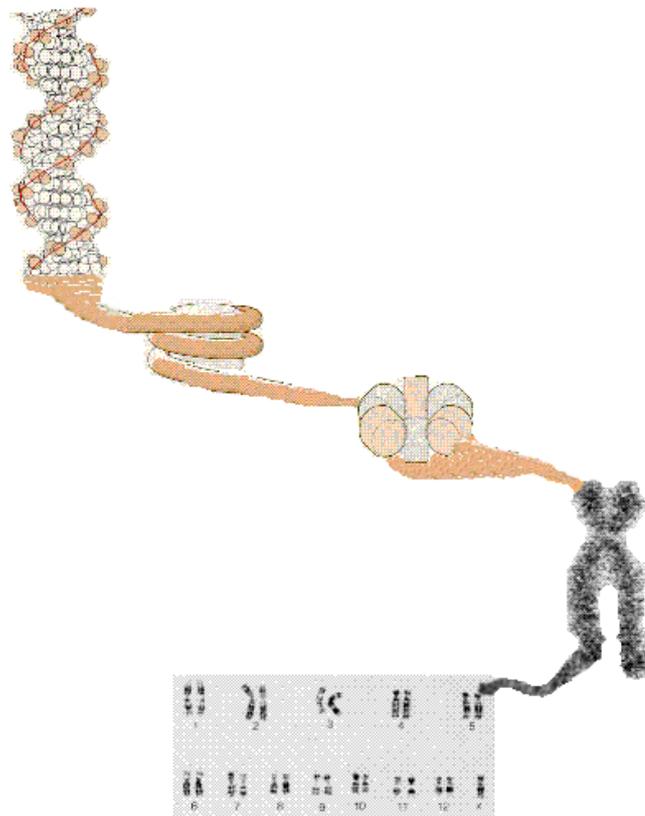


A GENÉTICA MOLECULAR EM HEMATOLOGIA: FERRAMENTA AUXILIAR NO DIAGNÓSTICO, AVALIAÇÃO DE PROGNÓSTICO E “FOLLOW-UP”



José Manuel Baptista Cabeda
Luciana Pinho

Porto , 1996

ÍNDICE

p.

PARTE 1 - CONCEITOS BÁSICOS DE GENÉTICA MOLECULAR 1

1 - O MATERIAL GENÉTICO	2
1.1 - ESTRUTURA QUÍMICA DO MATERIAL GENÉTICO	2
1.2 - A ESTRUTURA DA CROMATINA	6
1.3 - A ESTRUTURA DO CROMOSSOMA	9
1.4 - O CÓDIGO GENÉTICO	9
1.5 - TIPOS DE MUTAÇÕES	12
2 - A ANATOMIA DE UM GENE	13
2.1 - A ORGANIZAÇÃO DOS GENES NO GENOMA	13
2.2 - Estrutura do DNA, sua regulação e função	14
2.2.1- SEQUÊNCIAS REGULADORAS: OS PROMOTORES	15
2.2.1.1 - ORGANISMOS PROCARIOTAS	15
2.2.1.2 ORGANISMOS EUCARIOTAS	15
2.2.2 - SEQUÊNCIAS REGULADORAS: OS ENHANCERS	16
2.2.3 - SEQUÊNCIAS REGULADORAS: OS TERMINADORES	17
2.2.4 - EXONS E INTRONS: ORGANIZAÇÃO ESTRUTURAL	18
2.3 - ESTRUTURA DOS VÁRIOS TIPOS DE RNA, E SUA FUNÇÃO	19
3 - A FISILOGIA DO GENE	20
3.1 - A TRANSCRIÇÃO DO DNA	20
3.2 - A INSTABILIDADE DO mRNA	21
3.3 - TRANSCRIÇÃO REVERSA. OS RETROVIRUS	21
3.4 - A TRADUÇÃO (síntese proteica)	22
4 - A DIVERSIDADE IMUNOLÓGICA: REARRANJO GENÓMICO	25
4.1 - Introdução	25
4.2- Aspectos genéticos	25
4.2.1 - Estrutura somática dos genes do TCR	25
4.2.2 - Mecanismo de rearranjo somático dos genes do TCR	26
4.3 - O TCR na população normal	28
4.3.1 - Polimorfismos do TCR	28
5 - METODOLOGIA DE ESTUDO EM GENÉTICA MOLECULAR	31
5.1 - PREPARAÇÃO DE DNA E RNA	31
5.2 - ANÁLISE DE DNA	31
5.2.1 - ELECTROFORESE	31
5.2.2 - SHOUTHERN BLOT E HIBRIDAÇÃO	32

5.2.3 - DOT E SLOT BLOT	33
<hr/>	
5.4 - MANIPULAÇÃO ENZIMÁTICA DE DNA E RNA	33
5.4.1 - ENDONUCLEASES DE RESTRIÇÃO	33
<hr/>	
5.4.1.1 - Sistemas R-M tipo II	33
<hr/>	
5.4.1.2 - Sistemas R-M tipo IIs	34
<hr/>	
5.4.1.3 - Montar uma reacção de restrição	34
<hr/>	
5.4.2 - MODIFICAÇÃO E MARCAÇÃO DE ÁCIDOS NUCLEICOS	35
<hr/>	
5.4.2.1 - RADIOISÓTOPOS	35
<hr/>	
5.4.2.2 - POLIMERASES DO DNA	36
<hr/>	
5.4.2.3 - FOSFATASES E QUINASES	36
<hr/>	
5.4.3 - POLIMERASE CHAIN REACTION (PCR)	37
<hr/>	
5.4.3.1 - PRINCÍPIOS GERAIS DO PCR	37
<hr/>	
5.4.3.2 - RT-PCR	38
<hr/>	
5.4.4 ANÁLISE DE MUTAÇÕES	38
<hr/>	
5.4.4.1 - ANÁLISE DE MUTAÇÕES CONHECIDAS	38
<hr/>	
5.4.4.2 - ANÁLISE DE NOVAS MUTAÇÕES	39
<hr/>	

PARTE 2 - ÁREAS DE INTERVENÇÃO DA GENÉTICA MOLECULAR EM HEMATOLOGIA

	41
6 - INTRODUÇÃO	42
7 - HEMOPATIAS MALIGNAS	43
7.1 - Detecção de anomalias genéticas	43
7.1.1 - t(9;22) (q34;q11): BCR/ABL	44
7.1.2 - t(15;17) (q22;q21): PML/RARA	45
7.1.3 - O Linfoma de Burkitt e a translocação t(8;14)(q24;q32)	46
7.1.4 - t(14;18)(q32;q21): BCL2/IGH	46
7.1.5 - Mutações de p53 em CLL	47
7.2 - Detecção de monoclonalidade	47
8 - DETECÇÃO E CARACTERIZAÇÃO DE VÍRUS	49
8.1 - O vírus da Hepatite B (HBV)	49
8.2 - O Vírus da Hepatite C (HCV)	49
8.3 - O vírus Linfotrópico Humano (HTLV-I e HTLV-II)	49
8.4 - O vírus do síndrome da imunodeficiência adquirida (HIV)	50
8.5 - O vírus do Epstein bar (EBV)	50
9 - ESTUDO DE DOENÇAS GENÉTICAS	52
9.1 - DOENÇAS GENÉTICAS DO GLÓBULO RUBRO	52
9.1.1 - Anemias Não esferocíticas Congénitas	52
9.1.1.1 - Deficiência em Glucose-6-fosfato desidrogenase	52
9.1.1.2 - Deficiência em piruvato quinase	53
9.1.1.3 - Deficiência em δ -aminolevulinato sintetase (Anemia sideroblástica)	53
9.1.2 - Talassémias (anomalias das α e β -globinas)	54
9.1.3 - Esferocitose e eliptocitose hereditária	56
9.1.3.1 - Esferocitose Hereditária (deficiências de ankirina e banda 3)	56
9.1.3.2 - Anemia hemolítica hereditária (deficiência de proteína 4.2)	56
9.1.3.3 - Eliptocitose e poiquilocitose Hereditária (deficiências de α - e β -espectrina; proteína 4.1 e gricoforina C)	56
9.2 - DOENÇAS GENÉTICAS EM HEMOSTASE	58
9.2.1 - Resistência à proteína C Activada (mutação FV-Leiden)	58
9.2.2 - Doença de von Willebrandt (Mutações no gene do vWF)	59

9.2.3 - Trombose familiar (Mutações nos genes da Antitrombina III, Proteína C e Proteína S)	60
9.2.4 - Hemofilias (Mutações nos genes dos factores VIII e IX)	60
9.3 - HEMOCROMATOSE	61
9.3.1 - Estudos de marcadores genéticos no <i>locus</i> do HLA	62
9.3.2 - Estudos dos IRE e IRP	62
9.3.3 - Estudos Genéticos do Repertório da Célula T	63
PARTE 3 - MANUAL PRÁTICO	65
10 - HEMOPATIAS MALIGNAS	66
10.1 - Anomalias de Cariótipo. t(9;22) (q34;q11): BCR/ABL	66
11 - DETECÇÃO E CARACTERIZAÇÃO DE VIRUS	67
11.1 - Caracterização genotípica do HCV67	
12 - DOENÇAS GENÉTICAS	69
12.1 - Resistência à proteína C Activada (Mutação FV-Leiden)	69
12.2 - Caracterização da mutação de β -globina na β -talassémia	70

Índice de Figuras

Figura 1 - As bases azotadas que entram na composição dos nucleótidos.....	1
Figura 2 - As pentoses são o componente dos ácidos nucleicos que definem o seu tipo. A desoxirribose entra na composição do DNA, enquanto a ribose compõe o RNA	1
Figura 3 - Os nucleótidos unem-se por ligações fosfodiéster para dar origem aos ácidos nucleicos.....	1
Figura 4 - O DNA é formado por duas cadeias com orientação antiparalela, com as bases de cada uma das cadeias a hibridarem entre si.....	1
Figura 5 - Estruturas possíveis para a dupla hélice do DNA	1
Figura 6 - Tipos de cromatina nuclear	1
Figura 7 - a dupla hélice de DNA dá duas voltas ao núcleo central de proteínas do nucleosoma	1
Figura 8 - A organização do DNA nos nucleosomas coloca próximas sequências de DNA distantes na sequência linear.....	1
Figura 9 - Organização estrutural das histonas no nucleosoma	1
Figura 10 - A compactação das histonas na fibra de DNA de 10nm	1
Figura 11 - A organização dos nucleosomas na fibra de DNA de 30nm.	1
Figura 12 - Mutações frameshift e seus efeitos. Note-se que as inserções e as deleções podem anular-se mutuamente, fora da zona entre as duas mutações	1
Figura 13 - Diagrama simplificado da estrutura somática dos genes do TCR humano. Os elementos génicos representados por caixas brancas são pseudogenes (Adaptado de : * Toyonaga et al, 1985; ** LeFranc et al, 1989; * Yoshikai, 1991).....	2
Figura 14 - Equação padrão para a recombinação V(D)J. Os sinais de junção são indicados por triângulos e os segmentos codificantes por quadrados (Extraído de Lewis, 1994)	2
Figura 15 - Mecanismo proposto para a origem dos nucleótidos P (Adaptado de Lewis, 1994).....	2
Figura 16 - Sensibilidade de várias metodologias na detecção da Doença Residual Mínima (10 ¹ -10 ⁶ = detecção de 1 célula maligna em 10 - 10 ⁶ células normais)	4
Figura 17 - Os genes BCR e ABL normais, e as translocações que originam as proteínas p190 e p210 do gene químera BCR-ABL. A proteína p210 é característica da CML, sendo a p190 a proteína BCR-ABL encontrada na maioria dos casos de ALL (ver texto).....	4
Figura 18 - A localização cromossómica e estrutura normal dos genes PML e RAR α , e a translocação t(15;17)(q22;q21) que origina o gene químera PML-RAR α	4
Figura 19 - Diagrama mostrando os eventos genéticos geradores de uma das três translocações encontradas no Linfoma de Burkitt. O oncogene c-MYC está normalmente localizado no braço longo(q) do cromossoma 8. A translocação t(8;14) coloca este oncogene junto ao loc.....	4
Figura 20 - Diagrama mostrando os cromossomas 14 e 18 normais, bem como os cromossomas resultantes da translocação t(14;18)(q32;q21), a qual envolve os genes BCL-2 (18q21) e IgH (14q32).	4
Figura 21 - Gel de sequenciação do exon 5 do gene ALAS2, indicando uma transversoão C para A na posição 547, prevendo a substituição de uma Phe por uma Leu no aminoácido 165 da enzima	5
Figura 22 - Os clusters dos genes da globina nos cromossomas 11 e 16 (a). Durante a vida enbrionica, fetal e adulta, os genes activados e suprimidos são diferentes (b). As diferentes cadeias da globina são sintetizadas independentemente, associando então para for.....	5
Figura 23 - Posições das mutações do gene da β -globina mais frequentes na zona mediterrânica.	5
Figura 24 - Distribuição geográfica de algumas mutações na β -globina	5
Figura 25 - Estratégia de identificação da mutação FV-Leiden. O exon 10 do factor V é amplificado por PCR. O fragmento amplificado contém um local de restrição para Hind III no caso de existir a mutação, mas não no caso do gene não mutado.	5

Índice de Tabelas

Tabela 1 - Características dos tipos de hélice que o DNA pode tomar.....	1
Tabela 2 - Código genético: significado dos 64 codons	1
Tabela 3 - Exemplos de excepções à universalidade do código genético	1
Tabela 4 - Resumo de polimorfismos identificados em genes responsáveis por doenças hemorrágicas e trombóticas hereditárias	5

**CONCEITOS BÁSICOS
DE GENÉTICA
MOLECULAR**

1 - O MATERIAL GENÉTICO

1.1 - ESTRUTURA QUÍMICA DO MATERIAL GENÉTICO

Os ácidos nucleicos são a base da informação genética

Foi apenas em 1944 que Griffith demonstrou que a hereditariedade era transmitida pelos ácidos nucleicos. A experiência realizada demonstrou que a capacidade de matar um ratinho era conferida a uma estirpe bacteriana não virulenta, pelo DNA de uma outra estirpe bacteriana. Em 1952, Hershey e Chase demonstraram que o fago T2 (um vírus que infecta bactérias) transmite os seu código genético à bactéria infectada, através da injeção do seu DNA na bactéria, alargando assim o número de organismos que demonstradamente utilizam o DNA como registo genético. Sabemos hoje que com a excepção de alguns tipos de vírus, todos os organismos utilizam o DNA como portador da sua informação genética. Os vírus que fogem a esta regra utilizam o RNA para o mesmo efeito. Com a aceitação generalizada por volta dos anos 50 de que a informação genética residia no DNA, a grande questão passou a ser o mecanismo de armazenamento dessa informação. Com efeito, nesta altura não se compreendia como é que um polímero tão simples (apenas constituído por 4 tipos de unidades diferentes) e que se pensava ser homogéneo em toda a sua extensão podia codificar a enorme variedade de proteínas que compunham os organismos. Para tal, houve necessidade de elucidar de forma precisa a estrutura dos ácidos nucleicos.

Os ácidos nucleicos são polímeros de nucleótidos

Sabemos hoje que os ácidos nucleicos são polímeros de nucleótidos. Cada nucleótido contém um anel heterocíclico de carbono com 5 átomos de azoto (a base nitrogenada), 1 anel de 5 carbonos (uma pentose) e um grupo fosfato. As bases nitrogenadas são de 2 tipos: purinas e pirimidinas, sendo o número total de bases disponível de cinco (fig. 1). No entanto, cada tipo de ácido nucleico utiliza apenas 4 das cinco bases: o DNA contém Adeninas (A), Timidinas (T), Guaninas (G) e Citosinas (C); enquanto o RNA contém Adeninas (A), Uracilos (U), Guaninas (G) e Citosinas (C). As pentoses encontradas nos ácidos nucleicos são de 2 tipos: 2-desoxirriboses e riboses (Fig. 2) dando origem ao Ácido Desoxirribonucleico (DNA) e ao Ácido Ribonucleico (RNA).

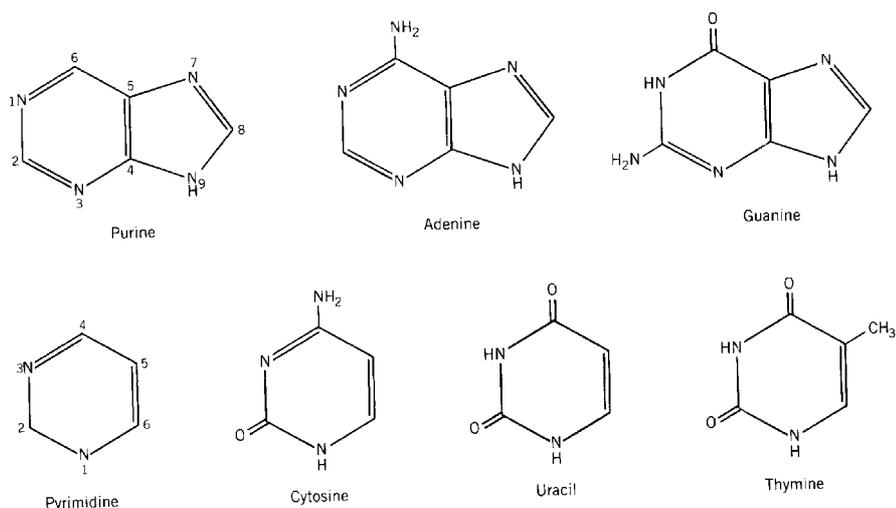


Figura 1 - As bases azotadas que entram na composição dos nucleótidos.

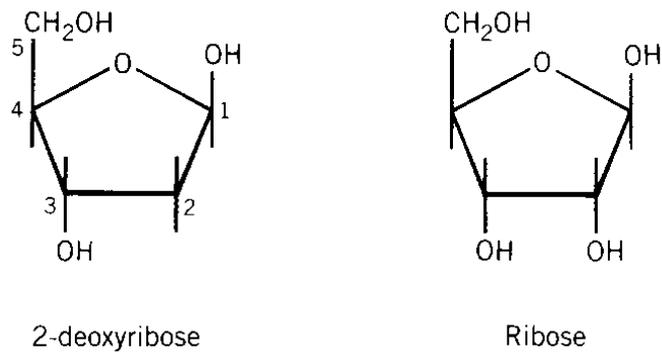


Figura 2 - As pentoses são o componente dos ácidos nucleicos que definem o seu tipo. A desoxirribose entra na composição do DNA, enquanto a ribose compõe o RNA .

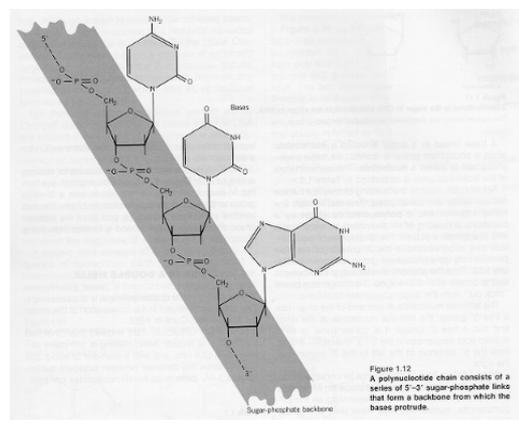


Figura 3 - Os nucleótidos unem-se por ligações fosfodiester para dar origem aos ácidos nucleicos.

Os nucleótidos são compostos por uma pentose, uma base nitrogenada e um grupo fosfato

Os nucleótidos são as unidades de construção dos ácidos nucleicos, sendo unidas por uma ligação 5'-3': o carbono 5' da pentose de um nucleótido une-se ao carbono 3' da pentose do nucleótido seguinte por uma ponte fosfodiester, ficando a base nitrogenada exterior ao esqueleto da ligação (Fig.3).

O DNA é uma dupla hélice

Em 1953, uma importante descoberta realizada por Watson & Crick transformou a visão do material genético. Dados de difracção de raios X mostraram que o DNA tem a forma de uma hélice regular. Das dimensões obtidas para a hélice na difracção de raios X, e da densidade do DNA, inferiu-se então que a hélice era composta por duas cadeias polinucleotídicas, com as bases de cada cadeia viradas para o interior da hélice. As bases de cada hélice emparelham de tal modo que uma purina se opõe sempre a uma pirimidina. Estes dados, conjugados com a observação anterior de Chargaff indicando que independentemente da quantidade de cada base, a proporção G:C e A:T é sempre a mesma no DNA, indicam que G emparelha com C e A com T na dupla hélice do DNA. Watson & Crick propuseram que o emparelhamento não se realizava por ligação covalente, mas por pontes de hidrogénio entre as bases nitrogenadas (Fig. 4). Para tal, as 2 cadeias devem orientar-se de modo antiparalelo (Fig. 4). Obteve-se assim para o DNA o modelo ilustrado na figura 5A.

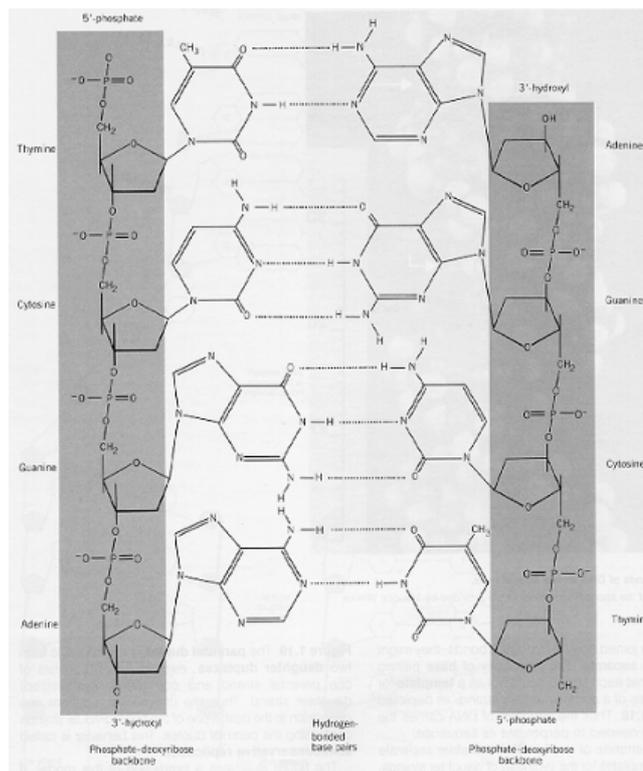


Figura 4 - O DNA é formado por duas cadeias com orientação antiparalela, com as bases de cada uma das cadeias a hibridarem entre si.

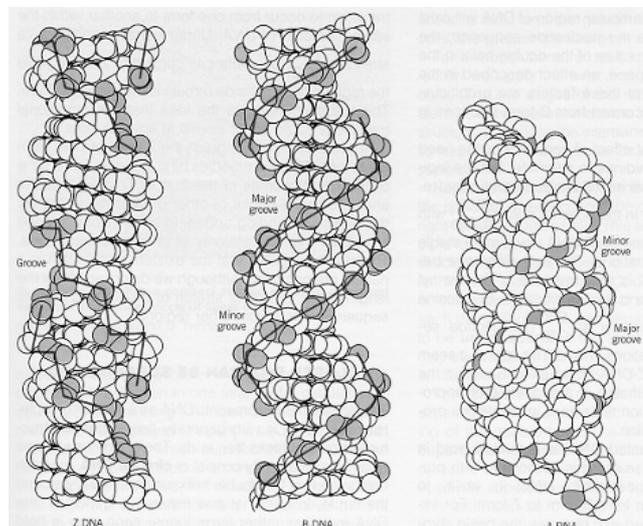


Figura 5 - Estruturas possíveis para a dupla hélice do DNA

A estrutura do DNA identificada por Watson & Crick, e ilustrada na figura 5A é a que em situações fisiológicas é mais frequente. No entanto nem todo o DNA da célula se encontra nesta estrutura, e certamente *in vitro* é possível manipular as condições do meio, favorecendo outras conformações. Na tabela 1 encontram-se sumariadas as características das 4 conformações teoricamente possíveis para a conformação dos ácidos nucleicos, podendo na figura 5B e C ver-se comparativamente a conformação prevista.

Tabela 1 - Características dos tipos de hélice que o DNA pode tomar

Tipo de Hélice	N.º Bases por Volta	Rotação por par de Bases	Elevação por par de bases	Diâmetro da hélice
A	11	+32.7°	2.56 Å	23 Å
B	10	+36°	3.38 Å	19 Å
C	9.33	+38.6°	3.32 Å	19 Å
Z	12	-30°	3.71 Å	18 Å

1.2 - A ESTRUTURA DA CROMATINA

A dimensão do material genético no Homem coloca o problema de como conseguir compactar 1,8m de DNA num núcleo que pode ser tão pequeno como 6µm ($6 \times 10^{-6}m$). Este empacotamento tem ainda que ser flexível já que deve mudar ao longo do ciclo celular, aumentando durante as mitoses de tal modo que os cromossomas se tornam individualizados e visíveis ao microscópio óptico.

A cromatina é composta pela eucromatina e pela heterocromatina

Durante a maior parte do ciclo celular a cromatina pode ser dividida em dois tipos de material genético (Fig. 6): a **Eucromatina** é a que ocupa a maior região do núcleo, sendo composta por material muito menos compactado que os cromossomas; a **Heterocromatina** é composta por material muito compactado, formando fibras (encontra-se num estado intermédio entre a compactação dos cromossomas e a relativa descompactação da eucromatina). a heterocromatina e a eucromatina não representam fibras de DNA diferentes, já que as mesmas fibras passam pelas duas zonas do núcleo. Constituem assim partes das fibras com diferentes estados de condensação: a heterocromatina é constituída por regiões do DNA que não são habitualmente expressas na célula em causa, enquanto os genes expressos se localizam na eucromatina (muito embora os genes na eucromatina não estejam todos a ser expressos).

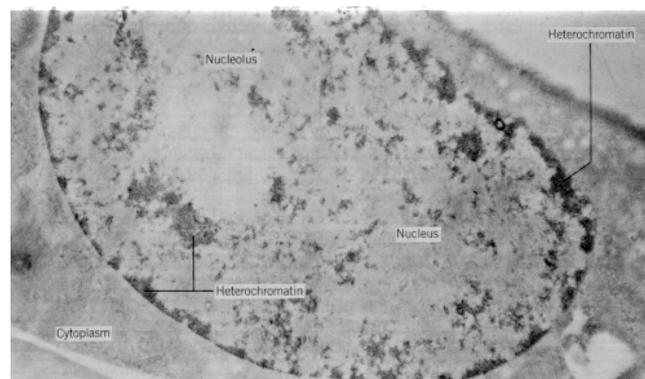


Figura 6 - Tipos de cromatina nuclear

Os nucleosomas são a estrutura básica da cromatina

Em 1974, foi descoberta a estrutura básica de organização da cromatina em todos os eucariotas. Esta subunidade organizativa básica (nucleosoma) contém cerca de 200 bp de DNA, organizados por um octâmero de proteínas pequenas e básicas (histonas) numa estrutura tipo rosário em que o DNA se localiza no exterior das

“contas”, e as proteínas no seu interior (Fig.7). Esta organização explica porque os locais de ligação a proteínas se encontram por vezes tão espaçados na sequência do DNA (Fig. 8). O octâmero de histonas é constituído por 2 cadeias de cada uma das histonas H1, H2A, H2B e H3, existindo ainda, por vezes, uma 5ª histona (H1) a estabilizar as 2 voltas de DNA ao octâmero (Fig. 9).

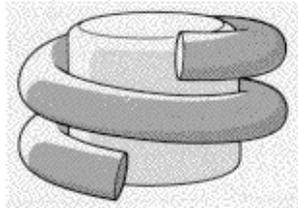


Figura 7 - a dupla hélice de DNA dá duas voltas ao núcleo central de proteínas do nucleossoma

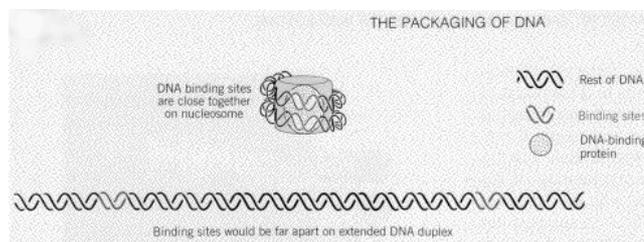


Figura 8 - A organização do DNA nos nucleossomas coloca próximas sequências de DNA distantes na sequência linear

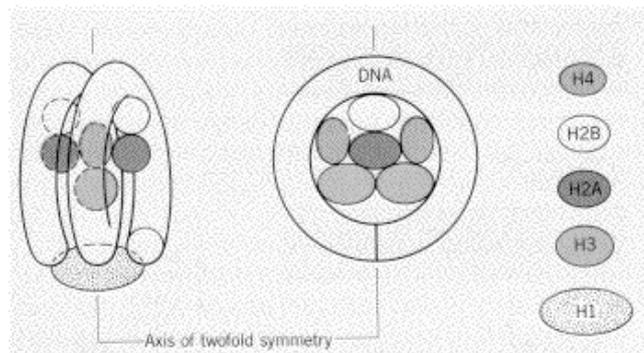


Figura 9 - Organização estrutural das histonas no nucleossoma

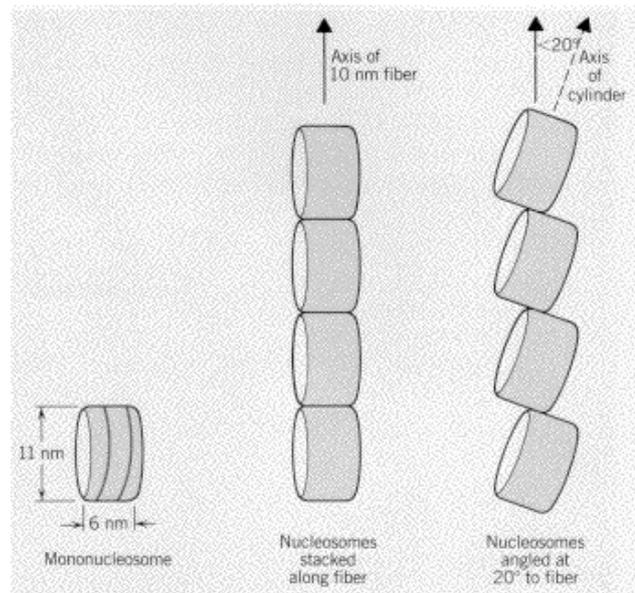


Figura 10 - A compactação das histonas na fibra de DNA de 10nm

A cromatina possui fibras de 10nm e de 30nm

A análise da cromatina ao microscópio electrónico revelou a existência de 2 tipos de fibras: a fibra de 10nm e a 30nm. A fibra de 10nm é essencialmente 1 sequência continua de nucleosomas (Fig. 11). Esta fibra ocorre em condições de baixa força iónica, e na ausência de histonas H1. Em condições de alta força iónica e na presença da histona H1, forma-se a fibra de 30nm, a qual é essencialmente um enrolamento de 6 nucleosomas por volta (Fig. 11).

O empacotamento do DNA e a transcrição

A transcrição (cópia dos genes em mRNA), como veremos na secção 3.1, envolve a deslocação no DNA de uma complexa maquinaria enzimática, e inclui a abertura da dupla cadeia do DNA. Este facto, não é compatível com um elevado grau de empacotamento das fibras do DNA, pelo que se compreende que os genes transcripcionalmente activos se localizem na eucromatina. No entanto, os resultados experimentais indicam que a estrutura dos genes transcripcionalmente activos envolve o empacotamento em nucleosomas, ainda que seja necessário admitir que durante a transcrição estes sejam temporariamente “desmontados” pela maquinaria enzimática.

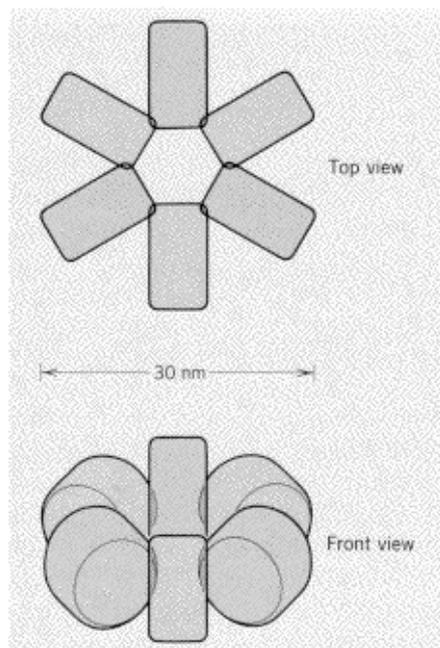


Figura 11 - A organização dos nucleossomas na fibra de DNA de 30nm.

1.3 - A ESTRUTURA DO CROMOSSOMA

Um cromossoma, quando na conhecida forma de X é composto por um par de cromátídeos (cada metade do X constitui um cromátídeo). Cada cromátídeo é composto por uma única fibra de DNA altamente compactado. Assim, em cada cromossoma estão representadas as 2 cópias dos genes do cromossoma (1 em cada cromátídeo). Como a mitose envolve a duplicação de cada cromossoma, para que as células filhas possuam também elas 2 cópias de cada *locus* génico, durante a mitose podem ser observados dois cromossomas de cada tipo (logo nesta fase do ciclo celular a célula possui 4 cópias de cada *locus* génico: estado $4n$). Após a segregação de cromátídeos que ocorre na fase final da meiose, cada núcleo filho regressa ao estado normal $2n$.

Os cromossomas são constituídos por 2 braços: um braço curto (p) e um braço longo (q), entre os quais existe o centrómero, o qual permite ao cromossoma ligar-se aos microtúbulos que durante a mitose vão organizar a segregação dos cromátídeos.

O telómero

Uma outra característica essencial em todos os cromossomas é o telómero. Esta estrutura situa-se na extremidade de cada cromossoma, contendo um elevado número de curtas sequências repetitivas, possivelmente para facilitar uma eficiente duplicação de todo o material genético. Certos autores especulam que a progressiva danificação dos telómeros em cada divisão celular constitui um relógio biológico de contagem do número de divisões.

1.4 - O CÓDIGO GENÉTICO

Dogma central da genética: um gene, um polipéptido

A descoberta do código genético pretendeu responder à questão já por nós formulada (Cap. 1.1) sobre o mecanismo que permite aos ácidos nucleicos, com uma estrutura baseada em apenas quatro tipos de nucleótidos, conter a informação que codifica um enorme número de proteínas, as quais possuem 20 tipos de

aminoácidos. A elucidação do código genético pretendeu ainda explicar como é que a expressão génica é regulada. No entanto, antes de esta questão poder ser estudada, era necessário estabelecer definitivamente a veracidade do dogma central da genética: Um gene - uma cadeia polipeptídica. Uma característica essencial do DNA é que a sua estrutura básica é independente da sequência (ao contrário das proteínas cuja conformação é directamente dependente da sequência). Assim, a sequência do DNA não parece ser importante devido à conformação, mas porque codifica uma sequência bem definida de aminoácidos. Note-se que o próprio conceito de que uma proteína contém sequências bem definidas de aminoácidos data dos anos 50 (a caracterização da insulina por Sanger), e portanto é estabelecida sensivelmente na mesma altura que se estuda a informação genética. A esta relação entre a sequência do DNA e a sequência proteica correspondente chamou-se código genético.

O código genético é degenerado

Como vimos, a sequência nucleotídica tem que conter informação suficiente para codificar aminoácidos diferentes. Como só há quatro tipos de nucleótidos no DNA, um cálculo simples indica que são necessários 3 nucleótidos (um triplete ou codão) para codificar um aminoácido. As combinações possíveis com três nucleótidos são $4^3=64$, pelo que o código genético é degenerado, isto é, mais do que um triplete deve codificar o mesmo aminoácido (Tabela 2).

Degenerância da 3ª base

Podem agrupar-se os codões segundo o aminoácido que codificam (Tabela 2). Quando tal é realizado, pode observar-se que com frequência, a base na 3ª posição não é significativa, porque os 4 codões com as mesmas 1ª e 2ª bases codificam o mesmo aminoácido (Tabela 2). Por vezes apenas distingue entre uma pirimidina e uma purina a 3ª posição. A esta especificidade reduzida na 3ª base chama-se degenerância da 3ª base. Esta característica, em conjunto com a tendência para aminoácidos semelhantes (isto é polares, hidrofóbicos, etc.) serem codificados por codões relacionados minimiza o efeito das mutações.

Tabela 2 - Código genético: significado dos 64 codons

		SEGUNDA BASE			
		U	C	A	G
PRIMEIRA BASE	U	UUU } <i>Phe</i> UUC } UUA } <i>Leu</i> UUG }	UCU } UCC } <i>Ser</i> UCA } UCG }	UAU } <i>Tyr</i> UAC } UAA } <i>STOP</i> UAG }	UGU } <i>Cys</i> UGC } UGA → <i>STOP</i> UGG → <i>Trp</i>
	C	CUU } CUC } <i>Leu</i> CUA } CUG }	CCU } CCC } <i>Pro</i> CCA } CCG }	CAU } <i>His</i> CAC } CAA } <i>Gln</i> CAG }	CGU } CGC } <i>Arg</i> CGA } CGG }
	A	AUU } AUC } <i>Ile</i> AUA } AUG → <i>Met</i>	AAU } AAC } <i>Thr</i> AAA } AAG }	AAU } <i>Asn</i> AAC } AAA } <i>Lys</i> AAG }	AGU } <i>Ser</i> AGC } AGA } <i>Arg</i> AGG }
	G	GUU } GUC } <i>Val</i> GUA } GUG }	GCU } GCC } <i>Ala</i> GCA } GCG }	GAU } <i>Asp</i> GAC } GAA } <i>Glu</i> GAG }	GGU } GGC } <i>Gly</i> GGA } GGG }

Três codões não codificam aminoácidos. Como se pode observar na tabela 2, estes codões (UUA, UAG e UGA) indicam o fim da sequência génica, sendo por isso chamados de codões stop.

O código genético foi inicialmente estudado na bactéria *E.Coli*, pelo que a universalidade deste necessitou de estudo aturado. Sabemos hoje, que genericamente o código genético é similar em todos os organismos vivos estudados. As exceções conhecidas são representadas por pequenas alterações em algumas espécies de microorganismos, e no código genético mitocondrial, o qual possui algumas particularidades em alguns organismos (Tabela 3).

Tabela 3 - Exemplos de exceções à universalidade do código genético

Organismo	Codon	Significado Provável na mitocondria	Significado habitual
Todos	UGA	Triptofano	Terminação
Levedura	CUA	Treonina	Leucina
Mosca da fruta	AGA	Serina	Arginina
Mamíferos	AGA	Terminação	Arginina
	AUA	Metionina	Isoleucina

Mutações frameshift

1.5 - TIPOS DE MUTAÇÕES

Uma vez que o código genético é lido em tripletos não sobreponíveis, a inserção ou remoção de um nucleótido causa uma alteração na fase de leitura, alterando os codões subsequentes. Este tipo de mutação é denominado em Inglês “frameshift”. Mutações deste tipo são susceptíveis de reverterem através da mutação inversa, isto é, se a primeira mutação foi uma inserção e a segunda uma deleção, ou vice-versa, apenas a zona do gene situada entre as duas mutações se encontra mutada. A segunda mutação, é denominada supressora, já que suprime o efeito da primeira, limitando a zona atingida (fig. 12)

Mutações pontuais

As **mutações pontuais** são mutações que ocorrem devido à substituição de um nucleótido por outro. A forma mais frequente de mutações pontuais é a **transição**, a qual ocorre quando uma pirimidina é substituída por outra, ou uma purina por outra. A **transversão** é menos frequente, e implica a substituição de uma pirimidina por uma purina, ou vice-versa. As mutações pontuais podem ser de 3 tipos, de acordo com o efeito que provocam no aminoácido codificado. Se não afectam o aminoácido codificado são chamadas **silenciosas**, se mudam o aminoácido codificado são chamadas **missense**, e se transformam o codão num codão stop são chamadas **nonsense**.

As mutações pontuais foram durante muito tempo consideradas as principais causas de mutações. Sabe-se no entanto hoje, que as deleções são também muito frequentes, representando uma significativa porção das mutações identificadas.

Polimorfismos

As mutações podem ser vantajosas, desvantajosas ou neutras, segundo as consequências funcionais que provocam. As mutações neutras, apesar de causarem alteração na sequência não ocasionam mudança funcional. Neste caso, deve falar-se em polimorfismo e não em mutação.

Selvagem	GCU	GCU	GCU	GCU	GCU	GCU	GCU	GCU	GCU	GCU
	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala
Inserção (+)	GCU	GCU	AGC	UGC	UGC	UGC	UGC	UGC	UGC	U
	Ala	Ala	Ser	Cys	Cys	Cys	Cys	Cys	Cys	Cys
Deleção (-)	GCU	GCU	GCU	GCU	GCU	_CUGCUG	CU			
	Ala	Ala	Ala	Ala	Ala	Leu	Leu			
Duplo mutante (+-)	GCU	GCU	AGC	UGC	UGC	_UCUGCU	GCU	GCU		
	Ala	Ala	Ser	Cys	Cys	Ser	Ala	Ala		
triplo mutante (+++)	GCU	GAC	UGC	AUG	CUG	CAU	GCU	GCU	GCU	
	Ala	Asp	Cys	Met	Leu	His	Ala	Ala	Ala	
triplo mutante (---)	GCU	_CUG	CU_C	UGC	U_CU	GCU	GCU			
	Ala	Leu	Leu	Cys	Ser	Ala	Ala			

Figura 12 - Mutações frameshift e seus efeitos. Note-se que as inserções e as deleções podem anular-se mutuamente, fora da zona entre as duas mutações

2 - A ANATOMIA DE UM GENE

2.1 - A ORGANIZAÇÃO DOS GENES NO GENOMA

Tipos de sequências genéticas

O genoma pode, de uma forma genérica, ser classificado em DNA não repetitivo e DNA repetitivo. A abundância relativa dos dois tipos de DNA podem ser experimentalmente determinados, tendo por base a diferente cinética de re-hibridação (DNA repetitivo encontra mais rapidamente uma sequência complementar com quem pode hibridar). O DNA não repetitivo representa sequências únicas, ou seja genes de cópia única no genoma. O DNA repetitivo é constituído por DNA moderadamente repetitivo, representando genes com várias cópias no genoma, e DNA altamente repetitivo. A função do DNA altamente repetitivo permanece até ao momento uma incógnita. Como já vimos, um exemplo deste tipo de DNA é o que existe nos telómeros, onde provavelmente tem a função de estabilizar o cromossoma. Existem no entanto, repetições de pequenas unidades de sequências espalhadas pelo genoma (mini e microssatélites), os quais constituem em pequenas sequências, repetidas um determinado número de vezes. O número de repetições é em muitos casos altamente polimórfico, pelo que estas sequências têm sido utilizados como marcadores no mapeamento genético.

Sequências moderadamente repetitivas: Nos genomas eucarióticos, os genes que existem em cópia única são poucos. Na maior parte dos casos, existem sequências com alguma similaridade, algumas das quais não funcionais (os pseudogenes). A vantagem da existência de mais que uma cópia dos genes é óbvia já que assim os organismos podem conservar uma cópia intacta do gene, mutando a outra, numa tentativa de evoluir. Neste processo de evolução, algumas cópias ficam com a sua funcionalidade comprometida, tornando-se **pseudogenes**. No entanto, uma vez que uma outra cópia funcional existe, nenhum efeito nefasto daí ocorre para o organismo.

Gene-clusters

Um conjunto de genes que descende por duplicação e variação de um gene ancestral é chamado de família génica. Os seus membros podem estar arrançados em grupos sequenciais (“gene clusters”), dispersos no genoma (muitas vezes mesmo em cromossomas diferentes), ou numa combinação de ambos os arranjos. Os “gene clusters” podem conter desde 2 até centenas de genes idênticos, alinhados em sequência. A dispersão dos genes ocorre por translocação de um gene após a duplicação. Os membros de um “gene cluster” têm função similar, mas podem ser expressos em tipos celulares diferentes ou em diferentes condições (Ex. Gene da globina). Em alguns casos, o gene cluster responde à grande necessidade de proteínas ou de RNA (ex.: rRNA e histonas).

DNA satélite

Sequências altamente repetitivas: As sequências altamente repetitivas tomam a forma de sequências muito curtas, repetidas muitas vezes em sequência. Formam-se assim blocos de material genómico, consistindo cada bloco em longas repetições de uma unidade. Em alguns casos as unidades são rigorosamente iguais, noutros são relacionadas. A repetição sequencial de unidades de sequência forma blocos de DNA com características físicas distintas do resto do genoma, o que pode ser utilizado para as isolar. Uma das propriedades físicas do DNA que depende da sequência é a densidade, a qual depende do conteúdo GC. A densidade e

habitualmente determinada mediante a centrifugação do DNA num gradiente de Cloreto de Césio (CsCl). O DNA forma assim bandas correspondentes a sua própria densidade. Quando este procedimento é realizado para DNA genómico eucariota, forma um pico algo largo, consistindo numa mistura de sequências com densidades próximas (a banda principal). Por vezes forma-se ainda um ou mais picos adicionais, de menor intensidade. A este material chama-se o DNA satélite.

O DNA satélite existe no genoma de várias espécies eucariotas, pode ter uma densidade superior ou inferior a banda principal, mas representa habitualmente menos de 5% do genoma total.

O DNA satélite encontra-se frequentemente localizado na heterocromatina, não sendo habitualmente possível encontrar as suas sequências entre o RNA.

Minisatelites e Microsatelites

Nos mamíferos, as sequências que compõem cada satélite mostram divergência apreciável entre as repetições de cada. Habitualmente existem sequências curtas predominantes, mas outras relacionadas com estas, mas contendo adições, substituições e deleções formam o restante satélite. Frequentemente, pode observar-se uma hierarquia nas repetições dos satélites, com uma sequência base a repetir-se, a qual por vezes sofre modificações, as quais por sua vez se repetem também de forma mais ou menos cíclica. Este facto originou uma hierarquia de nomenclatura: DNA satélite, minisatelites, microsatelites.

Colinearidade entre o DNA e a proteína

2.2 - Estrutura do DNA, sua regulação e função

A comparação directa entre a sequência do DNA de um gene, e a sequência da proteína respectiva, permite determinar se o gene e a proteína são ou não colineares: se a sequência do gene corresponde exactamente a sequência de aminoácidos da proteína. Nas bactérias e vírus, a equivalência é perfeita: cada gene contém uma sequência contínua de nucleótidos, cuja sequência e comprimento esta directamente relacionada com a da proteína. Quando falamos em correspondência entre o gene e a proteína, estamos no entanto a simplificar o que realmente se passa. Como veremos mais tarde, um gene não codifica directamente uma proteína, já que a informação tem que passar por um estado intermédio: o RNA. Assim, mesmo o mais simples dos genes tem que conter sequências de vários tipos:

- **sequências reguladoras ou não codificantes:** sequências que permitem a célula controlar que genes estão activos em cada momento, possibilitando assim uma resposta diferenciada dependente das necessidades de cada momento. As sequências reguladoras podem existir em cada extremidade do gene, e em alguns casos estar mesmo bastante distanciadas das sequências codificantes.
- **sequências codificantes:** sequências que são directamente transcritas para RNA, e deste codificadas em proteínas. Note-se que enquanto o DNA é de cadeia dupla, o RNA é de cadeia simples, pelo que apenas uma das cadeias do DNA pode ser idêntica a do RNA (codificante ou +), sendo a outra cadeia complementar do RNA (-).

Como acima foi dito, o gene não é no entanto tão simples nos eucariotas. Ao contrário das bactérias e vírus, nos organismos eucariotas, os genes e as proteínas não são colineares, isto é, a região codificante dos genes (exons) é interrompida a espaços irregulares por sequências não codificantes (introns). Este facto faz com que nos eucariotas, a expressão genica envolva um passo adicional: o splicing do RNA, ou processamento do RNA (com exons e introns) em mRNA.

2.2.1- SEQUÊNCIAS REGULADORAS: OS PROMOTORES

Promotores

Um promotor é uma sequência de DNA, habitualmente na extremidade 5' de um gene, com a função de se ligar a proteínas, e controlar a iniciação da transcrição. As proteínas a que um promotor se deve ligar, são várias, disso dependendo a sua dinâmica funcional. genericamente pode falar-se de proteínas repressoras, proteínas activadoras, e da RNA polimerase. As proteínas repressoras, ao ligar-se ao promotor impedem a ligação da RNA polimerase, impedindo assim o início da transcrição, enquanto a ligação das proteínas activadoras tem o efeito inverso. As propriedades do promotor que lhe conferem afinidade para as diversas proteínas dependem da sua sequência, pelo que esta varia de gene para gene, conferindo aos diversos genes características de regulação diferentes. No entanto, a ligação a polimerase do RNA é universalmente necessária, pelo que deve ser possível encontrar uma sequência "consenso" para os promotores. Esta sequência consenso consiste na sequência mínima comum entre os vários promotores, e deve incluir a sequência absolutamente necessária para a ligação a polimerase do RNA.

2.2.1.1 - ORGANISMOS PROCARIOTAS

Para os procariotas foi possível definir a região 44-50bp "upstream" do ponto de iniciação até 20bp "downstream" com sendo a região que interacciona com a polimerase do RNA, tendo sido definida uma sequência consenso consistindo de vários padrões:

Pribnow Box

Pribnow box ou sequência -10- imediatamente upstream do ponto de iniciação (-18 a -12) existe uma região com a sequência T₈₀A₉₅T₄₅A₆₀A₅₀T₉₆ (os números representam a frequência com que as bases ocorrem). A função desta sequência parece ser a de permitir que após a ligação da polimerase do RNA esta possa iniciar a sua evolução ao longo do gene, possivelmente por permitir a iniciação da abertura da cadeia do DNA (o facto de ter alto conteúdo AT facilita a abertura da dupla hélice).

Sequência -35

Sequência de reconhecimento ou Sequência -35 - O seu nome deriva do facto de esta ser parte da sequência que a polimerase tem que reconhecer, mas que não fica fortemente ligada a esta. A sequência consenso é: T₈₂T₈₄G₇₈A₆₅C₅₄A₄₅. A função desta região parece ser a de conferir a capacidade de ligação a polimerase do RNA.

2.2.1.2 ORGANISMOS EUCARIOTAS

No caso de organismos eucariotas, o estudo dos promotores é bem mais complexo, já que existem não uma RNA polimerase, mas três. A acrescentar a esta

dificuldade, esta o facto de não se conhecer com precisão todos os componentes da maquinaria de transcrição eucariota, pelo que os estudos *In vitro* ficam comprometidos.

A partida 2 particularidades existem nos eucariotas, relativamente ao que se passa nos procariotas: 1) o promotor da polimerase III fica localizado downstream do gene; 2) não é possível conhecer as particularidades do promotor da polimerase I, já que esta transcreve apenas os genes dos rRNA os quais são todos idênticos.

No entanto o promotor da RNA polimerase II, a responsável pela transcrição da maioria dos genes nos eucariotas são conhecidos com alguma profundidade. As principais sequências consenso identificadas nos promotores da RNA polimerase II dos eucariotas são:

TATA Box

TATA BOX - sequência consenso: $T_{82} A_{97} \frac{A_{63}}{T_{37}} A_{83} \frac{A_{50}}{T_{37}}$ Também conhecida por

Hogness box. Trata-se de uma sequência quase universalmente presente em mamíferos, aves, anfíbios e insectos. Posiciona-se a uma distância do ponto de iniciação entre 19 e 27bp. Como pode ver-se da sequência consenso, a TATA Box é constituída quase exclusivamente por AT, sendo as mutações que inserem um GC muito raras. Esta sequência é habitualmente rodeada por sequências ricas em GC, o que pode ser importante para a sua função.

CAAT Box

CAAT BOX - sequência consenso $GG \frac{T}{C} CAATCT$. Esta sequência está

presente em alguns promotores, mas não em todos. A sua distância ao ponto de iniciação ronda os 70 a 80bp.

As análises *In vitro* identificaram uma estrutura semelhante ao promotor bacteriano, imediatamente upstream do ponto de iniciação. No entanto, estudos *In vivo* revelaram a dependência de zonas ainda mais upstream da TATA box. Este componente pode consistir em duas regiões, uma entre -80 e -110 e a outra entre -50 e -70. esta última pode ou não conter a CAAT box. Juntos, estas duas regiões têm uma forte influência na frequência de iniciação, possivelmente por influência a ligação da RNA polimerase II.

Junto ao ponto de iniciação, em redor da TATA box existe um componente que parece não ter influência na frequência de iniciação, antes determinando o ponto de iniciação. Na ausência deste elemento, a transcrição tem uma iniciação errática.

2.2.2 - SEQUÊNCIAS REGULADORAS: OS ENHANCERS

Enhancers

Os promotores eucarióticos são bem mais complexos que dos procariotas. Ao contrário dos promotores procarióticos, e contrariamente ao que até agora assumimos, um promotor eucariótico não funciona só. A sua actividade é enormemente aumentada de acordo com a regulação efectuada por outro tipo de sequências reguladoras: os “enhancers”. Estas sequências são distinguíveis dos promotores devido a duas características essenciais:

- a sua posição relativamente ao promotor é muito variável, podendo ser considerável, e funcionando em qualquer sentido (“upstream” ou “downstream”) e orientação.
- Um enhancer não actua apenas num promotor, podendo interactuar com qualquer promotor colocado na sua área de influência.

Vários vírus contêm enhancers. Destes, os mais perigosos para a célula que o vírus infecta são os enhancers presentes nos retrovírus. Como estes vírus se integram no genoma da célula infectada, a presença de enhancers pode levar à inactivação de um ou mais genes celulares que de outra forma estariam silenciosos. Desta forma, os retrovírus podem de forma indirecta, originar patologias, mesmo no seu estado “dormente”, já que mesmo na ausência de transcrição viral, podem induzir alterações no programa genético da célula infectada.

O modo de funcionamento dos enhancers permanece desconhecido. Foram no entanto colocadas várias possibilidades, entre as quais:

- **Formação de estrutura no DNA em cadeia Z** (ver figura 5). Os enhancers contêm habitualmente uma sequência alternada de pirimidinas-purinas. Esta sequência tem elevada probabilidade de formar uma estrutura em z-DNA. Se, por um lado, o modo como esta estrutura poderia afectar a transcrição não está esclarecido, por outro lado, este mecanismo poderia explicar porque os enhancers funcionam independentemente da sua orientação.
- Ligação do DNA a uma estrutura como a matriz nuclear
- ligação directa à polimerase

2.2.3 - SEQUÊNCIAS REGULADORAS: OS TERMINADORES

Terminadores e anti-terminadores

Quando a polimerase do RNA inicia a transcrição, este prossegue com o complexo enzimático a percorrer o DNA, até que a enzima encontra um sinal para cessar a actividade. Neste ponto, a enzima para de adicionar nucleótidos, liberta a cadeia de RNA nascente, e dissocia-se do DNA. Assim, a terminação envolve a quebra de todas as pontes de hidrogénio entre o DNA e o RNA, e a reassociação da dupla hélice do DNA. A sequência de DNA que dá o sinal para que este processo ocorra chama-se terminador (ou abreviadamente **t**). Em alguns genes procarióticos, existem factores denominados anti-terminadores, que permitem à polimerase continuar a transcrição passando por um terminador, num processo chamado de “read-through”). Assim, a terminação não constitui simplesmente uma forma de terminar a transcrição, mas também uma forma de controlar esta, já que a existência dos anti-terminadores pode determinar a transcrição ou não de determinados genes que se encontrem após o terminador.

Pouco se sabe dos terminadores dos genes eucarióticos. A principal dificuldade no estudo dos terminadores em eucarióticos é a incerteza quanto ao local de terminação da transcrição. Ainda que a maior parte das espécies de mRNA

eucarióticas conhecidas possuam extremidades 3' bem definidas, é muito difícil saber se esta extremidade foi produzida por terminação ou por processamento. No caso dos produtos da polimerase II, o problema é exacerbado pelo extenso processamento que ocorre com a adição da cauda poli-A. Pelo menos em alguns casos foi possível determinar que a extremidade 3' observada no RNA é de facto originada por corte de uma cadeia de RNA mais longa.

Estudos efectuados com sequências de histonas (não poliadeniladas), permitiram verificar que o mRNA termina numa estrutura semicircular ("stem-loop"). Com efeito, mutações que impeçam a formação desta estrutura, impedem a terminação, enquanto que outras mutações que revertam a mesma estrutura, embora com uma sequência diferente, restauram a terminação. Assim, a estrutura parece mais importante que a sequência que a determina.

2.2.4 - EXONS E INTRONS: ORGANIZAÇÃO ESTRUTURAL

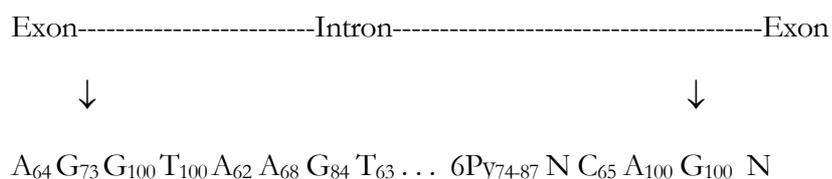
Splicing

Os genes Eucarióticos e procarióticos diferem numa característica essencial. Ao contrário dos genes procarióticos, os genes dos organismos eucarióticos não são contínuos, mas interrompidos. Significa isto, que no meio das sequências codificantes, surgem sequências que têm que ser retiradas do RNA, antes de este poder servir de molde à construção das proteínas. Este processo de transformação que o RNA sofre nos organismos eucarióticos é chamado de processamento, ocorre no núcleo, e como veremos envolve não apenas a remoção das sequências extra ("splicing") como outras transformações químicas.

Os genes eucarióticos são assim formados por dois tipos de sequências transcritas (isto é copiáveis para RNA) os exons e os introns (também chamados de intervening sequences). Os primeiros compõem as sequências que estarão presentes no RNA maduro, sendo os segundos as sequências que serão removidas durante o splicing.

A comparação das sequências nucleotídicas nas extremidades dos exons permite descrever as suas características:

- Não existe homologia entre extensa entre as duas extremidades de um intron, o que exclui a possibilidade da formação de uma estrutura secundária que determine os pontos de corte.
- As junções possuem uma sequência consenso conservada mas curta, a qual pode estar envolvida no processo de splicing:



2.3 - ESTRUTURA DOS VÁRIOS TIPOS DE RNA, E SUA FUNÇÃO

Na célula existem várias formas de RNA, as quais possuem estruturas e funções diferentes:

- **mRNA** - O mRNA ou RNA mensageiro, é a espécie de RNA que transporta a informação para a síntese das proteínas no ribossoma. O mRNA é formado no núcleo na transcrição do DNA, passando ainda por uma fase de processamento antes de atingir o citoplasma na forma madura (mRNA). O processamento efectuado inclui o “Splicing”, isto é a remoção das sequências não codificantes ou introns. Outras alterações introduzidas no processamento que ocorre no núcleo consistem na adição de uma cauda poli-adenina à extremidade 3', e metilação CAP da extremidade 5'. A estrutura CAP resulta da ligação de um G à purina com que a transcrição habitualmente se inicia, ficando este G na orientação inversa, e ligado pelo trifosfato deixado livre pela purina:



G sofre então uma ou mais metilações.

- **tRNA** - o tRNA ou RNA de transporte é um tipo de RNA que se encontra covalentemente ligado a um aminoácido, tendo como função o transporte do aminoácido para o ribossoma, onde este vai ser posicionado com precisão, sempre que o ribossoma estiver a ler um codão complementar do triplete que o tRNA possui (anticodão). As 64 espécies de tRNA (correspondentes aos 64 codões), possuem uma estrutura básica semelhante.
- **rRNA** - trata-se do RNA ribossomal, o qual é como o nome indica um dos componentes dos ribossomas. O rRNA constitui a maior parte da massa do ribossoma, e provavelmente todas as proteínas do ribossoma se associam ao rRNA. Assim, o rRNA forma como que o esqueleto do ribossoma, determinando a posição das várias subunidades proteicas.

3 - A FISILOGIA DO GENE

3.1 - A TRANSCRIÇÃO DO DNA

O RNA é uma espécie central na genética molecular

O RNA é a espécie de ácido nucleico com um papel mais alargado na genética molecular dos organismos. Não só o RNA tem o papel mais “mediático” de mensageiro, mas é também a espécie que assegura a descodificação da informação genética (o rRNA e o tRNA). Para além destes papéis centrais em todos os organismos, existem ainda vírus que utilizam o RNA como material de armazenamento de informação genética (os retrovírus). A produção do RNA tem habitualmente uma origem comum: a transcrição do DNA. No caso do mRNA, o produto formado é um intermediário cuja função requer ainda a tradução. No caso do tRNA e do rRNA, o produto formado é o efector da função a que se destina.

A transcrição: papel na regulação genica

A transcrição é talvez o passo por excelência para a regulação da expressão génica. A decisão principal na regulação de um gene, é habitualmente a decisão de transcrever ou não esse mesmo gene. O que se traduz possivelmente numa necessidade de economia de energia e materiais por parte da célula.

Mecanismo da transcrição

A transcrição é catalisada pela RNA polimerase, e envolve a síntese de uma cadeia de RNA complementar da cadeia molde do DNA (a outra cadeia do DNA é a imagem do RNA, isto é a sua sequência é equivalente à do RNA, excepto no facto de em vez de possuir Uracilos possui Timidinas). A transcrição ocorre pelo processo habitual de emparelhamento de bases num processo altamente regulado e encadeado. Em primeiro lugar, a polimerase deve ligar-se ao DNA de cadeia dupla. Em seguida, as duas cadeias do DNA devem ser separadas (abertura da dupla hélice), para tornar a cadeia complementar acessível à maquinaria de transcrição. A abertura da hélice do DNA é um processo localizado, e à medida que a transcrição prossegue, novas zonas do DNA vão ficando acessíveis, enquanto as zonas já transcritas se vão emparelhando de novo, por forma a preservar a dupla hélice. A fase de **iniciação da transcrição** envolve assim, o reconhecimento do DNA pela polimerase, a abertura da hélice do DNA, e a incorporação do primeiro nucleótido na cadeia do RNA nascente. O local do gene onde se processa todo este processo é naturalmente o **promotor**. O local da incorporação do primeiro nucleótido é designado “**start site**” ou “**startpoint**”.

Depois da fase de iniciação inicia-se a **fase de alongação**, a qual produz a extensão da cadeia de RNA nascente, originando um híbrido de emparelhamento DNA-RNA. No entanto, à medida que a alongação se processa, a polimerase caminha para novas regiões do DNA, abrindo a hélice noutras zonas do gene, e fechando nas regiões já transcritas, o que implica o desemparelhamento DNA-RNA.

A **terminação** envolve o reconhecimento de um sinal indicando que não devem ser adicionados mais nucleótidos. Nesta fase, termina a ligação DNA-RNA da cadeia nascente, com libertação da polimerase e da molécula de RNA.

Componentes enzimáticos da transcrição

Desta descrição se pode inferir que a polimerase do RNA (a enzima que catalisa a adição de nucleótidos à cadeia de RNA nascente) não funciona só, necessitando de um conjunto de outros componentes com funções essencialmente reguladoras e assessórias. Assim, quer a iniciação quer a abertura do DNA, quer a terminação são exemplos de processos em que intervêm outros factores para a progressão organizada e controlada da expressão génica. A maquinaria de transcrição das células eucarióticas é mais complexa e menos bem definida que a dos procariotas. Existem 3 polimerases nucleares, as quais ocupam diferentes locais do núcleo, e são cada qual composta por várias subunidades. Para complicar ainda mais o problema, existem ainda outras polimerases do RNA em mitocôndrias e cloroplastos.

Tipos de polimerases do RNA

A maior parte da actividade de polimerase do RNA é realizada, nos eucariotas, pela RNA polimerase I, a qual se encontra no nucléolo, e é responsável pela transcrição dos genes codificando os rRNA (cerca de 50-70% do RNA total sintetizado). A segunda enzima, é a RNA polimerase II (20-40% da actividade total de síntese de RNA), e é responsável pela síntese do RNA heterogéneo (hnRNA), o precursor do mRNA. A RNA Polimerase III é responsável pela restante actividade de produção de RNA (até 10% do total), tem localização nucleoplasmática e é responsável pela produção dos tRNA e muitos dos “small nuclear RNA” (snRNA).

Instabilidade do RNA: necessidade para a regulação genética

3.2 - A INSTABILIDADE DO mRNA

Nenhum mecanismo de controlo da transcrição pode ser uma forma eficaz de controlar a expressão génica, se o produto da transcrição (o mRNA) não tivesse uma vida curta. Se assim não fosse, previsivelmente ocorreria uma acumulação de mensageiro, ou pelo menos o mensageiro formado permaneceria activo tanto tempo que não seria possível parar de sintetizar a respectiva proteína. Na realidade, a instabilidade do mRNA é muito acentuada. As duas formas de determinar a instabilidade do DNA baseiam-se ambas no bloquear da síntese *de novo* do mRNA (transcrição), medindo então a sua capacidade para servir na síntese proteica (semi-vida funcional), ou a sua capacidade para hibridar com uma sonda (semi-vida química). De modo geral, a semi-vida funcional é ligeiramente inferior à semi-vida química, o que sugere que pequenas degradações como um simples corte poderão ser suficientes para a inactivação biológica do mRNA. Verifica-se que este primeiro passo inicial é seguido da degradação do mRNA nos seus nucleótidos componentes, de forma mais ou menos sequencial na direcção 5'→3'.

Os retrovírus e o dogma central da genética molecular

3.3 - TRANSCRIÇÃO REVERSA. OS RETROVIRUS

O dogma central da genética molecular afirma que os genes são unidades que se perpetuam a si próprios, e que funcionam através da sua expressão em proteínas, através de um intermediário de RNA. Note-se que o dogma, na sua versão original define um paradigma que considera que a informação genética é transmitida unidirecionalmente: DNA→RNA→Proteína.

Hoje em dia, sabemos que a restrição do dogma central não é absoluta. Efectivamente, a informação genética pode ser transmitida de forma diferente da acima prevista. Alguns vírus de RNA, utilizam o RNA para a propagação da sua informação genética. Se esta pode parecer uma extensão relativamente pequena do dogma central, já a existência nos retrovírus (vírus de RNA de cadeia simples que

utilizam o DNA de cadeia dupla como intermediária na sua replicação) de transcriptases reversas constitui uma grande mudança no paradigma da genética molecular. As transcriptases reversas são enzimas que catalisam a síntese de um DNA de cadeia simples a partir de uma cadeia de RNA. Esta cadeia de DNA pode então ser utilizada para sintetizar DNA de cadeia dupla, utilizando a maquinaria habitual da célula,, efectivamente revertendo um dos passos acima indicado: RNA→DNA. Este facto tem implicações profundas não só na forma de pensar a genética, mas também na biologia da infecção viral, já que este DNA de cadeia dupla formado, e que é uma cópia do RNA viral, vai agora integrar-se no genoma da célula, fazendo com que a infecção se propague de forma mais ou menos inofensiva à progenia da célula infectada (a integração no genoma celular é uma parte normal do ciclo de vida do vírus sendo necessária à transcrição dos genes virais). Uma outra implicação deste mecanismo é a possibilidade de uma infecção de vírus deste tipo poder mediar a inserção de mRNA celular no genoma, como se de RNA viral se tratasse, originando duplicação génica, e/ou inserção de uma cópia do gene sob a acção de um promotor diferente, efectivamente alterando o programa genético da célula. Uma outra implicação da infecção por este tipo de vírus, foi já por nós abordada aquando da discussão da existência de enhancers, e constitui na possibilidade de colocar genes celulares sob a acção de enhancers virais, uma vez mais alterando o programa genético da célula infectada.

Os tipos de retrovírus de que existe mais informação disponível são os que originam as **partículas tipo C** em aves e mamíferos. Estes vírus contêm duas cópias de RNA em cada virião. Assim, quando uma célula é infectada por dois viriões diferentes, podem-se originar viriões heterozigóticos, o que pode ser importante na aquisição de sequências celulares por parte do vírus, já que mesmo que em contrapartida perca algumas sequências do seu genoma, a restante cópia do RNA viral permite-lhe continuar a ser capaz de efectuar uma infecção eficaz.

3.4 - A TRADUÇÃO (síntese proteica)

A síntese proteica efectua-se no citoplasma, envolvendo uma complexa maquinaria genética centrada no ribossoma. Esta maquinaria genética pode ser vista como migrando ao longo do mRNA, lendo-o e utilizando a informação nele contida para alinhar com precisão cada aminoacil-tRNA, promovendo a ligação peptídica entre este e a cadeia peptídica nascente. O próprio ribossoma é assim um altamente elaborado e preciso complexo enzimático com diversificados componentes e vários centros activos, que requer vários cofactores para a sua actividade, e que obtém a energia química que necessita com a hidrólise de GTP.

Fases da síntese proteica

A síntese proteica pode ser dividida em várias fases:

Iniciação: envolve as reacções que precedem a formação da ligação peptídica. Requer a ligação do ribossoma ao mRNA, a formação de um complexo de iniciação contendo o primeiro aminoacil-tRNA. Trata-se de um processo relativamente lento em comparação com as restantes fases da síntese proteica.

Elongação: inclui todas as reacções desde a síntese da primeira ligação peptídica, até à adição do último aminoácido da cadeia polipeptídica. Os aminoácidos são

adicionados um a um, naquele que constitui o processo mais rápido da síntese proteica.

Terminação: inclui todos os passos necessários para a libertação da cadeia polipeptídica formada, bem como a dissociação do ribossoma do mRNA. Este é um processo lento, em comparação com o tempo necessário para adicionar um aminoácido na fase de elongação.

O ribossoma

Um ribossoma é composto por duas unidades (60S e 40S nos eucariotas) as quais, apesar de funcionarem em conjunto medeiam reacções diferentes na síntese proteica. O mRNA associa-se à subunidade menor, ficando associado a este por cerca de 30-40 nucleótidos. Apenas 2 moléculas de tRNA se podem associar ao ribossoma em cada momento, pelo que apenas 2 dos cerca de 30 codons associados ao ribossoma se encontram a ser processados em cada momento.

Cada tRNA liga-se ao ribossoma num local diferente deste, tendo cada um dos dois locais de ligação propriedades diferentes. Apenas o Local A (local de entrada) pode receber um aminoacil-tRNA. Antes da entrada do aminoacil-tRNA, este local expõe o codon a ser descodificado. O último dos codons já descodificados encontra-se no local P (local dador), sendo este local ocupado pelo peptidil-tRNA (um tRNA contendo o aminoácido já covalentemente ligado por uma ligação peptídica à restante cadeia polipeptídica nascente). Quando estes locais (A e P) estão ambos ocupados ocorre a formação da ligação peptídica com transferência do polipéptido nascente para o tRNA do local A. O ribossoma desloca-se então no mRNA libertando o tRNA do local P e transferindo para este local o peptidil-tRNA do local A, e expondo um novo codon no local A.

Iniciação

Nos eucariotas a iniciação começa com a ligação de GTP a um factor de iniciação denominado eIF-2 (eucariotic initiation factor 2). De seguida efectua-se a ligação de um N-formil-metionil-tRNA ao este complexo. É o conjunto de factores assim formado e que é denominado por complexo terciário, que se liga então à subunidade 40S do ribossoma, a qual com o auxílio de outros factores de iniciação reconhece então a extremidade 5' do mRNA (na qual se encontra a estrutura conhecida como CAP) por parte da subunidade 40S do ribossoma. A subunidade 40S migra então no mRNA até encontrar um codon de iniciação. Neste ponto, liga-se a subunidade 60S, após a remoção de eIF-2 do complexo de iniciação.

Elongação

Assim que a subunidade 60S se liga ao complexo de iniciação, o ribossoma fica pronto a iniciar a elongação. Para tal necessita de aminoacil-tRNA, o qual entra o local A, num processo mediado pelo factor eEF-1 (eucariotic elongation factor 1). Assim que o aminoacil-tRNA se encontra correctamente posicionado no local A, a peptidil transferase (uma função da subunidade 60S) catalisa a formação da ligação peptídica entre os aminoácidos dos locais P e A.

O último passo na elongação é a translocação, processo em que o ribossoma avança três nucleótidos de forma concertada(e que requer o factor adicional eEF-2). Este processo envolve a libertação do tRNA do local P, a passagem do peptidil-tRNA do local A para o local P, e a exposição do próximo codon no local A agora vazio.

Terminação

Dos 64 tripletos, apenas 61 codificam para aminoácidos, sendo os restantes três codons stop, ou de terminação. Qualquer destes três codons (UAG, UAA e UGA) é suficiente para terminar a síntese proteica.

Aos codons de terminação não corresponde nenhum tRNA, sendo estes reconhecidos directamente pelo factor proteico eRF (eucariotic release factor).

A reacção de terminação envolve a libertação do polipéptido do ultimo tRNA, a expulsão do tRNA do ribossoma, e a dissociação deste do mRNA.

Localização celular da síntese proteica

A célula eucariótica é uma estrutura finamente organizada, cujas funções são efectuadas em locais celulares definidos. A síntese proteica não constitui excepção, podendo os polirribossomas ser classificados em 2 tipos (livres e ligados a membranas), aos quais corresponde a síntese de diferentes grupos de proteínas . Os polirribossomas livres sintetizam proteínas que não interagem com membranas, enquanto os que se encontram associados às membranas sintetizam proteínas cuja futura localização depende da sua capacidade para se ligarem às membranas. Note-se no entanto que a denominação polirribossomas livres não significa que estes se encontrem livres em solução no citoplasma. Estes polisomas encontram-se associados ao citoesqueleto para o que provavelmente dependem do mRNA.

Os polirribossomas tendem a estar localizados perto de núcleos, nos locais de entrada do mRNA no citoplasma. A maior parte das proteínas sintetizadas são solúveis, e uma vez libertadas rapidamente difundem para longe do local de síntese. As proteínas que irão compor o citoesqueleto, tendem a integrar-se neste num local não muito distante do ponto de síntese.

As proteínas sintetizadas pelos ribossomas ligados a membranas têm vários destinos. Algumas são sequestradas em compartimentos celulares , outras são componentes membranares, e outras ainda são proteínas que se destinam a ser secretadas. Na maior parte dos casos das proteínas de membrana, a sua futura localização não depende da sequência da proteína madura, mas antes de uma sequência denominada “leader”, e que se localiza na zona terminal da cadeia polipeptídica nascente. Esta sequência, depois de ter determinado o destino da proteína será excisada do resto da proteína, originando a proteína madura.

4 - A DIVERSIDADE IMUNOLÓGICA: REARRANJO GENÓMICO

Como codificar num genoma limitado um numero ilimitado de receptores

4.1 - Introdução

Durante a evolução do sistema imunológico para funções de reconhecimento específico, a natureza teve que resolver o problema gigantesco de codificar num genoma limitado, um número suficiente de genes capaz de reconhecer o mundo exterior e interior do organismo. A solução encontrada é, se bem que económica, complexa, como o revela o facto de não ser ainda possível construir sistemas de recombinação *in vitro* isentos de células. Os receptores para antígenos são de 2 tipos: 1) as imunoglobulinas, capazes de reconhecer o antígeno na forma nativa, são produzidas pelos linfócitos B, e existem na forma solúvel e na forma membranar; 2) o receptor da célula T reconhece o antígeno depois de processado por células apresentadoras do antígeno, e apresentado no contexto do MHC dessa célula, existe fisiologicamente apenas na forma membranar.

Os genes do TCR

4.2- Aspectos genéticos

Dos dois tipos de linfócitos, a célula T é a responsável pela resposta imunológica dita celular. Para tal, estas células estão equipadas à sua superfície com um receptor para o antígeno (TCR do inglês T-Cell-Receptor), através do qual a célula madura recebe um estímulo de activação quando encontra o antígeno para o qual é específica. Este receptor é composto por um de dois tipos de heterodímeros ($\alpha\beta$ ou $\gamma\delta$). São portanto 4 os genes do TCR, dos quais apenas 2 estarão a ser transcritos em cada célula T. Cada um dos genes é composto por um máximo de 4 tipos de segmentos (V ou variável, D ou de diversidade, J ou de junção, C ou constante). Cada um destes segmentos é composto por mais que um elemento génico, dos quais cada clone celular escolherá um e apenas um para ser utilizado no TCR que irá expressar. As células T diferem assim de todas as restantes células do organismo (com excepção dos linfócitos B), pois o conteúdo genético da célula madura é diferente do de qualquer outra célula que não pertença ao mesmo clone.

Estrutura somática dos genes do TCR

4.2.1 - Estrutura somática dos genes do TCR

Os genes do TCR, tal como os das imunoglobulinas possuem uma configuração somática, igual em todas as células não linfóides. Nos linfócitos, a configuração destes genes é alterada no processo denominado recombinação, para dar origem a um gene funcional.

Os 4 genes do TCR existem em 3 locus cromossómicos, já que o gene δ está localizado no interior do gene α (Fig. 13). Os locus β e δ possuem 4 classes de segmentos (V,D,J,C), e os locus α e γ apenas 3 (V,J,C). Como se pode ver na Fig.13, no Homem, a organização básica dos locus do TCR é que tem sido denominada “extended”, em que os vários tipos de segmentos se organizam separadamente no genoma (V.V. (etc.) .D.D. (etc.) .J.J. (etc.)). No caso do locus $\alpha\delta$, uma variação a esta configuração permite ao gene δ partilhar segmentos V com o gene α (Lewis, 1994).

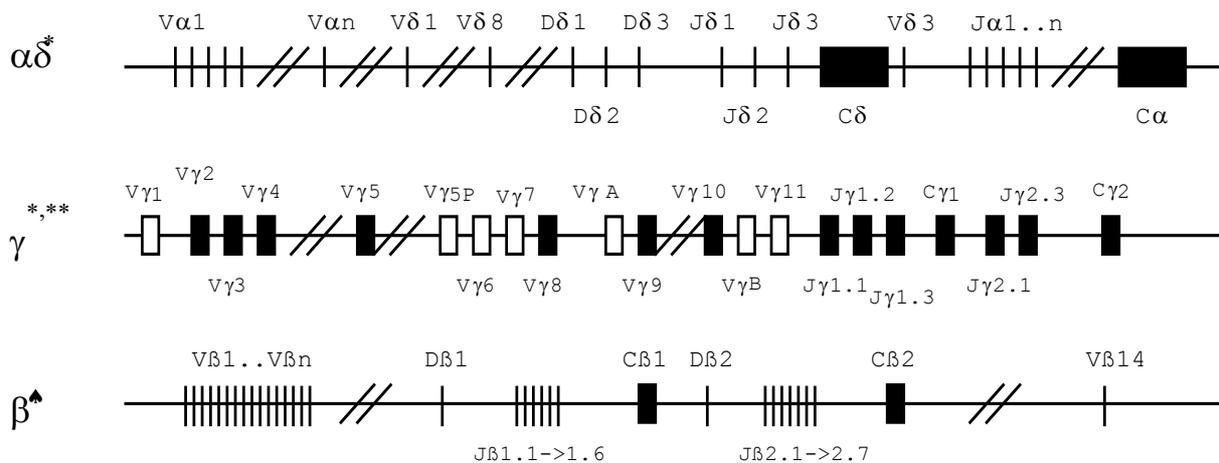


Figura 13 - Diagrama simplificado da estrutura somática dos genes do TCR humano. Os elementos genéticos representados por caixas brancas são pseudogenes (Adaptado de : ^a Toyonaga et al., 1985; ^{**} LeFranc et al., 1989; ^{*} Yoshikai, 1991).

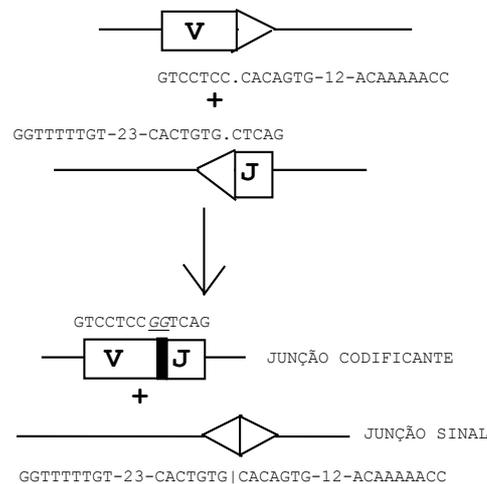


Figura 14 - Equação padrão para a recombinação V(D)J. Os sinais de junção são indicados por triângulos e os segmentos codificantes por quadrados (Extraído de Lewis, 1994)

Sinais de junção

4.2.2 - Mecanismo de rearranjo somático dos genes do TCR

O mecanismo de rearranjo somático dos genes do TCR não é diferente do observado para as imunoglobulinas. Na verdade, foi possível clonar células B com os genes do TCR rearranjado (O'Connor et al., 1985), sugerindo que ambos os receptores são substractos do mesmo conjunto de enzimas. O processo de recombinação quer do TCR quer das imunoglobulinas (doravante denominada recombinação V(D)J) depende primariamente de sequências sinal que flanqueando os segmentos a recombinar constituem todos os elementos necessários para indicar aos componentes enzimáticos onde efectuar a recombinação (Lewis et al., 1985; Akira et al., 1987; Hesse et al., 1987). Estes sinais de junção variam em sequência, mas seguem de muito perto o consenso heptâmero-espaçador-nonâmero, em que as sequências consenso do heptâmero e do nonâmero são respectivamente CACAGTG e AAAAAACC. O espaçador tem uma sequência muito variável, mas o seu comprimento tem 12 ou 23 pares de bases (bp) (Max et al., 1979; Sakano

et al., 1979,1981; Kurosawa et al., 1981). A regra base que dita a orientação dos rearranjos é a de que apenas podem rearranjar elementos com espaçadores diferentes, isto é, um elemento com uma sequência sinal composta por um espaçador de 12 bp apenas rearranja com uma outra cujo espaçador for de 23 bp e vice-versa. Desta forma rearranjos envolvendo elementos do mesmo grupo (V com V; J com J) são impedidos. O mecanismo molecular que origina esta restrição é no entanto ainda hoje desconhecido (Lewis, 1994).

**Junção
codificante e
junção sinal**

Quando dois segmentos génicos se envolvem no processo de recombinação, é feito um corte na fronteira entre a sequência sinal e a sequência codificante, em cada um. As quatro extremidades assim formadas são então ligadas formando uma “junção codificante”, e uma “junção sinal” (Fig. 14). Devido à configuração cromossômica, as sequências codificantes são retidas no genoma, sendo as “Junções sinal” excisadas sob a forma de DNA circular extracromossômico (Fujimoto et al., 1987; Okazaki et al., 1987).

**Resíduos N e
Resíduos P**

A junção codificante, não ocorre no entanto sempre numa posição fixa. Por um lado a quantidade de material genético com que cada elemento contribui pode variar em até 10 nucleótidos (Max et al, 1979; Sakano et al 1979; Weigert et al 1980). Por outro lado, resíduos extra não incluídos na configuração “germline”, podem ser incluídos (Sakano et al., 1981; Lafaille et al., 1989; McCormack et al., 1989). Estes resíduos extra podem ser de dois tipos fundamentais: os “resíduos N” (do Inglês Non-germline-regions) e os “resíduos P” (de Palindromicos).

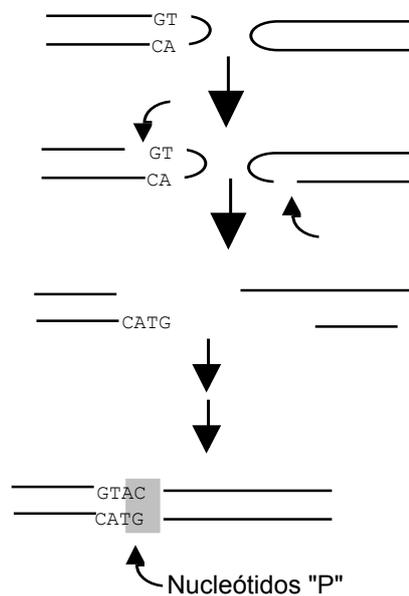


Figura 15 - Mecanismo proposto para a origem dos nucleótidos P (Adaptado de Lewis, 1994).

Os “resíduos N” têm tipicamente um elevado conteúdo G/C (Alt et al., 1982; Roth et al., 1989), não ultrapassam os 15 nucleótidos, e ocorrem mais frequentemente nas junções codificantes que nas junções de sinal (Lewis, 1994). Estes resíduos são adicionados pela enzima TdT (do Inglês Terminal deoxynucleotidil transferase) como o demonstram os modelos de animais transgênicos com inativação do gene desta enzima (Gilfillan et al, 1993; Komori et

al., 1993). No entanto, o facto de estes modelos resultarem em uma muito grande, mas não completa abolição da frequência de “resíduos N” parece indicar a existência de um mecanismo alternativo, independente da expressão de TdT (Lewis, 1994). A regulação de TdT na ontogenia, origina a menor frequência de “resíduos N” no período fetal ou neonatal, possivelmente para permitir o domínio de alguns receptores com especificidades necessárias numa fase mais precoce da ontogenia (Gu et al., 1990; Feeney, 1991, 1992).

Os “resíduos P” parecem ter origem numa molécula intermediária tipo “hairpin” gerada (após o corte na sequência sinal) pela ligação covalente das duas cadeias da dupla hélice do DNA, a qual seria posteriormente clivada num ponto diferente do inicial (Fig. 15.; Lieber, 1991; Roth et al., 1992).

Agentes de recombinação

O agente ou agentes de recombinação permanecem ainda largamente desconhecidos, ou incompletamente caracterizados e purificados (Lewis, 1994). A tendência actual é no entanto no sentido de aceitar que a recombinação V(D)J se realiza não por um factor, mas por uma colecção de factores com actividades pouco relacionadas. Os factores já identificados incluem RAG-1 e RAG-2 (do inglês Recombination activating Gene; Schatz et al., 1988, 1989; Oettinger et al., 1990;), NBP (do inglês nonamer binding protein; Halligan et al., 1987; Li et al., 1989), T-160 (Shirakata et al., 1991), Rc (Wu et al., 1993), RBP-Jk (Hamaguchi et al., 1989), Rp (do inglês recognition protein; Muegge et al., 1993). Dos factores identificados com base na sua capacidade para produzir cortes no DNA, nenhum apresentava a especificidade necessária (Desiderio et al., 1984; Kataoka et al., 1984; Hope et al., 1986). Apenas um factor foi identificado com base na sua actividade de ligase, tendo sido denominado VDJP (do inglês V(D)J Joining Protein; referido em Lewis, 1994). A actividade de ligase desta proteína só pôde ser observada em fragmentos contendo sinais de ligação, pelo que possui a especificidade necessária para estar envolvida na recombinação V(D)J (Lewis, 1994).

4.3 - O TCR na população normal

4.3.1 - Polimorfismos do TCR

Polimorfismos do TCR

As deleções de regiões variáveis foram dos primeiros polimorfismos a serem detectados no genoma do TCR, tanto em murganhos de laboratório (Behlke et al., 1986; Haqqi et al., 1989a., 1989b) como em murganhos selvagens (Pullen et al., 1990; Jouvin-Marche et al., 1989). Polimorfismos mais pontuais foram no entanto também detectados no gene de V β 17 de ratinho, verificando-se que as 2 substituições de aminoácidos afectavam a especificidade final do receptor (Cazenave et al., 1990).

No Homem, apenas uma deleção de V β foi documentada, consistindo na deleção de V β 6.2 (mas não de qualquer outro dos genes de V β testados) num único indivíduo venezuelano pertencente à tribo índia waraos (Concanon et al., 1987). No entanto os polimorfismos das regiões variáveis do TCR parecem ser quase universalmente representados, ainda que não frequentes na população (Concanon et al., 1987). Com efeito, uma busca sistemática por RFLP indicou a existência de polimorfismos em 12 das 14 famílias de V β estudadas (Concanon et al., 1987). Alguns destes polimorfismos podem constituir variações silenciosas,

como é o caso de um polimorfismo encontrado em V β 12.2 (Day et al, 1992), outras no entanto afectam a expressão do gene em linfócitos T maduros, como são os casos dos polimorfismos de V β 1 (Robinson, 1989), V β 18 (Charmley et al., 1993), V β 3 (Posnett et al., 1994a) e V β 6.7 (Posnett et al., 1986; Li et al., 1990; Prashar et al., 1991). Este último com a particularidade de ser detectável com um anticorpo (Posnett et al., 1986), o que permitiu mapear o epítotope de ligação do anticorpo numa zona de possível ligação a superantígenos (Prashar et al., 1991). Também o polimorfismo descrito para V β 3 é único, já que este polimorfismo se localiza no espaçador, constituindo assim, o único exemplo conhecido de uma mutação numa zona não codificante do TCR, que afecta a expressão do respectivo gene (Posnett et al., 1994a). Finalmente, a variação alélica identificada no V β 18 é a única que introduz um codão stop, originando um “buraco” no repertório presente em 11% dos indivíduos estudados (Charmley et al., 1993).

Estes dados indicam que mesmo variações moderadas de apenas 1 ou 2 pares de bases nas sequências codificantes ou não codificantes do genoma do TCR podem ter repercussões significativas no repertório do TCR (Vissinga et al., 1994).

LEITURA COMPLEMENTAR

- Akira S., et al *Science*. 1987; 238:1134-1138
- Alt F., et al *PNAS* 1982;79:4118-4122
- Aplan P.D., et al *Science* 1990; 250:1426-1429
- Bakhshi A., et al *PNAS* 1987; 84:2396-2400
- Behlke M., et al *PNAS* 1986; 86:767
- Boehm T. et al *FASEB J.* 1989; 3:2344-2359
- Brown L., et al *EMBO J.* 1990; 9:3343-3351
- Charmley P., et al *J.Exp.Med.* 1993; 177:135-143
- Concanon P. et al *J.Exp.Med.* 1987; 165:1130
- Day C.E., et al *Human Immunol.* 1992; 34:196
- Desiderio S., et al *Nature*. 1984; 308:860-862
- Feeney A.J. et al *Int.Rev.Immunol.* 1992; 8:113
- Feeney A.J. *J.Immunol.* 1991;147:4343
- Flug F., et al *PNAS* 1985; 82:3460-3464
- Fujimoto S., et al *Nature*. 1987; 327:242-243
- Furon L., et al *Blood* 1989; 73:559-565
- Gilfillan S., et al *Science* 1993; 261:1175
- Gilfillan S., et al. *Science* 1993; 261:1175
- Griesser H., et al. *Blood* 1989; 73:1402-1415
- Gu H., et al *EMBO J.* 1990; 9:2133-2140
- Halligan B.D., et al *PNAS* 1987; 84:7019-7023
- Hamaguchi Y., et al. *Nucleic Acids Res.* 1989; 17:9015-9026
- Haqqi T.M., et al *J.Exp.Med.* 1989a; 169:1903
- Haqqi T.M., et al. *Immunogenetics.* 1989b; 29:180
- Hare J., et al. *Blood* 1989; 73: 500-508
- Hesse J.E., et al. *Cell.* 1987; 49:775-783
- Hope T.J., et al *Science* 1986; 231:1141
- Jouvin-Marche E., et al *Eur. J. Immunol.* 1989; 19:1921-1926
- Kataota T., et al. *Nucleic Acids Res.* 1984; 12:5995
- Komori T., et al. *Science* 1993; 261,1171-1175
- Korsmeyer S.J. et al. *Annu.Rev.Immunol.* 1992;10:785-807
- Kurosawa Y., et al. *Nature.* 1981; 290:565-570
- Lafaille J.J., et al. *Cell.* 1989; 59:859-870
- Lee M.S., et al. *Science* 1987; 237:175
- LeFranc M.P., et al *TIBS.* 1989; 14/6: 214-218
- Lewis S., et al. *Science* 1985; 228:677-685
- Lewis S.M. *Advances in Immunology.* 1994; 56:27
- Li M., et al. *Genes Dev.* 1989; 3:1801-1813
- Li Y., et al *J.Exp.Med.* 1990; 171:221-230
- Lieber M.R. 1993; in I.R.Kirsh, ed. 9 239. CRC Press Ann Arbor
- Lieber M.R. *FASEB J.* 1991; 4:2934-2944
- Magrath I. *Cancer Res.* 1992; 52:5529s-5540s
- Max E.E., et al *PNAS.* 1979; 76:3450
- McCormack et al. *Cell* 1989; 56:785-791
- Minden M.D., and Mak. T. *Blood* 1986; 68:327-336
- Minden M.D., et al *PNAS* 1985; 82:1224-1227
- Muegge K., et al *PNAS* 1993; 90:4151-4155
- O'Connor N.T.J., et al. *The Lancet* 1985;8:1295-1297
- Oettinger M.A., et al. *Science* 1990; 248: 1517-1523
- Paslier D.L., et al. *Blood* 1987; 70:637-646
- Posnett D.N., *J.Exp.Med.* 1994. 179:1707-1711
- Rabbitts T.H. *Cell* 1991; 67:641-644
- Rabbitts T.H., et al *EMBO J.* 1985; 4:2217-2224
- Reis M.D., et al *Biochim. Biophys. Acta* 1991; 1072:177-192
- Roth D.B., et al. *Mol.Cell.Biol.* 1989; 9:3049-3057
- Roth D.B., et al *Cell* 1992; 70:983-991
- Royer H.D., and Reinherz E.L. *The New England J. of Medicine* 1987; 317: 1136-1142
- Sakano H., et al. *Nature* 1979; 280: 288-294
- Sakano H., et al. *Nature* 1981; 290:562-565
- Sawyers C.L., et al. *Cell* 1991; 64:337-350
- Schatz D.G., Baltimore D. *Cell.* 1988; 53:107-115
- Schatz D.G., et al., *Cell* 1989; 59:1035-1048
- Shirakata M., et al., *Mol.Cell.Biol.* 1991; 11:4528
- Smith J.L., et al., *J.Pathol.* 1992; 166:109-112
- Tawa A., et al., *Blood.* 1987; 70:1933-1939

- Toyonaga B., et al., *PNAS*. 1985; 82: 8624-8628
- Tycko B., and Skar J. *Cancer Cells*. 1990; 2:1-8
- Vissinga C.S., et al., *J. Immunol.* 1994; 152:1222
- Weigert M., et al., *Nature* 1980; 283:497-499
- Wu L.C., et al., *Nucleic Acids Res.* 1993; 21:5067
- Yoshikai Y. *Microbiol.Immunol.* 1991; 35:493-506

5 - METODOLOGIA DE ESTUDO EM GENÉTICA MOLECULAR

Passos na extração de DNA

5.1 - PREPARAÇÃO DE DNA E RNA

Qualquer análise em genética molecular requer, obviamente, o estudo do DNA ou do RNA, pelo que o primeiro passo em qualquer técnica genética consiste no isolamento e purificação de uma ou mesmo das espécies de ácidos nucleicos. As variadas técnicas disponíveis para esse efeito, as quais originam DNA ou RNA com diferentes propriedades de pureza e integridade física, possuem princípios semelhantes. Todas se iniciam com uma lise suave das células a estudar, seguida de ataques enzimáticos e/ou químicos para destruir os componentes proteicos da mistura. Finalmente, a purificação do DNA ou RNA faz-se por um de vários métodos, de acordo com os objectivos pretendidos.

DNAses e RNAses

A purificação e correcto manuseamento de RNA é bem mais difícil que para o DNA. Este facto não resulta de uma maior complexidade de procedimentos, mas da maior estabilidade das RNAses. Com efeito, ao contrário das DNAses, as RNAses são extremamente estáveis, e não necessitam de cofactores para funcionarem. Desta forma, não é possível inactivá-las com a adição de quelantes do magnésio (EDTA) como acontece para as DNAses. A inactivação das RNAses é eficiente com a utilização de dietilpirocarbonato (DEPC), mas a alta toxicidade deste composto, aliada à necessidade da sua eliminação por autoclavagem, impede a sua utilização em todas as soluções. Em reacções enzimáticas é possível utilizar inibidores específicos de RNAses (RNAse inhibitor ou abreviadamente RNAsin), como o extraído do tecido placentário, para inactivar as RNAses provenientes do material celular donde é extraído o RNA, mas este não é um método prático para o tratamento generalizado dos reagentes e material de plástico do laboratório. Se aliarmos a este facto, a presença em grandes quantidades de DNAses e RNAses nas mãos humanas, facilmente se compreende a imperiosa necessidade de utilizar luvas no laboratório de Genética Molecular, não para a protecção do operador, mas para proteger a amostra do ataque das RNAses e DNAses do manuseador da amostra e restantes materiais de laboratório.

5.2 - ANÁLISE DE DNA

5.2.1 - ELECTROFORESE

A grande maioria dos métodos de genética molecular requer num determinado momento o fraccionamento de ácidos nucleicos segundo o seu comprimento. Para tal utilizam-se as técnicas de electroforese, que consistem na separação dos ácidos nucleicos numa matriz porosa (habitualmente géis de agarose, ou acrilamida) sob a força de um campo eléctrico (os ácidos nucleicos têm carga negativa, pelo que migram em direcção ao polo positivo).

A matriz na qual os ácidos nucleicos devem ser separados depende essencialmente do tamanho dos fragmentos a separar, mas também do destino final a dar a estes uma vez separados. Existem diferentes tipos matrizes:

Géis de Agarose

- **agarose** - A agarose tem relativamente à acrilamida a vantagem de constituir uma matriz não tóxica de muito fácil preparação (basta solubilizar a agarose em pó num tampão (TAE ou TBE) a quente, e deixar então arrefecer. Várias agaroses existem, as quais permitem separar fragmentos com mais ou menos nucleótidos. A agarose normal, permite uma boa resolução para fragmentos relativamente grandes, utilizando-se em baixas concentrações (0.8-2%). Já a agarose Nusieve (FMC-bioproducts) permite separar com grande resolução fragmentos com menos de 1000bp, pelo que se adapta melhor aos fragmentos habitualmente obtidos por PCR. Uma variação destas duas agaroses (Nusieve 3:1), não é mais que uma mistura de 3 partes de Nusieve com uma parte de agarose normal, o que origina um gel com uma viscosidade aceitável em altas concentrações (normalmente até 4%, tal como a Nusieve), mas com um poder de resolução superior quer à Nusieve, quer à agarose normal. Existem ainda agaroses com baixa temperatura de fusão. Estas agaroses têm a desvantagem de ser mais sensíveis a aumentos de temperatura durante a electroforese, mas a vantagem de facilitarem a purificação do DNA separado.

Géis de acrilamida

- **Acrilamida** - Os géis de acrilamida baseiam-se na formação de uma matriz porosa, de polímeros de acrilamida. Para a fazer, utilizam-se monómeros de acrilamida, um reagente bifuncional (bis-acrilamida), e um gerador de radicais livres (iniciador da reacção de polimerização; normalmente peróxido de amónio), bem como um catalisador da polimerização (temed). A resolução dos géis de acrilamida é muito grande podendo facilmente separar fragmentos com apenas um nucleótido de diferença, pelo que habitualmente se utiliza como matriz nos géis de sequenciação. A grande desvantagem deste tipo de matriz consiste na sua grande fragilidade, toxicidade e complexidade de preparação.

5.2.2 - SOUTHERN BLOT E HIBRIDAÇÃO

Southern Blot e Northern Blot

Frequentemente, após a separação dos ácidos nucleicos, torna-se necessário identificar o fragmento de interesse, numa mistura complexa de fragmentos separados. A técnica de eleição para esse efeito é a técnica de Southern Blotting (ou Northern Blotting, conforme o ácido nucleico seja DNA ou RNA respectivamente), seguida de hibridação com uma sonda específica para o fragmento de interesse. Esta técnica consiste na passagem dos fragmentos de DNA (ou RNA) separados, depois de desnaturados, para uma membrana de Nylon ou celulose, por capilaridade ou por vácuo, seguida da fixação do ácido nucleico à membrana. Esta é então utilizada numa reacção de hibridação, em que é colocada uma sonda de cadeia simples de DNA (ou de RNA) em contacto com a membrana, em condições químicas e de temperatura que asseguram que a sonda hibrida apenas com o fragmento complementar. A ligação da sonda é então revelada por autorradiografia (no caso de sondas radioactivas), ou por quimoluminescência (ECL; Amersham).

5.2.3 - DOT E SLOT BLOT

O dot-blot, e a sua variante slot blot, constitui na aplicação do DNA directamente na membrana de Nylon ou de Celulose, sem a prévia separação electroforética. Esta transferência faz-se habitualmente por aplicação de um vácuo numa zona da membrana bem definida, em aparelhos especificamente construídos para o efeito. Consegue-se assim concentrar a amostra numa área bem definida, a qual pode ser um círculo (dot-blot) ou um fino rectângulo (slot-blot). A membrana assim conseguida pode então ser utilizada numa hibridação tal como no processo de southern-blot, o qual serve não para identificar segmentos, mas para verificar da existência ou não de uma determinada espécie de DNA na amostra inicial.

5.4 - MANIPULAÇÃO ENZIMÁTICA DE DNA E RNA

5.4.1 - ENDONUCLEASES DE RESTRIÇÃO

Funções biológicas dos sistemas R-M

Pensa-se que a função biológica das endonucleases de restrição e a protecção das células contra DNA externo a célula. Esta assunção tem como base o facto de a grande maioria das enzimas descobertas efectivamente cortarem o DNA extracelular que in vitro se insere. Simultaneamente o DNA endógeno é protegido devido a metilação específica de nucleótidos na sequência reconhecida, efectuada por uma metiltransferase específica. O conjunto da endonuclease de restrição e respectiva metiltransferase formam o que se designa por sistema de modificação de restrição (R-M system).

Existem pelo menos 4 tipos de sistemas R-M, distinguidos pela composição das suas subunidades, pelo tipo de sequências reconhecidas, e pelos cofactores necessários para a sua actividade. Cerca de 93% das enzimas caracterizadas pertencem a classe tipo II. Juntamente com as enzimas da classe tipo IIs (cerca de 5% das enzimas descritas) constituem o grosso das enzimas comercialmente disponíveis. As enzimas de tipo I (cerca de 1%) e as de tipo III (<1%) são relativamente pouco frequentes. Algumas outras enzimas existem que não podem ser incluídas em nenhuma destas classes.

5.4.1.1 - SISTEMAS R-M TIPO II

Restrição de sequências simétricas: sistemas R-M tipo II

As enzimas de tipo II são as mais simples. Reconhecem sequências de DNA simétricas, cortando entre as sequências, deixando um terminal 3' hidroxil e um terminal 5'fosfato. Requerem apenas magnésio para a sua actividade, e as metiltransferases respectivas requerem apenas s-adenosylmetionina. Reconhecem uma variedade de sequências quase ilimitada, mas poucas reconhecem sequências com menos de 4 ou mais de 8 bp.

Estas enzimas são habitualmente compostas por um homodímero, pelo que necessariamente interactivam com uma sequência de repetição invertida, já que cada subunidade reconhece o mesmo motivo em cadeias de DNA opostas.

As metiltransferases do tipo II são habitualmente compostas por um monómero, o que pode reflectir a necessidade de metilar apenas uma das cadeias de DNA nascentes (durante a duplicação do DNA, uma das cadeias já se encontra metilada), ao contrário das endonucleases de restrição que tem que reconhecer e cortar as duas cadeias do DNA.

As enzimas do tipo IIs utilizam geralmente os mesmos cofactores que as enzimas tipo II, mas as suas sequências de reconhecimento são assimétricas e ininterrompidas, tendo 4 a 7 bp. O local de corte não se situa no interior da sequência de reconhecimento, mas a uma distância de até 20bp num dos sentidos. Nestes sistemas, a metilação é efectuada por duas metiltransferases (uma para cada cadeia), sendo em alguns sistemas metiladas bases diferentes em cada cadeia do DNA.

5.4.1.3 - MONTAR UMA REACÇÃO DE RESTRIÇÃO

Estabilidade térmica

As enzimas de restrição (como todas as enzimas) devem ser sujeitas a menor variação térmica possível. A temperatura a que as enzimas são habitualmente conservadas é -20°C, pelo que quando se transportam para a bancada, devem permanecer em gelo, ou idealmente num congelador de bancada, os quais mantêm uma temperatura de -20°C durante cerca de 2 horas (depende do fabricante). Devido a instabilidade das enzimas a temperatura ambiente, estas devem ser os últimos componentes da mistura de reacção a adicionar, para minimizar quer o choque entre a composição do tampão de conservação e a da mistura, quer o tempo de permanência a temperatura ambiente.

Concentração de glicerol

Para aumentar o tempo de conservação, as enzimas de restrição são habitualmente conservadas em 50% de glicerol. No entanto, um excesso de glicerol na mistura de reacção (>5%) ocasiona um comportamento errático da enzima. Assim, o volume de enzima adicionado nunca pode exceder os 10% do volume total da reacção.

Tampão de reacção

As diferentes enzimas tem actividades diferentes em determinados tampões. Assim, as companhias que as fornecem estudaram um conjunto de tampões concentrados (habitualmente 10X), os quais são otimizados para a actividade das varias enzimas. Estes tampões devem sempre que possível ser utilizados com as respectivas enzimas, pois a actividade de uma enzima num tampão diferente do sugerido pode ser quase nula. Se a experiência obrigar a utilização de varias enzimas no mesmo tubo de reacção, deve ser escolhido o tampão que apresentar o melhor compromisso entre a actividade das duas enzimas (os fornecedores fornecem habitualmente uma tabela com a %actividade de cada enzima em cada um dos tampões que fornecem). Finalmente, algumas enzimas necessitam da adição de componentes extra aos tampões padrão (ex. BSA). Também neste caso, soluções concentradas destes compostos são fornecidas com a enzima.

Actividade das enzimas: Por definição, 1 unidade de enzima de restrição digere completamente 1µg de DNA num volume de 50µl, ao fim de 1 hora. No entanto esta actividade é apenas indicativa, já que tipos diferentes de DNA podem possuir conformações diferentes, e um numero diferente de locais de restrição. Assim, utiliza-se de modo geral 2 a 3 vezes mais enzima, e entre 3 a 16 horas de incubação. O volume da reacção não deve ser inferior a 50µl, já que aumentam os erros de pipetagem, e a probabilidade de a concentração de glicerol ser superior a 5%.

Um factor critico na boa execução de qualquer reacção enzimática é a homogeneidade da mistura. Deve-se homogeneizar a mistura de reacção por

inversão e pipetagem repetida, mas nunca utilizar o vortex já que a violência deste pode desnaturar a enzima, deixando-a inactiva.

Temperatura de reacção

A temperatura de incubação da maioria das endonucleases de restrição é de 37°C, mas algumas enzimas, isoladas de bactérias termofílicas necessitam de incubações entre 50-60°C (verifique a temperatura ideal para cada enzima, junto do fornecedor).

5.4.2 - MODIFICAÇÃO E MARCAÇÃO DE ÁCIDOS NUCLEICOS

5.4.2.1 - RADIOISÓTOPOS

³²P

O isótopo mais comumente utilizado para a marcação radioactiva de ácidos nucleicos é o ³²P. Este radioisótopo emite partículas β e tem uma actividade específica elevada (9200 Ci/mmol na sua forma pura) e um tempo de semi-vida relativamente curto (14 dias). Existem comercialmente disponíveis todas as espécies de trifosfatos de nucleótidos marcados com ³²P, e com variadíssimas actividades específicas. Note-se o átomo radioactivo do dNTP para a marcação de ácidos nucleicos deve ser o γ (os átomos α e β são libertados na formação da ligação com o nucleótido seguinte).

³³P (rever)

Recentemente foi lançado um novo radionucleótido com menor energia que o ³²P, mas idêntica capacidade de penetração: o ³³P. Devido à sua menor energia, este radionucleótido origina bandas mais bem definidas que o ³²P.

³⁵S

O ³⁵S emite partículas com energia ainda mais baixa que a do ³³P (a sua actividade específica é de 1500 Ci/mmol na forma pura), tendo no entanto um tempo de semi-vida mais longo (87 dias). Os nucleótidos marcados com ³⁵S possuem um grupo tiol em substituição de um oxigénio no grupo fosfato, o que pode inibir a actividade de algumas enzimas. Por outro lado, uma vez que a sua energia é menor, este radioisótopo induz menos danos no DNA, pelo que as sondas com ele marcadas são mais estáveis. A menor energia deste nucleótido permite também obter bandas ainda mais bem definidas que as obtidas com ³³P, muito embora possa levar mais tempo a imprimir o filme fotográfico. É de salientar ainda o facto de este radioisótopo ser menos nocivo para o operador de laboratório, pese no entanto o facto de também ser mais difícil de detectar contaminações com o auxílio de um contador Geiger.

³H

O trítio é o radioisótopo de menor energia que é utilizado para a marcação de ácidos nucleicos. Com a sua actividade específica de apenas 29 Ci/mmol na forma pura, e um tempo de semi-vida de 12 anos, é o isótopo mais fraco de todos os procedimentos autorradiográficos. A sua baixa capacidade de penetração torna-o também o isótopo menos perigoso no laboratório, mas também mais difícil de detectar com contadores portáteis tipo Geiger.

Outros radioisótopos

Ainda que tal aconteça com muito menos frequência, também é possível utilizar ¹⁴C e ¹²⁵I para a marcação de ácidos nucleicos.

5.4.2.2 - POLIMERASES DO DNA

Estão hoje em dia disponíveis uma vasta gama de polimerases do DNA, com propriedades e aplicações diferentes. Os principais factores a ter em conta na escolha da polimerase certa para cada tipo de trabalho são:

- **remoção de nucleótidos existentes:** da existência desta actividade pode depender a fidelidade do produto formado. Todas as polimerases cometem erros. Da sua capacidade de verificar o trabalho realizado, e remover os nucleótidos erroneamente incorporados depende a fidelidade do produto final. Obviamente, a existência desta actividade também resulta numa menor velocidade de reacção, o que pode dificultar a obtenção de produtos longos. Pode assim concluir-se que a opção pela existência ou não desta actividade na enzima escolhida deve ser realizada com base no resultado pretendido. Tipicamente, as reacções que se destinam a sequenciação, ou a clonagem devem ser sempre realizados por enzimas com esta actividade.
- **Estabilidade térmica:** também esta característica pode ser benéfica ou prejudicial, dependendo do objectivo e protocolo específicos a utilizar. Por exemplo, numa reacção de PCR, a utilização de enzimas termoestáveis evita a destruição da enzima no passo de desnaturação. No entanto, se o protocolo envolver a posterior inactivação da enzima pelo calor, é necessário escolher uma enzima termosensível.

5.4.2.3 - FOSFATASES E QUINASES

Estas enzimas são utilizadas com frequência na marcação radioactiva de zonas terminais de DNA. As fosfatases (por ex.. Fosfatase alcalina bacteriana (BAP) ou fosfatase intestinal de vitela (CIP)) são enzimas que procedem à hidrólise dos resíduos fosfato no nucleótido terminal 5' de uma cadeia de DNA ou RNA, ou ainda de nucleótidos isolados. O produto final da desfosforilação possui um grupo hidroxilo 5', o qual pode ser marcado radioactivamente com a T4 polinucleótido quinase, utilizando uma fonte de fosfato radioactiva ($[\gamma\text{-}^{32}\text{P}]\text{-ATP}$).

Ambas as fosfatases acima mencionadas requerem zn^{2+} para a sua actividade, mas a CIP é termosensível, sendo facilmente inactivada a 70°C ao fim de 10 min, possuindo ainda 10 a 20 vezes mais actividade nas mesmas condições experimentais.

5.4.3 - POLIMERASE CHAIN REACTION (PCR)

5.4.3.1 - PRINCÍPIOS GERAIS DO PCR

O PCR (Polimerase Chain Reaction) é um procedimento rápido para a amplificação enzimática *in vitro* de segmentos específicos de DNA. A descoberta desta tecnologia teve um enorme impacto na genética molecular, provocando uma revolução de tal ordem que em 1994 foi atribuído ao “inventor” do PCR um Prémio Nobel.

A base teórica do PCR é muito simples, baseando-se na propriedade das polimerases do DNA para catalisar a formação de uma cópia de uma cadeia de DNA, apenas quando encontram uma extremidade 3' livre. Desta forma, foi possível partir de uma molécula de DNA de cadeia dupla, desnaturá-la pelo calor, baixando de seguida a temperatura até um valor que permita a ligação específica de um oligonucleótido sintético, específico para a região 5' do segmento a amplificar. Depois desta hibridação específica, a polimerase inicia então a síntese da cadeia complementar ao “molde”. Entretanto, um processo semelhante deverá ter ocorrido em simultâneo para a restante cadeia da dupla hélice inicial, pelo que no fim deste ciclo, efectivamente foi duplicada a quantidade de DNA da zona de interesse. O processo prossegue com nova desnaturação pela temperatura, repetindo-se este ciclo um numero definido de vezes. Como em cada ciclo se duplica a quantidade de DNA de interesse que existia no início do ciclo, no final do processo amplificamos 2^n vezes o segmento de DNA em que se estava interessado. Na maior parte dos casos, utilizam-se entre 24 e 35 ciclos de temperatura, pelo que no final existem $2^{24}=16,777,216$ a $2^{35}=34,359,738,368$ vezes mais cópias do segmento de interesse que inicialmente.

Para implementar este procedimento é necessário incluir na mistura de reacção não só o DNA a estudar, mas também os oligonucleótidos específicos (primers), desoxinucleótidos trifosfatados (dNTP's), e uma polimerase do DNA termosensível (para resistir às flutuações de temperatura necessárias para realizar as várias fases da reacção). Na prática outros componentes são também adicionados, para que as condições de reacção serem as ideais para a enzima utilizada. Um dos componentes que todas as enzimas até agora descobertas utilizam é o $MgCl_2$, de cuja concentração dependente em larga medida a especificidade, e qualidade do DNA amplificado.

Os primers são utilizados num largo excesso relativamente ao DNA a ser amplificado, já que são necessárias pelo menos tantas moléculas de primer quantas as cadeias de DNA que se deseja formar. Os primers são desenhados por forma a que um tenha a sequência complementar invertida da extremidade 3' do segmento a amplificar (pelo que hibrida directamente com esta zona do molde, deixando livre uma extremidade 3' para que a polimerase inicie o seu trabalho produzindo uma nova cadeia na direcção $5' \rightarrow 3'$). O restante primer é desenhado por forma a ter a sequência da extremidade 5' do fragmento a amplificar. Desta forma este primer hibrida com esta extremidade da cadeia complementar do DNA, deixando livre uma extremidade para a cópia da respectiva cadeia. Como a polimerização só termina quando a temperatura se eleva para a fase de desnaturação, no primeiro ciclo produzimos 2 tipos de fragmentos, ambos com início bem definido, mas com

terminação incerta. No entanto, como os fragmentos produzidos no primeiro ciclo vão ser os moldes para a polimerização da cadeia complementar no ciclo seguinte, aos poucos, o produto preponderante tem ambas as extremidades determinadas pelos locais de ligação de ambos os primers.

5.4.3.2 - RT-PCR

A reacção de RT-PCR (reverse transcriptase-polimerase chain reaction), não é mais do que um PCR normal realizado a partir de um DNA sintetizado por transcrição reversa a partir de RNA. Assim, são necessários neste processo dois tipos de polimerases do DNA: primeiro uma DNA polimerase dependente do RNA (isto é a transcriptase reversa), e depois a polimerase normal. Este processo pode assim ser efectuado quer pela utilização sequencial de 2 enzimas diferentes, ajustando as condições da reacção à enzima, quer através da utilização de enzimas especiais que possuem ambas as actividades, embora em condições diferentes de reacção.

5.4.4 ANÁLISE DE MUTAÇÕES

Dois tipos básicos de análises podem ser efectuadas quando se pretende analisar as mutações de um dado gene: análises dirigidas à detecção de mutações conhecidas, ou análises que procuram identificar mutações que não são previamente conhecidas.

Os métodos dirigidos à identificação de mutações bem caracterizadas dirigem-se especificamente à zona do gene onde se sabe residir a mutação. Utiliza-se neste caso uma qualquer particularidade da sequência mutada para a distinguir da sequência normal, seja por acção enzimática, seja pela hibridação específica de uma sonda, seja ainda pela amplificação selectiva de genes quimera, formados pela junção de dois fragmentos genómicos habitualmente distantes.

No caso de a mutação que se procura não se encontrar ainda caracterizada, os métodos de detecção baseiam-se em características mais genéricas do comportamento dos ácidos nucleicos quer de cadeia simples, quer em cadeia dupla.

5.4.4.1 - ANÁLISE DE MUTAÇÕES CONHECIDAS

As tecnologias a seguir designadas permitem identificar com precisão mutações conhecidas:

- a) **RFPL** (Restriction Fragment Length Polymorphism): Se a mutação em causa originar a criação ou destruição de uma sequência reconhecida por uma enzima de restrição, pode-se utilizar esta enzima para identificar a presença ou ausência da mutação num fragmento amplificado por PCR. Esta mesma técnica pode ser utilizada com DNA genómico, mas neste caso, a detecção tem que ser realizada por Southern Blotting e hibridação, o que torna o processo bastante moroso.
- b) **ASO** (Allele Specific Oligonucleotide): Neste caso utiliza-se uma sonda constituída por um oligonucleótido sintético, que hibrida apenas com a sequência normal e não com a mutada (ou vice-versa). No caso de

mutações pontuais (as mais frequentes) pode ser difícil estabelecer condições de adstringência que permitam a distinção das duas sequências.

- c) **PASA** (PCR Allele Specific Amplification): Durante uma reacção de PCR utiliza um “primer” que hibrida apenas com o alelo mutante (ou vice-versa), pelo que só este é amplificado. Neste caso, a presença ou ausência de produto de PCR indica a presença ou ausência de mutação.
- d) **Amplificação de genes quimera**: Algumas mutações oncológicas consistem na troca de material genético entre cromossomas, ou no equivalente não visível microscopicamente, mas detectável por técnicas genéticas. Como resultado destas trocas de material genómico, sequências que numa célula normal se encontram muito distantes (até mesmo em cromossomas diferentes) ficam agora muito próximas, sendo passíveis de amplificação por PCR. Assim, a existência de produto de um PCR, em que os dois primers utilizados se ligam a genes normalmente muito afastados, permite identificar este tipo de mutações com grande precisão e sensibilidade.

5.4.4.2 - ANÁLISE DE NOVAS MUTAÇÕES

O despiste de novas mutações pode fazer-se mediante o recurso às seguintes tecnologias:

- a) **Corte de pontos “mismatch”**: Neste método a existência de um ponto de não complementaridade no emparelhamento de uma molécula híbrida DNA-RNA ou RNA-RNA (em que uma das cadeias representa o alelo normal), serve de substrato ao corte com RNase A, sendo o resultado da reacção visualizado por electroforese e autorradiografia. A presença de bandas indicando o corte indica a presença da mutação
- b) **Métodos baseados na temperatura de fusão¹**: Nestes métodos (DGGE - Denaturing Gradient gel electrophoresis; TGGE - Temperature Gradient Gel electrophoresis) utiliza-se a temperatura de fusão dos ácidos nucleicos para os separar numa electroforese.
- c) **Métodos conformacionais**: A mobilidade electroforética de espécies de DNA de cadeia simples depende da sua conformação. Como esta conformação é, pelo menos em parte, uma função da sequência, pode-se através deste método, nas condições certas distinguir 2 espécies de DNA com apenas um nucleótido de diferença.
- d) **sequenciação**: O mais trabalhoso método para a análise de mutações, mas também o que fornece informação mais precisa e completa sobre a mutação é a sequenciação.

¹ A temperatura de fusão do DNA é a temperatura a que as duas cadeias da dupla hélice se abrem, pela quebra das pontes de hidrogénio que as unem. Esta temperatura de fusão é uma função do conteúdo GC do DNA, já que existem 3 ligações de hidrogénio no emparelhamento GC, e apenas 2 no emparelhamento AT.

**ÁREAS DE
INTERVENÇÃO DA
GENÉTICA MOLECULAR
EM HEMATOLOGIA**

6 - INTRODUÇÃO

O advento da biologia molecular, e da genética molecular trouxeram consigo algumas descobertas com implicações no diagnóstico, terapêutica e “follow up” de doentes do foro das mais diversas especialidades médicas. A Hematologia, não sendo excepção beneficiou também dos avanços da genética molecular, em 3 áreas distintas: As doenças linfoproliferativas, o estudo das doenças víricas, e as doenças genéticas do foro hematológico. Os dois primeiros grupos são talvez aqueles em que o poder da genética molecular é mais pronunciado, já que permite não só detectar e caracterizar a doença, sendo um útil auxiliar de diagnóstico, mas disponibiliza ainda uma forma de avaliação da progressão da doença, proporcionando elementos de avaliação preciosos quanto ao tipo de resposta à terapêutica. No caso das doenças genéticas, o benefício maior ocorre nos estudos familiares, e de despiste em populações randomizadas, permitindo detectar indivíduos assintomáticos portadores da doença (quer sejam homozigóticos quer heterozigóticos), permitindo em muitos casos que o tratamento se inicie antes mesmo de existir doença clínica, prevenindo a formação de lesões. É ainda de salientar, no caso das doenças genéticas, a possibilidade de efectuar diagnóstico pré-natal. A possibilidade de trabalhar com pequenas quantidades de amostra introduzida pelas tecnologias de PCR, aliada à grande sensibilidade destas técnicas permite com segurança, e partindo de colheitas de material muito reduzidas, efectuar o diagnóstico pré-natal a fetos de alto risco.

7 - HEMOPATIAS MALIGNAS

7.1 - Detecção de anomalias genéticas

A possibilidade de detectar anomalias genéticas específicas de células malignas, fornece não só elementos importantes na caracterização da hemopatia maligna, como permite seguir o evoluir da doença, bem como a resposta à terapia. A doença residual mínima é definida como o nível mais baixo de doença detectável pelos métodos analíticos disponíveis. A sensibilidade dos métodos analíticos disponíveis tem variado muito ao longo do tempo. Tradicionalmente, as células malignas eram caracterizadas por métodos morfológicos ou citoquímicos. Mais recentemente foram utilizadas técnicas citogenéticas, as quais permitiram detectar alterações bem determinadas, e altamente específicas de células malignas. No entanto, todas estas técnicas envolvem o estudo e caracterização de um numero muito limitado de células, pelo que não permitem grande sensibilidade no estudo de doença residual mínima. O advento da genética molecular, trazendo de inicio um aumento da especificidade, mas sem grandes vantagens ao nível da sensibilidade (Southern Blot - ver fig.16), permitiu um grande avanço com a introdução da tecnologia de PCR, a qual permite em alguns casos identificar 1 célula maligna num universo de 10^6 células normais (fig. 16).

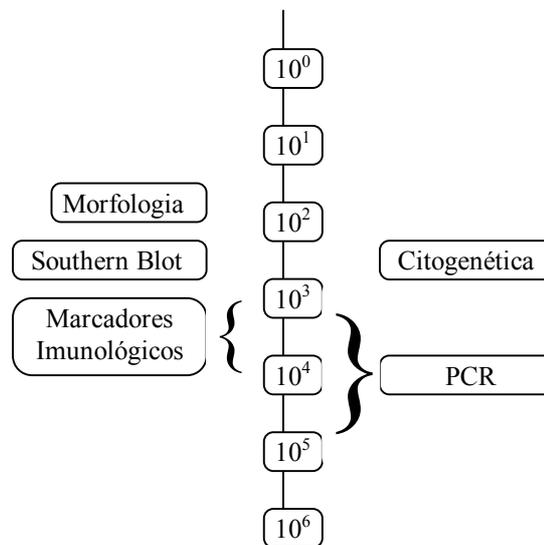
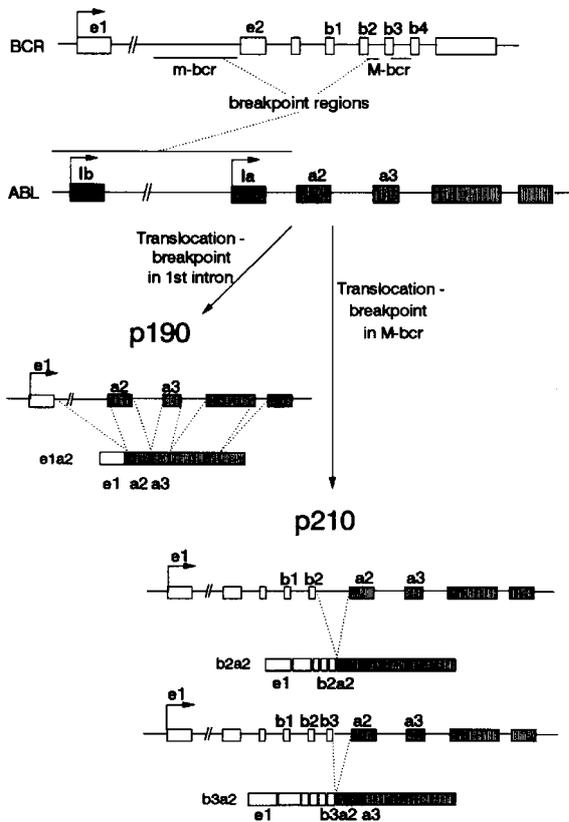


Figura 16 - Sensibilidade de várias metodologias na detecção da Doença Residual Mínima (10^1 - 10^6 = detecção de 1 célula maligna em 10^6 células normais)

7.1.1 - t(9;22) (q34;q11): BCR/ABL

A primeira anomalia cromossómica consistente em tumores humanos foi identificada por Nowell e Hungerford (19xx) na Leucemia Mielóide Crónica (CML) em 1960. Porque este achado foi realizado em Filadélfia, utilizou-se o nome desta cidade para designar esta translocação (t(9;22)(q34;q11) ou ainda cromossoma Ph. Estudos moleculares revelaram que esta translocação envolvia no cromossoma 9 o gene ABL (Abelson proto-oncogene) e no cromossoma 22 o gene BCR (Breakpoint Cluster Region gene) originando um gene químera codificando 1 proteína com capacidade oncogénica. O cromossoma Ph mais frequente, surge em cerca de 90% dos casos de CML, e variantes citogenéticas surgem em mais 5%. Dos restantes 5%, cerca de metade possui rearranjos do gene BCR não detectados por cariotipagem, mas visíveis por métodos moleculares, sendo os restantes 2.5% considerados Ph. O cromossoma Ph é ainda frequente em Leucemias Linfoblásticas Agudas (ALL; 5% das crianças e 20% dos adultos), e mais raramente em Leucemias Mieloblásticas Agudas (AML; cerca de 1%).

O cromossoma Ph é, como vimos, originado pela junção dos genes BCR e ABL. Esta junção, ocorre sempre no mesmo ponto no gene ABL, mas pode ocorrer em 3 locais diferentes do gene BCR. Estas diferentes junções, dão origem a 2 tipos de proteínas: a p190, resultante da junção do exon e1 do gene BCR com o a2 do gene ABL (transcrito e1a2) e a p210 resultante da junção do exon a2 do gene ABL com os exons b2 ou b3 do gene BCR (transcritos b2a2 e b3a2). A vasta maioria dos casos de CML (95%) expressam a proteína p210. Já na ALL, cerca de 70% dos casos de ALL expressam a p190, e os restantes 30% a p210.



Uma vez que o gene químera BCR-ABL é apenas expresso nas células malignas, a sua detecção molecular constitui um poderoso método de avaliar a progressão da doença. Com efeito, vários autores servindo-se da grande sensibilidade e especificidade da metodologia de PCR (reação em cadeia de polimerase), desenvolveram estratégias para avaliar a doença residual mínima, inferindo mesmo dados válidos na avaliação de prognóstico. Foi assim possível observar que se é frequente a detecção permanente ou intermitente de células residuais BCR-ABL⁺, vários meses após transplante de medula e remissão citogenética completa, já a sua detecção 1 ano após o transplante é indicadora de pior prognóstico que o dos casos em que se observe remissão por PCR. Não obstante estes dados, a validade da avaliação de prognóstico com base nos dados obtidos por PCR constitui, presentemente motivo de aceso debate e estudo.

Figura 17 - Os genes BCR e ABL normais, e as translocações que originam as proteínas p190 e p210 do gene químera BCR-ABL. A proteína p210 é característica da CML, sendo a p190 a proteína BCR-ABL encontrada na maioria dos casos de ALL (ver texto).

7.1.2 - t(15;17) (q22;q21): PML/RARA

A Leucemia Aguda Promielocítica (APL) é um subtipo da Leucemia Mieloblástica Aguda (FAB-M3), citogeneticamente caracterizada pela presença da translocação t(15;17)(q22;q21). Os genes envolvidos na translocação são o gene PML no cromossoma 15, e o gene RAR α (Receptor do Acido Retinóico) no cromossoma 17. Os pontos de rotura no *locus* RAR α não estão distribuídos ao acaso, mas localizados numa zona de 16 Kb do intron 2. De igual modo, os pontos de rotura no *locus* PML não são aleatórios, concentrando-se em apenas 3 regiões do gene: intron 3 (bcr3: 47% dos casos), exon 6 (bcr2: 4% dos casos), e intron 6 (bcr1: 49% dos casos). Como consequência da translocação, formam-se genes quimera (PML/RAR α e RAR α /PML).

O gene quimera PML/RAR α é transcricionalmente funcional, pelo que origina uma espécie de mRNA passível de detecção por RT-PCR. Foi assim possível determinar a presença deste transcrito em 100% dos casos de APL (em contraste com apenas 70% dos casos expressando o gene RAR α /PML), esta é uma tecnologia de grande valor na detecção de doença residual mínima nesta patologia. Se os estudos efectuados no final do tratamento parecem ter pouco valor prognóstico, já os estudos efectuados mais tarde parecem ter grande valor prognóstico, com resultados positivos em RT-PCR a indicarem uma recaída. Com efeito, estudos de doentes em remissão por períodos prolongados de tempo (4-12 anos) revelaram que a sobrevida está associada com a erradicação das células PML/RAR α , pelo que este deve ser o objectivo terapêutico. Desta forma, foi sugerido que um teste positivo deve ser indicativo da continuação do tratamento, ao passo que doentes com 2 testes negativos, e mais de 2 meses de remissão completa, podem ser poupados a sessões terapêuticas.

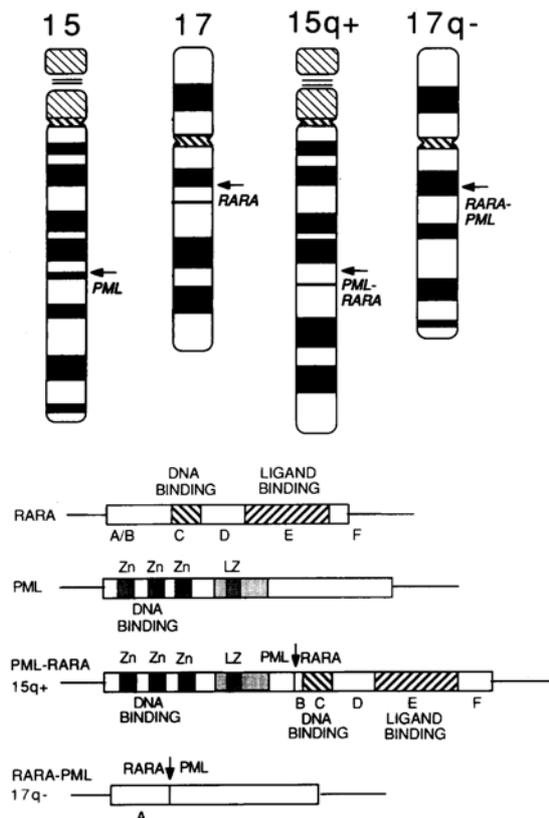


Figura 18 - A localização cromossômica e estrutura normal dos genes PML e RAR α , e a translocação t(15;17)(q22;q21) que origina o gene quimera PML-RARA.

7.1.3 - O Linfoma de Burkitt e a translocação t(8;14)(q24;q32)

No Linfoma de Burkitt, ocorrem 3 tipos de translocações recíprocas, todas envolvendo o gene MYC (8q24), e genes dos cromossomas 2 (IgK), 14 (IgH) e 22 (Igλ). A translocação mais frequente é a t(8;14)(q24;q32), a qual ocorre em mais de 75% destes linfomas. O gene MYC é conhecido pela sua importância na proliferação celular, mas não é expresso nas células B maduras. No entanto, estas translocações colocam este gene dependente de enhancers das imunoglobulinas, pelo que o gene passa a estar activo nas células B que possuem estas translocações, dando origem a níveis de mRNA semelhantes aos encontrados nas células normais em proliferação.

Estudos moleculares revelaram a existência de 2 mecanismos de geração da translocação t(8;14)(q24;q32). O primeiro, ocorre no Linfoma endémico da África equatorial, associado à infecção por EBV, o gene MYC não é rearranjado, encontrando-se intacto, se bem que próximo das regiões DH ou JH do gene IgH. Esta mutação ocorre no estágio celular pré-B, quando a maquinaria de recombinação dos genes das Imunoglobulinas está activa. O segundo mecanismo de geração desta translocação não está associado à infecção pelo EBV. Neste caso, a translocação ocorre imediatamente 3' do gene MYC, ou dentro deste, envolvendo ainda a região de "switch" do gene IgH. As células neste caso apresentam um fenótipo mais maduro, compatível com a ocorrência da mutação

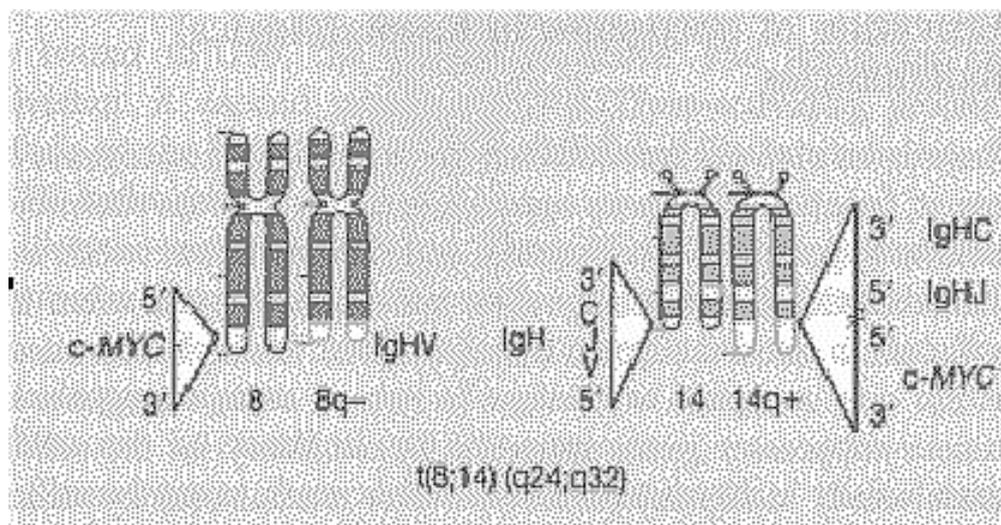


Figura 19 - Diagrama mostrando os eventos genéticos geradores de uma das três translocações encontradas no Linfoma de Burkitt. O oncogene c-MYC está normalmente localizado no braço longo(q) do cromossoma 8. A translocação t(8;14) coloca este oncogene junto ao loc

numa altura em que a célula efectuava o "switch" de imunoglobulinas.

7.1.4 - t(14;18)(q32;q21): BCL2/IgH

Em cerca de 85% dos "Follicular Lymphomas" (FL), e 25% dos Linfomas Difusos (DL), surge a translocação t(14;18)(q32;q21), envolvendo os genes BCL-2 (B-Cell Lymphoma/Leukemia-2 gene) no cromossoma 18 e um dos segmentos JH do gene da IgH no cromossoma 14. O gene BCL-2 codifica uma proteína que parece ter potencial oncogénico sendo importante na fase pré-B do desenvolvimento do

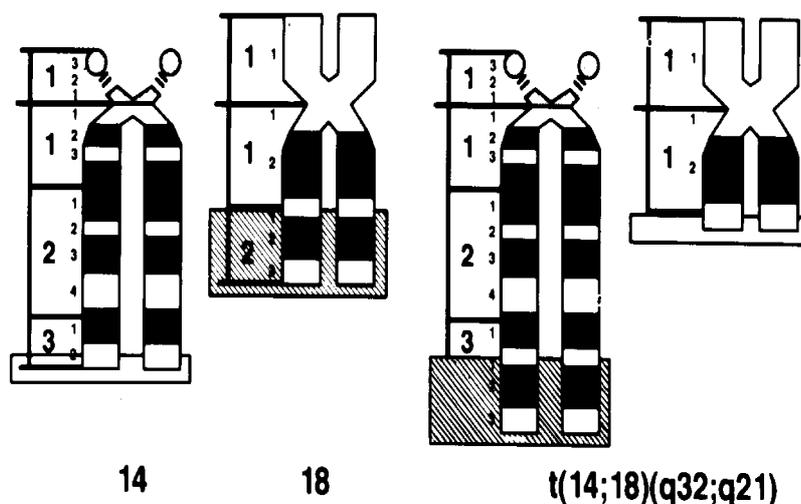


Figura 20 - Diagrama mostrando os cromossomas 14 e 18 normais, bem como os cromossomas resultantes da translocação t(14;18)(q32;q21), a qual envolve os genes BCL-2 (18q21) e IgH (14q32).

linfócito B, ao prevenir a morte celular por apoptose. Duas regiões de quebra foram identificadas no cromossoma 18: 2/3 das translocações envolvem uma região de 150 bp na zona 3' não traduzida do gene (o Major Breakpoint region ou mbr). As restantes translocações envolvem o Minor cluster region (mcr) localizado cerca de 20 Kb após o início do gene. A translocação parece não afectar a sequência do BCL-2, mas tão somente os níveis de mRNA deste gene, e ocorre presumivelmente por erro na maquinaria genética de recombinação das Ig. Esta interpretação parece ser suportada pela descoberta de regiões N na junção dos "breakpoints", bem como pela existência de mutações somáticas na mesma zona.

O advento do PCR transformou o estudo molecular das mutações envolvendo o gene do BCL-2, e particularmente a translocação t(14;18), tanto a nível do mbr como do mcr. Desta forma foi possível detectar 1 célula mutante num universo de 100,000 células normais, permitindo uma nova sensibilidade na detecção de doença residual mínima.

7.1.5 - Mutações de p53 em CLL

Apesar de anomalias citogenéticas envolvendo 17p13 serem raras, o gene p53 aí localizado encontra-se mutado (como determinado por SSCP) em 10-15% dos casos de CLL. Os doentes com deleções ou translocações que envolvem esta zona, possuem quase invariavelmente mutações deste gene. Existe ainda uma forte correlação entre a existência de mutações no gene p53, e um estadio avançado, resistente à quimioterapia, e curta sobrevida.

7.2 - Detecção de monoclonalidade

A detecção de monoclonalidade das células T e B, ainda que facilitada pela presença de genes rearranjados de forma específica para cada clone tem-se mostrado laboriosa, pelo que mais propicia ao laboratório de investigação que ao de rotina. Apesar disto, o interesse clínico pela definição laboratorial precisa de monoclonalidade permanece, já que este parece ser um critério de discriminação entre situações malignas e situações reactivas. O valor diagnóstico deste critério foi

no entanto recentemente posto em causa com a descoberta de expansões monoclonais em indivíduos sem qualquer expressão clínica. Estas expansões, cuja frequência parece aumentar com a idade, não podem assim ser interpretadas como evidência de malignidade. Ainda assim, não só a possibilidade de excluir monoclonalidade pode indicar a existência de situações reactivas, como a possibilidade de detectar quantidades pequenas de células com um determinado rearranjo permanece uma útil ferramenta no acompanhamento da acção terapêutica, dando indicações relativas à doença residual mínima.

A tecnologia de PCR, aumentando a sensibilidade e especificidade, relativamente a outras técnicas de genética molecular, abriu novas potencialidades de detecção de doença residual mínima, reduzindo o número mínimo de células malignas passíveis de detecção. No entanto, devido à complexidade dos *locus* génicos do Receptor da célula T (TCR) e das imunoglobulinas (Ig), só recentemente foi possível utilizar com sucesso esta técnica na detecção de expansões clonais de linfócitos, e ainda assim limitados à linhagem $\gamma\delta$. No entanto, trabalhos independentes com outros objectivos, colocaram à mercê do laboratório de Hematologia Clínica, as ferramentas necessárias para a montagem de estratégia semelhante na detecção de monoclonalidade de linfócitos $\alpha\beta$.

LEITURA COMPLEMENTAR

BCR/ABL

- Israeli S., Janssen JWG., Haas OA. et al. *Leukemia* 7:671-8, 1993
- Secker-Walker LM., Craig JM., Hawkins JM. et al. *Leukemia* 5:196-9,1991
- Cross NCP, Hughes TP, Lin F et al. *Brit.J.Haem.* 84:67-74, 1993
- Cross NCP, Lin F., Chase A. et al. *Blood* 82: 1929-36, 1993
- Lion T., Henn T., Gaiger A. et al. *Lancet* 341: 275-276, 1992
- Miyamura K., Tahara T., Tanimoto et al. *Blood* 81: 1089-93, 1993
- Roth MS, Antin JM, Ash R. et al. *Blood* 79: 276-282, 1992
- Cross NCP, Melo JV., Feng L. et al. *Leukemia* 8: 186-9, 1994
- Lin F., Goldman JM., Cross NCP. *Brit.J.Haem.* 86:683-5, 1994

PML-RAR α

- Sklar J., *J.Clin.Oncol.* 9: 1521-4,1991
- Longo I., Pandolfi PP, Biondi A. et al. *J.Exp.Med.* 172: 1571-5, 1990
- Borrow J., Goddard AD., Sheer D. et al. *Science* 249:1577-80,1990
- de Thé H., Zangrilli D., Pandolfi PP. et al. *PNAS* 88: 1977-81, 1991
- Diverio D., Lo-Coco F., D'adamo et al. *Blood* 79:3331-6, 1992
- Pandolfi PP., Alcalay M., Fagioli M. et al. *EMBO J.* 11:1397-1407, 1992
- Pandolfi PP., Grinani F., Alcalay M. et al. *Oncogene* 6: 1285-1292, 1991
- de Thé H., Lavau C., Marchio A. et al. *Cell* 66:675-684, 1991
- Kakizuka A., Miller WR., Umesono K. et al. *Cell* 66:663-674, 1991
- Biondi A., Rambaldi A., Pandolfi PP et al. *Blood* 80:492-497, 1992
- Alcalay M., Zangrilli D., Fagioli M. et al. *PNAS* 89: 4840-4, 1992
- Lo-Coco F. Diverio D., Pandolfi PP. et al. *Lancet* 340:1437-8, 1992
- Diverio D., Pandolfi PP., Biondi A. et al. *Blood* 82: 3556-9, 1993
- Chang K., Lu J., Wang G. et al. *Blood* (80 ou 81): 554-8,1992
- Alcalay M., Zangrilli D., Pandolfi PP. et al. *PNAS* 88:1977-81,1991

BCL-2/IgH

- Stetler-Stevenson M., Raffeld M., Cohen P. et al. *Blood* 72:1822-5, 1988

Monoclonalidade

- Botaro M., Barti E., Biondi A. et al. *Blood* 83:3271-78, 1994
- Bourguin A., Tung R., Galili N., Sklar J. *PNAS* 87:8536-40, 1990
- Casorati G., de-Libero G., Lanzavecchia A. et al. *J.Exp.Med.* :1521-35,1989
- Obata F., Tsunoda M., Ito K. et al. *Human Immunol.* 36:163-7, 1993

8 - DETECÇÃO E CARACTERIZAÇÃO DE VÍRUS

8.1 - O vírus da Hepatite B (HBV)

O título de vírus da Hepatite B necessário para causar infecção é reconhecidamente inferior ao título que é passível de detecção por marcadores serológicos. Assim, a detecção molecular do genoma viral, para além de constituir uma forma altamente específica de detecção do vírus, permite ainda a detecção de títulos virais mais baixos e patologicamente activos, bem como o acompanhamento da eficácia da acção terapêutica. Se à especificidade da reacção de hibridação, for ainda aliada a sensibilidade da reacção de PCR, obtém-se um método analítico extremamente sensível e específico, e portanto de grande utilidade clínica.

8.2 - O Vírus da Hepatite C (HCV)

O vírus da Hepatite C (HCV), isolado pela primeira vez em 1989, parece ser um importante factor causador de doença hepática crónica, cirrose e carcinoma hepatocelular, em todo o mundo. Hoje em dia foram já isoladas várias estirpes virais do HCV, o que permitiu desenvolver testes serológicos e de genética molecular (detecção de RNA viral) para a determinação da virémia.

O genoma do HCV possui 9379 nucleótidos, sendo constituído por uma cadeia simples de RNA com um único e longo “open reading frame”. O produto genético é uma proteína precursora com 3011 amino-ácidos, que por proteólise pós-traducional origina proteínas estruturais (core e envelope) e não estruturais (proteases, elicases, polimerases do RNA).

A detecção do HCV não é possível de ser realizada por testes padrão de detecção de antígenos no soro, já que as partículas virais circulam no soro em concentrações abaixo das detectáveis por imunoensaios. Assim, a maioria dos estudos epidemiológicos foram inicialmente baseados na prevalência de anticorpos contra o antígeno c100-3. Presentemente, é ainda testada a presença de anticorpos contra outros antígenos virais (testes ELISA de segunda geração). No entanto, estes testes originam um grande número de falsos positivos, pelo que se tornou necessário o desenvolvimento de testes confirmativos. Estes testes, baseando-se na detecção de RNA viral, após ou na ausência de amplificação por PCR.

Como já foi referido, várias estirpes do gene da Hepatite C foram já identificadas (estirpes 1a, 1b, 2, 2a, 2b, 3a, 4, 5), sendo neste momento incerto, se estirpes diferentes têm prognósticos e/ou repostas terapêuticas diferentes.

8.3 - O vírus Linfotrópico Humano (HTLV-I e HTLV-II)

O vírus HTLV tipo I foi o primeiro retrovírus humano descrito, sendo considerado o agente causador da Leucemia/linfoma da célula T de Adultos (ATLL). Este vírus tem ainda sido associado com uma família de doenças neurológicas incluindo “spastic paraparesis e Mielopatia associada a HTLV-I. Mais recentemente, também a polimiositose e a poliartrite têm sido associadas a infecções por HTLV-I. As manifestações clínicas, e epidemiologia do HTLV-II são no entanto menos claras, não havendo nenhum síndrome clínico especificamente associado à infecção por este vírus.

Devido à razão custo/benefício particularmente boa dos testes de detecção do HTLV baseados em ELISA, este é habitualmente o método de eleição para o screening de produtos sanguíneos. No entanto, e dado o elevado número de falsos positivos apresentados, um resultado positivo deve ser confirmado por um método diferente. Os métodos de detecção do genoma viral baseados em PCR e hibridação constituem testes poderosos complementares à ELISA. Devido à grande sensibilidade e especificidade do PCR, aliada à enorme sensibilidade da hibridação de DNA, este método permite com rapidez, sensibilidade e especificidade determinar a presença mesmo de níveis muito baixos de partículas virais, bem como seguir a resposta à terapêutica muito depois de os resultados por ELISA serem negativos.

8.4 - O vírus do síndrome da imunodeficiência adquirida (HIV)

O vírus da Imunodeficiência Humana (HIV) é demonstradamente o causador do Síndrome Humano de Imunodeficiência adquirida (SIDA). Sendo um retrovírus, é capaz de inserir uma cópia de DNA do seu genoma de RNA no genoma da célula infectada, podendo manter-se desta forma num estado latente.

Dos testes disponíveis para a detecção do HIV, o EIA ou o ELISA são habitualmente utilizados. No entanto, e devido ao elevado número de falsos positivos, um resultado positivo deve ser confirmado por uma técnica complementar, habitualmente Southern Blot. Os métodos baseados na reacção de PCR, ainda que de uso relativamente recente, constituem poderosos meios complementares ao ELISA, já que associam a enorme sensibilidade da reacção de PCR com a especificidade quer do PCR quer da hibridização. Estes métodos têm ainda a vantagem de detectar directamente o genoma viral, independentemente do estado imunológico ou replicativo do vírus, aumentando por isso a sua eficiência de detecção. Estes testes são ainda importantes na avaliação clínica da infecção em crianças nascidas de mães seropositivas, já que é independente da presença de anticorpos ou antigénios provenientes do sangue materno.

8.5 - O vírus do Epstein bar (EBV)

O vírus do Epstein Bar (EBV). é um vírus Herpes humano associado a tumores. Cerca de 95% dos adultos estão infectados por este vírus, sendo a infecção permanente. O EBV está classicamente associado ao Linfoma de Burkitt (BL), o qual é endémico em África e à doença linfoproliferativa do linfócito B (BLPD), característica dos indivíduos imunocomprometidos. Recentemente foram ainda identificadas associações entre o EBV e subtipos de Linfoma de Hodgkin e de Linfomas T.

O genoma viral é composto entre outros pelos genes de expressão latente: 6 genes de expressão nuclear (EBNA- Epstein Bar nuclear antigens), 3 genes codificando proteínas membranares (LMP e TP1 e 2). Todas as proteínas de expressão latente, com a excepção de EBNA1 são apresentadas pelo HLA, e reconhecidas pelas células citotóxicas, contribuindo para a manutenção da infecção num estado suportável pelo hospedeiro. Isto mesmo é exemplificado no facto de a maioria dos tumores associados a EBV que surgem nos doentes imunossuprimidos, regredirem quando a imunossupressão é reduzida ou retirada.

O Linfoma de Burkitt (BL) é fundamentalmente diferente da Doença Linfoproliferativa do Linfócito B, já que a primeira surge num indivíduo imunocompetente. Neste caso, o vírus parece escapar à vigilância do sistema imune, expressando apenas o gene EBNA1, o qual como vimos não é apresentado pelo HLA. No caso do BL, e como vimos (secção 2.1.3), existem ainda anomalias genéticas associadas à infecção pelo EBV, conduzindo à desregulação do c-myc e ou p53, o que contribui para o fenótipo maligno da célula infectada. Ainda que a implicação definitiva do EBV na etiologia dos linfomas que ocorrem em indivíduos infectados com o vírus da SIDA seja ainda motivo de investigação, o facto de a replicação do EBV nestas células ser muito alta, e de uma única estirpe parece favorecer a implicação do EBV.

Assim, parece ser importante controlar a replicação do EBV nos indivíduos afectados por estes tumores, sendo ainda necessário dispor de tecnologias laboratoriais para detectar a existência de baixos níveis de vírus replicativamente activos nos indivíduos em tratamento.

Leitura Complementar

HBV

- Zaaijer HL., ter-Borg F., Cuypers HTM. et al. *J.Clin.Microbiol.* 32:2088-91, 1994
- Carman WF., Hadziyannis S., Jacyna MR. et al. *Lancet* : 588-90, 1989
- Krogsgaard K. *Liver* 8: 257-83, 1988 (review)
- Hoofnagle JH. *New Engl.J.Med.* 323:337-9, 1990
- Paterlini P., Gerken G., Nakajima E. et al. *New Engl.J.Med.* 323:80-85, 1990
- Marcelini P., Martinot-Peignoux M. et al. *Ann.Intern.Med.* 112:227-8, 1990
- Lo ESF., Lo YMD. Tse CH, Fleming KA. *J.Clin.Pathol.* 45:689-92, 1992
- Larsen J., Hetland G. Skaug K. *Transfusion* 30:431-2, 1990
- Hoofnagle JH. *Transfusion* 30:384-6, 1990

HCV

- de-Vita S., Sansonno D., Dolcetti R. et al., *Blood* 86: 1887-92, 1995
- Alter H. *Blood* 85:1681-95, 1995 (Review)
- Damen M., Zaaijer HL., Cuypers HTM. et al. *Transfusion* 35:745-9, 1995
- Nalpas B., Romeo R., Pol S. et al. *Transfusion* 35:750-53
- Dusheiko GM. *Lancet* 345:562-6, 1995
- Shindo M., Arai K., Sokawa Y. et al. *Ann.Intern.Med.* 122:586-91, 1995
- Viazov S., Zibert A., Widell A., et al. *J.Vir.Methods* : , 1994
- Young KKY., Resnick RM., Myers T.W. *J.Clin.Microbiol.* 31:882-6,1993
- Pawlotsky JM., Roudot-Thoraval F. et al. *Ann.Intern.Med.* 122:169-73, 1995
- van-der-Poel CL., Cuypers HT., Reesink HW. *Lancet* 344:1475-9, 1994(review)
- Nousbaum JB., Pol S., Nalpas B. et al. *Ann.Intern.Med.* 122:161-8, 1995
- Imberti L., Cariani E., Bettinardi A. et al. *J.Virol.Methods* 34:233-43, 1991
- Lazizi Y., Elfassi E., Pillot J. *J.Clin.Microb.* 30:931-4, 1992
- Shibata M., Morishima T., Kudo T. et al. *Blood* 77: 1157-60, 1991
- Garson JÁ., Tedder RS., Briggs M. et al. *Lancet* 335:1419-22, 1990
- Shibata M., Kudo T., Kajiyama M. et al., *Am.J.Med.* 89:830-832, 1990

HIV

- Mellors JW., Kingsley LA., Rinaldo CR. et al. *Ann.Intern.Med.* 122:573-9, 1995
- Laure F., Rouzioux C., Veber F. et al. *Lancet* :538-40, 1988
- Imagawa DT., Lee MH., Wolinsky SM. et al. *New Engl.J.Med.* 320: 1458-62, 1989
- Rogers MF., Ou CY., Rayfield M. et al. *New Engl.J.Med.* 320:1649-54, 1989
- Katz SL., Wilfert CM. *New Engl.J.Med.* 320: 1687-8, 1989
- Perrin LH., Yerly S., Bachmann P. et al. *Blood* 76:641-5, 1990
- Brandt CD., Rakusan TA., Sison AV. et al. *J.Clin.Microb.* 30: 36-40, 1992
- Schochetman G & Shinsky JJ. in AIDS testing: Methodology and Management issues. Schochetman G & George R (eds). Springer-Verlag. p 90-110

9 - ESTUDO DE DOENÇAS GENÉTICAS

9.1 - DOENÇAS GENÉTICAS DO GLÓBULO RUBRO

9.1.1 - Anemias Não esferocíticas Congénitas

9.1.1.1 - DEFICIÊNCIA EM GLUCOSE-6-FOSFATO DESIDROGENASE

A deficiência em glucose-6-fosfato desidrogenase (G6PD) é uma anomalia genética muito frequente, sendo estimado que afecte cerca de 400 milhões de indivíduos em todo o mundo. A maioria dos portadores desta deficiência é assintomática, correndo no entanto o risco de desenvolver anemias hemolíticas agudas, quando expostos a certas infecções, drogas, ou ingestão de favas. Uma pequena porção dos portadores da deficiência sofre de uma doença mais pronunciada: anemia não-esferocítica congénita.

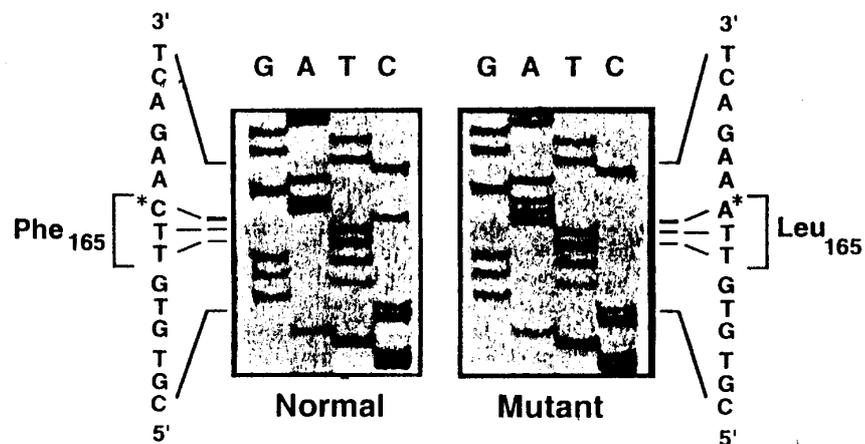


Figura 21 - Gel de sequenciação do exon 5 do gene ALAS2, indicando uma transversoão C para A na posição 547, prevenindo a substituição de uma Phe por uma Leu no aminoácido 165 da enzima.

A G6PD é um homodímero. O gene que codifica o respectivo polipéptido de 514 aminoácidos é extremamente conservado na escala evolutiva, estando localizado no homem no cromossoma X (Xq28). Este gene é expresso em todas as células do organismo, sendo essencial à viabilidade celular. Até ao momento foram identificadas 75 mutações de G6PD, correspondendo a mais de 100 variantes enzimáticas. As mutações constituem quase exclusivamente mutações Missense, causando portanto a substituição de um único aminoácido. Dois grupos de situações são de relevância clínica: a primeira é composta pelas mutações que afectando a actividade enzimática, deixam no entanto actividade suficiente para o metabolismo normal do eritrócito. Neste caso, os portadores são assintomáticos, enquanto não existirem factores externos propensores ao desenvolvimento de anemia hemolítica aguda. No segundo caso, a deficiência em G6PD é tão severa, que os indivíduos desenvolvem anemia não esferocítica congénita.

A caracterização das deficiências genéticas da G6PD permitiu verificar que mutações diferentes podem originar uma deficiência com semelhantes características bioquímicas, e vice-versa.

9.1.1.2 - DEFICIÊNCIA EM PIRUVATO QUINASE

A Deficiência em Piruvato Quinase (PKD) é a causa mais comum de Anemia Hemolítica Não esferocítica Hereditária. Geralmente assintomática nos portadores (heterozigóticos), a deficiência de Piruvato Quinase manifesta-se nos indivíduos homozigóticos (ou duplamente heterozigóticos) como anemia hemolítica crónica, com severidade variável.

No Homem existem dois genes de piruvato quinase: PKLR (codificando as isoenzimas L e R) e PKM2 (codificando as isoenzimas M1 e M2). o primeiro é o único habitualmente expresso em eritrócitos, sendo o responsável pela PKD nestas células.

A caracterização bioquímica da PK permitiu identificar cerca de 300 variantes enzimáticas, no entanto a caracterização molecular das mutações que afectam o gene da PK permitiu verificar uma heterogeneidade mais limitada, já que a mesma mutação parece estar associada a variedades bioquímicas diferentes (ex.: PK Nagasaki, PK Tóquio e PK Beirut resultam da mutação ¹¹⁵¹ACCG-ATG, e as variantes PK Fukushima e PK Maebashi resultam da mutação ³⁴⁹CAG - AAG).

As mutações encontradas gene PKLR, responsáveis pela PKD, incluem mutações tipo “*missense*”, “*nonsense*” e inserções. Estas mutações foram encontradas quer na região codificante quer na região promotora. De grande utilidade nos estudos familiares da PKD foi a descoberta de polimorfismos de microssatélites no intron 11, bem como de um polimorfismo C/A na posição 1705, os quais podem ser utilizados para seguir os haplótipos herdados dos progenitores.

9.1.1.3 - DEFICIÊNCIA EM δ -AMINOLEVULINATO SINTETASE (ANEMIA SIDEROBLÁSTICA)

A enzima δ -aminolevulinato sintetase (ALAS) catalisa o primeiro passo da síntese do grupo heme. Nos vertebrados existem 2 formas de ALAS: uma forma transcrita constitucionalmente em todas as células, e uma forma apenas existente nas células eritróides.

No Homem, a forma eritróide (ALAS2) é codificada num gene localizado no cromossoma X (Xp11.21), estando a sua actividade reduzida na anemia sideroblástica ligada ao cromossoma X. Que a origem da anemia sideroblástica está relacionada com este gene ficou claramente demonstrado quando em 1992 Bishop descreveu a primeira mutação neste gene, e a sua segregação ao longo de 8 gerações de uma família afectada. A identificação molecular desta mutação, permitiu ainda pela primeira vez desenhar métodos de diagnóstico e estudo de transmissão familiar para esta doença.

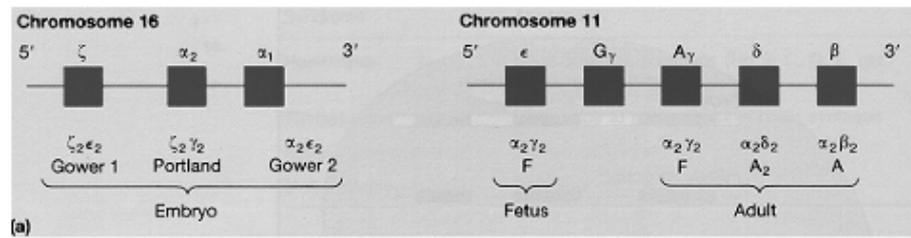
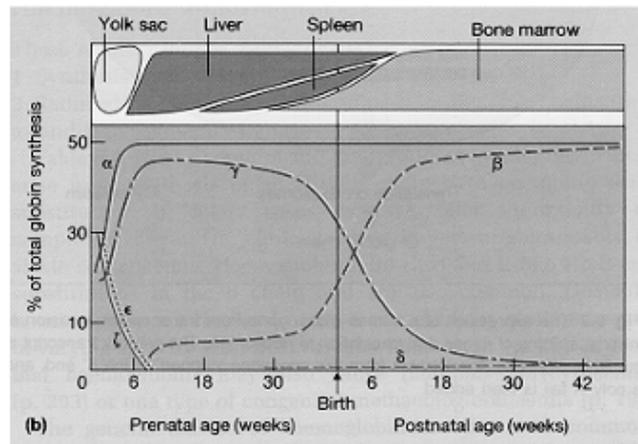


Figura 22 - Os clusters dos genes da globina nos cromossomas 11 e 16 (a). Durante a vida embrionica, fetal e adulta, os genes activados e suprimidos são diferentes (b). As diferentes cadeias da globina são sintetizadas independentemente, associando então para for



9.1.2 - Talassémias (anomalias das α e β -globinas)

Os *loci* genéticos das α - e β -globinas incluem vários genes arranjados em “clusters”. O cluster humano da α -globina, inclui um gene enbriónico (ζ), dois genes fetais/adultos (α_2 e α_1), e vários pseudogenes ($\psi\zeta_1$, $\psi\alpha_2$ e $\psi\alpha_1$) bem como um gene sem função conhecida (θ_1). Os genes da globina são expressos em níveis muito elevados nas células eritróides, mas não são expressos em nenhuma outra célula.

A α -talassémia resulta da produção deficiente das cadeias α da hemoglobina embriónica ($\alpha_2\epsilon_2$), fetal ($\alpha_2\gamma_2$) e adulta ($\alpha_2\beta_2$). As formas mais frequentes consistem na deleção de um ou ambos os genes α do cromossoma 16. Desta forma, os portadores de α -talassémia possuem 3 ($-\alpha/\alpha\alpha$) ou dois

($-\alpha/-\alpha$, $-/\alpha\alpha$) genes, ao passo que os doentes possuem apenas 1 gene ($-/-\alpha$). Os doentes com síndrome de “Hb Bart’s hydrops fetalis” não possuem genes α ($-/-$). Outras causas menos frequentes de α -talassémia são mutações pontuais no gene da α -globina, e muito raramente em elementos reguladores destes genes.

A β -talassémia caracteriza-se por uma síntese reduzida da β -globina, levando a um desequilíbrio da síntese das cadeias α /não α que é o factor maior na gravidade da doença.

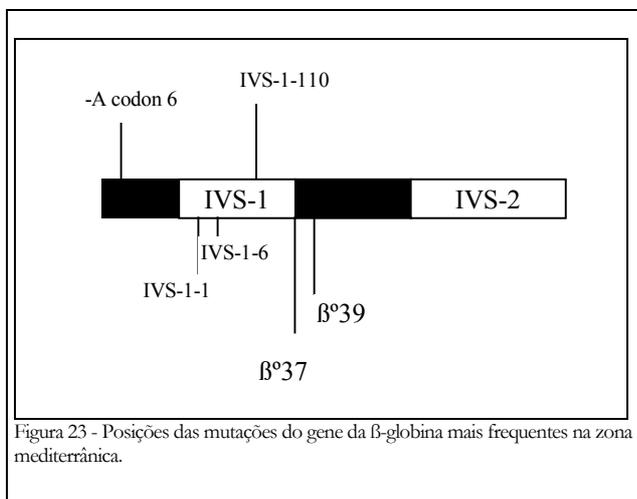


Figura 23 - Posições das mutações do gene da β -globina mais frequentes na zona mediterrânica.

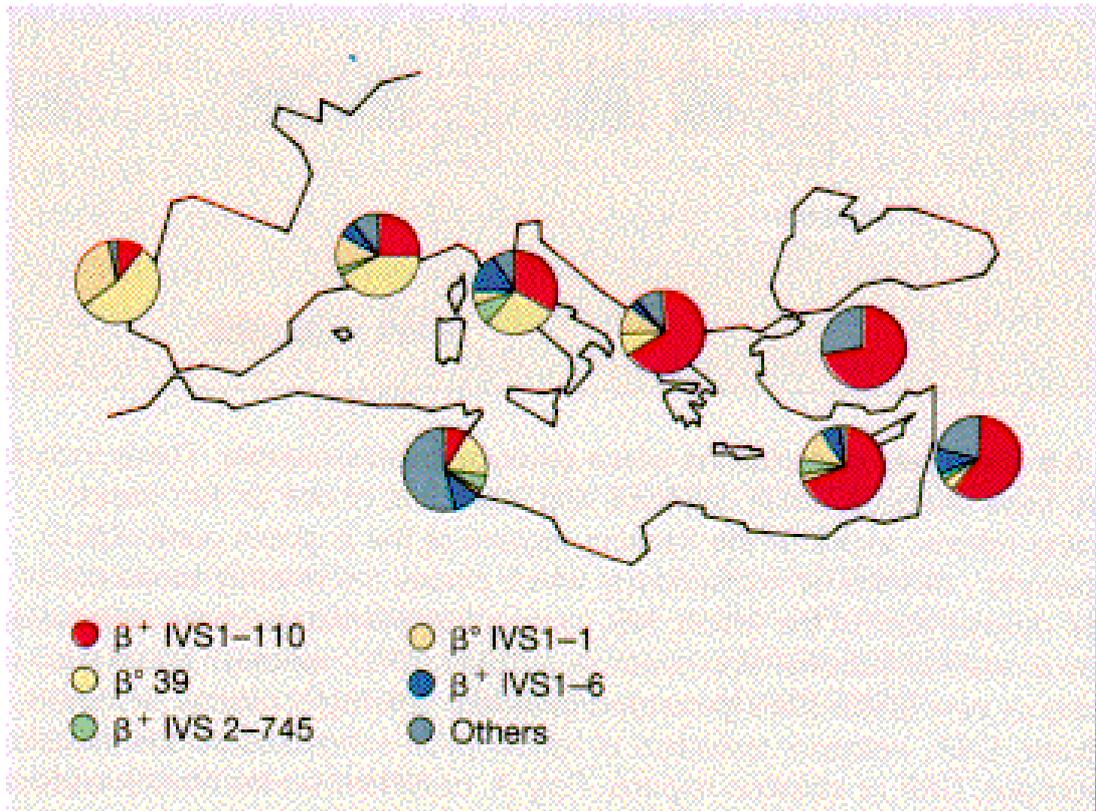


Figura 24 - Distribuição geográfica de algumas mutações na β -globina

Os genes da β -globina estão arranjados num cluster no braço curto do cromossoma 11, na ordem $5'-\epsilon-\gamma-A\gamma-\psi-\delta-\beta-3'$. O cluster contém muitos polimorfismos de restrição (RFLP). Uma característica destes polimorfismos, é que a sua associação às várias formas de β -globina não é ao acaso. Assim, em cada população foi encontrado um número limitado de haplótipos β (padrão de arranjos dos RFLP), pelo que a análise destes haplótipos fornece informação clínica (dependente do background genómico em que as mutações ocorrem). No início dos anos 80 esta era a estratégia de eleição para o screening da β -talassémia. Esta tecnologia era então complementada com a clonagem e sequenciação dos genes mutantes, completando-se assim a caracterização destes. Com o desenvolvimento da tecnologia de PCR em 1985, foi possível amplificar o gene da β -globina directamente a partir de DNA genómico, sequenciando directamente os produtos amplificados, facilitando deste modo a caracterização das mutações.

Ao contrário das mutações causadoras de α -talassémia, as β -talassémias são geralmente formas não deletionadas, podendo ocorrer na região codificante do gene, ou na região promotora. Existem mutações particularmente frequentes em determinadas comunidades, o que aliado à tecnologia de PCR em combinação com a utilização de reacções de restrição, simplifica grandemente a detecção e caracterização de mutações em estudos de screening, e de modo particular em estudos familiares.

9.1.3 - Esferocitose e eliptocitose hereditária

A esferocitose hereditária (HS), eliptocitose hereditária (HE) e a sua forma agravada Poiquilocitose hereditária (HP) são um conjunto heterogêneo de anemia hemolíticas congênitas. Estas patologias resultam de alterações nas proteínas da célula vermelha, e seu esqueleto membranar.

9.1.3.1 - ESFEROCITOSE HEREDITÁRIA (DEFICIÊNCIAS DE ANQUIRINA E BANDA 3)

Cerca de 40% dos casos de HS apresentam mutações afectando o gene da anquirina. Mutações tipo “*frameshift*” causam alterações estruturais muito grandes na proteína, causando formas dominantes (anquirina de estutgarda, e anquirina de Marburg). Alterações menos severas são causadas por mutações pontuais, e originam formas recessivas (anquirina de Düsseldorf, anquirina de Walsrode). Em alguns casos, as alterações só são visíveis no mRNA, já que resultam de junções *splicing* anormais (anquirina de praga, anquirina de Rakivnik).

Cerca de 20 a 30% dos casos de HS resultam de mutações no gene da proteína banda 3 (mutação de praga e mutação de Coimbra). Outro tipo de mutações (mutante de tuscaloosa e mutante de Montefiore), ocorrendo na zona citoplasmática da banda 3, causam perturbações na sua ligação à proteína 4.2, resultando num menor nível desta proteína na membrana eritrocitária.

9.1.3.2 - ANEMIA HEMOLÍTICA HEREDITÁRIA (DEFICIÊNCIA DE PROTEÍNA 4.2)

A total, ou quase total ausência de banda 4.2 resulta num quadro clínico diferente da HS. Os esferócitos não existem, e a fragilidade osmótica é pouco alterada. A hemólise é muito severa, e o padrão de transmissão recessivo. Foram identificados 4 mutantes do gene da proteína 4.2, os quais estão na origem deste quadro clínico (mutante de Nippon, mutante de Toseur, e mutante de Lisboa).

9.1.3.3 - ELIPTOCITOSE E POIQUILOCITOSE HEREDITÁRIA (DEFICIÊNCIAS DE α - E β -ESPECTRINA; PROTEÍNA 4.1 E GRICOFORINA C)

São conhecidas 25 mutações de α -espectrina, as quais se localizam em zonas onde perturbam o processo de auto-associação. Dependendo de o alelo não mutado ser de alta ou baixa expressão, assim a mutação pode ter uma representação suave ou mais acentuada.

Todas as mutações de β -espectrina conhecidas que originam HE localizam-se na repetição $\beta 17$, a qual contem o local responsável pela auto-dimerização. Mutações pontuais na hélice 2 ou mesmo na hélice 1 desta repetição possuem um padrão de transmissão recessivo. No entanto outras mutações, ocasionando truncagens de β -espectrina são transmitidas segundo um padrão dominante.

Cerca de 30% dos casos de HE resultam de alelos *Null* (não expressos ou não funcionais) do gene da proteína 4.1. Esta condição é clinicamente silenciosa na forma heterózigótica. Duas mutações foram descritas: uma mutação pontual no codon de iniciação “downstream” (o único existente no precursor do eritrócito), e uma deleção de um aminoácido no local de ligação do complexo actina-

espectrina. Outros exemplos de defeitos a este nível constituem os alelos que originam erros de *splicing*.

Mutações raras originam a falta de glicoforina C. Como a proteína 4.1 se liga à glicoforina, a falta desta origina a falta de proteína 4.1.

9.2 - DOENÇAS GENÉTICAS EM HEMOSTASE

A hemostase, constitui um riquíssimo campo de intervenção da genética molecular no diagnóstico, estudos familiares e rastreio populacional, já que o número de mutações e polimorfismos associados a doenças genéticas em hemostase é relativamente elevado (tabela 1). As técnicas utilizadas para o estudo das doenças genéticas em hemostase cobrem o espectro completo das técnicas de genética molecular, sendo portanto um ótimo exemplo da larga gama de tecnologias hoje disponíveis para estes estudos (tabela 1).

9.2.1 - Resistência à proteína C Activada (mutação FV-Leiden)

Recentemente, apenas 10% dos indivíduos que apresentavam tromboembolismo venoso possuíam uma anomalia genética predispondo à doença. Este grupo de doentes tinha uma deficiência absoluta ou funcional de um dos principais componentes dos mecanismos de regulação da coagulação (ATIII, Prot.C, Prot.S). Recentemente, um novo mecanismo foi encontrado para justificar esta patologia: a Resistência à Proteína C Activada (APCR). Este mecanismo, estando implicado na origem de cerca de 50% dos casos de trombose venosa (dependendo da selecção dos casos), parece ter origem numa mutação pontual do gene do Factor V (Factor V de Leiden).

É hoje habitualmente aceite que todos os casos de APCR têm origem nesta mutação, a qual está presente em cerca de 5% da população normal, sendo cerca de 20-50% dos doentes com tromboembolismo portadores heterozigóticos desta mutação. No entanto, não parece haver um aumento da frequência de portadores da mutação em doentes que tenham sofrido enfarte do miocárdio ou ataque cardíaco, sugerindo um papel limitado na doença arterial.

A presença da mutação, a qual pode ser facilmente detectada por PCR seguido de corte por enzima de restrição, aumenta em cerca de 8 vezes o risco de tromboembolismo. Este risco pode ainda ser aumentado por factores adicionais, que podem ser genéticos ou adquiridos.

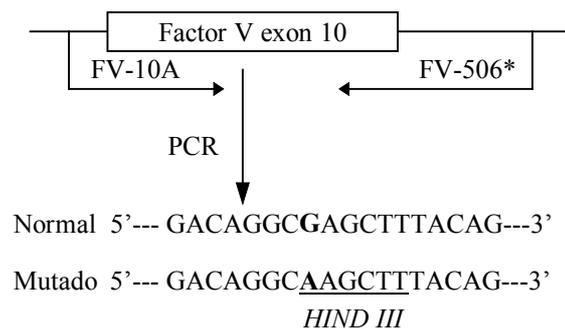


Figura 25 - Estratégia de identificação da mutação FV-Leiden. O exon 10 do factor V é amplificado por PCR. O fragmento amplificado contém um local de restrição para Hind III no caso de existir a mutação, mas não no caso do gene não mutado.

Tabela 4 - Resumo de polimorfismos identificados em genes responsáveis por doenças hemorrágicas e trombóticas hereditárias

Gene	Localização ^{a)}	Tipo de polimorfismo	Enzima de restrição	Nº Alelos	Heteroziguidade (%)
F VIII	Intron 13	Repetição CA	-	8	80
	Intron 18	RFLP	BclI	2	42
	Intron 19	RFLP	Hind III	2	42
	Intron 22	RFLP	XbaI	2	48
	Intron 22	Repetição CA	-	2	44
F IX	Intron 1	RFLP	DdeI	2	36
	Intron 3	RFLP	XmnI	2	41
	Intron	RFLP	TaqI	2	45
	Extr. 3'	RFLP	HhaI	2	48
vWF	Intron 2	RFLP	SmaI	2	45
	Intron 2	RFLP	HhaI	2	45
	Exon 14	RFLP	AccI	2	46
	Exon 18	RFLP	RsaI	2	45
	Intron 19	RFLP	MspI	2	44
	Exon 28	RFLP	HphI	2	50
	Exon 28	RFLP	BsteII	2	46
	Intron 40	Repetição(TCTA) _n	-	8	75
PC	7 Kb 5'	RFLP	MspI	2	42
	Exon 1	RFLP	RsaI	2	48
	Exon 6	RFLP	XbaI	2	48
	Exon 8	Sequência	-	2	45
PS	Exon 15	RFLP	BstXI	2	49
ATIII	Extr. 5'	Distância	-	2	37
	Exon 4	RFLP	PstI	2	37
		RFLP	DdeI	2	50
		Repetição (ATT) _n	-	>10	28

a) Nos introns a numeração corresponde ao nucleótido, nos exons a numeração corresponde ao aminoácido; Extr.=extremidade

9.2.2 - Doença de von Willebrandt (Mutações no gene do vWF)

A doença de von Willebrand (VWD) é a mais comum doença hemorrágica no homem, sendo causada por uma deficiência qualitativa ou quantitativa no factor de von Willebrand n(VWF). Esta doença pode ser transmitida segundo padrões dominantes ou recessivos, de acordo com a mutação em causa.

A clonagem do gene do VWF (180 Kb, contendo 52 exons, e localizado em 12p12) permitiu a identificação molecular das mutações responsáveis pela WWD. Os estudos iniciais permitiram a identificação de deleções neste gene. Contudo, e devido à dimensão do gene do VWF, a maior parte dos estudos posteriores centraram-se em zonas do gene importantes para funções determinadas da proteína, nomeadamente a dimerização e processamento intracelular (exons 1-16), ligação ao factor VIII (exons 17-25), ligação a colagénio (exons 28-34). Neste sentido, a classificação funcional do VWF constitui um valioso auxiliar no estudo genético de cada doente, já que permite concentrar esforços numa área restrita de um gene excessivamente longo para ser estudado por inteiro.

9.2.3 - Trombose familiar (Mutações nos genes da Antitrombina III, Proteína C e Proteína S)

Existem essencialmente 2 mecanismos inibidores da actividade das proteínases de serina envolvidas na coagulação: antitrombina III (ATIII), e o sistema Proteína C (PC)- Proteína S (PS)- trombosudolína (TM). A falha dos mecanismos inibidores predispõe para a trombose, e esta predisposição pode ser hereditária, como ficou demonstrado pela primeira vez em 1965 para a antitrombina (Egeberg et al., 1965). Desde então ficou demonstrada uma forte relação entre deficiências da PC e PS e tromboembolismo venoso.

A identificação de deficiências hereditárias a nível molecular da deficiência em antitrombina III ocorreu em 1984. Desde então, os avanços na genética molecular dos inibidores permitiu a identificação de outros defeitos genéticos, permitindo assim estudar a transmissão nas famílias afectadas.

O gene da antitrombina III está localizado em 1q23-25, tendo 13.5 Kb e 7 exons. A sequenciação dos exons permitiu a identificação de um numero considerável de mutações pontuais originando deficiência tipo I. A mutação mais comum consiste numa alteração da *fase de leitura* (Shift mutation), a qual resulta na presença de um codão stop prematuro. Outras alterações frequentes envolvem a formação de codões stop directamente resultantes de uma mutação pontual.

A deficiência tipo II parece resultar em todos os casos uma mutação pontual ocasionando a alteração de um aminoácido. Parece existir uma forte correlação entre a localização do aminoácido substituído, e a alteração funcional observada: alteração dos aminoácidos 24,27 e 129 (aminoácidos positivamente carregados), ou dos aminoácidos 41 e 99 (aminoácidos electricamente neutros) origina uma redução na afinidade para a heparina (negativamente carregada); mutações nos aminoácidos 382-394 resultam na deficiente inibição da trombina; mutações nos aminoácidos 402-407 e 429 produzem mutantes com alterações múltiplas.

O gene da PC (uma glicoproteína dependente da vitamina K) foi localizado em 2q13-14. Trata-se de um gene com 12 Kb e 9 exons. A sequenciação deste gene permitiu a identificação de várias mutações associadas à deficiência de Proteína C tipo I com um predomínio de substituições de aminoácidos.

O gene da PS (uma proteina do plasma dependente da vitamina K) está localizado no cromossoma 3 (3p11.1-11.2, possui cerca de 80Kb de DNA e contem 15 exons. Na sua vizinhança encontra-se ainda um pseudo-gene com uma estrutura muito homóloga. Os estudos para a identificação de mutantes nesta proteína são ainda poucos, mas foram já encontrados vários mutantes, responsáveis pelas deficiências de PS observadas.

9.2.4 - Hemofilias (Mutações nos genes dos factores VIII e IX)

A hemofilia A é uma das doenças hemorrágicas mais frequentes, sendo causada por uma deficiência do factor VIII (FVIII) circulante, cujo gene está localizado no cromossoma X (Xq28). A doença é heterogénea, tanto a nível molecular, como a nível da severidade clínica. A dimensão do gene do FVIII impediu até à pouco tempo atrás a análise genética das mutações causando o gene, dificultando a

caracterização da transmissão familiar. O uso de polimorfismos intragénicos ou extragénicos permitiu no entanto recentemente iniciar uma caracterização precisa do cromossoma X dos familiares de indivíduos afectados. Esta estratégia, não pode no entanto ser universalmente empregue, já que nem sempre existem marcadores polimórficos associados à mutação em causa. Nestes casos, apenas a determinação da mutação específica permite estudar com clareza os indivíduos em causa. Neste intuito, um esforço considerável foi desenvolvido no sentido de empregar métodos de identificação de alterações genéticas pontuais, obviando à sequenciação de grandes extensões genómicas. Com este propósito, técnicas como “single strand conformation polymorphism” (SSCP), “denaturing gradient gel electrophoresis” (DGGE) e chemical cleavage of mismatch (CCM) foram utilizadas com sucesso na identificação de mutações. Estes estudos permitiram determinar que a maioria dos doentes com doença moderada ou suave possuem mutações tipo missense, enquanto a maioria dos doentes graves possuem mutações tipo “frameshift”. Em alguns casos de doença grave foi ainda possível identificar um defeito no *splicing* dos exons 22-23.

A hemofilia B é uma doença hereditária de transmissão recessiva, devida à total ausência, ou quantidades reduzidas de factor IX (FIX). O gene do FIX foi totalmente sequenciado, contendo 8 exons. Cerca de 40% dos aminoácidos do FIX parecem ser essenciais para a função da proteína, pelo que não é de admirar que a doença seja causada por um largo espectro de mutações. Assim, e uma vez que não é possível dirigir o estudo para uma região específica do gene, as técnicas de estudo que têm sido utilizadas para identificar as mutações causadoras da Hemofilia têm sido essencialmente as mesmas técnicas acima descritas.

9.3 - HEMOCROMATOSE

A hemocromatose hereditária (HH) é uma doença genética de transmissão autossómica recessiva. Contrariamente ao que se pensava, sabemos hoje que a hemocromatose não é uma doença rara, já que a sua frequência é comparável à da anemia da “Sickle cell”, fibrose cística ou distrofia muscular (Nichols & Bacon, 1989). Foi estimado que só nos Estados Unidos devem existir entre 600,000 e 1,000,000 de doentes, bem como cerca de 27,000,000 de indivíduos portadores do gene (Nichols & Bacon, 1989).

A HH é caracterizada por uma falha na regulação da absorção do ferro da dieta, mantendo-se a absorção mesmo na presença de altos níveis de ferro armazenado (Alper et al., 1951; Cox & Peters, 1978; Williams et al., 1986; Lynch et al., 1989; Whittaker et al., 1989). Os indivíduos com HH possuem depósitos de ferro superiores a 4 g, não sendo raros indivíduos com depósitos superiores a 20 g, contrastando assim com os habituais 500-1000 mg nos indivíduos normais (Nichols & Bacon, 1989). O quadro completo de HH envolve a deposição de ferro em vários órgãos, nomeadamente: fígado, resultando em cirrose hepática; coração com diminuição da função cardíaca e perturbação do ritmo; articulações com formação de poliartropatia; pele, com formação de pigmentação dérmica característica; e glândulas endócrinas, originando falhas endócrinas como diabetes e gonadopatias (Milder et al., 1980).

A hemocromatose hereditária, não sendo uma doença para a qual exista uma cura é no entanto facilmente tratável por flebotomias (Bomford & Williams., 1976; Niederau et al., 1985). O tratamento ainda que seja meramente correctivo, e por isso exija intervenção durante toda a vida do doente, é bastante eficaz, consistindo numa primeira fase em 1 a 2 flebotomias semanais (tratamento intensivo), a que se seguem, após a depleção dos depósitos de ferro, flebotomias mensais ou trimestrais (tratamento de manutenção). Uma vez que a sobrecarga de ferro é passível de correcção, a hemocromatose hereditária, como modelo de estudo da interacção entre o metabolismo do ferro e o sistema imunológico, tem a grande vantagem de nos fornecer dados sobre a direcção das interacções. Com efeito, se a sobrecarga de ferro exercer um efeito fisiologicamente relevante no sistema imunológico destes indivíduos, a remoção da sobrecarga deverá corrigir esse defeito. Inversamente, se o sistema imunológico tiver uma acção relevante na homeostase do ferro, então esta deverá preceder a sobrecarga de ferro, não sendo corrigida com a remoção de ferro.

A origem precisa do erro na regulação da absorção do ferro na HH constitui ainda hoje um mistério, tal como o é a identidade do gene ou genes responsáveis pela doença. No entanto, e dada a aparente falta de relação entre a doença e perturbações da função imunológica clássica (ocorrência de infecções, a já descrita estreita associação entre o fenótipo HLA-A3 e o caractere hemocromatose (Simon, 1975, 1977a, 1977b) tem vindo a ser interpretado pelos geneticistas como indicador da estreita ligação física entre o *locus* HLA-A e o gene da hemocromatose (Simon, 1977a).

9.3.1 - Estudos de marcadores genéticos no locus do HLA

Devido à estreita associação entre o *locus* HLA-A no cromossoma 6 (6p21.3) o gene da hemocromatose, nos últimos anos tem-se assistido a um crescente ritmo de estudo de marcadores polimórficos tipo microssatélite nesta região do genoma humano, procurando-se definir marcadores mais estreitamente ligados à hemocromatose que o próprio HLA-A3. No momento presente estão disponíveis vários microssatélites, cuja segregação familiar permite definir haplótipos estendidos, dando maior rigor que a determinação HLA à classificação dos familiares como homozigóticos ou heterozigóticos, bem como na determinação de indivíduos potencialmente portadores do gene em estudos de “screening” da população. Deve no entanto notar-se que para o estudo da população normal, os estudos genéticos na hemocromatose são muito limitados, servindo apenas de elemento de apoio aos dados bioquímicos mais relevantes como os níveis séricos de ferritina, de transferrina, de ferro, e a taxa de saturação da ferritina

9.3.2 - Estudos dos IRE e IRP

Os IRE (do inglês Iron Responsive Elements) são sequências existentes nos mRNA do receptor da transferrina, ferritina, aconitase mitocondrial e eALAS (5-aminolevulinato sintetase). Trata-se de sequências muito conservadas na evolução (95% entre o homem, o rato e a galinha no caso do receptor da transferrina), e que possivelmente permitem a formação de 1 estrutura no mRNA em forma de ansa (Kühn, 1994). Estas sequências são as principais responsáveis pela regulação pós-translacional dos genes a que pertencem, conferindo-lhes a capacidade de se modularem dependendo da concentração intracelular de ferro (Kühn, 1994). Esta

regulação é mediada por uma proteína citoplasmática denominada IRP (do Inglês Iron Responsive Protein), mas previamente também conhecida como IRF, IRE-BP e FRP (Müllner et al., 1989; Rouault et al., 1988; Walden et al., 1988).

O efeito regulador do IRE (e conseqüentemente da IRP) depende da localização do IRE no mRNA. No mRNA do receptor da transferrina, os 5 IREs estão presentes na região 3' não traduzida, pelo que a ligação da IRP aumenta a estabilidade do mRNA, prolongando o seu tempo de semi-vida (Kühn, 1994). No caso do mRNA dos genes da ferritina, da aconitase mitocondrial e da eALAS, o IRE encontra-se situado na região 5' não traduzida do mRNA, a curta distância da sequência CAP, pelo que a ligação da IRP bloqueia a iniciação da tradução pelos ribossomas (Kühn, 1994).

A IRP é uma proteína bi-funcional (Klausner et al., 1993; Hirling et al., 1994). Na presença de ferro, a sua capacidade de ligar aos IRE é mínima, o que pode derivar da incorporação de 1 grupo 4Fe-4S na proteína. Nestas condições, a IRP é capaz de funcionar como aconitase citoplasmática, catalisando a formação de isocitrato (Emery-Goodman et al., 1993). Pelo contrário, na ausência de ferro, a IRP liga-se com grande especificidade aos IRE, perdendo a actividade de aconitase.

O sistema imunológico parece utilizar este duplo papel do IRP para controlar a sua actividade, e assim controlar o metabolismo intracelular do ferro.

Experiências descritas em 1993 por Drappier et al., comprovam a existência de efeitos reguladores de citoquinas, nomeadamente do IFN- γ e TNF- α na indução da ligação de IRP a IRE em macrófagos em cultura devido à modulação da síntese de óxido nítrico proveniente da via da L-arginina (Drappier et al., 1993).

O óxido nítrico, mercê da sua reduzida dimensão tem acesso directo ao núcleo 4Fe-4S da IRP, onde induz a sua degradação (Kühn, 1994). Assim, a presença de óxido nítrico (regulada pelo IFN- γ e TNF- α) é um factor tendente a aumentar a "pool" de IRP capaz de ligar ao IRE, diminuindo a "pool" capaz de funcionar como aconitase (Kühn, 1994).

9.3.3 - Estudos Genéticos do Repertório da Célula T

Os resultados do grupo da Prof. Maria de Sousa, revelando a existencia de uma correlação entre a razão CD4/CD8 e os niveis de ferro dos doentes, levou este mesmo grupo a procurar anomalias no repertório das células T nesta patologia. Este estudo revelou a existência de uma correlação entre anomalias na representação do TCR V β 6.7 entre as células CD8+ e a severidade da expressão clínica da doença, nomeadamente o desenvolvimento de cirrose hepática. Estes resultados, sugerindo intervenções do sistema imune na regulação de sistemas internos, reforçam a visão crescente do sistema imune como um sistema regulador, não só das potencialmente nocivas interferências externas, mas também das frequentemente fatais perturbações dos delicados equilíbrios homeostáticos internos do organismo. Esta interpretação faz com que estes resultados tenham implicações clínicas de grande relevo não só na hemocromatose hereditária mas também em todas as formas de sobrecarga de ferro (Porto, 1993; Cabeda, 1995).

LEITURA COMPLEMENTAR

Anemias Não Esferocíticas Congénitas

- Beutler E. *N.Engl.J.Med.* 324:169-174, 1991
- Beutler E. *Am.J.Hematol.* 42:53-58, 1993
- Vulliamy T., Beutler E. & Luzatto L. *Hum.Mutat.* 2:159-167, 1993
- Cell 67:641-4, 1991
- PNAS 89: 4028, 1992
- Cotter PD., Ruckmagel DL. & Bichop DF. *Blood* 84:3915-24, 1994
- Baronciani L.,Magalhães IQ. *Blood Cells, Molecules and Diseases* 21:49-55, 1995
- Miwa S., Kanno H., Fujii H. *Am.J.Hematol.* 42:31-35, 1993
- Noguchi T., Yamada K., Inoue H. et al. *J.Biol.Chem.* 262:14366-77, 1987
- Noguchi T., Inoue H., Tanaka T. *J.Biol.Chem.* 261: 13807-12, 1986
- Baronciani L., Beutler E. PNAS 90:4324-7, 1993
- Beutler E., Kuhl W., Gelbart T. et al. *J.Biol.Chem.* 266:4145-50, 1991
- Lenzner C., Jacobasch G., Reis A. et al. *Hum.Mol.Genet.* 3:523, 1994
- Kanno H., Fujii H., Hirono A. et al. *Blood* 79: 1347-50, 1992
- Kanno H., Wei DC., Chan LC. et al. PNAS 88:8218-21, 1991
- Kanno H. *Blood* 84:13a, 1994
- Neubauer B., Lakomek M., Winkler H et al. *Blood* 77: 1871-5, 1991
- Kanno H. Fujii H., Miwa S. *Blood* 81:2439-41, 1993
- Kanno H., Ballas SK., Miwa S. et al. *Blood* 83:2311-6, 1994
- Kanno H., Wei DCC., Miwa S. et al. *Blood* 82(suppl 1): 97a, 1993
- Lenzner C., Nürnberg P., Thiele BJ. et al. *Blood* 83:2817-22, 1994
- Bianchi P., Terragna C., Zappa M. et al. *Blood* 84: 14a, 1994
- Kanno H., Wei DC., Chan LC. et al. *Blood* 84: 3505-9, 1994
- Weiming X., Westwood B., Bartsocas CS. et al. *Blood* 85:257-63, 1995(review)
- Beutler E. *Blood* 84: 3613-36, 1994 (review)

Talassémias

- Higgs DR., *Baillière's Clinical Haematol.* 6: 117-50, 1993
- Grosveld F.Dillon N., Higgs D. *Baillière's Clinical Haematol.* 6:31-56, 1993
- Molina MA., Romero MJ., Abril E. et al. *Sangre* 39:253-256, 1994
- Lindeman R.Hu SP,Volpato F. et al. *Brit.J.Haematol.* 78:100-4,1991
- Spiegelberg R., Aulehla-Scholz C., Erlich H et al. *Blood* 73:1695-98, 1989
- Hazazian H.H & Boehm CD. *Blood* 72: 1107-16, 1988

Esfereocitose e Eliptocitose Hereditária

- Eber SW., Lux ML., Gonzales JM. et al. *Blood* 82:(s1), 308^a (abstract), 1993
- Jarolim P., Rubin HL., Liu SC. et al. *J.Clin.Invest.* 93: 121-130, 1994
- Alloisio N., Texier P., Forissier A. et al., *Blood* 82:(s1),4a(abstract), 1993
- Bouhassira EE., Schwatz RS., Yawata Y. et al., *Blood* 79: 1846-54, 1992
- Palek J. & Sahr KE. *Blood* 80: 308-30, 1992 (review)
- Delanay J & Dhemy D. *Semin.Hematol.* 30:21-33, 1993 (review)
- Wilmotte R., Maréchal J., Morlé L. et al. *J.Clin.Invest.* 91:2091-6, 1993
- Conboy JG., *Sem.Hematol.* 30: 58-73, 1993 (review)
- Dalla Venezia N., Gilsanz F. Alloisio N. et al. *J.Clin.Invest.*90:1713-17,1992
- Cartron JP, le-van-Kin C. & Colin Y. *Semin.Hematol.*30:152-68, 1993 (review)

von Wellebrandt Disease

- Eikenboom J.C.J., Reitsma PH., Peerinck K.M.J. et al. *Lancet* 341:982-6, 1993
- Ginsburg D., Sadler JE. *Thrombosis and Haemostasis* 69:177-84, 1993
- Sadler JE. *Thrombosis and Haemostasis* 71:520-5, 1994
- Zhang ZP., Falk G., Blombachm et al. *Human Molec. Genet.*1:767-8, 1992
- Sadler JE., Ginsburg D. *Thrombosis and Haemostasis* 69:185-91, 1993
- Inbal A., Kornbrot N., et al. *Thrombosis and Haemostasis* 73:324-9

Hemofiliias

- Tizzano EF., Domènech M. Baiget M. *Thrombosis and Haemostasis* 73:6-9, 1995
- de-la-Salle C. et al. *Thrombosis and Haemostasis* 70:370-376, 1993
- Nielsen LR., Scheibel E. et al. *Thrombosis and Haemostasis* 73:774-8, 1995
- Saad S., Rowley G. et al. *Thrombosis and Haemostasis* 71:563-70, 1994

FV-Leiden

- Dahlback B., Carlsson M. & Svensson P. PNAS 90:1004-8, 1993
- Heijboer H,Brandjes DPM, Buller HR et al. *N.Engl.J.Med.* 323:1512-6,1990
- Halbmayer W., Haushofer A., Schon R. et al. *Blood Coag.Fibrinolysis* 5:51-7,1994
- Griffin JH., Evatt B. Weideman C. et al. *Blood* 82:1989-93,1993
- Svensson PJ., Dahlback B. *N.Engl.J.Med.* 330:517-522, 1994
- Dahlback B & Hidelbrand B. PNAS 91:1396-1400, 1994
- Bertine R., Koeleman B., Koster T. et al. *Nature* 369:64-6, 1994
- Koster T. Rosendal F., de-Rond H. et al. *Lancet* 342:1503-6, 1993
- Ronde H., Bertina RM., *Thrombosis and hemostasis* 72:880-6, 1994
- Ireland H,Bayston T,Thompson E et al. *Thrombosis and hemostasis* 73:727-38,1995
- Gandrille S., Alhenc-Gelas M. & Aiach M. *Blood Coag. and Fibrinolysis* 6:245-8,1995
- Dahlback B. *Thrombosis and Hemostasis* 73:739-42,1995
- Zöller B., Svensson PJ., He X. et al. *J.Clin.Invest.* 94:2521-4,1994
- ,1995
- Peake IR., Bowen D., Bignell P. et al *Blood* 76:555-61, 1990

Trombose Familiar

- Lane DA., Olds RJ. et al. *Thrombosis and Haemostasis* 70:361-9, 1993
- Olds R., Lane DA., Chwdhury V. et al. *Biochemistry* 32:4216-24, 1993
- Lane DA., Olds R., Conard J. et al. *J.Clin.Invest.* 90:2422-33, 1992
- Reitsma PH., Poort SR., Allaart CF. et al. *Blood* 78:890-4, 1991
- van Amstel HKP., Teisma PH. et al. *Biochemistry* 29: 7853-61, 1990
- Sacchi E., Pinotti M., et al. *Thrombosis and Haemostasis* 73:746-9,1995
- Reitsma PH., Poort SR. et al. *Thrombosis and Haemostasis* 69:77-84, 1993
- Lind B., Schwartz M., Thorsen S. *Thrombosis and Haemostasis* 73:186-93, 1995
- Soria JM., Brito D., Barceló J. et al. *Thrombosis and Haemostasis* 72:65-9,1994
- Reitsma PH., Bernardi F. et al. *Thrombosis and Haemostasis* 73:876-99, 1995
- Lind B., van-Soling WW., Schwartz M. *Blood* 82:2423-32, 1993

Hemocromatose

- Nichols GM & Bacon BR.. *Am.J. Gastroenter.*1989;84:851-62
- Alper T, Savage DV, Bothwell TH. *J.Lab.Clin.Med.* 1951;37:665-75
- Cox TM & Peters TJ. *Lancet* 1978;1:123-4
- Williams R, Manenti F, Williams HS, Pitcher CS. *Br. J. Med.* 1986;2:78-81
- Lynch SR, Skikne BS, Cook JD. *Blood* 1989;74:2187-2193
- Whittaker P, Skikne BS, Covell AM. *J.Clin.Invest.* 1989; 83:261-7
- Milder MS, Cook JD, Stray S, Finch CA. *Medicine (Baltimore)* 1980;59:34-49
- Bomford A., & Williams R. *Q.J.Med.* 1976; 45:611-623
- Niederau C., Fischer R., et al. *N-Engl.J.Med.* 1985; 313:1256-62
- Simon M, Pawlotsky Y, Bourel M, et al. *Nouv. Presse Med.* 1975; 4:1432
- Simon M, Bourel M, Genetet B. et al. *N.Engl.J.Med.* 1977a; 297:1017-1021
- Simon MJ, Alerandre JL, Bourel M. *Clin. Genet.* 1977b; 11:327-341
- Kühn LC. *Baillière's Clinical Haematology* 1994; 7: 763-785.
- Müllner EW, Neupert B & Kühn LC. *Cell* 1989; 58:373-382
- Rouault TA, Hentze MW, Haile DJ et al. PNAS 1988; 86:5768-5772
- Walden WE, Daniels-McQueen S, et al. PNAS 1988; 85:9503-9507
- Klausner RD, Rouault TA & Harford JB. *Cell.* 1993;72:19-28
- Hirling H, Henderson BR & Kühn LC. *EMBO J.*1994; 13:453-461
- Emery-Goodman A, Hirling H. et al. *Nucleic Acids Research* 1993; 21: 1457-1461
- Drappier JC, Hirling H., Wietzerbin J., et al. *EMBO J.*1993; 12: 3643-3649
- Cabeda JMC. Tese de Doutorado. ICBAS, 1995
- Porto G. Tese de doutoramento, 1993

PARTE

3

MANUAL PRÁTICO

10 - HEMOPATIAS MALIGNAS

10.1 - Anomalias de Cariótipo. t(9;22) (q34;q11): BCR/ABL

Protocolo

BCR/Abl

a) Preparar a master mix da reacção de RT

componente	Volume (µl)			concentração final
	# de tubos→	1	5	
dd-H ₂ O		19	95	190
10x RT buffer		5	25	50
dNTP (10 mM cada)		5	25	50
random hexamers		5	25	50
Rnasin		1	5	10
RT enzyme		5	25	50
amostra		10	50	100
Volume total		50	250	500

Nota: A amostra de RNA é preparada como nos procedimentos gerais, depois de a master mix estar pronta

b) Incubar no termociclador (2h - 37°C; hold at 4°C)

c) preparar a master mix da reacção de PCR

componente	Volume (µl)			concentração final
	# de tubos→	1	5	
dd-H ₂ O		26.8	134	268
dNTP (10mM)		1	5	10
primer mix*		5.2	26	52
10x tampão		5	25	50
Genzyme		2	10	20
cDNA		10	50	100
Volume Total		50	250	500

* a primer mix é feita dissolvendo as 5AU em 1 ml de água, e misturando os primers em proporções iguais

d) incubar no termociclador:

35x (30"-96°C; 30"-64°C; 30"-72°C)

1 x (10"-72°C)

e) Fazer gel de 2% Nusieve agarose

f) carregar 10 µl de PCR product + 2µ loading buffer e correr ≈1h a 100V

11 - DETECÇÃO E CARACTERIZAÇÃO DE VIRUS

11.1 - Caracterização genotípica do HCV

Protocolo

- a) Fazer a master mix da reacção de RT (a master mix pode ser congelada sem a RNASin e o soro):

Genótipo HCV

componente	Volume (µl)				concentração final
# de tubos→	1	5	10	25	
dd-H ₂ O	4	20	40	100	
tampão RT (5x)	5	25	50	125	1x
dNTP (2.5 mM)	4	20	40	100	0.4 mM
Primer 1AS (10 pmol/µl)	5	25	50	125	2pmol/µl
RNASin	1	5	10	25	
NP40 (10%)	1	5	10	25	0.4%
Soro	5	25	50	125	
Volume Total	25	125	250	625	

- b) incubar no termociclador (progama 10):

1x (5' - 42°C; 30" - 92°C; 5' - 4°C)

- c) centrifugar

- d) adicionar 1 µl de RT

- e) Incubar no termociclador (programa 11):

1x (1h - 42°C; 10'-95°C; soak 4°C)

- f) centrifugar e colocar em gelo

- g) Fazer a mistura do 1º PCR:

componente	Volume (µl)				concentração final
# de tubos→	1	5	10	25	
dd-H ₂ O	41.5	207.5	415	1037.5	
10x tampão	10	50	100	250	1x
MgCl ₂ (15 mM)	10	50	100	250	1.5 mM
dNTP (2.5mM)	8	40	80	200	0.2 mM
primer 2AS (50 pmol/µl)	5	25	50	125	25 pmol/µl
Taq pol (2.5 U/µl)	0.5	2.5	5	12.5	1.25 U/100 µl
cDNA (obtido em f)	25	125	250	625	
Volume Total	100	500	1000	2500	

- h) incubar no termociclador (programa 16):

1x (5'-94°C)

34x (1' - 94°C; 1'-55°C; 2'-72°C)

1x (1' - 94°C; 1'-55°C; 7'-72°C)

soak at 4°C

- i) centrifugar e guardar no gelo

j) Fazer a master mix do 2º PCR:

componente	Volume (µl)				concentração final	
	# de tubos→	1	5	10		25
dd-H ₂ O		51.5	257.5	515	1287.5	
10x tampão		10	50	100	250	1x
MgCl ₂ (15 mM)		10	50	100	250	1.5 mM
dNTP (2.5mM)		8	40	80	200	0.2
primer 3S		5	25	50	125	
primer 4S		5	25	50	125	
primer 5AS		5	25	50	125	
Taq pol (2.5 U/µl)		0.5	2.5	5	12.5	1.25 U/100 µl
cDNA (obtido em f)		5	25	50	125	
Volume Total		100	500	1000	2500	

k) incubar no termociclador (método 19):

1x (5'-94°C)

24x (1' - 94°C; 1'-55°C; 2'-72°C)

1x (1'-94°C; 1'-55°C; 7'-72°C)

1x (5'-4°C; 10'-100°C; soak - 4°C) (desnaturação do produto do PCR)

l) retirar os tubos do termociclador minutos antes de serem necessários e centrifugar

m) distribuir 100µl de cada subtipo de sonda nas tiras e deixar O.N. a 4°C)

n) lavar. Adicionar 100µl de sol. de hibridação+10µl de produto de PCR

o) incubar 1h a 55°C.

p) preparar o a-DS-DNA (300µl anticorpo+14.7 ml diluente)

q) lavar. adicionar 100 µl d a-DS-DNA. incubar 1h a RT

r) preparar o enzime tracer (300µl tracer+14.7ml diluente)

s) lavar. adicionar 100µl de enzime tracer. incubatr 1h a RT

t) preparar o chromogen/substrate (9 ml de cada reagente). manter no escuro

u) lavar. adicionar 100µl de chromogen/substrate. incubar 30' a RT

v) adicionar 200µl de blocking reagent na mesma ordem e *velocidade* que o chromogen/substrate

w) ler a densidade optica a 430 e 630nm (resultado = OD₄₅₀-OD₆₃₀). Valores superiores a 0.15 indicam hibridação

12 - DOENÇAS GENÉTICAS

12.1 - Resistência à proteína C Activada (Mutaç o FV-Leiden)

Protocolo

Factor V - Leiden

a) preparar a master mix da reac o de PCR

componente	Volume (μ l)			concentra�o final
# de tubos→	1	5	10	
dd-H ₂ O	26.8	134	268	
dNTP (10mM)	1	5	10	0.25 mM
primer mix*	5.2	26	52	
10x tamp�o	5	25	50	1x
Genzyme	2	10	20	
DNA	10	50	100	
Volume Total	50	250	500	

* a primer mix   feita dissolvendo as 5AU em 1 ml de  gua, e misturando os primers em propor es iguais

Nota: A amostra de DNA   preparada como nos procedimentos gerais, depois de a master mix estar pronta

d) incubar no termociclador (m todo 55):

1x (5'-94 C)

30x (60''-93 C; 30''-62 C; 60''-72 C)

1 x (10'-72 C)

e) fazer a master mix da reac o de RE

componente	Volume (μ l)			concentra�o final
# de tubos→	1	5	10	
dd-H ₂ O	75	375	750	
Tamp�o NEBuffer2	10	50	100	1x
Mnl I	5	25	50	25U/100 μ l
Produto do PCR	10	50	100	
Volume Total	100	500	1000	

f) incubar pelo menos 2h a 37 C

e) Fazer gel de 2% Nusieve agarose

f) carregar 10 μ l de produto de PCR + 2 μ loading buffer e correr \approx 1h a 100V

12.2 - Caracterização da mutação de β -globina na β -talassémia

Protocolo

Genótipo β -talassémia

a) preparar a master mix da reacção de PCR

componente	Volume (μ l)		
	# de tubos→	1	5
dd-H ₂ O	61	305	610
dNTP (10mM)	2	10	20
primer (CD7)	6	30	60
primer (SR10)	6	30	60
10x tampão	10	50	100
MgCl ₂	3	15	30
Genzyme	2	10	20
DNA	10	5 x 10	100
Volume Total	100	500	1000

Nota: A amostra de DNA é preparada como nos procedimentos gerais, depois de a master mix estar pronta

d) incubar no termociclador:

1x (5'-94°C)

35x (30"-94°C; 30"-57°C; 30"-72°C)

1 x (10'-72°C)

1 x (hold at 4°C)

e) fazer a master mix das reacções de RE

componente	Volume (μ l) (por reacção)		
	BSA BI	BSU36I	AVA II
dd-H ₂ O	33.75	27.5	32.5
Tampão NEBuffer 2	5	-	-
Tampão NEBuffer 3	-	5	-
Tampão NEBuffer 4	-	-	5
Enzima (como indicado)	1.25	2.5	2.5
BSA	-	5	-
Produto do PCR	10	10	10
Volume Total	50	50	50

f) incubar pelo menos 2h (a reacção com BSA BI a 60°C, as restantes a 37°C)

e) Fazer gel de 2% Nusieve agarose

f) carregar 10 μ l de produto de PCR + 2 μ loading buffer e correr \approx 1h a 100V