

# Ciências Biológicas

## Biologia Molecular

Vânia Marilande Ceccatto



Geografia



História



Educação Física



Química



Ciências Biológicas



Artes Plásticas



Computação



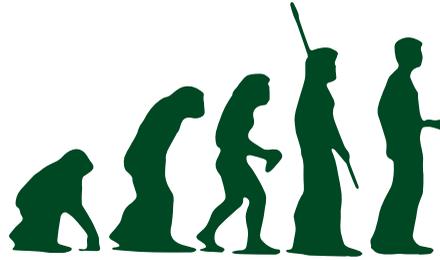
Física



Matemática



Pedagogia



# Ciências Biológicas

## Biologia Molecular

Vânia Marilande Ceccatto

2ª edição  
Fortaleza - Ceará



2015



Geografia



História



Educação  
Física



Pedagogia



Química



Ciências  
Biológicas



Artes  
Plásticas



Computação



Física



Matemática

Copyright © 2015. Todos os direitos reservados desta edição à UAB/UECE. Nenhuma parte deste material poderá ser reproduzida, transmitida e gravada, por qualquer meio eletrônico, por fotocópia e outros, sem a prévia autorização, por escrito, dos autores.

Editora Filiada à



<b>Presidenta da República</b> Dilma Vana Rousseff	<b>Conselho Editorial</b> Antônio Luciano Pontes
<b>Ministro da Educação</b> Renato Janine Ribeiro	Eduardo Diatary Bezerra de Menezes
<b>Presidente da CAPES</b> Carlos Afonso Nobre	Emanuel Ângelo da Rocha Fragoso
<b>Diretor de Educação a Distância da CAPES</b> Jean Marc Georges Mutzig	Francisco Horácio da Silva Frota
<b>Governador do Estado do Ceará</b> Camilo Sobreira de Santana	Francisco Josênio Camelo Parente
<b>Reitor da Universidade Estadual do Ceará</b> José Jackson Coelho Sampaio	Gisafran Nazareno Mota Jucá
<b>Vice-Reitor</b> Hidelbrando dos Santos Soares	José Ferreira Nunes
<b>Pró-Reitora de Graduação</b> Marcília Chagas Barreto	Liduína Farias Almeida da Costa
<b>Coordenador da SATE e UAB/UECE</b> Francisco Fábio Castelo Branco	Lucili Grangeiro Cortez
<b>Coordenadora Adjunta UAB/UECE</b> Eloísa Maia Vidal	Luiz Cruz Lima
<b>Direção do CCS/UECE</b> Gláucia Posso Lima	Manfredo Ramos
<b>Coordenadora da Licenciatura em Ciências Biológicas</b> Germana Costa Paixão	Marcelo Gurgel Carlos da Silva
<b>Coordenadora de Tutoria e Docência em Ciências Biológicas</b> Roselita Maria de Souza Mendes	Marcony Silva Cunha
<b>Editor da EdUECE</b> Erasmus Miessa Ruiz	Maria do Socorro Ferreira Osterne
<b>Coordenadora Editorial</b> Rocylânia Isídio de Oliveira	Maria Salette Bessa Jorge
<b>Projeto Gráfico e Capa</b> Roberto Santos	Sílvia Maria Nóbrega-Therrien
<b>Diagramador</b> Francisco José da Silva Saraiva	<b>Conselho Consultivo</b>
<b>Revisora Ortográfica</b> Fernanda Rodrigues Ribeiro Freitas	Antônio Torres Montenegro (UFPE)
	Eliane P. Zamith Brito (FGV)
	Homero Santiago (USP)
	Ieda Maria Alves (USP)
	Manuel Domingos Neto (UFF)
	Maria do Socorro Silva Aragão (UFC)
	Maria Lírida Callou de Araújo e Mendonça (UNIFOR)
	Pierre Salama (Universidade de Paris VIII)
	Romeu Gomes (FIOCRUZ)
	Túlio Batista Franco (UFF)

Dados Internacionais de Catalogação na Publicação  
Sistema de Bibliotecas  
Biblioteca Central Prof. Antônio Martins Filho  
Thelma Marylanda Silva – CRB-3 / 623  
Bibliotecário

C387b Ceccato, Vânia Marilande  
Biologia molecular / Vânia Marilande Ceccato. 2. ed.  
Fortaleza : EdUECE, 2015.  
139 p. : il. ; 20,0cm x 25,5cm . (Ciências Biológicas)  
Inclui bibliografia.  
ISBN: 978-85-7826-342-3  
1. Biologia molecular. I. Ceccato, Vânia Marilande.  
II. Título.

CDD 574.87

Editora da Universidade Estadual do Ceará – EdUECE  
Av. Dr. Silas Munguba, 1700 – Campus do Itaperi – Reitoria – Fortaleza – Ceará  
CEP: 60714-903 – Fone: (85) 3101-9893  
Internet: www.uece.br – E-mail: eduece@uece.br  
Secretaria de Apoio às Tecnologias Educacionais  
Fone: (85) 3101-9962

# Sumário

<b>Apresentação</b> .....	<b>5</b>
<b>Parte – Conhecendo o gene e o genoma</b> .....	<b>7</b>
<b>Capítulo 1 – Considerações iniciais e histórico</b> .....	<b>9</b>
1. Aspectos gerais .....	9
2. Breve histórico .....	11
<b>Capítulo 2 – Ácidos nucleicos</b> .....	<b>14</b>
1. Ácidos nucleicos .....	14
<b>Capítulo 3 – Núcleo eucariótico e cromossomos</b> .....	<b>18</b>
1. Genes e genomas .....	21
<b>Parte 2 – Mecanismos genéticos básicos</b> .....	<b>33</b>
<b>Capítulo 4 – Considerações iniciais e histórico</b> .....	<b>35</b>
1. Aspectos gerais .....	35
2. Breve histórico .....	36
<b>Capítulo 5 – Mecanismos genéticos básicos</b> .....	<b>38</b>
1. Replicação .....	38
2. Transcrição .....	40
3. Tradução .....	46
<b>Parte 3 – Polimorfismo gênico e evolução molecular</b> .....	<b>57</b>
<b>Capítulo 6 – Considerações iniciais e histórico</b> .....	<b>59</b>
1. Aspectos gerais .....	59
2. Breve histórico .....	60
<b>Capítulo 7 – Variação gênica</b> .....	<b>62</b>
1. Bactérias, genes e plasmídeos .....	62
2. Genomas extra nucleares: mitocôndrias e cloroplastos.....	63
3. Polimorfismo e marcadores moleculares .....	64
4. Alterações cromossômicas .....	66
5. O valor C em discussão .....	67
6. Elementos de transposição.....	70
7. Mutações.....	72
8. Recombinação gênica .....	76
<b>Parte 4 – Expressão gênica</b> .....	<b>87</b>
<b>Capítulo 8 – Considerações iniciais e histórico</b> .....	<b>89</b>
1. Aspectos gerais .....	89
2. Breve histórico .....	89

<b>Capítulo 9 – Expressão gênica.....</b>	<b>92</b>
1. Expressão e regulação em procariotos .....	92
2. Cromatina celular e epigenética .....	95
3. Em eucariotos: regulação da transcrição.....	97
4. Estabilidade do RNA como evento regulador .....	99
5. O desenvolvimento e seu controle.....	100
<b>Parte 5 – DNA recombinante .....</b>	<b>107</b>
<b>Capítulo 10 – Considerações iniciais e histórico .....</b>	<b>109</b>
1. Aspectos gerais.....	109
2. Breve histórico .....	110
<b>Capítulo 11 – Técnicas gerais .....</b>	<b>113</b>
1. Técnicas associadas: eletroforese e enzimas de restrição.....	113
<b>Capítulo 12 – Plasmídios: vilões e heróis .....</b>	<b>119</b>
1. 1ª Aplicação da técnica: bibliotecas genômicas e bibliotecas de cDNA.....	122
2. 2ª aplicação: transgenia .....	123
<b>Capítulo 13 – Reação em cadeia da polimerase: a revolução dentro da revolução .....</b>	<b>126</b>
1. Aplicação da técnica: <i>fingerprint</i> de DNA.....	129
<b>Sobre a autora .....</b>	<b>139</b>

# Apresentação

Estamos envolvidos numa revolução científica iniciada no século XX. É chamada de Revolução Biológica. Essa nova Biologia, segundo Waterman, T. H. em *“Revolution for biology”*, 1962, sucedeu a era atômica, resultando numa nova perspectiva humanística e ecológica. Essa revolução refere-se especialmente à integração das Ciências Biológicas com outras áreas e outros ramos científicos, nunca antes ocorridos. Assim, surgiram os “novos biólogos”: bioquímicos, biofísicos, bioestatísticos, biomatemáticos, bioinformatas, biólogos moleculares, filósofos, bioéticos.

Graças a técnicas espetaculares, alcançaram-se patamares nunca sonhados nos estudos moleculares. Na área de saúde, temos vacinas, medicamentos, tratamentos, diagnósticos e, sobretudo, conhecimentos biológicos básicos acumulados. As soluções práticas aplicadas envolvem um sem número de técnicas: a transgenia, células tronco, clonagem, o DNA forense etc. Na área gênica, em geral, nos regozijamos com avanços nas ômicas (incluindo a genômica e outras). A genética molecular, de plantas, de animais, de microrganismos, do ambiente e muitas outras áreas e subáreas tornam-se cada dia mais fascinantes para os profissionais envolvidos e para o público em geral. Por outro lado, ainda não faltam desafios. A vida humana (e, por conseguinte, toda a vida) é frágil e problemática. Estamos num planeta em plena crise ambiental. A cura para várias doenças como câncer (genericamente falando) e doenças crônico-degenerativas, doenças genéticas e a própria terapia gênica, apesar das perspectivas promissoras, ainda não avançou o esperado para este século.

Seria necessário esclarecer que a Biologia Molecular não se caracteriza como ciência em si própria, mas sim como ferramenta do profissional, disponível para esclarecer, num determinado nível, suas perguntas. A Biologia Molecular é sim, uma disciplina essencial, a qual se apropriou de várias outras disciplinas e técnicas para formar seu corpo. Esse corpo de conhecimento é praticamente ilimitado e, dessa forma, é difícil de ser definido. De uma forma muito simplificada, chamaríamos a Bioquímica e a Genética para compor uma base com  $n$  braços, da microbiologia, à imunologia, à citologia e muitas outras.

Buscamos uma abordagem mais atual possível para a disciplina. Assim, o gene e o genoma estão sempre em evidência. As ações bioquímicas estão implícitas nos conceitos, portanto, a Bioquímica é visceral para o nosso entendimento. Tenha sempre seu material de Bioquímica a mão para tirar dúvidas. Procure rever, especialmente, macromoléculas em geral, especialmente a estrutura de proteínas e mecanismo de ação das enzimas. Evidentemente não é possível a cober-

tura completa do conteúdo nem foi essa nossa pretensão, mas buscamos aqui um guia que possa incrementar e direcionar seus estudos. Outra preocupação é entender como os conhecimentos avançaram historicamente até o momento presente. Uma inspiradora frase de Carl Seagan diz que a Biologia está mais próxima da História do que da Física. Dessa forma, só é possível compreender o avanço da Biologia atual através do entendimento do passado.

Essa nova revolução também precisa de limites éticos, englobando perspectivas filosóficas inéditas para o profissional de biologia. Com o genoma, inúmeros questionamentos ainda não foram feitos, em termos de que chegará um momento onde não sabemos como será resguardada a individualidade pessoal, em razão das necessidades coletivas. No Brasil, a questão dos transgênicos é de suma importância, sendo um dos países de maior produção de *commodities* transgênicas do mundo. Células-tronco, DNA *fingerprint*, DNA *ancient*, questões forenses, farmacogenômica, são assuntos de interesse da academia e também do cidadão comum, mantendo um interesse contínuo pulsante na mídia.

O Brasil e o mundo precisam de biólogos e, mais que isso, de profissionais esclarecidos, de excelência, compromissados e ruidosos. Bons estudos!

**A autora**

**Parte**

**1**

**Conhecendo o  
gene e o genoma**



# Considerações iniciais e histórico

## Objetivo

- Relacionar a estrutura dos ácidos nucleicos aos conceitos de gene e genoma, além da integração destes à estrutura geral da célula.

## 1. Aspectos gerais

A sequência de DNA (Figura 1) é matéria prima para a **evolução**. Veja que uma nova sequência inédita de DNA não aparece naturalmente sendo formada na célula. Não surgem sequências randômicas novas, nenhum gene é totalmente novo. O DNA só surge a partir de outro molde (fita simples) de DNA predecessor. Então, de onde vêm as inovações? Na manutenção e na cópia da informação genética, ocorrem acidentes e erros aleatórios, alterando a sequência original dos nucleotídeos, criando **mutações**. Na divisão celular, essas mutações são passadas para a próxima geração de células. Numa competição com recursos limitados, é possível que estes mutantes sejam eliminados. Em raras condições ambientais especiais, esses mutantes podem permanecer, mantendo estas mutações e mais raramente ainda, podem ter uma vantagem evolutiva. Essas tentativas e erros aleatórios geram **diversidade**, que possibilita a manutenção das espécies.

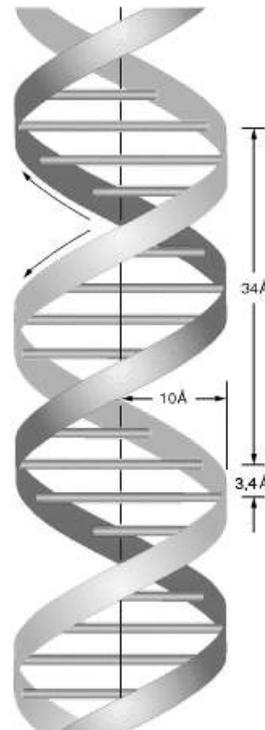


Figura 1 - Modelo simplificado da dupla fita de DNA. Os bastões mostram os pares de bases e as fitas mostram o esqueleto açúcar fosfato das duas fitas antiparalelas. As medidas são mostradas em Angstroms ( $1\text{Å} = 0,1\text{ nm}$ ).

Fonte: Griffiths et al. (1999).

**Evolução:** fenômeno biológico de diversificação genética, idealmente medida pela especiação. Está intimamente ligado à frequência dos alelos nas populações.

**Mutações:** alterações permanentes e transmissíveis na sequência ou composição de DNA ou de um cromossomo.

**Transcrição reversa:** produção de DNA a partir de fita simples de RNA, feito pela enzima transcriptase reversa. A técnica de laboratório que envolve a transcrição reversa é a produção de cDNAs (DNA complementar).

No contexto usado no texto, modelo biológico é um organismo estudado extensivamente, gerando dados que podem ser extrapolados pra outros organismos, geralmente para o mesmo grupo caracterizado taxonomicamente, ou servindo para explicar processos bioquímicos universais. São organismos que possuem linhagens homogêneas, geneticamente bem conhecidas, de fácil cruzamento, produtoras de alta prole, de facilidade de manuseio e manutenção barata.

As células **procarióticas** (células que não apresentam um núcleo distinto) possuem, em geral, um cromossomo formando um círculo fechado. A bactéria *Escherichia coli* é um organismo modelo para o estudo bioquímico e considera-se também como modelo procarionto em geral. Como **eubactéria**, em formato de bastão, vivendo no intestino humano e de outros vertebrados, pode crescer e reproduzir-se facilmente num meio simples com nutrientes, em placa de cultura. *E. coli* luta em condições químicas variáveis. Sua única molécula circular de DNA (tirando os plasmídeos) possui cerca de 4.639.221 pares de bases (linhagem k12), codificando cerca de 4.300 tipos diferentes de proteínas. Temos mais conhecimento sobre a *E. coli* do que sobre qualquer outro organismo vivo.

As complexas **células eucarióticas** possuem genomas maiores, formando organismos multicelulares, com células diferenciadas e altamente especializadas, as quais expressam diferentes conjuntos de RNA e de proteínas. Assim, o hepatócito é uma célula especializada com um conjunto único de genes expressos diferencialmente. Adipócitos, osteoblastos, eritrócitos, células epidérmicas, a princípio, parecem muito diferentes morfológicamente. Entretanto seu genoma é igual ao das outras células do organismo. A sua posição na estrutura multicelular é espantosamente bem regulada, coordenada não somente com as células vizinhas, mas em harmonia com sua idade, estado ambiental em que se encontra e funções a qual se destina, desde seu desenvolvimento embrionário. A assustadora complexidade eucariótica, tanto genética quanto fisiológica, exige o uso de modelo biológico, como a levedura *Saccharomyces cerevisiae*, serve como modelo **eucariótico** mínimo. *S. cerevisiae* é um fungo unicelular e cresce quase como uma bactéria, com toda complexidade de um eucarioto. *Arabidopsis thaliana* é um pequeno arbusto escolhido pelos biólogos como planta-modelo. Para as células animais, o verme nematódeo *Caenorhabditis elegans*, a mosca das frutas, *Drosophila melanogaster*, o rato *Mus musculus*, o camundongo *Rattus norvegicus* (Figura 2) são a escolha ideal para a pesquisa. Todos possuem genoma conhecido.



Figura 2 - Rato (*Rattus norvegicus* var. albino) caminhando em sistema experimental de comportamento.

Fonte: [http://www.unicamp.br/unicamp/unicamp\\_hoje/ju/agosto2004/imagens/260pag5e.jpg](http://www.unicamp.br/unicamp/unicamp_hoje/ju/agosto2004/imagens/260pag5e.jpg)

O genoma eucariótico, por sua vez, contém DNA não codificante, que, dependendo do organismo, pode ocorrer na ordem de milhares de vezes mais abundante que o genoma procaríoto. Cerca de 98% do genoma humano (11% em *E. coli*) é não codificante. Possui cerca de 3,5 bilhões de pares de bases, cerca de mil vezes mais nucleotídeos que o de *E. coli*. É possível, aliás, muito provável, que muito desse DNA não codificante possua funções regulatórias, além de outras, desconhecidas, a serem descobertas nos próximos anos.

O produto do gene é o RNA e de forma mais direta: a proteína. A síntese proteica não é dirigida pelo genoma diretamente, mas utiliza o RNA como intermediário molecular durante o processo chamado de transcrição. Na transcrição, uma das fitas do DNA serve como molde para que a enzima polimerizadora (RNA polimerase) produza a fita simples de RNA. Os vários RNAs serão, então, usados na tradução ou síntese proteica. Esse princípio fundamental foi chamado de “dogma central” da Genética ou da Biologia Molecular. O conceito de “dogma” ainda pode ser usado com reservas, pois muitas descobertas posteriores têm desafiado essa fundamentação, como a transcrição reversa. Esses conceitos serão revistos na Unidade 2.

## 2. Breve histórico

Apesar da estrutura clássica Watson-Crick do DNA ser conhecida desde 1953, os primeiros passos para tornar possível esse conhecimento foram gerados cerca de um século e meio antes, com os trabalhos de Mendel. Os genes, em 1865, seriam chamados de “fatores particulados”. Esses “fatores” passariam, sem serem modificados, do progenitor para a progênie. Em 1903, os cromossomos são visíveis nos microscópios da época e são reconhecidos como unidades hereditárias, vistas na divisão celular. Nos organismos diploides, que possuem dois conjuntos de cromossomos, uma cópia é herdada de cada progenitor. Dessa forma, os genes também seriam transmitidos para a próxima geração. Na década seguinte observou-se que os genes estão nos cromossomos e que o cromossomo contém arranjos lineares de genes.

De que forma isto aconteceu? Uma objeção séria ao trabalho de Mendel era a ausência de fatos concretos sobre o comportamento dos cromossomos nas fases mitótica e meiótica. Quando as Leis de Mendel foram redescobertas, em 1900, esses fatos já haviam sido confirmados, sendo utilizados, em 1903, por Walter S. Sutton, biólogo americano. Em seu trabalho, considerado clássico: “*The Chromosomes in Heredity*” - Os Cromossomos na Hereditariedade (publicado no periódico *Biological Bulletin*, n.4). Sutton destacou a informação de que o conjunto cromossômico é diploide e que, durante a meiose, cada gameta recebe apenas um cromossomo de cada par de homólogos. Dessa forma, os dados de Mendel podiam ser explicados concretamente; o gene

fazia parte do cromossomo. O artigo de Sutton possibilitou a reunião de conceitos que, na época, pareciam independentes, mas sabemos que hoje estão intrinsecamente ligados.

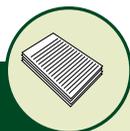
Quase na mesma época, em 1915, Thomas Hunt Morgan e seus alunos Alfred H. Sturtevant, Hermann J. Muller e Calvin B. Bridges, chamados de o “Grupo da Columbia University”, publicaram o livro *The Mechanism of Mendelian Heredity* - O Mecanismo da Herança Mendeliana, em que são relatadas, pela primeira vez, as bases cromossômicas da hereditariedade. Esse grupo designou localizações para mais de 85 genes mutantes em *Drosophila*. Os genes nos cromossomos foram localizados de forma estruturalmente linear, nunca ramificada. As distâncias relativas foram calculadas com bases em frequências obtidas em cruzamentos. As distâncias são medidas em unidades de mapa ou centimorgans. Pela primeira vez, um mapa genético foi feito.

A definição do DNA ficou, durante um tempo, “oculta” num assim chamado “princípio ou agente transformante” caracterizado através de um experimento aparentemente inexplicável, efetuado pelo microbiologista inglês Frederick Griffith, em 1928, trabalhando com duas linhagens de *Streptococcus pneumoniae*, as bactérias causadoras da pneumonia. Uma dessas linhagens era virulenta e outra, não virulenta. A linhagem não virulenta tornava-se virulenta (patogênica) quando era misturada a um extrato de bactérias virulentas mortas pelo calor. A informação genética da bactéria virulenta morta era lançada ao meio e atravessava a parede celular, incorporando-se ao genoma da bactéria não virulenta, transformando-a. Naquela época, os pesquisadores acreditavam que os genes (melhor dizendo, o “princípio transformante”) eram proteínas.

Qual a surpresa quando os biólogos Oswald Theodore Avery, Colin Munro MacLeod e Maclyn McCarthy mostraram que, após purificarem o mesmo extrato de bactérias virulentas mortas pelo calor, esse princípio era destruído pela enzima desoxirribonuclease, que degrada as moléculas de DNA.

Uma das ideias mais sensatas postuladas para o conceito de gene foi postulada por Archibald Edwardt Garrod em 1909, de que os genes afetavam a síntese de enzimas. Na época, os geneticistas mendelianos trabalhavam com organismos eucarióticos como milho, camundongo e *Drosophila*, inadequados para caracterizar a relação gene – proteína. Micro-organismos são organismos mais adequados para esse trabalho. Pesquisas com *Neurospora*, efetuadas por George W. Beadle e Edward Tatum, em 1940, geraram evidências fortes que sustentavam essa antiga hipótese: um gene – uma enzima. Essa evidência levou à hipótese um gene – uma proteína (mais adequado: **um gene – um polipeptídio**). A definição *funcional* que utilizamos no dia a dia, considera o gene como sequência de DNA que é capaz de produzir uma (ou mais de uma) cadeia polipeptídica, ou um RNA.

## Texto complementar



A Lei Nº 11.794, de 8 de outubro de 2008, chamada "Lei Arouca" regulamenta o uso de animais em pesquisa no Brasil. De acordo com a lei:

“Art. 1º A criação e a utilização de animais em atividades de ensino e pesquisa científica, em todo o território nacional, obedecem aos critérios estabelecidos nesta Lei.

§ 1º A utilização de animais em atividades educacionais fica restrita a:

I – estabelecimentos de ensino superior;

II – estabelecimentos de educação profissional técnica de nível médio da área biomédica.

§ 2º São consideradas como atividades de pesquisa científica todas aquelas relacionadas com ciência básica, ciência aplicada, desenvolvimento tecnológico, produção e controle da qualidade de drogas, medicamentos, alimentos, imunobiológicos, instrumentos, ou quaisquer outros testados em animais, conforme definido em regulamento próprio.

§ 3º Não são consideradas como atividades de pesquisa as práticas zootécnicas relacionadas à agropecuária.

Art. 2º O disposto nesta Lei aplica-se aos animais das espécies classificadas como filo Chordata, subfilo Vertebrata, observada a legislação ambiental.”

A Lei também cria o Conselho Nacional de Controle de Experimentação Animal – CONCEA, ao qual compete:

“I – formular e zelar pelo cumprimento das normas relativas à utilização humanitária de animais com finalidade de ensino e pesquisa científica;

II – credenciar instituições para criação ou utilização de animais em ensino e pesquisa científica;

III – monitorar e avaliar a introdução de técnicas alternativas que substituam a utilização de animais em ensino e pesquisa;

IV – estabelecer e rever, periodicamente, as normas para uso e cuidados com animais para ensino e pesquisa, em consonância com as convenções internacionais das quais o Brasil seja signatário;

V – estabelecer e rever, periodicamente, normas técnicas para instalação e funcionamento de centros de criação, de biotérios e de laboratórios de experimentação animal, bem como sobre as condições de trabalho em tais instalações;

VI – estabelecer e rever, periodicamente, normas para credenciamento de instituições que criem ou utilizem animais para ensino e pesquisa;

VII – manter cadastro atualizado dos procedimentos de ensino e pesquisa realizados ou em andamento no País, assim como dos pesquisadores, a partir de informações remetidas pelas Comissões de Ética no Uso de Animais - CEUAs, de que trata o art. 8º desta Lei;

VIII – apreciar e decidir recursos interpostos contra decisões das CEUAs;

IX – elaborar e submeter ao Ministro de Estado da Ciência e Tecnologia, para aprovação, o seu regimento interno;

X – assessorar o Poder Executivo a respeito das atividades de ensino e pesquisa tratadas nesta Lei.”

Conheça a Comissão de Ética para o Uso de Animais - CEUA/UECE no site: <http://www.uece.br/ceua>.

Fonte: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2007-2010/2008/lei/11794.htm](http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/11794.htm)

# Ácidos nucleicos

**Auto complementação:**

ligação entre pares de bases nitrogenadas complementares, sempre de fita simples.

**Esterioquímica:** ramo da

Química que estuda as relações espaciais entre os átomos de uma molécula.

**Tautômero:** isômeros ocorrem quando duas ou mais substâncias apresentam fórmulas moleculares idênticas, mas que diferem em suas fórmulas estruturais. A tautomeria é um caso particular de isomeria funcional, em que dois isômeros ficam em equilíbrio químico dinâmico.

**Ribozima:** são moléculas de RNA, que funcionam através de sua ligação à outra molécula de RNA alvo e, via complementação, fazem um corte específico na ponte fosfodiéster. A reação de *auto-splicing* ocorre por uma transesterificação, realizadas pelos introns do tipo I. O conceito pode ser aplicado aos tipos de RNA que apresentam atividade na tradução, como o RNAt.

## 1. Ácidos nucleicos

### 1.1 O DNA é um polidesoxiribonucleotídeo

Ácidos nucleicos (DNA e RNA), as proteínas e os polissacarídeos são polímeros produzidos pela adição repetitiva de subunidades monoméricas (monômeros) a uma das extremidades da cadeia em crescimento. A base molecular do DNA é o **nucleotídeo**, mais precisamente um desoxirribonucleotídeo. O nucleotídeo consiste em uma 2'-desoxirribose ladeada pelo lado direito (carbono 1') por uma base nitrogenada e do lado esquerdo (carbono 5'), por um grupo fosfato. A primeira ligação é chamada N-glicosídica, e a segunda, 5'-fosfo-monoéster. A ligação fosfodiéster consiste no grupo fosforil entre dois nucleotídeos adjacentes. O esqueleto de cada fita consiste em resíduos alternados de açúcar e fosfato, projetando-se para fora como o corrimão da dupla fita. Os degraus da escada são as bases nitrogenadas ligadas por pontes de hidrogênio (Figura 3). A sequência de bases nitrogenadas é irregular, caracterizando a informação genética, em evolução a alguns milhões de anos. O conjunto completo, haploide ou diploide, é chamado **genoma**. A revelação da estrutura da dupla hélice complementar de Watson e Crick mostrou que o DNA é uma estrutura simples em termos moleculares, guardando, em duas fitas, a informação genética, de tal forma que qualquer alteração pode ser corrigida a qualquer tempo.

Existem dois tipos de bases no DNA: purinas e pirimidinas (Figura 4). As purinas são a adenina e a guanina, e as pirimidinas são citosina e timina. As bases púricas estão ligadas à desoxirribose pela ligação glicosídica no nitrogênio 1 e as pirimídicas pelo N9. As bases apresentam dois estados tautoméricos alternativos, em equilíbrio entre si. Essa capacidade de formar tautômeros alternativos é uma frequente origem de erros durante a síntese do DNA, conforme será visto posteriormente no conceito “mutações” (veja item 7.7 - mutações).

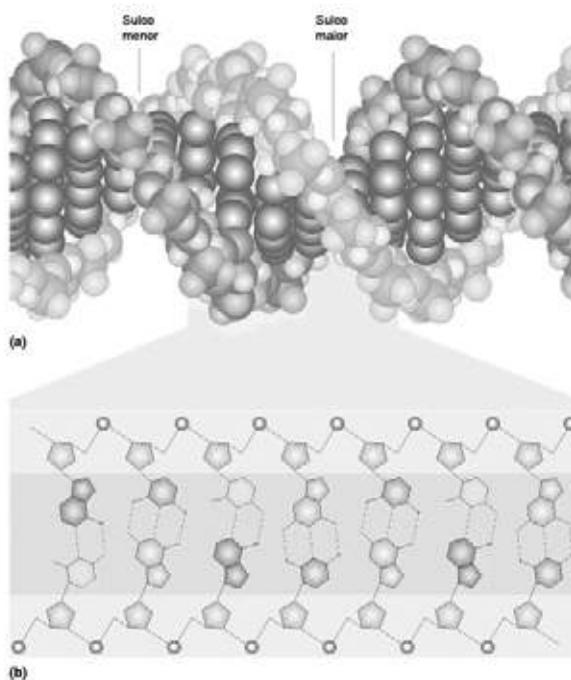
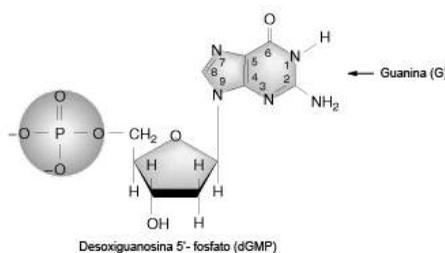
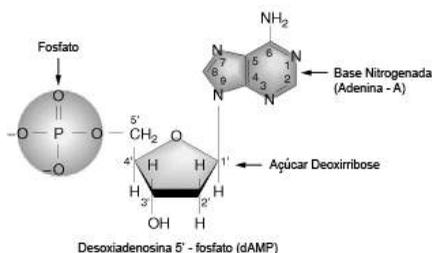


Figura 3 – (a) Modelo bola e bastão da dupla fita do DNA e (b) modelo químico.

Fonte: Griffiths et al. (1999).

Nucleotídeos de Purinas



Nucleotídeos de Pirimidina

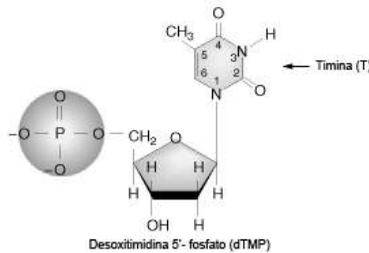
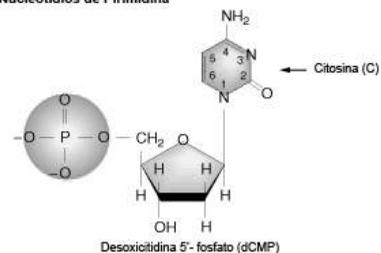


Figura 4 - Estrutura química dos quatro desoxirribonucleotídeos principais, blocos fundamentais do DNA.

Fonte: Griffiths et al. (1999).

O AZT (azido, 2', 3'-didesoxitimidina) é um fármaco usado na terapia contra o retrovírus responsável pela imunodeficiência humana. Na célula infectada, o RNA é copiado em DNA por uma transcriptase reversa, que é passível de erros por não apresentar as capacidades de revisão da DNA polimerase III. O AZT é azido-TTP, um análogo ao trifosfato de timidina, o dTTP (um dos desoxirribonucleotídeos incorporados ao DNA, além do dATP, dCTP e dGTP). A enzima transcriptase reversa incorpora o azido-TTP em lugar do dTTP, bloqueando a transcrição reversa porque o grupo 3'azido não pode formar uma ligação fosfodiéster com os nucleosídeos trifosfatos subsequentes.

A dupla hélice compõe-se de duas fitas ou cadeias polinucleotídicas mantidas unidas pelas ligações não covalentes fracas entre os pares de bases nitrogenadas. A adenina sempre se liga a uma timina e vice-versa, e a guanina sempre com uma citosina e vice-versa. A polaridade estrutural de uma fita é oposta à da outra fita, portanto, uma encontra-se na direção  $3' \rightarrow 5'$  e a outra,  $5' \rightarrow 3'$ , de acordo com os carbonos da ribose. Assim, se um lado apresenta  $5' - AGCTCA - 3'$  a sua correspondente complementar e oposta é  $3' - TCGAGT - 5'$ . Formam-se três pontes de Hidrogênio entre CG e duas entre AT. Os pares de bases formam pilhas entre os dois esqueletos açúcar-fosfato. As interações de empilhamento e as pontes de hidrogênio estabilizam **estereoquimicamente** a dupla fita.

O deslocamento de bases permite que uma base seja deslocada para fora da fita dupla, sendo que as enzimas responsáveis pelo reparo do DNA, ou metiladoras, atuam sobre uma base numa configuração extra helicoidal, permitindo o encaixe dessa base na cavidade catalítica da enzima (veja figura 7).

Algumas características da dupla fita incluem um empilhamento helicoidal de cerca de 10 pares de bases correspondentes a uma volta completa ao redor da hélice. Como resultado, torna-se um longo polímero estendido, formando duas cavidades ou sulcos laterais, uma maior e outra menor. Essa propriedade tem um caráter fisiológico importantíssimo. As cavidades guardam as extremidades de cada par de base, apresentando um padrão estereoquímico de doadores e receptores de pontes de hidrogênio, que é reflexo da sequência de pares de bases interna. Esse mecanismo de decodificação é reconhecido pelas cadeias laterais dos aminoácidos das enzimas metabolizadoras do DNA, reconhecendo e ligando a sequências de DNA específica, sem que seja necessário abrir a dupla fita.

## 1.2 O RNA é um oligoribonucleotídeo

O RNA, que possui fita simples, estruturalmente, é quimicamente muito parecido com o DNA. Entretanto, possui características próprias distintas, graças à sua capacidade de se dobrar e formar diferentes estruturas terciárias. Alguns RNAs são enzimas, trabalhando ativamente na produção de proteínas.

O RNA é composto por nucleotídeos, porém com uma ribose contendo uma hidroxila na posição 2'. A base nitrogenada timina é substituída por uracila. Outro ponto essencial é que ele possui fita simples. O RNA é um intermediário entre a informação genética do DNA e a síntese protéica, como RNAm. O RNA é estrutural quando na composição do ribossomo – **RNAr** e desempenha papel enzimático como RNA transportador – **RNAt** entre outros papéis relevantes. Ainda, pode ser considerado molécula reguladora da expressão gênica, chamado *interferente* ou silenciador da transcrição e da tradução.

Embora o RNA seja fita simples, ele exibe com frequência muitas regiões de dupla hélice, formando **auto complementação** em várias regiões, formando grampos, laços ou estruturas as mais complexas, como se vê na formação do ribossomo. Uma característica do RNA que o capacita a formar duplas hélices é um pareamento de bases adicional que não é do tipo Watson-Crick (tipo comum). É o par de bases G:U, com duas pontes de hidrogênio. Lembre-se de que a dupla fita de RNA é bastante diferente da hélice dupla regular e extensa do DNA. Dessa forma, o RNA pode formar grande variedade de estruturas terciárias, com alta complexidade. Essas diferentes conformações ocorrem, especialmente, pela liberdade de rotação no esqueleto de suas regiões não-pareadas. Proteínas auxiliam a formação de estruturas terciárias compostas por grandes moléculas de RNA, como as encontradas nos ribossomos. Estas proteínas revestem as cargas negativas dos esqueletos de fosfato, cujas cargas repulsivas desestabilizam a estrutura.

Funcionando como enzima clássica, contendo sítio ativo, sítio de ligação para o substrato e sítio ativo para ligação com co fator, o RNA enzimático é conhecido como **ribozima**. Uma das mais conhecidas é a RNase P, envolvida nas modificações pós-transcricionais do RNAt, a partir de longos RNAt precursores. Esta é composta por RNA e proteínas. Somente a parte composta por RNA participa da catálise. A enzima cobre cargas negativas do RNA, possibilitando que ele possa se ligar ao seu substrato (RNAt) negativamente carregado. Outros RNAs podem remover introns, de precursores de alguns mRNAs, tRNAs e RNAs ribossômicos, no processamento do RNA ou RNA *splicing*.

O fato do RNA ser capaz de processos catalíticos inspirou a possibilidade de que, num mundo primitivo, o RNA poderia ter atuado tanto como material genético e como maquinaria enzimática. Ele teria surgido antes que o DNA e possivelmente ter dado origem a este. O DNA, dessa forma, teria se liberado para o armazenamento da informação genética. A ribozima, como molécula catalítica, poderia ser considerada uma relíquia evolutiva de um mundo primitivo, onde o RNA teria um papel preponderante. As proteínas, mais diversas, teriam se especializado posteriormente na sua função catalítica.

### Símbolos usados em genética molecular e citogenética humana:

Grupos cromossômicos: A-G, segundo os seguintes critérios citológicos: Grupo A (cromossomos 1 a 3) grandes cromossomos aproximadamente centroméricos, Grupo B (cromossomos 4 e 5) grandes cromossomos com centrômeros submedianos (ou sub-metacêntricos), Grupo C (cromossomos 6 a 12 e o cromossomo X) cromossomos de tamanho médio com centrômeros submedianos (ou sub-metacêntricos), Grupo D (cromossomos 13 a 15) cromossomos de tamanho médio acrocêntricos. Cromossomo 13 possui um proeminente satélite no seu braço curto. Cromossomo 14 possui um pequeno satélite no seu braço curto. Grupo E (cromossomos 16 a 18) cromossomos curtos com centrômeros medianos ou submedianos (ou metacêntricos ou submetacêntricos). Grupo F (cromossomos 19 e 20) cromossomos curtos com centrômeros metacêntricos. Grupo G (cromossomos 21, 22 e o Y) cromossomos muito curtos acrocêntricos números de cromossomos autossomos 1 – 22, cromossomos sexuais: X e Y, braço mais curto do cromossomo p, braço mais longo do cromossomo q.

## Capítulo

# 3

## Núcleo eucariótico e cromossomos

Para compreender o conceito de **gene**, torna-se necessário inicialmente compreender a estrutura física do cromossomo e, em consequência, do núcleo eucariótico ou da célula procariótica onde se insere. Portanto, o DNA está sempre organizado como cromossomo. Experimentos mostraram que algumas estruturas do **cromossomo**, na verdade, são determinadas sequências de DNA em pontos localizados, são essenciais para a manutenção deste. São elas: telômeros, centrômeros e origens de replicação.

As condições do DNA no núcleo eucariótico são muito importantes para facilitar a expressão do gene. Imagine o DNA como a informação contida nos livros de sua biblioteca. Livros não usados, geralmente, ficam empacotados e localizados em condições mais difíceis de acessibilidade, enquanto que os livros muito usados ficam pelas mesas, abertos e prontos para serem lidos. À medida que a informação é requerida e maior é a demanda, desespiralizam-se as várias fibras de cromatina, possibilitando a sua leitura.

O termo **cromatina** foi cunhado como termo genérico, usado pelos primeiros microscopistas (cromo = cor; tina = substância) para definir o conteúdo do núcleo eucariótico. Alguns textos científicos referem-se às **fibras de cromatina** para designar partes do cromossomo que estão condensadas, em diferentes graus de empacotamento. De certa forma, pode-se dizer que a cromatina é a substância amorfa que preenche o núcleo interfásico (que não está em divisão). Nessa fase, a célula passa por intenso metabolismo de produção de proteínas e prepara-se para a divisão celular, replicando o seu DNA.

De que forma o DNA é empacotado nos núcleos eucarióticos? As histonas são subunidades proteicas usadas para empacotar o DNA. São proteínas altamente conservadas evolutivamente e carregadas positivamente, ligando-se ao DNA por interações iônicas.

O DNA faz duas voltas em torno de um **octâmero de histonas**. Esse octâmero é formado pelas histonas H2A, H2B, H3 e H4, duas cópias cada. A histona H1 ancora o DNA em torno desse arranjo. Esse conjunto é conhecido

como **nucleossomo** (Figura 5). Cada nucleossomo é ligado a outro por uma curta ligação de DNA nu, chamada de DNA ligador. Portanto, o conjunto de nucleossomos aparece como contas em um colar (... – o – o – o – o – ...). Essa configuração é identificada como **cromômero**. Dessa forma, o DNA eucariótico encontra-se em um estado estável como polinucleotídeo complexo com proteínas. O estado de contas no colar envolve então que o nucleossomo é a unidade de compactação do DNA. A partir desse estado, então, ocorre empacotamentos cada vez maiores.

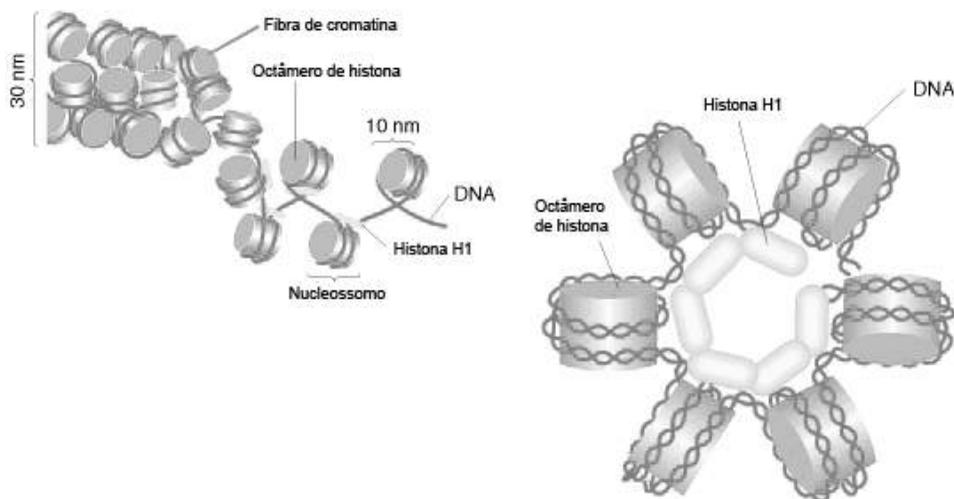


Figura 5 - Estrutura e arranjo dos nucleossomos com o DNA.

Fonte: Griffiths et al. (1999).

A cromatina, dessa forma, pode ser considerada um estado dinâmico do núcleo eucariótico. A **heterocromatina** é um estado de permanente condensação. Sabe-se que a heterocromatina é composta essencialmente de material não codificante, e sua inatividade transcricional é devida à sua alta condensação. Da mesma forma, a **euromatina** é a porção mais descondensada disponível para a leitura. Note que o conceito de cromatina e essas subdivisões são meio genéricas, facilmente aplicáveis, visíveis ao microscópio eletrônico, ao núcleo interfásico. Entretanto, à medida que os cromossomos se dividem e se espiralizam em direção à divisão celular, as fibras de cromatina se espiralizam formando estruturas cada vez mais condensadas e indisponíveis, de forma geral, para a transcrição. O máximo de condensação dos cromossomos corresponde ao cromossomo metafásico.

**Cromatina:** complexo de DNA e proteínas histonas. Conteúdo genérico do núcleo interfásico.

**Nucleossomo:** unidade estrutural da compactação da cromatina, consistindo em duas voltas de fita de DNA ao redor do octâmero de histonas.

**Regiões funcionais do cromossomo:** O centrômero é uma região constrita do cromossomo envolvida na união das cromátides-irmãs. Nele ocorre o cinetócoro, uma reunião de proteínas específicas em que se ligam os microtúbulos do fuso mitótico, na divisão celular. Os telômeros correspondem às extremidades dos cromossomos. A origem de replicação é uma sequência específica na qual se inicia a replicação do DNA. Um característica importante dessas sequências é que são formadas por sequências repetitivas **em tandem**, não codificantes, que, entretanto, são cruciais para o funcionamento do cromossomo. Veja também: *em tandem*.

Continua...

...Continuação

### Interação DNA-histonas:

o nucleossomo forma uma estrutura dinâmica, permitindo um acesso intermitente das proteínas que se ligam ao DNA. Portanto, o DNA envolvido com o nucleossomo não sofre qualquer dificuldade em ser replicado ou ser transcrito, graças à remodelação do nucleossomo.

Nucleossomos são formados imediatamente à replicação, deixando o DNA sempre empacotado.

### Sequências em tandem:

referem-se especialmente a sequências repetitivas, as repetições ficam lado a lado, como vagões num trem, às vezes, com uma pequena sequência variável comunicante.

### Acetiltransferases

(Figura 7) são enzimas que catalisam a adição de grupos acetil a grupos lisina das regiões N-terminais ("caudas") das histonas. Desacetilases são enzimas que removem estas adições. A acetilação está ligada à alta taxa de transcrição dessas regiões do DNA e, da mesma forma, as histonas desacetiladas, com áreas reprimidas do DNA. A metilação (feita por enzimas metiltransferases) deixa a cauda das histonas ligada mais fortemente ao DNA. As modificações nas caudas das histonas originam sítios de ligação para enzimas que fazem modificações na cromatina.

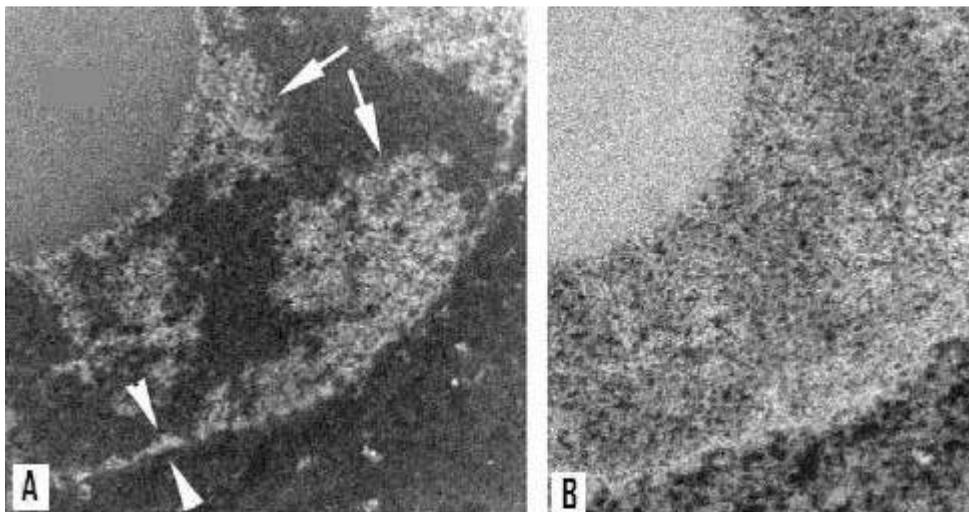


Figura 6 - Cromatina interfásica em embrião de rato, mostrando, em A: as setas indicam os locais onde a cromatina está mais concentrada próximo ao envelope nuclear. Em B: diferente tonalização. O círculo mais claro (parte superior) é o nucléolo.

Fonte: Ahmed et al. (2010).

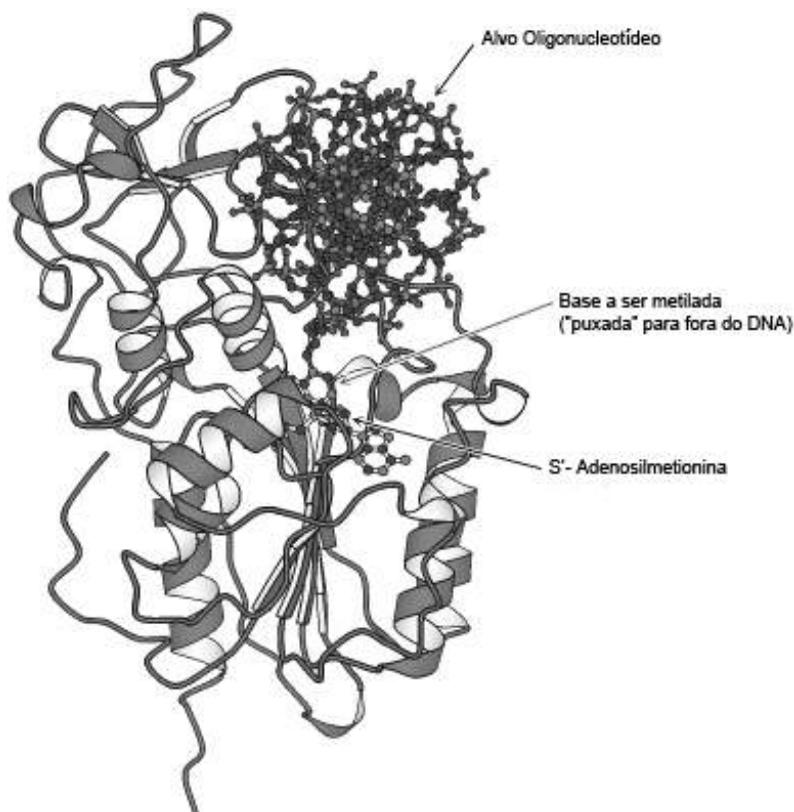


Figura 7 - Estrutura de uma DNA metilase ligada a um oligonucleotídeo mostrando a base a ser metilada.

Fonte: Berg et al. (1999).

# 1. Genes e genomas

## 1.1 Estrutura do gene

Uma definição **funcional** para o gene poderia ser: região do DNA transcrita como uma única unidade de transcrição, carregando informações sobre uma característica hereditária particular, correspondendo a uma única proteína (ou cadeias polipeptídicas relacionadas) ou a um único RNA (ou RNAs relacionados). Essa definição é essencial para o entendimento prático do gene, mas existem outras, mais recentes.

Os genes eucarióticos não têm a mesma estrutura dos genes procarióticos. Veja na transcrição do gene que o RNAm precursor, ou transcrito primário, resultante da transcrição pela enzima RNA polimerase, possui a mesma sequência do gene que o originou. Entretanto, isso não é verdade para os genes eucarióticos, os quais são genes interrompidos, ou seja, possuem sequências que interrompem a sequência pressuposta da proteína. Dessa forma, os **introns** são essas sequências intervenientes, não codificantes, as quais são removidas no processamento ou *splicing* do RNA. Os **exons** correspondem às sequências codificantes, que serão unidas no *splicing*, formando um RNAm maduro (alguns chamam transcrito secundário) pronto para ser traduzido em proteína no citoplasma. Por definição, o gene começa e termina em exons, iniciando pela extremidade 5' e terminando na 3'. Após o *splicing* os exons são sempre unidos na mesma ordem em que se encontra no DNA, de forma que existe **colinearidade** entre o gene e a sua proteína. É importante lembrar que nem todos os genes eucarióticos são interrompidos e, por outro lado, nem todos os procariotos estão livres de introns. Nos eucariotos superiores, a maioria dos genes são interrompidos, os introns são geralmente muito mais longos do que os exons. Portanto, a sequência do gene é muito maior que a da proteína que codifica.

A *complementaridade* entre fitas de DNA é determinada pelo pareamento Watson-Crick das bases AT e CG. Sequências distintas de DNA são ditas *similares* quando compartilham uma determinada proporção de pb (pares de bases) em comum. Caso a similaridade seja considerável, é possível que as sequências similares sejam *homólogas*, na qual se subentende que existe uma ancestralidade comum para as sequências estudadas. A separação das duas fitas complementares é chamada de **desnaturação** (de forma análoga ao visto em proteínas). Essa desnaturação ocorre quando há quebra das pontes de hidrogênio entre as bases, através do aumento de temperatura. A temperatura de fusão ( $T_m$ ) – lembre que fusão, neste caso, refere-se à separação de moléculas, como na água – é a temperatura média em que duas fitas de DNA se desnaturam, a qual é dependente de sua composição GC. Uma temperatura típica de fusão é em torno dos 87°C em genomas de mamíferos. A *renaturação* é o

**Splicing**: processo de retirada dos íntrons e união dos éxons, efetuado por snRNPs (“*small nuclear ribonucleoproteins*”) proteínas nucleares complexadas com RNA.

**Colinearidade**:

correspondência linear entre o gene e sua respectiva proteína.

**$T_m$**  – Temperatura de fusão (“*m*” de “*melting*”) é o ponto médio do intervalo de temperatura no qual o DNA é desnaturado.

**Genes housekeeping**:

são genes que codificam proteínas de funções metabólicas que ocorrem em todas as células.

As **ômicas** correspondem a processos e técnicas moleculares e técnicas que envolvem grande quantidade de informação como a genômica. Surgiram na sua continuidade, a transcriptômica, a proteômica, a metabolômica e outras ômicas como a peptidômica, a imunolômica. O sufixo “oma” - “conjunto de” - corresponderia ao conjunto de dados, constituindo um banco de dados produzido (como “genoma”), enquanto que “omica” seria a técnica em si (como “genômica”). As ômicas caracterizam-se por serem um corpo maciço de informações obtidas através de técnicas rápidas e automatizadas, de alto rendimento. Acrescentam-se ferramentas computacionais (bioinformática) que permitam integração, comparação de dados e anotação biológica.

### A anemia falciforme

consiste numa patologia que, em condições de baixa tensão de oxigênio, a cadeia  $\beta$  da hemoglobina A, a principal forma da hemoglobina no adulto, produz uma proteína de estrutura alterada e, conseqüentemente, altera significativamente sua função. A alteração de um único aminoácido faz com que a proteína se polimerize como um bastão, resultando em deformação da hemácia. A mutação que produz essa alteração é uma única alteração em um nucleotídeo, alterando o códon para glutamato  $\rightarrow$  valina. Essa mutação é chamada de não conservativa, pois o novo aminoácido contém características químicas diferentes do anterior. Raras mutações como essa são fenotipicamente importantes, pois interferem na sobrevivência do indivíduo.

processo contrário à desnaturação. Quando estão envolvidos ácidos nucleicos com origens diferentes, a renaturação é chamada de *hibridização*.

Sequências *regulatórias* aparecem no DNA, produzindo proteínas regulatórias ou se ligando a elas. Possuem função de regular a expressão de genes estruturais. O gene *estrutural* é aquele que tem um correspondente no RNAm maduro (que já passou por *splicing*). O gene estrutural pode ser definido como produtor de um RNAm intermediário de uma proteína. Genes chamados de constitutivos têm uma expressão constante (também chamados genes “*housekeeping*” – ou mantenedores da casa (veja Unidade 4).

As sequências de DNA podem ser encontradas em diversos tipos de complexidade. O DNA *não repetitivo* inclui as sequências únicas do genoma, codificantes. No DNA repetitivo, estão presentes em duas ou mais cópias, não codificantes, compondo-se de várias frações: *moderadamente e altamente repetitivo*. Compõe-se de famílias de sequências relacionadas, não necessariamente idênticas. O genoma de procariotos em geral, compõe-se apenas de DNA não repetitivo. Nos eucariotos inferiores, temos que o DNA, em sua maior parte, é não repetitivo. Em células animais e eucariotos, a porcentagem das frações repetitivas varia, mas ocupa uma grande parte do genoma. Portanto, genomas grandes apresentam altas quantidades de DNA repetitivo. O DNA moderadamente repetitivo está disperso por todo o genoma, geralmente na forma de sequências relativamente curtas. O DNA altamente repetitivo forma agrupamentos. A repetição *em tandem* de uma sequência curta pode criar uma fração do DNA com características específicas, geralmente com uma composição de bases única, características que são usadas para isolar essas frações. Esse DNA é chamado de DNA *satélite*, pelo fato de ficar separado do resto por ter densidade diferenciada. Um excelente exemplo de sequência repetitiva são as sequências **Alu** (veja Unidade 3). Essas sequências são exemplos de **SINES** (*Short Interspersed Elements*), ou repetições curtas. Temos também as sequências **LINES** (*Long Interspersed Elements*) ou repetições longas.

As sequências repetitivas e o DNA satélite são encontrados na heterocromatina. Em mamíferos, são comuns sequências que lembram satélites, formadas por repetições *em tandem* de uma unidade curta, mas que tem um tamanho total mais curto, consistindo, por exemplo, de 5-50 pb, com extenso polimorfismo entre indivíduos. Essas sequências são chamadas de *minisatélite* ou regiões VNTR (número variável de repetições *em tandem*) (veja a Unidade 5).

Comparando os RNAs procarióticos e eucarióticos, os RNAs procarióticos são *poliistrônicos* (produzem vários RNAs relacionados, cuja transcrição é controlada por um único processo regulatório, chamado *operon*). Os RNAs eucarióticos são *monocistrônicos*, produzindo um único RNAm e com diferentes níveis de regulação (veja a Unidade 4).

## 1.2 Estrutura genômica e evolução

O genoma de um organismo é seu conjunto completo de cromossomos. O mapa definitivo ou físico, do genoma consiste na sua sequência completa. A partir da sequência, torna-se possível identificar genes e sua localização.

Como os genomas evoluem? Os organismos diferem muito quanto ao tamanho do genoma, ao número de cromossomos, à abundância e ao tamanho dos introns e exons, à quantidade de DNA repetitivo e número de genes. No caso do número de genes, este se correlaciona muito grosseiramente com a complexidade dos organismos. Assim, o ser humano tem quase a mesma quantidade de genes da planta do arroz (cerca de 25 mil genes estruturais). Obviamente há uma correlação entre a complexidade de um organismo e o tamanho de seu genoma. Bactérias contêm cerca de 500 genes, humanos, cerca de 25 mil. Portanto, existe um complicado processo de evolução molecular. Este nem sempre pode ser conferido nas características fenotípicas ou morfo anatômicas. Para isso é preciso conferir os **marcadores moleculares**, além de alinhar e comparar inúmeras sequências de DNA.

Um ponto importante é que os genes em si, são mais conservados que a estrutura geral do genoma. Os genes diretamente ligados ao **metabolismo basal** da célula são sempre sujeitos a pressões evolutivas que não permitem alterações importantes. Mesmo assim, acumulam mutações sinônimas, ou neutras, como veremos mais adiante. Sequências repetitivas e introns têm maior liberdade para acumular mutações. As falhas nos mecanismos de replicação, recombinação e reparação do DNA, pela enzima DNA polimerase podem levar à simples falha de um nucleotídeo, como rearranjos genômicos em larga escala, como deleções, duplicações, inversões e translocações cromossômicas. Essas variações tomam forma nos polimorfismos de simples nucleotídeo (SNPs – *Single Nucleotide Polymorphism*, veja a Unidade 3), como pontos na sequência genômica onde os organismos podem variar, por apenas um nucleotídeo naquela posição. Os SNPs e outros polimorfismos nem sempre são visíveis no fenótipo, porém são herdáveis. Sítios polimórficos possuem um papel relevante na prática dos estudos evolutivos, médicos e forenses, pois podem caracterizar um táxon ou discriminar famílias ou indivíduos.

Um estudo promissor para compreender a evolução genômica são as **duplicações gênicas**, produzidas majoritariamente por *crossing-over* desigual. Ou seja, um dos gametas recebe uma sequência a mais e, se isso não perturbar a harmonia gênica, permanece no organismo a ser gerado, produzindo variação genômica. Genes duplicados então divergem, a cópia pode ser funcional ou não, por mutação de perda de função, gerando um **pseudogene**. Caso ambas as cópias permaneçam funcionais, podem então divergir quanto ao padrão de expressão. Dessa forma, duplicação seguida de divergência pode gerar inúmeras

**Marcadores moleculares:** sequência de DNA ou gene transmissível, detectável por técnica molecular, que tenha utilidade na análise genética.

**Metabolismo basal:** metabolismo essencial das células que possibilita sua sobrevivência.

**Pseudogene:** sequências de DNA derivadas de uma cópia funcional de um gene porém inativadas por mutação posterior.

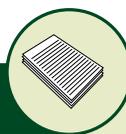
**Famílias de genes:** conjunto de genes relacionados, derivado de duplicação gênica ou outro processo multiplicador.

**Famílias de proteínas:** grupo de proteínas estruturalmente e funcionalmente similar, com identidade maior que 50% quando alinhada.

**Exon shuffling:** embaralhamento de exons, durante a transcrição, que pode resultar em novas proteínas. Veja também *splicing* alternativo.

famílias de genes, (consequentemente, **famílias proteicas**) com funções correlacionadas, aumentando a complexidade biológica. Um exemplo bem estudado de família gênica advinda de duplicação é a família das globinas. Comparando as diferentes proteínas globinas (como a hemoglobina), vemos homologies indiscutíveis nas suas sequências de aminoácidos e na sua estrutura, indicando que elas todas devem ter se originado de um ancestral comum. Pseudogenes globínicos também são encontrados. Portanto, é razoável dizer que sequências de aminoácidos homólogas em proteínas denotam funções correlacionadas e antecessores em comum. O fenômeno do parentesco gênico, que resulta em repetições de genes e pseudogenes é chamado de **redundância genética**.

Uma consequência disso é a existência de **domínios** proteicos, encontrados em diferentes famílias de proteínas. São caracterizados por uma repetição de uma estrutura proteica, em número variável, inseridos na estrutura tridimensional da proteína. Servem como elemento modular, produzindo “colchas de retalhos” que produzem novas proteínas. Os domínios proteicos geralmente estão inseridos nos exons. Imagine que juntando vários exons diferentes é possível formar novas proteínas, o que é chamado de embaralhamento de exons (*exon shuffling*).



## Textos complementares

### Registro da proteína insulina humana no GenBank

LOCUS AAA59172 110 aa linear PRI 12-FEB-2001  
 DEFINITION insulin [Homo sapiens].  
 ACCESSION AAA59172  
 VERSION AAA59172.1 GI:386828  
 DBSOURCE locus HUMINS01 accession J00265.1  
 KEYWORDS .  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
*Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
 Catarrhini; Hominidae; Homo.*  
 REFERENCE 1 (residues 1 to 110)  
 AUTHORS Bell,G.I., Pictet,R. and Rutter,W.J.  
 TITLE Analysis of the regions flanking the human insulin gene and  
 sequence of an Alu family member  
 JOURNAL Nucleic Acids Res. 8 (18), 4091-4109 (1980)  
 PUBMED 6253909  
 COMMENT On Aug 28, 1993 this sequence version replaced gi:186432.  
 Method: conceptual translation.

```

FEATURES             Location/Qualifiers
    source            1..110
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="11p15.5"
                     /tissue_type="liver"
                     /dev_stage="foetus"
    Protein           1..110
                     /product="insulin"
    sig_peptide       1..24
                     /gene="INS"
                     /note="G00-119-349"
    mat_peptide       25..110
                     /gene="INS"
                     /product="c peptide; G00-119-349"
    Region            26..110
                     /region_name="IGF_insulin_like"
                     /note="IGF_like family, insulin_like subgroup, specific
                     to vertebrates. Members include a number of peptides
                     including insulin and insulin-like growth factors I and
                     II, which play a variety of roles in controlling processes
                     such as metabolism, growth and...; cd04367"
                     /db_xref="CDD:58536"
    Site              order(32,35,37,108)
                     /site_type="other"
                     /note="putative receptor binding surface"
                     /db_xref="CDD:58536"
    CDS               1..110
                     /gene="INS"
                     /coded_by="join(J00265.1:2424..2610,J00265.1:3397..3542)"
                     /note="precursor"
                     /db_xref="GDB:G00-119-349"

ORIGIN
      1  malwmrllpl lallalwgpd paaafvnqhl cgshlvealy lvcgergffy
tpktrreaed
      61  lqvqvelgg gpgagslqpl alegslqkrq iveqcctsic slyqlenycn

```

Fonte: <http://www.ncbi.nlm.nih.gov/protein/AAA59172.1>

## Projeto Genoma

Uma ideia considerada absurda no seu início foi a ideia de se conhecer, letra por letra, a informação genética do ser humano, o chamado **Projeto Genoma Humano** (PGH). A técnica de sequenciamento de DNA, surgida por volta de 1985, tornou possível pensar estratégias que deveriam ser seguidas para um projeto desta magnitude. O maior responsável pelo desenvolvimento da tecnologia de sequenciamento por terminação do dideoxi, nos anos 70, foi F. Sanger, britânico, ganhador do prêmio Nobel por duas vezes. O objetivo primordial da iniciativa pública do governo americano para o PGH era produzir a sequência completa do genoma humano com taxas mínimas de erro. Formou-se assim o Consórcio Internacional de Sequenciamento do Genoma Humano (HGSC). A previsão inicial foi que, em 2005, após 15 anos, este seria completado. Nos

**Bioinformática:** consiste numa nova disciplina que congrega a Biologia e Informática. Podemos chamar de Bioinformática clássica, estudo ou ferramenta que permite obter, estocar, manipular e avaliar sequências de biomoléculas, especialmente DNA, RNA ou proteínas.

anos 80 e 90 apareceram plataformas robotizadas para sequenciamento em larga escala, acoplando o DNA a grupamentos químicos fluorescentes, detectáveis por laser. Assim, em janeiro de 2000, cerca de 20 centros de sequenciamento já geravam cerca de 1000 bases por segundo, 24 horas por dia, 7 dias por semana. Francis Collins assumiu a coordenação após o próprio J. Watson ter sido coordenador.

Em 1998, entretanto, surgiu um desafiante do PGH oficial: C. Venter, presidente da empresa privada *Celera Genomics*, em parceria com a empresa Perkin-Elmer Applied Biosystems, anunciou que iria sequenciar o genoma humano até 2001. Esta proposta alternativa concluiria o sequenciamento com 99,9% de cobertura com um erro inferior a 1:10.000 a um custo total final de 200 a 250 milhões de dólares. Números à parte, a grande preocupação da iniciativa pública foi que a empresa patenteasse o genoma humano, posteriormente, vendendo o direito de acesso aos dados.

Em fevereiro de 2001, foi publicado o primeiro rascunho do genoma humano como duas respectivas análises dos dois grupos nas famosas revistas científicas: *Nature* e *Science*. Esse rascunho possuía uma cobertura de 90% e 150.000 descontinuidades. Em 14 de abril de 2003, o genoma foi considerado concluído. A versão final contém 99% do genoma, com 99,99% de precisão, as descontinuidades ficaram em apenas algumas centenas. Uma série de conquistas adicionais incluem vários genomas eucarióticos como o do camundongo, do rato, de vários outros animais e vegetais-modelo para a pesquisa. A **Bioinformática** possibilitou a análise das sequências obtidas, seu registro autoral, anotação, acesso e aplicação de ferramentas, na comparação de sequências de nucleotídeos, aminoácidos, estrutura de proteínas, acesso e comparação de genomas, entre muitas outras. Aliada à rede mundial de computadores, possibilitou acesso público e gratuito a este conhecimento.

O estudo molecular de sequências de genes e proteínas, incluindo aspectos evolutivos que podem ajudar a classificar e caracterizar grupos taxonômicos de organismos pode ser considerado uma das maiores contribuições do PGH. O PGH possibilitou excelente progresso nas heranças monogênicas ou com conjunto limitado de genes, aplicável ao diagnóstico, tratamento e prognóstico. A identificação do gene e de suas mutações têm revelado inúmeras estratégias. Entretanto, doenças multifatoriais ou poligênicas, envolvendo inúmeras interações gene-gene e gene-ambiente, como por exemplo, o câncer, doenças psiquiátricas, diabetes e doenças cardiovasculares são um grande desafio para os pesquisadores, envolvendo estratégias ainda mais complexas para seu estudo. Outra questão importante é que o PGH foi feito a partir de um pequeno grupo de voluntários. Dessa forma, não inclui a herança genética que é única (no sentido parental), inerente a todos os organismos, mas representa apenas uma condição gênica universal. A análise genômica constatou a extraordinária *sin-tenia* (conservação da ligação genética entre táxons diferentes) entre centenas de grupos de organismos, o que vem ajudando sobremaneira os estudos evolutivos. Veja na Figura 8 a homepage do Projeto Genoma.

Figura 8 - Página inicial da ferramenta Entrez – Genome Project – apresentando centenas de links para genomas, finalizados e em andamento.

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

## Após dez anos, mapa dos genes dá poucas pistas para tratamentos médicos

Dez anos depois de o ex-presidente dos EUA Bill Clinton (1993-2001) ter anunciado que o primeiro esboço do genoma humano estava concluído, a medicina ainda não viu grande parte dos benefícios prometidos.

Para os biólogos, o genoma produziu uma surpresa significativa após outra. Mas o objetivo primário do genoma humano, de US\$ 3 bilhões, era descobrir as raízes genéticas de doenças comuns como câncer e Alzheimer, e criar tratamentos, no entanto, permanece em grande parte uma ilusão. De fato, uma década após o esforço e os geneticistas estão quase de volta ao início no conhecimento das origens das doenças comuns.

Ao anunciar, em 26 de junho de 2000, que o primeiro esboço do genoma humano tinha sido alcançado, Clinton disse que ele "revolucionaria o diagnóstico, a prevenção e o tratamento da maioria ou de todas as doenças humanas".

Em entrevista coletiva, Francis Sellers Collins, então diretor da agência do genoma dos Institutos de Saúde dos EUA, disse que o diagnóstico genético de doenças seria alcançado em dez anos e que os tratamentos começariam a surgir talvez cinco anos

A palavra **sintenia**, empregada em genética molecular, descreve uma situação em que diferentes genes ou diferentes *loci* gênicos estão ligados no mesmo cromossomo, em indivíduos ou espécies. Nos genes sintênicos, esta localização torna-se preservada, assim, rearranjos no genoma tais como translocações cromossômicas podem separar dois *loci*, resultando na perda de sintenia entre eles. De forma inversa, as translocações também podem reunir dois *loci* gênicos anteriormente separados, resultando em ganho de sintenia. A sintenia compartilhada pode refletir relações funcionais entre os genes sintênicos, como combinações de alelos que são vantajosos quando herdados juntos ou que estão compartilhando mecanismos de regulação. Etimologicamente, a palavra "sintenia" é um neologismo usado para descrever uma precisa preservação da ordem dos *loci* gênicos a partir de um ancestral comum. Os genes *Hox* nos animais são genes que são determinantes do plano corporal, interagindo de forma crítica entre si, são essencialmente conservados sintenicamente. A chamada *macrosintenia* seria a preservação de grandes porções de um cromossomo, *microsintenia*, preservação da sintenia, preservação em poucos genes por um tempo.

depois. "Em longo prazo, talvez em mais 15 ou 20 anos, vocês verão uma completa transformação da medicina terapêutica", ele acrescentou.

A indústria farmacêutica gastou milhões de dólares para desvendar segredos genômicos e está começando a colocar no mercado várias drogas orientadas pela genética. Enquanto os laboratórios continuam despejando muito dinheiro na pesquisa do genoma, ficou claro que a genética da maior parte das doenças é mais complexa do que se previa, e que são necessários muitos anos antes que novos tratamentos sejam capazes de transformar a medicina.

A última década trouxe uma enxurrada de descobertas de mutações causadoras de doenças no genoma humano. Mas na maioria dos casos, as descobertas só explicaram uma pequena parte do risco de contrair a doença. E alguns cientistas começam a temer que muitas variantes genéticas ligadas a doenças possam ser ilusões estatísticas.

O Projeto do Genoma Humano foi iniciado em 1989 com o objetivo de sequenciar, ou identificar, as 3 bilhões de unidades químicas no conjunto de instruções genéticas humano, encontrar as origens genéticas de doenças e, então, desenvolver tratamentos. Com a sequência na mão, o passo seguinte foi identificar variantes genéticas que aumentam o risco de doenças comuns, como câncer e diabetes.

Em 2002, os Institutos de Saúde iniciaram um projeto de U\$138 milhões chamado HapMap, para catalogar as variantes comuns dos genomas europeu, asiático do leste e africano.

Com o catálogo na mão, a segunda etapa seria ver se alguma das variantes era mais comum nos pacientes com determinada doença do que em pessoas saudáveis. Esses estudos exigiram grandes números de pacientes e custaram vários milhões de dólares cada um. Quase 400 deles tinham sido concluídos em 2009. O resultado é que centenas de variantes genéticas comuns hoje foram estatisticamente ligadas a diversas doenças.

Fonte: adaptado de Wade (2010).

## Síntese da Parte



Os ácidos nucleicos consistem de polinucleotídeos. O núcleo eucariótico contém DNA em várias condições de condensação, condição que afeta a transcrição dos genes. A cromatina é o conteúdo eucariótico nuclear. Genes e genomas são definidos de forma funcional, embora haja outras definições. Os genes codificam RNA ou polipeptídeos e os genomas abrangem o conjunto dos cromossomos. Discute-se o projeto genoma humano e dos experimentos que levaram à descoberta da dupla fita do DNA.

## Atividades de avaliação



1. Explique, em termos gerais, como funciona o fluxo da informação genética (replicação, transcrição e tradução). Inclua nesta questão: principais elementos envolvidos (enzimas e outros elementos) etapas básicas de cada processo, em sequência e o momento em que os processos ocorrem na célula.
2. Procure o conceito de transcrição reversa e quais organismos a produzem.
3. Desenhe um nucleotídeo com suas moléculas constituintes, indicando as ligações entre elas. No esquema, apresentar a diferença entre desoxiribonucleotídeo e ribonucleotídeo.
4. Procure e esquematize os diferentes tautômeros das bases nitrogenadas.
5. Procure as diferentes classes de íntrons e definição mais completa de ribozima.
6. Desenhe e legende um núcleo eucariótico com todas suas estruturas básicas: membrana nuclear, nucléolo, poros nucleares, elementos cromáticos. Analisar a função de cada um deles.
7. Recorde os ciclos celulares (mitose e meiose) e suas fases.
8. Procure nos livros e materiais didáticos: informações sobre cromossomos politênicos.
9. Procure na rede mundial de computadores: técnica de cariotipagem.
10. Procure na rede mundial de computadores: técnicas de visualização do núcleo e de seus componentes, como por exemplo, a imunofluorescência.
11. Revise a estrutura de proteínas: o que é estrutura primária, secundária, terciária e quaternária.
12. Desenvolva os conceitos relacionados: domínios estruturais proteicos e motivos estruturais proteicos.
13. Faça uma busca bibliográfica no GenBank: Visite o site do NCBI (<http://www.ncbi.nlm.nih.gov>). Seguir o roteiro de busca abaixo para o PubMed;
  - a) No campo "Search" procure "PubMed" ou "PubMed Central" (nessa opção somente são apresentados opções de literatura grátis para o download;
  - b) Digite a palavra-chave de seu interesse (veja sugestões a seguir);

c) Quantos resultados ("*results*") retornou a sua busca? Ou seja, quantas referências foram obtidas? Quantas revisões ("*reviews*") e quantos textos disponíveis grátis? ("*Free Full Text*")?

Por exemplo: "1 to 20 of 211181" significa "1 a 20 de 211.181 referências no total." Veja um resultado na página seguinte:

### The antitumor activity of the fungicide ciclopirox. (Título)

Zhou H, Shen T, Luo Y, Liu L, Chen W, Xu B, Han X, Pang J, Rivera CA, Huang S. (Autores)

Int J Cancer. 2010 Mar 11. [Epub ahead of print] (Publicação)

PMID: 20225320 [PubMed - as supplied by publisher]

Related articles

Clicando no link do título (a sua escolha) obtém-se novas informações, especialmente o Resumo.

d) Procure os trabalhos da bibliografia deste texto ou relacionados.

Sugestões de palavras-chave: patologias como *breast cancer* - câncer de mama (usado no exemplo), nome científico de organismos: "*Canis familiaris*" ou "*dog*", grupo geral: "bacteria", moléculas: *cytochrome*, etc.

Observação: ao fazer esta busca nos computadores em rede da sua Universidade, é possível obter mais material em "*download*". A rede CAPES está disponível somente nas redes de IES do país.

14. Leia os textos complementares e discuta as perspectivas do projeto genoma para as próximas décadas.

## Leituras, filmes e sites



### Livros

- SCHWARTZMAN, S. Um espaço para a ciência: a formação da comunidade científica no Brasil. Brasília: MCT/CNPq/CEE, 2001. 307 p.
- WILMUT, I.; CAMPBELL, K.; TUDGE, C. Dolly: a segunda criação. Rio de Janeiro: Objetiva, 2000. 394 p.
- BRODY, A. R.; BRODY, D. E. As sete maiores descobertas científicas da história. São Paulo: Companhia das Letras. 1999. 440 p.

## Filmes

- Gattaca: a experiência genética (Dir. Andrew Niccol, 1997) - fabuloso filme sobre questões do genoma e bioética.
- Uma verdade inconveniente (Dir. Davis Guggenheim, 2006) - aquecimento global.
- Sonhos tropicais (Dir. André Sturm, 2002) – a “revolta da vacina” no Rio de Janeiro.
- Avatar (Dir. James Cameron, 2009) - fabuloso filme que exhibe clonagem e exobiologia.
- Viagem fantástica (Dir. Richard Fleischer, 1966, refilmagem de Pieter Kuijpers, 2005 – Viagem fantástica II – rumo ao cérebro) – Fisiologia humana e ciência.

## Sites

### (Livros grátis para download)

- <http://bafanaciencia.blog.br/>
- <http://gigapedia.com/>
- <http://search.4shared.com/q/1/biologia>
- <http://bio-livros.blogspot.com>

### (Instituições e associações)

- <http://www2.abed.org.br/> (Associação Brasileira de Educação a Distância - ABED)
- <http://www.abep.org.br/usuario/GerenciaNavegacao.php> (Associação Brasileira de Estudos Populacionais - ABEP)
- <http://www.ancib.org.br/> (Associação Nacional de Pesquisa e PG em Ciência da Informação - ANCIB)
- <http://www.sbhc.org.br/> (Sociedade Brasileira de História da Ciência - SBHC)
- <http://www.sbeb.org.br/> (Sociedade Brasileira de Engenharia Biomédica - SBEB)

## Referências



AHMED, K.; DEHGHNI, H.; RUGG-GUNN, P.; FUSSNER, E.; ROSSANT, J.; BAZETT-JONES, D. P. Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. **Plos One**, Munich, v. 5, n. 5, p. 1-13, 2010.

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular biology of the cell**. 5. ed. New York: Garland Science, 2008. 1.261 p.

BERG, J. M.; TYMOCZKO, J. L.; STRYER, L. **Bioquímica**. 6. ed. Rio de Janeiro: Guanabara Koogan, 2008. 1.114 p.

ENGSTRÖM P. G; SUI, S. J. H.; DRIVENES, O; BECKER T. S; LENHARD, B. Genomic regulatory blocks underlie extensive microsynteny conservation in insects. **Genome Research**, New York, v. 17, n. 12, p. 1898–1908, 2007.

GRIFFITHS, A. J. F.; MILLER, J. H.; SUZUKI, D. T.; LEWONTIN, R. C.; GELBART, W. M. **An introduction to genetic analysis**. New York: W. H. Freeman and Company, 1996. 915 p.

JEFFREYS, A. J.; WILSON, V.; THEIN, S. L. Hypervariable “minisatellite” regions in human DNA. **Nature**, New York, v. 314, p. 67-73, 1985.

VENTER, J. C.; ADAMS, M. D.; MYERS, E. W. et al. The sequence of the human genome. **Science**, Washington, v. 291, p. 1304-1351, 2001.

WADE, N. Após dez anos, mapa dos genes dá poucas pistas para tratamentos médicos. **Folha de São Paulo**, São Paulo, 21 jun. 2010. Caderno The New York Times, p. 1-2.

WATERMAN, T. H. Revolution for biology. **American Scientist**, North Carolina, v. 50, p. 458, 1962.

**Parte**

**2**

# **Mecanismos genéticos básicos**



# Considerações iniciais e histórico

## Objetivo

- Apresentar os mecanismos genéticos básicos da célula referentes ao fluxo da informação genética e as suas implicações.

## 1. Aspectos gerais

A informação genética em todas as células segue, a princípio, uma via de mão única: DNA especifica a síntese de RNA, e RNA especifica a síntese de polipeptídeos, que, posteriormente, formam proteínas. Por causa de sua universalidade, o fluxo de informação genética DNA → RNA → polipeptídeo (proteína) é discutido como o dogma central da Biologia Molecular, conforme foi visto no Capítulo 1. Neste capítulo iremos detalhar um pouco mais essas importantes etapas. Atenção, pois esses processos são a base para nosso futuro entendimento das técnicas de DNA recombinante.

A primeira etapa, a síntese de RNA usando uma DNA polimerase dependente de RNA, é descrita como a **transcrição** e ocorre no núcleo das células eucarióticas e, em certa medida, nas mitocôndrias e cloroplastos, únicas organelas que apresentam uma capacidade genética, além do núcleo eucariótico. Podemos chamar então o RNA (qualquer tipo) de **transcrito**, pois ele é sempre produto de transcrição. A segunda etapa, a síntese de polipeptídeos, é conhecida como **tradução** e ocorre pela ação dos ribossomos, grandes complexos de proteínas e RNA os quais são encontrados no citoplasma e também em mitocôndrias e cloroplastos. As moléculas de RNA que contêm a informação específica para um polipeptídeo é o RNA mensageiro (RNAm).

A expressão da informação genética segue um princípio de colinearidade: a sequência linear de nucleotídeos no DNA é decodificada para dar uma sequência linear de nucleotídeos de RNA, que pode ser decodificado, por sua vez, em grupos de três nucleotídeos (códon) para dar uma sequência linear de aminoácidos no produto polipeptídeo. A tradução envolve um conceito

importante, que é o código genético. Corresponde à chave que é usada para transferir a informação de um alfabeto de 4 letras (AGCT) para um alfabeto de 20 letras (conjunto de aminoácidos envolvidos na síntese protéica). Apenas uma pequena porcentagem de todo o DNA das células são sempre transcritas. De acordo com as suas necessidades, células diferentes transcrevem diferentes segmentos do DNA (unidades de transcrição), que são **unidades discretas**, espaçadas irregularmente ao longo da sequência do DNA. No entanto, a grande maioria do DNA celular nunca é transcrita em qualquer célula.

No contraponto desse raciocínio, está o processo de replicação. Todo o DNA é *sempre* replicado. A replicação é um ramo do dogma envolvido na produção do DNA a partir do próprio DNA, portanto, DNA *sempre* é formado a partir de outro DNA.

O que esses processos têm em comum?

São todos processos envolvidos na produção de novas macromoléculas, todas envolvem polimerizações (formação de polímeros). DNA é um polímero de desoxiribonucleotídeos; RNA é um oligômero de ribonucleotídeos; proteína, polímero de aminoácidos. São processos do **anabolismo** celular. Outro ponto em comum é que possuem três etapas básicas: (i) iniciação, quando um complexo de iniciação é formado, contendo as enzimas e partículas iniciais, “montadas” no início do processo (geralmente no início do molde); (ii) alongamento: etapa repetitiva onde ocorre rápido crescimento do polímero; por fim, ocorre a (iii) finalização, a qual também apresenta um complexo de finalização, também com enzimas e partículas de finalização (no final do molde).

## 2. Breve histórico

As décadas de 1940 e 1950 foram marcadas por estudos essenciais na Biologia Molecular. George Beadle e Edward Tatum examinaram, em 1941, as bases bioquímicas do metabolismo em *Neurospora*, a relacionando o genótipo ao fenótipo, ou genes a proteínas. O paciente método desenvolvido por eles, por meio do qual mutantes deficientes nutricionalmente, chamados *auxotróficos*, não crescem em meio mínimo, mas crescem em meio com o nutriente do qual o mesmo é negativo. Assim, se o fungo tinha uma mutação que o tornava incapaz de produzir arginina, este só crescia em meio mínimo contendo arginina. Classificando os mutantes em grupos, foi possível caracterizar quais eram os requerimentos necessários para cada grupo, determinando qual enzima seria necessária para cada etapa, até chegar ao produto final que é a arginina. Esse método foi usado por Adrian Srb e Norman H. Horowitz, revelando o ciclo de produção da arginina. Esses experimentos foram a base da teoria **um gene, uma enzima**. Quando se verificou a formação multimérica de muitas proteínas, o modelo foi modificado para **um gene, um polipeptídeo**.

A partir da descoberta da dupla hélice (1953), sabia-se, então, que essa molécula era a transmissora das informações genéticas. Entretanto, nada se sabia sobre como as proteínas podiam ser formadas a partir dela. Somente dez anos depois (em 1961), foi que ocorreu a descoberta de um código genético presente no DNA que organiza a montagem das proteínas, através dos experimentos de Marshall Warren Nirenberg e Johann Heinrich Matthaei, usando RNAs sintéticos produzidos por uma enzima polinucleotídeo fosforilase, a qual consegue produzir um RNA sem um molde para guiá-la. Dessa forma, produziram homopolímeros, moléculas de RNA formadas por um único nucleotídeo, por exemplo: 5'...UUUUUU...3' ou RNA poli(U). Este RNA poli(U) foi então adicionado a 20 tubos cada um contendo um sistema de células de síntese de proteínas e mais os 20 aminoácidos, cada um deles separadamente em cada tubo. Obviamente somente o tubo contendo o aminoácido fenilalanina apresentou proteínas recém sintetizadas radioativas. Novos experimentos, similares a esse puderam ser feitos, e o significado de todos os códons foi concluído em 1965. Nirenberg recebeu o Premio Nobel em Fisiologia ou Medicina em 1968. Nirenberg faleceu em 15 de janeiro de 2010.

# Mecanismos genéticos básicos

## Unidades discretas:

unidades descontínuas, pontuais e ordenadas, existentes em cada gene. **dATP, dCTP, dGTP, dTTP**: desoxi adenosina trifosfato, desoxi citosina trifosfato, desoxi guanosina trifosfato, desoxi timidina trifosfato. De forma genérica, são chamados de dNTPs ou desoxirribonucleotídeos trifosfato. São adicionados à fita de DNA nascente, perdendo dois fosfatos inorgânicos. **dNMP**: desoxinucleotídeo monofosfato.

**Primer**: Trecho curto e temporário de RNA hibridizado a uma fita simples de DNA molde. Sua função é fornecer um grupo 3' – OH para a ligação de um novo nucleotídeo.

**Fita atrasada** ou descontínua (filamento *lagging*): filamento de DNA replicado de forma descontínua, a qual é sintetizada na direção contrária à da abertura da forquilha de replicação.

## 1. Replicação

Durante o processo de síntese de DNA (replicação do DNA), as duas fitas de DNA de cada cromossomo são desenroladas por uma enzima **helicase**, e cada fita de DNA dirige a síntese de uma fita de DNA complementar para gerar dois duplex DNA, cada um dos quais é idêntico à molécula. A replicação é descrita como **semiconservativa** (Figura 1), pois cada nova fita duplex contém um filamento da molécula-mãe e uma fita de DNA recém sintetizada. A **DNA polimerase** (Figura 2) é a principal enzima, a qual catalisa a síntese de DNA, usando as fitas parentais como molde utilizando os quatro desoxirribonucleotídeos (**dATP, dCTP, dGTP, dTTP**) como precursores de nucleotídeos.

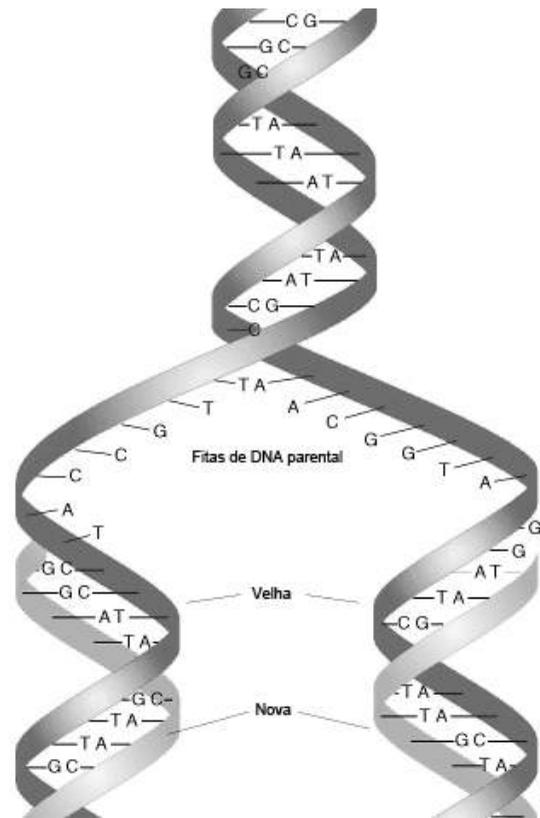


Figura 1 - Modelo semiconservativo da replicação do DNA. As novas fitas são produzidas sempre na direção 5'→3', adicionando novas complexidades à um mecanismo cheio de detalhes. As pontes de H são abertas pelas enzimas helicases. Fonte: Griffiths et al. (1999).

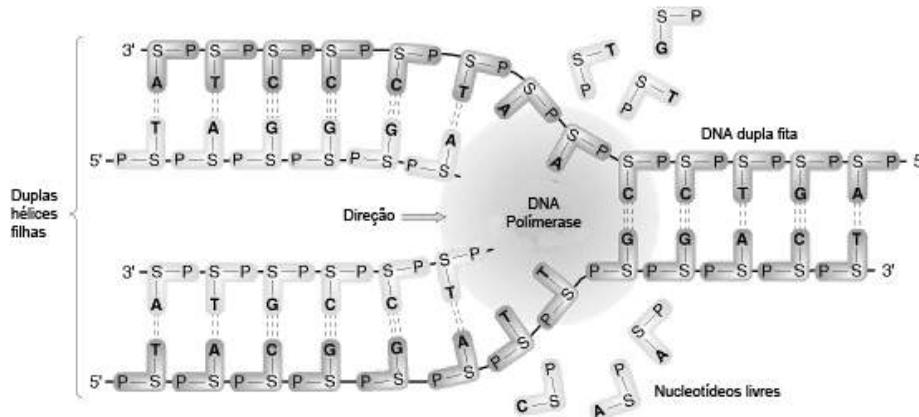


Figura 2 - Esquema da replicação apresentando a forquilha de replicação, a direção da replicação e outros detalhes. Fonte: Griffiths et al. (1999).

As reações moleculares catalisadas pelas DNA polimerases ocorrem na extremidade 3'-OH exposta do DNA em crescimento adicionando a essa extremidade, um dNMP fornecido por um precursor dNTP, o dNMP é adicionado e os dois resíduos de fosfato distal, ou seja, os resíduos  $\beta$  e  $\gamma$  são clivados e resultam num grupo pirofosfato (PPi), o qual é descartado. Dessa forma, a energia necessária para o processo é fornecida inerentemente pelos próprios substratos.

Após a abertura das fitas, a enzima DNA primase (ou iniciase) adiciona *primers* de RNA a esses inícios de replicação para que possa formar a extremidade 3'-OH. Esses *primers* de RNA serão retirados após a polimerização, pela própria DNA polimerase (uma outra isoforma). À medida que retira, adiciona desoxiribonucleotídeos corretos, tapando esses pequenos buracos. A enzima DNA ligase faz o fechamento final das pontes fosfodiéster, que não são fechadas pela DNA polimerase.

A replicação do DNA é iniciada em pontos específicos, que são chamados de "origens de replicação". A partir dessa origem, o início ocorre em uma forquilha ou bolha de replicação, em que o DNA parental é aberto e bifurca-se. Como as duas cadeias do duplex DNA parental são antiparalelas, cada uma funciona individualmente como modelo para a síntese de uma fita complementar antiparalela. As duas fitas correm e crescem em direções opostas, cada uma em direção à abertura da forquilha de replicação. Entretanto, a direção do crescimento das duas novas cadeias, feita pela enzima DNA polimerase ocorre sempre na direção 5'  $\rightarrow$  3'. Uma das fitas demora mais para ser feita, chamada de fita atrasada (ou descontínua), e é exatamente aquela que não está com a direção 5'  $\rightarrow$  3' da sua fita parental em direção à abertura da forquilha de replicação. A fita oposta é chamada de líder (ou contínua).

#### Fita líder ou contínua (filamento *leading*):

Filamento de DNA replicado continuamente, com poucas interrupções de *primers*, sintetizado na direção da abertura da forquilha de replicação.

#### Fragmentos de Okasaki:

Trechos curtos e temporários de DNA recém-sintetizado, iniciados a partir dos *primers* e finalizados pela enzima DNA ligase. Note que são temporários, já que, no final da replicação, não se nota qualquer diferença na fita de DNA.

#### A transcrição reversa

foi encontrada por meio do estudo de tumores de galinhas, gatos e camundongos, onde vírus com RNA foram implicados na produção do tumor. Em 1970, David Baltimore, Howard Temin e Satoshi Mizutani descobriram uma DNA polimerase dependente de RNA, chamada de transcriptase reversa, a qual permite a cópia do RNA em DNA. A transcrição reversa é responsável pela inclusão, nos genomas de muitos tipos de sequências de DNA que foram retrotranscritas, convertendo o RNA unifilar em DNA bifilar.

A enzima **telomerase** faz a replicação dos telômeros. Nos telômeros humanos, temos a sequência TTAGGG repetida *em tandem*. A enzima estende a ponta 3' do filamento molde. Sem a telomerase, as pontas dos cromossomos podem ficar cada vez mais curtas após cada replicação. Essas perdas podem caracterizar perda de genes, o que pode ser letal. Uma característica excepcional é que ela contém um molde de DNA incluso. Estudos mostram uma correlação entre tamanho do telômero e envelhecimento humano, em distúrbios chamados *progerias*, doenças genéticas caracterizadas por envelhecimento tardio. Os pacientes de progeria têm telômeros curtos. Acredita-se que uma diminuição do telômero contribui normalmente, para o processo de envelhecimento. O Prêmio Nobel de Fisiologia ou Medicina do ano de 2009 foi atribuído conjuntamente a Liz Blackburn, Carol Greider e a Jack Szostak pela descoberta da enzima telomerase e dos telômeros como protetores das extremidades dos cromossomos (Figura 3).

Essa exigência introduz uma assimetria no processo de replicação do DNA: só a fita líder terá o grupo 3'-OH livre no ponto de bifurcação e é sintetizada continuamente. Nesse lado, ocorre a adição sequencial de nucleotídeos e alongamento contínuo na mesma direção que o movimento de abertura da forquilha de replicação. No entanto, a síntese da cadeia atrasada tem que ser feita de forma fragmentada, pois a DNA polimerase tem que "esperar" que a fita se abra. Esses fragmentos possuem de 100-1000 nucleotídeos de tamanho e são referidos como **fragmentos de Okasaki**. Esses fragmentos serão posteriormente unidos pela enzima DNA ligase, de forma que nenhuma ponte fosfodiéster deixará de ser formada. No final, toda a molécula é uniforme, sem sinal ou defeito aparente.

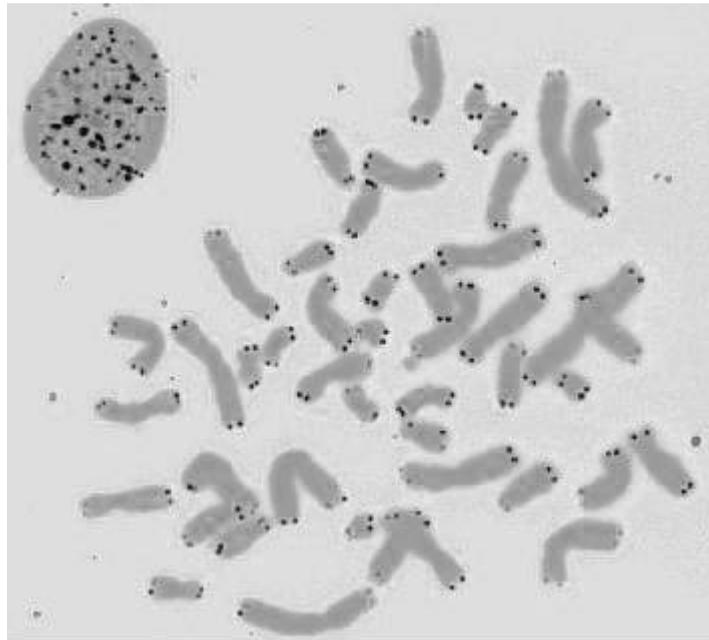


Figura 3 - FISH em cromossomos humanos, evidenciando os telômeros, com anticorpos fluorescentes. Cromossomos na fase metafásica. (Veja FISH na Unidade 5). Fonte: Arnoult et al. (1991).

## 2. Transcrição

As informações genéticas, tanto nos procariotos quanto nos organismos eucariotos, passa pela molécula de RNAm para depois ser traduzido em um polipeptídeo funcional. Em procariotos, uma única RNA polimerase é a enzima responsável pela produção de todos os tipos de RNAs: RNAt, RNAm e RNAr. Nos eucariotos encontramos cerca de três RNAs polimerases: I, II e III. A RNA polimerase I é responsável pela síntese de RNAs ribossômicos (menos do RNAr 5S), RNA polimerase II que é a principal responsável pela síntese dos RNAs mensageiros e outros RNAs ligados à recomposição (ou *splicing* do

RNA), e a RNA polimerase III sintetiza RNAr 5S, tRNAs e outros **RNAs de baixo peso molecular**.

A transcrição, portanto, é o processo de produção de qualquer RNA, que também pode ser chamado de transcrito. Como, na replicação, uma das fitas do DNA é usada como molde, porém, diferente da replicação, nem todo o DNA é transcrito, mas somente genes (no sentido funcional, que já comentamos) os quais codificam polipeptídeos ou RNA. Na replicação, todo o DNA é replicado, em tese, sem faltar uma base. Na transcrição não há *primers*, primase, helicase e ligase. A enzima RNA polimerase é responsável por todo o trabalho, mas somente quando todas as subunidades da enzima se agrupam sobre o DNA, numa condição chamada formação da enzima **RNA polimerase holoenzima**.

A enzima possui cinco subunidades polipeptídicas:  $2\alpha$ ,  $\beta\beta'$  e  $\sigma$  (duas alfas, beta e beta linha e sigma). A holoenzima forma-se sobre o DNA, mais precisamente sobre a *sequência promotora*, ou *promotor*. O promotor é a sequência onde se liga inicialmente a subunidade sigma. Quando isso acontece, outras subunidades se juntam e formam então a holoenzima, num chamado, assim, *complexo fechado*. Assim que se inicia a transcrição, justamente a partir do códon de iniciação, temos um *complexo aberto*. Abre-se, então, uma bolha de transcrição de aproximadamente 19 nucleotídeos. A bolha abre-se na frente e fecha-se atrás, sendo conduzida pela holoenzima em direção ao final do gene.

O RNA possui uma ponta 5' e uma ponta 3'. Os nucleotídeos também são adicionados nesta direção 5'  $\rightarrow$  3' pela enzima RNA polimerase. Geralmente, a ponta 5' é desenhada do lado esquerdo e a ponta 3' do lado direito, como quem lê. À medida que a molécula de RNA, progressivamente, vai sendo produzida, a ponta 5' é deslocada do molde e a bolha de transcrição fecha-se atrás da polimerase. Várias polimerases movem-se ao longo do gene, produzindo várias moléculas de RNA ao mesmo tempo em diferentes estágios.

Uma questão importante é que a RNA polimerase copia uma das fitas, a **fita molde**. Portanto a fita do RNA recém sintetizado é igual à **fita não molde**, lembrando que o RNA possui uracilas em que de timinas. Sequências de bases, quando mostradas na literatura, como as que você vê nesta publicação e em outras, são sempre iguais ao RNA ou ao filamento não-molde, também chamado de *codificante*.

Como foi dito anteriormente, a transcrição possui três fases: iniciação, alongamento e término. Existem diferenças importantes nesses estágios, tanto em procariotos quanto em eucariotos. A iniciação tem um local iniciador que é a sequência promotora ou promotor, situada um pouco antes do códon de iniciação do gene (Figura 4). Essa parte *upstream* ao gene também

A **replicação** é sempre semiconservativa, iniciando-se nas origens de replicação, de onde saem forquilhas de replicação para os dois lados. Utiliza *primers* de RNA. A fita é produzida sempre na direção 5'  $\rightarrow$  3'. Os nucleotídeos utilizados para a construção do novo DNA são dNTPs, produzindo dois fragmentos concomitantemente: *lagging* e *leading*. Tipos diferentes de enzimas participam da replicação, em ordem e sequência bem estabelecidas, finalizando com a retirada dos *primers*, reparação pela DNA polimerase e DNA ligase. Todo o DNA é replicado, incluindo as sequências teloméricas. A replicação ocorre na fase S da interfase. Após a replicação, histonas imediatamente já se ligam e formam a cromatina.

A numeração que é colocada no gene: ...-3-2-1+1 +2 +3... diz respeito somente à **posição** do nucleotídeo no gene. O códon de iniciação é a base dessa numeração (ATG = +1+2+3). Os números negativos à esquerda referem-se às sequências *upstream* ao gene. Cuidado: essa numeração não tem qualquer relação com ionização ou propriedades químicas do gene.

#### RNA polimerase holoenzima:

a reunião de todos os polipeptídeos que compõem a enzima RNA polimerase.

#### RNA de baixo

##### peso molecular:

pequenas moléculas de RNA nuclear usadas na recomposição (*splicing*) do RNAm.

**Fita molde** (para a transcrição): uma das fitas que servem de molde para a RNA polimerase. Qualquer das fitas pode ser considerada molde, entretanto, nem todas são. Nos procarionotos, é comum ver genes, de ambos os lados da fita.

#### Fatores gerais de

**transcrição:** pequenas partículas protéicas que são usadas na transcrição eucariótica. Sua função é aumentar a afinidade da enzima RNA polimerase ao promotor.

**Fatores específicos** são identificados com um determinado gene e ajudam na sua regulação, induzindo ou reprimindo um gene específico.

contêm *sequências reguladoras* (veja a Unidade 4), sendo que a região promotora é uma das mais importantes e estudadas. Por convenção, o primeiro nucleotídeo do códon de iniciação (normalmente a sequência ATG) recebe a numeração +1. Os próximos nucleotídeos à direita continuam a numeração: +2, +3, +4, assim até o fim do gene. *Upstream* (anterior ao gene) os nucleotídeos recebem numeração negativa (-1, -2, etc). Nas bactérias, como a *E. coli*, encontramos as caixas (ou *boxes*) que são sequências consenso (Veja a Unidade 4). O fator ou subunidade sigma é importante na ligação com o promotor. Nas bactérias também temos vários tipos de fatores sigma, como  $\sigma^{70}$ , (o expoente 70 foi atribuído porque ela possui 70 kDa). Os diferentes fatores sigma reconhecem sequências promotoras diferentes e, assim, podem transcrever diferentes conjuntos de genes. Nos eucariotos, a transcrição é regulada bem no ponto de início pela presença dos fatores de transcrição. Os **fatores gerais de transcrição** mais a enzima RNA polimerase II formam o complexo de pré-iniciação (veja também a Unidade 4).

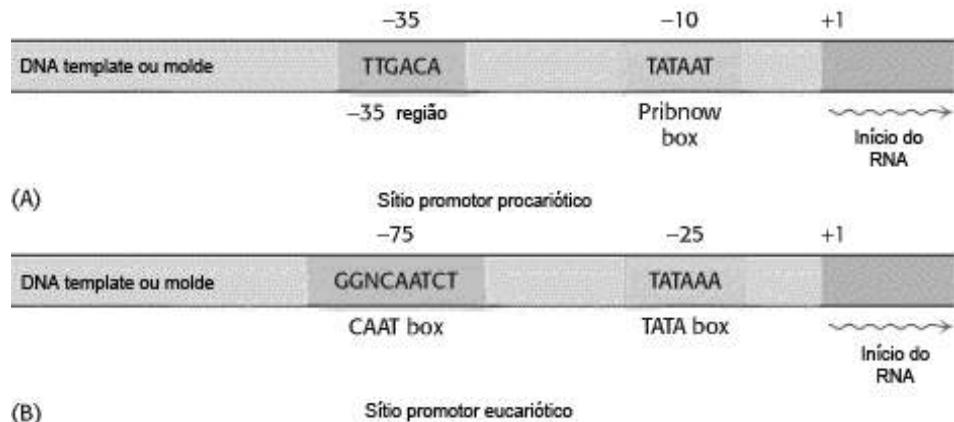


Figura 4 - Sítios promotores da transcrição. (A) procarionóticos e (B) eucarióticos. Fonte: Berg et al. (2008).

No alongamento, em bactérias como também em eucariotes, temos a bolha de transcrição (Figura 5), expondo o filamento molde. Nos eucariotos, entretanto, os íntrons são retirados pelo processo de *splicing*, ou recomposição. A recomposição reúne os éxons formando um RNAm maduro ou transcrito secundário. Portanto, só as sequências codificantes são levadas para o citoplasma para serem traduzidas.

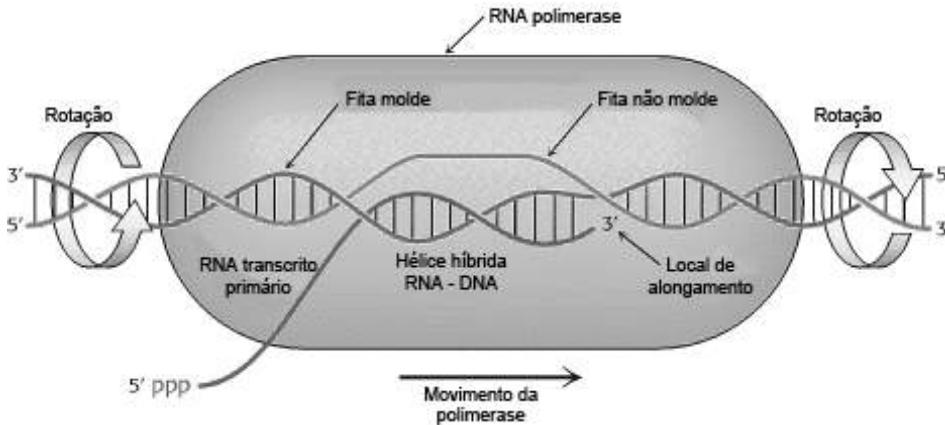


Figura 5. Bolha de transcrição apresentando o alongamento da fita de RNA. Fonte: Berg et al. (2008).

Nos procaríotos encontra-se somente uma enzima transcrevendo todos os genes. Nos eucariontes existem três RNA polimerases diferentes: (I, II e III), sendo que somente a II transcreve RNAm, originando proteínas.

### 2.1 Recomposição e processamento

Conforme já comentamos, a remoção dos íntrons é chamada de *recomposição* ou *splicing* (Figura 6). Pequenas partículas formadas por proteínas e RNA, chamadas snRNPs (*small nuclear ribonucleoproteins*) ou mesmo chamadas snRNA (alguns autores chamam de “*snurps*” devido à pronúncia em inglês) removem os íntrons com precisão. As junções éxon-íntron possuem um papel importante nessa questão. São formadas por sítios nos quais ocorrem as reações de recomposição. Nas pontas 5' dos íntrons geralmente possuem GU e AG nas pontas 3', chamadas de regras GU – AG. Outro ponto comum, sempre encontrado nos íntrons é uma A observada na posição entre 15 e 45 nucleotídeos antecedente ao sítio de corte 3'. Nucleotídeos que flanqueiam aqueles altamente conservados também mostram conservação, porém em menor grau. Esses nucleotídeos conservados no transcrito primário são reconhecidos por cinco snRNPs (U1, U2, U3, U4 e U5). Essas proteínas e mais outras fazem um complexo chamado de **spliceossomo** (Figura 7).

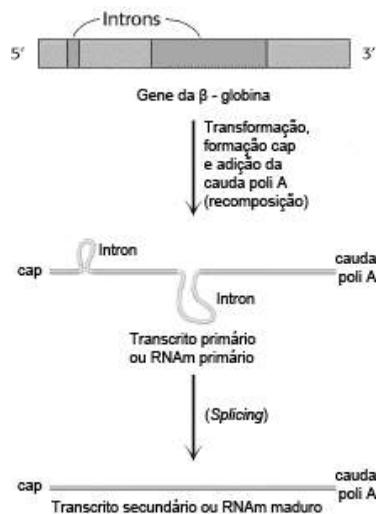


Figura 6 - Transcrição e recomposição do gene da B-globina. Fonte: Berg et al. (2008).

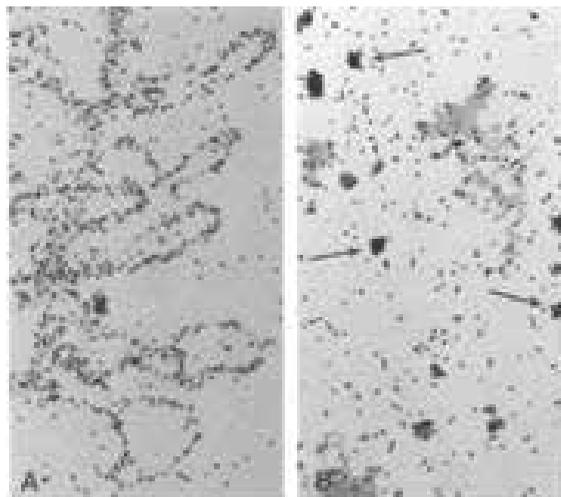


Figura 7 - Autoradiografia de *loops* cromossômicos do tritão *Notophthalmus viridescens* após hibridização com uma sonda antisense para a snRNA U1, sonda tritiada (marcada com o isótopo  $^3\text{H}$ ). (A) snRNA U1 são vistas como glóbulos. (B) As setas mostram spliceossomos. Fonte: Wu et al. (1991).

A notícia da ocorrência de auto *splicing* no gene de rRNA do protozoário ciliado *Tetrahymena*, no ano de 1981, foi um choque para o estudo molecular. Esse transcrito tinha capacidade de remover em si próprio um íntron de 413 nucleotídeos, sem nenhuma outra proteína para ajudar. Posteriormente foram encontrados outros íntrons (**íntrons tipo I e tipo II**) com essa propriedade, porém continuam sendo exceção à regra geral. Nos últimos estudos, percebe-se que é o componente ribonucleico, e não a proteína, o responsável pela retirada do íntron. Esses achados e o papel das ribozimas fornecem um novo paradigma para o papel do RNA na célula, que muitos vêm chamando de *Mundo do RNA*. Essa nova abordagem inclui a teoria de que o RNA teria sido predecessor do DNA por sua versatilidade catalizadora e capacidade genética (Figura 8).

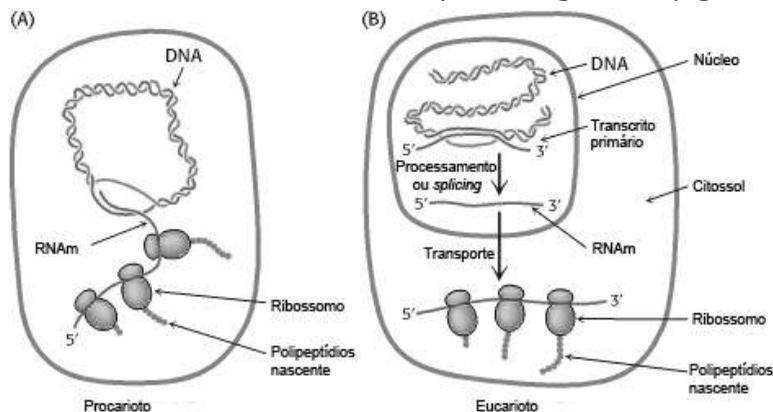


Figura 8. A transcrição e a tradução são acoplados nos procaríotos (A) e espacialmente e temporalmente separados nos eucariotos (B). Fonte: Berg et al. (2008).

O processamento das pontas 5' e 3' (Figura 9) ocorre quando o RNAm emerge da transcrição. Consiste em revestimentos químicos protetores das pontas 5' e 3'. A ponta *cap* é adicionada ao lado 5' por várias proteínas. O revestimento consiste em uma 7-metilguanossina ligada ao transcrito por três grupos fosfato. A cauda poli A é adicionada ao lado 3', depois que o RNAm é solto da enzima.

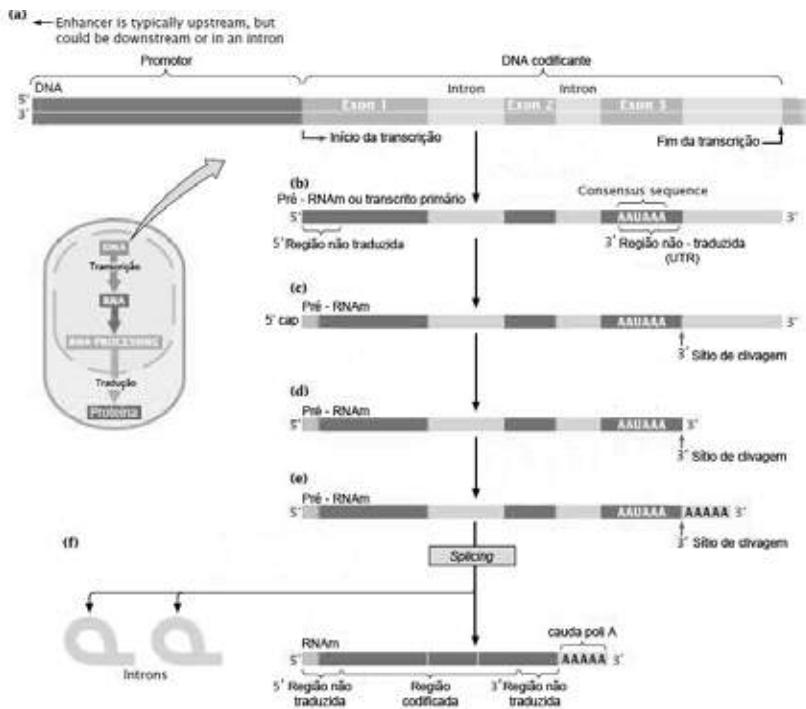


Figura 9 – Processamento, recomposição (*splicing*) e produção do RNAm maduro em eucariotos. Note a presença do promotor *upstream* ao gene, a presença de introns e éxons. Fonte: Pierce (2004).

O processamento também tem como objetivo a sinalização para o ribossomo. Essa sinalização é feita por sequências sinalizadoras como a **sequência 5' não traduzida ou 5' UTR**, que não codifica aminoácidos. Em bactérias essa região contém a sequência chamada **Shine-Dalgarno**, sítio de ligação ao ribossomo.

A finalização da transcrição só ocorre muitos nucleotídeos após o fim do gene, em sequências finalizadoras. Dois tipos de finalização são conhecidos: a chamada dependente da proteína *rô* e a independente de *rô* (ou intrínseca). A sequência finalizadora, na verdade antecipa o ponto de término. A transcrição não termina abruptamente no final do gene, como um carro no final de uma estrada. O finalizador é como um quebra-molas, que diminui a velocidade. Na finalização dependente de *rô*, a proteína auxilia, correndo ao longo da fita de RNAm recém sintetizada e movendo-se em direção à ponta

A **transcrição** possui propriedades importantes: é seletiva, somente algumas sequências do DNA são transcritas em um determinado momento. Somente o molde é transcrito. O transcrito é complementar ao molde e de polaridade inversa. Os nucleotídeos (ATP, GTP, CTP e UTP) contêm três fosfatos, sendo que dois deles são retirados, liberando energia para a síntese. A fita em crescimento é produzida sempre na direção 5' → 3' (portanto, a leitura é feita na direção 3' → 5'). A transcrição não precisa de *primer*. A RNA polimerase é uma enzima multimérica que necessita de um fator sigma, específico, que localiza o promotor. O promotor contém sequências consenso essenciais para a ligação com a enzima. A enzima possui atividade revisora. A transcrição começa em alguns poucos nucleotídeos antes do códon de iniciação e termina pouco depois da sequência finalizadora.

**Íntrons tipo I e II:** íntrons autorremovíveis, ou seja, independem de proteínas para a sua recomposição (*splicing*). São encontrados em alguns genes de rRNA, em alguns genes codificantes de proteínas e em mitocôndrias e cloroplastos.

**Sequências 5' e 3'**

**UTR:** de “*Untranslated Region*”, sequências não codificantes, localizadas nas extremidades dos genes, as quais são transcritas, mas não traduzidas.

**Sequência de Shine-Dalgarno:**

sequência *upstream* ao gene, a qual é encontrada no RNAm. Reconhecida pelo ribossomo, faz hibridização com o RNAr, para que o primeiro códon seja colocado no sítio P. Encontrada em bactérias, possui análogos em eucariotos como a sequência Kozak.

**Proteína rô:** proteína que auxilia na finalização da transcrição, liberando a RNA polimerase de sua ligação com o DNA, através de sua atividade de helicase.

3', seguindo o movimento da RNA polimerase. Ao encontrar o grampo de RNA ela para e, com atividade de helicase, abre o híbrido de RNAm/DNA, liberando a nova fita de RNA.

No segundo caso, o término é direto, sem a ajuda de uma proteína. A sequência de término contém cerca de 40 nucleotídeos, terminando num trecho rico em GC que é seguido por uma fileira de seis ou mais A. O RNAm transcrito, portanto, é rico em GC. Essas bases formam complementação, produzindo um grampo no RNAm transcrito primário. O grampo faz com que a RNA polimerase pare a transcrição.

### 3. Tradução

A tradução é feita nos ribossomos. Estes são considerados verdadeiras máquinas que coordenam todo o processo. Alguns fatos têm que ser apresentados antes de compreendermos os passos essenciais da tradução. A tradução é um dos processos mais complexos da maquinaria celular eucariótica (e procaríótica também). Os três passos essenciais já discutidos anteriormente: início, meio e fim, são acrescidos de mais dois: pré-início (ativação dos aminoácidos) e pós-final (modificações pós-traducionais, nas proteínas eucarióticas). A ativação dos aminoácidos (Figura 10) consiste na ligação específica de um aminoácido e de seu tRNA específico por uma enzima **aminoacil tRNA sintetase**. Existem 20 dessas enzimas, cada uma específica para cada “casamento” aminoácido e RNAt. O RNAt é carregado com um aminoácido ligado à sua ponta 3'. O reconhecimento dos tRNAs pela sintetase depende de sequências diferentes de nucleotídeos dos tRNAs. O aminoácido alanina (Ala) liga-se ao seu tRNA específico (tRNA<sup>Ala</sup>), originando a seu aminoacil-tRNA (Ala-tRNA<sup>Ala</sup>).

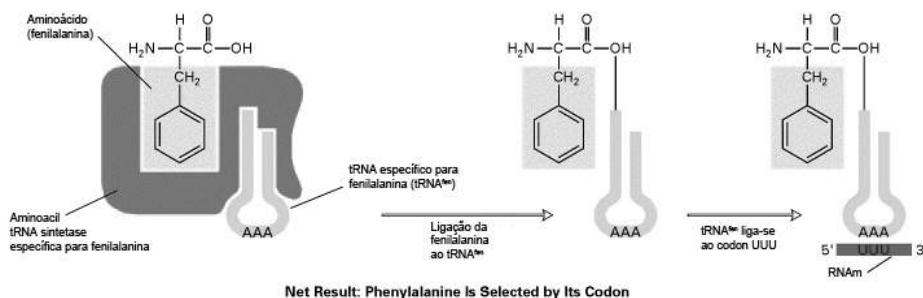


Figura 10 - Ativação dos aminoácidos: ligação do RNAt ao seu aminoácido específico pela enzima aminoacil-t-RNA sintetase. Note, no final, a ligação códon/anticódon. Fonte: Lodish et al. (1999).

As partículas envolvidas no processo são muitas e devem ser listadas inicialmente: o ribossomo, contendo duas subunidades: 30S e 50S (em bactérias); o complexo ribossômico de iniciação 70S ocorre quando os dois estão

juntos, o RNAm, conjunto de três proteínas, os fatores de iniciação (IF I, II e III), o tRNA iniciador (fMet-tRNA<sup>fMet</sup>) e trifosfato de adenosina (ATP). As duas subunidades ribossômicas ligam-se e separam-se em equilíbrio dinâmico. O complexo ribossômico 70S ocorre quando os dois estão juntos. O ribossomo completo contém três sítios ou cavidades internas: sítios A, P e E (A de aminoacil, P de peptídeo e *Exit* de saída). No alongamento, temos o fator de alongamento Tu (EF – Tu), fator de alongamento Ts (EF – Ts) e GTP. A enzima peptidil transferase faz a ligação peptídica. Fatores de liberação são proteínas que, em *E. coli*, são responsáveis pela finalização da tradução, liberando o polipeptídeo. Existem três fatores de liberação que ligam-se ao ribossomo: RF1, 2 e 3, de acordo com codons específicos de liberação.

Inicialmente, considerando os passos principais: iniciação, alongamento e término, temos uma sequência típica de passos. O RNAm deve se ligar à subunidade menor do ribossomo (30S). Nas bactérias, a subunidade menor liga-se à sequência de Shine-Dalgarno no RNAm, posicionando o ribossomo no códon inicial. Nos eucariotos, a ponta 5' é revestida pelo cap, com a ajuda de fatores de iniciação. O cap é reconhecido pelo ribossomo e também pela sequência Kozak. A iniciação eucariótica requer mais fatores de iniciação. IF3 liga-se à subunidade menor do ribossomo e impede a ligação da subunidade maior durante a iniciação. A partir da junção de GTP, IF2 e IF1, a fMet-tRNA<sup>fMe</sup> no códon de iniciação que já está posicionado no sítio P. A partir daí, o GTP é hidrolisado em GDP e Pi, os fatores de iniciação IF1, IF2 e IF3 são liberados para que a subunidade maior possa se firmar. Assim, temos o complexo de iniciação 70S (Figura 11 A).

A partir desse momento, dá-se início ao alongamento, em que temos três etapas. Na primeira, o tRNA carregado é levado para o sítio A, fazendo companhia para o fMet-tRNA<sup>fMet</sup>. Junto com ele entra EF-Tu e EF-Ts e GTP para ser hidrolisado. Nessa fase, o sítio P contém o fMet-tRNA<sup>fMet</sup>, e o sítio A contém o próximo aminoacil – tRNA. Após o tRNA carregado ser colocado no sítio A, o GTP é hidrolisado em GDP, liberando o complexo EF-Tu-GTP. O EF-Tu sai do complexo, liberando GDP + Pi, sendo que EF-Ts regenera o complexo EF-Tu-GTP, que está, então, pronto para se combinar com outro tRNA carregado. Os dois aminoácidos iniciais estão próximos, e a enzima peptidil transferase produz um dipeptídeo. Quando isso acontece, o primeiro aminoácido (fMet-tRNA<sup>fMet</sup>) fica liberado do seu transportador, que fica descarregado. Essa foi a segunda etapa. A próxima e última etapa do alongamento é a translocação do ribossomo todo ao longo do RNAm no sentido 5' → 3'. Essa etapa posiciona o ribossomo no códon seguinte e requer o fator de alongamento G (EF – G) e a hidrólise de GTP em GDP. Os tRNAs ainda estão ligados pelas pontes de hidrogênio entre o códon/anticódon. O que ocorre é o deslocamento dentro

dos sítios E, A e P. O tRNA<sup>Met</sup> cai dentro do sítio de saída (*Exit*), o segundo tRNA – aminoácido vai para o sítio P, e o sítio A fica novamente liberado para o próximo tRNA carregado. O ciclo se repete, adicionando aminoácidos um de cada vez. O término ocorre quando se chega ao códon de finalização. O fator 1 reconhece os códons UAA e UAG, RF2 reconhece UGA e UAA. O fator 3 forma complexo com GTP e liga-se ao ribossomo. Estes fatores de liberação então promovem o corte do tRNA no sítio da cadeia polipeptídica. O GTP que forma o complexo com RF1 é hidrolisado em GDP e Pi. Outros fatores adicionais podem ajudar a liberação de todos os componentes, RNAm, RNAt e ribossomos, os quais estão livres para começar novamente (Figura 11B). Veja, na Figura 12, a comparação entre a organização gênica, transcrição e tradução entre procaríotos e eucariotos.

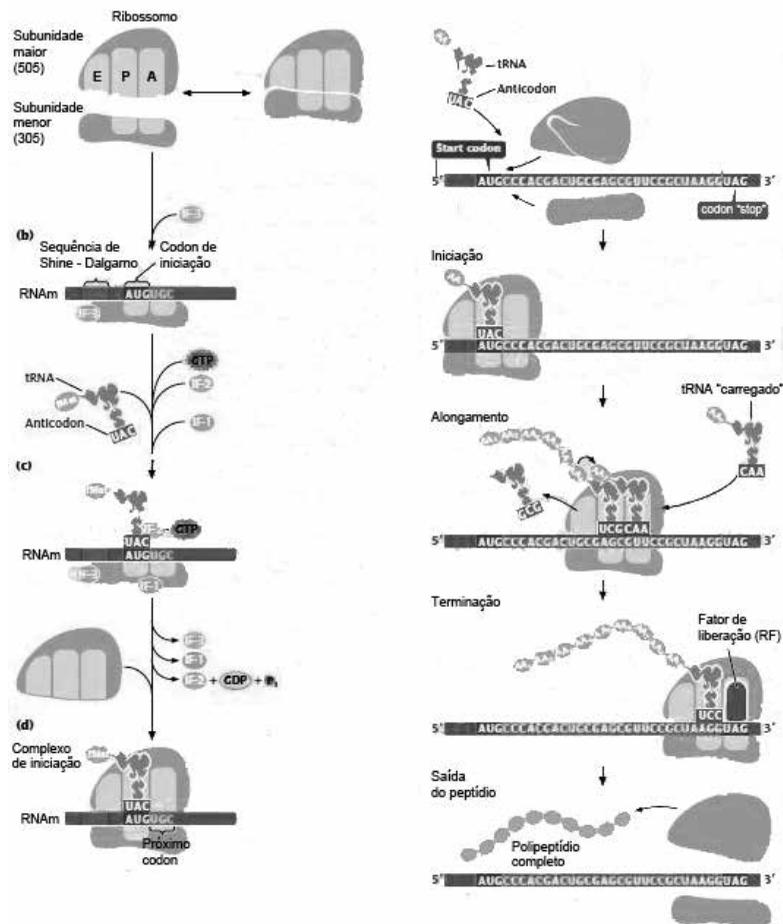


Figura 11 – (A) Início da tradução com a formação do complexo de iniciação. (B) Resumo das etapas da tradução.

Fonte: Pierce (2004).

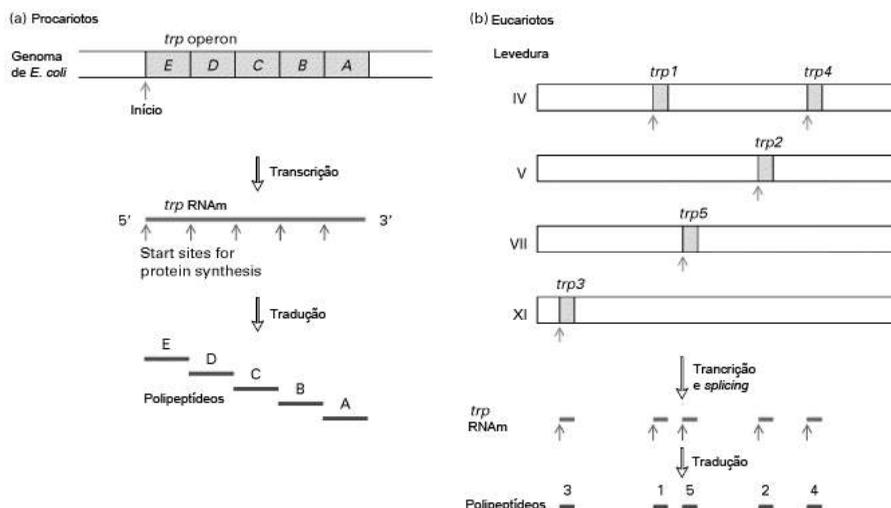


Figura 12 - Comparação entre a organização gênica, transcrição e tradução entre procarionotos e eucariotos. Em A: o óperon triptofano mostra-se como um segmento contínuo do cromossomo de *E. coli*, contendo cinco genes e que codificam as enzimas necessárias, passo a passo para a síntese do triptofano. Esse tipo de RNAm é chamado policistrônico em comparação com o RNAm eucariótico monocistrônico. Em B verifica-se que os mesmos genes encontram-se distribuídos em cromossomos diferentes, e cada um forma um RNAm isolado. Observe que cada gene eucariótico possui um promotor específico.

Fonte: Lodish et al. (1999).

### 3.1 Código genético

Observando com atenção o Quadro 1, que apresenta a conversão dos códons de DNA em aminoácidos, nota-se que alguns aminoácidos correspondem a mais de um códon. Existem 6 trinças que codificam para arginina (CGT, CGC, CGA, CGG, AGA, e AGG), ou seja, se no ribossomo estiver presente a trinca CGT ou a trinca AGG, ocorrerá a adição de uma arginina na cadeia polipeptídica. Outros aminoácidos possuem um número menor de códons, por exemplo: a valina corresponde a quatro códons; a isoleucina é codificada por três códons, e a fenilalanina corresponde a duas trinças. A existência de códons que são sinônimos (determinam a introdução do mesmo aminoácido na cadeia polipeptídica) faz com que o código genético seja redundante (ou *degenerado*). O único aminoácido que possui apenas um códon é a metionina, e o códon desse aminoácido (ATG) é sinal para o início do processo de tradução.

Os **eventos pós traducionais** ocorrem juntamente com o dobramento da proteína em sua correta forma tridimensional. Isso ocorre com a ajuda de proteínas **chaperonas**, uma classe encontrada em quase todos os organismos, as quais formam complexos com outras proteínas, fazendo que elas dobrem ou voltem a se dobrar (como no caso de eventos de desnaturação proteica). Outras modificações consistem em: modificação das cadeias laterais como a **fosforilação**, adição de grupos fosfato por enzimas cinases, aos grupos hidroxila dos aminoácidos serina, treonina e tirosina, enquanto que as enzimas fosfatases retiram esses grupos. A **ubiquitinação** consiste na adição de várias cadeias de uma proteína ubiquitina, que marca a proteína para degradação por uma protease chamada **proteossomo**. **Sequências sinais** correspondem a um pequeno polipeptídeo, o qual direciona a proteína para uma organela como o retículo endoplasmático, por exemplo. Se a proteína for de membrana, esta é direcionada para a membrana, passando por poros do retículo endoplasmático. Em sua maioria, as proteínas eucarióticas são inativas, necessitando de modificações promovendo ativação ou degradação.

Código genético padrão											
TTT	F	Phe	TCT	S	Ser	TAT	Y	Tyr	TGT	C	Cys
TTC	F	Phe	TCC	S	Ser	TAC	Y	Tyr	TGC	C	Cys
TTA	L	Leu	TCA	S	Ser	TAA	*	Ter	TGA	*	Ter
TTG	L	Leu	TCG	S	Ser	TAG	*	Ter	TGG	W	Trp
CTT	L	Leu	CCT	P	Pro	CAT	H	His	CGT	R	Arg
CTC	L	Leu	CCC	P	Pro	CAC	H	His	CGC	R	Arg
CTA	L	Leu	CCA	P	Pro	CAA	Q	Gln	CGA	R	Arg
CTG	L	Leu	CCG	P	Pro	CAG	Q	Gln	CGG	R	Arg
ATT	I	Ile	ACT	T	Thr	AAT	N	Asn	AGT	S	Ser
ATC	I	Ile	ACC	T	Thr	AAC	N	Asn	AGC	S	Ser
ATA	I	Ile	ACA	T	Thr	AAA	K	Lys	AGA	R	Arg
ATG	M	Met	ACG	T	Thr	AAG	K	Lys	AGG	R	Arg
GTT	V	Val	GCT	A	Ala	GAT	D	Asp	GGT	G	Gly
GTC	V	Val	GCC	A	Ala	GAC	D	Asp	GGC	G	Gly
GTA	V	Val	GCA	A	Ala	GAA	E	Glu	GGA	G	Gly
GTG	V	Val	GCG	A	Ala	GAG	E	Glu	GGG	G	Gly

Código genético de mitocôndrias de vertebrados											
TTT	F	Phe	TCT	S	Ser	TAT	Y	Tyr	TGT	C	Cys
TTC	F	Phe	TCC	S	Ser	TAC	Y	Tyr	TGC	C	Cys
TTA	L	Leu	TCA	S	Ser	TAA	*	Ter	TGA	W	Trp
TTG	L	Leu	TCG	S	Ser	TAG	*	Ter	TGG	W	Trp
CTT	L	Leu	CCT	P	Pro	CAT	H	His	CGT	R	Arg
CTC	L	Leu	CCC	P	Pro	CAC	H	His	CGC	R	Arg
CTA	L	Leu	CCA	P	Pro	CAA	Q	Gln	CGA	R	Arg
CTG	L	Leu	CCG	P	Pro	CAG	Q	Gln	CGG	R	Arg
ATT	I	Ile i	ACT	T	Thr	AAT	N	Asn	AGT	S	Ser
ATC	I	Ile i	ACC	T	Thr	AAC	N	Asn	AGC	S	Ser
ATA	M	Met i	ACA	T	Thr	AAA	K	Lys	AGA	*	Ter
ATG	M	Met i	ACG	T	Thr	AAG	K	Lys	AGG	*	Ter
GTT	V	Val	GCT	A	Ala	GAT	D	Asp	GGT	G	Gly
GTC	V	Val	GCC	A	Ala	GAC	D	Asp	GGC	G	Gly
GTA	V	Val	GCA	A	Ala	GAA	E	Glu	GGA	G	Gly
GTG	V	Val i	GCG	A	Ala	GAG	E	Glu	GGG	G	Gly

Diferenças para o código padrão:		
AGA	Ter *	Arg R
AGG	Ter *	Arg R
AUA	Met M	Ile I
UGA	Trp W	Ter

Quadro 1 - Códigos genéticos: padrão e mitocôndria de vertebrados.

Legenda: i - iniciação; ter - terminação.

Fonte: <http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/index.cgi?chapter=tgencodes#SG1>

No código genético três códons (ATT, TAG, TGA) não correspondem a aminoácidos; essas trincas são sinais de final de tradução (códons de parada). O estudo do processo de tradução em várias espécies de organismos revelou que o código usado para conversão de nucleotídeos em aminoácidos é o mesmo. São poucas as exceções (veja na parte final do quadro) e, por isso, o código genético é chamado de *universal*.

Outras propriedades do código genético: a *redundância* ou *degeneração*, a qual é o fato de um aminoácido ser codificado por mais de um códon. Códons sinônimos codificam para o mesmo aminoácido (veja mutações sinônimas na Unidade 3). A *oscilação* corresponde ao fato de que muitos códons sinônimos diferem na terceira posição. O anticódon é encontrado no RNAt, o qual tem que ser complementar ao códon. Em 1966, Francis Crick desenvolveu a hipótese da oscilação, que propôs que pareamentos não padrões de bases podiam ocorrer na terceira base, ficando mal pareados ou *oscilantes*. Assim, de 30 a 50 tRNAs são suficiente para a cobertura dos 61 códons do código genético. Os tRNAs são chamados, assim, de *isoaceptores*, pois, apesar de ter diferenças na terceira base de seu anticódon, levam o mesmo aminoácido.

A matriz de leitura corresponde aos conjuntos (de três nucleotídeos) possíveis, de códons para cada sequência. Qualquer sequência de nucleotídeos possui potencialmente três matrizes de leitura. Lembre-se de que quem determina a matriz é o códon iniciador, o qual é o primeiro códon a especificar um aminoácido. O primeiro códon geralmente é o AUG (metionina), assim, todas as proteínas iniciam sua cadeia com uma metionina (apesar de ser retirada posteriormente). Nas bactérias, temos um tipo modificado de metionina que inicia as proteínas bacterianas, a N-fomilmetionina, sendo que esse grupo formil pode também ser retirado posteriormente.

Finalmente, conforme já comentado, o código genético é, geralmente, *não-superposto*. Alguns vírus mostram código superposto, assim, duas proteínas podem ser codificadas na mesma sequência.

**Código genético** não deve ser usado para designar a informação genética. Muitas vezes código genético aparece na mídia como sinônimo para DNA, genoma, sequência de nucleotídeos, teste do DNA, projeto genoma humano etc. O código genético consiste na chave de codificação do alfabeto de 4 letras do DNA para as 20 letras das proteínas. Veja esta citação: “*Em jogo está a glória de ser o primeiro a decifrar o código genético do homem, o chamado genoma humano.*” revista *Veja* (20 de outubro de 1999 - p. 106 seção *Ciência*). Comentário: O código genético foi decifrado em 1965.



## Texto complementar

### Estará a sociedade preparada?

Nota: O texto abaixo foi publicado em 1967.

Novas informações estão sendo obtidas no campo da genética bioquímica numa taxa extremamente rápida. Até agora, este conhecimento teve relativamente pouco efeito sobre o homem. Mais informação deve ser obtida antes que aplicações práticas sejam possíveis sendo que os problemas técnicos a serem resolvidos são formidáveis. Entretanto, quando estes obstáculos forem removidos este conhecimento vai influenciar de forma marcante o futuro do homem. Tal poder pode ser usado sabiamente ou não, para melhoramento ou detrimento da humanidade.

Salvador Luria disse "o progresso da ciência é tão rápido que cria um desequilíbrio entre o poder nas mãos dos homens e as condições sociais nas quais este poder é exercido". Então, nem o aviso dos cientistas, a extensão da informação pública ou o desejo dos cidadãos podem compensar pelas inadequações das instituições de lidarem com as novas situações.

O público entende em alguma extensão os recentes avanços na genética bioquímica, mas tem somente uma vaga noção do que pode se esperar para o futuro, a despeito dos esforços de muitos cientistas em informar o público sobre prováveis derivações futuras.

Onde estamos hoje? A linguagem genética agora é conhecida e está claro que a maioria, se não todas as formas de vida do planeta, usa esta mesma linguagem, com menores variações. Mensagens genéticas simples agora podem ser quimicamente sintetizadas. Cirurgia genética, aplicada à microorganismos é uma realidade. Genes podem ser preparados de uma estirpe bacteriana e inseridos em outra, a qual pode ser mudada geneticamente. Tais mudanças são herdáveis. Até agora isto não foi possível ser feito em células mamíferas.

O que podemos esperar para o futuro? Mensagens genéticas curtas mas significativas podem ser sintetizadas quimicamente. Desde que as instruções sejam escritas na linguagem reconhecível pelas células, a mensagem poderá ser usada para o programa celular. As células poderão carregar estas instruções e o programa poderá ser herdado. Eu não sei quanto tempo teremos antes que seja possível programar células com mensagens quimicamente sintetizadas. Certamente, os obstáculos experimentais são formidáveis. Entretanto, eu tenho poucas dúvidas que os mesmos eventualmente serão transpostos. A única questão é quando. Meu palpite é que as células serão programáveis com mensagens sintéticas dentro de 25 anos. Se os esforços nestas linhas forem intensificados, as bactérias poderão ser programáveis dentro de 5 anos.

O ponto que merece uma ênfase especial é que o homem pode ser capaz de programar suas próprias células com informações sintéticas antes que seja capaz de avaliar adequadamente as consequências de longa data destas alterações, antes de ser capaz de formular objetivos e antes de resolver problemas éticos e morais que ainda serão levantados. Quando o homem tornar-se capaz de instruir suas próprias células ele deve refletir se ele tem desejo suficiente para usar este conhecimento para o benefício da humanidade. Afirmo este problema na necessidade de resolvê-lo, porque decisões concernentes a aplicação do conhecimento deve ser feita em última análise pela sociedade, e somente uma sociedade informada pode tomar tais decisões sabiamente.

Fonte: Nirenberg (1967).

## Síntese da Parte



Na replicação, a dupla hélice é desenrolada na forquilha de replicação, os nucleotídeos são polimerizados pela DNA polimerase, juntamente com a ação de várias outras enzimas. Na transcrição, temos a enzima RNA polimerase a qual faz a cópia em RNA dos genes. Os principais componentes da maquinaria de tradução são os três tipos básicos de RNA: RNA<sub>t</sub>, RNA<sub>r</sub> e RNA<sub>m</sub>. A precisão da tradução deriva da especificidade de códon com anticódon, da ligação do RNA<sub>t</sub> com seu aminoácido e do início da tradução, em que o primeiro códon deve ser reconhecido pelo ribossomo. Um complexo de iniciação, o qual reúne todos os componentes, com exceção feita a alguns fatores, reúne as duas subunidades do ribossomo aos RNAs citados. O alongamento é feito com o avanço do ribossomo no RNA<sub>m</sub> até chegar a um códon de finalização.

## Atividades de avaliação



1. Procure em livros de Bioquímica os conceitos de anabolismo e catabolismo. Dê exemplos.
2. Busque o papel das diferentes RNAs polimerases de eucariotos. Faça um contraponto com os tipos e funções de DNAs polimerases em procariotos.
3. Busque o passo a passo das ligações das snRNPs na formação do spliceossomo.
4. Na leitura complementar apresentada *“Estará a sociedade preparada?”* de 1967, o pesquisador Marshall W. Nirenberg diz *“Meu palpite é que as células serão programáveis com mensagens sintéticas dentro de 25 anos.”* Após a leitura da Unidade 5, comente sobre as previsões de Nirenberg.
5. Ainda sobre esse texto, o que é possível dizer sobre a afirmação de Nirenberg: *“está claro que a maioria, se não todas as formas de vida do planeta, usa esta mesma linguagem, com menores variações?”*
6. Observando o quadro que apresenta os códigos genéticos, que diferenças encontramos entre os dois códigos apresentados? Por que ainda consideramos o código genético *universal*? Verifique em outros livros o código genético escrito com as letras do RNA<sub>m</sub>. Que diferenças são verificadas?

## Leituras, filmes e sites



### Livros

- STRATHERN, P. Crick, Watson e o DNA em 90 minutos. Rio de Janeiro: Jorge Zahar, 2007. 96 p.
- CHEDIAK, K. Filosofia da biologia. Rio de Janeiro: Jorge Zahar, 2008. 84 p.
- DAWKINS, R. A grande história da evolução. São Paulo: Companhia das Letras, 2009. 472 p.

### Filmes

- Criação (Dir. Jon Amiel, 2009) - história da "Origem das espécies" de Darwin.
- O vento será sua herança (Dir. Stanley Kramer, 1960 e também refilmagem de Daniel Petrie de 1999) - famoso "julgamento do macaco" que ocorreu no Tennessee – EUA, em 1925, sobre o ensino do darwinismo nas escolas.
- O pesadelo de Darwin (Dir. Hubert Sauper, 2004) - Ecologia e ambiente.

### Sites

#### (Museus e exposições científicas):

- <http://www.museunacional.ufrj.br/> (Museu Nacional)
- <http://www.cienciaviva.org.br/> (Espaço Ciência Viva)
- <http://www.eciencia.usp.br/> (Estação Ciência da USP)
- <http://www.museufritzplaumann.ufsc.br/> (Museu Entomológico Fritz Plaumann)
- <http://www.mos.org/> (Boston Museum of Science - EUA)
- <http://www.miamisci.org/> (Miami Science Museum - EUA)
- <http://www.mnh.si.edu/> (Museu de História Natural do Smithsonian Institute - EUA)

- <http://www.ufrgs.br/ceclimar/> (Museu do Centro de Estudos Costeiros, Limnológicos e Marinhos - CECLIMAR - UFRGS)
- <http://www.fiocruz.br/museudavida/> (Museu da Vida - FIOCRUZ)
- <http://www.museuhistoriconacional.com.br/> (Museu Histórico Nacional - MHN)
- <http://www.museudomar.com.br/> (Museu do Mar – Santos/SP)

## Referências



ARNOULT, N.; SCHLUTH-BOLARD, C.; LETESSIER, A.; DRASCOVIC, I.; BOUARICH-BOURIMI, R.; CAMPISI, J.; KIM, S.; BOUSSOUAR A.; OTTAVIANI, A.; MAGDINIER, F.; GILSON, E.; LONDONO-VALLEJO, A. Replication timing of human telomeres is chromosome arm-specific, influenced by sub-telomeric structures and connected to nuclear localization. **PLoS Genetics**, Detroit, v. 6, n. 4, p. 1-15, 2010.

BERG, J. M.; TYMOCZKO, J. L.; STRYER, L. **Bioquímica**. 6. ed. Rio de Janeiro: Guanabara Koogan, 2008. 1.114 p.

LODISH, H.; BERK, A.; ZIPURSKY, S. L.; MATSUDAIRA, P.; BALTIMORE, D.; DARNELL, J. E. **Molecular Cell Biology**. 4. ed. New York: W. H. Freeman and Company, 1999. 1.084 p.

NIREMBERG, M. W. Will society is prepared? **Science**, New York, v. 157, n. 3789, 1967.

PIERCE, B. **Genética**: um enfoque conceitual. 1. ed. Editora Guanabara Koogan: Rio de Janeiro, 2004. 788 p.

WU, Z. A., MURPHY, C.; GALL, J. G. Small nuclear ribonucleoproteins and heterogeneous nuclear ribonucleoproteins in the amphibian germinal vesicle: loops, spheres, and snurposomes. **Journal of Cell Biology**, New York, v. 113, n. 3, p. 465-483, 1991.



**Parte**

**3**

**Polimorfismo gênico e  
evolução molecular**



# Considerações iniciais e histórico

## Objetivo

- Discutir variações genéticas e moleculares dentro do contexto evolutivo.

## 1. Aspectos gerais

Durante a maior parte do século vinte, o foco das pesquisas moleculares foi o reconhecimento dos fenótipos, identificando os produtos dos genes, as proteínas. A partir do sequenciamento gênico vimos que, surpreendentemente, somente uma parte do genoma eucariótico é destinada a essas proteínas. O genoma é um sistema dinâmico, sempre em reparo e remodelamento. É mutável durante o ciclo de vida do indivíduo e durante décadas e milênios, durante a evolução de uma espécie, claro, dentro dos limites. As variações encontradas no genoma variam de pessoa para pessoa, desde mudanças de uma base até rearranjos cromossômicos em larga escala. Dessa forma, não existe um “genoma humano”, genérico. Não existe o “normal” em qualquer sequência. O chamado “tipo selvagem” ou *wild type* indica a versão mais comum ou mais frequente, de uma característica genética na população, por exemplo, o olho vermelho da *Drosophila* em relação ao olho branco, raro, o qual seria o “mutante”. O olho é um fenótipo facilmente verificável, porém muitas variações não têm fenótipo detectável, portanto, as variantes podem ser silenciosas, de tal forma que não alteram visivelmente o fenótipo. A evolução molecular está em curso, porém ainda não chegamos a compreender totalmente suas características.

Na agricultura e na pecuária, o conhecimento gênico é um grande aliado, pois permite cruzamentos seletivos que podem aumentar ou diminuir a frequência gênica de determinadas características genéticas desejáveis de plantas e animais. Essa *seleção artificial* proporcionou enormes vantagens de sobrevivência para o ser humano (embora estejamos pagando o altíssimo preço ambiental desse estilo de vida). Essa seleção não é aplicável ao ser humano, pois não é ético interferir nas liberdades individuais. O aconselhamento genéti-

**Substâncias tóxicas:** Uma substância é considerada tóxica quando causa dano ao organismo. A injúria tóxica, nesse sentido, refere-se a uma perigosa alteração, real ou potencial, da célula ou dos tecidos constituintes e os processos, os quais podem ou não ser detectados por sintomas clínicos bem claros. Uma substância pode não ser, na realidade, tóxica, mas as condições são tais que ela pode apresentar risco à saúde. Nem todas as substâncias tóxicas são sintéticas, muitas são naturais, resultado dos processos metabólicos dos organismos como a solanina. A solanina é uma toxina encontrada em batatas e solanáceas. É um esteroide, glicoalcaloide. As batatas contêm de 2 a 15 mg/100 g de peso úmido de batatas. Ela é um inibidor da enzima acetilcolinesterase, do sistema nervoso. É necessária uma dose muito alta para atingir efeitos tóxicos, pois ela tem baixa absorção pelo trato gastrointestinal. Há poucos estudos sobre sua mutagenicidade. O que define a toxicidade não é a substância, mas a quantidade ingerida. A ricina, extraída da planta mamona, é uma das três substâncias que acredita-se serem as mais tóxicas. Ela é o legendário veneno do guarda-chuva, possivelmente usado por espões. Um mg de ricina pode matar um homem adulto. As outras duas substâncias mais potentes são o antrax e a toxina botulínica. Pesquise sobre esse interessante assunto.

co, efetuado por especialistas médicos, é um serviço de grande alcance social, que busca evitar o sofrimento dos pacientes com anomalias. Existem defensores da ideia de que os seres humanos deveriam evitar determinados casamentos. Porém, existem extremistas que defendem, mesmo de forma velada, na “melhoria” da espécie humana para características consideradas “desejáveis” como inteligência, beleza, altura etc. Por outro lado, há correntes que defendem que indivíduos suspeitamente propensos à violência, indivíduos de determinados grupos étnicos, homo ou bissexuais, pessoas com desvios sexuais tidos como indesejáveis (como a pedofilia) etc, sejam impedidos de procriar e sejam até, castrados, ou mesmo mortos. A chamada *eugenia* buscava aumentar a frequência das “boas” características através de métodos absurdos, de castração até genocídio em massa, como ocorreu durante o nazismo.

Portanto, a humanidade não existe para ser “melhorada” e, como disse Jacob Monod, não parece existir planos evolutivos que estejam dirigindo a evolução (humana) para esta ou aquela direção. Mesmo que a pesquisa biológica futura indique, indubitavelmente, que certos comportamentos complexos (e mesmo perigosos) sejam fortemente determinados no genoma, a ética deve sim, preponderar em todos os sentidos em nossos estudos e atitudes. Muitas características são claramente deletérias; outras, nem tanto, e outras podem ser até vantajosas em alguns casos. Quem deve ter a última palavra e determinar o que é adequado ou não adequado? O uso do *determinismo genético*, ou seja, a ideia de que somos o que nossos genes determinam, precisa e deve ser evitado a todo custo, especialmente à luz do conhecimento filosófico, da ética e da moderna Biologia.

## 2. Breve histórico

O estudo das mutações está estreitamente ligado ao estudo químico, histórico, de **substâncias tóxicas**. Especialmente de uso militar, essas substâncias tóxicas podiam matar ou prejudicar o ser humano, além de animais e plantas. Dessa forma, eram sistematicamente estudados nos laboratórios, para verificar seu potencial mutagênico. As mutações podem surgir espontaneamente, mas sempre se pergunta se alterações ambientais não teriam causado tal mutação. Algumas substâncias químicas e radiação possuem ação mutagênica, embora seus efeitos variem consideravelmente em relação às doses aplicadas, à concentração dos reagentes, intensidades utilizadas e outros fatores variáveis.

Em 1927, o pesquisador americano Joseph Hermann Muller demonstrou que as mutações em moscas podiam ser induzidas por radiações, como os raios X. Os raios X aumentavam muito o número de mutações nas gerações posteriores, em todos os organismos estudados. Junto com Thomas Hunt Morgan, foi um dos redescobridores do mendelismo, publicando o livro *The Mechanism of Mendelian Heredity* conforme já contamos.

A geneticista Charlotte Auerbach, nascida em Berlim, em 1899, migrou para a Inglaterra, onde conduziu pesquisas sobre mutações em *Drosophila*. Na Inglaterra encontrou Muller, que mostrou que a radiação induz mutações. Charlotte, então, começou a tentar induzir mutações em *Drosophila*, com substâncias químicas, porém sem muito sucesso. Na época havia muitos pesquisadores desenvolvendo o gás mostarda (utilizado na Primeira Guerra Mundial). Colaborando com o farmacologista James Michael Robson. Eles mostraram que o gás mostarda era um potente mutagênico, reduzindo a viabilidade dos gametas e aumentando o número de mutações nas gerações de moscas. Ela experimentou outros agentes químicos sobre *Drosophila*, *Neurospora* e leveduras. Suas experiências, comparando os raios X e as mutações induzidas quimicamente, mostraram que os raios X causaram mutações em todo o corpo, enquanto que os produtos químicos causaram mutações em apenas *algumas* partes do corpo. Outro resultado mostrava que os raios X produziram mais alterações cromossômicas do que os mutagênicos químicos. A explicação para isso, segundo Auerbach, foi que os raios X produzem alterações genéticas quase que imediatamente. Os produtos químicos são mais atrasados, e apenas alguns genes foram afetados após várias divisões celulares.

Em agosto de 1945, ocorreu o lançamento da bomba atômica sobre as cidades de Hiroshima e de Nagasaki, no Japão. Nesse momento, foram liberadas imensas quantidades de radiação. O acidente no reator nuclear de Chernobyl, no nordeste da Ucrânia, ocorreu em 1986. Estima-se que ocorreu a liberação de cerca de 100 milhões de Curies, o equivalente radioativo a uma bomba atômica, a qual seria pouco menor que as lançadas anteriormente. Evidentemente, a ciência e a tecnologia dos anos 80 possibilitaram um melhor entendimento das patologias e das alterações em decorrência dessa tragédia. Considera-se que a taxa de câncer de tireoide de crianças em toda a Ucrânia seja cerca de dez vezes maior que os níveis pré-explosão. Evidentemente, esses lamentáveis acontecimentos lançaram inúmeras questões sobre o papel das radiações nas frequências populacionais relacionadas a diversas patologias, tanto congênitas quanto não congênitas. As sequências de DNA dessas populações foram examinadas, evidenciando o potencial nocivo das radiações.

Caso fosse possível concluir algo sobre esses terríveis acontecimentos, vemos que, apesar do avanço científico que essas técnicas propiciaram, do outro lado da moeda temos sérios questionamentos éticos que o mau uso delas causou. Cientistas criaram a bomba atômica e sabe-se, que mesmo hoje em dia, suspeita-se do desenvolvimento de armas químicas e biológicas. Considere a ideia de que o cientista é neutro, imparcial, incorruptível, um ser quase divino, acima do bem e do mal como uma falácia, um mito.

**Efeitos biológicos das radiações:** As células, quando expostas à radiação, sofrem ação de fenômenos físicos, químicos e biológicos. A radiação causa ionização dos átomos, que afeta moléculas, que poderão afetar células, que podem afetar tecidos, que poderão afetar órgãos, que podem afetar a todo o corpo. Os linfócitos (glóbulos brancos) e das células que produzem sangue estão em constante reprodução e são as mais sensíveis às radiações. As células reprodutivas e gastrointestinais não se reproduzem tão rápido, portanto, são menos sensíveis. Células nervosas e musculares são as mais lentas e, portanto, as menos sensíveis. Os efeitos somáticos classificam-se em imediatos e retardados com base num limite adotado por convenção, de 60 dias. O mais importante dos efeitos imediatos das radiações após exposição do corpo inteiro a doses relativamente elevadas é a Síndrome Aguda de Radiação (SAR). O efeito retardado de maior relevância é a formação de tumores, os quais, geralmente, só aparecem vários anos após a irradiação. A SAR varia conforme a dose e a absorção diferencial dos órgãos do corpo.

# Variação gênica

## 1. Bactérias, genes e plasmídeos

As bactérias são organismos-chave para o entendimento da Biologia Molecular. São nelas que os mais importantes princípios bioquímicos e moleculares foram descobertos. As bactérias, como *E. coli*, possuem um ou dois cromossomo circulares. Os plasmídeos são elementos genéticos extracromossômicos, pequenas moléculas de DNA circular. Os plasmídeos possuem tamanho variável, de poucos milhares a mais de cem mil pares de bases. Em comparação ao genoma bacteriano, como *E. coli*, a qual possui cerca de 4,6 milhões de pares de bases, são relativamente pequenos. As bactérias possuem um ou mais plasmídeos. Em geral, não possuem genes essenciais para o funcionamento metabólico da célula bacteriana, porém são essenciais para a reprodução, pois alguns promovem a reprodução da bactéria, outros contêm genes para eliminar outras bactérias, portanto, possuem uma função na sobrevivência e muitos contêm genes de resistência a antibióticos. A chamada *competência* bacteriana é uma característica de algumas espécies que são mais capazes de captar o DNA exógeno. Sob condições apropriadas, qualquer DNA pode ser transferido para células competentes, mesmo sendo não bacteriano. As células que recebem o DNA são chamadas *transformantes*.

Como possui sua própria origem de replicação, o DNA se reproduz independentemente do cromossomo bacteriano. *Epissomos* são plasmídeos, capazes de se replicar independentemente ou de se integrar ao cromossomo bacteriano. As bactérias podem transferir e recombinar informações genéticas através de três mecanismos diferentes: conjugação, transformação e transdução. A *conjugação* é a transferência direta de plasmídeos através de uma conexão; na *transformação*, ocorre incorporação do DNA do meio, sendo este injetado dentro da bactéria e podendo sofrer troca de genes, ou recombinação com o cromossomo bacteriano. Na *transdução*, os vírus bacterianos, como os bacteriófagos, levam DNA de uma bactéria para outra. Desta forma genomas bacterianos são remodelados e permite evolução em um organismo que não possui atividade sexual.

### Teoria endossimbiótica:

considera que as mitocôndrias e os cloroplastos já foram bactérias de vida livre que se tornaram endossimbiontes de células eucarióticas primitivas.

### Heteroplasmia e

**homoplasmia:** a presença de mutações no DNA plastídico é comum.

Quando numa célula temos mitocôndrias contendo DNA mutante e mitocôndrias contendo DNA não mutante, chamamos esta condição de "heteroplásmica".

Quando todas são iguais (todas mutantes ou todas não mutantes), chamamos 'homoplasmia'.

## 2. Genomas extra nucleares: mitocôndrias e cloroplastos

Sabemos que, além do genoma nuclear, existem outros genomas, nas mitocôndrias e nos cloroplastos. A teoria **endossimbiótica** diz que um precursor da célula eucariótica (proto eucarioto) endocitou uma possível proto bactéria ou célula procariótica. As mitocôndrias e os cloroplastos teriam sido formados pela evolução dessas estruturas (Figura 1). Acredita-se que este evento tenha ocorrido entre 1 e 1,5 bilhão de anos atrás. As bactérias fotossintéticas e as células eucarióticas também levariam à evolução dos cloroplastos.

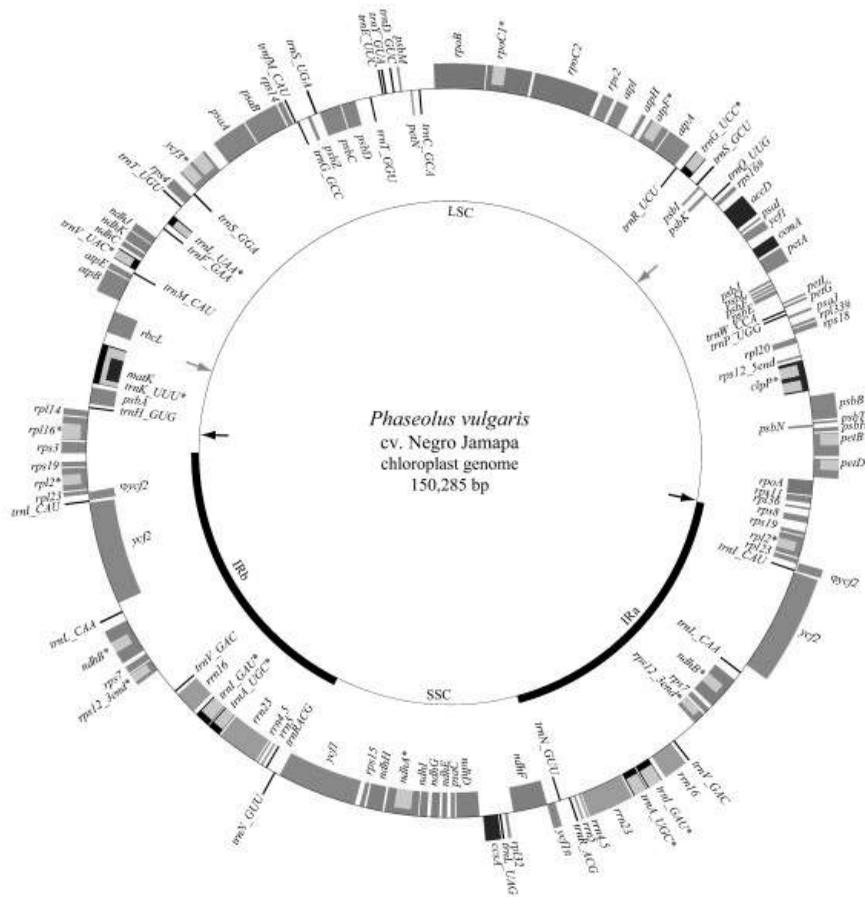


Figura 1 - Mapa esquemático do genoma do cloroplasto de *Phaseolus vulgaris* (*plastoma*). Os genes que aparecem fora do mapa são transcritos na direção do relógio e na parte interna contra. Íntrons são indicados por um asterisco.

Fonte: Guo et al. (2007).

O genoma mitocondrial das células atuais é muito pequeno. A maioria das proteínas necessárias para o metabolismo da mitocôndria é obtida a partir do genoma nuclear. Comparativamente às células bacterianas atuais, as quais possuem poucos Megabase de DNA e algumas centenas ou milhares

**Fosforilação oxidativa:** processo que gera ATP nas mitocôndrias. Ocorre na membrana interna da mitocôndria, requerendo vários transportadores de elétrons diferentes, proteínas codificadas pelo mtDNA e genes nucleares.

**Radicais livres:** é um termo usado para designar átomo ou molécula com existência independente, contendo um ou mais elétrons não pareados, nos orbitais externos. São altamente reativos, capazes de reagir com qualquer composto situado próximo à sua órbita externa, passando a ter função oxidante ou redutora de elétrons. Organelas que metabolizam o oxigênio, o nitrogênio e cloro, que geram grande quantidade de metabólitos. São as grandes produtoras de radicais livres.

**Padrões isoenzimáticos:**

enzimas que apresentam padrões de isoformas enzimáticas e, portanto, na eletroforese de proteínas formam muitas bandas, ou seja, são altamente polimórficas e dessa forma, permitem que se verifique a variabilidade genotípica.

**Insuficiência cardíaca e polimorfismo gênico:**

apesar dos avanços terapêuticos no último século, a insuficiência cardíaca continua sendo uma das principais causas de mortalidade, de internação hospitalar e de limitação funcional em todo o mundo. Nas últimas décadas, a influência dos fenótipos vem sendo utilizada para a avaliação do prognóstico dos pacientes, porém, em um grande número de casos, pacientes com fenótipos semelhantes evoluem de formas distintas. Há, portanto, a necessidade de se pesquisar os fatores genéticos que possam influenciar em um prognóstico individualizado. Os principais polimorfismos envolvidos são os genes do receptor beta adrenérgico, o da enzima conversora de angiotensina, da aldosterona, da angiotensina, do fator de necrose tumoral (TNF- $\alpha$ ) e do óxido nítrico.

de genes, considera-se que esses genes teriam sido transferidos para o genoma da célula hospedeira. O genoma mitocondrial consiste em uma única molécula de DNA circular, altamente helicoidizada. As plantas contêm mitocôndrias que comportam um conjunto complexo de múltiplas moléculas circulares de DNA. Como foi comentado anteriormente, os genomas mitocondriais são muito pequenos em relação ao núcleo celular. Não possui histonas, da mesma forma que as bactérias. Outra particularidade é que possui conteúdo GC diferente do genoma, nuclear de forma que o genoma mitocondrial pode ser separado por uma centrifugação diferencial.

As mitocôndrias e os cloroplastos presentes no citoplasma são, em geral, herdados de um só genitor (*herança uniparental*). Nos animais, o mtDNA é herdado quase que exclusivamente da mãe, embora a literatura acuse exceções. As células podem conter dezenas a centenas de organelas, sendo que cada uma pode conter muitas cópias do genoma dessa organela. Uma célula de fígado de rato, por exemplo, pode conter 5 a 10 moléculas de mtDNA em cerca de 1000 mitocôndrias, e o mtDNA equivale a 1% do DNA celular. É possível que haja populações de organelas com mutações diferentes, convivendo na mesma célula, uma condição chamada **heteroplasmia**. Quando essa célula se divide, sua prole pode apresentar fenótipos diferentes.

O mtDNA humano é uma molécula circular com 16.569 pb que codificam dois rRNAs, 22 tRNAs e 13 proteínas. Essas proteínas auxiliam na fosforilação oxidativa, mas não suprem totalmente a demanda de proteínas para essas reações, as quais são complementadas pelos genes nucleares. Essa relação é essencial para se compreender porque certas patologias e, até mesmo, o próprio envelhecimento do organismo, causam diminuição na quantidade de ATP. Estudos mostram que o mtDNA, com a idade, acumula mutações, especialmente pela proximidade da produção de **radicais livres** produzidos pelos processos citados. Os tecidos, como músculo e cérebro, necessitam de ATP em grande quantidade e com a idade, e certas patologias associadas podem apresentar sintomas como deficiência no movimento, devido à demanda de ATP pela musculatura esquelética, diminuição no raciocínio e na memória, devido à demanda nervosa e cerebral, além dos rins, fígado e pâncreas.

**3. Polimorfismo e marcadores moleculares**

Observando o genoma, torna-se possível verificar vários tipos de polimorfismos: polimorfismo de **comprimento de fragmentos de restrição** (RFLP), tratado na Unidade 5, consiste de um polimorfismo que depende do genoma ser cortado com enzimas de restrição, portanto, baseia-se totalmente na técnica. O polimorfismo de curtas repetições em *tandem* são repetições de dezenas a centenas de bases, sendo que o número de polimorfismos é herdável. A

aplicação mais conhecida é o teste de paternidade. As repetições podem ser detectadas por *Southern Blot* ou por PCR (veja também a Unidade 5). O polimorfismo de um único nucleotídeo (SNP) foi comentado na Unidade 1.

Um problema importante em biologia é a discriminação de genótipos. Uma área que chama a atenção dos biólogos, porém relativamente confusa em termos de definição, os *marcadores moleculares* é uma expressão coletiva para um grupo de recursos destinado a várias funções em Bioquímica, Biologia Molecular e Genética. Assim, em razão do seu nível de polimorfismo, podemos ter sequências (proteínas e DNA) que podem discriminar grupos, subgrupos ou mesmo indivíduos em particular. Os marcadores vieram em função da necessidade de novos descritores, além dos utilizados tradicionalmente pelos biólogos. Os descritores morfológicos são características morfológicas descritas pelos geneticistas, por exemplo, quando precisam descrever ou cultivar uma nova variedade.

A década de 70 viu o surgimento de descritores baseados em enzimas e proteínas. Várias enzimas contêm múltiplas **isofomas** (isoenzimas) que aparecem especificamente em determinados genótipos, influenciando na atividade e caracterizando diferentes valores quantitativos e qualitativos. Dez anos após os primeiros trabalhos com descritores de proteínas e enzimas, surgiram os de DNA, com capacidade de detectar variação genética adicional. O avanço principal que essas técnicas trazem é a possibilidade de acessar diretamente o genótipo de um indivíduo, evitando, assim, a expressão do fenótipo e a influência do ambiente sobre ele. Também, enquanto os descritores de proteína amostram apenas as regiões ativas na expressão gênica, os marcadores de DNA permitem uma ampla amostragem do genoma de um indivíduo. Mutações que ocorrem em regiões não codificadoras de genes podem ser identificadas com análise de DNA, mas não com a análise morfológica ou de proteínas. Entretanto, marcadores bioquímicos ainda são muito utilizados.

Em tese, qualquer sequência de DNA pode ser chamada de marcador molecular (veja *Fingerprint*, na Unidade 5). Essas sequências são, geralmente, muito polimórficas e sempre são herdáveis. Existe um número quase ilimitado de descritores de DNA disponíveis, possibilitando o acesso da variabilidade genética utilizáveis para quaisquer tipos de organismo. O problema a ser resolvido no laboratório é como visualizar e quantificar o polimorfismo, para que seja possível fazer as devidas comparações entre os genótipos.

As técnicas incluem o uso de sondas, reação em cadeia da PCR, eletroforese, análise de fragmentos de restrição, comparações via sequenciamento do DNA, enfim, praticamente quase todos os recursos disponíveis (veja Unidade 5).

O homem domesticou, na sua existência, somente cerca de cem a duzentas, de milhares de espécies vegetais. Destas, menos de 15 atualmente suprem a maior parte da dieta humana. Essas 15 espécies podem ser agrupadas nas seguintes classes: cereais (arroz, trigo, milho, sorgo e cevada); raízes e caules (beterraba, cana-de-açúcar, batata, mandioca e inhame); legumes (feijão, soja e amendoim) e frutas (citros, banana, manga etc). Quando o melhoramento de plantas teve seu início? O milho e outras as culturas mostram detalhes que atestam que ele começou na mais remota antiguidade e não era um trabalho cientificamente dirigido. O melhoramento executado pelo homem primitivo resultava da simples procura de tipos mais adequados para satisfazer a suas necessidades. O milho é originário do Novo Mundo, onde foi domesticado pelos índios. Há milhares de anos, espalhou-se pelas Américas, a partir do México (provável centro de origem), e, quando aqui aportaram, os europeus encontraram centenas e centenas de variedades conservadas pelas várias tribos indígenas. Principalmente a partir do século XIX, os melhoristas eram apenas pessoas práticas que tinham habilidade de selecionar, dentre muitas outras, as plantas que apresentavam diferenças e que podiam ser de interesse econômico.

### 3.1 Aplicação: melhoramento genético

O melhoramento genético envolve uma série de técnicas genéticas clássicas, como o cruzamento e a seleção das progênes. Para fazer isso, é necessário que o pesquisador tenha à sua disposição um *banco de germoplasma*. Uma coleção, contendo a maioria ou, pelo menos, o máximo número das raças, das variedades, dos tipos e dos subtipos de plantas e animais de interesse econômico, na necessidade de fazer cruzamentos direcionados. Universidades, institutos e empresas particulares sempre buscam agregar novos acessos ao seu plantel. É assim com o café, o trigo, os caprinos, o arroz, as frutas tropicais e muitos outros, até mesmo flores e plantas ornamentais.

Na França, desde 1990, os dados de padrões isoenzimáticos são requeridos para a identificação e o registro de novas linhagens de milho, havendo planos de solicitar os padrões de marcadores de DNA. Mesmo assim, o uso de marcadores moleculares para registro de novas variedades ainda é limitado, mesmo em países onde a proteção de cultivares já há praticada a bastante tempo. No entanto, com a recente disponibilidade de tecnologias de DNA que revelam alto nível de polimorfismo e consistência nos resultados, é provável que os marcadores moleculares passem a ser incluídos no registro de novos genótipos. No Brasil, com a Lei de Proteção dos Cultivares, os marcadores tornaram-se essenciais na identificação e no reconhecimento de germoplasma em plantas (veja o texto complementar).

### 4. Alterações cromossômicas

A *poliploidia* ocorre em células e organismos quando possuem mais de dois pares de conjuntos homólogos de cromossomos. Mudanças no número de cromossomos podem ser de dois tipos: *euploidia aberrante* e *aneuploidia*. No primeiro caso, temos mudanças em conjuntos completos de cromossomos e, no segundo, somente em partes dos cromossomos (Veja o item 7.7 mutações). A poliploidia é especialmente comum em plantas (Figura 2), também ocorre em animais como *Carassius auratus auratus* (peixe dourado), no salmão e nas salamandras. A poliploidia, entretanto, raramente ocorre em humanos, restringindo-se a alguns tecidos, como o fígado e o tecido muscular. Algumas condições patológicas provocam trissomias, tetrassomias e outras condições de variação do número de cromossomos, mas que não se comparam ao verificado em termos de ploidia em outros organismos. Organismos com múltiplos de conjuntos cromossômicos básicos (genoma) são chamados de *euploides*.

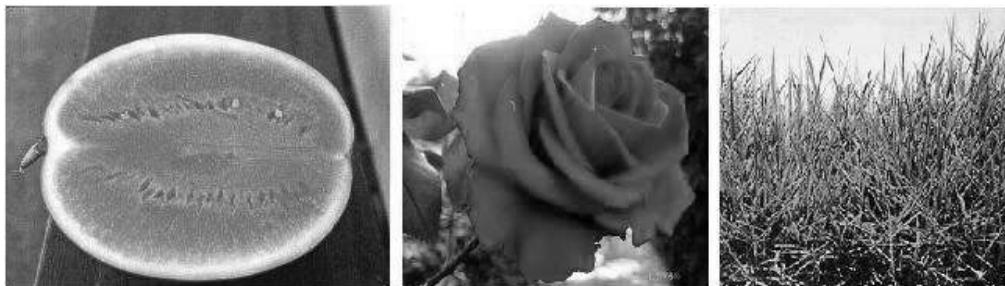


Figura 2 - Melancia sem sementes, rosa gigante e forrageira mais produtiva. A melancia sem semente é uma variedade triploide japonesa, resultado do cruzamento de plantas diploides ( $2n$ ) com tetraploides ( $4n$ ), produzindo sementes triploides ( $3n$ ) inviáveis. Tem mais sabor e polpa, porém necessita de maior quantidade de mão-de-obra para fazer os cruzamentos entre as plantas  $2n$  e  $4n$ . Rosas e gramíneas tetraploides.

Fonte: [http://www.google.com.br/search?hl=pt-BR&q=Melancia+sem+semente&btnG=Pesquisar&meta=&aq=f&aqi=&aql=&oq=&gs\\_rfai=](http://www.google.com.br/search?hl=pt-BR&q=Melancia+sem+semente&btnG=Pesquisar&meta=&aq=f&aqi=&aql=&oq=&gs_rfai=)

Nos organismos poliploides, é possível distinguir os *autopoliploides*, que têm conjuntos cromossômicos múltiplos com origem dentro de uma espécie, e *alopoliploides*, que têm conjuntos de duas ou mais espécies diferentes. Os *alopoliploides* formam-se apenas entre espécies proximamente relacionadas. No caso de conjuntos cromossômicos diferentes, são chamados *homeólogos* (parcialmente homólogos).

Mudanças na estrutura dos cromossomos devidos às quebras no DNA também são conhecidas como *rearranjos*. Por exemplo, um segmento cromossômico pode ser perdido, constituindo uma deleção, ou duplicado (não confundir com duplicação gênica) ou invertido. A translocação move um segmento para outro cromossomo. As causas e como ocorrem esses rearranjos não serão discutidos aqui, mas são facilmente revisáveis nos bons livros de estudos genéticos ou na rede mundial de computadores.

## 5. O valor C em discussão

Em 1948, Roger e Colette Vendrely apresentaram um trabalho que mostrava que o conteúdo do DNA nuclear das células em todos os indivíduos é constante, dentro de uma dada espécie animal. Essa afirmativa demonstrava que o DNA, não a proteína, era a substância na qual os genes são compostos. A quantidade de DNA dentro de uma espécie e dentro de um mesmo indivíduo (ressalvadas condições especiais da presença, em um mesmo organismo, de células com diferentes níveis de ploidia) é aceita como *Constante*. O **valor C** define a quantidade de DNA do genoma haploide, não replicado, de um indivíduo (Figura 3). Esse termo foi sugerido pelo biólogo canadense T. R. Gregory em 2000. Assim, a quantidade de DNA é expressa em picogramas (pg), 10-12g

**Organismos geneticamente modificados** contêm uma sequência nova de DNA ou contêm DNA recombinante. Todo transgênico é um OGM. Existem diferenças sutis entre os dois conceitos. Nem todo OGM é um transgênico, ou seja, um OGM pode ter um gene overexpresso, ou subexpresso, pode não ter sido introduzido um gene, mas feitas alterações que podem aumentar a expressão de um gene. Portanto, o conceito de OGM engloba o de transgênico.

ou em mega pares de bases de nucleotídeos (Mb = 106 pares de bases), sendo que 1pg corresponde a 965Mb. O valor C pode ser muito importante no estudo comparativo de grupos de organismos em várias áreas do conhecimento biológico, como a fitogeografia e a sistemática.

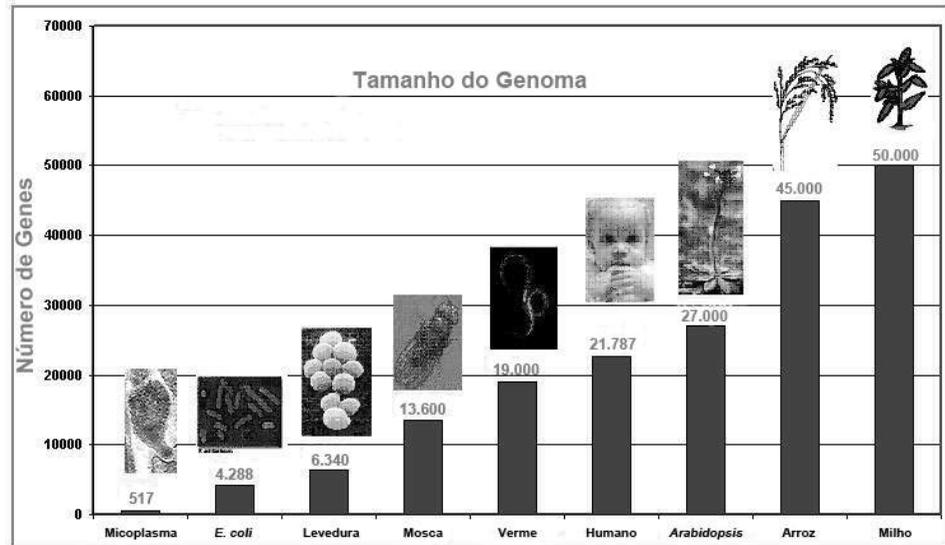


Figura 3 - Número de genes em vários organismos.

Fonte: <http://www.bio.miami.edu/~cmallery/150/gene/hisbiotech/hisbiotech.htm>

Alguns autores têm desafiado a premissa do valor constante, citando casos isolados que apresentam alterações na quantidade de DNA que poderia ser influenciada por fatores ambientais, mas salientam a necessidade de repetições e amplas discussões sobre o assunto, o que poderia atualizar o conceito da constância do valor C.

Ao se examinar as grandes linhas evolutivas, é possível perceber uma relação geral entre aumento da quantidade de DNA e complexidade evolutiva. Leveduras possuem 14 Mb; *Caenorabiditis elegans* possui 100 Mb; humanos, 3500 Mb. Profundamente chocante, porém, foi a descoberta que, ao examinarmos os eucariotos, vemos que essa quantidade pode variar em até 80.000 vezes, que essa variação não se condiz com a complexidade dos organismos nem com o número de genes. Veja a Figura 4, apresentando caso do genoma mitocondrial. O caso dos anfíbios é o mais destacado, pois possuem muito mais DNA nuclear do que os mamíferos. Salamandras possuem 40 vezes mais DNA que os humanos. Samambaias, consideradas plantas primitivas, apresentam valores altos em relação às angiospermas. Esse chamado **paradoxo do valor C** aparentemente não tem ainda explicação adequada: Presença de íntrons, transposons, DNA carregado passivamente pelos cromossomos, DNA com funções não gênicas, entre outras.

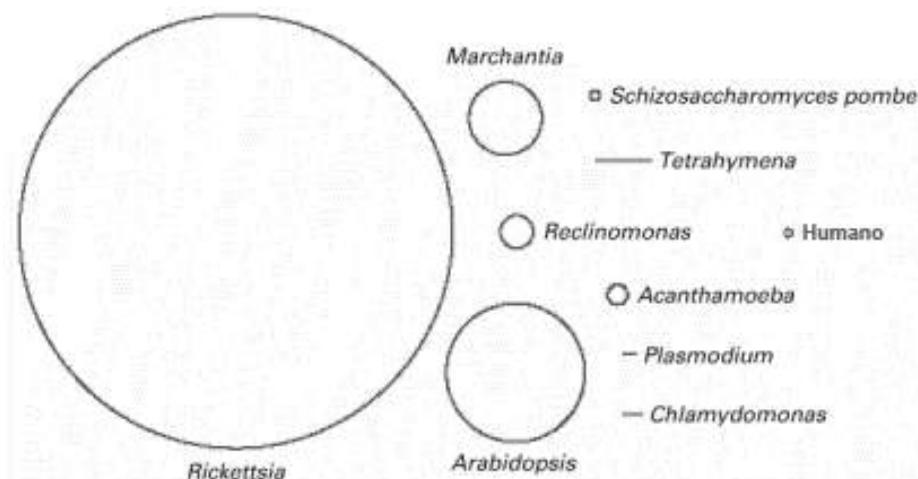


Figura 4 - Mais de 200 genomas mitocondriais já foram determinados. Os tamanhos de alguns desses mtDNA são mostrados em escala. Círculos representam genomas circulares e linhas, lineares. O círculo maior mostra o genoma de *Rickettsia pro-wazekii*, uma pequena bactéria patogênica cujo genoma é o que mais lembra uma mitocôndria. O número de proteínas codificadas não corresponde ao tamanho dos genomas, enquanto o mtDNA humano codifica 13 proteínas; o DNA de *Arabidopsis* codifica somente 32 proteínas (apenas 2,5 vezes mais), sendo 22 vezes maior. O mtDNA do protozoário *Reclinomonas americana* tem 97 genes, mais que qualquer mitocôndria de qualquer organismo analisado.

Fonte: Alberts et al. (2008).

Para obtenção dos valores C, os pesquisadores vêm usando técnicas como microdensitometria por Feulgen, a qual se baseia na ligação específica do DNA ao corante *Feulgen* (ou reagente de *Schiff*). Outra técnica que vem sendo utilizada é a citometria de fluxo, a qual foi inicialmente desenvolvida para contagem e análises rápidas de células sanguíneas. A partir da década de 80, começou a ser adaptada para células vegetais. Essa técnica envolve a análise das propriedades óticas de partículas em fluxo, movendo-se em relação ao ponto de medida. As partículas estão hidrodinamicamente contidas no centro de um fluxo estreito de líquido e passam por um foco de luz intensa, o que permite a excitação dos fluorocromos presentes nas partículas. Os pulsos de luz e de fluorescência são colhidos por um sistema de detecção ótica, separados por filtro e convertidos em pulsos elétricos por sensores óticos.

Um excelente local da rede mundial de computadores para dados sobre o tamanho de genomas (animais) é o site *Animal Genome Size Database*: <http://www.genomesize.com>.

## 6. Elementos de transposição

Barbara McClintock nasceu em 1902 e estudou botânica na Cornell University – EUA. Tentou entrar para a pós-graduação em Genética, mas não foi aceita, por ser mulher. Conseguiu entrar na pós-graduação em Botânica e estudou os cromossomos de milho. McClintock permaneceu na Universidade, estudando ligação entre genes, *crossing over*, mapeamento gênico com rearranjos cromossômicos etc. Um problema da época era a presença de grãos chamados de *variegados* ou multicoloridos. Os grãos eram incolores ou pigmentados, mas nesse caso, tinham faixas de cor. Rollins Adams Emerson, contemporâneo de McClintock, sugeriu que essas faixas coloridas eram resultado de uma mutação instável, porém não havia maiores explicações. McClintock sugeriu, em 1948, que essa mutação era resultado de um gene que se movia, apresentando a hipótese de que havia inclusive genes controladores, mas que também se moviam juntos. Alfred Sturtevan, destacado geneticista da época, comentou: “eu não compreendi uma palavra do que ela disse, mas se ela diz que é assim, deve ser mesmo”. Seu trabalho somente foi reconhecido e devidamente interpretado quando se verificou a presença de transposons nos anos 70 e 80 em praticamente todos os organismos estudados. Ela ganhou o prêmio Nobel em 1984.

Estamos acostumados a sequências de DNA que ocupam seus *loci* gênicos fielmente nos seus respectivos cromossomos nucleares e nos genomas das mitocôndrias e dos cloroplastos. Entretanto, existem sequências que, por incrível que possa parecer, levam vidas particulares, independentes da replicação do cromossomo, ou seja, não estão ligadas restritamente ao ciclo celular para se reproduzir. Nas bactérias e em alguns eucariotos, encontramos pequenas moléculas circulares que se replicam independentemente, nossos conhecidos *plasmídeos*. Como vimos, os episômos se comportam como plasmídeos, porém se integram ao DNA hospedeiro, como um vírus, incluindo os próprios fagos que infectam as bactérias. Da mesma forma, ocorre a dinâmica dos elementos transponíveis ou transposons (já chamados de “genes saltadores”). Geralmente inserem cópias de si mesmos em outras sequências ou, mesmo, ele próprio é capaz de sair de um lugar no genoma e inserir-se em outro. Os processos de transposição ainda não são bem esclarecidos, porém o pouco que sabemos é um pouco estarrecedor, a princípio, porém, a estranheza costuma desvanecer-se à medida que conhecemos os processos de remodelamento do genoma. Assim, veremos que a transposição tem suas semelhanças com a recombinação por *crossing over*, por exemplo.

Alguns elementos transponíveis carregam genes com efeitos fenotípicos adaptativos, como exemplo clássico, genes para resistência a antibióticos. Entretanto, tem sido visto também que muitos elementos móveis podem transportar somente genes essenciais para sua transposição, especialmente *transposases*, as quais são enzimas para a transposição. Quase 50% do ge-

O transposon **Mutator** (*Mu*) foi originalmente identificado em milho e depois utilizado para a identificação de genes. Possui dois genes que são transcritos em sentido inverso, a partir de promotores que estão contidos nas repetições terminais invertidas chamadas de *mudrA* e *mudrB*. A proteína codificada de *mudrA*: MURA tem homologia com as transposases bacterianas, sendo encontrada em diversas espécies vegetais, enquanto que o outro gene *mudrB* codifica para a proteína MURB, cuja função ainda não está esclarecida. O sistema Mutator compreende também sequências denominadas *MuLES* (*Mutator-like elements*), encontrados também em arroz, algodão, soja e *Arabidopsis*.

noma humano é derivado de sequências produzidas por transposição, (veja a Figura 5). Note que os estudos mostram que a maioria dessas sequências estão inativas, incapazes de mais transposição, possivelmente há cerca de 50 milhões de anos. Os estudos sobre algumas patologias têm mostrado atividade dessas sequências. Veja a presença da sequência *Alu* na figura 6.

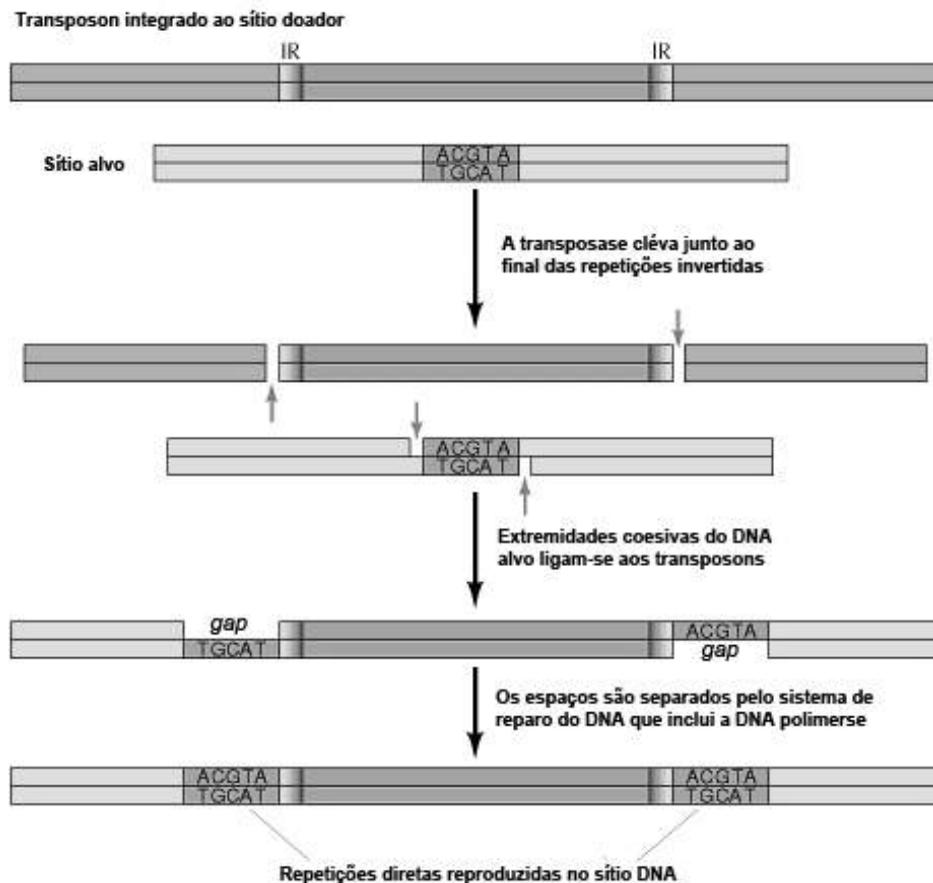
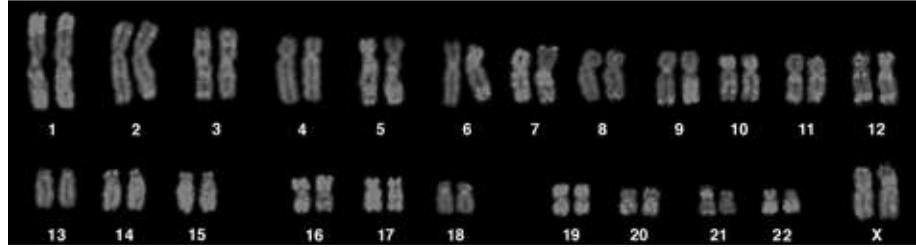


Figura 5 - Transposição simples. A transposição simples não envolve replicação do transposon. A enzima transposase cliva em ambos os lados do transposon e introduz o elemento através do corte em extremidades coesivas. Os *gaps* são consertados pela reparação do DNA feita pela DNA polimerase.

Fonte: Cooper e Hausman (2009).



```

GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGG
GCGGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGA AAC
CCGTCTCTACTAAAAATACAAAATTAGCCGGGGCGTGGTGGCGCGCGCCTGTA
ATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCGG
AGGTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCG
AGACTCCGTCTC

```

### Sequência Alu

Figura 6 - Exemplo de elemento de transposição – Sequência *Alu* - encontrada em mais de um milhão de cópias no genoma humano, constituindo cerca de 11% do nosso DNA. Na figura, vemos um cariótipo humano feminino, mostrando em tons mais claros, os locais de concentração. A seguir, temos a sequência. Note que, dentro destes 11%, encontramos muitas alterações devido a mutações. Portanto, elas não são sempre iguais. A sequência *Alu* é um exemplo de SINE (veja Unidade 1). Várias desordens herdáveis têm sido atribuídas à recombinação mediada por sequências *Alu*, incluindo a diabetes tipo II insulinoresistente, síndrome de Lesch-Nyhan, doença de Tay-Sachs, deficiência de componente C3, hipercolesterolemia e  $\alpha$ -talassemia, incluindo o Sarcoma de Ewing, o câncer de mama e a leucemia mielócita aguda.

Fonte: Batzer e Deininger (2002).

## 7. Mutações

O DNA é a fonte de informação genética, e a célula investe muito na manutenção dessa informação. Existem vários mecanismos altamente eficientes que proporcionam reparação ao DNA modificado ou danificado. Para que as células e os organismos possam ter chances de sobrevivência, as sequências de DNA devem ser passadas exatas e inalteradas para as próximas gerações. Altas taxas de mutações significam alterações metabólicas que podem acarretar distúrbios agudos ou crônicos e morte. Taxas menores podem perturbar o metabolismo de forma a ainda manter certo grau de sobrevivência do indivíduo, porém com possíveis repercussões na reprodução da célula ou do indivíduo. Com a morte da estrutura ou do indivíduo antes de sua reprodução, as mutações não são incorporadas na população, desaparecendo antes de deixar descendentes mutantes que a perpetuem.

Erros na replicação e lesões químicas espontâneas do DNA podem alterar bases e lesões causadas por mutagênicos e até romper as pontes fosfodiéster. Essas alterações podem ser transmitidas ou, em determinados casos, inviabilizar a célula ou o próprio organismo a ser derivado por essa célula. As mutações podem ocorrer em qualquer gene e em qualquer local do gene, no entanto, são mais frequentes nos seguintes casos: a) em regiões com sequências de DNA *repetitivas ou simétricas*, os chamados pontos quentes (*hotspots*), nos quais aumenta o risco de uma cadeia de DNA emparelhar consigo própria durante a replicação; b) em genes de maior tamanho, que, assim, têm uma maior probabilidade de sofrer alterações na sua sequência de bases; c) em genes do genoma mitocondrial que não têm mecanismos de reparação do DNA.

Percebe-se, então, um paradoxo relativo ao metabolismo do DNA e dos ácidos nucleicos e a sua fiel perpetuação. Se o DNA fosse mantido com fidelidade perfeita, como seria possível a evolução? Como seria possível ocorrer variabilidade genética sem que ocorressem alterações gênicas que se mantivessem e pudessem ser submetidas à pressão de seleção? Portanto, o desafio biológico é produzir e manter variabilidade, entretanto, compatível com a vida de forma geral e com a sobrevivência individual.

As mutações podem ser

a) **gênicas** – alteram a sequência de nucleotídeos do DNA, por substituição, adição ou remoção de bases. Podem conduzir à modificação da molécula de RNAm, que é transcrita a partir do DNA e, conseqüentemente, à alteração da proteína produzida, o que tem, geralmente, efeitos no fenótipo.

b) **cromossômicas** – traduzem-se numa alteração da estrutura (mutação cromossômica estrutural) ou do número (mutação cromossômica numérica) de cromossomos. Podem afetar uma determinada região de um cromossomo, um cromossomo inteiro ou todo o complemento cromossômico de um indivíduo.

Também podem ser **somáticas** ou **germinativas**, dependendo do tipo de tecido que afetam. Ocorre durante a replicação do DNA que precede uma divisão mitótica. No primeiro caso, todas as células descendentes são afetadas, mas podem localizar-se apenas numa pequena parte do corpo. As mutações somáticas estão na origem de alguns tipos de câncer. Não são transmitidas à descendência. As mutações nas células germinativas – ocorre durante a replicação do DNA que precede a meiose. A mutação afeta os gametas e todas as células que deles descendem após a fecundação – é transmitida à descendência.

Vemos a seguir os tipos mais comuns de mutações moleculares:

**Mutação silenciosa:** substituição de uma base do DNA por outra (no 3º nucleotídeo de cada códon), porém não ocorre alteração de aminoácido, devido à redundância do código genético. São muito comuns e responsáveis pela diversidade genética, que não é expressa fenotipicamente.

**Mutação com perda de sentido:** substituição de uma base do DNA por outra; como consequência ocorre a substituição de um aminoácido por outro na proteína codificada. A conformação da proteína pode ser alterada (ex: anemia falciforme).

**Mutação sem sentido (*nonsense*):** substituição de uma base do DNA de tal modo que, no RNAm, um códon que especifica um aminoácido é alterado para um códon de parada ou o contrário. Origina uma proteína inativa mais curta ou mais longa do que a proteína normal.

**Deleção:** nesse caso podem ser removidas uma única base do DNA ou milhares delas. A remoção de um número de bases que não seja múltiplo de três altera completamente a mensagem do gene.

**Inserção:** o número de bases adicionadas ao DNA pode variar. A adição de um número que não seja múltiplo de três altera completamente a mensagem do gene. Quando é inserida uma sequência igual a outra, ocorre uma duplicação.

As mutações podem ser espontâneas ou induzidas. Mutações espontâneas podem ocorrer porque as quatro bases nucleotídicas podem existir sob duas formas diferentes, uma comum e outra muito rara (chamada *tautomérica*). As principais são a *desaminação* e a *depurinação* (Figura 7) e a formação de *dímeros de pirimidina* (Figura 8). Quando uma base adquire, temporária ou permanentemente, a sua forma rara, pode emparelhar-se com uma base diferente. Outra causa de mutações espontâneas são erros na replicação do DNA motivados pela DNA polimerase. Quase sempre esses erros são reparados

durante o processo de replicação do DNA. Contudo, alguns persistem. Finalmente, há erros na meiose ou mitose (*não disjunção* de homólogos ou cromátides, tendo como consequência a formação de células com excesso ou falta de cromossomos).

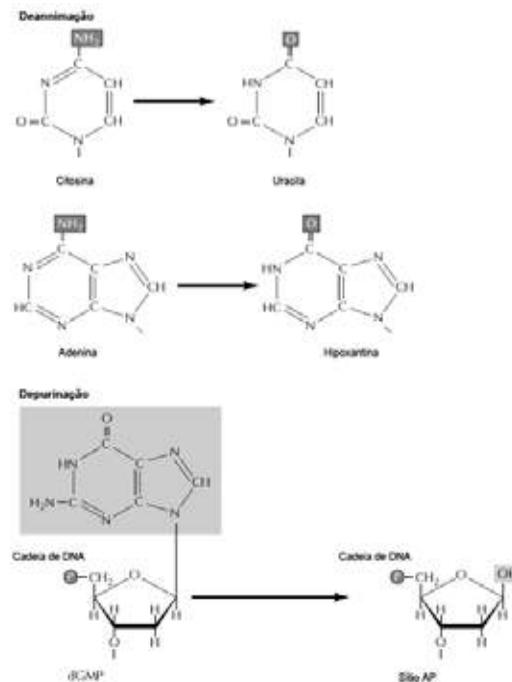


Figura 7 - As duas formas predominantes de danos espontâneos ao DNA: (A) desaminação da adenina, citosina e guanina, (B) depurinação (perda da base purina) resultante da clivagem da ponte entre a base purina e a ribose, produzindo um sítio apurínico (AP) dGMP = deoxiguanosina monofosfato.

Fonte: Cooper e Hausman (2009).

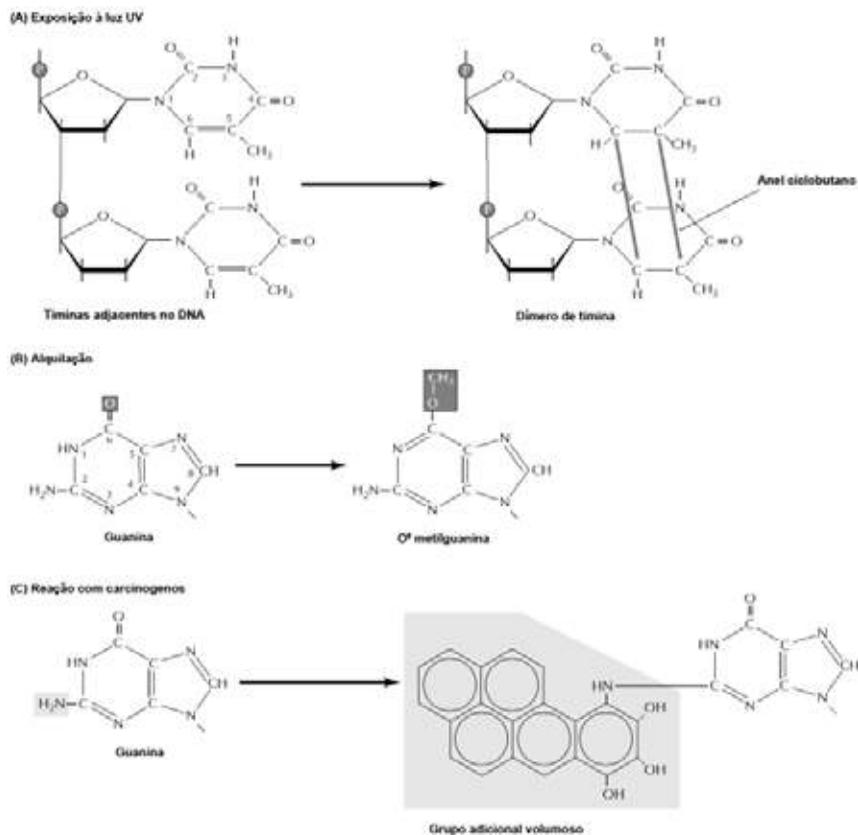


Figura 8 - (A) A luz UV induz a formação de dímeros de pirimidina, nos quais duas pirimidinas adjacentes (como exemplo, duas timinas) são unidas numa estrutura de anel em ciclobutano. (B) A alquilação é a adição de grupos metil ou etil para várias posições nas bases do DNA. Nesse exemplo, a alquilação da posição O6 da guanina resulta na formação de uma O6 - metilguanina. (C) muitos carcinogênicos (como o benzo-(a)pireno reagem com as bases de DNA, resultando na adição de grupos químicos à molécula de DNA.

Fonte: Cooper e Hausman (2009).

As mutações **induzidas**, por sua vez, são causadas por agentes mutagênicos e radiações. Conforme vimos, muitas descobertas foram feitas como consequência de testes militares, como o gás mostarda e as armas atômicas. Como resultado das pesquisas físicas e químicas, as quais tiveram grande avanço nas últimas décadas, apareceram uma infinidade de produtos capazes de interagir com o DNA. Por exemplo, os *análogos de bases*, como o 5-bromouracil, análogo à timina, e também a 2-aminopurina, análogo à adenina. Os chamados *agentes alquilantes* adicionam grupos metil e etil às bases nitrogenadas. O gás mostarda também é um agente alquilante. **Reações oxidativas** são causadas pelas formas reativas do oxigênio, como os radicais (livres) superóxido, peróxido de hidrogênio e radicais hidroxilas. Esses danificam o DNA e induzem mutações. Temos

também os *agentes intercalantes*, como acridina, brometo de etídio e dioxina. Estes se intercalam entre as bases do DNA, causando inserções e deleções.

### 7.1 Mecanismos de reparação

Existem vários mecanismos de reparo para o DNA. Conforme já vimos, a DNA polimerase exerce esse papel durante a sua polimerização, mantendo fiel a informação genética durante a replicação do DNA. Após a replicação, novos grupos enzimáticos entram em ação e procuram erros, diminuindo ainda sensivelmente a taxa de erros que “escapam” após a divisão celular. A maioria dos mecanismos de reparo envolve a remoção de nucleotídeos, numa via comum com quatro etapas básicas: a) *detecção do erro*, reconhecendo o DNA danificado, b) *excisão*, em que as endonucleases de reparo do DNA cortam o arcabouço de fosfodiéster em um ou ambos os lados do dano ao DNA, c) *nova polimerização*, em que a DNA polimerase adiciona nucleotídeos, substituindo os danificados e d) a *DNA ligase* fecha os cortes da ligação fosfodiéster.

No sistema de reparo de malpareamento, ocorre um mecanismo adicional da célula de reparar erros que escaparam da revisão da DNA polimerase III. Esses malpareamentos são temporários, pois são eliminados no próximo ciclo de replicação, quando aparecem, resultam nas mutações. Esse nucleotídeo adicionado de forma errônea deve ser retirado do lado recém-sintetizado, e não o seu molde do lado parental. Esse reparo tem que ocorrer de forma rápida, antes que a mutação fique incorporada em uma das fitas.

## 8. Recombinação gênica

Se a mutação é a fonte primária de variabilidade, a recombinação gênica é que efetivamente realiza a "mistura" entre os diferentes genes dos seres vivos. A evolução seria extremamente lenta se não houvesse um modo de reunir características de diferentes indivíduos formando um recombinante. A mutação em si é um fenômeno bem raro e, se a evolução dependesse apenas da mutação, certamente não teríamos hoje em dia organismos tão bem adaptados ao seu ambiente. A mutação e a recombinação agem em conjunto: a mutação modifica o DNA, e a recombinação realiza uma mistura entre as partes modificadas de dois organismos.

A recombinação gênica acontece durante a meiose. A partir do momento que os gametas se unem durante a fertilização, cada um deve conter apenas metade do número de cromossomos que outras células do corpo possuem. Caso contrário, a célula fertilizada teria cromossomos a mais. Dentro das células germinativas os cromossomos homólogos ficam pareados. Enquanto eles são comprimidos, os cromossomos podem quebrar, e cada um pode trocar uma porção do seu material genético pela porção correspondente de seu par. Essa forma de recombinação é o

nosso conhecido **crossing-over** (Figura 9). Quando os cromossomos se “grudam” de volta e se separam, cada um obtendo um novo material genético do outro. As versões de genes que o cromossomo apresenta agora são diferentes da original. Como esse processo só ocorre com uma das duas cópias do cromossomo, o conjunto das informações iniciais não é totalmente perdido. O resultado é o aumento da variabilidade. Uma das proteínas envolvidas é a proteína RecA (Figura 10).

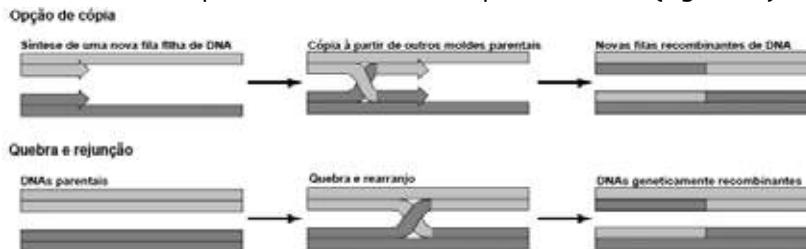


Figura 9 - Modelos de recombinação - Uma molécula de fita simples, parenta, liga-se à outra, por replicação ou simples troca, através de um **quiasma** (ou **junção de Holliday**), resultando em recombinação. Cometário: o DNA recombinante, neste caso refere-se ao produto do **crossing over**. Não confundir com o DNA recombinante obtido por Engenharia genética.

Fonte: Cooper e Hausman (2009).

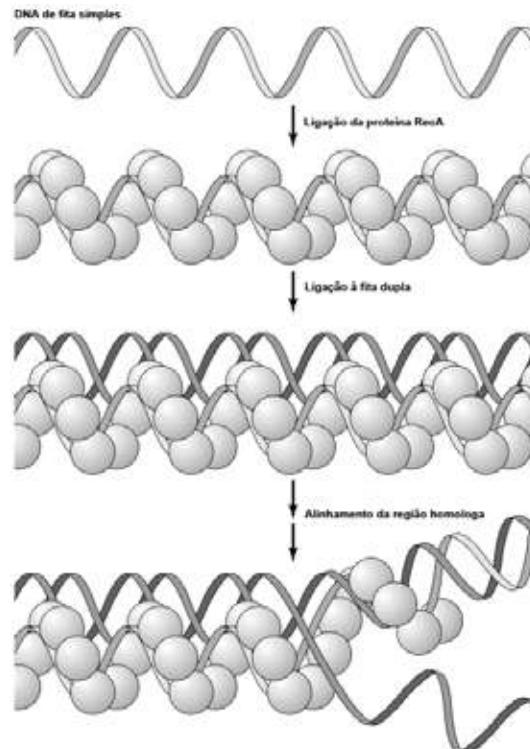
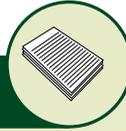


Figura 10 - Função da proteína RecA: a proteína RecA inicialmente liga-se à fita simples de DNA, para formar um filamento de proteína - DNA. A proteína RecA, então, liga-se à segunda para formar o complexo, formando uma região heteroduplex.

Fonte: Cooper e Hausman (2009).



## Texto complementar

### Para remover a palavra raça dos prontuários médicos no Brasil

Os avanços da genética molecular e o sequenciamento do genoma humano permitiram um exame detalhado da correlação entre a variação genômica humana, a ancestralidade biogeográfica e a aparência física das pessoas e mostraram que os rótulos previamente usados para distinguir raças não têm significado biológico. Pode parecer fácil distinguir fenotipicamente um europeu de um africano ou de um asiático, mas tal facilidade desaparece completamente quando procuramos evidências destas diferenças raciais no genoma das pessoas. Uma plethora de linhas independentes de pesquisa molecular fornece evidências científicas incontrovertíveis de que raças humanas não existem do ponto de vista genético ou biológico. Apesar disto, o conceito de raças persiste como construção social e cultural em nossa sociedade.

**RAÇAS HUMANAS NA MEDICINA BRASILEIRA** - o problema na medicina é que as raças continuam a ser vistas como verdades biológicas e não como meras construções sociais! Como exemplo, basta ler a bula no medicamento Cozaar, um bloqueador do receptor da angiotensina, vendido no Brasil pela Merck Sharp & Dohme, que contém o seguinte alerta: "Com base no estudo LIFE (Losartan Intervention For Endpoint Reduction in Hypertension — Intervenção com losartan para redução de desfechos na hipertensão), os benefícios de COZAAR® (Losartan potássico, MSD) na morbidade e mortalidade cardiovascular comparados aos do atenolol não se aplicam a pacientes negros com hipertensão e hipertrofia ventricular esquerda". Este não é um exemplo isolado, pois sabemos que Estados Unidos, bulas de 8% (15 entre 185) dos novos medicamentos introduzidos de 1995 a 1998 continham advertências sobre diferenças "raciais" em sua eficácia ou efeitos colaterais. Em 2005 a Food and Drug Administration (FDA) aprovou o medicamento BiDil® para tratamento da insuficiência cardíaca congestiva, mas somente para negros.

Com base nos critérios de autoclassificação do censo do IBGE, em 2000 a população brasileira era composta por 53,4% de brancos, 6,1% de pretos e 38,9% de pardos. O que representam esses números em termos de ancestralidade genética? Essa é a pergunta que tentamos responder usando as ferramentas da genética molecular. Os nossos estudos demonstraram que no povo brasileiro existe um alto índice de mistura genética que torna as características de aparência física, como cor da pele, olhos, cabelos, formatos dos lábios e do nariz, em indicadores muito pobres da origem geográfica dos ancestrais de um determinado indivíduo.

A fauna e flora dependem muito da geografia. Assim, diferentes origens geográficas significam exposições diferentes a elementos tóxicos do meio ambiente e a diferentes dietas. Como consequência, emergiram polimorfismos genéticos nos genes dos sistemas enzimáticos de detoxificação, representando adaptações seletivas ao padrão geográfico de elementos tóxicos da flora e dieta locais. Essas variações estão hoje em dia sendo detectadas pelos seus efeitos farmacogenéticos, já que o nosso corpo usa os mesmos sistemas enzimáticos para metabolizar medicamentos. Podemos usar em medicina clínica a aparência física, e mais particularmente a cor da pele, como um substituto dos testes genômicos específicos? A resposta é não, especialmente no Brasil, onde nossos dados mostram que a correlação entre cor e ancestralidade é muito imperfeita. Espera-se que possamos em breve ter à disposição testes genômicos que nos permitam traçar o perfil farmacogenético do paciente e usar este

perfil na decisão clínica. Até lá, devemos nos refrear de usar raça como substituta dos testes farmacogenéticos.

No consultório médico, o que está sendo examinado é um paciente individual e não um grupo populacional. A autoclassificação ou a avaliação médica do grupo racial de um paciente não tem nenhum valor em decisões sobre o diagnóstico, o tratamento farmacológico ou outras terapias. Isto não quer dizer que a pigmentação da pele do (a) paciente e outras características físicas não devem ser observadas e devidamente anotadas. Elas são partes integrais do exame físico e podem ter implicações clínicas. O que queremos defender é a retirada da expressão raça dos prontuários médicos no Brasil.

Alguns autores têm argumentado que as raças podem constituir-se em sub-rogadas de variáveis não-genéticas sociais e culturais, e que abrir mão dessa classificação significaria perda de correlações ambientais e prejuízo para a medicina e a população. Mas essa utilidade hipotética da classificação racial deve ser considerada no contexto dos seus possíveis riscos, sob uma ótica de relação risco-benefício, como tudo em medicina. Temos de tomar cuidado de não dar legitimidade a falsos indicadores e fazer todo o esforço para abandonar essas tênues correlações, atacando com firmeza as verdadeiras variáveis genéticas e ambientais que afetam saúde e doença.

PELO BANIMENTO DO CONCEITO - assim, o conceito de raça lembra uma casca de banana: vazio e perigoso. Vazio, porque sabemos que raças humanas não existem como entidades biológicas. Perigoso, porque no passado, a crença de que raças humanas possuíam diferenças biológicas substanciais e bem demarcadas contribuiu para justificar discriminação, exploração e atrocidades, mesmo no contexto médico.

Existe consenso entre geneticistas e antropólogos que a única divisão biologicamente coerente da espécie humana é em bilhões de indivíduos e não em um punhado de raças. Esse fato deve ser absorvido pela sociedade e incorporado às suas convicções e atitudes morais. Vimos acima que independente da cor, a vasta maioria dos brasileiros tem simultaneamente um grau significativo de ancestralidade africana, europeia e ameríndia. O genoma de cada brasileiro é um mosaico altamente variável e individual formado por contribuições das três raízes ancestrais. Assim, não faz sentido falar em afrodescendentes ou eurodescendentes porque a maior parte dos brasileiros tem uma proporção significativa de ascendência africana, europeia e ameríndia. Além disso, por causa da pobre correlação entre cor e ancestralidade, não faz sentido falar sobre "populações" de brasileiros brancos ou de brasileiros negros. Assim, a única maneira de lidar eticamente com a variabilidade genética dos brasileiros é individualmente, como seres humanos únicos e singulares nos seus genomas mosaicos e nas suas histórias de vida. Do ponto de vista médico, esta conscientização nos leva a propor que o conceito de raça deveria ser banido da medicina brasileira e que a palavra raça deve ser eliminada dos nossos prontuários clínicos.

Fonte: Pena (2007).

## Saiba mais



A Lei de Proteção de Cultivares, Lei nº 9456, foi sancionada em 25/04/1997, a qual consiste na proteção intelectual dos direitos de criação do pesquisador desenvolvedor de pesquisa genética. A Lei proíbe o uso, pelo produtor de sementes, de uma cultivar protegida, somente poderá ser feito mediante prévia autorização do criador da cultivar, que poderá ou não exigir o pagamento de "royalties" pela sua exploração

comercial. De acordo com o Art. 3º considera-se, para os efeitos desta Lei:

I – melhorista: a pessoa física que obtiver cultivar e estabelecer descritores que a diferenciem das demais;

II – descritor: a característica morfológica, fisiológica, bioquímica ou molecular que seja herdada geneticamente, utilizada na identificação de cultivar;

III – margem mínima: o conjunto mínimo de descritores, a critério do órgão competente, suficiente para diferenciar uma nova cultivar ou uma cultivar essencialmente derivada das demais cultivares conhecidas;

IV – cultivar: a variedade de qualquer gênero ou espécie vegetal superior que seja claramente distinguível de outras cultivares conhecidas por margem mínima de descritores, por sua denominação própria, que seja homogênea e estável quanto aos descritores através de gerações sucessivas e seja de espécie passível de uso pelo complexo agroflorestal, descrita em publicação especializada disponível e acessível ao público, bem como a linhagem componente de híbridos;

V – nova cultivar: a cultivar que não tenha sido oferecida à venda no Brasil há mais de doze meses em relação à data do pedido de proteção e que, observado o prazo de comercialização no Brasil, não tenha sido oferecida à venda em outros países, com o consentimento do obtentor, há mais de seis anos para espécies de árvores e videiras e há mais de quatro anos para as demais espécies;

VI – cultivar distinta: a cultivar que se distingue claramente de qualquer outra cuja existência na data do pedido de proteção seja reconhecida;

VII – cultivar homogênea: a cultivar que, utilizada em plantio, em escala comercial, apresente variabilidade mínima quanto aos descritores que a identifiquem, segundo critérios estabelecidos pelo órgão competente;

VIII – cultivar estável: a cultivar que, reproduzida em escala comercial, mantenha a sua homogeneidade através de gerações sucessivas;

IX – cultivar essencialmente derivada: a essencialmente derivada de outra cultivar se, cumulativamente, for:

a) predominantemente derivada da cultivar inicial ou de outra cultivar essencialmente derivada, sem perder a expressão das características essenciais que resultem do genótipo ou da combinação de genótipos da cultivar da qual derivou, exceto no que diz respeito às diferenças resultantes da derivação;

b) claramente distinta da cultivar da qual derivou, por margem mínima de descritores, de acordo com critérios estabelecidos pelo órgão competente;

c) não tenha sido oferecida à venda no Brasil há mais de doze meses em relação à data do pedido de proteção e que, observado o prazo de comercialização no Brasil, não tenha sido oferecida à venda em outros países, com o consentimento do obtentor, há mais de seis anos para espécies de árvores e videiras e há mais de quatro anos para as demais espécies;

X – linhagens: os materiais genéticos homogêneos, obtidos por algum processo autogâmico continuado;

XI – híbrido: o produto imediato do cruzamento entre linhagens geneticamente diferentes;

XII – teste de distinguibilidade, homogeneidade e estabilidade (DHE): o procedimento técnico de comprovação de que a nova cultivar ou a cultivar essencialmente derivada são distinguíveis de outra cujos descritores sejam conhecidos, homogêneas quanto às suas características em cada ciclo reprodutivo e estáveis quanto à repetição das mesmas características ao longo de gerações sucessivas;

XIII – amostra viva: a fornecida pelo requerente do direito de proteção que, se utilizada na propagação da cultivar, confirme os descritores apresentados;

XIV – semente: toda e qualquer estrutura vegetal utilizada na propagação de uma cultivar;

XV – propagação: a reprodução e a multiplicação de uma cultivar, ou a concomitância dessas ações;

XVI – material propagativo: toda e qualquer parte da planta ou estrutura vegetal utilizada na sua reprodução e multiplicação;

XVII – planta inteira: a planta com todas as suas partes passíveis de serem utilizadas na propagação de uma cultivar;

XVIII – complexo agroflorestal: o conjunto de atividades relativas ao cultivo de gêneros e espécies vegetais visando, entre outras, à alimentação humana ou animal, à produção de combustíveis, óleos, corantes, fibras e demais insumos para fins industrial, medicinal, florestal e ornamental.” Fonte: <http://www.planalto.gov.br/ccivil/leis/L9456.htm>

## Síntese da Parte



Nesta unidade verifica-se que existem várias formas de alterar o genoma, desde os variados tipos de mutações, recombinação, transposons e ploidias. Muitas dessas alterações são mais conhecidas e verifica-se que possuem funções bem definidas, como na recombinação, em que existe aumento na variabilidade. Porém muitas, aparentemente, não trazem sentido ao genoma, como os transposons e possivelmente podem ser deletérias, em alguns casos patológicos. Os genomas das bactérias, mitocôndrias e cloroplastos, mais simples, são mais fáceis de serem estudados e compreendidos e fornecem algumas pistas para compreender a evolução molecular.

## Atividades de avaliação



1. Leia atentamente o resumo de congresso que segue e responda as perguntas:

**Título:** Detecção de cinco novas mutações em pacientes brasileiros com gangliosidose GM1

URI: <http://hdl.handle.net/10183/6718>

Autor: Vieira, Matheus Barbosa

Orientador: Coelho, Janice Carneiro

Co-orientador: Matte, Ursula da Silveira

Data: 2006

Nível: Mestrado

Instituição: Universidade Federal do Rio Grande do Sul. Instituto de Ciências Básicas da Saúde. Programa de Pós-Graduação em Ciências Biológicas: Bioquímica.

Palavra-chave: Erros inatos do metabolismo, Gangliosidose gm1, Mutações

**Resumo:** A Gangliosidose GM1 é um Erro Inato do Metabolismo (EIM) causado pela deficiência da enzima B-galactosidase ácida. Essa doença é caracterizada pelo acúmulo de metabólitos não degradados, principalmente gangliosídeo GM1, nos lisossomos de vários tipos celulares. Baseado na idade de início e na atividade residual da enzima, a Gangliosidose GM1 é classificada em três diferentes tipos: infantil, juvenil e adulto. O gene da B-galactosidase ácida (GLB1, GeneBank M27507) está situado no cromossomo 3 e possui mais de 60 kb, contendo 16 exons. Cerca de 50 mutações associadas à doença estão descritas na literatura. No sul do Brasil, há uma alta frequência dessa doença (1:17.000 nascidos vivos). Neste trabalho, vinte pacientes diagnosticados no Hospital de Clínicas de Porto Alegre (Brasil) tiveram o gene GLB1 investigado por SSCP (Single Strand Conformational Polymorphism) usando DNA extraído de sangue periférico. Através desta triagem foram encontradas 52 alterações de mobilidade do DNA, indicando a presença de mutações. As amostras relativas aos exons 2 e 15 foram submetidas a sequenciamento direto com seqüenciador ABI310(Applied Biosystems) utilizando kit BigDye 3.1. Cinco novas mutações no gene GLB1 (**F63Y, R38G, Y36S, Y64F e R59C**) e duas mutações já descritas (**R59H** e 1622-1627insG) foram encontradas. Este trabalho possibilitou a genotipagem completa de 6 pacientes e parcial de 5, e direcionou a investigação de mutações, contribuindo diretamente no diagnóstico da enfermidade e permitindo a realização de estudos de correlação genótipo/fenótipo destes pacientes. Fonte: <http://www.lume.ufrgs.br/handle/10183/6718>

- a) o que são erros inatos do metabolismo?
- b) procure informações nos livros e na rede mundial de computadores sobre as gangliosidoses.
- c) o que é frequência gênica?
- d) procure as enzimas e as funções dos lisossomos.
- e) o resumo relata que os pesquisadores encontraram mutações nos pacientes. A seguir, temos a sequência traduzida em proteína, do gene da enzima beta-D-galactosidase (EC 3.2.1.23) humana citada. Localize as posições em que se verificaram alterações, veja o exemplo:

Mutação **F63Y**: conte até a posição 63 e encontrará o aminoácido F (Fenilalanina) – nos pacientes estudados, este foi alterado para Y (Tirosina).

```
MPGFLVRILL LLLVLLLLLP TRGLRNATQR MFEIDYSRDS FLKDGQPFYR
ISGSIHYSRV PRFYWKDRLL KMKMAGLNAI QTYVPWNFHEP WPGQYQFSED
HDVEYFLRLA HELGLLVILR PGPYICAEWE MGGLPAWLLK KESILLRSD
PDYLAADVWK LGVLLPKMKP LLYQNGGPVI TVQVENEYGS YFACDFDYLR
FLQKRFRHHL GDDVVFLETTD GAHKTFLKCG ALQGLYTTVD FGTGSNITDA
FLSQRKCEPK GPLINSEFYT GWLDHWGQPH STIKTEAVAS SLYDILARGA
SVNLYMFIGG TNFAYWNGAN SPYAAQPTSY DYDAPLSEAG DLTEKYFALR
NIIQKFEKVP EGPIPPSTPK FAYGKVTLEK LKTVGAALDI LCPSGPIKSL
YPLTFIQVKQ HYGFVLYRTT LPQDCSNPAP LSSPLNGVHD RAYVAVDGIP
QGVLERNNVI TLNITGKAGA TLDLLVENMG RVNYGAYIND FKGLVSNLTL
SSNILTDWTI FPLDTEDAVR SHLGGWGHRD SGHHDEAWAH NSSNYTLPAF
YMGNFSIPSG IPDLPQDTFI QFPGWTKGQV WINGFNLGRY WPARGPQLTL
FVPQHILMTS APNTITVLEL EWAPCSSDDP ELCAVTFVDR PVIGSSVTYD
HPSKPVEKRL MPPPPQKNKD SWLDHV
```

No caso da alteração **1622-1627insG**:

```
1561 atatactcac ggactggacg atctttccac tggacactga ggatgcagtg
cgcagccacc
```

```
1621 tgggggggctg gggacaccgt gacagtggcc accatgatga agcctgggcc
cacaactcat
```

```
1681 ccaactacac gctcccggcc
```

aparece uma inserção de guaninas no gene na posição de 1622 a 1627.

para ver o gene completo e esta tradução:

site: [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov), digite o no do registro: M27507 (veja no resumo) na opção "SEARCH": *Nucleotide* e tecla *Enter*. (ou pelo link direto: <http://www.ncbi.nlm.nih.gov/nuccore/179400>)

f) Compare as técnicas para verificação de mutações na Figura 11.

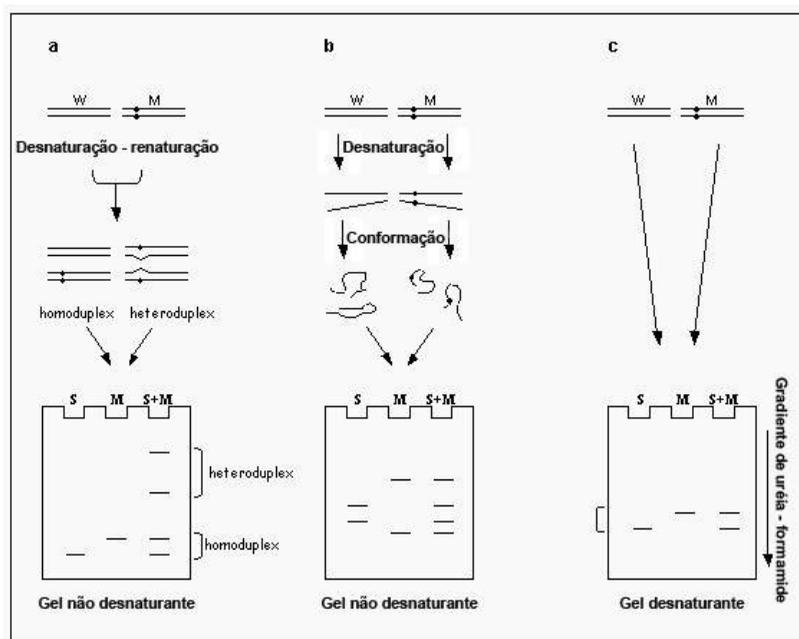


Figura 11 - Esquemas de três diferentes técnicas eletroforéticas para verificação de mutações. Estas técnicas permitem a visualização de alterações mínimas em duas fitas de DNA, a selvagem W (ou *wild type*) e a mutante M. (a) Ensaio de corrida de fita dupla ou HTA (*Heteroduplex Tracking Assay*): as duas fitas são desnaturadas/renaturadas e posteriormente misturadas. Note que as hélices homoduplex W e M separam-se muito pouco. Através da formação de heteroduplex, as diferenças ficam mais conspícuas, através do aumento da diferença de migração entre elas; (b) Polimorfismo de diferentes estruturas de fita simples ou SSCP (*Single Stranded Conformation Polymorphism*): as mutações criam pequenas alterações tridimensionais na fita simples, causando diferença na migração e (c) eletroforese de gel em gradiente desnaturante ou DGGE (*Denaturing Gradient Gel Electrophoresis*), W: wild type e M: mutante. Gel desnaturante: o gel de agarose contém um gradiente de concentração de um desnaturante químico. Durante a migração, a molécula encontra concentrações cada vez mais altas, alterando dramaticamente o tempo de migração, aumentando o polimorfismo entre as diferentes fitas.

## Leituras, filmes e sites



### Leituras

- PENA, S. D. J. **Humanidade sem raças?** 1. ed. São Paulo: Publifolha, 2008. 72 p.
- BARBUJANI, G. **A invenção das raças:** existem mesmo raças humanas? diversidade e preconceito racial. São Paulo: Contexto, 2007. 176 p.

## Filmes

- Homo sapiens 1900 (Dir. Peter Cohen, Suécia, 1998) - documentário sobre racismo e eugenia.
- O ponto de mutação (Dir. Bernt Capra, 1990) - existencialismo e ciência, baseado na obra de Fritjof Capra.

## Sites

### (Sociedades científicas brasileiras – *filie-se!*)

- <http://www.abc.org.br> (Academia Brasileira de Ciências - ABC)
- <http://www.abcmc.org.br> (Associação Brasileira de Centros e Museus de Ciência - ABCMC)
- <http://www.aoceano.org.br> (Associação Brasileira de Oceanografia - AOCEANO)
- <http://www.fesbe.org.br> (Federação de Sociedades de Biologia Experimental)
- <http://www.sbac.org.br> (Sociedade Brasileira de Análises Clínicas - SBAC)
- <http://www.sbbq.org.br> (Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq)
- <http://www.sbfic.org.br> (Sociedade Brasileira de Ficologia)
- <http://www.sbfis.org.br> (Sociedade Brasileira de Fisiologia – SBFIS)
- <http://www.sbg.org.br> (Sociedade Brasileira de Genética - SBG)
- <http://www.sbi.org.br> (Sociedade Brasileira de Imunologia - SBI)
- <http://www.sblimno.org.br> (Sociedade Brasileira de Limnologia - SBL)
- <http://www.sbma.com.br> (Sociedade Brasileira de Malacologia - SBMA)
- <http://www.sbmicrobiologia.org.br> (Sociedade Brasileira de Microbiologia - SBM)
- <http://www.sbpcnet.org.br> (Sociedade Brasileira para o Progresso da Ciência - SBPC)
- <http://www.sbpflap.cjb.net> (Sociedade Brasileira de Parasitologia - SBP)
- <http://www.sbv.org.br> (Sociedade Brasileira de Zoologia - SBZ)
- <http://acd.ufrj.br/geologia/sbp/sbp.htm> (Sociedade Brasileira de Paleontologia - SBP)

- <http://sbfte.org.br> (Sociedade Brasileira de Farmacologia e Terapêutica Experimental -SBFTE)
- <http://www.sbmm.org.br/> (Sociedade Brasileira de Microscopia e Microanálise - SBMM)

## Referências



ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular biology of the cell**. 5. ed. New York: Garland Science, 2008. 1.261 p.

BATZER, M. A.; DEININGER, P. L. Alu repeats and human genomic diversity. **Nature Reviews**, New York, v. 3, p. 370-381, 2002.

BENNET, M. D.; LEITCH, I. J. Nuclear DNA amounts in angiosperms and their modern uses - 807 new estimates. **Annals of Botany**, Bristol, v. 86, p. 859-909, 2000.

COOPER, G. M.; HAUSMAN, R. E. **The cell: a molecular approach**. 5. ed. Sunderland: Sinauer Associates, 2009. 771 p.

GAO, L.; YI, X.; YANG, Y-X.; SU, Y-J.; WANG, T. Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. **BMC Evolutionary Biology**, Kansas City, v. 9, n. 130, p. 1-14, 2009.

GREGORY, T. R. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. **Biological Reviews**, Cambridge, v. 76, n. 1, p. 65-101, 2001.

GUO, X.; CASTILLO-RAMÍREZ, S.; GONZÁLEZ, V.; BUSTOS, P.; FERNÁNDEZ-VÁZQUEZ, J. L.; SANTAMARÍA, R. I.; ARELLANO, J.; CEVALLOS, M. A.; DÁVILA, G. Rapid evolutionary change of common bean (*Phaseolus vulgaris* L) plastome, and the genomic diversification of legume chloroplasts. **BMC Genomics**, London, v. 8, p. 228, 2007.

PENA, S. D. J. Para remover a palavra raça dos prontuários médicos no Brasil. **Ciência e Cultura**, São Paulo, v. 59, n. 1, p. 4-5, 2007.

**Parte**

**4**

**Expressão gênica**



# Considerações iniciais e histórico

## Objetivo

- Apresentar aspectos da expressão gênica em procariotos e eucariotos.

## 1. Aspectos gerais

Como as células expressam seus genes? Toda sequência de DNA pode ser considerado um gene? Procariotos e eucariotos expressam seus genes da mesma forma? Vimos que nem toda molécula de DNA pode ser considerada um gene, pois a definição funcional diz que o gene deve ser codificador de algum RNA ou polipeptídeo. Pelo menos na maioria dos eucariotos, temos muito mais DNA não codificante do que codificante. Obviamente, existem muitas diferenças entre a expressão eucariótica e a procariótica. Procariotos são unicelulares, apresentam uma expressão mais econômica e menos complexa, por isso, podem ser considerados modelos biológicos para estudo da expressão. Os eucariotos são pluricelulares, possuem células diferenciadas e especializadas. Os seres eucariotos partem de uma única célula que determina (num momento exato do desenvolvimento) o início de diferenciação em folhetos germinativos que irão se transformar em pele, ossos, cabelos, bastonetes e hepatócitos. O uso das células-tronco embrionárias é um exemplo prático de que as células eucarióticas são induzíveis à diferenciação nos seus primórdios. O estudo da diferenciação celular eucariótica ainda está nos seus primórdios, sendo considerado um dos maiores desafios da era pós-genômica. Há mais a ser descoberto ainda do que informação acumulada sobre o assunto da expressão gênica.

Nos grandes genomas eucarióticos, somente uma pequena porção de DNA codifica proteínas. Na levedura, temos 70% de expressão de genes pelo genoma, havendo um gene a cada 2 Kb de sequência. *Saccharomyces cerevisiae* foi o primeiro organismo eucariótico a ter o seu genoma sequenciado. A levedura possui 5.885 genes codificantes de proteínas. Em humanos, menos de 5% (alguns autores colocam 1% do genoma) é codificante. Nos humanos ocorre um gene para cada 130 Kb de sequência.

A galactosemia é um distúrbio metabólico extraordinário no qual lactentes são incapazes de utilizar a galactose encontrada no leite. A galactose é importante além de sua capacidade de fornecer glicose. Também é importante componente dos cerebrosídeos, gangliosídeos e glicoproteínas. Estes são os constituintes das membranas celulares. Os cerebrosídeos e gangliosídeos formam uma porção significativa dos lipídeos cerebrais. Nos pacientes portadores de galactosemia, a galactose não é transformada e acumula-se em vários tecidos, ocasionando retardo mental severo e catarata. O fígado mostra dilatação e, frequentemente, a doença causa a morte.

## 2. Breve histórico

O biólogo francês Jacques Lucien Monod nasceu em 1910, estudou Medicina em Paris até 1940. Com a guerra, fugiu para a Inglaterra. Ele lutou na África, onde foi ferido. Seu desejo era ser médico cirurgião. Doutorou-se em 1947, mas não podia exercer a profissão de cirurgião devido às extensas feridas de guerra. Buscou por várias vezes, sem sucesso, emprego com André Lwoff, com quem tinha lutado na Segunda Guerra, em seu laboratório, no Instituto Pasteur e sempre foi recusado. Um dia Lowoff perguntou a ele se queria trabalhar com bacteriófagos. Apesar de não ter a menor idéia do que Lowoff sugeriu, ele aceitou e, assim, dedicou-se à biologia dos fagos, no Instituto Pasteur. Em 1958, François Jacob, começou a trabalhar com Jacques Monod em pesquisas sobre os mecanismos de transmissão da informação genética, como os conceitos de RNA mensageiro e genes reguladores da célula controlam a síntese de proteínas. Posteriormente, aprofundou os problemas da genética bacteriana, em foco no estudo dos mecanismos de sua reprodução (*Sexualidade e Genética de Bactérias*, 1961). Jacob, Monod e André Lwoff receberam o Prêmio Nobel de Medicina de 1965, por essa pesquisa. É importante lembrar que essas descobertas não tinham as facilidades que temos hoje em termos de tecnologia; elas foram deduzidos inteiramente com estudos genéticos feitos em bactérias, os quais posteriormente foram confirmados com isolamento de genes e de proteínas utilizando equipamentos e técnicas modernas.

Esse trabalho tinha como premissa o conhecimento de que as células bacterianas e outras poderiam responder a condições externas, regulando os níveis de produção de suas principais enzimas metabólicas e/ou a atividade dessas enzimas. Por exemplo, se uma bactéria encontra-se em um caldo contendo lactose, em vez da glicose, que é um açúcar mais simples, a bactéria tem que adaptar-se à necessidade de: 1) transporte da lactose pela membrana; 2) decomposição da lactose em seus componentes glicose e galactose e 3) conversão da galactose em glicose.

Conhecia-se a produção das enzimas responsáveis por essas etapas, mas, quando havia glicose no meio, essas enzimas eram ausentes. Os estudos foram então evoluindo pela teoria da **ação alostérica** das enzimas, em que postula que pequenas moléculas podiam ligar e desligar o processo. Jacob e Monod foram os descobridores teóricos e experimentais que demonstraram como o sistema da lactose em *E. coli* funcionava, através de proteínas específicas dedicadas a reprimir a transcrição do DNA em RNA, o repressor lac.

Esse repressor (o repressor *lac*) é produzido em todas as células, liga-se diretamente ao DNA de genes que controla e fisicamente impede o **aparelho de transcrição** de ter acesso ao DNA. Na presença de lactose, o repressor liga lactose, tornando-se não poder ligar ao DNA, bem como, a re-

pressão da transcrição é levantado. Dessa forma, um ciclo de realimentação é construído, permitindo que o conjunto de proteínas que metabolizam lactose seja produzido quando necessário.

Jacob Monod publicou "*La logique du vivant*" - Histórias da Biologia, em 1970, uma interessante análise histórica do conceito de herança. Outro trabalho de Monod: *O acaso e a necessidade* publicado também em 1970, na França, tornou-se um dos mais celebrados trabalhos da moderna Biologia. Monod diz em seu livro: "*A tese que apresentarei aqui é a de que a biosfera não contém uma classe previsível de objetos ou de fenômenos, mas constitui um acontecimento particular, de certo compatível com os primeiros princípios, mas não dedutível desses princípios. Portanto, essencialmente imprevisível.*" Para Monod, a evolução não tem projeto ou seja, você pode, avaliando todo o sistema biológico, estudar e ver de onde cada ser evoluiu. Porém, se estivesse lá no começo de tudo, jamais poderia prever o aparecimento do ser humano.

# Expressão gênica

**Ação alostérica:** geralmente um fator que se liga a uma proteína, fazendo com que ela mude de configuração tridimensional, permitindo que ocorra alguma ação.

Por exemplo, o aminoácido triptofano liga-se ao repressor, ativando-o alostericamente. Dessa forma, ele se liga ao DNA.

**Repressor do triptofano:** proteína regulatória produzida pelo gene regulador (*trpR*), a qual inativa o operador do triptofano quando este está em alta concentração.

**Operador:** parte de um ôperon que controla a expressão de um ou mais genes estruturais, servindo como sítio de ligação para uma ou mais proteínas regulatórias.

**AMP cíclico:** adenosina 3,5-monofosfato, o AMP cíclico é um nucleotídeo sintetizado a partir de ATP por ação da adenilato ciclase, enzima localizada na face interna da membrana plasmática. O cAMP exerce seus efeitos intracelulares ativando um tipo particular de proteína quinase chamada “proteína quinase dependente de cAMP”. Ele liga-se às subunidades reguladoras provocando dissociação das subunidades catalíticas que, então, tornam-se ativas.

## 1. Expressão e regulação em procariotos

As diferenças entre procariotos e eucariotos algumas vezes parecem superficiais e, em outras vezes, parecem marcantes. Vimos que, estruturalmente, os procariotos apresentam falta de vários componentes celulares, como de citoesqueleto, da própria carioteca, falta de partículas, como as histonas, de cap 5' e cauda poli A 3'; de sequências nucleotídicas, como os íntrons, DNA repetitivo etc. Apesar de ser relativamente simples, as bactérias são seres de rápida resposta ao ambiente, pois disso depende sua sobrevivência. Lembre-se dos conceitos da metabolização de energia (glicogênese, vias glicolíticas etc) vistos em Bioquímica. A dependência de energia (diga-se açúcar) da bactéria (*E. coli*) é o exemplo perfeito para o estudo de suas características de regulação gênica.

As bactérias nadam em um mar de nutrientes. Esses nutrientes variam em qualidade e quantidade, aumentando e diminuindo, de acordo com as questões ambientais envolvidas. As *E. coli* intestinais, após uma “bela refeição”, podem estar presentes numa sopa de lactose, glicose, galactose ou xilose, dependendo do tipo de alimento ingerido. Lembre-se de que cada tipo diferente precisa de uma proteína específica sinalizadora e de uma via metabólica. Qual das duas opções seguintes seria a mais interessante? 1) A célula bacteriana fica de prontidão, com todo o arsenal de enzimas e proteínas simultaneamente disponível para degradar todo e qualquer tipo diferente de açúcar; ou 2) a célula bacteriana, inicialmente, identifica o açúcar preponderante e assim ativa as cascatas metabólicas necessárias à sua degradação (e em consequência, desliga outros).

Imagine o quanto a bactéria iria gastar se fosse manter continuamente, um arsenal disponível para produção contínua dessas proteínas, enzimas e manutenção de vias metabólicas complexas. A opção correta é a 2. Metaforicamente, a célula atua com interruptores ou mecanismos que atuam ligando e desligando genes necessários e desnecessários, para cada ocasião. A opção 1, entretanto, não é de todo errada, porque os mecanismos de sinalização, os quais são necessários para saber (rapidamente) o que está

acontecendo no ambiente e fornecer um gatilho para cada situação são permanentemente disponibilizados, ou seja, são genes ligados, o tempo todo.

Uma das proteínas regulatórias mais estudadas em *E. coli* é o repressor do triptofano. Em bactérias, proteínas como a proteína repressora do triptofano, desliga genes, reprimindo a transcrição. De que forma? Ligando-se a um local específico *upstream* ao gene chamado *operador*. Esse operador é uma pequena sequência de DNA com cerca de 15 nucleotídeos, fica localizado dentro da região promotora do gene (onde se liga à enzima RNA polimerase). Essa ligação da proteína ao operador bloqueia a transcrição pelo simples fato de impedir a RNA polimerase de se ligar ao promotor. Entretanto, esse repressor deve ser ativado pela presença de triptofano no meio. O meio, nesse caso, pode ser o intestino do mamífero, que acabou de se alimentar. O repressor é uma proteína alostérica; a ligação do triptofano a ela produz uma ligeira mudança conformacional que permite, então, a ligação ao DNA. Portanto, liga-se ao DNA o complexo repressor-ativador (no caso, o triptofano).

O gene que é desligado, no entanto, é um conjunto de genes, todos próximos entre si, responsáveis pela produção de triptofano, chamado *óperon do triptofano*. Obviamente, a célula bacteriana prefere adquiri-lo pronto do meio ambiente. O sistema óperon é exclusivo de bactérias, os eucariotos regulam individualmente cada gene (Figura 1).

**Óperon:** um grupo de genes que constitui uma unidade reguladora ou controladora.

**Repressão catabólica:** redução mediada por glicose nas taxas de transcrição de óperons que especificam enzimas envolvidas nas vias metabólicas.

**Região de controle de locus:** Região próxima aos genes que são controladoras destas regiões.

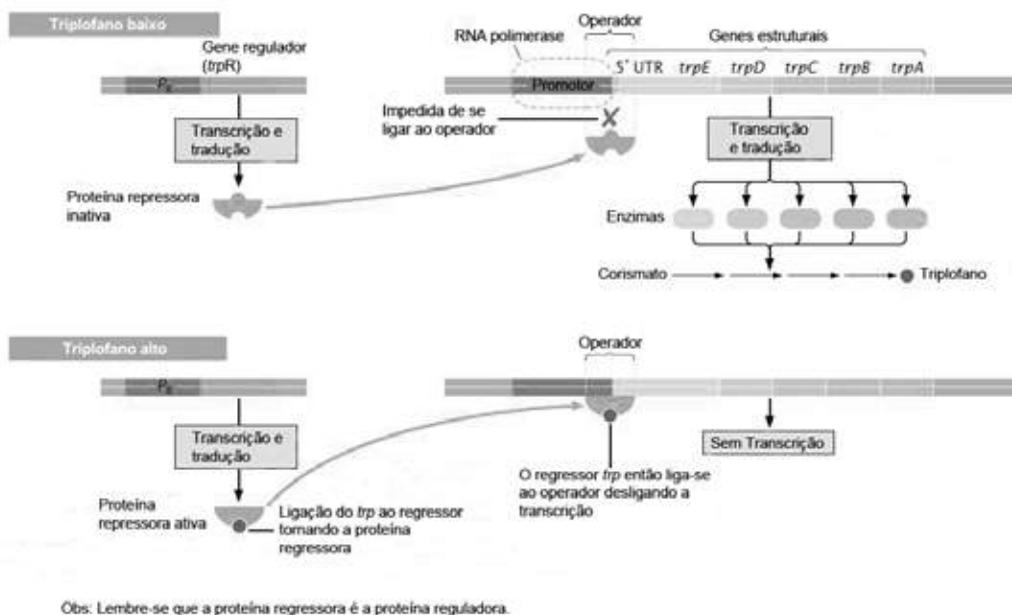


Figura 1 - Óperon triptofano em procariotos. Note que os genes estruturais (*trpE*, *trpD*, *trpC*, *trpB*, *trpA*) estão próximos ao gene e ao promotor comum, formando um RNAm comum.

Fonte: Pierce (2004).

Os **erros inatos do metabolismo** resultam da deficiência enzimática com dano direto numa via metabólica específica, bastante raras como a fenilcetonúria (deficiência da enzima fenilalanina hidroxilase), homocistinúria e hiperomocistinúria (deficiência da enzima cistationina B-sintase), a qual produz níveis elevados de homocistina e metionina na urina. A maioria dos tratamentos para essas patologias envolve restrição dietética do substrato (geralmente aminoácidos), reposição do produto final, depleção da substância de armazenamento, substituição da proteína mutante (como no caso da galactosidase), transplante (rim, no caso de cistinose e outras, como a doença de Crushing, fígado, no caso da tirosinemia) e remoção cirúrgica, como remoção da tireoide (síndrome do carcinoma medular da tireoide), baço e cólon.

Esse tipo de regulação coletiva de genes em bactérias só é possível pelos seguintes motivos: a) os genes de um determinado metabolismo bacteriano estão arranjados próximos, de forma grupal, sob regulação de um único promotor; b) uma vez ativado, produz um RNA coletivo chamado de policistrônico (veja transcrição na Unidade 2), que, depois de formado, vai dar origem às várias proteínas desse metabolismo e c) a proteína regulatória geralmente está acoplada à presença ou à ausência de outra substância que pode ser seu ativador.

O repressor (ou o ativador) pertence a uma classe de proteínas que pertence ao metabolismo basal da célula. A presença da proteína regulatória, nesse caso, é sempre constante. Por isso, a bactéria pode responder rapidamente a um aumento ou uma diminuição do triptofano. Esse tipo de expressão é chamada de *constitutiva*. Inúmeros genes do metabolismo fazem parte do conjunto constitutivo de genes de uma célula, pois estão sempre sendo expressos, independentemente dos fatores ambientais. Outros genes são induzíveis (ativados) ou reprimíveis (inativados), mas podem ter também uma expressão mínima constante.

Ativadores bacterianos, como a proteína ativadora do metabolismo (CAP), ligam-se ao AMP cíclico (cAMP), antes que ela possa se ligar ao óperon. Os genes ativados por CAP são, então, dependentes da concentração de cAMP na célula. O aumento de glicose na célula torna os níveis de AMP cíclico baixos, e o cAMP tem menos probabilidade de se ligar a CAP. Portanto, em *E. coli*, a concentração de cAMP é inversamente proporcional ao nível de glicose. Quando a glicose está baixa na célula bacteriana, os níveis de cAMP são altos. O complexo CAP-cAMP liga-se ao DNA junto ao operador lac. A glicose é sempre preferida porque sofre glicólise e requer menos energia metabólica que outros açúcares.

Na presença de glicose, todos os outros genes de metabolismo de açúcares são reprimidos, num processo chamado de *repressão catabólica*. Essa repressão coletiva é resultado do controle positivo da célula em resposta à glicose.

Uma questão interessante na regulação eucariótica: se todo gene precisa de reguladores específicos, a expressão dessas proteínas regulatórias também necessitaria de outras proteínas regulatórias e assim por diante. Finalmente, um número infinito de genes seria necessário para a expressão de poucos genes. Como isso é possível? A resposta é que, além da expressão constitutiva, existe controle pelos sinais de dentro e fora da célula, que ajustam o programa transcricional às necessidades fisiológicas da célula. Além disso, são usadas na forma de combinações e afetando mutuamente a sua atividade. Uma grande fração do genoma, cerca de 10%, codifica proteínas regulatórias.

## 2. Cromatina celular e epigenética

Ao analisarmos a estrutura da cromatina, imaginamos se os nucleossomos, como estrutura, podem afetar a expressão do gene, a replicação, a reparação etc. Estudos sugerem que o nucleossomo é modelado de forma que as enzimas possam atravessá-lo, desde que ocorram alguns desencaixes entre os dímeros de histona. À medida que ocorre progressivo empacotamento da cromatina, na divisão celular, o gene torna-se cada vez mais inacessível. O empacotamento máximo dos cromossomos ocorre durante a metáfase, a qual é uma fase transcricionalmente inativa. Para resumir, a ideia é a de que os nucleossomos não impediriam as enzimas metabolizadoras (RNA e DNA polimerases), mas o empacotamento, esse sim, impede a transcrição.

O exemplo clássico é a família da  $\beta$ -globina, expressa em *células eritroides*, estudada especialmente em humanos e em galinhas. Possui cerca de cinco genes expressos durante o desenvolvimento, os quais são ativados/desativados por um mecanismo de alteração da estrutura da cromatina. Essas mudanças surgem de uma região chamada de **região de controle de locus** (LCR). Pacientes com *talassemia* (um tipo de anemia) apresentam mutações nessa região específica. As regiões de genes das globinas, nesses pacientes, mostram-se resistentes ao tratamento com uma DNase, mostrando que eles se encontram inacessíveis e, portanto, inativados, pela cromatina.

A *metilação* é um recurso usado para regular os genes, geralmente de forma inibitória ou inativadora, pois a adição de metilas (-CH<sub>3</sub>) aos nucleotídeos impede o perfeito ajuste do substrato DNA às enzimas que o metabolizam. A metilação é efetuada por uma enzima metilase. Nos vertebrados, ela ocorre restritamente aos nucleotídeos de citosina, na sequência CG. Esse padrão de metilação é herdado pelas células-filhas, após a replicação do DNA.

A literatura sugere que a metilação do DNA é usada em vertebrados para inativar o gene ou o segmento cromossômico e assegurar que ele permaneça nesse estado. É comum perceber a presença de metilações em genes não expressos, de forma que a metilação CG foi adotada como modo de obstrução ao início de transcrição em segmentos inativos do genoma. Porém, quando o gene retoma sua ativação (e caso isso ocorra), como, durante o desenvolvimento do organismo, ele acaba perdendo suas metilações. O cromossomo X inativo das fêmeas de vertebrados demonstra o papel da metilação no silenciamento gênico, pois, quando este é metilado, sua tendência é ficar permanentemente condensado.

Enquanto a metilação inibe, a *acetilação de histonas* (-CH<sub>3</sub>CO) relaxa a estrutura dos nucleossomos, permitindo que os fatores de transcrição se liguem ao DNA, possivelmente pela neutralização das cargas positivas das histonas. Desacetilases promovem a retirada desses grupos acetil das histonas e

A *recomposição alternativa* ou *splicing* alternativo é uma junção diferencial de éxons para formar mais de uma variante do mesmo gene. O significado deste mecanismo é que ele aumenta a diversidade dos produtos genéticos. Um exemplo proeminente no nível de DNA é a troca da classe de imunoglobulina. No nível do RNA, verifica-se o exemplo da distrofina, onde ocorre um gene mutado nas distrofias musculares de Duchenne e Becker. Na distrofia ocorrem sete promotores que são utilizados para gerar expressões específicas em diferentes tecidos.

Evidências científicas sobre a **epigenética** mostram que alterações nos genes não são suficientes para responder determinadas questões. Trata-se de um novo campo da biologia que desvenda os mistérios de como o ambiente pode influenciar o comportamento das células sem modificar o código genético. O hábito de fumar, o uso de esteróides e certas drogas também influenciam as funções genéticas por interagir química ou fisicamente com o DNA. Gêmeos idênticos (possuidores do mesmo genoma) apresentam diferenças em seu comportamento e fisiologia que poderiam ser explicados dessa forma.

restauram a repressão pela cromatina. Torna-se muito provável a associação entre os dois processos: metilação e desacetilação. Da mesma forma, a desmetilação do DNA permitiria a adição de grupos acetil pelas acetiltransferases, alterando a estrutura do nucleossomo.

A **epigenética** é um campo recente da Genética Molecular que estuda mudanças reversíveis e herdáveis no genoma funcional que não alteram a sequência de nucleotídeos do DNA. Existem três mecanismos principais de alterações epigenéticas: metilação do DNA, modificações de histonas e ação de RNAs não codificadores. As modificações de histonas mais bem estudadas são as acetilações, fosforilações e ubiquitinações, formando o chamado código de histonas, determinando a conformação da cromatina. Já a ação de RNAs não codificadores (ver capítulo 5) está relacionada ao silenciamento pós-transcricional de genes através do mecanismo de *RNA de interferência* (RNAsi). Veja mais a diante onde ocorre o bloqueio da tradução ou degradação do RNAm alvo. Além da ação bloqueadora da transcrição, os RNAsi podem ser associados à metilação de sequências de DNA. A pesquisa na área da epigenética alcança implicações na Agricultura, na Biologia e doenças humanas, incluindo o entendimento sobre células-tronco, câncer e envelhecimento. Inclui o estudo de como os padrões de expressão são passados para os descendentes; como ocorre a mudança de expressão espaço-temporal de genes durante a diferenciação de um tipo de célula e como fatores ambientais podem mudar a maneira como os genes são expressos. Todos esses mecanismos parecem estar interligados para a organização estrutural da cromatina, tornando-a mais acessível ou não aos fatores de transcrição.

O ambiente influencia de forma importante as mudanças epigenéticas. Qualquer alteração ambiental, como ataque de patógenos, **xenobióticos**, tipo de alimentação, pode acarretar mudanças epigenéticas. O estresse ambiental, incluindo a hibridação e a **poliploidização**, são determinantes na ocorrência de variações epigenéticas. Sendo assim, a epigenética está intimamente relacionada com o aumento de variabilidade fenotípica dos indivíduos resultando em relevante importância para a evolução.

Existem **genes estruturais** que codificam polipeptídios que participam do metabolismo ou na biossíntese. Os genes ditos **regulatórios** produzem RNA ou proteínas que interagem com outras sequências, induzindo ou reprimindo a expressão gênica. As proteínas regulatórias são proteínas com capacidade de se ligar ao DNA (não confundir com as polimerases), reconhecendo sequências específicas (ou **sequências regulatórias**), encaixando-se geralmente ao sulco maior do DNA, reconhecendo átomos expostos nesse local.

As proteínas de ligação ao DNA possuem domínios de ligação ao DNA. Nesses domínios alguns aminoácidos fazem contato geralmente através de pontes de H com bases nitrogenadas. Essas proteínas de ligação são agrupa-

das de acordo com esses motivos encontrados dentro dos seus domínios de ligação. Os motivos de ligação ao DNA são bem conhecidos, são estruturas simples, como as alfa-hélices. Esses motivos se agrupam geralmente em *dímeros* que se dispõem face a face (Figura 2).

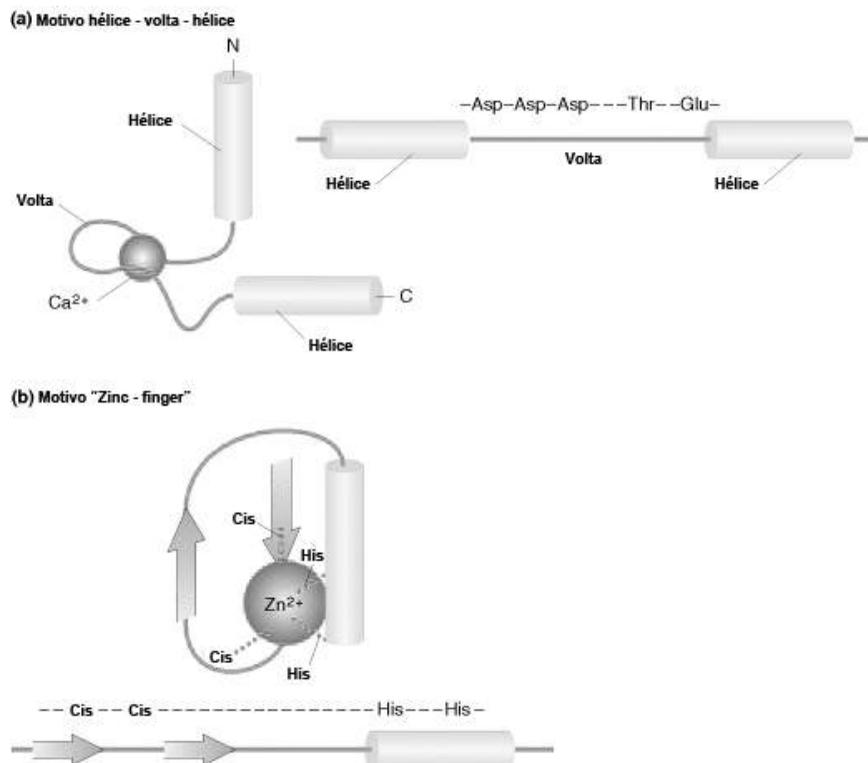


Figura 2 - Estruturas em bola e em bastão para motivos proteicos de ligação ao DNA (a) motivo em hélice – volta – hélice. (b) motivo *zinc finger* (ou dedos de zinco). (c) motivo dedos de zinco fixando-se a dupla hélice de DNA formando um complexo dedos de zinco-DNA. Note que os bastões são estruturas proteicas em alfa-hélice, as linhas são estruturas proteicas ao acaso. Fitas são estruturas proteicas em folha-beta. As esferas são íons.

Fonte: Griffiths et al. (1999).

### 3. Em eucariotos: regulação da transcrição

“O que é verdade para *E. coli* também é verdade para o elefante”, disse Monod, porém existem alterações, no caso da expressão gênica. As bactérias mostram economia na sua regulação. O mesmo não ocorre em eucariotos, nos quais a regulação é complexa, e um gene pode responder de muitas formas a vários sinais diferentes. Retornando ao tema da transcrição de genes, a transcrição eucariótica contém várias proteínas que se ligam à maquinaria de início da transcrição e possibilitam que ela seja mais eficiente ou não. O início da transcrição é caracterizado pela formação de um aparelho basal de transcrição.

**Xenobióticos:** (do grego, xenos = estranho) são moléculas químicas estranhas a um organismo ou sistema biológico. Considera-se que não é normalmente produzido. Também se aplica a substâncias presentes em concentrações muito mais elevadas que o nível normal. Em específico, medicamentos tais como antibióticos são xenobióticos em humanos porque o corpo humano não os produz nem as substâncias fazem parte da dieta.

**Poliploidização:** aumento do conjunto de cromossomos (2n diplóide) ou genomas – como triploide (3n), tetraploide (4n), pentaploide (5n), etc.

**Genes estruturais:** genes que especifica a síntese de um polipeptídeo.

**Genes regulatórios:** genes que controlam a taxa de expressão de outro gene.

### Outras definições de genes

**Gene repressor:** gene que codifica uma proteína regulatória repressora.

**Gene segmentar:** gene que controla a formação de segmentos adjacentes no corpo de *Drosophila*, gene homeiótico (veja o item “O desenvolvimento e seu controle”).

**Gene seletor:** um gene que influencia o desenvolvimento de segmentos corpóreos específicos em *Drosophila*, gene homeiótico.

**Gene supressor tumoral:** um gene cujo produto está envolvido na repressão da divisão celular.

**Genes homólogos:** genes que evoluíram de um gene ancestral comum.

**Genes ortólogos:** genes homólogos presentes em espécies diferentes.

**Genes parálogos:** genes homólogos presentes dentro da espécie.

Este consiste de RNA polimerase, uma série de fatores gerais de transcrição e proteínas acessórias repressoras e ativadoras. Esse aparelho basal de transcrição reúne-se sobre a sequência promotora (ou promotor) *upstream* ao gene a ser transcrito, e sua função é melhorar a afinidade da enzima RNA polimerase ao promotor. Lembre-se de que o promotor não é uma proteína, mas uma sequência de DNA *upstream* ao gene, ao qual se ligam todas estas partículas.

A RNA polimerase desempenha o papel principal. Em geral, temos três RNAs polimerases nos organismos eucarióticos: I, II e III. Para a produção do pré-RNA<sub>m</sub> (ou transcrito primário: proteínas em geral), alguns **snRNAs** e **snoRNAs**, a polimerase II é a responsável. Quando esta se fixa ao promotor, os fatores de transcrição são liberados, e a enzima começa a fazer seu trabalho de polimerização.

O passo a passo para a afinidade da RNA polimerase II pelo seu promotor é um processo que se inicia num local específico do promotor, chamado TATA box. Essa sequência é rica em nucleotídeos T e A, localizado a cerca de 25 nucleotídeos *upstream* ao início do gene. Os promotores merecem atenção especial pela sua importância na transcrição e na regulação do gene. Os promotores são encontrados em todos os organismos, porém são mais conhecidos em bactérias. O exame dos promotores mostra que estes variam de sequência, trechos curtos, mas são comuns a muitos promotores. Esses trechos curtos são comuns e, como todos os trechos conservados nos organismos, são *sequências consenso*. Uma das sequências mais comuns está na posição 10, chamada **Pribnow box** ou **TATA box** (5' TATAAT 3'). Essa é uma sequência consenso; portanto, representa os nucleotídeos mais comumente encontrados.

Um fator de transcrição TFIID liga-se ao TATA box e promove uma torção no DNA que é essencial para o início do processo. Após o TFIID, liga-se de forma adjacente o TFIIB, os próximos fatores também se ligam (TFIIE, TFIIIF e TFIIH). TFIIH utiliza ATP, fosforilando a enzima RNA polimerase II, mudando sua estrutura conformacional de forma que ela se pode liberar desse complexo de fatores e proteínas e possa caminhar em direção ao gene e transcrevê-lo. O sítio de fosforilação é uma cauda peptídica da RNA polimerase, em que a TFIIH contém uma enzima com atividade quinase em suas subunidades.

As proteínas ativadoras e co-ativadoras (acentuadores ou *enhancer*) do aparelho basal de transcrição (Figura 3) promovem a transcrição, facilitando a montagem sobre o promotor. Alguns deles acumulam a função de acetiltransferase (adicionam acetilas) e facilitam mais ainda a transcrição pela alteração da estrutura da cromatina. As proteínas repressoras, ao contrário, inibem a transcrição, silenciando-a pela competição que fazem pelos sítios de ligação dos acentuadores. Uma questão que chega a ser insólita é o fato

de os acentuadores serem capazes afetar a transcrição em promotores distantes. Sugere-se que o DNA se dobre de forma a posicionar o promotor e o acentuador de forma próxima. *Insuladores* ou elementos limítrofes, como a própria palavra diz, insulam ou isolam o efeito dos acentuadores, de forma que eles não fiquem acentuando a transcrição de qualquer gene. Porém, ainda pouco se conhece sobre essa ação, mas exemplos já são conhecidos.

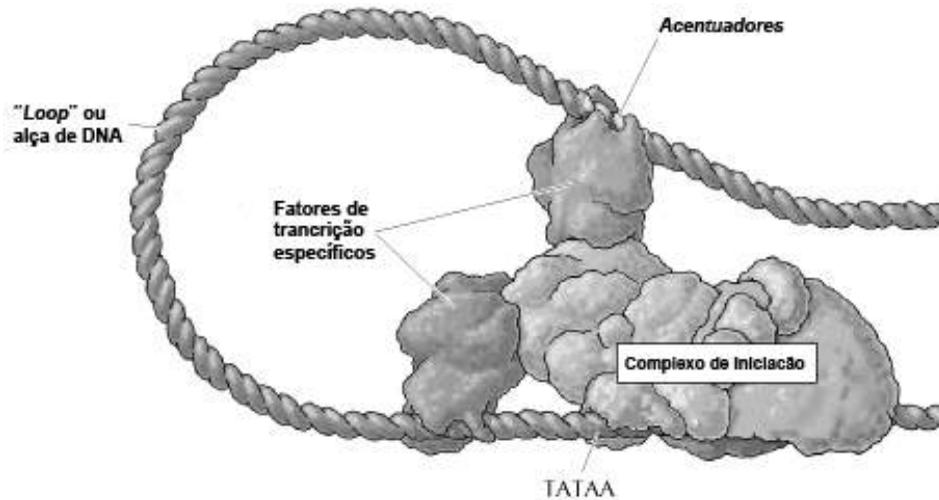


Figura 3 - Fatores de transcrição ligam-se à acentuadores ou *enhancers* os quais são capazes de interagir com fatores de transcrição gerais no promotor devido à formação de *loops* no DNA.

Fonte: Cooper e Hausman (2009).

#### 4. Estabilidade do RNA como evento regulador

Uma abordagem diferente e relativamente recente que vêm chamando a atenção, é o uso do RNA como regulador, devido ao fato que existem eventos de regulação gênica dependentes do RNA. Imagine que a taxa de síntese e degradação do RNAm regula a quantidade de RNAm disponível, que, por sua vez, determina a quantidade de proteína a ser produzida. Destaca-se o silenciamento do RNA (RNAsi – *short interfering RNA*, RNAi ou RNA de interferência). Essas pequenas moléculas de RNA são intermediárias de um processo conhecido como RNA *interference* ou interferência mediada por RNA. Esse fenômeno foi descrito pela primeira vez no nemátodo *Caenorhabditis elegans*, em 1998, quando se observou que a presença de apenas algumas poucas moléculas de RNA dupla-fita (RNAds) eram capazes de abolir a expressão de um determinado gene com sequência similar àquele RNAds.

O *silenciamento gênico* corresponde a uma interação entre sequências homólogas de DNA e RNA; portanto, é dependente de homologia entre essas sequências. Podemos definir dois tipos. Primeiramente, silenciamento gênico

pós-transcricional, em que ocorre a degradação de RNAs mensageiros homólogos no citoplasma, levando à inibição da tradução. Outro seria o silenciamento gênico transcricional relacionado com o bloqueio na transcrição, induzida por um RNA antisense derivado do próprio DNA, metilando a região promotora.

O RNAm possui meia vida curta, porém variável, de minutos a horas, dias ou meses. Ele é degradado por vários tipos diferentes de ribonucleases. O mRNAm bacteriano é o mais facilmente degradável, justamente por não possuir elementos protetores como o cap 5' e a cauda poli A 3', tipicamente durando apenas alguns minutos. O silenciamento do RNA aparece em fungos, plantas e animais, mas cogita-se muito mais que isso. O papel do silenciamento gênico tem intrigado a ciência. Muitos postulam que seja defesa contra ácidos nucléicos estranhos, proteção contra transposons eventuais, em plantas, a expressão gênica de multigenes ou genes duplicados.

## 5. O desenvolvimento e seu controle

Morgan foi um dos primeiros descobridores das mutações em moscas. Ele e seus alunos, como Calvin Bridges, nas décadas de 10 e 20, encontraram dezenas de mutantes em *Drosophila*. Essas alterações mostravam que o desenvolvimento embrionário deveria ser regulado por alguns genes então desconhecidos. Por exemplo, na mutação chamada *antennapedia*, pernas apareciam no lugar de asas (**gene homeótico**). Uma abordagem para o estudo do desenvolvimento pode ser o isolamento e a caracterização de mutantes com o estudo das proteínas responsáveis que são codificadas e que produzem tal defeito.

Uma das perguntas essenciais é: quais genes são responsáveis pelo crescimento e pelo desenvolvimento? Já vimos que existem genes do tipo *housekeeping* ou constitutivos, que mantêm o estado basal de metabolismo celular. Não são esses os conjuntos de genes que nos referimos, pois buscamos genes envolvidos com a formação de órgãos e tecidos, especificando tipos celulares, plano corpóreo, número, identidade e outros padrões corporais.

As mutações mais interessantes do ponto de vista científico são as que causam algum pequeno defeito facilmente perceptível na progênie, produzindo segmentos diferentes, apêndices etc. Pesquisas mostraram em moscas existem alguns genes, chamados complexo *Hox*, encontrados em oito diferentes *loci* gênicos situados no terceiro cromossomo de *Drosophila*. A mutação *antennapedia* é produzida por cinco genes *Hox*. No embrião em desenvolvimento, os genes *Hox* (Figura 4) são expressos em locais específicos, formando padrões específicos de desenvolvimento. Hoje sabemos da existência do *homeobox*, um curto domínio proteico encontrado nessas proteínas *Hox*. O homeobox codifica um homeodomínio de 60 aminoácidos. Essa sequência é muito comum nas pro-

teínas Hox. As similaridades desse homeodomínio com proteínas repressoras (motivo hélice-giro-hélice) mostrou que as proteínas Hox ligam-se ao DNA, funcionando nos mesmos princípios já vistos em relação à ativação e a desativação de genes. Os genes Hox foram encontrados também em grande número de animais, mostrando sua presença em ancestrais distantes e evidenciando que a maioria de genes ligados ao desenvolvimento são comuns aos grupos animais.

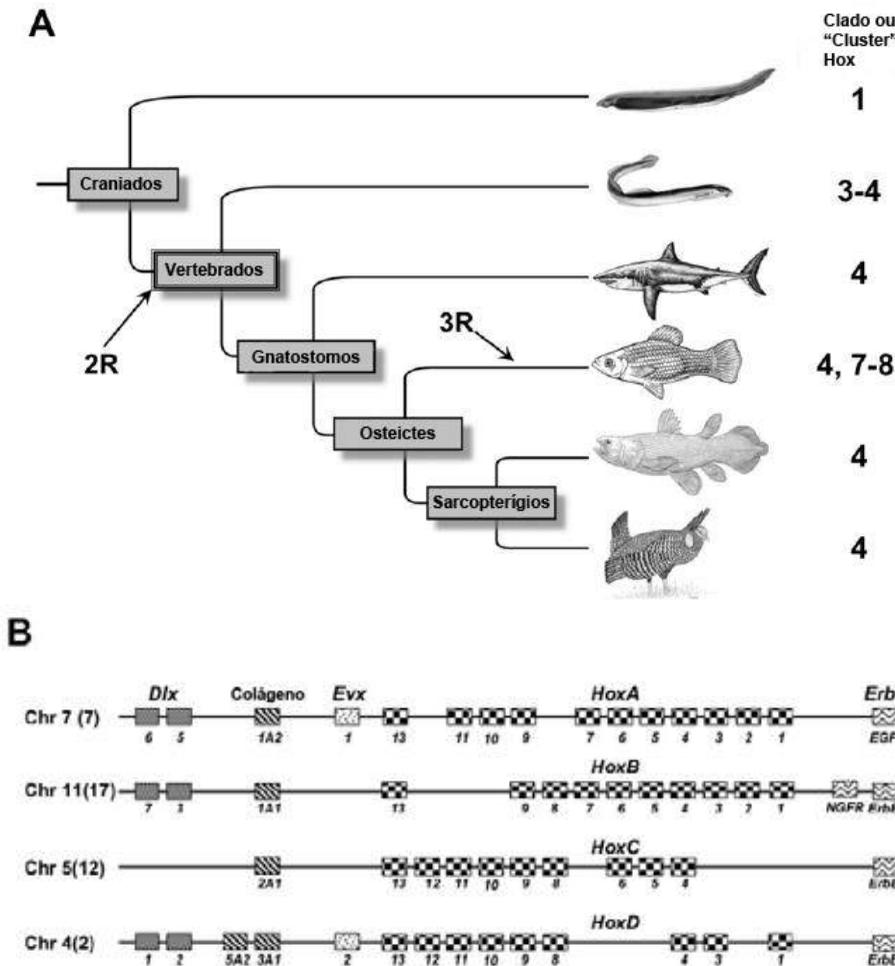


Figura 4 – (A) Filogenia do grupo craniata mostrando localizações filogenéticas com duplicações gênicas para o gene *Hox*. (B) Representação esquemática da família *Hox* em vertebrados.

Fonte: Lynch e Wagner (2009).

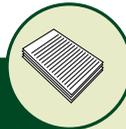
Em humanos, aproximadamente 3% de todas as gestações humanas terminam no nascimento de uma criança com um problema genético. A criança, nesses casos, é perfeitamente sadia, apresentando defeitos isolados como a fenda labial, problemas cardíacos ou defeito do tubo neural, os quais, apesar de natureza multifatorial, em boa parte, são corrigidos com relativa facilidade com

O processo de *splicing* alternativo modificou a noção do gene e levou ao entendimento de que um gene poderia produzir múltiplas proteínas diferentes a partir de uma mesma sequência. Nesse novo contexto, a definição ficou resumida na expressão conhecida de "um gene, várias proteínas", que representava a noção de uma sequência de DNA que possui éxons como segmentos independentes e que esses éxons, dependendo do contexto celular poderiam fazer proteínas diferentes. Lembre-se do éxon *shuffling* ou embaralhamento de éxons eucariótico.

cirurgia. Outras anomalias derivam de uma alteração cromossômica ou mesmo de alterações em genes isolados, quando o indivíduo pode apresentar certa gama de sintomas ou *síndrome*, com várias anomalias congênitas. O desenvolvimento normal de um indivíduo envolve processos de replicação, migração e diferenciação rigidamente controlados temporal e espacialmente, numa sequência pré-programada de ativação e desativação de genes.

A ciência da *dismorfologia* examina nos pacientes diferenças, tanto aquelas facilmente visíveis quanto aquelas mais sutis, buscando estabelecer padrões que podem ser reconhecidos durante o diagnóstico. Muitas anomalias envolvem sistemas orgânicos, defeitos cardíacos congênitos, malformações intestinais etc., que precisam de avaliação genética precisa.

### Texto complementar



#### Células tronco embrionárias

Linhagens de células tronco embrionárias (células ES - do inglês *embryonic stem*) são células não diferenciadas, derivadas do botão embrionário de blastócistos, que têm como característica principal pluripotência. Ou seja, quando reintroduzidas em um blastocisto, as células ES possuem capacidade de retomar o desenvolvimento normal, colonizando diferentes tecidos do embrião, incluída a linhagem germinativa. Somente aquelas linhagens de células ES capazes de colonizar a linhagem germinativa de um embrião são consideradas altamente pluripotentes.

Quando cultivadas em condições específicas, as células ES podem ser mantidas em seu estado não diferenciado por múltiplas divisões celulares. Por outro lado, essas células podem ser induzidas a iniciar um programa de diferenciação *in vitro*. Por exemplo, quando cultivadas em suspensão, as células ES formam espontaneamente agregados de células diferenciadas chamados “corpos embrioides” (Ebs “*embryoid bodies*”), que simulam o desenvolvimento de um embrião pré-implantado. Através de análises morfológicas, imuno-histoquímicas e moleculares, encontra-se uma grande variedade de linhagens embrionárias dentro dos EBs – hematopoética, neuronal, endotelial, cardíaca e muscular.

Essas propriedades das células ES levaram ao seu uso como modelo *in vitro* de desenvolvimento embrionário precoce. Nelas, podem ser estudados os mecanismos de diferenciação celular, o processo de inativação do cromossomo X, e os efeitos de substâncias tóxicas e biologicamente ativas no desenvolvimento embrionário *in vitro*.

**Modelos animais para doenças genéticas humanas:** nos últimos anos, células ES vêm sendo utilizadas para produzir camundongos transgênicos e modelos animais de doenças genéticas humanas. Esses modelos animais são importantes ferramentas para o estudo das patologias associadas com uma doença específica, além de servirem como um sistema no qual podem ser testadas, novas terapias. O recente desenvolvimento de técnicas de manipulação do genoma de camundongos faz do camundongo transgênico o melhor sistema disponível para o estudo de doenças genéticas humanas.

Fonte: adaptado de Kerkis et al. (2001).

## Síntese da Parte



A regulação gênica consiste na expressão do gene mediada por diversos e variados mecanismos em organismos procarióticos e eucarióticos. O conhecimento da expressão gênica é um pouco limitado em eucariotos porque cada gene possui uma regulação específica. Proteínas regulatórias são essenciais nesse processo e possuem domínios de ligação ao DNA. Sua ação é mais intensa na iniciação da transcrição, através de interação com a RNA polimerase. Outros mecanismos incluem a regulação por RNA, entre outros.

## Atividades de avaliação



1. Procure, em livros de embriologia e desenvolvimento, como ocorre a diferenciação dos tecidos multicelulares. Faça um resumo.
2. A clonagem de animais de companhia, como cães e gatos, vêm se tornando um bom negócio nos EUA e na Europa. Entretanto, como a epigenética pode “atrapalhar” a clonagem? Dica: lembre que a epigenética depende do ambiente gênico.
3. Pesquise doenças genéticas que ocorrem devido a alterações na expressão do gene, por exemplo, a talassemia.
4. Procure em livros os domínios e motivos de ligação ao DNA.
5. Procure resumir como o câncer é afetado pelas alterações na regulação normal dos genes.
6. Procure o significado de “*microarrays*” ou “microarranjos” e como eles podem ser usados para análise de expressão de genes.

## Leituras, filmes e sites



### Leituras

- DAWKINS, R. **O gene egoísta**. 1. ed. Belo Horizonte: Itatiaia, 2001. 230 p.
- SAGAN, C. **Variedades da experiência científica**. 1. ed. São Paulo: Companhia das Letras, 2008. 304 p.

### Filmes

- A guerra do fogo (Dir. Jean-Jacques Annaud, 1981) - evolução humana.
- O curandeiro da selva (Dir. John McTierman, EUA, 1992) - farmacologia e meio ambiente.
- A ilha (Dir. Michael Bay, EUA, 2005) - clonagem humana.
- Quanto vale ou é por quilo? (Dir. Sérgio Bianchi, 2005) - sociologia e miséria social.
- Escola da vida (Dir. William Dear, 2005) - educação e biologia.
- O tráfico humano (Dir. Christian Duguay, 2005) – sociologia e denúncia.

### Sites

#### Bases de dados públicas

- <http://www.sciencenet.com.br>
- <http://www.isiknowledge.com>
- <http://www.scielo.br>
- <http://www.dominiopublico.gov.br>
- <http://www.prossiga.br/basesdedados>

#### Revistas e sites de divulgação científica

- <http://www.biotecnologia.com.br> (Revista Biotecnologia)
- <http://www.sciam.com.br/> (*Scientific American* - versão nacional)
- <http://cienciahoje.uol.com.br/> (Ciência Hoje)

- <http://www.uenf.br/index.html/jornalonline> (Jornal UENF universitário online)
- <http://www.agencia.fapesp.br/> (Agência de Notícias da FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo)
- <http://www.abjc.org.br/> (Associação Brasileira de Jornalismo Científico)
- <http://www.usp.br/agen/> (Agência USP de notícias)
- <http://www.comciencia.br/comciencia/> (Revista eletrônica de jornalismo científico)

## Referências



COOPER, G. M.; HAUSMAN, R. E. **The cell: a molecular approach**. 5. ed. Sunderland: Sinauer Associates, 2009. 771 p.

GRIFFITHS, A. J. F.; MILLER, J. H.; SUZUKI, D. T.; LEWONTIN, R. C.; GELBART, W. M. **An introduction to genetic analysis**. New York: W. H. Freeman and Company, 1996. 915 p.

KERKIS, A.; SOUKOIAN, M.; KERKIS, I.; MERKEL, C.; MELLO, M. R. B.; PEREIRA, L. V. Células tronco embrionárias. **Biotecnologia Ciência & Desenvolvimento**, Brasília, n. 20, p. 20-25, 2001.

LYNCH V. J.; WAGNER, G. P. Multiple chromosomal rearrangements structured the ancestral vertebrate hox-bearing protochromosomes. **PLoS Genetics**, Detroit, v. 5, n. 1, p. 1-11, 2009.

PIERCE, B. **Genética: um enfoque conceitual**. 1. ed. Editora Guanabara Koogan: Rio de Janeiro, 2004. 788 p.



**Parte**

**5**

**DNA recombinante**



# Considerações iniciais e histórico

## Objetivo

- Compreender as técnicas básicas do DNA recombinante, habilitar a discussão das bases teóricas dessas técnicas e suas conexões entre si e com a atual perspectiva biotecnológica.

## 1. Aspectos gerais

O homem sempre foi capaz de alterar a paisagem à sua maneira. A agricultura surgiu no sudoeste asiático por volta de 10.000 a.C. e do pastoreio, a domesticação de carneiros, cabras e ovelhas acerca de 3.200 a.C. Com o advento do sedentarismo humano, a moradia torna-se fixa em aldeias, próximas a rios e inicia-se a domesticação de animais, como o dromedário, boi, cão, cabra etc. Utilizam também roupas de tecido, lã, linho e algodão. Em uma região do Oriente próximo, chamada crescente fértil, ocorre a invenção da agricultura, com cereais como o trigo e a cevada. Começa, então, a economia de produção, visando o lucro. A economia humana forçou a escolha das melhores raças animais, a agricultura torna-se negócio. Nas prateleiras dos estabelecimentos comerciais, o homem esmera-se em apresentar os melhores e mais suculentos frutos de seu trabalho. Portanto, não deveria haver surpresa quando se fala em *engenharia genética*. Nossos pais e nossos avós a faziam bem antes de nós, quando escolhiam as “melhores” sementes para plantar ou arranjavam “um bom reprodutor” para sua fazenda. As plantas e os animais dos quais nos alimentamos ou que usamos para produzir os bens de consumo são muito diferentes daquelas que encontramos na natureza. Os cruzamentos são intencionais, como o que fazemos com os cães, gerando uma variedade de cores, formas e características desejáveis para os seres humanos. O cão doméstico é muito diferente do seu ancestral: o lobo cinzento. Hoje temos tecnologia para acelerar algumas alterações no laboratório de Biologia ou Genética Molecular.

Atualmente, podemos dizer que é quase impossível separar a Biologia Molecular do próprio estudo genético, devido à própria natureza das moléculas da hereditariedade. A Genética Molecular é uma área de interface geralmente mais voltada para a área da saúde, abrangendo o estudo de patologias genéticas. A genética clássica envolve o conhecimento epistemológico acumulado durante anos de estudo antes do advento da estrutura do DNA e do início da Genética Molecular. Fruto de grandes pensadores como Mendel, Darwin, Sutton, Morgan, Fischer, Haldane e Wright, esses pesquisadores e outros fundaram a genética clássica. Esta envolve o estudo de cruzamentos, de frequências gênicas, mapas de ligações, mapas genéticos, entre outros. Aplicações mais evidentes: a genética de populações, o melhoramento genético, usos em sistemática, em ecologia etc.

**FISH** (*Fluorescence In Situ Hybridization*) é uma técnica citogenética desenvolvida por Christoph Lengauer usada para detectar e localizar a presença ou ausência de sequências de DNA em cromossomos. FISH usa sondas fluorescentes (ou radioativas) que se ligam às partes do cromossomo em que as sondas se hibridizam, em alto grau de similaridade. Para a visualização, é necessário o uso de microscopia de fluorescência. Pode ser usado também para localizar mRNAs específicos dentro de amostras de tecidos, apresentando o perfil específico espacial e temporal de expressão de genes dentro de células e dos tecidos. Algumas doenças importantes diagnosticadas com FISH incluem a síndrome de Prader-Willi, Síndrome de Angelman, leucemia mielógena crônica, leucemia linfoblástica aguda, Cri-du-chat, síndrome velocardiofacial e síndrome de Down. Seu principal problema é o altíssimo custo laboratorial e de mão-de-obra especializada. FISH é uma das técnicas mais proeminente da citogenética molecular, a combinação da Biologia Molecular com a Citogenética.

Mesmo dentro dessa perspectiva, há muita história para se contar. O trabalho com o DNA *in vitro* sempre foi muito difícil e dependeu também do avanço das técnicas bioquímicas, que, a partir dos anos 50, ganharam grande avanço. Muitas descobertas, entretanto, tiveram que aguardar que a tecnologia pudesse resolver suas indagações. Os anos 70 foram palco de uma revolução dentro de uma revolução, que foi o advento da reação em cadeia da polimerase, ou PCR. A partir desse momento, a biologia dos ácidos nucleicos escapou dos muros da academia e tomou de assalto a vida do cidadão comum. A partir desse momento, foi impossível refrear as inúmeras aplicações, desde resultados forenses até pesquisas ecológicas, paleontológicas etc. Resultados de testes de DNA inundam a mídia hoje em dia. O cidadão de nível médio talvez tenha problemas se for indagado sobre o que seja uma proteína, porém certamente saberá o peso que um teste de DNA pode fazer a sua vida e à sociedade. Com base nele, condenados à morte são libertos, seguros podem ser pagos, herdeiros são reconhecidos. Entretanto, o biólogo deve reconhecer na PCR e em outras técnicas as reais possibilidades como ferramentas de trabalho.

## 2. Breve histórico

Como fazer cortes na molécula de DNA? A resposta foi oferecida por Werner Arber, que, em 1968, anunciou a **endonuclease de restrição** de *E. coli*. Na época verificou-se em experimentos de transformação induzida, que bactérias sempre degradavam o DNA de outras bactérias introduzidos experimentalmente nelas. Hoje sabemos que distintas classes de bactérias contêm distintas classes de **enzimas de restrição**. Uma questão interessante é que a bactéria protege o seu próprio DNA, através de **metilações**. Arber, então, encontrou uma nova classe de enzimas, as enzimas de restrição (do verbo “restringir”) (Figura 1). Atualmente se conhecem várias classes de enzimas de restrição, como as do tipo I e II. As do tipo II são utilizadas nos laboratórios para fazer cortes nas moléculas de DNA. Na época, Arber encontrou as enzimas de restrição do tipo I, as quais cortam o DNA em locais distantes do local desejado, portanto, inadequadas para o uso no laboratório. Foi ganhador do prêmio Nobel em 1978. As enzimas de restrição do tipo II, essas sim, são endonucleases de restrição específicas, o que foi confirmado pelo geneticista microbiano Hamilton Smith e seus colegas, Wilcox e Kelley, na Universidade Johns Hopkins.

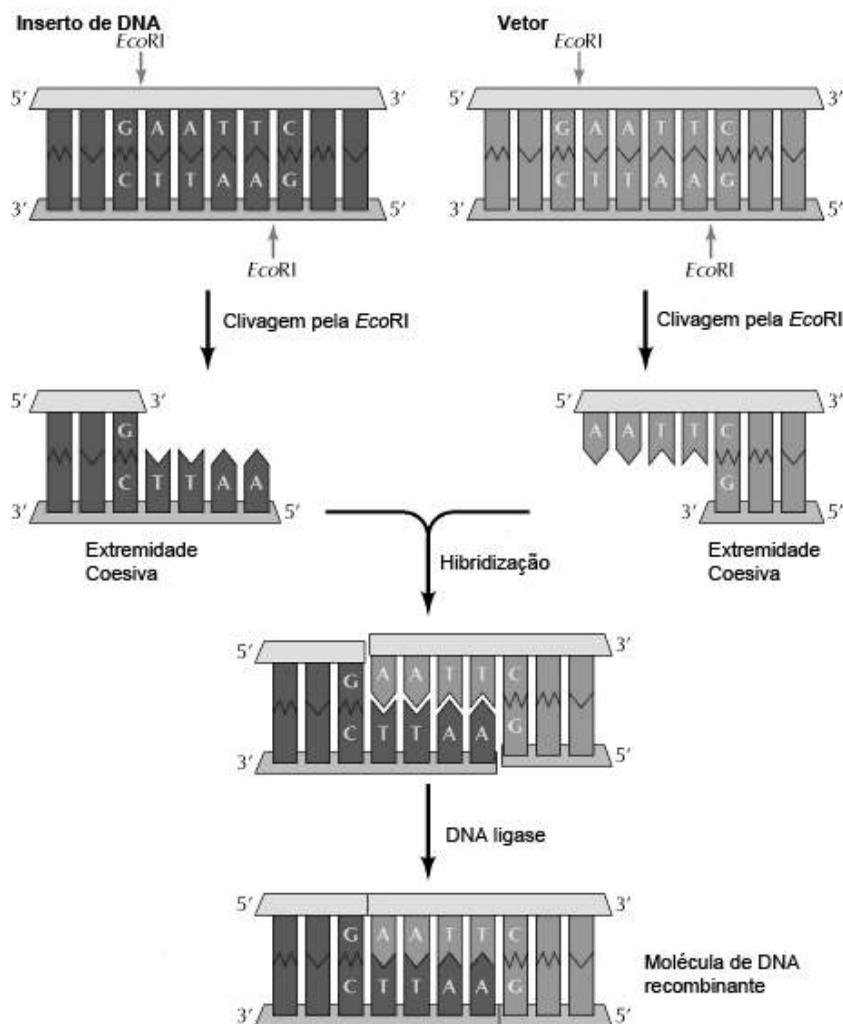


Figura 1 - O DNA recombinante é feito pela digestão de uma endonuclease de restrição, que gera extremidades de fitas simples, as quais se tornam complementares. A finalização é feita pela DNA ligase.

Fonte: Cooper e Hausman (2009).

A reação da PCR foi desenvolvida por Kary Mullis, em 1983, nos Estados Unidos. Alguns trabalhos pioneiros foram feitos em 1971, por Gobind Khorana, que descreveu o princípio básico de replicar um DNA usando dois *primers*. No processo original de Mullis, o DNA dupla fita foi separado em duas fitas simples pelo aquecimento a 96°C. Nessa temperatura, no entanto, a DNA polimerase obtida de *E. coli* foi destruída, de forma que, a cada novo ciclo, o pesquisador deveria reabastecer a reação de enzima, tornando o processo muito ineficiente. Esse problema só foi resolvido cerca de dez anos mais tarde, com o uso da *Taq* DNA polimerase. O registro inicial dessa enzima, purificada da bactéria *Thermus aquaticus*, obtida em fontes termais, foi publicada em 1976.

**Identificando genes causadores de patologias:** na comparação entre genes de determinadas patologias, obtidos de pessoas com fenótipo normal em relação com os obtidos com pacientes, é possível observar deleções, inserções e outras mutações. Dessa forma, é possível fornecer pistas de qual (quais) gene(s) são responsáveis. Para isso são usados heredogramas familiares para identificar os possíveis doadores das sequências. Utilizando essas sequências, em pacientes e familiares com fibrose cística nos quais se constatou que, as proteínas preditas fornecidas por esse genes são proteínas de transporte de íons. Para confirmar, essas sequências foram usadas para transformar linhagens celulares obtidas de pacientes, cultivadas in vitro, as quais tiveram seu sistema normalizado.

Hoje sua patente na utilização da PCR vale centenas de milhões de dólares. Torna-se interessante como algumas descobertas um pouco obscuras, anos mais tarde, tornam-se de importância prática imensa. Uma de suas desvantagens é uma taxa relativamente alta de erro na polimerização. Outras enzimas estão em utilização, como a Pfu polimerase, extraída de *Pyrococcus furiosus*, com uma das menores taxas de erro de polimerização. A enzima Vent é extraída de *Thermococcus litoralis*, também conhecida como *Tli* polimerase. Estas são usadas somente em alguns casos específicos.

# Técnicas gerais

## 1. Técnicas associadas: eletroforese e enzimas de restrição

Imagine uma fita de DNA cortada em vários pedaços. Como seria possível visualizá-los? A **eletroforese** pode ser uma solução adequada, a princípio. Com ela, o DNA migra através de uma rede de polímeros de polissacarídeos, puxado por uma corrente elétrica, em direção ao polo positivo. Ao caminhar, as fitas menores vão mais rápidas; as maiores, mais lentas, se enredam e param mais rapidamente; as intermediárias se separam e, com o tempo, não conseguem mais caminhar, formando *bandas*. Cada banda contém várias fitas de DNA do mesmo tamanho, mas não significa que sejam iguais, apenas têm o mesmo peso. Para proteínas, a eletroforese segue o mesmo protocolo analítico. Depois de separados, é possível isolar as bandas, cortando-as do gel, retirando-se o gel com lavagens e centrifugação e usar a sequência.

Para acompanhar a corrida eletroforética, sempre se utiliza um *marcador de peso molecular*. No caso do DNA, utiliza-se uma coleção de pedaços de ácidos nucleicos com peso molecular conhecido. Eles se distribuem e permitem visualizar, por comparação do perfil eletroforético da amostra, o seu peso molecular. Na verdade não se considera como “peso molecular”, mas sim *massa molecular relativa*. Outra vantagem é que o uso do marcador de peso permite verificar se a corrida foi adequada ou não.

O DNA não é corável, ou marcado, com os mesmos corantes que se utilizam para a proteína. As bandas no gel de agarose são invisíveis. Para vê-las é possível usar vários tipos de marcação, como isótopos radioativos para marcar o DNA, como o  $^{32}\text{P}$  (Figura 2). O fósforo incorpora-se ao DNA e emite uma partícula beta que é detectável, usando filmes fotográficos (autorradiografia). Entretanto, é mais simples usar agentes intercalantes, como o brometo de etídio. Essa substância é um fluoróforo que absorve a cerca de 300 nm e libera a energia a 590 nm, na cor laranja ou amarela, tomando as bandas de DNA visíveis num transiluminador.

**Eletroforese** é o movimento de moléculas carregadas em um campo elétrico. É um método importante para separação de moléculas biológicas porque, normalmente, não afeta a estrutura nativa de biopolímeros e devido ser altamente sensível a pequenas diferenças em ambas as cargas e massa. O meio ideal de apoio para eletroforese é quimicamente inerte, de fácil manuseio e tem porosidade controlável. Embora o papel seja o suporte escolhido para algumas aplicações, como aminoácidos e pequenos peptídeos, há limitações porque a porosidade não pode ser controlada. Géis com porosidade controlável ou variável são ideais para separação de componentes de alto peso molecular (proteínas e ácidos nucleicos), porque eles permitem fácil movimentação de macromoléculas enquanto previnem convecção e minimizam difusão.

Géis também adicionam a possibilidade de combinação de separação eletroforética com efeito de peneiramento molecular – permeação do gel. Géis com efeitos peneira podem ser preparados com tamanhos de poros aproximadamente do mesmo tamanho de moléculas a serem separadas pelo controle da concentração do gel. Três tipos de géis são comumente usados – amido, poliacrilamida e agarose. Para proteínas, são usados mais comumente a acrilamida e bisacrilamida. Estes podem ser polimerizados na presença de radicais livres para gerarem géis de variados, os tamanhos de poros podem ser controlados. Isso permite ao meio eletroforético ter um efeito direto na separação eletroforética. Se o tamanho dos poros for aproximadamente igual aos das moléculas, moléculas pequenas se moverão livremente num campo elétrico e moléculas maiores serão restringidas na migração delas.

O genoma bacteriano, como vimos anteriormente, foi um dos primeiros a ser sequenciado, o qual contem cerca de um milhão de correto. Os trabalhos iniciais em Biologia Molecular do DNA eram feitos com *Drosophila* e outros genomas pequenos, utilizando métodos indiretos, como a construção de mapas genéticos (feitos através de cruzamentos entre indivíduos e computando as frequências) e citogenética (cariótipos feitos em lâminas histológicas). Genomas grandes, como os eucarióticos, tornavam a pesquisa difícil para os pesquisadores. A tarefa de isolar uma única sequência era quase impossível. É possível usar nucleases e cortar o DNA, porém, elas cortam em locais inespecíficos (digerindo os nucleotídeos a partir das extremidades – *exonucleases* ou cortando de forma inespecífica, pedaços de DNA – *endonucleases*). Pode-se ainda usar clivagem mecânica como a própria agitação do tubo, sonificadores (ultra som) etc. Entretanto, esses métodos, com o tempo do experimento, acabam por digerir totalmente o DNA e o cortam de forma aleatória, produzindo pedaços de igual tamanho, portanto, pouco úteis para se isolar um gene ou, mesmo, qualquer sequência.

Uma boa resposta para facilitar os trabalhos apareceu com a descoberta das endonucleases de restrição ou simplesmente **enzimas de restrição** (Figura 2). Essas enzimas cortam o DNA em locais específicos, sempre os mesmos para cada enzima, chamados *sítios de restrição* (por ex. GAATTC é o sítio de restrição da enzima *EcoRI*). Portanto, uma fita de DNA digerida por endonucleases de restrição fica em pedaços, contendo, em cada uma de suas extremidades, uma parte do sítio de restrição. Esse sítio no DNA compreende 4 a 8 nucleotídeos, distribuindo-se de forma aleatória no DNA.

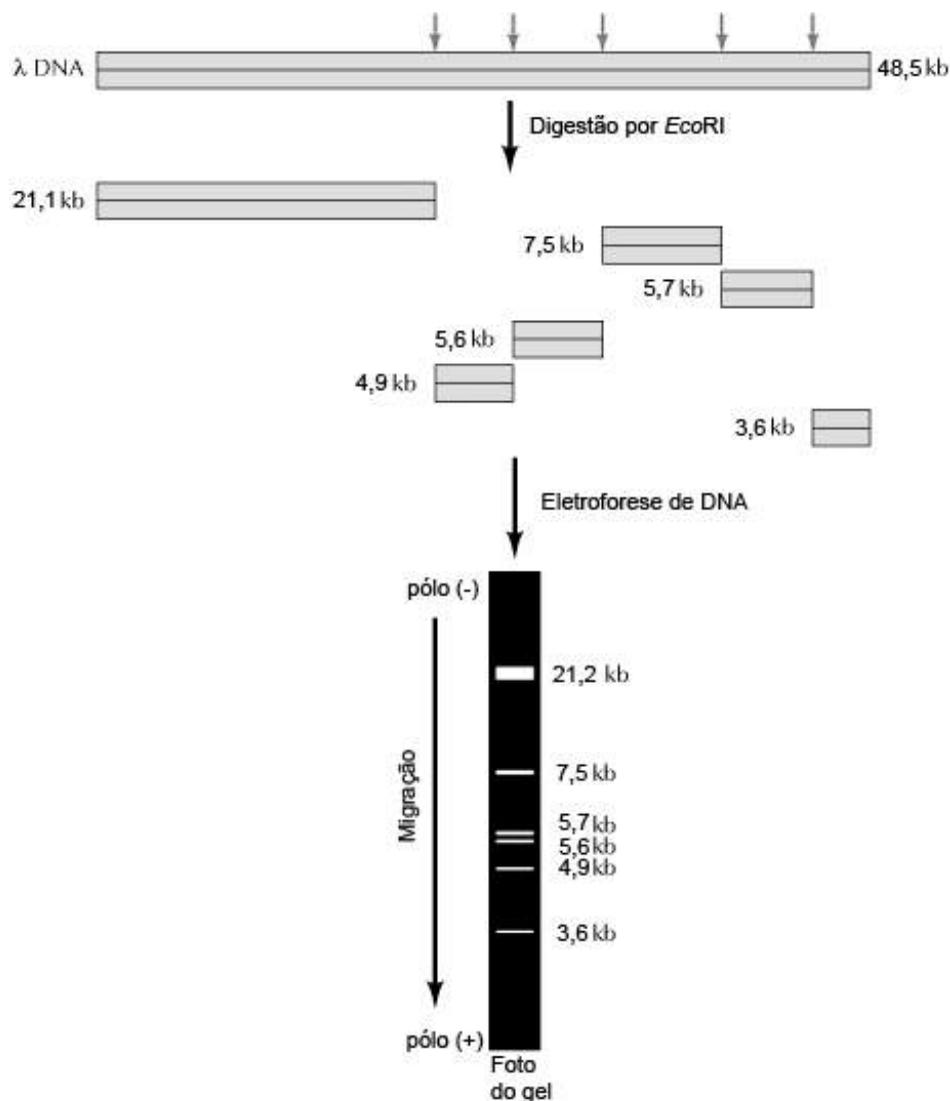


Figura 2 - Digestão de uma sequência de DNA com *EcoRI* e visualização da corrida eletroforética.

Fonte: Cooper e Hausman (2009).

### 1.1 Primeira aplicação da técnica: mapas de restrição ou RFLPs

Uma proposta muito utilizada que utiliza tanto a eletroforese quanto as enzimas de restrição é a produção de mapas de restrição. Também conhecido como RFLP ou *Restriction Fragment Length Polymorphism* ou polimorfismo de tamanho de fragmento de restrição. O princípio é relativamente simples. Para entendê-lo, vamos imaginar uma atividade. Cinco ou seis alunos recebem, cada um, uma fita de papel, todas quase do mesmo tamanho. Cada um deles faz pedaços da tira, em quantidades e tamanhos que quiser. Para

As **enzimas de restrição** são proteínas importantes para o trabalho molecular, especialmente aquelas que formam *pontas adesivas* (ou *extremidades coesivas*). Os segmentos onde elas cortam, ou sítios de restrição, formam geralmente um *palíndromo* de DNA, como o reflexo num espelho. Para a formação do DNA recombinante, ocorre, então, a hibridização entre as fitas simples. Na natureza, essas enzimas são encontradas ocorrendo naturalmente em diversos procaríotos. Como todas as enzimas, existem em poucas cópias, tornando difícil seu isolamento no laboratório. A maioria das enzimas são produzidas em laboratórios farmacêuticos que produzem reagentes para pesquisa, utilizando, por exemplo, expressão da proteína de interesse por sistemas heterólogos. O site mais bem estruturado, contendo milhares de enzimas de restrição é o REBASE (*Restriction Enzymes Base*), um repositório de dados exclusivo sobre enzimas de restrição, contendo ferramentas o reconhecimento de sítios de restrição em sequências de nucleotídeos. Apresenta também sua correspondente enzima comercialmente disponível (veja em <http://rebase.neb.com/rebase/rebase.html>).ttv

analisar, cada aluno precisa fazer, numa superfície plana, como uma mesa, a organização dos pedaços em ordem crescente de tamanho, de forma que seja possível comparar todos os conjuntos ao mesmo tempo. Verifica-se de forma um pouco grosseira, dois a dois, quais alunos formaram conjuntos mais parecidos. Talvez seja até possível, numericamente, numa matriz, organizar quais grupos são mais parecidos entre si e com outros grupos.

Da mesma forma, o DNA de várias espécies pode ser cortado por enzimas de restrição, formando um perfil eletroforético singular para aquela espécie, este pode ser comparado com outros perfis e proporcionar uma comparação evolutiva, na qual os organismos podem ser analisados através do polimorfismo dos fragmentos cortados por enzimas de restrição.

Esse tipo de trabalho é muito adequado para genomas pequenos ou para produtos gerados por PCR (um único fragmento em muitas cópias). Especialmente na microbiologia, variantes bacterianas são determinadas por métodos fenotípicos. O polimorfismo apresentado por moléculas informacionais, como proteínas e ácidos nucleicos, é utilizado como discriminador entre **variantes biológicas**. Nesses métodos, bactérias são submetidas a testes morfológicos e bioquímicos, de sensibilidade a fagos e a bacteriocinas, perfis imunológicos (gerando **sorotipos**) e perfis de susceptibilidade a agentes antimicrobianos. Esses testes são muito úteis até certo ponto, pois conseguem determinar gênero e espécie. Entretanto, não são aplicáveis para discriminar subespécie (estirpes e variantes). Como estão ligados ao fenótipo, e não ao genótipo, os resultados obtidos por esses métodos dependem fortemente das condições ambientais e são aplicáveis na tipagem de apenas algumas espécies.

O genótipo não é alterado em função das alterações ambientais e fisiológicas, por isso são muito usados no estudo da diversidade microbológica, produzindo *genotipos* para diferenciar de *sorotipos* diferentes. Organismos geneticamente relacionados partilham características genéticas, bioquímicas e, muitas vezes, fatores de virulência. Para a microbiologia clínica e na epidemiologia, a genotipagem torna-se crítica, pois é possível rastrear a origem e a disseminação de uma estirpe que se tornou causadora de um surto infeccioso episódico, ou seja, é possível saber onde se iniciou uma contaminação.

## 1.2 Segunda aplicação da técnica: sondas moleculares

A fita simples de ácido nucleico é capaz de se ligar com alta seletividade com uma segunda fita complementar. A desnaturação do DNA (separação das duas fitas complementares) ocorre em altas temperaturas e volta a se renaturar quando a temperatura cai. Esses fenômenos opostos, de desnaturação e renaturação ocorrem pela propriedade de hibridização do DNA. Fitas duplas

híbridas podem se formar com DNA-DNA, DNA-RNA (os *primers* adicionados durante a replicação) e RNA-RNA. O processo de hibridização é a restauração das pontes de hidrogênio entre os pares de base. A hibridização será mais estável quanto maior for a complementaridade de bases entre as duas fitas. Uma fita simples de DNA de pouco mais de 17 a 20 nucleotídeos poderia ser usada para localizar uma determinada sequência, por exemplo, num gel de agarose, após a corrida. Evidentemente, ela necessita ser marcada com uma substância radioativa ou um fluoróforo para que seja possível sua localização. Essa pequena fita simples, marcada, é chamada de sonda molecular, ou *probe* (em inglês).

Um pequeno problema ocorre nesse ponto: como pode a sonda de nucleotídeos hibridizar, se os pedaços de DNA estão imersos e inacessíveis no polímero do gel de eletroforese? A resposta é a técnica é conhecida como **Southern Blot**. A palavra *blot* tem o significado de *transferência*, por isso, significa transferência de *Southern*, conforme podemos ver no passo a passo abaixo:

1. Endonucleases de restrição são usadas para o corte de DNA de alto peso molecular em fragmentos menores.
2. Os fragmentos passam por eletroforese em gel de agarose, onde são separados e formam bandas.
3. Uma membrana de nitrocelulose é colocada sobre o gel ou abaixo, dependendo da direção da transferência.
4. Através de eletroporação (atração das moléculas de DNA por corrente elétrica) ou por ação capilar (água ou solução alcalina), o DNA é transferido do gel para a membrana. Note que o padrão eletroforético de bandas não é perdido no processo.
5. A membrana é seca em vácuo ou exposta à luz ultravioleta para fixar o DNA à membrana.
6. A membrana é exposta à sonda, que, então, se hibridiza ao DNA. Para evitar ligações não específicas, utilizam-se previamente reagentes bloqueadores.
7. Após a exposição, a membrana é lavada e procede-se à revelação, dependendo do tipo de marcação que a sonda carrega. Faz-se a autoradiografia (no caso de isótopos), revelação (fluoróforos ou substâncias cromogênicas).
8. O resultado é uma membrana contendo uma ou mais bandas marcadas, as quais correspondem à sequência procurada. Pode-se voltar ao gel, recortar a banda correspondente e proceder a outras técnicas.

**Estringência:** conjunto de condições que interfere na associação das fitas de ácidos nucleicos. Alguns fatores, como as concentrações de sais e da sonda e a temperatura de hibridização, influenciam significativamente nas condições de estringência. A alta estabilidade térmica encontrada nos híbridos RNA-RNA em relação aos híbridos DNA-RNA e DNA-DNA favorece condições de alta estringência, e isso aumenta a especificidade da reação, ou seja, a probabilidade de a sonda hibridizar apenas com o RNA-alvo.

**Epítipo:** definido como parte de uma molécula, especialmente relevos de superfície tridimensional, que é reconhecido, encaixando-se com precisão pelas células do sistema imunológico, como os anticorpos, as células B ou as células T. Estes não precisam ser necessariamente exógenos, mas do próprio corpo. A parte de um anticorpo que reconhece o epítipo se chama paratopo.

**Sorotipos:** a definição formal mostra que são microorganismos diferentes no tipo de antígenos de superfície, ou o tipo de microorganismo, determinado pela sua componente de antígenos baseado em uma de várias reações diferentes antígeno-anticorpo, ou uma subdivisão taxonômica. Por exemplo: existem quatro sorotipos de vírus da dengue, e cada um desses sorotipos é associado a uma variedade de manifestações clínicas.

**Variantes biológicas:**

são características mais complexas ou mais simples, que podem definir ou compor a estrutura de um organismo, definindo-o como uma nova variação em relação à outra. Portanto, os organismos mutantes apresentam uma variação (definida pelo observador) em relação a um organismo ancestral ou a seu tipo selvagem ou *wild type*. As diferenças biológicas entre os seres humanos, variações maiores ou menores entre genes, pode apresentar fenótipos que dão às pessoas as cores de cabelo diferentes, tornam os indivíduos mais propensos a certas doenças e determinar como as pessoas reagem a medicamentos, na maioria dos casos, como resultado de fatores hereditários e influência dos recursos naturais e também dos ambientes sociais. Há grande diversidade genética dentro de todas as populações humanas. Infelizmente, muitas vezes explicações biológicas podem ser usadas para sustentar relações racistas entre "raça", a biologia e a doença. Um outro exemplo na área molecular: citocinas (proteínas imunológicas envolvidas em doenças inflamatórias como a IL-6 e TNF-alfa) aparecem com grande variação, mesmo em indivíduos saudáveis, mas também são afetadas por parâmetros como idade, sexo, tabagismo etc.

9. A temperatura de hibridização é muito importante, pois é muito fácil obter falsos positivos. Aumentando a temperatura, ou a **estringência**, somente as hélices hibridizadas mais estáveis permanecerão.
10. Apesar de serem rotineiros em muitos laboratórios, os *blots* são geralmente usados apenas para complementar ou confirmar resultados obtidos em outras técnicas. A ocorrência de falsos positivos, a difícil reprodutibilidade dos perfis, os custos e as dificuldade no manuseio, o que demanda mão de obra experiente, contribuem para isso. Apesar desses fatores, a hibridização com sondas, quando bem executada, é definitiva em diagnósticos de doenças genéticas, devido à alta seletividade das sondas.

O método de *Southern blot* é referência ao pesquisador Edwin Southern, biólogo britânico, que o desenvolveu em 1975. Como homenagem e brincadeira com o biólogo, apareceram posteriormente protocolos derivativos dessa técnica, como o *Western blot* (usado para localizar proteínas específicas no gel de eletroforese, com anticorpos), *Northern blot* (usado para RNA), *Eastern blot* (para modificações pós-traducionais), *Southwestern blot* (para proteínas que se ligam ao DNA). Para o *Eastern blot*, múltiplas técnicas reivindicam esse nome. São usados para lipídeos, proteínas, glicoconjugados e **epítomos** de carboidratos. Coletivamente, são modificações pós-traducionais. As proteínas transferidas são analisadas para as modificações pós-traducionais, pela detecção de lipídeos, carboidratos, fosforilação ou qualquer outra modificação proteica. Todos eles usam os mesmos princípios similares, utilizam uma membrana de nitrocelulose (nylon ou outro material) para receber a amostra vinda de uma eletroforese, sem, contudo, perder o padrão eletroforético. O próximo passo é a hibridização com uma sonda. Para proteínas, essa sonda é, na verdade, um anticorpo, produzido especificamente para localizar o antígeno. Em outros casos, pode ser utilizado um reagente, como cromogênios, fluoróforos, etc.

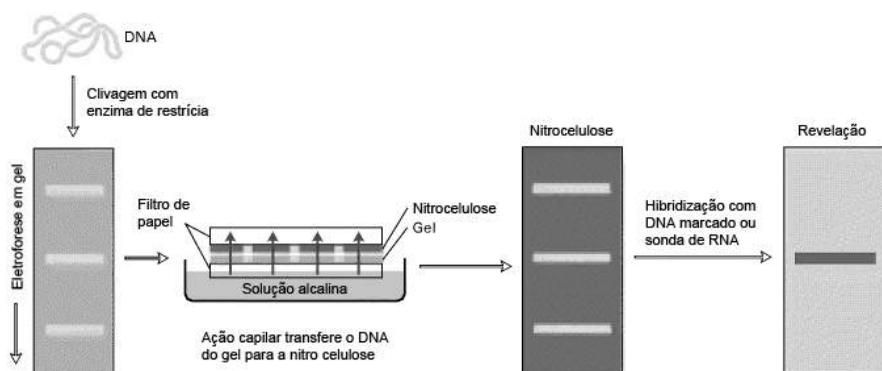


Figura 3 - Técnica de transferência de **Southern** ou **Southern blot**. A mistura de fragmentos de restrição é hibridizada (após a transferência do gel para membrana de hibridização) com uma sonda específica e lida como positivo ou negativo para a hibridização.

Fonte: Lodish et al. (1999).t

## Plasmídios: vilões e heróis

Os plasmídeos bacterianos são elementos extracromossômicos bem conhecidos (ver Unidade 3). A conjugação bacteriana envolve a transferência de plasmídeos entre duas bactérias. São moléculas muito pequenas de dupla fita de DNA, circulares, contendo poucos genes. Possuem origem de replicação e, com isso são replicados junto com a bactéria, porém, independentemente do cromossomo bacteriano. Genes que são carregados por plasmídios geralmente estão envolvidos com a resistência a antibióticos. Espalhando rapidamente essa resistência, quando estão sob pressão seletiva para esse fator ambiental, os plasmídeos são chamados de “vilões nosocomiais”. Obviamente, o ser humano é que é culpado da resistência bacteriana, especialmente pelo uso indiscriminado e incorreto dos antibióticos.

Graças a essas características, o plasmídeo foi eleito como ferramenta biogenética pelos biólogos moleculares. Os plasmídeos, na Engenharia Genética, são chamados de vetores plasmidiais, e comumente são usados para multiplicar ou expressar genes particulares. Para isso, estão disponíveis comercialmente, específicos para esses usos. Os sítios de restrição são características do vetor, disponibilizados nos plasmídeos pelos próprios fabricantes. Inicialmente o plasmídeo é cortado com uma enzima de restrição (em um sítio de restrição equivalente). O gene a ser inserido é também cortado com a mesma enzima de restrição, produzindo em ambos extremidades coesivas e complementares, que permitem que se unam e formem uma molécula só, ainda circular. O gene a ser replicado pode ser qualquer sequência, de qualquer ser vivo. O plasmídeo recombinante é, então, inserido e naturalmente absorvido pela bactéria, ocorrendo transformação bacteriana. O processo de transformar bactérias através de um vetor é chamado *clonagem molecular*. Cuidado, porque esse processo não tem nenhuma ligação com a clonagem gênica (como o caso da ovelha Dolly) e também não é relacionado ao processo natural de conjugação que ocorre em bactérias. Não existe interesse em transformar as bactérias em si, buscando, por exemplo, transformar bactérias patogênicas em não patogênicas. Não pense que as bactérias e os plasmíde-

**YAC, BAC e MAC** - *Yeast artificial chromosome* (YACs) ou cromossomo artificial de levedura. Cada YAC inclui os três elementos essenciais de um cromossomo: telômeros, centrômeros e origem de replicação. Os YACs podem ser transferidos para células de camundongo. BACs e MACs são os equivalentes para cromossomos artificiais de bactéria e de mamíferos, respectivamente. Veja as Figuras 4, 5 e 6.

**Fago  $\lambda$**  (lambda) – bacteriófago lambda, o qual infecta *E. coli*. No genoma do fago, genes não essenciais são retirados. Ao serem retirados, deixam pontas curtas não coesivas que servem para receber o novo inserto de DNA. O cromossomo lambda possui pontas curtas unifilamentares chamadas sítios *cos* que são necessárias para embalar o DNA completo na cabeça do fago. Assim, são empacotados em capas proteicas e adicionados à *E. coli*. Os fagos injetam seu DNA na célula hospedeira em que poderá ser replicado. Interessante é que o DNA não será empacotado em uma capa de fago lambda a menos que tenha o tamanho adequado (40 a 50 Kb). Apenas os fragmentos de tamanho certo contendo os genes essenciais para o empacotamento serão embalados nas capas dos fagos e, assim, compõem um sistema “natural” de seleção dos recombinantes.

os são usados como reagentes biológicos com o objetivo de replicar (ou expressar) o gene de interesse, sendo possível mantê-lo indefinidamente, disponível para os trabalhos.

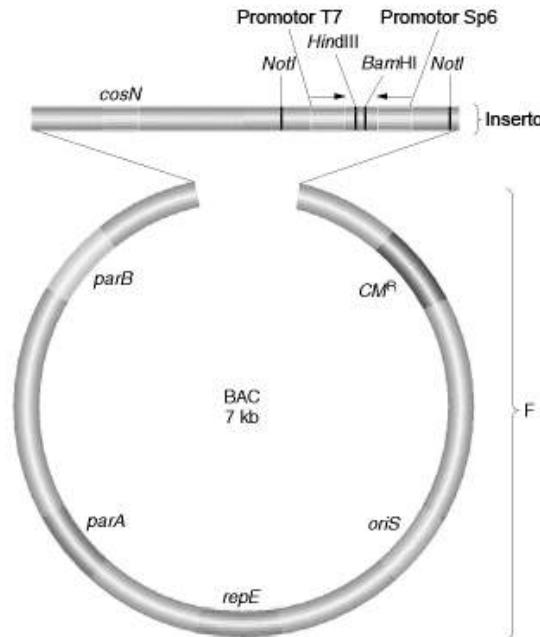


Figura 4 - Estrutura do cromossomo artificial de bactérias – BAC – apresentando os genes e os locais de inserção do gene de interesse (nos sítios de restrição). **CMR** - gene marcador para a resistência ao antibiótico cloranfenicol. *oriS*, *repE*, *parA* e *parB* são genes F para replicação e regulam o número de cópias. **cosN** é o sítio *cos* do fago lambda (veja o glossário). **HindIII** e **BamHI** são sítios de restrição nos quais é possível inserir o gene de interesse. Existem dois promotores: T7 e Sp6. Dois sítios **NotI** são usados para retirar posteriormente os fragmentos.

Fonte: Griffiths et al. (1996).

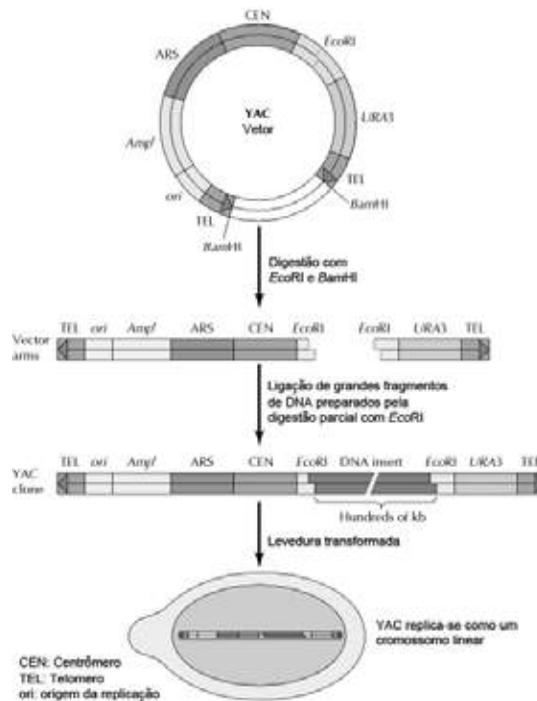


Figura 5 – YAC: **Yeast Artificial Chromosome** ou cromossomo artificial de levedura. O vetor contém uma origem de replicação (**ori**) e um gene que confere resistência à ampicilina (**Ampr**), podendo ser propagado em também em bactérias *E. coli*. Em adição, esse vetor contém uma origem de replicação de levedura (**ARS**), um centrômero (**CEN**), um gene marcador que permite o crescimento em ausência de uracila (**URA3**) e dois telômeros (**TEL**). O vetor é digerido com **EcoRI** e **BamHI**, produzindo dois braços que terminam com telômeros e sítios **EcoRI**. Os fragmentos maiores de DNA preparados com a digestão com **EcoRI** são ligados aos braços do vetor. As células de levedura são, então, transformadas com os vetores recombinantes, os quais irão replicá-los da mesma forma que os cromossomos lineares.

Fonte: Cooper e Hausman (2009).

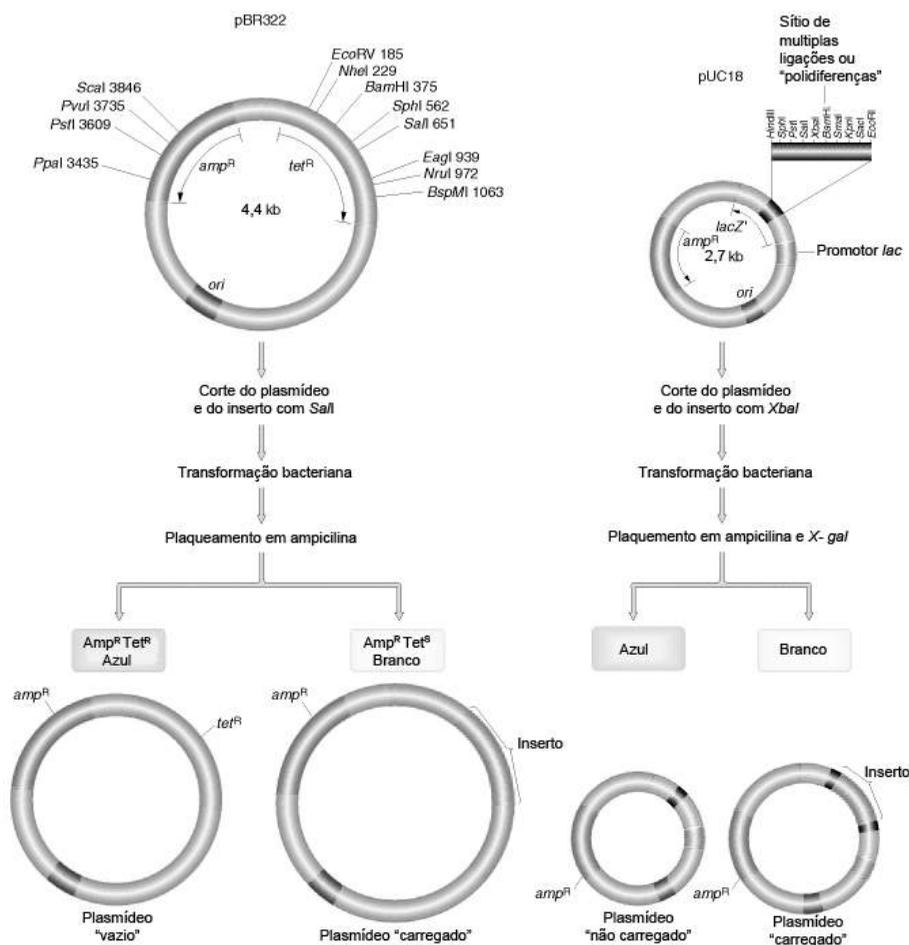


Figura 6 - Dois plasmídeos muito usados na clonagem molecular, apresentando a sua estrutura geral e a sítios de restrição (pBR322 e pUC19). O sucesso da inserção de genes em pBR322 é detectável pela inativação de um gene de resistência à tetraciclina (*tet<sup>R</sup>*). A inserção no plasmídeo pUC18 é detectado pela inativação da  $\beta$ -galactosidase indicado pela função do gene *LacZ*, a qual resulta na inabilidade de converter o substrato artificial X-Gal (veja o glossário) em um corante azul.

Fonte: Griffiths et al. (1996).

Para produzir a clonagem molecular, apresentamos a seguir os passos básicos.

1. Isolamento da sequência a ser inserida pelo corte com enzima de restrição.
2. Corte do plasmídeo com a mesma enzima de restrição.
3. Reação de ligação entre o plasmídeo e a sequência: produção de um plasmídeo recombinante.
4. A DNA ligase é usada para fechar as ligações fosfodiéster.
5. Transformação da bactéria com o plasmídeo recombinante.

O gene *lacZ*, plasmidial, é usado como um gene marcador, ou repórter, em combinação com o meio de crescimento de bactérias (usadas na clonagem gênica) contendo o açúcar X-gal, uma galactose. O gene *lacZ* codifica para a enzima  $\beta$ -galactosidase, que quebra o açúcar X-gal, produzindo um pigmento azul. Assim, as bactérias, com um gene *lacZ* íntegro, produzem colônias azuis. A estratégia é que o gene encontra-se no plasmídeo usado para inserir o gene de interesse. Este gene deve ser clonado dentro daquele gene, inativando-o. Assim, as colônias recombinantes são brancas, e não azuis, sendo assim, facilmente reconhecidas e postas para repicagem. Estas são os chamados "clones bacterianos", contendo os genes de interesse, produzindo bibliotecas genômicas e outras aplicações.

6. Recuperação das bactérias recombinantes.
7. As colônias de bactérias recombinantes são, então, armazenadas e podem ser cultivadas, a qualquer momento, em meio líquido, gerando milhões de cópias do gene inserido.
8. O plasmídeo recombinante pode então ser retirado por centrifugação do meio onde cresceram as bactérias, gerando muitas cópias que podem ser usadas em sequenciamento, produção de transgênicos, vacinas gênicas, transfectados para células eucarióticas etc.

Os **vetores de clonagem** têm por objetivo a clonagem do gene, sendo, então, possível armazenar o gene. Em geral, precisam de alguns requerimentos essenciais para funcionar como tal: origem de replicação (para que possam ser replicados pela bactéria); sítios únicos de restrição (onde será cortado para a inserção do gene); genes de seleção (usualmente de resistência a antibióticos, para proporcionar um meio de se recuperar os recombinantes).

Os **vetores de expressão** têm por objetivo que a sequência inserida seja expressa em proteína. Para tanto, é necessário que além dos requerimentos mencionados anteriormente, também estejam presentes sequências regulatória e promotor, *upstream* ao gene.

## 1. Primeira aplicação da técnica: bibliotecas genômicas e bibliotecas de cDNA

**Bancos de sequências expressas** são conhecidos como **ESTs** ou *Expressed Sequence Tags* (ou rótulos de sequências expressas). Consiste na identificação (rótulo) de proteínas expressas em um tecido, por exemplo, no hepatócito. Esse tipo de abordagem pode ser muito útil, através de comparação de tecidos. Comparando-se as sequências expressas, por exemplo, de um tecido normal com tecido tumoral, é possível encontrar genes que são expressos diferencialmente (tanto quantitativa quanto qualitativamente) e aí traçar estratégias para uma terapêutica.

A clivagem de um genoma eucariótico inteiro, com uma enzima de restrição, pode gerar literalmente milhões de fragmentos. Cada fragmento é inserido em um plasmídeo e usado para transformar uma bactéria. As múltiplas colônias armazenadas formam um conjunto global chamado de “biblioteca genômica”. Cada colônia recombinante da biblioteca é chamada “clone”. Cada fragmento é sequenciado, resultando no genoma completo do organismo, disponível para análise. Existem inúmeras bibliotecas genômicas de centenas de organismos, facilitando muito a vida dos pesquisadores, que têm disponível, na rede, em portais de dados, as sequências do genoma.

Pode-se produzir uma biblioteca a partir de outros tipos de fragmentos. Uma das melhores contribuições da era pós-genômica é buscar a expressão do gene de forma massiva, na forma de um transcriptoma (coleção de transcritos de mRNAs) ou biblioteca de cDNA (DNA complementar).

9. Células de um tecido são usadas para a extração do mRNA. Para garantir que somente RNAs maduros sejam utilizados, lança-se mão de colunas contendo caudas expostas de oligo-dTs. O RNAm maduro eucariótico possui cauda 5'poli A, sendo então capturado pela coluna.
10. O RNAm é eluído da coluna, estando pronto para ser utilizado.

11. Inicialmente um primer de oligo dT liga-se à cauda poli A, possibilitando que uma polimerase adicione nucleotídeos. No caso essa polimerase é a transcriptase reversa, capaz de usar um molde de RNAm para criar uma fita híbrida RNAm/DNA.
12. O RNAm é degradado com uma solução alcalina; deixa-se uma parte final que se autocomplementa, resultando num terminal em alça.
13. A DNA polimerase finaliza produzindo uma fita nova DNA/DNA.
14. A alça é degradada com uma nuclease.
15. O resultado é uma cópia de cDNA complementar ao mRNA inicial.
16. Quando todos os fragmentos estiverem complementados, entra em ação a clonagem, produzindo uma biblioteca com os fragmentos.
17. A biblioteca pode ser sequenciada, apresentando uma coleção de sequências expressas do DNA naquele tecido.

## 2. Segunda aplicação: transgenia

A introdução de DNA exógeno via vetor de clonagem ou expressão, numa célula eucariótica animal ou vegetal é o primeiro passo para a produção de animais e plantas transgênicos (Figura 7). O organismo transgênico é um organismo permanentemente alterado geneticamente através de tecnologia do DNA recombinante. Entretanto, não é tão simples assim obter a inserção definitiva de um gene em um cromossomo da célula hospedeira. É desejável que esse gene esteja presente nas células filhas. Portanto, o início da produção de um transgênico envolve a inserção do gene de interesse no vetor, posteriormente, a transfecção do gene na célula, a seleção da célula recombinante através de genes de seleção ou genes marcadores ou repórter. O gene inserido deve, então, ser transcrito, traduzido pelos ribossomos e por outras partículas, produzindo uma proteína perfeita, no tecido certo e no tempo certo do desenvolvimento do organismo para que possa ser útil ao metabolismo. A célula transfectada deve produzir um organismo completo, transgênico e, sabemos, quase sempre isso não é possível.

**Eletroporação:** a célula passa por um pulso breve de alta voltagem, o que a torna transitoriamente mais permeável à moléculas de DNA no meio extracelular, aumentando a frequência de células competentes.

**Microinjeção de DNA:** uso de uma agulha muito fina, com a qual um hábil manipulador injeta DNA nas células uma a uma. O método é muito eficaz, porém trabalhoso e rende poucas células competentes por vez.

**Vetores virais:** vírus introduzem seu material genético nas células. Um retrovírus introduz seu RNA nas células hospedeiras, onde a sua própria transcriptase reversa converte seu RNA em DNA e então este é integrado aos cromossomos do hospedeiro. Na técnica, genes virais são trocados por DNA de interesse. O método possui alta infetividade, e é mais trabalhoso.

**Terapia gênica:** terapia pela qual se busca a expressão de genes induzida em seres humanos ou em animais de experimentação. Busca-se a introdução eficiente de DNA num número suficiente de células somáticas (é proibido o uso de células germinativas), fazendo com que estas expressem uma proteína. Esta pode ser uma proteína em déficit no organismo ou que possa atacar células cancerosas.

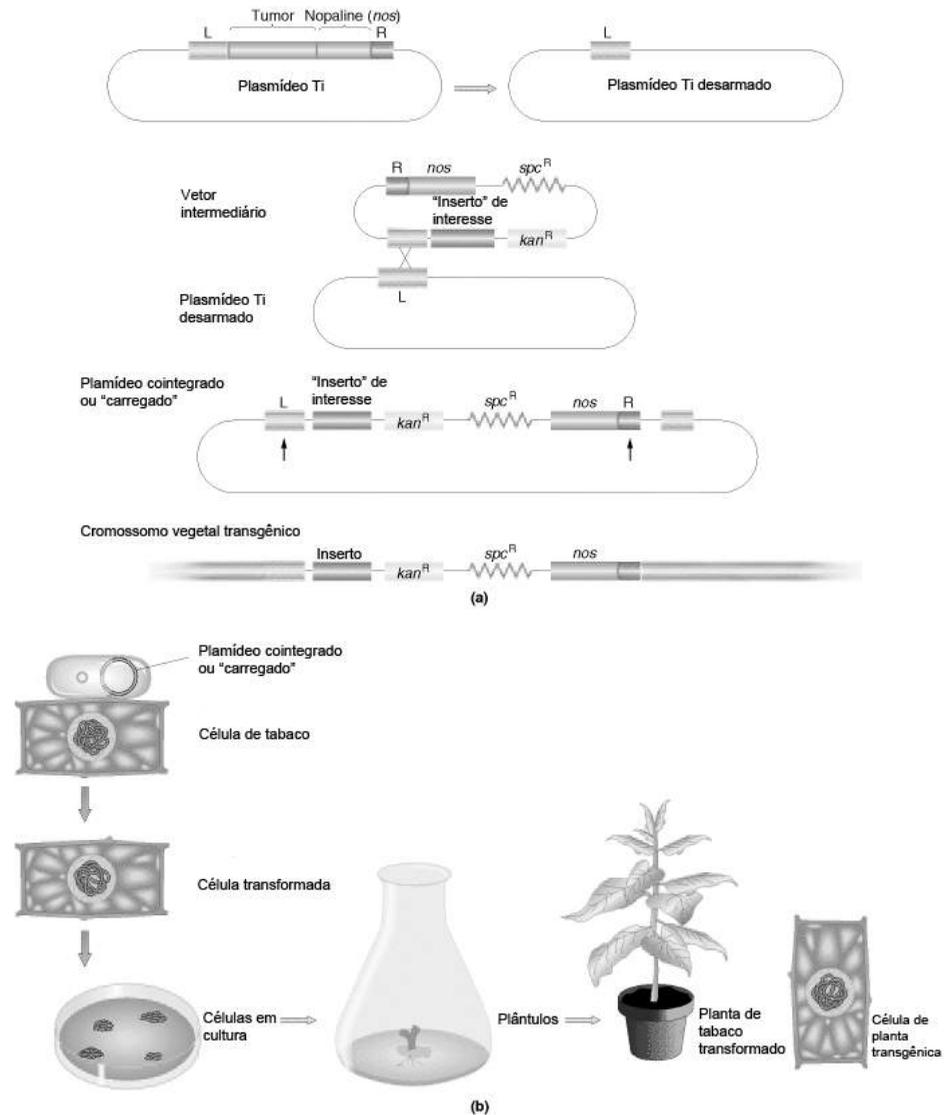


Figura 7 - Modo de produção de plantas transgênicas. (a) um vetor intermediário é usado para clonar o segmento de interesse. O vetor intermediário é recombinado com um plasmídeo Ti desarmado, para gerar uma estrutura coiterada contendo o inserto de interesse e um gene de seleção marcador de recombinantes que, nesse caso, é um gene de resistência à kanamicina, entre as bordas do T-DNA. (b) protocolo de geração da planta transgênica através da transformação com T-DNA.

Fonte: Griffiths et al. (1996).

Plantas são mais hábeis para transgenia, pois de qualquer célula vegetal, em tese, podemos produzir um organismo completo (chamada *totipotência vegetal*). O mesmo benefício não se pode fazer o mesmo com célula animal. As células animais são *pluripotentes*, de uma célula embrionária. Em tese, seria possível derivar qualquer tecido (veja a polêmica do uso de células-

-tronco ou embrionárias). Para contornar esse “problema” várias técnicas são usadas para produzir animais-modelo transgênicos, especialmente ratos e camundongos com transgenia para várias doenças e que, assim, tornam-se disponíveis para testes diagnóstico e novas terapias.

A utilização de células animais para a inserção de material genético exógeno envolve centenas de técnicas e protocolos experimentais. Esses experimentos também permitem compreender como funciona o genoma animal e gera também informações para a clonagem animal. Métodos mais usuais para inserção de genes nos cromossomos animais envolvem, por exemplo, **a eletroporação, microinjeção de DNA**, além de **vetores virais**. Esses vetores também vêm sendo estudados para uso em **terapia gênica**.

Os animais transgênicos são definidos de várias formas. Para a *Federação Européia das Associações de Ciência em Animais de Laboratório*, transgênico é “um animal que possui seu genoma modificado artificialmente pelo homem, quer por meio da introdução, quer da alteração ou da inativação de um gene - uma sequência definida de DNA. Esse processo deve culminar na alteração da informação genética contida em todas as células desse animal, até mesmo nas células germinativas (óvulos e espermatozoides), fazendo com que essa modificação seja transmitida aos descendentes”. O uso de ratos e camundongos transgênicos, modelos para várias patologias vêm sendo muito expandido nos laboratórios experimentais. Um dos métodos mais utilizados é a microinjeção de DNA no núcleo de ovos fertilizados, esperando-se que ocorra uma integração cromossômica adequada. Os ovos são, então, introduzidos na fêmea onde se induz uma pré-gravidez, com uso de machos estéreis. Faz-se, então, testes na progênie resultante para saber quais animais estão efetivamente expressando o gene. Os animais então são cruzados entre si para obtenção de homozigose para o gene inserido, mantendo-se então essa nova variedade transgênica pelo mesmo princípio do endocruzamento. Os genes inseridos podem ser induzidos por algum componente da dieta, injeção de um fármaco ou xenobiótico.

Nas plantas, um aliado essencial tem se mostrado imprescindível na transformação das células vegetais. Uma bactéria do solo, *Agrobacterium tumefaciens*, invade células vegetais em locais feridos, produzindo um tumor, conhecido popularmente como “galha”. O plasmídeo típico de *A. tumefaciens* é um grande plasmídeo chamado Ti. Quando a bactéria contacta uma parte da célula vegetal, um segmento chamado DNA T é transferido do plasmídeo Ti para o vegetal. O DNA T, então, fica incorporado ao cromossomo vegetal. Esse é um raro exemplo de engenharia genética natural.

## Reação em cadeia da polimerase: a revolução dentro da revolução

**Primers degenerados:** são misturas de *primers* semelhantes, mas não iguais. O desenho desses *primers* advém de uma proteína conhecida, e não de um gene conhecido. Como vários códons podem codificar um aminoácido, muitas vezes é difícil deduzir qual dos códons possíveis é utilizado naquele gene em particular. Um exemplo: se o aminoácido é isoleucina, podemos ter um 5'-ATH – 3', onde A é adenina, T é timina e H é adenina, timina ou citosina. Veja Quadro 13.1.

**Paleogenética:** relação entre paleontologia e genética para estudo de múmias, fósseis, roupas, materiais preservados antigos, ovos fossilizados, coprólitos etc. Utiliza sequências marcadoras nucleares e mitocondriais e *primers* degenerados para estudos genéticos.

Poucas vezes a tecnologia foi tão importante para o avanço da ciência como o que aconteceu em 1983, com o aparecimento da reação em cadeia da polimerase. Uma técnica única, de extrema sensibilidade, facilidade de manuseio, simplicidade, rapidez e custo relativamente baixo, com aplicações monumentais (Figura 8).

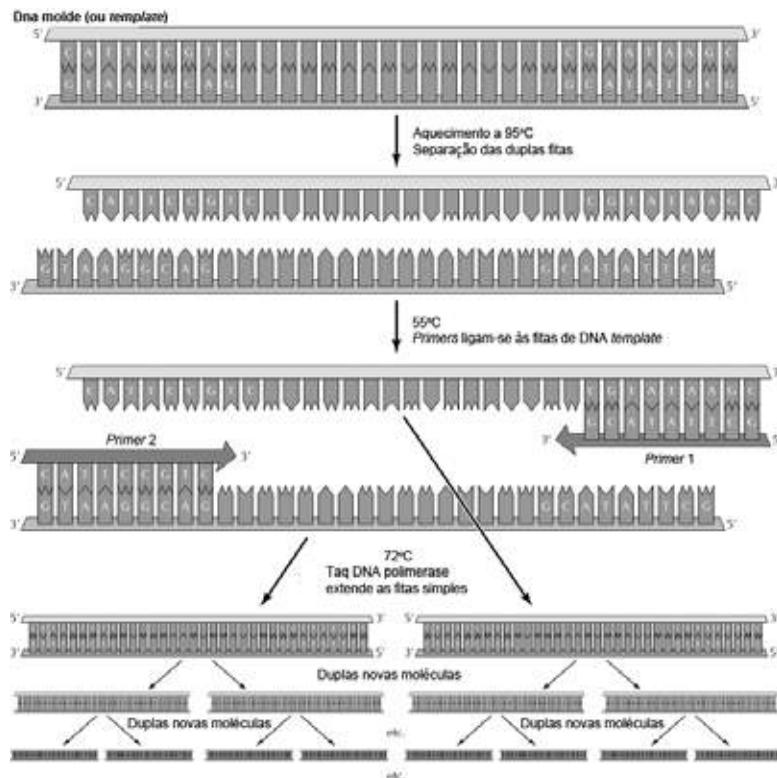


Figura 8 - Amplificação por PCR - A região do DNA a ser amplificada é flanqueada por duas sequências usadas para o início da síntese. A dupla fita é separada nas duas fitas simples pelo calor. Os *primers* se ligam (oligonucleotídeos de 15 a 20 bases) de cada lado da sequência alvo. A DNA polimerase de *Thermus aquaticus* (Taq polimerase) é usada para sintetizar novo DNA a partir dos *primers*, formando duas novas moléculas. O processo pode ser repetido inúmeras vezes, resultando na replicação controlada da sequência alvo.

Fonte: Cooper e Hausman (2009).

Resumo do protocolo:

- Abertura da fita de DNA que servirá de molde por desnaturação térmica (etapa com duração entre 30s e 1min à temperatura de 92 - 96°C);
- Pareamento de oligonucleotídeos sintéticos, *primers* que funcionam como os iniciadores da reação de polimerização, a cada uma das fitas do DNA molde, à região complementar da fita que sofrerá a duplicação (duração de 30s a 1min à temperatura entre 58 e 65°C);
- Polimerização, através de uma enzima polimerase, das novas fitas de DNA a partir de cada um dos iniciadores, utilizando cada um dos quatro dNTP como substrato da reação de polimerização (duração entre 45 segundos e 1 minuto, a 72°C). Cada ciclo é repetido em torno de 60 vezes e promove a amplificação da região alvo determinada conforme afinidade das sequências iniciadoras (*primers*). Assim, os *primers* reconhecem, por complementaridade, o local de início do local a ser amplificado, efetua a ligação e sinaliza para a polimerase o início da sequência a ser replicada.

O grande problema inicial foi a desnaturação seguida da enzima polimerase, que, obtida de *E. coli*, não suportava a temperatura para abertura das fitas de ácidos nucleicos. Assim, a cada ciclo, uma nova quantidade de enzima deveria ser adicionada.

Em 1988, contudo, Randall Saiki e colaboradores substituíram a polimerase de *E. coli* pela já conhecida **Taq polimerase**, polimerase da bactéria *Thermus aquaticus*, que vive em águas quentes de gêiseres e vulcões submersos e possui um ponto de crescimento ótimo entre 70 e 75°C.

Então, qual o contexto de utilização da clonagem molecular e da PCR? Existe inicialmente um contexto histórico. O DNA recombinante, no sentido de criação de novas sequências de DNA, surgiu com a enzima de restrição. Trata-se de obter a primeira resposta técnica, inserir o gene em vetores, fazer a alteração de um genoma bacteriano. A clonagem molecular permite manter sequências de DNA, de diferentes tamanhos, por tempo indeterminado, intactas, podendo ser recuperadas e multiplicadas a qualquer momento, bastando fazer o repique dessas bactérias, ou leveduras, recombinantes, após anos ou décadas, guardadas em freezer (Veja também a chamada estratégia **shotgun** de genoma) A clonagem molecular é adequada à produção de projeto genoma. A PCR, por sua vez, amplifica o gene apenas. O produto da PCR é a sequência amplificada, utilizável em dezenas de projetos, mas só pode ser guardada por alguns dias, com perda de material. A fundamental vantagem do PCR é que não necessita de tantos passos quanto a clonagem, o único segmento que será amplificado é aquele entre dois *primers*, o que acarreta especificidade. Em tese, basta uma cópia do *template* (molde de DNA) para

**DNA barcode:** ou “código de barras” - metáfora para código identificador baseado em DNA, como um rótulo que seria dado a todos os organismos baseado numa sequência de cerca de 648 nucleotídeos na maioria dos grupos. O DNA *barcode* seria diferente entre indivíduos da mesma espécie, porém num grau menor. O objetivo seria o estudo e a preservação da biodiversidade, devido a vantagens como:

facilidade na identificação da vida, pois independe de estágios vitais, distinção de espécies parecidas ou quase indistinguíveis, democratização do acesso, entre outros. Veja o projeto “*Consortium for the barcode of life (CBOL)*” - <http://www.barcoding.si.edu/>

**Shotgun:** estratégia usada para cortar um genoma em milhares de pedaços usando uma, duas ou mais enzimas de restrição, gerando coleções de pedaços que se justapõem. Assim, após o sequenciamento, são montadas em sequências consenso, cobrindo todo o genoma com pedaços justapostos. As sequências de leitura superpostas são chamadas de sequências contigs (contíguas, ou que se tocam). Espaços não sequenciados de genoma, que porventura ocorram após a junção dos *contigs* (feito através de bioinformática) são então completados posteriormente, usando *primers* específicos ou mesmo outras abordagens.

**PCR Hot Start:** técnica de PCR utilizada para diminuir a possibilidade de formar produtos inespecíficos, o que pode ocorrer em uma PCR comum. Certas combinações de *primers* e DNA-molde. Na reação tradicional produtos inespecíficos podem ocorrer devido ao anelamento inespecífico dos *primers* (favorecido por baixa temperatura) em sequências de fita simples presentes na reação como DNA parcialmente degradado. Para evitar isso, a técnica busca evitar o contato da polimerase na primeira etapa de desnaturação

**PCR Touch-down:** nesta técnica, o termociclador é programado para diminuir de forma sequencial a temperatura de anelamento dos *primers*. Nos ciclos iniciais da PCR (nos primeiros dez ciclos) até determinada temperatura, a qual é supostamente a ideal, é mantida nos ciclos restantes. A temperatura de anelamento mais alta (maior estrincência) evita a formação de produtos inespecíficos, permitindo que o produto esperado seja o predominante.

que a amplificação ocorra, sendo portanto, é muito sensível. Especificidade, sensibilidade e rapidez no trabalho traduzem-se em praticidade na utilização diária e em aplicações variadas. Um crime pode ser investigado se houver uma molécula de evidência, como poucas células foliculares de um único fio de cabelo do suspeito ou células sob as unhas da vítima. Um conceito relativamente difícil de compreender é que a PCR em si não apresenta a identificação do suspeito. Ninguém tem seu nome tatuado na molécula. É preciso um conhecimento *a priori* para a produção e análise do PCR.

No caso anterior, a eletroforese do *fingerprint* (impressão digital) de DNA fornece algo como um código de barras (**barcode**). Esse código fornece um perfil que necessita de base comparativa, como a pré-existência de um banco de perfis de DNA de pessoas condenadas por crimes sexuais, por exemplo. Na falta ou na deficiência do mesmo, suspeitos do crime (ou num outro caso, supostos parentes de uma vítima não identificada) devem contribuir com seu DNA para fornecer essa base de comparação. Por isso, o trabalho molecular nunca exclui a investigação policial. Caso se queira produzir uma sequência específica, por exemplo, o gene da insulina, torna-se essencial o conhecimento *a priori* do gene da insulina, facilmente obtível através dos bancos de dados disponíveis na rede mundial de computadores ou mesmo na própria literatura. No PCR específico, temos que saber que sequência específica queremos amplificar e desenhar os *primers* específicos para ela. No PCR *inespecífico*, *primers* ditos **degenerados** (Quadro 13.1) são usados para amplificar qualquer sequência.

Quadro 13.1

Código IUPAC para desenho de primers degenerados					
Símbolo	Descrição	Bases representadas			
A	Adenina	A			
T	Timina		C		
C	Citosina			T	
G	Guanina				G
U	Uracila			U	
W	Weak	A		T	
S	Strong		C	T	
M	Amino	A			
K	Keto			T	G
R	Purina	A			G
Y	Pirimidina		C	T	
B	Menos A		C	T	G
D	Menos C	A		T	G
H	Menos G	A	C		G
V	Menos T	A	C	T	
N	Qualquer nucleotídeo	A	C	T	G

## 1. Aplicação da técnica: *fingerprint* de DNA

Quem nunca ouviu falar no “teste do DNA”? Um recurso intensamente explorado hoje em dia pela mídia, o teste do DNA na verdade é chamado tecnicamente de *fingerprint* (impressão digital) e permite uma enorme gama de aplicações forenses e criminais. Investigações civis também são beneficiadas, como a paternidade, o reconhecimento de herdeiros, o reconhecimento de restos mortais, os estudos de cadáveres mumificados (paleogenética), os estudos históricos e ecológicos.

A técnica do *fingerprint* é baseada em polimorfismo de marcadores moleculares. O marcador molecular para a execução da técnica de *fingerprint* são os VNTRs (*Variable Number of Tandem Repeats* ou número variável de repetições *em tandem*) (Figura 9).

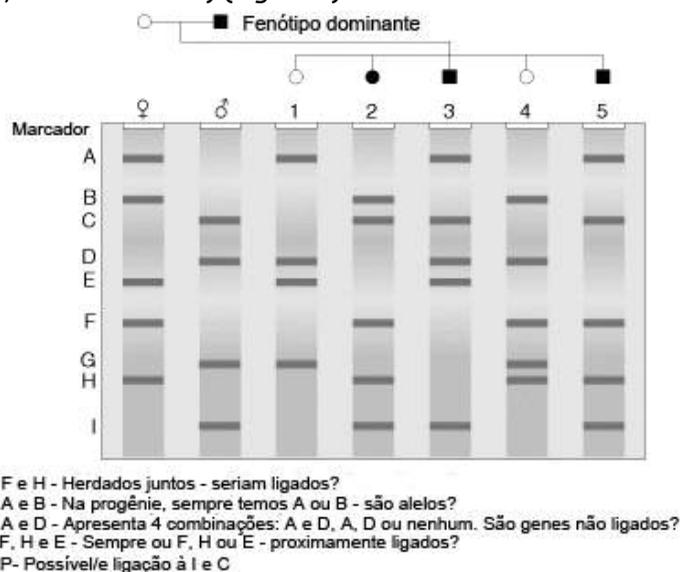


Figura 9 - Utilizando o *fingerprint* de DNA como marcadores moleculares, em mapeamento por eletroforese. Alelos de dominância podem ser localizados por análise de ligação.

Fonte: Griffiths et al. (1999).

Estas VNTRs são sequências curtas, que, como qualquer sequência, são herdáveis. Para identificar estes VNTRs, torna-se possível usar PCR e/ou sondas específicas para estes. O passo a passo para a técnica do *fingerprint* é visto a seguir:

1. O DNA é extraído do doador;
2. O DNA é cortado com uma ou mais enzimas de restrição;
3. Os fragmentos são separados por eletroforese;
4. Faz-se a transferência de *Southern*;
5. Utilizando várias sondas específicas de VNTRs, procede-se à incubação com elas;

**PCR multiplex:** é uma reação em que várias regiões diferentes de DNA são amplificadas ao mesmo tempo e no mesmo tubo devido à utilização simultânea de vários pares de *primers* específicos para vários *loci*. Permite maior rapidez na obtenção do resultado, com economia de reagentes.

**Nested PCR:** nesta técnica, o *template* é produto de amplificação anterior numa segunda amplificação, com *primers* que se anelam internamente ao produto da primeira amplificação.

**RT-PCR:** *Reverse transcriptase* – PCR. Numa etapa anterior, o *template* de DNA é produzido por meio da transcrição reversa aplicada a uma fita de RNA<sub>m</sub> obtida de um tecido (veja o item Bibliotecas – Bibliotecas de cDNA). Esse trabalho pode ser feito num termociclador comum ou no real time, que é um termociclador quantitativo, mais adequado a esse tipo de trabalho, pois o interesse é quantificar a expressão de genes.

### Extração do DNA

O protocolo básico de extração de DNA exige que inicialmente as células sejam rompidas para que os constituintes celulares sejam liberados. Para tecidos vegetais utilizam-se brotos ou folhas jovens e, para tecidos animais, sangue, fígado ou tecidos macios. Pode-se congelar o tecido previamente e depois macerá-lo em almofariz e pilão. Em pequena escala, pode-se usar um bastão de vidro e um tubo de microcentrifugação, sem o uso de nitrogênio líquido. Utilizam-se detergentes como o SDS (dodecil sulfato de sódio) ou CTAB (brometo de cetiltrimetilamônio), para rompimento das membranas. O tampão de extração possui pH ao redor de 8,0, já que as DNases (degradadoras do DNA) possuem atividade ótima ao redor de 7,0. O uso de EDTA (ácido etileno diamino tetracético) no tampão de extração inibe a ação de DNases que têm como cofatores metais, já que o EDTA é um agente quelante. A separação do DNA das proteínas é um passo importante, através de várias extrações com fenol e/ou clorofórmio, os quais são desnaturantes de proteínas, tornando-as solúveis em fase aquosa. Os ácidos nucleicos se encontram nessa fase, então as proteínas ficam floculadas e formam uma fase separada na solução. Os polissacarídeos também devem ser separados dos ácidos nucleicos, pois tornam a solução viscosa, o que dificulta a migração dos ácidos nucleicos em gel de eletroforese. O detergente também é usado para essa função.

6. As sondas hibridizam-se com DNA da membrana;
  7. Faz então a autorradiografia ou faz-se a revelação do reagente usado na marcação da sonda;
  8. Obtém-se então um padrão de bandas, comparável com outros padrões.
- Ou
9. O DNA é extraído do doador;
  10. Uma alíquota é retirada e submetida a PCR com *primers* específicos para os VNTRs;
  11. O produto do PCR passa por eletroforese;
  12. Faz-se revelação comum com brometo de etídeo;
  13. Obtém-se então um padrão de bandas, comparável com outros padrões;
  14. O DNA pode ser sequenciado (Figura 10).

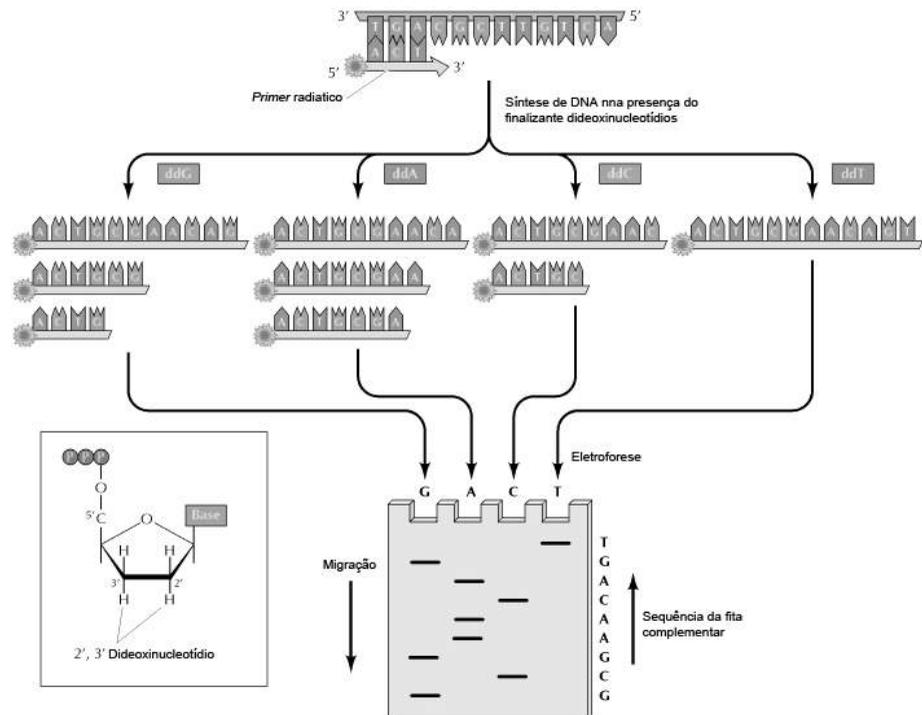
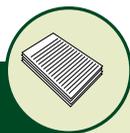


Figura 10 - Sequenciamento de DNA pelo método de Sanger. Dideoxineucleotídeos, os quais são desprovidos de grupos OH nos carbonos 3' e 2', são usados para terminar a síntese de DNA em bases específicas. Essas moléculas são incorporadas normalmente pelas fitas em crescimento. No momento que os dideoxineucleotídeos são incorporados, encerra-se a reação de PCR. Quatro reações separadas são feitas, cada um contendo um dos dideoxineucleotídeos misturados com os desoxineucleotídeos normais. Assim, cada reação produz somente fragmentos finalizados com o seu dideoxi específico. Os produtos são separados por eletroforese, e a leitura, então, pode ser feita, de forma a produzir a sequência de DNA.

Fonte: Cooper e Hausman (2009).

## Textos complementares



### A Lei de Biosegurança - Lei nº 11.105, de 24 de março de 2005.

“Art. 3o Para os efeitos desta Lei, considera-se:

I – organismo: toda entidade biológica capaz de reproduzir ou transferir material genético, inclusive vírus e outras classes que venham a ser conhecidas;

II – ácido desoxirribonucléico - ADN, ácido ribonucléico - ARN: material genético que contém informações determinantes dos caracteres hereditários transmissíveis à descendência;

III – moléculas de ADN/ARN recombinante: as moléculas manipuladas fora das células vivas mediante a modificação de segmentos de ADN/ARN natural ou sintético e que possam multiplicar-se em uma célula viva, ou ainda as moléculas de ADN/ARN resultantes dessa multiplicação; consideram-se também os segmentos de ADN/ARN sintéticos equivalentes aos de ADN/ARN natural;

IV – engenharia genética: atividade de produção e manipulação de moléculas de ADN/ARN recombinante;

V – organismo geneticamente modificado - OGM: organismo cujo material genético – ADN/ARN tenha sido modificado por qualquer técnica de engenharia genética;

VI – derivado de OGM: produto obtido de OGM e que não possua capacidade autônoma de replicação ou que não contenha forma viável de OGM;

VII – célula germinal humana: célula-mãe responsável pela formação de gametas presentes nas glândulas sexuais femininas e masculinas e suas descendentes diretas em qualquer grau de ploidia;

VIII – clonagem: processo de reprodução assexuada, produzida artificialmente, baseada em um único patrimônio genético, com ou sem utilização de técnicas de engenharia genética;

IX – clonagem para fins reprodutivos: clonagem com a finalidade de obtenção de um indivíduo;

X – clonagem terapêutica: clonagem com a finalidade de produção de células-tronco embrionárias para utilização terapêutica;

XI – células-tronco embrionárias: células de embrião que apresentam a capacidade de se transformar em células de qualquer tecido de um organismo.

§ 1º Não se inclui na categoria de OGM o resultante de técnicas que impliquem a introdução direta, num organismo, de material hereditário, desde que não envolvam a utilização de moléculas de ADN/ARN recombinante ou OGM, inclusive fecundação *in vitro*, conjugação, transdução, transformação, indução poliplóide e qualquer outro processo natural.

§ 2º Não se inclui na categoria de derivado de OGM a substância pura, quimicamente definida, obtida por meio de processos biológicos e que não contenha OGM, proteína heteróloga ou ADN recombinante.

Art. 4o Esta Lei não se aplica quando a modificação genética for obtida por meio das seguintes técnicas, desde que não impliquem a utilização de OGM como receptor ou doador:

I – mutagênese;

II – formação e utilização de células somáticas de hibridoma animal;

III – fusão celular, inclusive a de protoplasma, de células vegetais, que possa ser produzida mediante métodos tradicionais de cultivo;

### CTNBio

Comissão Técnica Nacional de Biossegurança foi criada através da lei nº 11.105, de 24 de março de 2005 e tem a finalidade de prestar apoio técnico consultivo e assessoramento ao Governo Federal na formulação, atualização e implementação da Política Nacional de Biossegurança relativa a OGM, bem como no estabelecimento de normas técnicas de segurança e pareceres técnicos referentes à proteção da saúde humana, dos organismos vivos e do meio ambiente, para atividades que envolvam a construção, experimentação, cultivo, manipulação, transporte, comercialização, consumo, armazenamento, liberação e descarte de OGM e derivados. Fonte: <http://www.ctnbio.gov.br/>

IV – autoclonagem de organismos não-patogênicos que se processe de maneira natural.

Art. 5o É permitida, para fins de pesquisa e terapia, a utilização de células-tronco embrionárias obtidas de embriões humanos produzidos por fertilização *in vitro* e não utilizados no respectivo procedimento, atendidas as seguintes condições:

I – sejam embriões inviáveis; ou

II – sejam embriões congelados há 3 (três) anos ou mais, na data da publicação desta Lei, ou que, já congelados na data da publicação desta Lei, depois de completarem 3 (três) anos, contados a partir da data de congelamento.

§ 1o Em qualquer caso, é necessário o consentimento dos genitores.

§ 2o Instituições de pesquisa e serviços de saúde que realizem pesquisa ou terapia com células-tronco embrionárias humanas deverão submeter seus projetos à apreciação e aprovação dos respectivos comitês de ética em pesquisa.

§ 3o É vedada a comercialização do material biológico a que se refere este artigo e sua prática implica o crime tipificado no art. 15 da Lei no 9.434, de 4 de fevereiro de 1997.

Art. 6o Fica proibido:

I – implementação de projeto relativo a OGM sem a manutenção de registro de seu acompanhamento individual;

II – engenharia genética em organismo vivo ou o manejo *in vitro* de ADN/ARN natural ou recombinante, realizado em desacordo com as normas previstas nesta Lei;

III – engenharia genética em célula germinal humana, zigoto humano e embrião humano;

IV – clonagem humana;

V – destruição ou descarte no meio ambiente de OGM e seus derivados em desacordo com as normas estabelecidas pela CTNBio, pelos órgãos e entidades de registro e fiscalização, referidos no art. 16 desta Lei, e as constantes desta Lei e de sua regulamentação;

VI – liberação no meio ambiente de OGM ou seus derivados, no âmbito de atividades de pesquisa, sem a decisão técnica favorável da CTNBio e, nos casos de liberação comercial, sem o parecer técnico favorável da CTNBio, ou sem o licenciamento do órgão ou entidade ambiental responsável, quando a CTNBio considerar a atividade como potencialmente causadora de degradação ambiental, ou sem a aprovação do Conselho Nacional de Biossegurança – CNBS, quando o processo tenha sido por ele avocado, na forma desta Lei e de sua regulamentação;

VII – a utilização, a comercialização, o registro, o patenteamento e o licenciamento de tecnologias genéticas de restrição do uso.

Parágrafo único. Para os efeitos desta Lei, entende-se por tecnologias genéticas de restrição do uso qualquer processo de intervenção humana para geração ou multiplicação de plantas geneticamente modificadas para produzir estruturas reprodutivas estéreis, bem como qualquer forma de manipulação genética que vise à ativação ou desativação de genes relacionados à fertilidade das plantas por indutores químicos externos”.

Fonte: [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2004-2006/2005/lei/L11105.htm](http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/lei/L11105.htm)

### **Knock-out, genética reversa e RNAi**

A perda de função ou knock-out é um tipo de experimento utilizado dentre as estratégias para se conhecer a função de um determinado gene. Busca-se impedir que o gene produza a proteína de forma que se possa analisar as consequências ou o fenóti-

po decorrente desta inibição em um organismo, parte dele, ou em células cultivadas. Deriva-se então, por via indireta, em quais vias o produto deste gene está envolvido, vias de sinalizações, metabolismos, possíveis regulações, etc. Da mesma forma, este tipo de abordagem, da proteína para o gene é chamada de “genética reversa”, diferentemente da abordagem clássica (direta).

Uma das técnicas mais usadas para os estudos de genética reversa é a de *knock out* de genes específicos. Com modificações apropriadas, o *knock out* pode ser utilizado para identificar funções tecido-específicas de um determinado gene. Este método é difícil, laborioso, especializado e custoso. Entretanto é possível fazer um semi *knock out* de genes em experimentos de genética reversa, impedindo que ocorra a tradução do RNA mensageiro (mRNA). Os métodos de silenciamento gênico baseados em mRNA poderão ser utilizados no tratamento de doenças causadas pela expressão de genes deletérios. É possível utilizar agentes *antisense*. Esta estratégia consiste em introduzir um oligonucleotídeo complementar ao mRNA-alvo dentro de uma determinada célula, por exemplo. O híbrido DNA-RNA formado pode bloquear a ligação do mRNA ao ribossomo ou ativar a degradação do mRNA-alvo por uma RNase H. Apesar da simplicidade conceitual, as tecnologias antisense são, de certa forma, limitadas tanto pela ocorrência de inibição inespecífica de outros genes quanto pela dificuldade em se introduzir o agente antisense produzido.

O objetivo de se usar RNAi é silenciar de forma específica um determinado gene. Portanto, deve-se selecionar cuidadosamente uma região do mRNA alvo e evitar que a região escolhida tenha similaridade com regiões pertencentes a outros RNAs não relacionados. Aconselha-se fazer uma busca nos bancos de dados nos quais estão depositadas as seqüências do genoma para verificar se a seqüência escolhida é específica. Ao escolher um pequeno fragmento de RNA, recomenda-se evitar as primeiras 75 bases após o códon de início da transcrição, seqüências com mais de 50% de G + C, bem como, as regiões 5' e 3' não traduzidas, pois assume-se que proteínas regulatórias se liguem a essas regiões. Estudos recentes mostram que critérios bioquímicos, termodinâmicos e estruturais devem ser considerados quando se “desenham” siRNAs. Sugere-se que a extremidade 5' da fita antisense seja rica em AU, o que confere uma baixa estabilidade a essa extremidade, aumentando a probabilidade de que essa fita seja incorporada ao complexo RISC. Por outro lado, aconselha-se que a extremidade 5' da fita sense tenha um G ou C, de forma que apresente maior estabilidade interna. A utilização de diferentes dsRNA pequenos para um mesmo mRNA pode gerar resultados mais ou menos eficazes, dependendo da estrutura secundária da molécula alvo. Um siRNA é considerado eficiente quando provoca uma redução de pelo menos 90% no nível protéico.

Fonte: Barbosa e Lin (2004).

## Síntese da Parte



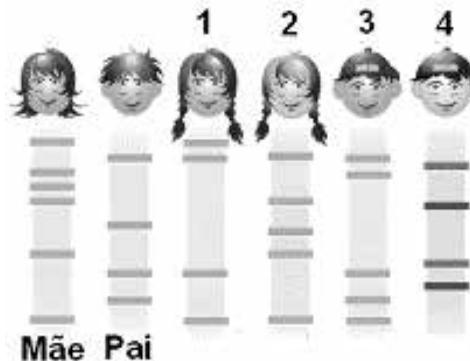
O DNA recombinante é feito com o uso de enzimas de restrição. Essas enzimas cortam o DNA em locais específicos. A transformação de bactérias, animais e plantas é possível com o uso das enzimas de restrição e de vetores que fazem a ponte entre o inserto e o organismo a ser transformado ou trans-

fectado. A reação da PCR possibilitou a ampliação controlada de sequências, permitindo novas aplicações, como a produção de bibliotecas, *fingerprint* etc.

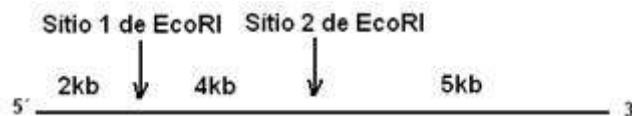
## Atividades de avaliação



1. Procure, leia e comente a lei de patentes e a Lei de Biosegurança. O que elas determinam e o que têm em comum?
2. Por que os cruzamentos comuns efetuados com variedades de plantas e animais não são considerados DNA *recombinante*? As raças animais e as variedades comuns de plantas (especialmente de importância econômica, como milho, feijão, arroz) que são lançadas pelas empresas agropecuárias ou institutos agrários têm que passar pelas avaliações do CTNBio e pelas questões da Lei de Biosegurança? Avalie.



3. No *fingerprint* de DNA acima, 1, 2, 3 e 4 são filhos do casal. Entretanto, um deles é filho (a) de um primeiro casamento (da mãe ou do pai) e um dos filhos (as) é adotivo (a). Identifique os dois casos mencionados e o Justifique essas afirmações.
4. O mapeamento de restrição de um trecho linear de DNA revela os seguintes **sítios de restrição** para a enzima *EcoRI*. Os fragmentos resultantes são separados por eletroforese em gel, e o gel é corado. No espaço para eletroforese abaixo, desenhe as bandas correspondentes: (Kb = Kilobases) a cada poço.



- a) Esse pedaço de DNA é digerido por *EcoRI*.
- b) Se ocorrer uma mutação que altera o sítio 1 de corte da *EcoRI*.

- c) Se ocorrer uma mutação que altera os sítios 1 e 2 de *EcoRI*.
- d) Se ocorrer uma inserção de 1000 pb entre os dois sítios de restrição.
- e) Se ocorrer uma deleção de 500 entre os dois sítios de restrição.



5. A partir da sequência polipeptídica a seguir, desenhe um **primer** degenerado, usando o código genético e o Quadro 13.1, de acordo com o exemplo:

Aminoácido	Asn	Tir	Tre	Tri	Lis	Tre	Val	Ala	Asn	Trp	Arg
Código genético	AAC	UAC	ACA								
	AAU	UAU	ACC								
			ACG								
			ACU								
Primer	AAS	TAS	ACN								

Arg (Arginina); Asn (Asparagina); Trp (Tryptofano); Tre (Treonina); Ala (Alanina); Gli (Glicina); Met (Metionina); Val (Valina); Leu (Leucina); Tir (Tirosina).

6. Verificação de enzimas de restrição e como cortar uma sequência: <http://rebase.neb.com/rebase/rebase.html>

Clique em "Rebase Tools"

Clique em "**Theoretical digests with all REBASE prototypes...**"  
(Digestão teórica com todos os prototipos encontrados na Rebase)

Digite ou cole uma sequência de DNA ou de proteínas no campo

Clique em *Submit* (Submeter)

Verifique os resultados

## Leituras, filmes e sites



### Leitura

- ANDRADE, A.; PINTO, S. C.; OLIVEIRA, R. S. (Orgs.) **Animais de laboratório: criação e experimentação**. 1. ed. Rio de Janeiro: FIOCRUZ, 2002. 388 p.
- CUMMINGS, K. **Concepts of genetics**. Prentice Hall: New Jersey, 1997. 703 p.
- DUPAS, G. **Ética e poder na sociedade da informação**. 2. ed. São Paulo: Unesp, 2004. 135 p.
- GRIFFITHS, A. J.F.; MILLER, J. H.; SUZUKI, D. T.; LEWONTIN, R. C.; GELBART, W. M. **An introduction to genetic analysis**. W.H. New York: Freeman and Company, 1996. 915 p.
- ROTHWELL, N. V. **Understanding genetics: a molecular approach**. New York: Wiley-Liss, 1993. 656 p.

### Sites

- Livros disponíveis na web (on-line): Bookshelf no site Entrez PubMed (<http://www.ncbi.nlm.nih.gov/PubMed/>):

Introduction to genetic analysis. 7th ed. Griffiths, Anthony J.F.; Gelbart, William M.; Miller, Jeffrey H.; Lewontin, Richard C. New York: W H Freeman & Co; c1999.

Modern genetic analysis. Griffiths, Anthony J.F.; Gelbart, William M.; Miller, Jeffrey H.; Lewontin, Richard C. New York: W H Freeman & Co; c1999.

Developmental biology. 6th ed. Gilbert, Scott F. Sunderland, Massachusetts: Sinauer Associates, Inc.; c2000. ISBN 0 87893 243 7.

### Cursos de Genética

- a) *Primer on Molecular Genetics* ([http://www.ornl.gov/sci/techresources/Human\\_Genome/publicat/primer/toc.html](http://www.ornl.gov/sci/techresources/Human_Genome/publicat/primer/toc.html))
- b) *Intermediate Genetics* (<http://www.ndsu.nodak.edu/instruct/mcclean/plsc431/>)
- c) *Plant Molecular Genetics* (<http://www.ndsu.nodak.edu/instruct/mcclean/plsc731/index.htm>)

- d) *Genetics*, Emporia State University (<http://www.emporia.edu/biosci/genetics/genetics.htm>)

**Banco de dados (genomas, sequências de DNA, proteínas, artigos, material didático, etc.), Periódicos On-Line, etc.**

- Entrez PubMed: <http://www.ncbi.nlm.nih.gov/PubMed/>
- The WWW Virtual Library: <http://vlib.org/Overview.html>
- Links: O mundo da Genética: [http://www.ornl.gov/TechResources/Human\\_Genome/links.html](http://www.ornl.gov/TechResources/Human_Genome/links.html)
- Projeto Genoma Humano: <http://www.ornl.gov/hgmis/>

### Gregor Mendel

- MendelWeb: <http://www.mendelweb.org/>
- Trabalhos de Mendel: <http://www.esp.org/foundations/genetics/classical/gm-65.pdf>

### Charles Darwin

- Trabalhos de Charles Darwin: <http://pages.britishlibrary.net/charles.darwin/>

### Filmes

- E a banda continua a tocar (Dir. Roger Spottiswoode, EUA, 1993) - descoberta da Aids.
- Nas montanhas dos gorilas (Dir. Michael Apted, EUA, 1988) - antropologia e meio ambiente.
- A ilha das flores (Dir. Jorge Furtado, Brasil, 1989) - crítica ao consumismo.

## Referências



BARBOSA, A. S.; LIN, C. J. Alternativamente ao knock-out de genes em experimentos de genética reversa, podem ser adotadas estratégias que impeçam a tradução do RNA mensageiro (mRNA). **Arquivos Brasileiros de Endocrinologia & Metabologia**, São Paulo, v. 48, n. 5, 2004.

COOPER, G. M.; HAUSMAN, R. E. **The cell: a molecular approach**. 5. ed. Sunderland: Sinauer Associates, 2009. 771 p.

GRIFFITHS, A. J. F.; GELBART, W. M.; MILLER, J. H.; LEWONTIN, R. C. **Modern genetic analysis**. New York: W. H. Freeman and Company, 1999. 589 p.

GRIFFITHS, A. J. F.; MILLER, J. H.; SUZUKI, D. T.; LEWONTIN, R. C.; GELBART, W. M. **An introduction to genetic analysis**. New York: W. H. Freeman and Company, 1996. 915 p.

LODISH, H.; BERK, A.; ZIPURSKY, S. L.; MATSUDAIRA, P.; BALTIMORE, D.; DARNELL, J. E. **Molecular Cell Biology**. 4. ed. New York: W. H. Freeman and Company, 1999. 1.084 p.

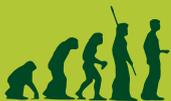
SILVA-VALENZUELA, M. G.; ALMEIDA, F. C. S.; MATIZONKAS-ANTONIO, L. F.; LIBÓRIO, T. N.; ACQUAFREDA, T.; CAZAL, C.; FERRAZ, A.; NUNES, F. D. Hibridização in situ com sonda não-radioativa para mRNA: princípios e aplicações em patologia. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, Rio de Janeiro, v. 42, n. 3, 2006.

## Sobre a autora

**Vânia Marilande Ceccatto**: é bióloga, licenciada e bacharelada pela Universidade Estadual Paulista Júlio de Mesquita Filho, a UNESP. Fez Mestrado em Ciências pela Universidade de São Paulo – USP e Doutorado em Bioquímica pela Universidade Federal do Ceará. É professora de Biologia Molecular desde 1998 na Universidade Estadual do Ceará. Está ligada ao Mestrado Acadêmico em Ciências Fisiológicas, também da UECE.



A não ser que indicado ao contrário a obra **Biologia Molecular**, disponível em: <http://educapes.capes.gov.br>, está licenciada com uma licença **Creative Commons Atribuição-Compartilha Igual 4.0 Internacional (CC BY-SA 4.0)**. Mais informações em: [http://creativecommons.org/licenses/by-sa/4.0/deed.pt\\_BR](http://creativecommons.org/licenses/by-sa/4.0/deed.pt_BR). Qualquer parte ou a totalidade do conteúdo desta publicação pode ser reproduzida ou compartilhada. Obra sem fins lucrativos e com distribuição gratuita. O conteúdo do livro publicado é de inteira responsabilidade de seus autores, não representando a posição oficial da EdUECE.



Ciências Biológicas

**F**iel a sua missão de interiorizar o ensino superior no estado Ceará, a UECE, como uma instituição que participa do Sistema Universidade Aberta do Brasil, vem ampliando a oferta de cursos de graduação e pós-graduação na modalidade de educação a distância, e gerando experiências e possibilidades inovadoras com uso das novas plataformas tecnológicas decorrentes da popularização da internet, funcionamento do cinturão digital e massificação dos computadores pessoais.

Comprometida com a formação de professores em todos os níveis e a qualificação dos servidores públicos para bem servir ao Estado, os cursos da UAB/UECE atendem aos padrões de qualidade estabelecidos pelos normativos legais do Governo Federal e se articulam com as demandas de desenvolvimento das regiões do Ceará.



UNIVERSIDADE ESTADUAL DO CEARÁ

