição na regulação da actividade dos genes. In idem anterior pagina 495.



**INTRODUÇÃO À BIOLOGIA MOLECULAR  
FISIOLÓGICA E PATOLÓGICA**

## ÍNDICE

<b>1 - <u>Ritmos circadianos nos mamíferos</u></b>	444
1.1 – Aspectos de biologia molecular dos ritmos circadianos mamíferos	444
1.2 - Trama molecular do relógio circadiano nos mamíferos	447
1.3 - Vias de input-fótico	449
1.4 - Output do oscilador circadiano dos mamíferos	450
1.5 - Genoma e relógios circadianos	450
<b>2 - <u>Transdução de sinais visuais</u></b>	454
2.1 - Proteínas envolvidas no ciclo visual e cascatas	454
2.2 - Visão colorida nas células em cone	460
<b>3 - <u>Famílias de receptores da olfacção e vias de sinalização</u></b>	461
3.1 - Detecção de cheiros por receptores de sete hélices acoplados a proteínas G	461
3.2 - Detecção de cheiros pelas guanilil-ciclases	465
<b>4 - <u>Stress e resposta inflamatória. Quimocinas. Citocinas</u></b>	468
4.1 - Stress e outras agressões que induzem alterações metabólicas	468
4.2 - Resposta inflamatória	469
4.3 - Quimocinas e seus receptores	474
4.4 – Citocinas	482
<b>5 - <u>Necrose e apoptose</u></b>	495
5.1 - Sinalização por diversas vias	495
5.2 – Caspases	501
5.3 - Proteínas adaptadoras	502
5.4 - Família de receptores do factor de necrose tumoral (TNF-R ou receptores de morte)	504
5.5 - Família Bcl2	505
5.6 - Sinalização pelos receptores de morte CD95, CD120a e DR3	507
<b>6 - <u>Ciclo celular</u></b>	513
6.1-Ciclínas e CdK	513
6.2-Regulação e desregulação (oncogenese) do ciclo celular	518
<b>7 – <u>Oncogénese</u></b>	523
7.1 – Envolvimento de factores de transcrição na carcinogenese	523
7.2 – Oncogénese e suas proteínas, sua classificação e suas características	526
7.3 – Transdução de sinais mediados por oncogenes em células epiteliais mamárias de modelos transgênicos de rato	530
<b>Bibliografia</b>	533

## **1. - Ritmos Circadianos nos Mamíferos**

### **1.1 -Aspectos de biologia molecular dos ritmos circadianos nos mamíferos**

Os ritmos circadianos são as oscilações durante as 24 horas, de diversos processos biológicos ocorridos nos animais e que são controladas por relógios internos, (centrais ou periféricos). Estes ritmos produzidos internamente não funcionam isolados do seu meio envolvente sendo desencadeados por fontes externas sobretudo pela luz (pelos recorrentes amanhecer e anoitecer).

Estes ritmos circadianos têm três componentes a considerar ou sejam as vias de entrada ( input ), os relógios propriamente ditos (osciladores circadianos) que produzem os ritmos biológicos e as vias de saída ( output ) para o resto do organismo.

A base molecular dos osciladores circadianos nos animais é constituída por um conjunto de genes conservados ou sejam os Clock, Bmal 1, Period e Timeless que produzem um mecanismo autoregulador negativo retroinibidor de transcrição - translação, que contem os elementos essenciais que produzem os ritmos circadianos.

Os relógios circadianos resultam pois da expressão e actividade destes genes ( genes do relógio ) e dos seus produtos.

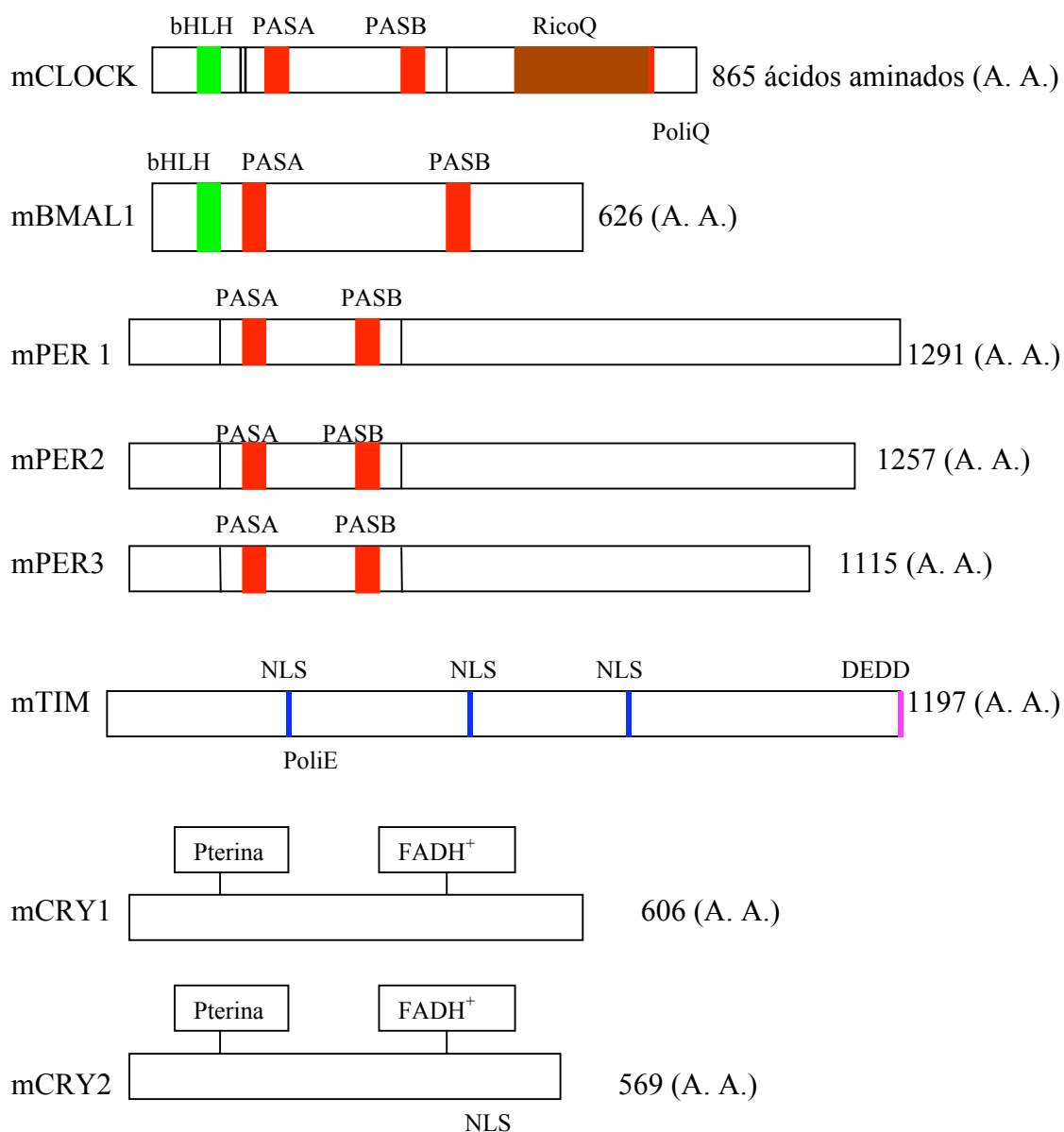
O relógio em si próprio consiste assim deste conjunto de genes que alternadamente activam e inibem a expressão de um subconjunto específico destes genes.

A expressão de um ou mais dos genes do “relógio” induz passado tempo e devido às actividades das proteínas expressas, a inibição da expressão desses mesmos genes, originando assim a renovação da transcrição dos genes do relógio oscilatórios. Isto decorre em cerca de 24 horas ou seja o ritmo circadiano.

A maquinaria do relógio circadiano em si é bem conhecida desde 1997, ao contrário do que sucede com os elementos envolvidos nos seus input e outputs.

Parece certo que oito genes estão implicados como componentes dos relógios circadianos nos mamíferos (ver figuras das páginas seguintes).

### Proteínas do relógio circadiano mamífero e seus domínios



m = Rato

Diagramas esquemáticos das oito proteínas componentes do relógio circadiano

bHLH = hélice-ansa-hélice básica

PAS = famílias com domínio PER-ARN-SIM

RicoQ = região rica em glutamina

PoliQ = região de repetição da glutamina

NLS = sinal de localização nuclear

PoliE = região de repetição do glutamato

DEDD = sequência conservada Asp-Glu-Asp-Asp

Pterina e FADH<sup>+</sup> são dois cromóforos/cofatores que se ligam ao CRY1 e CRY2 e possivelmente responsáveis pela mediação dos efeitos da luz sobre estas proteínas.

Contudo outros genes adicionais pontualmente envolvidos nas vias de output e input fotonico estão identificados nos relógios circadianos mamíferos.

Nos relógios circadianos ou osciladores os sistemas compostos pelas proteínas CLOCK , BMAL1 e CRY são habitualmente designados pela abreviatura CCO correspondendo ao oscilador circadiano central ou âmago (core circadian oscillator).

Existem espalhados pelos organismos animais múltiplos relógios ou osciladores independentes que são controlados por diferentes “inputs” e que são responsáveis pelo controlo de diversos “outputs”. Assim enquanto alguns destes relógios respondem à luz, outros são insensíveis à luz, outros respondem a variações alimentares ao contrário de outros que são insensíveis aos alimentos.

Há uma série de estímulos e vias de “entrada” que têm acesso aos relógios circadianos impondo uma oscilação diária da maioria das funções biológicas, inclusivé hormonas circulantes, funções cardíacas e circulatórias e temperatura do corpo.

As diversas fases metabólicas ocorridas nas células interrelacionam a percepção do ambiente com o aparelho circadiano oscilatório.

O núcleo supraquiasmático (SCN) do hipotálamo contem um pacemaker circadiano central sendo um regulador chave que controla directamente as oscilações fisiológicas registadas através dos organismos animais, nos osciladores periféricos, e na execução do comportamento rítmico.

Contudo diversos tecidos e diversos tipos de células utilizam relógios celulares oscilatórios independentes.

As células da maior parte dos tecidos estudados até agora, possuem este relógio CCO funcional que dirige a expressão circadiana dos genes Per e Cry assim como a de genes envolvidos na fisiologia circadiana específica dos tecidos, originando oscilações cíclicas destes relógios periféricos.

Este mecanismo central responsável pela regulação dos ritmos circadianos é um ciclo transcriptional de retroinibição negativa.

Na maioria dos tecidos a componente positiva da ansa de retroinibição CCO é constituída pelo heterodímero CLOCK:BMAL1.

Como já referimos o núcleo supraquiasmático (SCN) do hipotálamo, rapidamente desencadeia variações aos estímulos luz/escuridão ao contrário do que sucede com o CCO do fígado. Contudo a alimentação é um potente condicionador do ciclo circadiano hepático. Em outros órgãos periféricos estudados, rins, pâncreas, músculos do esqueleto e pulmões, verifica-se que cada um destes tecidos mantem um CCO funcional que condiciona a expressão rítmica dos genes circadianos como os Per e Cry.

No interrelacionamento entre relógios osciladores independentes parece admitir-se que o SCN, pacemaker chave nos ritmos circadianos, se interrelacionaria com esses órgãos periféricos através de um sinal difusível e independente do contacto directo célula-célula.

Refere-se mesmo que o ácido retinóico seria uma molécula com estas características, capaz de interactuar com receptores e proteínas CLOCK e outras proteínas.

Também a hormona adrenocorticotrófica (ACTH), a hormona libertadora da corticotrofina (CRH) parecem exhibir um perfil circadiano.

Os relógios circadianos controlam os níveis de diversos metabolitos celulares ou em circulação, inclusive os energéticos.

Como já referimos os relógios periféricos operando no coração, pulmões, músculos e outros órgãos, exibem um ritmo circadiano da expressão dos seus genes controlado pelo CCO.

No fígado está assinalado que um grande número de genes está sob controlo do CCO, inclusive genes codificadores de enzimas do metabolismo, inclusive citocromo P 450s, enzimas envolvidos na biossíntese do heme e no funcionamento das mitocôndrias.

No cérebro parece que a proporção dos cofactores NAD oxidados e reduzidos tem um papel central na regulação do CCO.

Tem sido proposto que os NADH e NADPH, estimulam a actividade de ligação ao DNA dos heterodímeros CLOCK:BMAL1 e NPAS2: BMAL1, enquanto as formas oxidadas dos dois cofactores NAD<sup>+</sup> e NADP<sup>+</sup> inibem a activação do factor de transcrição da CCO.

## **1.2 - Trama molecular do relógio circadiano mamífero**

Na figura seguinte desenha-se o relógio circadiano mamífero com os seus quatro componentes ângulo da ansa mecânica autoreguladora negativa retroinibidora de transcrição-translação.

Os dois componentes positivos da ansa de retroinibição são o gene Clock, e um outro gene o Bmal1 cada um dos quais contem a sequência proteica prevista denominadas CLOCK e BMAL1 respectivamente, ambas contendo domínios de dimerização, hélice-ansa-hélice básica (bHLH) e PAS (PER-ARNT-SIM).

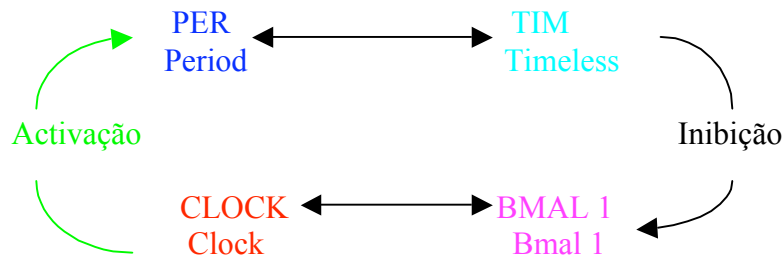
As proteínas codificadas pelos dois genes Clock e Bmal1, dimerizam-se por interacção dos domínios HLH e/ou PAS, ligam-se a sequências reguladoras do DNA chamadas caixas E (E-boxes) e activam a transcrição dos outros dois componentes ou sejam os genes period e timeless, que correspondem ao lado negativo da ansa retroinibidora.

No caso do period os mamíferos têm três parálogos denominados Per1, Per2 e Per3.

As proteínas expressas por estes genes as PERIOD e TIMELESS translocam-se para o núcleo após fosforilação de ambas as proteínas e sua heterodimerização. Uma vez chegadas ao núcleo condicionam negativamente a transcrição dos seus próprios genes que antes tinham sido activados pelo heterodímero CLOCK-BMAL1.

Passado tempo as proteínas PERIOD e TIMELESS sofrem turnover e a sua retroinibição negativa é reduzida e os heterodímeros CLOCK-BMAL1 começam a transcrição de novo o que constitui um ciclo circadiano do relógio. O tempo completo desta ansa de retroinibição é de cerca de 24 horas.

Foi possível identificar em diversos tecidos animais os perfis de expressão do mRNA dos genes dos relógios circadianos mamíferos sendo comprovada a importância fundamental do núcleo supraquiasmático (SCN) na condução da expressão dos ritmos circadianos.



Também o tecido da retina nos mamíferos contém um oscilador circadiano autônomo parecendo possuir todos os constituintes necessários para constituir um relógio circadiano.

O gene Clock codifica um novo membro da família dos factores de transcrição a hélice-ansa-hélice básica (bHLH) / PAS (PER-ARNT-SIM).

A proteína CLOCK do rato contém 865 ácidos aminados e domínios para ligação ao DNA e para interacção com outra proteína, e ainda um domínio C terminal rico em glutamina.

O mRNA para o gene Clock existe em grande quantidade no SCN e nos olhos, locais de residência de osciladores circadianos, embora o mRNA do Clock seja também expresso no resto do cérebro sobretudo no hipocampo, cerebelo e bulbo olfatório e ainda em múltiplos tecidos dos organismos tais como, coração, pulmão, fígado, rins, testículos e ovário.

Os parálogos period parecem altamente conservados, e as proteínas PERIOD contêm também domínios PAS. Também há indicações de que o mRNA do Per1 de rato é muito abundante no SCN e na retina e expresso ainda noutras regiões do cérebro de uma forma um tanto similar à que sucedia com a expressão do gene Clock, sendo expresso inclusivé no coração, pulmão, fígado, músculos esqueléticos, rins e testículos. As seqüências de ácidos aminados do PER1 e PER2 são muito similares e todas contêm domínios PAS.

Também a proteína expressa pelo gene Bmal 1 interacciona com a proteína CLOCK.

As proteínas bHLH-PAS funcionam tipicamente como factores de transcrição após heterodimerização. Há diversas proteínas interaccionantes com proteínas CLOCK duas destas são proteínas bHLH-PAS.

O Bmal 1 é expresso de uma forma semelhante à do gene Clock e Per1 de rato parecendo que a proteína BMAL 1 é a parceira normal para ligação à proteína CLOCK.

A timeless identificada nos mamíferos é o único membro deste ângulo da ansa retroinibidora cuja seqüência em ácidos aminados não inclui um domínio PAS, contendo no entanto outras seqüências que lhe permitem funcionar, e tem sido detectada no SCN, nos olhos, embora com níveis de expressão relativamente baixos.

Tem sido observada uma expressão dos mRNA que exhibe oscilação circadiana tal como o mesmo sucede relativamente à concentração das proteínas respectivas.

Assim no que se refere às oscilações dos mRNA destes genes as mais elevadas oscilações têm sido observadas nos period paralogos. O pico desta oscilação aparece cerca de seis horas após o nascer do sol, dentro do SCN, enquanto na retina ocorre algumas horas depois. Estas oscilações foram também observados noutros tecidos do cérebro e do corpo.

Também resultados preliminares das variações circadianas da expressão do gene clock e Bmal 1 sugerem que no rato as suas oscilações não são muito marcadas durante as 24 horas do dia, mas ao nível do SCN do rato parecem estas oscilações ocorrerem com intensidade.

A oscilação parece ser incompleta ao nível do mRNA do gene timeless.

Contudo nem todos os componentes do relógio circadiano necessitam oscilar para que o relógio funcione.

No entanto a interacção entre os genes implicados e as proteínas por eles expressas é essencial para que possa funcionar a ansa retroinibidora da transcrição-translação.

O Bmal 1 é um parceiro normal de ligação dos Clock. Os factores de transcrição bHLH ( inclusivé os que possuem domínios PAS ) ligam-se caracteristicamente a um palindromo de seis pares de bases da sequência do DNA chamada caixa E ( E-box ) cuja sequência consenso é CACGTG.

As proteínas CLOCK e BMAL 1 tem um efeito positivo sobre a transcrição do gene Per1 de rato.

### **1.3 - Vias de input fótico**

Parece que a luz afecta e constitui uma via de input do relógio circadiano principal no SCN sincronizando o relógio interno dos organismos com o meio envolvente.

Experiências realizadas (Foster 1998) incluindo a remoção dos olhos, levam a arritmidade circadiana. As vias precisas da fototransdução bem como as células envolvidas continuam no entanto por conhecer.

Parece que as vias de transdução fóticas que desencadeiam os ritmos circadianos devem ser distintas daquelas subjacentes à percepção visual da luz.

Inclusivamente foi possível verificar que a destruição de células em bastonete e em cone fotoreceptoras do olho não impede a fase desencadeadora do relógio circadiano nos mamíferos. Outros fotopigmentos, inclusivé os criptocromos de luz azul referidos abaixo podem desempenhar um papel nas respostas circadianas á luz.

Além dos seis genes (clock, Bmal 1, timeless e três period-paralogos) que constituem o centro do oscilador circadiano mamífero estão recentemente assinalados nos mamíferos genes criptocromos, o criptocromo 1, (Cry 1) e o criptocromo 2 (Cry 2) implicados no desencadear e na produção do ritmo circadiano em múltiplos organismos, inclusivé mamíferos (vide figura anterior).

Estes genes Cry embora homólogos das fotoliasas (enzimas de reparação do DNA activadas pela luz) parecem actuar nos mamíferos como fotoreceptores da luz azul, uma vez que não foi detectada actividade fotoliase nas proteínas expressas. Estes genes são expressos no rato com intensidade no SCN e



na retina, e também através de diversos tecidos do organismo sendo tentador admitir que possuem os fotorreceptores circadianos.

Logicamente as vias de input fóticas podem convergir sobre elementos do mecanismo do relógio circadiano, e hoje sabe-se que alguns componentes deste relógio são regulados pela luz.

#### **1.4 - Output do oscilador circadiano dos mamíferos**

A forma como o pacemaker circadiano contido no SCN transmite a informação circadiana ao longo do tempo para o resto do organismo constitui motivo de estudo intenso, tendo sido proposto que um oscilador circadiano baseado na transcrição-translação possa regular o output através da regulação da via circadiana nos seus genes integradores.

No entanto sabe-se que os osciladores periféricos requerem o input cíclico do relógio principal situado no SCN para manter ritmicidade. Genes locais que controlam o relógio circadiano localmente nos órgãos periféricos fornecem a flexibilidade necessária para controlar os ritmos circadianos presentes nestes tecidos.

#### **1.5 - Genoma e relógios circadianos**

As famílias de genes do relógio circadiano nos mamíferos constituem o âmago deste mecanismo circadiano e decorrem com vias retroactivas positivas (activação) e negativas (inibição) de transcrição / translação possuindo dois factores de transcrição contendo [(bHLH)/PAS] os CLOCK e BMAL 1 que interactuam ou emparelham através dos seus domínios PAS.

Estes domínios PAS encontram-se em diversas famílias de proteínas (PER, ARNT e SIM).

Os heterodímeros CLOCK-BMAL1 conduzem activando (vide figura seguinte) o ritmo da transcrição de três genes de período (mPer1 – mPer3 no rato) e dois genes criptocromo (mCry1 e mCry2) todos eles identificados nos seres humanos através do estudo das sequências dos respectivos “drafts”.

Há medida que as proteínas mPER e mCRY são traduzidas, elas formam complexos PER-CRY que são translocados do citoplasma para o núcleo. Aqui as proteínas mCRY actuam como reguladores negativos ao interactuarem directamente com CLOCK ou com BMAL1, ou com ambos, inibindo a transcrição que estes desencadeavam, formando uma via de feedback negativa (inibição).

Ao mesmo tempo mPER2 contribui para a transcrição de Bmal 1, a qual é ritmicamente expressa com uma fase em pico (oposta aquela do mPer/mCry) formando uma via de feedback positivo (activação).

A acção de activação / inactivação das vias de feedback positivo e negativo perpetuam a natureza sustida por ela própria, do relógio circadiano.

A fosfatação e a proteólise das proteínas dos relógios podem determinar a sua localização celular e estabilidade e são elementos importantíssimos na regulação dos tempos dentro do mecanismo molecular das 24 horas. A Ckle (uma caseína cinase I epsilon dos mamíferos) fosfata as PER influenciando o seu turnover.

As fosfatações são pois acontecimentos importantes para os relógios circadianos.

Concluindo o relógio circadiano central nos mamíferos é o núcleo supraquiasmático (SCN) do hipotálamo anterior, que orienta as variações diárias de diversos processos fisiológicos e comportamentais tais como o sono, ritmo de vigília e variações diárias da temperatura corporal, níveis hormonais, conhecimento e memória.

O alvorecer e o crepúsculo coordenam ou arrastam o relógio circadiano através de vias neurais conectando a retina com o SCN de forma que o relógio circadiano central e os seus ritmos de saída não decaiam ao longo das 24 horas, mas permaneçam alinhados com o dia solar.

Os mecanismos básicos do relógio circadiano dos mamíferos são pois compostos por ansas interactuantes positivamente e negativamente sobre a transcrição e translação.


Os componentes essenciais deste mecanismo conforme vimos atrás são os dois factores de transcrição CLOCK e BMAL1 contendo domínios hélice-ansa-hélice básica (bHLH) e PAS.

Os CLOCK e BMAL1 emparelham entre si por interacção dos respectivos domínios PAS.

Estes domínios PAS encontram-se numa série de famílias de proteínas (como por exemplo PER, ARNT e SIM) e são comuns aos factores de transcrição que são essenciais para a maquinaria dos relógios circadianos dos fungos, insectos e mamíferos.

Os heterodímeros CLOCK-BMAL1 conduzem a transcrição rítmica dos três genes Period (mPer1 – mPer2 no rato) e dos dois genes Cryptochrome (mCry1 e mCry2).


#### Legenda da figura seguinte

As proteínas CLOCK (C) e BMAL1 (B) dimerizam-se e ligam-se a caixas E (CACGTG) no DNA (  ) activando a transcrição dos genes period e timeless ou cry (ansa + na direita da figura). Os heterodímeros CLOCK-BMAL 1 regulam a transcrição rítmica de três genes Period (m Per1-m Per3) e de dois genes Cryptochrome (m Cry1 e m Cry2).

À medida que as proteínas P do mPER e do mCRY (C) são traduzidas, portanto

no citoplasma, elas formam complexos PER-CRY (C) P-2 (á esquerda na figura) que são

translocados para o núcleo e aqui as proteínas mCRY podem funcionar como uma regulação negativa na medida em elas interactuam directamente com CLOCK ou com BMAL 1 ou com ambas inibindo a transcrição da PER e da CRY.

As proteínas expressas PERIOD e TIMELESS ou CRY translocam-se para o núcleo após fosfatação (  ) e a sua heterodimerização, e como dissemos anteriormente, no núcleo inibem a sua própria transcrição interactuando com as proteínas CLOCK e BMAL ou ambas (ansa de feedback negativa em baixo na figura).

A fosfatação e a proteólise das proteínas do relógio circadiano são importantíssimas pois são responsáveis pela sua localização e estabilidade cada 24 horas.

Portanto estes genes period e timeless ou cry podem ser activados pelas proteínas CLOCK e BMAL 1 e inactivados pelas suas proteínas expressas PERIOD e TIMELESS ou CRY.

Mas estas proteínas PERIOD e TIMELESS ou CRY com a passagem do tempo sofrem o seu turnover o que diminui conseqüentemente o seu poder inibidor e por outro lado os heterodímeros CLOCK:BMAL1 começam a transcrição de novo. Ao mesmo tempo o mPER2 parece contribuir para a transcrição do Bmal 1 que é expresso ritmadamente com uma fase em pico oposta aquela do mPer/m Cry formando uma ansa de feedback positiva.

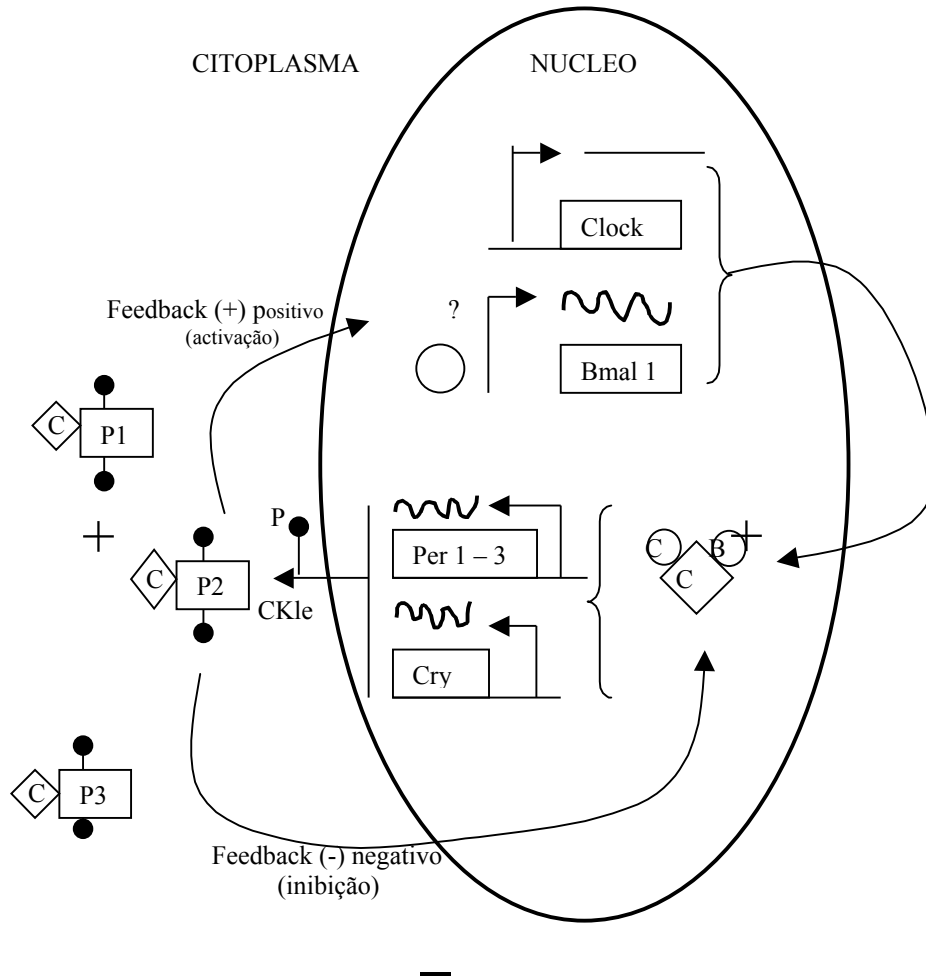
A PER1 pode transmitir a informação da luz às ansas moleculares, pois o seu RNA pode ser induzido pela luz ao amanhecer e ao anoitecer.

A PER3 pode contribuir para a transdução da oscilação para os sistemas de saída (output).

Todo este conjunto constitui o relógio circadiano e o tempo da ansa de retroinibição positiva (activação e em cima à esquerda na figura) e negativa ( inibição em baixo à direita na figura) é de cerca de 24 horas.

Este jogo da activação e inibição perpetua a natureza do relógio circadiano se sustentando a si próprio.

Modelo do trabalho do relógio circadiano dentro do SCN



CLOCK (C)

BMAL 1 (B)

CRY (C)

PER (P1-P3)

mRNA (wavy line)

Fosfatos (black dot)

Cinase = CKle

## 2 - Transdução de sinais visuais

### 2.1 – Proteínas envolvidas no ciclo visual e cascatas

Os sinais visuais para atingirem a retina, que contem o aparelho sensor visual, entram nos olhos passando progressivamente através da córnea, da câmara anterior do olho que contem o humor aquoso, cristalino e do corpo vítreo que consiste de humor vítreo.

São pois os olhos órgãos extraordinariamente complexos com o seu próprio metabolismo inerente ao acto visual.

A transdução destes sinais luminosos envolve a conversão de acontecimentos fotoquímicos em acontecimentos bioquímicos e eléctricos.

Esses sinais luminosos (fotões) penetrando no olho através do cristalino atravessam as fibras do nervo óptico, os neurónios ganglionares e os bipolares, e os núcleos de células em bastonete e em cone antes de atingirem o segmento mais externo dos bastonetes e cones onde começa a transdução de sinal. A este nível, nos segmentos mais externos das células em bastonete e em cone, os fotões são absorvidos por fotorreceptores desencadeando aqui um acontecimento bioquímico e que é a isomerização do pigmento visual retinal que passa da forma 11-cis para a forma all-trans. Daqui resulta que a outra parte do fotorreceptor, a parte proteica, modifica a sua conformação o que desencadeia uma mudança do potencial de membrana da célula ou seja ocorre a conversão do sinal luminoso (fotão) num acontecimento bioquímico (isomerização) seguido por um sinal eléctrico (potencial de membrana alterado) que é depois conduzido para o cérebro pelo nervo óptico.

Os fotorreceptores são pois pigmentos contidos nos discos achatados existentes no interior das células em bastonete e em cone, pigmentos estes que são heteroproteínas, a rodopsina nas células em bastonete e pigmentos vermelho, verde e azul nas células em cone. Na heteroproteína rodopsina, a opsina é o seu grupo prostético, e nas heteroproteínas que formam nas células em cone os pigmentos vermelho, verde e azul, as proteínas são diferentes umas das outras e diferentes também da opsina.

Um único fotão pode excitar uma célula em bastonete da retina, e o seu fotorreceptor a rodopsina que é um membro da família de receptores membranários de sete hélices, desencadeando a excitação visual pela fotoisomerização do 11-cis-retinal activando uma cascata de transdução de sinal em que intervem uma proteína G que induz a hidrólise de GMP cíclico. Esta última hidrólise vai desencadear o encerramento de canais específicos de cationes originando um sinal nervoso. O abaixamento da concentração de cálcio induzida pela luz coordena a recuperação e adaptação.

Vejamos em pormenor como tudo isto decorre ou melhor como uma cascata de transdução de sinal converte a luz num movimento atómico e depois num sinal nervoso.

Os dois tipos de células receptoras já referidos (bastonetes e em cone) funcionam nos vertebrados, as primeiras para a luz sem brilho não identificando as cores, e as segundas funcionando com a luz brilhante e identificando as cores.

Na retina existem milhões destes tipos de células formando sinapses com as células bipolares e estas conectando com outras células nervosas da retina.

Detenhamo-nos um pouco mais nas características destas células em bastonete.

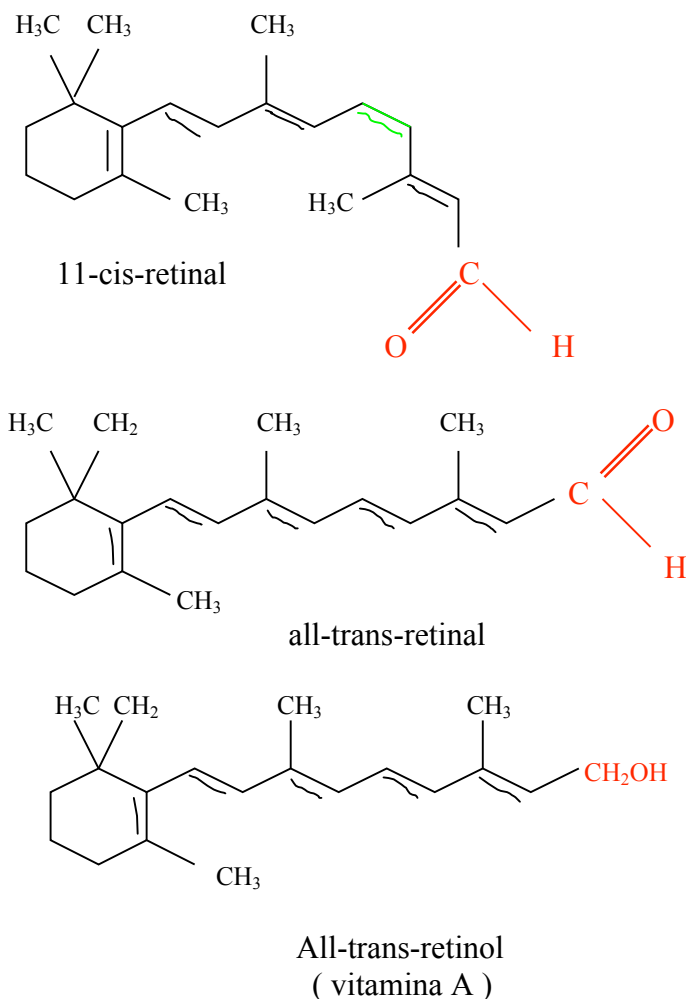
Revelam um alto grau de compartimentação e a sua porção mais externa é que está adaptada para a recepção da luz, existindo aqui cerca de 1.000 discos empilhados contendo as heteroproteínas fotorreceptoras e que estão continuamente a ser renovadas. Na porção mais interna destas células em bastonete, situa-se o núcleo e um corpo sináptico contendo varias vesículas cheias de transmissores. Na plasma membrana destas células existem canais específicos de catiões que se podem encontrar abertos (no escuro) ou fechados (acção da luz).

No escuro os canais abertos deixam passar rapidamente iões sódio para o segmento externo das células uma vez que o gradiente electroquímico é grande. Com efeito é criado pelas bombas  $\text{Na}^+\text{-K}^+\text{-ATPase}$  situadas no segmento interno das células em bastonete.

Quando da incidência da luz estes canais específicos dos catiões são encerrados (um simples fóton encerra centenas de canais), diminuindo a entrada de sódio e tornando a plasma membrana hiperpolarizada em cerca de 1mV (mais negativa no interior).

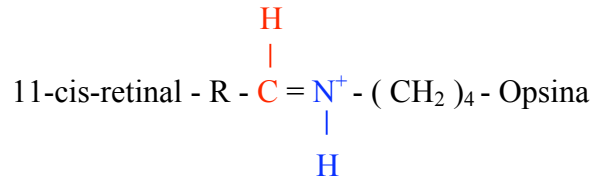
A luz ao desencadear esta hiperpolarização induz depois a transmissão desta hiperpolarização da plasma membrana do segmento externo para o corpo sináptico no segmento interno da célula em bastonete e conduzida para outros neurónios na retina.

Nos pigmentos fotorreceptores (nos cromóforos) contidos nos discos empilhados das células em bastonete a heteroproteína rodopsina, contem uma porção proteica a opsina, e um grupo prostético o 11-cis-retinal (vide figura seguinte). O precursor deste 11-cis-retinal é o all-trans-retinol (vitamina A) que não é biossintetizada pelos mamíferos, levando a sua falta à cegueira nocturna e eventualmente à deterioração do segmento externo das células em bastonete.



A cor da rodopsina e a sua capacidade de resposta à luz depende da presença de 11-cis-retinal.

O 11-cis-retinal está ligado à proteína opsina através de uma ligação em base de Schiff (o grupo aldeído do 11-cis-retinal liga-se covalentemente a grupo  $\epsilon$ -aminado de um resíduo de lisina 296 da opsina na sétima hélice).



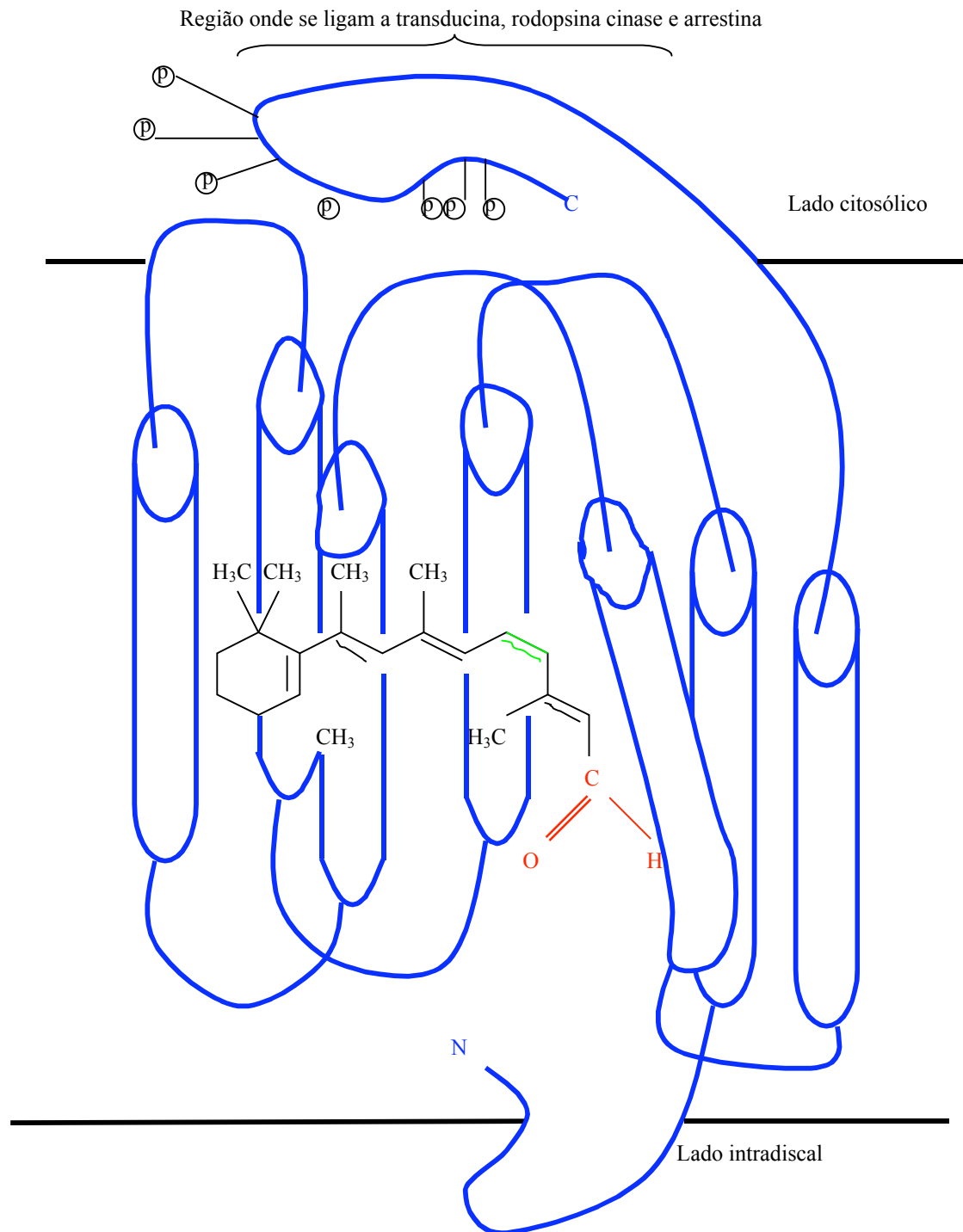
Como já tivemos oportunidade de referir a rodopsina é uma heteroproteína transmembranária com sete hélices que se encontra representada na figura seguinte.

Sendo uma proteína de 40 kDa, apresenta do lado citosólico da membrana dos discos uma extremidade C carboxílica e a outra extremidade N-terminal encontra-se no interior do disco (lado intradiscal). O grupo prostético da rodopsina ou seja o 11-cis-retinal situa-se numa “bolsa” que se encontra a meio da membrana entre o lado citosólico e o lado intradiscal.

O sinal é transmitido para uma cascata enzimática a partir deste lado citosólico que é a região onde interactivam, a transducina, a rodopsina cinase e a arrestina.

Salienta-se que esta molécula receptora de sete hélices a rodopsina transmembranária possui na sua extremidade carboxílica (lado citosólico) resíduos de serinas e treoninas que quando fosfatadas correspondem ao estado desactivado deste receptor.

## Estrutura da rodopsina dos vertebrados



Conforme vimos anteriormente a excitação visual promove a isomerização do 11-cis-retinal em all-trans-retinal (vide figura atrás) e altera por outro lado a conformação da cadeia polipeptídica (opsina) correspondente tudo isto em pico-segundos, e devido á deslocação da base de Schiff (entre o grupo aldeido do 11-cis-retinal e o grupo -aminado da lisina na posição 296 da sétima hélice da opsina). Esta deslocação da base de Schiff de cerca de 5Å da estrutura cíclica da extremidade corresponde a conversão da energia do fóton em movimento de átomos.



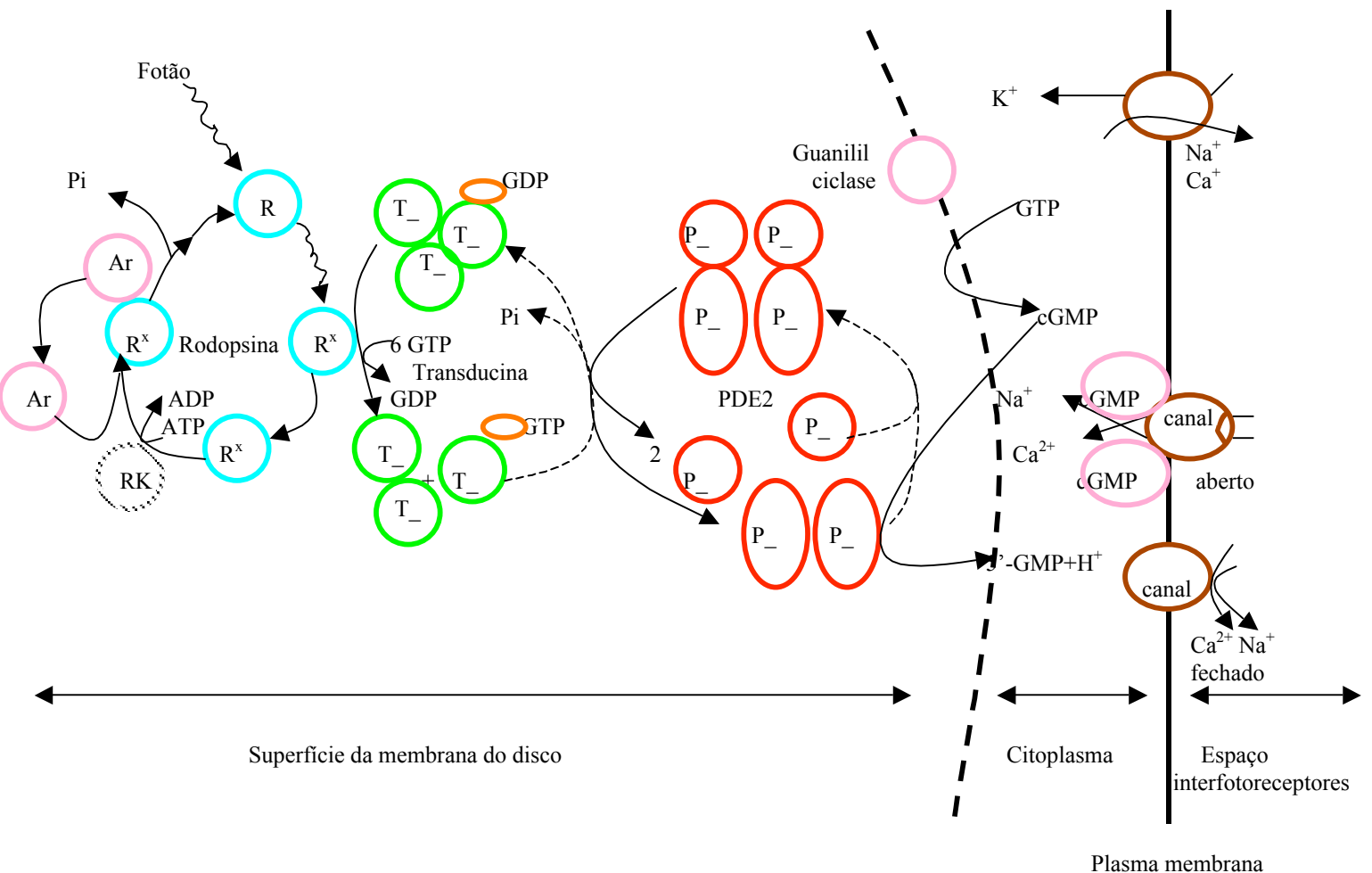
Estas modificações da conformação do retinal e da opsina progridem gerando intermediários fotolíticos da fotólise da rodopsina, que acaba por se dissociar dando opsina e all-trans-retinal, embora essa rodopsina depois acabe por ser regenerada.

Neste intermediários fotolíticos, a rodopsina fotoexcitada ( $R^*$ ), vide figura adiante, induz a activação de uma cascata em que intervem uma proteína G e que conduz à hidrólise do GMP cíclico, do que resulta o encerramento dos canais específicos para os catiões e hiperpolarização.

Está referido que numa célula em bastonete adaptada ao escuro, a absorção de um simples fóton bloqueia o fluxo de mais de um milhão de iões sódio.

Conforme se pode ver na figura seguinte ocorrem três ciclos entre a conversão da energia luminosa e os impulsos nervosos.

### Cascata de reacções bioquímicas no ciclo visual



Um dos ciclos corresponde às reacções envolvendo a rodopsina, outro ciclo corresponde às reacções envolvendo a transducina, e o último ciclo envolve a fosfodiesterase (PDE), resultando da interligação entre estes ciclos a hiperpolarização da plasma membrana das células em bastonete ou das células em cone.

No ciclo da esquerda da figura referida esquematiza-se a activação e inactivação da rodopsina.

No ciclo envolvendo a transducina que é uma proteína G clássica trimérica (constituída pelas subunidades  $T_{\alpha}$ ,  $T_{\beta}$  e  $T_{\gamma}$ ) que interacciona com a rodopsina do primeiro ciclo formando um complexo que gera um sinal que vai a fosfodiesterase do 3º ciclo, e esta fosfodiesterase hidroliza o GMP cíclico (à direita na figura) o que diminui o número de canais catiónicos abertos, tornando o potencial de membrana mais negativo promovendo portanto a hiperpolarização da plasma membrana (à direita na figura).

Vejamos em mais pomenor esta cascata de reacções bioquímicas no ciclo visual servindo-nos da figura em apreço.

No escuro os canais específicos dos catiões da plasma membrana (à direita na figura) são mantidos parcialmente abertos pelo cGMP formado pela acção da guanilil ciclase. A concentração deste GMP cíclico pode ser afectada de diversas formas dentro das células em bastonete ou em cone condicionando assim o número de canais de  $Na^{+}$  que estão abertos e por tal o potencial de membrana.

A rodopsina ( $R$ ) activada pela energia fotónica da luz ou seja  $R^{*}$  sobretudo o intermediário da fotólise metarodopsina II forma um complexo com a transducina do segundo ciclo. Este complexo  $R^{*}$ -transducina, transducina formada pelas sub-unidades  $\alpha$ ,  $\beta$  e  $\gamma$ , sendo portanto uma proteína G trimérica, altera a sua conformação o que possibilita a troca de GDP ligado por GTP, e quando ocorre esta troca a subunidade  $\alpha$  da transducina dissocia-se das duas subunidades  $\beta$  e  $\gamma$  da transducina o que permite que a sub-unidade  $\alpha$  ( $T_{\alpha}$ ) interacciona com a fosfodiesterase ( $PDE$ ) do chamado por nós 3º ciclo. Esta PDE nas células em bastonete é composta por quatro subunidades proteicas (é pois um heterotetramero), uma subunidade  $\alpha$  ( $P_{\alpha}$ ), outra  $\beta$  ( $P_{\beta}$ ) ambas as subunidades catalíticas, e duas subunidades  $\gamma$  ( $P_{\gamma}$ ) reguladoras.

Quando a subunidade  $T_{\alpha}$  interacciona com a PDE activa esta o que promove é a hidrólise do GMP cíclico (cGMP) dando 5'-GMP.

Esta diminuição da concentração de cGMP diminui também o número de canais abertos (à direita e em baixo na figura) o que como já referimos torna mais negativo o potencial de membrana que se hiperpolariza.

A activação da PDE consiste portanto na dissociação das suas subunidades catalíticas ( $P_{\alpha}$  e  $P_{\beta}$ ) das suas subunidades reguladoras ( $2 P_{\gamma}$ ).

Por outro lado e voltando um pouco atrás a  $T_{\alpha}$  tem actividade GTPase hidrolizando portanto o GTP que lhe está ligado originando GDP e fosfato inorgânico e daí a dissociação das sub-unidades  $T_{\alpha}$  (da transducina) das subunidades reguladoras  $\gamma$  da PDE o que vai permitir que estas se reassociem com as subunidades  $P_{\alpha}$  e  $P_{\beta}$  da PDE (subunidades catalíticas) inibindo assim a actividade da PDE.

Nas células em cone a PDE é composta por duas subunidades  $\alpha$  em vez da  $\alpha$  e  $\beta$  das células em bastonete, mas as cascatas de reacções bioquímicas no ciclo visual é idêntica nos dois tipos de células.

A concentração intracelular de  $Ca^{2+}$  condiciona as concentrações do cGMP.

Nas células em bastonete e no escuro o cálcio entra nestas células pelos canais de sódio (à direita na figura) atingindo valores de 500 nM o que permite uma actividade da guanilato ciclase baixa.

Quando os canais de  $\text{Ca}^{2+}$  são encerrados e deixa de entrar cálcio nas células os permutadores de cálcio-sódio ( à direita na figura) continuando a funcionar levam a uma diminuição da concentração intracelular de  $\text{Ca}^{2+}$  o que induz a activação da guanilato ciclase com o aumento subsequente de c GMP. Na paragem das reacções do ciclo visual são muito importantes a ressíntese de cGMP e a hidrólise do  $\text{T}_\alpha\text{-GTP}$ , tal como a inactivação da rodopsina  $\text{R}^x$  activada.

A rodopsina activada  $\text{R}^x$  é fosfatada por uma cinase específica (RK) e a rodopsina fosfatada tem alta afinidade para uma proteína citosólica, a arrestina (AR) com a qual forma um complexo incapaz de interactuar com a transducina, podendo parar a cascata de reacções.

Resumem-se no quadro seguinte as principais proteínas envolvidas no ciclo visual:

Proteína	Relação com a membrana	Massa molecular	Concentração citoplasmica ( M)
Rodopsina	Intrínseca	39	
Transducina ( + + )	Periférica ou solúvel	80	500
Fosfodiesterase	Periférica	200	150
Rodopsina cinase	Solúvel	65	5
Arrestina	Solúvel	48	500
Guanilato ciclase	Ligada ao citoesqueleto	?	?
Canal activado pelo cGMP	Intrínseca	66	?

## 2.2-Visão colorida nas células em cone

A capacidade para distinguir as cores reside nas células em cone que existem em muito menor número do que as células em bastonete, nos seres humanos. Os cães têm ainda menos células em cone e as aves têm muito mais.

Consoante o tipo de pigmentos contidos pelas células, azul, verde ou vermelho assim os tipos de células em cone. Normalmente em cada célula há apenas um tipo de pigmento.

Tal como nas células em bastonete cada um destes pigmentos das células em cone possuem proteínas fotorreceptoras de sete hélices com o grupo prostético 11-cis -retinal que se isomeriza em all-trans-retinal por activação pela luz.

Cores diferentes daquelas dos pigmentos visuais são distinguidas por estimulação graduada dos diferentes cones e análise comparativa pelo cérebro. A visão cromática é tricromática.

Os mecanismos gerais através dos quais a luz estimulam as células em cone são exactamente iguais aos promovidos nas células em bastonete.

Nos pigmentos fotorreceptores das células em cone o grupo prostético também se liga através de uma base de Schiff protonada com as proteínas, e os ácidos aminados que se dispõem à volta deste local de ligação influenciam o nível de energia e dão diferentes espectros de absorção consoante os pigmentos.

As sequências de ácidos aminados de algumas destas proteínas dos pigmentos são bem conhecidas havendo graus de homologia variável entre elas.

Os genes responsáveis por estes pigmentos visuais estão mapeados nos seres humanos, o gene da rodopsina por exemplo localiza-se no cromossoma 3, o gene responsável pelo pigmento azul no cromossoma 7 e os genes para o vermelho e para o verde no cromossoma X, podendo mutações nestes genes originar visão anormal das cores.

Salienta-se também que nas células em bastonete a absorção de um simples fóton origina resposta de corrente cem vezes mais intensa do que nas células em cone, mas o tempo de resposta é quatro vezes mais rápido nas células em cone.

### **3-Famílias de receptores da olfação e vias de sinalização**

#### Receptores olfactivos e sua sinalização

Os mamíferos podem discriminar mais de dez mil cheiros distintos o que parece devido à sua capacidade para que os seus neurónios olfactivos possam exprimir um grande número de receptores de sete hélices que não se encontram organizados espacialmente no seu epitélio olfactivo mas sim como projecções dos axónios formando distintos modelos dentro do bulbo olfatório.

Estes receptores de sete hélices são como que proteínas Gs (Golf) que activam uma isoforma específica de adenil ciclase (tipo III) que desencadeia a subida da concentração de cAMP que vai activar canais franqueados ou accionados por nucleótidos cíclicos, podendo o canal ser também accionado por cGMP.

No entanto estão assinalados em roedores uma outra grande família de guanilil ciclases neuronais específicas sediadas na plasma membrana dos cílios dendríticos contendo um domínio extracelular com características estruturais de receptores, constituindo uma segunda família de receptores de cheiros e de feromonas.

Nos roedores o epitélio olfactivo contém mais de 1000 receptores de cheiros sendo a maior família de receptores existente nestes animais.

A percepção olfactiva é desencadeada quando um cheiro se liga ao receptor acoplado à proteína G desencadeando uma cascata que origina cAMP acabando com uma despolarização do neurónio respectivo.

Outra via para percepção olfactiva pode ser a indução da elevação de cGMP em vez de cAMP.

#### **3.1 - Deteção de cheiros por receptores de sete hélices acoplados a proteínas G**

Esta família de receptores está identificada em ratos, humanos, peixes, Drosophila melanogaster, etc..

Neurónios olfactivos individuais são distintos, cada neurónio exprimindo um único gene para um receptor olfactivo, com uma região transmembranária altamente variável que é responsável pela ligação

ao cheiro podendo a simples substituição de um único ácido aminado nesta região alterar a especificidade de ligação à molécula odorífica.

No entanto moléculas de receptores olfactivos individuais podem reconhecer múltiplos cheiros com comprimentos de cadeias similares (diferindo em apenas dois ou quatro carbonos) e com similares grupos alifáticos na “cabeça”.

Também está assinalado que alguns cheiros são reconhecidos por vários receptores havendo uma correlação positiva entre o comprimento da cadeia da molécula do cheiro e o número de receptores que ele estimula.

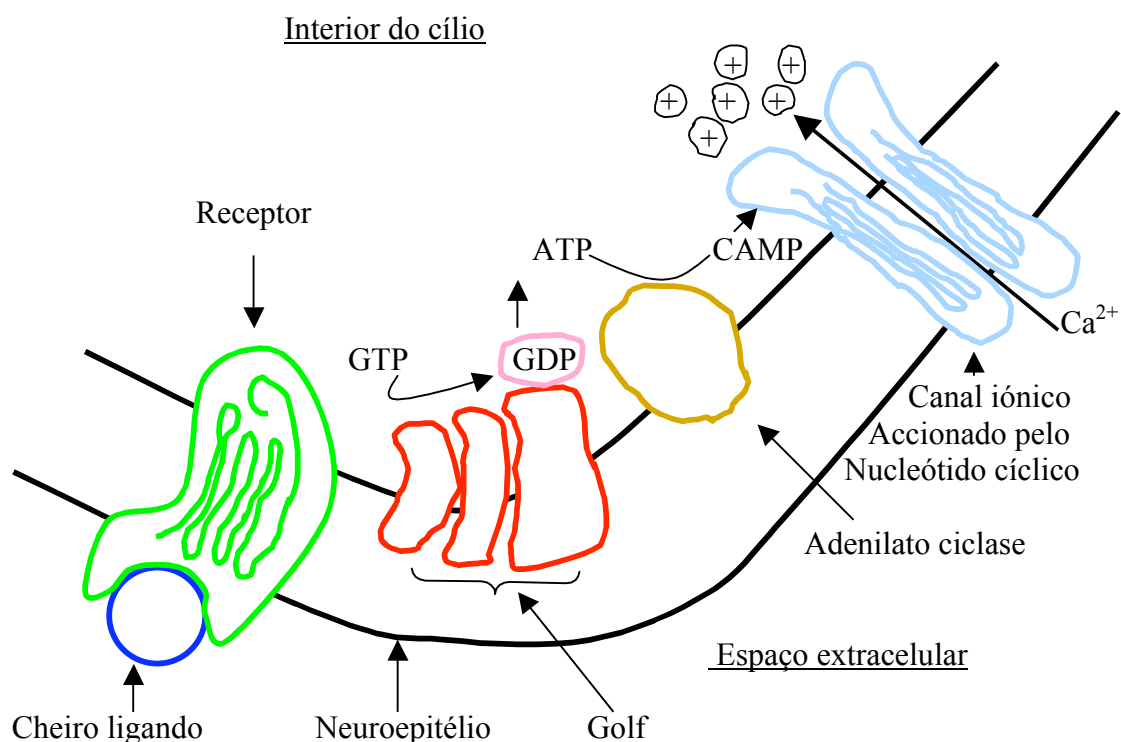
No órgão vomeronasal está assinalado um conjunto único de receptores de sete hélices que detectam feromonas que regulam o comportamento sexual parecendo que cada neurónio sensor neste órgão exprime apenas um receptor para feromona.

Nas vias de sinalização os cheiros extracelulares ligam-se a um receptor específico no neuroepitélio e o respectivo neurónio despolariza-se. Na G proteína heterotrimérica (Golf) ocorre a troca de GTP pelo GDP na subunidade  $\alpha$  que é semelhante à proteína estimuladora Gs da adenilil ciclase e esta forma da Golf com GTP ligado activa a adenilil ciclase nos neurónios olfactivos o que parece ser uma etapa que ocorre em todos os neurónios olfactivos.

Encontram-se nove isoformas da adenilil ciclase, sendo o tipo III abundantemente expresso no tecido olfatório. Os canais de catiões franqueados por nucleótidos cíclicos foram primeiro identificados nas células em bastonete e em cone da retina do olho, sendo depois identificados nos neurónios olfatórios.

Estes canais accionados por nucleótidos cíclicos são oligómeros compostos por sub-unidades  $\alpha$  e  $\beta$  e cada sub-unidade consistindo de seis segmentos transmembranários, com um domínio N-terminal muito variável e com um domínio C-terminal de 130 ácidos aminados homólogo aos domínios da proteína cinase A e G que ligam nucleótidos cíclicos.

Transdução de sinal num neurónio olfatório  
utilizando receptor transmembranário de sete hélices



Estes canais na forma aberta não são selectivos e deixam passar catiões inclusivé  $\text{Na}^+$  e  $\text{Ca}^{2+}$ , abrindo e fechando espontaneamente estes canais na ausência de nucleótidos cíclicos, mas estes têm maior afinidade para os canais abertos.

Quando os cheiros induzem a elevação das concentrações de cAMP intracelular o equilíbrio no canal (forma aberta, forma fechada) desloca-se para a forma aberta permitindo a entrada de  $\text{Ca}^{2+}$  e outros catiões despolarizando a conductancia através da membrana celular e transmitindo um sinal eléctrico para o bulbo olfativo.

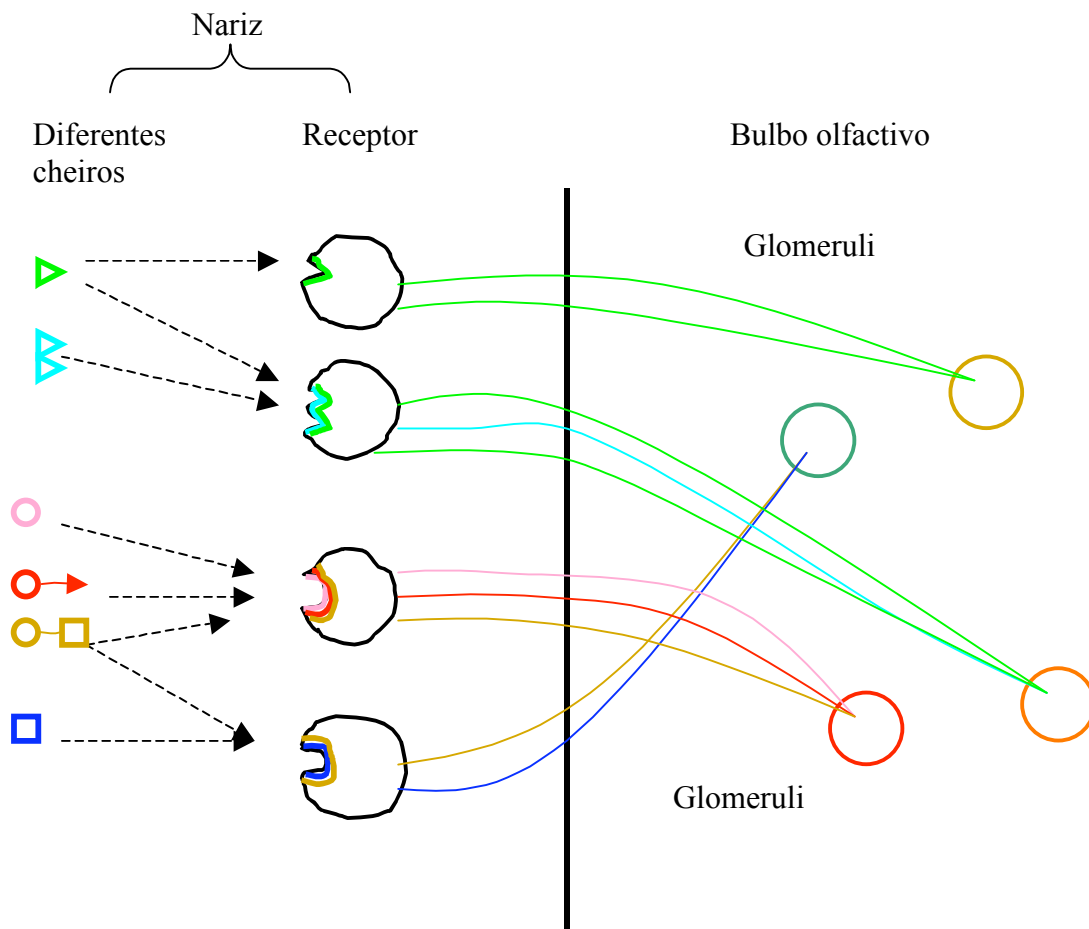
Há alguns investigadores que referem que alguns cheiros activam uma via fosfoinositideo nos neurónios olfativos, pois alguns cheiros aumentam extraordinariamente a concentração de cAMP nos cílios olfatórios, enquanto outros cheiros falham nestes aspectos.

Parece que a formação enzimática de  $\text{IP}_3$  induzida pelos cheiros é rápida e transitória não se acompanhando da subida de cAMP.

Admite-se pois que possam existir vias assinalantes de cheiros sem cAMP.

Os receptores de cheiros individuais são expressos em apenas 0,1% dos neurónios olfactivos. Estas moléculas receptoras são expressas dentro de quatro zonas topográficas amplas no epitélio nasal, havendo pontos de convergência de neurónios, ou glomérulos, nos neurónios sensores nasais o que permite formar um mapa topográfico dos receptores de cheiro. (vide figura seguinte).

### Modelo para mapa combinatório de cheiros no bulbo olfactivo



#### Legenda da figura

Sinais específicos mandados aos glomeruli dentro do bulbo olfactivo quando um cheiro se liga a um receptor.

Como todos os neurónios que exprimem um dado receptor convergem para um ou dois glomeruli específicos e os receptores podem ligar moléculas cheirosas com estruturas diferentes mas relacionadas, os mesmos glomeruli podem ser activados por diferentes cheiros.

Também cada cheiro liga-se a diferentes combinações de receptores e depois resulta um perfil diferente de glomeruli activados dentro do bulbo olfactivo.

Qualquer molécula odorífica individualmente pode ligar-se a vários receptores distintos e “activar” assim um diferente conjunto de glomeruli no bulbo olfáctico. Este conjunto combinatório constitui provavelmente a base para reconhecimento dos cheiros e memória olfatória.

A discriminação dos cheiros pode ser explicada através da combinação de receptores tal como se indica na figura antes referida.

### 3.2 - Detecção de cheiros pelas guanilil ciclases

Apesar da maioria dos cheiros serem reconhecidos por receptores de sete hélices, há estímulos olfatórios que activam vias alternativas a partir de uma família de receptores guanilil ciclase que podem modelar as concentrações de cGMP em diversos tipos de células. Esta actividade guanilil ciclase tem sede em formas moleculares solúveis e em formas ligadas a membranas.

As guanilil ciclase solúveis (vide figura seguinte) nos mamíferos têm quatro sub-unidades, duas subunidades  $\alpha$  ( $\alpha_1$  e  $\alpha_2$ ) e duas subunidades  $\beta$  ( $\beta_1$  e  $\beta_2$ ) encontrando-se normalmente na forma de heterodímeros de subunidades  $\alpha$  e  $\beta$  formando domínios catalíticos e domínios reguladores. As regiões C-terminais das subunidades  $\alpha$  e  $\beta$  são homólogas do domínio catalítico da adenilil ciclase e a especificidade da guanilil ciclase pode ser mudada para ATP em vez de GTP por mudança de apenas três ácidos aminados que se encontram invariavelmente no seu domínio catalítico.

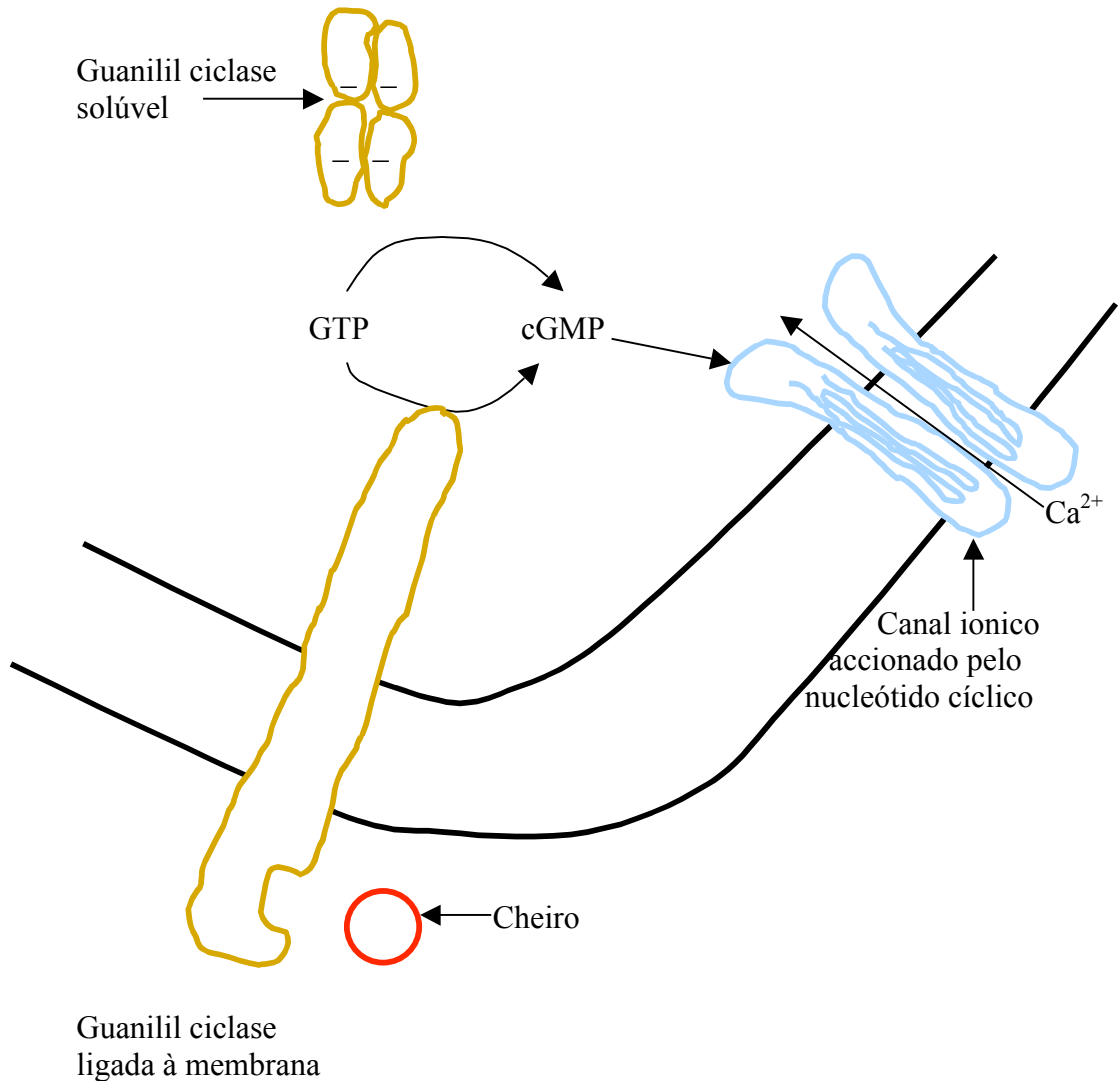
As guanilil ciclase ligadas a membranas são (vide figura seguinte) constituídas por uma única cadeia polipeptídica de cerca de mil e cem ácidos aminados, com um domínio extra-celular onde se liga o cheiro, um segmento transmembranário, e um domínio intracelular “like cinase”, e um domínio C-terminal com actividade ciclase. Os domínios extra-celulares variam um tanto ao contrário dos outros segmentos destes receptores. Diversos receptores guanilil ciclase estão identificados nos mamíferos. Designados pela ordem da sua descoberta (de GC-A até GC-G) estes receptores parecem ligar um tipo de ligando cada, servindo para diferentes funções em diversos tecidos (vide quadro seguinte).

Distribuição e função de guanilil ciclases mamíferas ligadas a membranas

<u>Receptor</u>	<u>Ligando</u>	<u>Distribuição pelos tecidos</u>	<u>Função</u>
GC-A	ANP, BNP	Cérebro, rim supra-renais tecido adiposo ileum terminal	Medeia natriurese/diurese ou é o receptor primário no rim para substancia libertada pelo coração para regular o volume de sangue
GC-B	CNP	Cérebro, placenta, hipófise, ileum terminal	Regula o crescimento celular e pressão sanguínea
GC-C	STa guanilina	Intestino, epitélio olfatório, supra-renais, cérebro, testículos, placenta e fígado	Serve como receptor intestinal para um peptideo enterotoxico e bacteriano
GC-D	?	Tecido olfatório	Potencial receptor quimo sensor
GC-E	?	Células fotoreceptoras glândula pineal	Restaura os níveis de cGMP após fotoestimulação
GC-F	?	Células fotoreceptoras	Restaura os níveis de cGMP após fotoestimulação
GC-G	?	Pulmão, intestino, músculo esqueleto	



Transdução de sinal num neurónio olfactório utilizando receptor guanilil ciclase ligado a membrana ou guanilil ciclase solúvel



O receptor GC-A é um receptor importante dos rins e dos vasos do tecido muscular liso para péptidos que controlam o volume de sangue.

O GC-B é um receptor fisiológico para o péptido natriurético do tipo C.

Os receptores GC-D, GC-E e GC-F, são expressos em neurónios sensitivos, os primeiros (GC-D) expressos no neuroepitélio do nariz e o GC-E e GC-F expressos nas células fotorreceptoras da retina de humanos, rato, galinhas e vacas.

O GC-D é expresso sobretudo nos cílios neurónios olfactivos.

As células fotorreceptoras e neurónios sensitivos do nariz diferem no facto da fotoestimulação desencadear uma hiperpolarização das células em bastonete e em cone na retina, enquanto os cheiros induzem a despolarização dos neurónios olfactivos.

No entanto os mecanismos de sinalização são similares envolvendo G-proteínas, receptores de sete hélices, e canais franqueados por nucleótidos cíclicos.

Os canais olfactivos accionados por nucleótidos cíclicos têm uma alta afinidade para o cGMP e para o cAMP e em algumas espécies, canais franqueados por nucleótidos cíclicos ligam o cGMP com maior afinidade.

Ocorre um persistente influxo de cálcio nos neurónios olfactivos quando se encontra presente cGMP.

Os canais franqueados por nucleótidos cíclicos são altamente sensíveis a variações na concentração de cGMP, e a regulação das guanilil ciclases nos neurónios olfatórios é muito importante.

Nestes neurónios olfatórios também está assinalada a presença de uma fosfodiesterase activada pelo cGMP, ( expressa com os receptores GC-D num subconjunto de neurónios olfactivos), embora com localização celular diferente (a fosfodiesterase ao longo de todo o neurónio e o receptor GC-D localizado sobretudo nos cílios dendríticos). Existe ainda nestes neurónios olfactivos uma outra molécula que é uma proteína cinase dependente do cGMP cuja activação tem um papel por ora ainda não esclarecido.

Na regulação das guanilil ciclases ligadas a membranas dos neurónios olfactivos os domínios extracelulares destas guanilil ciclases (GC-D, -E, -F) têm um alto grau de idendidades na sua estrutura primária mantendo resíduos de cisteína altamente conservados nesta zona, parecendo que as pontes dissulfureto formadas por elas são fundamentais para a interacção com o ligando.

Na regulação das guanilil ciclases solúveis nos neurónios olfactivos, foi observado que o cGMP ou activadores das guanilil ciclase solúveis podem diminuir a sensibilidade dos neurónios olfatórios para os cheiros de uma forma que depende do cálcio.

As guanilil ciclases ligadas a membranas são directamente activadas pela ligação com cheiros. Pelo contrário as guanilil ciclases solúveis geram uma corrente “steady state” variável que contribui para a adaptação ao cheiro é a dessensibilização.

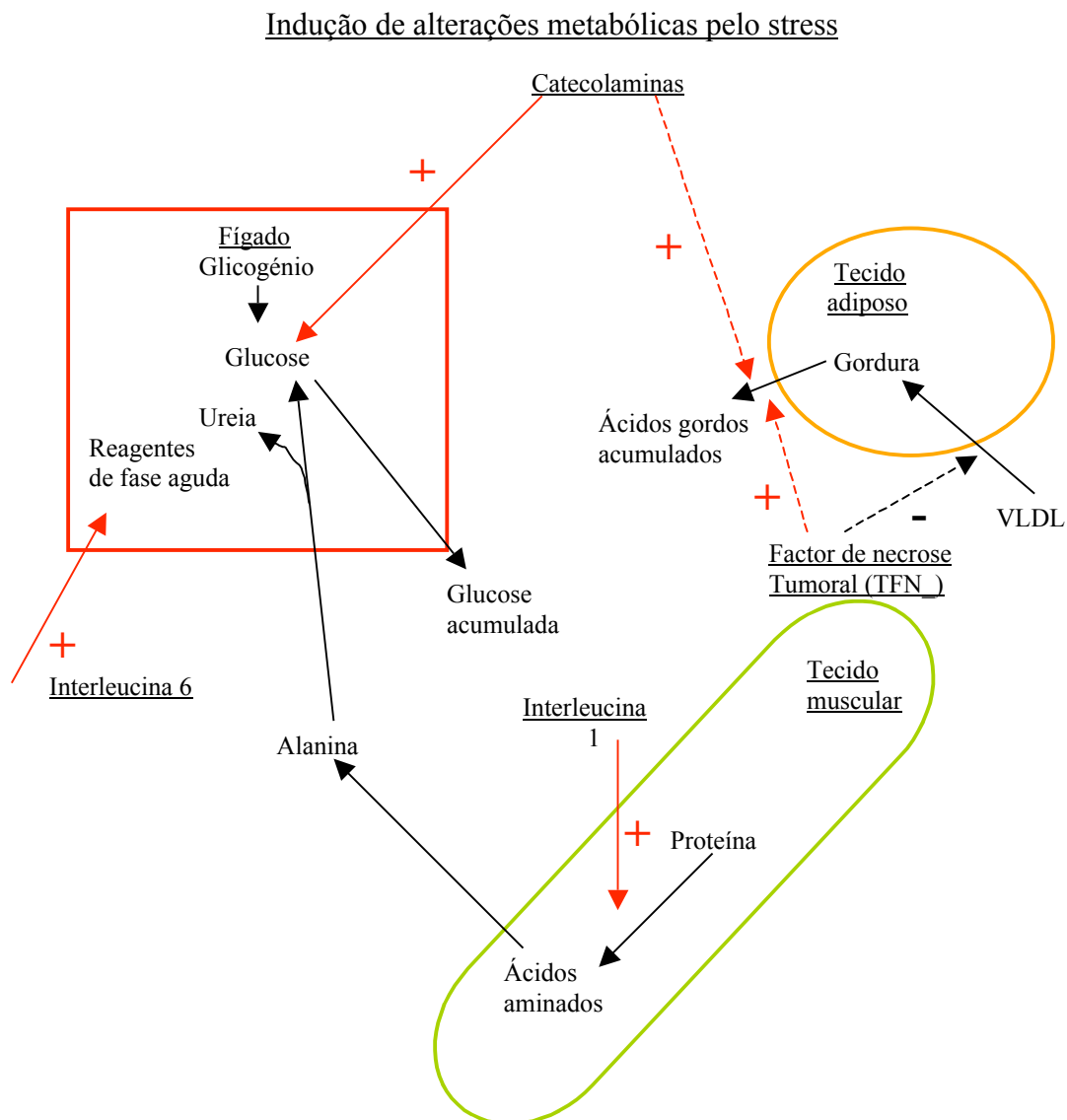
Em conclusão a interacção dos cheiros com diferentes classes de receptores (clássicos receptores de sete hélices transmembranárias ou classe de guanilil ciclases) podem iniciar vias assinalantes quer cGMP quer cAMP que convergem em canais olfatórios franqueados por nucleótidos cíclicos para produzir potenciais de acção.

## 4 - Stress. Resposta inflamatória. Quimocinas. Citocinas

### 4.1 - Stress e outras agressões como factores que induzem alterações metabólicas

O stress fisiológico, as agressões cirúrgicas, as falhas renais, queimaduras, e infecções induzem a subida no sangue do cortisol, glucagina, catecolaminas e hormona do crescimento, observando-se ainda nessas situações resistência à insulina e elevadas taxas de glicémia e de ácidos gordos livres no sangue, bem como reduzida síntese proteica e acentuado aumento da quebra das proteínas.

Este balanço azotado negativo é mediado por proteínas dos monocitos e linfócitos, tais como as interleucinas 1 e 6 (IL1 e 6), e pelo factor de necrose tumoral (TNF  $\alpha$ ). Estas citocinas originam febre e uma série de outras alterações metabólicas. A interleucina 1 activa a proteólise nos músculos do esqueleto e a interleucina 6 estimula a síntese pelo fígado das chamadas proteínas reagentes de fase aguda onde se inclui o fibrinogénio, proteínas do complemento, alguns factores de coagulação como a  $\alpha_2$ -macroglobulina, etc. (vide figura anexa).



O TNF $\alpha$  suprime a biossíntese da gordura no adipócito e impede a tomada da gordura circulante pela inibição da lipoproteína lipase ao mesmo tempo que estimula a lipólise e inibe a libertação da insulina.

Estas citocinas parecem ainda ser a causa dos muitos danos promovidos pelas infecções crónicas.

Alguns tumores sintetizam e segregam péptidos activos como o ACTH, factor do crescimento dos nervos, factores de crescimento “insulina – like”, que podem modificar a regulação endócrina do metabolismo energético. Por outro lado o organismo pode responder ao tumor libertando interleucinas –1 e –6 e TNF $\alpha$  pelas células do sistema imunitário.

O TNF $\alpha$  – também chamado caquexina porque produz desgaste orgânico pode actuar de uma forma paracrina tal como a IL-1 estimulando febre, proteólises, lipólise e síntese de proteínas reagentes de fase aguda pelo fígado.

#### **4.2 – Resposta inflamatória**

Uma infecção local ou uma injúria em qualquer tecido, rapidamente atrai glóbulos brancos sanguíneos para a região afectada o que constitui parte da resposta inflamatória, que ajuda a combater a situação.

A inflamação é um conjunto de reacções que ocorrem após uma série de injúrias tissulares, ou infecção ou estimulações imunológicas, que visam defender os organismos contra substâncias endógenas alteradas ou contra substâncias estranhas, sendo essas reacções de natureza bioquímica e celular e tendo uma intensidade que depende da extensão do trauma inicial.

A reparação desse trauma inicial é um processo complexo e prolongado de reparação e remodelação dos tecidos injuriados.

Em esquema pode representar-se este processo de reparação e remodelação da forma referida no esquema seguinte.

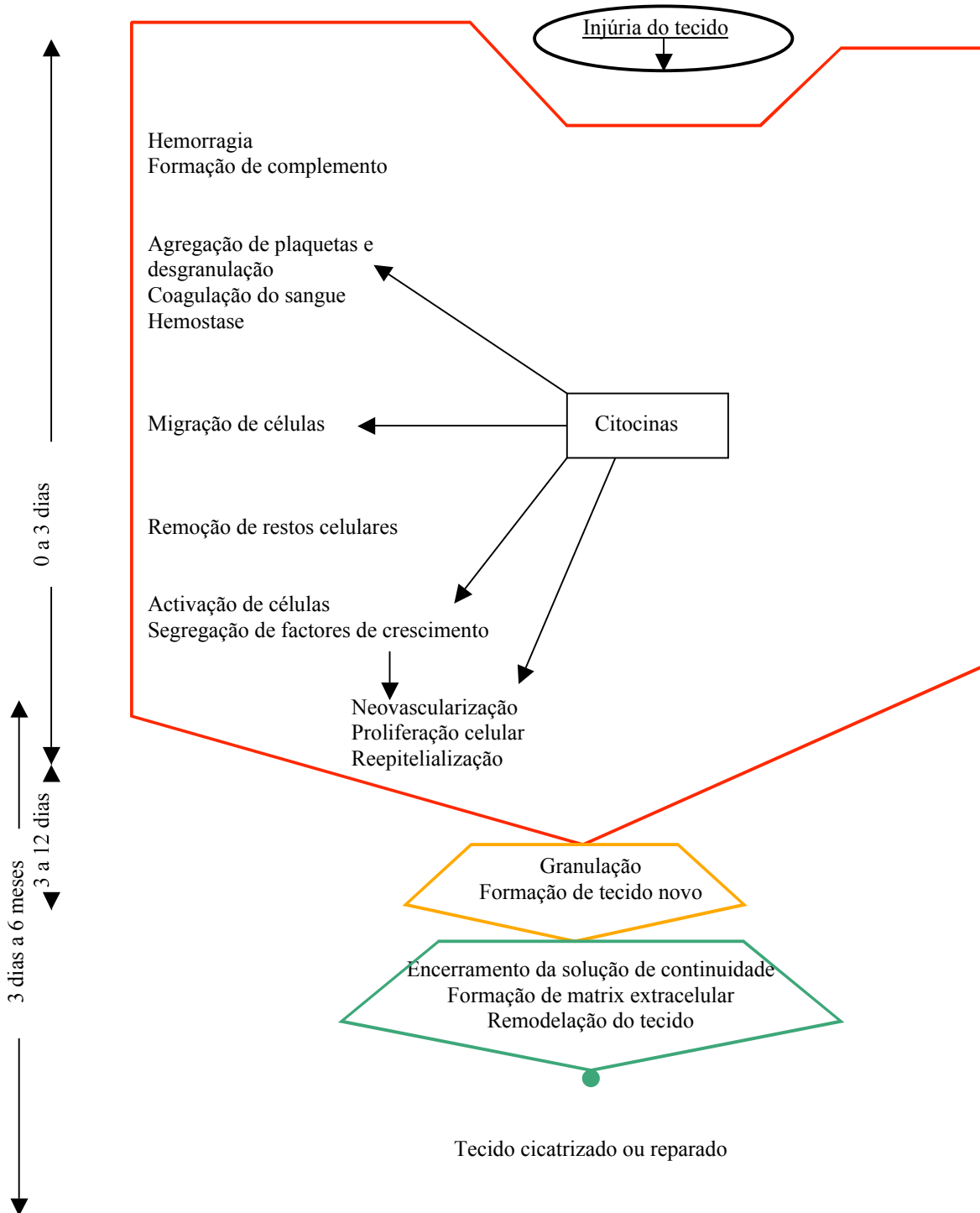
As manifestações sistémicas da inflamação são a hipertermia, e a reacção de fase aguda que leva à biossíntese de proteínas desta fase aguda, no fígado.

A reacção de fase aguda envolve uma série de mudanças endócrinas, metabólicas ou neurológicas muito complexas, que têm lugar localmente ou sistemicamente.

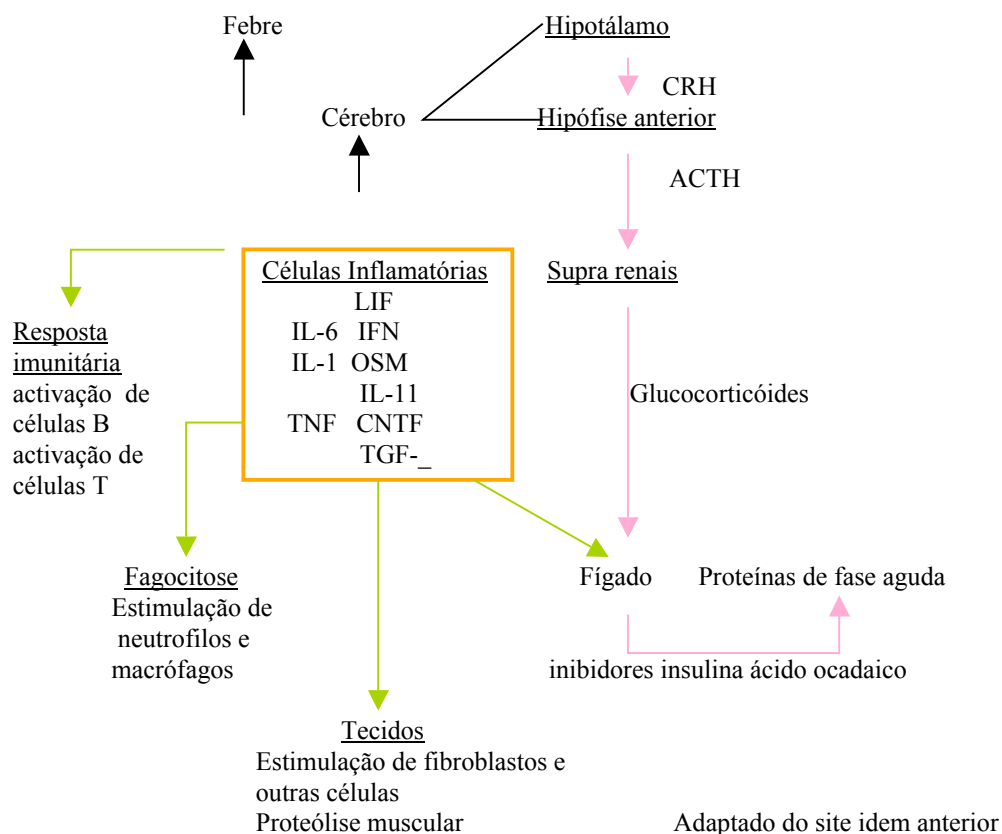
Na reacção da fase aguda aumentam também o número de leucócitos periféricos sobretudo neutrófilos e seus precursores.

Esta reacção de fase aguda é iniciada e mediada por uma série de citocinas segregadas por diversos tipos celulares (leucócitos polimorfonucleares, fibroblastos, células endoteliais, monocitos, linfócitos, etc.).

Esquematiza-se seguidamente a regulação destas reacções de fase aguda bem como a síntese das proteínas de fase aguda.



Adaptado do site  
<http://www.copewithcytokines.de/>



Citocinas inflamatórias como as IL-6, IL-1, TNF, TGF, IFN e LIF são produzidas pelas células inflamatórias, induzindo reacções locais e sistémicas, levando à activação de células como leucócitos, fibroblastos, células endoteliais e células dos músculos lisos originando a formação de mais citocinas além de realizarem acções directas nos hepatócitos .

O principal mediador das reacções de fase aguda é a IL-6 que por seu turno é regulada pela IL-1.

As proteínas de fase aguda são cerca de 30 proteínas diferentes e a sua concentração no soro sanguíneo pode ser aumentada (reagentes positivos de fase aguda) ou diminuída (reagentes negativos de fase aguda) dentro de 1,5 hora após se instalar a reacção inflamatória.

### Proteínas de fase aguda da inflamação e sua função

	<u>Função</u>
_1 glicoproteína ácida	Interactua com o colagénio Promove o crescimento do fibroblasto Liga certos esteroides
_1 anti-quimotripsinogenio	Inibidora de proteólise
_1 antitripsina	
_2 antiplasmina	Modula a cascata de coagulação
_2 macroglobulina	Inibe várias proteínases
Antitrombina-3	Modula a cascata coagulação
C1 a C9	Componente do complemento
Proteína C-reactiva	Liga-se a fosforilcolina da membrana Activa o complexo Interage com as células T e B
Ceruloplasmina	Transporta cobre
Factor VIII	Coagulação do sangue
Factor B	Componente do complemento
Ferritina	Transporta ferro
Fibrinógeno	Formação de coágulos
Fibronectina	Formação de coágulos
Haptoglobina	Remoção de hemoglobina
Caliceína	Permeabilidade vascular e vasodilatação
Heme oxigenase	Degradação do heme
LPs proteína de ligação	Activação de macrófago
Superoxidodesmutase manganês	

Há diferenças nas concentrações das diversas proteínas de fase aguda consoante as espécies animais estudadas (humano, coelho, rato, ratazana).

As proteínas de fase aguda são sintetizadas sobretudo nos hepatócitos que têm capacidade para produzir todo o espectro destas proteínas, e libertá-las para o sangue, havendo aqui um perfil característico destas proteínas.

A elevada concentração de certas proteínas de fase aguda tem relevância para efeitos de diagnóstico e também prognóstico.

Em condições normais esta reacção de fase aguda não é observável em perturbações funcionais que não resultem de um processo inflamatório o que permite fazer a distinção entre estas duas situações.

A reacção inflamatória no local da injúria é caracterizada por um aumento inicial do fluxo sanguíneo no local, favorecimento da permeabilidade vascular, e influxo ordenado e dirigido com acumulação selectiva de diferentes células efectoras do sangue periférico no local da injúria, além da exudação do plasma para a lesão.

As células inflamatórias aumentam o número de moléculas de adesão celular à sua superfície, as células endoteliais activadas, em ordem a permitir o tráfego preciso de leucócitos da circulação para os locais inflamatórios (extravasamento).

O resultado final de uma resposta inflamatória é determinada pelo equilíbrio entre as citocinas pró-inflamatórias e as citocinas anti-inflamatórias.

As principais citocinas pró-inflamatórias são a IL-1 $\alpha$ , IL-1 $\beta$ , IL-6 e TNF- $\alpha$ . Outros mediadores pró-inflamatórios englobam o LIF, IFN- $\gamma$ , OSM, CNTF, TGF- $\beta$ , GM-CSF, IL-11, IL-12, IL-17, IL-18.

As principais citocinas anti-inflamatórias englobam a IL-4, IL-10, IL-13, IL-16, IFN- $\gamma$ , TGF- $\beta$ , IL-1 $\alpha$ , IL-1 $\beta$ , GM-CSF.

Também as prostaglandinas são mediadores naturais da inflamação. A administração da PGE $_1$  e PGE $_2$  induzem sinais de inflamação que inclui rubor, e calor (devido à vasodilatação arteriolar) e “inchaço” e edema em consequência da permeabilidade capilar aumentada.

A PGE produzida pelos macrófagos, mast cells, células B (tecidos imunitários) desencadeia a quimocinese das células T.

Os agentes pirogênicos (indutores de febre) activam a biossíntese das prostaglandinas originando a libertação do PGE $_2$  no hipotálamo onde a temperatura do corpo é regulada.

Também os ácidos monohidroxieicosatetraenóicos (HETE) compreendendo a via da lipoxigenase são potentes mediadores envolvidos na inflamação.

Os mono-HETEs e LTB $_4$  (leucotrienos B $_4$  ou ácido 5, 12, dihidroxieicosatetraenóico) estimulam a migração (quimotaxia) dos eosinófilos e neutrófilos, tornando-os nos principais mediadores da infiltração de PMN (polimorfonucleares) leucócitos nas reacções inflamatórias.

A resposta inflamatória é pois complexa e mediada por uma série de moléculas assinalantes produzidas localmente pelas “mast cells”, terminação dos nervos, plaquetas e glóbulos brancos assim como pela activação do complemento. Algumas destas moléculas sinais actuam sobre os capilares próximos induzindo as células endoteliais a afrouxarem as suas interligações nas suas superfícies adesivas em ordem à passagem dos glóbulos brancos que se escapam assim dos vasos por entre as células endoteliais e através da lamina basal, com a ajuda de enzimas digestivas. Outras moléculas quimioatraentes para determinados tipos específicos de glóbulos brancos intervêm também favorecendo a polarização destas células para o local interessado o que origina um grande número de glóbulos brancos na zona afectada.

Outras moléculas assinalantes produzidas no decurso da resposta inflamatória, entram no sangue e levam à estimulação da medula óssea para produzir mais leucócitos. A medula óssea é capital neste contexto, pois com excepção dos linfócitos e de alguns macrófagos, a maioria dos tipos de glóbulos brancos nos mamíferos adultos são produzidos apenas na medula óssea.

O sistema complemento assim chamado porque ele complementa e amplifica a acção dos anticorpos, é o principal meio através das quais os anticorpos defendem os vertebrados contra a maioria das infecções bacterianas. Consiste num sistema de proteínas do soro sanguíneo que podem ser activadas por complexos antigénio-anticorpo ou por microorganismos desenvolvendo uma cascata de reacções proteolíticas cujo resultado final é a formação de complexos de ataque às membranas dos microrganismos onde formam “buracos” que acabam por destruí-los.



### **4.3 – Quimocinas e seus receptores**

Por quimiotaxia entende-se o movimento dirigido das células, em consequência de um dado estímulo químico.

As quimocinas ou sejam citocinas quimiotáticas, têm capacidade de atrair subconjuntos específicos de leucócitos para os locais onde ocorrem inflamações, podendo também as quimocinas estar implicadas em diversos outros processos tais como a maturação dos linfócitos, a angiogénese e o crescimento de tumores.

Uma série de vírus e protozoários patogénicos para os mamíferos também utilizam as vias das quimocinas e seus receptores para infectar os seus hospedeiros e interferir nas reacções imunológicas consequentes.

A deslocação dos leucócitos para os seus diversos locais de acção, é um processo que decorre em diversas etapas que implicam a expressão e coordenação sequencial de selectinas, integrinas e quimocinas.

As quimocinas são um grupo de pequenas moléculas relacionadas entre si (com cerca de oito a catorze KDa) sobretudo básicas que regulam o tráfego de diversos tipos de leucócitos através de interacção com receptores de sete domínios transmembranários acoplados com G proteínas.

As quimocinas desempenham papéis fundamentais no desenvolvimento, homeostase e funcionamento do sistema imunitário actuando também sobre as células do sistema nervoso central, e sobre as células endoteliais implicadas na angiogénese ou angioestase.

As pequenas citocinas, intercrinas ou quimocinas são proteínas que parecem ter actividades mitogénicas, quimiotáticas ou inflamatórias sendo factores de baixo peso molecular com 70 a 100 resíduos de ácidos aminados segregados por células inclusivé fibroblastos, macrófagos e células endoteliais em resposta a diversos estímulos como por exemplo factores de crescimento, interferões e produtos bacterianos e virais.

A interleucina 8 foi a primeira quimocina a ser caracterizada há cerca de 10 anos, mas hoje são conhecidas mais de 40 quimocinas nos humanos. As quimocinas são a maior família de citocinas da imunofisiologia dos mamíferos. Elas são reguladoras da inflamação e actuam principalmente sobre os leucócitos induzindo a sua migração e respostas respectivas.

Algumas destas proteínas e respectivos receptores estão envolvidos na inflamação, em doenças auto-imunes e na infecção pelo HIV-1.

A migração de células inflamatórias e não inflamatórias desencadeada pelas quimocinas é canalizada para os tecidos apropriados ou compartimentos dentro dos tecidos.

A maioria das quimocinas formam dímeros, mas estes dissociam-se após diluição, e os monómeros é que constituem a forma activa. A estrutura tridimensional de cada monómero é virtualmente idêntica, mas a estrutura quaternária nas quimocinas varia consoante a subfamília.

Nas quimocinas a porção N-terminal das suas estruturas proteicas interactua nos leucócitos com os receptores para essas quimocinas, enquanto a porção C-terminal das quimocinas tem uma alta afinidade para as glicosaminoglicanas e outras porções hidrocarbonadas carregadas negativamente, o que origina uma ligação intensa das quimocinas com as superfícies das células endoteliais e com as glicoproteínas da

matrix extracelular dos tecidos conectivos, formando-se um gradiente de concentrações de quimocinas ligadas, em relação às quimocinas solúveis livres ou não ligadas. O movimento dos leucócitos e a sua localização ocorre para as mais elevadas concentrações de quimocinas ligadas.

O principal domínio de ligação ao receptor de todas as quimocinas situa-se próximo da terminação  $\text{NH}_2$  do polipéptido e antagonistas podem ser obtidos truncando ou realizando substituições nesta região.

A região N-terminal das quimocinas tem um papel dominante na sua actividade biológica. Os efeitos das quimocinas sobre a angiogénese e o crescimento dos tumores estão bem assinalados embora os mecanismos de actuação sejam desconhecidos.

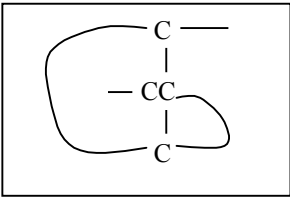
Claro que a capacidade para que um dado tipo de leucócito específico responda a um dado tipo de quimocina depende da expressão de um receptor apropriado para essa quimocina, na superfície desses leucócitos.

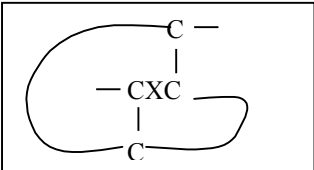
As quimocinas são constituídas por uma simples cadeia polipeptídica com cerca de 70-100 ácidos aminados, e com uma homologia entre as diversas quimocinas de 20 a 95% incluindo cisteínas conservadas que ajudam a identificar quatro subfamílias, as C, CC, CxC e Cx<sub>3</sub>C, a segunda e terceira contendo quatro cisteínas (duas adjacentes em CC ou separadas por um ácido aminado em CxC). Cada subfamília contém diversos membros, assim na subfamília “CxC” estão assinaladas dez proteínas diferentes (de plaquetas, monocitos, macrófagos, etc.) com diversas propriedades assim como subfamílias C-C estão referidas com 20 proteínas diferentes (de monocitos, macrófagos, timo, etc.), todas elas incluindo quatro cisteínas conservadas.

A fractalcina é a única quimocina que tem um domínio transmembranário e que pode ser ligada à superfície celular.

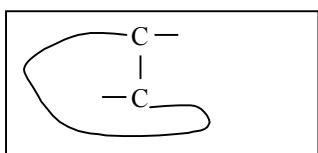
No quadro seguinte referem-se as quatro famílias de quimocinas, alguns tipos de quimocinas de cada família, e os seus receptores respectivos.

Família de quimocinas                      Quimocinas                      Receptores

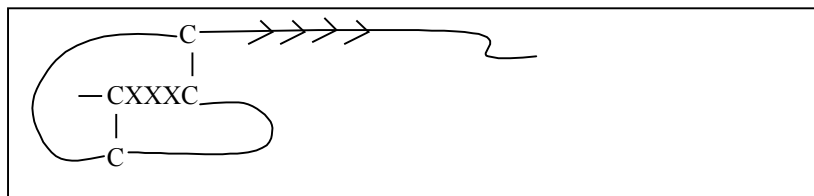
CC		MiP-1 <sub>1</sub> ; MCP-3; RANTES; HCC-1; MPIF McP-1,-2,-3,-4,-5 Eotaxina-1,2; MCP-2,-3,-4; RANTES TARC, MDC MiP-1 <sub>1</sub> , 1 <sub>2</sub> ; RANTES; MCP-2,-3 MiP-3 <sub>1</sub> MiP-3 <sub>2</sub> ; SLC I-309 MiP-1b; RANTES; MCP-2,-3,-4	CCR1 CCR2 CCR3 CCR4 CCR5 CCR6 CCR7 CCR8 CCR9
----	---	--	--

CXC		IL-8; GCP-2 IL-8 ;GCP-2 ;GRO-a,-b,-g ;ENA-78 NAP-2 IP-10; MIG; I-TAC SDF-1 BCL-1	CXCR1 CXCR2 CXCR2 CXCR3 CXCR4 CXCR5
-----	---	---	--

C	Linfotactina	CR1
---	--------------	-----



CX3C	Fractalcina	CX3CR1
------	-------------	--------



As quimocinas são pois uma grande família de citocinas quimotáticas que actuam através de receptores acoplados a G proteínas e que regulam o tráfego de leucócitos, a angiogénese, a hematopoiese e a organogénese supondo-se que sejam benéficas a ajudar os organismos contra os agentes infecciosos e perturbadoras nas doenças marcadas por inflamações patológicas.

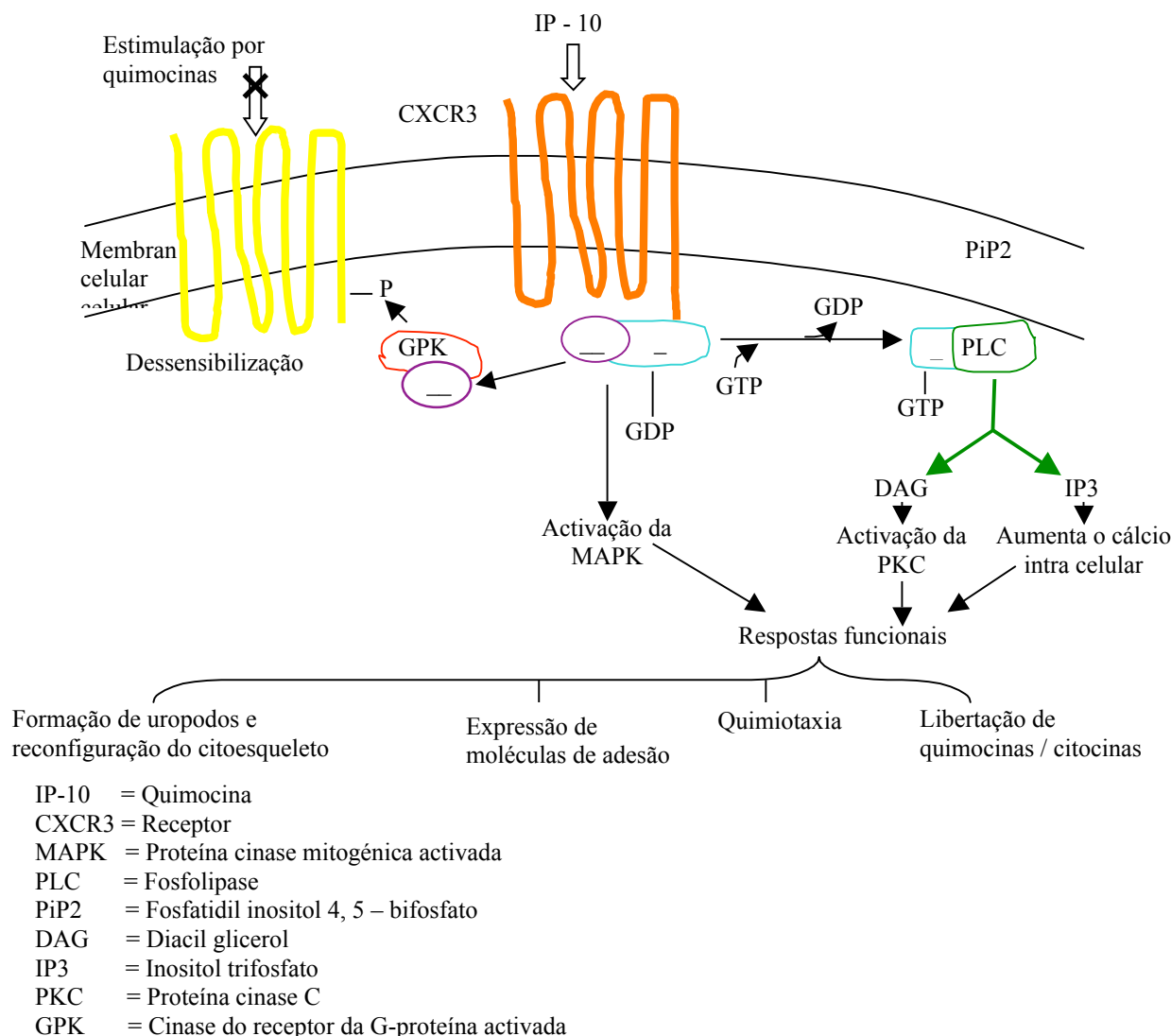
Conforme se pode ver na figura seguinte os receptores para as quimocinas têm uma porção N-terminal extracelular, que interactua com a quimocina, três ansas extracelulares e outras três ansas intracelulares, e uma porção C-terminal intracelular.

Os receptores para as quimocinas, tal como as quimocinas podem ser agrupados em quatro grupos (CR, CCR, CXCR e CX3R) consoante a estrutura primária da respectiva proteína, podendo ocorrer dentro de cada grupo grande semelhança destas sequências.

As ansas intracelulares dos receptores das quimocinas estão ligadas com proteínas G compostas por sub-unidades  $\alpha$ ,  $\beta$  e  $\gamma$  que intracelularmente se ligam a GTP.

Na figura seguinte esquematizam-se vias assinalantes (ou de transdução de sinal) intracelulares desencadeadas pelas quimocinas IP-10 (figura adaptada de J. B. Rottman, Vet. Pathol. 36:357-367-1999)

Via assinalante desencadeada pela quimocina IP-10 e a dessensibilização do respectivo receptor por fosfatação da porção C-terminal do receptor pelo complexo \_\_ – GPK



Na figura anterior uma quimocina (IP-10) ao interagir com o respectivo receptor (CXCR3) activa este desencadeando ao nível da subunidade G<sub>βγ</sub> da proteína ligadora de GTP (G-proteína) uma troca de GDP por GTP o que origina na G-proteína uma dissociação da subunidade G<sub>βγ</sub> das subunidades G<sub>α</sub>. Estas subunidades G<sub>βγ</sub> interagem e activam a fosfolipase (PLC) que activada hidrolisa o fosfoinositidilinositol 4,5 bifosfato (PIP2) originando os mediadores inositol trifosfato (IP3) e diacilglicerol (DAG), o primeiro difundindo-se para o citosol onde origina a mobilização de cálcio intra e extracelular aumentando a concentração deste cálcio intracelular. Por seu turno o DAG activa uma proteína cinase C (PKC) que fosfata uma série de proteínas intracelulares, inclusive proteínas cinases activadoras de mitogénios (MAPKs) activando factores de transcrição, fosfolipases A2 e outras cinases, contribuindo para a reorganização de elementos do citoesqueleto.

A resultante polarização dos receptores das quimocinas parece implicada no movimento directo das respectivas células através do contexto de gradiente de concentração de quimocina.

A porção C-terminal do receptor das quimocinas, após a interação da quimocina com o receptor activando este, esta porção C-terminal dizíamos,liga-se com o complexo G \_\_ dissociado da proteína ligadora de GTP antes referido (G-proteína) que entretanto se tinha complexado com cinases de receptores activados por G-proteínas (GPKs), desencadeando a fosfatação de resíduos de serina e treonina da porção C-terminal do receptor das quimocinas,o que origina a dessensibilização do receptor para mais estimulação pela quimocina, o que pode estar na base do “depósito” e retenção dos leucócitos respectivos dentro dos tecidos inflamados.

Na figura em apreciação referem-se uma série de respostas funcionais desencadeadas na célula pela interação da quimocina com o respectivo receptor.

Há dezassete subtipos de receptores de quimocinas conhecidos que ligam cada um deles diversas quimocinas, sendo esses receptores designados pelos nomes da sub-classe de citocina seguido por um número (por exemplo CC R1 – 10, CxC R1 – 5, etc.).

Os diversos receptores tem uma simples cadeia polipeptídica com 25 a 80% de homologia entre si.

No quadro seguinte( adaptado da fonte do quadro anterior) refere-se a diversidade de receptores de quimocinas expressos pelos diversos tipos de leucócitos.

	CCR1	CCR2	CCR3	CCR4	CCR5	CCR6	CCR7	CCR8	CCR9	CCR10	CXCR1	CXCR2	CXCR3	CXCR4	CXCR5	CR1	CX3CR1
Neutrófilos											+	+					
Eosinófilos	+		+														
Basófilos		+	+														
Monocitos	+	+			+				+					+			+
Células T(naive)														+			
Células T (memória)	+	+	+	+	+	+	+	+					+				+
(Killer) C. citotóxicas naturais		+			+								+				+
C.dendríticas	+	+	+	+	+	+	+	+						+			

Estes receptores são expressos pelos leucócitos e por uma série de outras células tais como neurónios, astrocitos, microglia, epitélios e endotélios.

A expressão dos receptores de quimocinas nos linfocitos varia com o estado de diferenciação de cada célula. Conforme se pode ver no quadro anterior os linfocitos T ainda sem memória (naive) exprimem apenas os tipos CXCR4, mas as células T com memória exprimem pelo menos nove receptores.

É pois a presença de determinados receptores à superfície das células leucocitárias que determina a que tipo de quimocina essas células responderão.

Na inflamação e na resposta do hospedeiro à infecção a atracção dos leucócitos para os tecidos interessados é essencial e este processo é controlado por quimocinas, que são como referimos citocinas quimotáticas (quimocina é um curto neologismo para “citocinas quimotaticas”).

Como dissemos mais de 40 quimocinas estão assinaladas, sendo proteínas de 8 a 14 KDa com 20 a 95% de homologia nas suas sequências de ácidos aminados, sendo subdivididas em famílias na base da posição relativa dos seus resíduos de cisteína.

As quimocinas regulam o movimento dos leucócitos, sendo segregadas nos locais da inflamação e da infecção pelas células residentes nos tecidos, leucócitos residentes e recrutados, e células endoteliais activadas por citocinas.

As quimocinas são retidas no local sobre a matriz extracelular e nas proteoglicanas heparana sulfato à superfície das células estabelecendo-se um gradiente de concentrações das quimocinas à volta da zona inflamada assim como à superfície do endotélio que a rodeia.

Os leucócitos rolam sobre o endotélio através de um processo mediado por selectinas e entram em contacto com quimocinas retidas nos locais antes referidos.

As quimocinas activam as integrinas dos leucócitos e levam a uma aderência firme e ao extravasamento destes.

Os glóbulos vermelhos através de receptores para quimocinas também ajudam a manter um gradiente de concentração de quimocinas nos tecidos e na corrente sanguínea.

Nas várias doenças inflamatórias ocorre uma acumulação selectiva de diversos sub-grupos de leucócitos, através de um processo controlado pela expressão de algumas quimocinas.

Cada doença tem um infiltrado inflamatório característico no qual as concentrações de cada quimocina e do seu RNAm são finamente reguladas.

Já referimos diversas famílias de quimocinas, na base da posição relativa dos resíduos de cisteína na proteína madura, assim como os principais tipos de quimocinas produzidos, o tipo de receptor que com elas interacciona e ainda o tipo de células para as quais as quimocinas são quimiotácticas.

No quadro seguinte referem-se para algumas doenças inflamatórias humanas, o tipo de sub-grupos de leucocitos acumulados selectivamente no infiltrado em cada processo e as quimocinas que pela sua expressão controlam cada processo.

<b><u>Doença inflamatória</u></b>	<b><u>Tipo de infiltrado</u></b>	<b><u>Quimocinas expressas</u></b>
Síndrome agudo respiratório	Neutrófilos	Interleucina-8 GRO- $\alpha$ , $\beta$ , $\gamma$ ENA-78
Asma	Eosinófilos Células T Monocitos Basófilos	MCP-1, -4 MIP-1 $\alpha$ Eotaxina RANTES
Pneumonia bacteriana	Neutrófilos	Interleucina-8 ENA-78
Sarcoidosis	Células T Monocitos	IP-10
Glomerulonefrite	Monocitos Células T Neutrófilos	MCP-1 RANTES IP-10
Artrite reumatoide	Monocitos Neutrófilos	MiP-1 $\alpha$ Interleucina-8 MCP-1 ENA-78
Osteoartrite		MiP-1 $\alpha$
Aterosclerose	Células T Monocitos	MCP-1, -4 IP-10
Doenças inflamatórias dos intestinos	Monocitos Neutrófilos Células T Eosinófilos	MCP-1 Interleucina-8 MiP-1 $\alpha$ Eotaxina IP-10
Psoríase	Células T Neutrófilos	MCP-1 GRO- $\alpha$ IP-10 Interleucina-8 MIG
Meningite bacteriana	Neutrófilos Monocitos	Interleucina-8 GRO- $\alpha$ MCP-1 MiP-1 $\alpha$ , -1 $\beta$
Meningite viral	Células T Monocitos	MCP-1 IP-10

GRO – growth-regulated oncogene

ENA-78 – péptido 78 ativando neutrófilos derivado de células epiteliais

MCP – proteína atraente de monocitos

MiP – proteína inflamatória de macrófagos

RANTES – regulada por activação expressa e segregada por células T normais

IP-10 – proteína 10 induzível pelo interferon

As quimocinas são pois selectivas para subtipos específicos de leucócitos quimioatraindo-os especificamente e possuem consoante os seus tipos, selectividades de banda larga ou estreita, sendo pois os orientadores chave do tráfego específico dos leucócitos em condições básicas e de emergência na inflamação e na imunidade.

Algumas quimocinas funcionam ainda, em contextos específicos, como factores de crescimento, por exemplo na angiogénese e na hematopoiese.

Algumas quimocinas suprimem infecções “in vitro” ao bloquear a interação viral (por exemplo do HIV-1) com receptores específicos da quimocina que são indispensáveis para a penetração dos vírus nas células.

Todas as quimocinas actuam como sinais intercelulares, mas as suas origens, alvos e regulação são diversos, sendo produzidas, a maioria por diversos tipos celulares, embora algumas sejam produzidas só por um tipo celular, sendo produzidas de uma forma constitutiva induzida ou não.

Algumas quimocinas são reguladas por citocinas pró-inflamatórias como o TNF (factor de necrose tumoral), IL-1 (interleucina 1), o interferão gama etc.

A maioria das quimocinas induzíveis são reguladoras ao nível da transcrição.

Os alvos das quimocinas (largos ou estreitos) variam. Assim as quimocinas cujo alvo são os neutrófilos encontram-se sobretudo na subfamília CxC, enquanto os monocitos / macrófagos, eosinófilos, e basófilos são sobretudo atraídos pelas quimocinas da subfamília CC. As BCA – 1 são uma quimocina CxC específica para os linfócitos B e ambas as famílias CC e CxC também contem membros específicos dos linfócitos T.

Os contactos dos leucócitos com os endotélios pode ser transitório, reversível e independente de qualquer activação. Mas se se mantêm esses contactos as células rolam através da superfície endotelial por interação de selectinas com contra-adesinas independentemente de quimocinas.

Nos locais onde tem sede inflamação, os leucócitos entram numa segunda fase que envolve activação de integrinas  $\alpha_2$  dependentes de quimocinas, estabelecendo-se uma forte ligação com os receptores das células endoteliais, parecendo que os leucócitos interactuaram com as quimocinas que se encontram imobilizadas à superfície por proteoglicanos, o que evita a “lavagem” pela corrente sanguínea desta zona e que conduz os leucócitos para a zona da inflamação. Todo este processo favorece a fagocitose, a produção de superóxidos, libertação de grânulos e actividade bactericida “in vitro”. In vivo as quimocinas podem induzir a acumulação de leucócitos sem activação.

Admite-se através da presença constitutiva de quimocinas nos gânglios linfáticos, no timo e na medula óssea que elas têm um papel regulador na produção normal de leucócitos e sua distribuição.

Perfis complexos da expressão das quimocinas têm sido correlacionados com diversas doenças inflamatórias.

Também estão assinaladas quimocinas angiogénicas e angioestáticas. Apesar das quimocinas e seus receptores terem provavelmente surgido como factores antimicrobianos, alguns são utilizados por agentes infecciosos para facilitar a infecção, quer explorando os receptores para penetrarem nas células (patogénicos intracelulares), quer subvertendo os mecanismos de acção das quimocinas.

Por outro lado será a expressão das quimocinas que poderá ditar a composição celular e a posição tri-dimensional das células leucocitárias dentro de cada tipo de lesão.

A composição do meio em quimocinas é que vai definir o fenótipo da lesão e por neutralização de certos tipos de quimocinas representadas no meio será possível modificar completamente esta composição celular ao nível da lesão.



Quando uma dada célula entra num dado tecido em consequência de um gradiente quimiotático gerado por uma dada quimocina, essa célula viajará através desse gradiente até atingir um valor desse gradiente de concentração da quimocina em que o receptor da célula para essa quimocina seja dessensibilizado, não respondendo mais à estimulação por essa quimocina. Neste momento a célula permanece nessa determinada localização dentro da lesão.

Uma dada quimocina ou receptor para quimocina pode ser determinante para a formação inicial de uma lesão, o que sugere que cada lesão inflamatória pode ter uma quimocina ou receptor para quimocina dominante, e que se forem bloqueados poderão impedir a formação da lesão.

No entanto as quimocinas além de condicionarem o recrutamento de linfócitos para os tecidos inflamados também condicionam o seu processo normal de maturação, estando bem assinaladas diferentes expressões de receptores para quimocinas nos tipos 1 e 2 (Th 1 e Th2) das células “helper”.

Os linfócitos T podem ser distribuídos em diferentes conjuntos consoante os tipos de citocinas produzidas por eles. Os Th1 controlam respostas mediadas por células e estão implicados em inflamações crónicas, enquanto os Th2 estão implicados em doenças alérgicas. Os Th1 exprimem receptores para quimocinas CCR5 e CXCR3, enquanto os TH2 exprimem CCR3, CCR4 e CCR 8.

As quimocinas além de mediar a inflamação e a maturação e polarização das células T, também medeiam como já referimos a angiogénese que é muito importante no crescimento dos tumores, podendo também essa angiogénese ser regulada por um equilíbrio muito delicado entre as quimocinas angiogénicas e as quimocinas angioestáticas que podem condicionar assim os processos de crescimento dos tumores e as suas metástases.

Antagonistas das quimocinas CXC que têm um papel importante na angiogénese podem ser importantes no tratamento do cancro, enquanto pelo contrário agonistas dessas quimocinas podem ser úteis como favorecedores de cicatrização ou recuperação de lesões.

Também é conhecido que certos vírus e protozoários podem utilizar vias que integrem quimocinas/receptores para infectar os seus hospedeiros.

Pode resumir-se portanto que as quimocinas e os seus receptores estão implicados num grande conjunto de doenças, e que potentes antagonistas destes receptores poderão ser úteis para tratamento de uma série de situações patológicas tais como, inflamações crónicas, neoplasias e doenças infecciosas.

#### **4.4-Citocinas**

As citocinas são pois numerosas e pequenas proteínas segregadas que se ligam a receptores da superfície de certas células, desencadeando a diferenciação e/ou proliferação dessas células. As linfocinas (englobam interleucinas) que também fazem parte das citocinas, regulam a intensidade e a duração da resposta imunológica.

A hematopoiese, o processo através do qual as células estaminais, “stem cells” pluripotentes originam diversos tipos de células sanguíneas maduras envolve uma série muito complexa de factores de crescimento hematopoiéticos.

As “stem cells” hematopoiéticas dão origem a todos os tipos de células sanguíneas maduras quer de série linfóide (células B- e células T-) quer da série mielóide (monocitos, granulocitos, eritrocitos, megacariócitos).

A proliferação, diferenciação e actividades funcionais das células hematopoiéticas são pois reguladas por um grupo muito variado de proteínas designadas colectivamente como citocinas.

As citocinas produzem múltiplos efeitos biológicos que dependem muitas vezes do estado de diferenciação das células alvo, e isto é feito através da activação de variadíssimas vias de transdução de sinais.

Há citocinas que promovem o crescimento e sobrevivência de mais de uma estirpe de células ao contrário de outras que actuam essencialmente sobre um determinado tipo de linha celular.

O número de citocinas até hoje assinalado é enorme, basta referir que apenas para as interleucinas estão assinaladas mais de vinte membros, sendo as interleucinas numeradas de acordo com sequência temporal da sua identificação, sendo portanto citocinas com diversos alvos celulares mas cujo nome reflecte a sua descoberta em interacções leucócito-leucócito.

As citocinas sendo péptidos sintetizados por muitos tipos celulares e libertados também pelos glóbulos brancos e pelos macrófagos dos tecidos estimulam ou suprimem a actividade funcional dos linfócitos, monocitos, neutrófilos, fibroblastos e células endoteliais.

As citocinas controlam o desenvolvimento da resposta imunitária, sendo por vezes chamadas “hormonas do sistema imunitário” pois elas modulam a diferenciação e divisão das “Stem cells” células mãe hematopoiéticas e a activação dos linfócitos e fagócitos. Recorda-se que as “Stem cells” se podem dividir sem limite, dando lugar a uma descendência diferenciada.

As citocinas são habitualmente proteínas segregadas extracelularmente habitualmente com menos de 80KDa e algumas glicosiladas, produzidas por muitos tipos celulares e envolvidas nas interacções célula a célula, actuando através de receptores localizados na superfície das células alvo normalmente em pequeno número actuando de uma forma parácrina, autócrina ou endócrina.

Autócrino é um mecanismo de controle do crescimento em que uma célula segrega um factor solúvel que se liga ao seu receptor o qual é expresso também pela célula que produz esse factor, criando assim uma ansa autogena no qual o produto actua de volta sobre a mesma célula que o produz.

Parácrino é um mecanismo de controle do crescimento envolvendo a síntese de um factor solúvel por uma célula que influencia o crescimento e actividade funcional das células próximas exprimindo o receptor para este factor.

O mecanismo difere daquele mediado por factores endócrinos que envolvem o transporte do factor para o local de acção, através da circulação (acção endócrina das hormonas).

Justacrínico é um mecanismo de controlo do crescimento e de comunicação intercelular envolvendo contactos específicos célula a célula.

As citocinas desencadeiam como já referimos nas células alvo diversas funções celulares tais como activação, proliferação, quimiotaxia imunomodulação, libertação de outras citocinas ou

mediadores, crescimento e diferenciação celular, apoptose, etc.. Refere-se seguidamente uma classificação global de citocinas.

### Classificação global das citocinas baseadas na sua estrutura tridimensional

#### A – Com cadeias peptídicas em hélices alfa

##### A1 – Superfamília com “feixes” de 4 hélices de cadeias longas

- Somatofina. Prolactina
- Factor estimulador de colónias de grunulocitos (G-CSF)
- Factor de crescimento mielomonocítico
- Interleucina ~6 (IL-6)
- Interleucina ~3 (IL-3)?
- Interleucina ~7 (IL-7)
- Factor inibitório da leucemia (LiF)
- Oncostatina M
- Factor neurotrófico ciliar (CNTF)
- Factor de diferenciação colinérgico (CDF)

##### A2 – Superfamília com “feixes” de 4 hélices de cadeias curtas

- Interleucina 2 (IL-2)
- Interleucina – 4 (IL-4)
- Interleucina –13 (IL-13)
- Interferão alfa (IFN- $\alpha$ )?
- Interleucina – 5 (IL-5)
- Factor estimulador de colónias granulocitomacrófago (GM-CSF)
- Interleucina – 3 (IL-3)
- Factor estimulador de colónias macrófagos (M-CSF)

##### A3 – “Feixes” de 4 hélices dímero-dímero

- Interferão gama (IFN- $\gamma$ )
- Interleucina – 10 (IL-10)
- Interferão beta (IFN- $\beta$ )

#### B – Com cadeias peptídicas de tipo beta

##### B1 – Beta em trevo

- Interleucina 1 – alfa
- Interleucina 1 – beta
- Factor de crescimento fibroblasto (FGF)

##### B2 – Beta em sandwich

- Factor de necrose tumoral alfa (TNF- $\alpha$ )
- Factor de necrose tumoral beta (TNF- $\beta$ )
-

### B3 – EGF – like

- Factor de crescimento transformante alfa (TGF- $\alpha$ )

### B4 – Domínios com dimerização de “laçadas” de cistina

- Gonadotrofina
- Factor de crescimento dos nervos (NGF)
- Factor de crescimento derivado das plaquetas (PDGF)
- Factor de crescimento transformante beta (TGF- $\beta$ )

### C – com cadeias mistas alfa – beta

- IL-8
- IP-10
- Factor das plaquetas 4 (PF-4)
- bTG
- GR $\alpha$ , 9E3, HLA-A<sub>2</sub>
- MiP-1 alfa – proteína 1 inflamatória dos macrófagos alfa
- MiP-1 beta – proteína 1 inflamatória dos macrófagos beta
- M GSA (actividade estimuladora de crescimento do melanoma)

Em animais modelos utilizados para estudo da asma tem sido possível classificar as numerosas citocinas envolvidas neste processo nos seguintes grupos genéricos:

- Linfocinas IL-2, IL-3, IL-4, IL-5, IL-13, IL-15, IL-16, IL-17
- Citocinas pr-inflamatórias IL-1, TNF, IL-6, IL-11, GM-CSF, SCF
- Citocinas anti-inflamatórias IL-10, IL-Ira, IFN- $\gamma$ , IL-12, IL-18
- Citocinas quimiotáticas (quimocinas) RANTES, MCP-1, MCP-2, MCP-3, MCP-4, MCP-5, MiP-1 $\beta$ , eotaxina, IL-8
- Factores de crescimento PDGF, TGF- $\alpha$ , FGF, EGF, IGF
- 

IL – interleucina

TNF – factor de necrose tumoral

GM-CSF – factor estimulador de colónias granulocito-macrófago

IFN – interferão

PDGF – factor de crescimento derivado de plaquetas

TGF – factor de crescimento transformante

FGF – factor de crescimento fibroblasto

EGF – factor de crescimento epidérmico

IGF – factor de crescimento like-insulina

IL-Ira – interleucina 1 antagonista

As citocinas também têm um papel importante na apresentação de antígenos e podem favorecer ou suprimir a capacidade dos macrófagos como células apresentadoras de antígenos, tal como exercem um importante efeito regulador sobre a expressão de moléculas de adesão, e tendo também um papel fundamental no recrutamento de células inflamatórias.

As citocinas implicam na sua acção a interacção com receptores os quais são agrupados nas seguintes superfamílias:

1 – Superfamília de receptores das citocinas engloba as IL-2R<sup>α</sup> e cadeias  $\beta$ , IL-4R, IL-3R<sup>α</sup> e cadeias  $\gamma$ , IL-5<sup>α</sup> e  $\beta$ , IL-6R, gp130, IL-12R e GM-CSFR. As suas regiões extracelulares têm diversos tipos de domínios e podem ter só uma cadeia polipeptídica.

2 – Superfamília imunoglobulina que possui nas suas sequências extracelulares domínios do tipo imunoglobulina, incluem a IL-1R, IL-6R, PDGFR e M-CSFR.

3 – Superfamília receptor proteína cinase com um domínio extracelular glicosilado e um domínio intracelular catalítico de tirosina cinase, inclui receptores para os PDGF, EGF e FGF.

4 – Superfamília receptor interferão com receptor o IFN- $\alpha$  R e IFN- $\beta$  receptor e IL-10R. A transdução de sinal envolve fosfatação e activação da JAK e TYK2 proteínas tirosina cinase.

5 – Superfamília do factor de crescimento dos nervos que engloba receptores para o NGFR, TNFR-1 (p55) e TNFR-II (p75) cujo modo de transdução de sinal não é ainda conhecido.

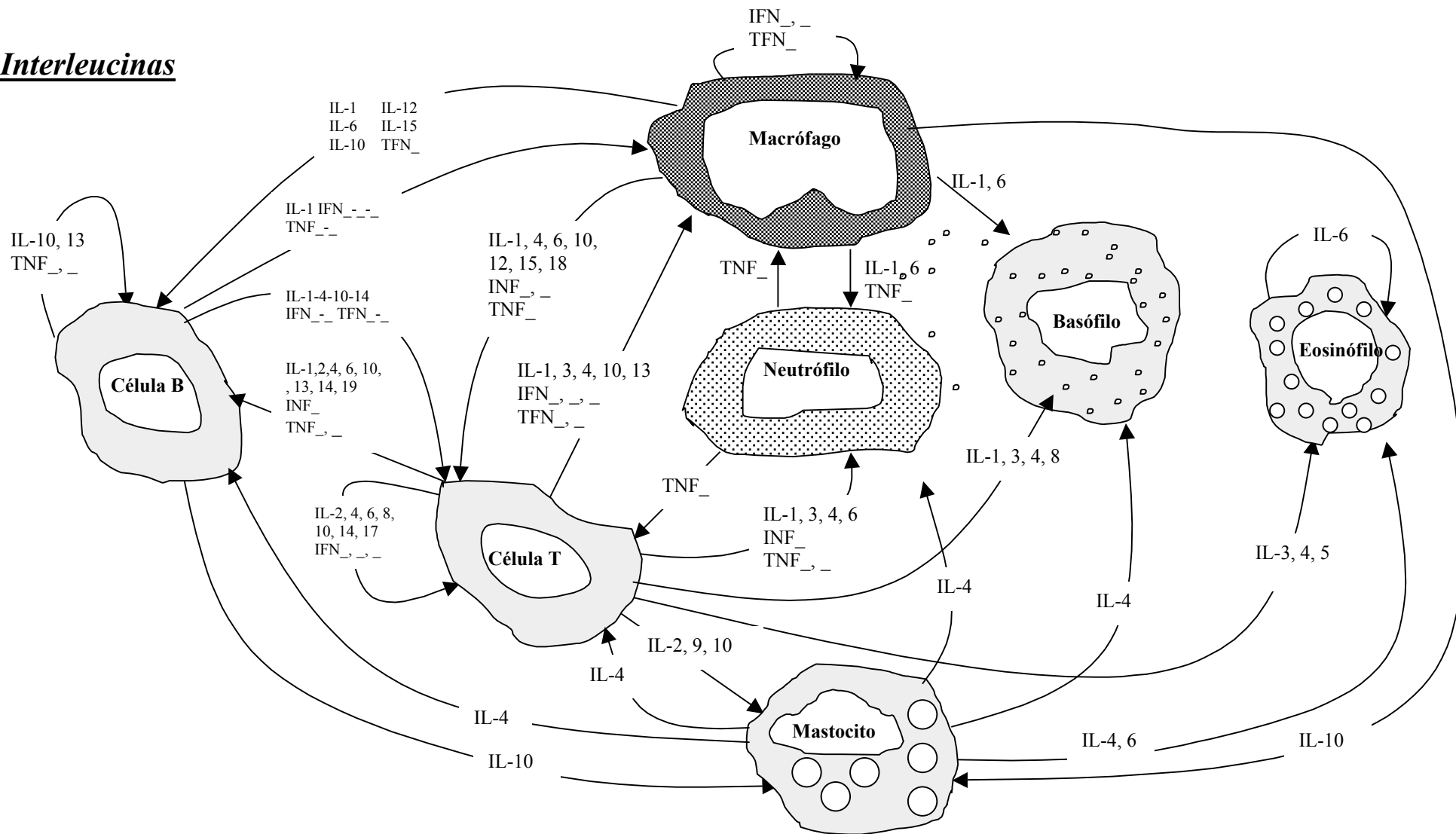
6 – Superfamília de receptores (acoplados a proteína G) de sete domínios transmembranários que engloba os receptores para as quimocinas. Os receptores nas suas porções intracelulares estão acoplados a proteínas heterotriméricas ligando GTP que induzem hidrólise do fosfatidilinositol fosfato e activam cinases fosfatases e canais iónicos.

Mostra-se no esquema da figura seguinte algumas interacções entre diversos tipos celulares com indicação das interleucinas intervenientes em cada uma das vias.

Nos quadros e diagramas sinópticos que referimos a seguir esquematizamos os seguintes aspectos de envolvimento de diversas citocinas:

1. Mecanismo de resposta inflamatória
2. Algumas cascatas bioquímicas induzidas pela activação celular
3. Citocinas, alguns efeitos biológicos
4. Transdução de sinais das citocinas
5. Citocinas hematopoiéticas (mecanismos de acção)
6. Alguns factores de crescimento e sua acção
7. Alguns factores estimuladores de colónias (CSF) que influenciam a formação de células sanguíneas
8. Linfócitos. Locais de formação, diferenciação e compartimentação
9. Linfócitos. Funções e tipos de resposta
10. Activação linfocitária
11. Algumas linfocinas segregadas pelas células T em resposta a antígenos
12. Propriedades de algumas interleucinas

# Interleucinas



## 1 - Mecanismo da resposta inflamatória

Estímulo

Citocinas

Hepatocito

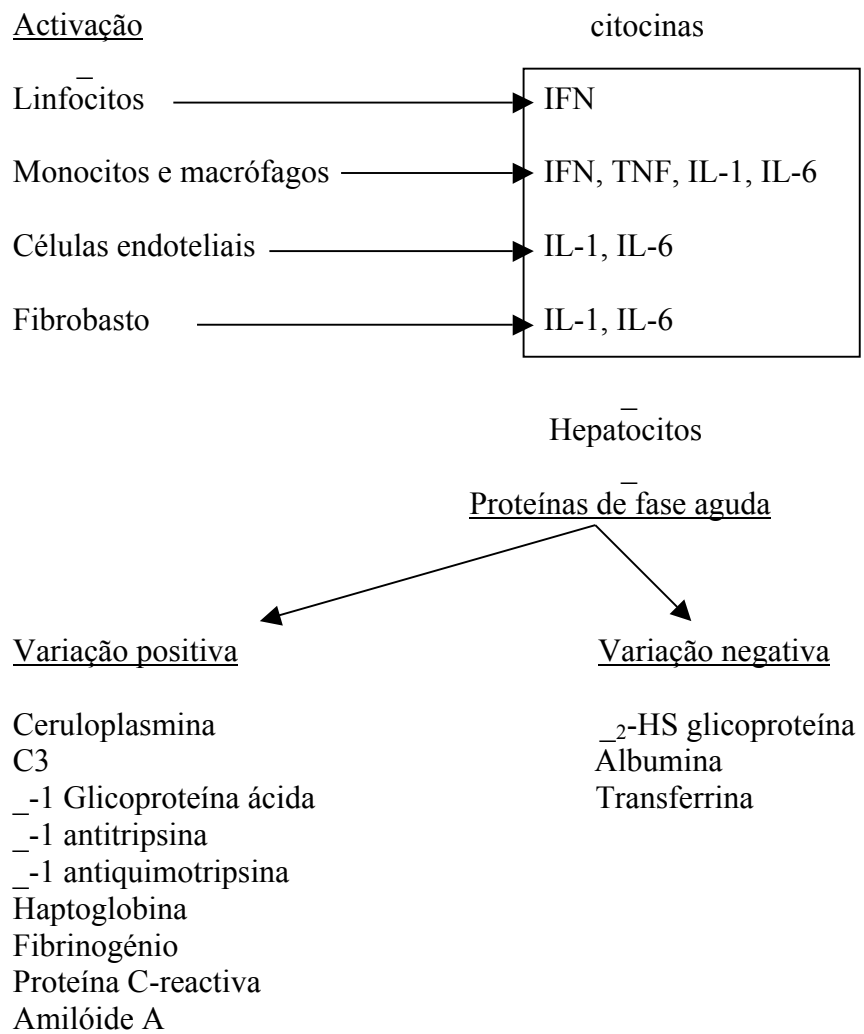
Proteínas de fase aguda  
( \_ síntese e segregação )



Outras células alvo \_ outros efeitos

Reacção imunológica  
Reacção fagocitária  
Activação dos fibroblastos  
Proteólise muscular  
Reacção cerebral  
(febre/sono)

2 - Algumas cascatas bioquímicas induzidas pela activação celular





### 3 - Citocinas

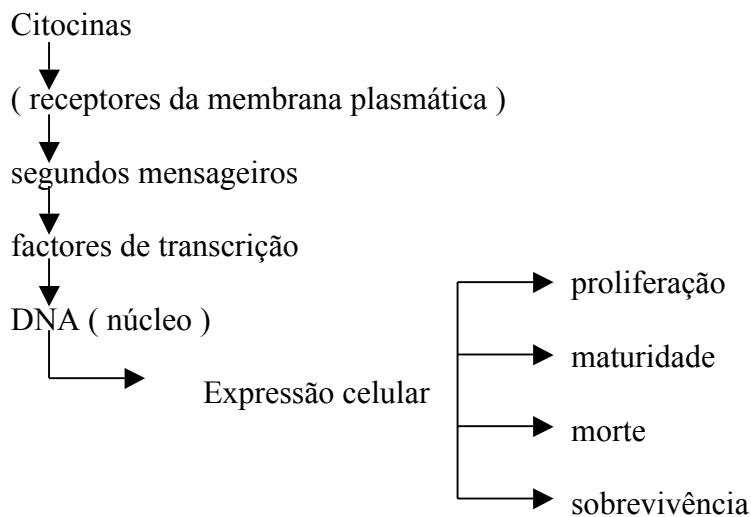
#### Alguns efeitos biológicos (autócrinos e / ou parácrinos)

- Influência no metabolismo e nas actividades celulares
- Controlo da proliferação, diferenciação e do fenótipo celular
- Controlo da hematopoiese
- Controlo da resposta inflamatória fagocitária e citotóxica
- Controlo da resposta imunológica e das defesas contra vírus e parasitas
- Remodelação óssea
- Reparação de feridas

#### 4 - Transdução de sinal das citocinas ( principais mecanismos de acção intracelular )

- Proteínas G-AMP cíclico – proteína cinase A
- Tirosina cinases
- Fosfolipases A<sub>2</sub>, C e D
- Inositol fosfato-cálcio intracelular
- Diacilglicerol-proteínas cinase C
- Araquidonato-eicosanoides

#### 5 - Citocinas hematopoiéticas Mecanismo de acção



#### 6 - Alguns factores de crescimento e sua acção

<b><u>Factor</u></b>	<b><u>Composição</u></b>	<b><u>Actividades representativas</u></b>
PDGF (factor de crescimento derivado das plaquetas)	AA, AB ou BB / cadeia A = 125AA / cadeia B = 160AA	Estimula a proliferação das células do tecido conectivo e celular da neuroglia
EGF (factor de crescimento epidérmico)	53AA	Estimula a proliferação de vários tipos de células
IGF-I (factor de crescimento like-insulina I)	70AA	Colaboram com o PDGF e EGF; estimulam a proliferação de células adiposas e células do tecido conectivo
IGF-II (factor do crescimento like-insulina II)	73AA	
TGF- $\alpha$ (factor do crescimento transformante)	Duas cadeias, cada com 112AA	Potencia ou inibe a resposta da maioria das células ou outros factores de crescimento, dependendo do tipo celular; regula a diferenciação de alguns tipos celulares
FGF (factor de crescimento fibroblástico)	Acídico:140AA Básico: 146AA	Estimula a proliferação de vários tipos celulares inclusivé fibroblastos, células endoteliais e mioblastos
Interleucina 2 – (IL-2)	153AA	Estimula a proliferação de linfocitos T
NGF (factor de crescimento dos nervos)	Duas cadeias cada com 118AA	Promove o crescimento dos axónios e sobrevivência de neurónio simpáticos e sensorios e do CNS
IL-3, GM-CSF, M-CSF, G-CSF, eritropoietina (factores de crescimento da célula hematopoiética)		

AA – ácidos aminados

7 - Alguns factores estimuladores de colónias (CSF) que influenciam a formação de células sanguíneas

Factor	Tamanho (no rato)	Células alvo	Células produzidas
Eritropoietina	51.000 daltons	CFS-E	Células renais
Interleucina 3 (IL-3)	25.000 daltons	Células estaminais pluripotentes, células progenitoras, células diferenciação	Linfocitos T células epidérmicas
Granulócito / macrófago CSF (GM-CSF)	23.000 daltons	Células progenitoras GM	Linfocitos T células endoteliais fibroblasto
Granulócito CSF (G-CSF)	25.000 daltons	Células progenitoras GM e neutrofilos	Macrófagos fibroblasto
Macrófago CSF (M-CSF)	70.000 daltons (dímeros)	Células progenitoras GM e macrófagos	Fibroblasto macrófago células endoteliais

## 8 - Linfocitos – Locais de formação, diferenciação e compartimentação

### Primários (diferenciação)

- Medula óssea
- Timo

### Secundários (resposta imunitária)

- Baço
- Gânglios
- Subepitélio gastrointestinal
- Sangue e linfa

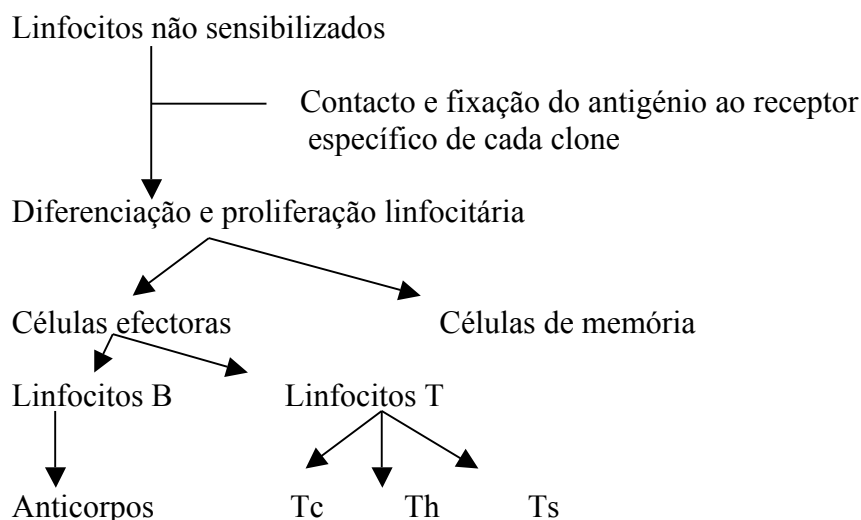
## 9 - Linfocitos – funções e tipos de resposta

### Funções – Defesa do organismo contra substâncias estranhas

### Tipos de resposta

- Celular (Linfocitos T)
  - Maturação no timo, circulantes, resposta tardia
- Humoral (Linfocitos B e plasmocitos)
  - Origem e maturação na medula óssea
  - Fixação no tecido linfático secundário
  - Resposta imediata
- Produção de anticorpos (M, A, G, D e E)

## 10 - Activação linfocitária



As células linfocitárias  $T_H$  (helper) respondem aos antígenos segregando linfocinas, factores que estimulam o crescimento e maturação dentro dos linfócitos das células T e das células B, através de receptores específicos em células alvo. Nem todas as células  $T_H$  segregam todos os factores: diferentes subconjuntos ( $T_{H1}$  e  $T_{H2}$ ) orquestram as respostas a diferentes tipos de infecção através de linfocinas particulares segregadas.

Algumas das linfocinas são chamadas interleucinas (IL).

## 11 - Algumas linfocinas segregadas pelas células T em resposta a antígeno

<b><u>Linfocina</u></b>	<b><u>Função</u></b>
IL-2	Estimula o crescimento das células T Estimula a segregação de Ac pelas células B
Interferão $\gamma$ (IFN $\gamma$ )	Activa macrófagos Destaca classes de Ig das células B
IL-4	Activação das células B Destaca classes de Ig nas células B Estimula o crescimento de células T Activa macrófagos
IL-3	Estimula a diferenciação das células sanguíneas
GN-CSF	Estimula o crescimento e diferenciação de granulócitos e macrófagos
IL-6	Estimula a segregação de Ac pelas células B Outro efeito sobre células não sanguíneas
IL-7	Estimula e crescimento das pré-B e pré-T células

As CTL (células “killer”) ou linfócitos T citotóxicos respondem aos antígenos de duas formas. Numa delas segregam proteínas e proteases capazes de formar canais na membrana da célula alvo,

provavelmente activando estas proteínas a apoptose nestas células alvo. Na outra forma activam o receptor Fas (vide adiante) nas células alvo o que desencadeia a activação da via da apoptose. Estas células citotóxicas T também segregam outras linfocinas.

## 12 - Propriedades de algumas interleucinas (IL)

IL	Nome alternativo	Peso molecular aproximado	Fonte	Alvo	Acção
IL-1		15.000	Células apresentadoras de antigénio	Células helper T	Ajuda a activação
IL-2	Factor de crescimento da célula T	15.000	Algumas células helper T	Todas as células T activadas	Estimulação proliferação
IL-3	Multi-CSF	25.000	Algumas células helper T	Várias células hematopoiéticas	Estimula a proliferação
IL-4	Factor-1 (BSF-1) estimulador das células B	20.000	Algumas células helper T	Células B, células T, « mast cells »	Ajuda a activar e promover a proliferação: aumenta a classe II de moléculas MHC nas células B
IL-5	Factor 2 (BC GF-2) do crescimento das células B	50.000 (dímero)	As mesmas células helper T que fazem IL-4	Células B eosinófilos	Promove proliferação e maturação
IL-6	Factor-2 (BSF-2) estimulador das células B	25.000	Algumas células helper T e macrófagos	Células B activadas, células T	Promove a maturação das células B para células segregadoras de Ig
Interferão		25.000 (dímero)	As mesmas células helper T que fazem IL-2	Células B, macrófagos, células endoteliais	Induz moléculas da classe II MHC e accionam macrófagos

A interleucina 1 (IL-1) tem diversas funções biológicas tais como a de ser um factor produtor de hipertermia (pirogénio), induzir a síntese de prostaglandinas, estar envolvidas na activação e proliferação de linfócitos T, assim como dos linfócitos B através das interleucinas 2.

Existem duas formas diferentes da IL-1 a  $\alpha$  e a  $\beta$  que têm contudo sequências cerca de 25% idênticas. Ambas estas formas são sintetizadas como proteínas precursoras de cerca de 270 resíduos de ácidos aminados que são depois processados post-translacionalmente por cisão da sequência terminal de cerca de 115 resíduos.

O perfil de assinatura destas citocinas é uma região conservada na sua porção C-terminal.

A interleucina 6 é uma importante hormona polipeptídica produzida por diversos tipos celulares que modela a resposta imunitária e está envolvida no controlo da temperatura do corpo e que regula no fígado a síntese das proteínas de “fase aguda”.

O receptor da IL-6 consiste de duas cadeias, uma que liga o ligando e uma glicoproteína indispensável para a transmissão transmembranária do sinal.

A activação das células helper T é pois um processo complexo que envolve diversas proteínas interleucinas segregadas que actuam localmente como mediadores químicos. Algumas células helper T activam macrófagos segregando interferões gama ou seja uma interleucina.

## **5 – Necrose e apoptose**

### **5.1 – Sinalização por diversas vias**

A apoptose ou morte programada das células é uma forma ubíqua de morte celular muito importante durante a embriogénese, desenvolvimento e homeostase celular no estado adulto.

A alteração deste processo de morte celular fisiológica normal, pode originar quer apoptose excessiva quer apoptose insuficiente o que pode desembocar em vários estádios de doenças e patologias.

A maioria das células contem maquinaria para levar à apoptose e também inerentes mecanismos de repressão desta. A apoptose é regulada e programada para a remoção ordenada de células supérfluas, envelhecidas ou alteradas.

Em cada segundo, milhões de células do corpo humano e animal sofrem apoptose. Uma anormal resistência à apoptose está correlacionada com malformações, doenças auto- imunes ou cancro devido à persistência de células supérfluas, imunocitos self-específicos ou células mutadas, respectivamente.

Pelo contrário, uma apoptose favorecida participa em patologias agudas (infecção por toxinas microbianas, isquémias ou enfartes) ou doenças crónicas (neurodegenerativas, neuromusculares, SIDA).

Enquanto a apoptose é necessária para a saúde e para a doença a necrose é sempre consequência de injúrias, severas e agudas.

As causas clássicas de necrose incluem a hipertermia, a inibição da fosfatação oxidativa, da glicólise ou do ciclo de krebs, autólise, hipoxia, complemento e diversas toxinas.

A apoptose, contrariamente à necrose, necessita da acção regulada de enzimas catabólicas (proteases e nucleases) dentro dos limites de uma membrana plasmática quase intacta, sendo acompanhada por alterações da morfologia celular (condensação da cromatina, picnose, cariorexis), fragmentação da cromatina com libertação de monómeros e oligómeros de 200 pares de bases, e activação de caspases. Subtis alterações na plasma membrana ocorrem antes da sua rotura, com libertação do citosol e seus organelos para o espaço inter celular.

Durante a apoptose, as mitocôndrias não manifestam anormalidades estruturais importantes. Pelo contrário a necrose não envolve a degradação regular do DNA e das proteínas e é acompanhada pela dilatação de todo o citoplasma (oncosis) e da matriz mitocondrial o que ocorre pouco antes da rotura da membrana celular.

A morte celular fisiológica (apoptose) e em alguns casos a morte celular acidental (necrose) implicam um processo em duas etapas.

Na primeira etapa numerosos estímulos fisiológicos e alguns patológicos desencadeiam um aumento da permeabilidade mitocondrial que leva à libertação de factores aptogénicos através da membrana externa mitocondrial e a uma dissipação do gradiente electroquímico do folheto interno mitocondrial. Esta transição da permeabilidade mitocondrial (PT) é o acontecimento limitante do processo da morte celular.

Na segunda etapa as consequências do disfuncionamento mitocondrial [colapso do potencial transmembranario do folheto interno da mitocôndria (que em condições normais é quase impermeável) desacoplação da cadeia respiratória, superprodução de aniões superóxido, rotura da biogénese mitocondrial, perda do cálcio da matriz e da glutatona], desencadeia uma catástrofe bioenergética que leva à rotura da integridade da plasma membrana (necrose) e / ou à activação de caspases apoptogénicas pelas proteínas mitocondriais que saem para o citosol (citocromo C, factor indutor da apoptose) com activação secundária das endonucleases (apoptose).

A velocidade relativa destes dois processos ( catástrofe bioenergenética relativamente à activação proteasica e endonucleasica ) determina quando uma célula sofre necrose primária ou apoptose.

As diferenças fundamentais entre apoptose e necrose referem-se no quadro seguinte.

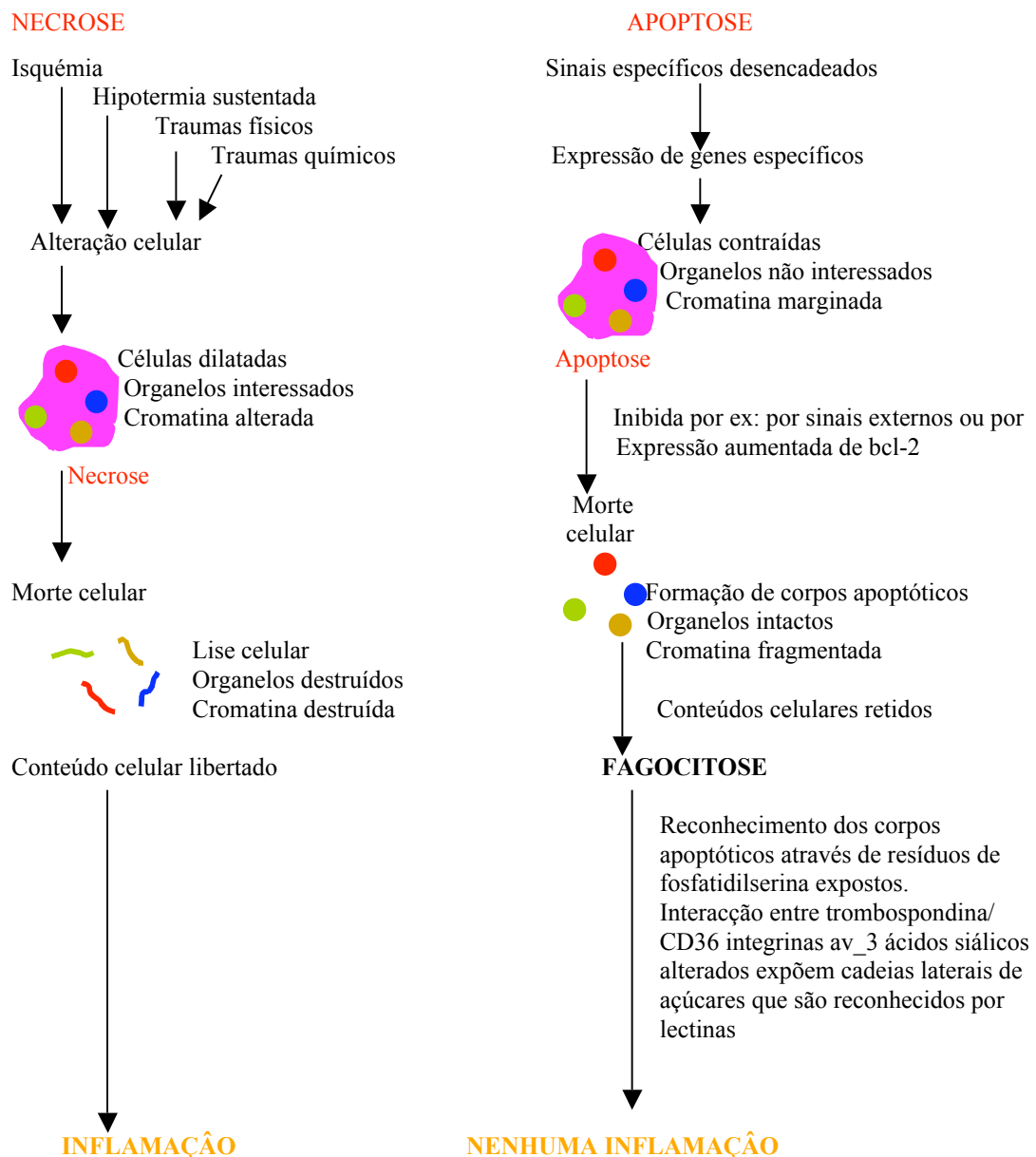
#### Comparação da apoptose com a necrose

<u>Apoptose</u>	<u>Necrose primária</u>
Fisiológica ou patológica (alteração subnecrotica) Susceptibilidade finamente regulada Plasma membrana sensivelmente intacta até tarde Eliminação heterofágica Sem perda do conteúdo celular, pouca ou nenhuma inflamação Enzimas celulares induzem picnose, cariorexis, endonucleose Degradação selectiva de proteínas por proteases específicas (caspases) Pequenas alterações na plasma membrana (perda da assimetria antes da perda da integridade da membrana) “Encolhimento” celular nenhuma dilatação mitocondrial	Acidental Sempre patológica Não regulada ou regulada pobremente Plasma membrana destruída cedo Perda de contactos celulares Inflamação Dilatação de todo o citoplasma (oncosis) Dilatação mitocondrial  <div style="text-align: center;"><u>Necrose secundária</u></div> Citólise secundária à apoptose quando células mortas são removidas por heterofagia
Os aspectos morfológicos de ambos os modos de morte celular (–mitocôndrias normais ou mitocôndrias condensadas na apoptose e mitocôndrias dilatadas na necrose )– não devem levar à concepção errónea de que as mitocôndrias são as principais intervenientes na necrose mas são irrelevantes na apoptose. Com efeito as alterações nas membranas mitocondriais são críticas para a apoptose e para a necrose.	

A apoptose difere portanto da necrose sendo esta última consequência por exemplo de injúrias (inflamação, feridas) ataques pelo complemento, hipoxia severa, hipotermia, infecções por vírus líticos, ou exposição a toxinas que eventualmente desencadeiam a lise celular.

A necrose é pois passiva, sendo um processo catabolico .

### Comparação da morte celular por necrose e por apoptose



A apoptose ao contrario da necrose, necessita de um sistema funcional produtor de energia.

O balanço entre a morte por apoptose e por necrose parece depender da intensidade da agressão e do ATP intracelular disponível.



Os primeiros sinais da morte das células apoptóticas são alterações morfológicas das células tais como condensação da cromatina, desaparecimento do nucleolo e alteração das superfícies celulares onde aparecem vesículas (“blebs”). A estes sinais seguem-se a marginalização da cromatina para a superfície interna da membrana nuclear, eventualmente a alteração de uma série de nucleases que fragmentam o DNA, primeiro em grandes fragmentos e depois em fragmentos do tamanho dos nucleossomas.

As células que morrem por apoptose (ao contrário de que sucede na necrose) contraem-se e eventualmente quebram-se em corpos apoptóticos, não libertando os conteúdos intracelulares não ocorrendo portanto inflamação ao contrário do que sucede na necrose.

A necrose é um processo passivo e a indução da apoptose é um processo activo regulado geneticamente que necessita de expressão coordenada de diversos genes, e uma vez posto em movimento é irreversível.

Apoptose é um programa de suicídio celular intrínseco envolvido no turnover normal dos hepatócitos e de outros tipos celulares.

A apoptose envolve a condensação do conteúdo celular, a quebra da membrana nuclear e a formação de corpos apoptóticos que são pequenas vesículas envolvidas por membrana e que são fagocitadas pelas células vizinhas.

A apoptose é pois um processo fisiológico para a morte das células, e que é crítico para o desenvolvimento normal e para o funcionamento dos organismos pluricelulares, originando as anormalidades no controlo da morte das células doenças como por exemplo: cancro, desordens de auto-imunidade e degeneratórias.

Os sinais para a apoptose podem ser intracelulares ou extracelulares, e todos eles convergem activando um grupo de proteases com cisteínas específicas da apoptose, as chamadas caspases que quebram os seus substratos após os resíduos de ácido aspártico.

O dismantelamento e a remoção das células condenadas é realizada por proteólise dos constituintes celulares, com degradação do DNA, e fagocitose pelas células vizinhas.

Portanto a condensação da cromatina a fragmentação do DNA em fragmentos nucleossomais, a quebra da membrana nuclear e a formação dos corpos apoptóticos devem-se à actividade das caspases.

Vejamos os mecanismos bioquímicos implicados na apoptose (ver figura seguinte).

A fragmentação do DNA é mediada por um factor heterodimérico de 40 e 45Kda (nos humanos chamado DFF e nos ratos CAD existindo nestes um inibidor do CAD chamado ICAD) que não têm qualquer homologia na sua sequência com outras proteínas de função conhecida.

Nas células apoptóticas a DFF45 que tem dois locais caspases de quebra, é cindida em três fragmentos mais pequenos, dissociando-se depois a DFF45 da DFF40 e induzindo a oligomerização da DFF40 formando-se um complexo com actividades de Dnase.

Nas células vivas as caspases existem como zimogénios inactivos, como o DFF, que são activadas por cisão proteolítica através de duas vias (vide figura anexa).

Numa via que envolve receptores na superfície celular que assinalam morte celular, como é o caso do receptor Fas e do receptor para o factor de necrose tumoral (TNF), é induzida a activação intracelular da caspase 8.

Os ligandos Fas e TNF que habitualmente se encontram na forma de trímeros, ligam-se e activam os seus receptores induzindo a trimerização destes.

Estes receptores activados por seu turno recrutam moléculas adaptadoras tais como as FADD/MORT1 (Fas-associated protein with death domain) as quais recrutam procaspase 8 para o complexo receptor, onde sofrem uma activação autocatalítica. A caspase 8 activada quebra e activa outras caspases como a 3,6 e 7 constituindo a principal actividade caspase das células apoptóticas.

Na outra via para activação das caspases ocorre a libertação do citocromo C das mitocôndrias. Este citocromo é uma proteína solúvel exclusivamente localizada no espaço intermembranário mitocondrial e ele é libertado durante a apoptose, devido ao facto da membrana externa (folheto externo) das mitocôndrias se tornar permeável a ele, e quando libertado liga-se á Apaf1 (apoptotic protease – activating factors).

A Apaf1 é um monómero citosólico de 130Kda consistindo de três domínios distintos. A indução da apoptose leva o Apaf1 a formar um complexo multimérico com o citocromo C, e estes complexos recrutam e activam a procaspase 9, que activada liberta-se do complexo e activa as caspases 3, 6 e 7.

A libertação do citocromo C é regulada pela família de proteínas Bcl-2, família esta que contem membros pré-apoptóticos (Bak, Bim, Bad e Bax) e membros anti-apoptóticos (Bcl-2 e Bcl-xl).

A super expressão do anti-apoptótico Bcl-2 ou do seu homólogo Bcl-xl bloqueia a libertação do citocromo C induzida por uma série de estímulos apoptóticos. Pelo contrário o Bax, Bak e Bid libertam directamente o citocromo C “in vivo” e “in vitro”.

Sinais de morte celular intrínsecos ou extrínsecos podem ser transmitidos para os mitocôndrias a partir de diferentes compartimentos celulares.

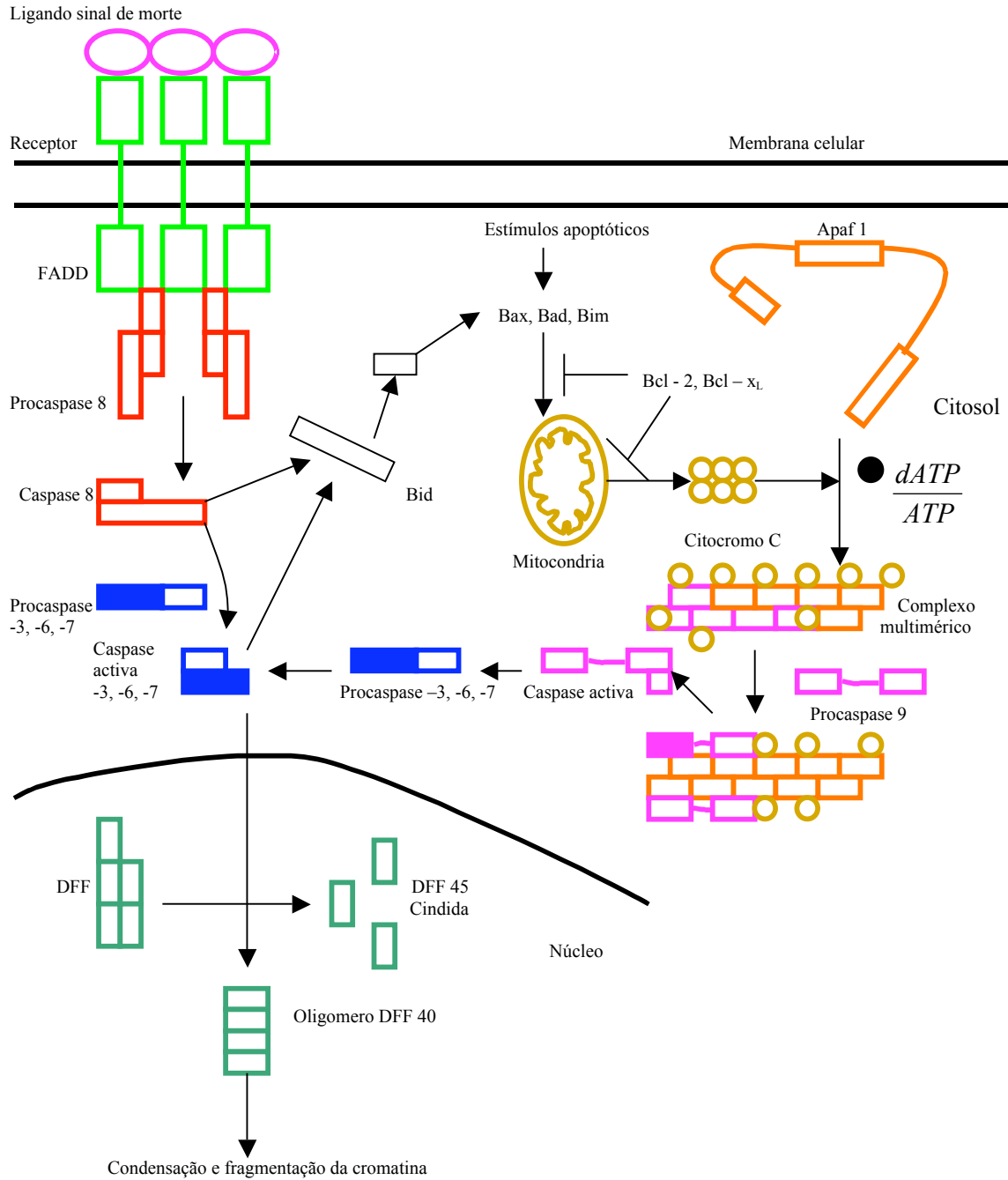
Sinais extra celulares com os ligandos Fas ou TNF activam a caspase 8 intracelularmente e esta activada cinde e activa a Bid (um membro pró-apoptótico da família Bcl-2) que é translocada para a mitocôndria onde induz a libertação do citocromo C ampliando o sinal de activação das caspases.

Sinais de sobrevivência extra celulares inibem a apoptose activando a via fosfatidilinositol-3-cinase/Akt levando à fosfatação da Bad. Esta Bad fosfatada liga 14-3-3 proteína e é sequestrada no citosol, enquanto desfosfatada a Bad transloca-se para a mitocôndria. O cálcio ( $Ca^{2+}$ ) pode induzir a apoptose activando uma fosfatase dependente de calcineurina que desfosfata a Bad. Outros sinais intrínsecos de morte celular podem regular a translocação da Bim. Nas células vivas normais a Bim liga-se á LC8, um componente do citoesqueleto.

Também a Bax se transloca do citoplasma para as mitocôndrias durante a apoptose regulando a libertação de citocromo C e a activação das caspases.

## Apoptose por duas vias diferentes para activação das caspases

- Fas / FADD / caspase 8
- Citocromo C / Apaf 1 / caspase 9



Nos timocitos e fibroblastos embrionários, a caspase 9 e o seu adaptador Apaf 1 são necessários para a morte celular, mas já são dispensáveis para a apoptose induzida pelo CD 95 (Fas / APO 1) e pelo TNF R1.

Pelo contrário a caspase 8 e o seu adaptador o FADD são necessários para o CD 95 e TNF R1 transduzirem a apoptose, mas são dispensáveis para outras vias indutoras da morte celular.

Parece pois que os mamíferos têm dois mecanismos distintos para activar as caspases efectoras. Um é iniciado pelos sinais induzidos pelo stress dentro das células, requer Apaf 1 e caspase 9 e é regulada pela família de proteínas Bcl-2. O outro é activado pelo CD95 e receptores relacionados, requer FADD e caspase 8 e não pode ser bloqueado pelas proteínas da família Bcl-2.

Parece que a via CD 95 \_ FADD \_ caspase 8 é de importância fisiológica.

Não é bem claro ainda se as três vias mais especulativas de assinalamento pelo CD 95 ou seja a CD 95 \_ RIP \_ RAIDD \_ caspase 2, a CD 95 – Daxx \_ ASK 1 \_ JNK \_ caspase desconhecida e a CD 95 \_ FLASH \_ caspase desconhecida ou adaptador são de importância fisiológica.

Um grande número de compostos químicos e biológicos podem desencadear a apoptose pelo menos em alguns tipos de células e em certas condições.

A morte celular é crítica para o desenvolvimento dos animais.

A perda celular por apoptose ocorre em diversos tecidos e em múltiplos estádios da diferenciação celular.

Também é sabido que apoptoses anormais podem desempenhar um papel no desenrolar das doenças.

Em cada órgão o seu número de células depende das proporções entre a migração celular, a divisão celular e a morte celular.

A morte das células nos organismos pluricelulares encontra-se sob controlo genético havendo muitos estudos recentes sobre os receptores para sinais de morte e os ligandos de sinais para este efeito, estando também identificadas vias de transdução de sinais distintas, bem como os componentes desta maquinaria efectora e respectivas funções bioquímicas.

Na apoptose intervêm quatro grupos funcionais de moléculas que são as caspases, as proteínas adaptadoras que controlam a activação das caspases iniciadoras, membros da superfamília de receptores (TNF-R) dos factores de necrose tumoral (TNF) e membros da família Bcl – 2 de proteínas.

Vejamos algumas das características mais importantes destes quatro grupos de moléculas.

## **5.2 – Caspases**

São cisteína proteases que podem ser detectadas em todas as células sofrendo apoptose, independentemente da sua origem ou do estímulo que a desencadear, estando identificadas catorze caspases nos mamíferos. Estas enzimas reconhecem motivos tetrapeptídicos e cindem os seus substratos no lado carboxílico de um resíduo de aspartato.

As caspases são responsáveis não só pela degradação celular dos substratos durante a fase final da apoptose mas são também importantes reguladores da iniciação da morte celular.

Cada caspase tem diferentes especificidades que dependem da sequência de ácidos aminados acima do (local) da cisão (as posições P2 – P4).

As caspases são produzidas na forma de zimogénios quase inactivos, sendo as enzimas plenamente activas constituídas por heterotetrameros compostos por duas sub-unidades idênticas com cerca de ~20Kda mais duas sub-unidades idênticas de ~10Kda.

As chamadas caspases iniciadoras (por exemplo a caspase 8 e caspase 9) iniciam uma avalanche de crescente actividade das caspases, ao processar e activar as chamadas caspases efectoras.

Algumas caspases, sobretudo as efectoras quebram e inactivam proteínas vitais das células como é o caso das enzimas de reparação do DNA, laminina, gelsolina, e uma proteína cinase C6.

Durante a apoptose, o i CAD é cindido por caspases o que origina a libertação de uma endonuclease activa que quebra o DNA internucleossomas.

As caspases 2, 9, 8 e 10 são iniciadoras de apoptose enquanto as caspases 3, 6, 7 são caspases efectoras.

Nas caspases 8 e 10 existem dois domínios efectores de morte (DED) e nas 2,9,1 e 4 existem domínios de recrutamento de caspases (CARD).

### **5.3 - Proteínas adaptadoras**

Estas proteínas adaptadoras ou de acostagem ou “docagem” tal como os ligandos e receptores têm uma estrutura em módulos, podendo diversas proteínas de docagem ligar-se umas às outras através de módulos específicos distintos daqueles outros módulos que lhes permitem ligar-se a outras moléculas assinalantes.

Estas proteínas ligam os efectores de morte celular ou sejam as caspases e os reguladores de morte celular, os receptores de morte, e os membros da família Bcl – 2, formando associações físicas entre si.

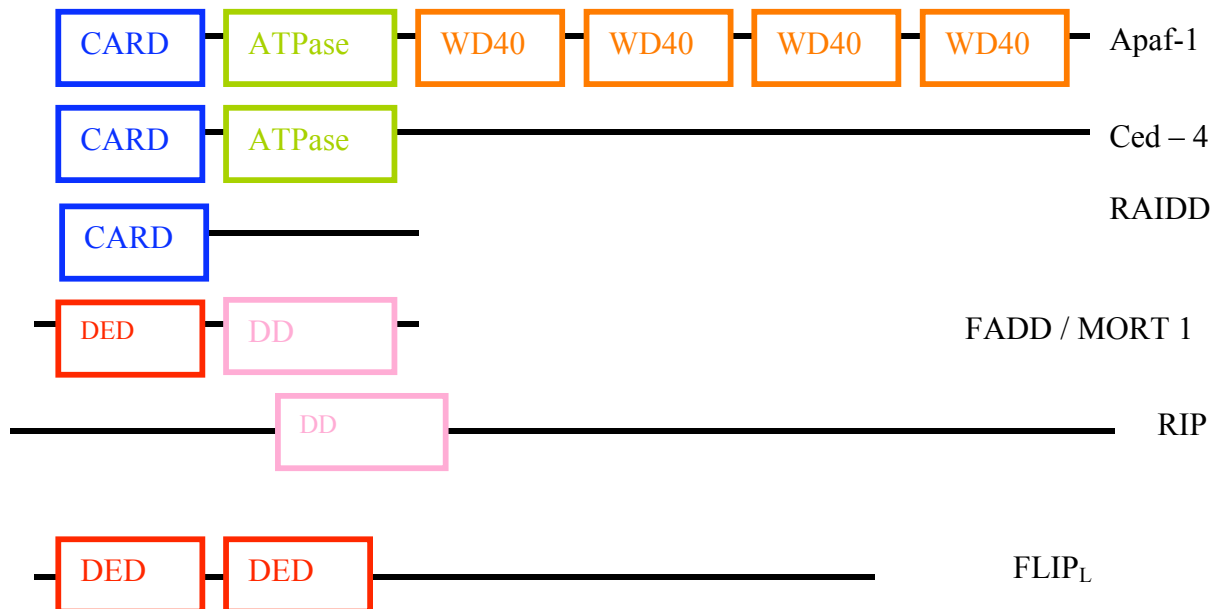
Assim formam-se associações entre as proteínas adaptadoras e caspases ou membros da família de TNF – R por interacção homotípica entre os domínios chamados domínios de morte (DD), o domínio efector de morte (DED) e o domínio de recrutamento de caspases (CARD) ( vide figura seguinte ).

O motivo DD que serve como local de docagem é composto por seis regiões anfipáticas \_ – helicoidais arranjadas de uma forma antiparalela umas em relação às outras e são transdutoras de alterações de conformação.

Acima e abaixo das regiões DD encontram-se outras regiões consoante as moléculas com propriedades distintas e específicas. A região DD encontra-se no lado citoplásmico do CD95 (ou Fas/APO 1, é um receptor da família de receptores TNF) e de outros membros relacionados com a família dos TNF-R (receptores do factor de necrose tumoral), assim como em moléculas adaptadoras como é o caso da FADD/MORT1 (Fas – associating death domain protein / mediator of receptor – induced

toxicity), TRADD (TNF-R1 – associated death domain protein) e RIP (proteína interagindo com o receptor).

### Proteínas adaptadoras



DD = Domínio de morte

DED = Domínios efectores de morte

CARD = Domínios de recrutamento de caspases

Apaf-1 e Ced4 têm domínios ATPase conservados

Apaf-1 tem repetições WD40 na extremidade C

FLIP<sub>L</sub> inibe a FADD

Após interacção de um membro da família TNF-R contendo um domínio DD, com uma molécula adaptadora contendo também um domínio DD é viabilizada a agregação de caspases e sua activação. Este recrutamento e agregação das caspases pode ser mediada ainda por outros domínios, os DED existentes nas moleculares adaptadoras.

Repetições DED em tandem encontram-se também nos zimogénios das caspases 8 e 10.

As interligações cruzadas do receptor CD95 originam também a agregação dos pró-caspase 8 e sua activação através da molécula adaptadora FADD.

A via assinalante receptor de morte \_ FADD - caspase 8 pode ser bloqueada por moléculas de FLIP que evitam o recrutamento e activação da pro-caspase 8.

Nem todas as caspases iniciadoras contêm domínios DED e daí nem todas elas serem activadas através das interligações cruzadas do TNF-R.

As pró-caspases 9 e 2 dos mamíferos contêm domínios CARD que também existem nos seus adaptadores específicos Apaf-1 e Ced-4.

#### 5.4 - Família de receptores do factor de necrose tumoral (TNF-R ou receptores de morte)

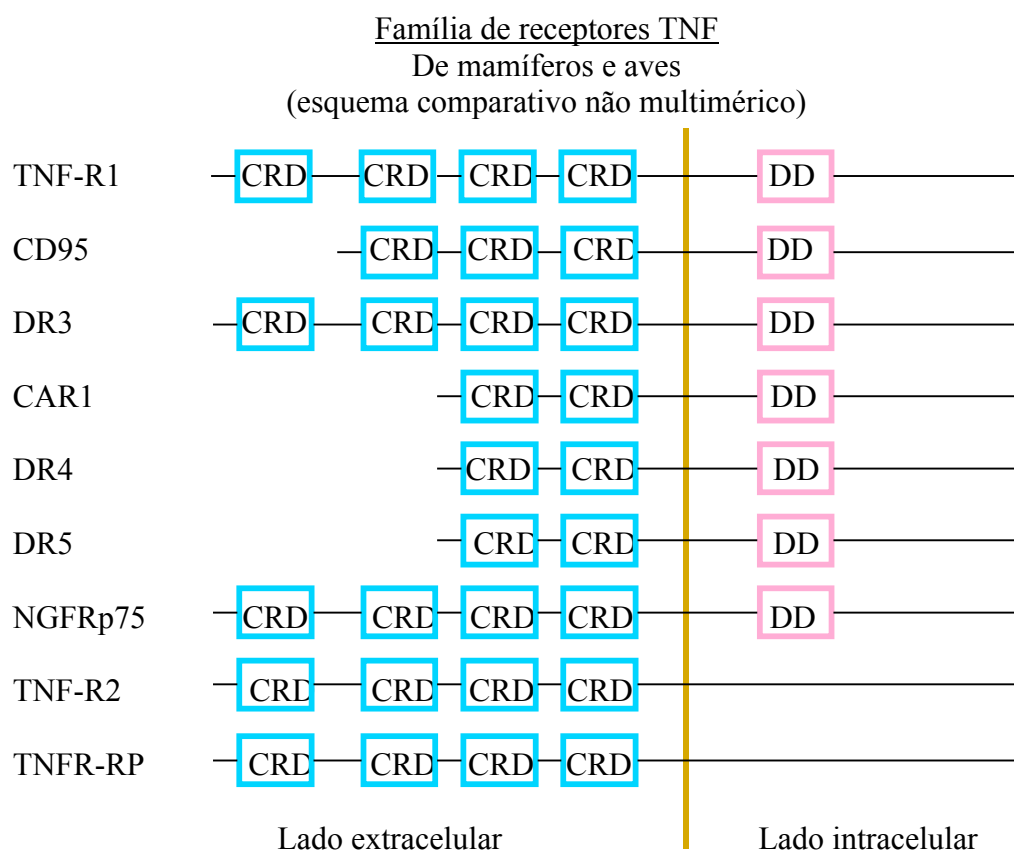
Esta família tem acção pleiotrópica, podendo desencadear proliferação celular, sobrevivência, diferenciação ou morte tudo dependendo do tipo de célula e dos sinais que ela recebe.

Estes receptores são activados por uma série de ligandos relacionados no que se refere à sua estrutura (vide figura seguinte).

A maioria destes ligandos são sintetizados na forma de trímeros ligados à membrana celular e parece que para que o sinal seja transmitido é necessário que se estabeleça uma extensa interligação cruzada dos receptores.

Os membros da família de receptores TNF funcionam como trímeros e multímeros de trímeros.

Os receptores CD95 (também chamado Fas ou APO-1), o TNF-R1 e outros membros desta família contem uma região DD no lado citoplásmico que é fundamental para a indução da apoptose.



CRD = Repetições ricas em cisteína

DD = Domínios de morte

Os receptores da família TNF podem também activar reguladores da transcrição como é o caso do NF – KB e da Jun cinase (JNK).

O NF – KB é um grupo de factores de transcrição que controla a expressão de diversos genes que participam na inflamação e na resposta imunitária. Na maioria das células estes factores encontram-se latentes pois são inibidos por um grupo de proteínas chamado I-KB que após fosfatação é degradado libertando a acção da NF-KB. As proteínas controladas pelo NF – KB incluem várias que contribuem para a acção próinflamatória do TNF –  $\alpha$ .

A transmissão de sinais através da TRADD, cinases RIP, cinase induzida pela NF – KB (NIK), cinase IKB (IKKa) e IKKB culmina na fosfatação e degradação da família I - KB dos inibidores da NF – KB. Estes sinais promovem sobrevivência celular em vez de apoptose o que apoia o papel anti-apoptótico das vias NF – KB e JNK(vide mais adiante).

## **5.5 - Família Bcl – 2**

Esta família é o protótipo do inibidor da apoptose e engloba uma série de membros tais como as proteínas mamíferas Bcl-2, Bcl-X<sub>L</sub>, Bcl-W, A1/B/i-1, Mcl-1 e Boo/Diva e uma série de proteínas homólogas víricas que promovem a sobrevivência celular e têm três ou quatro regiões com uma extensa semelhança nas sequências em ácidos aminados com a Bcl-2 (vide figura seguinte).

Outros membros da família mais distantes no relacionamento estrutural englobam a Bax, Bcl-xs, Bak e Bok/Mtd que contem duas ou três regiões de homologia Bcl-2 ( BH )e favorecem a apoptose em condições de stress.

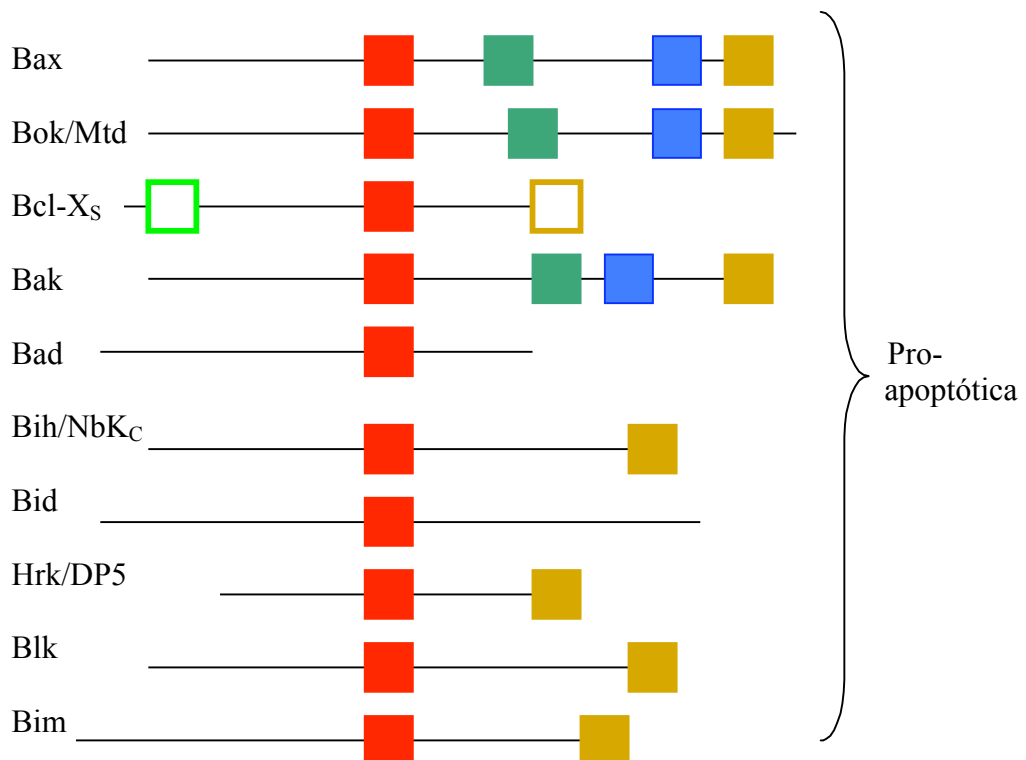
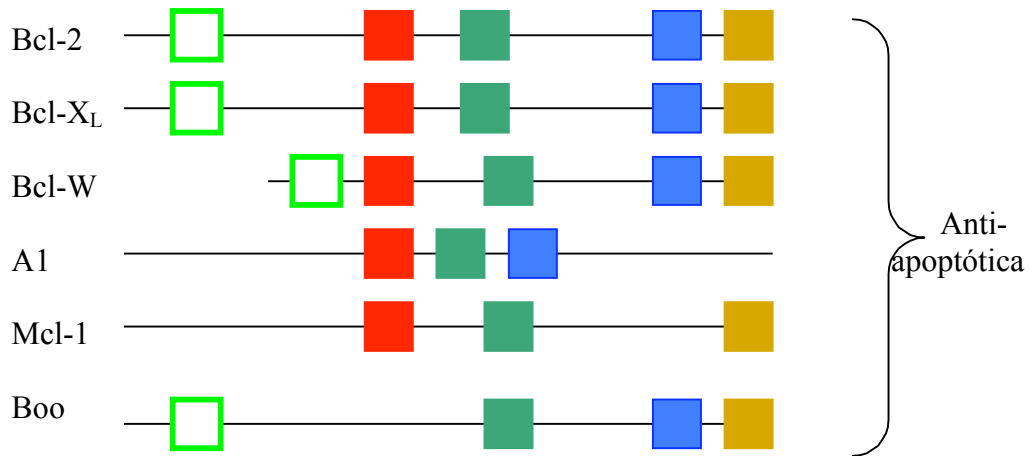
Os mais potentes indutores de morte celular são o grupo que possui apenas uma região BH3 como é o caso da Bad (vide figura), Bik/Nbk, Bid, Hrk/DP5, Blk e Bim.







Os membros pró-apoptóticos e anti-apoptóticos desta família Bcl-2 podem interactuar entre si.



Família Bcl-2  
Comparação estrutural

Proteína



-  = Domínios transmembranários
-  = Domínios de homologia Bcl-2( BH2)
-  = Domínios de homologia Bcl-2( BH1)
-  = Domínios de homologia Bcl-2( BH3)
-  = Domínios de homologia Bcl-2( BH4)
-  = cadeia polipeptídica

## 5.6 – Sinalização pelos receptores de morte CD95, CD120a e DR3

Certas células têm na sua superfície sensores únicos, chamados receptores de morte que detectam a presença de sinais de morte extracelulares e que respondem ligando imediatamente um mecanismo celular intrínseco de apoptose.

A apoptose desempenha nos animais um papel muito importante no seu desenvolvimento e na sua homeostase.

As células morrem durante a morfogénese ou sinaptogénese no embrião em desenvolvimento e nos animais adultos durante o “turnover” dos tecidos ou no fim de uma resposta imunitária.

Os receptores de morte pertencem à superfamília de genes para os receptores do factor de necrose tumoral (TNF) sendo caracterizados por possuírem domínios extracelulares similares ricos em cisteína, contendo ainda uma sequência homóloga citoplásmica chamada “domínio da morte” (DD).

Estes domínios da morte ligam os receptores de morte com a maquinaria apoptótica da célula mas em alguns casos podem desencadear acções opostas ou que contrariam a apoptose. Algumas moléculas que transmitem sinais a partir de receptores de “morte” contêm elas próprias domínios de morte.

Os receptores de morte melhor conhecidos são o CD95 (Fas ou Apo1) e o TNF-R1 (p 55 ou CD120a) e ainda o receptor 3 de morte DR3 ou Apo3.

Os ligandos que activam estes receptores, com excepção do NGF (nerve growth factor) são moléculas relacionadas estruturalmente que pertencem à superfamília de genes TNF.

Quatro membros da família de ligandos do factor de necrose tumoral (TNF), os TNF<sub>1</sub>, LT<sub>1</sub>, LT<sub>2</sub> e LIGHT interactuam com quatro receptores da família TNF factor de crescimento dos nervos, o p55 TNF receptor (CD120a), o p75 TNF receptor (CD120b) e o receptor beta linfotóxina (LT<sub>2</sub>R) e o mediador da entrada do vírus herpes (HVEM) para controlar uma larga gama de funções de resposta da imunidade inata e adquirida, sobretudo a indução e regulação da morte celular e o processo inflamatório.

Com excepção do ligando LT<sub>2</sub> que é segregado pelas células todos os outros membros desta família de ligandos são formados como proteínas transmembranárias do tipo II (que tem a extremidade N do lado citoplásmico e a região C-terminal estendendo-se para o lado extracelular) podendo assim actuar de uma forma justacrínica.

Alguns são no entanto sujeitos a proteólise.

Os membros da família TNF parecem actuar na forma de trímeros, na maioria dos casos como homotrímeros.

A actividade assinalante da família de receptores TNF é desencadeada pela justaposição dos seus domínios intracelulares após ligação das moléculas ligandos de estrutura trimérica aos seus domínios extracelulares induzindo translocações e alterações conformacionais.

O ligando CD95 (CD95L) liga-se ao CD95, o TNF e a linfotóxina<sub>1</sub> ligam-se ao TNF-R1 e o ligando Apo3 (Apo3L ou TWEAK) liga-se ao DR3.

Vejamos com algum detalhe as vias assinalantes pelo CD95, pelo TNF-R1 e pelo DR3 (vide figuras seguintes).

### Sinalização pelo CD95

O CD95 e o CD95L têm um papel importante em três tipos de apoptose fisiológica a saber:

- 1) Na detecção periférica das células T maduras activadas quando se chega ao fim da resposta imunitária.
- 2) Na morte por células citotóxicas T de células alvo como sucede nas células infectadas por vírus ou nas células cancerosas.
- 3) Na morte de células inflamatórias em locais de resposta imunitária privilegiada como é o caso dos olhos.

O CD95L é uma molécula trimérica como outros membros da família TNF, que se parece ligar a três moléculas de CD95 originando um “aglomerado” de domínios de morte dos receptores.

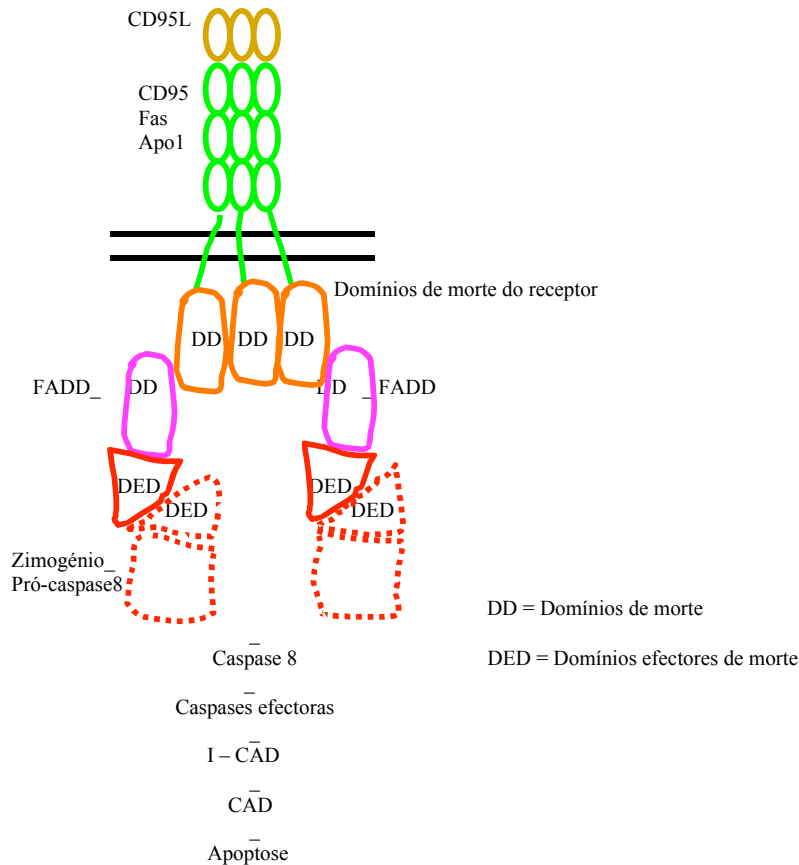
Uma proteína adaptadora FADD (Fas-associated death domain) ou Mort1 liga o seu próprio domínio de morte (DD) ao aglomerado dos domínios de morte dos receptores. O FADD por outro lado contém um domínio efector de morte (DED) que se liga a um domínio análogo repetido em tandem que existe no zimogénio da caspase 8 (também chamado FLICE ou MACH), o que leva a oligomerização da caspase 8 que conduz a sua activação por cisão de si própria.

A caspase 8 activa depois outras caspases como a 9 que por ser turno activa a CAD (uma enzima que participa na cisão do DNA) por inactivação de uma proteína inibidora (ICAD) que mantinha a CAD no citoplasma e que depois permite a translocação da CAD para o núcleo.

O Fas/Apo1 (CD95), um receptor da família de receptores TNF activado por um ligando distinto, induz a morte celular através de mecanismos comuns aos do CD120a (vide adiante).

O CD95 além de induzir a morte das células pode também estimular o crescimento celular e induzir a síntese das citocinas interleucinas 6 e 8.

## Apoptose induzida pelo CD95



## Sinalização pelo TNFR1 ou CD120a

O TNF é o principal mediador das reacções inflamatórias, imunológicas e patofisiológicas.

O TNF é produzido sobretudo por macrófagos activados e por células T em resposta as infecções e é também expresso nas células endoteliais e outros tipos de células, e através do receptor respectivo TNFR1, activa os factores de transcrição NF-KB e AP-1 originando a indução de genes pró-inflamatórios e imunomoduladores, raramente desencadeando apoptose a não ser que a síntese proteica se encontre bloqueada.

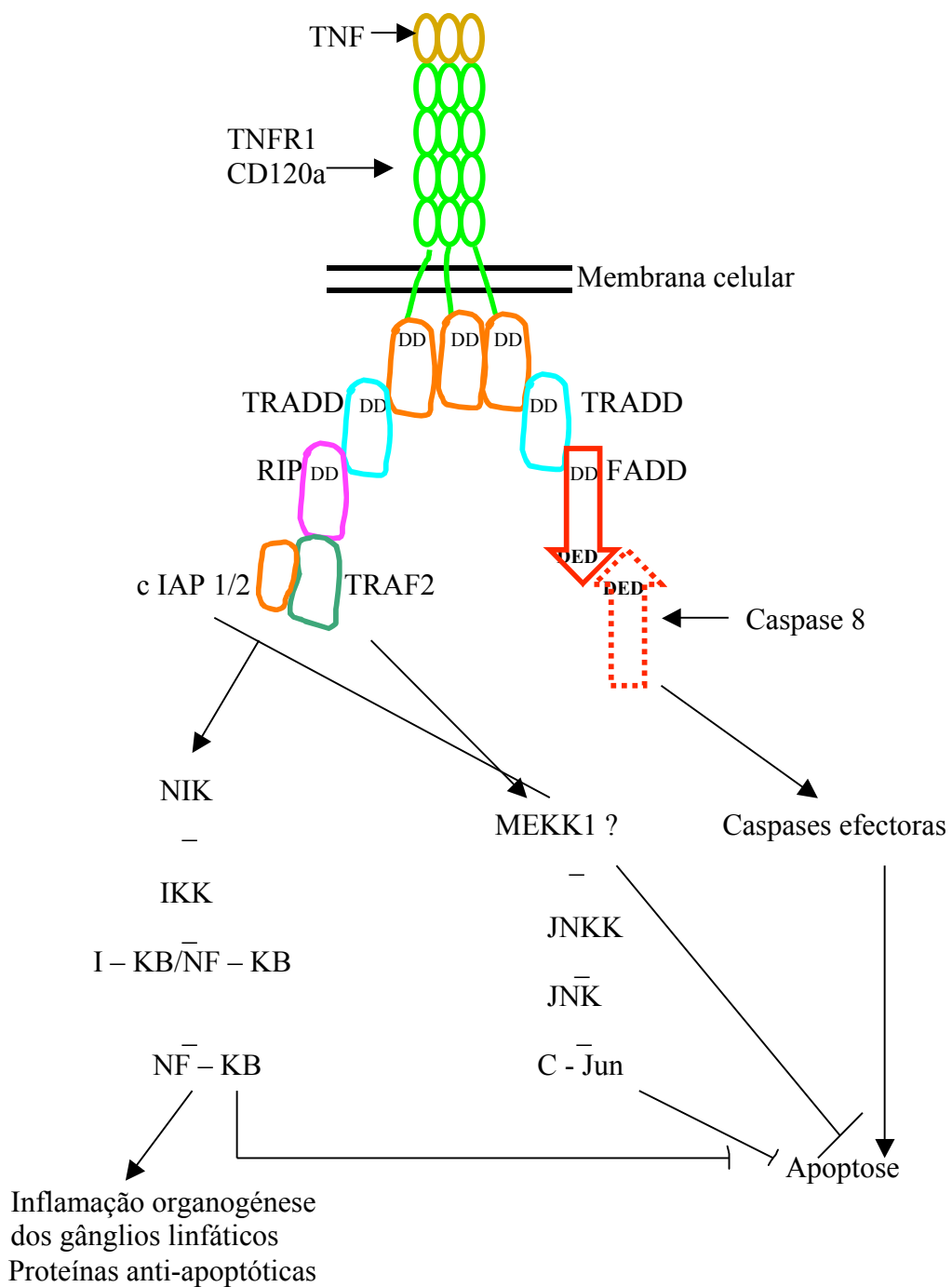
Também o TNF trimeriza os receptores TNFR1 levando à associação dos domínios de morte (DD) no lado citoplásmico dos receptores e posteriormente uma proteína adaptadora a TRADD (TNFR – associated death domain) liga-se através dos seus próprios domínios DD aos domínios DD do receptor. A TRADD recruta depois diversas moléculas assinalantes para o receptor activado como é o caso da TRAF-2 (TNFR – associated factor 2) e a RIP (receptor interacting protein) que estimulam a via que leva à activação do NF –KB e da JNK/AP-1, enquanto a FADD medeia a activação da apoptose (vide figura anexa).

O NF – KB consiste de duas subunidades (p50 e p65) e existe na forma de um complexo com o IKB ( seu inibidor ) nas células em repouso.

A TRAF2 e a RIP activam a NIK (uma cinase indutor do NF – KB) que por seu turno activa a IKK que é inibidor I – KB do complexo cinase. A IKK fosfata a I – KB levando esta à degradação pelo proteosoma permitindo depois a translocação do NF – KB para o núcleo para activar a transcrição. A via do TRAF2 e RIP para a JNK envolve uma cascata complexa que leva à activação do C – Jun um protooncogene que codifica parte do factor de transcrição AP1 um “enhancer” de muitos genes.

O CD120a controla os mecanismos de defesa coordenando o processo inflamatório e ao nível de todo o organismo animal, induz mudanças tais como a febre, a perda de apetite ou a elevação das proteínas de fase aguda do soro sanguíneo. Este receptor controla ainda a defesa ao nível da célula individual, induzindo a morte das células atacadas por agentes patogénicos.

Sinalização pró-apoptótica e antiapoptótica do TNFR1



### Sinalização pelo DR3

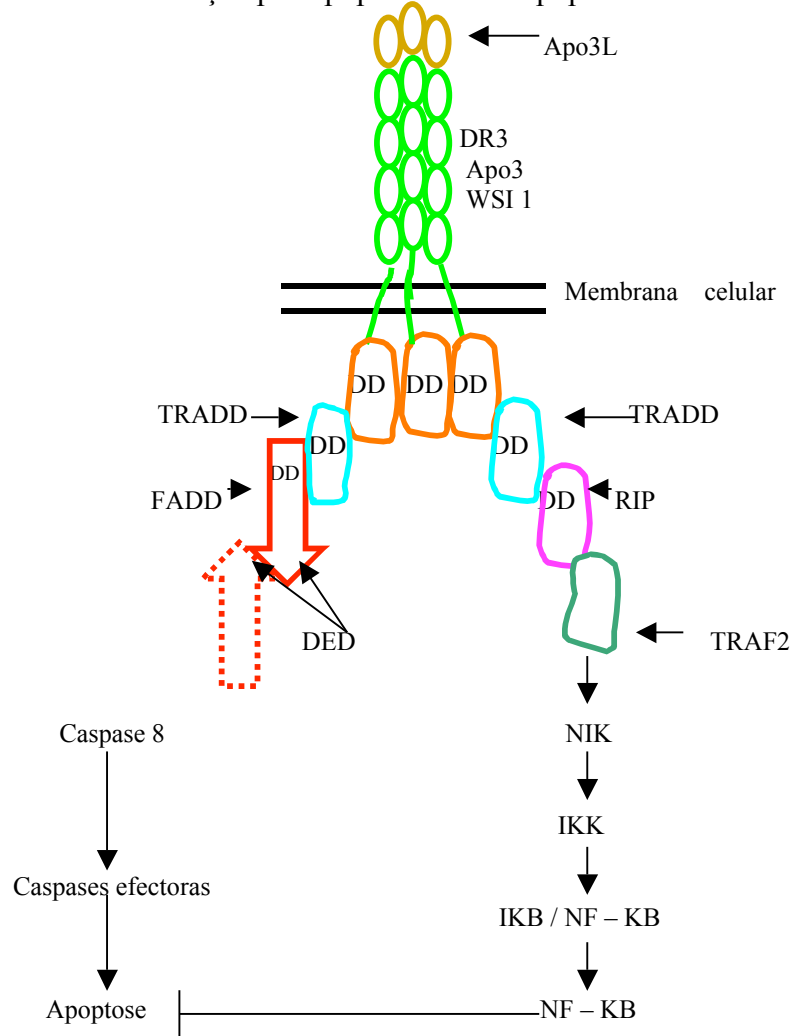
O DR3 desencadeia respostas semelhantes às desenvolvidas pelo TNFR1, sobretudo a activação do NF – KB e a apoptose.

Na activação do NF – KB fá-lo através da TRADD, TRAF2 e RIP como o TNFR1 e na apoptose através do TRADD, FADD e caspase 8 (vide figura seguinte).

A expressão TNF ocorre sobretudo nos macrófagos activados e nos linfocitos enquanto o mensageiro RNA do Apo3L é expresso constitutivamente em vários tecidos.

O TNFR1 é expresso de forma ubíqua enquanto os transcritos de DR3 se encontram sobretudo no baço, timo, e sangue periférico.

Sinalização pró-apoptótica e antiapoptótica e DR3



## 6-Ciclo celular

### 6.1 - Ciclina e CdK

O ciclo celular origina a duplicação e a transmissão da informação genética contida em cada célula para a geração celular seguinte, o que necessita de que o DNA seja replicado com fidelidade e as cópias dos cromossomas idênticos distribuídas pelas duas células filhas.

O ciclo celular é um processo complexo implicado no crescimento e proliferação das células, no desenvolvimento dos organismos, na regulação da reparação das alterações do DNA, na hiperplasia dos tecidos em consequência de certas injúrias, e ainda em diversas doenças como por exemplo nos cancros.

O ciclo celular eucariota compreende quatro etapas a G1, S, G2 e M, sendo a primeira uma fase hiato “gap” durante a qual as células preparam a replicação do DNA e integram sinais mitogénicos e de inibição do crescimento, decidindo quando o ciclo celular é para prosseguir, pausar ou é interrompido.

Na etapa S ocorre a síntese do DNA, enquanto a etapa G2 é uma segunda fase hiato “gap” durante a qual a célula prepara o processo de divisão, divisão esta ou mitose ou etapa M, em que se replicam os cromossomas e se acantonam em núcleos separados, formando-se duas células filhas.

A etapa Go corresponde a células que “saem” do ciclo celular e se tornam quiescentes.

Num dado ciclo celular cada etapa do ciclo, entra ou sai consoante a síntese ou degradação de determinadas ciclina específicas. Regra geral antes de uma célula passar à etapa seguinte do seu ciclo celular, é degradada a ciclina apropriada que actua na etapa anterior e sintetizada a ciclina da etapa seguinte do ciclo celular.

Os factores de crescimento actuam primeiramente sobre as células nas etapas Go e G1. A remoção daqueles factores de crescimento no início da etapa G1 originará um retorno para a etapa Go. Mas se essa remoção for feita na parte final da etapa G1 depois das células ultrapassarem o ponto de restrição que existe nesta etapa, as células continuarão a evoluir para a fase ou etapa S. Este ponto de restrição é muito regulado pelo Rb (retinoblastoma).

A paragem do ciclo celular é desencadeada por diversos factores intrínsecos e extrínsecos (como por exemplo a nutrição das células) que podem afectar diferentes “checkpoints”.

Carências nutritivas das células podem induzir a saída do ciclo celular e entrada na etapa Go.

A paragem do ciclo na etapa G1 permite a reparação do DNA antes da replicação desse mesmo DNA, enquanto a paragem na etapa G2 permite a reparação antes de ocorrer a separação dos cromossomas, na mitose.

A paragem na etapa G1 depende de p53, proteína esta que tem diversas funções tais como reconhecer e ligar-se a diversos tipos de alterações do DNA, inclusivé de DNA de uma só hélice, inserção/delecção de emparelhamentos incorrectos ( “ mismatches”),e a terminações livres de DNA.A p53 tem interacção com sequências específicas do DNA, e possui actividade de transcrição. Após alteração do DNA, a p53 estimula a transcrição da p21 e esta inibe as CdKs da etapa G1.

O papel da p53 na paragem da etapa G2 é mal conhecido.



O ciclo celular também pode ser alterado pela acção de diversos agentes víricos.

O ciclo celular é por vezes desregulado nas neoplasias, devido a alterações quer dos oncogenes que indirectamente afectam o ciclo celular, quer de genes supressores de tumores ou sejam oncogenes que regulam directamente o ciclo celular tais como os pRb, p16, ciclina D1 e mdm-2.

Na etapa G1 a proteína retinoblastoma (pRb) é um alvo importante das CdKs.

As CdKs (cinases dependentes de ciclinas) são reguladas por numerosas proteínas tais como as p23, p21 e cdc 25.

Nas etapas G1 e S os mais importantes inibidores são as famílias p16 e p21, incluindo-se nas p16 as p16, p18 e p19 que inactivam as CdK 4 e CdK6 evitando a fosfatação e inactivação da pRb. A família p21 inclui a p21, p27 e p57 que se ligam a ciclinas evitando a fosfatação da pRb pela CdK induzindo pois a dissociação da E2F.

O ciclo celular eucariota é regulada pela síntese e destruição periódicas das ciclinas que se associam e activam cinases dependentes das ciclinas ( COLK ).

As ciclinas são quaisquer das diversas proteínas afins cuja concentração sobe e desce durante o decurso do ciclo celular das células eucariotas. As ciclinas são subunidades reguladoras que formam complexos com as proteínas cinases dependentes das ciclinas, desta forma activando e determinando a especificidade do substrato através do ciclo celular.

As proteínas cinase ciclina dependentes (CdK) são proteínas cinases que são cataliticamente activas apenas quando estão ligadas à ciclina. Vários complexos CdK-ciclina fazem avançar as várias etapas do ciclo das células eucariotas ao fosfatar proteínas alvo específicas.

Nas células mamíferas estão assinaladas nove CdKs de 1 a 9 ( cinases dependentes de ciclinas ) estando também identificadas 16 ciclinas nos mamíferos ( vide quadro seguinte ) tendo todas uma região homóloga comum chamada caixa ciclina que é um domínio onde se ligam e activam as CdKs.

<b>Ciclinas nos mamíferos</b>		
<b>Tipos de ciclinas</b>	<b>CdK associada</b>	<b>função</b>
A	Cdk1( cdc2 ), cdk2	Entrada na fase S e transição, crescimento dependente da ancoragem
B1, B2	Cdk1	Saída da G2, mitose
C	Cdk8	Regulação da transcrição transição Go para a fase S
D1, D2, D3	Cdk4, cdk6	Transição da fase Go para S
E	Cdk2	Transição G1 para S
F	?	Transição G2 para M
G1, G2	Cdk5	Resposta a alteração do DNA
H	Cdk7	Activação CdK, regulação da transcrição, reparação do DNA
K	?	
I	?	Regulação da transcrição, CdK activação
T1, T2	Cdk9	Regulação da transcrição

Contudo nem todas as ciclinas e CdKs funcionam na regulação do ciclo celular, podendo intervir ainda na regulação da transcrição, reparação do DNA, diferenciação e apoptose.

A actividade das CdK durante o ciclo celular é regulada a outros níveis além da ligação às ciclinas. Assim a fosfatação das subunidades CdK pode regular positivamente ou negativamente a actividade cinase.

As CdKs são reguladas (inibidas ou estimuladas) de diversas formas, como por exemplo por fosfatação de resíduos de serina e treonina.

As ciclinas contêm sequencias PEST, motivos ricos em prolina(P), glutamato (E ), serina (S) e treonina (T) que são alvos para a sua degradação por ubiquinação em momentos específicos.

Também a associação com inibidores das CdK ( CKI ) é um importante meio de regulação das CdK.

As ciclinas do tipo D são as primeiras ciclinas a ser induzidas quando as células na fase Go são estimuladas para entrar no ciclo celular . As suas concentrações ( das ciclinas do tipo D ) em geral não oscilam durante o ciclo celular ( ao contrário de muitas outras ciclinas ) mas são controladas por factores de crescimento.

As ciclinas de tipo D associam-se e activam cdk4 e cdk6 ( vide quadro anterior e figura adiante ).

O substrato primário para estas cinases dependentes de ciclinas do tipo D é a proteína supressora do tumor retinoblastoma ( Rb ) que é uma proteína crítica na regulação da progressão G1, ligando-se e regulando um grande número de proteínas celulares, inclusivé factores de transcrição da família E2F que regulam a expressão de genes de proteínas implicadas na progressão do ciclo celular e da síntese de DNA,

inclusive ciclinas E e A, cdc2 ( cdk1 ), B-myb, dihidrofolato reductases, timidina cinase e DNA polimerase \_.

A ciclina A, que também é regulada em parte pelo E2F acumula-se na transição fase G1/S e persiste ao longo da fase S.

Nos inibidores das cinases dependentes de ciclinas, os p21 e p16 desempenham papéis importantes no controlo do ciclo celular coordenando sinais internos e externos e impedindo a proliferação em diversos pontos chave de controlo ( checkpoints ).

A síntese do DNA nas células eucariotas copia milhões de pares de bases ao longo de numerosos cromossomas, no momento exacto do ciclo de divisão celular.

Esta cópia tem início em centenas de origens ( ORC ) diferentes, algumas postas a funcionar no início da fase S do ciclo celular, enquanto outras são postas a funcionar mais tarde.

A iniciação da replicação do DNA, isto é o complexo ORC ( origin of replication complex ) não actua sozinho no controlo desta iniciação.

Uma ou mais proteínas adicionais ligam-se a estas origens da iniciação “mais tarde na mitose” e permanecem aqui ligados até começar a fase S (vide figura adiante). Estas proteínas são as cinases dependentes de ciclinas (CdKs) e operam em associação com substratos proteicos específicos (ciclinas).

A actividade das CdK é governada por um complexo de sub-unidades reguladoras e por fosfatações cujos efeitos sobre a sua conformação são hoje conhecidos.

Ambas as CdK e ciclinas podem controlar o ciclo celular, empurrando as células para a fase S e para a síntese de iniciação do DNA.

A degradação das CdK remove o sinal que inibe a divisão celular e o ciclo celular desloca-se de novo para a mitose.

Parece pois que a adição e remoção de ORC ciclinas e CdKs ocorrem de uma forma cíclica.

A mesma enzima activa primeiro a replicação do DNA e depois após uma ronda de replicação do DNA ter começado, inibe a nova formação do complexo de pré-replicação, constituindo um arranjo eficiente para a coordenação do início da síntese do DNA.

Em 1980 ficou claro que os processos moleculares que regulam a replicação dos cromossomas e a divisão celular são fundamentalmente similares em todas as células eucariotas.

A replicação celular é controlada pela regulação de dois acontecimentos críticos no ciclo celular: a iniciação da síntese nuclear do DNA e a iniciação da mitose.

Um pequeno número de proteínas cinases são os controladores chave destes acontecimentos. Elas regulam a actividade de várias proteínas implicadas neste processo, fosfatando-as em locais específicos, activando umas, inibindo outras e coordenando assim as suas actividades. Estas cinases são proteínas heterodiméricas sendo as suas sub-unidades reguladoras chamadas ciclinas porque as suas concentrações ciclam ( ou oscilam ) em fase ou consonância com o ciclo celular.

As sub-unidades catalíticas são chamadas cinases dependentes de ciclinas ( CdKs ) porque não têm actividade cinásica a não ser que associados com uma ciclina.

Cada sub-unidade CdK pode associar-se com diferentes ciclinas e é esta que determina quais as proteínas que vão ser fosfatadas pelo complexo CdK - ciclina ( vide quadro anterior ).

A maioria dos mecanismos moleculares regulam o início da fase S ou iniciação da mitose.

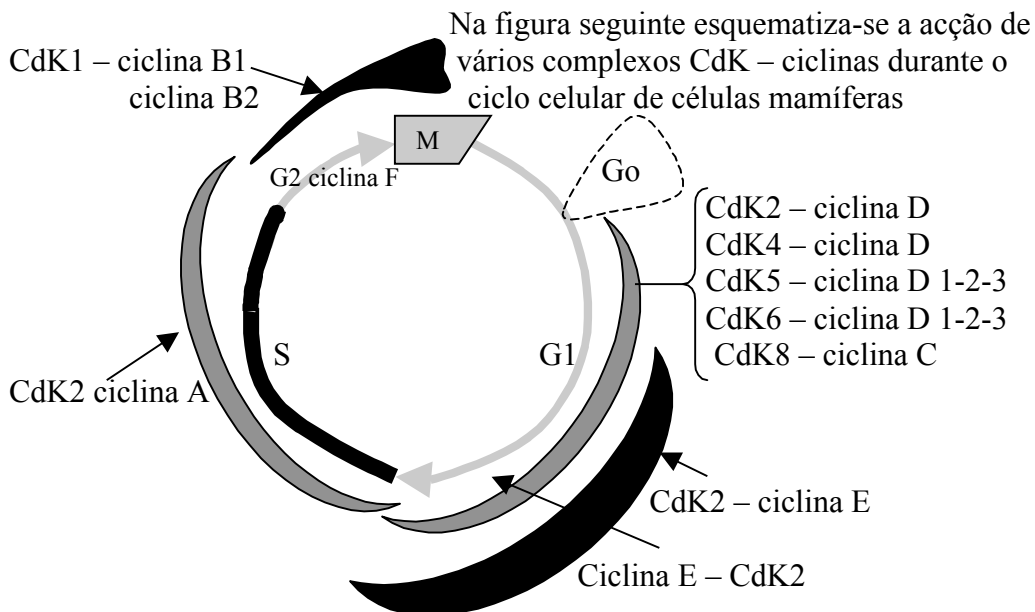
A actividade das subunidades catalíticas CdK é regulada por outras proteínas cinases e por fosfatases que regulam também e coordenam a divisão celular, e ainda por inibidores. Tudo isto nas células mamíferas é controlado por regulação dos genes e da sua transcrição.

As células mamíferas contêm algumas ciclinas diferentes que funcionam no período G 1, promovendo a entrada na fase S.

A maioria delas não se exprime em células não proliferantes que saíram do ciclo celular na fase G1, entrando numa fase de paragem chamada Go. Quando são adicionados a estas populações na fase Go factores de crescimento são induzidos uma série de genes chamados genes de resposta-precoce , e as proteínas por eles induzidas são factores de transcrição que induzem a expressão de um segundo conjunto de genes, os genes de resposta retardada, onde se incluem as proteínas por eles produzidas as G1 ciclinas e as CdKs.

As células mamíferas possuem 3 CdKs principais, cujas concentrações sobem e descem durante o decurso de cada ciclo celular.

Os complexos G1, CdK-ciclinas activam factores de transcrição sempre presentes na célula que estimulam a transcrição de genes codificadores de enzimas da replicação de DNA e síntese dos desoxiribonucleótidos.



As CdKs associadas com várias ciclinas medeiam a passagem através de várias etapas do ciclo celular em particular a entrada na fase S e a entrada na mitose ( fase M ).

Como se disse a especificidade para o substrato das CdKs é determinada pela ciclina associada, e a actividade dos complexos CdK-ciclina é regulada por fosfatação em locais específicos de activação e em sítios de inibição por ligação com proteínas inibidoras.

Para facilidade de interpretação inclusivé da figura anterior repetem-se no quadro abaixo as principais CdK nas células mamíferas, as ciclinas que lhes estão associadas e a fase em que intervêm no ciclo celular.

#### CdKs nas células mamíferas

<u>CdK</u>	<u>Ciclinas associadas</u>	<u>Fase de acção no ciclo celular</u>
CdK2	Ciclina D1, D2, D3	G1
	Ciclina E	G1 S transição
	Ciclina A	S
CdK4	Ciclina D1, D2, D3	G1
CdK6	Ciclina D1, D2, D3	G1
( CdK1 )	Ciclinas B1 e B2	G2_M transição; M inicial

A CdK3 encontra-se em baixas concentrações e a sua função não é clara. Duas outras ciclinas ( C e F ) foram clonadas mas as suas funções não estão esclarecidas.

Outras proteínas CdK mamíferas e ciclinas estão para ser identificadas. Resumindo as flutuações na actividade da família de proteínas cinases CdK orientam o ciclo celular eucariota , cujas características dominantes seguintes:

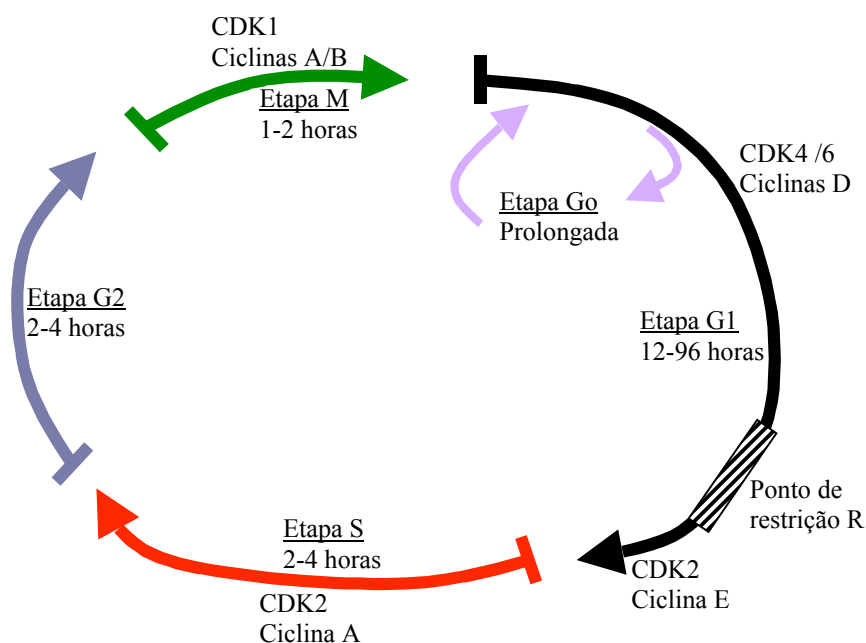
- A -A sua actividade é regulada por fosfatação, ou por ligação de inibidores estequiométricos das cinases ou ainda pela proteólise regulada das ciclinas.
- B-A actividade da CdK é regulada por acontecimentos dentro ou fora das células e controla por outro lado a replicação do DNA, a segregação dos cromossomas e a divisão celular.
- C-A família CDK/ciclinas contem cinco CdKs ( CdK1, 2, 3, 4 e 6 ) e quatro classes de ciclinas ( A, B, D e E ) que regulam o ciclo celular; três CdKs ( CdK7, 8 e 9 ) e três classes de ciclinas ( C, H e T ) que regulam a transcrição e outras CdK e classes de ciclinas têm funções pouco conhecidas.
- D-As sequências das ciclinas são menos bem conservadas que as das CdKs.
- E-A sequenciação do genoma humano revelou pouco sobre o ciclo celular.

## **6.2 - Regulação e desregulação (oncogénese) do ciclo celular**

A regulação do ciclo celular deve assegurar que os acontecimentos ocorridos em cada etapa do ciclo celular, estejam completados antes de se passar à etapa seguinte do ciclo, o que é realizado pela monitoragem da integridade do DNA em pontos chave de controlo (checkpoint) situados na parte final (no último terço) da etapa G1 (ponto de restrição R) e na interface entre a etapa G2 e a etapa M. Desta forma procura-se evitar a progressão e a propagação de células que sofreram mutações ou alterações.

## Ciclo celular

No esquema seguinte referem-se as diversas etapas que o integram, com os tempos muito arbitrários, que podem corresponder a cada uma destas etapas, bem como alguns factores e pontos intervenientes.



Na etapa G1 aumentam a expressão das ciclinas D (D1, D2, D3) que se associam com Cdk4 e Cdk6 e que levam à fosfatação e activação das Cdk.

Estas Cdk activadas fosfatam a proteína retinoblastoma (Rb).

A Rb hiperfosfatada permite a libertação das proteínas de transcrição E2F e a ultrapassagem do ponto de restrição R antes referido.

A Rb é desfosfatada depois da mitose estar completa (vide figura anexa).

Existe pois nesta etapa G1 um ponto de restrição R muito importante.

Na ultima parte da etapa G1 aumenta a expressão de ciclina E. O complexo ciclina E/ Cdk2 é necessário para a transição G1—S.

Na transição G1/S aumenta a expressão da ciclina A que persiste através da fase S.

A interacção ciclina A e Cdk2 acompanha a síntese de DNA.

Na última parte da etapa S a ciclina A associa-se com Cdh1.

Na etapa G2 existe um ponto chave de controlo que responde a DNA alterado ou incompletamente sintetizado, retardando nestas circunstâncias a entrada em mitose para o DNA ser reparado ou o ciclo ser abortado.

A mitose M é impulsionada por níveis aumentados de ciclinas A e B complexadas com Cdk1.

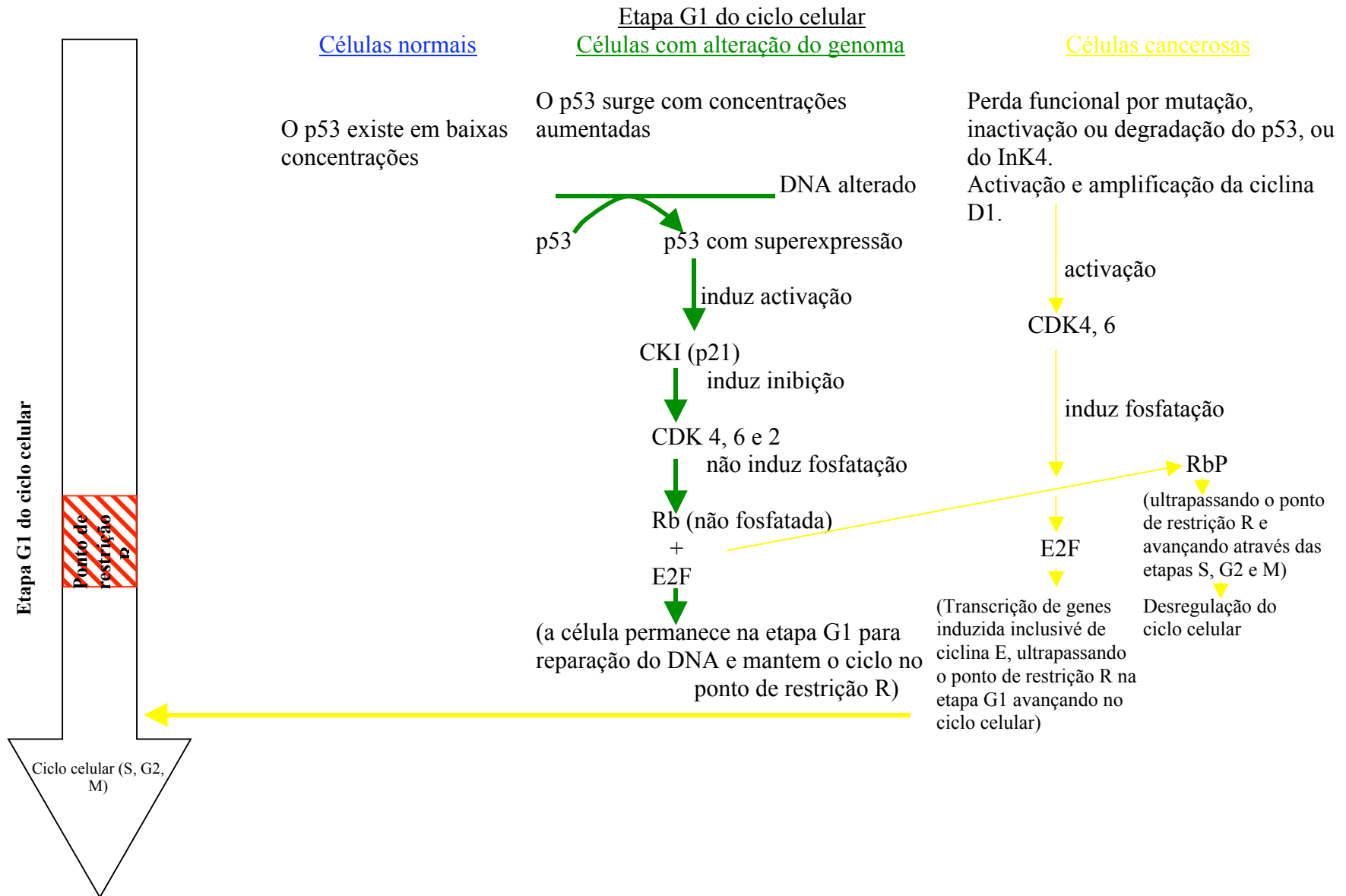
As células respondem a estímulos externos (mitogénicos ou factores de crescimento) desencadeando cascatas de fosfatações intracelulares, através da regulação da expressão de ciclinas que se associam com cinases dependentes de ciclinas (CdKs) e que fazem as células entrar em ciclo celular.

Na ausência de sinais mitóticos as células podem sofrer diferenciação, apoptose ou entrar numa etapa ou estado de quiescência (G<sub>0</sub>).

O ciclo celular é pois impulsionado por uma série de proteínas cinases (complexos CdKs com as respectivas ciclinas).

As ciclinas sofrem portanto oscilações ora sendo biossintetizadas e depois sofrendo degradação proteolítica, assegurando a transição entre as diversas etapas do ciclo celular.

A integridade do genoma da célula é controlada pelo factor de transcrição p53. Havendo alterações ao nível do DNA o p53 interrompe o ciclo celular para permitir a reparação do DNA o que é acompanhado pela inibição pela p53 da fosfatação da Rb.





Em células em replicação normal as concentrações de p53 são baixas e o ciclo celular progride normalmente. O p53 por seu turno também é regulado negativamente pela proteína MDM2.

Na presença de DNA alterado o p53 liga-se a uma sequência específica do DNA induzindo um aumento substancial da síntese dessa proteína p53 que é depois activada por fosfatação subsequente, interrompendo o ciclo celular para ocorrer a reparação do DNA.

O p53 controla o ciclo celular através da regulação do inibidor p21 da CdK (CKI) que é um inibidor activo das CdK4,6 e 2. Esta inibição das cinases impede a fosfatação da Rb e daí a célula permanecer na etapa G1 para que ocorra a reparação do DNA.

Há duas famílias de inibidores CdK implicados na regulação do ciclo celular, a família Cip/Kip (alvos os CdK4,6 e 2) que inclui os inibidores p21 e p27 e a família INK4 (inibidores da cinase 4 dependente de ciclina) que codifica a p16INK4a e a p19 ARF.

Se a alteração do DNA ultrapassa a capacidade da sua reparação o p53 leva a célula a entrar em apoptose induzindo a expressão da proteína pró-apoptótica Bax.

As células cancerosas permanecem em ciclo celular independentemente de controlos internos ou externos.

A proliferação descontrolada das células tumorais e cancerosas está na dependência da alteração dos genes que regulam directamente os seus ciclos celulares.

Em condições normais as células na etapa G1 são sensíveis a sinais de crescimento externos e sujeitas a regulação interna e após ultrapassarem nesta etapa G1 o ponto de restrição R avançam sem problemas através das etapas S e M.

Com as células anormais não se passam as coisas assim pois estas células têm capacidade para se escaparem ao ponto de restrição R na etapa G1, não se submetendo aos controlos de regulação interna habituais e proliferando na ausência de estímulos vindos do exterior das células.

Os processos oncogénicos exercem os seus maiores efeitos através do interessamento de alvos que são regulados para a progressão da etapa G1 do ciclo celular.

Alterações estruturais sobretudo e mutacionais que levem ao silenciamento das proteínas p53, Rb ou INK4 desregulam o ciclo celular, tal como a superexpressão de ciclinas, sobretudo das ciclinas D que promovem a fosfatação da proteína Rb e a consequente libertação do factor de transcrição E2F, tudo isto aumentando o risco da formação de tumores.

A p53 desregula a entrada na etapa S, a Rb afecta o controlo na transição G1/S, a p16 INK49 bloqueia a fosfatação da Rb e a ciclina D1 promove a fosfatação da Rb.

O grupo p53 nos cancros humanos é o gene mais frequentemente mutado. A proteína p53 tem habitualmente uma semi-vida muito curta, mas é estabilizada e acumula-se nas células com DNA alterado ou nas células que respondem a certas formas de stress.

A inactivação da Rb pode ocorrer por fosfatação ou por alteração genética directa do gene Rb.

A ciclina D1 pode ser superexpressa em muitos cancros humanos, devido a uma amplificação do respectivo gene ou a translocações do locus D, podendo assim contribuir para a oncogénese.

Estão identificados mais de cem proto-oncogenes e genes supressores de tumores, e a maioria deles funciona na transdução de sinais, (imitando os efeitos de estímulos mitogénicos persistentes) o que leva a desarticular as células dos seus controlos respectivos.

As vias de transdução destes sinais nas células cancerosas convergem no controlo da passagem através da etapa G1, induzindo as ciclinas desta etapa G1, ultrapassando os inibidores das CdK, evitando a saída do ciclo celular e perturbando os pontos de controlo do respectivo ciclo celular.

Nas células cancerosas as vias de transdução de sinais mais frequentemente alteradas são aquelas em que intervêm os genes supressores de tumores, RB e p53.

## **7 – Oncogénese**

### **7.1 – Envolvimento de factores de transcrição na carcinogénese**

De uma maneira geral as situações de cancro nos animais resultam da transformação de células que adquirem capacidade para crescer desordenadamente e para metastizar, o que está na dependência de certos genes designados oncogenes. Com efeito certos genes sofrendo mutações ou exprimindo exageradamente certas proteínas podem contribuir para a oncogénese e daí a sua designação de oncogenes.

Há outro grupo de genes designados proto-oncogenes que são genes celulares normais análogos de oncogenes, e cujos produtos da sua expressão são componentes de diversas vias que regulam o crescimento e diferenciação das células normais estando implicados em cascatas de transdução de sinais promovidos por hormonas podendo influenciar a própria transcrição. Quando esses proto-oncogenes sofrem mutações são convertidos em oncogenes celulares.

#### O ciclo celular e a oncogénese

Nas neoplasias um tecido anormal cresce, por proliferação celular, mais rapidamente que o tecido normal e continua a crescer mesmo que o estímulo desencadeador dessa situação desapareça ou pare de actuar.

Os oncogenes e genes supressores de tumores estão envolvidos em certo aspectos com a transdução de sinais mitogénicos ou com o controlo do ciclo celular em “checkpoints”.

A transdução de sinais mitogénicos implica diversas oncoproteínas já hoje conhecidas, que vão desde os factores de crescimento segregados extracelularmente, até aos seus receptores, passando depois pelas cascatas citoplásmicas de transdução desses sinais até aos factores de transcrição.

Nas células normais essas oncoproteínas transduzem sinais da superfície celular para o núcleo, originando as transcrições de proteínas implicadas na iniciação do ciclo celular.

Nas células anormais, estas vias de transdução de sinais permanecem em actividade contínua ou então os mecanismos que as podem inibir são desactivados.

Muitas oncoproteínas ou sejam proteínas oncogénicas expressas por oncogenes ou genes supressores de tumores não são no entanto reguladores imediatos do ciclo celular. Nas oncoproteínas reguladoras imediatos do ciclo celular referem-se várias no quadro seguinte.

Mutações no p53 estão assinaladas em mais de 50% das neoplasias malignas.

A mdm-2 liga-se com a p53 estabilizando esta e inibindo a actividade transcripcional. A mdm-2 também se liga á pRb interferindo com a interacção com a E2F.

A p21 tem um papel oncogénico mal conhecido.

A proteína pRb tal como a p53 é supressora de tumores.

A família de proteínas p16 são inibidores do ciclo celular na etapa G1 inactivando os complexos ciclina-CdK, impedindo que as proteínas pRb continuem hiperfosfatadas.

As ciclinas D1 são oncogenes sendo reguladoras do ponto de restrição do ciclo celular, formando complexos com as CdK4 e CdK6 que fosfatam e inactivam as pRb.

As cdc25A e cdc25 B são oncogenicas pois desfosfatam as Cdks nos seus centros catalíticos. Pela sua superexpressão ou activação das cdc25 fosfatases, as CdKs são mantidas num estado activado.

#### Oncoproteínas reguladoras do ciclo celular nas neoplasias

<u>Oncoproteína/ supressora de tumores</u>	<u>Função no ciclo celular normal</u>	<u>Alterações oncogénicas</u>
p 53	Promove a paragem na etapa G1 após alteração do DNA	Mutação
p Rb	Reguladora do ponto de restrição	Delecção
mdm-2	Ultrapassa a actividade transcripcional da p53	Amplificação do gene ou favorecimento da translação do mRNA
P 16	Inactiva os complexos Ciclina-CdK na etapa G1	Delecção
Ciclina D1	Inicia o ciclo celular	Amplificação do gene
cdc 25A , cdc 25B	Activa as CdK nas G1/S	Super expressão

Nos animais em crescimento a multiplicação celular em geral é mais importante que a morte celular, atingindo-se mais tarde um estadio nos animais adultos em que as duas situações se equilibram o

que não significa que pontualmente e em certos tipos de células não continue a ocorrer um rápido turnover ao contrário de outros. Cita-se o exemplo das células do epitélio do intestino delgado, e de alguns leucócitos que têm uma renovação em poucos dias, enquanto os glóbulos vermelhos podem ter uma semi-vida de 100 dias (consoante as espécies animais) e as células hepáticas saudáveis raramente morrem, e as células do cérebro adulto também parecem ter pouca substituição celular.

Estes ritmos da multiplicação celular e o seu equilíbrio relativo com a morte celular tem características muito próprias consoante as espécies animais e sobretudo as espécies pecuárias envolvidas na produção animal.

No entanto ocorrem situações anormais, patológicas em que ocorre desregulação da multiplicação celular. Neste tipo de situações as células em causa deixam de obedecer à respectiva regulação, e podem entrar em crescimento e divisão desencadeando a formação de “clones” de células que se multiplicam indefinidamente gerando a formação de massas tumorais.

As causas destas desregulações podem ser factores genéticos e fisiológicos, embora os factores genéticos implicados não sejam geralmente transmitidos através dos gâmetas mas correspondendo antes a alterações no DNA das células somáticas. Estas alterações do DNA convertem os genes pré-existentes em oncogenes, e estes por sua vez exprimem produtos responsáveis pelo crescimento e multiplicação celular desordenadas.

Esta ocorrência pode ter sede nos vários tipos de células que integram um organismo animal e enquanto a célula normal de um dado tipo fica num estado quiescente, a célula cancerosa do mesmo tipo tende a invadir os tecidos vizinhos, por vezes através da lâmina basal que define o limite dos tecidos, invadindo o organismo animal, o que constitui a metastização.

As células normais e as células cancerosas diferem em vários aspectos tais como: controlo do crescimento celular, morfologia, interações célula a célula, propriedades das membranas, estrutura do citoesqueleto, segregação de proteínas, expressão dos genes e na mortalidade.

A conversão de uma célula normal numa célula cancerosa pode ser devida à expressão de um ou mais genes. Na maior parte dos casos os oncogenes actuam cooperativamente, não bastando um só oncogene para criar uma situação cancerosa, sendo alterações genéticas múltiplas necessárias para esse efeito.

A mutação num oncogene promotor de cancro pode ocorrer por alteração do segmento codificador de proteínas do gene ou então por alteração da regulação da formação e função das proteínas.

Esta conversão do proto-oncogene em oncogene pode resultar de alterações na estrutura do DNA, como por exemplo mutações pontuais, deleção, etc, translocações cromossomais ou por reduplicação extensa da região onde está contido o oncogene.

Os oncogenes podem pois ser desencadeados nas próprias células ou ser para estas transportados por exemplo através de vírus, de produtos químicos carcinogénicos ou através de radiações.

Os vírus implicados nestas situações podem ser vírus com DNA ou retrovírus contendo RNA.

Os produtos químicos carcinogénicos podem ser de variada natureza química mas têm a característica de eles próprios serem electrofílicos ou tornarem-se electrofílicos depois de metabolizados

pelos organismos. Esta activação metabólica é promovida pelo sistema citocromo P-450 gerando um carcinogénio que altera o DNA convertendo assim um proto-oncogene num oncogene.

As radiações atómicas, por Raios X e ultra-violetas, todas afectam o DNA ,alterando os proto-oncogenes.

Ainda na indução do cancro, certos produtos químicos podem potenciar a actividade electrofílica dos carcinogénios, e daí a sua denominação de promotores.

São os genes que controlam o crescimento e a divisão celular, bem como aqueles que codificam um dos cinco tipos ou classes de proteínas referidos seguidamente, a maioria dos oncogenes conhecidos:

- I- Factores de crescimento
- II- Receptores
- III- Transdutores intracelulares de sinais
- IV- Factores de transcrição nucleares
- V- Proteínas que controlam o ciclo celular.

É conhecido que a divisão celular é desencadeada por um sinal gerado pela acção de um ligando, um factor de crescimento que com o seu respectivo receptor desencadeia uma cascata de transdução de sinais que vai condicionar a expressão de uma série de genes dessas células. É sabido que todo este processo é coordenado por proteínas de controlo do ciclo celular como é o caso da Rb e da p53 que são expressas por genes supressores de tumores ou sejam anti-oncogenes.

A inactivação destes genes anti-oncogenes por exemplo por mutações, em ambos os cromossomas favorece como se compreende a acção dos oncogenes e a carcinogénese.

As oncoproteínas intracelulares ou sejam as proteínas expressas pelos oncogenes são produtos implicados na transdução de sinais. Estão neste caso desde tirosina cinases como por exemplo o src do vírus do Sarcoma de Rous e proteínas G de ligação ao GTP e de hidrólise (ras oncogenes) até factores de transcrição nucleares.

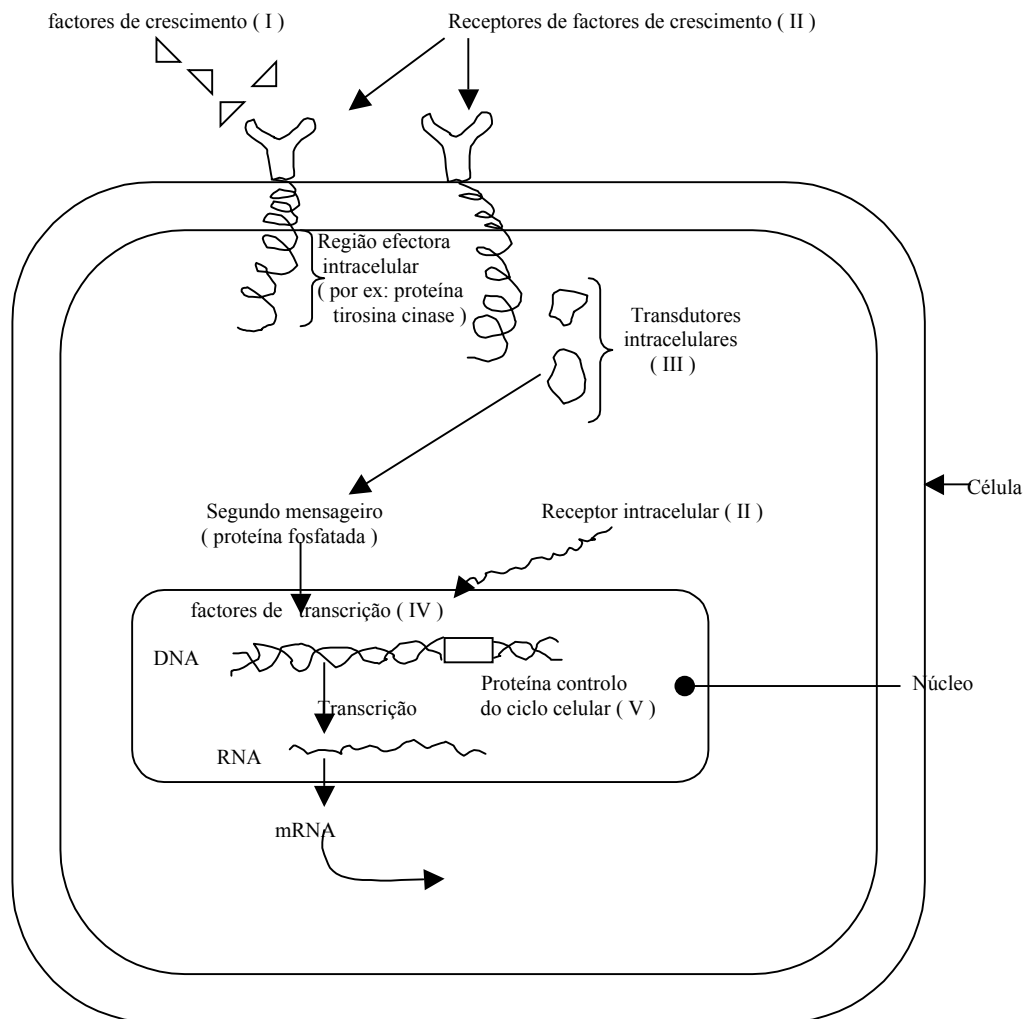
## **7.2 - Oncogenes e suas proteínas, sua classificação e suas características**

Vejamos como as oncoproteínas podem transformar as células normais em células malignas. As oncoproteínas causam cancro quer porque não existiam antes e passam a estar presentes ou então porque passam a ser expressas em concentrações muito superiores ao normal.

Um proto-oncogene e o seu oncogene são designados por três letras em itálico (por exemplo: *src*), sendo utilizado o mesmo código sem ser em itálico e com letra grande na primeira letra para identificar as suas proteínas produzidas (por exemplo: Src). O oncogene de um vírus é assinalado como prefixo V (por exemplo: V-*src*) e o correspondente proto-oncogene celular tem o prefixo C (por exemplo: C-*src*).

No controlo do crescimento das células participam cinco tipos de proteínas como já referimos: os factores de crescimento (I), os receptores para os factores de crescimento e hormonas (II), os transdutores intracelulares de sinais (III), os factores de transcrição nucleares (IV), e as proteínas de controlo do ciclo celular (V) (vide figura seguinte).

Controlo do crescimento celular e tipos de proteínas participantes cujas formas mutantes originam cancro



As mutações que alteram a estrutura ou expressão dos grupos I, II, III e IV dão lugar a oncogenes activos.

O grupo V actua sobretudo como supressor de tumores e as mutações nestes genes (ciclina, cinases dependentes de ciclina, p53 e RB) libertam as células dos respectivos controlos aumentando a sua tumorização.

As oncoproteínas são modificadoras das transcrições (activadores ou repressores) e como tal alteram a composição em proteínas da célula (proteoma), afectando o controlo do crescimento celular de diversas maneiras.

No quadro seguinte referem-se os oncogenes mais representativos bem como as proteínas expressas pelos seus proto-oncogenes nos cinco tipos (I, II, III, IV e V) de produtos dos genes que participam no controlo do crescimento.

Os oncogenes raramente provêm de genes codificadores de fatores de crescimento (classe I).

Os genes responsáveis pela expressão de receptores para hormonas e factores de crescimento (Classe II) podem tornar-se oncogenes quando devido a uma mutação, o receptor assume tal conformação (mesmo na ausência de ligando interagindo com o receptor) que continua sempre activo. Na figura seguinte mostram-se dois exemplos de proteínas receptoras de proto-oncogenes a serem transformadas num processo oncogénico em proteínas receptoras oncogénicas.

No exemplo representado á esquerda na figura, uma alteração no codão que especifica um dado acido aminado (valina) no domínio transmembranário da proteína receptora codificada pelo gene neu produz o oncogene neu também chamado her-2 em virtude da substituição da valina por glutamina.

No outro exemplo representado à direita na figura parte do domínio extracelular expresso pelo proto-oncogene que se une ao ligando é suprimida em virtude da deleção ocorrida no gene correspondente o que leva á formação do oncogene ( como no Erb B).

Já tivemos oportunidade de referir que alguns oncogenes sozinhos não transformam as células em malignas necessitando da cooperação de outros genes para se instalar esse efeito.

A classe III que envolve os transdutores intracelulares de sinais é a maior classe de oncogenes conhecida.

O mecanismo utilizado por alguns oncogenes para promover a transformação carcinogénica ,recorre a fosfatações de uma ou de várias proteínas alvo como pode ser o caso de fosfatações desencadeadas sobre resíduos de tirosina dessas proteínas cinases.

Os oncogenes da classe IV (factores de transcrição nucleares) exercem efeitos directos sobre a transcrição ligando-se ao DNA e modificando as velocidades de iniciação da transcrição.

Nas proteínas da classe V (proteínas que controlam o ciclo celular) as ciclinas, cinases dependentes de ciclinas associadas com o síndrome de Li-Fraumeni, e a p53 e a RB (gene do retinoblastoma) WT1, WT2 (genes do tumor de Wilm's) ou APC (gene da colipolipose familiar) são reguladoras deste controlo, e a regulação alterada da expressão de pelo menos uma ciclina assim como mutação do p53 e RB (genes supressores de tumores ou anti-oncogenes) podem ser oncogénicos.

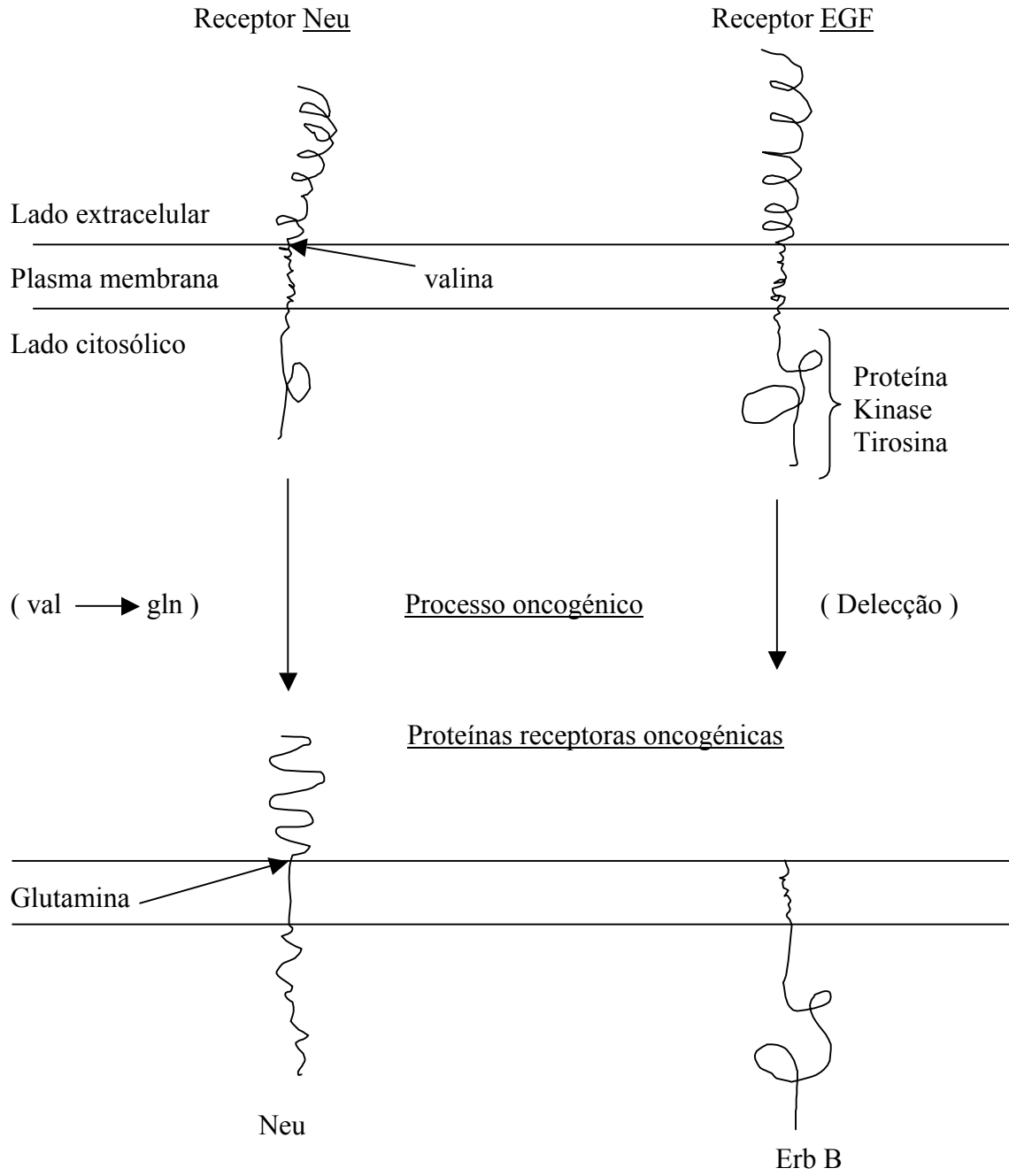
A apoptose ou suicídio induzido das células é um mecanismo de protecção contra o cancro, através de macrófagos que englobam e degradam essas células.

## Oncogenes e proteínas expressas pelos seus proto-oncogenes

<u>Oncogene</u>	<u>Localização do oncogene</u>		<u>Localização sub-célular da proteína</u>	<u>Proteína codificada pelo proto-oncogene</u>	
	<u>Retrovírus amimais</u>	<u>Tumor não viral</u>			
<b>Factor de crescimento ( I )</b>					
sis	Sarcoma dos símios		Segregado	Factor de crescimento derivado das plaquetas (PDGF)	
<b>Receptores ( II )</b>					
<u>a) Receptores da superfície celular com actividade de proteína tirosina cinase</u>					
Fms	Sarcoma felino de Mc Donough		Plasma membrana	CSF-1 receptor (factor estimulante de colónias)	
erb B	Eritoblastose aviária		Plasma membrana	Receptor do EGF (factor do crescimento epidérmico)	
neu ( ou erb B-2 )		Neuroblastoma	Plasma membrana	Relacionado com o receptor da EGF	
Ros	UR II sarcoma aviaria		Plasma membrana	Relacionado com o receptor da insulina	
<u>b) Receptores intracelulares</u>					
erb A	Eritroblastose aviaria		Nuclear	Receptor da hormona da tiróide	
<b>Transdutores intracelulares ( III )</b>					
<u>a) Proteína Tirosina-Cinase</u>					
src	Sarcoma de Rous aviário		Citoplasma	} Proteínas cinases que fosfatam resíduos de tirosina	
yes	Sarcoma aviario Yamaguchi		Citoplasma		
fips ( fes )	Sarcoma aviaria Fujinami e Sarcoma felinos		Citoplasma		
abl	Leucina murina	Leucemia mielogénica-cronica	Citoplasma	} Proteínas Cinases específicas para a serina e treonina	
met		Osteosarcoma murino	Núcleo		
<u>b) Proteína cinase serina/treonina</u>					
mos	sarcoma murino Moloney		Citoplasma	} Proteínas Cinases específicas para a serina e treonina	
raf (mil)	sarcoma murino 3611		Citoplasma		
<u>c) Ras proteínas</u>					
Hã-ras	Sarcoma murino Harvey	Bexiga, mama, rim (carcinomas)	Plasma membrana	Proteínas que ligam nucleótidos de guanina com actividade GTPase	
Ki-ras	Sarcoma murino de kirsten	Carcinomas do pulmão e colon	Plasma membrana	. Proteínas que ligam nucleótidos de guanina com actividade GTPase	
N-ras		Neuroblastomas e leucomas	Plasma membrana		
<u>d) Adaptadores</u>					
crk	Vírus sarcoma aviario		citoplasma	Contem proteínas com domínios SH <sub>2</sub> e SH <sub>3</sub> mas sem actividade catalítica	
<b>Factor de transcrição nucleares (IV)</b>					
jun	Vírus 17 sarcoma aviário		Núcleo	. Factor de transcrição AP1	
fos	Osteosarcoma FBJ		Núcleo		
myc	Myclocitomatose aviária mc29		Núcleo		
N-myc		Neuroblastoma leucemia	Núcleo	. Proteínas que regulam a transcrição	
myb	Mieloblastose aviária		Núcleo		
ski	Aviária sk/v770		Núcleo		
rel	Reticuloendoteliose aviária		Núcl. e citoplasma	. Supressores de tumores	
<b>Proteínas que controlam o ciclo celular (V)</b>					
RB		Retinoblastoma e osteosarcoma	Núcleo		
p53		A maioria dos tipos de cancro humano	Núcleo		



Proteínas receptoras de proto-oncogenes



### **7.3 - Transdução de sinais mediados por oncogenes em células epiteliais mamárias de modelos transgênicos de rato**

A resposta das células epiteliais mamárias aos factores de crescimento depende do número de receptores específicos nessas células para esses factores e da sua acoplação com uma série de vias de transdução de sinal intracelulares.

O desenvolvimento, a maturação e a diferenciação das células epiteliais mamárias dependem do conjunto de interacções das hormonas e dos factores de crescimento.

O crescimento da glândula mamária implica a proliferação celular, a diferenciação e a morte programada das células (apoptose).

Após a gravidez, ocorre um rápido desenvolvimento lóbulo alveolar do epitélio mamário levando à indução da diferenciação terminal e à lactação quando do crescimento.

Após a desmama ocorre a involução rápida do epitélio mamário através da apoptose.

A regulação de todo este ciclo da proliferação, diferenciação e apoptose depende do equilíbrio entre os factores de crescimento solúveis, as hormonas e as interacções célula-substrato.

Esta interacção dos factores de crescimento é feita através dos seus receptores e respectivas vias de transdução de sinais.

Tomando o exemplo da família dos receptores para o EGF (factor de crescimento epidérmico) estes receptores interactuam com uma série de moléculas citoplásmicas para transdução de sinal, e a estimulação da actividade tirosina cinase intrínseca destes receptores (RTK) origina a sua autofosfatação e a partir daqui proteínas com domínios específicos de homologia 2 src (SH<sub>2</sub>) ou de ligação/interacção com as fosfotirosinas (PTB) podem ser induzidas para activação.

A proteína fosfatidilinositol cinase 3i (PI-3i) contendo domínios SH<sub>2</sub> é uma molécula de transdução de sinal a partir das RTK, e quando associada com factores de crescimento activados origina lípidos fosfoinositideo 3i.

A estimulação da PI-3i cinase pode originar por outro lado a heterodimerização de erbB-2 e erbB-3 e ainda activar proteínas serina cinases.

Algumas destas proteínas serina cinases activadas tem muito impacto sobre a sobrevivência celular, evitando por exemplo a apoptose.

A activação da PI-3i cinase tem pois um papel importante na morfogénese e na sobrevivência das células mamárias.

Outra via de transdução de sinal importante no controlo da proliferação de epitélio mamário é a que envolve a Ras. Proteínas contendo domínios SH<sub>2</sub> e PTB como a Grb<sub>2</sub> e Shc formam complexos com os receptores de factores de crescimento e activam vias de transdução de sinais em que intervêm moléculas Ras.

A interacção da proteína adaptadora Grb2 através dos seus domínios SH<sub>2</sub> com receptores activados de factores de crescimento estimula a permuta do nucleótido de guanina SOS convertendo-a da sua forma inactiva que contem GDP na sua forma activa contendo GTP.

Este complexo Grb2/SOS e os receptores dos factores de crescimento activados podem ligar-se por associação com as proteínas adaptadoras Shc através dos locais de fosfatação das tirosinas específicas dentro desta Shc.

Há também sugestões de uma outra via independente Ras em que interviria o factor de transcrição C-Myc.

A activação da via de transdução de sinal envolvendo Ras pode originar a estimulação de uma cascata de serina cinases tais como Raf, MEK e MAP cinase.

Estas cascatas de cinases activadas têm como alvo factores de transcrição nuclear tais como proteínas Fos, Jun e Ets.

Parece que os receptores tirosina cinase (RTK) tais como o erb B-2 podem activar vias Ras através de diversas proteínas adaptadoras.

Qualquer um dos quatro locais de autofosfatação de tirosinas na erb b-2 (tirosina 1144, 1201, 1227 e 1253) podem independentemente mediar transformações oncogénicas através da activação de uma via Ras, estando assinaladas as tirosina 1144 como parceiro para ligação com a proteína adaptadora Gbr2 e a tirosina 1227 como parceiro para ligação com a proteína adaptadora Shc. A autofatação de tirosina 1028 regula negativamente a transformação mediada pela erb B-2.

A activação dos receptores dos factores de crescimento é acompanhada pela estimulação de Src das famílias das tirosinas cinases.

Receptores tirosina cinase como o Met e o erb B-2 associam-se com cinases da família Src através de resíduos específicos de fosfotirosina dentro do receptor aumentando a sua actividade o que pode influenciar a progressão do ciclo celular através da estimulação da ciclina D1 mediando assim a proliferação celular e migração celular.

Estas vias estão bem assinaladas como implicadas na transdução de sinais para proliferação e diferenciação celular mas quanto ao seu papel na formação de tumores mamários tudo é menos claro.

No cancro do seio humano está bem assinalada a ampliação e elevada expressão do oncogene erb B-2/neu.

## **Bibliografia**

### **1 - Ritmos circadianos**

- Clayton, J. D., Kyriacou, C. P. & Reppert, S. M. (2001) Keeping time with human genome. Nature 409: 829-831.
- Reppert, S. M. & Weaver, D. R. Comparing clockworks: mouse versus fly. J. Biol. Rhythms 15, 357-364
- Young, M. W. 2000 Circadian rhythms. Marking time for a kingdom. Science 288, 451-453
- Shearman, L. P. et al 2000 interacting molecular loops in the mammalian circadian clock, Science 288, 1013-1019
- Dunlap, J. C. 1999 Molecular bases for circadian clocks. Cell 96, 271-290
- Hogenesch, J. B. et al. 2000 The basic helix-loop-helix-PAS protein MOP9 is a brain-specific heterodimeric partner of circadian and hypoxia factors, J. Neurosci, 20, RC83
- Gotter, A. L. et al 2000 A time-less function for mouse Timeless. Nature Neurosci, 3, 755-756
- Katzemberg, D. et al 1998 A CLOCK polymorphism associated with human diurnal preference. Sleep 21, 569-576
- King, D. D. & Takahashi, J. S. (2000) – Molecular genetics of circadian rhythms in mammals – Annu. Rev Neurosci. 23 : 713-742.
- Reppert, S. M. & Weaver, D. R. (2001) – Molecular analysis of mammalian circadian rhythms – Annu. Rev Physiol 63 : 713-742

### **2 - Transdução de sinais visuais**

- Devlin, T.M. .The eye:metabolism and vision-Textbook of biochemistry and clinical correlations, Wiley-Liss,1997
- Stryer, L. Biochemistry, W.H. Freeman and Company, New York, 1995
- Fung, B.-K., Hurley, J.B., & Stryer, L., (1981). Flow of information in the light-triggered cyclic nucleotide cascade of vision. Proc. Nat. Acad. Sci. 78:152-156.
- Noel, J.P., Hamm, H., & Sigler, P.B., (1993). The 2.2 Å crystal structure of transducin-complexed with GTP. S. Nature, 366:654-663.
- Nathans, J., (1989). The genes for color vision. Sci. Am. 260(2):42-49.
- Wald, G., (1968). The molecular basis of visual excitation. Nature, 219:800-807.
- Palczewski, K., (1994) Is vertebrate phototransduction solved? New insights into the molecular mechanism of phototransduction. Invest. Ophthalmol. Vis. Sci. 35:3577.
- Stryer, L., (1991). Visual excitation and recovery. J Biol. Chem., 266:10711
- Zigler, J. S. Jr., & Goosey, J. (1981). Aging of protein molecules: lens crystallins as a model system. Trends Biochem. Sci. 7:133.

### **3 - Famílias de receptores da olfacção e vias de sinalização**

- Gibson, A. D. & Garbers, D. L. (2000) – Guanylyl cyclases as a family of putative odorant receptor – Annu. Rev Neurosci 23 : 417-439
- Jones, D. T., Masters S. B., Bohrn H. R., Reed RR (1990) – Biochemical characterization of the stimulatory GTP – binding proteins. The large and small forms of Gs and the olfactory – specific G protein, golf, J. Biol. Chem 265 : 2671 – 76
- Malnic, B., Hirono J., Sato, T, Buck L. B. (1999). Combinatorial receptor codes for odors. Cell 96 : 713-23

- Zhao, H. Ivic, L. Otak, JM. et al (1998) Functional expression of a mammalian odorant receptor Science 279 : 237-42

#### 4 - **Stress e resposta inflamatória.Quimocinas.Citocinas**

- Locat, M. e Murphy P. M. (1999) – Chemokines and chemokine receptor biology and clinical relevance in inflammation and AIDS. Annu. Rev Med. 50 : 425-40
- Furie, M. B., Randolph, GY (1995) Chemokines and tissue injury. Am J. Pathol 146 : 1287-30
- Prescott, S. M., Zimmerman, G. A. Stafforini, D. e McIntyre, T. M. (2000) – Platelet-activating factor and related lipid mediators – Annu. Rev Biochem 69 : 419-45
- Garbers, D. L. e Dubois, S. K. (1999) – The molecular basis of hypertension Annu. Rev. Biochem 68 : 127-155

#### 5 - **Necrose e Apoptose**

- Nijhawan, D., Honarfour, N. E Wang, X., (2000) – Apoptosis in neural Development and disease Annu. Rev. Neurosci 23 : 73-87
- Ashkenazi A, Dixit V. M., (1998) Death receptors signaling and modulation Science 281 : 1305-8
- Kroemer G, Zanoni, N., Susina 3<sup>a</sup> (1997) Mitochondrial control of apoptosis Immunol Today 18 ; 44-51
- Nagata, S., (1997) Apoptosis by death factor Cell 88 : 355-65
- Wallach, D., Varfolomeev, E. E., Malinin, N. L., et al (1999) Tumor necrosis factor receptor and Fas signaling mechanisms Annu. Rev. Immunol 17 : 331-67
- Wallach, D. Boldin, M., Varfolomeev, E., et al (1997) – Cell death induction by receptors of the TNF family towards a molecular understanding FEBS lett 410 : 96-108
- Van ostade X, Tavernier J, Fiers W., (1994) Structure- activity studies of human tumor necrosis factor Protein Eng 7 : 5-22
- Williams Abbot L, Walter B. N., cheung T. C., Goh C. R., Porter A. G., Wan C. F., (1997) The lymphotoxin-alpha (L $\alpha$ ) subunit is essential for the assembly, but not for the receptor specificity of the membrane-anchored LT  $\alpha$ 1  $\beta$  2 heterotrimeric ligand J. Biol. Chem 272 : 19451-56
- Grell M, Douni E, Wajant H, Löhden M, et al (1995) The transmembrane form of tumor necrosis factor is the prime activating ligand of the 80 kDa tumor necrosis factor receptor Cell 83 : 793-802
- Kapas L, Hong L, Cady A. B., et al (1992) Somnogenic, pyrogenic and anorectic activities of tumor necrosis factor-alpha and TNF-alpha fragments Am J. Physiol 263: R708- 13
- Rathlen D. A., Ferrante A., Aston R. (1993) Differential effects of small tumor necrosis factor alpha peptides on tumor cell cytotoxicity, neutrophil activation and endothelial cell procoagulant activity Immunology 80: 293-99
- Boldin M. P., Met I. L., Varfolomeev E. E., et al (1995) Self-association of the “death domains” of the p55 tumor necrosis factor (TNF) receptor and Fas/Apo1 prompts signalins for TNF and Fas/Apo1 effects J. Biol. Chem. 270: 387-91
- Naismith J. H., Devine T. Q., Brandhuber B. J., Sprang S. R., (1995) Crystallographic evidence for dimerization of unliganded tumor necrosis factor receptor J. Biol. Chem. 270: 13305-7
- Nicholson D. W., Thornberry N. A., (1997) Caspases: killer proteases Trends Biochem Sci. 22: 299-306

- Wallach, D. (1997) Apoptosis. Placing death under control (news; comment) Nature 388: 123, 125-26
- Wallach D. (1997) Cell death induction by TNF a matter of self control Trends Biochem Sci 22: 107-46
- Bazan J. F., (1993) Emerging families of cytokines and receptors Curr. Biol 3:1603-6
- Kolesnick, R. N., e Krönke, M. (1998) Regulation of ceramide production and apoptosis Annu. Rev. Physiol 60: 643-65
- Adam-Klages S., Adam D., Wiegmann K., et al (1996) Fan, a novel WD repeat protein, couples the p53 TNF-receptor to neutral sphingomyelinase Cell 86: 937-47
- Kidd, V. J., (1998) Proteolytic activities that mediate apoptosis Annu. Rev. Physiol 60: 533-73
- Strasser, A., O'connor, L., e Dixit V. M., (2000) Apoptosis signaling Annu. Rev. Biochem 69: 217-45
- Kroemer, G., Dallaporta, B., Resche-Rigon, M., (1998) The mitochondrial death/life regulator in apoptosis and necrosis Annu. Rev. Biochem 60: 619-42
- Thompson, E. R., (1998) The many roles of c-Myc in apoptosis Annu. Rev. Physiol 60: 575-600
- King, K. L., e Cidlowski, J. A., (1998) Cell cycle regulation and apoptosis Annu. Rev. Physiol 60: 601-17
- Kroemer, G., Zamzani, N. E Sugia A. Santos (1997) Mitochondrial control of apoptosis Science 18: 1-44-50
- Ashkenazi, A., e Dixit V. M., (1998) Death receptors Signaling and modulation Science 281: 1305-1308
- Nagata, S., (1997) Apoptosis by death factor Cell 88: 355-65
- Grell, M., Douni, E., Wajant, H., et al (1995) The transmembrane form of tumor necrosis factor is the prime activating ligand of the 80 kDa tumor necrosis factor receptor Cell 83: 793-807

## 6 - Ciclo celular

- Sher, C. J. (1996). Cancer cell cycles. Science, 274-1372 a 1677
- Israels, E. D. & Israels L. C. (2001) – The cell cycle, Stem Cells, 19, 1, 88-91
- Murray A.W. e Marks, D. ( 2001 ) Can Sequencing shed light on cell cycling? Nature 409: 844-846
- Johnson, D.G. e Walker, C.L. ( 1999) Cyclins and cell cycle checkpoints Annu. Rev. Pharmacol. Toxicol 39: 295-312
- Murray, A.W. e Marka, D. ( 2001 ) Can sequence shed light on cell cycling? Nature 409? 844-846
- Morgan, D.O. Cyclin-dependent kinases:engines, clocks, and microprocessors. Annu. Rev. Cell. Dev. Biol. 13: 261-291 ( 1997 )
- Amon, A. The spindle checkpoint. Curr. Opin. Genet. Dev. 9: 69-75 ( 1999 )

## 7 – Oncogénese

- Varmus, H., e Weinberg, R. A., (1993) Genes and the Biology of cancer Scientific American Library
- Barbacid, M., (1987) ras Genes Annu. Rev. Biochem 56: 779-827
- Bojushi, M. S., e Molormick, F., (1993) Proteins regulating Ras and its relatives Nature 266: 643-654
- Famifi, A., Harrington, E. A., e Evan, G. I., (1992) Cooperative interaction between c-myc and bcl-2 proto-oncogenes Nature 359: 554-556
- Fleg, L. A., (1993) The many roads that lead to Ras Science 260: 767-768
- Hunter, T., (1993) Editorial overview: oncogenes and the cell cycle Curr. Opin. Genet Dev 3: 1
- Motokura, T., e Arnold, A., (1993) Cyclin D and oncogenesis Curr. Opin Genet Dev. 3: 5
- Nishida, E., e Gotoh, Y., (1993) the Map kinase cascade is essential for diverse signal transduction pathways Trends Biochem Sci. 18: 128-131
- Futreal, P. <sup>a</sup>, Kasprzyk, A., et al (2001) Cancer and genomics Nature 409: 850-852
- Jimenez-Sanchez, G., Childs, B., e Valle, D., (2001) Human diseases genes Nature 409: 853-855
- Higginson, J., Muir, C. & Munoz, H., 1992 Human cancer : Epidemiology and Environmental causes (Cambridge Monographs on cancer Research, Cambridge, UK)
- International Human Genome Sequencing Consortium, (2001) Initial sequencing and analysis of the human genome, Nature 409, 860-921
- Kaclin, W. G. Jr.(1999) The p53 gene family. Oncogene 18, 7701-7705
- Rowley, J. D., The critical role of chromosome translocations in human leukemias. Annu. Rev. Genet. 32, 495-519 (1998)
- Bell, R. S., Wunder, J. & Andrullis, I. (1999) Molecular alterations in bone and soft-tissue sarcoma. Can. J. Surg. 42, 259-266
- Jimenez-Sanchez, G., Childs, B. & Valle, D. In The Metabolic and Molecular Bases of Inherited Disease (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 167-174 (McGraw-Hill, New York, 2001)
- Antonarakis, S. E. & McKusick, V. A. (2002) OMIM passes the 1.000 disease gene mark. Nature Genet 25, 11
- International Human Genome Sequencing Consortium, (2001) Initial sequencing and analysis of the human genome. Nature 409, 860-921
- The Gene Ontology Consortium, (2000) Gene ontology: Tool for the unification of biology. Nature Genet 25, 25-29
- Rubin, G. M., et al, (2000) Comparative genomics of the eukaryotes. Science 287, 2204-2215
- Brusilow, S. W., Valle, D. & Arn, P. H., in Current Therapy in Neonatal Perinatal Medicine (ed Nelson, N. M.) 164-169 (B. C., Decker, Philadelphia, 1989)
- Tracy, E., Childs, B. & Scriver, C. R. (1995) Response to treatment in hereditary metabolic disease: 1993 survey and 10 year comparison, Am. J. Hum. Genet. 56, 359-367

## Iconografia

As figuras e quadros (bem como os textos que lhes dão apoio) indicados na respectiva iconografia foram adaptados e redesenhados a partir das seguintes fontes:

Página 3-Proteínas do relógio circadiano mamífero e seus domínios. In King,D.P. & Takahashi-Molecular genetics of circadian rhythms in mammals.Ann. Rev. Neurosci., 2000, 23. p.719.

Página 5- Idem anterior pagina 718.

Página 10-Modelo do trabalho do relógio circadiano dentro do SCN. In Clayton, J. D. et all..Keeping time with the human genome. Nature, 2001, 409, p.829.

Página 13- 11-cis-retinall-trans-retinall-trans-retinol. In Stryer, L.. Biochemistry.1995 4<sup>th</sup> edition,p.333.

Página 14- Estrutura da rodopsina nos vertebrados, In idem anterior pag. 334.

Página 16- Proteínas do ciclo visual. In Devlin, T. M.,- Textbook of biochemistry with clinical correlations.1997-4 th ed.,p.943.

Página 18- In idem anterior, pag.944.

Página 21 – Transdução de sinal num neurónio olfatório utilizando receptor transmembranário de sete hélices. In Gibson, A.D. & Garbers, D.L.-Guanylyl cyclases as a family of putative odorant .Ann.Rev.Neurosci. , 2000, 23, p.431.

Página 22- Modelo para mapa combinatório de cheiros no bulbo olfativo. In idem anterior pag. 432.

Página 24- Distribuição e função de guanilil ciclases mamíferas ligadas a membranas. In idem anterior pag.423.

Página 25- Transdução de sinal num neurónio olfatório utilizando receptor guanilil ciclase ligado a membrana ou guanilil ciclase solúvel. In idem anterior pag.629.

Página 28- Indução de alterações metabólicas pelo stress. In Devlin, M. T., Textbook of biochemistry with clinical correlations.Wiley-Liss, 1997, p.554.

Página 30- <http://www.copewithcytokines.de/>

Página 31-In idem anterior

Página 32- Proteínas de fase aguda da inflamação e sua função.In idem anterior

Página 43- 44-Classificação global das citocinas baseada na sua estrutura tri-dimensional.In [http://cmbi.bjmu.edu.cn/cmbidata/cgf/CGF\\_Database/cytweb/cyt\\_strucs/index.html](http://cmbi.bjmu.edu.cn/cmbidata/cgf/CGF_Database/cytweb/cyt_strucs/index.html)

Página 47- Interleucinas. In MED Systems.Diagnostic GMBH,Vienna.Austria.

Página 48- Mecanismo da resposta inflamatória e algumas cascatas bioquímicas induzidas pela activação celular. In Martins Silva, A.. Apontamentos de química fisiológica.Faculdade de Medicina.Lisboa.

Página 50- Citocinas,transdução de sinal das citocinas e citocinas hematopoiéticas. In idem anterior.

Página 51- Alguns factores de crescimento e sua acção In Alberts, B. et all.-Molecular Biology of the Cell.2th ed.,1989,Garland Publishers Inc.,p.747.

Página 52- Alguns factores estimuladores de colónias(CSF) que influenciam a formação de celulas sanguineas. In idem anterior pag. 981.

Página 53- Algumas linfocinas segregadas pelas celulas T em resposta a antigenio. In Lodish, H. et all.Molecular Cell Biology.Scientific American Books, W. H. Freeman and Company, New York, 1995, p.133

Página 54- Propriedades de algumas interleucinas(IL). In Alberts,B. et all. referido anteriormente, pag.1049.

Página 56- Comparação da apoptose com a necrose. In Kroemer, G. et all.- The mitochondrial death/life regulator in apoptosis and necrosis.Ann.Rev. Physiol.,1998,60,p.621.

Página 59- Apoptose por duas vias diferentes. In Nijhawan, D. et all.-Apoptosis in neural development and disease.Ann.Rev. Neurosci.,2000,23,p. 79.

Página 62- Proteínas adaptadoras. In Strasser, A. et all.-Apoptosis signaling,Ann.Rev. Biochem., 2000, 69, p.221.



- Página 63-Família de receptores TNF. In idem anterior pag.223.
- Página 65- Família Bcl-2. In idem anterior pag.224.
- Página 68- Apoptose induzida pelo CD 95- In Ashkenazi, A. &Dixit, V. M.-Death receptors:signaling and modulation.Science,281, pag.1306.
- Página 69- Sinalização pró-apoptótica e antiapoptótica do TNFR1. In idem anterior .
- Página 70- Sinalização pró-apoptótica e antiapoptótica e DR3. In idem anterior.
- Página 84 –Controlo do crescimento celular e tipos de proteínas participantes cujas formas mutantes originam cancro.In Lodish, H. et all. anteriormente referido, pag.1259.
- Página 86- Oncogenes e proteínas expressas pelos seus proto-oncogenes .In idem anterior pag. 1262 e 1263.
- Página 87- Proteínas receptoras de proto-oncogenes. In idem anterior, pag.1261.

# **BIOLOGIA MOLECULAR NA PRODUÇÃO ANIMAL**

## ÍNDICE

<b>1 - <u>Glândula mamária e segregação láctea</u></b>	<b>541</b>
1.1 - Introdução à biologia da glândula mamária	541
1.2 - Factores de crescimento e diferenciação na glândula mamária	544
1.3 – Puberdade	548
1.4 - Involução mamária	550
1.5 - Somatostatina, opióides progesterona e progestinas no desenvolvimento da glândula mamária	552
1.6 - Papéis da prolactina no desenvolvimento da glândula mamária	553
1.7 - Factores de transcrição e cascatas de transdução de sinais da glândula mamária	562
1.8 - Sinergias e antagonismos interactivos dos factores de transcrição na regulação da expressão dos genes das proteínas do leite	565
1.9 - Caspases e plasmina na biologia da glândula mamária involutiva	586
1.10 - Estrutura e funções dos genes das proteínas do leite	588
1.11 - Fracções proteicas do leite e seus polimorfismos genéticos	598
1.12 - Regulação da fracção lipídica do leite	602
Bibliografia	605

## **1 - Glândula mamária e segregação láctea**

### **1.1 - Introdução à biologia da glândula mamária**

Em todos os animais as glândulas mamárias são muito semelhantes, apesar de ocorrerem particularidades próprias das diversas espécies animais, inclusive no que se refere à composição do leite por elas segregado.

O tecido secretor ou parênquima e o tecido conectivo ou estroma são os principais constituintes da mama, embora as quantidades de cada um deles varie muito consoante as circunstâncias.

É conhecido nas várias espécies animais que o peso e respectiva capacidade secretora aumentam até os animais atingirem a maturidade, o que no caso das vacas leiteiras ocorre à volta dos seis anos.

Dando mais atenção à anatomia do úbere bovino recorda-se que as suas duas metades estão separadas por um ligamento suspenso médio que vem da túnica abdominal e que é constituído por duas lamelas de tecido conectivo elástico amarelo.

Lateralmente estes úberes são contidos por ligamentos suspensores laterais constituído em grande parte por fibras não elásticas. Destes ligamentos laterais emanam numerosas lamelas para dentro do úbere e que se ligam com o tecido intersticial deste.

Na superfície ventral destes úberes bovinos as lamelas dos ligamentos laterais e médios antes referidos, formam uma espécie de saco que contem o úbere.

No que se refere ao sistema de recolha do leite ao nível do úbere recorda-se que a organização do sistema de colecta de leite também é muito semelhante nas diversas espécies animais. O tecido secretório constituído por alvéolos e ácinos, comunica através de grandes canais com a cisterna da glândula e esta por seu turno, abre para a cisterna da teta que por seu turno, através do canal da teta abre para o exterior.

Existe uma estrutura que irradia para cima a partir do canal da teta para a teta e que é a roseta de Furstenberg constituída por uma série de dobras da membrana mucosa, dobras estas que têm também uma série de dobras secundárias, ajudando este conjunto à retenção do leite o que é sobretudo devido, nos bovinos, ao esfíncter muscular que envolve o canal da teta.

Nos ovinos este esfíncter é essencialmente constituído por tecido elástico e nos suínos há habitualmente dois canais da teta em cada teta que drenam áreas distintas.

A natureza do epitélio que reveste todo este sistema de recolha do leite nos animais é um tanto distinto. Assim ao nível dos alvéolos encontra-se uma simples camada de epitélio e é aqui que é formado o leite a segregar. O epitélio da pele que reveste a teta é escamoso e estratificado, e o epitélio que reveste interiormente o canal da teta é transicional, enquanto a teta e as cisternas da glândula são atapetadas por um epitélio de duas camadas de células cubóides que têm sobrepostas uma camada de células altamente cilíndricas.

O crescimento mamário que ocorre em determinadas fases da vida da fêmea tem diversas características consoante essas fases. Assim durante a gestação ocorre uma hiperplasia celular mamária. No início da lactação ocorre uma hipertrofia celular mamária e durante a involução mamária que acompanha o declínio da lactação ocorre morte celular.

Durante esta involução post-lactação ocorre pois um declínio muito acentuado da massa de tecido secretório da mama, enquanto nas primeiras semanas da lactação aumenta o crescimento deste tecido secretório da mama.

Também o perfil de crescimento mamário difere entre a primeira e a segunda lactação sendo mais acentuado nesta última. Contudo este tamanho da mama vai crescendo com os sucessivos ciclos ou períodos de lactação.

Está provado que durante as primeiras semanas de lactação ocorre alguma proliferação celular e na mama em lactação a ordenha frequente e o tratamento com BST (hormona do crescimento bovino) induzem hipertrofia celular, hiperplasia celular ou reduzida morte celular.

O parênquima mamário acompanha normalmente o perfil da produção de leite, sendo maior quanto maior a produção de leite ou seja o tecido secretório mamário está correlacionado com a produção de leite.

Após paragem da ordenha o parênquima mamário declina, embora sem voltar ao seu estado virgem inicial, e no segundo período de lactação seguinte aumenta mais do que durante o primeiro período de lactação.

O que temos vindo a referir traduz sobretudo o que se passa nos bovinos, embora se saiba que nas cabras o crescimento da mama durante a última parte da segunda gestação é muito mais rápido do que na primeira gestação e daí mais parênquima mamário e mais leite na segunda lactação.

A diferenciação celular do tecido mamário tem lugar em várias etapas. Assim ela começa algumas semanas antes do parto, passando depois por uma fase de pouca diferenciação até ao parto, acelerando-se depois até à fase do pico da lactação.

Após se atingir este pico de lactação o estado de diferenciação do tecido secretor mamário estabiliza, sendo posteriormente o declínio da lactação devido à diminuição do número de células secretoras da mama.

Pode no entanto afirmar-se que durante a gestação o principal elemento que contribui para o crescimento do parênquima mamário é a proliferação celular (hiperplasia) enquanto entre o parto e o pico da lactação predomina a diferenciação celular (hipertrofia).

A diferenciação celular do tecido secretório mamário parece ainda aumentar em consonância com a frequência das ordenhas, o mesmo parecendo suceder com o número de células novas, havendo no entanto a hipótese de ao mesmo tempo haver uma diminuição de morte celular.

Nas vacas esta situação está bem caracterizada havendo com o aumento da frequência das ordenhas, além do aumento da diferenciação celular, um maior diâmetro dos alvéolos secretórios e maior tamanho também celular.

A diminuição do número de ordenhas promovem uma involução mamária com a conseqüente quebra da produção de leite que parece devida a uma actividade celular mais reduzida sem diminuição do número de células.

A produção da hormona do crescimento bovino (BST) por processos biotecnológicos está hoje industrializada promovendo a sua administração aumentos da produção leiteira rondando os 20%,

observando-se também que a sua administração a novilhas púberes e pré-púberes aumenta o volume do parênquima mamário. Nas vacas adultas a administração de BST durante o período seco, desencadeia um aumento da produção de leite apenas nos trinta dias da lactação subsequente.

Os efeitos desencadeados pela BST parece envolverem uma distribuição homeorretica diferente dos nutrientes na glândula mamária, não parecendo aumentar a síntese de DNA na mama, nem promover efeitos mitógenicos nem de proliferação celular, embora pareçam reduzir a morte celular no parênquima mamário. A acção da BST desencadeia os seus efeitos sobre a glândula mamária através de IGF-1 (Insulin-like growth factor I), promovendo um aumento da sua concentração sanguínea e também um aumento da sua produção local (forma parácrinica da IGF-1).

As células da glândula mamária possuem receptores para a IGF-1 embora o seu número não seja condicionado pela frequência das ordenhas.

A acção da BST ao nível da glândula mamária é sobretudo consequência de um aumento da longevidade celular devido a uma morte celular reduzida. As populações celulares podem sofrer alterações em consequência de um desequilíbrio entre a proliferação celular e a morte celular.

No entanto sabe-se que o declínio da lactação resulta mais na perda de células do que da diminuição da diferenciação celular. Mais, esta involução do tecido mamário que acompanha o declínio da lactação implica a remodelação do tecido mamário com uma próteolise intensa, atribuída a activação da plasmina (proteases do soro), cuja concentração no leite aumenta correlativamente, na vaca, com a involução mamária (mais adiante voltaremos a este assunto).

A BST administrada inibe o aumento de plasmina no leite.

Refira-se que a morte celular pode ser devida a uma resposta patológica a uma “agressão” de qualquer tipo, ou então é uma morte celular programada (apoptose) correspondendo a um processo fisiológico controlado de um ciclo de mudança do desenvolvimento celular, parecendo ser um processo que depende de hormonas (provavelmente pela diminuição da PRL).

Refira-se desde já, embora mais adiante se desenvolva mais este assunto, que a prolactina (PRL) revela grande importância nos ruminantes no início da lactação (lactogénese) ao contrário do que sucede depois durante o restante período de lactação, tendo portanto um papel oposto ao da BST.

Mais, se se inibir a secreção de PRL em vacas e cabras a produção de leite não parece ser afectada, nem a administração de PRL exógena produz qualquer efeito sobre a produção de leite o que não é contudo contestado por experiências realizadas.

Também a frequência de ordenhas condiciona nas vacas, o número de receptores mamários para a PRL. Três ordenhas diárias aumentam este número e ordenhas incompletas reduzem-no.

Se a existência no epitélio mamário de receptores para a PRL não oferece dúvidas, o mesmo não sucede com os receptores para a GH (hormona do crescimento), apesar de estarem caracterizados RNA mensageiros para estes receptores de GH em bovinos e noutras espécies animais.

## **1.2 - Factores de crescimento e diferenciação na glândula mamária**

A glândula mamária constitui um órgão particular pois a maior parte do seu desenvolvimento ocorre após o nascimento dos animais, e depois ao longo da sua puberdade, e ainda durante as sucessivas gravidezes e correspondentes períodos de lactação.

### **Factor de crescimento fibroblástico e suas características no desenvolvimento mamário durante a gravidez.**

Os factores de crescimento dos fibroblastos (FGF) são pleiotróficos podendo consoante as células desencadear cascatas de transdução de sinal que possam redundar em efeitos mitógenicos de espectro largo, ou promovendo migração celular ou ainda modelando a diferenciação celular.

Os FGF como sinal são importantes no desenvolvimento do lóbulo alveolar do parênquima mamário.

Estão identificados sete membros na família FGF.

Genericamente a estimulação desencadeada pelos FGF aumenta o pH intracelular e as concentrações de  $Ca^{2+}$ , o turnover PI (fosfoinositol), assim como a fosfatação de uma série de proteínas intracelulares tais como PLC-, Raf-1, ERK-1 e ERK2 cinases e SG cinase ribossomal (vide cascatas de transdução de sinais).

Estes ligandos, os FGF segregados, constituem um sinal que se liga aos receptores da superfície celular, com alta afinidade, induzindo a dimerização destes e a subsequente activação dos seus domínios tirosina cinase do lado citoplásmico dos receptores.

A estrutura destes receptores para os FGF ou sejam os FGFR, possuem duas ou três ansas semelhantes às imunoglobulinas e que são o local que interactua com o ligando sinal no seu domínio extracelular, havendo depois um elemento que atravessa a plasmamembrana celular e que se continua no domínio citoplásmico dotado da actividade já referida.

Há quatro genes para os receptores FGF (FGFR1, 2, 3 e 4) nos mamíferos, havendo no entanto processos de “splicing” alternativos dos FGFR 1, 2 e 3 que originam sete protótipos de receptores com diferentes especificidades para os FGF e com diferentes distribuições celulares.

### **O factor de crescimento hepatocito e a neuregulina na morfogénese celular da glândula mamária**

O epitélio da glândula mamária é controlado no seu crescimento e morfogénese por interacções entre o mesênquima e as células epiteliais o que é feito através da intervenção de diversos factores.

Estes factores interactuam com receptores específicos do tipo tirosina cinases contidos nas células epiteliais, através de sistemas de sinalização parácrinos, pois partem de células mesênquimatosas vizinhas.

Dois destes factores muito importantes “in vitro” para a morfogénese e diferenciação das células epiteliais são o factor de desenvolvimento do hepatocito/scatter factor (HGF/SF) e a neuregulina.

Nestes sistemas assinalantes parácrinos são referidos o HGF/SF, a neuregulina e membros da família de receptores c-erbB, o receptor c-met, os receptores c-ret, o factor neurotrófico derivado de linhas celulares gliais (GDNF) e ainda os FGF antes referidos e os seus receptores.

Destes sistemas o primeiro, o HGF/SF induz a formação de estruturas tubulares na mama e o segundo a neuregulina induz a formação de estruturas alveolares.

Qualquer um destes dois factores são expressos na glândula mamária a partir do mesênquima que envolve as células epiteliais que por seu turno são dotadas de receptores para estes factores.

Durante a puberdade o factor HGF/SF induz a formação de ramificações dos canais inibindo por outro lado a diferenciação terminal inclusivé a expressão das proteínas do leite. A neuregulina durante a gravidez promove a formação lóbulo-alveolar inclusivé a expressão de componentes do leite como a  $\alpha$ -caseína.

As cascatas de transdução de sinal induzidas por estes dois factores também são diferentes pois o primeiro através do seu receptor c-met envolve Gab 1 e PI cinase (Gab 1 é uma subunidade  $\beta$ 1 de um receptor gama aminobutirico; c-met é um receptor do factor de crescimento do hepatocito que tem uma actividade de proteína tirosina cinase constituído por duas subunidades  $\alpha$  (50kDa) e  $\beta$  (145kDa) ligadas por ponte dissulfureto).

Na neuregulina parece que um receptor c-erbB2 (receptor tirosina cinase que pertence à família dos receptores EGF mas que não é activado nem pelo EGF, nem pelo TGF- $\beta$  nem pela anfiregulina) é suficiente para desencadear os seus efeitos biológicos admitindo-se que intervenha uma MAPK cinase (Mitogen-activated protein kinase).

No tecido mamário tanto o desenvolvimento dos canais como o desenvolvimento dos alvéolos podem ser influenciados pelo EGF (factor de crescimento epidérmico) e pelo TGF- $\beta$  e TGF- $\alpha$ , que derivam das células epiteliais, sobretudo após o parto, e actuam de forma autócrina.

#### Expressão e localização durante o desenvolvimento das glândulas mamárias bovinas, de alguns factores de crescimento.

O desenvolvimento e a lactação implicam a actuação de uma série de entidades hormonais, bem como diversos factores de crescimento que induzem o crescimento e a diferenciação da glândula mamária, o que ocorre ciclicamente, durante as várias etapas.

É sobejamente conhecido que hormonas esteróides e peptídicas controlam todo este processo que a nível celular é mediado ainda pela intervenção de vários factores de crescimento actuando de forma autócrina, justacrinica ou paracrinica.

Vejamos como se exprimem nas glândulas mamárias bovinas durante a mamogénese, lactação e involução vários factores de crescimento IGF 1 ou insulina-like growth factor I, FGF-1 ou factor de crescimento fibroblástico 1, FGF-2 ou factor de crescimento fibroblástico 2 e TGF- $\beta$  ou factor de crescimento transformante, bem como o FGFR ou seja o receptor do factor de crescimento fibroblástico e o GHR ou receptor da hormona do crescimento.



No quadro 1 seguinte, inspirado em Sinowatz et al. referem-se os dados semi-quantitativos para diferentes factores de crescimento em tecido mamário bovino em diferentes estádios de desenvolvimento dos úberes, e o perfil de distribuição histoquímico da IGF-1 e GHR.

Legenda do quadro 1

I – Mamogénese

I.1 – Mamogénese com desenvolvimento ductal (novilhas não grávidas)

I.2 – Mamogénese com desenvolvimento lobulo-alvéolar durante a 1<sup>a</sup>

Gravidez

I.2.1 – 194 a 213 dias de gravidez

I.2.2 – 255 a 272 dias de gravidez

II – Lactogénese: 5 a 11 dias após o parto

III – Galactopoiese

III.1 – Pico da lactação

III.2 – Fim da lactação

IV – Involução (3 a 4 semanas após secagem, sem gravidez)

Quadro I

Estádios de desenvolvimento do úbere bovino

	Mamogénese			Lactogénese	Galactopoiese		Involução		
	I.1 Novilhas não gravidas (ductos)	I.2.1 194 a 213 dias de gravidez (lobulo- alveolar)	I.2.2 255 a 272 dias de gravidez (lobulo- alveolar)	II Lactogénese 5 a 11 dias após parto	III.1 Pico da lactação	III.2 Fim da lactação	IV Secagem 3 a 4 semanas após a secagem, não grávida		
Dados semi-quantitativos	IGF-1	+++	+	++	-	-	-	+---	
	TGF $\beta$	++	+	+	-	-	-	++	
	FGF-1	+++	++	++	-	-	-	+	
	FGF-2	++	+	+++	+	+	-	+	
	FGFR (receptor)	++	+	++	+	+	±	+	
	Perfis histoquímicos	<u>IGF-1</u>							
		<u>Epitélio</u>							
		Ductos	+	+	±	±	±	+	++
		Alvéolos	Não	-	-	±	-	+	++
		Mioepitélio	-	-	-	-	-	-	-
<u>Estroma</u>									
Fibroblasto		+	+	±	+alguns	+alguns	+alguns	+	
Adipocitos		++	++	++	++	Não	Não	++	
Endotélio		-	-	-	-	-	-	-	
Músculo liso		++	++	+---	+	+	+---		
Macrofagos		+++	+++	+++	+++alguns	+++alguns	+++	+++vários	
<u>GHR</u>									
<u>Epitélio</u>									
Ductos		++	+---	+---	+---	++	++	+/-	
Alvéolos	Não	+	+	+---	++	++	+/-		
Mioepitélio	+/-	-	-	-	-	-	-		
<u>Estroma</u>									
Fibroblasto	10- 20%+	10- 20%+	+	+	+	+	+alguns		
Adipocitos	++	+---	+---	+---	Não	Não	+		
Endotélio	+	++++	+---	+	+	+	+		
Músculo liso	++	++	++	++	++	++	++		
Macrofagos	+++	+++	+++	+++alguns	+++alguns	+++	+++vários		

Todos os factores de crescimento referidos no quadro anterior têm uma expressão relativamente elevada nas novilhas virgens, tendem depois a reduzir a sua expressão (IGF-1 e TGF- $\beta$ ) ou a nivelar-se (FGF-1 e FGF-2) durante a gravidez (I.2.1 e I.2.2) passando depois essa expressão a ser negativa ou mais fraca durante a lactogénese (II) e a galactopoiese (III.1 e III.2) aumentando depois de novo durante o período da involução (IV).

O epitélio dos ductos revela uma fraca ocorrência de IGF-1 durante a mamogénese.

O epitélio dos alvéolos é negativo para o IGF-1 durante a mamogénese, lactogénese e galactopoiese mas denota actividade durante a involução.

No estroma, no que se refere a IGF-1 verifica-se alguma expressão no citoplasma dos adipócitos e células vasculares dos músculos lisos, revelando os macrófagos uma forte expressão deste factor.

Não se conhece bem o papel fisiológico da GH na mamogénese e na lactação, tal como o papel do respectivo receptor GHR.

No quadro anterior em apreço foi possível verificar que o epitélio dos ductos apresenta expressão destes receptores durante a maioria dos estádios do desenvolvimento mamário, enquanto os epitélios alveolares contêm uma quantidade moderada de GHR durante a gravidez que aumenta depois durante a lactação e, galactopoiese, enquanto nas vacas secas diminui ou não existe mesmo.

Apesar de se conhecer o importante papel da GH na manutenção da produção de leite nos ruminantes o seu mecanismo preciso ao nível da glândula mamária não é claro tal como a falta de GHR neste órgão.

Tem sido sugerido que a GH possa desempenhar papéis homeostáticos, tais como variações na partilha de energia, admitindo-se mesmo que a GH aumente a produção de leite de uma forma indirecta através da estimulação da produção de IGF-1 quer ao nível do fígado quer ao nível do úbere e seria este que aumentaria a produção de leite.

Parece provado o papel directo da GH através do seu receptor pois foi demonstrada a presença de uma quantidade distinta de mRNA codificador de GHR e da sua proteína receptora durante a lactação e a galactopoiese, enquanto ao mesmo tempo não se verificava a presença no epitélio glândular de IGF-1 ou era muito baixa e no estroma só era fracamente expressa.

### **1.3 - Puberdade**

A puberdade das novilhas no que toca ao crescimento e desenvolvimento da glândula mamária condiciona fortemente a potencial capacidade produtiva leiteira dos animais quando atingirem a maturidade.

Como já se disse também nesta fase da puberdade as hormonas e factores de crescimento intervêm interagindo uns com os outros quer promovendo activações quer inibições.

O papel dos estrogénios e da GH são fundamentais, e os factores de crescimento intervêm quer de uma forma sistémica quer de uma forma específica.

Contudo com células epiteliais mamárias cultivadas “in vitro” as hormonas estrogénicas e a GH pouca acção estimuladora revelam sobre o seu crescimento sendo neste aspecto muito mais importantes factores sistémicos ou de origem local, tais como: a insulina (INS) e os factores de crescimento like-insulina (IGF-1 e IGF-2); o factor de crescimento epidérmico (EGF); os factores de crescimento fibroblástico (FGF-1 e FGF-2); o factor de crescimento transformante  $\alpha$  (TGF  $\alpha$  e TGF  $\alpha$  1,2 ou 3); anfíregulina (AR); factores decrescimento derivados das plaquetas (PDGF) e factor 1 de crescimento derivado da glândula mamária (MDGF-1).

De todos eles os IGF-1 e a sua variante natural des (1-3), os IGF II e a INS têm um efeito mitogénico com a correlativa síntese de DNA, sobretudo os dois primeiros.

Contudo a acção mitogénica promovida pela IGF-1 pode ser bloqueada pelas proteínas que se ligam especificamente a elas ou sejam as IGFBP-2 e 3 o que é extrapolável para o que acontece ao nível da glândula mamária.

Estes factores IGF e IGFBP podem ser formados ao nível do fígado ou localmente no tecido mamário ou noutros tecidos, encontrando-se portanto no soro sanguíneo, no colostro, no leite, linfa e fluido folicular.

Durante a puberdade a proliferação do epitélio mamário pode também ser desencadeada por factores de crescimento da família EGF, TGF- $\alpha$  e AR embora mais tenuamente do que a promovida pela família IGF. Estão identificados dois tipos de receptores para as IGF:

A família de factores de crescimento epidérmicos engloba a EGF, a heparina ligante da EGF, o factor  $\alpha$  transformante do crescimento (TGF- $\alpha$ ), a amipiregulina e várias heregulinas que contêm um ou mais módulos EGF, e actuam através do receptor da EGF (Erb B-1), situado na plasmamembrana, muito específico e de alta afinidade para os ligandos.

A família FGF consiste de mais de 15 polipéptidos relacionados, muito distribuídos através dos organismos e desencadeando os seus efeitos através de pelo menos cinco receptores FGF tirosina cinases (FGFRs), que apesar de expressos nos bovinos durante o desenvolvimento, a lactação e a involução, as suas concentrações são maiores nas novilhas virgens ou nas primíparas ou ainda durante a involução mamária.

Nas glândulas mamárias dos ruminantes as respectivas células são reguladas no seu crescimento por três FGF.

O FGF-1 é ácido (aFGF) o FGF-2 é básico (bFGF) e o FGF-7 é um factor ceranocítico (KGF).

A família TGF- $\alpha$  contendo diversas proteínas multifuncionais condicionam a formação da matriz extracelular podendo também afectar a diferenciação e crescimento celular consoante o tipo de células. A TGF- $\alpha$  1 é segregada por quasi todas as células havendo três isoformas a TGF- $\alpha$  1,  $\alpha$  2 e  $\alpha$  3, tendo a primeira delas três tipos de receptores.

Assim a TGF- $\alpha$  1 estimula a proliferação das células epiteliais mamárias (em baixas concentrações) e pelo contrário inibem essa proliferação em altas concentrações sendo factores importantes durante a puberdade.

Outros factores reguladores do crescimento e desenvolvimento mamário são o factor 1 de crescimento derivado da mama (MDGF-1) que é um potente agente mitogénico e tem sido encontrado no

leite bovino e de outros animais, e o PDGF (factor de crescimento derivado das plaquetas que actua indirectamente sobre as células epiteliais através do estroma.

#### **1.4 - Involução mamária**

Durante a involução mamária, como já referimos, ocorre um declínio muito acentuado da massa de tecido secretório da mama.

Por outro lado a prolactina (PRL) revela grande importância nos ruminantes, no início da lactação ao contrário do que sucede depois durante o restante período de lactação e involução, de forma um tanto oposta da da hormona do crescimento (GH) que desencadeia os seus efeitos sobre a glândula mamária através do IGF 1 (factor de crescimento like-insulina).

Os efeitos tanto da GH e da PRL ao nível da glândula mamária implicam a IGF-1. A GH aumenta a concentração deste factor e a PRL inibe a expressão da proteína 5 de ligação ao factor de crescimento like-insulina (IGFBP-5) a partir do epitélio mamário.

Isto significa que se a GH aumenta a IGF-1 a PRL tem uma actividade no mesmo sentido na medida em que inibe a expressão de uma proteína (IGFBP-5) que bloqueia a actividade da IGF.

Durante a involução mamária é conhecido o declínio da PRL no soro sanguíneo e daí a remoção da inibição da expressão da IGFBP-5, o que tem como consequência os seus mais elevados teores bloquearem a IGF-1 impedindo a interacção deste com os respectivos receptores resultando de tudo isto a apoptose da célula epitelial mamária.

O IGF-1 é um factor importante de sobrevivência das células do epitélio mamário, e o IGFBP-5 existindo em elevadas concentrações na glândula mamária em involução, inibe portanto as acções do IGF-1 a este nível.

A IGFBP-5 tem ainda a particularidade de interactuar com a  $\kappa$ -caseína e esta caseína favorece a conversão do plasminogénio em plasmina enzima. Isto é possível porque a  $\kappa$ -caseína interactua com o plasminogénio e com o seu activador tissular o tPA sendo a plasmina por seu turno um iniciador importante para remodelação da matriz extracelular (ECM) durante a involução mamária.

Também se admite que a IGFBP-5 ao bloquear a IGF-1 activa a apoptose, e sequestra também o PAI-I que é o inibidor 1 do activador do plasminogénio, resultando de tudo isto a quebra do plasminogénio e libertação da plasmina. A IGFBP-5 interactua ainda com uma série de componentes da matriz extracelular.

Por outro lado a PRL e a GH na glândula mamária inibem a actividade do tPA e consequentemente a conversão do plasminogénio em plasmina o que contribuiria para inibir a morte celular e a remodelação da matriz extracelular (ECM), logo o declínio da PRL durante a involução mamária favorece a morte celular, através do vector IGF e da  $\kappa$ -caseína do leite.

O leite em todas as espécies animais estudadas contém IGFs e IGFBPs e os teores dos primeiros são mais elevados no colostro do que no leite “maduro”.

As IGFbps são uma família de G proteínas que interactivam com as IGF-1 e 2. As IGFbps podem ser activadoras ou inibidoras, e actuarem também independentemente das IGF, tudo dependendo da sua concentração, estado de fosfatação e proteólise.

O colostro bovino contem IGFbp-1, 2, 3 e 4 em elevados teores, o 3 predominando nas secreções pré-parto, sendo reduzido após o parto, enquanto o IGF-1 existe no soro sanguíneo em pequenas concentrações.

Conforme já referimos o IGFbp-5 existe em elevadas concentrações na glândula mamária em involução, e pelo menos na ratinha o notável aumento de concentração de IGFbp-5, ocorre 24 horas após a indução da involução mamária por remoção da sucção pelas crias, ocorrendo também nestas circunstâncias aumentos mais modestos de IGFbp-2 e 4, mas não dos IGFbp 1,-3 e -6 todos eles expressos nas células epiteliais mamárias.

O leite porcino contem IGFbp-1, 2, 3 e 4 parecendo pelo menos o 2 e 3 serem expressos na glândula mamária, em vez de provir do sangue que chega à mama.

Estão identificados nos IGFbp-1 e 2 sequências RGD, uma característica das moléculas que interactivam com as integrinas da superfície das células.

As moléculas de IGFbp-5 têm locais que podem ser fosfatados por cinases adequadas, e envolvidos assim em interacções com outras proteínas como por exemplo a  $\kappa$ -caseína, sobretudo na forma dimérica. Este complexo como já referimos, interactiva com o plasminogénio e com o tPA (activador tissular) levando à formação da plasmina uma enzima que cinde uma série de pró-enzimas como a prostomelina e procolagenases, principiando a degradação da ECM que ocorre no fim do período de lactação.

Também se admite que a IGFbp-5 como já referimos interactiva com o PAI-I (inibidor 1 do activador do plasminogénio t-PA) bloqueando-o. O PAI-I não bloqueado ligando-se ao t-PA inibe as acções deste.

Haveria pois um duplo papel do IGFbp-5, por um lado ligando o IGF-1 para inibir as acções deste, induz apoptose, por outro ligando-se ao PAI-I bloqueia a acção deste e activa o t-PA originando a remodelação da ECM.

Teríamos pois associada a morte celular e a remodelação tissular da matriz ECM num processo altamente coordenado.

Independentemente das evoluções sofridas pela GH e pela PRL ao longo dos diversos períodos de lactação é possível referir o seguinte:

- a) A prolactina e a GH suprimem a activação das plasminas tal como inibem a apoptose
- b) A prolactina inibe a expressão da IGFbp-5 e como tal a IGF-1 pode desencadear as suas acções
- c) A GH aumenta a produção de IGF-1 que é um factor de sobrevivência celular

d) A prolactina inibindo a síntese de IGFbp-5 está a aumentar a disponibilidade de PAI-I suprimindo assim a produção de plasmina.

## **1.5 - Somatostatina, opióides, progesterona e progestinas no desenvolvimento da glândula mamária**

No tecido mamário está assinalada a existência de somatostatina e opióides, que se comportariam como reguladores negativos importantes no funcionamento da glândula mamária quer na secreção hormonal quer na regulação da proliferação celular.

A somatostatina produzida sobretudo no hipotálamo e no pâncreas, é um decatetrapéptido que se encontra contudo em diversos outros tecidos normais e cancerosos, existindo nos líquidos biológicos em duas formas a 14 e a 28. Interactuam as somatostatinas com receptores específicos transmembranários de sete hélices, dos quais existem cinco tipos diferentes (SSTR 1-5) encontrando-se em tecidos mamários normais e malignos, com a particularidade de nos tecidos mamários normais a actividade da somatostatina apenas residir no estroma, mas nos tecidos cancerosos essa actividade também existe nas células epiteliais.

O papel da somatostatina consiste em acções directas e indirectas, encontrando-se nas primeiras a inibição da proliferação celular (em função das concentrações de somatostatina) e nas segundas o seu efeito sobre a secreção de GH e a concentração de IGF-1.

Quanto aos opióides endógenos eles derivam de três proteínas precursoras distintas a próencefalina A e B (Pródinorfina) e a pró-opiomelanocortina (POMC) que podem originar diversos péptidos opióides com características distintas inclusivé com diversas afinidades para os respectivos receptores que são também receptores transmembranários de sete hélices e dos quais existem três tipos, delta, mu e kapa.

Outros derivados opióides estão descritos, provenientes por exemplo da hemoglobina, do glúten, e das caseínas \_ e \_, todos eles diminuindo a proliferação do tecido mamário.

Estes receptores de sete hélices por interacção com a somatostatina e opióides endógenos induzem uma diminuição do AMP cíclico, por inibição da adenilato ciclase e os segundos ligandos ou sejam os opióides podem ainda interactuar com elementos do citoesqueleto.

Estes receptores de sete hélices da somatostatina e dos opióides endógenos possuem no entanto uma série de homologias tornando possível interacções com efeitos aditivos com outros tipos de ligandos (quer ao nível do ligando quer ao nível do receptor). Estas interacções ocorrem por exemplo com a somatostatina e o EGF, e com opióides e receptores adrenérgicos ou da dopamina.

Também a hormona esteróica progesterona é fundamental para o desenvolvimento lóbulos-alveolar do epitélio mamário, interactuando a progesterona com receptores intracelulares do tipo A e do tipo B que por sua vez, após dímerização, interactuam com elementos específicos de resposta ao nível do DNA, havendo ainda proteínas específicas co-activadoras ou co-repressoras que são essenciais para a interacção com o complexo geral de transcrição.

A GH é produzida pela glândula pituitária mas também a glândula mamária o pode fazer por indução promovida por progestinas, pelo menos nos canídeos, embora se saiba que este fenómeno não é exclusivo desta espécie animal.

Sabe-se que a expressão do gene da GH mamária é regulada. No entanto o mRNA da GH mamária, transcrito, é idêntico ao produzido na pituitária, inclusivé o 5'UTR, parecendo que o promotor é idêntico.

Neste gene promotor da GH está reconhecido o elemento de resposta Pit-1 ou POU 1F1, que é essencial, embora na glândula mamária a GH seja expressa na ausência de expressão do Pit-1.

Contudo sabe-se que são necessários para a expressão de GH mamária, factores de transcrição adicionais.

A GH mamária expressa pode desencadear acções ao nível da própria mama onde induzem efeitos autócrinos e parácrinos sobre a proliferação e diferenciação do epitélio mamário, ou então essa GH mamária é libertada para o leite desencadeando portanto efeitos exócrinos.

Na figura 1 seguinte refere-se a contribuição da GH mamária para os efeitos endócrinos promovidos pela GH da pituitária

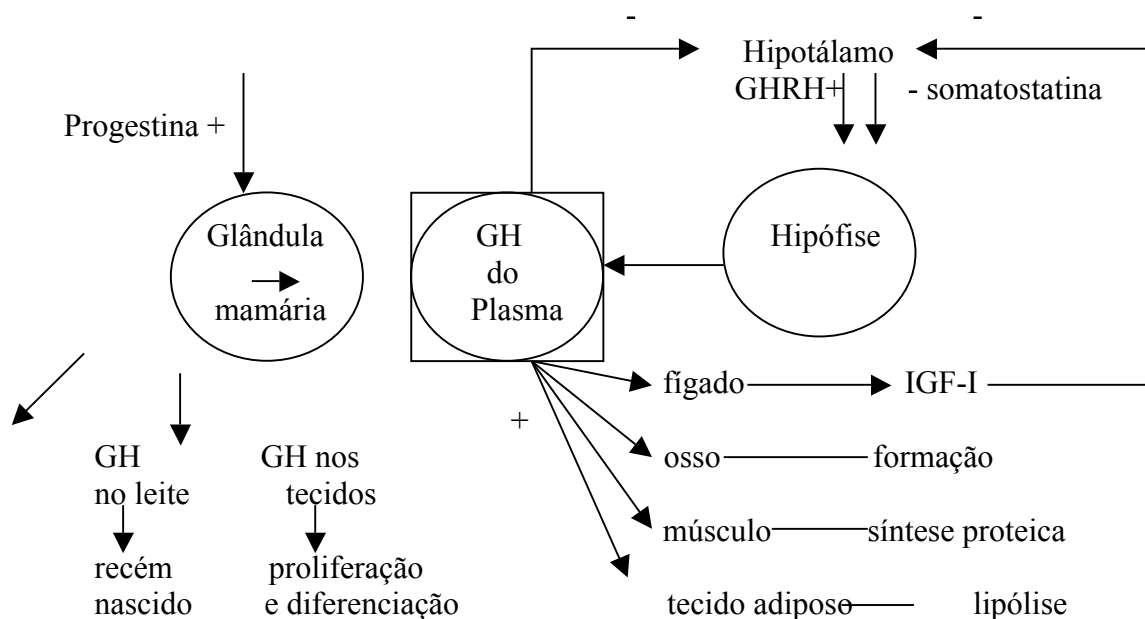
### 1.6 - Papeis da prolactina (PRL) no desenvolvimento da glândula mamária

A PRL produzida na hipófise e também em diversos outros tecidos não hipofisários, embora nestes últimos em menor quantidade, condiciona muito mais processos fisiológicos do que todas as outras hormonas da hipófise juntas, estando assinalado que mais de 300 efeitos biológicos distintos são desencadeados pela PRL.

A PRL tem um papel importantíssimo sobre a glândula mamária e também um pleotrofismo funcional muito grande, quer na reprodução, quer na osmoregulação, quer no comportamento e imunorregulação.

Figura 1

Contribuição da GH mamária para os efeitos endócrinos da GH produzida na pituitária, (efeitos locais parácrinos /autócrinos nos tecidos e efeitos exócrinos via leite



Convirá nesta altura fazer uma breve revisão da organogénese da glândula mamária.



Em ratinhas no momento do nascimento são visíveis rudimentos indistinguíveis da arquitectura mamária dos canais galactófaros.

Ocorre depois durante a puberdade o desenvolvimento mamário posterior que é regulado pelas hormonas da hipófise e das gónadas, que actuam directa e indirectamente e originando a formação da origem dos canais (TEBs ou terminal end buds) com a subsequente proliferação, iniciando-se também a formação dos epitélios dos ductos.

À medida que a puberdade avança formam-se as ramificações lobulares a partir dos sistemas de canais ou ductos por acção da PRL e progesterona completando-se depois a organogénese.

Após a gravidez ou melhor até meio da gravidez são evidentes o alongamento dos canais, a sua ramificação e o aumento do número de lóbulos, ao mesmo tempo que ocorre uma grande proliferação do epitélio lóbulo-alveolar tudo induzido pela PRL, pelos lactogénios placentários, progesterona e factores de crescimento locais.

Aquando do parto este efeito lóbulo-alveolar é transformado num fenótipo secretor, com a biossíntese das proteínas do leite e enzimas lactogénicas.

Depois no fim do período de lactação, ocorre a involução do sistema lóbulo-alveolar, em consequência da estase do leite e da diminuição da concentração dos lactogénios sistémicos.

Durante estas diferentes fases, aparecimento lobular, expansão lóbulo-alveolar durante a gravidez, diferenciação lactacional, manutenção da secreção láctea, a PRL é um regulador obrigatório de tudo isto, ou seja da organogénese, do crescimento lóbulo-alveolar e da diferenciação funcional.

A PRL também denominada hormona lactogénica ou hormona luteotrófica contém 199 ácidos aminados, sendo bem conhecidas a sequência do cDNA da PRL em diversas espécies animais.

A prolactina aparece numa multiplicidade de isoformas com várias formas postranslacionais, e com uma série de modificações químicas tais como fosfatações e glicosilações.

Os estímulos que promovem a libertação da PRL não são apenas o estímulo da amamentação, mas também a luz, a audição, a olfação e o stress têm esse papel.

A inibição da sua secreção não é só devida à dopamina de origem hipotalâmica, mas também a outros factores dentro do cérebro, hipófise, e órgãos periféricos.

O gene da PRL além de se exprimir na hipófise anterior encontra-se ainda no cérebro, decidua, miométrio, glândulas lacrimais, timo, baço, linfócitos circulantes, células linfóides da medula óssea, fibroblastos da pele, glândulas sudoríperas, encontrando-se o produto da sua expressão no soro, líquidos cefalo-raquidiano, amniótico, fluido folicular e suor.

A PRL é também sintetizada no epitélio mamário sendo segregada pelo leite em concentrações significativas e absorvida pelo intestino do recém-nascido induz a maturação do sistema hipotalâmico neuroendócrino.

A PRL sintetizada na hipófise segue a via endócrina clássica e a PRL sintetizada noutros tecidos não hipofisários actua de uma forma mais directa como um factor de crescimento, neurotransmissor ou

imunomodelador, de uma forma autócrina (sobre as próprias células que a segregam) ou parácrina (sobre células adjacentes) não influenciando portanto as concentrações de PRL hemáticas.

A PRL interaccua com o seu receptor específico que pertence à superfamília da classe I de receptores citocina, sendo o sinal de transdução mediado pelo menos por duas classes de moléculas assinalantes, as Janus cinases e as STAT (signal transducers and activators of transcription).

Contudo a translocação das hormonas polipeptídicas para dentro dos núcleos das células, e aí desencadeando a sua acção, tem nos últimos anos sido muito apontada sobretudo para a PRL.

O receptor para a PRL ou seja o PRLR encontra-se em muitos tecidos, inclusivé no epitélio mamário, nos linfocitos T e B e nos macrófagos, e nas células lúteas, aumenta o seu teor durante a gravidez, desencadeando a interacção PRL-PRLR proliferação e diferenciação celular actuando de forma endócrina, parácrina e autócrina, e ainda no tecido mamário activando a transcrição dos genes da proteína do leite durante a lactação.

Este PRLR apresenta diversas formas em diversos tecidos de diversas espécies animais, devido a diversos processos de “splicing”.

Da interacção PRL-PRLR resulta a dimerização dos receptores através da justaposição de domínio intracitoplásmico dos PRLR iniciando-se a cascata de transdução de sinal (Figura 2).

Ainda na região intracelular dos PRLR existem domínios que contêm as caixas 1 e 2 (são regiões homólogas à de outros receptores de citocinas) e estas caixas contêm motivos implicados na ligação com Jak 2 consistindo de ácidos aminados hidrofóbos prolina e hidroxiprolina e acídicos.

A caixa 1 é indispensável para o funcionamento do PRLR.

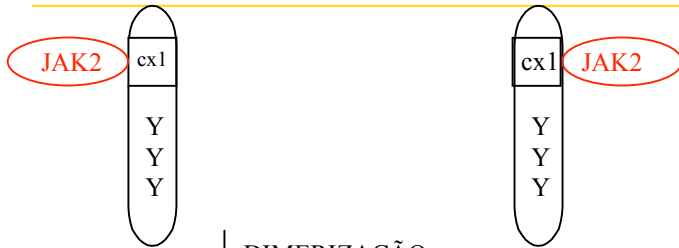
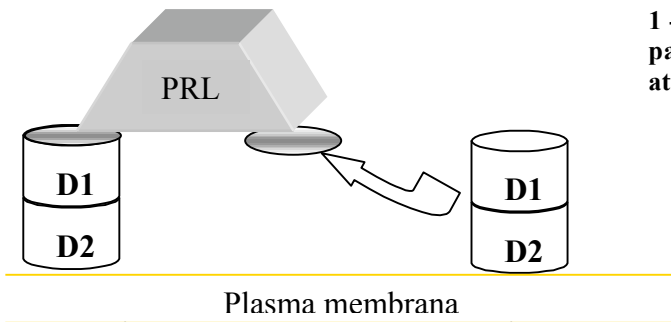
Ainda dentro da porção C-terminal do PRLR de rato os resíduos de tirosina que aí se encontram podem também contribuir para interacção com o STAT 5 e activação da Jak 2.

O domínio extracelular das isoformas do PRLR consiste de subdomínios D 1 e D 2, ambos análogos com moléculas de fibronectina do tipo III.

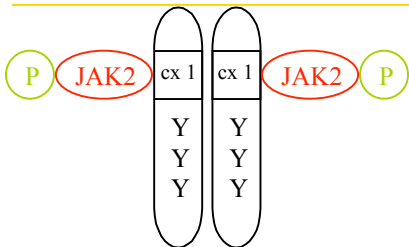
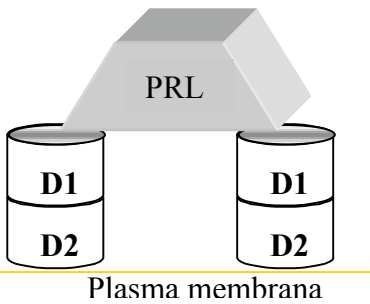
O domínio intracelular das isoformas do PRLR diferem no comprimento e na composição, possuindo no entanto duas regiões conservadas as chamadas caixas 1 e 2, além de quatro resíduos de tirosina.

A fosfatação da Jak 2 é importantíssima na activação de todas as isoformas do PRLR, contudo nas isoformas das formas curtas do PRLR, apesar de elas conterem os quatro resíduos de tirosina no seu domínio intracelular, estes não são fosfatados.

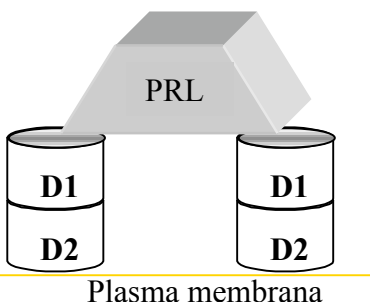
1 - Interação do ligando PRL (contendo dois locais para interaccionar) com uma molécula de PRL-R através do domínio NH2 terminal D1 desta.



2 - Indução da interacção do ligando PRL através do 2º lugar de interacção com uma segunda molécula de PRL-R formando um dímero que leva a activação da tirosina cinase Janus cinase 2 citosólica (associada com o domínio intracelular da PRL-R) que se transfosfata uma á outra.



↓ Fosforilação da JAK e do PRL-R



3 - As JAK2 cinases fosfatam as tirosinas do próprio PRL-R

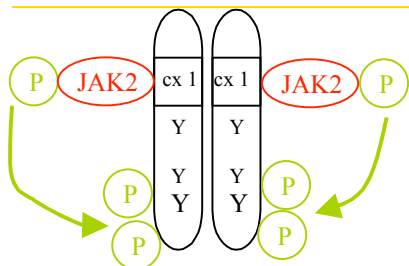


Figura 2

Da dimerização dos PRLR resulta a fosfatação rápida dos seus domínios assinalantes antes referidos os quais vão activar cinases com elas associadas como é o caso da Jak 2, da Fyn, da Shc-Grb 2-Sos, e da Bag-1/Bcl-2 que desencadeiam diversas cascatas de transdução de sinal que acabam por induzir a expressão de genes implicados na proliferação (como por exemplo o IRF-1, a ciclina B e a histona 3) e no fenótipo mamário diferenciado (como por exemplo proteínas do leite como a  $\alpha$ -caseína).

Vejam os em maior detalhe estas vias de transdução de sinais desencadeadas pela activação da PRLR.

Nas vias de transdução de sinais iniciados pela activação do PRLR há que considerar os seguintes aspectos. (vide figura 3 seguinte)

1) Há uma via Jak/STAT em que estão envolvidas moléculas transdutoras de sinal as STAT 1, 3, 5a e 5b.

As moléculas STAT contêm um domínio para interactuar com o DNA, um domínio SH3-like, um domínio SH2-like, e um domínio NH2 e um domínio COOH terminal de transactivação.

A fosfatação de tirosina na isoforma longa do PRLR desencadeia uma interacção com o domínio SH2 de uma molécula de STAT, que enquanto “docado” ao PRLR é fosfatado pela Jak cinase que está associada ao receptor.

Depois a STAT fosfatada dissocia-se deste complexo e homo ou heterodimeriza-se através dos domínios SH2 com outra molécula de STAT fosfatada, sendo este dímero translocado para o núcleo onde activa um local no promotor de um gene alvo chamado esse local GAS ( $\gamma$ -interferon activated sequence).

A tirosina das isoformas curtas do PRLR não são fosfatadas pela Jak 2, mas esta pode servir como local para “docagem” da STAT 1.

2) Há uma outra via de transdução de sinal a MAPK cascata que envolve uma proteína cinase (MAPK) que activa a mitogenese a qual está implicada na fosfatação e consequente activação de uma grande gama de factores de transcrição de genes de resposta imediata.

As tirosinas fosfatadas de isoformas longas do PRLR activadas interactuam com proteínas adaptadoras como é o caso das Shc/Grb-2/Sos interligando o receptor com a cascata Ras/Raf/MAPK.

3) Observações recentes sugerem que ocorrem comunicações entre a via Jak/STAT e a via MAPK.

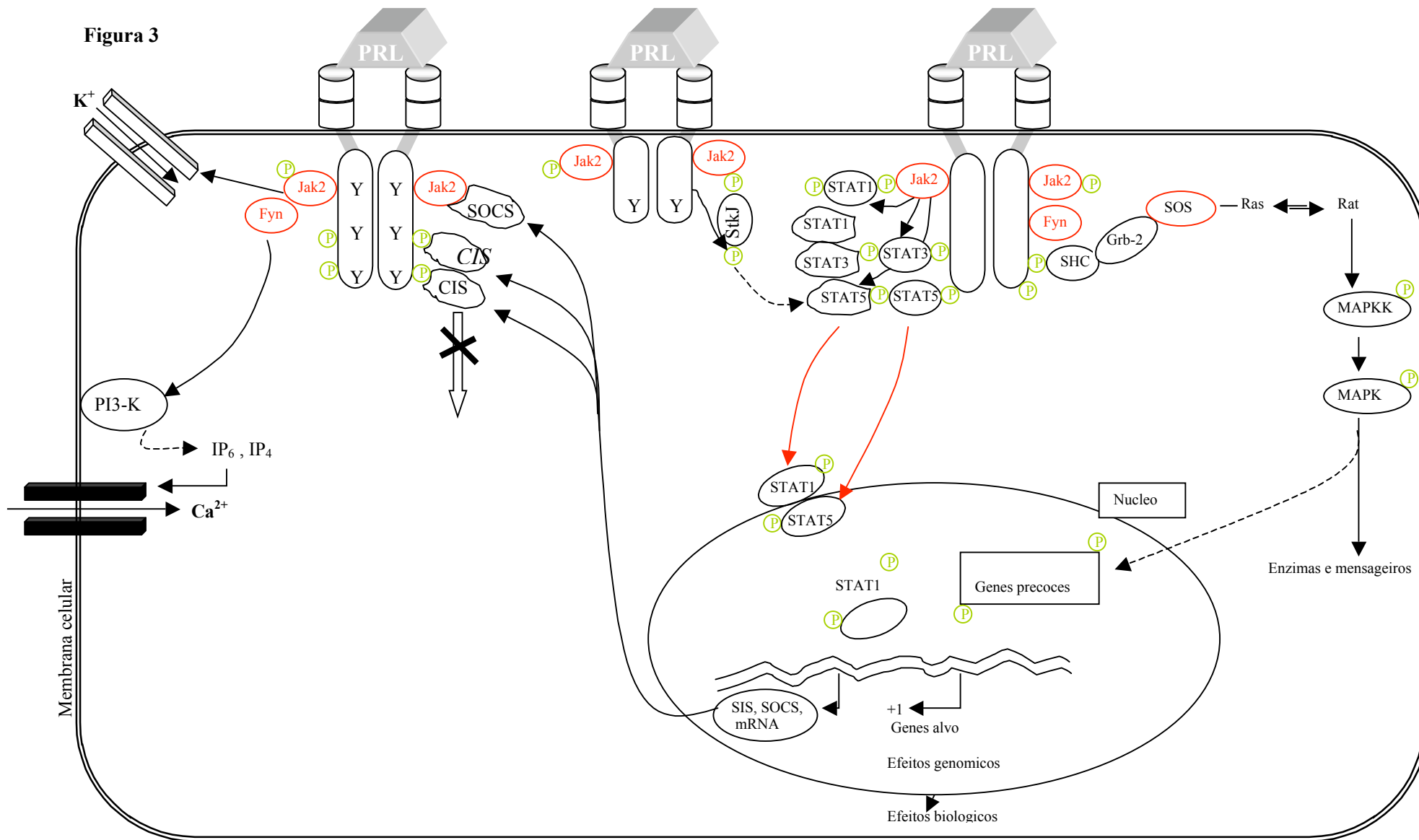
4) Quanto a canais iónicos, a caixa 1 do domínio intracelular do PRLR, através da Jak2 activaria tirosina cinases dependentes de canais sensíveis ao cálcio e potássio.

Por outro lado a região COOH terminal do PRLR estará implicada na produção de mensageiros intracelulares como é o caso do IP4 (inositol 1,3,4,5 tetrafosfato) e do IP6 (inositolhexacisfosfato) que abrem os canais de  $Ca^{2+}$  não franqueados pela voltagem.

5) As cinases Src também são activadas pela PRL, como é o caso da c-Src e da Fyn que estão implicadas na fosfatação da tirosina de cinases do fosfatidilinositol (P13K).

Na regulação da via Jak/STAT, esta via pode ser inibida pelos SOCS (supressors of cytokine signaling) que inibe as Jak cinases, ou inibida pela CIS (cytokine-inducible SH2-containing protein) que compete com os STAT nos locais de “docagem” no PRLR.

Figura 3



Está verificado em ratinhas deficientes em STAT 5<sup>a</sup> que o crescimento mamário lóbulo-alveolar durante a gravidez é inibido, não havendo lactação após o parto visto não ocorrer diferenciação terminal.

Também fêmeas STAT 5b<sup>-/-</sup> não têm desenvolvimento mamário, embora as proteínas do leite sejam expressas mas são insuficientes.

A fertilidade de fêmeas deficientes em STAT 5b está muito comprometida já assim não sucedendo se for o STAT 5<sup>a</sup>.

Também a ausência do gene da ciclina D 1 impede o desenvolvimento da glândula mamária.

No desenvolvimento da glândula mamária os PRLR e as cascatas de transdução de sinal por eles desencadeadas são essenciais não parecendo os receptores para o estradiol e para a progesterona ter um papel tão importante.

No desenvolvimento do tecido mamário além da PRL muitos outros factores de crescimento têm papéis parácrinos em diferentes fases do desenvolvimento mamário, como é o caso do EGF (factor de crescimento epidérmico), da NGR (neuregulina), dos produtos do gene Wnt, do KGF (factor de crescimento ceranócítico), do HGF (factor de crescimento do hepatocito) e do IGF-1 (factor de crescimento insulina-like).

A NGR activa o desenvolvimento alveolar e a actividade secretora, sendo este factor expresso no estroma da mama durante a gravidez.

Durante a puberdade o crescimento dos canais da glândula mamária são estimulados por estrogénios que interactuam os receptores Er<sub>α</sub> ao nível das células epiteliais mamárias e com os receptores Er<sub>β</sub> no estroma.

Referem-se na figura 4 seguinte as acções esquemáticas da PRL sobre as células epiteliais mamárias.

Acções esquemáticas da PRL  
sobre as células epiteliais mamárias

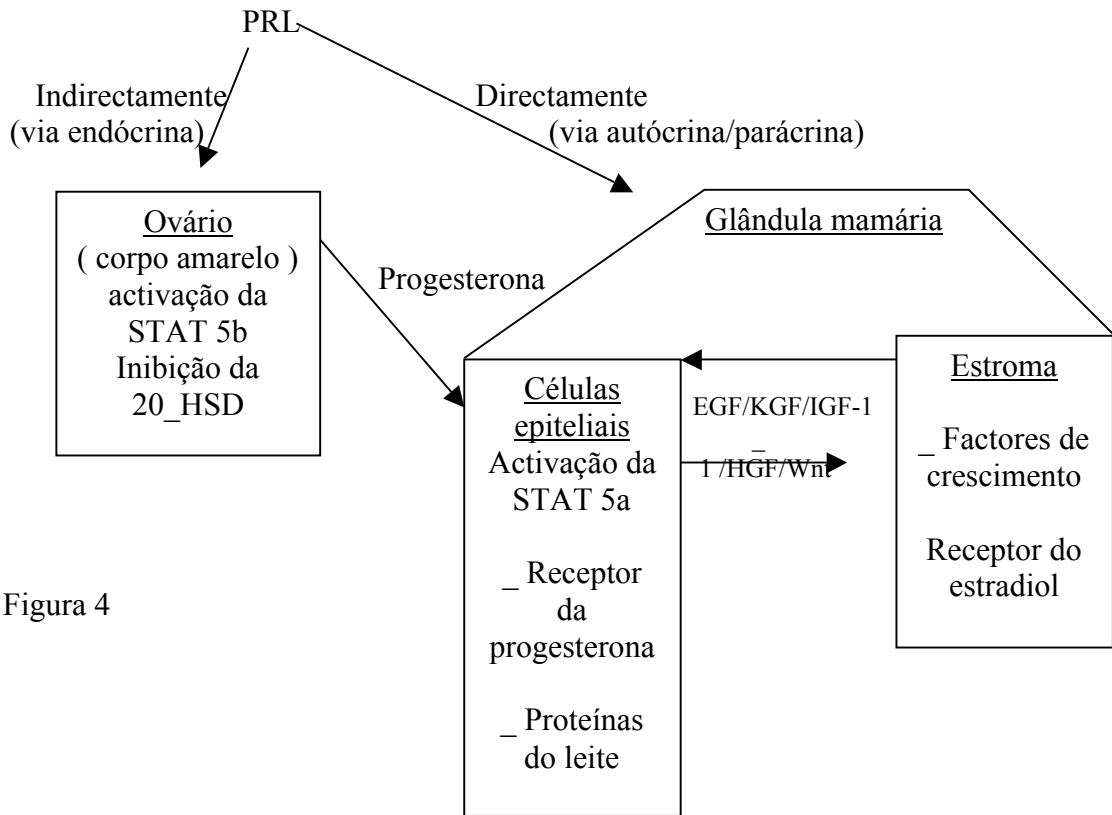


Figura 4

Dada a sua importância voltamos à translocação nuclear do complexo PRL-PRLR (internamento) que parece ocorrer rapidamente e devida a motivos di-leucina existentes dentro dos domínios intracelulares do PRLR. Este internamento parece ocorrer através da via endossoma/corpo multivesicular/lizossomas, sendo parte da PRL degradada.

Parece que a PRL é um factor que durante o ciclo celular favorecerá a entrada na fase S, sendo a ausência de PRL impeditiva da expressão de genes de ciclinas e histonas necessárias para a entrada nesta fase S. Estão assinaladas dentro do núcleo das células 10 a 20% da PRL intracelular total.

O internamento de proteínas do meio extracelular para o citosol e para o núcleo, através da via trans-Golgi/ER é chamado retrotransporte e está verificado que na ausência de estimulação comitogénica promovida pela IL-2 e pelo EGF esta retrotranslocação da PRL para o núcleo não ocorre.

Também a EGF, a NGF, a PDG-F, a insulina, o FGF sofrem esta retrotranslocação para o núcleo.

É possível que este retrotransporte para o núcleo seja ajudado por chaperones, sendo referidos a peptidil-prolil-isomerase (PPI) e a cidofilina (Cyp) como chaperones para a retrotransposição para o núcleo da PRL.

A Cyp existe no sangue e no leite e pode interagir “in vivo” e “in vitro” com a PRL e GH, não tendo a Cyp um papel mitogénico, nem alterando a afinidade da PRL para o PRLR e a activação da Jak2 associada a este último.



No entanto a CYP<sub>17</sub> aumenta notavelmente a retrotranslocação para o núcleo, da PRL, podendo assim desencadear um mecanismo directo de acção somatolactogénica dentro do núcleo.

## **1.7 - Factores de transcrição e cascatas de transdução de sinais da glândula mamária**

### Factor de transcrição PKA

As proteínas cinases do tipo PKA, nas células epiteliais mamárias modulam a velocidade de segregação da caseína. Por outro lado as PKA e as AKAPs (A-kinase anchoring proteins) suas inibidoras, encontram-se associadas com membranas dos elementos das vias secretórias.

Com efeito nas células epiteliais mamárias as proteínas cinases podem interagir no interior das células com outras proteínas, de uma forma específica, e em determinados locais sub-celulares o que origina uma selectividade funcional da sua acção ou seja uma dada enzima pode manifestar diversas especificidades consoante as células, ou de local para local da mesma célula, ou em diferentes períodos de um programa de crescimento e diferenciação celular.

A PKA possui diversos isoenzimas e é uma proteína cinase activada pelo AMP cíclico proporcionalmente à concentração deste, actuando a PKA como transdutor primário de sinais reguladores dentro das células.

Vejamos como os seus efeitos biológicos podem ser modulados nas células epiteliais mamárias.

Para isso recordamos algumas características da PKA.

No seu estado inactivo esta proteína cinase PKA é um heterotetramero com duas subunidades catalíticas (C) e duas subunidades reguladoras (R), existindo múltiplas isozimas de ambas as subunidades C e R, sendo codificadas por diferentes genes tais como RI<sub>1</sub>, RI<sub>2</sub>; RH<sub>1</sub>, RH<sub>2</sub>; C<sub>1</sub>, C<sub>2</sub>.

Também as subunidades C revelam diversas variantes resultantes de “splicing” alternativos.

As subunidades R interagindo com as subunidades C anulam a actividade catalítica da enzima, sendo a interacção com o sinal AMPc em dois locais de cada subunidade R, que leva à dissociação-libertação das subunidades C com actividade catalítica que vão fosforilar uma série de proteínas substrato (desde enzimas de biossíntese, degradação ou interconversões metabólicas, a reguladores de transcrição, canais iónicos, receptores transmembranários, etc.)

A acção de fosfodiesterases hidrolizando o AMPc inactivam as PKA activas.

Como já referimos a activação da PKA é proporcional à concentração de AMPc a qual varia dentro da mesma célula em consonância com os locais subcelulares.

Um processo de regular a actividade das PKA pode ser o de restringir a sua distribuição dentro das células o que pode ser alcançado com a disponibilidade de outras moléculas de proteínas que se possam ligar com as subunidades R. Estas moléculas expressas chamadas AKAPs (A-kinase anchoring proteins) podem interagir com as PKA subunidades R e com outra proteína travando a actividade enzimática.

São conhecidas diversas AKAPs com vários pesos moleculares (129.000, 106.000 e 75.000) e cuja distribuição entre o citosol e as membranas não é afectada pela activação da PKA nas células intactas.

Estas diversas AKAP são expressas pelas células diferentemente consoante a fase do ciclo celular e consoante o estágio de diferenciação.

Mas não são apenas as subunidades R que têm uma distribuição subcelular atípica, sucede o mesmo com as subunidades C das PKA.

Em membranas do tecido mamário de ratinhas a presença de PKA é evidente e a estimulação da adenilato ciclase aumenta nestas membranas a actividade total da PKA e a concentração total de subunidades C, da qual parece haver dois tipos de população. Num tipo as subunidades C estão associadas estequiometricamente com as subunidades R e estão provavelmente ancoradas às membranas ER através de AKAP 129.000 e no outro tipo as subunidades C encontram-se livres.

### Alguns outros factores de transcrição e vias assinalantes na diferenciação mamária e seu funcionamento

A família de factores de transcrição ETS codificados pelos genes ets nos mamíferos está implicada em processos de desenvolvimento e também de tumorigenese.

Uma das características essenciais desta família ETS é possuir um domínio ETS que se liga a locais do DNA contendo um motivo central 5'-GGAA/T-3'.

Os genes alvo destes factores de transcrição ETS são múltiplos.

Nesta família ETS dos mamíferos foram caracterizados dois novos factores de transcrição o ELF 3 e o ELF 5 que se julga estarem restringidos às células epiteliais, desempenhando um papel na diferenciação mamária normal e no seu funcionamento e ainda nas neoplasias mamárias.

Estes ELF 3 e 5 regulam positivamente a transcrição das proteínas acídicas do soro do leite as WAP.

A expressão desta WAP é muito fraca até meio da gravidez, mas depois aumenta milhares de vezes devido a acções combinadas dos glucocorticóides e PRL e também à ligação de factores a um local do factor activador da célula mamária (MAF), que é um elemento conservado ETS-like presente no promotor proximal da WAP.

A ELF 3 presente na mama “virgem” e durante o inicio da gravidez, declina depois, para ressurgir depois após a desmama quando da apoptose da maioria das células epiteliais mamárias.

O ELF 5 parece ter uma expressão muito diferente do ELF 3.

A expressão dos genes específicos da mama é regulada por diferentes mecanismos hormonais ou independentemente das hormonas, nenhum deles sendo totalmente específico do tecido mamário, parecendo que a interacção complexa de factores de transcrição seja responsável pela expressão mamária.

## Vias de sinalização Jak/STAT no epitélio mamário

Apesar de já anteriormente nos termos debruçado sobre este assunto, voltamos a ele numa perspectiva mais concisa.

Os STAT são uma família de factores de transcrição latentes activados por citocinas e factores de crescimento e responsáveis pelo crescimento, diferenciação, sobrevivência e apoptose das células mamárias.

Esta família de factores de transcrição tem sete membros, situados aos pares em três cromossomas diferentes (o STAT 5<sup>a</sup> e o 5b encontram-se num gene).

O STAT 3 é activado na involução da mama caracterizada pela apoptose das células epiteliais, sendo portanto um factor de morte para o tecido mamário diferenciado.

O STAT 5 é um factor de sobrevivência do epitélio mamário.

Os STAT 3 e 5 têm pois papéis antagónicos.

Outro factor de transcrição desta família STAT é o MPBF (for milk protein binding factor) ou MGF (for mammary gland factor).

Os STAT latentes e inactivos do citoplasma são activados após ligação de uma citocina ou de um factor de crescimento com um receptor que induzindo a dimerização destes desencadeiam uma cascata de transdução de sinal com fosfatação de tirosinas nas Jak cinases citosólicas associadas com os receptores que acabam também por ser fosfatados. Criam-se desta forma conformações daquele complexo que permitem a acostagem do STAT e a sua consequente fosfatação em tirosinas o que leva à homo ou heterodimerização dos STAT, sua translocação para o núcleo e aqui interactua no DNA com promotores específicos dos genes alvo sobretudo genes das proteínas do leite.

Claro que há especificidade dos STATs para os receptores consoante o tipo de células, e um único ligando pode activar diferentes STATs.

Os STAT 1 e 3 são expressos a níveis quasi constantes, o 4 é pouco detectavel, o 5<sup>a</sup> é necessário para a mamopoiese e para a eficiente expressão das proteínas do leite e para a lactogénese, sendo o 5b pior conhecido.

O STAT 5<sup>a</sup>/b é o único na sua acção indutora durante a gravidez.

O STAT 2 não tem sido detectado no tecido mamário.

O STAT 6 admite-se que possa ter um papel na glândula mamária.

## **1.8 - Sinergias e antagonismos interactivos de factores de transcrição na regulação da expressão dos genes das proteínas do leite**

### **Organização do nucleossoma e remodelação da cromatina**

### **Integrinas. Suas contribuições para a regulação da expressão dos genes das proteínas do leite**

#### **Sinergias e antagonismos interactivos dos factores de transcrição na regulação da expressão dos genes das proteínas do leite**

Para o gene da  $\kappa$ -caseína é bem conhecido o sinergismo cooperativo entre a prolactina (mediada pelo STAT 5) e os glucocorticóides (mediado pelo receptor glucocorticóide ou GR) para a activação da transcrição deste gene.

Também é bem conhecido para este mesmo gene da  $\kappa$ -caseína um efeito antagónico de interacção em que o factor de necrose tumoral  $\alpha$  (TNF  $\alpha$ ) inibe a sinalização desencadeada pela PRL através do factor nuclear NF-KB.

O factor de transcrição NF-KB pertence a uma família que consiste de cinco proteínas relacionadas quanto à sua estrutura.

Como veremos mais adiante em figura esquemática esta activação ou inibição resulta de interacções que ocorrem a diversos níveis.

No caso da sinergia activadora entre o GR e o STAT 5, ocorre interacção directa entre estas duas proteínas o que leva a uma acção de cooperação para a sua ligação ao DNA. Salienta-se no entanto que os glucocorticóides podem desencadear também efeitos indirectos modelando a expressão de outros genes ditos secundários e responsáveis por exemplo pela fosfatação favorecida das tirosinas do STAT 5 induzida pela PRL.

O factor de transcrição NF-KB activado pode desencadear o seu efeito antagónico sobre a expressão do gene da  $\kappa$ -caseína, ao inibir a fosfatação das tirosinas do STAT 5, admitindo-se noutra hipótese que a inibição de ligação ao DNA da STAT 5 possa ser devida à sobreposição no promotor do gene da  $\kappa$ -caseína, dos locais para interacção com o STAT 5 e o NF-KB.

Verifica-se pois que estes efeitos sinérgicos ou antagónicos entre o NF-KB e o STAT5 resultam de mecanismos cruzados que ao fim e ao cabo influenciam a activação do STAT5 e a capacidade de interacção dos factores de transcrição com o DNA.

Refira-se desde já que na activação da expressão destes genes das proteínas do leite ao nível das células epiteliais mamárias, há também sinais desencadeados pelos componentes da matriz extracelular.

Com efeito nas células epiteliais da glândula mamária a expressão dos genes das proteínas do leite pode ainda ser regulada por numerosos estímulos extracelulares, a maioria dos quais induzem a inibição dessa expressão ao nível do início da transcrição proteica.

No quadro 2 seguinte referem-se sinais extracelulares e factores de transcrição implicados nesta regulação da expressão dos genes das proteínas do leite ao nível do epitélio mamário, bem como o seu efeito sobre a transcrição.

A activação pela insulina do STAT5 apenas está assinalado em tecido não mamário, embora admitamos mais adiante que ela possa também ocorrer neste tecido.

O EGF tem um papel duplo pois activa a expressão do gene da  $\alpha$ -caseína nas células epiteliais mamárias, mas reprime a acção das hormonas lactogénicas nas células terminais diferenciadas.

Também a proteína C/EBP  $\beta$  (CCAAT/enhancer binding protein beta) numa versão truncada, que é um importante factor de transcrição activador de uma série de genes, é expresso nas células epiteliais mamárias, estando nestas células assinalados outros factores de transcrição que medeiam as respostas de hormonas esteróides, da prolactina, da insulina, da interleucina 4 (IL-4) e do TNF $\alpha$ .

Na figura 5 seguinte esquematizam-se no promotor do gene da  $\alpha$ -caseína de roedores, ruminantes e humanos, locais de ligação de factores de transcrição em duas regiões reguladoras definidas uma a região proximal LHRR de resposta à hormona lactogénica e uma região distal favorecedora (enhancer)

## Quadro 2

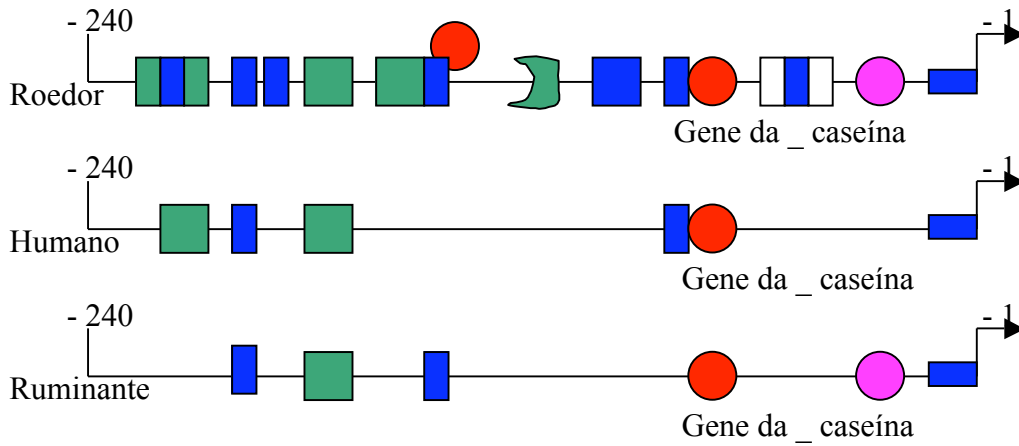
### Sinais extracelulares e factores de transcrição que regulam a transcrição dos genes das proteínas do leite

<u>Sinal extracelular</u>	<u>Factor de transcrição alvo</u>	<u>Efeito sobre a transcrição</u>
Prolactina	STAT5	Activação
Interleucina4	STAT6	Activação
Glucocorticoides	Receptor glucocorticoides	Activação
Progestinas	Receptor da progesterona	Repressão
Insulina	STAT5 (?)	Activação
Factor crescimento epidérmico ( EGF )	STAT5	Repressão/activação
Factor de necrose tumoral $\alpha$ (TNF- $\alpha$ )	NF-KB	Repressão
Transforming growth factor $\beta$ (TGF- $\beta$ )	?	repressão
Matrix extracelular	STAT5, NF-1?	Activação
?	NF-1	Activação/repressão
?	Proteínas de ligação ao enhancer CCAAT	?
?	Proteínas domínio ETS	Activação
?	Factor octamerico	Activação
?	YY-1	Repressão

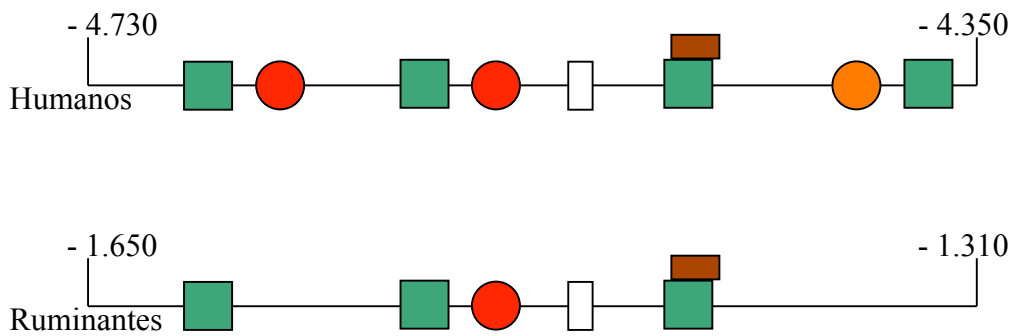
Genes da \_caseína

Figura 5










Próximal LHRR



Distal enhancer



Legenda

- |   |                           |   |  |
|---|---------------------------|---|--|
|  | TATA box                  |  | STAT5 / STAT6                                      |
|  | Oct                       |  | YY1  |
|  | GR half sites palindromas |  | C/EBP (versão truncada da proteína enhancer CCAAT) |
|  | NFKB                      |  | NF1  |
|   |                           |  | Ets  |

Na figura anterior -1 indica o nucleótido que antecede o local de início de transcrição e os -240, -4.350, -4.730, -1.310 e -1.650 indicam os nucleótidos e as posições respectivas que ocupam relativamente ao local de início da transcrição.

Como se verifica na figura na LHRR há múltiplos locais de ligação para as proteínas enhancer de ligação CCAAT e pelo menos também um local de ligação para o STAT5, parecendo que a actividade funcional desta LHRR é menor nos humanos e ruminantes do que nos roedores, o que pode ser devido á falta de mais STATs5.

Nos ruminantes os locais de ligação referidos na figura anterior são idênticos nos ovinos, caprinos e bovinos.

Nestas espécies a activação da expressão dos genes da proteína do leite depende da região enhancer distal onde existem vários locais C/EBP e STAT5.

O YY1 é um factor que participa na repressão da transcrição do gene da  $\alpha$ -caseína ao ligar-se ao LHRR proximal, embora pareça que dada a sua proximidade dos locais de ligação, o STAT5 iniba a ligação de YY1.

A sinergia entre o receptor glucocorticóide (GR) e a STAT5 parece resultar da interacção proteína-proteína entre estes dois factores, que não depende da ligação da GR ao DNA do promotor, promovendo a GR activada a fosforilação da tirosina no STAT5.

Na figura 6 seguinte do lado esquerdo mostra-se a interacção da GR com a STAT5 a dois níveis.

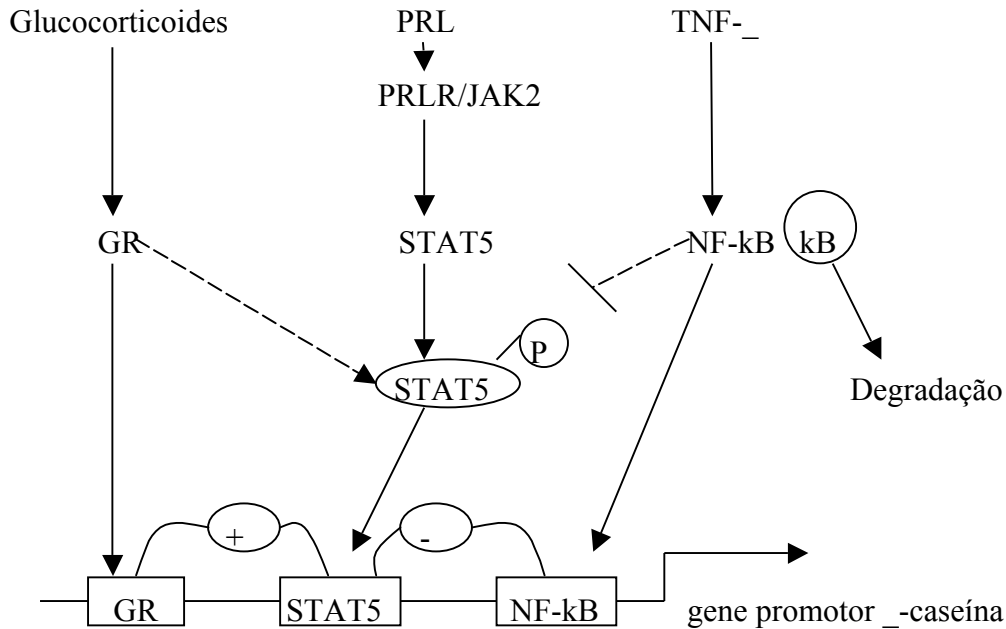
Na figura 6 indica-se que além das interacções directas da GR e STAT5 ao nível do gene promotor da  $\alpha$ -caseína, também a GR promove a fosfatação da tirosina da STAT5.

O factor  $\alpha$  de necrose tumoral (TNF- $\alpha$ ) é um regulador multifuncional do desenvolvimento da glândula mamária tendo uma acção inibidora sobre a expressão do genes da proteína do leite sendo um activador da NF-kB.

Provavelmente mais do que um mecanismo está implicado no antagonismo entre o STAT5 e o NF-kB, tais como na competição entre o STAT5 e o NF-kB para se ligarem ao gene promotor da  $\alpha$ -caseína e pela inibição da fosfatação da tirosina no STAT5.

Figura 6

Vias de transdução de sinais nas células epiteliais mamárias e mecanismos inter cruzados entre glucocorticóides, prolactina (PRL) e factor  $\alpha$  da necrose de tumores (TNF- $\alpha$ )



As interações dos factores de transcrição implicados na regulação da transcrição do gene da proteína do leite parecem ser operativas a dois níveis, ao nível da região próxima LHRR onde a organização desta estrutura e dos respectivos locais de ligação dos factores é determinante da via de transdução de sinal operante, podendo-se assim originar diferenças na força e estado de dependência das interações entre os factores de transcrição e os genes e por outro lado podem ocorrer interações entre vários factores sem ser necessário a sua ligação prévia ao DNA.

Dada a importância da regulação da expressão dos genes das proteínas do leite não queremos deixar de abordar agora novos desenvolvimentos sobre este assunto, em que surgem algumas discordâncias em relação a afirmações feitas anteriormente.

Os genes que codificam as caseínas e as proteínas do soro do leite são reguladas por hormonas peptídicas e esterólicas, sobretudo as hormonas lactogénicas, prolactina, insulina, e hidrocortisona e ainda nessa regulação intervêm as interações célula a célula e célula substrato matriz extracelular.

As proteínas do soro do leite, a proteína ácida do soro ou WAP, a  $\alpha$ -lactoglobulina e a  $\beta$ -lactalbumina são codificadas por um gene de cópia única relativamente pequeno.

As caseínas são codificadas por um “cluster” ou grupo de genes de cópia única.

Nos bovinos no cromossoma 6 e numa extensão de 250 kb agrupam-se os genes da  $\alpha$ -s-1-,  $\beta$ -,  $\alpha$ -s-2 e k-caseínas.



Os genes  $\alpha$ -s-1 e  $\alpha$ -caseínas estão ligados próximo um do outro e arranjados com uma orientação 5'---3' e 3'---5' respectivamente.

Os três genes codificadores das proteínas sensíveis ao cálcio ou sejam a  $\alpha$ -s-1,  $\alpha$  e  $\alpha$ -s-2 têm motivos reguladores comuns nas regiões flanqueadoras proximal e distal 5'. Os genes da  $\alpha$ -caseína e da WAP, pelo menos na rata, necessitam para a regulação hormonal, de um elemento complexo de DNA chamado elemento de resposta compósito ou CoRE.

Os elementos de resposta compósito ou CoRE contêm pois múltiplos locais para interação com vários factores de transcrição mediando a regulação por hormonas da expressão dos genes das proteínas do leite.

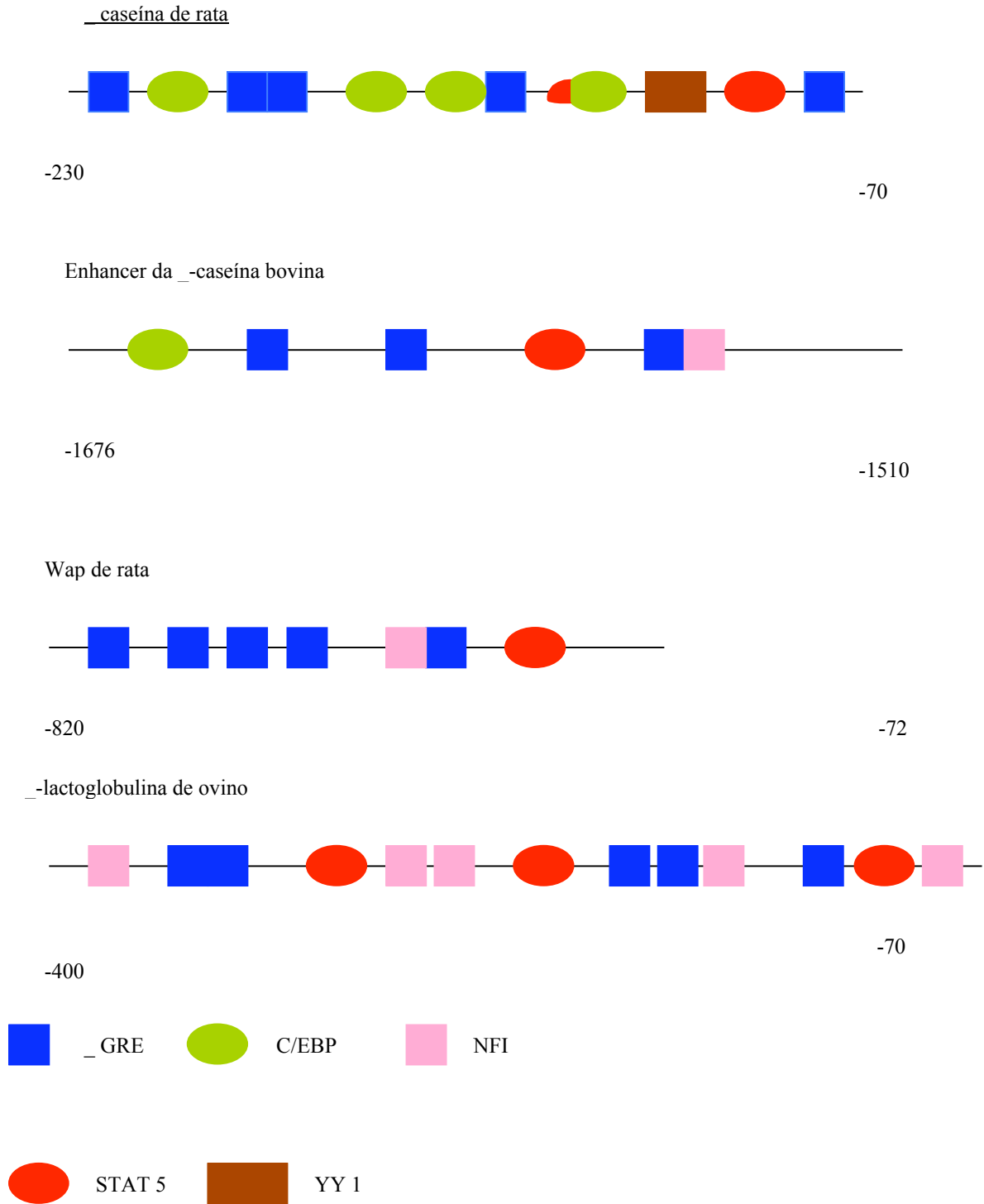
Os CoRE são pois grupos de locais, que interactuam com factores de transcrição, contendo elementos reguladores positivos e negativos, que fazem parte das vias de transdução de sinais através de interacções proteínas/DNA e proteínas/proteínas.

Os níveis de activação da transcrição promovida pelos CoRE são muito superiores aqueles provenientes de um só factor de transcrição actuando sozinho.

Os genes das proteínas do leite são pois controlados na sua expressão por regiões promotoras contendo CoREs.

Figura 7

Genes promotores de proteínas do leite



Referiremos seguidamente algumas características de elementos de resposta compósita já anteriormente abordados e que regulam a expressão das proteínas do leite.

### Elementos de resposta compósita (Figura 7)

O STAT5 é um factor de transcrição que responde à PRL na glândula mamária, (inicialmente conhecido como MGF-mammary gland factor - ou MPBF-milk protein binding factor).

O STAT5 é expresso em variadíssimos tecidos e na glândula mamária ele não está limitado apenas à lactação, sendo activado pela PRL, diversas hormonas e factores de crescimento e citocinas.

O STAT5a e STAT5b têm cerca de 95 % de ácidos aminados comuns, e ambos apresentam diversas isoformas resultantes de diferentes “splicing”. Para a expressão dos genes da -caseína, e -s-1-caseína é necessária a interacção com STAT5 resultante da indução hormonal.

Para a expressão da WAP são necessários um ou mais locais para a STAT5, tal como para a -lactoglobulina, pelo menos em ratas transgenicas.

Vejamos mais algumas proteínas que intervêm em respostas compósitas.

### C/EBPs

São factores de transcrição contendo motivos leucina zipper básicos (bZIP) na sua extremidade carbóxilica, motivos estes que medeiam dimerização e interacção com o DNA:

Desta família, o C/ERP - e - são muito importantes no desenvolvimento da glândula mamária e na expressão das proteínas do leite.

As várias C/EBPs, a -, - e - parecem exibir em diversos tecidos, uma expressão em cascata, a - e - exprimindo-se primeiro e relacionando-se com a proliferação celular, e a - exprimindo-se depois.

NF1 (nuclear factor) são proteínas diversas possuindo todas no entanto uma porção N-terminal que interactua com o DNA e um domínio para dimerização. A extremidade carbóxilica é muito variável consoante a proteína NF1.

Há dezenas de isoformas dentro desta família de proteínas.

As isoformas NF1 específicas do tecido mamário são componentes do CoRE que regula a expressão das WAP, e de outras proteínas do soro do leite

Durante a lactação existem duas formas destes NF1 a 46 e 68 kd enquanto durante a involução mamária apenas surge uma isoforma 74 kd.

### GR (receptor glucocorticoide )

A maioria dos genes que exprimem proteínas do leite, não contém elementos de resposta aos glucocorticóides (GRES).

No entanto o promotor da proteína WAP do leite contém um conjunto de meios-palíndromos (half-palindromic, GRE) GRÉS nas duas regiões a dos nucleótidos reguladores distais (nucleótidos -853 a -729) e a dos nucleótidos reguladores proximais (nucleótidos -160 a -70).

Também o promotor proximal da  $\kappa$ -caseína contém vários half-GREs.

### Yin Yang (YY)-1 e região da caixa do leite

O promotor da  $\kappa$ -caseína contém uma região reguladora negativa entre os nucleótidos -150 e -110, reprimindo este elemento bipartido a transcrição na ausência de hormonas lactogénicas.

Os promotores proximais dos genes da  $\alpha$  e da  $\beta$ -caseína, na rato, contém também locais de ligação putativos YY-1.

A YY1 é uma proteína multifuncional que pode activar ou reprimir a transcrição e que é expressa ubiquamente.

Nas células epiteliais mamárias a concentração da proteína YY1 não é afectada pelas hormonas lactogénicas.

Os diversos factores de transcrição actuam em cooperação uns com os outros, sendo as interacções entre eles, (proteína com proteína) possivelmente tão importantes como a interacção de cada factor de transcrição com o seu elemento de resposta.

A activação da transcrição a partir de um CoRE é geralmente maior do que aquela que resulta da soma de cada uma das suas partes componentes.

### Interacções de proteínas com proteínas que influenciam a estrutura da cromatina do promotor da $\kappa$ -caseína consequentemente a sua activação ou repressão (Figuras 8 e 9)

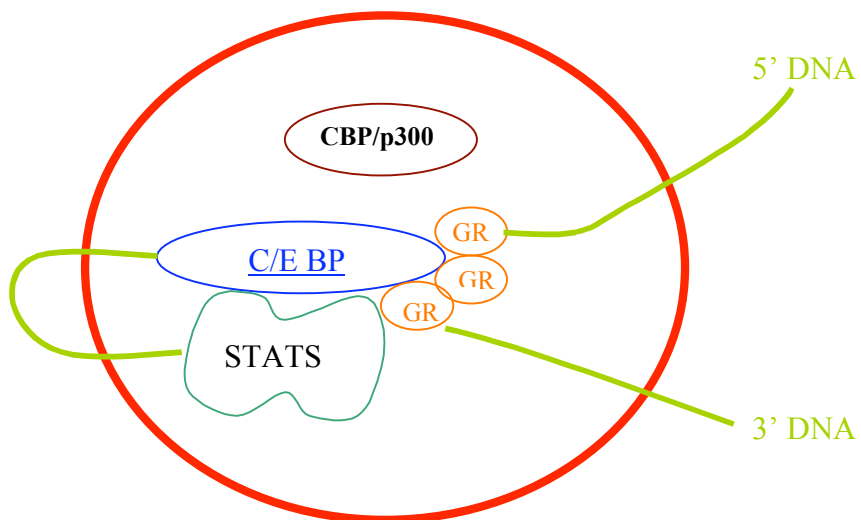
#### Estado activado para transcrição

Interacções entre a STAT 5, a C/EBP  $\beta$  (CCAAT/enhancer binding protein  $\beta$ ) e GR (receptor glucocorticoid) podem gerar um complexo de activação estável.

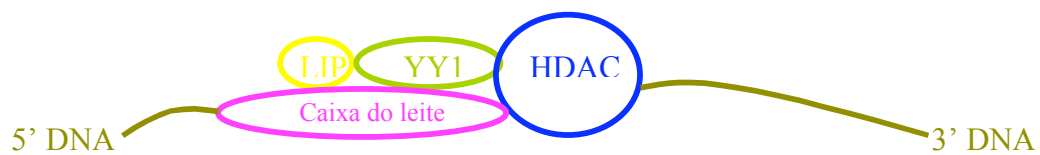
Com efeito estas três proteínas recrutam a p300/CBP para o promotor da  $\kappa$ -caseína, originando a acetilação das histonas e correlativa activação da transcrição.

Figura 8 e 9

Estado activado para transcrição



Estado reprimido para transcrição



### Estado reprimido para transcrição (Figura 9)

No estado reprimido, a proteína LIP (liver-enriched inhibitory protein) pode travar a proteína YY1 (Yin Yang) em ordem a enfraquecer a ligação desta proteína com o promotor da  $\kappa$ -caseína, podendo então “recrutar” a enzima desacetilase das histonas (HDACs), o que corresponde a desacetilações das histonas e consequente repressão da transcrição.

Como temos vindo a referir a expressão dos genes das proteínas do leite é regulada em sinergia pelos glucocorticóides (hormonas lactogénicas essenciais) e PRL e insulina.

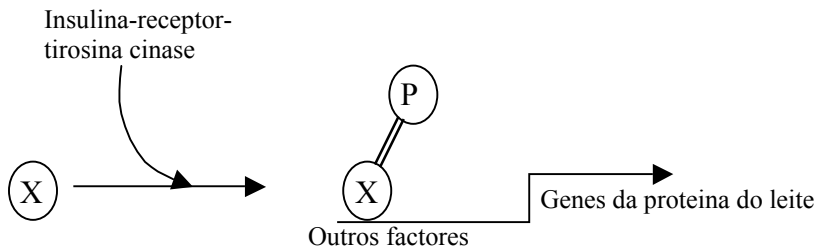
A insulina é indispensável para a lactação estimulando a síntese de fosfoproteínas mamárias, parecendo que o efeito da insulina sobre a caseína é mediado ao nível da transcrição dos seus genes.

Também a IGF-1 e a IGF-II estimulam a expressão dos genes das proteínas do leite, embora com uma potência cerca de dez vezes menor que a insulina.

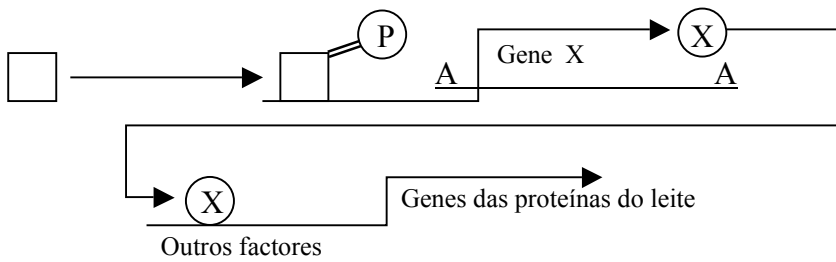
A activação da transcrição das proteínas do leite induzida pela insulina parece envolver diversos mecanismos (quatro) que se referem seguidamente (Figura 10).

Figura 10

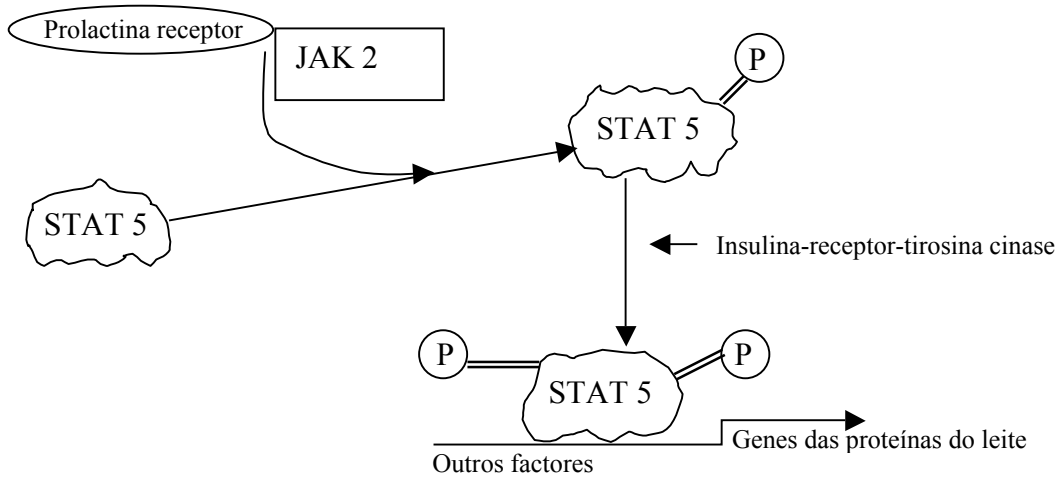
- 1 - Activação do factor de transcrição X pela insulina a qual se liga e activa directamente os genes das proteínas do leite



- 2 - Activação indirecta pela insulina da transcrição de um gene A que codifica o factor de transcrição X



- 3 - Modulação pela insulina da activação dependente da prolactina e via JAK/STAT



- 4 - Expressão de um gene do factor de transcrição X dependente da insulina, factor X que é directamente activado por nutrientes como a glucose e uma vez activado pode activar directamente a expressão das proteínas do leite

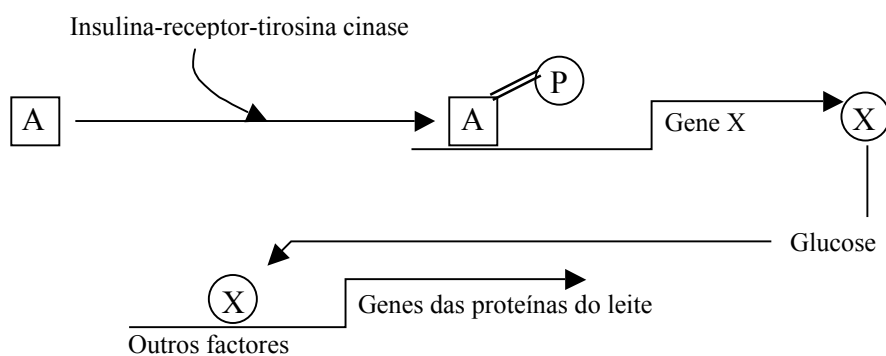


FIGURA 10



## Organização do nucleossoma e remodelação da cromatina na regulação

Conforme se pode verificar na figura 11 seguinte a cromatina transcriptionalmente incompetente limita a acessibilidade ao DNA pelos factores reguladores da transcrição.

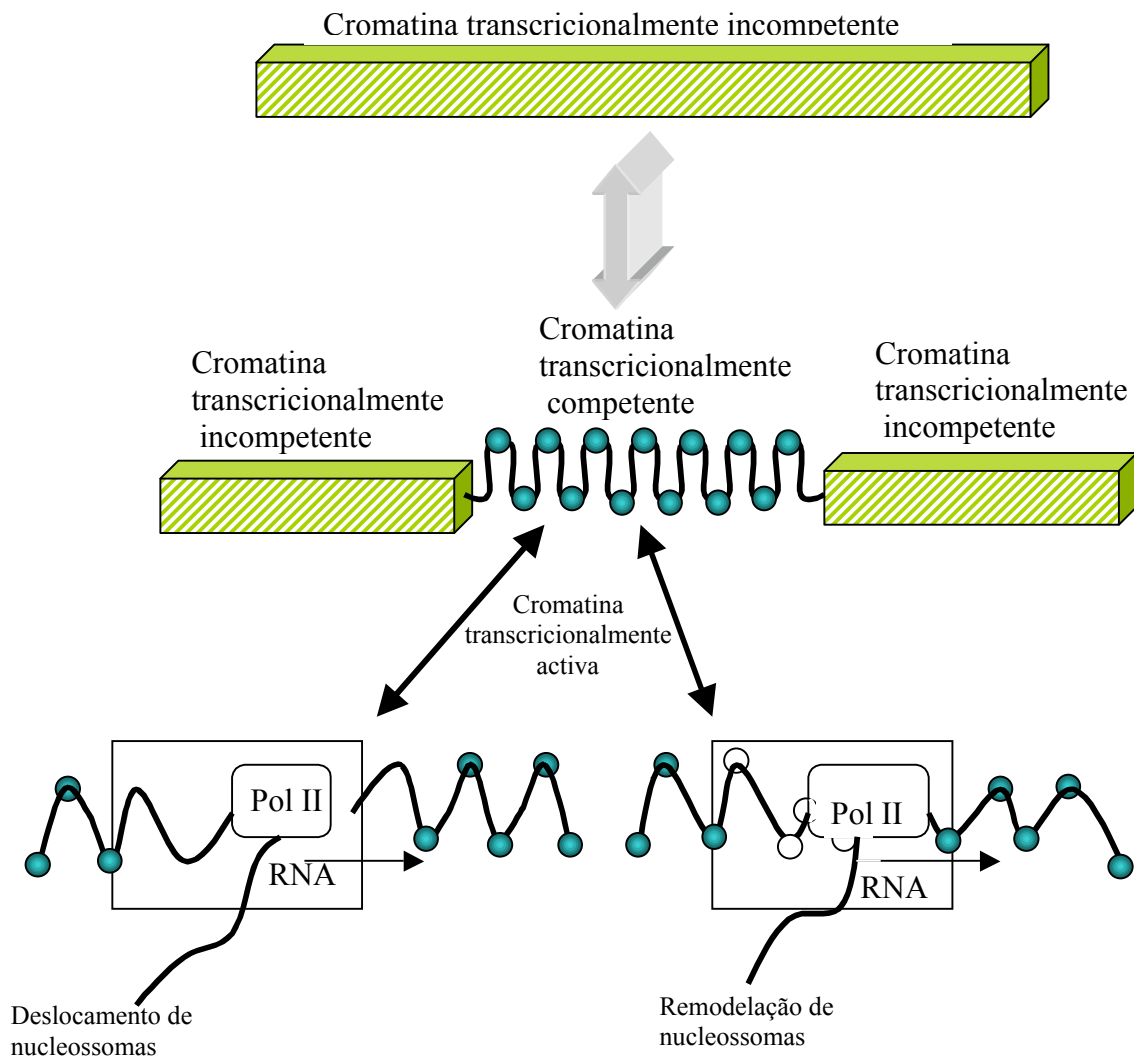
A remodelação destas regiões reguladoras são portanto essenciais para o controlo da expressão dos genes que nelas estão contidos.

É conhecido que a estrutura da cromatina é dinâmica e as suas modificações estão correlacionadas com a repressão ou com a activação da transcrição, tendo essa estrutura cromatínica um papel fundamental não só na transcrição mas também na replicação e reparação do DNA.

Como se compreende a diferenciação celular corresponderá a uma remodelação de grandes porções da cromatina em ordem a tornar competente grandes porções do genoma nela contido.

Ocorre ainda que a transcrição de genes localizados nestas regiões implicarão outras alterações adicionais da estrutura da cromatina localizada dentro de pequenas regiões reguladoras.

Figura 11



Na vida de cada gene, a regulação precisa da sua expressão, ora “ligando” (exprimindo), ora “desligando” (não exprimindo) é central, e pode-se dizer que tem um perfil específico para cada gene, no que se refere ao local e ao momento da vida biológica, sendo hoje conhecido como muitos factores de transcrição regulam esta expressão dos genes.

No que se refere à expressão dos genes dos tecidos mamários que representam um tecido diferenciado, ela é bastante complexa na medida em que são expressos uma série de genes em resposta a sinais hormonais, celulares e da matriz extracelular.

A mama é um órgão especial pois a maioria do seu desenvolvimento ocorre no estado adulto dos animais, e é ainda sujeito a uma série de ciclos sucessivos de desenvolvimento e de regressão ou involução.

A  $\alpha$ -lactoglobulina ( $\alpha$  Lg) é a principal proteína do soro do leite da maioria dos mamíferos, constituindo mesmo um marcador específico do tecido mamário.

Ora é conhecido que a organização dos nucleossomas do gene da  $\alpha$  Lg regulam temporariamente a expressão deste gene.

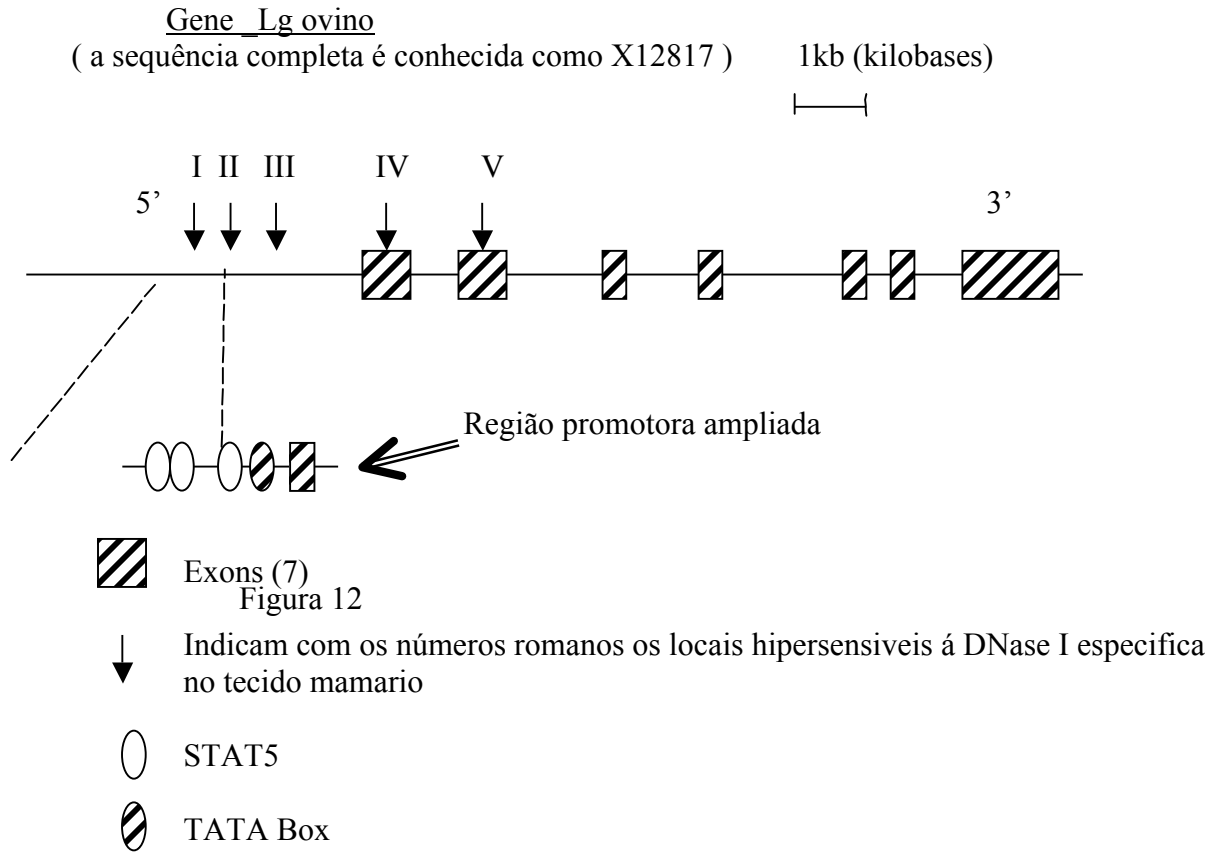
A expressão do gene da  $\alpha$  Lg é regulada por uma complexa interacção de hormonas e factores de crescimento em associação com interacções célula a célula e célula - matriz extracelular.

Assim sabe-se que a prolactina que é o principal estimulador lactogénico activa a expressão da  $\alpha$  Lg através do factor de transcrição STAT5 numa cascata de transdução de sinal.

Pela utilização da desoxirribonuclease I (DNase I) verificou-se que dentro do gene  $\alpha$  Lg ovino existem elementos que regulam temporariamente a expressão deste gene.

Dentro do gene promotor da  $\alpha$  Lg ovina existem três locais para interactuarem STAT5 tal como se esquematiza na figura 12 seguinte.

Figura 12



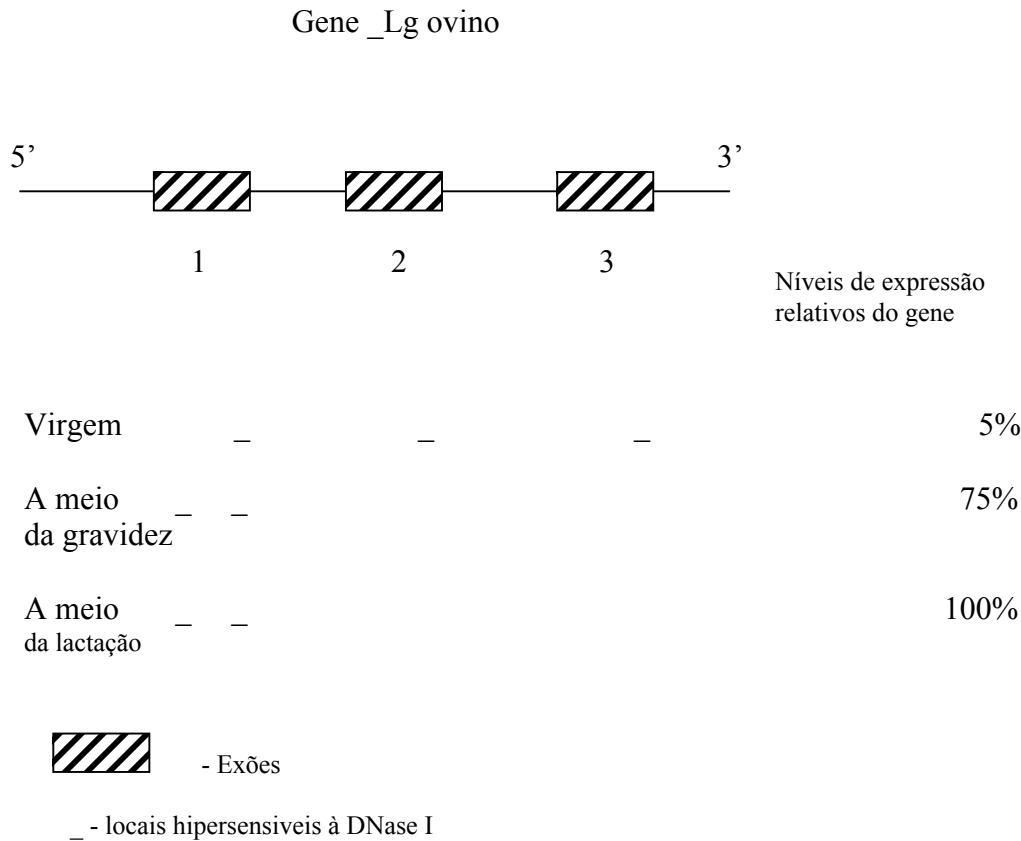
Esta região da figura anterior contém vários locais de ligação para os factores de transcrição NF-1 e STAT5 sendo este último activado pela cascata de transdução de sinal induzida pela prolactina a principal hormona lactogénica.

A activação da expressão do gene \_Lg desencadeia alterações na estrutura de cromatina que envolve a região promotora.

Por exemplo: no fígado onde este gene não é expresso não há locais hipersensíveis para a DNase I, mas na glândula mamária da ovelha em lactação existem locais distintos para este efeito, havendo ainda outros locais hipersensíveis adicionais dentro dos dois primeiros intrões para a transcrição completa da \_Lg. Só o local hipersensível localizado no promotor e chamado HS III é essencial para a expressão nas células mamárias.

Na figura 13 seguinte mostram-se a formação de locais hipersensíveis à DNase I no gene da \_Lg ovina durante a gravidez e a lactação.

Figura 13



Como se verifica no animal virgem a expressão do gene Lg é apenas detectável (5%), mas durante a gravidez a expressão do gene aumenta gradualmente até ao 110º dia da gravidez, continuando depois a aumentar através do parto até atingir o máximo durante a lactação.

A hipersensibilidade do local HS III á DNase I é paralela a esta última evolução da expressão do gene Lg [ hipersensibilidade mínima (mínimo corte) antes do meio da gravidez e cortes substanciais depois, durante a lactação .

Também a evolução dos STAT5a e b acompanham estas evoluções, havendo um marcado aumento da STAT5a em relação ao STAT5b durante a gravidez, parecendo que o STAT5a é o principal mediador do sinal lactogénico

Integrinas. Suas contribuições para a regulação da expressão dos genes das proteínas do leite

“In vivo” no controlo da proliferação, apoptose, diferenciação e manutenção da polaridade baso-apical do epitélio mamário o funcionamento normal das 1 integrinas é fundamental, uma vez que as interacções células mamárias-matriz extracelular regulam o crescimento e apoptose das células, podendo a ECM regular o fenótipo do epitélio mamário.

Está assinalado que a adesão das células epiteliais mamárias às proteínas laminina da membrana “basement” regulam a expressão do gene da  $\alpha$ -caseína, enquanto o factor de transcrição STAT5 também é regulado pelas interacções ECM-células.

Como referiremos as integrinas são os principais receptores das células para a ECM, sendo heterodímeros transmembranários compostos por subunidades  $\alpha$  e  $\beta$  não ligados covalentemente e possuindo um grande domínio extracelular que interactua com diversos ligandos como por exemplo proteínas da ECM ou a outros receptores da superfície das células, enquanto o domínio intracitoplásmico dessas integrinas interactua com complexos do citoesqueleto podendo assim fazer transdução de sinais.

Referem-se seguidamente (Figura 14) as integrinas expressas na bicamada da glândula mamaria

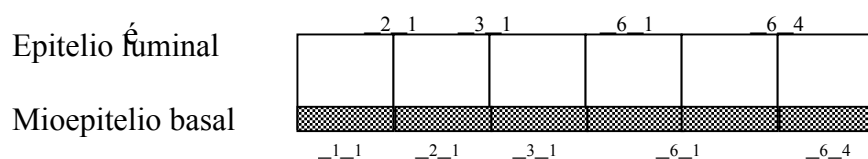


Figura 14

Vários dados experimentais sugerem que a expressão dos genes da  $\alpha$ -caseína nas células epiteliais mamárias necessitam da interacção com a laminina mediada pela  $\alpha_1$ - integrina e que a actividade de ligação ao DNA do STAT5 (regulador da expressão do gene da proteína do leite depende da adesão á membrana basement.

A polaridade das células epiteliais luminiais com as suas diferentes propriedades é estabelecida através da interacção das células com a sua membrana basal, requerendo o funcionamento de  $\alpha_1$ - integrinas normais.

A matriz extracelular (ECM) desempenha diversas funções muito importantes durante todas as fases de desenvolvimento da glândula mamária, inclusivé durante a lactação ao regular a expressão dos genes das proteínas do leite, estimulando a expressão da maioria destes genes, estando isto bem caracterizado para a  $\alpha$ -caseína e  $\alpha$ -lactoglobulina.

A laminina, proteína da matriz extracelular induz um sinal que na presença de hormonas lactogénicas promove a expressão dos genes das proteínas do leite.

Está identificado um elemento de 160-bp (entre os nucleótidos 1677 e 1517 no gene da  $\alpha$ -caseína bovina) ou seja o denominado BCE ( $\alpha$ -casein element 1) que é específico da mama sendo um enhancer dependente da matriz extracelular. Neste BCE existem dois locais para interacção com a C/EBPs e com o STAT5.

Também os primeiros 406 bp do promotor da  $\alpha$ -lactoglobulina respondem à matriz extracelular possuindo locais para interactuar com o STAT5 e com o NF1.

Também se conhece que na glândula mamária a fosfatação do receptor da prolactina, a fosfatação/ativação da Jak2, a fosfatação da STAT5 e a interacção com o DNA originada pela prolactina dependem todos da matriz extracelular, regulando assim a expressão dos genes das proteínas do leite.

A matriz extracelular modula ainda a estrutura da cromatina como vimos anteriormente.

As integrinas  $\alpha$  são uma grande família de receptores da superfície das células que medeiam interacções célula a célula e a adesão da célula à matriz.

Algumas integrinas reconhecem a sequência R-G-D no seu ligando proteico extracelular na matriz.

As integrinas são como dissemos um dímero de cadeias  $\alpha$  e cadeias  $\beta$ .

Cada subunidade tem um largo domínio extracelular N-terminal e um outro domínio intracitoplásmico e ligando os dois, um domínio transmembranário.

Alguns receptores têm uma cadeia  $\beta$  comum revelando diferenças na cadeia  $\alpha$ .

A integrina  $\alpha_1$  associa-se com a  $\beta_1$  para formar um receptor para a laminina que é uma proteína da membrana “basement”.

Todas as cadeias  $\alpha$  das integrinas possuem quatro repetições de uma região com 40 ácidos aminados na extremidade C-terminal do domínio extracelular. Cada uma destas repetições tem oito cisteínas.

Nas integrinas  $\alpha$  as subunidades alfa são por vezes cindidas após a translação dando cadeias pesadas e leves ligadas por uma ponte dissulfureto.

A cadeia  $\alpha_1$  (VLA-1) (CD49a) com a cadeia  $\beta_1$ , actuam como receptores para a laminina e colagénio.

As cadeias  $\alpha_2$  (VLA-2) (CD49b) com a cadeia  $\beta_1$  actuam como receptores e ligam colagénio.

As cadeias  $\alpha_3$  (VLA-6) com as cadeias  $\beta_1$ , formam um receptor plaqueta laminina.

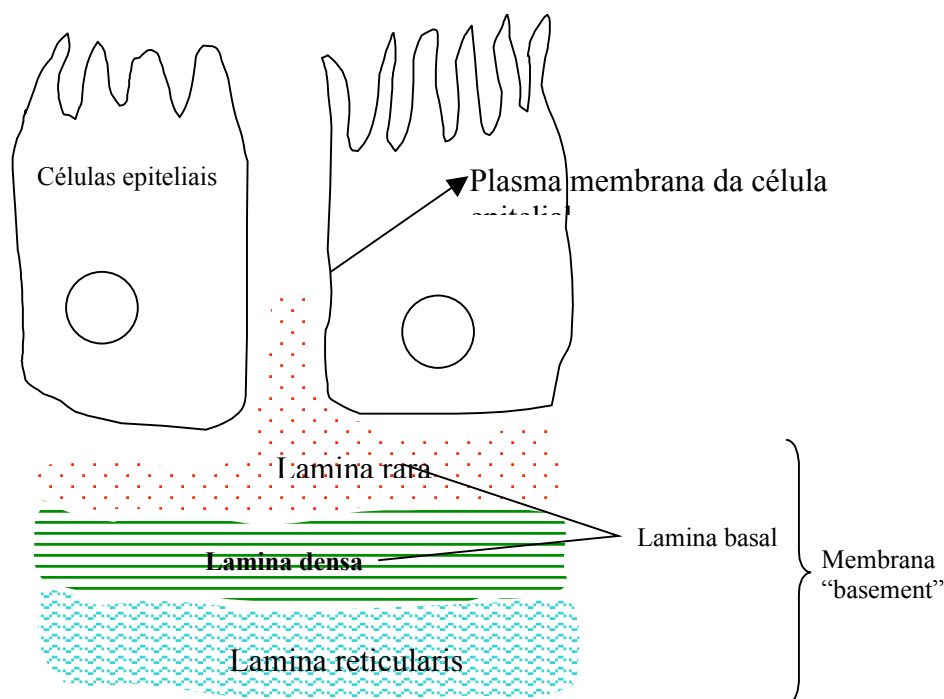
A laminina é uma glicoproteína complexa, consistindo de três cadeias polipeptídicas distintas ( $\alpha$ ,  $\beta$  e  $\gamma$ ) que se ligam entre si através de pontes dissulfureto.

As lamininas têm uma localização extracelular, encontrando-se em membranas “basement”(o principal componente) e formam estruturas “coiled coil”.

A lamina basal é pois uma matriz extracelular especializada, composta sobretudo de colagénio do tipo IV, proteoglicanas e laminina (Figura 15).

Figura 15

**Desenho esquemático da lamina basal subjacente a uma célula epitelial**



As lamininas cadeias  $\alpha 1$  ligam-se às células através de um receptor de alta afinidade. As lamininas medeiam a ligação, migração e organização das células em tecidos.

Interacções

Os efeitos e interacções entre as células e a matriz extracelular (ECM) são claramente evidenciadas ao nível da glândula mamária, inclusive a forma como a remodelação da ECM condiciona o funcionamento da glândula mamária.

No tecido epitelial mamário murino a expressão dos genes específicos deste tecido, está interrelacionado com a estrutura e composição da ECM, participando as integrinas na expressão das proteínas do leite que inclusivamente são reguladas “in vivo” pela remodelação da ECM.

Esta remodelação da ECM conforme veremos mais adiante é consequência de modificações na expressão de proteinases da ECM e dos seus inibidores.

A regulação da expressão das proteínas do leite é um tanto variável consoante as proteínas em causa. No caso da  $\kappa$ -caseína a expressão do seu gene é regulada pela interacção célula-ECM parecendo ocorrer ao nível da transcrição.

Portanto os contactos células epiteliais mamárias com a ECM regulam a expressão dos respectivos componentes. Mais, o fenótipo secretório das células epiteliais mamárias durante a lactação e a involução mamária é condicionado pela remodelação da ECM.

Como já se referiu as células epiteliais mamárias, durante a gravidez proliferam e depois após o parto desenvolvem-se os alvéolos secretores que se mantêm durante a lactação, constituídos por células epiteliais polarizadas unidas por “tight junctions” e contactando com uma membrana “basement” uniforme. Esta estrutura alveolar secretora após a desmama sofre uma involução tal como a expressão das proteínas do leite que é inibida, assumindo a glândula mamária uma estrutura similar à dos animais virgens.

Durante todo este período de lactação, há um perfil próprio de proteinases e respectivos inibidores que são secretados, tal como o perfil das proteínas do leite também varia durante este período.

Assim por exemplo mRNA da  $\kappa$ -caseína é expresso generosamente durante a fase de lactação, ao contrário do que sucede com os mRNA das proteinases e seus inibidores, sucedendo o inverso durante o fim da lactação ou seja durante a involução mamária.

Durante esta involução a chamada membrana “basement” (que engloba a lamina basal e a lamina reticularis) é degradada pelas proteinases e as células epiteliais mamárias que deixaram de exprimir os genes das proteínas do leite acabam por morrer.

A degradação da ECM envolve a cooperação entre as células epiteliais e as células mesênquimatosas, e quando a involução mamária começa a primeira alteração é a indução dos inibidores das proteinases (TIMPs), mas depois a expressão destes declina substancialmente enquanto aumenta a expressão das proteinases (MMPs).

As MMP ou matrizmetaloproteinases, ou matrixinas, são enzimas extracelulares (E.C. 3.4.24.-) dependentes do zinco, segregadas numa forma inactiva (zimogénio) que difere da forma madura pela presença de um própeptido N-terminal.

Há uma região das matrixinas envolvidas na sua autoinibição (a cisteína “switch”).

A maioria das metaloproteinases dependentes do zinco, têm uma estrutura primária similar o que leva também a designar esta superfamília como metzincinas.

Os TIMP ou inibidores tissulares das metaloproteinases, são uma família de proteínas que formam complexos com as MMP (como por exemplo colagenases) inactivando-as irreversivelmente.

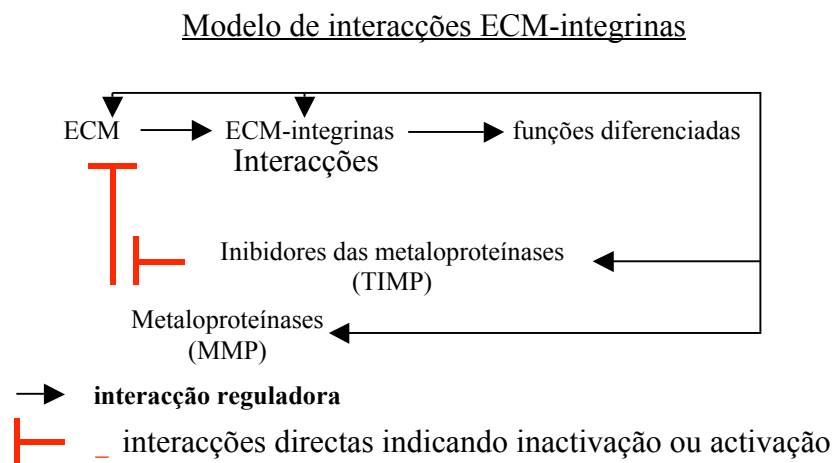
É o equilíbrio entre MMPs e TIMPs que modula o estado de diferenciação das células mamárias, sendo os contactos ECM-integrinas regulado pelo controlo da renovação (turnover) da membrana “basement”.

Os sinais transduzidos pelas integrinas parecem ser necessários para o controlo da expressão dos genes específicos do tecido mamário e para determinar a viabilidade celular.



Na figura 16 seguinte esquematizam-se as interacções entre a ECM e as integrinas, mostrando-se as interacções directas entre as proteínas para efeitos de activação e inactivação (por exemplo MMP-----ECM ou TIMP-----MMP ou ECM---ECM-Integrinas) e as diversas interacções reguladoras.

Figura 16



O equilíbrio entre as metaloproteínases e os inibidores das proteínases determina a estabilidade dos componentes da matriz extracelular e assim afecta a natureza dos ligandos que têm acesso às integrinas e a outros receptores da superfície celular.

A sinalização mediada pelas integrinas regula o fenótipo diferenciado das células e modula a expressão ou actividade das metaloproteínases, dos seus inibidores, das integrinas e dos componentes do matriz extracelular.

### 1.9 - Caspases e plasmina na biologia da glândula mamária involutiva

Na glândula mamária, após a desmama ou seca, as células epiteliais alveolares morrem por apoptose (sendo depois removidas por fagocitose), um processo de morte programada das células, desencadeada por vias de sinalização que parecem envolver a activação de uma família de aspartato específico cisteína proteases, ou sejam as caspases com localização intraepitelial.

Esta apoptose caracteriza-se morfológicamente pela vesiculação da membrana plasmática das células, condensação da cromatina e produção de corpos apoptóticos ou sejam restos de fragmentos celulares envolvidos por membranas.

A apoptose é em muitos casos carente de energia, havendo ainda necessidade da expressão de certos genes, assim como ocorre uma fragmentação do DNA genómico sem ser ao acaso.

As caspases, ou sejam os principais efectores da apoptose possuem duas subunidades o que leva à formação de um complexo enzimático.

A involução mamária post-lactação é pois acompanhada pela apoptose dessas células epiteliais secretoras, contribuindo o leite acumulado transitoriamente na mama para esse efeito.

A falta de estímulo resultante da não sucção do leite origina uma diminuição das hormonas lactogénicas e que favorece a involução mamária.

As caspases podem ser activadas por diversas vias de transdução de sinais, podendo partir por exemplo de receptores de morte como o Fas/APO-1 que no entanto não é essencial para a involução da glândula mamária.

Durante esta involução mamária as taxas de RNA mensageiro das caspases 1, 3 e 9 aumentam o que significa que aumentam estas e aumentam as enzimas activadas.

### Papel dos activadores do plasminogénio na mama bovina

Durante a fase final da lactação e durante a involução mamária post-lactação ocorre a conversão aumentada do plasminogénio em plasmina.

É conhecido que o pico da produção leiteira nos bovinos ocorre durante o segundo mês de lactação e que após isso se entra numa involução gradual da produção leiteira até a lactação terminar o que está correlacionado com a involução da massa de tecido secretor da mama.

O sistema plasmina-plasminogénio desempenha um papel importante nesta involução e sabe-se inclusivamente noutros sistemas celulares que ele está implicado também na proliferação celular.

A conversão do plasminogénio em plasmina leva à activação desta enzima que é uma serina protease capaz de degradar a maioria das proteínas da matriz extracelular e daí a sua importância nas interacções células –ECM.

Vejamos o controlo desta conversão.

A conversão plasminogénio-plasmina é controlada por activadores do plasminogénio (PA) e pelos seus inibidores (PAI).

Há dois tipos de activadores PA, a urocinase-PA (U-PA) e o activador tissular PA (t-PA) sendo estes dois enzimas expressos por genes distintos.

A actividade t-PA parece sobretudo envolvida na fluidez do meio extracelular (trombolise) e o U-PA parece mais implicado na remodelação dos tecidos que ocorre na glândula mamária durante a sua involução.

Por outro lado estes activadores PA podem ser controlados pelos seus inibidores PAI, dos tipos PAI-1 e PAI-2.

Existe ainda um receptor na superfície celular, específico para a U-PA (urocinase-PA) ou seja o U-PAR que pode interagir com o U-PA com elevada afinidade, activando e acelerando a conversão plasminogénio em plasmina.

Nas células epiteliais e mioepiteliais a expressão do U-PA, PAI e U-PAR está bem assinalada e está firmemente ligada essa expressão com o ciclo celular, havendo uma correlação entre a proliferação celular e a expressão dos genes relacionados com o PA:

No entanto a expressão dos genes relacionados com a PA não está directamente relacionada com a mitogénese.

No leite à medida que a lactação avança, aumentam as taxas de plasminogénio e plasmina, embora a relação plasminogénio/plasmina decline com o avançar do período de lactação o que indica que durante este está favorecida a conversão plasminogénio---plasmina.

Na fase final da lactação nas vacas, a lactação e a involução mamária podem ocorrer simultaneamente.

Nas mamites a taxa de plasmina no leite aumenta coexistindo com a actividade proteolítica aumentada.

A administração de bST evita a formação de plasmina admitindo-se pois que a bST retarda a involução gradual na glândula mamária.

### **1.10 - Estrutura e funções dos genes das proteínas do leite**

A expressão dos genes é regulada, inclusivé na glândula mamária, a diversos níveis nos animais superiores como se refere seguidamente:

- 1) Ao nível da cromatina aberta que permite a acessibilidade dos factores reguladores aos genes.
- 2) Ao nível da modulação dos promotores acessíveis com a correspondente velocidade de transcrição.
- 3) Ao nível do processamento do RNA através do “splicing” e da poliadenilação.
- 4) Ao nível da exportação para o citoplasma dos transcritos maduros.
- 5) Ao nível da estabilidade e translação dos mensageiros.

No leite dos ruminantes mais de 95% das suas proteínas são sintetizadas a partir de seis genes estruturais,  $\alpha$ -lactalbumina e  $\alpha$ -lactoglobulina (as principais proteínas do soro do leite) e as caseínas  $\alpha$ -s1,  $\beta$ ,  $\gamma$ -s2 e k codificadas estas ultimas quatro por quatro genes firmemente unidos e aglomerados por esta ordem, num fragmento genómico de DNA de 250 kb com sede no cromossoma 6 das vacas e no 4 das cabras. A  $\alpha$ -lactalbumina está assinalada no cromossoma 5 e a  $\alpha$ -lactoglobulina no cromossoma 11 na vaca

A activação destes genes requer, as caixas TATA e CAAT assinalantes do inicio de transcrição pela RNA polimerase II, requerendo ainda além destas caixas locais para interacção com efectores mixtos implicados na especificidade do estágio dos tecidos, e da indução e modulação da expressão da unidade de transcrição.

Como já referimos é conhecido que a indução da expressão dos genes das proteínas do leite envolve a prolactina, glucocorticóides, insulina, e vários outros factores lactogénicos inclusive a GH e ainda interacções entre as células epiteliais mamárias e a matriz extracelular.

Embora existam diferenças nas distâncias que separam os genes das caseínas uns dos outros, a organização genómica global está bem conservada entre os diversos mamíferos tal como a presença de elementos reguladores cis-actantes determinantes.

Na figura 17 seguinte esquematizam-se a organização genómica dos locus das caseínas nas vacas/cabras bem como as suas características e a organização estrutural das quatro unidades de transcrição das caseínas bovinas e da de duas proteínas do soro do leite.

Figura 17

Organização do genoma

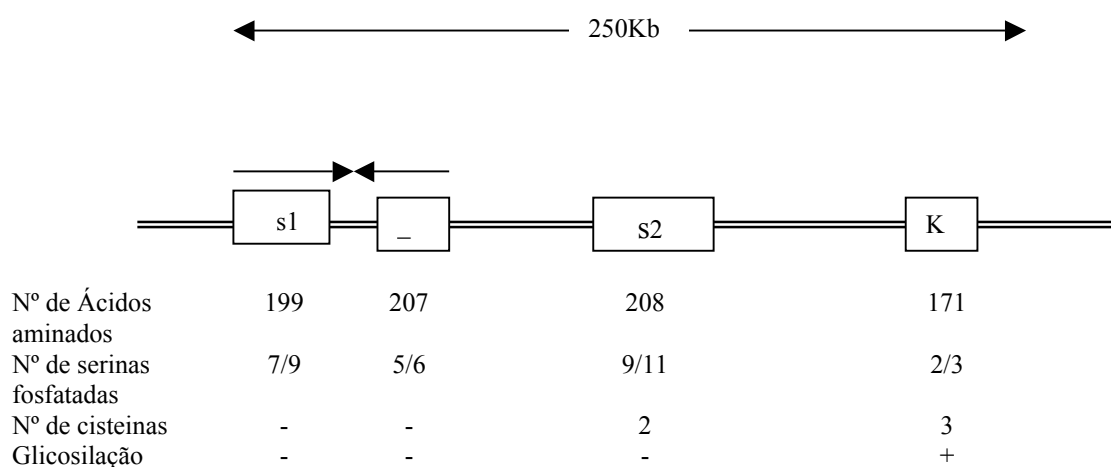
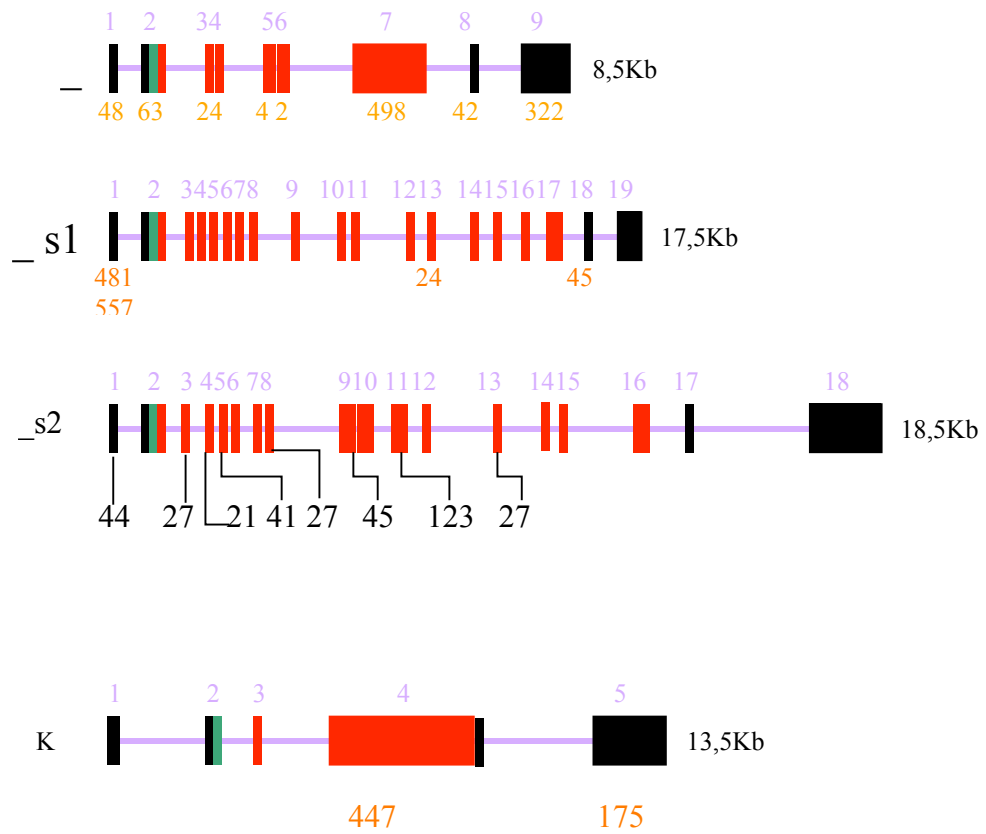


Figura 17 A

## Organização estrutural das quatro caseínas



Parte do exão codificando o péptido sinal

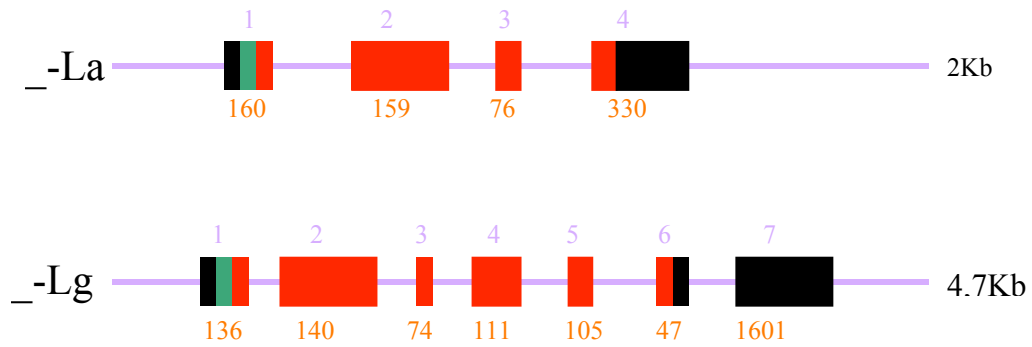
Intrões

Exões e partes de exões codificando proteínas maduras

Regiões 5' e 3' não transcritas

Figura 17 B

Organização estrutural das unidades de transcrição codificando as duas principais proteínas do soro do leite



#### Intrões

■ Exões ou parte de exões codificando proteínas maduras

■ Regiões 5' e 3' não translidas

■ Parte do exão codificando o péptido sinal

1, 2, etc. = nº do exão

160, 159, etc. = nº de pares de bases de cada exão

Figura 17 C

Os quatro genes das caseínas são coordenados nas suas expressões e os três genes codificando caseínas sensíveis ao cálcio ( $\alpha_{s1}$ ,  $\alpha_{s2}$  e  $\kappa$ ) revelam motivos reguladores comuns na região flanqueadora 5' proximal, enquanto a organização da região flanqueadora 5' proximal da  $\kappa$ -caseína é diferente. Apesar disso o seu perfil de expressão parece idêntico ao de outros genes das caseínas.

É necessário um conjunto comum de factores de transcrição para a expressão dos genes das proteínas do leite na maioria das espécies mamíferas.

Como dissemos os genes das três caseínas sensíveis ao cálcio ( $\alpha_{s1}$ ,  $\alpha_{s2}$  e  $\kappa$ ) têm motivos comuns na região promotora e contêm sequências similares a codificar o péptido sinal e múltiplos locais de fosfatação.

Os genes das caseínas  $\alpha_{s1}$  e  $\alpha_{s2}$  encontram-se divididos em diversos pequenos exões, que sofrem operações de splicing complexas, e as caseínas “deleted” provêm dos exões “skipping” (saltitões).

Os quatro genes das caseínas bovinas encontram-se aglomerados em cerca de 250Kb do cromossoma seis ou talvez quatro como mapeamento recente assinala (banda q 32) (numeração de acordo com a ISCNDA 89 standard).

Pseudogenes da  $\alpha$ -lactalbumina e  $\alpha$ -lactoglobulina estão assinalados nos ruminantes.

Está provada a ocorrência de importantes motivos reguladores na região flanqueadora proximal 5' inclusive um local para reconhecimento por um factor nuclear mamario específico.

A heterogeneidade dos quatro tipos de caseínas ( $\alpha_{s1}$ ,  $\alpha_{s2}$ ,  $\beta$  e  $\kappa$ ) resulta da O-fosfatação incompleta das caseínas, O-glicosilação da  $\kappa$ -caseína, da proteólise parcial pela plasmina e do polimorfismo genético.

A ordem relativa dos três locus das caseínas  $\alpha_{s1}$ ,  $\alpha_{s2}$  e  $\kappa$  foi postulada já em 1973.

Os genes das três caseínas sensíveis ao cálcio ( $\alpha_{s1}$ ,  $\alpha_{s2}$  e  $\kappa$ ) têm similares múltiplos locais de fosfatação e peptideos sinal.

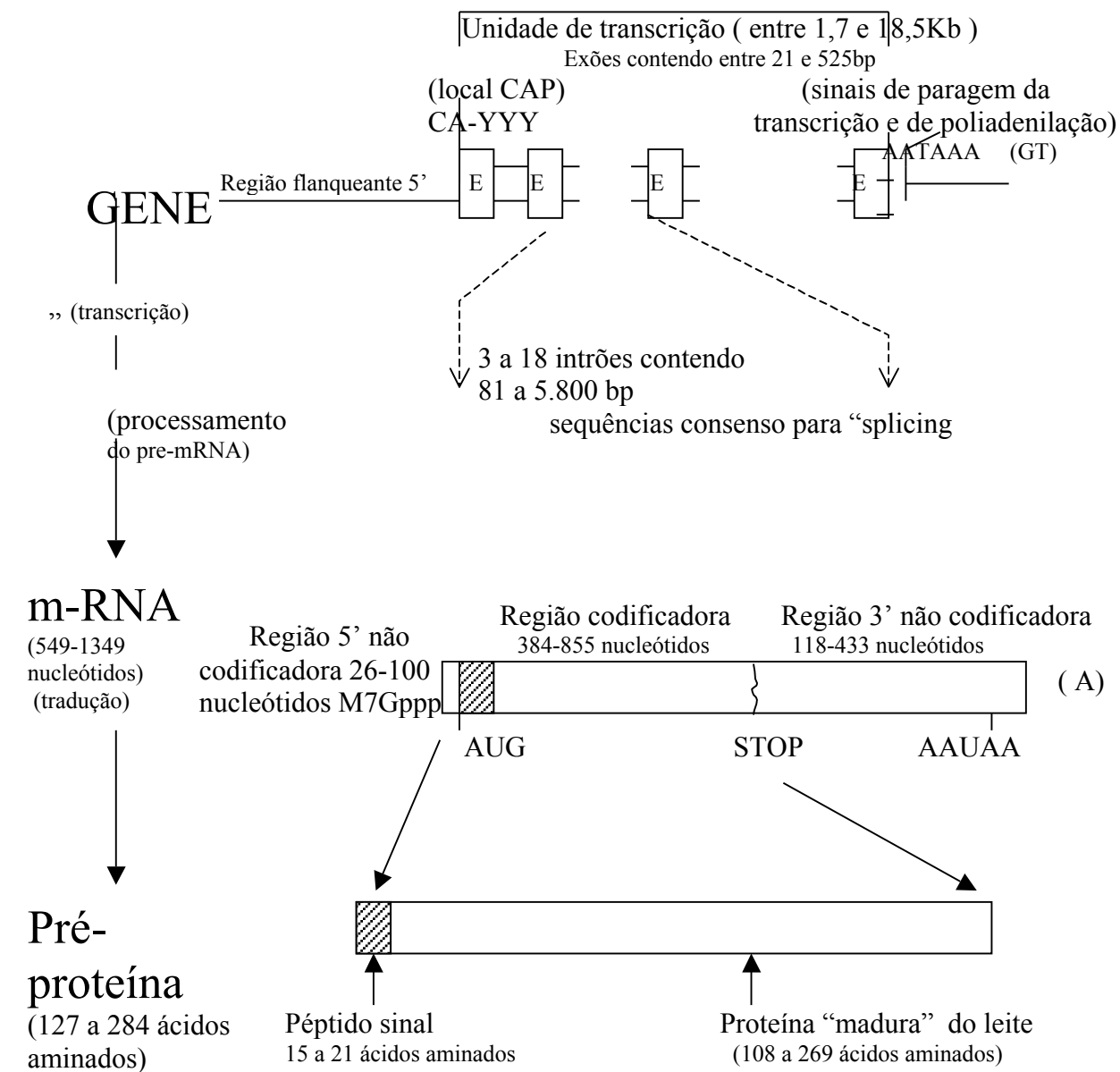
O mRNA codificador das proteínas do leite aumenta notavelmente, nas células epiteliais mamarias desde o meio da gravidez até á lactação, e nesta fase da lactação o mRNA das proteínas do leite representa cerca de 60 a 80 % da totalidade do mRNA existente o que parece dever-se a um aumento da velocidade da transcrição (2 a 4 vezes) e a um aumento da sua semi-vida (17 a 25 vezes).

Figura 18

Representação geral e esquemática das unidades de transcrição e mRNA codificando as principais proteínas específicas do leite

A unidade de transcrição tem tamanhos variando entre 1,7 e 18,5 Kb e contém 3 a 18 intrões constituídos por 81 a 5.800 bp. Os intrões por vezes contêm sequências repetitivas.

Os exões (E) na sua maioria são um tanto curtos e os seus tamanhos oscilam entre 21 e 525 bp.



Y = Pirimidinas



Estes genes e os m-RNA codificam as quatro caseínas, a  $\alpha$ -lactoglobulina, a  $\beta$ -lactalbumina e a proteína acídica do soro (WAP).

Na figura 18 anterior o m-RNA contem uma extremidade protegida M7Gppp que parece proteger esta região 5' não traduzida contra a degradação enzimática e facilitando a ligação da sub-unidade ribossomal 40 S e Met tRNA do complexo dos factores de iniciação contendo ainda o mRNA uma porção que é traduzida delimitada por codões de iniciação e de terminação, e uma região 3' não traduzida com um local de reconhecimento para a poliadenilação localizada 13 a 20 nucleótidos acima da cauda poly (A).

Cerca de 20% pelo menos do mRNA traduzível da proteína do leite tem uma cauda poly (A) muito curta, ou até não existente.

Os mRNA codificando as caseínas,  $\alpha$ -lactoglobulinas e  $\beta$ -lactoalbumina e WAP têm de 549 a 1.349 nucleótidos e a região codificante representa em média 60 a 70% do mRNA.

No estudo das sequências nucleotídicas interespecies do cDNA codificante das caseínas, na região especificando o péptido sinal, foi possível verificar que a maioria das substituições nucleotídicas observadas na região codificante do péptido sinal não induzem a substituição de qualquer ácido aminado. Mas pelo contrário na região codificadora da proteína “madura” ocorrem numerosas substituições nas três posições de diversos codões o que sugere que a transferência eficiente das mais abundantes proteínas do leite através das membranas do retículo endoplasmico necessitam da integridade estrutural do péptido sinal transitório para interactuar com a partícula reconhedora do sinal ou com a sequência receptora do sinal.

Na estrutura dos genes codificadores das proteínas do leite referida e em figura mais adiante, a unidade de transcrição contem de 5' a 3' as sequências consenso correspondentes ao sítio do “cap”, aos numerosos locais para o “splicing”, e o sinal de poliadenilação AATAAA seguido de uma extensão nucleotídica rica em GT, assinalando o fim da transcrição. A região flanqueadora proximal 5' contem uma caixa TATA e possivelmente outros motivos que assinalam o local de transcrição para a RNA polimerase II e para o ubíquo factor de transcrição TF II D.

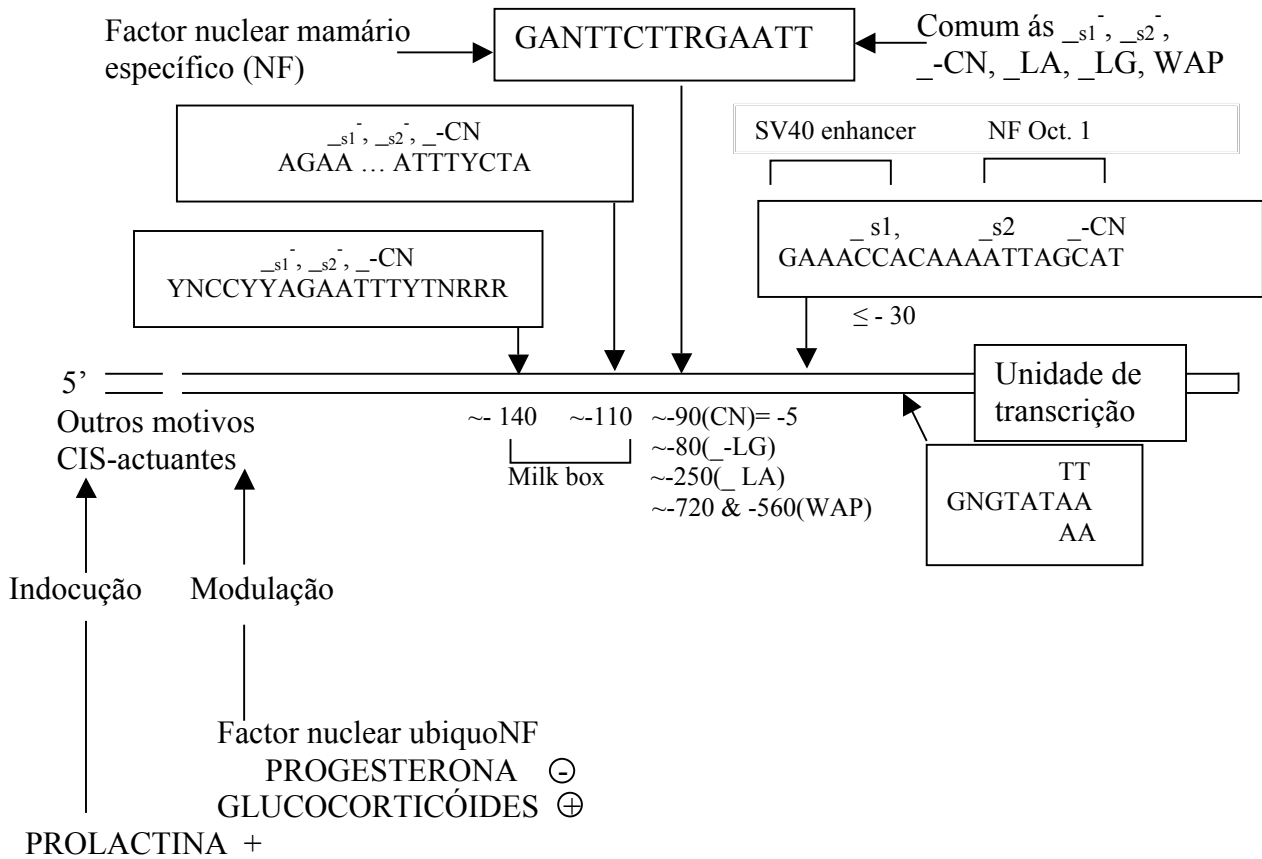
A activação do gene necessita ainda do reconhecimento por diversos efectores de sequências consenso de conformações locais implicadas na indução e modulação da expressão.

Muitos destes motivos cis-actuaentes são proximais 5' da unidade de transcrição, mas alguns podem ser 5' distal ou mesmo ocorrer dentro da unidade de transcrição ou da região flanqueadora 3'. (Figura 19).

Outros elementos distais, regiões de ligação a matriz estão implicados na partilha do genoma em domínios funcionais topologicamente distintos e podem ser importantes na expressão dos genes.

Figura 19

Alguns motivos estruturais importantes identificados nas regiões flangeadoras 5' dos genes codificando as principais proteínas específicas do leite em diversas espécies animais



A caixa TATA e possivelmente outros sinais indicam o início da transcrição pelo complexo polimerase II-factor de transcrição II D.

Outras caixas referem motivos estruturais comuns aos genes codificadores das três caseínas sensíveis ao cálcio ( $\text{CN}$ ,  $\text{LA}$ ,  $\text{LG}$ ).

A caixa superior é comum a estes três genes e ainda dos da  $\text{LG}$ ,  $\text{LA}$  e  $\text{WAP}$ , sendo este motivo reconhecido pelo factor nuclear mamário específico.

Estão identificadas as localizações de diversos locais de potencial reconhecimento para receptores glucocorticóides e progesterona e para outros efectores, localizados sobretudo na região flangeadora 5'.

A organização nos bovinos dos quatro genes codificadores da caseína está hoje bem conhecida. Os genes da caseína parecem ser estruturalmente muito diferentes em virtude dos tamanhos das suas

unidades de transcrição que varia de 8,5 a 18,5 Kb e o número de intrões que oscila entre 4 e 18. Contudo subsiste a ideia de que têm um precursor comum as  $\_s1^-$ ,  $\_s2^-$  e  $\_CN$ .

O gene da K-caseína não parece ter qualquer “pattern” comum com outros genes das caseínas, postulando-se que pode estar relacionado com o fibrinogénio, quanto à sua evolução.

Comparação entre estes vários genes das caseínas de diversas espécies animais revelam maiores divergências entre os intrões homólogos do que entre os exões, divergências estas que têm sede sobretudo nas diferentes sequências nucleotídicas e no seu tamanho é ainda a ocorrência de sequências repetitivas de diferentes tipos nas regiões flanqueadoras.

Nos bovinos a localização cromossomal dos genes das caseínas e a organização destes genes revelam que a sua ordem deva ser  $\_s1 - \_ - \_s2 - K$  dentro de um locus aglomerado de caseína de 250Kb o qual deve depender de uma região controladora do locus. Os genes da caseína parecem ocorrer com uma cópia simples por cromossoma haplóide.

Os loci caseínicos estão assinalados no cromossoma 5 do rato, 12 da coelha 4 nos humanos e nos ovinos e 6 nos bovinos segundo uns, 4 segundo outros (vide quadro 3).

É difícil discriminar os cromossomas 4 e 6 nos ruminantes domésticos.

#### Localização cromossomal das proteínas do leite

	<b>Caseínas</b>	<b>Lactalbumina</b>	<b>Lactoglobulina</b>	<b>WAP</b>
Vaca	6	5	11	-
Cabra	4	-	11	-
Ovelha	4	3	3	-
Humano	4	12	-	-
Rato	5	-	-	11
Coelho	12	-	-	-

Quadro 3

Como já referimos no funcionamento dos genes que expressam as caseínas a acção sinérgica das hormonas lactogénicas e da matriz extracelular sobre a expressão do gene da  $\_$ -caseína está bem estabelecida.

Em cultura de células, os glucocorticóides na presença de insulina são necessários para a rápida e forte indução do promotor da  $\_$ -caseína pela prolactina. A primeira hormona pode actuar indirectamente ou regulando os genes sensíveis aos glucocorticóides produzindo efectores “trans-acting” sobre o gene da  $\_$ -caseína ou rompendo nucleosomas no locus da  $\_$ -caseína o qual contem sequências consenso para complexos receptores glucocorticóides.

Locais de ligação para proteínas, semelhantes a motivos que existem nos genes da  $\_s2^-$  caseína estão assinalados na região flanqueadora 5' do gene  $\_$ -caseína, sugerindo uma desrepressão da transcrição.

Dados sobre elementos reguladores controlando outros genes de outras caseínas são escassos. No entanto está assinalado um factor nuclear especificamente mamário e um factor octamero 1 (Oct.1) (ver figura anterior).

Na expressão dos genes das proteínas do soro do leite, baixas concentrações de mRNA da WAP existem no tecido mamário das virgens e nos animais em início de gravidez mas depois ocorre um aumento de quase mil vezes a meio da lactação em consequência da proliferação e diferenciação das células epiteliais.

A manutenção e indução da expressão do gene endógeno da WAP depende da acção sinérgica das hormonas lactogénicas (prolactina, glucocorticóides e insulina) e interacção célula-célula e célula-matriz extracelular.

Parece que os elementos que respondem às hormonas lactogénicas se distribuíram ao longo de 2,5Kb na região flanqueadora 5' do gene.

Quanto ao gene da  $\alpha_2$ -lactoglobulina, nas espécies ovinas, ele é expresso a meio da gravidez e durante a gestação a sua taxa de expressão é mais alta que a da mRNA das caseínas, aumentando lentamente até ao parto e depois deste aumenta mais rapidamente. A expressão do gene da  $\alpha_2$ -lactoglobulina nos ovinos parece menos dependente das hormonas lactogénicas do que os genes das caseínas.

A expressão do gene da  $\alpha_2$ -lactalbúmina no rato requer a acção sinérgica da insulina e da prolactina e é inibida nos mamíferos pela progesterona e o AMP cíclico também é um regulador negativo.

O mRNA maduro susceptível de tradução no citoplasma deriva do transcrito primário através de uma série de processamentos complexos em que recordamos: 1) adição de um nucleótido metilado (capping) na extremidade 5' do pré-mRNA nascente; 2) a poliadenilação da terminação encurtada 3' do transcrito primário; 3) fixação de proteínas ribonucleicas e metilação dos resíduos de adenosina interiores; 4) excisão dos intrões e "splicing" dos exões pelo equipamento spliceossoma; 5) desadenilação parcial.

O splicing dos 19 exões do pré-mRNA da  $\alpha_{s1}$  e 18 exões do pré-mRNA  $\alpha_{s2}$  é muito complexo e daí apareceram diversas variantes das caseínas.

Os conhecimentos actuais sobre a estrutura e funcionamento dos genes das principais proteínas do leite pode ser utilizado para seleccionar animais com genótipos leiteiros interessantes e para criar animais transgénicos utilizados como modelos para estudar carcinomas mamários, ou para produzir proteínas exógenas de alto valor como produtos farmacêuticos.

Análise por fenotipagem do DNA tem sido utilizada para seleccionar bovinos possuindo o alelo B da K-caseína, a qual está associada com uma velocidade maior de coagulação do leite e com a obtenção de uma coalhada mais firme.

Também caprinos machos possuindo os alelos A, B ou C da  $\alpha_{s1}$  caseína estão associados nas fêmeas suas descendentes com mais elevadas produções de  $\alpha_{s-1}$  e caseínas totais.

O gene da K-caseína ou um gene ligado muito próximo a este afecta a percentagem de proteína do leite e o gene da  $\alpha_2$ -lactoglobulina e um gene ligado muito próximo a este, afecta a percentagem de gordura. É importante identificar alguns alelos associados com características leiteiras de interesse. A sua utilidade na análise fenotípica do DNA de recém nascidos ou do esperma é óbvia.

## **1.11 - Fracções proteicas do leite e seus polimorfismos genéticos**

Nos humanos o “locus” da família de genes da caseína compreende pelo menos quatro genes de caseína, pela ordem seguinte  $\alpha_1$ ,  $\alpha_2$  e  $\kappa$ , tendo um tamanho aproximado de 350Kb havendo sugestões de que este aglomerado caseínico esteja localizado dentro dos 700Kb do aglomerado gene da albumina que parece localizado em 4q13.

As caseínas são membros de uma família de vários genes pelo menos na vaca e nos humanos, sendo o grupo de proteínas mais rapidamente divergentes, apesar de conterem uma série de resíduos de ácidos aminados bem conservados ao longo de todo o seu comprimento.

Também o leite bovino contém quatro caseínas, duas  $\alpha$ , uma  $\beta$  e uma  $\kappa$ .

Os ácidos aminados N-terminais fosfoserina/fosfotreonina da  $\alpha$ -caseína são cruciais para o funcionamento biológico da molécula e variação no seu número pode influenciar a qualidade total do leite.

Como já vimos a transcrição do gene da  $\alpha$ -caseína (CSN2) é controlado primariamente por um elemento de respostas compósitas (CoRE) que integra sinais vindos das hormonas lactogénicas, prolactina (PRL), insulina e hidrocortisona nas células epiteliais mamárias. Este CoRE contém locais para ligação do STAT5 e para o C/EBP-beta ou TCF5 e diversos meios locais para receptores glucocorticoides (GR).

Foi observada transactivação cooperativa quando todos os três factores eram expressos, mas não entre o STAT5 e o C/ERP-beta na ausência de GR activo.

O gene da Kappa-caseína é muito similar nos humanos e bovinos, estendendo-se ao longo de 8.821 nucleotídeos consistindo de cinco exões indo de 33 a 496 nucleótidos separados por intrões de 1.146 a 2.942 nucleótidos.

Nos bovinos, os quatro genes de caseína,  $\alpha_1$ ,  $\alpha_2$ ,  $\beta$  e  $\kappa$  estão fisicamente ligados dentro de uma região de 200 a 300 Kb e parece que arrançados pela ordem  $\alpha_1$ ,  $\alpha_2$  e Kappa caseína, sendo a distância entre  $\alpha_1$  e  $\beta$  e entre  $\alpha_1$  e  $\kappa$  de cerca de 10 e 300Kb respectivamente.

O gene da  $\alpha$ -caseína está orientado na direcção oposta à dos genes  $\alpha_1$  e  $\kappa$ .

As caseínas são classificadas em duas famílias, a primeira constituída de  $\kappa$ -caseína e a segunda agrupando as  $\alpha_1$ ,  $\alpha_2$  e  $\beta$ -caseína.

As caseínas  $\alpha/\beta$  são contudo uma família de proteínas que tem divergido muito rapidamente, embora duas regiões estejam conservadas, um aglomerado de resíduos de serina fosfatados e a sequência sinal.

A marca distintiva desta família de proteínas é baseada nos últimos oito resíduos da sequência sinal.

As caseínas não são compactas, sendo moléculas flexíveis com muitos resíduos acessíveis a solventes, com muitas partes das cadeias polipeptídicas (as não fosfatadas e não glicosiladas) assumindo nos ruminantes conformação alfa-hélice-loop-alfa-hélice na qual a região “loop” contém um conjunto de locais para fosfatação.

A  $\alpha$ -lactalbumina, é uma sub-unidade reguladora da lactose sintetase que na glândula mamária muda a especificidade do substrato da galactosil-transferase da N-acetilglucosamina para glucose.

Está assinalada por Perez et al (1994), a reduzida concentração de caseína no leite de cabra devido à ocorrência de uma sequência LINE na 3' UTR do alelo codificador da  $\alpha$ -s1 caseína E.

### Polimorfismos

Nas regiões 5' flanqueadoras dos genes das proteínas do leite estão identificadas várias sequências putativas cis-reguladoras e locais potenciais de ligações de factores reguladores trans-actantes.

Estão assinalados no tecido mamário nestas regiões 5' flanqueadoras dos genes das proteínas do leite, locais para interacção com os factores de transcrição, MGF/STAT5a (mammary gland factor/signal transducer and activator of transcription), MAF (mammary cell-activating factor), MPBF (milk protein binding factor, identificado como STAT5a), PMF (pregnancy-specific mammary nuclear factor), assim como locais para factores de transcrição ubíquos como os AP1, CREB, CTF/NF1, OCT-1, SP1, NF-kB, GR, TR, YY1, HSE.

A maioria destas sequências reguladoras encontram-se nas regiões proximais dos promotores dos genes das proteínas do leite, dentro das 200-250 bp adjacentes ao local de início da transcrição, embora no gene da  $\alpha$ -caseína bovina um elemento regulador distal o bce1 se localize 1.500 bp acima do local de início da transcrição

Estão assinalados, recentemente, polimorfismos dentro destas regiões 5' flanqueadoras dos genes das proteínas do leite ou seja dos elementos cis-reguladores.

As variações nas sequências nucleotídicas das regiões promotoras dos genes das caseínas do leite bovino podem alterar a afinidade destas regiões para os factores de transcrição, e influenciar assim a composição das proteínas do leite.

Vejamos por tipos de genes de proteínas do leite os polimorfismos respectivos em elementos cis-reguladores:

#### 1) $\alpha$ -Lactalbumina

Esta proteína é essencial para a síntese da lactose na glândula mamária.

Ela parece ser monomórfica nas raças bovinas europeias.

Contudo na sequência 5' flanqueadora do gene da  $\alpha$ -lactalbumina da mama de bovino Holstein estão identificados três polimorfismos nas posições +15, +21, e +54 em relação ao ponto de início da transcrição do mRNA.

As vacas com o alelo A deste gene produzem mais leite, mais proteína, e mais gordura enquanto no alelo B há maior percentagem de proteína e gordura.

#### 2) $\alpha$ -Lactoglobulina

Estão identificados doze polimorfismos nos bovinos sendo mais frequente o A e o B.

O alelo B está associado com mais elevado teor de caseína e gordura no leite de vaca.

### 3) k-Caseína

Estão assinaladas duas variantes genéticas nos bovinos a A e a B, sendo a última produtora de leite de qualidade superior nos bovinos.

O alelo B está também associado com tempos de coagulação do leite mais curtos.

### 4) Outras caseínas.

Na  $\kappa$ -caseína de bovino está assinalada uma simples deleção de um par de bases na posição -516 do local de início da transcrição.

Nas caseínas sensíveis ao cálcio, nos bovinos, estão assinaladas 34 variantes em cerca de 1.2kb das regiões flanqueadoras dos genes destas caseínas 17 destas 34 mutações (50%) situam-se dentro de locais conhecidos como potencialmente reguladores, sugerindo o seu possível papel na interação com factores de transcrição.

No polimorfismo genético dos genes das proteínas do leite estão assinaladas no quadro seguinte 40 variantes genéticas diferentes nos loci das proteínas do leite bovino sendo encontradas grandes diferenças nas frequências dos alelos consoante as raças bovinas.

O impacto destes polimorfismos, na composição do leite e na sua capacidade para a produção de queijo é proeminente na  $\alpha$ -lactoglobulina ( $\alpha$ -Lg A/B) e na K-CN A/B nos bovinos e  $\alpha$ <sub>s1</sub>-CNE nas cabras.

As seis principais proteínas do leite dos ruminantes são codificadas por quatro genes para as caseínas (CN), o gene para  $\alpha$ -lactoglobulina ( $\alpha$ -Lg) e o gene da  $\alpha$ -lactoalbumina ( $\alpha$ -LA).

Como se disse nos bovinos os quatro genes para as caseínas ( $\alpha$ <sub>s1</sub>-CN,  $\alpha$ -CN,  $\alpha$ <sub>s2</sub>-CN e K-CN) são agrupados no cromossoma seis ocupando uma região de cerca de 250Kb.

A ordem dos genes bem como a sua orientação transcricional são altamente conservadas entre os ruminantes.

Os genes da caseína nos ruminantes encontram-se num simples loci, enquanto existem pseudogenes para a  $\alpha$ -LG dos bovinos e ovinos, encontrando-se nos bovinos para a  $\alpha$ -LA um pseudogene flanqueado por duas sequências LINE directamente repetidas.

O comprimento da unidade de transcrição e o número e comprimento dos exões varia muito consoante os genes das proteínas do leite.

São conhecidas as sequências genómicas de todos os genes das lactoproteínas, bem como as diferenças alélicas a nível molecular, o que permite compreender os mecanismos reguladores implicados na expressão dos genes das proteínas do leite.

Nos bovinos, caprinos e ovinos a sequência e posição das diferenças em ácidos aminados em quase todas as variantes de lactoproteínas é conhecida. Ao lado de simples substituições de ácidos aminados, o que constitui o tipo mais comum de mutação, alguma deleção e variação no número de grupos fosfatos estão identificadas.

Nos bovinos ,no que se refere às variantes das seis principais lactoproteínas do leite eles são mais polimórficos do que nos caprinos e nos ovinos, estando assinaladas quarenta variantes genéticas destas seis lactoproteínas (vide quadro 4).

A deleção de 13 ácidos aminados na s1-CN bovina é devida ao exão 4\_saltitão (skipping) devido a uma mutação pontual dentro do local dador do corte e não por deleção do DNA genômico. Variantes mais curtas da mesma caseína s1-CN de ovino e caprino são devidas ao “skipping” do exão 16 e a diferente “splicing” do exão 11, 13 e 16 do pré m-RNA.

Existem polimorfismos específicos nos alelos A e B da região promotora proximal do gene da -LG, que podem originar diferentes expressões de ambos os alelos.

Nos bovinos os alelos A e B da K-CN podem ser expressos diferentemente mas não existem polimorfismos específicos nos alelos na região promotora próxima.

A distribuição das frequências dos alelos das lactoproteínas varia muito consoante as raças de bovinos. Para a maioria das raças os locus s2-CN e LA são considerados monomórficos.

Na maioria das raças os alelos s1-CNB e CNA1 e A2 são os mais frequentes.

Quadro 4  
Alelos mais comuns dos genes das seis lactoproteínas dos bovinos

<u>Locus</u>	<u>Alelos</u>
<u>s1</u> -CN	A: deleção dos Ácidos Aminados 14-26
	B: alelo de referência
	C: 192 Glu <u>Gly</u>
	D: 53 Ala <u>Thr P</u>
	E: 59 Gln <u>Lys</u> ; 192 Glu <u>Gly</u>
	F: 66Ser P <u>Leuc</u>
	G: 371 bp inserção no exão 19

<u>Locus</u>	<u>Alelos</u>
<u>s2</u> -CN	A: alelo de referência
	B: não determinado
	C: Ser P <u>Gly</u>
	D: Delec. A A 51 <u>59</u> , exão VIII skipping

<u>Locus</u>	<u>Alelos</u>
<u>-CN</u>	A1: 67Pro <u>His</u>
	A2: alelo de referência
	A3: 106 His <u>Gln</u>
	B: 67 Pro <u>His</u> ; 122 Ser <u>arg</u>
	C: 35 Ser P <u>Ser</u> ; 37 Glu <u>Lys</u> 67 Pro <u>His P</u>
	D: 18 Ser P <u>Lys</u>
E: 36 Glu <u>Lys</u>	



<u>Locus</u>	<u>Alelos</u>
K-CN	A: alelo de referência
	B: 136Thr _ Ile; 148 Asp - Ala
	C: 81Asn _ Asp; 97 Arg _ His; 136 Thr _ Ile; 148 Asp – ala
	E: 155 Ser _ Gly
	F: 10 Arg _ His
	G: 97 Arg _ Lys

<u>Locus</u>	<u>Alelos</u>
-LG	A: 64 Gly _ Asp; 118 Ala-Val
	B: alelo de referência
	C: 59 Glu _ His
	D: 45 Glu _ Gln
	E: 158 Glu _ Gly
	F: 50 Pro _ Ser; 129 Asp _ Tyr 158 Glu _ Gly
	G: 78Ile _ Met; 158 Glu _ Gly
	H: não determinado
	I:106 Glu _ Gly
	J: 126 Pro _ Leu
	W: 56 Ile _ Leu
	X: não determinado

<u>Locus</u>	<u>Alelos</u>
-LA	A: 10 Arg _ Glu
	B: alelo de referência
	C: não determinado

As variantes genéticas das lactoproteínas afectam a composição do leite e as suas propriedades queijeiras, embora existam resultados muito discordantes a este propósito.

A idade dos animais, período de lactação, raças, modelos estatísticos, etc. podem condicionar fortemente os resultados.

Associados com mais alta produção de leite, gordura e proteína parecem estar os alelos s1-CNB e -CNA, bem como provavelmente a K-CN BB e -LG AA sobre a produção de proteína.

Efeitos positivos significativos sobre a percentagem de gordura, sólidos totais e conteúdo em caseína foram identificados com os alelos K-CN B e -LGB.

Também um efeito aditivo positivo foi encontrado entre o alelo -LG B e o conteúdo em caseína e o número de caseínas.

Leite contendo a variante K-CN B revelou melhor capacidade para a coagulação, maior firmeza da coalhada, melhor produção de queijo e melhor recuperação de gordura.

Parece pois que os efeitos mais marcantes das variantes genéticas das lactoproteínas dos bovinos se encontram no K-CN alelo B o qual se encontra associado com coagulação mais rápida do leite, e coalhada mais firme e para o alelo  $\beta$ -LG B.

O principal efeito sobre o conteúdo em caseína na cabra parece dever-se ao  $\beta$ <sub>s1</sub>-CN onde o alelo E é expresso a um nível três vezes inferior

### **1.12 - Regulação da fraccção lipídica do leite**

A regulação e manipulação nutritiva no síndrome da baixa de gordura do leite, tem grande importância pois a modificação da composição em ácidos gordos do leite pode contribuir para tornar o leite mais saudável como alimento.

A gordura do leite é sobretudo constituída por triglicéridos e nos ruminantes os ácidos gordos do leite dependem de duas fontes, ou provêm de novo da síntese a partir do acetato e do  $\beta$ -hidroxibutirato ou então provêm da tomada a partir dos ácidos gordos de cadeia longa em circulação provenientes da dieta ou das reservas lipídicas da gordura animal.

Os ácidos gordos contidos nos triglicéridos do leite têm comprimentos de cadeia variados, sendo aquelas nos ácidos gordos de C4 a C14 e mesmo uma parte dos C16 provenientes da síntese de novo, e os restantes C16 e outros de cadeia mais longa provenientes da tomada de ácidos gordos pré-formados.

O facto de existir uma alta correlação genética entre a produção de leite e a produção de gordura deste torna difícil a utilização da selecção para alterar a gordura do leite, independentemente de outros constituintes do leite.

Mas a nutrição e a dieta podem rapidamente (em alguns dias) condicionar a produção e composição da gordura do leite, podendo algumas dietas reduzirem acentuadamente a síntese de gordura desencadeando até o síndrome da depressão da gordura do leite (MFD).

Há sobretudo dois tipos de dietas comerciais que podem provocar este síndrome MFD, o primeiro tipo é o que veicula grandes quantidades de carboidratos facilmente digestíveis e reduzidas quantidades de constituintes fibrosos, como é o caso das dietas conterem muitos cereais e poucas forragens. No segundo tipo encontram-se dietas suplementadas com certos óleos vegetais ou muitas sementes donde são extraídos estes ou ainda com óleos de pescado, que contêm muitos ácidos gordos insaturados.

Além dos tipos de dietas antes referidas como causadoras de MFD, outros factores também estão implicados nesta situação, como é o caso da preparação física das dietas (por exemplo o tamanho das partículas e dos pellets) as práticas de manejo ( por exemplo a frequência de administração das dietas) e factores do próprio animal (por exemplo fase da lactação, nível de ingestão).

As dietas desencadeadoras de MFD surtem efeitos em poucos dias e esses efeitos são específicos, podendo a gordura do leite ser reduzida em 50% ou mais, com pouca variação da produção de leite ou de outros componentes do leite, produzindo uma diminuição da produção de todos os ácidos gordos, (embora o declínio destes seja maior ao nível dos ácidos gordos sintetizados de novo) o que leva que a composição da gordura do leite também se altera diminuindo as proporções de ácidos gordos de cadeia curta e média e aumentando os de cadeia comprida. Está também assinalado no síndrome MFD um aumento marcado dos ácidos gordos trans. Sabe-se que o tamponamento do pH do rúmen reduz ou elimina este síndrome desencadeado por dietas de muitos cereais (grão) e pouca forragem, tal como o

“bypassing” do rúmen por infusão no abomaso reduz ou elimina a MFD devida a suplementos com óleos, o que permite concluir que a indução de MFD implica produtos formados pelas bactérias do rúmen, que dependem das mudanças ocasionadas nas fermentações no rúmen provocado pelas dietas.

Para explicar o síndrome MFD tem sido aventadas duas hipóteses: numa delas considera-se que a diminuição da gordura se deve a um menor fornecimento de precursores lipídicos à glândula mamária e noutra atribui-se a causa da diminuição de gordura à inibição directa na glândula mamária de uma ou mais etapas da biossíntese lipídica dos ácidos gordos poli-insaturados que sofreram uma incompleta ou invulgar biohidrogenação no rúmen, sobretudo ácidos gordos trans e metabolitos com eles relacionados.

É sabido que o estado endócrino pode alterar a partilha de nutrientes pelos tecidos. A insulina tem efeitos agudos sobre a velocidade lipogénica do tecido adiposo, que estimula, e da lipólise que inibe, mas a glândula mamária de bovino não responde às variações de insulina em circulação.

Dietas com elevadas concentrações de grãos (cereais) originam uma elevada produção de propionato no rúmen e elevadas taxas de gluconeogénese hepática o que desencadeia um aumento de secreção de insulina pelo pâncreas. Esta subida da concentração de insulina que chega a atingir aumentos de 2 a 5 vezes ao nível do tecido adiposo favorece a tomada de precursores lipogénicos (lipogénese) e diminui a lipólise. Esta canalização preferencial de nutrientes lipogénicos para o tecido adiposo originaria a sua falta no tecido mamário e daí MFD.

Contudo experiências concludentes parecem negar esta teoria glucogénica-insulina para explicar o síndrome MFD.

A teoria da inibição directa da síntese da gordura do leite pelos trans ácidos gordos, trans-C<sub>18:1</sub> foi proposta há três dezenas de anos e hoje parece a mais plausível.

Os ácidos gordos trans-C<sub>18:1</sub> são intermediários que se formam durante a biohidrogenação dos ácidos gordos polinsaturados.

Numa grande gama de dietas há uma estreita relação entre a diminuição da gordura do leite e o seu aumento em trans-C<sub>18:1</sub>.

Uma dieta de ácidos gordos insaturados com uma fermentação no rúmen perturbada levando a uma biohidrogenação incompleta são duas razões necessárias para surgir o síndrome MFD.

Está demonstrado que a MFD está relacionada com uma mudança do perfil de isómeros trans C<sub>18:1</sub> para concentrações aumentadas de trans-10C<sub>18:1</sub>.

Na biohidrogenação do ácido linoleico, a primeira etapa é uma isomerização para formar Cis-9, trans-11 ácido linoleico conjugado (CLA), seguindo-se duas reduções que dão trans-11C<sub>18:1</sub> e depois estearato.

O Cis-9, o trans-11CLA e o trans-11C<sub>18:1</sub> são os principais isómeros dos ácidos gordos no processo típico da biohidrogenação que ocorre na gordura do leite.

Em algumas circunstâncias a biohidrogenação no rúmen pode-se alterar e formar-se o trans-10C<sub>18:1</sub> observado no MFD. A via para esta formação tem sido proposta e na sua etapa inicial pode formar-se um trans-10, Cis-12CLA.

Na gordura do leite o conteúdo de trans-10 C<sub>18:1</sub> e trans-10, Cis-12CLA são ambos inversamente relacionados com o teor de gordura do leite, sendo o mecanismo desta inibição motivo de profunda investigação.

A propósito da fracção lipídica do leite refere-se que o “tamanho” do depósito de gordura no organismo animal dos bovinos parece ser regulado por uma via de transdução de sinais em que participaria a leptina.

A leptina pode actuar directamente ou indirectamente sobre o sistema nervoso central em ordem a inibir a ingestão de alimentos e a regular o consumo de energia como parte do mecanismo de homeostase visando manter constante a massa adiposa.

A leptina é também produzida nas células epiteliais mamárias dos bovinos sendo a sua produção aqui regulada por via hormonal. Assim ocorre um rápido aumento do mRNA da leptina bovina em resposta à insulina ou ao factor de crescimento insulina-like (IGF-1).

Admite-se que a leptina que se encontra presente também no leite, e nas células epiteliais mamárias dos bovinos possa ser regulada por factores que alteram o funcionamento mamário e a repartição de nutrientes, o que sugere que a leptina é um sinal autócrino/parácrino na glândula mamária bovina.

## **Bibliografia**

### **1.1 - Introdução á biologia da glândula mamária**

Colter, R. J., Mc Grath, M. F., Byatt, J. C. e Zurfluh, L. ( 1993 ) Regulation of bovine mammary growth by peptides hormones: involvement of receptors, growth factors and binding proteins. Livestock Production Science, 35: 21-33

Knight, C. H. e Wilde C. J. ( 1993 ). Mammary Cell Changes During Pregnanay and Lactation. Livestock Production Science, 35: 3-19

Plath, A., Einspanier, R. et al. ( 1998 ). Expression and localization of members of the fibroblast growth factor fammily in the bovine mammary gland J. Dairy Science, 81: 2604-2613

Sandowsk, Y. Peri I., e Gertler, A. (1993 ). Partial purification and characterization of putative paracrine/autocrine bovine mammary epithelium growth factors. Livestock Production Science 35: 35-48

### **1.2 - Factores de crescimento e diferenciação na glândula mamária**

Fantl, V., Creer, A. et al. ( 2000 )- fibroblast growth factor signal ling and cyclin D1 function are necessary for normal mammary gland development during pregnancy A transgenic mouse approach in Biology of the Mammary Gland Advances in Experimental Medicine and Biology, vol. 480 Eds Jan Mol e Roger Clesg pp 1-7

Niemann, G., Brinkmann, V., e Birchmoirs, W. ( 2000 ). Hepatocyte growth factor and neuregulin in mammary gland cell moepogenesis, in Biology of the Mammary Gland

Sinowatz, F., Schams, D., Plath, A., e Köelle, S. ( 2000 ). Expression and localization of growth factors during mammary gland development, in Biology of the Mammary Gland

Plath, A., Peters, F. e Einspanier, R. ( 1996 ). Detection and quantitation of specific mRNA by ribonuclease protection assay using denaturing horizontal polyacrylamide gel electrophoresis: a radioactive and nonradioactive approach. Electrophoresis 17: 471-472

Sinowatz, F., Schams, D., Plath, A., e Köelle, S. ( 2000 ). Expression and localization of growth factors during mammary gland development, in Biology of the Mammary Gland.

### **1.3 - Puberdade e 1.4 - Involução**

Jones, J. S., clemmons D. B. ( 1995 ). Insulin-like growth factors andpheir binding proteins: biological actions. Endocrine Rev. 16: 3-34

Purup, S., Vestegaard, M., e Sejrsen, K. ( 2000 ) Involvement of growth factors in the regulation of puberlat mammary growth in catle. In .Biology of the Mammary Gland.

Tonner, E., Allan, G., Skkreta, L., et al. ( 2000 ). Insulin-like growth factors binding protein-5 ( IGFBP-5 ) potentially regulates programmed cell death and plasminogen activation in the mammary gland, in .Biology of the Mammary gland.

### **1.5 - Somatostatina, opióides, progesterona e progestinas no desenvolvimento da glândula mamária**

Clarke, C. L., Sutherland, R. L., (1990 ). Progestin regulation of cellular proliferation. Endon. Rev. 11: 266-301

Graham, J. D., Clarke, C. L., ( 1997 ). Physiological action of progesterone in target tissues. Endocrino. Rev. 18: 502-519

Hatzoglou, A., Bakogeorgou, E. et al. ( 2000 ). Somatostatin and opioid receptors in mammary tissue. Role in cancer cell growth in .Biology of the Mammary Gland.

Humphreys, R. C., Lydon, J., O' Malley, S. W., Rosen, J. M. ( 1997 ). Mammary gland development is mediated by both stromal and epithelial progesterone receptors. Mol. Endocrinol 11: 801-811

Mol, J. A., Van Garderen, E., Rutteman, G. R., Rijnberk, A., ( 1996 ). New insights in the molecular mechanism of progestin-induced proliferation of mammary epithelium; induction of the local biosynthesis of growth hormone ( GH ) in the mammary gland of dogs, cats and humans. J. Steroid Biochem Mol. Biol. 57: 67-71

### **1.6 - Papéis da prolactina no desenvolvimento da glândula mamária**

Bazan, J. F., ( 1990 ). Haematopoietic receptors and helical cytokines. Immunology Today 11: 350-354

Binart, N., Ormandy, C. J. e Kelly, P. A. ( 2000 ). Mammary gland Development and the prolactin receptor. In .Biology of the Mammary Gland.

Clevenger, C. V., Chang, W. P., Ngo, W., Pasha, T. L. M. et al. ( 1995 ). Expression of prolactin and prolactin receptor in human breast carcinoma. Evidence for an autocrine/paracrine loop. Am. J. Pathol 146: 1-11

Ben-Jonathan, N., Mershon, J. L., Allen, D. L., and Steinmetz, R. W., 1996, Extrahypothalamic prolactin: distribution, regulation, functions, and clinical aspects. Endocr. Rev. 17: 639-669

Clevenger, C. V., Russel D. H., Appasamy, P. M. e Prystowsky, M. B. ( 1990 ). Regulation of IL.2 drives T-Lymphocyte proliferation by prolactin. Proc. Natl. Acad. Sci. USA, 87: 6460-6464

Jones, F. E., Jerry, D. J., Guarino, B. C., Andrews, G. C., and Stern, D. F., 1996, Heregulin induces in vivo proliferation and differentiation of mammary epithelium into secretory lobuloalveoli. Cell Growth Differ. 7: 1031-1038

Kacsoh, B., Veress, Z., Toth, B. E., Avery, L. M., and Grosvenor, C. E., 1993, Bioactive and immunoreactive variants of prolactin in milk and serum of lactating rats and their pups. J. Endocrinol 138: 243-257

Kelly, M. A., Rubinstein, M., Asa, S. L., Zhang, G., Saez, C., Bunzow, J. R., Allen, R. G., Hnasko, R., Ben-Jonathan, N., Grandy, D. K., and Low, M. J., 1997, Pituitary lactotroph hyperplasia and chronic hyperprolactinemia in dopamine D2 receptor-deficient mice. Neuron 19: 103-113

Krane, I. M. and Lader, P., 1996, NDF/heregulin induces persistence of terminal end buds and adenocarcinomas in the mammary glands of transgenic mice. Oncogene 12: 1781-1788

Lebrun, J. J., Ali, S., Goffin, V., Ullrich, A. e Kelly, P. A., ( 1995 ) a single phosphotyrosine residue of the prolactin receptor is responsible for activation of gene transcription. Proc. Natl. Acad. Sci. USA, 92: 4031-4035

Li, C. H., Dixon, J. S., Lo, T. B., Pankov, Y. M., and Schmidt, K. D., 1969, Amino acid sequence of ovine lactogenic hormone. Nature 224: 695-696

Liu, X., Robinson, G. W., Wagner, K. U., Garret, L., Wynshaw-Boris, A., and Hennighausen, L., 1997, Stat5a is mandatory for adult mammary gland development and lactogenesis. Genes Dev. 11: 179-186

Maher, D. W., Lee, B. A., and Donoghue, D. J., 1989, The alternatively spliced exon of the platelet-derived growth factor A chain encodes a nuclear targeting signal. Molecular and Cellular Biology 9: 2251-2253

Neuville, M. C. and Daniel, C. W., 1987, The Mammary Gland: Development regulation and function ( New York: Plenum Press ), pp 383-438

Nicoll, C. S., Mayer, G. L., and Russel, S. N., 1986, Structural features of prolactins and growth hormones that can be related to their biological properties. Endocr. Rev. 7: 169-203

Rakowiez-Szulezyska, E. M., Rodeck, U., Herlyn, M., and Koprowski, H., 1986 Chromatin binding of epidermal growth factor, nerve growth factor, and platelet-derived growth factor factor in cells bearing the appropriate surface receptors. Proceedings of the National Academy of Science USA 83: 3728-3732

Udy, G. B., Towers, R. P., Shell, R. G., Wilkins, R. J., Park, S. H., Ram, P. A., Waxman, D. J., and Davey, H. W., 1997, Requirement of Stat5b for sexual dimorphism of body growth rates and liver gene expression. Proc. Natl. Acad. Sci. USA 94: 7239-7244

Yang, Y., Spitzer, E., Meyer, D., Sachs, M., Niemann, C., Hartman, G., Weidner, K. M., Birchmeier, C., and Birchmeier, W., 1995, Sequential requirement of hepatocyte growth factor and neuregulin in the morphogenesis and differentiation of the mammary gland. J. Cell Biol. 131: 215-226

## **1.7 - Factores de transdução e cascatas de transdução de sinais da glândula mamária**

Beebe, S. J., 1994, The cAMP-dependent protein kinases and cAMP signal transduction. Sem. Cancer Biol. 5: 285-294

Beebe, S. J., Salmonsky, P., Jahnsen, T., and Li, Y., 1992, The C<sub>ε</sub> subunit is a unique isoenzyme of the cAMP-dependent protein kinase. J. Biol. Chem. 267: 25505-25512

Chapman, R. S., Lourenço I. P., Tonner, E., et al. ( 2000 ). The role of Stat3 in apoptosis and Mammary Gland Induction. Conditional Deletion of STAT3. In Biology of the Mammary Gland.

Feliciello, A., Rubin, C. S., Avvedimento, E. V., and Gottesman, M. E., 1998, Expression of A kinase anchor protein 121 is regulated by hormones in thyroid and testicular germ cells. J. Biol. Chem. 273: 23361-23366

Houslay, M. D., and Milligan, G., 1997, Tailoring cAMP-signalling responses through isoform multiplicity. TIBS 22: 217-224

Liu, X. W., G. W. Robinson and L. Hennighausen. 1996. Activation of Stat5a and Stat5b by tyrosine phosphorylation is tightly linked to mammary gland differentiation. Mol. Endocrinol. 10: 1496-1506

Liu, X. W., G. W. Robinson, K. U. Wagner, L. Garrett, A. WynshawBoris, and L. Hennighausen. 1997. Stat5a is mandatory for adult mammary gland development and lactogenesis. Genes & Dev. 11: 179-186

McKnigh RA, Spencer M, Dittmer J, Brady JN, Wall RJ and Hennighausen L, 1995, Mol. Endo. 9: 717-24

Meinkoth, J. L., Alberts, A. S., Went, W., Fantozzi D., Taylor, S. S., Hagiwara, M., Montminy, M., Feramisco, J. R., 1993. Signal-transduction through the cAMP-dependent protein kinase. Mol. Cell. Biochem. 128: 179-186

Pawson, T., and Scott, J. D., 1997, Signalling through scaffold, anchoring, and adaptor proteins. Science 278: 2075-2080

Phil, J. C., T. G. Burdon, and C. J. Watson. 1996. Differential activation of STATs 3 and 5 during mammary gland development. FEBS lett. 396: 77-80

Phil, J. C., T. G. Burdon, and C. J. Watson. 1996. Differential activation of STATs 3 and 5 during mammary gland development. FEBS lett. 396: 77-80

Teglund, S., C. McKay, E. Schuetz, J. M. vanDeursen, D. Stravopodis, D. M. Wang, M. Brown, S. Bodner, G. Grosveld, and J. N. Ihle. 1998. Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. Cell 93: 841-850

Thomis, D. C., Floyd-Smith, G., and Samuel, C. E., 1992, Mechanism of interferon action-cDNA structure and regulation of a novel splice-site variant of the catalytic subunit of human protein kinase A from interferon-treated cells. J. Biol. Chem. 267: 10723-10728



Welle T, Garimorth K, Philipp S, Jennewein P, Huck C, Cato ACB and Doppler W, 1994, Eur. J. Biochem. 223: 997-1006

### **1.8- Sinergias e antagonismos interactivos dos factores de transcrição na regulação da expressão dos genes das proteínas do leite**

Alam T, An MR, Mifflin RC, Hsieh C-C, Ge X, Papaconstantinou J. trans-Activation of the  $\alpha_1$ -acid glycoprotein gene acute phase responsive element by multiple isoform of C/EBP and glucocorticoid receptor. J Biol Chem 268, 15681-15688, 1993

Chen J, Sadowski HB, Kohanski RA, Wang LH. Stat5 is a physiological substrate of the insulin receptor. Proc Natl Acad Sci USA 94, 2295-2300, 1997

Dedhar, S. and Hannigan, G. E., 1996, Integrin Cytoplasmic interactions and bidirectional transmembrane signalling. Curr. Op. Cell Biol., 8: 657-669

Doppter, W, Germeyer, S. e Weinich, H.G. (2000). Synergistic and antagonistic interactions of transcription factors in the regulation of milk protein gene expression. Mechanisms of cross-talk between signalling pathway. In Biology of the Mammary Gland.

Doppler W, Welte T, Philipp S. CCAAT/enhancer-binding protein isoforms  $\alpha$  and  $\beta$  are expressed in mammary epithelial cells and bind to multiple sites in the  $\alpha$ -casein gene promoter. J Biol Chem 270, 17962-17969, 1995

Faraldo, M. M., Deugnier, M. A., Thiery, J. P., e Gkunnova, M. A. (2000). Development of mammary gland requires normal  $\alpha$ -integrin function. In Biology of the Mammary Gland.

Giancotti, F. G., 1997, Integrin signaling: specificity and control of cell survival and cell cycle progression. Curr. Op. Cell Biol., 9: 691-700

Glukhova, M., Koteliansky, V., Sastre, X. and Thiery, J. P., 1995, Adhesion systems in normal breast and in invasive breast carcinoma. Am. J. Pathol. 146: 706-716

Groner B, Altiok S, Meier V. Hormonal Regulation of Transcription Factor Activity in Mammary Epithelial Cells. Mol Cell Endocrinol 100, 109-114, 1994

Groner B, Altiok S, Meier V. Hormonal Regulation of Transcription Factor Activity in Mammary Epithelial Cells. Mol Cell Endocrinol 100, 109-114, 1994

Harris S, McClenaghan M, Simons JP, Ali S, Clark AJ. Developmental regulation of the sheep beta-lactoglobulin gene in the mammary gland of transgenic mice Dev Genet 12, 299-307, 1991

Hennighausen L, Robinson WG, Wagner KU, Liu X, Prolactin signaling in mammary gland development. J Biol Chem 272, 757-7569, 1997

Hynes NE, Taverna D, Harwerth IM, Ciardiello F, Salomon DS, Yamamota T, Groner B. Epidermal growth factor receptor, but not c-erbB-2 activation, prevents lactogenic hormone induction of the  $\alpha$ -casein gene in mouse mammary epithelial cells in culture. Mol Cell Biol 10, 4027-4034, 1990

Ip MM, Shoemaker SF, Darcy K. Regulation of rat mammary epithelial cell proliferation and differentiation by tumor necrosis factor- $\alpha$ . Endocrinology 130, 2833-2844, 1992

Lee YHP, Lee WH, Kaetzel CS, Parry G, Bissell ML. Interaction of mouse mammary epithelial cells with collagen substrata: regulation of casein gene expression and secretion. Proc Natl Acad Sci USA 82, 1419-1423, 1985

Li S, Rosen JM. Nuclear factor I and mammary gland factor (STAT5) play a critical role in regulating rat whey acidic protein gene expression in transgenic mice. Mol Cell Biol 15, 2063-2070, 1995

Liu X, Robinson GW, Wagner KU, Garret L, Wynshaw-Boris A, Hennighausen L. Stat5a is mandatory for adult mammary gland development and lactogenesis. Genes Dev 11, 179-186, 1997

Lubon H, Hennighausen L. Conserved region in the rat  $\alpha$ -lactalbumin promoter is a target site for protein binding in vitro. Biochem J 256, 391-396, 1988

Meier VS, Groner B. The nuclear factor YY1 participates in repression of the  $\alpha$ -casein gene promoter in mammary epithelial cells and is counteracted by mammary gland factor during lactogenic hormone induction. Mol Cell Biol 14, 128-137, 1994

Meredith, J. E., Winitz, S., McArthur Lewis, J., Hess, S., Ren, X.-D., Renshaw, M. W. and Schwartz, M. <sup>a</sup>, 1996, the regulation of growth and intracellular signaling by integrins. Endocrine Rev. 17: 207-220

Moriggl R, Berchtold S, Friedrich K, Standke GJ, Kammer W, Wissler M, Stocklin E, Gouilleux F, Groner B. Comparison of the transactivation domains of Stat5 and Stat6 in lymphoid cells and mammary epithelial cells. Mol Cell Biol 17, 3663-3678, 1997

Mieth M, Boehmer FK, Ball R, Groner B, Grosse R. Transforming growth factor- $\beta$  inhibits lactogenic hormone induction of  $\alpha$ -casein expression in HC11 mouse mammary epithelial cells. Growth Factors 4, 9-15, 1990

Myers CA, Schmidhauser C, Mellentinmichelotti J, Fragoso G, Roskelley CD, Casperson G, Mossi R, Pujuguet P, Hager G, Bissell MJ. Characterization of BCE-1, a transcriptional enhancer regulated by prolactin and extracellular matrix and modulated by the state of histone acetylation. Mol Cell Biol 18, 2184-2195, 1998

Philp JAC, Burdon TG, Watson CJ. Differential activation of STATs 3 and 5 during mammary gland development. FEBS Lett 396, 77-80, 1996

Raught B, Liao WSL, Rosen JM. Developmentally and hormonally regulated CCAAT/enhancer-binding protein isoforms influence beta-casein gene expression. Mol Endocrinol 9, 1223-1232, 1995

Raught B, Khursheed B, Kazanasky A, Rosen J. YY1 represses  $\alpha$ -casein gene expression by preventing the formation of a lactation-associated complex. Mol Cell Biol 14, 1752-1763, 1994

Schmidhauser C, Bissell MJ, Myers CA, Casperson GF. Extracellular matrix and hormones transcriptionally regulate bovine  $\alpha$ -casein 5' sequences stably transfected mouse mammary cells. Proc Natl Acad Sci USA 87, 9118-9122, 1990

Schmitt-Ney M, Doppler W, Ball RK, Groner B.  $\kappa$ -casein gene promoter activity is regulated by the hormone-mediated relief of transcriptional repression and a mammary-gland-specific nuclear factor. Mol Cell Biol 11, 3745-3755, 1991

Sonnenberg, A., Daams, H., van der Valk, M. A., Hiken, J. and Hilgers, G., 1986, Development of mouse mammary gland: identification of stages of differentiation of luminal and myoepithelial cells using monoclonal antibodies and polyvalent antiserum against keratin. J. Histochem. Cytochem. 34: 1037-1046

Streuli CH, Edwards GM, Delcommenne M, Whitelaw CBA, Burdon TG, Schindler C, Watson CJ, Stat5 as a target for regulation by extracellular matrix. J Biol Chem 270, 21639-21644, 1995

Streuli, C. H., Schmidhauser, C., Bailey, N., Yurchenko, P., Skubitz, A.P.N., Rockelley, C., Bissell, M. J. 1995, Laminin mediates tissue-specific gene expression in mammary epithelia. J. Cell Biol., 129: 591-603

Topper YJ, Freeman CS, Multiple hormone interactions in the developmental biology of the mammary gland. Physiol Rev. 60: 1049-1106, 1980

Varela LM, Ip MM. Tumor necrosis factor- $\alpha$ : a multifunctional regulator of mammary gland development. Endocrinology 137, 4915-4924, 1996

Wakao H, Gouilleux F, Groner B, Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. EMBO J 13, 2182-2191, 1994

Watson CJ, Burdon TG. Prolactin signal transduction mechanisms in the mammary gland: the role of the Jak/Stat pathway. Rev. Reprod 1-5, 1996

Watson CJ, Gordon KE, Robertson M, Clark AJ, Interaction of DNA-binding proteins with a milk protein gene promoter in vitro: identification of a mammary gland-specific factor. Nucl Acids Res 19, 6603-6610, 1991

Weite T, Philipp S, Cairns C, Gustafsson J-A, Doppler W. Glucocorticoid receptor binding sites in the promoter region of milk protein genes. J Steroid Biochem Molec Biol 47, 75-81, 1993

Welte T, Garimorth K, Philipp S, Doppler W. Prolactin-dependent activation of a tyrosine phosphorylated DNA binding factor in mouse mammary epithelial cells. Mol Endocrinol 8, 1091-1102, 1994

Whiteaw, C.B.A. (2000). Nucleosome organisation of the  $\kappa$ -lactoglobulin gene transcription complex formation. In .Biology in the Mammary Gland.

Whitelaw CBA. Regulation of ovine beta-lactoglobulin gene expression during the first stage of lactogenesis. Biophys Biochem Res Comm 209, 1089-1093, 1995

Winklehner-Jennewein P, Geymayer S, Lechner J, Welte T, Hansson L, Geley S, Doppler W, A distal enhancer region in the human beta casein gene mediates the response to prolactin and glucocorticoid hormones. Gene 217, 127-139, 1998

Wyszomierski SL, Yeh J, Rosen JM. Glucocorticoid receptor/signal transducer and activator of transcription 5 (STAT5) interactions enhance STAT5 activation by prolonging STAT5 DNA binding and tyrosine phosphorylation. Mol Endocrinol 13, 330-343, 1999

### **1.9 - Caspases e plasmina na biologia da glândula mamária**

Askenazi, A., and V. M. Dixit: Death receptors: signaling and modulation. Science 1305-1308 (1998)

Feng, Z., A. Marti, B. Jehn, H. J. Altermatt, G. Chicaiza, and R. Jaggi: glucocorticoid and progesterone inhibit involution and programmed cell death in the mouse mammary gland. J. Cell Biol. 131, 1095-1103 (1995)

Knight, C. H., and Wilde, C. J., 1993, Mammary cell changes during pregnancy and lactation. Livest. Prod. Sci. 35: 3-19

Marti, A., Graber, H., et al. (2000). Caspases Decoders of apoptotic signals during mammary involution. Caspases activation during involution. In Biology in the Mammary Gland.

Politis, J. (2000). The role of plasminogen activator in the bovine mammary gland. In Biology in the Mammary Gland.

Politis, I., 1995, Plasminogen activator system: implications for mammary cell growth of involution. J. Dairy Sci. 79: 1097-1107

Salvesen, G. S., and V. M. Dixit: Caspases: intracellular signaling by proteolysis. Cell 91, 443-446 (1997)

Strange, R., Li, S. Saurer, <sup>a</sup> Burkhardt, and R. R. Friis: Apoptotic cell death and tissue remodelling during mouse mammary gland involution. Development 115, 49-58 (1992)

Saksela, <sup>o</sup>, and Rifkin, D. B. 1988, Cell-associated plasminogen activation: regulation and physiological significance. Annu. Rev. Cell Biol. 4: 93-118

Thornberry, N. A., and Y. Lazebnik: Caspases: enemies within. Science 281, 1312-1316 (1998)

Topper, Y. J., and C. S. Freeman: Multiple hormone interactions in the developmental biology of the mammary gland. Physiol. Rev. 60, 1049-1106 (1980)

Walker, N. I., R. E. Bennett, and J. F. Kerr: Cell death by apoptosis during involution of the lactating breast in mice and rats. Am. J. Anat. 185, 19-32 (1989)

### **1.10 - Estrutura e funções dos genes das proteínas do leite**

Bovenhuis, H., van Arendonk, J. A. M., Korver S. 1992. Associations between Milk Protein Polymorphism and Milk Production Traits. J. Dairy Sci. 75: 2549-2559

Dove, P. 820009. Genetic polymorphisms in milk protein genes and their impact on milk composition. In ..Biology of the Mammary Gland.

Ferranti, P., Addeo, F., Malorni, A., Chianese, L., Leroux, C., Martin, P. 1997. Differential splicing of pre-messenger RNA produces multiple forms of mature caprine  $\alpha_{s1}$ -casein. Eur. J. of Biochem. 249: 1-7

Grosclaude, F. Joudrier, P., Mahe, M. F. 1978. Polymorphisme de la caseine  $\alpha_{s2}$  bovine: étroite liaisons du locus  $\alpha_{s2}$ -Cn avec les loci  $\alpha_{s1}$ -Cn,  $\beta$ -Cn et K-Cn; mise en évidence d'une délétion dans le variant  $\alpha_{s2}$ -CN D. Ann. Génét. Sél. Anim. 10: 313-327

Jakob, E. and Puhán, Z. 1992. Technological Properties of Milk as Influenced by Genetic Polymorphism of Milk Proteins. Int. Dairy Journal 2: 157-178

Lum, L. S., Dove, P., Medrano, J. F. 1997. Polymorphisms of Bovine  $\beta$ -Lactoglobulin Promoter and Differences in the Binding Affinity of Activator Protein-2 Transcription Factor. J. Dairy Sci. 80: 1398-1397

Mohr, U., Koezan, D. Linder, D., Hobom, G., Erhardt, G. 1994. A single point mutation results in A allele-specific exon skipping in the bovine  $\alpha_{s1}$ -casein mRNA. Gene 143: 187-192

Lunden, <sup>a</sup>, Nilson, M., Janson, L. 1997. Marked Effect of  $\beta$ -Lactoglobulin Polymorphism on the Ratio of Casein to Total Protein in Milk. J. Dairy Sci. 80: 2996-3005

Mercier, J.C. and Vilotte, J.L. 1993. Structure and function of milk protein genes. J. Dairy Sci. 76:3079-3098

Ng-Kwai-Hang, K. F., Hajes, J. F., Moxley, J. E., Monardes, H. G. 1986. Relationships between Milk Protein Polymorphism and Major Milk Constituents in Holstein- Friesian Cows. J. Dairy Sci. 69: 22-26

Passey, R., Glenn, W., Mackinlay, A. 1996. Exon skipping in the ovine  $\alpha_{s1}$ -casein gene. Comp. Biochem. Physiol. 114B: 389-394

Perez, M. J., Leroux, C., Bonastre, A. S., Martin, P. 1994. Occurrence of a LINE sequence in the 3' UTR of the goat alpha s1-casein E-encoding allele associated with reduced protein synthesis level. Gene 147: 179-187

Rijnkels, M., Kooiman, P. M., de Boer, H. A., Pieper, F. R. 1997. Organization of the bovine casein gene locus. Mammalian Genome 8: 148-152

Ron, M., Yoffe, <sup>o</sup>, Ezra, E., Medrano, J. F., Weller, I. J. 1994. Determination of effects of milk protein genotype on production traits of Israeli Holsteins. J. Dairy Sci. 77: 1106-1113

### **1.11 - Fracções proteicas do leite e seus polimorfismos genéticos**

Bovennuis, H. van Arendonk, J. A. M. e Korver, S. (1992) associations between milk protein polymorphisms and milk production traits. J. Dairy Sci. 75: 2544-2559

Ikonen, T., Ahlfors, K., Kemp, R., Ojala, M., e Ruoltinen, ° (1999). Genetic parametro for the milk coagulation properties and prevalence of non coagulating milk in finish dairy cows. J. Dairy Sci. 82: 205-214

Mercier, J. C., e Vilolte, J. L. (1993). Structur and function of milk protein genes. J. Dairy Sci. 76: 3079-3098

Palmquist, D. C., Beaulieu, A. D. e Barbano, D. M. (1993). ADSA foundation Symposium: milk fat synthesis and Modification. Feed and animal factors influencing milk far composition. J. Dairy Sci. 76: 1753-1771

Perez, M.J., Leroux, C., Bonastre, A.S. Martin, P. (1994). Occurence of a LINE sequence in the 3'UTR of the goat alpha s-1 casein E-encoding allele associated with reduced protein synthesis level. Gene 147:179-180

Sutton, J. D. (1989). Altering milk composition by feeding J. Dairy Sci 72: 2801-2814

Van Feenennaam, A. e Medrano, J. F. (1991). Milk protein polymorphisms in california dairy cattle. J. Dairy Sci. 74: 1730-1742

Walker, N. I., R. E. Bennett, and J. F. Kerr: Cell death by apoptosis during involution of the lactating breast in mice and rats. Am. J. Anat. 185, 19-32 (1989)

### **1.12 - Regulação da fraccão lipídica do leite**

Barber, M. C., Clegg, R. A., Travers, M. T., and Vernon, R. G., 1997, Lipid metabolism in the lactating mammary gland. Biochim. Biophys. Acta. 1347: 101-126

Baumgard, L. H., Corl, B. A., Dwyer, D. A., S\_bp, A., and Bauman, D. E., 1999, Identification of the conjugated linoleic acid isomer which inhibits milk fat synthesis. Am. J. Physiol. (in press)

Bauman, D. E., Griinar, J. M. (2000)- Regulation and nutritional manipulation of milk fat Low-fat milk syndrome. In Biology of the Mammary Gland.

Davis. C. L., and Brown, R. E., 1970, Low-fat Milk Syndrome. In Physiology of Digestion and Metabolism in the Ruminant (A.T. Phillipson, ed.), Oriel Press Limited, Newcastle upon Tyne, UK. Pp. 545-565

Erdman, R., 1996, Milk fat depression: some new insights. Tri-State Dairy Nutrition Conference, Fort Wayne, IN, pp. 1-16

German, J. B., Morand, L., Dillard, C. J., and Xu, R., 1997, Milk Fat Composition: Targets for Alteration of Function and Nutrition. In Milk Composition and Biotechnology (R.A.S. Welch, D.J.W. Burns, S. R. Davis, A. I. Popay, and C. G. Prosser, eds.), CAB international, New York, NY, pp. 35-72

Griinari, J. M., Dwyer, D. A., McGuire, M. A., Bauman, D. E., Palmquist, D. L., and Nurmela, K.V.V., 1998, Trans-octadecenoic acids and milk fat depression in lactating dairy cows. J. Dairy Sci. 81: 1251-1261

Griinari, J. M., and Bauman, D. E., 1999, Biosynthesis of CLA and incorporation into Milk Fat. In: Advances in Conjugated Linoleic Acid Research, Vol. I (M.P. Yurawecz, M.M. Mossoba, J.K.G. Kramer, M.W. Pariza, and G.J. Nelson, eds.), AOCS Press, Champaign, IL, pp. 180-200

Griinari, J. M., Nurmela, K., Dwyer, D. A., Barbano, D. M. and Bauman, D. E., 1999, Variation of milk fat concentration of conjugated linoleic acid and milk fat percentage is associated with a change in ruminal biohydrogenation. J. Anim. Sci. 77 (Suppl. 1): 117-118 (abstr.)

Palmquist, D. L., Beaulieu, A. D., and Barbano, D. M., 1993, Feed and animal factors influencing milk fat composition. J. Dairy Sci. 76: 1753-1771

Griinari, J. M., Nurmela, K., Dwyer, D. A., Barbano, D. M. and Bauman, D. E., 1999, Variation of milk fat concentration of conjugated linoleic acid and milk fat percentage is associated with a change in ruminal biohydrogenation. J. Anim. Sci. 77 (Suppl. 1): 117-118 (abstr.)

Sutton, J. D., 1989, Altering milk composition by feeding. J. Dairy Sci. 72 2801-2814

<b>2. - <u>Miogénese, músculos do esqueleto e produção animal</u></b>	618
2.1 – Miogénese	618
A - Embriogénese, interacção célula-célula e célula- matriz extracelular e factores _ de transformação do crescimento	618
B - Genes miogénicos	621
C - Génese e regeneração dos músculos do esqueleto	624
D - Hipertrofia muscular bovina	629
E - Ratos transgénicos sem miostatina	632
F - Expressão da miostatina	633
G - Regulação da miostatina e crescimento muscular	634
H - Hipertrofia muscular ovina	635
I - Factor de crescimento do tecido conectivo	635
J - Hipertrofia das células dos músculos do esqueleto induzidas por inibidores de metaloproteases: miostatina como um potencial mediador	636
L - Comparação da GH, IGF1 e testosterona com os mRNA de receptores e da miostatina em músculos do esqueleto de homens idosos	636
2.2 - Factores que afectam o crescimento e desenvolvimento dos músculos do esqueleto	637
Insulina,hormona do crescimento,hormonas da tiróide e outras hormonas	637
Esteróides sexuais naturais e produção animal	638
Efeitos do exercício e nutrientes	639
Alguns outros efeitos dos agentes anabolizantes	640
2.3 - Endocrinologia dos peixes e produção animal	643
2.4 – Bibliografia	647



## 2. - Miogénese, músculos do esqueleto e produção animal

### 2.1 - Miogénese

- A - Embriogénese e interacção célula-célula e célula matriz-extracelular, factores \_ de transformação do crescimento
- B - Genes miogénicos
- C - Génese e regeneração dos músculos do esqueleto
- D - Hipertrofia muscular bovina
- E - Ratos transgénicos sem miostatina
- F - Expressão da miostatina
- G - Regulação da miostatina e crescimento muscular
- H - Hipertrofia muscular ovina
- I - Factor de crescimento do tecido conectivo
- J - Inibidores de metaloproteases e miostatina
- L - GH, IGF1, testosterona, miostatina

#### A - Embriogénese e interacção célula-célula e célula matriz –extracelular, factores \_ de transformação do crescimento (TGF )

Durante o desenvolvimento embrionário dos animais superiores as interacções entre as células desempenham um papel fundamental determinando a organização das células e dos tecidos com uma dada população de células induzindo a formação das células vizinhas.

Neste papel indutivo a superfamília TGF\_ de glicoproteínas segregadas (citocinas, factores de crescimento e transformação) e respectivos receptores que são cinases serina-treonina transmembranárias, desempenha um papel amplo quer nos vertebrados quer nos invertebrados.

Estas TGF\_ controlam as diferentes modelações das células ao longo do eixo dorso ventral, induzindo os tecidos da mesoderme e induzindo diferentes modelações consoante a sua concentração.

Nos mecanismos celulares do desenvolvimento a endoderme dá origem ao intestino e órgãos associados tais como os pulmões e fígado.

A mesoderme origina os tecidos conectivos, músculos, e sistemas vascular e urogenital.

A ectoderme origina a epiderme e o sistema nervoso.

O tubo neural é formado por alterações coordenadas no formato celular.

Blocos de células mesodérmicas desacoplam-se para formar somitos em qualquer lado do eixo do corpo.

Recorda-se que durante o desenvolvimento o genoma permanece constante mas o perfil da expressão dos genes modifica-se.

Durante a embriogénese as interacções locais entre as células e entre estas e a matriz extracelular regulam o desenvolvimento dos organismos multicelulares, sendo o desenvolvimento embriológico

dirigido pela indução ou seja o mecanismo através do qual uma dada população celular influencia o desenvolvimento das células vizinhas.

Os factores \_ de transformação do crescimento (TGF\_) têm numerosos efeitos indutivos nos vertebrados e invertebrados e constituem uma superfamília que engloba uma série de moléculas assinalantes extracelulares relacionadas.

Alguns membros desta superfamília têm efeitos antiproliferativos em certos tecidos e efeitos proliferativos sobre outros tecidos. Estes factores também promovem a produção de moléculas de adesão celular, outros factores de crescimento e moléculas da matriz extracelular.

Na figura esquemática seguinte mostra-se que a forma activa madura destes factores contem 110-140 ácidos aminados que resultam por proteólise da região C-terminal da proteína precursora.

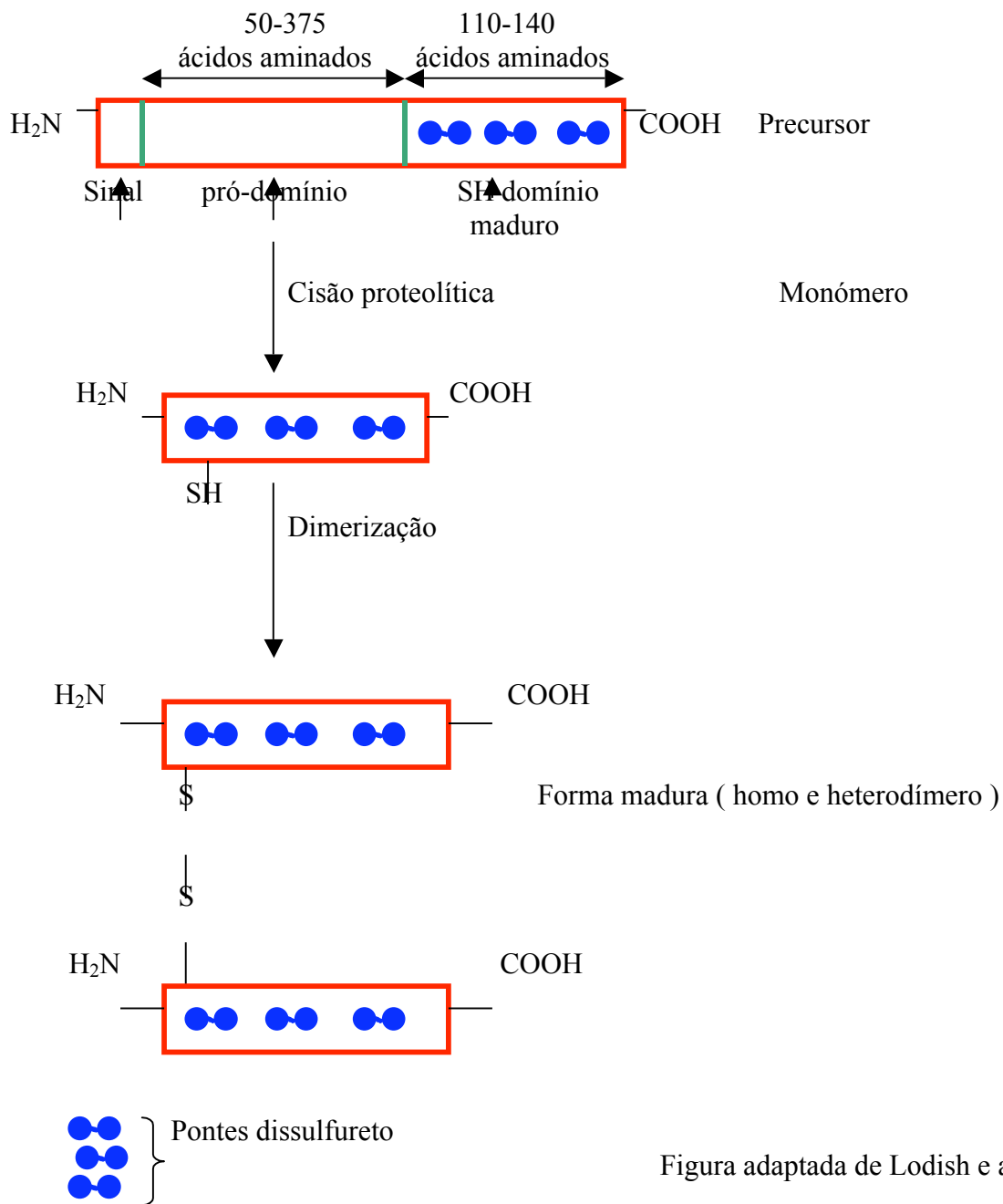
Este precursor, além do péptido sinal N-terminal contem um pródomínio de 50-375 ácidos aminados.

Na estrutura do precursor ocorrem três pontes dissulfureto, e uma cisteína N-terminal em cada monómero o que permite a ligação de dois monómeros de TGF\_ para formar um dímero, homo ou heterodímero, aumentando esses últimos a diversidade funcional destas proteínas.

Entre as diferentes proteínas desta superfamília TGF\_ encontram-se grandes diferenças de sequências dos ácidos aminados nas regiões N-terminais.

Existem à superfície das células, receptores para os TGF\_ dos tipos I, II e III, sendo os dois primeiros cinases serina/treonina transmembranárias sendo a transdução de sinal efectuada de forma ainda desconhecida.

Formação de formas maduras da superfamília TGF<sub>β</sub> de factores de crescimento



A função destes receptores TGF<sub>β</sub> pode ser modificada de diversas formas, por exemplo por diversos tipos de “splicing” ou dimerização.

Também proteínas solúveis ou componentes da matriz extracelular podem ligar-se ao TGF<sub>β</sub> ligando, e sequestrar este numa forma que não lhe permite interaccionar com os receptores.

As interacções indutivas entre as células epiteliais e mesenquimatosas na regulação da formação de órgãos internos como os rins, pulmões, pâncreas e intestino, estão bem esclarecidas parecendo que os contactos directos célula com célula e as interacções célula-matriz estão ambas implicadas na regulação dessa indução.

Estas interações controlam não só a especificação de cada tipo celular mas também a modelação dos tecidos.

Uma série de acontecimentos indutivos regulam o desenvolvimento precoce do embrião, havendo uma sequência determinada destes efeitos indutivos. Na formação de órgãos internos e na organização dos tecidos, a lâmina basal é essencial para a diferenciação de diversas células epiteliais, tal como o contacto directo célula-célula regula a indução do tecido renal, e tal como a proteína indutiva chamada Hedgehog organiza a modulação por exemplo da perna de frango ou da asa da *Drosophila*.

Os organismos multicelulares consistem de diferentes tipos de células exibindo diferentes funções e morfologias, tudo dependendo da expressão de conjuntos dos seus genes específicos.

O desenvolvimento dos músculos (miogénese) nos mamíferos tal como a neurogénese na *Drosophila* revelou o papel dos factores de transcrição do tipo HLH (hélice - ansa-hélice) na especificação de ambos os tipos celulares.

O desenvolvimento dos mioblastos (células precursoras das células musculares) depende das Myo D e Myf5 (expressos por genes para proteínas específicas do músculo) que são proteínas miogénicas todas membros da família HLH factores de transcrição que se ligam ao DNA.

Um aspecto importante do desenvolvimento embrionário animal é o processo de regionalização controlado por quatro conjuntos de genes diferentes; três regulam a modelagem ao longo do eixo anterior/posterior, e um regula a modelagem ao longo do eixo dorsoventral. O sistema anterior controla o desenvolvimento da cabeça e tórax. O estabelecimento de estados diferentes das células na região anterior pelo menos no embrião da *Drosophila* depende de forma crítica do factor de transcrição bicóide, localizando-se o mRNA bicóide maternal no extremo anterior do embrião precoce. Após translação a proteína bicóide difunde-se para o extremo posterior do embrião produzindo um gradiente de concentrações, que origina outros gradientes de proteínas que determinam a expressão de domínios de vários outros genes.

A proteína Nanos derivada maternal está localizada na região posterior do embrião precoce, e funciona como determinante nesta região, inclusive inibindo a translação de proteínas da região anterior.

A determinação de diferentes evoluções celulares ao longo do eixo dorso-ventral utiliza outros meios para regulação da expressão dos genes, como por exemplo o toll mRNA.

O papel de genes selectores na determinação da identidade de certas regiões (por exemplo a localização das estruturas adultas) tem sido bastante estudado ao longo do eixo anterior-posterior nos mamíferos. Estes genes que ocorrem em complexos de “clusters” codificam factores de transcrição contendo domínios “homeobox”, sendo nos mamíferos estes genes contíguos, formando um “cluster” chamados Hox.

Os mamíferos têm quatro “clusters” Hox localizados em quatro cromossomas diferentes.

## **B - Genes miogénicos**

Os diferentes tipos de células que compõem os animais realizam uma grande diversidade de funções e morfologias, devido a diferenças qualitativas e quantitativas na expressão dos seus genes específicos.

No entanto ainda não é possível para os organismos multicelulares referir qual o conjunto completo de moléculas reguladoras que caracteriza este ou aquele tecido ou que é responsável pelas diferenças entre dois dados tipos de células, apesar dos consideráveis progressos realizados no âmbito das proteínas que se ligam ao DNA e controlam a transcrição de genes específicos de um dado tipo de células.

Parece no entanto que a expressão dos genes correspondentes a um tipo específico de células reflecte a actividade de factores de transcrição específicos ou de uma combinação única deles.

O desenvolvimento dos músculos do esqueleto nos mamíferos é um bom sistema para estudar o papel dos factores de transcrição no controlo específico de cada tipo de células.

As células musculares esqueléticas são fibras alongadas multinucleadas, podendo a miogénese nos mamíferos ser dividida em três etapas: a determinação mioblástica, a migração e a diferenciação em tecido muscular, sendo os mioblastos provenientes dos somitos, as suas precursoras.

Os somitos são blocos de células mesodérmicas localizadas no embrião ao lado do tubo neural.

Os mioblastos não estão portanto ainda diferenciados e migram dos somitos para as regiões onde se formarão músculos, onde formam sincícios diferenciando-se depois. A diferenciação é um processo em que estão implicadas mudanças na expressão dos genes e que levam a que uma célula precursora origine ou se transforme numa célula especializada distinta.

O gene mio D (myogenic determination gene D) está implicado no desenvolvimento muscular onde desempenha um papel chave, estando identificados ainda três outros genes, a miogenina, ou myf4, o myf5 e o mrf4 que também funcionam no desenvolvimento muscular. O myf3 é o homólogo humano do myo D1 de rato. O Myo D ou factor miogénico 3, ou Myf3 é o antigénio 1 de diferenciação miogénica ou Myo D1, que são símbolos de designações alternativas descritas com o número 159970 OMIM.

Estas quatro proteínas miogénicas – Myo D, Myf5, miogenina e Mrf4 são todas membros da família dos factores de transcrição que se ligam ao DNA, através de estruturas com HLH (helix-loop-helix).

Estas proteínas HLH formam homo e heterodímeros estes com maior afinidade que ligam ao DNA numa mesma sequência CANNTG chamada E-BOX e que se encontra em múltiplas cópias (calcula-se que 256 em 256 nucleótidos) na maioria dos “enhancers” específicos do músculo.

A Myo D pode ligar-se apenas a uma simples E-BOX mas ela apenas é activadora quando se liga a dois ou mais locais.

A afinidade do Myo D para o DNA é muito maior como já se disse se for na forma de heterodímero complexada com uma proteína E2A, proteína esta que é também uma proteína HLH que se exprime em muitas células e é necessária também para a miogénese.

Os domínios do E2A e da Myo D que se ligam ao DNA são similares mas não idênticos e ambos reconhecem a sequência E-BOX mas a activação do gene miogénico depende de ácidos aminados específicos na Myo D.

Esta ainda descrito que uma proteína Id (inibidor do DNA) pode inibir a actividade da Myo D ao ligar-se com a Myo D e E2A impedindo a formação de dímeros e a ligação ao DNA.

A expressão de qualquer uma das proteínas miogénicas pode induzir a diferenciação de células em músculo, *in vitro*. Tal como as proteínas leucina zipper, as proteínas HLH têm papéis muito importantes na diferenciação e desenvolvimento celular, sendo a sua actividade modulada pela formação de heterodímeros como dissemos

Também está descrita uma proteína celular que contém uma região HLH e que aparentemente actua como um regulador negativo formando heterodímeros inactivos com os Myo D.

A região básica de ligação ao DNA (13 ácidos aminados) é necessária para reconhecimento do local de ligação consenso do Myo, para activação da transcrição e para a conversão dos fibroblastos em músculo.

Pelo contrário, a substituição de três resíduos de ácidos aminados na bHLH da proteína E12 (um dos diferentes produtos de splicing do gene E2A não específico de tecidos) pode permitir a ligação ao local de união Myo D mas não induzir a miogénese.

A expressão da Myo D num grande número de tipos celulares leva à conversão destes em músculo. Apenas 68 ácidos aminados do Myo D constituem o domínio bHLH, e são necessários e suficientes para a conversão miogénica.

Como dissemos o Myo D é um factor de transcrição específica dos músculos que se liga à maioria das sequências “enhancer” específicas do músculo, através dos motivos específicos CANNTG (N = subconjuntos específicos de nucleótidos) para activar a expressão dos genes específicos do músculo.

A ligação ao DNA é no entanto um acontecimento separável da activação da transcrição subsequente, embora a região básica do Myo D seja necessária para as duas coisas.

A análise da sequência e o perfil de expressão do MYF5 em músculo humano revelou que esta proteína apesar de relacionada estruturalmente com a Myo D1, são proteínas diferentes, embora ambas induzam fenótipo miogénico em fibroblastos C3H de ratos embrionários.

No rato, os genes miogénicos básicos HLH são activados sequencialmente durante o desenvolvimento muscular, o mRNA para o Myf5 é detectado primeiro nos somitos com 8 dias sendo depois marcadamente reduzido aos 14 dias.

Ratos com mutantes dos genes Myf5 e Myo D cruzados, não possuindo estes genes, nascem imóveis e morrem rapidamente com completa ausência histológica dos músculos do esqueleto o que parece indicar que quer o Myf5 quer o Myo D são necessários para a determinação dos mioblastos do esqueleto e sua propagação, durante o desenvolvimento embrionário, sendo importantíssimos na miogénese.

Nos bovinos o Myf5 está situado no cromossoma 12 q 21 e a sua proteína expressa possui 255 ácidos com estrutura primária conhecida revelando o domínio compreendido entre os resíduos 96 e 135 (40AA) a componente HLH.

O Myo D1 (miogenina) no rato *Mus musculus*, têm 224 ácidos aminados com estrutura primária conhecida com o domínio HLH entre os resíduos 94 e 133 (40AA).

O Myo D1 nos ovinos têm 319 ácidos aminados com o domínio HLH entre os resíduos 122 e 161 (40AA).

O Myo D1 nos suínos tem 319 ácidos aminados com o domínio HLH entre os resíduos 122 e 161 (40AA).

As estruturas primárias do Myo D1 de ovinos e suínos é quase inteiramente idêntica (homologia).

O Myf5 bovino tem um domínio básico entre os resíduos 83 e 95 que se liga ao DNA enquanto o domínio entre os resíduos 96 e 135 é um motivo HLH (por similaridade).

### **C - Gênese e regeneração dos músculos do esqueleto**

Os músculos englobam uma série de tipos celulares, todos eles especializados para a contração mas possuindo também funções diferentes uns dos outros.

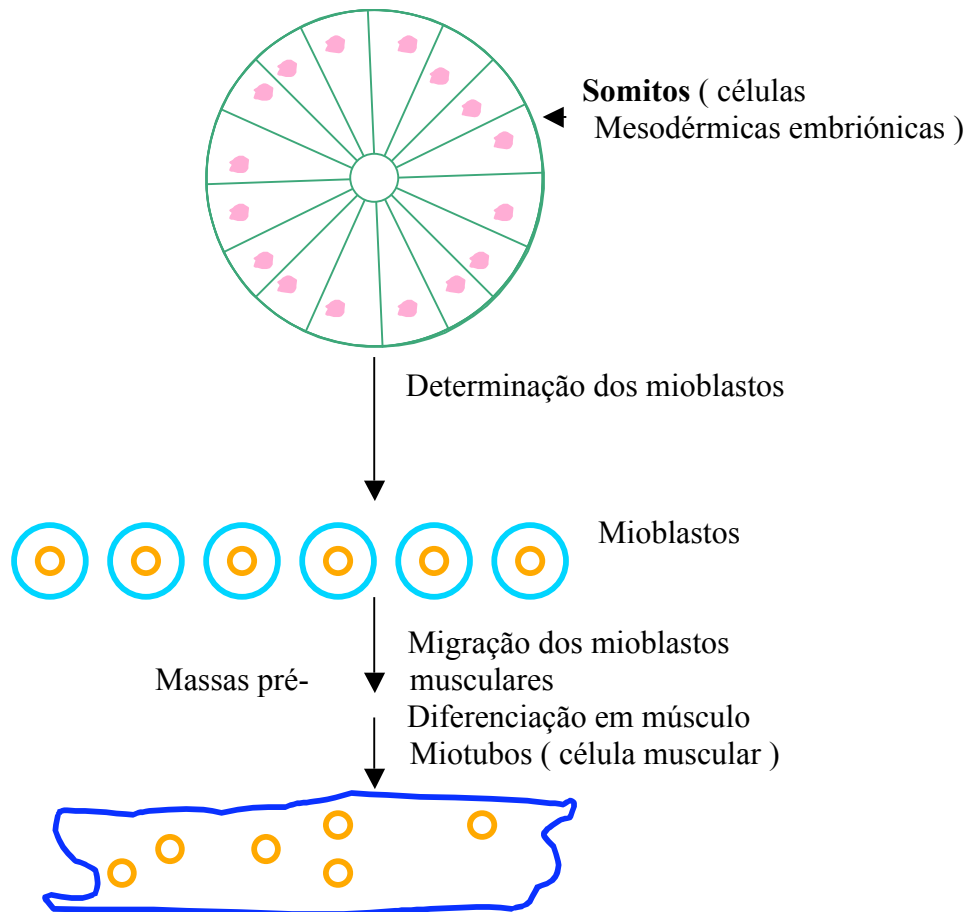
Um sistema contráctil envolve actina e miosina como componentes básicos das células animais.

Os mamíferos possuem quatro categorias principais de células especializadas em contração e que são as células dos músculos do esqueleto, dos músculos cardíacos, dos músculos lisos e as células mioepiteliais, que diferem na função, estrutura e desenvolvimento. Todas têm sistemas organizados em filamentos baseados na actina e na miosina, embora estas moléculas sejam um tanto diferentes na sua estrutura, arranjo e associação com outros tipos de proteína, consoante os tipos de células.

As células de fibras musculares dos músculos do esqueleto, que são sincícios contendo vários núcleos dentro do mesmo citoplasma (os outros tipos de células musculares têm apenas um simples núcleo) diferem um tanto das células do músculo cardíaco, das células dos músculos lisos (não estriadas) e das células mioepiteliais (não estriadas, e ao contrário das outras células musculares, têm sede no epitélio e derivam da ectoderme).

As novas células dos músculos do esqueleto formam-se por fusão dos mioblastos que são portanto precursores das células dos músculos do esqueleto e que são originárias dos somitos do embrião vertebrado (vide figuras anexas). As proliferações dos mioblastos são estimuladas por factores de crescimento como o FGF, mas uma vez fundidos não se dividem mais. É sabido que uma simples proteína reguladora de um gene chave, como é o caso da proteína Myo D pode converter um fibroblasto num mioblasto . Após um período de proliferação os mioblastos fundem-se uns com os outros para formar células dos músculos do esqueleto multinucleadas sofrendo à medida que ocorre essa fusão uma marcada mudança do seu fenótipo o que depende da activação coordenada de um conjunto de genes específicos do músculo.

Diagrama esquemático das três etapas de desenvolvimento dos músculos do esqueleto dos mamíferos

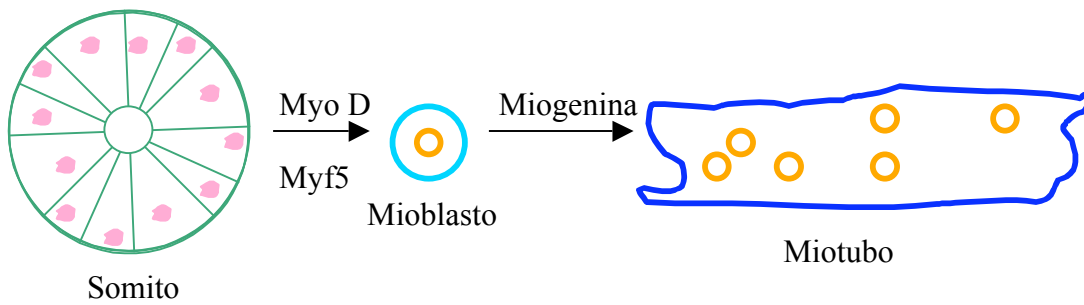


Quando o crescimento dos mioblastos (células precursoras musculares) em culturas é parado, as células fundem-se para produzir miotubos com as características estriações transversais das células musculares diferenciadas.

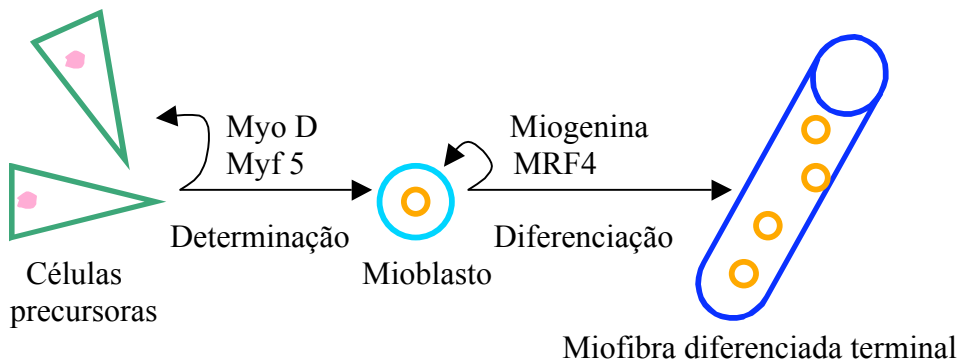
Os miotubos assemelham-se a células musculares diferenciadas sintetizando várias se não todas as proteínas especializadas implicadas na contração.



## Controlo genético dos músculos do esqueleto dos mamíferos



## Papel de membros da família Myo D na miogénese dos músculos do esqueleto



As células dos músculos dos esqueletos mamíferos são pois formadas pela fusão de diversos mioblastos. Uma célula muscular madura distingue-se das outras células pelo seu grande número de proteínas características (tipos específicos de actina, miosina, tropomiosina, troponina, creatinina fosfocinase, receptores de acetilcolina etc.).

Nos mioblastos proliferantes estas proteínas específicas dos músculos e os respectivos mRNA não existem, ou apenas se encontram em muito baixas concentrações. Mas à medida que os mioblastos se começam a fundir uns com os outros a produção de todas estas populações moleculares aumenta até cerca de 500 vezes.

A Myo D, normalmente é apenas expressa nos mioblastos e nas células dos músculos do esqueleto e tem capacidade para subverter o controlo normal dos genes dos fibroblastos e converter estes numa célula muscular.

A proteína Myo D, encontra-se concentrada no núcleo celular e tem capacidade para se ligar a regiões reguladoras de diversos genes específicos dos músculos.

Após um período de proliferação os mioblastos fundem-se uns com os outros para formar as células multinucleadas dos músculos do esqueleto.

Após a fusão os núcleos nunca mais replicam o seu DNA. Esta fusão implica o reconhecimento mútuo e específico entre os mioblastos, pois eles não se fundem com células não musculares adjacentes.

Também é conhecido que as células musculares podem modular as suas propriedades por mudança das isoformas das suas proteínas.

Uma célula do músculo do esqueleto, uma vez formada subsiste depois durante toda a vida do animal, mas durante este período, cresce, amadurece e modela as suas características consoante as exigências das respectivas funções.

Uma série de variantes proteicas (chamadas isoformas) podem ser produzidas dentro dos componentes do aparelho contráctil ,para este efeito, através de diversos tipos de “splicing”, produzindo vários tipos de células dos músculos do esqueleto (como por exemplo vermelhas e brancas). A expressão global dos genes dos músculos vermelhos do esqueleto difere da dos músculos brancos do esqueleto.

O estudo em ratos da expressão diferencial de cerca de 3.000 mRNA entre estes dois tipos de músculos, permitiu verificar que as sequências de 49mRNAs são diferentemente expressas entre os músculos vermelhos e brancos.

Isto leva a admitir uma complexa regulação do fenótipo entre os diferentes tipos de fibras dos músculos.

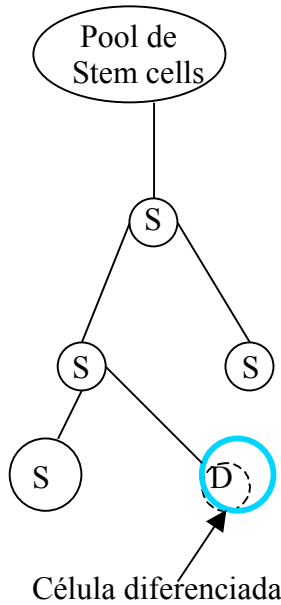
A expressão dos genes das células musculares pode ser influenciada artificialmente com relativa facilidade.

No entanto nos adultos alguns mioblastos persistem como células estaminais (stem cells) quiescentes.

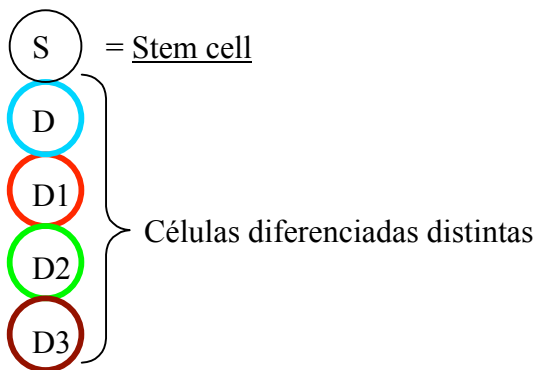
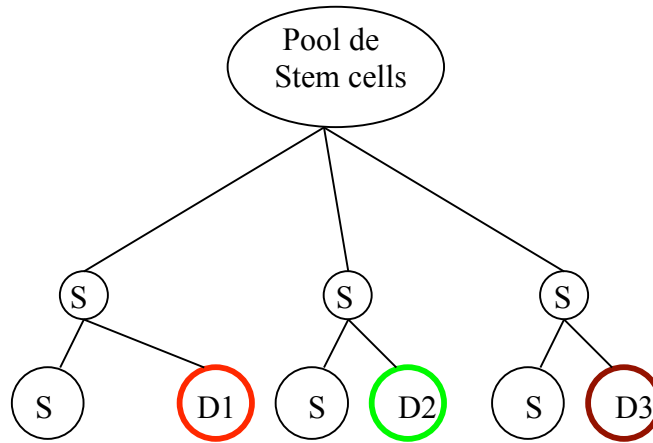
Stem cell é uma célula que se renova a si própria, que se divide para dar uma célula com um potencial de desenvolvimento idêntico e/ou uma outra com um potencial de desenvolvimento mais limitado.

Produção de células diferenciadas a partir de stem cells ( S )

Célula unipotencial



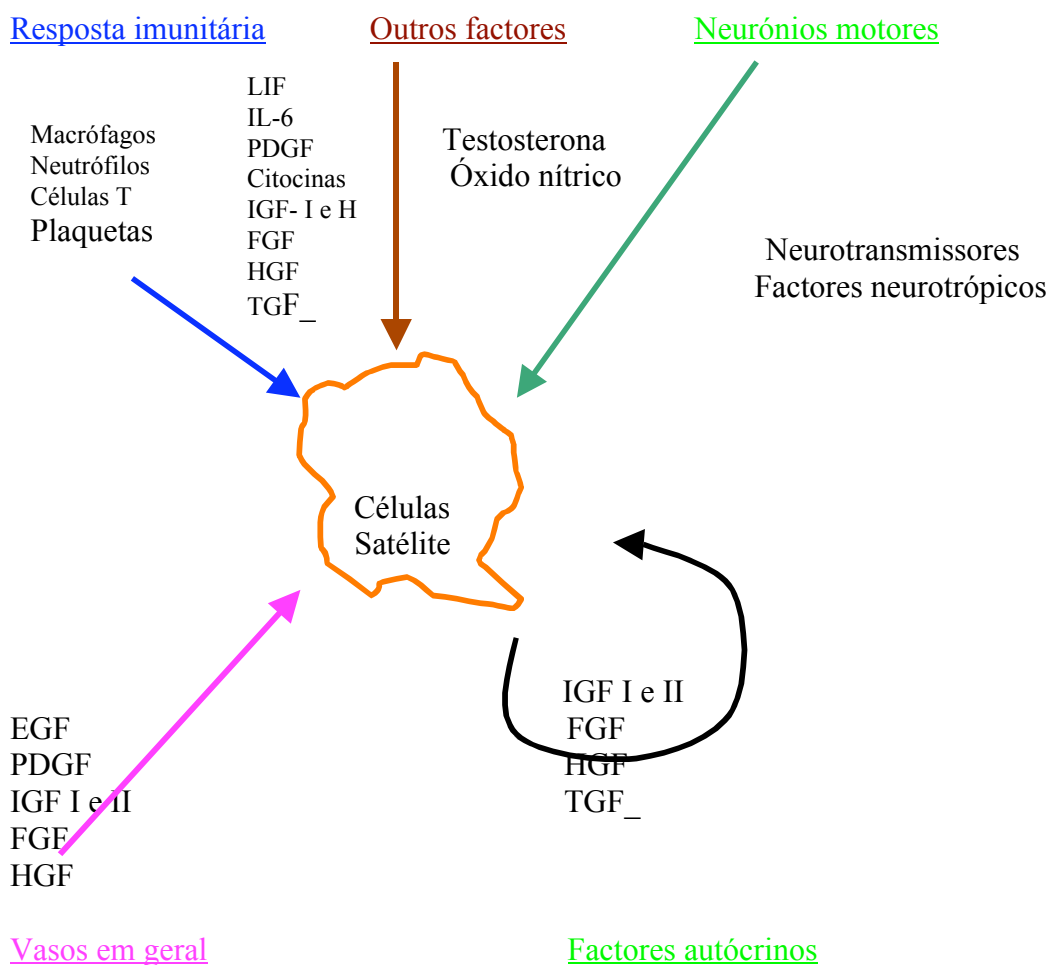
Célula pluripotencial



Recorda-se que as células dos músculos do esqueleto são incapazes de se dividir após a fusão da mioblastos, e no adulto o número de células polinucleadas dos músculos do esqueleto é atingido muito cedo, (nos humanos antes do nascimento) sendo o subsequente aumento enorme do volume muscular alcançado pelo alargamento de cada célula. O aumento de comprimento depende do recrutamento de mais mioblastos para fusão das suas extremidades com as células multinucleadas existentes aumentando também o número de núcleos. O aumento transversal ou do perímetro transversal do músculo depende do aumento do tamanho e do número de miofibrilhas contrácteis que cada célula muscular contém mais do que da alteração no número de células musculares ou de número dos seus núcleos.

Contudo nos adultos persistem alguns mioblastos em contacto com as células musculares maduras e essas células satélites quiescentes e em reserva digamos assim podem ser activadas para proliferação e a sua descendência fundir-se e formar novas células musculares.

## Factores que modelam a actividade das células satélite quiescentes



### **D - Hipertrofia muscular bovina**

Esta hipertrofia muscular (mh) corresponde ao que designamos por “garupa dupla” e que ocorre com elevada frequência nas raças de bovinos Belgian Blue e Piedmontese, corresponde a um locus mh autosomal recessivo mapeado no cromossoma dois destes bovinos.

A hipertrofia muscular descrita nos bovinos e ovinos consiste num aumento anormal de tecido muscular originado inteiramente por um alargamento das células musculares existentes e não pela hiperplasia muscular na qual o aumento do tecido muscular é devido á formação e crescimento de novas células musculares normais.

Contudo há referências que apontam para que a “garupa dupla” nos bovinos é caracterizada por uma hiperplasia muscular com um aumento do número de fibras musculares mais do que pelo aumento do seu diâmetro individual (<http://www.ncbi.nlm.nih.gov/htbin-post/omim/dispim?601788>).

Está demonstrado que uma mutação no gene bovino MSTN (Mp locus) que codifica a miostatina é responsável pelo fenótipo da “garupa dupla”. Grobet et al (1997. Nature Genet 17 : 71-74) verificam que uma deleção de 11bp na sequência codificante para o domínio bioactivo C-terminal desta proteína origina a hipertrofia muscular.

Nos bovinos de “garupa dupla” (ou hipertrofia muscular Mh) ocorre um aumento da massa muscular de cerca de 20%, conseqüentemente com maior produção de carne magra e mais tenra.

Esta característica é autossomal recessiva e o respectivo locus cromossomal é designado por mh ocorrendo com alta frequência nos bovinos Piedmontese e Azul belga, e ainda noutras raças, e apesar das suas vantagens óbvias obriga a intervenções cesarianas frequentes como regra, quando do nascimento das crias, além de outras desvantagens.

O locus mh em cruzamento Belga azul x Frísia localiza-se na terminação centromérica do cromossoma 2 cerca de 3,1cM do microsátélite TCLA44.

O gene desta miostatina é um membro da superfamília de péptidos factores de crescimento transformante beta1 (TGF\_1).

Em ratos “knocked out” para este gene da miostatina foi observado um fenótipo muito semelhante ao da “garupa dupla”.

A sequenciação do DNA da miostatina em animais normais homozigotos e em bovinos de “garupa dupla” revelou uma deleção de 11bp (dos nucleótidos 821 a 831) originando uma “frameshift” (ou seja uma mutação que leva ao aparecimento de um ácido aminado não relacionado e que geralmente pára o codão) com a subseqüente terminação prematura no domínio carboxil-terminal bioactivo.

Esta é uma região que se encontra altamente conservada nesta família de péptidos TGF.

Também a mesma mutação está assinalada na raça Astuniana de “garupa dupla”.

Mas em bovinos de garupa dupla Piedmonteses, está assinalada, pelo contrário, uma transição G-A que converte um resíduo de cisteína em tirosina na mesma região altamente conservada do gene.

Numa triagem de 35 bovinos de “garupa dupla” de dez raças europeias, foram encontradas sete seqüências variantes diferentes (alelos) na região codificadora do gene da miostatina, podendo cinco delas originar uma deficiência de miostatina, como se refere seguidamente:

- a) a deleção de 11bp
- b) a inserção/deleção na qual dez bases não relacionadas são inseridas no lugar de sete bases que foram deleccionadas no nucleótido 418
- c) a transição C-T na nucleótido 610
- d) a transição G-T no nucleótido 676
- e) a transição G-A no nucleótido 938

A maioria das raças de animais com garupa dupla são homozigotos para uma das cinco mutações perigosas, ou são heterozigotos compostos de duas mutantes.

Há uma considerável heterogeneidade genética nos animais de “garupa dupla”, e as mutações não são todas únicas para uma raça, duas sendo possuídas por mais do que uma raça.

Por outro lado, duas raças a Limousin e a Blond d’Aquitane têm “garupa dupla” mas não têm qualquer das cinco mutações perigosas antes referidas.

As mutações no gene da miostatina têm pois profundas implicações na produção de carne sugerindo novos caminhos para esta.

Parece que a expressão molecular da miostatina e da Myo D é maior nos fetos bovinos de “garupa dupla” do que nos animais de musculatura normal segundo Oldham, J. M.

A miostatina nos bovinos ou factor 8 de crescimento/diferenciação (precursor expresso) pelo gene GDF8 ou MSTN ou MH actua especificamente como um regulador negativo do crescimento do músculo esquelético, sendo um homodímero (ligado por dissulfuretos por similaridade) e é especificamente expresso nos músculos esqueléticos em desenvolvimento e nos adultos.

As mais elevadas concentrações aparecem nos músculos semimembranosos e biceps femoris dos membros posteriores, encontrando-se mais baixas concentrações noutros músculos destes membros.

Defeitos no gene GDF8 originam o fenótipo da garupa dupla, uma situação autossómica recessiva frequente no gado bovino Azul Belga e Piedmontese sendo caracterizada, segundo alguns autores por um aumento do número de fibras musculares (hiperplasia), o que origina um aumento de massa muscular de 20-25%.

A miostatina pertence à família do TGF-Beta sendo portanto uma glicoproteína sinal, factor de crescimento englobável nas citocinas.

Esta superfamília TGF- engloba um grande número de factores de crescimento e diferenciação que desempenham um papel importante na regulação do desenvolvimento embrionico e na manutenção da homeostase dos tecidos nos animais adultos.

Garzaliz – Cadavid et al (1998) *Proc. Nat. Acad. Sci.* 95: 14938-14943 verificaram que a expressão da miostatina se correlaciona inversamente com a massa muscular isenta de gordura nos humanos, sendo detectável nos músculos humanos do esqueleto nas fibras de tipo 1 e de tipo 2.

Zimmers et al (2002) *Science* 296.: 1486-1488 demonstraram que a miostatina é sintetizada como uma préproteína que é activada por duas cisões proteolíticas. A remoção da sequência sinal (resíduos 1 a 18) é seguida pela quebra num local de processamento tetrabásico, originando um pró-péptido aminoterminal de 26KD e um péptido carboxi-terminal de 12,5KD, um dímero onde se situa a porção biologicamente activa da proteína.

Zimmers et al demonstraram ainda que a miostatina circula no sangue do rato adulto numa forma latente que pode ser activada por tratamento ácido, para uma forma similar ao TGF-beta.

A superexpressão sistémica da miostatina no rato adulto induz uma perda profunda do músculo e da gordura sem diminuição do “intake” de nutrientes o que parece semelhante ao que se passa nos síndromas de caquexia humana e sugere que a miostatina pode ser um utensílio farmacológico muito útil em situações clínicas como a caquexia, em que se deseja um crescimento muscular.

A miostatina bovina com 375 ácidos aminados de comprimento (precursor não processado) com um peso molecular correspondente de 425551 Da tem os seguintes aspectos estruturais:

Sinal – Do resíduo 1 a 18 (potencial)

Pró-péptido – Do resíduo 19 a 266 (potencial)

Cadeia – Do resíduo 267 a 375 (factor 8 de crescimento/diferenciação)

Dissulfureto – 281 e 340  
309 e 372  
313 e 374  
Oligossacáridos ligados – 47  
71  
variantes – 94 - [F\_L no MH raça Piedmonteses]  
313 - [C\_Y no MH raça Piedmonteses]  
14 - [W\_T no MH raça Piedmonteses]

A porção N terminal do GDF8 bovino “maduro” situa-se antes do resíduo 281 cerca do resíduo 267.

A miostatina dos ovinos (GDF8 ovinos) tem, tal como a dos bovinos, 375 ácidos aminados com o PM de 42827 Da (no precursor não processado) e com características estruturais e estruturas primárias com alto grau de homologia.

Os fragmentos de miostatina de suínos estudados têm estruturas primárias muito parecidas com as miostatinas ovinas e bovinas.

Gonzalez – Cadavice et al (1998) anteriormente citados referem que a estrutura do gene MSTN contém três exões e tem três locais putativos de iniciação da transcrição.

O Mh locus nos bovinos está situado na mesma região que o gene do colagénio COL3A1 mapeado nos bovinos em 2q12-q22 ou 2q14.q15 segundo outros.

O estudo com modelos animais por McPherron et al (1997) *Nature* 387: 83-90 permitiu verificar que animais (ratos) sem Gdf8 eram maiores que os animais controlos e revelavam um notável aumento da massa muscular esquelética, com os músculos individuais dos animais mutantes a pesarem 2 a 3 vezes mais que os controle, parecendo que este aumento da massa muscular se devia a uma combinação da hiperplasia da célula muscular e a hipertrofia, o que sugere que as funções do Gdf8 são especificamente a de regulador negativo do crescimento do músculo do esqueleto. A produção e segregação de leptina está também reduzida nestes animais mutantes o que está associado com a reduzida deposição de gordura.

### **E - Ratos transgênicos sem miostatina**

Como já se referiu ratos transgênicos a que falta o gene funcional da miostatina revelam maior massa dos músculos do esqueleto em consequência da hipertrofia e hiperplasia (Mc Pherron et al *Nature* 387: 83-90, 1997) o que sugere que nos ratos normais a miostatina actua como um regulador negativo da massa muscular.

Foi possível verificar em ratos, num estudo comparativo efectuado com a expressão do gene da miostatina em músculos lentos e rápidos, que a expressão da miostatina não parece associada com a atrofia muscular originada por falta de exercício, embora a miostatina esteja fortemente associada com a expressão de isoformas de cadeias pesadas da miosina no músculo normal.

## **F - Expressão da miostatina**

Na expressão da miostatina em tecidos de suínos foi possível verificar a ocorrência de dois transcritos mRNA (1,5 e 0,8kb) para a miostatina em fetos com 21 e 35 dias ocorrendo um aumento notável aos 49 dias.

Após nascimento dos animais esta expressão parece declinar.

Foi possível verificar que a expressão da miostatina nos músculos do esqueleto atinge um máximo antes do nascimento e que uma maior expressão desta miostatina está ligada com um menor peso à nascença.

Por outro lado a expressão da miostatina na glândula mamária indica o seu possível papel no desenvolvimento da glândula mamária e/ou na lactação.

O excesso de musculatura nos bovinos de garupa dupla resulta de mutações no gene da miostatina, mas nos animais normais o papel da miostatina não está completamente clarificado.

A miostatina e o Myo D aumentam durante a formação das fibras musculares, indicando um papel coordenado na diferenciação terminal e ou na fusão dos mioblastos.

No entanto a expressão molecular da miostatina e do Myo D são maiores nos fetos bovinos de “garupa dupla” do que nos fetos normais.

O que sugere que ocorra um mecanismo de feedback nesta regulação que seja rompido nos casos dos animais de garupa dupla.

Foi possível verificar em células C2H<sub>12</sub> de músculo do esqueleto de rato, utilizando miostatina recombinante que há inibição, dependente da dose, da proliferação celular (contagem de células e ensaio de Formazan) da síntese do DNA (por incorporação de timidina H<sup>3</sup>) e da biossíntese proteica (incorporação da leucina 1C<sup>14</sup>).

O efeito inibidor é maior sobre os miotubos do que sobre os mioblastos, não se observando qualquer efeito significativo sobre a degradação proteica ou sobre a apoptose.

Tudo isto sugere que a miostatina pode controlar as massas musculares através da inibição do crescimento muscular ou da sua regeneração.

Por outro lado é conhecido que mutações na miostatina no rato e nos bovinos estão ligadas a um aumento da massa muscular o que sugere que a miostatina inibe o crescimento muscular.

A super-expressão da miostatina em células C<sub>2</sub>C<sub>12</sub> do músculo de rato, inibe o processo miogénico e a diferenciação miogénica desregulando os níveis de mRNA musculares dos factores Myo D e miogenina.

A miostatina, um membro da superfamília TGF<sub>β</sub> (factor <sub>β</sub> de transformação do crescimento) é um regulador negativo da miogénese, que controla a proliferação dos precursores das células musculares, evitando a progressão dos mioblastos da fase G<sub>1</sub> para a fase S durante o ciclo celular.



Os mioblastos, células precursoras musculares, são induzidos para se fundir formando miotubos que se assemelham às células musculares multinucleadas diferenciadas.

Há indicações de que a miostatina especificamente regula a p21Waf1/Cip1, um inibidor das cinases dependentes das ciclinas e que diminui os níveis de actividade da Cdk2 nos mioblastos, parecendo também induzir a formação da forma hipofosfatada da Rb que se acumularia. Tudo isto levaria à paragem dos mioblastos na fase G1 e daí a desregulação da sua proliferação.

### **G - Regulação da miostatina e crescimento muscular**

A proteína miostatina purificada a partir de células mamíferas consiste de um complexo formado não covalentemente pelo pró-peptido N-terminal e um dímero de fragmentos C-terminal ligados por pontes dissulfureto.

O dímero purificado C-terminal de miostatina tem capacidade para se ligar a receptores do tipo II da activina (membros da superfamília TGF $\beta$  induzidos da mesoderme como os quatro tipos de TGF $\beta$ ) mais ao ActRIIB e menos ao ActRIIA.

Esta ligação da miostatina ao ActRIIB pode ser inibida pela proteína follistatina ligando à activina e em mais elevadas concentrações, pelo pró-peptido da miostatina.

Foi possível verificar experimentalmente que estas moléculas propeptídeo, follistatina e outras moléculas podem bloquear esta via de transdução de sinal e como tal serem meios úteis para favorecer o crescimento muscular com utilidade evidente em produção animal e até terapêuticas.

O miostatina é um factor de crescimento segregado que é proteolísado nos mioblastos para originar uma miostatina madura COOH-terminal e um péptido latente-associado NH<sub>2</sub> terminal.

Nos bovinos Piedmontese de “garupa dupla” (o que ocorre com alta frequência) é expressa uma miostatina na qual a cisteína (313) é substituída pela tirosina.

Foi verificado nestes bovinos Piedmontese de garupa dupla que havia um aumento da expressão do mRNA da miostatina e da proteína miostatina precursora, nos músculos do esqueleto, embora os níveis de miostatina madura em circulação nestes animais não variasse.

É conhecido que a proteína miostatina dos animais normais inibe a proliferação de mioblastos C<sub>2</sub>C<sub>12</sub>, mas a miostatina destes bovinos Piedmontese de garupa dupla não inibe a proliferação dos mioblastos o que indica que esta última miostatina, em que a cisteína foi substituída por tirosina, altera a estrutura da ponte dissulfureto com a consequente perda da sua actividade biológica.

Esta mutação parece afectar o processamento e a estabilidade da miostatina madura mas não a sua segregação.

A miostatina (GDF8) factor de crescimento e diferenciação é um factor essencial para a regulação da massa de músculos do esqueleto do rato .

Foram assinaladas em 1997 por Mc Pherron sequências de miostatina de nove outras espécies de vertebrados e identificadas mutações na sequência codificadora da miostatina em duas raças de bovinos

de garupa dupla. O Belgian Blue revelou uma sequência da miostatina com uma deleção de 11 nucleótidos no terceiro exão o que elimina virtualmente toda a região madura, activa da molécula.

### **H - Hipertrofia muscular ovina**

Foi observada em ovinos domésticos (*ovis aries*) uma mutação que origina uma hipertrofia muscular, simultaneamente com carne menos gorda e com melhorados índices de conversão alimentar.

Há indicações de que seja responsável por isto um gene autossomal dominante. Utilizando marcadores de DNA de Bovidae foi assinalado o locus callipyge (*clpg*) correspondente aquela situação, no cromossoma 18, utilizando uma bateria de marcadores do cromossoma 21 de bovinos.

A hipertrofia muscular em ovinos consiste num aumento anormal de tecido muscular também originado inteiramente pelo alargamento das células existentes e não por hiperplasia muscular, (na qual um aumento anormal dos tecidos musculares é consequência da formação e crescimento de células musculares novas, normais).

Esta situação foi primeiramente descrita por Cockett et al (1994) *Proc. Nat. Acad. Sci. USA* 91: 3019-3023 que revelou que o locus desta mutação potencialmente vantajosa tem sede no cromossoma 18, próximo da terminação não centromérica. O alelo mutante parece levar a um aumento da síntese de proteína e a uma diminuição da degradação proteica, originando maior rendimento da carcaça, menos gordura e marmoreado da carne apesar do peso da carcaça não ser afectado.

A conformação da carcaça é alterada com um aumento da massa muscular a ter sede na perna e lombos, e com melhores índices de conversão embora com o mesmo ritmo de crescimento em relação a animais normais, embora com produção de lã bruta menor, e comprimento do fio também menor.

### **I - Factor de crescimento do tecido conectivo**

O precursor do factor de crescimento do tecido conectivo (CTGF bovino) tem por principal função ser um mitogénio mitoatraente do tecido conectivo sendo segregado pelas células do endotélio vascular, promovendo a proliferação e diferenciação dos condrócitos. Medeia ainda a adesão celular e favorece o crescimento dos fibroblastos sendo um factor indutor da síntese de DNA.

Parece actuar como um monómero encontrando-se na forma solúvel na matriz extracelular.

Pertence à família de proteínas ligantes de factores de crescimento insulina - like (IGF) (IGFBP8).

A sua estrutura de 349 ácidos aminados (do precursor não processado) com um P. M. de 37924 Da, tem um sinal entre os resíduos 1 e 26 e o factor de crescimento do tecido conectivo entre os resíduos 27 e 349 com um domínio VWFC entre os resíduos 101 e 167 e outro domínio CTCK entre os 256 e 330 e com várias pontes dissulfureto.

Nakanishi et al (2001) *Biochem Biophys Res Commu* 281\_: 678-681 produziram ratos transgénicos que super-expressavam CTGF sob o controlo do promotor do colagénio tipo XI do rato. O crescimento embrionário e neonatal ocorreu normalmente mas passados alguns meses após nascimento revelaram nanismo e densidade óssea diminuída, parecendo a super expressão do CTGF afectar certas etapas da ossificação endocondrial sendo também os testículos mais pequenos e a fertilidade desses ratos transgénicos afectada.

O precursor do factor do crescimento do tecido conectivo (CTGF suíno) também com 349 ácidos aminados e 37924 Da de P. M. revela características estruturais inclusivé estrutura primária muito homologa à de bovino.

### **J - Hipertrofia das células dos músculos do esqueleto induzidas por inibidores de metaloproteases: miostatina como um potencial mediador**

O crescimento e diferenciação celular são controlados em muitos tecidos por factores parácrinos os quais muitas vezes necessitam de processamento proteolítico para serem activados.

As metaloproteínas da família “metzincin” tais como metaloproteases da matriz estão envolvidas no desprendimento de factores de crescimento, citocinas e receptores.

Inibidores destes metaloproteases podem aumentar a fusão de mioblastos C<sub>2</sub>C<sub>12</sub> in vitro e provocarem hipertrofia dos miotubos, não parecendo faze-lo através de proteínas que regulam o crescimento do músculo in vitro como é o caso da IGF1, calcineurina e factor de necrose tumoral \_.

Por outro lado a maturação proteolítica da miostatina parece ser reduzida nestas células tratadas com inibidores das metaloproteases, pois apenas se encontra neste miotubos hipertrofiados precursor da miostatina não processado.

Parece pois que as metaloproteases estão implicadas na regulação do crescimento e diferenciação dos músculos do esqueleto, e que a maturação proteolítica da miostatina nas células C<sub>2</sub> C<sub>12</sub> está ligada a estas metaloproteases inibíveis, sendo possível que a falta de processamento da miostatina, devido à presença dos inibidores das metaloproteases possa explicar esta hipertrofia dos miotubos.

### **L - Comparação da GH, IGF1 e testosterona com os mRNA de receptores e da miostatina em músculos do esqueleto de homens idosos**

Nos humanos a GH, IGF1 e testosterona (T) são mediadores importantes na biossíntese proteica dos músculos, e assim da massa muscular, todas elas declinando com a idade.

A concentração destas hormonas em circulação pode em princípio estar relacionada com os níveis de transcrição dos seus receptores e relacionados negativamente com a expressão do gene da miostatina. Contudo o que se verifica é que não há relações significativas entre a idade, massa de carne magra do organismo ou percentagem de gordura com os níveis de transcrição dos receptores para a GH, IGF1, testosterona ou miostatina.

Não há também correlação significativa entre as concentrações no soro sanguíneo, de GH, IGF1 e T com os seus correspondentes níveis de mRNA nos músculos do esqueleto.

Contudo o receptor para a GH está correlacionado negativamente com os níveis de mRNA da miostatina.

Tudo isto sugere que os declínios de GH e T ao longo da idade das pessoas possa levar a um aumento da miostatina e a uma dissociação nos efeitos autocrinos da IGF1 sobre a biossíntese proteica muscular, contribuindo assim para a sarcopénia associada com o avançar da idade.

## **2.2 - Outros factores que afectam o crescimento e desenvolvimento dos músculos do esqueleto**

### **Insulina, hormona do crescimento, hormona da tiróide e outras hormonas**

Alguns factores que afectam o crescimento muscular após o nascimento dos animais, englobam hormonas como a insulina, a hormona do crescimento, as hormonas da tiróide, as glucocorticóides, e as hormonas sexuais, mas também o grau de exercício físico está implicada neste desenvolvimento tal como os regimens alimentares.

A insulina é provavelmente o factor mais importante na regulação do balanço proteico nos músculos do esqueleto. O seu papel na iniciação da transcrição e tradução em proteínas durante a biossíntese destas parece limitado aos músculos de contracção rápida (alta capacidade glicolítica e alta capacidade oxidativa nos músculos vermelhos de contracção rápida e alta actividade glicolítica e baixa capacidade oxidativa nos músculos brancos de contracção rápida).

Nos músculos do coração e no músculo soleus a insulina não tem efeito sobre a iniciação da cadeia peptídica.

Em todos os músculos a falta de insulina leva a uma redução do RNA (+/-) 15% parece que devido a uma redução do número de ribossomas, parecendo alterar o perfil polissomal.

A insulina estimula ainda o transporte de alguns ácidos aminados para o interior dos músculos.

A insulina é também activa na redução da degradação das proteínas, operando talvez através dos lisossomas e tendo poucos efeitos sobre as proteínas com semi-vida curta, sendo provavelmente nos músculos a actina e miosina mais degradadas por acção da insulina, uma vez que têm uma semi-vida mais longa que as proteínas sarcoplásmicas.

A hormona do crescimento (GH) parece implicada no turnover proteico muscular, estando bem caracterizada a redução da biossíntese proteica (50%) nos animais a que foi removida a hipófise tal como a redução da degradação proteica (30%), a diminuição de RNA e provavelmente o número de ribossomas.

Algumas das acções da GH são mediadas por somatomedinas.

As hormonas da tiróide diminuem a biossíntese proteica e de DNA no músculo e têm poucos efeitos sobre a degradação proteica.

Nos esteróides sexuais o papel dos androgénios é obvio pois o crescimento muscular nos machos é superior ao das fêmeas e a castração diminui o crescimento muscular.

Os esteróides anabólicos aumentam a deposição proteica nos ruminantes e suínos, mas nos ratos, equídeos e aves os seus efeitos não são consistentes.

Ambos, os androgénios e estrogénios podem ter efeitos anabólicos.

Efeitos das glândulas endócrinas e suas segregações sobre o crescimento e metabolismo das  
proteína e gordura nos músculos

<u>Factor</u>	<u>Efeitos Endócrinos</u>
Adenohipófise ( hormona do crescimento GH )	Aumenta o transporte de ácidos aminados para o interior das células e a biossíntese do DNA, RNA e proteínas: efeitos anabólicos globais; efeitos negativos sobre a deposição de gordura.
Tiride ( T4 e T3 )	Os efeitos dependem das doses ; doses baixas são anabólicas e altas, catabólicas, o hipotiroidismo e hipertiroidismo provocam diminuição muscular e efeitos de mobilização da gordura.
Pâncreas ( insulina )	Aumenta a síntese proteica e o transporte de ácidos aminados; diminui a quebra de proteínas; aumenta a deposição de gorduras.
( Glucagon )	Diminui a síntese proteica; aumenta a gluconeogénese e mobilização de gordura
Cortéx supra-renal ( glucocorticóides )	Diminuem a síntese de DNA, RNA e proteínas; aumenta a quebra de proteínas; aumenta a gordura em doses baixas, mas ocorre mobilização da gordura
Medular da supra-renal ( adrenalina )	Mobiliza a gordura
Testis ( androgénios )	Aumenta a eficiência da síntese do DNA, RNA e proteína, aumentando o peso do músculo e a massa proteica; efeito negativo sobre o metabolismo lipídico.
Ovário ( estrogénios )	Efeitos anabólicos nos ruminantes; doses baixas mobilizam a gordura e altas afectam negativamente o metabolismo proteico e positivamente as gorduras; nos ratos são geralmente catabólicas.

**Esteróides sexuais naturais e produção animal**

Os agentes anabólicos em geral promovem o crescimento dos animais afectando o metabolismo (anabolismo e catabolismo) em ordem a produzir um aumento da deposição proteica favorecendo assim o crescimento. Este é obtido por um aumento da retenção azotada (balanço azotado positivo) e pelo consumo da gordura corporal dos animais, simultaneamente com a obtenção de um melhor índice de conversão dos alimentos utilizados.

Este aumento da deposição proteica resulta de um aumento da biossíntese proteica sem variação ou com diminuição da degradação proteica, e tem sede sobretudo ao nível dos músculos do esqueleto. Daí serem também designados estes agentes anabólicos como agentes miotróficos, assinalando assim o seu local principal de actuação, ou seja os seus órgãos alvos.

No que se refere aos estrogénios naturais utilizados para efeitos de melhorar a produção animal o estradiol não é geralmente utilizado como promotor do crescimento, embora dê bons resultados quando associado com androgénios e gestagénios.

A influência dos estrogénios sobre o crescimento dos animais e seu turnover proteico depende de vários factores, diferindo os ruminantes dos não ruminantes, pois enquanto nos primeiros é induzido um aumento do peso vivo, noutras espécies animais sucede o contrário.

Baixas doses de estrogénios são anabólicas, mas altas doses são lipogénicas e catabólicas, mas tudo depende das espécies animais. Nas aves o estradiol aumenta a deposição de gordura com ou sem ganho de peso.

Nos androgénios naturais o mais importante é a testosterona, sendo os androgénios verdadeiros agentes anabólicos pois actuam directamente sobre os músculos, e nos músculos do esqueleto a testosterona não pode ser convertida em dihidrotestosterona, nem a sua acção é mediada por outras hormonas endógenas. As concentrações de testosterona não têm nada a ver com as características do colagénio.

Os receptores para a testosterona encontram-se em muitos tecidos e sobretudo nos músculos do esqueleto e daí os androgénios terem efeitos directos sobre as células musculares.

A castração diminui o número destes receptores enquanto a administração de testosterona a fêmeas aumenta os receptores para os androgénios.

Recorda-se que os receptores para os androgénios são diferentes dos receptores para os estrogénios.

A resposta nos androgénios depende ainda dos grupos musculares.

Qualidade da carne e qualidade da carcaça são coisas diferentes, sendo a segunda a qualidade da produção animal, enquanto a primeira é a qualidade do produto final que depende de vários factores tais como os músculos (vermelhos, brancos ou mistos) a partir dos quais a carne é obtida (o que depende por seu turno da utilização ou não de novas tecnologias tais como a estimulação eléctrica, a velocidade de queda do pH, encolhimento pelo frio, estirpes genéticas, etc.).

### **Efeitos do exercício e nutrientes**

Nos efeitos desencadeados pelo exercício dos animais é notório a hipertrofia muscular, embora esta hipertrofia não envolva alteração nas respostas dos músculos às hormonas, nem tenha reflexos na biossíntese proteica parecendo que a degradação das proteínas é consideravelmente reduzida.

Ultimamente tem havido sugestões de que a hipertrofia muscular assemelhando-se ao crescimento muscular normal, terá mais elevadas taxas de síntese e degradação proteicas.

No que se refere à disponibilidade de nutrientes sobre o crescimento muscular a síntese proteica é afectada pela concentração de ácidos aminados disponíveis; aumentando esta até certos valores aumenta também aquela. A omissão de um ácido aminado dispensável, com excepção da prolina, não afecta praticamente a biossíntese proteica, mas a ausência de um ácido aminado essencial ou indispensável (excepto a histidina) reduz notavelmente a velocidade da biossíntese proteica.

Em circunstâncias normais, a concentração dos ácidos aminados não determinaram a velocidade de síntese proteica nos músculos.

Em condições adversas nutricionais, a síntese proteica é geralmente afectada.

Nos animais jovens mal alimentados com um carácter agudo, os efeitos são menos notórios nos músculos do que noutros tecidos.

É pois muito difícil definir um perfil de concentração de ácidos aminados no plasma sanguíneo indicativo da velocidade da biossíntese proteica, pois ele, perfil no plasma, pode ser alto ou baixo independentemente da velocidade da síntese proteica. Esta não é pois controlada pela concentração das aminoacidémias.

O papel da glucose, ácidos gordos voláteis, ácidos gordos livres e cetonas não é muito claro embora se saiba que o seu aumento melhora a retenção azotada.

### **Alguns outros efeitos de agentes anabolizantes**

<b><u>Factor</u></b>	<b><u>Efeito Nutritivo</u></b>
Diminuição da entrada de proteína	Diminui a síntese de DNA, RNA e proteína e também diminui os seus catabolismos.
Diminuição da entrada de energia	Diminui a síntese de DNA, RNA e proteína, mas aumenta a sua quebra.
Ácidos aminados essenciais	Aumenta a síntese proteica e diminui a quebra proteica.
<b><u>Factor</u></b>	<b><u>Efeitos mistos</u></b>
Deservação	Diminui a síntese proteica e aumenta a quebra de proteínas.
Exercício	Aumenta o transporte de ácidos aminados e a síntese proteica mas diminui a quebra de proteínas.

A síntese proteica nos músculos representa cerca de \_ da síntese total de todo o organismo.

A regulação hormonal do metabolismo proteico muscular está associado com o “turnover” da proteína (síntese e proteólise).

Assim nos humanos os hormonas anabólicas estimulam o crescimento muscular aumentando a síntese proteica (a hormona do crescimento, factores de crescimento “insulina<sup>++</sup>like” e testosterona) ou diminuindo a proteólise (a insulina). Ao contrário dos animais em crescimento, o principal efeito anabólico da insulina nos humanos adultos sobre a proteína muscular é uma inibição da proteólise.

A síntese proteica é estimulada apenas na presença de aportes de altos níveis de ácidos aminados.

Uma combinação de hormonas de stress (glucagina, glucocorticóides e catecolaminas) originam catabolismo muscular.

Ainda nos humanos a hormona tiróide que é essencial durante o crescimento, em excesso ou em deficiência causam desgaste muscular por mecanismo ainda desconhecidos.

Nos efeitos dos agentes anabólicos sobre os metabolitos do sangue e hormonas endógenas, é conhecido que o estradiol 17- $\beta$  aumenta as concentrações sanguíneas de GH e de insulina nos ruminantes, sendo também conhecido que os estrogénios e androgénios deprimem a tiroxina sobretudo os primeiros.

A regulação da GH nos animais homeotérmicos é controlada por péptidos hipotalâmicos hipofisiotróficos tais como a GRF (factor de libertação da GH) e a SRIF (factor inibidor da libertação da GH), embora toda esta regulação global seja modulada por factores nutricionais e outros factores ambientais e seja diferente consoante as espécies animais homeotérmicas e poiquilotérmicas.

A GH é um promotor do crescimento potencial para os suínos.

As concentrações sanguíneas de GH são mais baixas em vitelas de garupa dupla do que em vitelas normais de todas as idades.

A GH ao nível dos músculos altera o transporte dos ácidos aminados e o metabolismo do RNA ao nível ribossomal, aumentando a biossíntese proteica mas alterando a degradação proteica.

Há autores (Steel e tal., 1993) que admitem que as acções da GH a nível muscular são modeladas através do cAMP, afectando também a segregação das somatomedinas A, B e C, sobretudo a primeira e última, estando também assinalado que afecta a segregação de insulina, hormona esta que se sabe ter efeitos proteoanabólicos ao nível das células dos músculos do esqueleto.

De forma análoga a insulina no soro sanguíneo é mais baixa nos vitelos de garupa dupla do que nos vitelos normais. A idade tem um efeito linear e quadrático sobre as concentrações de insulina. Uma dieta de elevada energia após a desmama duplica as concentrações de insulina.

Os efeitos anabólicos dos esteróides dependem muito das espécies animais e por vezes dos sexos. Bovinos e ovinos respondem positivamente a estes agentes anabólicos, mas os suínos já revelam resultados variáveis tal como as aves.

As hormonas da tiróide são essenciais para o normal crescimento dos animais mamíferos, estando assinalado que contribuem para o aumento da síntese proteica (a T3 e a T4).

Os efeitos das hormonas tiróides, como já referimos dependem das doses, doses baixas são anabólicas e doses altas são catabólicas, neste caso aumentando mais a quebra de proteínas do que a sua síntese.

Está assinalada a presença de receptores para as hormonas tiróides (T4 e T3) no citoplasma, mitocôndrias e núcleos das células dos tecidos alvo.

As hormonas tiróides activam as RNA polimerases I e II e a translação e transcrição são assim influenciadas.



A insulina afecta o transporte dos ácidos aminados e da glucose, aumenta a biossíntese proteica e diminui a quebra de proteínas nos músculos do esqueleto, sendo estes efeitos modulados por receptores específicos, reversíveis e de alta afinidade, da plasma membrana das células dos músculos do esqueleto. Algumas das acções da insulina são mediadas por alterações de alguns mensageiros intracelulares como é o caso do cAMP segundo Hazelwood, R. L. (1993). Contudo os seus efeitos sobre os ácidos nucleicos e proteínas não é mediado através do cAMP, parecendo que as acções da insulina sobre a biossíntese proteica e degradação proteica são independentes uma da outra.

Os glucocorticóides das cápsulas supra-renais são esteróides catabólicos, tendo as hormonas da adrenocortical em geral um efeito inibidor sobre as várias facetas do crescimento, inibindo directamente a síntese de DNA e das proteínas e a replicação celular, originando a desintegração de alguns tipos celulares.

Eles inibem a segregação de GH e têm efeitos anti-insulina, havendo provas de que os glucocorticoides também reduzem a produção do mRNA que exprime a IGF-1, ou somatomedina C.

Os corticoides causam também a síntese de proteínas específicas que têm um efeito inibidor do metabolismo, tendo também um efeito negativo sobre culturas de células musculares.

Nos bovinos e ovinos está assinalada uma correlação negativa entre os corticoesteróides e o crescimento.

Nos mecanismos de acção dos agentes anabólicos estrogénios há que considerar os efeitos directos e indirectos.

O mecanismo de acção dos estrogénios é pior conhecido do que o dos androgénios, sobretudo no que se refere ao crescimento dos músculos do esqueleto.

O efeito directo dos estrogénios é mediado ao nível da transcrição pela expressão dos genes.

Os receptores para os estrogénios são diferentes dos dos androgénios nos músculos. Apesar de haver uma grande sobreposição competitiva no local de ligação entre os esteróides que se ligam aos receptores da progesterona, androgénios, glucocorticóides e mineralocorticóides, os compostos que revelam mais alta afinidade para estes receptores revelam pouca ou nenhuma afinidade para os receptores dos estrogénios.

Também há provas de que os estrogénios se ligam aos receptores dos androgénios mas com reduzida afinidade (cinco a dez vezes menos).

Os efeitos indirectos dos estrogénios são mediados através da segregação de GH e de insulina que passam a ser mais elevadas nos animais tratados com estrogénios.

O mecanismo de acção da progesterona é similar ao das outras hormonas esteróides, a molécula hidrófoba atravessa a plasma membrana celular e no interior da célula liga-se reversivelmente à proteína receptora da progesterona, originando uma mudança da conformação deste complexo que se liga ao DNA regulando a transcrição dos genes. Este receptor para a progesterona parece ser diferente de todas as outras proteínas receptoras para esteróides sexuais

Para efeitos meramente supletivos no quadro seguinte revêm-se alguns factores de crescimento polipeptídicos que se se tem verificado estimularem a proliferação de células em cultura.

### Alguns factores de crescimento polipeptídicos

<u>Factor</u>	<u>Fonte</u>	<u>Peso molecul ar</u>
Factor de crescimento epidérmico ( EGF )	Submaxilar de rato	6045
Factor de crescimento dos nervos ( NGP )	Urina humana Submaxilar do rato	6201 13.254
Factor de crescimento dos fibroblastos ( FGF )	Pituitária bovina	13.400
Factor de crescimento like-insulina ( IGF )-1	Cérebro bovino Plasma humano	13.000 7.649
Idem IGF-2	Plasma humano	7.471
Somatomedina A ( SM )	Plasma humano	7.000
Somatomedina C(SM)( possivelmente idêntica a IGF-1)	Plasma humano	8.567
Actividade estimuladora de multiplicação ( MSK ) ( 95% de homologia com a IGF-2 humana )	Cultura de células hepáticas de rato Soro de rato	7.484
Factor de crescimento derivado das plaquetas ( PDGF )	Plaquetas humanas Soro humano	31.000

### **2.3. - Endocrinologia dos peixes e produção animal**

Nos peixes teleosteos a hormona do crescimento (GH) tem um papel regulador muito importante em acontecimentos fisiológicos muito diversos, tais como a reprodução, a regulação osmótica ou iónica, o metabolismo, o crescimento e o desenvolvimento.

A segregação da GH é regulada por factores neuroendócrinos hipotalâmicos que actuam directamente sobre as células somatotróficas da hipófise, as modulam, bem como a segregação ou actividade de outros factores neuroendócrinos. Tudo isto é influenciado pelo estado nutricional e reprodutivo dos peixes havendo diferenças notáveis consoante as espécies.

Nos peixes os neuropéptidos, hormona de libertação da somatotrofina, hormona de libertação da GH, hormona de libertação da tirotrofina, neuropéptido Y, serotonina e polipéptido activador da adenilato ciclase da pituitária, estimulam todos a GH.

Contrariamente a somatostatina tem uma acção fortemente inibidora da libertação da GH na dourada e na carpa mas é menos eficiente nos salmões e nas trutas.

A segregação da hormona do crescimento (GH) nos salmonídeos e noutros peixes é controlada pela SRIF (somatostatina) que é o principal factor inibidor, e pela GRF (factor de libertação da hormona do crescimento) que é o principal factor estimulador. A GH nos salmonídeos estimula a segregação de IGF-1 e a retro regulação da segregação da GH podem ser reguladas através de variações das concentrações de IGF-I.

A GH promove a adaptação à água do mar nos salmonídeos e nas tilápias.

A transcrição da hormona do crescimento (GH) nos vertebrados depende entre outros factores, de factores de transcrição específicos da pituitária (tais como Pit-1) e da hormona tiróide (TH) e dos receptores para a hormona tiróide.

A GH é a principal hormona reguladora do crescimento somático nos salmonídeos.

O efeito estimulador do crescimento da GH é provavelmente integrado com o da IGF-1 (insulin-like growth factor I) tal como sucede nos vertebrados.

A GH estimula a síntese de proteínas e melhora a conversão dos alimentos durante o crescimento, promovendo também a quebra dos lípidos e do glicogénio assim como a gliconeogénese, funções que são provavelmente de grande importância durante o jejum quando as concentrações de GH sobem.

Durante o desenvolvimento e transformação dos juvenis as concentrações de GH em circulação parecem ser reguladas pelo ambiente, durante o aumento dos dias na primavera as concentrações de GH sobem e a temperatura modula a regulação do fotoperíodo da GH.

O papel da GH na adaptação dos salmonídeos juvenis à água de mar é complexa.

Na água doce a GH melhora a capacidade hipoosmoreguladora estimulando a actividade Na super (+), K super (+) ATPases nas brânquias, e provavelmente actuando também ao nível dos rins e intestinos. A seguir à entrada na água do mar, as concentrações de GH e o seu “turnover” aumentam transitoriamente.

A GH está ainda envolvida na regulação da maturação sexual.

A GH aumenta o apetite mas não se sabe se isto sucede por acção directa sobre o sistema nervoso central ou indirectamente através de alterações metabólicas.

A GH aumenta a actividade natatória.

A GH é uma hormona importante e multifuncional nos salmonídeos sendo um mediador central das alterações sazonais na fisiologia e comportamento destes peixes.

Como se disse a segregação da GH nos teleósteos é estimulada por uma série de factores neuroendócrinos inclusivé o factor de libertação da GH, a hormona libertadora da gonadotrofina (GnRH), dopamina, neuropéptido Y, hormona libertação da tirotrofina (TRH) e colecistocinina.

Os esteróides sexuais também influenciam a resposta dos somatotrofos aos factores neuroendócrinos.

A utilização de péptidos GnRH ou esteróides como promotores do crescimento dos peixes é promissora embora esteja ainda numa fase muito principiante.

Na regulação da hormona do crescimento (GH) em Anguilla anguilla estão provados os fortes efeitos inibidores da somatostatina (SRIF) e do IGF-I (insulin-like growth factor I) tal como parece suceder em diversos vertebrados.

A hormona do crescimento (GH) é uma das principais hormonas que regulam o crescimento dos vertebrados, mas o seu papel no desenvolvimento das larvas dos peixes é incerto.

A GH também é um importante regulador da alimentação.

A somolactina (SL) é uma nova hormona identificada na pituitária nos teleósteos pertencente à família hormona do crescimento-prolactina.

A hipófise de uma série de teleósteos contem duas gonadotrofinas ( GtHs ) que são produzidas em populações celulares distintas.

Em diversas espécies estas gonadotrofinas são expressas em concentrações variadas ao longo do ciclo reprodutivo.

Admite-se que a síntese destas gonadotrofinas é regulada por hormonas hipotalâmicas e esteróides das gónadas.

Durante a ovulação e desova a GH do plasma aumenta tal como as gonadotrofinas.

Nas tilapias o efeito das GnRH (hormona libertadora da gonadotrofina) sobre as gonadotrofinas parece ser mediado por duas vias a cAMP-PKA e a PKC.

A somatostatina (SRIF) não altera as taxas de transcrição da GH quer nas tilapias quer nas trutas arco-íris mas pode alterar a síntese da GH por modulação da translação.

Os esteróides das gónadas afectam a transcrição das gonadotrofinas diferentemente, nas tilapias a testosterona eleva-a primeiro mas mais tarde diminui-a.

Parece pois que o promotor do gene das gonadotrofinas nos salmonídeos contem vários elementos de resposta dos estrogénios (EREs).

As concentrações de gonadotrofinas (GtHs) na pituitária e soro de carpas vulgares (*Cyprinus carpio*) variam com a idade, sendo significativamente mais baixas nos jovens relativamente aos adultos, havendo também variações significativas sazonais no soro (não na pituitária )mais elevadas nos peixes maduros sexualmente e nos em pré-ovulação.

As concentrações de GH na pituitária e soro também variam consoante a idade sendo mais altos nos jovens.

As IGF (insulin-like growth factors) encontram-se na circulação e ligam-se a proteínas com altas afinidades de ligação para elas e que são as IGFBP (insulin-like growth factor binding proteins) que parecem encontrar-se em todos os vertebrados.

A GH de tilapia e a prolactina de tilapia (tPRL) aumentam significativamente os níveis da IGFBP.

A IGF-1 tem sido sugerida como um sinal potencial ligando o crescimento e a puberdade nos mamíferos.

Em salmonídeos imaturos foi possível verificar que elevadas concentrações de leptina humana não altera o seu crescimento, nem os níveis plasmáticos de IGF-1, ou insulina, de GH ou de tiroxina.

Diversos aspectos do desenvolvimento, diferenciação e metabolismo dos vertebrados são regulados pelas hormonas tiróides e retinóides que actuam através de uma família de proteínas receptoras intracelulares que actuam directamente sobre os genes alvo.

Há poucas dúvidas de que as hormonas tiróides são essenciais para o desenvolvimento dos peixes e estão assinalados abundantes receptores para elas nas larvas do sargo do mar.

O polipéptido intestinal vasoactivo (VIP) pertence á família dos péptidos que têm uma sequência conservada e englobam o glucagon, a secretina, o péptido libertador da insulina dependente da glucose, o factor de libertação da GH, péptidos tendo histidina N-terminal e isoleucina C-terminal (PHI) e os péptidos de répteis helodimina e helospectinas.

O VIP e péptidos relacionados têm sido isolados de diversas espécies de teleósteos e revelam uma sequência em ácidos aminados similar à encontrada nos mesmos péptidos mamíferos.

O VIP tem sido assinalado no hipotálamo de teleósteos estando implicado no controlo da segregação de prolactina.

Na regulação neuroendócrina do apetite nos peixes, a bombesina e o sistema peptídico colecistocinina do cérebro têm acções poderosas na saciedade das douradas tendo ainda proeminentes acções na regulação neuroendócrina da segregação da GH nas douradas.

Após uma refeição, as douradas tem um curto aumento da concentração de GH no soro e depois uma diminuição para valores mais baixos do que aqueles que existiam antes da refeição.

O neuropéptido Y que tem um papel importante como estimulador da ingestão de alimentos nos mamíferos encontra-se também no cérebro neuroendócrino e nos centros cerebrais reguladores do apetite nas douradas e noutros teleósteos.

## Bibliografia

### 2.1 - Miogénese

#### B - Genes miogénicos

- Barth, J.L., Worrel, R.A., Crawford J.M., Morris, J., Ivanić, R. ( 1993 ) Isolation, sequence and characterization of the bovine myogenic factor-encoding gene myf5 Gene 127: 185-191
- Chang, K.C., Fernandes, K., Chantler P.D. ( 1995 ) Cloning and in vivo expression of the pig Myo D gene J. Muscle Res. Cell Motil 16: 243-247
- Davis, R.L., e Weintraub, H., ( 1992 ) Acquisition of myogenic specificity by replacement of three amino acids residues from Myo D into E12 Science 256: 1027-1030
- Huynem, L., Bass, J., Gardne, R.C., Bellamy A.B. ( 1992 ) Nucleotide sequence of the sheep Myo D1 gene Nucleic Acids Res. 20: 374-37
- Lee, S.J. McPherron, A.C., ( 2001 ) Regulation of myostatin activity and muscle growth Proc. Nat Acad. Sci. USA, 98,16: 9306-9311
- Oldman, J.M. et al ( 2001 ) Molecular expression of myostatin and Myo D is greater in double-muscled than normal-muscled cattle fetuses Am. J. Phys. Regulatory Integrative Comparative Physiology 280: R1488-R1493
- Rescan, P.Y., Jutel, I., Rallicre, C., ( 2002 ) Two myostatin gene are differentially expressed in myotomal muscles of the trout ( *Oncorhynchus mykiss* ) J. Exp. Biol. 204: 3523-3529
- Rudnicki, M.A., Schnegelsberg, D.N.J., Stead, R.H., Braun, T., Arnold, H.H., Jaenisch, r., ( 1993 ) Myo D or Myf5 is required for the formation of skeletal muscle Cell 75: 1357-1359
- Wehling, M., Cai, B., Tidball, J.G. ( 2000 ) Modulation of myostatin expression during modified muscle use FASEB J-14: 103-110

#### C - Gênese e regeneração dos músculos do esqueleto

- Campbell, W.G., Gordon, S.G., Carlson, C.J. et al ( 2001 ) Differential global gene expression in red and White skeletal muscle Am. J. Physiol. Cell Physiol 280: C763-C760

#### D - Hipertrofia muscular bovina

- Charlier, C., et al ( 1995 ) The mh/gene causing double-muscling in cattle maps to bovine chromosome 2. Mammalian Genome 6: 788-792
- Grobet, L., Marlin, L.J.R., Poncelet, D., Pirottin, D., Browwers, B., Riquet, J, et al ( 1997 ) A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle Nature Genetics 17: 71-74
  - Kambadur, R., Sharma, M., Smith, T.P.L., Bass, J.J. ( 1997 ) Mutations in myostatin ( GDF8 ) in double-muscled Belgian Blue and Piedmontese cattle Genome Research 7: 910-916

- Listrat, A., Picard, B., Geasy, Y. ( 1999 ) Age-related changes and location of type I, III, IV, V and VI collagens during development of four fetal skeletal muscles of double-muscled and normal bovin animals Tissue & Cell 31: 17-27
- Listrat, A., et al ( 1999 ) Insulin-like growth factor II ( IGF II ) mRNA expression during skeletal muscle development of double muscled and normal bovine foetuses Reproduction, Nutrition, Development 39: 113-124
- McPherron, A.C., Lawler, A.M., Lee, S.J. ( 1997 ) Regulation of skeletal muscle mass in mice by a new TGF-beta superfamily member Nature 387: 83-90

#### E - Ratos transgênicos sem miostatina

- Carlson, C.J., Booth, F.W., Gordon, S.E. ( 1999 ) Skeletal muscle myostatin mRNA expression is fiber-type specific and increases during hindlimb unloading Am. J. Physiol-Regulatory, Integrative and Comparative Physiology 277: R601-R606
- Lin, J., Arnold, H.B., et al ( 2002 ) Myostatin Knockout in mice increases myogenesis and decreases adipogenesis Biochem. Biophys. Res. Commun 291: 701-706
- Szabo, G., et al ( 1998 ) A deletion in the myostatin gene causes the compact ( Cmpt ) hypermuscular mutation in mice Mammalian Genome 9: 671-672

#### F - Expressão da miostatina

- Edmondson, D.G., Olson E.N. ( 1989 ) A gene with homology to the myc similarity region of Myo D1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program Genes Dev. 3: 628-640
- Jeanplong, F., Sharma, M., Somers, W.G., Bass, J.J., Kambadur, R. ( 2001 ) Genomic organization and neonatal expression of the bovine myostatin gene Mol. Cell. Biochem. 220: 31-37
- Rios, R., Carneiro, S., Arce, V. ^M. E Deusa, J., ( 2002 ) Myostatin is an inhibitor of myogenic differentiation Am. J. Physiology-cell Physiology 282,5: C993-C999
- Shaoquan, Ji, Losinski, R.L., Cornelius, S.G., et al ( 1998 ) Myostatin expression in porcine tissues: tissue specificity and developmental and postnatal regulation Am. J. Physiol-Regulatory, Integrative and Comparative Physiology 275: R1265-R1273
- Stratil, A., Kopecny, M. ( 1999 ) Genomic Organization, sequence and polymorphism of the porcine myostatin ( GDF8; MSTN ) gene Anim. Genet 30: 468-470
- Thomas, M., Langley, B., et al ( 2000 ) Myostatin, a negativ regulator of muscle growth, functions by inhibiting myoblast proliferation J. Biol. Chem. 275,51: 40235-40243

#### G - Regulação da miostatina e crescimento muscular

- Berry, C., Thomas, M., Langley, B., Sharma, M. e Kambadur, R., ( 2002 ) Single cysteine to tyrosine transition inactivates the growth inhibitory function of Piedmontese myostatin Am. J. Physiol. Cell Physiology 283:1: C135-C141
- Dickman, S. ( 1997 ) Agricultural genetics-gene mutation provides more meat on the hoof ( vol.277,pg 1922. 1997 ) Science278: 2037
- McPherron, A.C., Lee, S.J. ( 1997 ) Double muscling in cattle due to mutation in the myostatin Gene Proc. Nat. Acad. Sci. USA 94: 12457-12461

#### H - Hipertrofia muscular ovina

- Bidwell, C.A., Shay, T.L., et al ( 2001 ) Differential expression of the GTL2 gene within the callipyge region of ovine chromosome 18 Aminal Genetics 32: 248-256
- Charlier, C., Segers, K., Wagenaar, D., et al ( 2001 ) Human-ovine comparative sequencing of a 250kb imprinted domain encompassing the Callipyge ( clpg ) locus and identification of six imprinted transcripts: DLK1, DAT, GTL2, PEG11, antiPEG11, and MEG8 Genome Research 11: 850-862
- Cockett, W. E., Jackson, S.P., et al ( 1994 ) Chromosomal localization of the callipyge gene in sheep ( ovis aries ) using bovine DNA markers Proc. Nat. Acad. Sci. USA 91: 3019-3023

#### I - Factor de crescimento do tecido conectivo

- Brigstock, D.R., Steffen C.L., Kim G.Y., veganta R.K., Diehl J.H., Handing P.A. ( 1997 ) – Purification and Characterication of noul heparin-binding growth factors in uterine secretory fluids Identification as heparin-regulated Mr 10.000 forms of connective tissue growth factor J. Biol. Chem. 272: 20275-20282
- Taylor, W.E., Shalender,B., Artaza, J., et al ( 2001 ) Myostatin innibits cell proliferation and protein synthesis in C<sub>2</sub>H<sub>12</sub> muscle cells Am. J. Physiol. Endocrinal Metab. 280: E221-E228

#### J - Inibidores de metaloproteinases e miostatina

- Huet, C.Li, Z.F., Liu, H.Z. et al ( 2001 ) Skeletal muscle cell hypertrophy induced by inhibitors of metalloproteases; myostatin as a potential mediator American Journal of Physiology-Cell Physiology 281: 5-C1624-C1634

#### L - GH,IGF, testosterona e miostatina

- Marcell, T.J., Harman, S.M., Urban R.J. et al ( 2001 ) Comparison of GH, IGF-1 and testosterone with mRNA of receptors and myostatin in skletal muscle in older man American Journal of Physiology. Endocrinology and metabolism 281: 6-E1159-E1164
- Thomas, S.M. e Brugge, J. ( 1997 ) Cellular functions regulated by SRC family Kinase Ann.Rev. Cell Dev. Biol. 13: 513-609



## 2.2 - Factores que afectam o crescimento e desenvolvimento dos músculos do esqueleto

- Lone, K.P., ( 1997 ) Natural sex steroids and their xenobiotic analogs in animal production: Growth, Carcass quality, Pharmacokinetics, metabolism, mode of action, residues, methods and Epidemiology in Critical Reviews in Food Science and Nutrition, 37 ( 2 ): 93-209

## 2.3 - Endocrinologia dos peixes e produção animal

- Bjornsson, B.Th. ( 1997 ) the biology of salmon growth hormone: From daylight to dominance Fish Physiology and Biochemistry 17,1-6: 9-24
- Bjornsson, B.T., Valur Stefanson, G., Berge, A.I. et al ( 1998 ) Circulating growth hormone levels in atlantic salmon smolts following seawater transfer: effects of photoperiod regime, salinity, duration of exposure and season Aquaculture 168: 121-137
- Gomez, J.M., Weil, C., et al ( 1999 ) Growth hormone ( GH ) and Gonadotrofin subunit Gene expression and pituitary and plasma changes during spermatogenesis and oogenesis in rainbow trout ( *Oncorhynchus mykiss* ) General and comparative endocrinology, 113, 3; 413-428
- Holloway, A.C., Leatherland, J.F. ( 1998 ) Neuroendocrine regulation of growth hormone secretion in teleost fishes with emphasis on the involvement of gonadal sex steroids Reviews in Fish Biology and Fisheries 8,4: 409-429
- Huang, Y.S., Rousseau, K. Belle, N.L., et al ( 1998 ) Insulin-like growth factor I ( IGF-1 ) stimulates gonadotrophin production from eel pituitary cells: a possible metabolic signal for induction of puberty Journal of Endocrinology 159-1
- Lin, Hao-Ran ( 1994 ) Neuroendocrine application in Aquaculture. Processing of the International Symposium on Biotechnology Applications in aquaculture-December 5-10<sup>th</sup> 1994 Taipei Taiwan
- Melamed, P., Rosenfeld, H., Elizer, A., e Yaron, Z., ( 1998 ) Endocrine regulation of gonadotrofin and growth hormone gene transcription in fish Comparative-Biochemistry and Physiology 119C:3: 325-338
- Melamed, P., Yaron, Z., ( 1999 ) calcium ionophores lead to apoptotic-like changes in Tilapia pituitary cells General and Comparative Endocrinology-114-1: 19-27
- Oommen, O.V., Johnson, B., ( 1998 ) Metabolic effects of ovine growth hormone in a teleost, *Anabrus testudineus*. Trends in Comparative Endocrinology and Neurobiology: from molecular to integrative Biology Vaudry H., Tononn, M.C., Roules E.W., e Loop A., ( eds ) Annals of the New York Academy of Sciences, 1998 vol.839 pag. 380-381
- Park, R., Shepherd, B.S., et al ( 2000 ) Effects of homologous pituitary hormone treatment on serum insulin-like growth factor-binding proteins ( IGF-BPs ) in hypophysectomized tilapia, *Oreochromis mossambicus*, with special reference to a novel 20kDa IGF-BP General and Comparative Endocrinology 117:3: 404-412
- Sundell, K., Dellefors, C., Björnsson, B.T., ( 1998 ) Wild and Patchery-reared brown trout, *Salmo Trutta*, differ in smolt related characteristics during parr-smolt transformation Aquaculture 167: 53-65

- Rousseau, K., LeBelle, N., Vidal, B., et al ( 1999 ) Neuroendocrine regulation of growth hormone ( GH ) release in teleosts : Phylogenetic conservations and variating Reproductive Physiology of Fish Norberg, B., Kjesby, U.S., Tarange, G.L., Anderson, E., Stefanson, S.O., ( eds ) University of Bergen Dept. Of Fish and Mari. Biol. 2000 pag.59
- Seidelin, M., Maolsem, S.s. ( 1999 ) Endocrine control of Na super (+), K super(+) ATPase and chloride cell development in brown trout ( salmo trutta ) interaction of insulin like growth factor I eith prolactin and growth hormone Journal of Endocrinology 162-1
- Sundell, K., Dellefors, C., Björnsson, B.T., ( 1998 ) Wild and Patchery-reared brown trout, Salmo Tturtla, differ in smolt related characteristics during parr-smolt transformation Aquaculture 167: 53-65