

# BIOLOGIA MOLECULAR

Dr<sup>a</sup> Angela Adamski da Silva Reis

## **Apresentação**

Prezado(a) Cursista, bem vindo à disciplina de Biologia Molecular que compõe o quarto período do Curso de Licenciatura em Ciências Biológicas PARFOR-EAD-ICB-UFG. O progresso no conhecimento em nível de material genético propiciou a junção das disciplinas de Bioquímica e Genética, gerando a Biologia Molecular, a qual é o estudo em nível molecular da estrutura e função do material genético, bem como a compreensão dos produtos gênicos. A Biologia Molecular estuda a associação entre DNA, RNA e proteínas gerando conhecimento em nível de sistemas celulares procarióticos e eucarióticos. Abrange diversas áreas com aplicações das técnicas moleculares no diagnóstico de doença genéticas e infecciosas, estudos de evolução, medicina forense, dentre outras. O conteúdo que iremos utilizar neste roteiro irá introduzir o assunto no âmbito da Biologia Molecular, não substituindo os livros texto apresentados no conteúdo programático da disciplina. O material que utilizaremos é um roteiro para que, de forma didática, possamos compreender a Biologia Molecular aplicada à formação de professores.

## **1. A natureza molecular e a organização do material genético**

### **1.1 A natureza molecular**

A inter-relação entre as moléculas de informação, o ácido desoxirribonucléico (DNA) e o ácido ribonucléico (RNA) promovem a expressão gênica, a qual na

maior parte está envolvida com a síntese de polipeptídeos. Assim, as proteínas são o principal produto funcional da matriz de DNA, desempenhando várias funções celulares como enzimas, receptores, proteínas de reserva ou transporte, fatores de transcrição, moléculas sinalizadoras, hormônios, etc.

O DNA é uma dupla hélice bifilamentar que encontra-se nas células eucarióticas compactado nos cromossomos no núcleo celular, nas mitocôndrias e também nos cloroplastos das células vegetais. O açúcar no DNA é a desoxirribose, um açúcar de 5 carbonos. Os resíduos de açúcar ligam-se uns aos outros por ligações fosfodiéster.

**A molécula de DNA é composta por um polímero longo linear formado por açúcar, grupamento fosfato e bases nitrogenadas.** Assim, a unidade básica da molécula de DNA é composta por nucleotídeo (desoxirribose, fosfato e base nitrogenada). E um nucleosídeo é formado por desoxirribose e base nitrogenada.

As bases nitrogenadas ligam-se covalentemente ao átomo do carbono 1 de cada resíduo de açúcar e possuem anéis heterocíclicos de átomos de carbono e nitrogênio, os quais são hidrofóbicos e ficam orientados para o interior da dupla hélice. As bases são classificadas em púricas (Adenina e Guanina), que possuem dois anéis heterocíclicos ou pirimídicas (Citosina e Timina), as quais possuem apenas um anel heterocíclico (Figura 1).

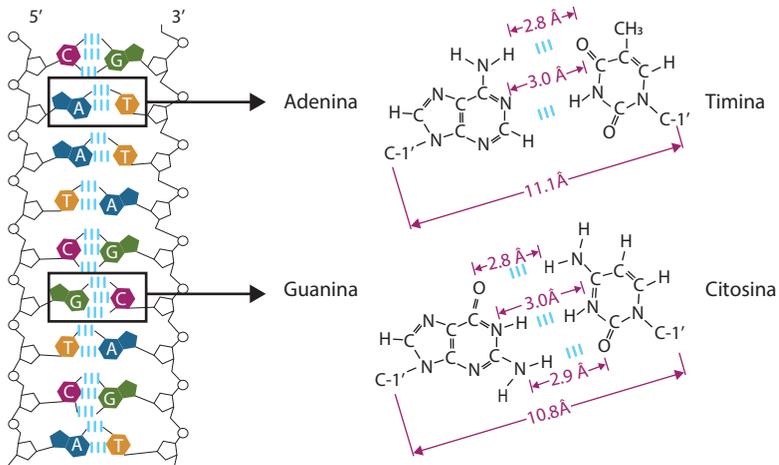


Figura 1. Estrutura do DNA: dupla hélice antiparalela bifilamentar

A ligação entre as bases nitrogenadas ocorre por pontes de hidrogênio, onde uma base púrica sempre liga-se a uma base pirimídica. Nesta ligação Adenina (A) pareia-se com Timina (T) através de duas pontes de hidrogênio e Guanina (G) pareia-se com

Citosina (C) através de três pontes de hidrogênio, sendo esta mais forte do que AT. A formação dessas ligações promove efeitos hidrofóbicos no interior da dupla hélice e hidrofílicos na região das curvaturas da dupla hélice, promovendo assim o pareamento coordenado entre as bases, as quais proporcionam a estabilidade em função da configuração espacial e eletrônica (Figura 1.1). O empilhamento das bases ocorre em função das forças de Van der Waals entre os anéis aromáticos de bases adjacentes.

A **Regra de Chargaff** demonstra que a relação molar entre A/T seja igual a 1,0, o mesmo ocorrendo com relação C/G, embora as concentrações molares entre A/T e C/G variem em função da sequência de nucleotídeos. Essa propriedade garante a replicação da molécula de DNA de forma precisa, como também na transmissão das informações genéticas

A estrutura do DNA é uma dupla hélice antiparalela, sendo que cada uma das fitas de DNA é nomeada de acordo com a posição do grupamento fosfato no carbono da pentose (desoxirribose). Assim, o grupamento fosfato presente no carbono 5 do nucleotídeo forma a fita 5' (sense -5'>3') e o grupamento hidroxila presente no carbono 3 do nucleotídeo forma a fita 3' (anti-sense -3'>5') [Figuras 1.1. e 1.2.). As ligações glicosídicas no DNA, entre as desoxirriboses e as bases nitrogenadas, não estão diretamente opostas na dupla hélice, e assim geram duas curvaturas desiguais em seu contorno, uma maior e outra menor (Figura 2.).

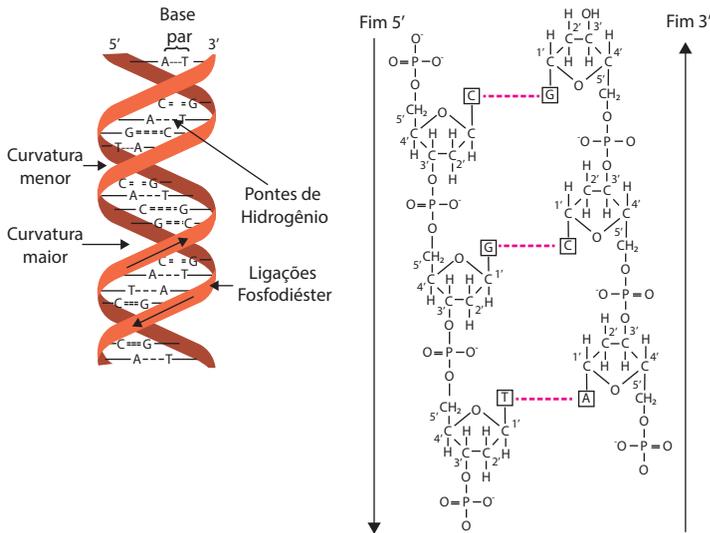


Figura 2. A estrutura do DNA é uma hélice antiparalela

Os fenômenos físicos que ocorrem no DNA são fundamentais para os processos de replicação, transcrição e recombinação. Leia sobre renaturação e desnaturação, bem como temperatura de melting.

A composição da molécula de RNA é semelhante ao DNA, com a diferença no resíduo de açúcar, a ribose, que em vez de timina, temos a base nitrogenada uracila (U). Existem semelhanças entre DNA e RNA quanto a estrutura, ambos são polímeros lineares de subunidades ligadas entre si por pontes fosfodiéster 5'->3'. Entretanto, na molécula de RNA, a Timina (T) é substituída por U, as outras três bases A, C e G também estão presentes. O RNA normalmente é de fita simples, embora os pareamentos entre as bases nitrogenadas possa ocorrer, formando estruturas secundárias. Diferentemente do DNA, os tipos de RNA possuem funções específicas nas células, sendo classificados de acordo com a localização e função celular (RNA mensageiro, RNA transportador e RNA ribossomal). Outra diferença entre DNA e RNA, é que o RNA é mais reativo que o DNA devido a presença da hidroxila livre no carbono 2 da ribose.

Leia sobre a descoberta da estrutura do DNA em 1950 e como os experimentos de Rosalind Franklin foram importantes para a elucidação desta molécula.

## 1.2 A anatomia do genoma nos organismos celulares (organização gênica)

Os organismos celulares procarióticos e eucarióticos apresentam diferença na organização de seus genomas. O genoma de ambos os organismos possuem a função básica de servir como um repositório da informação codificada. O genoma de procariotos são relativamente mais simples que genomas eucarióticos. O genoma de um organismo procarioto apresenta um cromossomo circular que consiste de uma molécula de DNA de fita dupla covalentemente fechada.

O **genoma dos procariotos** apresentam características peculiares. Praticamente todo o genoma possui função regulatória ou codificante (genes estruturais) para a síntese de proteínas com pequeno espaço entre os genes. Os genes estão agrupados em unidades transcricionais policistrônicas, ou seja os genes estão localizados adjacentes (Operon), na região onde um gene termina, outro inicia. Assim, todos os genes são expressos como uma única unidade.

Esse tipo de organização gênica é comum nos procariotos. Como exemplo podemos citar o **Operon lactose** (*Operon lac*), o qual contém três genes estruturais, *Z*, *Y* e *A* que codificam as enzimas  $\beta$ -galactosidase, lactose-permease e tiogalactosídeo-transacetilase envolvidas na conversão do dissacarídeo lactose nas suas unidades de monossacarídeo - glicose e galactose. É importante ressaltar que estes organismos possuem também unidades genéticas acessórias extra cromossomais, como os plasmídeos, bacteriófagos e elementos genéticos transponíveis.

O **genoma** da maioria dos eucariotos são maiores e mais complexos do que os genomas de procariotos. No entanto, o tamanho do genoma não está relacionado à complexidade genética. Estes genomas que tem basicamente em comum o fato de se encontrarem envolvidos por uma membrana nuclear, apresentam uma considerável variação individual de organismo para organismo. Apesar disso, apresentam alguns aspectos comuns na organização gênica. O genoma apresenta genes com sequências codificantes (**éxons**) que são separados por sequências não-codificadores (**íntrons**). O gene inteiro é transcrito para produzir uma longa molécula de RNA mensageiro (RNAm) e os íntrons são então removidos por um mecanismo denominado *splicing* de forma que somente os éxons são incluídos no RNAm para a síntese de proteínas.

Os íntrons são responsáveis por uma fração substancial no tamanho do genoma dos eucariotos superiores, nos eucariotos complexos, os íntrons são responsáveis por cerca de dez vezes mais DNA do que os éxons. Embora a maioria dos íntrons não tenha função conhecida, sugere-se que alguns codificam RNAs funcionais. Existem evidências de que os íntrons podem ter auxiliado a acelerar a evolução, facilitando a recombinação entre regiões codificadoras de proteínas de diferentes genes - um processo conhecido como “embaralhamento de éxons”, mas não há relação entre a reunião inicial das sequências codificadoras de proteínas antes da divergência evolutiva entre as células procarióticas e eucarióticas.

Pesquise sobre a expressão de RNAs não codificadores intrônicos no câncer.

Muitos genes eucarióticos estão presentes em cópias múltiplas, denominadas **família gênicas**. Alguns membros dessas famílias funcionam em diferentes tecidos ou em diferentes estágios do desenvolvimento. Outros membros das famílias gênicas (pseudogenes) foram inativados por mutações e não mais representam genes funcionais. Sabe-se que aproximadamente 40% do DNA de mamíferos é composto de sequências de DNA altamente repetitivo, algumas das quais estão presentes em 105 a 106 cópias por genoma.

O **DNA repetitivo** não é codificante e é composto por sequências curtas e similares, que encontram-se em tandem (agrupadas) ou dispersas pelo genoma. Aquelas

em tandem compõem DNA satélite, microsatélite e minissatélite. Aquelas sequências espalhadas pelo DNA genômico são classificadas como elementos dispersos curtos (SINE) do inglês, *short interspersed elements* ou elementos dispersos longos (LINE) do inglês, *Long interspersed elements*.

Os genomas da maioria dos eucariotos são muito mais complexos do que os dos procariotos, como o DNA das células eucarióticas também apresentam-se de forma diferente do DNA nas células procarióticas. Nos eucariotos, o genoma é composto por múltiplos cromossomos, cada um contendo molécula de DNA linear compactada. Embora o número e o tamanho dos cromossomos sejam muito variáveis entre as diferentes espécies, sua estrutura básica é a mesma para todos os eucariotos.

O DNA das células eucarióticas está firmemente ligado a pequenas proteínas básicas (**histonas**) que empacotam o DNA de modo ordenado no núcleo celular (Figura 3.). Essa propriedade de compactação do DNA via histonas torna-se importante em função do conteúdo de DNA do eucariotos. O tamanho total do DNA estendido em uma célula humana, é de quase 2 metros, mas esse DNA deve encaixar-se em um núcleo com um diâmetro de 5 a 10 $\mu$ m. Nos procariotos o mecanismo de compactação do DNA é distinto dos eucariotos, mas ainda não é bem compreendido.

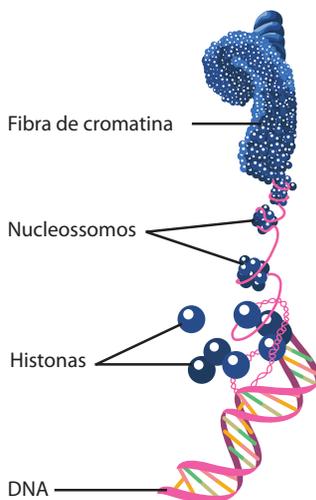


Figura 3. Empacotamento do DNA nos cromossomos exige hierarquias múltiplas de histonas

Vimos que nas células eucarióticas, o DNA encontra-se associado a uma série de proteínas nucleares específicas (histonas) que formam um complexo denominado **cromatina** (Figura 1.3.). As histonas contêm uma grande proporção de aminoácidos básicos (arginina e lisina), os quais facilitam a ligação à molécula de DNA que é carregada negativamente. Existem cinco tipos de histonas H1, H2A, H2B, H3 e H4. A cromatina além das histonas, possui proteínas não histônicas envolvidas na compactação do DNA

A unidade estrutural básica da cromatina é denominada **nucleossomo** (Figura 4.), na qual o DNA está enrolado em torno de histonas, formando uma partícula central que é selada pela histona H1 que contém de 150 a 250 pares de bases associado a um octâmero de proteínas que é constituído por duas cópias das histonas centrais: H2A, H2B, H3 e H4 (Figura 5.). O empacotamento do DNA em nucleossomos produz uma fibra de cromatina de aproximadamente 10 nm de diâmetro. A cromatina é condensada novamente, por enrolamento e uma fibra de 30 nm, contendo seis nucleossomos por volta.

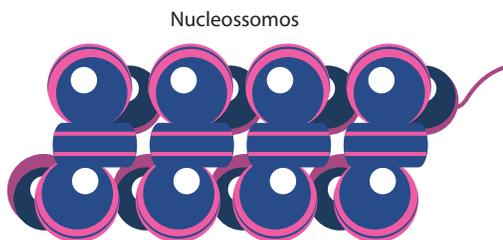


Figura 4. Estrutura dos nucleossomos

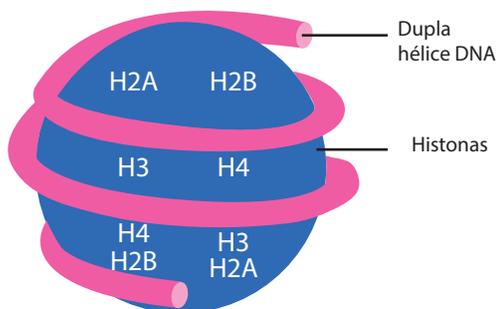


Figura 5. Histonas associadas a dupla hélice

Acesse: <http://www.youtube.com/watch?v=gbSIBhFwQ4s> para a visualizar a animação da compactação do material genético

O grau de condensação da cromatina varia durante o ciclo celular. Em células em interfase a cromatina é denominada eucromatina, e está relativamente descondensada e distribuída por todo o núcleo. Em contraste, cerca de 10% da cromatina interfásica é denominada heterocromatina, forma altamente condensada que é semelhante à cromatina das células em mitose. A heterocromatina não é transcrita e contém sequências de DNA

altamente repetitivas, tais como as sequências presentes nos centrômeros e telômeros dos cromossomos.

A medida que a célula entra em mitose, seus cromossomos tornam-se extremamente condensados para que possam ser distribuídos para as células filhas. Acredita-se que as alças das fibras de cromatina dobram-se sobre si mesmas para formar o cromossomo metafásico compacto das células mitóticas, nas quais o DNA está compactado quase 10.000 vezes. Na cromatina condensada não há síntese de RNA, assim, a transcrição cessa durante a mitose.

Os cromossomos só podem ser visualizados nas células em divisão. A introdução de técnicas de bandeamento permitiu a identificação individual de cada cromossomo. O estudo e identificação dos cromossomos humanos foi uniformizado, em uma conferência em Denver (EUA) em 1960, estabelecendo uma classificação adotada mundialmente. Como resultado desta convenção, os cromossomos humanos são hoje classificados e ordenados em sete grupos, de acordo com seu tamanho e morfologia. Os grupos são designados por letras de A a G e os cromossomos, numerados de 1 a 22 (os autossomos) por ordem decrescente de tamanho, e os cromossomos sexuais designados por X e Y. Todos os cromossomos são distinguidos individualmente através de um padrão de bandas claras e escuras, além de seu tamanho e da localização do centrômero. Podem ser recortados manualmente a partir de fotomicrografias ou – com o uso de câmeras digitais – selecionados através de programas de edição de imagens, sendo então agrupados aos pares para formar o que chamamos de cariótipo. Os tipos bandeamentos podem revelar bandas e sub-bandas em posições específicas no cromossomo. A Figura 6. demonstra um exemplo para nomenclatura dos cromossomos.

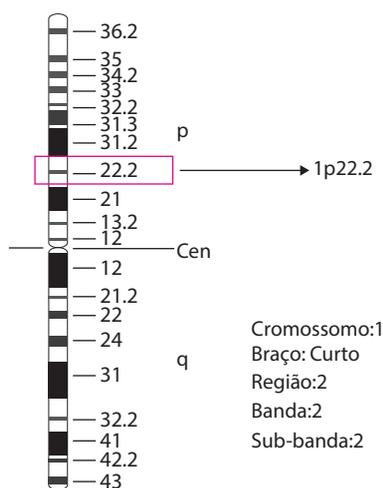


Figura 6. Conferência de *Denver* para nomenclatura dos cromossomos

## 2. A replicação do DNA

### 2.1 A replicação do DNA associada ao ciclo celular

O ciclo de divisão da maioria das células consiste em quatro processos coordenados: crescimento celular, replicação do DNA, distribuição dos cromossomos duplicados às células-filhas e divisão celular. Nas bactérias, o crescimento celular e a replicação do DNA se dão durante a maior parte do ciclo celular, e os cromossomos duplicados são distribuídos às células-filhas associados a membrana plasmática.

Entretanto, nos eucariotos, o ciclo celular é mais complexo e consiste em quatro fases distintas. Embora o crescimento celular seja normalmente um processo contínuo, o DNA é sintetizado somente durante uma fase do ciclo celular, e os cromossomos duplicados são então distribuídos aos núcleos-filhos por uma série complexa de eventos que precedem a divisão celular. O processo entre estes estágios do ciclo celular é controlado por um sistema regulador conservado, que não só coordena os diferentes eventos que controlam a proliferação celular.

**A replicação do DNA é mediada por enzimas que copiam a molécula com fidelidade.** As pesquisas envolvendo a replicação de DNA bacteriano e de suas enzimas auxiliaram a estabelecer inúmeras propriedades básicas que se mostraram aplicáveis à síntese de DNA em outros organismos. Existem evidências apontando para o papel de uma proteína agindo como molécula iniciadora do processo de replicação e também da existência de uma ligação física entre o DNA e a membrana celular. Essa ligação foi demonstrada com a região do DNA que contém a **origem da replicação** do cromossomo de *E. coli* (*OriC*) e extratos de membranas celulares bacterianas.

Origens de replicação já foram isoladas de organismos procariotos e de alguns eucariotos, apresentando a característica comum de serem ricas em sequências AT. O DNA de células procarióticas, plasmídeos e vírus possuem somente uma origem de replicação, enquanto que várias origens são encontradas nos organismos eucarióticos. O mecanismo da replicação do DNA em eucariotos ainda não é bem esclarecido.

### 2.2 O mecanismo da replicação mediado por enzimas

Durante o processo de síntese do DNA, as duas fitas de DNA parental de cada cromossomo são desenroladas, e cada uma das fitas dirige a síntese de uma fita complementar a si. Ou seja, cada fita do DNA serve como molde para a síntese de uma nova fita, produzindo duas moléculas novas do DNA, cada uma formada por uma fita nova e uma fita antiga. Assim, denominamos, que a **replicação do DNA é semiconservativa** (Figura 7).

A partir da origem, a replicação prossegue ao longo da fita de DNA em uma ou em ambas as direções, sequencialmente até o término. Na unidirecional, uma forquilha da

replicação parte da origem e segue replicando o DNA em uma só direção. Na bidirecional, duas forquilhas de replicação deixam a origem em direções opostas.

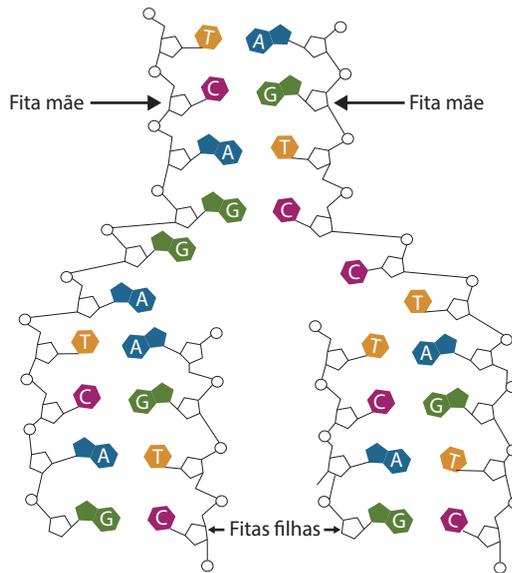


Figura 7. O mecanismo da replicação é baseado no pareamento da dupla hélice

O experimento de Meselson-Stahl demonstrou que a replicação do DNA é semiconservativa. No experimento, células de *E. coli* cultivadas na presença de  $^{15}\text{N}$ , foram incubadas em  $^{14}\text{N}$  e foi observado que o DNA fita dupla após uma geração continha uma fita com isótopo pesado ( $^{15}\text{N}$ ) e a outra com o isótopo leve ( $^{14}\text{N}$ ). Na segunda geração, a progênie era constituída de DNA híbrido ( $^{14}\text{N}/^{15}\text{N}$ ) ou de moléculas com as duas fitas contendo  $^{14}\text{N}$ .

O processo da replicação do DNA é baseado no pareamento da dupla hélice, onde o início se dá em *OriC*. O primeiro evento que ocorre é a ligação de uma proteína iniciadora no DNA de *Ori*, levando o desenrolamento parcial do DNA molde. O DNA continua sendo desenrolado pela ação da **helicase** (rompimento de pontes de hidrogênio entre as bases nitrogenadas) e **proteínas SSB** (proteínas de ligação ao DNA de fita simples), atuam continuamente no desenrolamento da dupla fita e de exposição da fita molde, sendo os *primers* de RNA sintetizados pela **primase**. As duas forquilhas de replicação formadas na origem movem-se em direções opostas ao longo da molécula circular de DNA em procariontos.

A medida que as fitas do DNA parental são desenroladas, as regiões do DNA à frente da forquilha da replicação são forçadas a rotar. Esse processo levaria a molécula ficar torcida sobre si mesma e eventualmente bloqueando a replicação. Neste sentido, as topoisomerasas, enzimas que catalisam a quebra e a reunião reversível das fitas do DNA conferem a estabilidade da fita parental aberta durante o processo de síntese de DNA.

Iniciado na origem da replicação, o processo de síntese de DNA resulta na forquilha da replicação em que um duplex de DNA parental se bifurca em duas fitas. A formação das novas fitas de DNA deve-se à atividade da DNA polimerase que adiciona nucleotídeos para a formação da fita filha de acordo com a fita molde. A DNA polimerase I (pol I) de *E. coli* foi descrita no trabalho de Arthur Kornberg em 1955. Com o advento das pesquisas outras DNA polimerases foram descobertas em *E. coli*.

Todas as DNA polimerases compartilham duas propriedades fundamentais com implicações na replicação do DNA. Há cinco classes descritas de DNA polimerase em mamíferos, inclusive uma que atua na replicação do genoma mitocondrial. Como descrito anteriormente, nos eucariotos a replicação do DNA é bidirecional, formando forquilhas da replicação a partir dos múltiplos pontos de iniciação. O início da replicação na fase S do ciclo celular, dá-se em momentos diferentes nas diversas origens, mas eventualmente, forquilhas da replicação podem fundir-se.

A polimerase requer uma fita única, não pareada para atuar como molde e um iniciador, que é um segmento de fita complementar ao molde com uma extremidade 3' OH livre possibilitando o crescimento da cadeia somente no sentido 5'→3', no qual os nucleotídeos são adicionados pela enzima. Cada nucleotídeo que chega é selecionado pelo pareamento de bases com o nucleotídeo apropriado da fita molde. O produto da reação tem uma nova extremidade 3' livre, o que permite a adição de outro nucleotídeo.

A reação da atividade da polimerase envolve a transferência de nucleotídeo trifosfato (dNTP) na extremidade 3' com hidroxila livre, promovendo o alongamento da cadeia (Figura 7.). Ocorre a formação da ligação fosfodiéster entre os resíduos de açúcar, ligações glicosídicas entre as desoxirriboses e as bases nitrogenadas e ligações do tipo pontes de hidrogênio entre as bases nitrogenadas.

As reações catalisadas pela DNA polimerase incluem a adição do substrato de dNMP, fornecido pelo dNTP. Como as fitas do DNA são antiparalelas e o crescimento da cadeia de DNA se dá sempre na direção 5'→3', isso gera uma assimetria no processo de replicação de DNA. Assim, em ambas as fitas, a síntese de DNA ocorre no sentido 5'→3' (Figura 8.). A fita sense (líder) é sintetizada de maneira ininterrupta na direção do movimento da forquilha da replicação, com a necessidade de apenas um iniciador (primer). No entanto, a fita antisense (descontínua) é sintetizada em pequenos pedaços, onde são gerados vários fragmentos, denominados fragmentos de Okazaki, e o crescimento da cadeia ocorre no sentido 5'→3' contrariamente em relação à direção

geral da forquilha da replicação. Na fita antisense é necessário a síntese de mais de um primer, assim, a cada novo primer sintetizado, o sentido do deslocamento da primase torna-se oposto ao da forquilha da replicação, formando um alça.

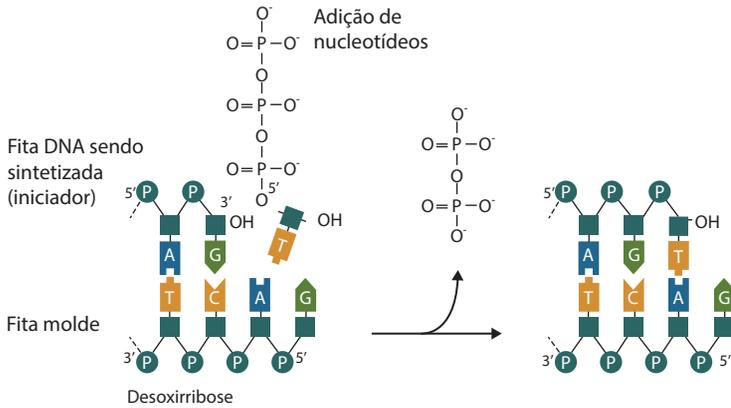


Figura 8. Alongamento da cadeia de DNA

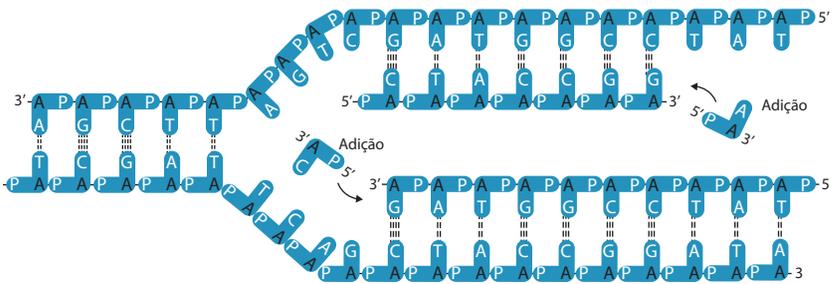


Figura 9. A síntese de DNA ocorre no sentido 5'→3" com adição de desoxinucleotídeos trifosfatados (dATP, dGTP, dCTP e dTTP)

Para formar uma fita contínua a partir de uma fita de síntese descontínua, os *primers* de RNA devem ser removidos dos fragmentos de Okazaki e substituídos por nucleotídeos que serão complementares a fita molde. Em *E. coli*, os *primers* são removidos pela ação combinada da RNase H, uma enzima que degrada RNA dos híbridos RNA-DNA, e da atividade de exonucleases 5'→3' da DNA polimerase I, a qual remove os ribonucleotídeos das extremidades 5' e preenche com nucleotídeos a falha entre os fragmentos de Okazaki, tais fragmentos são unidos pela DNA ligase que garante o crescimento da cadeia.

A Figura 9. resume a atividade das enzimas (replissomos) envolvidas no processo da replicação do DNA. É importante ressaltar que em procarionotos há também tipos diferentes de DNA polimerases atuando no processo da replicação.

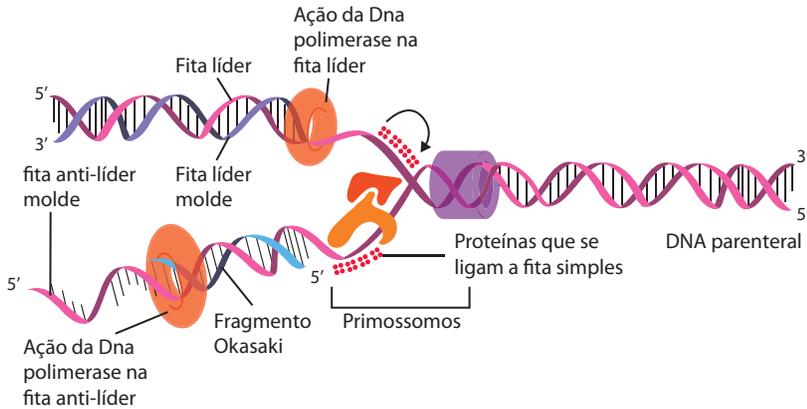


Figura 10 A assimetria da síntese das fitas (5'→3' e 3'→5') de DNA e enzimas envolvidas no processo da replicação do DNA

Na replicação do DNA de *E. coli*, as duas forquilhas de replicação durante o movimento bidirecional, encontram-se na região terminal TER localizada a 180° a partir de OriC. O término da replicação em genomas circulares em geral não apresenta problemas, pois as forquilhas de replicação devem se encontrar em algum ponto do DNA, no caso de replicação bidirecional, ou unidirecional, retornam à origem da replicação.

Acesse: <http://www.youtube.com/watch?v=teV62zrm2P0> para ver a animação envolvendo a atuação dos replissomos no processo da replicação

### 3. Transcrição e Processamento de RNA

#### 3.1 Síntese de RNA

Os mecanismos básicos pelos quais a transcrição é regulada foram igualmente elucidados em experimentos pioneiros feitos com *E. coli*, na qual a expressão gênica permite que a célula responda às variações no ambiente, tal como as mudanças na disponibilidade de nutrientes. Portanto, a compreensão da transcrição em *E. coli*

forneceu o embasamento para estudos dos mecanismos mais complexos que regulam a expressão gênica em células eucarióticas.

A **transcrição** é o processo pelo qual uma molécula de RNA é sintetizada a partir da informação contida na molécula de DNA fita dupla. Apenas uma das fitas de DNA é utilizada como molde durante a síntese de RNA, que segue as regras de pareamento que a replicação, exceto pelo pareamento de uracila (U) com adenina (A). Assim, a molécula de RNA sintetizada é complementar à fita de DNA que lhe deu origem e é idêntica à outra, sendo as timinas substituídas por uracila. A principal enzima responsável pela síntese de RNA é a RNA polimerase, que catalisa a polimerização de ribonucleotídeos 5' trifosfatados (NTPs), direcionados pelo molde de DNA.

A **RNA polimerase** tem múltiplas atividades: reconhecem e ligam-se a sequências específicas de DNA; desnaturam o DNA expondo a sequência de nucleotídeos a ser copiada; mantém as fitas de DNA separadas na região da síntese; mantém estável o híbrido DNA:RNA; renaturam o DNA na região imediatamente posterior de síntese e sozinhas ou com auxílio de proteínas específicas, terminam a síntese do RNA. A síntese de RNA é similar ao de DNA, e assim, como a DNA polimerase, a RNA polimerase catalisa o crescimento de cadeia de RNA sempre na direção 5'→3' (Figura 11).

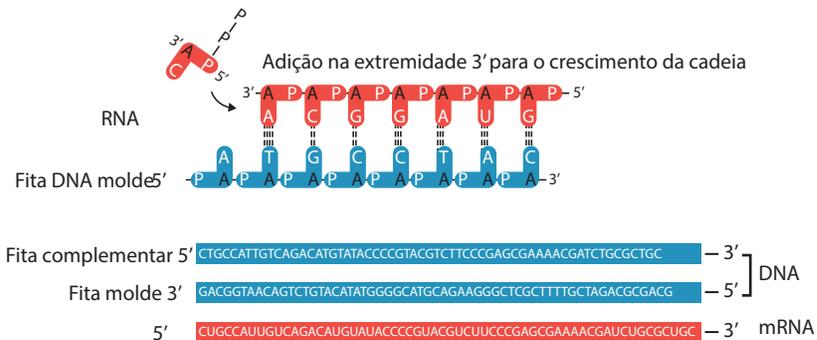


Figura 11. Mecanismo de transcrição do RNA

Contrariamente à DNA polimerase, a RNA polimerase não requer a presença de um *primer* pré-formado para iniciar a síntese do RNA. A RNA polimerase de *E. coli* é uma enzima complexa composta de múltiplas cadeias polipeptídicas. A enzima completa (core enzima) consiste em quatro diferentes tipos de subunidades:  $\alpha_2$ ,  $\beta$ ,  $\beta'$  e  $\sigma$ . A subunidade  $\sigma$  (sigma) está ligada fracamente e pode ser separada das outras subunidades gerando um núcleo que consiste em duas subunidades  $\alpha$ , uma  $\beta$  e uma  $\beta'$ . O subunidade  $\sigma$  está envolvida com o reconhecimento das sequências de DNA que indicará a partir de qual desoxirribonucleotídeo a síntese de RNA deverá ser iniciada.

A sequência de DNA na qual a RNA polimerase liga-se para iniciar a transcrição de um gene é chamada promotor, o qual sinaliza exatamente onde a síntese do RNA deve ser iniciada. Na ausência da subunidade  $\sigma$ , a RNA polimerase liga-se inespecificamente e com baixa afinidade ao DNA. O papel da subunidade  $\sigma$  é o de direcionar a polimerase aos promotores pela ligação às sequências de região -35 e -10, levando a iniciação de transcrição no começo do gene.

A síntese de RNA prossegue até que a polimerase encontre um sinal de término, onde o RNA é liberado da polimerase e a enzima se dissocia do molde de DNA. O tipo mais simples e comum de sinal de terminação em *E. coli* consiste em uma região repetida e invertida, simétrica, rica em GC, seguida de estrutura estável semelhante a um “grampo” (hairpin) de cabelo, por complementaridade de bases. A formação de tal estrutura auto complementar no RNA capaz de formar uma associação com o molde de DNA e termina a transcrição.

O processo de transcrição em eucariotos é bem complexo e ainda não está totalmente elucidado. Embora a síntese de RNA em procariotos e eucariotos seja muito semelhante, o processo como um todo apresenta diferenças. Em eucariotos há cinco tipos diferentes de RNA polimerases descritas que estão envolvidas na síntese de RNAs específicos.

As RNAs polimerases eucarióticas sempre necessitam de proteínas auxiliares, os fatores de transcrição. Os promotores são inicialmente reconhecidos por fatores de transcrição que, ligados ao DNA, interagem com outros fatores, formando um complexo ao qual as RNA polimerases se associam.

As sequências promotoras de muitos genes transcritos pela RNA polimerase do tipo II são identificadas pela proteína de ligação a TATA, a qual recruta fatores de transcrição adicionais e a RNA polimerase para o promotor. Os promotores eucarióticos são em geral sequências que estão próximo ao sítio de início da transcrição e possuem sequências consenso TATA (T/A)A(A/T), denominadas **TATA box**. Os **enhancers**, sequências regulatórias, contêm, tipicamente, sítios de ligação para proteínas múltiplas que atuam conjuntamente para regular a expressão gênica.

Acesse: <http://www.youtube.com/watch?v=WsofH466lqk> para visualizar o processo de síntese de RNA

### 3.2 Processamento do RNA

A molécula de RNAm recém sintetizada no processo de transcrição é chamada de transcrito primário. O processamento do transcrito primário ocorre tanto em procariotos,

como em eucariotos, bem como nos diferentes tipos de RNAs encontrados nas células. A maioria dos transcritos primários de RNAm de eucariotos apresenta uma particularidade, o RNAm que é sintetizado no núcleo deve, primeiro ser transportado para o citoplasma antes que possa ser usado como molde para síntese proteica. Esse RNAm precursor sofre várias alterações bioquímicas antes de serem exportados do núcleo. Desta forma, o processamento do pré-RNAm inclui a modificação de ambas as extremidades da molécula, bem como a remoção de íntrons de seu interior.

Uma modificação na extremidade 5' do pré-RNAm se deve a adição de um resíduo de 7-metilguanossina ligado ao resíduo 5' terminal do RNAm por meio de uma ligação 5',5'-trifosfato. Essa estrutura é denominada de **CAP**, a qual é decorrente da metilação na posição 7 da guanina, catalisada pela enzima guanina-7-metil-transferase, resultando no nucleotídeo 7-metilguanossina (Figura 12. A). O CAP é adicionado antes de a síntese do transcrito primário estar completa.

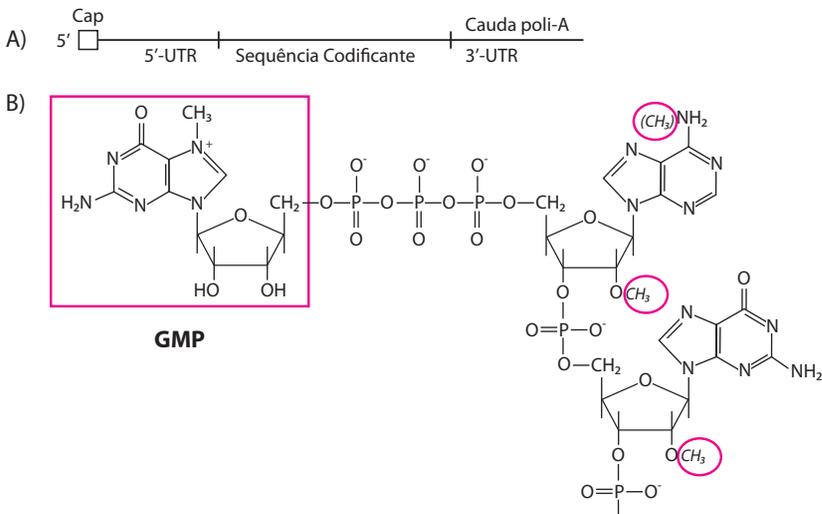


Figura 12. Processamento do RNA: A) o transcrito primário de organismo eucarióticos possuem uma 7-metilguanossina adicionada à extremidade 5' por meio de uma ligação 5'-5' trifosfato; B) Na extremidade 3', o transcrito primário recebe um conjunto de 80 a 250 resíduos de A, compondo a cauda poli(A).

Os RNAm de eucariotos possuem uma sequência de 80 a 250 resíduos de A na sua extremidade 3', a qual é denominada cauda poli (A) [Figura 12. B]. Essa cauda serve como sítio de ligação para uma ou mais proteínas associadas provavelmente ajudam a proteger o RNAm da destruição enzimática. Depois que o RNAm poliadenilado chega ao citoplasma, a cauda poli (A) vai diminuindo com o decorrer do tempo e função da atividade de nucleases presentes no citoplasma. A poliadenilação precede a excisão de íntrons.

A cauda poli (A) é adicionada em um processo de múltiplos passos. O transcrito é alongado além do sítio onde a cauda poli (A) será adicionada, sendo então clivado no sítio de adição da poli (A) por uma endonuclease componente de um grande complexo enzimático da RNA polimerase II. A poliadenilato-polimerase sintetiza uma cauda poli (A) de 80 a 250 nucleotídeos de comprimento, começando no sítio de clivagem.

Tanto os íntrons quanto os éxons são transcritos do DNA para o RNA. Os íntrons no transcrito primário são então removidos por *splicing*, sendo os éxons reunidos para formar um RNA maduro e funcional. A retirada de um íntron do transcrito primário ocorre após duas reações de transesterificação (transferência da ligação fosfodiéster) para unir os éxons, liberando o íntron na forma de laço (Figura 13).

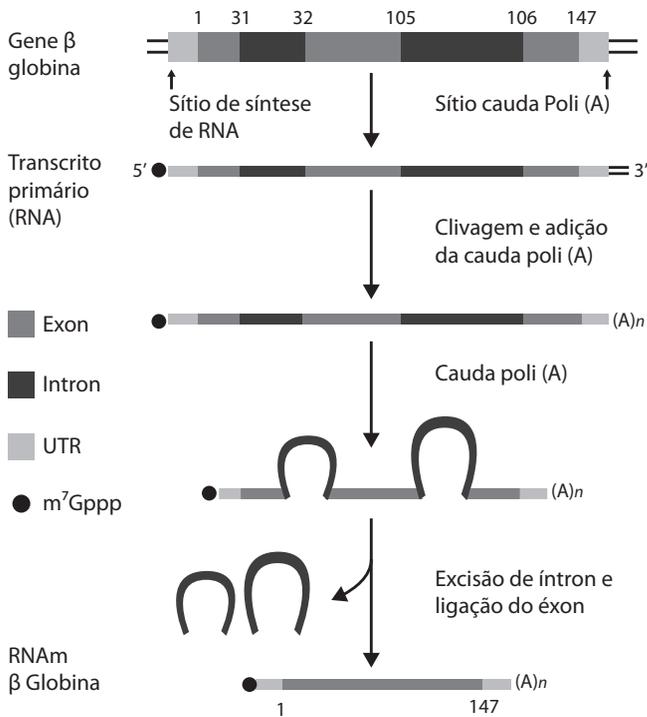


Figura 13. Mecanismo de splicing para remoção dos íntrons

No *splicing*, os Pré-RNAm nucleares acontece em grandes complexos chamados spliceossomos, compostos de proteínas e RNAs nucleares pequenos (RNAsn). Estes últimos identificam sequências junto aos sítios de *splicing* dos Pré-RNAm e catalisam a reação de *splicing*. Alguns RNAs mitocondriais, de cloroplastos e bacterianos sofrem *auto-splicing*, processo no qual a reação de *splicing* é catalisada por sequências de íntrons.



de nucleotídeos sejam lidos de maneira sucessiva e sem sobreposição na informação contida no RNAm. As quatro letras do código de DNA (A, T, C e G) em grupos de três, podem gerar  $4^3=64$  combinações diferentes, o que é insuficiente para codificar os 20 aminoácidos descritos na Figura 14., sendo que um aminoácido é codificado por mais de um códon, exceto metionina e triptofano. Desta forma, o código genético é considerado degenerado. Os códons que apresentam o mesmo aminoácido são considerados sinônimos.

Os códons são a chave para a tradução da informação genética, direcionando a síntese de proteínas específicas. O quadro de leitura é estabelecido quando uma molécula de RNAm é iniciada e é mantido à medida que a maquinaria de síntese vai lendo a mensagem sequencialmente, de um triplete para outro. Vários códons apresentam funções especiais.

O códon de iniciação AUG é o sinal mais comum para o início de um polipeptídeo em todas as células, além de codificar resíduos de metionina (Met) nas posições internas dos polipeptídeos. Os códons de terminação (UAA, UAG, e UGA) são denominados *stop* códons ou códons sem sentido, geralmente sinalizam o fim da síntese do polipeptídeo e não codificam qualquer aminoácido conhecido.

O código genético é considerado universal. Com exceção de pequenas variações nas mitocôndrias, em algumas bactérias e alguns eucariotos unicelulares. Os códons dos aminoácidos são idênticos em todas as espécies descritas. Portanto, parece que todas as formas de vida possuem ancestral comum na escala evolutiva, cujo código genético foi preservado ao longo da evolução biológica.

## 4.2 A maquinaria para síntese de proteínas

Cada aminoácido é codificado na sequência de DNA como um códon, contendo uma sequência de três nucleotídeos. Devem existir, portanto, moléculas adaptadoras que transfiram a informação contida no genoma à sequência de aminoácidos, nas proteínas. Os adaptadores são moléculas de **RNAs transportadores** (RNAt), que são sequências pequenas de nucleotídeos, sendo que há um RNAt para cada aminoácido.

Os aminoacil-RNAts (RNAt) servem então como adaptadores que alinham aminoácidos no molde de RNAm. A primeira base do códon no RNAm (lido na direção 5'→3') pareia-se com a terceira base do anticódon. Ou seja, os anticódons são orientados no sentido 3'→5' (Figura 15). As aminoacil-RNAt-sintetase catalisam a ligação entre o RNAt aos aminoácidos que então, ligam-se aos códons no RNAm pelo pareamento de bases complementares no anticódon. Primeiramente, nesta reação, o aminoácido é unido ao AMP, formando uma aminoacil AMP intermediária. Posteriormente, o aminoácido ativado é transferido para a 3' terminal do RNAt. A alta fidelidade do reconhecimento de aminoácidos resulta em parte de uma revisão de leitura, pela qual as aminoacil AMP incorretas são hidrolisadas em vez de ser ligadas ao RNAt no segundo passo da reação.

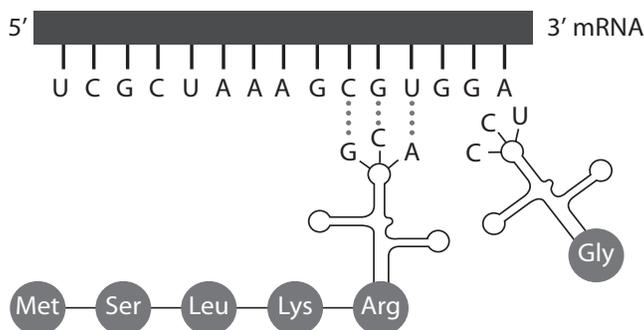


Figura 15. Pareamento entre códon e anticódon pelo aminoacil-RNA-t

Várias moléculas de RNAt foram isoladas a partir de organismos procarióticos e eucarióticos, apresentando uma estrutura secundária semelhante a folha de trevo. Leia sobre a estrutura do RNAt.

Os ribossomos são os sítios de síntese proteica, tanto em células procarióticas quanto em eucarióticas. Apresentam uma estrutura compacta de ribonucleoproteínas, consistindo de duas subunidades. Cada subunidade é formada por proteínas associadas a molécula de RNA ribossômico (RNAr). Normalmente as células contêm muitos ribossomos. Assim, como os RNAts, os RNAr formam estruturas secundárias características pelo pareamento de bases complementares.

Estruturas gerais dos ribossomos procarióticos e eucarióticos são similares, embora diferem em alguns detalhes. Pesquise sobre estas diferenças.

Os ribossomos contêm sítios específicos que tornam-se capazes de se ligar ao RNAm, aos RNAt e a outros fatores protéicos específicos, todos necessários a síntese proteica. Cada subunidade dos ribossomos possui vários sítios ativos, sendo os sítios A e P os relacionados à síntese de proteínas. O **Sítio A** (de aminoacil) é o sítio de entrada do aminoacil-RNAt que transporta o aminoácido. O peptidil-RNAt que transporta a cadeia de polipeptídeos crescente se liga ao **Sítio P** (de peptidil). Cada novo aminoácido é adicionado pela transferência da cadeia em crescimento ao novo aminoacil-RNAt, formando uma nova ligação peptídica. O RNAt que cedeu seus aminoácidos é então liberado do sítio A, e o ribossomo se move um códon de

distância ao longo do RNAm, transferindo o novo peptidil-RNAt ao sítio P e deixando o sítio A livre para o próximo aminoacil-RNAt (Figura 16).

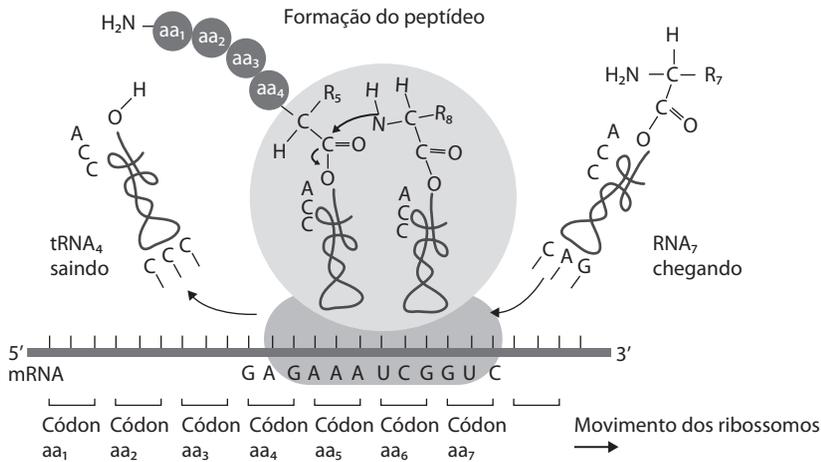


Figura 16. Adição de aminoácidos pelo aminoacil-RNA-t de acordo com a sequência de códons para formação do peptídeo

O ribossomo se movimenta, e o RNAt vazio é posicionado no sítio E antes de ser descartado. O **sítio E** (de saída, do inglês Exit) contém um RNAt (ele não leva mais um aminoácido) que está pronto para ser liberado pelo ribossomo. Esse ciclo de reações é repetido cada vez que um aminoácido é adicionado à cadeia polinucleotídica, fazendo com que um aminoácido seja adicionado e sua extremidade amino rumo à sua extremidade carboxila até o momento em que um códon de término seja encontrado.

### 4.3 A tradução

Embora os mecanismos da síntese proteica em células procarióticas e eucarióticas sejam similares, também existem diferenças, particularmente nos sinais que determinam as posições nas quais a síntese de uma cadeia polipeptídica é iniciada com base em um molde de RNA. A **tradução** começa em sítios específicos de iniciação. Tanto os RNAs mensageiros procarióticos quanto os eucarióticos contêm regiões não-traduzidas (UTRs) nas suas extremidades 5' e 3'. Os RNAs mensageiros eucarióticos usualmente codificam somente uma única cadeia polipeptídica, mas muitos RNAs procarióticos codificam múltiplos polipeptídeos que são sintetizados independentemente a partir de sítios de iniciação distintos. Os detalhes da expressão gênica entre procariotos e eucariotos será discutido posteriormente.

A tradução é geralmente dividida em três estágios: iniciação, alongamento e término. Tanto em procariotos como em eucariotos, o primeiro passo do estágio de iniciação é a ligação de um iniciador específico de RNAt metionil e do RNAm à subunidade ribossomal pequena. A subunidade grande então, liga o complexo, formando um ribossomo funcional, no qual o alongamento da cadeia polipeptídica prossegue até o ribossomo encontrar um códon de término no RNAm. Uma variedade de fatores não ribossomais é necessária para a iniciação, alongamento e terminação da tradução.

Acesse: <http://www.youtube.com/watch?v=Ikq9AcBcohA&feature=fvwrel> para visualizar as animações do mecanismos de síntese de proteínas

Quanto a **regulação da tradução**, RNAs específicos sofrem ação de proteínas repressoras. A atividade geral da tradução pode ser regulada também pela modificação de fatores de iniciação, particularmente IF2. A própria poliadenilação também foi considerada um dos mecanismos importante da tradução durante o desenvolvimento inicial. Uma variedade de RNAm é estocada na forma não traduzida de oócitos; a tradução desses RNAm é ativada pela fertilização ou estágios tardios do desenvolvimento.

## 5. Controle da expressão gênica

### 5.1 Controle da expressão gênica em procariotos

Essencialmente, a expressão gênica é dividida em transcrição, processamento e tradução, e pode ser controlada em qualquer um dos estágios. As bactérias possuem necessidade de regular a expressão de seus genes. Quando há mudança ambiental, a célula bacteriana possui condições de se adaptar à nova situação, por mudança na atividade metabólica. Essa atividade pode ser regulada pelo controle da síntese de enzimas e outras moléculas.

As enzimas que participam do metabolismo de açúcar são necessárias para decompor fontes diferentes de carbono para produzir energia. Existem muitos tipos diferentes de compostos que as bactérias utilizam como fonte de carbono, incluindo açúcares como a lactose, glicose, galactose e xilose. Uma enzima diferente é necessária para permitir que cada um destes açúcares seja metabolizado. Neste sentido, para cada substrato é necessário uma enzima. Se caso uma célula bacteriana fosse sintetizar ao mesmo tempo todas as enzimas de que possivelmente necessitaria, isto iria lhe custar muito mais energia para expressar os genes necessários ao metabolismo do que obteria no metabolismo destas fontes de carbono.

Portanto, a célula bacteriana criou mecanismos moleculares para regular a transcrição de seus genes que codificam enzimas de acordo com as mudanças ambientais e suas necessidades fisiológicas. Os fatores determinantes que controlam a quantidade em que cada proteína é sintetizada e mantida em uma célula bacteriana são concentrações de RNAm correspondentes às suas proteínas. Ou seja, a expressão gênica é controlada em nível de transcrição, tradução e pós-tradução.

As proteínas podem ser divididas em duas grandes classes quanto a sua forma de expressão: as que são expressadas constitutivamente e aquelas cuja a expressão é ativada, seletivamente induzidas, de acordo com as necessidades da célula, as quais são controladas por mudanças ambientais e genéticas.

Os procariotos possuem genes que codificam enzimas em uma mesma rota metabólica e estão dispostos linearmente no cromossomo bacteriano e são controlados coordenadamente. Esses genes estão agrupados em unidades transcricionais policistrônicas (Operon). Assim, um **operon** é considerado uma unidade de expressão gênica incluindo genes estruturais e elementos que controlam sua expressão. Operons são comuns em bactérias, mas não são encontrados em eucariotos, onde os genes são transcritos e regulados individualmente.

Um operon é composto por uma região promotora, pelo operador e os genes estruturais. Um operon é transcrito em um único RNAm policistrônico, o que significa que codifica mais de uma cadeia polipeptídica com todos os genes transcritos, ou quando nenhum deles é transcrito, resulta em um sistema de controle de energia na célula bacteriana.

A atividade de uma proteína regulatória controla a expressão gênica ligando-se ao operador próximo ao promotor dos genes que serão regulados. As proteínas regulatórias normalmente ligam-se a sequências regulatórias de DNA próximas de onde a RNA polimerase se liga e então ativam ou reprimem a transcrição do gene. A partir do momento em que a proteína repressora se desliga do operador, há o início da transcrição dos genes. Os indutores ligam-se ao repressor e modificam sua conformação, diminuindo sua afinidade pelo operador.

As características da regulação na transcrição em procariotos pode ser vista na regulação da expressão das enzimas necessárias para o metabolismo de lactose em bactéria *E. coli*, o operon *lac*. O sistema existente passou por um longo processo evolutivo, o qual foi selecionado para operar de uma maneira ótima para a eficiência de energia da célula bacteriana. Presumivelmente, por causa de considerações de eficiência energética, duas condições ambientais têm de ser satisfeitas para que as enzimas metabólicas de lactose sejam expressas.

Uma condição é que a lactose tem de estar presente no meio ambiente. Esta condição faz sentido, porque seria ineficiente para a célula para produzir as enzimas metabólicas para lactose em circunstâncias em que não há nenhum substrato para metabolizar. A outra condição é que a glicose não pode estar presente no ambiente da célula. Porque o metabolismo da glicose produz mais energia utilizável para a célula

do que o faz o metabolismo de lactose. A repressão da transcrição dos genes para o metabolismo da lactose na presença de glicose é um exemplo de repressão catabólica.

### 5.1.1 Operon *lac*

O **Operon lactose** (*Operon lac*) contém três genes estruturais, Z, Y e A. O metabolismo de lactose em procariontes requer duas enzimas: uma permease para transportar a lactose na célula e  $\beta$ -galactosidase para clivar a molécula de lactose para produzir glicose e galactose.  $\beta$ -galactosidase e Permease são codificadas pelos genes, Z e Y, respectivamente. Um terceiro gene, o gene A, codifica uma enzima adicional, denominada transacetilase, que hidrolisa os  $\beta$ -galactosídeos que não foram hidrolisados, estes então são acetilados pela transacetilase, e uma vez acetilados podem se difundir através da membrana celular e então serem eliminados pela célula. Assim, as transacetilases estão envolvidas com a acetilação dos compostos não degradados pela  $\beta$ -galactosidase. Todos os três genes são transcritos em uma única molécula de RNAm.

Um quarto gene, o I, que codifica uma proteína repressora, responsável pelo bloqueio da expressão dos três genes Z, Y e A. Esse gene está mapeado próximo aos genes policistrônicos. A proteína expressa pelo gene I, denominada repressora se liga a região do operador impedindo que a RNA polimerase ligada ao promotor transcreva os genes estruturais adjacentes. Quando a lactose está presente, ela se liga ao repressor e modifica sua conformação de modo que o repressor não se liga mais ao operador, de modo que o repressor perde a afinidade ao operador. Assim, a RNA polimerase é capaz de transcrever os três genes estruturais (Z, Y e A) ocorrendo a expressão das três enzimas (Figura 17). Os derivados da lactose que inativam o repressor e levam a expressão dos genes lac são denominados de indutores.

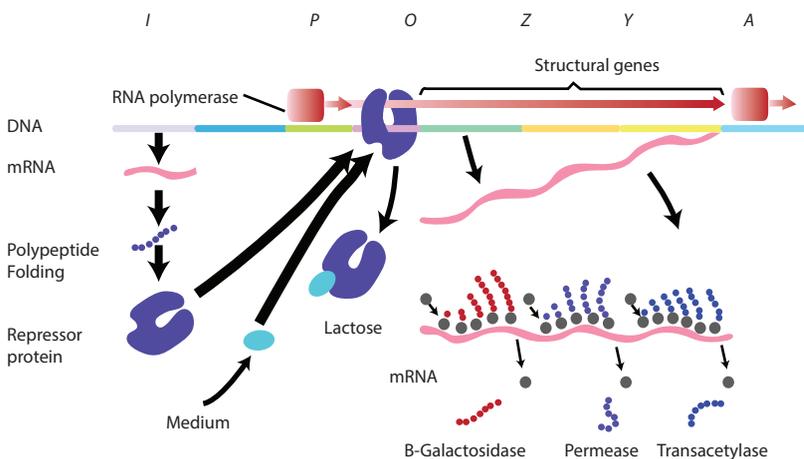


Figura 17. Operon *lac* - modelo de regulação gênica em procariontes.

Quando o repressor se liga ao operador, o repressor impede o início da transcrição pela RNA polimerase. Tal repressão da transcrição dos genes policistrônicos no operon *lac*, se deve a baixa concentração de lactose no meio. Assim, o repressor se liga ao operador impedindo a expressão dos genes *Z*, *Y* e *A*. Normalmente, a RNA polimerase liga-se a regiões específicas, no promotor, de modo que possa ocorrer a transcrição dos genes policistrônicos no operon *lac* quando na presença de lactose (Figura 17).

Há um sistema adicional de controle superposto ao sistema repressor-operador em operon *lac*, repressor catabólico. Se tanto a lactose quanto a glicose estiverem presentes, a síntese de  $\beta$ -galactosidase não é induzida até que a glicose tenha sido usada. Quando a glicose está presente em altas concentrações, a concentração AMPc é baixa, induzindo a repressão catabólica. À medida que a concentração de glicose diminui, a concentração de AMP<sup>+</sup> aumenta, o que se faz necessário para a ativação do operon *lac*.

Outros operons foram descritos em *E. coli*, como o operon triptofano, operon arabinose e operon maltose. Leia mais sobre estes operons para melhor compreensão da expressão gênica em procariotos. Acesse: <http://www.youtube.com/watch?v=oBwtxdI1zvK> para visualizar a animação do operon *lac*.

## 6. Controle da expressão gênica em eucariotos

### 6.1 Controle da expressão gênica em eucariotos

Nos organismos procarióticos, a expressão gênica depende fundamentalmente de condições nutricionais e ambientais. Embora a transcrição atue através dos mesmos mecanismos em todas as células, ela é mais complexa nos organismos eucarióticos. Diferentemente dos procariotos, a expressão gênica em eucariotos envolve múltiplas RNA polimerases diferentes que transcrevem classes distintas de genes. Tais RNA polimerases necessitam de proteínas adicionais para, de maneira específica, iniciarem a transcrição.

A expressão de genes eucarióticos é controlada primariamente em termos de iniciação da transcrição, embora em alguns casos, a transcrição possa ser atenuada e regulada em etapas posteriores. Assim, como nos procariotos, a transcrição em células eucarióticas é controlada por proteínas que se ligam especificamente às sequências regulatórias e modulam a atividade da RNA polimerase. A expressão gênica nos vários tipos celulares diferenciados dos organismos multicelulares é desempenhada primariamente pelas ações combinadas de múltiplas proteínas transcricionais regulatórias

diferentes. Além disso, o empacotamento do DNA na cromatina e sua modificação por metilação acrescenta um nível ainda maior de complexidade no controle da expressão gênica em eucariotos.

O DNA de todas as células eucarióticas está firmemente ligado a histonas, formando a cromatina. A transcrição de um gene eucariótico é fortemente reprimida quando o DNA está condensado dentro da cromatina. Os genes ativamente transcritos são encontrados em cromatina descondensada correspondendo às fibras de cromatina estendidas. Contudo, a descondensação da cromatina não é suficiente para tornar o DNA um molde acessível para a transcrição.

Neste sentido, o remodelamento da cromatina permite a ativação transcricional de genes que é mediado pela movimentação dos nucleossomas e por acetilação/metilação de histonas. A modificação das histonas por acetilação aumenta o acesso do DNA nucleossomal aos fatores de transcrição, e essas modificações na cromatina estão intimamente ligadas à regulação transcricional. As enzimas que catalisam a acetilação das histonas estão associadas com ativadores transcricionais, enquanto as histonas desacetilases estão associadas com repressores da transcrição. A metilação nas extremidades do promotor está associada à ausência de transcrição. Este é um dos diversos eventos de regulação que influenciam a atividade de um promotor.

A maioria dos genes é controlada em nível de transcrição. A ativação de genes específicos em células eucarióticas superiores tem importante papel durante o desenvolvimento embrionário e a diferenciação tissular. Em geral, o ligar e desligar de grupos de genes não é reversível, ao contrário da situação em bactérias, onde a atividade gênica é continuamente modulada em resposta à condições fisiológicas. Entretanto, existem exceções nos eucariotos. Por exemplo, as leveduras respondem a mudanças de condições ambientais.

Fatores acessórios são necessários para a iniciação da transcrição com o reconhecimento da região promotora. Três classes de RNA polimerase atuam na transcrição de eucariotos e possuem estruturas semelhantes. A RNA polimerase I transcreve RNAr; a RNA polimerase II transcreve RNAm e a RNA polimerase III transcreve RNAt e outros pequenos RNAs. Nenhuma das três enzimas reconhece seus promotores diretamente, é necessário a atuação de fatores de transcrição. Por exemplo, o fator TBP (do inglês, *TATA binding protein*) é necessário para a iniciação por todas as três enzimas de polimerização.

A ligação bem sucedida da RNA polimerase no seu promotor e a consequente iniciação da transcrição, requer a ação de outras proteínas, de três tipos. A primeira classe compreende os fatores de transcrição basais requeridos para os promotores da RNA polimerase II. A segunda inclui os transativadores de ligação ao DNA; que se ligam aos intensificadores ou UAS (sequências ativadoras "acima") e facilitam a transcrição. A terceira inclui os co-ativadores, que agem indiretamente, não pela ligação ao DNA, e são requeridas para comunicação entre os transativadores de ligação ao DNA,

e o complexo de transcrição basal. Além disso, proteínas repressoras podem interferir nesta comunicação.

A RNA polimerase II (RNA pol II) para obter máximas taxas de transcrição necessita de promotores, elementos proximais a promotores e *enhancers* (acentuadores). Cada um desses elementos estão associados ao controle positivo. Os promotores e elementos proximais a promotores funcionam próximos ao ponto de iniciação, enquanto que os acentuadores estão a distância deste ponto.

O TBP reconhece a região iniciadora no DNA, que inclui o sítio da transcrição TATA *box*. Suas subunidades são denominadas TAFs (do inglês, *Associated factors* = fatores associados à TBP). Depois da ligação de TBP, fatores de transcrição e a RNA polimerase associam-se a região promotora, formando um complexo basal entre a RNA pol II e os fatores gerais de transcrição. Esse complexo de iniciação é montado sobre os promotores para a RNA pol II por uma sequência ordenada de associação entre os fatores de transcrição.

Assim, a maioria das transcrições requer a presença de co-ativadores que formarão um elo entre os transativadores e a polimerase. O co-ativador mais bem caracterizado é o TFIID (do inglês, *Transcription Factor II D*), que inclui o TBP e as TAFs, que podem interagir com os transativadores, facilitando a transcrição. Um outro co-ativador importante é o mediador encontrado na levedura que se liga ao CTD (do inglês, *C-Terminal Domain*) da RNA polimerase II.

A ligação de fatores a sequências específicas é seguida por interações do tipo proteína-proteína com componentes gerais do complexo de transcrição. Os fatores de transcrição geralmente possuem uma construção modular, na qual domínios independentes responsáveis pela ligação no DNA e ativação da transcrição. A função principal do domínio de ligação no DNA talvez seja a de “prender” o domínio ativador nas proximidades do complexo de iniciação.

Os motivos de domínios de ligação ao DNA dos fatores eucarióticos de transcrição pertence a várias famílias estruturais. Os transativadores possuem um domínio estrutural para ligação específica no DNA e um ou mais domínios adicionais para ativação transcricional e para interação com co-ativadores. A interação entre proteínas reguladoras é mediada por domínio contendo motivos zíperes de leucina ou hélice-alça-hélice.

Tanto zíperes de leucina como as proteínas hélice-alça-hélice possuem papéis importantes na regulação da expressão gênica tecido-específica, sendo a formação de dímeros entre diferentes membros dessas famílias é um aspecto crítico do controle de suas funções. Os receptores de hormônios esteróides interagem com transativadores transcripcionais. A estrutura deste tipo de receptor apresenta uma região N-terminal variável, e uma região central conservada, constituída de dois dedos de zinco, que possibilitam a ligação ao DNA. No tecido alvo os transativadores difundem-se através da membrana e ligam-se ao receptor específico no núcleo. O complexo liga-se a sequência elemento de

resposta hormonal e alteram a expressão gênica. A ligação altera a conformação do receptor de forma que ele torna-se capaz de interagir com fatores de transcrição adicionais, aumentando ou diminuindo a expressão de genes adjacentes.

Embora a expressão gênica de eucariotos seja predominantemente regulada em nível de transcrição, em determinados estágios do desenvolvimento ou em situações especiais a regulação pode acontecer em nível de tradução. O desenvolvimento de um organismo multicelular apresenta o desafio regulatório mais complexo. O destino das células no início é determinado pelo estabelecimento de gradientes de proteínas, anterior posterior e dorso ventral, que atuam como transativadores ou repressores, regulando genes requeridos para o desenvolvimento de estrutura apropriada para uma parte específica do organismo. Assim, conjunto de genes regulatórios atuam em sucessão temporal e espacial, transferindo certas áreas de uma célula ovo em estruturas predizíveis no organismo adulto.

## 7. Mutações e reparo

### 7.1 Mutações

Desde o início do projeto genoma humano e mais recentemente, da proteômica, tem havido grande fluxo de informações a cerca dos genes e seus produtos. Consequentemente, existe hoje, um melhor entendimento da História Natural da Doença, no que se refere ao conjunto de inter-relações que envolvem o suscetível, a etiologia e o ambiente, conferindo o binômio saúde-doença.

O nosso genoma é constituído de 3 bilhões de nucleotídeos, dos quais cerca de 2% codificam proteínas. Grande porcentagem do material genético restante corresponde a regiões repetidas do DNA incluindo sequências derivadas de elementos de transposição.

O DNA não é uma molécula estática. Frequentemente suas bases estão expostas a agentes físicos (radiação ionizante ou por luz ultravioleta), químicos ou biológicos, ou por erros de cópia do material genético durante a divisão celular que provocam modificação na sua estrutura ou composição química. Tais alterações são denominadas de mutações. Uma mutação é definida como qualquer alteração na sequência de bases ou arranjo do material genético, o ácido desoxirribonucléico (DNA). Costuma-se distinguir três categorias de mutações:

- 1) aquelas que afetam o número dos cromossomos;
- 2) aquelas que afetam a estrutura dos cromossomos;
- 3) aquelas que alteram genes individuais.

As duas primeiras categorias de mutações são objetos de estudo da citogenética, e a última identificável por técnica de biologia molecular. As mutações são ditas deletérias

ou patogênicas, quando causam doenças. Em alguns casos, a presença apenas de um alelo mutado (heterozigose) é suficiente e, em outros, há a necessidade de que os dois alelos de determinado gene estejam alterados (homozigose), para que ocorra a manifestação clínica. No entanto, a maioria das alterações presentes no genoma de uma pessoa não causa um efeito drástico sobre suas características fenotípicas, sendo assim consideradas neutras.

Há três mecanismos básicos pelos quais uma mutação gênica pode desencadear alterações fenotípicas:

**Perda de função:** resulta em um produto com função reduzida ou sem função, seja de forma quantitativa (deficiência na produção), seja qualitativa (o produto gênico gerado em quantidades normais, não é capaz de desencadear o efeito esperado). Mutações desse tipo frequentemente causam doenças recessivas, pois níveis inferiores a 50% são suficientes para a grande parte dos produtos gênicos (principalmente enzimas). As manifestações fenotípicas sensíveis a dosagem gênica são o resultado de mutações que ocorrem nos genes que codificam ou receptores ou mais raramente enzimas, cujas funções são limitadoras de taxa, por exemplo hipercolesterolemia familiar.

**Ganho de função:** resulta em um produto com atividade exacerbada ou com nova função. Geralmente, causa fenótipos dominantes, por que a presença do alelo normal não evita o comportamento anormal do alelo mutado. As mutações que alteram a especificidade do tecido de expressão de um gene também podem ser consideradas como sendo mutações de ganho de função. Os exemplos incluem os rearranjos cromossômicos que resultam na combinação de sequências de dois genes diferentes vistas em tumores específicos, como a leucemia mielóide crônica e linfoma de Burkitt. Onde a nova função do gene quimérico promove o processo neoplásico. As mutações de ganho são herdadas dominantemente, e os raros casos de mutações de ganho que ocorrem no estado homozigoto são em geral associados a um fenótipo muito grave, que geralmente é um distúrbio letal pré natalmente, por exemplo acondroplasia homozigótica.

**Efeito dominante – negativo:** o produto alterado, prejudica a função do produto codificado pelo alelo normal, apresentando portanto, herança dominante. Assim, é uma mutação na qual um gene mutante no estado heterozigoto resulta na perda da atividade da proteína ou funcionamento, como consequência do produto gênico mutante interferindo na função do produto gênico normal do alelo correspondente. As mutações dominantes negativas são particularmente comuns nas proteínas que são dímeros ou multímeros, por exemplo proteínas tais como os colágenos, as quais podem resultar em osteogênese imperfeita.

Do ponto de vista da genética molecular, poderíamos chamar de mutação toda alteração do código genético que tiver por consequência final uma doença. Uma sequência de DNA está sujeita a, praticamente, todos os tipos de mutações imagináveis. Podendo haver uma perda de códon(s) inteiro(s)- deleção sem desvio de quadro de

leitura; ou deleções que interrompam as sequências triplete de um códon – deleção com desvio do quadro de leitura da polimerase; mutação que leve ao ganho do material genético extra (inserção); troca de uma única base nitrogenada por outra, fazendo com que o triplete codifique para outro aminoácido- mutação de ponto; troca de base(s) nitrogenada(s) para outra(s), fazendo com que o triplete codifique para um código de início ou parada da transcrição.

Tanto o tipo de mutação como a localização desta dentro de um éxon de um gene, e que tipo de alteração polipeptídica essa mutação vai causar, poderão gerar variabilidade de expressão de uma determinada doença. Uma mesma mutação que desvie o quadro de leitura da polimerase terá, possivelmente, um efeito fenotípico muito diferente se ocorrer no primeiro códon de um gene de 3.000 códon, ou se ocorrer no último. Mutações podem ocorrer em vários momentos do ciclo biológico, atingindo etapas da replicação, transcrição ou tradução. Podem atingir todas as células, quando são herdadas, ou parte das células do corpo quando adquiridas durante a vida.

Na ausência de qualquer evidência de um efeito mutagênico, uma mutação de ocorrência natural é considerada como sendo espontânea. A amplitude de uma mutação pode variar de uma mudança em um único par de bases (mutação de ponto) a uma alteração de uma grande região de um cromossomo (anomalia cromossômica). As mutações podem ocorrer em qualquer ponto no DNA total de um organismo.

Em humanos, aproximadamente 95% do DNA não codifica produtos gênicos. Como resultado, muitas mutações não apresentam efeito fenotípico, pois elas estão localizadas em regiões do genoma que não tem impacto em funções celulares. Em contraste, uma mutação dentro de um éxon de um gene estrutural pode alterar o funcionamento do produto gênico e causar uma grande alteração fenotípica. Finalmente, a compreensão da base biológica de um distúrbio genético depende da caracterização do gene responsável pela condição e determinação das consequências das mutações deste gene.

Quando as mutações gênicas ocorrem em células que não originam gametas, elas são chamadas de mutações somáticas, e não tem chance de serem passadas de um genitor para sua prole. Por outro lado, as mutações no DNA das células da linhagem germinativa (mutações de linhagem germinativa) podem ser transmitidas pelos gametas para a geração seguinte.

Uma terminologia simples descreve os tipos de substituições de nucleotídeos que podem ocorrer em nível de DNA como mutações de genes estruturais. As bases guanina e adenina tem as mesmas estruturas químicas fundamentais, enquadra-se na classe de compostos chamados purinas. Similarmente, as bases citosina e timina são quimicamente relacionadas uma à outra e a outros compostos, chamados pirimidinas. Qualquer substituição de uma purina por outra purina diferente ( $G < > A$ ) ou de pirimidina por outra pirimidina diferente ( $T < > C$ ) em uma molécula de DNA é uma mutação de transição. Uma mutação de transversão é qualquer substituição de uma

purina por pirimidina ou vice versa (G <> T; A <> T; G <> C; A <> C). As transições ocorrem mais frequentemente que as transversões.

Uma substituição dentro da região codificante de um gene estrutural pode mudar um códon de mRNA e leva a inserção de um aminoácido diferente na proteína. Pode surgir alguma confusão quando são descritos estes tipos de mutações, pois as mudanças de nucleotídeos nas sequências de DNA e RNA devem ser consideradas. As consequências de mutação de um códon do DNA dependem de qual par de bases é mudado, da natureza da substituição, da especificidade do novo códon e da localização. Em geral, as mutações de códons do DNA são classificadas como:

**1) Silenciosas:** ocorre quando há uma mudança em um códon do DNA, mas o aminoácido que é inserido na proteína não é alterado. Para muitos aminoácidos, existem vários códons. Em vários destes casos, a diferença entre códons para um único aminoácido está na terceira posição, onde U, G, C ou A podem estar presentes sem alterar a especificidade do códon. Assim, uma mutação silenciosa não tem impacto no funcionamento da proteína.

**2) Neutras:** representa uma mudança de nucleotídeo em nível de DNA que altera um códon, de modo que outro aminoácido é incorporado na proteína sem aparente perda de função. Este tipo de mudança é tolerada se a substituição ocorrer em uma parte da proteína que não é importante para seu funcionamento, ou se o aminoácido alternativo tiver propriedades físico – químicas similares ao original.

**3) Sentido trocado:** é quando uma substituição de par de bases que produz um códon que especifica outro aminoácido. A gravidade deste tipo de mutação depende da natureza do aminoácido substituído, e se o aminoácido original tem papel essencial na função da proteína. Assim, uma mutação neutra pode ser considerada de sentido trocado sem consequências óbvias.

**4) Sem sentido:** ocorre quando a substituição de um nucleotídeo muda um códon que especifica um aminoácido para um códon finalizador. A presença de um códon finalizador dentro de um mRNA causa a produção de uma proteína incompleta (truncada). Se a mutação sem sentido for perto do N-terminal, é provável que seja sintetizado um fragmento proteico funcional. Entretanto, se for perto da ponta C, pode ser produzida uma cadeia polipeptídica incompleta tendo uma atividade limitada.

**5) Mudança de matriz de leitura:** quando um par de base é ou inserido ou deletado de uma região codificante do gene estrutural, a sequência de códons após uma destas mutações pode ser alterada, de modo que os novos códons são traduzidos em uma sequência de aminoácidos completamente nova, sem nenhuma semelhança com a proteína original. Esse tipo de mutação tem efeito devastador no funcionamento de uma proteína, pode ser retida uma capacidade limitada.

**6) Sítios de recomposição mutados:** a remoção dos íntrons do transcrito primário de RNA depende de nucleotídeos específicos nos limites de uma extremidade de um éxon e o começo do outro íntron (sítio de corte 5') e o final de um

íntron e começo do éxon seguinte (sítio de corte 3'). A maquinaria de recomposição une os nucleotídeos de um sítio de corte em 5' ao sítio de corte 3' seguinte. Se os sítios de nucleotídeos ou os nucleotídeos vizinhos de um sítio de corte no DNA são alterados por mutação, então, após a transcrição, ocorre uma recomposição anormal. Especificamente, se um sítio de corte em 3' está mutado, a maquinaria de corte pulará este sítio e complementarará o corte no sítio 3' seguinte apropriado. Neste caso, um éxon flanqueado por dois íntrons é removido de um transcrito primário de RNA, o mRNA processado tem um éxon faltando. Esta forma aberrante de recomposição é chamada de saltar de éxon. Similarmente, se o sítio de corte em 5' é modificado, a maquinaria de corte irá ignorar este sítio no transcrito primário de RNA, e um íntron será incluído como parte do mRNA processado. Em ambos os casos, também é possível um evento alternativo de corte. Entretanto, quando ocorre uma mutação no sítio de corte em 3' ou 5', independente do evento de recomposição, nenhuma das moléculas processadas de mRNA tem a sequência correta de nucleotídeos para a tradução da proteína autêntica.

De todas as mutações nos genes causadores de doenças. As mutações de sentido trocado /sem sentido são 3 a 6 vezes mais frequentes que as deleções e os defeitos de recomposição respectivamente. A tabela 7.1. sumariza os tipos de mutações e suas respectivas frequências

<b>Tabela 7.1.Frequência das classes de mutações</b>	
<b>Frequências de mutações em genes causadores de doenças</b>	
<b>Tipo de mutação</b>	<b>Frequência(%)</b>
Deleção	21,8
Inserção/duplicação	6,8
Rearranjos complexos	1,8
Recomposição	9,8
Reguladora	0,8
Sentido trocado/sem sentido	58,9
Sequências repetidas	0,1

## 7.2 Reparo do DNA

As células dos organismos vivos possuem diversas enzimas envolvidas no mecanismo de reparo do DNA, impedindo erros de cópias, quebras e alterações que ocorrem na estrutura desta molécula responsável pela informação genética. Assim, durante a evolução foram selecionadas diversas estratégias para minimizar os efeitos deletérios das lesões em nível de DNA.

As lesões em nível de DNA podem bloquear a replicação ou transcrição podendo levar a um alta taxa de mutações. Assim, para manter a integridade do genoma, as células desenvolveram mecanismos para reparar essas lesões. Esses mecanismos podem ser divididos em duas classes: reversão direta da reação química responsável pelo dano e remoção das bases alteradas seguida por substituição com DNA recém-sintetizado. Onde a reparação do DNA falha, mecanismos adicionais desenvolvem-se para possibilitar que as células lide como o dano provocado. A Figura 7.1 demonstra o mal pareamento entre as bases nitrogenadas.

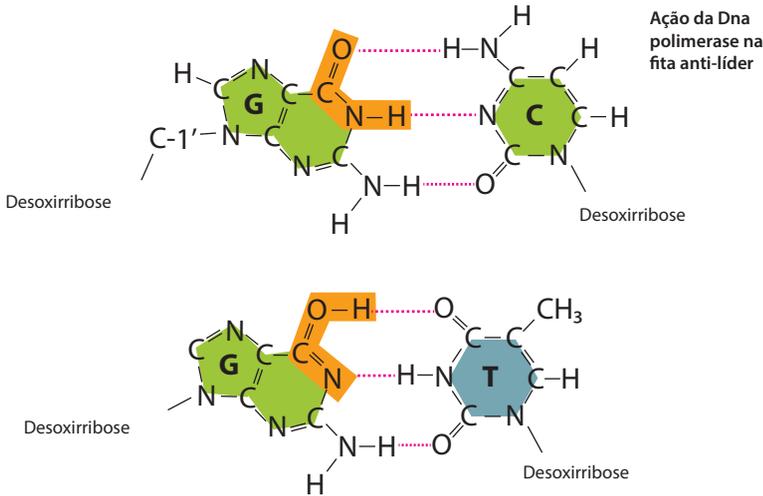


Figura 18. Mal pareamento de bases

### 7.2.1 Reversão da lesão no DNA

A reversão da lesão do DNA é o mecanismos mais simples pelo qual p DNA lesado deve ser reparado. Somente alguns poucos tipos de lesões no DNA são reparados desta maneira, particularmente os dímeros de timina que resultam da exposição da luz ultravioleta (UV) e resíduos de guanina alquilados que foram modificados pela adição de grupos metil ou etil na posição O<sup>6</sup> do anel purínico.

A luz UV induz a formação de dímeros de pirimidina, nos quais duas pirimidinas adjacentes (como duas timinas) são unidas por um anel ciclobutano. Esta formação distorce a estrutura da cadeia de DNA e bloqueia a transcrição ou a replicação para além do local da lesão. Os dímeros de timina podem ser reparados por fotorreativação, processo no qual a energia da luz visível é usada para quebrar as ligações que formam o anel ciclobutano.

Outro exemplo de reversão da lesão, envolve danos ao DNA por reação de agentes alquilantes. Um tipo particular de lesão é a metilação de O<sup>6</sup> da guanina, resultando em O<sup>6</sup>-metilguanina, formando o pareamento entre timina e citosina. A enzima O<sup>6</sup>-metilguanintransferase remove o grupo metil não inativo da guanina e o transfere para uma cisteína interna da proteína.

### 7.2.2 Reparo por excisão

Embora eficiente, a reversão direta das lesões é limitada, devido à sua especificidade de enzima / substrato. Assim, os organismos desenvolveram mecanismos de reparo mais gerais, que removem bases lesadas do genoma e as substituem por sequências de nucleotídeos. O DNA lesado é identificado e removido como bases livres ou como nucleotídeos. A falha resultante é, então, preenchida através da síntese de uma nova fita de DNA, usando a fita complementar intacta como molde. A uracila pode ser incorporada no DNA ocasionalmente no lugar durante a replicação. A excisão da uracila é catalisada pela DNA glicosilase, a qual deixa o açúcar a desoxirribose no DNA sem a base. A falha gerada é preenchida pela DNA polimerase e selada pela ligase, promovendo a incorporação da base correta.

### 7.2.3 Reparo pós-replicação

A replicação de DNA lesado pode resultar na síntese de fitas filhas contendo falhas na região oposta à lesão na fita molde parental. Tais falhas no RNA recém sintetizado podem ser preenchidas pela recombinação com uma fita parenteral intacta ou pela reparação sujeita a erros, usando a fita parenteral como molde.

### 7.2.4 Reparo por recombinação

O reparo por recombinação homóloga é utilizado para reparar quebras duplas presentes logo após a replicação do DNA, e para a sua atuação é necessária uma segunda cópia intacta disponível. Portanto, as células empregarão esta via de reparo em fase específicas do ciclo celular. O correto posicionamento das cromátides irmãs por coenzimas provavelmente facilita a identificação da sequência homóloga. Após a identificação da sequência, a cópia dupla fita é utilizada para selar as extremidades quebradas.

### 7.2.5 Reparo SOS

O produto do gene *recA* pode estar envolvido em rotas de reparação, uma capacidade importante é a indução a resposta SOS. Os elementos reguladores são a proteína

RecA e o repressor LexA. LexA inibe a expressão dos genes SOS, e sua remoção do DNA é necessária para resposta. O repressor LexA é inativado pela sua autoclivagem, que requer a proteína RecA. A proteína RecA é o elo entre a lesão no DNA e a indução SOS. Lesões no DNA levam a vários sinais na fita simples (durante a replicação) e RecA liga-se a fita simples. A ligação ativa RecA levando a clivagem do LexA e a indução SOS.

## 8. Tópicos em genética molecular

### 8.1 A aplicabilidade das técnicas moleculares

As principais “ferramentas” utilizadas para análise de material genômico no âmbito do diagnóstico são as técnicas baseadas reação em cadeia da polimerase (PCR, do inglês, *Polymerase Chain Reaction*), southern blot e o uso de enzimas de restrição. Com o aumento do número de estudo de genomas, as sequências de genes foram exploradas para o desenvolvimento de testes diagnósticos baseados nas sequências gênicas na técnica da PCR. Técnica molecular proposta por Kary Mullis no final da década de 1980, sendo atribuído o Prêmio Nobel de Química em 1993. Esta técnica ampliou as possibilidades da análise de DNA e fez com que a biologia molecular encontrasse novas aplicações.

A metodologia requer, primeiramente, o conhecimento, pelo menos parcial, do DNA alvo para o desenvolvimento de oligonucleotídeos iniciadores (*primers*) que irão hibridizar-se especificamente com a sequência alvo. A técnica permite a amplificação de DNA *in vitro* pela variação de temperatura “assemelhando-se” a replicação do DNA, onde a polimerização das novas fitas de DNA inicia-se a partir de cada um dos iniciadores que se ligam complementarmente a fita molde e através da enzima DNA polimerase que utiliza os quatro dNTPs (Adenina, Timina, Guanina e Citosina) ocorre a amplificação do DNA alvo em cada ciclo.

Na atualidade, temos algumas variantes da PCR, a que se destaca é o sistema PCR em tempo real, promovendo um significativo avanço tecnológico no diagnóstico molecular. O sistema da PCR em tempo real é baseado no uso de corantes ou sondas fluorescentes que permitem o monitoramento do produto amplificado como, por exemplo, o *SYBR-Green I*, que se ligam as fitas duplas de DNA geradas durante a amplificação. Um outro método é o uso de uma sonda dirigida especificamente a uma região interna da sequência que se deseja amplificar, o sistema *TaqMan*. À medida que vai ocorrendo a amplificação, a sonda vai sendo hidrolisada e há a liberação de um fluorocromo que absorve energia e após hidrólise emite luz. A análise da emissão de luz é feita por um detector de sinal luminoso e um amplificador de sinal, que traçam um gráfico com a absorção obtida após cada ciclo da PCR. O ciclo em que o limite

de negatividade é ultrapassado e está diretamente relacionado à quantidade de DNA amplificado. As ampliações são coletadas aos logo dos ciclos da PCR (desnaturação, anelamento e extensão) ao invés da quantidade de alvo acumulado após um número fixo de ciclos.

A PCR em tempo real possibilita a eliminação da etapa laboriosa pós-amplificação (preparo do gel para eletroforese), convencionalmente necessária para visualização do produto amplificado. As vantagens da PCR em tempo real em relação a PCR convencional são inúmeras e incluem: velocidade, reprodutibilidade e capacidade de quantificação. Essa tecnologia, que é altamente sensível, já está sendo desenvolvida para fazer o acompanhamento de inúmeras doenças.

Com o genoma humano descrito, as sequências depositadas facilitaram os sistemas de identificação da variação do DNA e de sua correlação subsequente à saúde humana. A forma mais comum de variação genética no genoma humano é uma mudança de um par de base nitrogenada, o polimorfismo de nucleotídeo único (SNP, do inglês, *single-nucleotide polymorphism*). Neste contexto, a identificação de SNPs por métodos moleculares é utilizada como método de diagnóstico para determinar uma doença ou a demonstrar a susceptibilidade individual a uma determinada patologia.

Adicionalmente, a detecção de SNPs apresenta grande relevância no diagnóstico clínico e as técnicas descritas também podem ser utilizadas. Podemos descrever um grande avanço na medicina na área de transplantes, em função da avaliação molecular do *locus* de Antígenos Leucocitários humanos (HLA), facilitando a verificação de compatibilidade entre doador e receptor em transplantes de órgãos, principalmente nos transplante de medula óssea. Bem como, os diversos polimorfismos associados com doenças cardiovasculares e desordem de coagulação sanguínea.

Mundialmente, as doenças trombóticas constituem um problema de saúde de origem multifatorial e multigênica. As trombozes são caracterizadas pela formação aguda de trombos em veias e artérias. Dentre as doenças cardiovasculares, a trombose venosa destaca-se pela associação entre fatores de risco adquiridos, como imobilização prolongada, cirurgias, fraturas, gestação e etc. e fatores genéticos, mutações nos genes da proteína C, Proteína S e antitrombina.

A resistência hereditária à proteína C ativa (RPCA) tem sido identificada como principal causa da maioria dos casos de trombose venosa, 95% dos casos de RPCA estão associados a mutação pontual G1691A no éxon 10 do gene do fator V da coagulação, o fator V Leiden (FVL). Esse polimorfismo não-sinonímico causa a perda da clivagem do fator V, gerando um quadro de hipercoagulopatia e, conseqüentemente aumentando o risco de trombose venosa. Em indivíduos homozigotos FVL, o risco de trombose venosa é 50 a 100 vezes maior que em pacientes homozigotos normais, enquanto em pacientes heterozigotos o risco é de 5 a 10 vezes. Baseado na necessidade de avaliação e acompanhamento de pacientes com casos de trombose venosa e prevenção de seus respectivos familiares, o diagnóstico clínico pode ser feito utilizando a

discriminação alélica do fator V da coagulação (G1691A) por PCR em tempo real e / ou PCR-RFLP (PCR-Restriction Fragment Length Polymorphism).

A fibrose cística (FC) é um distúrbio autossômico recessivo e é conhecida como mucoviscidose. Apresenta fenótipo diversificado, com diferentes sistemas orgânicos sendo afetados em graus variados. A grande maioria dos portadores de (FC) sofre de insuficiência pancreática devido a ductos entupidos das glândulas secretoras do pâncreas. Esse bloqueio impede a secreção de enzimas digestivas no duodeno. Além, dos ductos pancreáticos, as passagens dos pulmões do fígado, das vias reprodutivas masculinas e do intestino geralmente estão bloqueadas. Também, ocorre o dano pulmonar resultante do acúmulo de muco e infecções bacterianas das vias aéreas, causando a principal causa de sofrimento contínuo e até morte, por volta dos 30 anos de idade.

A base genética da doença foi estabelecida quando um gene clonado, depois chamado de gene regulador transmembranar de condutância na fibrose cística (CFTR), foi encontrado com a mesma mutação em várias pessoas com FC. Esta mutação é uma deleção de três pares de base do códon da fenilalanina (F508del;  $\Delta$ F508) que codifica o aminoácido na posição 508 da proteína CFTR. A proteína CFTR é um canal multifuncional e multidomínio de íons cloreto das células epiteliais, que não só regula a concentração intracelular de íons cloreto, conduzindo-os para fora da célula, mas regula outros canais de proteínas transportadoras. A falta de uma correlação significativa entre os genótipos CFTR e os fenótipos de FC sugere que outros genes (genes modificadores) e / ou fatores ambientais determinem o resultado fenotípico.

Estudos populacionais, com grande número de fibrocísticos, comprovam a relação genótipo-fenótipo principalmente para as manifestações pancreáticas, mas muito pouco para as pulmonares. Portanto, as características genéticas da relação genótipo-fenótipo nas manifestações pulmonares permanecem pouco esclarecidas. Sabe-se que mesmo indivíduos homozigotos para a mutação de maior prevalência ( $\Delta$ F508) apresentam grande variabilidade no comprometimento e na evolução da doença pulmonar. Embora influências ambientais possam interferir nas manifestações clínicas pulmonares, tem sido descrita a possibilidade de suma variação genética adicional, como a presença de genes modificadores, contribuindo para a expressão clínica final em cada paciente. Alguns estudos têm relatado que polimorfismos em genes que não o CFTR podem modificar a gravidade da doença pulmonar na FC.

Genes modificadores fora do gene CFTR podem influenciar a gravidade do fenótipo dos fibrocísticos através de vários mecanismos:

1) podem modular o fenótipo alternando a condução de cloro; podem regular o splicing e a expressão do gene CFTR;

2) podem modular a susceptibilidade à infecção por bactérias e a resposta inflamatória nos pulmões. Além disso, o quadro pulmonar dos fibrocísticos pode ser modificado por genes associados com clearance mucociliar e com os danos e reparos do tecido epitelial.

O conceito de múltiplos modificadores genéticos na doença mendeliana, como na FC, é diferente do conceito de múltiplos variantes genéticos na doença não-mendeliana, como na asma. Nas doenças genéticas complexas, como na asma, variantes genéticas múltiplas interagem entre si e com o ambiente, causando a doença. Ao contrário, a FC, por ser uma doença mendeliana, é causada por mutações no gene *CFTR*, e há variações genéticas não ligadas ao gene *CFTR* que podem ser desfavoráveis ou benéficas e que modificam a gravidade do fenótipo em conjunto com fatores ambientais. De fato, polimorfismos genéticos, com ou sem efeitos em sujeitos saudáveis, podem ser modificadores na FC.

A identificação da consequência da ação de genes modificadores permitirá um melhor entendimento dos aspectos fisiopatológicos, da correlação genótipo-fenótipo e maximizará o tratamento da FC. Estudos descrevem a associação entre os polimorfismos do éxon 1 (códon 52, 54 e 57) e a região promotora (haplótipos HY, LY e LX) do gene *MBL2*, o polimorfismo T869C no gene *TGF-β1* e o polimorfismo C-159T no gene *CD14* com a gravidade do quadro pulmonar de pacientes com FC. As análises podem ser feitas por PCR-RFLP e em tempo real.

O estudo de variações individuais na capacidade de metabolização de xenobióticos já identificou inúmeros polimorfismos genéticos, sendo a maioria destes de relevância clínica. Os polimorfismos metabólicos que têm sido associados de forma mais consistente ao aumento do risco de câncer incluem as enzimas da superfamília citocromo P450 (CYP450), as glutatona S-transferase (GST) e as N-acetiltransferase (NAT).

A maioria dos agentes carcinogênicos requer ativação metabólica antes de se ligarem às biomoléculas (DNA, RNA e proteínas) e as variações nos processos de ativação e detoxificação de compostos químicos desempenham papel crucial na tumorigênese ambiental. Assim, vários polimorfismos de genes responsáveis pela codificação de enzimas envolvidas na biotransformação de carcinógenos também se relacionam com o desenvolvimento do câncer. Entre os fatores envolvidos com a carcinogênese têm-se características genéticas relacionadas com o gênero, idade e síndromes de susceptibilidade familiar ao câncer.

Considerada a principal enzima detoxificante da fase II, a GST desempenha um papel fisiológico na iniciação da detoxificação de potenciais agentes alquilantes, incluindo xenobióticos tóxicos e produtos reativos de processos intracelulares, como a peroxidação de lipídios. As GSTs formam uma família multigênica subdivididas com base em seu ponto isoelétrico e na sequência de aminoácidos:  $\alpha$ ,  $\mu$ ,  $\pi$  e  $\theta$ . Cada uma contendo produtos de diferentes *locus* gênico. Recentes estudos demonstram que os genótipos nulos de *GSTM1* e *GSTT1* (polimórfico), separadamente ou combinados, são fatores genéticos de alto risco para o desenvolvimento de tumores. O genótipo *GSTM1* nulo já foi associado ao desenvolvimento de diversos tipos de cânceres, como o de pulmão; bexiga; mama e de cabeça e pescoço. O câncer de pulmão é o mais

frequentemente associado à ausência do gene *GSTM1*, por este gene se relacionar com a detoxificação de certos epóxidos dos HPAs (Hidrocarbonetos Policíclicos Aromáticos) encontrados na fumaça do cigarro e em outros produtos de combustão.

Vários relatos na literatura sugerem que fatores inerentes ao próprio indivíduo, incluindo os polimorfismos genéticos podem conferir diferenças individuais na ocorrência do câncer, como também a outras doenças complexas, como insuficiência renal crônica, diabetes e doenças cardíacas.

Nos últimos anos, os “testes de DNA” ganharam o conhecimento do público leigo. Diversos casos criminais foram solucionados diante do DNA fingerprinting. Um caso que teve grande repercussão mundial e que determinou a popularização dos testes moleculares foi o caso O. J. Simpson nos EUA. Do ponto de vista social, a determinação da identidade genética pelo DNA constitui um dos produtos mais revolucionários da moderna biologia molecular. Em menos de 20 anos ela tornou-se uma ferramenta indispensável em investigação criminal.

A determinação da identidade genética pelo DNA é uma técnica muito superior à todas as técnicas preexistentes de medicina forense, inclusive às clássicas impressões digitais. O DNA pode ser encontrado em todos os fluidos e tecidos biológicos humanos. Além disso, os estudos dos polimorfismos de DNA nos permite construir um perfil genético absolutamente indivíduo-específico. Ao contrário das proteínas, quantidades ínfimas de DNA podem ser amplificadas bilhões de vezes através da PCR.

Finalmente, características moldadas ao longo da história evolutiva dos seres vivos adaptaram o DNA para ser uma molécula informacional com baixíssima reatividade química e grande resistência à degradação. Esta robustez da molécula de DNA, conjuntamente com o fato de que ele contém informação necessária fazem com que o DNA seja ideal como uma fonte de identificação resistente à passagem do tempo e às agressões ambientais frequentemente encontrada em cenas de crime.

Assim, a determinação de identidade genética pelo DNA pode ser usada para demonstrar a culpabilidade dos criminosos, exonerar os inocentes, identificar corpos e restos humanos em desastres aéreos e campos de batalha, determinar paternidade com confiabilidade praticamente absoluta, elucidar trocas de bebês em berçários e detectar substituições e erros de rotulação em laboratórios de patologia clínica.

A evolução da determinação de identidade genética até se tornar hoje, dentro do complexo ambiente do sistema criminal da justiça, um procedimento robusto, sensível, muito informativo e altamente objetivo. É importante ressaltarmos o histórico das técnicas utilizadas e analisarmos a evolução dos métodos durante os últimos anos e discutir os problemas potenciais que podem advir como consequência da rápida aplicação da tecnologia do DNA recombinante na sociedade humana.

## Referências

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular Biology of the Cell**, 4th Edition. Garland Science, New York, NY, 2002.

COOPER, G.M. **A célula – uma abordagem molecular**, 2ª Edição. Editora Artmed, Porto Alegre, RS, 2005.

LEWIN, B. **Genes VIII**. Editora Pearson Prentice Hall, New Jersey, 2004.

NELSON, D.L. & COX, M.M. **Lehninger Principles of Biochemistry**, 5a. ed. Ed.Worth, 2009.

LEHNINGER, A. L. **Princípios de Bioquímica**, 4ª Edição. Editora Sarvier, São Paulo, SP, 2007.

WATSON JD, BAKER TA, BELL SP, GANN A, LEVINE M, LOSICK R. **Biologia Molecular do Gene**. 5ª ed. Editora Artmed, Porto Alegre, RS , 2006

ZAHA, A. PASSAGLIA, L.P., BUNSELMAYER, H.F.(Org). **Biologia Molecular Básica**. 4ª edição. Editora Artmed, Porto Alegre, 2012.

MIR, L. **Genômica**. Editora Atheneu. São Paulo, 2004.

NUSSBAUM, R.L.; MCINNES, R.R.; WILLARD, H.F. **Genética Médica**. Editora Guanabara Koogan. Rio de Janeiro, 2002.

NUSSBAUM, R.L.; McINNES, R.R.; WILLARD, H.F. **Genética Médica**. Editora Elsevier. Rio de Janeiro, 2008.

PASTERNAK, J.J. **Genética Molecular Humana: Mecanismos das Doenças Hereditárias**. Editora Manole. São Paulo, 2002.

GRIFFITHS, A.J.F; GELBART, W.M.; MILLER, J.H.; LEWONTIN, R.C. **Uma Introdução à Genética**. Editora Guanabara Koogan, Rio de Janeiro, 2009.

JORDE, B.L.; CAREY, J.C.; BAMSHAD, M.J; WHITE, R.L. **Genética Médica**. Editora Elsevier. Rio de Janeiro, 2004.