

# Genética Molecular





UNIVERSIDADE FEDERAL  
DE SANTA CATARINA

**BIOLOGIA**  
licenciatura a distância

# Genética Molecular

*Maria Cecília Menks Ribeiro*



Ministério  
da Educação



Florianópolis, 2009.

## **Governo Federal**

**Presidente da República** Luiz Inácio Lula da Silva

**Ministro de Educação** Fernando Haddad

**Secretário de Ensino a Distância** Carlos Eduardo Bielschowky

**Coordenador Nacional da Universidade Aberta do Brasil** Celso Costa

## **Universidade Federal de Santa Catarina**

**Reitor** Alvaro Toubes Prata

**Vice-Reitor** Carlos Alberto Justo da Silva

**Secretário de Educação à Distância** Cícero Barbosa

**Pró-Reitora de Ensino de Graduação** Yara Maria Rauh Muller

**Pró-Reitora de Pesquisa e Extensão** Débora Peres Menezes

**Pró-Reitor de Pós-Graduação** Maria Lúcia Camargo

**Pró-Reitor de Desenvolvimento Humano e Social** Luiz Henrique Vieira da Silva

**Pró-Reitor de Infra-Estrutura** João Batista Furtuoso

**Pró-Reitor de Assuntos Estudantis** Cláudio José Amante

**Centro de Ciências da Educação** Carlos Alberto Marques

## **Curso de Licenciatura em Ciências Biológicas na Modalidade a Distância**

**Diretora Unidade de Ensino** Sonia Gonçalves Carobrez

**Coordenador de Curso** Maria Márcia Imenes Ishida

**Coordenador de Tutoria** Zenilda Laurita Bouzon

**Coordenação Pedagógica** LANTEC/CED

**Coordenação de Ambiente Virtual** Alice Cybis Pereira

**Comissão Editorial** Viviane Mara Woehl, Alexandre Verzani Nogueira, Milton Muniz

## **Projeto Gráfico Material impresso e on-line**

**Coordenação** Prof. Haenz Gutierrez Quintana

**Equipe** Henrique Eduardo Carneiro da Cunha, Juliana Chuan Lu, Laís Barbosa, Ricardo Goulart Tredezini Straioto

## **Equipe de Desenvolvimento de Materiais**

**Laboratório de Novas Tecnologias - LANTEC/CED**

**Coordenação Geral** Andrea Lapa

**Coordenação Pedagógica** Roseli Zen Cerny

## **Material Impresso e Hipermídia**

**Coordenação** Laura Martins Rodrigues, Thiago Rocha Oliveira

**Adaptação do Projeto Gráfico** Laura Martins Rodrigues, Thiago Rocha Oliveira

**Diagramação** Jessé Antunes Torres, Eduardo Santaella Malaguti, Ana Flávia Maestri, Laura Martins Rodrigues

**Ilustrações** Jessé Antunes Torres, Eduardo Santaella, Marc Bogo, Thiago Rocha Oliveira, Rafael Naravan, Karina Silveira, Rafael de Queiroz Oliveira

**Revisão gramatical** Isabel Maria Barreiros Luclktenberg

## **Design Instrucional**

**Coordenação** Isabella Benfica Barbosa

**Design Instrucional** Vanessa Gonzaga Nunes

## **Colaboração**

Heloisa Helena Zaccaron Milioli, Tiago Moretti

Copyright © 2009 Universidade Federal de Santa Catarina. Biologia/EaD/UFSC  
*Nenhuma parte deste material poderá ser reproduzida, transmitida e gravada sem a prévia autorização, por escrito, da Universidade Federal de Santa Catarina.*

R484g Ribeiro, Maria Cecilia Menks  
Genética molecular / Maria Cecilia Menks Ribeiro.— Florianópolis :  
BIOLOGIA/EAD/UFSC, 2009.

158p.

ISBN 978-85-61485-12-2

1. Histórico e estrutura do DNA. 2. Replicação. 3. Fluxo de informação. 4. Mutação. I. Título.

CDU 575.1

Catálogo na fonte elaborada na DECTI da Biblioteca Universitária da Universidade Federal de Santa Catarina.

# Sumário

<b>Apresentação.....</b>	<b>7</b>
<b>1 Histórico e Estrutura do DNA.....</b>	<b>11</b>
1.1 O Princípio Transformante e o DNA.....	14
1.2 Estrutura Molecular .....	17
1.3 Propriedades .....	22
1.4 Tipos de Seqüência .....	24
1.5 Organização Celular.....	30
1.6 Funções Biológicas.....	31
Resumo.....	33
Referências .....	34
<b>2 Replicação do DNA.....</b>	<b>37</b>
2.1 A Replicação do DNA é central para toda a Biologia.....	40
2.2 Visualização Celular da Replicação: Autorradiografia .....	43
2.3 Origem Única e Replicação Bidirecional .....	44
2.4 DNA Polimerase e Síntese <i>in vitro</i> do DNA .....	44
2.5 O “Paradoxo do Ponto de Crescimento” e a Síntese Descontínua de DNA .....	46
2.6 Iniciação e Problema do “Primer”.....	48
2.7 A Complexidade do Sistema de Replicação Completo .....	49
2.8 Replicação Semiconservativa em Cromossomos Eucarióticos.....	52
2.9 Múltiplos Replicons por Cromossomo .....	52
2.10 O Início da Replicação .....	55
2.11 O Término da Replicação: A questão dos telômeros.....	55
2.12 Componentes do Aparelho de Replicação dos Eucariotos.....	58
Resumo.....	59
Referências .....	59

### **3 O Fluxo da Informação ..... 63**

3.1 A estrutura do RNA.....	67
3.2 O RNA e o fluxo da informação.....	70
3.3 Síntese do RNA.....	73
3.4 RNA Polimerases.....	74
3.5 RNA Replicases Virais.....	75
3.6 Transcrição do RNA.....	75
3.7 Tipos de RNA.....	82
3.8 Tradução.....	101
3.9 Diferenças entre Procariotos e Eucariotos.....	111
3.10 Inibidores de Síntese de Proteínas.....	113
Resumo.....	115
Referências.....	117

### **4 Variabilidade Genética ..... 119**

4.1 Mutação.....	122
4.2 Reparo do DNA.....	137
4.3 Recombinação.....	146
4.4 Recombinação Homóloga ou Generalizada.....	146
4.5 Recombinação Sítio Específica.....	148
4.6 Transposição.....	148
4.7 Retrovírus e Retrotransposons.....	153
4.8 Transformação (Exp. GRIFFITH).....	155
Resumo.....	156
Referências.....	158

# Apresentação

*O século XX caracterizou-se por uma nova visão da Biologia, construída a partir da compreensão dos eventos moleculares intracelulares, registrando estruturas celulares em dimensões submicroscópicas. Impulsionada por um novo aparato instrumental, foi possível caracterizar as moléculas biológicas com uma precisão sem precedentes.*

*Essa nova compreensão da Biologia foi fundamental para elucidar diversos aspectos do funcionamento celular e principalmente da hereditariedade, descrevendo um universo surpreendente e fascinante. Atualmente a Genética Molecular é uma das áreas mais efervescentes da ciência.*

*A compreensão da diversidade biológica como um conjunto de variações sobre um mesmo tema, a vida, leva naturalmente à busca de atributos compartilhados pelos seres vivos. A pesquisa biológica revelou, de fato, uma unidade desconcertante na diversidade da vida, na forma da universalidade do DNA como memória genética (com exceção dos vírus de RNA, se forem considerados vivos), na quase universalidade do código genético, no compartilhamento de vias metabólicas básicas etc.*

*Esses conhecimentos vêm gerando um grande número de novas tecnologias com aplicações que se estendem a diversas áreas da atividade humana como a identificação de indivíduos, a detecção de mutações associadas a doenças hereditárias e as diversas vertentes da biotecnologia.*

*Os projetos genoma concluídos e em andamento geram uma grande quantidade de informações, que são continuamente agregadas e atualizadas. O armazenamento e o acesso a essas informações se dão por meio de um banco de dados com acesso livre, o genbank ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)).*

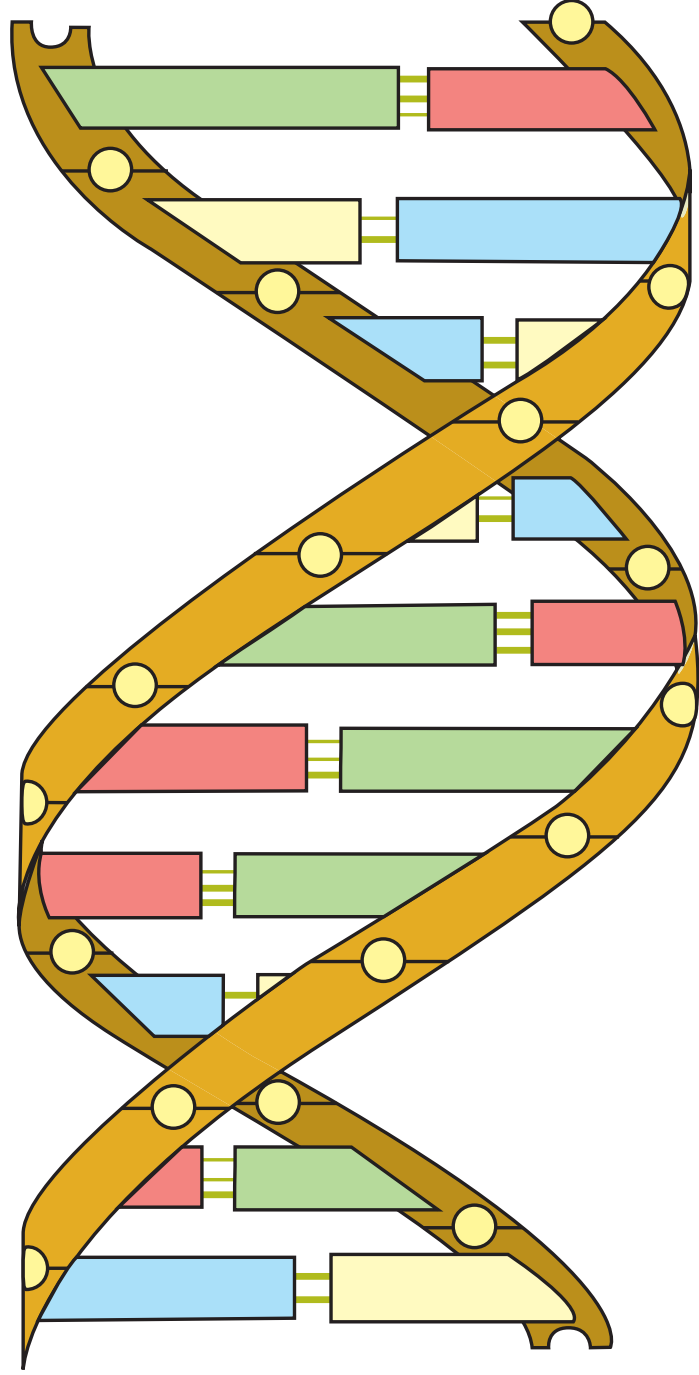
*O seqüenciamento, o isolamento e a clonagem de genes, a organização e a caracterização de alterações no genoma, o desenvolvimento de organismos transgênicos, o isolamento e a purificação de outras biomoléculas e a determinação da estrutura de proteínas representam algumas das inúmeras questões que só foram detectadas e/ou esclarecidas por meio de metodologias que essa ciência possibilitou desenvolver com a finalidade de desvendar a natureza mais íntima dos processos biológicos, impulsionando pesquisas em várias outras áreas do conhecimento. A Genética Molecular se fortaleceu pela colaboração de bioquímicos e de biólogos moleculares, e vem se consolidando como uma das principais áreas científicas do século XXI.*

**Maria Cecília Menks Ribeiro**





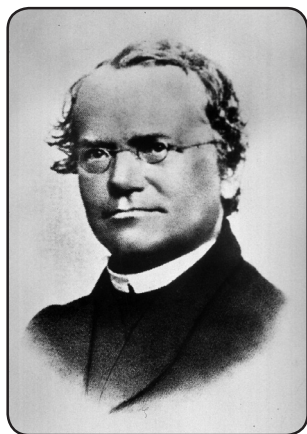
# CAPÍTULO 1



## Histórico e Estrutura do DNA

*Este capítulo resgata aspectos históricos da descoberta do material genético até a identificação do DNA e caracteriza sua estrutura molecular e a relação com suas funções biológicas.*





*Gregor Mendel (1822–1884) publicou seu famoso estudo em 1866. Apesar de ter enviado cópias de seu trabalho para diversos cientistas eminentes da época, permaneceu ignorado pela comunidade científica por mais 34 anos, sendo somente em 1900 seus trabalhos redescobertos por três geneticistas botânicos interessados em problemas similares, como os das ervilhas de Mendel.*

Os conceitos fundamentais de genética começaram a se definir com Mendel, em 1868, com suas conclusões acerca da perpetuação das características hereditárias. A partir de experimentos realizados com ervilhas, deduziu que os fatores responsáveis pela hereditariedade ocorriam aos pares. Verificou também que as proporções observadas nos descendentes eram compatíveis com a distribuição binomial.

Outra contribuição importante foi de Friedrich Miescher, que, em 1869, investigando a composição química de células humanas (obtidas de curativos purulentos – linfócitos), isolou uma substância proveniente do núcleo que denominou **nucleína**, caracterizando sua constituição química (oxigênio, nitrogênio e fósforo). Mas a função celular dessa molécula não foi esclarecida.

Paralelamente, estudos citológicos, realizados pelo próprio Miescher, caracterizaram o núcleo como um componente importante para a continuidade da célula. Também foram reconhecidas estruturas dinâmicas nas células, denominadas cromossomos, que se individualizam na divisão celular, estando associados à perpetuação da célula. Foram caracterizados os processos de mitose e meiose.

A partir de 1900, quando as leis de **Mendel** foram redescobertas, diversas características com herança mendeliana foram iden-

### História da Nucleína

Em 1869, Friederich Miescher, um bioquímico suíço que trabalhava na Alemanha, isolara a partir de bandagens impregnadas de pus fornecidas por um hospital local uma substância que denominou “nucleína”. Como o pus é primordialmente formado por glóbulos brancos, que possuem um núcleo e, portanto, cromossomos contendo DNA, ao contrário dos glóbulos vermelhos que são anucleados, proporcionou uma boa fonte de DNA. Mais tarde, quando descobriu que a “nucleína” só era encontrada nos cromossomos, compreendeu a importância de sua descoberta.

tificadas em diversos organismos. Também começaram a surgir evidências de que as características mendelianas estariam relacionadas a enzimas, fundamentando o conceito “um gene, uma enzima”. O interessante é que esses conceitos poderiam ser aplicados a qualquer organismo, sendo um ponto de convergência importante para a Biologia. A descoberta dos microorganismos veio potencializar essa investigação, proporcionando a oportunidade de caracterizar o material hereditário.

A contribuição de Archibald Garrod (1902) foi fundamental para a genética humana. Esse médico inglês interessou-se por uma doença rara, a alcaptonuria, **cujo sintoma mais pronunciado era a cor da urina**, que se tornava escura quando exposta ao ar, sugerindo tratar-se de um erro no metabolismo, pela disrupção de uma via bioquímica. Embora essa condição fosse rara na população, era freqüente entre filhos de pais consangüíneos, sendo assim explicada pelas leis de Mendel. Essa foi a primeira conexão causal entre genes e seu efeito fisiológico. Os genes de certa forma controlam processos metabólicos, de maneira que um erro – mutação – pode resultar numa rota metabólica alterada. Outros exemplos logo foram descritos. Atualmente mais de 5.000 condições com herança mendeliana são conhecidas em humanos e estão agrupadas na base de dados **OMIM**.

A relação entre deficiência metabólica e características genéticas foi demonstrada em experimentos realizados na levedura *Neurospora*, por Beadle e Tatum em 1941, corroborando a idéia de que os genes deveriam ser enzimas.

: A ciência da época atribuí-  
: tal escurecimento a uma  
: substância produzida  
: pelas bactérias que vivem  
: no intestino, mas Garrod  
: argumentava que o  
: aparecimento da urina preta  
: em recém-nascidos, cujo  
: intestino não possui colônias  
: de bactérias, significaria que  
: a substância era produzida  
: pelo próprio corpo.

www.ncbi.nih.gov/omim

## 1.1 O Princípio Transformante e o DNA

Em 1928 Frederick Griffith descobriu a transformação bacteriana. Esse médico, estudando o agente causador da pneumonia, o *Streptococcus pneumoniae*, evidenciou que ocorria a “transformação” de uma cepa não patogênica (R) em uma cepa patogênica (S) (veja a Figura 1.1). A transformação estava associada a um componente celular que foi denominado “princípio transformante”, que deveria ser o material genético, estando assim associado à virulência da bactéria.

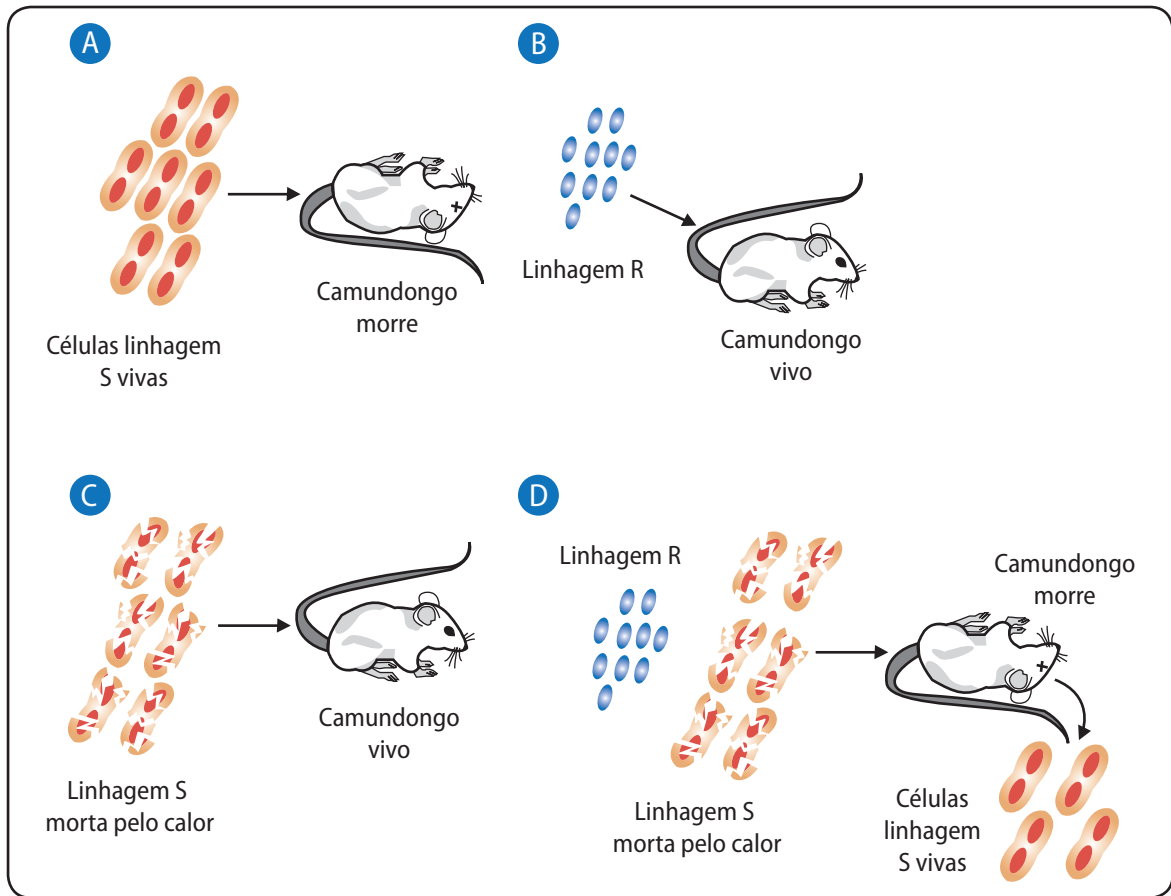


Figura 1.1 - Experimento de Griffith demonstrando a transformação bacteriana

O curioso é que o “princípio transformante” parecia ser uma substância resistente ao calor, uma propriedade não observada nas proteínas.

Esse caráter do armazenamento e da transmissão da informação biológica como característica do processo vital suscitou diversas especulações sobre a natureza das unidades hereditárias. Essa identificação da natureza química dos genes tornaria possível compreender os fenômenos biológicos através das leis da Física e da Química. Esse pensamento foi formalizado pelo físico Erwin Schrödinger, em seu livro “O que é vida?”, publicado em 1944, influenciando diversos protagonistas das descobertas subsequentes.

Desta forma, a caracterização do material hereditário tornou-se um desafio científico. Por um lado parecia óbvio que as **proteínas** seriam o material hereditário, já que eram as únicas moléculas suficientemente complexas para essa função tão importante. E o fato

*Proteínas são cadeias moleculares formadas de 20 componentes básicos diferentes, os aminoácidos. Como as permutações na ordem dos aminoácidos ao longo da cadeia são quase infinitas, os biólogos acreditavam que elas poderiam facilmente codificar as informações que sustentam a extraordinária diversidade da vida.*

de serem componentes dos cromossomos, as estruturas celulares relacionadas à hereditariedade, corroborou essa idéia. No entanto, a maneira como essas moléculas eram replicadas e transmitidas de uma célula para outra era insuspeitável.

Paralelamente, a caracterização da natureza química e funcional das moléculas biológicas avançava continuamente, principalmente com o advento de novas metodologias. Assim, a composição química da nucleína começou a ser elucidada, sendo denominada ácido nucléico, identificados os seus componentes químicos básicos e evidenciada sua natureza polimérica. Cada unidade do polímero foi denominada nucleotídeo e composta de bases nitrogenadas, açúcar e fosfato. Esse açúcar é uma pentose, denominada ribose. Quando tem uma perda do oxigênio no carbono 2', é denominada desoxirribose. Embora esses dois tipos de açúcar sejam encontrados nos ácidos nucléicos, não ocorrem simultaneamente na mesma molécula. Desta forma são reconhecidos dois tipos de ácidos nucléicos: ácido ribonucleico (RNA), presente no citoplas-

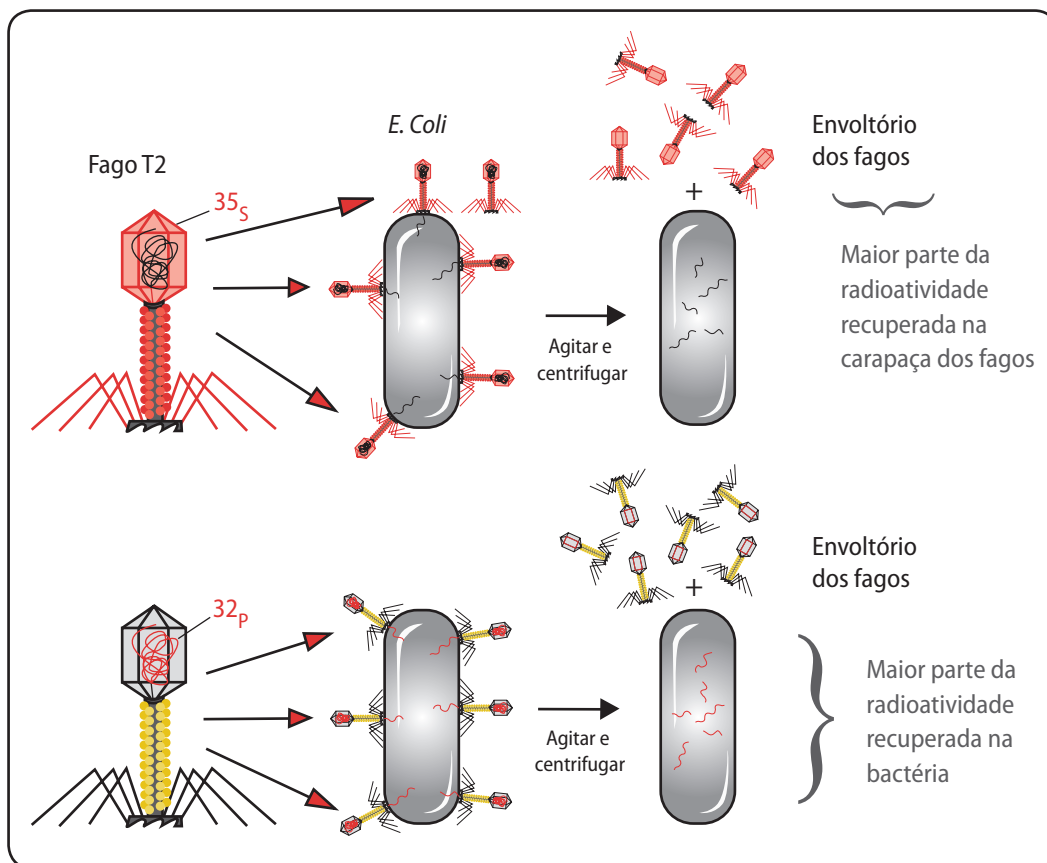


Figura 1.2 - Experimento com isótopos radioativos realizado por Hershey e Chase



ma e núcleo, e ácido desoxirribonucleico (DNA), encontrado majoritariamente no núcleo. No entanto, o DNA foi considerado um componente periférico dos seres vivos. A uniformidade de sua estrutura e a regularidade da sua composição sugeriam uma função estrutural, sem uma atividade metabólica relevante.

Foi o grupo liderado por Oswald Avery (1944) que caracterizou o princípio transformante, realizando experimentos de transformação similares aos realizados por Griffith, mas com os componentes do extrato celular purificados. Qual não foi a surpresa ao constatar que o princípio transformante era o DNA. A confirmação definitiva que o DNA era o material hereditário foi obtida por Alfred Hershey e Martha Chase (1952), demonstrando que o componente viral que infecta as células é o DNA (veja a Figura 1.2).

## 1.2 Estrutura Molecular

Embora com essas evidências fosse inquestionável que o DNA era o material hereditário, ainda permanecia obscura a maneira como os genes transmitiam suas características às gerações.

Admitia-se que se descobrindo a estrutura molecular do DNA seria possível elucidar sua função biológica e compreender a natureza do código genético, e diversos cientistas consagrados, como Linus Pauling, buscavam elucidar a estrutura molecular do DNA.

A caracterização da estrutura molecular do DNA era um desafio por sua complexidade. Em 1938 William Astbury havia obtido *imagens cristalográficas* do DNA que evidenciaram uma alta densidade da molécula e uma estrutura helicoidal, determinando também o intervalo entre a repetição de suas subunidades (veja a Figura 1.3).

No entanto, a correlação entre a composição química e a estrutura não era óbvia, tampouco parecia ter um significado funcional e muito menos proporcionar a diversidade esperada para uma molécula com esse atributo.

Erwin Chargaff demonstrou que havia diferenças entre os organismos na composição das bases, sugerindo os

A difração de raios X é uma maneira de estudar a estrutura atômica de qualquer molécula que possa ser cristalizada. O cristal é bombardeado com raios X, que espalham seus átomos e dispersam-nos. O padrão de dispersão fornece informações sobre a estrutura da molécula, embora por si só não baste para elucidá-la.

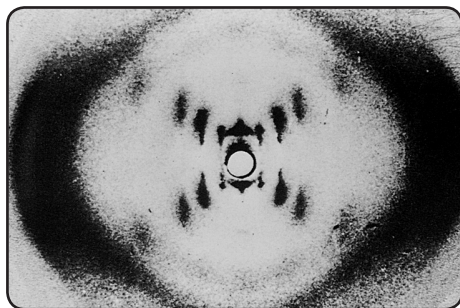


Figura 1.3 - Imagem cristalográfica do DNA obtida por William Astbury

primeiros indícios da diversidade no DNA (veja a Tabela 1). No entanto, havia proporcionalidade entre o conteúdo de purinas (adeninas e guaninas) e pirimidinas (citosinas e timinas), assim como entre as adeninas (A) e timinas (T) e citosinas (C) e guaninas (G).

**Tabela 1 - Concentração molar das bases\* no DNA de várias origens**

Organismo	Tecido	Adenina	Timina	Guanina	Citosina	A+T/G+C
<i>Escherichia coli</i> (K12)	–	26,0	23,9	24,9	25,2	1,00
<i>Diplococcus pneumoniae</i>	–	29,8	31,6	20,5	18,0	1,59
<i>Mycobacterium tuberculosis</i>	–	15,1	14,6	34,9	35,4	0,42
Levedura	–	31,3	32,9	18,7	17,1	1,79
<i>Paracentrotus lividus</i> (ouriço-do-mar)	Esperma	32,8	32,1	17,7	18,4	1,85
Peixe	Esperma	27,8	27,5	22,2	22,6	1,23
Rato	Medula óssea	28,6	28,4	21,4	21,5	1,33
Homem	Timo	30,9	29,4	19,9	19,8	1,52
Homem	Fígado	30,3	30,3	19,5	19,9	1,53
Homem	Esperma	30,7	31,2	19,3	18,8	1,62

\* Definida como moles de constituintes de nitrogênio por 100g de átomos fosfato hidrolisados

Já no laboratório liderado por Maurice Wilkins no King's College em Londres, Rosalyn Franklin buscava determinar as medidas cristalográficas da molécula de DNA por meio do padrão de difração dos raios X para subsidiar um modelo estrutural para a molécula.

Foram os cientistas James Watson e Francis Crick que em 1953 conseguiram decifrar a estrutura do DNA, reunindo as diferentes informações existentes e inspirados na metodologia utilizada por Linus Pauling para a elaboração de modelos de proteínas.

Esse modelo apresenta uma simplicidade desconcertante e ainda sugere indícios de seu funcionamento.

De acordo com o modelo proposto, os nucleotídeos (veja a Figura 1.4), monômeros da molécula, ligar-se-iam em longas cadeias por meio do grupo fosfato e da hidroxila da pentose (desoxirribose).

Desta forma, a ordem possível dos nucleotídeos seria ilimitada.

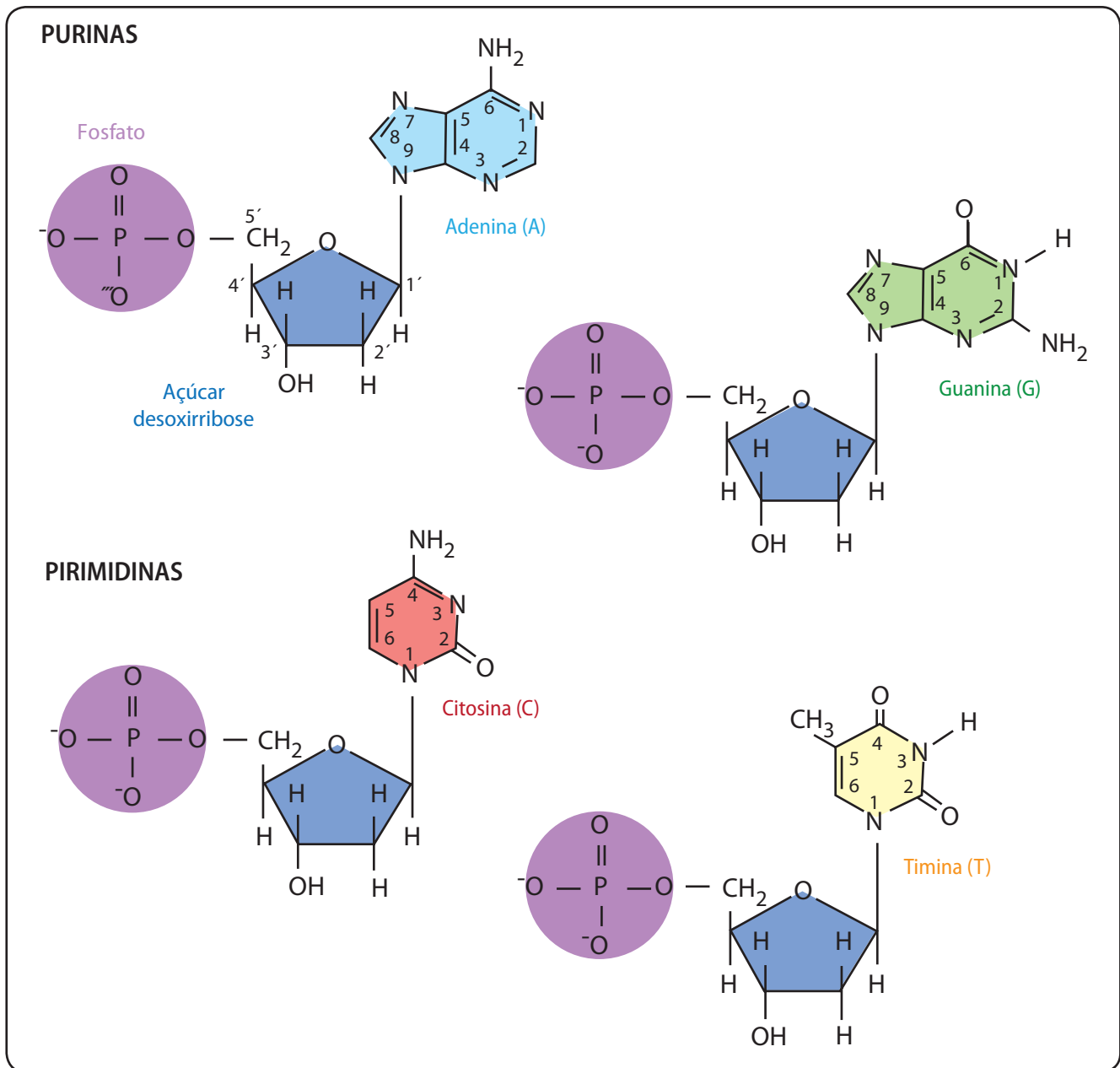


Figura 1.4 - Estrutura dos nucleotídeos do DNA

*<<http://biomodel.uah.es/model1j/dna/inicio.htm>>  
(neste link a molécula de DNA pode ser visualizada em diversos ângulos)*

A alta densidade da molécula sugeriu a presença de mais de uma cadeia, e pelo **modelo estrutural** foi demonstrado que a particularidade mais interessante do DNA é o fato de tratar-se de uma fita dupla, sendo as cadeias de orientações opostas (antiparalelas). As duas cadeias manter-se-iam unidas por pontes de hidrogênio, havendo complementaridade entre os nucleotídeos das duas cadeias, sendo A complementares às T e as C complementares à G. Desta forma, a informação na molécula de DNA estaria na seqüência de nucleotídeos.

Com esse arranjo, a distância entre as fitas seria uniforme, medindo 20Å (Angstroms), esse conjunto apresenta uma volta completa a cada 34Å. Essa proposta se ajusta às dimensões estabelecidas pelas imagens cristalográficas (veja a Figura 1.5).

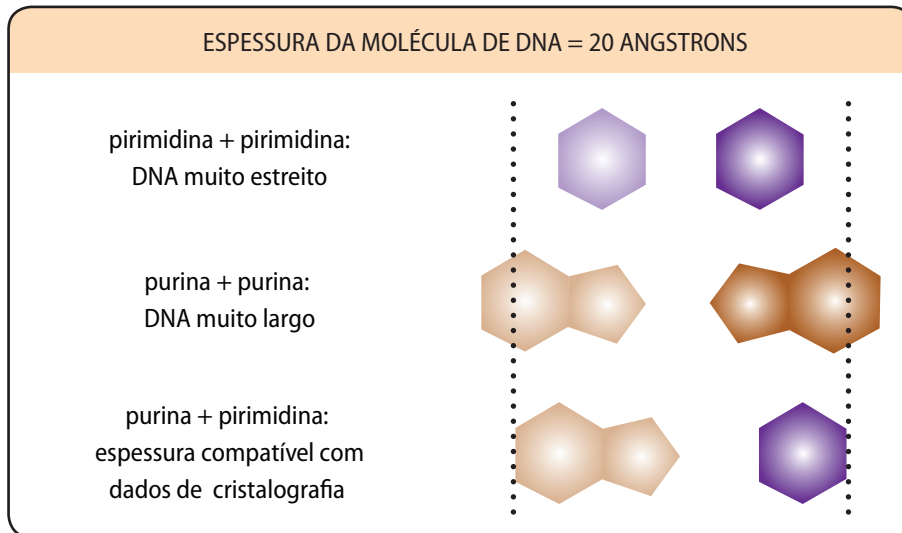


Figura 1.5 - Espessura da molécula de DNA e sua relação com o pareamento de bases

A ligação entre os nucleotídeos na cadeia se dá por meio de uma ligação fosfodiéster, conferindo uma grande resistência química à molécula e possibilitando assim a continuidade da seqüência, sendo a seqüência de nucleotídeos a particularidade de cada molécula de DNA (veja a Figura 1.6).

O modelo proposto para a estrutura molecular do DNA era compatível com a sua função biológica, a informação poderia ser armazenada na seqüência de nucleotídeos. E pelo modelo proposto não havia nenhuma restrição quanto à seqüência, ainda a presença de uma cadeia complementar sugeria um possível mecanismo de cópia.

O modelo proposto por Watson e Crick, apesar de fundamentado em análises estruturais, foi um modelo teórico, não proporcionou a possibilidade instrumental de averiguar a seqüência de bases. Nos anos subseqüentes diversas alternativas metodológicas foram implementadas com o intuito de se obterem informações a respeito da seqüência de bases e do sistema operacional.

Atualmente grande parte dessas questões estão sendo esclarecidas, a obtenção de **seqüências de DNA** e a comparação de seqüências são obtidas por diversos instrumentos, sendo as informações agregadas num banco de dados de acesso universal, o genbank.

Podem ocorrer diferenças no gene responsável por uma característica. Denomina-se alelo cada forma alternativa do gene, que, por sua vez, vai determinar uma característica diferente na pessoa. Como exemplo, podemos citar o sistema ABO de grupos sanguíneos. As pessoas podem ser, então, do grupo sanguíneo A, B, AB ou O. Dessa maneira, pode-se observar que elas, embora apresentem uma estrutura de DNA semelhante, podem ter características diferentes fenotípicas de acordo com a variabilidade das seqüências em seus genes. Acesse: <[www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)>.

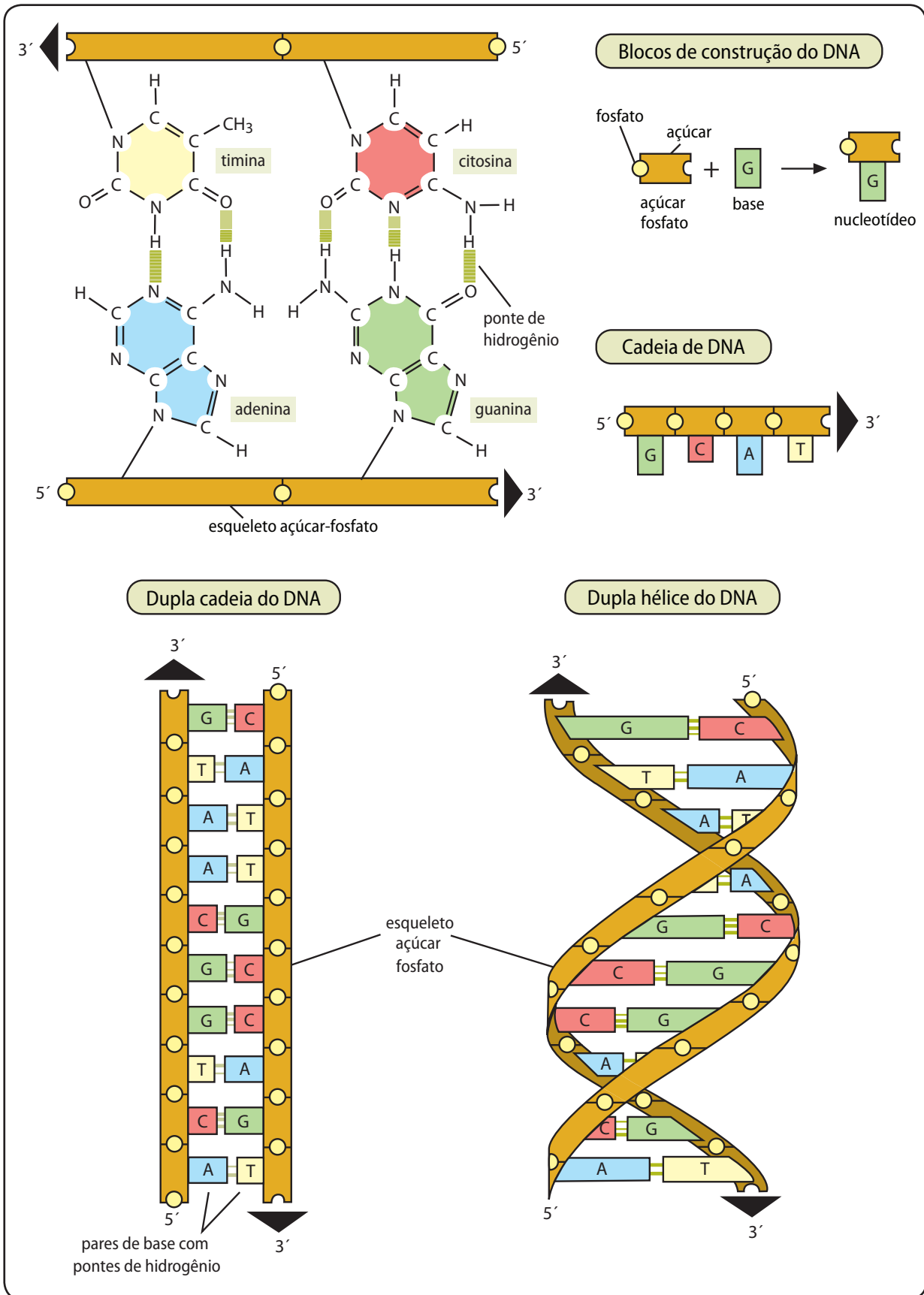


Figura 1.6 - Modelo para estrutura da molécula de DNA

Uma das principais questões levantada é de que a seqüência de bases possibilitaria o armazenamento da informação, poderia ser respondida por meio da análise e da comparação das seqüências de DNA de diferentes organismos.

Esse pensamento pode parecer simples, mas, ao se verificar o tamanho do genoma de diferentes organismos, pode-se ter uma idéia da dimensão do problema.

O próximo desafio seria a obtenção de seqüências de DNA do maior número possível de organismos e a construção de bancos de dados acessíveis que possibilitassem acessar e processar informações com agilidade.

### 1.3 Propriedades

As ferramentas para caracterizar o DNA foram desenvolvidas inicialmente a partir das propriedades físico-químicas do DNA. O DNA é uma molécula estável quimicamente, diversos métodos laboratoriais possibilitam a obtenção e a purificação do DNA genômico. Essas moléculas uma vez purificadas podem ser mantidas em laboratório com relativa facilidade. Essas características também possibilitam a obtenção de DNA a partir de vestígios biológicos e mesmo a partir de fósseis de animais e plantas já extintos.

Uma vez que as cadeias são mantidas unidas por pontes de hidrogênio, ligações químicas fracas, é relativamente fácil separar as duas cadeias – processo chamado de desnaturação – experimentalmente pelo calor ou pelo tratamento com álcalis. Uma vez restituídas as condições adequadas ocorre a renaturação, ou seja, a reunião das cadeias, respeitando-se a complementaridade de bases, processo denominado hibridação, que proporciona uma grande gama de aplicações. Essa hibridação pode ocorrer com o DNA original ou com qualquer outra molécula de DNA que apresente **homologia** (veja a Figura 1.7).

A **eletroforese** é uma metodologia que, pela natureza ácida do DNA, possibilita a separação de moléculas de acordo com o seu tamanho (veja a Figura 1.8).

#### • Eletroforese

- Método para separação de
- macromoléculas (DNA, RNA
- e proteínas) em uma matriz
- na qual se aplica um campo
- elétrico. A migração de cada
- fragmento depende de sua
- carga elétrica, dessa forma, o
- DNA, por apresentar cargas
- negativas devido à presença
- do ácido fosfórico, possui
- suas moléculas atraídas
- para o pólo positivo e
- repelidas do pólo negativo.

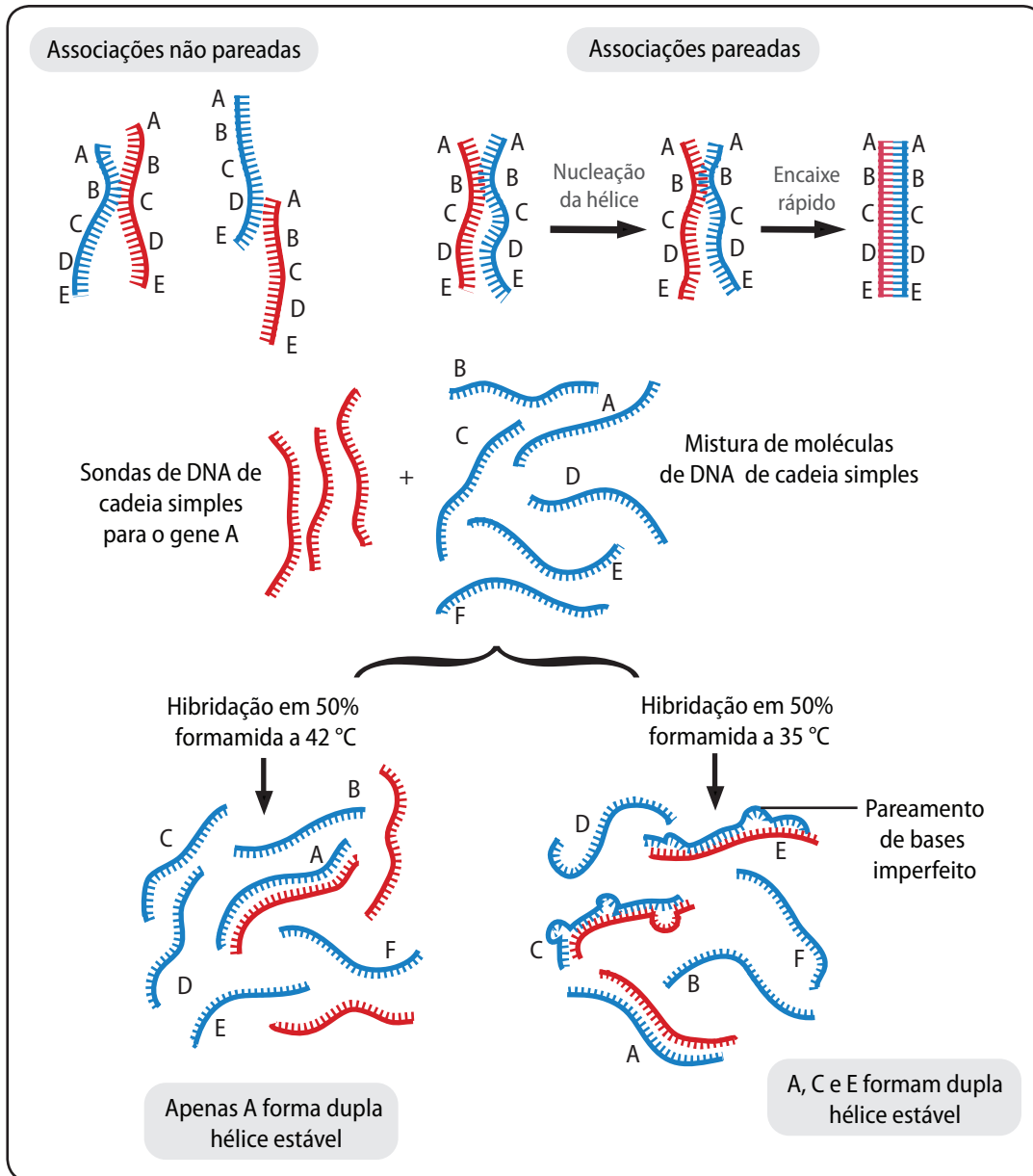
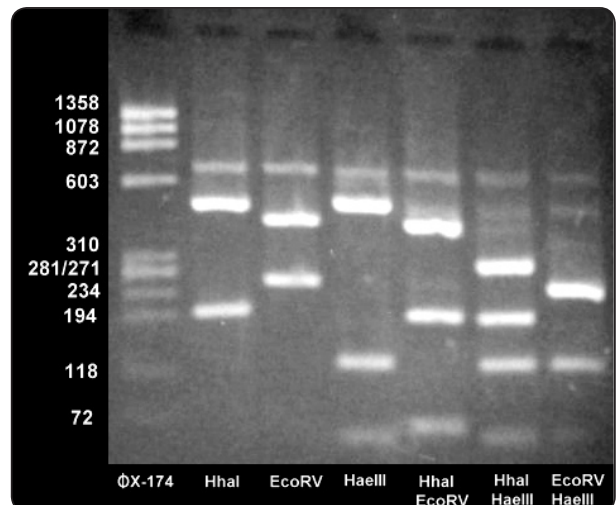


Figura 1.7 - Desnaturação e reassociação de moléculas de DNA (hibridação). As duplas hélices reassociam a partir das cadeias separadas numa reação que depende da colisão ao acaso de duas cadeias complementares de DNA. A grande maioria das colisões são não produtivas, enquanto que raras resultam numa pequena região com complementaridade, formando pares de bases

Figura 1.8 - Eletroforese de moléculas de DNA em gel de agarose



## 1.4 Tipos de Seqüência

Existem duas características importantes para individualizar o DNA, a seqüência de bases e o tamanho da molécula, sendo utilizada a unidade pb (pares de bases) para quantificar o **DNA**.

Para mais informações acesse:  
<[http://nobelprize.org/educational\\_games/medicine/dna\\_double\\_helix](http://nobelprize.org/educational_games/medicine/dna_double_helix)>.

Por meio dessas abordagens preliminares foi possível caracterizar o conteúdo de DNA de diversos organismos e também iniciar a caracterização da seqüência (Figura 1.9).

O conteúdo de DNA é virtualmente constante em todas as células do organismo, mas pode diferir entre organismos congêneres, sendo denominado genoma.

Também se começou a esclarecer o significado de outro ácido nucléico, o RNA, cuja quantidade é variável, estando relacionado à atividade metabólica da célula.

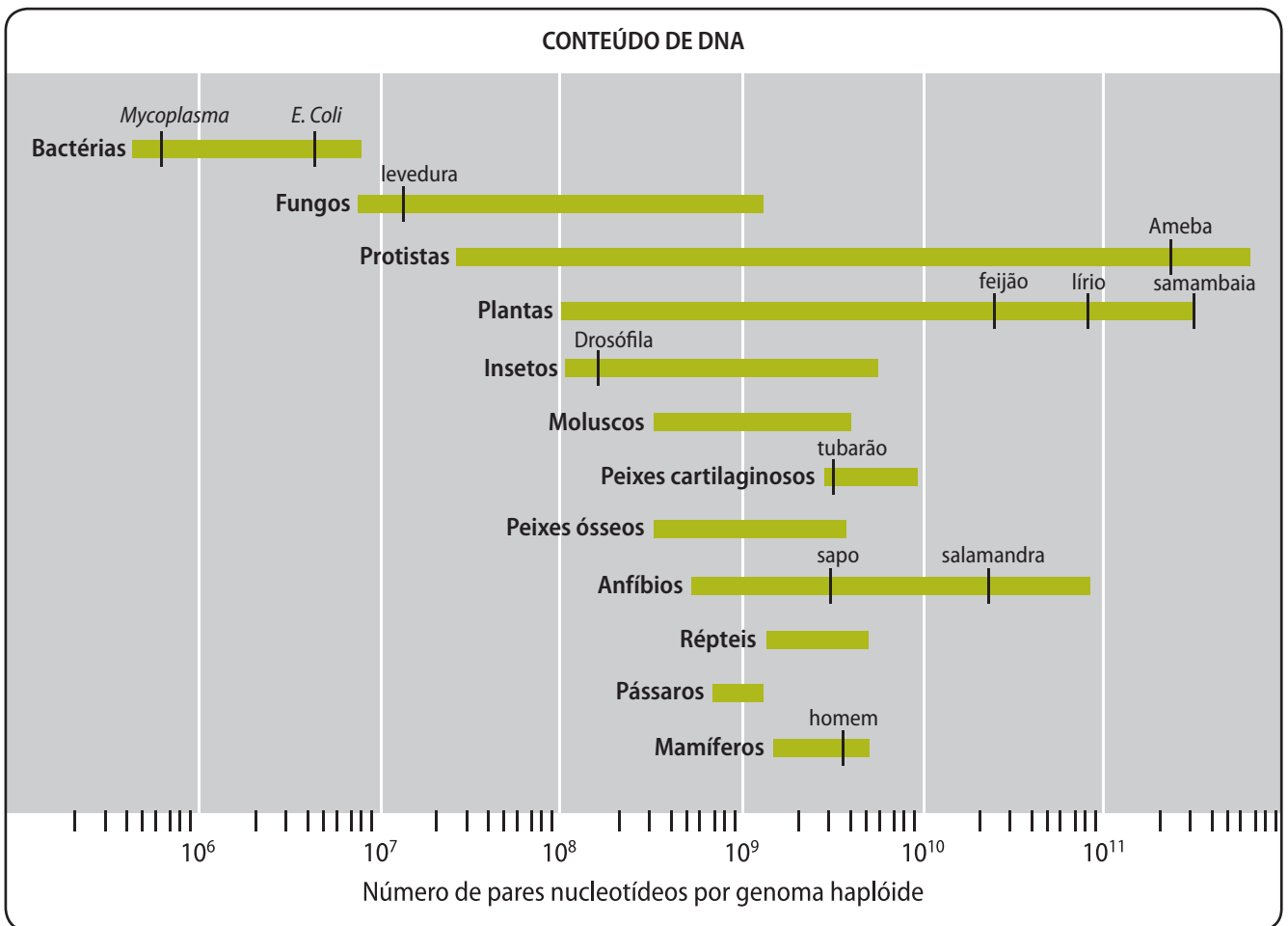


Figura 1.9 - Conteúdo de DNA em diferentes organismos



Adicionalmente, esses estudos possibilitaram a comparação inicial do DNA de diferentes organismos. Quanto maior a homologia existente, maior o índice de reassociação existente e a homologia diminui à medida que os organismos se distanciam evolutivamente. Recentemente, foi demonstrado que algumas regiões do genoma humano e do chimpanzé têm até 99,4% de homologia (veja a Tabela 2).

**Tabela 2 - Similaridades entre a seqüência de nucleotídeos no DNA de organismos de diferentes reinos**

Organismo	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>Saccharomyces cerevisiae</i> , 18S	–	0,29	0,33	0,05	0,06	0,08	0,09	0,11	0,08	0,11	0,11	0,08	0,08
<i>Lemna minor</i> , 18S	0,29	–	0,36	0,10	0,05	0,06	0,10	0,09	0,11	0,10	0,10	0,13	0,08
L cell, 18S	0,33	0,36	–	0,06	0,06	0,07	0,07	0,09	0,06	0,10	0,10	0,09	0,07
<i>Escherichia coli</i>	0,05	0,10	0,06	–	0,24	0,25	0,28	0,26	0,21	0,11	0,12	0,07	0,12
<i>Cholorbium vibrioforme</i>	0,06	0,05	0,06	0,24	–	0,22	0,22	0,20	0,19	0,06	0,07	0,06	0,09
<i>Bacillus firmus</i>	0,08	0,06	0,07	0,25	0,22	–	0,34	0,26	0,20	0,11	0,13	0,06	0,12
<i>Corynebacterium diphtheriae</i>	0,09	0,10	0,07	0,28	0,22	0,34	–	0,23	0,21	0,12	0,12	0,09	0,10
<i>Aphanocapsa</i> 6714	0,11	0,09	0,09	0,26	0,20	0,26	0,23	–	0,31	0,11	0,11	0,10	0,10
Chloroplast (Lemna)	0,08	0,11	0,06	0,21	0,19	0,20	0,21	0,31	–	0,14	0,12	0,10	0,12
<i>Methanobacterium thermoautotrophicum</i>	0,11	0,10	0,10	0,11	0,06	0,11	0,12	0,11	0,14	–	0,51	0,25	0,30
<i>M. Ruminantium</i> strain M-1	0,11	0,10	0,10	0,12	0,07	0,13	0,12	0,11	0,12	0,51	–	0,25	0,24
<i>Methanobacterium sp.</i> , Cariaco isolate JR-1	0,08	0,13	0,09	0,07	0,06	0,06	0,09	0,10	0,10	0,25	0,25	–	0,32
<i>Methanosarcina barkeri</i>	0,08	0,07	0,07	0,12	0,09	0,12	0,10	0,10	0,12	0,30	0,24	0,32	–

Os dados recentes do seqüenciamento do genoma humano mostram que todos os seres humanos são incrivelmente semelhantes – 99,9%, geneticamente falando. Somos todos de uma única espécie, a *Homo sapiens*. Não há nenhuma base genética ou biológica para estabelecer diferenças entre os seres humanos. *As pessoas aparentadas são ainda mais similares*. Isto porque as crianças herdam

o genoma de seus pais, metade do pai e metade da mãe. Quanto mais próximo o parentesco, mais semelhante a seqüência de DNA do genoma.

Ao se reassociar o DNA após desnaturação, parte do DNA se reassocia rapidamente e outra parte tem reassociação mais lenta. Esse fenômeno está relacionado à presença de seqüências denominadas seqüências repetitivas, quanto maior o número de seqüências repetitivas, mais rápida é a reassociação. As seqüências que demoram mais para se reassociar são as seqüências presentes um número menor de vezes no genoma, as seqüências moderadamente repetitivas e as seqüências únicas.

O DNA de seqüências únicas possui complementaridade com o RNA mensageiro, sendo considerado o DNA codificante e sua relação com a informação para proteínas está bem caracterizada. Outras seqüências, denominadas moderadamente repetitivas, também contêm informações para RNAs funcionais como o RNA ribossômico (rRNA) e o RNA transportador (tRNA). A função biológica do DNA altamente repetitivo ainda é desconhecida, paradoxalmente em alguns organismos esse tipo de DNA pode representar até 90% do DNA genômico. Desta forma, o conteúdo de DNA não está relacionado diretamente à complexidade do organismo (veja a Figura 1.10).

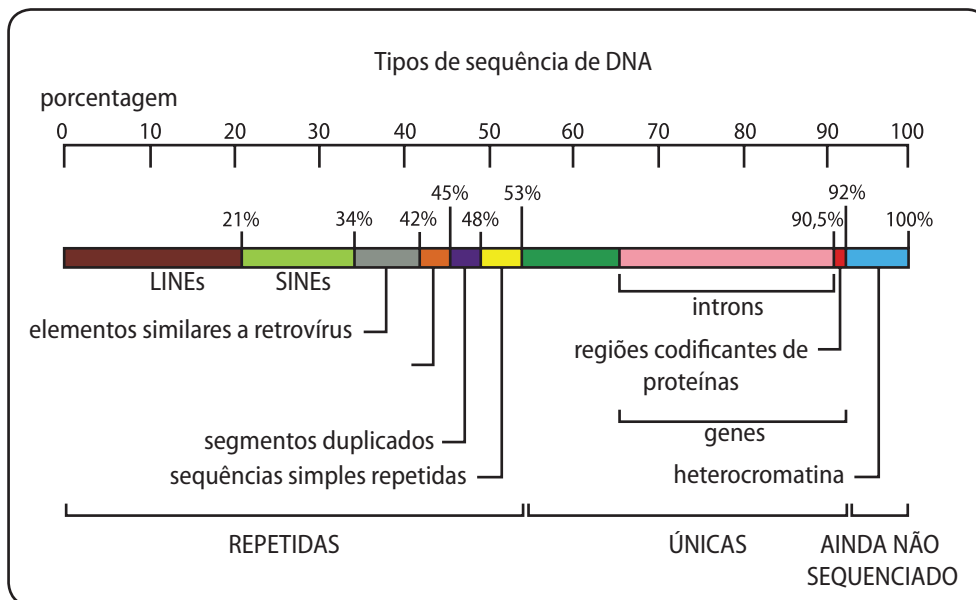


Figura 1.10 - Tipos de seqüência do DNA

Na verdade, à medida que os diferentes tipos de seqüências de DNA foram sendo caracterizados, constatou-se que nos eucariotos apenas uma pequena parte do DNA (~1,5%) codificava efetivamente proteínas.

Além do DNA genômico, outras estruturas celulares também contêm DNA, o que corrobora a teoria endossimbionte. O DNA é encontrado nas mitocôndrias e nos cloroplastos, mais recentemente também foi reconhecida a presença de DNA nos centríolos de alguns protozoários, o kDNA.

A árdua tarefa de caracterizar a seqüência de DNA começou a ser amenizada com a descoberta das **enzimas de restrição**.

Essas enzimas obtidas de bactérias que se mostravam capazes de restringir infecções virais têm a capacidade de produzir clivagens em seqüências específicas do DNA. Com essas enzimas era possível fragmentar o DNA de forma reprodutível. A amplitude dessa fer-

### Descoberta das enzimas de restrição

Na década de 1960, um bioquímico suíço chamado Werner Arber deu uma contribuição inesperada à tecnologia do DNA recombinante. Arber não estava interessado nas grandes questões referentes à base molecular da vida, e sim num aspecto enigmático da história natural dos vírus. Ele estudara o processo pelo qual alguns DNA virais são destruídos depois de inseridos em células bacterianas hospedeiras. Ele observou que algumas células hospedeiras identificavam certos DNA virais como estranhos e seletivamente o atacavam. As enzimas de restrição, ou endonucleases como também são chamadas, são naturalmente produzidas em algumas bactérias com a finalidade de protegê-las contra infecção viral. Quando um vírus injeta sua molécula de DNA na bactéria a fim de realizar seu ciclo reprodutivo, algumas vezes a bactéria reconhece esse DNA como estranho e destrói essa molécula utilizando as enzimas de restrição para cortá-la em inúmeros fragmentos. O nome dessas enzimas refere-se à sua função, uma vez que elas **restringem**

a multiplicação do vírus que poderia interferir no funcionamento da bactéria.

Mas, se essas enzimas são capazes de destruir DNA viral, por que não degradam também o DNA da bactéria, uma vez em que todo o mundo natural o DNA é a mesma molécula básica, não importando se está em bactéria, vírus, planta ou animal? O que impedia uma bactéria de atacar o seu próprio DNA ao mesmo tempo que investia contra o DNA do vírus?

Essa foi a segunda grande descoberta de Arber. Ao produzir uma enzima de restrição que visa a uma seqüência específica que normalmente varia de 4 a 6 bases, a bactéria também produz uma segunda enzima que modifica quimicamente a **mesma** seqüência em seu **próprio** DNA. Toda vez que aquela seqüência ocorrer, essa seqüência aparece ligada a uma molécula protetora, na verdade, um grupo metil (CH<sub>3</sub>). Isso impede que a enzima de restrição atue no DNA da bactéria, prevenindo sua destruição, enquanto a enzima de restrição pode atacar livremente o DNA do vírus.

ramenta foi sensacional, sendo possível inferir diferenças e trechos da seqüência de DNA. Os RFLP (polimorfismos de tamanho de fragmentos de restrição – Restriction Fragment Length Polimorphism) e as metodologias congêneres permitiram identificar diferenças individuais nos genomas, com diversas aplicações.

As diferenças individuais não estão limitadas à seqüência do DNA, mas também ao número de repetições de algumas seqüências.

Basicamente existem dois tipos de variações no DNA de indivíduos normais: aquele que afeta um simples par de bases e o que envolve seqüências repetidas. Na primeira a alteração de um par de bases na seqüência do DNA pode criar ou abolir um sítio de restrição em um determinado loco do genoma. Na segunda, a variabilidade resulta na observação do número de seqüências repetidas uma do lado da outra, *em tandem*, por exemplo, (AATGCGGTACTGAGCC)<sub>n</sub>, repetidas em números diferentes em cada indivíduo, dando-lhe uma característica única e gerando inúmeros alelos em função do número de vezes (n) em que essa estrutura é repetida ao longo daquele loco.

Algumas dessas enzimas produzem clivagens assimétricas nas duas fitas, gerando extremidades denominadas coesivas, por possuírem um pequeno segmento de fita simples (veja a Figura 1.11).

Uma vez que não existem diferenças químicas entre o DNA de diferentes organismos, verificou-se que esses fragmentos de DNA poderiam ser unidos por uma outra enzima também purificada das células, a DNA ligase. Essa metodologia foi denominada tecnologia do DNA recombinante, e por meio dela foi possível criar moléculas de DNA quiméricas. Inicialmente houve um temor que essa metodologia pudesse ser uma ameaça, pelo perigo potencial que poderia representar, além de aspectos morais e éticos de tal manipulação. Essas moléculas uma vez inseridas em células podem ser armazenadas, sendo criadas diversas bibliotecas de DNA

A identificação genômica possibilita detectar diferenças genéticas entre as espécies ou entre indivíduos da mesma espécie. Desta forma, a análise do DNA para a identificação dos indivíduos baseia-se no fato de que cada ser humano tem sua composição genética única, com exceção dos gêmeos univitelinos. Por outro lado, o DNA de um dado indivíduo é exatamente igual em qualquer célula do seu corpo, quer tenha sido extraído da raiz do cabelo, do sangue ou do esperma. A identificação genômica encontra aplicações em diversas áreas como determinação de paternidade duvidosa, identificação de suspeitos de crimes, identificação celular em transplantes de medula e pesquisa de populações humanas antigas e de animais extintos, fornecendo dados importantes para estudos arqueológicos e evolutivos.

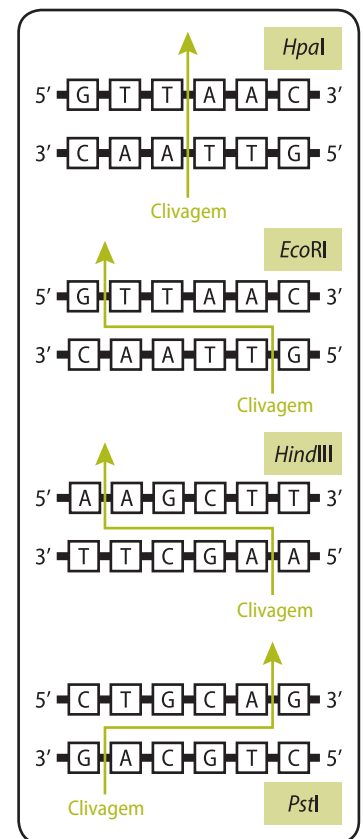


Figura 1.11 - Clivagem na molécula de DNA produzida por diferentes enzimas de restrição

com moléculas selecionadas e também fornecidas informações adicionais, possibilitando uma abordagem funcional inovadora que produziu desdobramentos fantásticos.

Mas é interessante notar que, à medida que esse conhecimento é construído, são desenvolvidas diversas aplicações que têm proporcionado grandes benefícios e também muitos temores. Sempre houve polêmica em relação aos aspectos éticos e morais dessa manipulação. Esse temor é sempre justificável em se tratando do tipo de sociedade em que vivemos, onde a ganância pelos lucros transcende qualquer princípio do bom senso.

### Obtenção de insulina por DNA recombinante

A engenharia genética promete acelerar o processo de melhoramento dos produtos agrícolas e de animais domésticos que, durante séculos, tem sido desenvolvido pelo homem por meio de cruzamentos dirigidos e da seleção dos descendentes. Essas situações ilustradas podem aperfeiçoar mecanismos biológicos já existentes na natureza ou mesmo criar novas alternativas úteis ao homem, como no caso de substâncias de uso terapêutico produzidas por engenharia genética.

Um exemplo disso é a insulina, a primeira droga aprovada para ser produzida em escala industrial por meio de técnicas de engenharia genética. A insulina é um hormônio muito importante produzido nas células do pâncreas e excretado na corrente sanguínea que regula o metabolismo do açúcar.

Um dos tipos de diabetes (tipo I) ocorre devido a uma incapacidade das células do pâncreas de produzir insulina ou produzir em insuficiência (tipo II), entretanto, os sintomas da doença podem ser evitados com injeções diárias desse hormônio. Antes do advento da engenharia genética, a insulina era obtida a partir de pâncreas bovino ou suíno. Embora a insulina animal seja biologicamente ativa em humanos, sua seqüência de aminoácidos é levemente diferente, o que pode provocar reações imunes nos pacientes. Entretanto, a insulina produzida em um sistema de expressão a partir de clonagem do gene humano não apresenta tais efeitos. A insulina produzida por engenharia genética encontra-se no mercado desde 1984 e, embora não apresente grandes vantagens em relação aos custos se comparada com outras formas de produção, para o paciente é muito mais adequada do que a insulina de origem animal. São estimados oito milhões de diabéticos somente nos Estados Unidos. A insulina se tornou um paradigma da indústria da biotecnologia, quando foi anunciada a sua obtenção por meio do DNA recombinante, as ações da Genetech subiram de US\$ 35 para US\$ 89, em poucos minutos, sendo a mais rápida elevação da história de Wall Street.

Por outro lado, atingem outro símbolo, que é a força vital, soterrando o vitalismo, razão pela qual surgem imensas polêmicas éticas e religiosas. Essa mudança de paradigma coloca o DNA no centro das atenções, emergindo um conceito unificador na Biologia a partir do qual diversas questões poderiam ser esclarecidas.

## 1.5 Organização Celular

O DNA celular possui uma organização adicional que permite a estabilização da molécula. Nos procariotos o genoma consiste de uma molécula de DNA circular, podendo existir moléculas adicionais menores denominadas plasmídios, também circulares. Os plasmídios em geral conferem propriedades adicionais às células que os possuem. Essas estruturas consistem numa poderosa ferramenta para a manipulação do DNA.

Nos eucariotos, a maior parte do DNA está compartimentalizada no núcleo da célula, consistindo de moléculas lineares organizadas em estruturas denominadas **cromossomos**, cada espécie possui uma constituição cromossômica característica (cariótipo).

Durante o ciclo celular o DNA é organizado no núcleo em diversos níveis de compactação, atingindo o máximo de compactação na divisão celular, numa fase denominada metáfase, quando os cromossomos individualizados podem ser visualizados ao microscópio. Nas demais fases os cromossomos não são vistos individualmente, mas mantêm um padrão de organização, ocupando regiões específicas no núcleo denominadas territórios cromossômicos.

O primeiro nível de organização consiste no nucleossomo, onde uma estrutura composta de oito unidades de proteínas básicas denominadas **histonas** formam um arranjo em que o DNA se enovela. Esse conjunto denominado **nucleossomo** é formado pelas histonas H2a, H2b, H3 e H4 e um segmento de DNA de cerca de 200 pares de bases. A seguir essas estruturas são organizadas em forma de solenóide por meio da histona H1, formando uma fibra de cromatina de cerca de 30 nm (nanômetros) de espessura. Esse padrão de organização é bastante conservado evolutivamente, assim como as histonas (veja a Figura 1.12). Os níveis subseqüentes

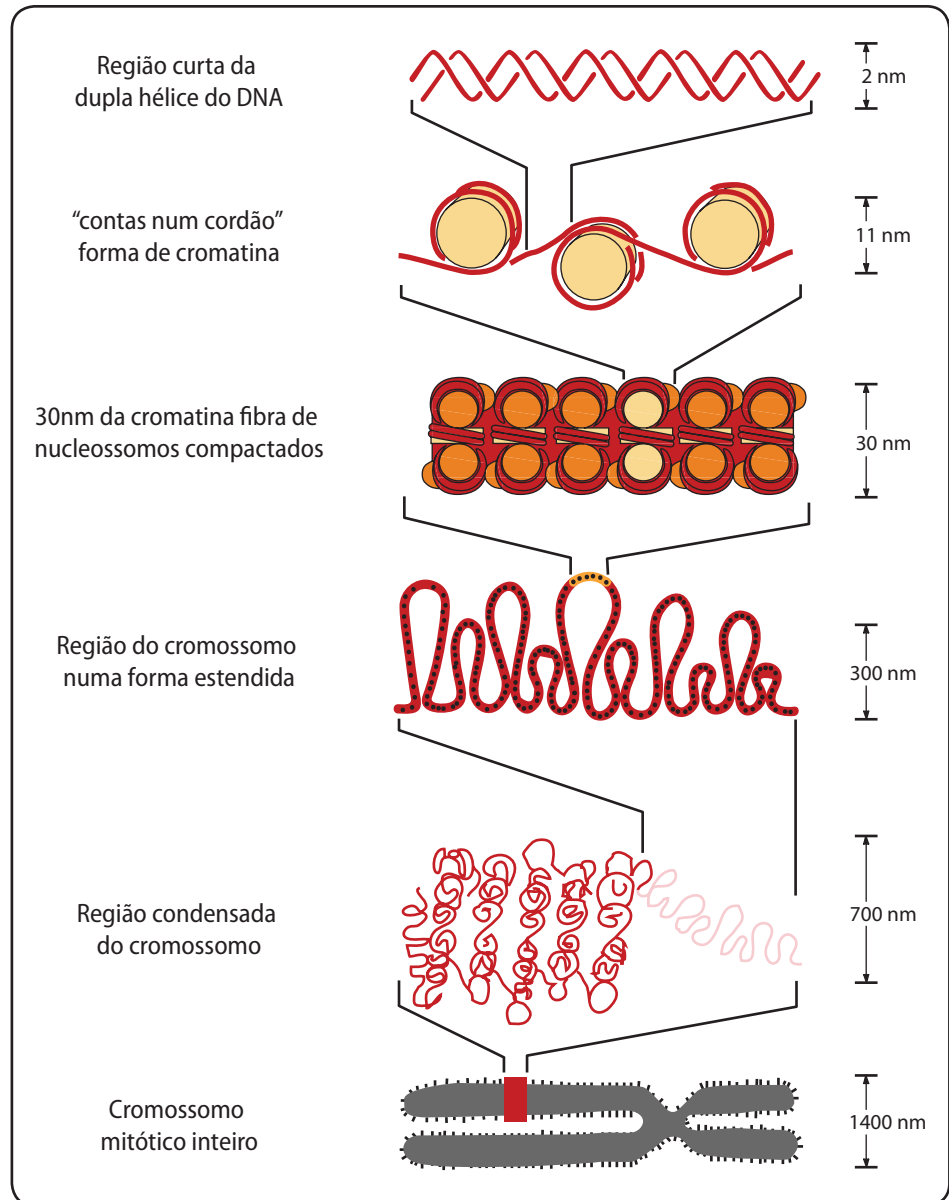


Figura 1.12 - Organização do DNA nos cromossomos

de compactação ainda não são totalmente compreendidos, mas envolvem a espiralização dessa fibra e sua ligação a uma matriz cromossômica constituída majoritariamente de **topoisomerases**.

## 1.6 Funções Biológicas

Ainda não estava esclarecido como o DNA exerce sua influência hereditária. É crucial estabelecer como o material hereditário, seja qual for a sua natureza, pode influenciar as características da célu-

la, como a maquinaria celular decodifica as mensagens das moléculas de DNA. A questão da continuidade da informação pode ser inferida do modelo. Foi natural sugerir que a seqüência de bases do DNA seria de alguma forma um código para a seqüência de aminoácidos nas cadeias polipeptídicas.

A partir desse modelo da estrutura do DNA, diversos aspectos desse funcionamento foram sendo esclarecidos e novas particularidades evidenciadas, transformando o DNA num ícone da ciência contemporânea.

### **Funções do material genético:**

#### **1. Função de genótipo ou replicação**

O material genético deve ser capaz de armazenar a informação genética e transmitir essa informação fielmente dos pais para os descendentes, de geração após geração.

#### **2. Função de fenótipo ou expressão gênica**

O material genético deve controlar o desenvolvimento do fenótipo do organismo, seja esse um vírus, uma bactéria, uma planta ou um animal, ou seja, o material genético deve dirigir o crescimento e a diferenciação do organismo a partir do zigoto unicelular até o adulto maduro. Para controlar esse processo, o material genético deve não apenas se expressar acuradamente, mas cada gene deve agir no momento e no local precisos, garantindo que o fígado seja formado por células hepáticas, o sistema nervoso, por células nervosas e assim por diante.

#### **3. Função de adaptação ou mutação e recombinação**

O material genético deve possibilitar plasticidade ao genoma, permitindo a variabilidade genética e, portanto, a capacidade de adaptação ao ambiente. O DNA é uma molécula reativa, sendo alvo de alterações espontâneas e induzida por agentes físicos, químicos e biológicos que produzem alterações na molécula do DNA. Os sistemas de reparo do DNA atuam para corrigir essas alterações, preservando a integridade do genoma. Os mecanismos de recombinação possibilitam o reagrupamento da informação, gerando novas combinações de seqüências codificantes e individualidade aos organismos.



## Resumo

Após a redescoberta das leis de Mendel, no início do século XX, o grande desafio no meio científico era a identificação da composição química dos genes.

Devido ao fato de os cromossomos serem constituídos por ácidos nucléicos e proteínas, postulava-se que uma dessas macromoléculas poderia ser a matéria-prima dos genes. Como o DNA é uma molécula extremamente repetitiva e estável, enquanto as proteínas apresentam uma grande variedade estrutural, muitos pesquisadores acreditavam que eram as proteínas que carregavam a informação genética. Experimentos realizados por Griffith, Avery e McLeod durante as décadas de 1940 e 1950 mostraram elegantemente que o DNA é o material genético das células. Em 1953, Watson e Crick apresentaram a estrutura de dupla hélice do DNA baseada em dados de difração de raios X.

As características peculiares da estrutura da molécula de DNA, como a regularidade de sua estrutura, a seqüência linear de bases e o pareamento de bases entre as duas fitas componentes da molécula estão associadas à sua função biológica.

A similaridade estrutural entre o DNA de todos os seres vivos e a possibilidade de variabilidade no conteúdo e na seqüência de nucleotídeos é um conceito unificador na Biologia, tornando-se a caracterização e a análise das seqüências uma ferramenta indispensável para estudos filogenéticos e funcionais e possibilitando, também, a manipulação do DNA e a construção de moléculas híbridas com DNA proveniente de diferentes organismos.

Nos eucariotos, o DNA está organizado em estruturas denominadas cromossomos. Esse padrão de organização é conservado evolutivamente.

No genoma dos eucariotos são identificadas seqüências únicas e seqüências repetitivas. Enquanto o DNA de seqüências únicas contém informações para a construção de proteínas, a função do DNA de seqüências repetitivas ainda não é conhecida, recebendo a denominação de DNA lixo.

## Referência Comentada

### O DNA: uma sinopse histórica

*Talles Henrique Gonçalves de Oliveira, Neusa Fernandes dos Santos e Leila Maria Beltramini*

Uma revisão cuidadosa e bem elaborada, compreendendo os fatos históricos científicos mais importantes que permitiram o entendimento atual da estrutura e da funcionalidade da molécula de DNA.

OLIVEIRA, T. H. G., SANTOS, N. F. D., BELTRAMINI, L. M. O DNA: uma sinopse histórica. **Revista Brasileira de Ensino de Bioquímica e Biologia Molecular**, Artigo 1, Edição 01/2004, 24 fev. 2004. Disponível em: <<http://www.ib.unicamp.br/lte/bdc/visualizarMaterial.php?idMaterial=153>>. Pág. A1-A16.

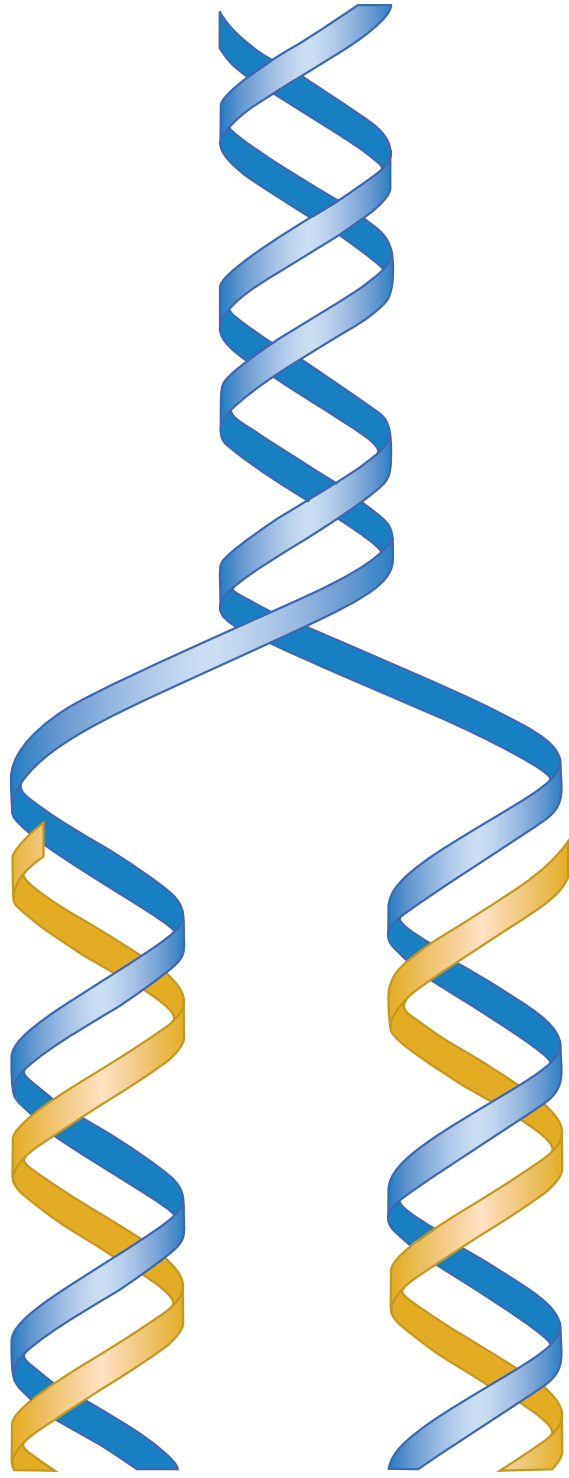
## Referências Bibliográficas

ALBERTS, B. et al. **Molecular Biology of the Cell**. 4<sup>th</sup>. ed. New York: Garland Pub. Inc., 2002.

WATSON, J. D. et al. **Biologia molecular do gene**. 5. ed. Porto Alegre: Artmed, 2006.



# CAPÍTULO 2



## Replicação do DNA

*Neste capítulo, vamos reconhecer a estrutura de duas cadeias do DNA, com as bases complementares pareadas como um mecanismo lógico de replicação, a dinâmica e os componentes da maquinaria celular envolvidos nessa atividade. Compreender como o material genético é capaz de armazenar a informação genética e transmitir essa informação fielmente dos pais aos descendentes, de geração após geração, e as peculiaridades desse processo.*



Para ver a divisão celular no microscópio de contraste de fase, acesse: <<http://youtube.com/watch?v=DD3IQknCEd>> ou <<http://youtube.com/watch?v=s1yUTbXyWU>>.

Visto ao microscópio, o **processo de mitose**, pelo qual uma célula se divide, e se torna duas, é um dos espetáculos mais fascinantes de toda a biologia. Ninguém que observe o fenômeno desenrolado num filme em câmera acelerada pode deixar de sentir-se excitado e temeroso. Como demonstração dos poderes de organização dinâmica de que é dotada a matéria viva, o ato de divisão é assaz expressivo, porém ainda mais perturbador é o aparecimento de dois conjuntos idênticos de cromossomos onde antes existia apenas um (CRICK, 1954).

Os organismos vivos se perpetuam pela reprodução. A reprodução pode ser uma simples duplicação (fissão celular) como em bactérias ou envolver mecanismos complexos de reprodução sexual como ocorre nas plantas superiores e nos animais. Em todos os casos, a reprodução envolve a transmissão fiel da informação genética armazenada no DNA.

A replicação do DNA ocorre num momento preciso do ciclo celular denominado fase S, que sucede outra fase denominada G1. Ainda não são claros os processos que atuam na transição da fase G1 e S. A consequência da replicação do DNA é a divisão celular, quando são geradas novas células-filhas com o mesmo conteúdo genômico da célula precursora (veja a Figura 2.1).

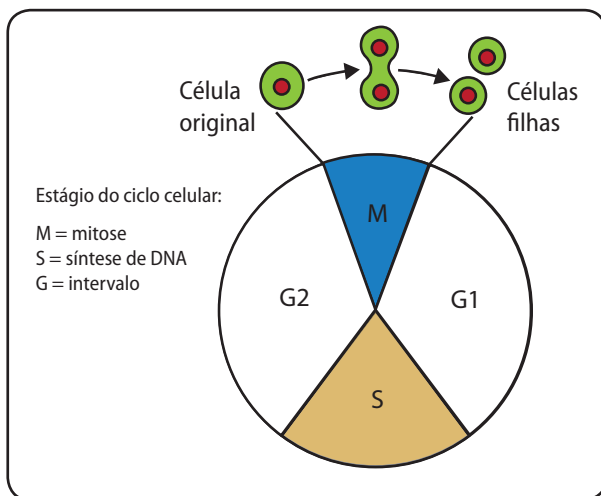


Figura 2.1 - Ciclo celular

## 2.1 A Replicação do DNA é central para toda a Biologia

Uma das características do modelo de Watson e Crick é permitir uma maneira de replicação do DNA aparentemente simples.

Anteriormente, imaginava-se que a replicação deveria envolver dois estágios. No primeiro, a molécula deveria criar uma forma intermediária ou um “negativo”. A seguir, o molde negativo seria utilizado para criar uma nova molécula.

A estrutura de duas cadeias do DNA, com as bases complementares pareadas ao longo da molécula, significa que ela possui o seu próprio negativo e encurtou em um estágio o ciclo de replicação hipotético.

Com a demonstração de que as cadeias podem se separar e cada cadeia permanecer intacta, a maquinaria química celular não teria muita dificuldade em reunir uma nova cadeia, pareando a base apropriada ponto por ponto ao longo de cada cadeia simples, criando uma cadeia complementar e produzindo duas cadeias onde existia apenas uma (veja a Figura 2.2).

Esse modo de replicação tem uma implicação dramática: embora cada cadeia da dupla hélice se separe durante a mitose e construa uma nova cadeia, nenhuma das cadeias originais é destruída. Em todas as divisões celulares que ocorrem desde a concepção para criar um ser humano a partir de uma célula-ovo fertilizada, as cadeias originais dos cromossomos de cada um dos pais nunca são destruídas, apenas separadas e reformadas diversas vezes. Algumas células do organismo adulto ainda podem conter as moléculas de DNA originais herdadas dos pais, não apenas cópias, mas os mesmos átomos.

Esse raciocínio lógico pode ser viabilizado tecnicamente de formas diferentes. Na replicação **conservativa**, a cadeia dupla total

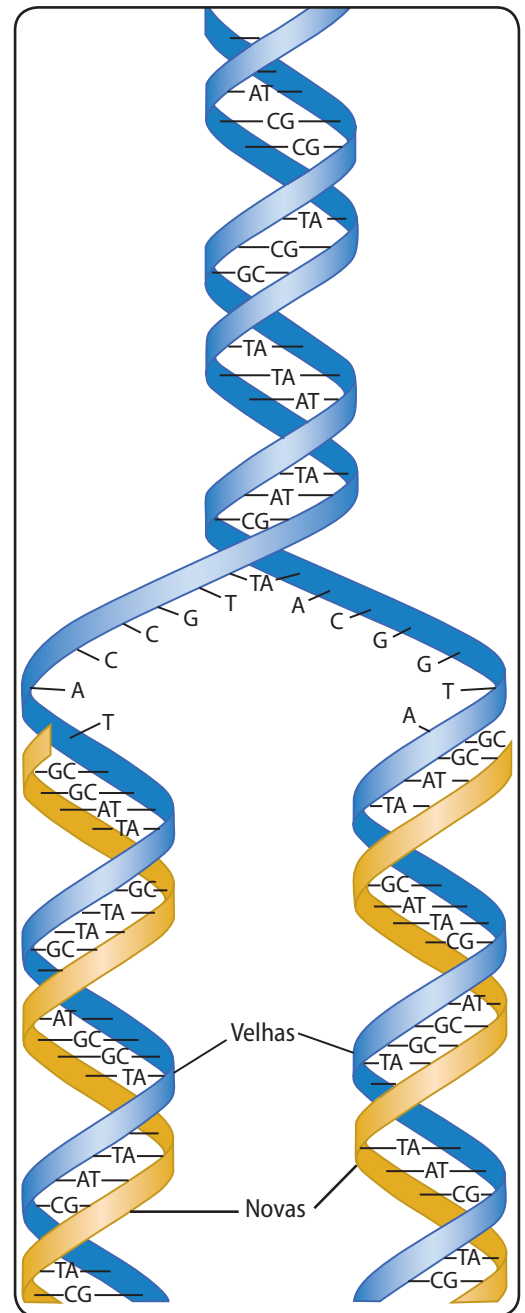


Figura 2.2 - Replicação do DNA: As duas cadeias da dupla hélice parental se separam, e cada uma serve de molde para a síntese de uma nova cadeia pelas regras do pareamento das bases.



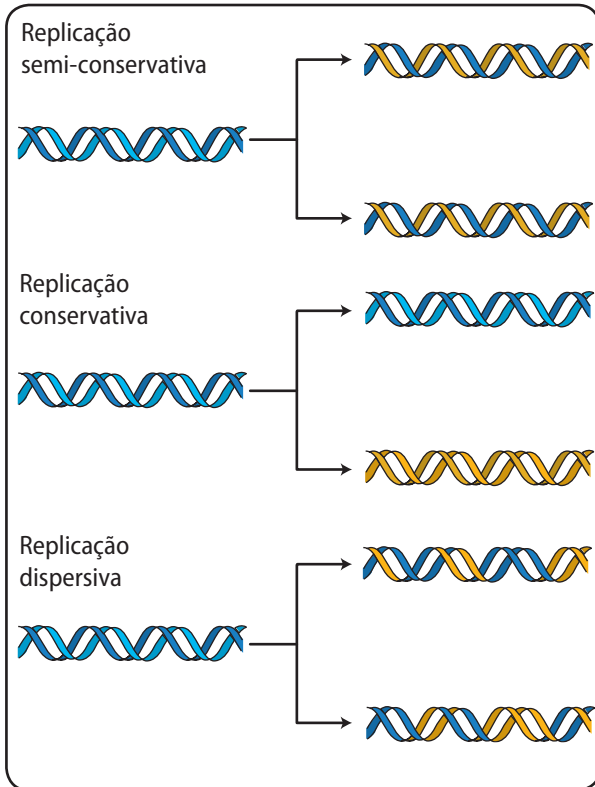


Figura 2.3 - Possíveis mecanismos de replicação do DNA

seria copiada e uma célula-filha herdaria a nova cópia e a outra célula-filha herdaria a molécula original. Outra forma possível seria a replicação **dispersiva**, na qual a molécula original seria fragmentada em pequenas unidades antes de ser copiada, com os segmentos da molécula original sendo divididos entre as células-filhas. Seria também possível a replicação semiconservativa, quando uma das cadeias contém a fita original e a cadeia complementar é recém-sintetizada (veja a Figura 2.3).

Foram Matthew Meselson e Franklin Stahl que, utilizando uma metodologia inovadora, conseguiram distinguir essas possibilidades. A primeira etapa foi encontrar uma maneira de distinguir as cadeias originais do DNA em uma colônia de células em crescimento das cadeias novas construídas a partir dos precursores – o alimento que as células precisam para crescer.

A alternativa escolhida foi alimentar as células com um isótopo pesado (radioativo) de um dos elementos essenciais. Após várias alternativas optaram pelo nitrogênio pesado, o  $^{15}\text{N}$ . As células escolhidas para esse estudo foram de *E. coli*. A segunda parte do experimento foi medir as diferenças de densidade entre as moléculas de DNA das diferentes gerações celulares. Isto foi possível utilizando centrifugação em gradiente de densidade de cloreto de céσιο (veja a Figura 2.4).

***E. coli***

A *Escherichia coli* é uma bactéria bacilar Gram-negativa, que, juntamente com o *Staphylococcus aureus*, é a mais comum e uma das mais antigas bactérias simbiotes do Homem.

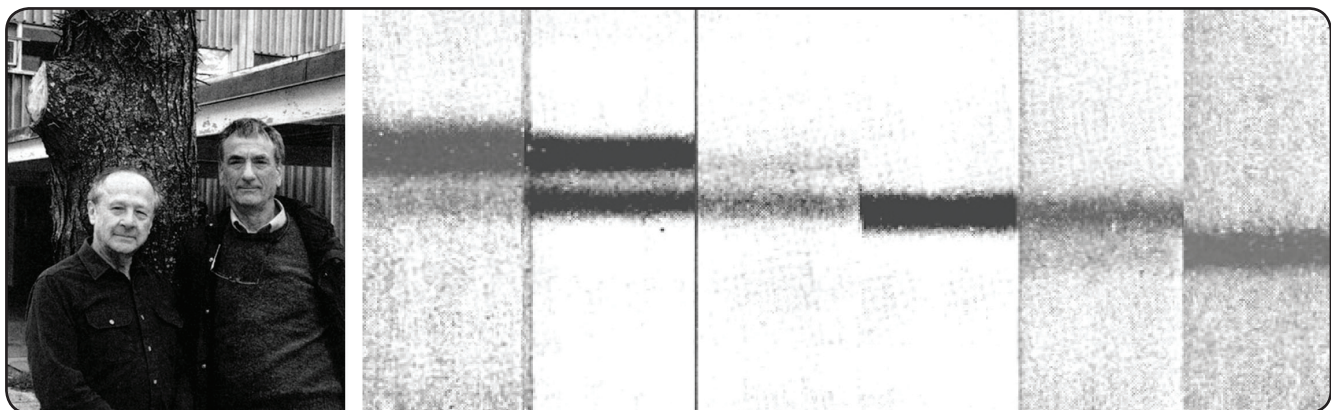


Figura 2.4 - Experimento de Meselson e Stahl demonstrando a replicação semiconservativa

Quando tudo estava pronto, o experimento foi relativamente simples: primeiro uma colônia de *E. coli* foi cultivada em meio contendo  $N^{15}$ . Após várias gerações, todas as moléculas de DNA das bactérias vivas na colônia conteriam, entre outras coisas, nitrogênio pesado e, quando analisadas pela técnica da ultracentrifugação, apresentavam uma banda escura no fundo do tubo. A seguir as colônias de bactérias foram removidas do meio contendo nitrogênio pesado e alimentadas com o isótopo ordinário,  $^{14}N$ . O resultado foi novamente uma banda escura, mas numa posição que corresponde à metade da densidade entre o DNA marcado e o DNA normal. A única explicação seria de que na nova geração de *E. coli* a molécula de DNA consiste de uma cadeia pesada herdada diretamente do ancestral e de uma cadeia leve manufaturada com o nitrogênio leve. Na segunda geração, como previsto pela replicação semiconservativa, foram observadas duas bandas nos tubos, correspondendo aos dois tipos de DNA, um híbrido como o DNA da primeira geração e outro formado apenas com nitrogênio leve (veja a Figura 2.5).

### Organismo modelo *Escherichia coli*

Essa bactéria é um organismo modelo não só para a genética de procariontes mas de certo modo para toda a genética, por apresentar muitas características adequadas para a pesquisa genética, uma das quais é a sua fácil obtenção, pois vive no tubo digestivo dos humanos e de outros animais de sangue quente. A *E. coli* tem um único cromossomo circular com 4,6 megabases de tamanho. Seus 4.000 genes não contêm íntrons e cerca de 35% são de função desconhecida.

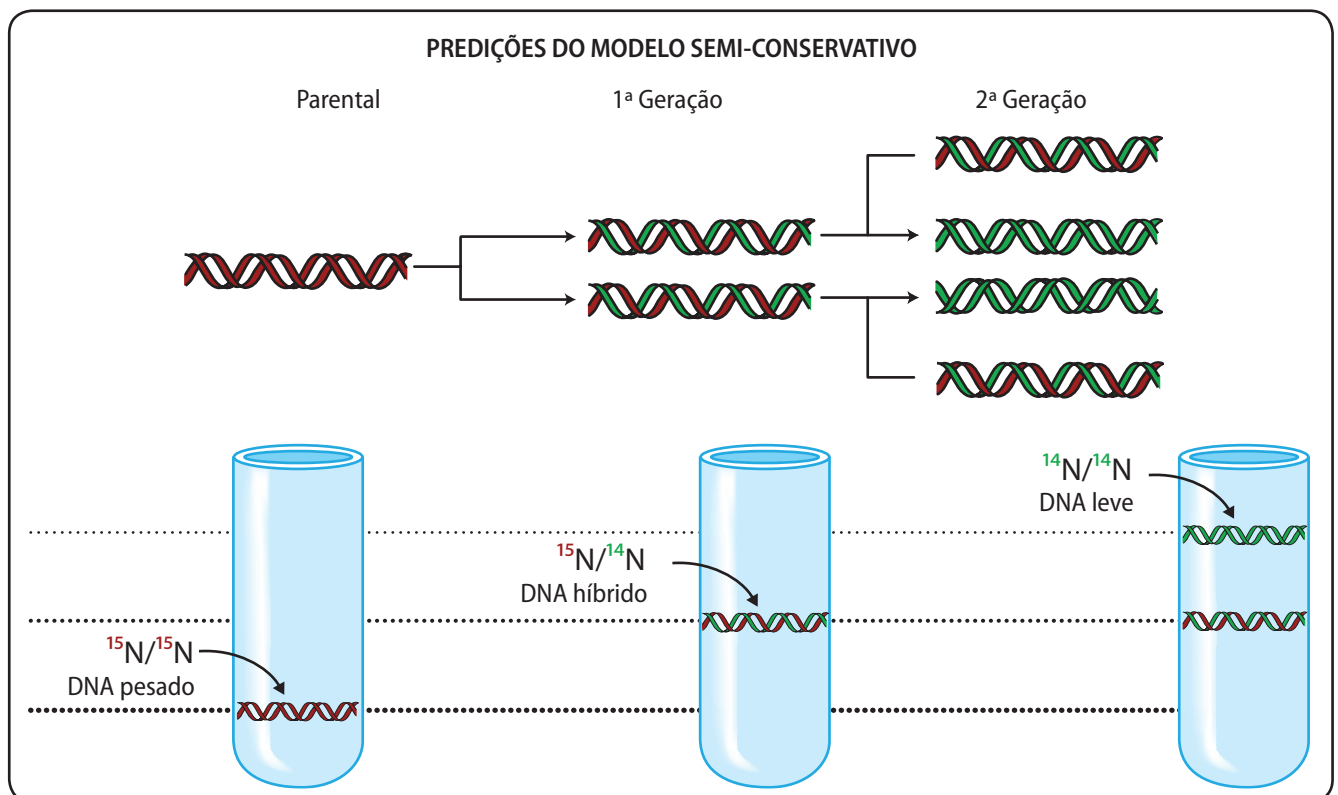


Figura 2.5 - Constituição das moléculas do DNA após a replicação semiconservativa

## 2.2 Visualização Celular da Replicação: Autorradiografia

A visualização dos cromossomos replicando foi possível por meio de uma técnica chamada autorradiografia. Esse método permite localizar isótopos radioativos em preparações citológicas ou macromoléculas pela exposição a uma emulsão fotográfica sensível à radiação de baixa energia. A autorradiografia é particularmente útil no estudo do metabolismo do DNA porque o DNA pode ser marcado especificamente nas células em divisão com timidina  $^3\text{H}$ , o deoxirribonucleotídeo da timina. A timidina é incorporada exclusivamente no DNA, não estando presente nos outros constituintes principais da célula.

Desta forma, no experimento realizado por Cairns (1963), as células de *E. coli* foram cultivadas em meio contendo timidina  $^3\text{H}$  por diferentes períodos, a seguir as células foram cuidadosamente lisadas de forma a não quebrar os cromossomos e esses foram coletados em

membranas filtrantes. Essas membranas foram fixadas em lâminas de vidro e cobertas com emulsão sensível a partículas  $\beta$  (os elétrons de baixa energia emitidos durante o decaimento do trítio) e mantidas no escuro por um tempo suficiente para o decaimento radioativo. As autorradiografias observadas quando a emulsão foi revelada mostraram que os cromossomos de *E. coli* são estruturas circulares que existem na forma  $\theta$  intermediária durante a replicação. Essas autorradiografias sugeriram que o desemparelhamento das duas cadeias complementares parentais e a replicação semiconservativa ocorrem simultaneamente ou estão estreitamente associados (veja a Figura 2.6).

A interpretação de Cairns destas autorradiografias foi de que a replicação semiconservativa inicia em um sítio no cromossomo, o qual foi denominado origem, e prossegue seqüencialmente e unidirecionalmente ao

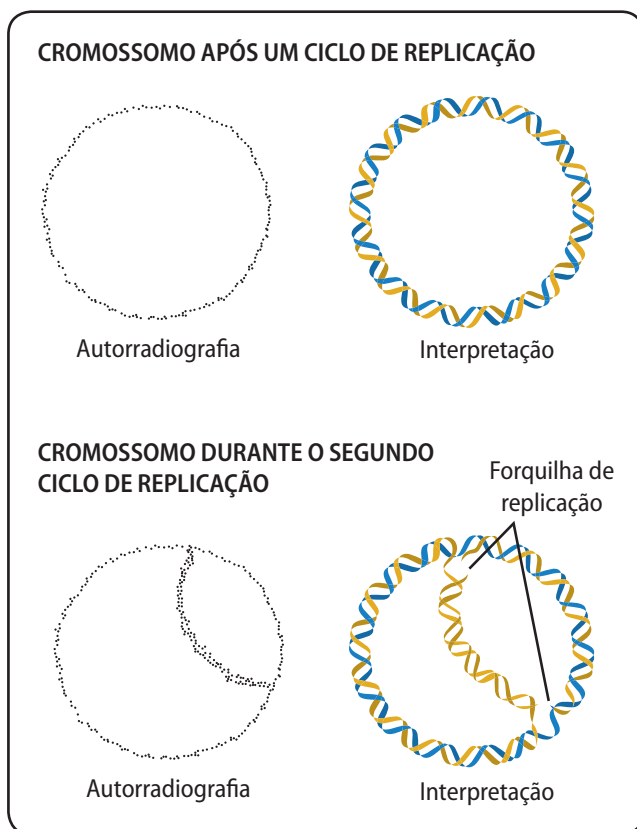


Figura 2.6 - Autorradiografia da replicação do DNA

redor da estrutura circular. No entanto, evidências subseqüentes demonstraram que sua interpretação estava incorreta em um ponto: a replicação ocorre bidirecionalmente, não unidirecionalmente. Cada estrutura em forma de Y é uma forquilha de replicação e as duas **forquilhas de replicação** se movem em direções opostas ao longo do cromossomo circular (veja a Figura 2.7).

### 2.3 Origem Única e Replicação Bidirecional

Os experimentos de Cairns não forneceram nenhuma evidência a respeito da **origem de replicação**, que poderia ocorrer em um sítio específico ou ao acaso no cromossomo. Posteriormente surgiram evidências diretas demonstrando que em *E. coli* e em diversos organismos a replicação ocorre bidirecionalmente a partir de uma origem específica.

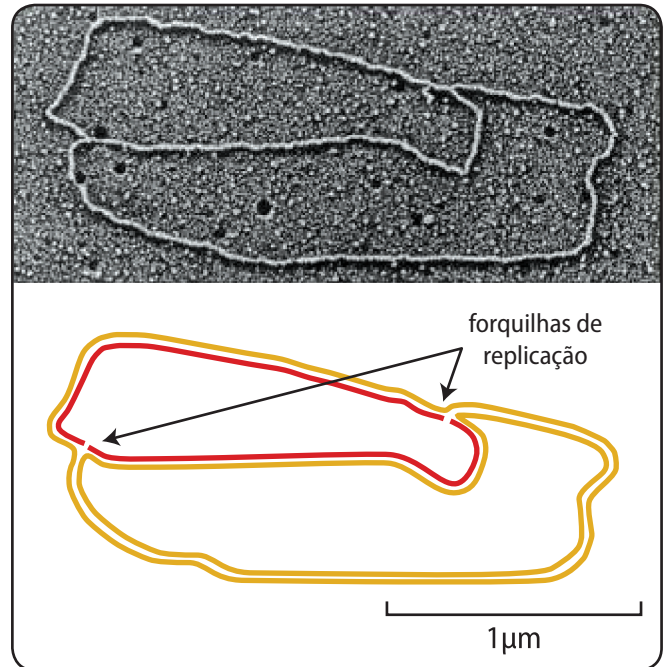


Figura 2.7 - Replicação do DNA bidirecional

### 2.4 DNA Polimerase e Síntese *in vitro* do DNA

As evidências citológicas são inconclusivas em relação ao processo bioquímico relativo a esses eventos.

Muito do que foi aprendido sobre os mecanismos moleculares envolvidos nos processos biológicos se deu por meio do fracionamento de células em várias organelas, macromoléculas e outros componentes e depois reconstituindo os sistemas em tubos de ensaio, os assim chamados sistemas *in vitro*. Esses sistemas *in vitro* podem ser dissecados bioquimicamente mais facilmente que os sistemas "*in vivo*".

A síntese *in vitro* do DNA foi obtida por Arthur Kornberg e colaboradores em 1957. Kornberg, que recebeu o prêmio Nobel em 1959 por esse trabalho, isolou uma enzima de *E. coli* (chamada DNA polimerase) que catalisa a adição covalente de nucleotídeos a cadeias preexistentes de DNA. Essa enzima requer 5'-trifosfatos

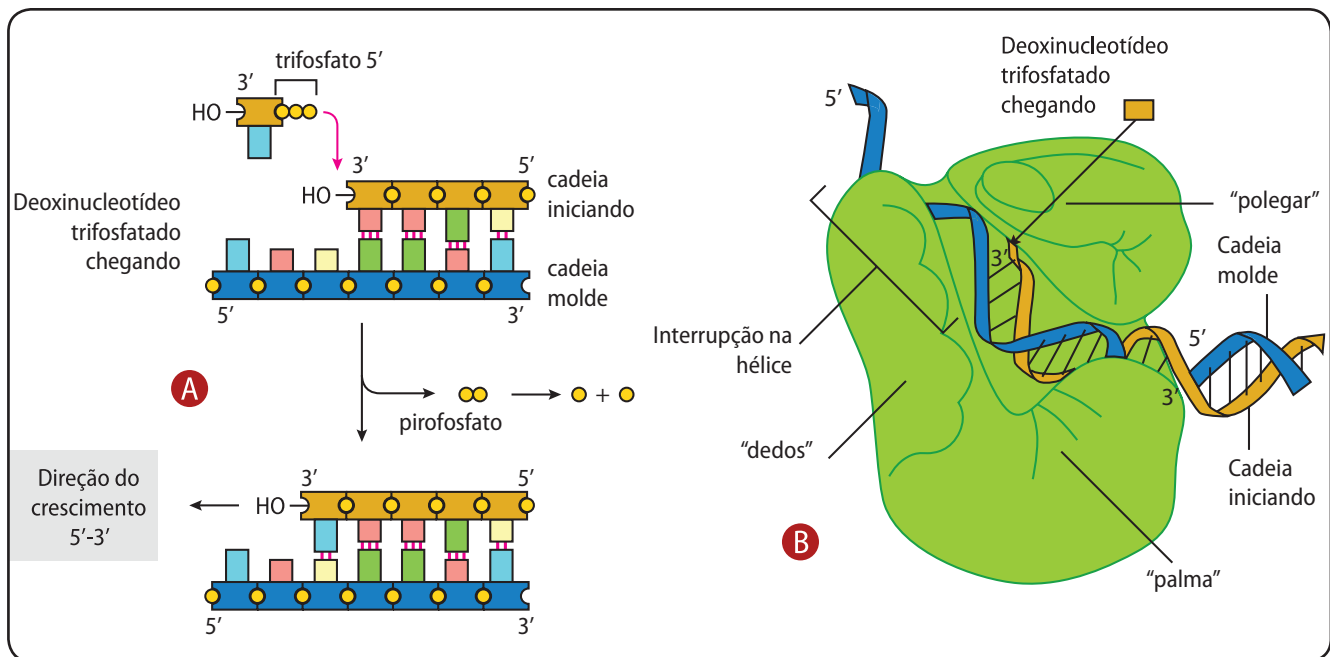


Figura 2.8 - (A) Reação catalizada pela DNA polimerase. (B) Modelo da DNA polimerase

dos 4 deoxinucleotídeos e é ativa apenas na presença de íons  $Mg^{++}$  e DNA preexistente. Esse DNA deve fornecer dois componentes essenciais para o processo, um para a função de iniciador (primer) e outro para a função de molde (template) (veja a Figura 2.8).

1. DNA iniciador: A DNA polimerase não pode iniciar a síntese da cadeia de DNA de novo. Tem **uma necessidade absoluta de um 3'-hidroxil livre em uma cadeia preexistente**. A DNA polimerase catalisa a formação da ligação fosfodiéster entre a extremidade 3'-OH do último nucleotídeo da cadeia do DNA e o 5'-fosfato do deoxinucleotídeo trifosfatado subsequente. A direção da síntese é sempre  $5' \rightarrow 3'$ . A energia para essa reação é proveniente da quebra do trifosfato.
2. DNA molde: A DNA polimerase não possui especificidade de seqüência, isto é, a enzima requer uma cadeia de DNA molde cuja seqüência, por meio das regras do pareamento de bases, dirige a síntese de uma seqüência complementar de bases na cadeia sintetizada. O pareamento incorreto causa diminuição da atividade catalítica da DNA polimerase, que também consegue distinguir os deoxinucleotídeos e os ribonucleotídeos trifosfatados.

Essa enzima, hoje melhor caracterizada, possui diversos sítios funcionais, apresentando uma alta fidelidade na replicação do DNA e uma alta processividade (hoje em dia são conhecidas DNA

polimerases que atuam em velocidades de até 100.000 nucleotídeos por segundo).

Após a descoberta de Kornberg, diversas DNA polimerases foram isoladas e caracterizadas a partir de diferentes organismos. A função precisa de algumas dessas DNA polimerases ainda não são claras. Todas as DNA polimerases até hoje identificadas atuam no sentido  $5' \rightarrow 3'$ . A presença de íons metálicos é imprescindível para a atividade da enzima.

## 2.5 O “Paradoxo do Ponto de Crescimento” e a Síntese Descontínua de DNA

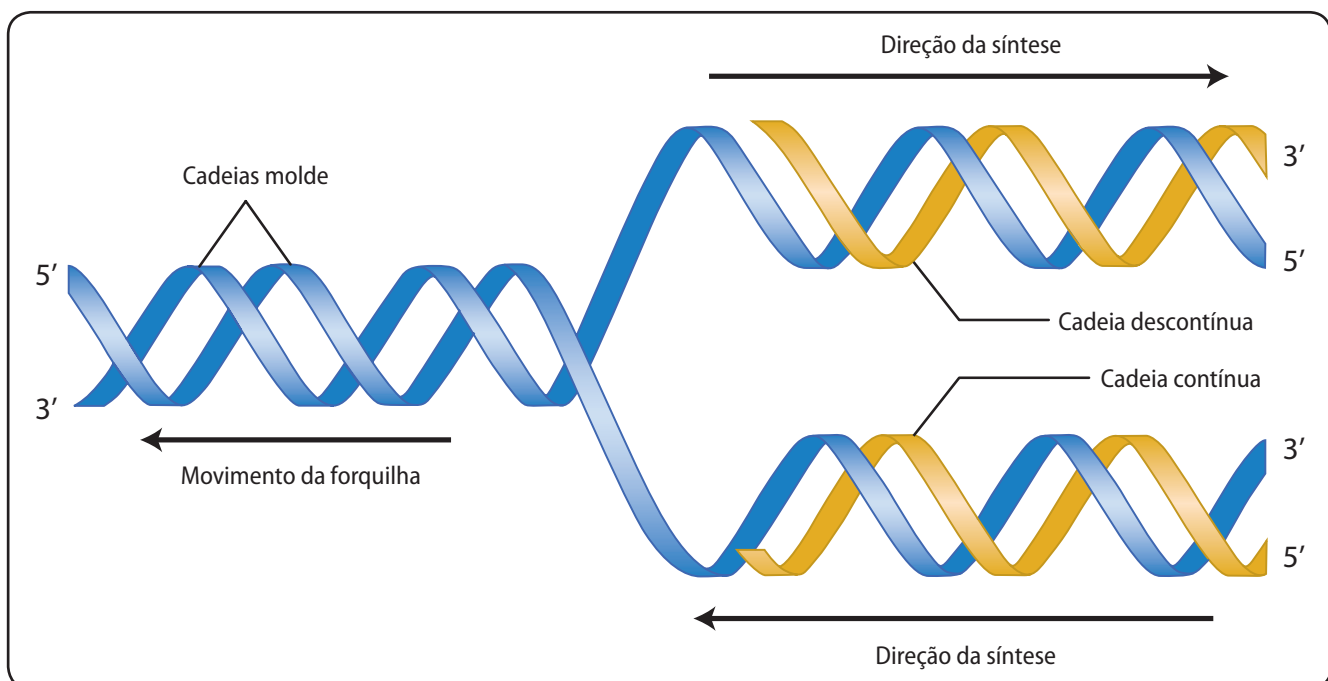
Os estudos de moléculas de DNA replicantes por meio das autorradiografias e da microscopia eletrônica indicam que as duas cadeias-filhas sintetizadas em cada forquilha de replicação se estendem na mesma direção, ao menos no nível macromolecular. Visto que as cadeias complementares da dupla hélice têm polaridades opostas, isto significa que a síntese está ocorrendo no sentido da terminação  $5'$  (ou  $3' \rightarrow 5'$ ) e da terminação  $3'$  (ou  $5' \rightarrow 3'$ ). No entanto, todas as DNA polimerases conhecidas têm necessidade absoluta de um  $3'$ -hidroxil livre, o que possibilita apenas a síntese  $5' \rightarrow 3'$ .

Esse paradoxo existiu por vários anos, enquanto os bioquímicos procuravam por polimerases capazes de realizar a síntese no sentido  $3' \rightarrow 5'$ . Nenhuma polimerase com essa capacidade foi até o momento encontrada. Ao contrário, fortes evidências foram se acumulando indicando que **toda síntese ocorre na direção  $5' \rightarrow 3'$** . A resolução desse paradoxo foi pela demonstração que a síntese de uma das cadeias é **descontínua** (veja a Figura 2.9).

Pela autorradiografia e pela microscopia eletrônica demonstrou-se que as cadeias nascentes do DNA sintetizadas em cada forquilha de replicação são estendidas na **mesma direção no nível macromolecular**. Entretanto, no nível molecular, a síntese ocorre em **direções opostas**. Ambas as cadeias são sintetizadas na direção  $5' \rightarrow 3'$ . As cadeias que evoluem na direção  $3' \rightarrow 5'$  crescem pela síntese de pequenos segmentos (sintetizados  $5' \rightarrow 3'$ ) e pela subsequente reunião desses pequenos segmentos pela enzima **DNA ligase**. A evidência

do modo **descontínuo** de replicação do DNA foi obtida de estudos nos quais intermediários da síntese de DNA foram detectados por autorradiografia após o cultivo das células durante períodos curtos em meio contendo timidina  $^3\text{H}$  (marcação de pulso). Quando células de *E. coli* foram marcadas por pulsos de 15 segundos, por exemplo, toda marcação foi encontrada em pedaços pequenos, com 1.000 a 2.000 nucleotídeos. Esses pequenos segmentos de DNA, frequentemente denominados **fragmentos de Okazaki**, são menores nos eucariotos, 100 a 200 nucleotídeos. Quando são utilizados períodos de marcação de pulso mais longos, a marcação é observada majoritariamente em moléculas grandes de DNA – provavelmente do tamanho de moléculas contendo todo o DNA presente em cromossomos intactos. Nos períodos curtos de marcação de pulso, a radioatividade presente nos pequenos “fragmentos” de DNA se tornam incorporados nas moléculas do tamanho dos cromossomos durante o crescimento subsequente das células em meio contendo timidina não radioativa. Isto é importante porque indica que os fragmentos de Okazaki são efetivamente intermediários na síntese de DNA e não um tipo de subproduto metabólico. Evidências inequívocas demonstram que a síntese de DNA é contínua na cadeia crescendo no sentido  $5' \rightarrow 3'$  (denominada cadeia “leading”) e descontínua na cadeia crescendo na direção  $3' \rightarrow 5'$  (denominada cadeia “lagging”).

Figura 2.9 - Evolução da replicação



## 2.6 Iniciação e Problema do “Primer”

Como dito anteriormente, todas as DNA polimerases conhecidas têm uma necessidade absoluta de uma extremidade 3'-OH, um primer e uma cadeia molde apropriada para a sua atividade. Assim, **não existe nenhuma DNA polimerase conhecida que possa iniciar a síntese de uma nova cadeia de DNA**. Visto que a síntese de cada fragmento de Okazaki requer um evento de iniciação, um mecanismo eficiente de iniciação de cadeia é essencial para a replicação do DNA. A RNA polimerase, uma enzima complexa que catalisa a síntese de moléculas de RNA a partir de moldes de DNA, é capaz de iniciar a síntese de novas cadeias de RNA em locais específicos do DNA. Quando isso ocorre, é formado um híbrido RNA-DNA no qual o RNA nascente é ligado por pontes de hidrogênio ao DNA. Desde que as DNA polimerases foram capazes de estender cadeias polinucleotídicas contendo um iniciador de RNA (primer) com uma extremidade 3'-OH livre, cientistas em diversos laboratórios começaram a testar a idéia de que a síntese de DNA é iniciada por **primers de RNA**. Atualmente existem evidências definitivas apoiando a proposta de que a síntese de DNA é iniciada por pequenos segmentos de RNA. Esses primers de RNA são substituídos posteriormente por DNA, para isso uma enzima com atividade de exonuclease, a **RNase H**, degrada o RNA que está pareado com o DNA, exceto o último nucleotídeo que está ligado ao primeiro deoxinucleotídeo, que é removido por outra exonuclease. A seguir as lacunas são preenchidas pela DNA polimerase completando a fita do DNA. Os fragmentos resultantes são unidos por uma ligação covalente pela **DNA ligase** (veja a Figura 2.10).

A síntese dos primers de RNA é catalisada por enzimas denominadas **primases**, que apresentam atividades distintas das demais RNA polimerases. Nos procariotos, os primers de RNA possuem 10-60 nucleotídeos de tamanho. Em eucariotos são menores, com aproximadamente 10 nucleotídeos. O uso de primers de RNA é o mecanismo mais comum utilizado para iniciar a síntese de DNA. No entanto, alguns vírus parecem utilizar mecanismos diferentes para o início da síntese de DNA.



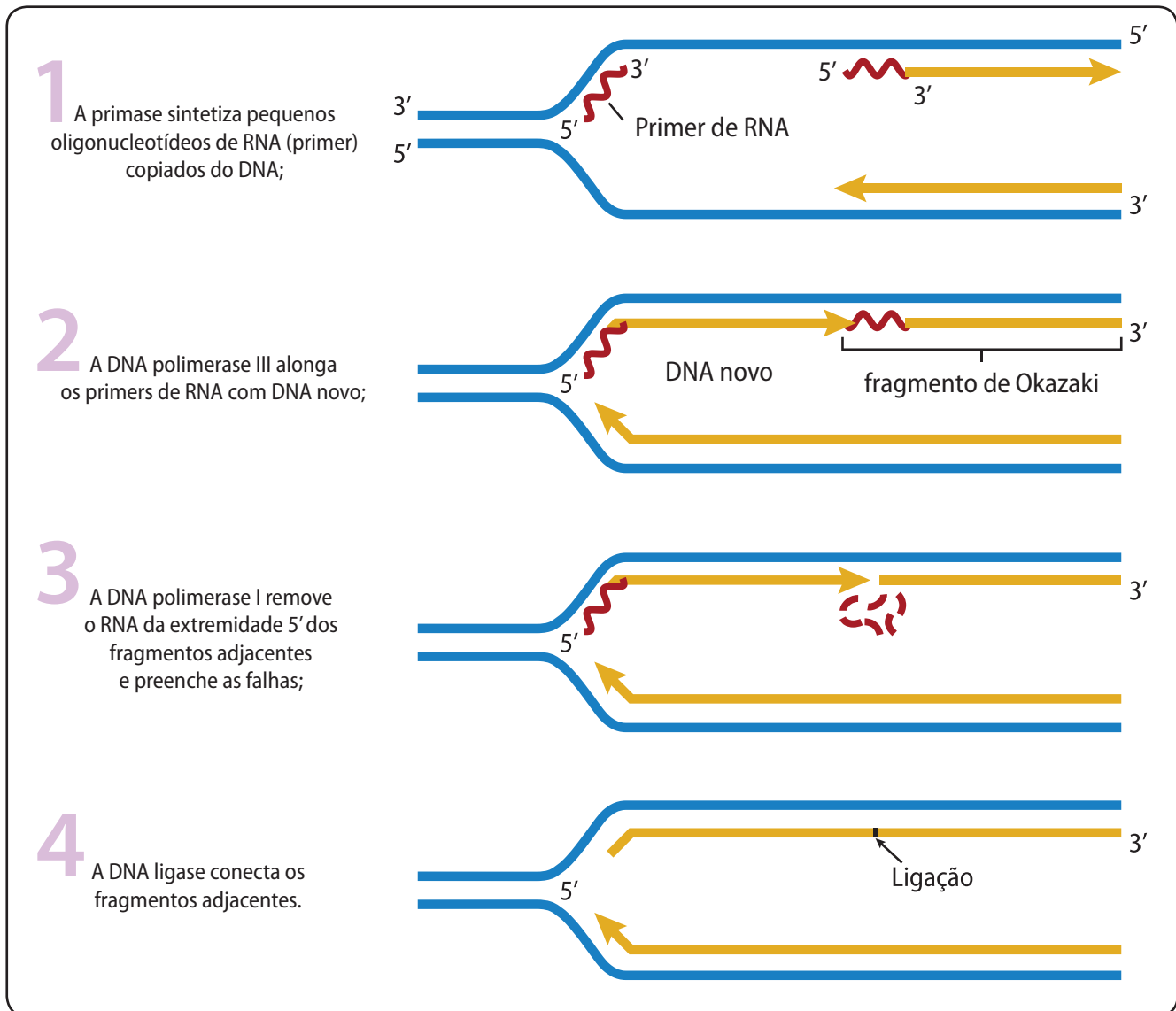


Figura 2.10 - Síntese da cadeia descontínua

## 2.7 A Complexidade do Sistema de Replicação Completo

Quando Watson e Crick formularam o modelo da estrutura em dupla hélice do DNA, imediatamente reconheceram que a natureza complementar das duas cadeias possibilitava uma base simples e fidedigna para a duplicação do material genético. A demonstração por Meselson e Stahl da replicação semiconservativa sedimentou o conceito de que ocorre a separação das duas cadeias e a seguir cada uma é utilizada como molde para a síntese das cadeias complementares. Assim, a dupla hélice original dirige a síntese de uma

progênie de duas duplas hélices idênticas. A purificação por Kornberg de uma enzima com capacidade de sintetizar DNA *in vitro*, a DNA polimerase I, forneceu a evidência final da ligação do que foi pensado ser um mecanismo elegantemente simples para a replicação do material genético – mas esse não foi o caso. Atualmente, os cientistas ainda tentam compreender os detalhes do mecanismo de replicação do DNA.

**A replicação do DNA é muito complexa.** É realizada por um complexo multienzimático denominado sistema de replicação ou **replissomo**. Nos eucariotos, os componentes da maquinaria de replicação ainda não estão totalmente identificados. Mesmo nos procariotos, a replicação do DNA necessita de diversas proteínas diferentes, e os detalhes do funcionamento dessas proteínas ainda estão sendo esclarecidos.

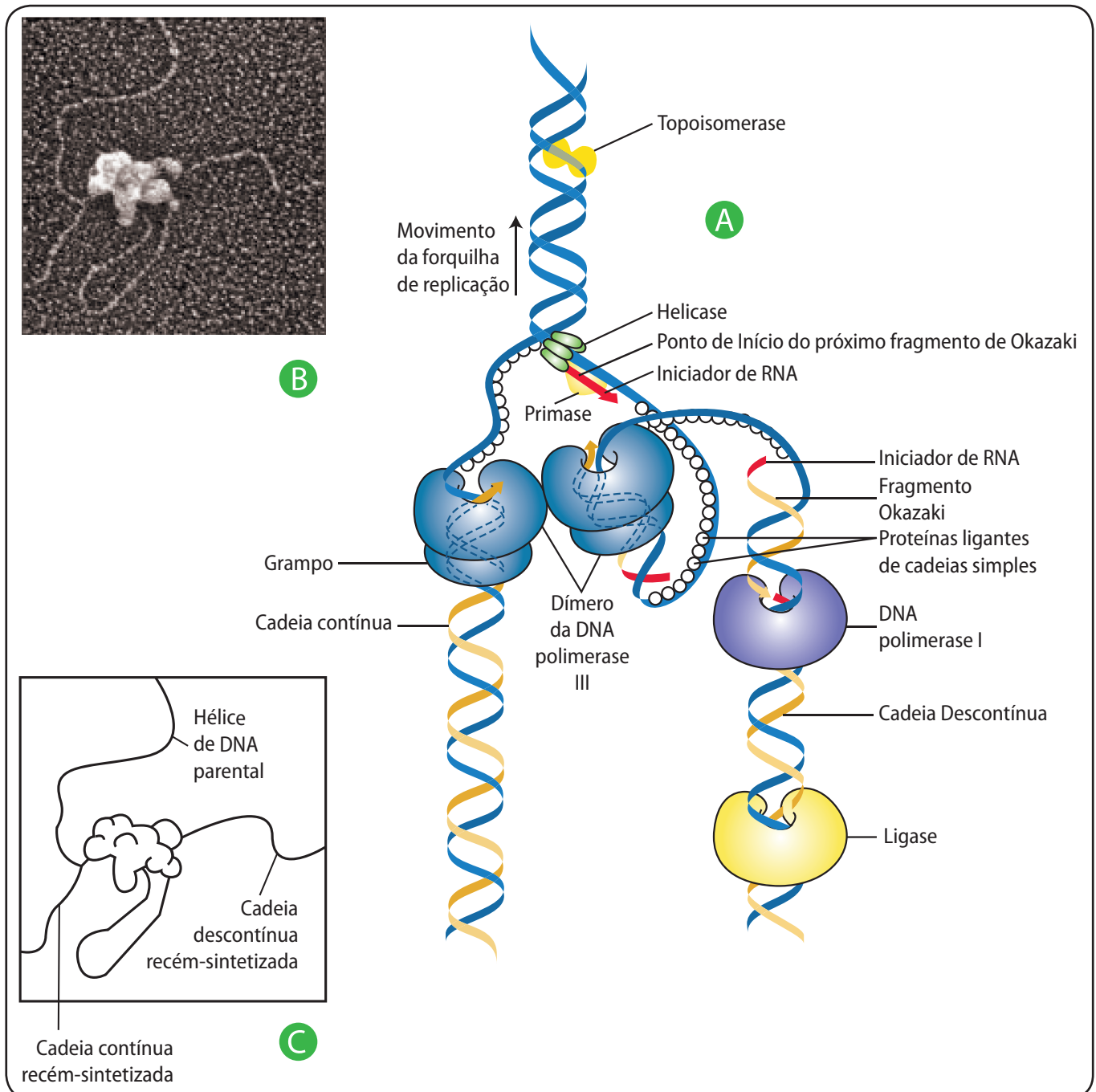
Inicialmente, as cadeias da dupla hélice parental precisam se separar de forma que cada uma sirva de molde para a síntese de uma nova cadeia. A separação das cadeias e o movimento da forquilha de replicação ocorrem progressivamente com as cadeias sendo temporariamente separadas adiante da forquilha de replicação, enquanto esta se move ao longo do cromossomo. Três proteínas diferentes estão envolvidas na separação das cadeias da dupla hélice:

1. proteínas separadoras de DNA ou **helicases** – diretamente envolvidas na catálise da separação das duplas hélices;
2. proteínas ligantes de cadeias simples de DNA (**SSBP** – Single Strand Binding Proteins) – se ligam fortemente às regiões de DNA de fita simples produzidas pela ação das **helicases** e ajudam a estabilização necessária para a polimerização. As SSBP se ligam ao DNA como tetrâmeros. O DNA complexado com as SSBP se replica *in vitro* 100 vezes mais rapidamente que o DNA não complexado. Provavelmente, as cadeias simples não complexadas formam estruturas secundárias que interferem no movimento das DNA polimerases e de outros componentes do complexo de replicação ao longo da molécula;
3. **DNA girases**, que catalisam a formação de enrolamentos negativos no DNA, são essenciais para a replicação e supõe-se que desempenham um papel fundamental no processo de desenro-

lamento, mas seu funcionamento não é conhecido. O superenrolamento é decorrente do processo de separação das fitas (veja a Figura 2.11).

As cadeias de DNA nascentes são iniciadas pelo uso de primers de RNA. A síntese dos primers de RNA é catalisada por uma classe especial de enzimas denominadas **primases**. A atividade das primases requer a formação de um complexo de primase e ao me-

Figura 2.11 - (A) Complexo atuando na forquilha de replicação (replissomo). (B) e (C) imagem do replissomo em microscopia eletrônica e sua interpretação



nos outras seis proteínas, esse complexo é chamado **primossomo**. O primossomo realiza a primeira reação iniciadora na cadeia **leading** (contínua) e a seguir a iniciação repetitiva na síntese dos fragmentos de Okazaki na cadeia **lagging** (interrompida).

A extensão covalente das cadeias de DNA durante a replicação do cromossomo da *E. coli* é realizada pela DNA **polimerase III**. Subseqüentemente à extensão pela DNA **polimerase III** na forquilha de replicação, a DNA **polimerase I** catalisa a remoção dos primers de RNA pela ação conjunta de ação das atividades de exonuclease 3'→5' e polimerase 5'→3', e a **DNA ligase** catalisa a ligação covalente dos fragmentos de cadeia simples resultantes.

Diversos dos componentes essenciais para a replicação do DNA em *E. coli* foram identificados geneticamente, isto é, por meio de linhagens portadoras de mutações que impossibilitam a replicação do DNA em determinadas condições. Espera-se que a função exata de diversos produtos gênicos envolvidos na replicação do DNA seja conhecida nos próximos anos. No entanto, as tentativas de isolar replissomos funcionais intactos não têm sido bem-sucedidas.

## 2.8 Replicação Semiconservativa em Cromossomos Eucarióticos

Os cromossomos eucarióticos, da mesma forma que os procarióticos, replicam-se de maneira semiconservativa. Isto foi demonstrado por Taylor, Woods e Hughes em 1975, por meio de autorradiografia de cromossomos de *Vicia faba*. Entretanto, na época desse experimento ainda não se sabia que cada cromossomo contém uma única molécula de DNA. Experimentos subseqüentes em diversos eucariotos demonstraram em todos os casos replicação semiconservativa.

## 2.9 Múltiplos Replicons por Cromossomo

Quando células eucarióticas são marcadas por pulsos de timidina  $^3\text{H}$  e o DNA é extraído e analisado por autorradiografia, são obser-

A palavra “*tandem*” significa um atrás do outro.

vados arranjos em *tandem*, sugerindo que macromoléculas únicas de DNA contêm diversas origens de replicação. Quando o pulso de marcação é seguido por um período curto de crescimento em meio não radioativo, esses arranjos contêm regiões centrais de alta densidade de grãos com caudas de densidade de grãos decrescente em cada término. Esse resultado demonstra que a replicação nos eucariotos é bidirecional da mesma maneira que nos procariotos. As caudas com densidade de grãos decrescente resultam da diluição gradual dos pools intracelulares de timidina  $^3\text{H}$  pela timidina  $^1\text{H}$ , enquanto a forquilha de replicação se move bidirecionalmente da origem central em direção ao término da replicação. Um segmento do cromossomo cuja replicação está sob controle de uma origem e dois términos é denominado replicon (veja a Figura 2.12). Nos procariotos, o cromossomo inteiro é um único replicon.

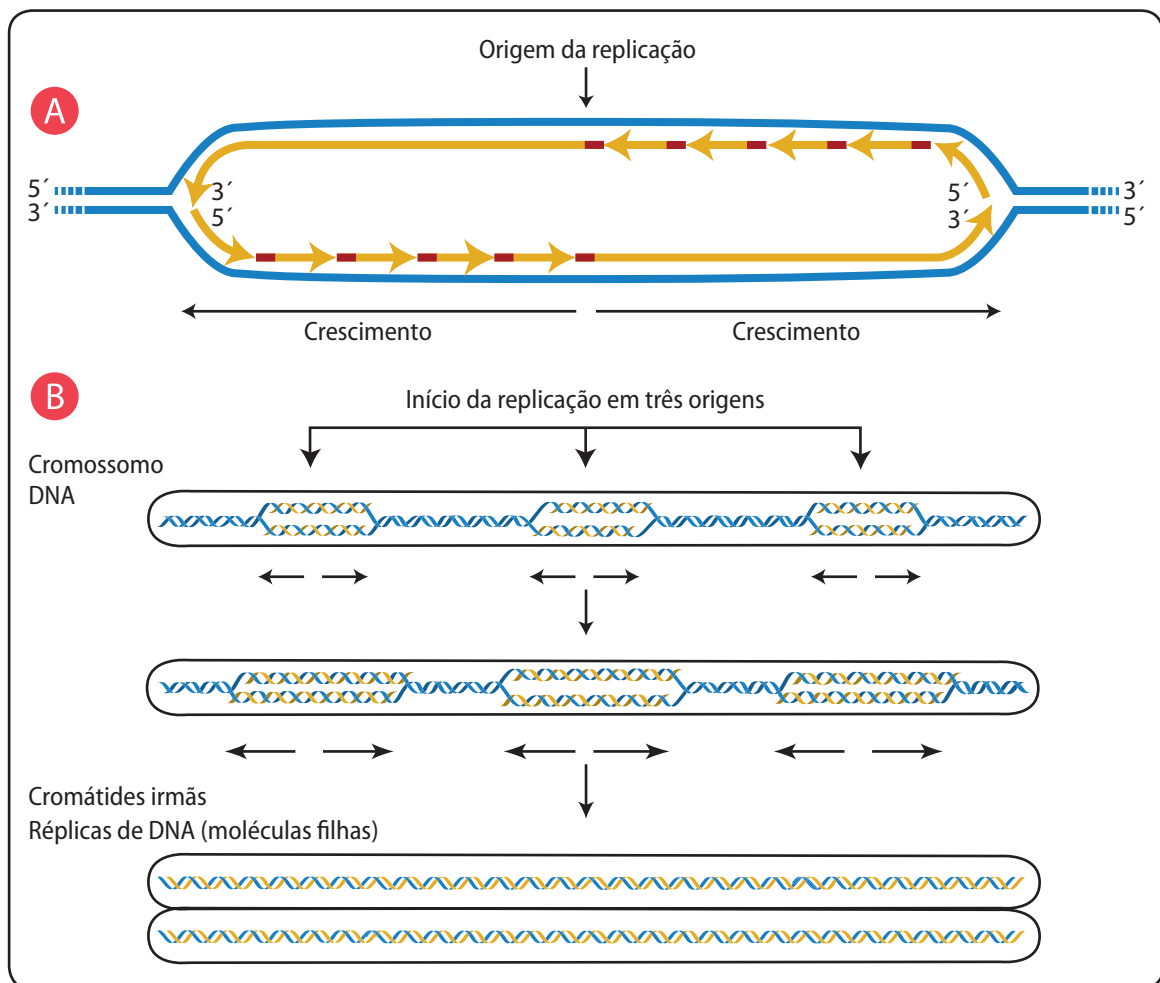


Figura 2.12 - Diversas origens de replicação nos cromossomos eucarióticos

A existência de diversas unidades de replicação por cromossomo eucariótico, cada uma com sua própria origem foi demonstrada por microscopia eletrônica (veja a Figura 2.13). Algumas evidências indicam que a replicação pode se iniciar em um número maior de locais durante as divisões celulares muito rápidas do início da embriogênese em relação aos estágios posteriores do desenvolvimento. Um aspecto que está claro é de que o DNA recém-replicado é rapidamente empacotado em nucleossomos.

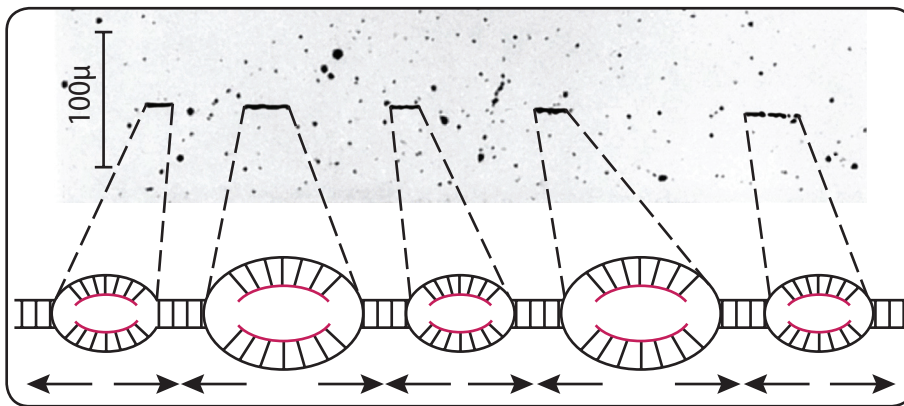


Figura 2.13 - Autorradiografia evidenciando diversas origens de replicação

As múltiplas origens de replicação são necessárias para possibilitar a replicação completa das moléculas de DNA durante o ciclo celular. Por exemplo, na *Drosophila melanogaster*, as moléculas dos maiores cromossomos contêm aproximadamente  $6,5 \times 10^7$  pares de nucleotídeos. A velocidade de replicação do DNA a  $25^\circ\text{C}$  é aproximadamente 2.600 pares de nucleotídeos por minuto. Uma única forquilha de replicação demoraria  $17\frac{1}{2}$  dias para atravessar uma molécula inteira. Com duas forquilhas de replicação movendo-se bidirecionalmente a partir de uma origem central, a molécula poderia se replicar em apenas  $8\frac{1}{2}$  dias. Os cromossomos de embriões de Drosófila se replicam a cada 3-4 minutos e o núcleo se divide a cada 9-10 minutos durante as clivagens iniciais do zigoto. Portanto, a replicação completa do DNA dos cromossomos maiores em  $3\frac{1}{2}$  minutos necessita de 7.000 forquilhas de replicação distribuídas em intervalos iguais ao longo das moléculas.

O número de origens de replicação ativas varia durante os diferentes estágios do desenvolvimento. Na Drosófila existem aproximadamente 10 vezes mais origens ativas por núcleo durante as

clivagens iniciais do que nas células adultas. Ainda não se sabe o que determina, quantas e quais origens estão ativas em células diferentes e em diferentes estágios do desenvolvimento.

## 2.10 O Início da Replicação

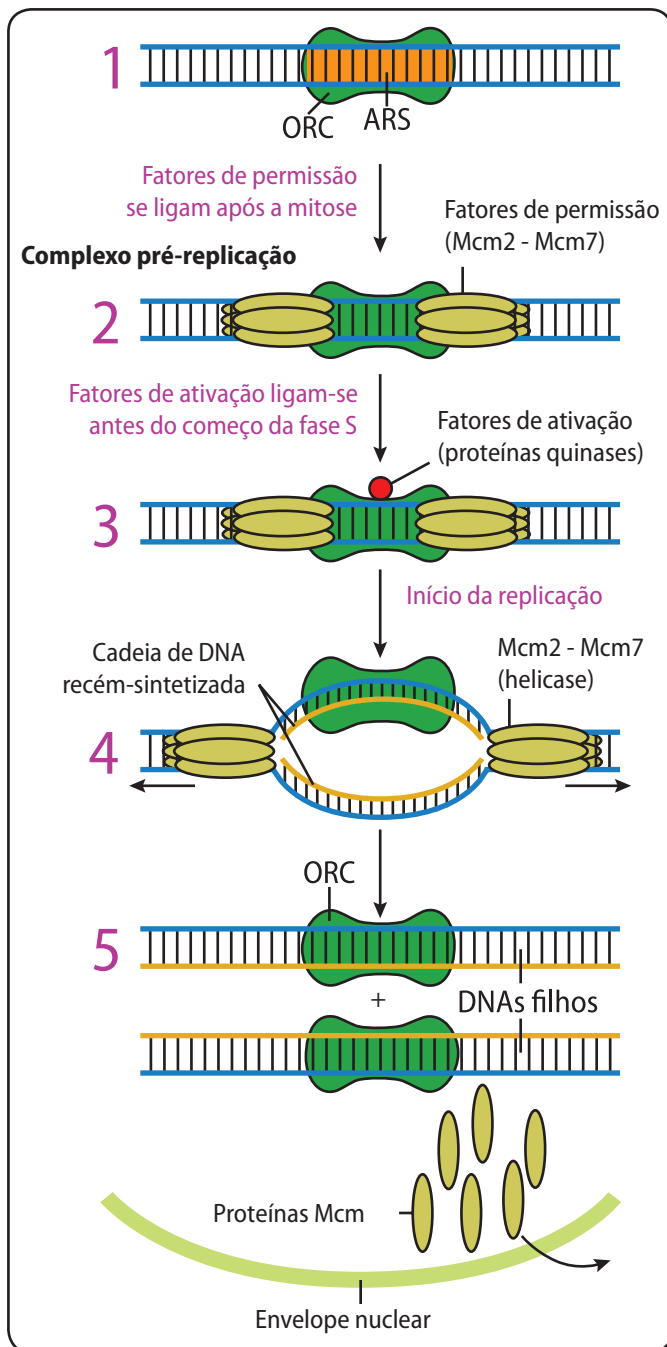


Figura 2.14 - Reconhecimento da origem de replicação

Em todos os organismos é fundamental que o início da replicação seja controlado para possibilitar que o número de cromossomos e o número de células permaneçam equilibrados.

A formação da forquilha de replicação requer a separação das fitas de DNA, gerando um sítio de desenrolamento que se move ao longo do DNA durante a extensão e o início da nova cadeia. Os sítios específicos onde esse processo se inicia são denominados origens de replicação, onde se ligam proteínas específicas que produzem a separação das cadeias e o início do processo (veja a Figura 2.14).

## 2.11 O Término da Replicação: A questão dos telômeros

O fato de o sistema de replicação do DNA ocorrer diferencialmente nas duas fitas acarreta um problema adicional para os cromossomos lineares.

Nos cromossomos circulares, com apenas uma origem de replicação, a finalização do processo evolui bidirecionalmente até que haja o encontro das duas forquilhas e a religação das fitas; as moléculas resultantes são separadas pela ação da topoisomerase.

Nos cromossomos lineares, cada fita evolui independentemente. No entanto, esse sistema não pode finalizar a replicação nas extremidades. Após a retirada do primer de RNA na fita descontínua, um pequeno trecho não é replicado. Com isso, a cada ciclo de replicação há um encurtamento de uma das duas moléculas-filhas de DNA. Isto acarreta a perda de informação dessas regiões (veja a Figura 2.15).

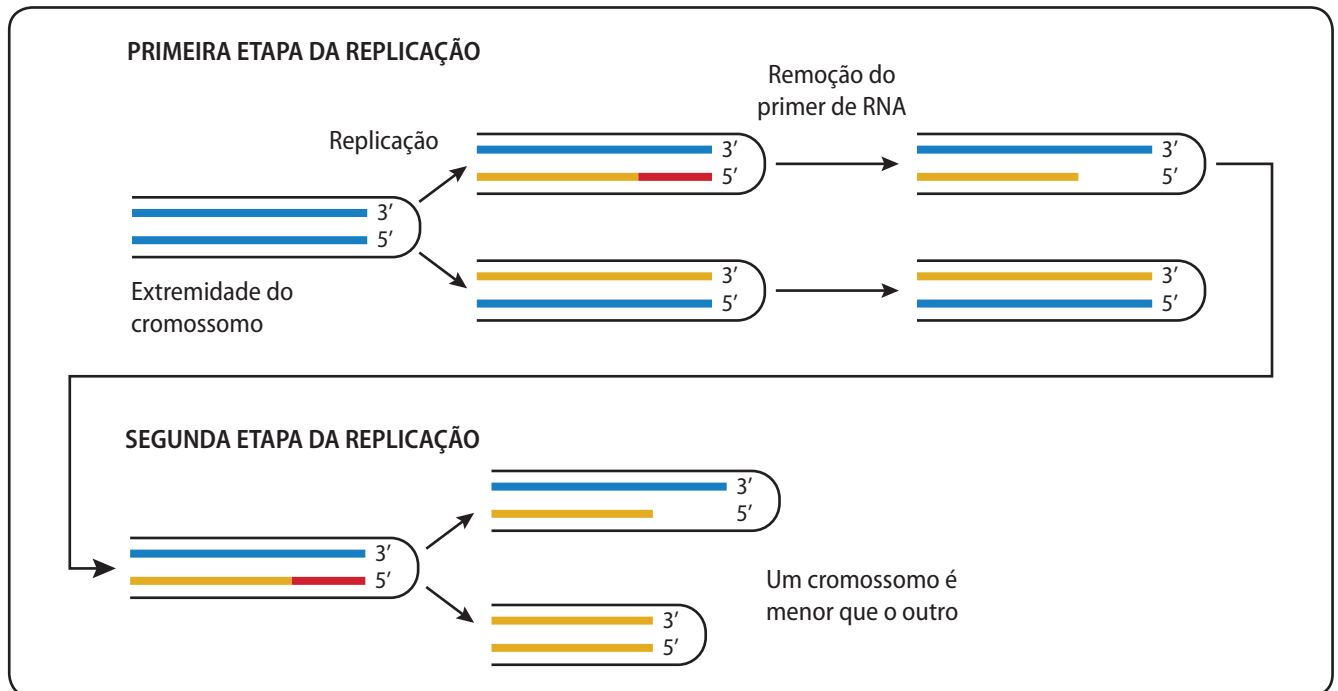


Figura 2.15 - Erosão dos telômeros

Apesar dessa dificuldade, os cromossomos lineares são mais interessantes evolutivamente, pois possibilitam a realização de **crossing-over**.

As regiões terminais dos cromossomos são denominadas **telômeros** e são essenciais para a estabilidade dos cromossomos. Essas regiões contêm uma seqüência de DNA repetida diversas vezes em *tandem*; essa seqüência é bastante conservada evolutivamente. Em humanos a seqüência repetida é 5'-TTAGGG-3'.

Na replicação do DNA, a remoção do último primer acarreta encurtamento da seqüência telomérica. Isto acarreta uma limitação quanto ao número de divisões celulares possíveis para uma

### Crossing-over

É um evento que ocorre na meiose e acarreta troca de partes entre cromossomos homólogos, resultando no aumento da variabilidade genética pelo embaralhamento dos genes maternos e paternos no mesmo cromossomo, sendo os cromossomos resultantes diferentes dos originais.



**Senescência celular**

Senescência é o processo natural de envelhecimento no nível celular ou o conjunto de fenômenos associados a esse processo.

célula. Quando a perda dessa seqüência é significativa, ocorre a erosão dos telômeros, causando instabilidade cromossômica na célula. Esse fenômeno está associado à **senescência celular**.

No entanto, algumas células possuem a capacidade de replicar os telômeros, não ocorrendo perda destas seqüências. Isto se deve

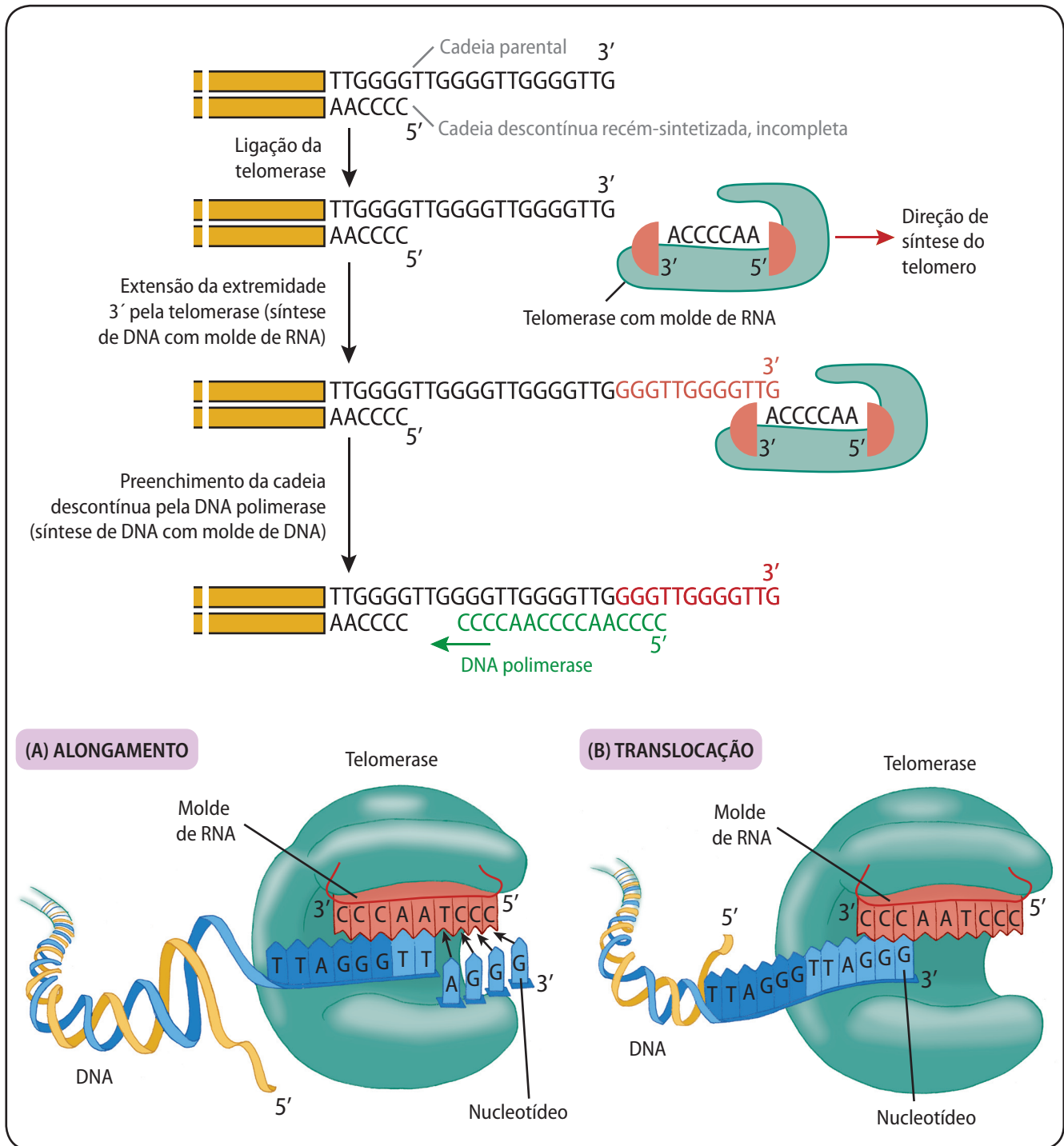


Figura 2.16 - Funcionamento da telomerase

a uma enzima celular denominada **telomerase**. Essa enzima possui tanto componentes protéicos quanto RNA. Esse componente de RNA serve de molde para a replicação da seqüência telomérica na extremidade 3'. Essa DNA polimerase, extraordinariamente, utiliza o RNA como molde, característica presente em uma classe de enzimas denominada transcriptase reversa. Para separar o molde do produto, há um componente da enzima com atividade de RNA-DNA helicase (veja a Figura 2.16). O controle da atividade da telomerase se dá por proteínas que se ligam à região de fita dupla do telômero.

## 2.12 Componentes do Aparelho de Replicação dos Eucariotos

A importância da replicação para a manutenção da informação requer acuracidade e processividade das enzimas envolvidas.

A *E. coli* possui pelo menos cinco tipos de DNA polimerases: a DNA polimerase I, que preenche espaços entre fragmentos de Okazaki na cadeia descontínua (lagging), sendo a enzima principal para preencher as lacunas (gaps) durante o reparo do DNA; a DNA polimerase III, que é a principal enzima na replicação de DNA, composta de diversas subunidades, com mais de 600kDa, sendo seu centro constituído das subunidades  $\alpha$ ,  $\epsilon$ , e  $\theta$ . As duas subunidades  $\beta$  têm a função de impedir que o centro da polimerase se desligue do DNA molde, deslizando com ele. As demais DNA polimerases atuam no reparo do DNA.

Em eucariotos o sistema é mais complexo e as células possuem múltiplas DNA polimerases; uma célula normal apresenta mais de 15 polimerases diferentes. Em mamíferos as principais DNA polimerases são a DNA polimerase  $\alpha$ , que atua na síntese da cadeia descontínua, a DNA polimerase  $\beta$ , que atua no reparo do DNA, a DNA polimerase  $\gamma$ , que está envolvida na replicação do DNA mitocondrial, a DNA polimerase  $\delta$ , que é responsável pela síntese da cadeia contínua, e a DNA polimerase  $\epsilon$ , que participa do reparo do DNA.

A alta processividade é obtida com o auxílio de outras proteínas que atuam na forquilha de replicação, podendo atuar numa velocidade de 1.000 nucleotídeos por segundo.

A identificação dos sinais que regulam a replicação do DNA durante o crescimento e a diferenciação de organismos multicelulares como plantas superiores e animais proporcionará um grande avanço para a compreensão desse processo.

## Resumo

A replicação do DNA é central para toda a Biologia. Esse processo ocorre de forma semiconservativa. A DNA polimerase é a principal enzima utilizada para a replicação e atua na direção 5'-3', utilizando como substrato a molécula de DNA molde e os nucleotídeos trifosfatados, possuindo alta processividade e acurácia. A replicação ocorre bidirecionalmente e simultaneamente nas duas fitas, em uma das fitas a replicação é descontínua. O sistema de replicação completo é complexo, com enzimas que realizam atividades distintas, como DNA helicases, primase, SSBP, DNA ligase e topoisomerasas. Os cromossomos de eucariotos possuem diversas origens de replicação. As extremidades dos cromossomos apresentam uma dificuldade intrínseca de replicação, a cada ciclo de replicação há um encurtamento de uma das duas moléculas-filhas de DNA. Em algumas células a enzima telomerase possibilita a recomposição desse encurtamento.

## Referência Comentada

### Biologia molecular do gene

*James D. Watson, Michael Levine, Alexander Gann, Richard Losick, Tania A. Baker, Stephen P. Bell*

*Capítulo 8:* Texto abrangente, compreendendo os detalhes moleculares, com ilustrações que facilitam a compreensão e exemplos que possibilitam a contextualização.

WATSON, J. D. et al. **Biologia molecular do gene**. 5. ed. Porto Alegre: Artmed, 2006.

## Referências Bibliográficas

ALBERTS, B. et al. **Molecular Biology of the Cell**. 4th. ed. New York: Garland Pub. Inc., 2002.

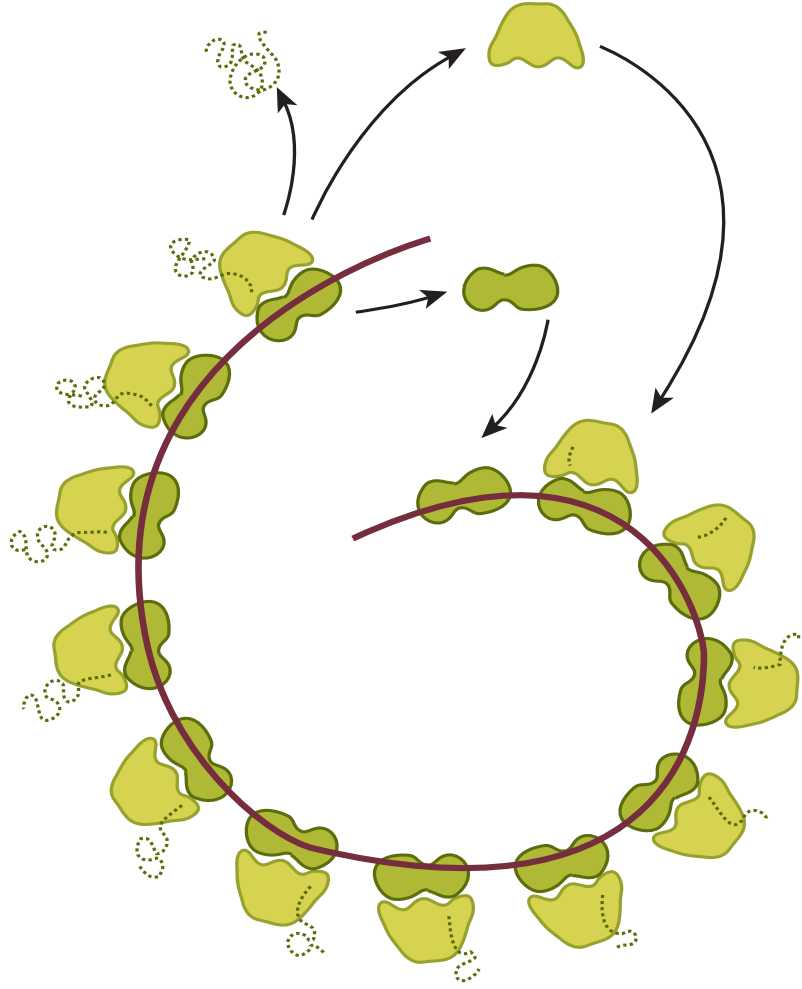
GRIFFITHS, a. j. f. et al. **Introdução à Genética**. 8. ed. Rio de Janeiro: Guanabara-Koogan, ,2006.

LEWIN, B. **Genes VII**. Porto Alegre: Artmed, 2001.

PURVES, W. K. et al. **Vida: a ciência da Biologia**. Célula e hereditariedade. 6. Ed. Porto Alegre: Artmed, 2002. v. I.



# CAPÍTULO 3



## O Fluxo da Informação

*O material genético deve controlar o desenvolvimento do fenótipo do organismo. A compreensão desse processo iniciou-se a partir de pressupostos teóricos, considerando-se a importância das proteínas para os sistemas biológicos, a relação entre a seqüência de nucleotídeos dos ácidos nucléicos e a seqüência de aminoácidos, e a universalidade do código genético.*

*Este capítulo tem como objetivo sistematizar o conhecimento das etapas distintas do fluxo da informação: a síntese de RNA a partir do DNA, denominada **transcrição**, e a síntese da cadeia polipeptídica a partir do molde de RNA, denominada **tradução**. Também vamos caracterizar a maquinária celular envolvida nessas atividades e estabelecer as diferenças entre procariotos e eucariotos, e considerar a repercussão destas diferenças .*





A elucidação do processo de síntese protéica pelas células é um dos maiores desafios da Biologia, sendo indiscutível a importância das proteínas para os sistemas biológicos. A relação entre a seqüência de nucleotídeos dos ácidos nucléicos e a seqüência de aminoácidos das proteínas parece óbvia, mas a caracterização do processo funcional é um desafio persistente. A partir dos anos 1950 e 60 estabeleceu-se intensa colaboração entre bioquímicos e biólogos moleculares que resultou em descobertas importantes sobre o processo de síntese protéica.

### A hipótese da seqüência e o Dogma Central em 1957

Meus pensamentos (e de diversos colegas) são baseados em dois princípios gerais, chamados de hipótese da seqüência e Dogma Central. A evidência experimental para ambos é deficiente, mas acho de grande ajuda para elucidar esses problemas complexos. Sua natureza especulativa é enfatizada pela denominação. Um exercício instrutivo seria tentar construir uma teoria sem utilizá-los. Pela hipótese da seqüência, assume-se que a especificidade de um segmento de ácido nucléico seria expressa apenas pela seqüência de bases, de forma que sua seqüência seria um código (simples) para a seqüência de aminoácidos de uma determinada proteína. Essa hipótese é amplamente aceita. Sua virtude é que unifica diversos princípios: a importância bioquímica das proteínas e a função do gene,

especialmente de seu ácido nucléico; a linearidade das moléculas de proteína (consideradas covalentemente); e a linearidade genética do gene funcional. Diversos laboratórios, incluindo o nosso, trabalham para obter evidências experimentais dessa hipótese. O Dogma Central estabelece que uma vez que a informação tenha sido passada para proteína não há mais retorno. A transferência de informação do ácido nucléico para ácido nucléico, ou de ácido nucléico para proteína, seria possível, mas a transferência de informação de proteína para proteína, ou de proteína para ácido nucléico, seria impossível. A informação significa a determinação precisa da seqüência, tanto na seqüência de bases quanto dos aminoácidos na proteína.

CRICK, F. H. C. On Protein Synthesis. Symp. Soc. Exp. Biol. XII: 138-163. quoted in: JUDSON, H. F. **The Eighth Day of Creation**. Expanded Edition (1979, 1996) 1958.

Inicialmente a abordagem foi especulativa, quando foram cogitadas as possibilidades a partir do conhecimento disponível. A partir daí foram conduzidos experimentos laboratoriais que demonstraram experimentalmente os pressupostos teóricos, sendo as lacunas preenchidas paulatinamente.

Após a elucidação da estrutura molecular do DNA e a constatação de que a informação deveria estar codificada na seqüência de DNA, começaram as especulações de qual seria a relação entre a seqüência de nucleotídeos do DNA e a seqüência de aminoácidos nas proteínas, e o sistema operacional para a célula processar essa informação.

Em 1954 George Gamow, um renomado físico, sugeriu que as proteínas deveriam ser sintetizadas diretamente no DNA, sendo os aminoácidos alinhados na ordem correta por meio de cavidades com formas diferentes, específicas para cada tipo de aminoácido, formadas pela seqüência de DNA. Esse modelo mostrou-se limitado por restringir as possíveis seqüências de aminoácidos nas proteínas. Também se mostrou paradoxal, uma vez que a dimensão molecular dos nucleotídeos é muito maior do que a dimensão dos aminoácidos. Outro paradoxo é o fato de que nos eucariotos o DNA está localizado no núcleo, enquanto a síntese de proteínas ocorre no citoplasma.

Parecia plausível que existisse uma molécula intermediária. Os estudos preliminares de Brachet (1944) e Caspersson (1947) sugeriram que havia uma íntima conexão entre RNA e síntese de proteínas. Uma vez que as evidências experimentais demonstravam que a síntese protéica era proporcional à quantidade de RNA (isto estava de alguma forma equivocado, visto que o RNA observado não era o RNA mensageiro, mas o RNA ribossômico), sugere-se que o RNA poderia ser esse intermediário. Citologicamente essa proposta era possível, porque o RNA está presente tanto no núcleo quanto no citoplasma. Além do mais, o RNA possui propriedades compatíveis com essa função:

1. polinucleotídeo;
2. composição de bases reflete a composição de bases do DNA;

3. heterogêneo em tamanho;
4. sintetizado e degradado rapidamente.

Essas considerações resultaram na hipótese conhecida como o Dogma Central da Biologia, segundo o qual o fluxo da informação ocorreria no seguinte sentido:

DNA → RNA → Proteína

Assim, existem duas etapas distintas no fluxo da informação: a etapa da síntese de RNA a partir do DNA, denominada transcrição, e a síntese da cadeia polipeptídica a partir do molde de RNA, denominada **tradução**.

A informação genética seria codificada na seqüência linear das bases no DNA. Esse também seria o caso para o outro ácido nucleico, o RNA. Existem quatro tipos de bases no RNA que são idênticas às encontradas no DNA, com apenas uma exceção, o uso da **uracila no RNA** em lugar da timidina (5-metil-uracila). O RNA também pode ser o portador da informação genética, como é o caso de alguns vírus de RNA. Mas outra função é mais evidente – a seqüência de bases na maioria dos RNAs é determinada pela seqüência de bases do DNA. Assim, o RNA pode atuar como um intermediário portador de informação. Particularmente, a maneira pela qual a informação genética é expressa como proteínas é sempre por meio de um **RNA intermediário**. A seqüência de aminoácidos nas proteínas é determinada pela seqüência de bases do RNA mensageiro. As seqüências dos outros RNAs componentes da maquinaria de síntese protéica são também determinadas pelo DNA, mas essa informação não é traduzida diretamente em proteína.

### 3.1 A estrutura do RNA

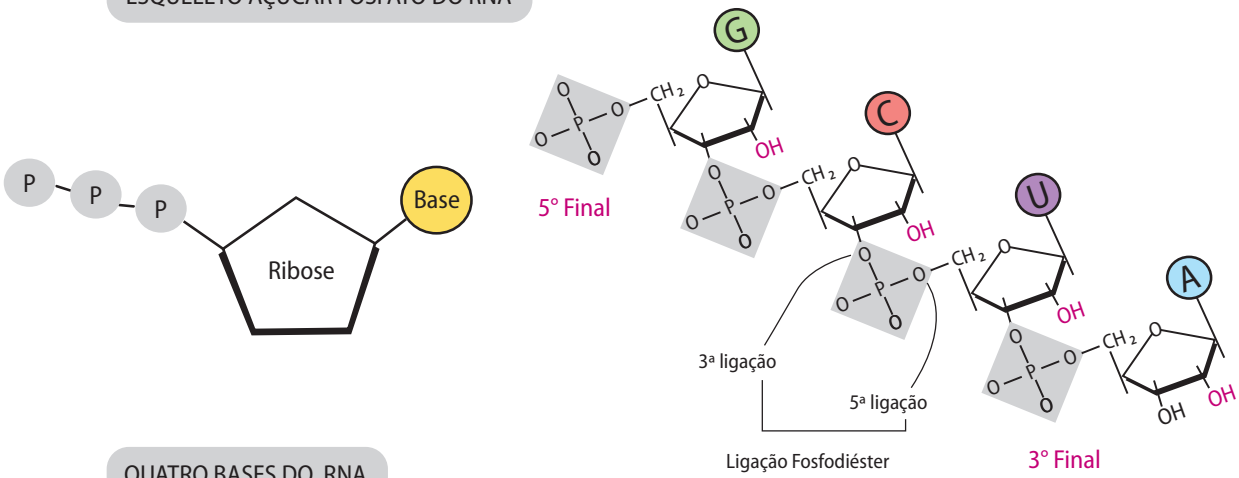
O RNA é uma molécula monocatenária (de uma única cadeia), embora possa ter regiões de dupla hélice decorrente do **parea-**

**mento intramolecular**, e em alguns vírus é de fita dupla. Assim, as análises de difração dos raios X não puderam ser interpretadas logicamente, não se verificando a mesma regularidade observada no DNA. O RNA difere do DNA quanto à composição das bases, sendo utilizada a uracila no lugar da timidina, e quanto à pentose, sendo utilizada a **ribose**. A cadeia mostra similaridade ao DNA na união dos nucleotídeos, apresentando também a orientação 5' e 3', a presença de uma hidroxila na posição 2' torna a molécula mais reativa, portanto mais instável que o DNA. Na maioria das moléculas de RNA celulares, as razões de A/U e C/G diferem de um, e também quando aquecido o RNA, não se desnatura em duas moléculas polinucleotídicas separadas. Assim, concluiu-se que o RNA não tem uma estrutura regular. No entanto, as curvas de desnaturação demonstraram a existência de algum tipo de pareamento de bases no RNA, partes da molécula poderiam formar ganchos antiparalelos. A presença de regiões de dupla hélice foi confirmada por estudos de rotação óptica e difração de raios X. A importância dessas alças não foi considerada até ser confirmado que, em alguns casos, o RNA tem estruturas secundária e terciária definidas e precisas, com importantes conseqüências biológicas. Como o pareamento de bases é intramolecular, esse é facilmente revertido. Embora a estrutura de pareamento de bases seja revertida pela temperatura, a cadeia permanece unida pelas ligações fosfodiéster (veja a Figura 3.1).

#### Características da molécula de RNA

- molécula linear/regiões de dupla hélice;
- nucleotídeos A, U, C e G;
- pentose – ribose;
- localizado no núcleo e no citoplasma;
- quantidade variável – proporcional à síntese de proteínas.

ESQUELETO AÇÚCAR FOSFATO DO RNA



QUATRO BASES DO RNA

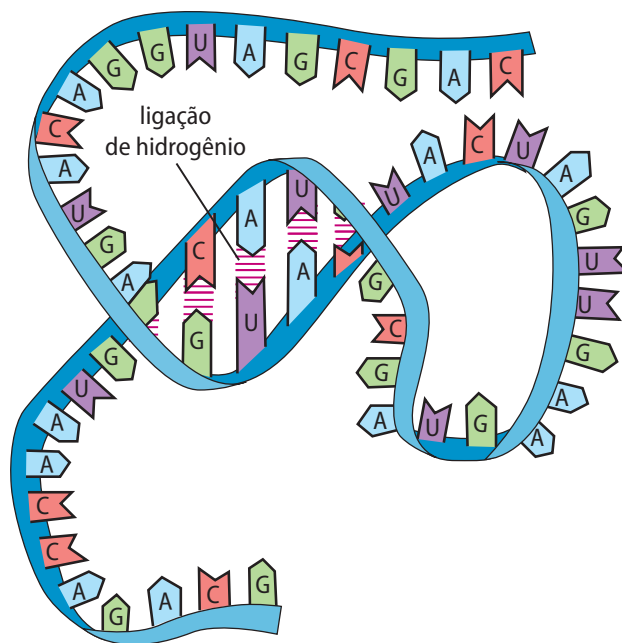
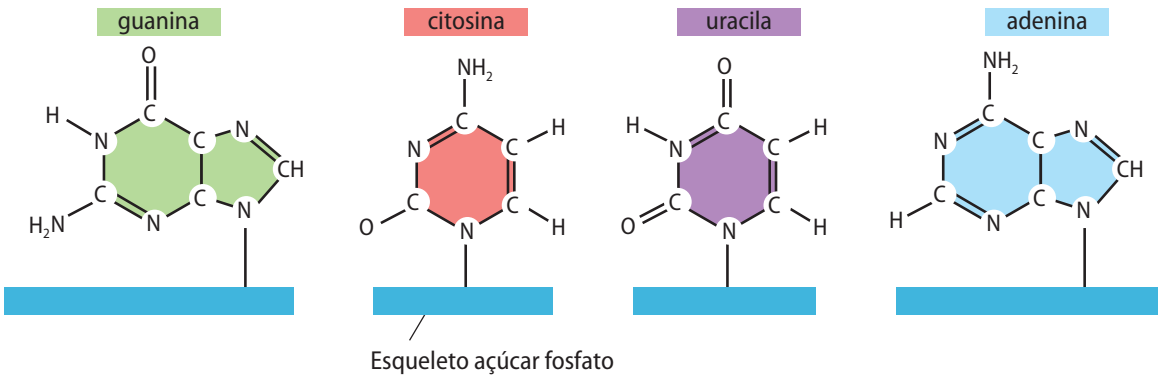


Figura 3.1 - Estrutura da molécula de RNA

## 3.2 O RNA e o fluxo da informação

A partir da hipótese do Dogma Central, os detalhes do fluxo da informação foram esclarecidos paulatinamente por meio de uma gama de procedimentos experimentais.

Para isto, os sistemas de síntese protéica *in vitro* foram fundamentais. Inicialmente, em 1953, Paul Zamecnik e colaboradores demonstraram que o sítio onde ocorria a síntese de proteínas era uma estrutura celular denominada **ribossomo**. A seguir Zamecnik e Hoagland (1958) demonstraram que os aminoácidos, antes de serem incorporados à cadeia polipeptídica, ligavam-se a pequenas moléculas de RNA, denominadas **RNA transportadores** (tRNA).

Inicialmente, pensava-se que todo RNA celular teria a função de molde. No entanto, a caracterização de diferentes formas de RNA evidenciou uma complexidade adicional. Embora em termos quantitativos sejam predominantes os RNAs encontrados nos ribossomos (rRNA), a regularidade do tamanho e da composição do rRNA seria incompatível com a complexidade esperada, também são moléculas altamente resistentes, difíceis de degradar. Outros estudos mostravam que os moldes para a síntese de proteínas em bactérias tinham vida curta. Também era estranho que as seqüências de bases dos RNA ribossômicos tivessem pouca correlação com o DNA cromossômico.

A resolução desse paradoxo ocorreu em 1961 com a descoberta de uma terceira forma de RNA, o **RNA mensageiro** (mRNA) por Brenner, Jacob e Meselson. Este mostrou-se o verdadeiro molde da síntese protéica, estabelecendo-se assim, os princípios do sistema funcional.

Os ribossomos atuam como fábricas moleculares de polipeptídios, que deslizam ao longo da molécula de mRNA, onde os aminoácidos são alinhados adequadamente por meio do RNA transportador para depois serem unidos por ligações peptídicas.

Desta forma, dever existir uma regra para traduzir a informação da seqüência de nucleotídeos na seqüência de aminoácidos. As primeiras especulações teóricas sugeriam que não era possível uma relação direta. Considerando-se combinações entre os nucleotídeos, seriam necessárias combinações com no mínimo

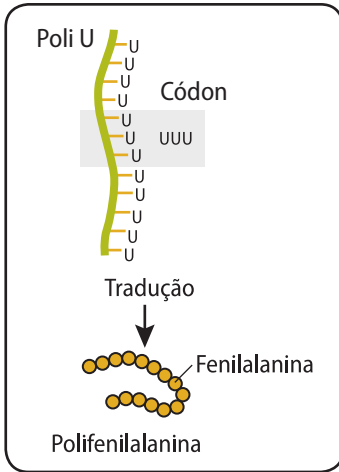


Figura 3.2 - Síntese polipeptídica com poli U

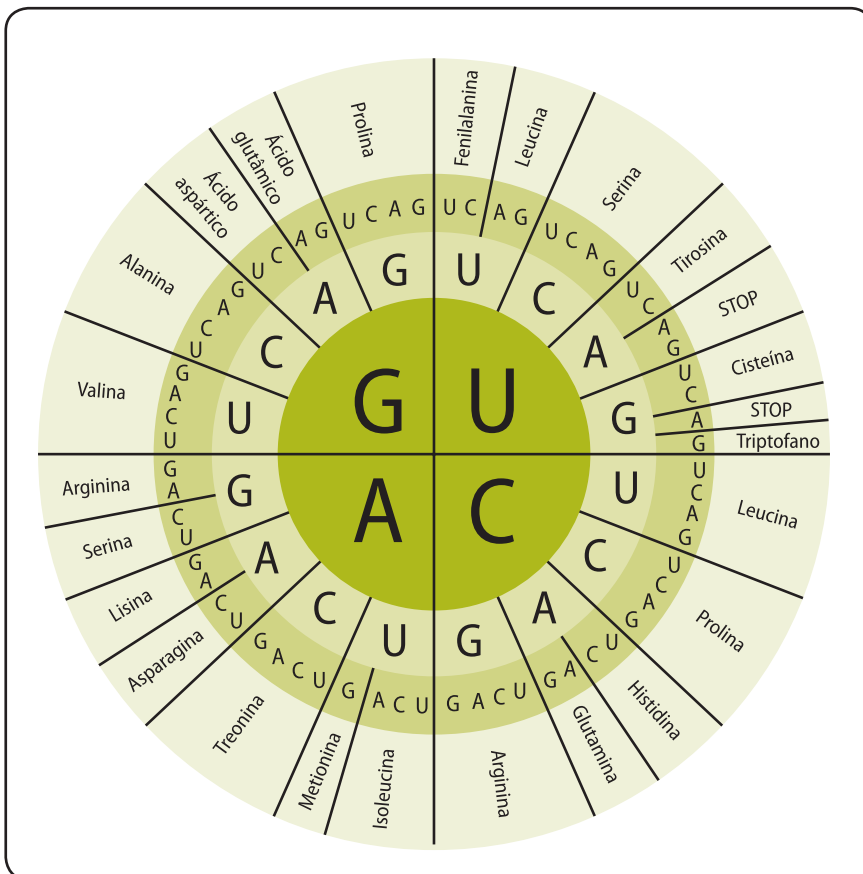
três nucleotídeos para produzir a diversidade necessária ( $4^3=64$ ), mas isto produzia uma diversidade excedente, possibilitando uma redundância. A natureza do código em triplete foi confirmada experimentalmente por Brenner e Crick (1961).

Embora a dinâmica da informação fosse compreendida, o código ainda não havia sido decifrado, sendo o triplete de nucleotídeos denominado códon.

A solução para essa questão foi possível com o uso de RNAs mensageiros sintéticos para síntese protéica *in vitro*. O experimento pioneiro realizado por Nirenberg e Matthaei (1961) utilizou um mRNA constituído apenas de uracila (poli U), o que resultou num peptídeo composto de fenilalanina, sugerindo que o códon UUU codificava o aminoácido fenilalanina (veja a Figura 3.2). A partir desse sistema, outros códons foram sendo decifrados e as 64 possíveis combinações elucidadas.

Essa interpretação denominada **código genético** mostrou-se linear, sem sobreposições, dos 64 tripletes possíveis, 61 codificam

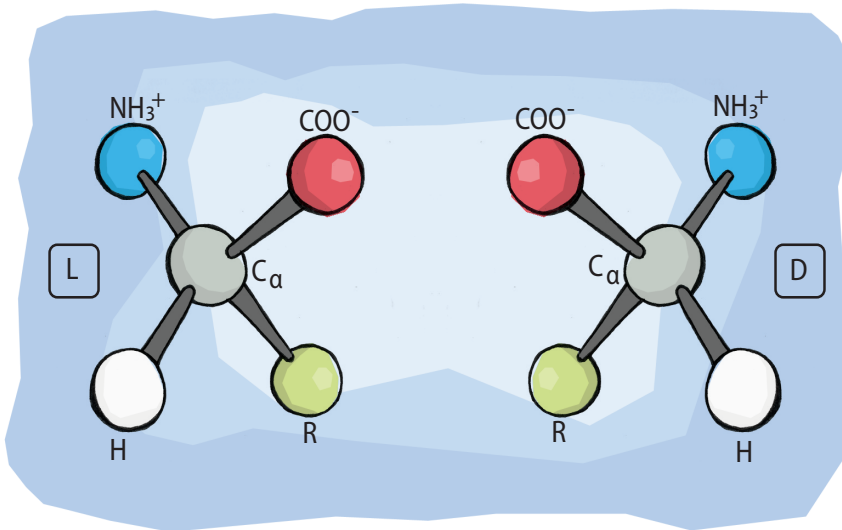
Figura 3.3 - Código genético



aminoácidos. Dos 20 aminoácidos, apenas dois são codificados por apenas um códon, metionina e triptofano, os demais são codificados por mais de uma combinação (veja a Figura 3.3). Os códons que não codificam aminoácidos foram identificados como códons de terminação. Essa interpretação é utilizada universalmente pelos seres vivos. Curiosamente, as células utilizam apenas os aminoácidos levogiros (veja a Figura 3.4) para a síntese de proteínas.

### Isômeros Ópticos

O átomo de carbono- $\alpha$  é assimétrico, o que permite a formação de imagens especulares (ou estéreo) L e D.



Proteínas consistem exclusivamente de aminoácidos-L.

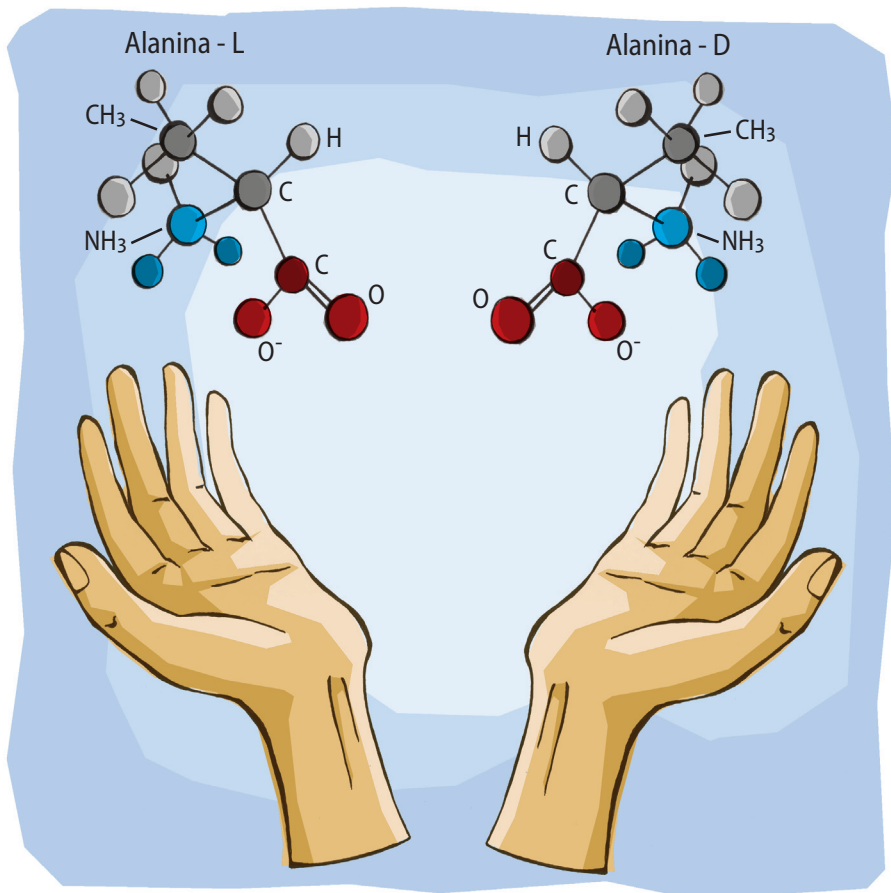


Figura 3.4 - Isomeria dos aminoácidos



### 3.3 Síntese do RNA

O processo de síntese de RNA constitui a etapa inicial do fluxo da informação e é denominada **transcrição**. A caracterização desse processo é indispensável para a compreensão do funcionamento celular.

Por meio do fracionamento celular e da autorradiografia após a incorporação de precursores radioativos nas moléculas de RNA, foi possível averiguar a cinética da síntese de RNA em células e tecidos, demonstrando que a maioria dos precursores são incorporados inicialmente no núcleo e a seguir o RNA recém-sintetizado é transportado para o citoplasma, embora uma fração desse RNA permaneça no núcleo. A localização da síntese do RNA no núcleo

forneceu a evidência experimental de que o DNA seria o molde para a síntese de RNA (veja a Figura 3.5).

O processo bioquímico da síntese de RNA começou a ser esclarecido quando S. Weiss e J. Hurwitz (1961) identificaram, em extratos celulares de bactérias e de núcleos de mamíferos, atividade enzimática

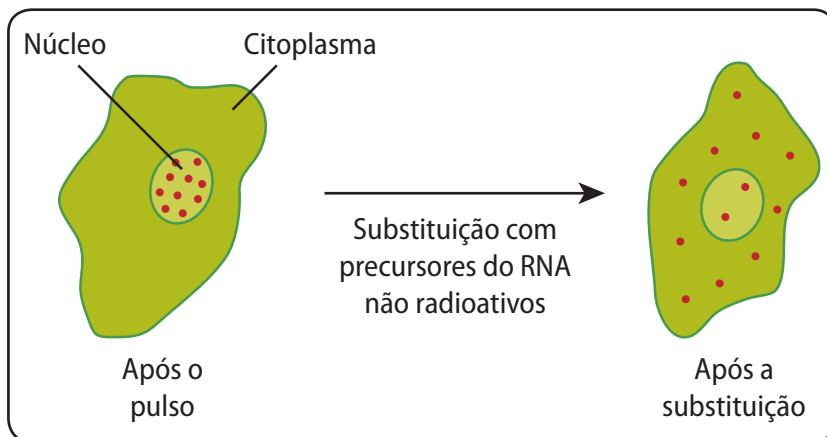


Figura 3.5 - Cinética da síntese de RNA

de síntese de RNA, numa reação dependente da presença de DNA. Essa reação enzimática utiliza como precursores nucleotídeos trifosfatados (ATP, CTP, GTP e UTP), liberando pirofostato inorgânico concomitantemente à formação de ligações fosfodiéster entre os nucleotídeos monofosfatados incorporados à cadeia de RNA. A enzima envolvida nesse processo, denominada RNA polimerase, é específica para ribonucleotídeos trifosfatados. No entanto, a natureza do RNA produzido é dependente do DNA adicionado.

Os estudos histológicos e autorradiográficos sugerem que a duplicação do DNA e a síntese de RNA no núcleo ocorrem em períodos diferentes durante o ciclo de divisão nuclear. Num determinado momento, durante a fase S uma seqüência determinada atua como molde para a replicação da molécula de DNA, em outra

etapa, geralmente fora da fase S, a mesma seqüência é utilizada como molde para a transcrição da molécula de RNA.

A compreensão bioquímica do processo de síntese de RNA ainda é precária, apesar da lógica inerente ao sistema de informações, algumas peculiaridades do processo evidenciam sua complexidade.

### 3.4 RNA Polimerases

A proteína que na *E. coli* catalisa a síntese de RNA utilizando DNA como molde é a RNA polimerase. Essa holoenzima é composta de cinco polipeptídios ligados não covalentemente: duas subunidades  $\alpha$  com peso molecular de 40.000 Daltons, uma subunidade  $\beta$  com peso molecular de 150.000 Daltons, uma subunidade  $\beta'$  com peso molecular de 160.000 Daltons e uma subunidade  $\sigma$  com peso molecular de 90.000 Daltons (veja a Figura 3.6). A polimerase principal, com capacidade de sintetizar as ligações fosfodiéster, não contém a subunidade  $\sigma$  (sigma) e sua composição é  $\alpha_2\beta\beta'$ . A subunidade  $\sigma$  é necessária para o início da transcrição, enquanto outra proteína, denominada fator rho ( $\rho$ ), é importante para o término da transcrição.

Nas células eucarióticas três tipos principais de RNA polimerases são encontrados, cada um com grande complexidade de subunidades. Essas polimerases podem ser distintas por sua localização celular e pela sensibilidade a uma toxina de fungos, a  $\alpha$ -amanitina. A RNA polimerase I é encontrada no nucléolo, é resistente à  $\alpha$ -amanitina e transcreve os rRNA grandes. A RNA polimerase II é encontrada no nucleoplasma, é sensível a baixas concentrações de  $\alpha$ -amanitina e é a principal enzima responsável pela síntese do RNA heterogêneo nuclear, que é posteriormente processado em mRNA. A RNA polimerase III é também encontrada no nucleoplasma, é sensível a altas concentrações de  $\alpha$ -amanitina e sintetiza o tRNA e rRNA 5S. Também são encontradas RNA polimerases nas mitocôndrias e nos cloroplastos (veja a Tabela 3.1). Existem outras duas enzimas que catalisam a síntese de polímeros de RNA, mas não utilizam DNA como molde.

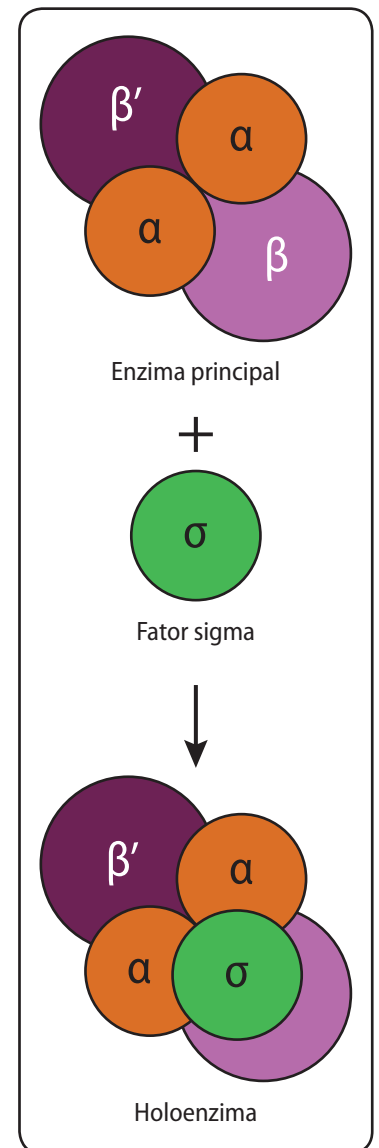


Figura 3.6 Estrutura da RNA polimerase bacteriana

**Tabela 3.1- Enzimas que atuam na transcrição em procariotos e eucariotos**

Procariotos	Eucariotos			
	Tipo	Localização	Síntese	Concentração
RNA polimerase	RNA pol I	Nucléolo	rRNA	50-70%
	RNA pol II	Nucleoplasma	mRNA, snRNA(U)	20-40%
	RNA pol III	Nucleoplasma	RNA pequenos, tRNA, RNA 5S, snRNA U6	10%

### 3.5 RNA Replicases Virais

Alguns bacteriófagos (vírus que infectam bactérias) com RNA icosaédricos possuem genomas com uma única cadeia de RNA linear. Esse RNA é sintetizado por uma enzima grande, a RNA replicase, composta de proteínas virais e do hospedeiro. Curiosamente, na *E. coli* as enzimas utilizadas pela replicase do bacteriófago de RNA estão também envolvidas na síntese de proteínas.

### 3.6 Transcrição do RNA

O processo de transcrição é um processo altamente seletivo que produz repercussões importantes em todas as atividades celulares. Esse processo não ocorre espontaneamente, mas é determinado por uma lógica ainda pouco compreendida. Alguns aspectos desse funcionamento parecem ser programados mostrando convergências importantes com o padrão de desenvolvimento. O desenvolvimento só é possível por esse padrão diferenciado de transcrição. No entanto, há também uma influência de fatores ambientais nesse processo. É mais fácil descrever esse processo em bactérias porque é mais simples que a transcrição em eucariotos.

Tecnicamente, a transcrição é dividida em três etapas:

1. **início** – ligação da RNA polimerase aos promotores e aos iniciadores de cadeia, determina o início da transcrição;
2. **alongamento** – RNA polimerase copia a seqüência do DNA, liberação do fator  $\sigma$ , polimerização 5'→3', velocidade 30 nucleotídeos/segundo, fita molde de DNA 3'→5'; e
3. **término** – sinal de terminação, liberação das moléculas de RNA, liberação da RNA polimerase, molde de DNA conservado.

### 3.6.1 Início

O início da síntese de RNA é um processo complexo que envolve ao menos três estágios: reconhecimento, separação das cadeias e catálise pela formação das ligações fosfodiéster.

A seqüência do DNA que sinaliza o início da transcrição é denominada **promotor**. A RNA polimerase reconhece o promotor como uma estrutura do DNA de fita dupla. Embora possa ocorrer no DNA separação local das cadeias, não é suficiente para expor o número de bases necessárias para o sinal específico de início da síntese de RNA. Assim, a polimerase possivelmente acessa a seqüência promotora por meio das alças da estrutura do DNA. A subunidade  $\sigma$  na *E. coli* tem uma função decisiva nessa etapa do reconhecimento, interagindo covalentemente com o DNA sobre aproximadamente 12 pares de bases (veja a Figura 3.7).

Após a polimerase encontrar a seqüência correta, em virtude da ligação não covalente das alças da estrutura do DNA entre as bases e a enzima, as cadeias devem ser separadas para possibilitar o pareamento dos precursores no molde. Esse processo envolve a formação de ligações iônicas entre os aminoácidos básicos da enzima e os grupos fosfato acídicos do DNA. Utilizando a energia liberada pela formação das ligações iônicas, as cadeias são mantidas separadas. Estima-se que aproximadamente 7 pares de bases, próximos ao local da interação do molde com o precursor, são separados pela polimerase.

Uma seqüência de consenso é o protótipo da seqüência de bases que sumaria as bases encontradas mais freqüentemente em cada posição na análise das seqüências individuais de diversos promotores.

- **Promotores**
- São combinações de curtos elementos de seqüência, normalmente localizados na região imediatamente a montante do gene, freqüentemente a 200 pb do sítio de iniciação da transcrição, e servem para iniciar a transcrição.

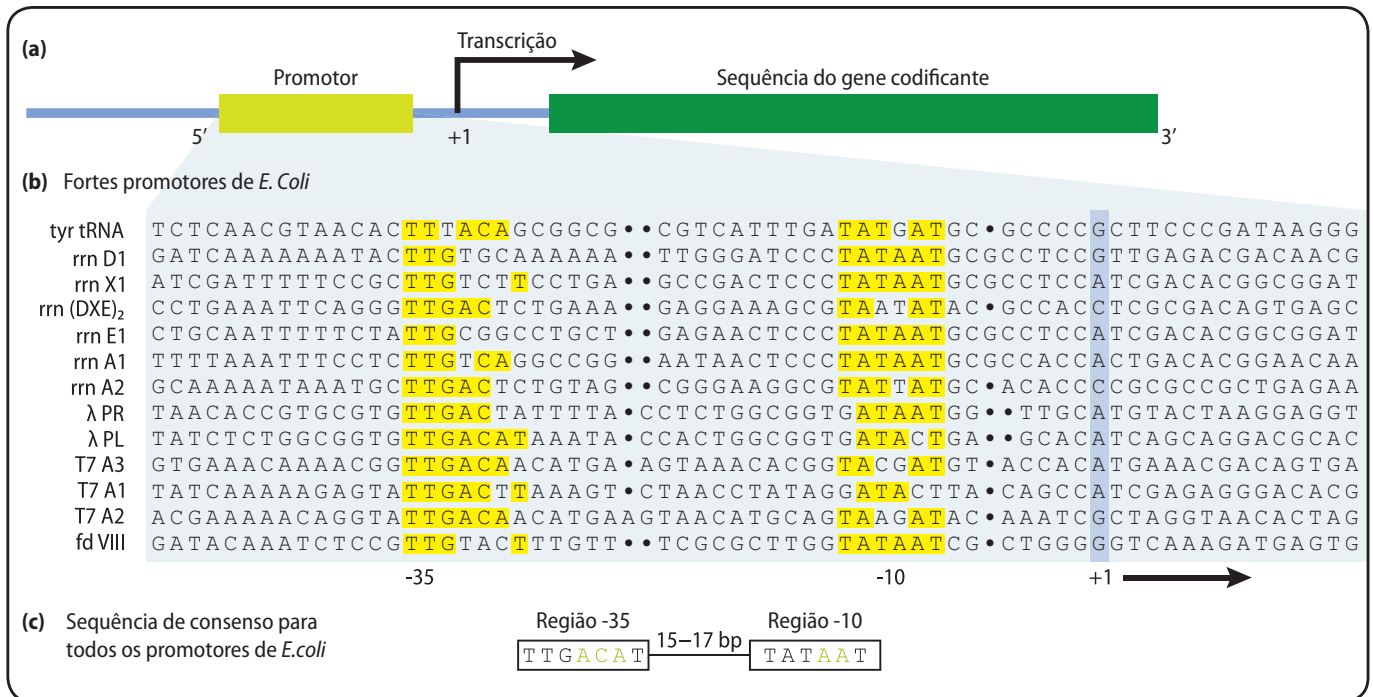


Figura 3.7 - Sequências promotoras de diferentes genes bacterianos

Nos **procariontos**, são reconhecidas duas regiões localizadas, -10 e -35 nucleotídeos antes do início da transcrição (+1).

Nos **eucariotos**, a posição é um pouco diferente, mas a sequência de consenso também apresenta homologia. As sequências de consenso não são idênticas, mas possuem características comuns, com diferenças de poucas bases.

A posição da polimerase no promotor é tal que a primeira ligação fosfodiéster é formada usualmente entre uma purina e o próximo nucleotídeo, ambos ligados por pontes de hidrogênio ao molde. A cadeia nascente tem desta maneira uma terminação 5' trifosfato; nenhum primer é necessário. À medida que o alongamento prossegue, a subunidade  $\sigma$  é dissociada e o restante da síntese é realizado pela polimerase principal.

Nos eucariotos no lugar do fator sigma, a ligação da RNA polimerase ao DNA é mediada por diferentes proteínas denominadas **fatores de transcrição**, sendo altamente seletivas e críticas para o controle do funcionamento gênico.

Enquanto os procariontos possuem um único tipo de RNA polimerase, os eucariotos têm um repertório mais amplo, com RNA polimerases diferenciadas para cada tipo de RNA celular, conforme já mencionado.

A transcrição depende da seqüência promotora. Há necessidade de ativação do promotor por uma proteína, como o fator sigma nos procariotos, para que ocorra a transcrição. Desta forma, a razão de transcrição depende da quantidade de proteína ativadora disponível, quanto maior sua concentração mais rapidamente encontrará e se ligará ao promotor.

A razão de transcrição depende também da potência do promotor. Se a seqüência promotora se encaixa perfeitamente ao sítio de ligação, o gene será transcrito freqüentemente. Por outro lado, se a seqüência promotora é similar ao sítio ideal de ligação, mas não há um encaixe perfeito, a seqüência será transcrita menos freqüentemente, porque a ligação do ativador será mais tênue. A seleção atua a favor da afinidade adequada – nem todos os promotores são sítios de ligação ideais, porque nem todos os genes precisam ser transcritos na razão máxima.

### 3.6.2 Alongamento

O alongamento da cadeia ocorre no grupo hidroxil livre 3'. A informação para a síntese é fornecida pelo DNA molde e o alongamento é efetuado pela enzima RNA polimerase (veja a Figura 3.8).

A relação espacial entre a RNA polimerase e o DNA molde não é conhecida em detalhes. A holoenzima, quando ligada ao DNA,

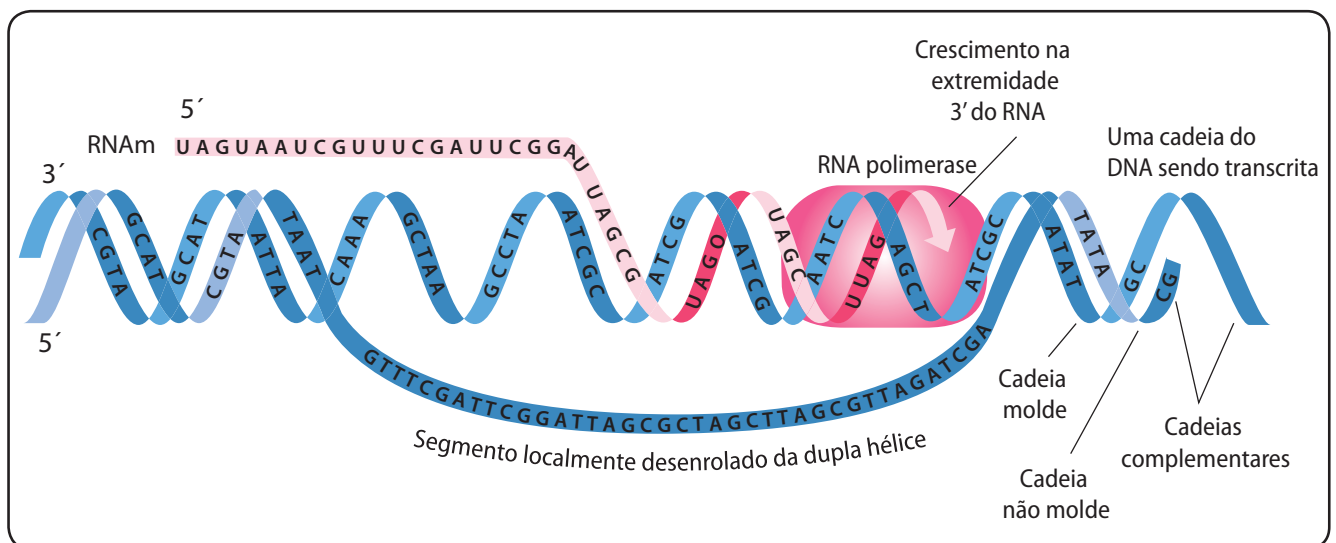


Figura 3.8 - Alongamento da cadeia de RNA

pode proteger da digestão com nucleases um segmento de aproximadamente 65 pares de bases.

O mecanismo de ação de diversos antibióticos se dá por meio do bloqueio na síntese de RNA. Alguns, como a actinomicina-D, bloqueiam a transcrição, intercalando-se entre os pares de bases sucessivos no DNA molde. Outros agentes que atuam desta maneira são a aflatoxina-B e o brometo de etídeo, que em baixas concentrações mostra especificidade para o DNA mitocondrial.

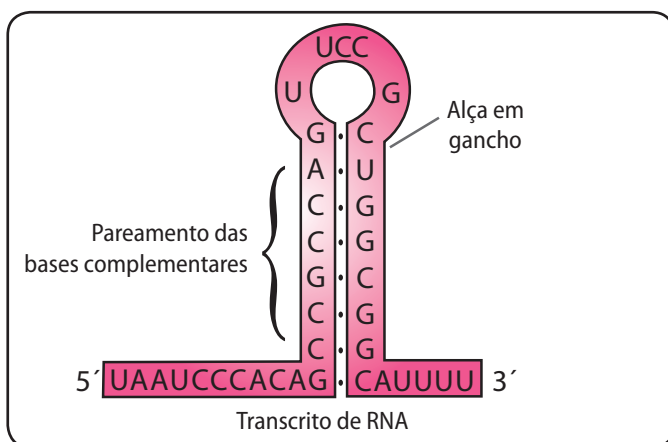
### 3.6.3 Término

Para prosseguir com a síntese de RNA, a polimerase principal e a cadeia nascente de RNA devem manter contato satisfatório com a seqüência de DNA. Quando um segmento de nucleotídeos do DNA com pouca afinidade (seqüência de terminação) é encontrado, há uma tendência para a dissociação da polimerase e do RNA e o término da transcrição. Dois elementos têm uma função importante nesse processo. O primeiro é a estrutura secundária da cadeia nascente de RNA. A maioria das seqüências de término tem uma repetição invertida que permite ao produto a formação de uma estrutura em gancho. Isto provavelmente atua como um freio para a polimerase, causando uma diminuição ou pausa no seu progresso ao longo do DNA molde.

Outro elemento presente nas células bacterianas é a proteína rho ( $\rho$ ), sua função é possibilitar o término quando a seqüência de terminação não é suficiente. Assim, a terminação rho-dependente é verificada nos terminadores fracos (veja a Figura 3.9). A maneira pela qual a proteína rho interage com a RNA polimerase central ainda não foi elucidada.

O processo de transcrição é assim um evento conservativo que não produz modificações na molécula de DNA que serviu de molde, sendo a molécula de RNA resultante complementar à molécula de DNA utilizada como molde para a sua síntese. A cadeia molde tem orientação 3'-5', enquanto a cadeia complementar do DNA

Figura 3.9 - Término da transcrição



é denominada cadeia não molde e tem orientação 5'-3'. Além disso, a cadeia não molde possui a mesma seqüência de nucleotídeos da cadeia de RNA (substituindo-se T por U), sendo também denominada cadeia sense, enquanto a cadeia molde é denominada anti-sense (veja a Figura 3.10).

RNA- 5'→3'

DNA- 3'→5' FITA ANTI-SENSE (fita molde)

DNA- 5'→3' FITA SENSE (possui seqüência similar ao RNA)

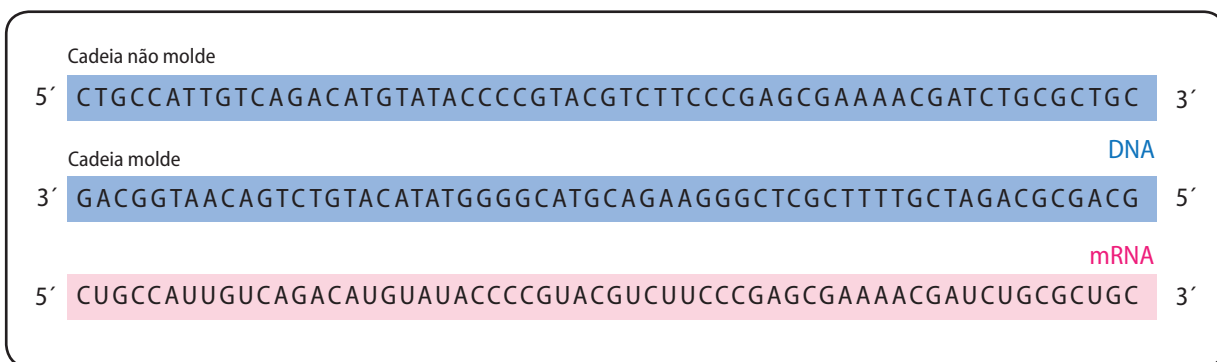


Figura 3.10 - DNA transcrito e RNA resultante

Vale lembrar que apenas uma das fitas serviu de molde para a síntese de RNA. A fita complementar também pode conter informações, mas o conteúdo dessa informação é diferente, o que aumenta o potencial de armazenamento da informação pelo DNA. Os vírus exploram bastante essa potencialidade (veja a Figura 3.11).

Desta forma, a unidade do DNA funcional ou UNIDADE DE TRANSCRIÇÃO é definida como um segmento de DNA (gene ou genes) transcrito em uma única molécula de RNA pela RNA polimerase. Essa unidade é mais ampla que a seqüência transcrita, os promotores fazem parte dessa unidade. Outros elementos genômicos também podem influenciar a freqüência com que cada unidade de transcrição é ativada, como os potenciadores, os operadores e os silenciadores (veja a Figura 3.12).



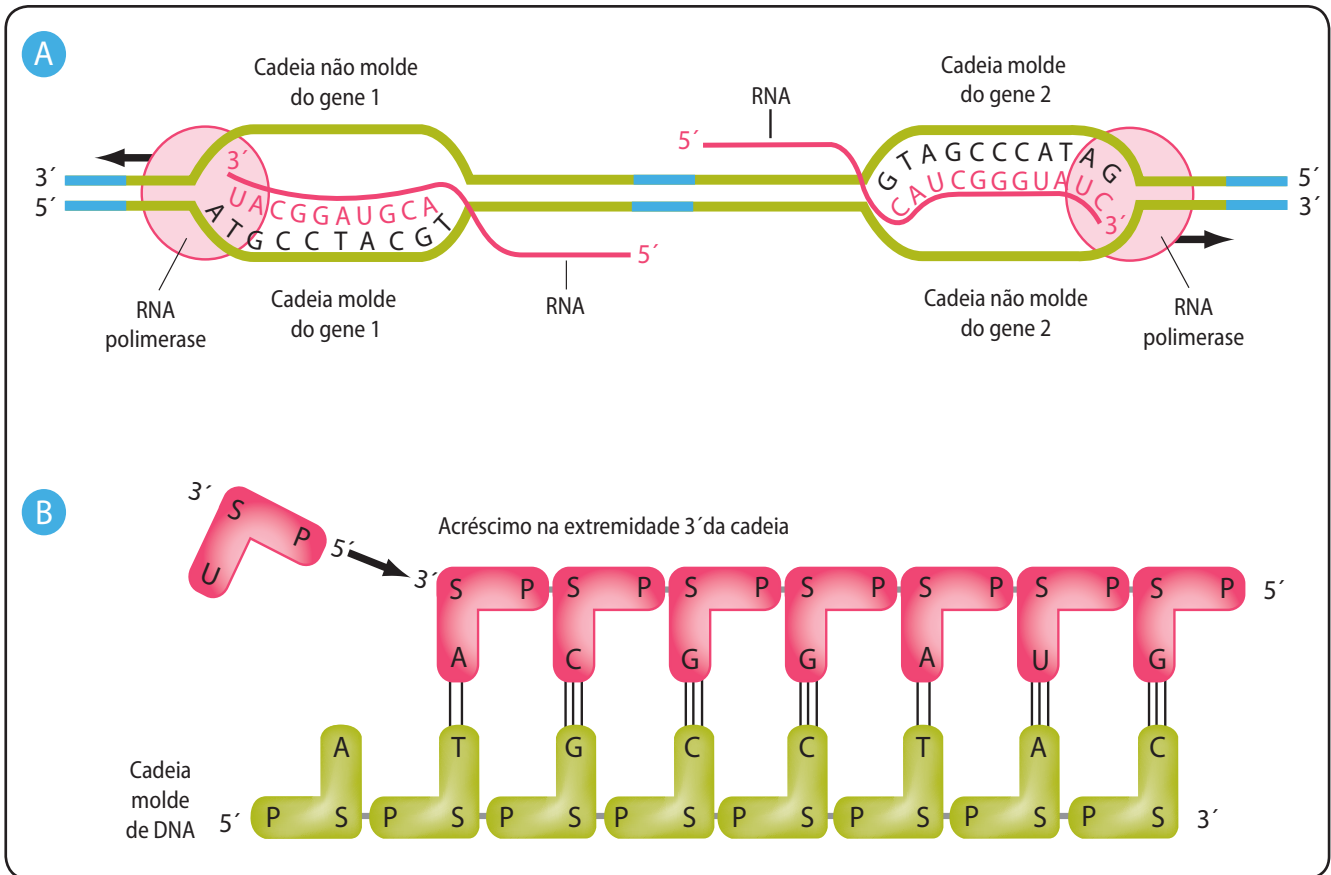


Figura 3.11- Ao longo da molécula de DNA os genes podem apresentar orientações opostas, a informação das duas fitas pode ser transcrita separadamente.

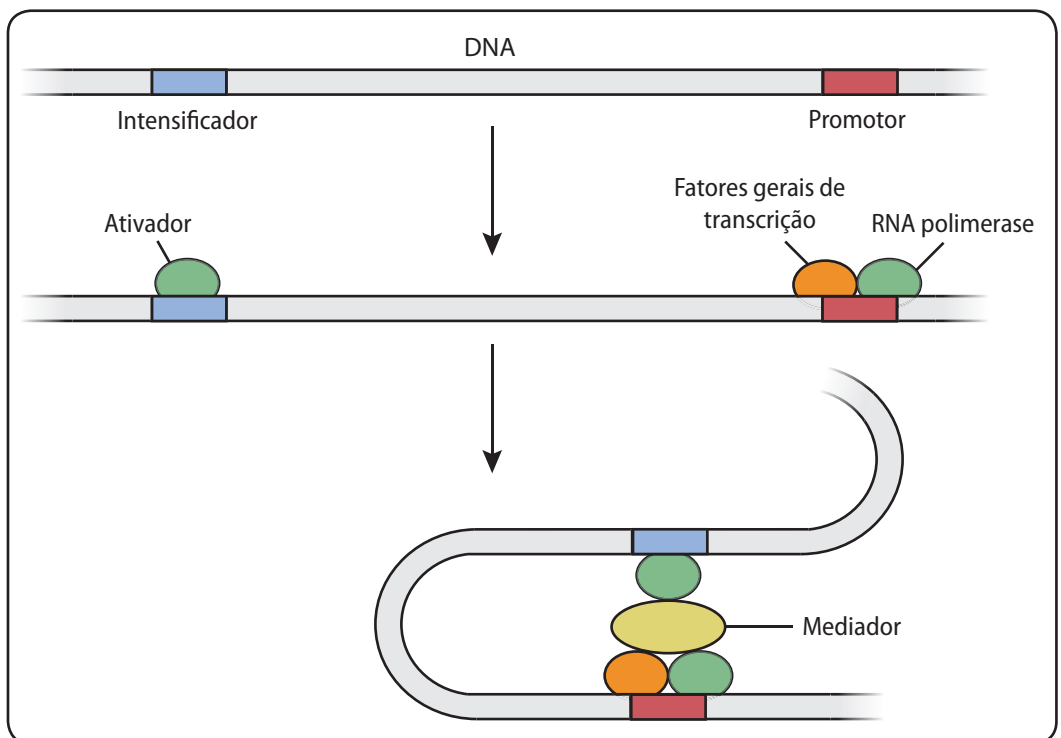


Figura 3.12 - Outros fatores que influenciam a transcrição

### 3.7 Tipos de RNA

Embora quimicamente o RNA seja similar, funcionalmente são reconhecidos diversos tipos de RNA. Essa diversidade evidencia a importância dessas moléculas para o funcionamento da célula, desempenhando atividades estruturais e metabólicas, além do seu conteúdo informacional. Assim, admite-se que os sistemas biológicos se originaram a partir de moléculas de RNA, evoluindo para um sistema de armazenamento de informação mais estável, o DNA (veja a Figura 3.13).

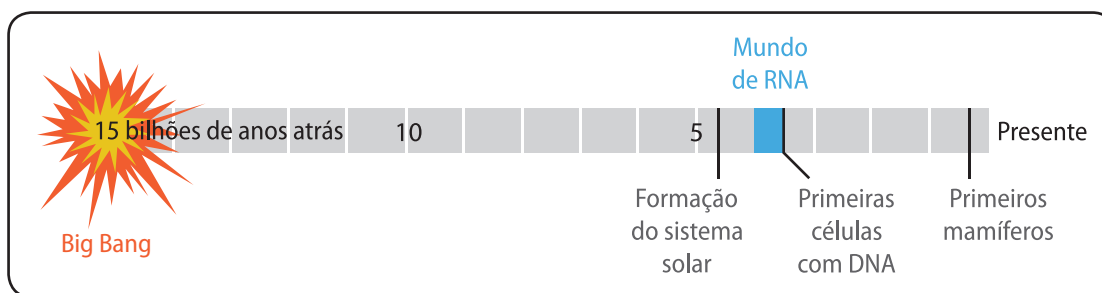


Figura 3.13 - Origem da vida a partir do RNA

O DNA pode ser considerado um tipo de molécula que na maioria das vezes armazena a informação e a torna acessível para a transcrição no RNA. Isso não ocorre com o RNA, que pode ser dividido em diversas classes com base na sua função e nas propriedades físicas. A evidência dessas diferentes categorias foi primeiro obtida pela diferença nas propriedades estruturais.

Todo RNA presente no citoplasma de células eucarióticas é sintetizado no núcleo. No entanto, a maior parte do RNA produzido nunca deixa o núcleo e, portanto, não participa diretamente da síntese de proteínas. Embora a estrutura secundária e terciária das moléculas de RNA seja importante para sua função, essas características estruturais são determinadas pela seqüência de nucleotídeos, ou seja, a estrutura primária, como no caso as proteínas.

Alguns dos principais tipos de RNA são:

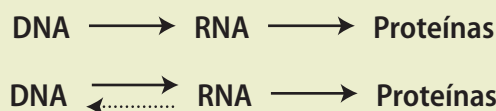
1. **RNA viral** – material genético da maioria dos vírus de plantas e alguns vírus de bactérias e animais, RETROVÍRUS. Ocorrência de transcrição reversa RNA →DNA;

2. **RNA mensageiro (mRNA)** – estrutura heterogênea e muito ativa metabolicamente, com reciclagem rápida. Específica para a síntese de uma determinada proteína;
3. **RNA ribossômico (rRNA)** – tipo de RNA que se sedimenta pela centrifugação de alta velocidade, parte da partícula ribossômica. Presente nos ribossomos, nas organelas celulares compostas de RNA e nas proteínas onde ocorre a síntese de proteínas, e na interação entre tRNA e mRNA;
4. **RNA transportador (tRNA)** – permanece no sobrenadante após sedimentação dos ribossomos, é relativamente pequeno, com 70 a 90 nucleotídeos, e estável quimicamente; e
5. **RNA de interferência** – atua no controle de diversos RNAs.

Ao se analisarem os RNAs celulares e comparar sua homologia ao DNA, verificou-se que o RNA não corresponde exatamente ao DNA molde, evidenciando a ocorrência de processamento do transcrito primário. A amplitude e a extensão desse processamento são diferentes nos diversos tipos de RNA, sendo também diferenciadas em procariotos e eucariotos. O significado desse processamento ainda não é compreendido integralmente.

### 3.7.1 RNA Viral

O RNA de fita simples é o material hereditário da maior parte dos vírus de vegetais e diversos vírus de bactérias e animais. Dois grupos de vírus animais denominados reovírus apresentam RNA de fita dupla. A replicação do RNA viral utiliza uma grande variedade de mecanismos, os quais refletem as diversas maneiras de utilizar o RNA diretamente como genoma, o mecanismo não usual de transcrição de RNA em DNA – com o auxílio de uma enzima denominada transcriptase reversa – e a formação de uma cadeia dupla de DNA. Essa observação causou sensação quando foi descoberta, modificando o Dogma Central da Biologia, com respeito ao fluxo de informação entre macromoléculas:



A seta do RNA para o DNA é interrompida porque o fluxo reverso da informação do RNA para o DNA não ocorre freqüentemente. A transcrição reversa também ocorre em células eucarióticas não infectadas, mas a sua função normal ainda não é conhecida.

Considera-se que o uso do RNA, nos vírus de RNA, para armazenamento da informação pode representar um estágio evolutivamente primitivo de transferência de informação. Parece plausível supor que um sistema primitivo de síntese de proteínas poderia envolver uma forma auto-replicante de RNA e que a dupla hélice de DNA, com a vantagem da estabilidade, representaria um estágio posterior.

### 3.7.2 RNA Mensageiro

No final da década de 50, quando se evidenciou que os ribossomos estavam envolvidos na síntese de proteínas, surgiu a questão de como os ribossomos continham informação, ou seja, se cada tipo específico de proteína seria sintetizada por um ribossomo especial com a informação necessária para a síntese daquela proteína. Experimentalmente, verificou-se que um minuto após a infecção por um bacteriófago, as células de *E. coli* interrompem a síntese das proteínas bacterianas e iniciavam a síntese das proteínas necessárias para a replicação viral. Parecia improvável que num período tão curto pudessem ser sintetizados ribossomos suficientes para possibilitar a síntese maciça das proteínas envolvidas na replicação viral.

Por essas e outras razões, F. Jacob e J. Monod em 1961 propuseram que o DNA deveria atuar como molde para a síntese de um RNA mensageiro especial que se liga ao ribossomo, possibilitando assim a especificidade da síntese protéica. O experimento que confirmou esta hipótese foi realizado no mesmo ano por S. Brenner, F. Jacob e M. Meselson, que demonstraram com o uso de precursores radioativos que após a infecção pelo fago não eram sintetizados ribossomos novos. Ao contrário, era sintetizado um RNA de vida curta que se unia aos ribossomos preexistentes.

Esse fenômeno também foi demonstrado em células de *E. coli* não infectadas por meio da indução enzimática, em que a bactéria

sintetiza rapidamente as enzimas necessárias para metabolizar a lactose quando expostas a esse açúcar em meio de cultura. Novamente, foi a tradução do mRNA recém-sintetizado nos ribossomos preexistentes que deu origem às novas enzimas.

O mRNA no citoplasma é normalmente conectado aos ribossomos. Um dos problemas encontrados para purificar o mRNA é a sua instabilidade. Atualmente existem evidências de que nas células eucarióticas o mRNA é mais estável. Existem também diferenças estruturais entre os mRNA de procariotos e eucariotos.

O mRNA de procariotos varia consideravelmente em tamanho por diversas razões. A primeira razão é que as cadeias polipeptídicas variam em tamanho ( $5 \times 10^3 - 5 \times 10^5$  D), assim, o tamanho do RNA que as codifica deve variar na proporção de 10 D de RNA para cada Dalton de proteína. Segundo, o seqüenciamento de diversos RNAs mensageiros demonstrou que há um segmento iniciador na extremidade 5' da molécula (a tradução se processa no sentido 5' para 3') e esse iniciador varia em tamanho e seqüência nos diferentes tipos de mRNA. Terceiro, existem vários casos em que uma molécula de mRNA codifica simultaneamente várias cadeias polipeptídicas – RNA policistrônico. Normalmente, as proteínas codificadas por uma única molécula de mRNA têm funções metabólicas relacionadas. Por exemplo, uma molécula de mRNA de  $4 \times 10^6$  Daltons de massa codifica dez enzimas envolvidas no metabolismo da histidina.

Embora o mRNA das células eucarióticas tenha estabilidade suficiente para possibilitar sua purificação, é ainda muito menos estável que os outros tipos de RNA. Entretanto, em células metabolicamente inativas como ovócitos não fertilizados, ou em células especializadas que realizam síntese protéica em razões constantes como células do fígado e glóbulos vermelhos imaturos, as moléculas de mRNA podem ser estáveis por vários dias.

Assim como nas células procarióticas, o tamanho do mRNA citoplasmático é bastante variável. No entanto, é raro encontrar um mRNA que codifique mais de uma proteína. Também ocorrem iniciadores na extremidade 5' e trailers na extremidade 3' não traduzidos. Todos os mRNA citoplasmáticos contêm na extremidade 3' uma cadeia de aproximadamente 100-300 resíduos de adenina,

a cauda poli A. A função da **cauda poli A**, sintetizada no núcleo, não é conhecida, embora a inibição da adição da cauda poli A possa impedir o transporte do mRNA do núcleo para o citoplasma. A estabilidade de alguns mRNA quando injetados nas células é influenciada pelo tamanho da cauda poli A. Entre os mRNA que não possuem cauda poli A estão aqueles que codificam as histonas.

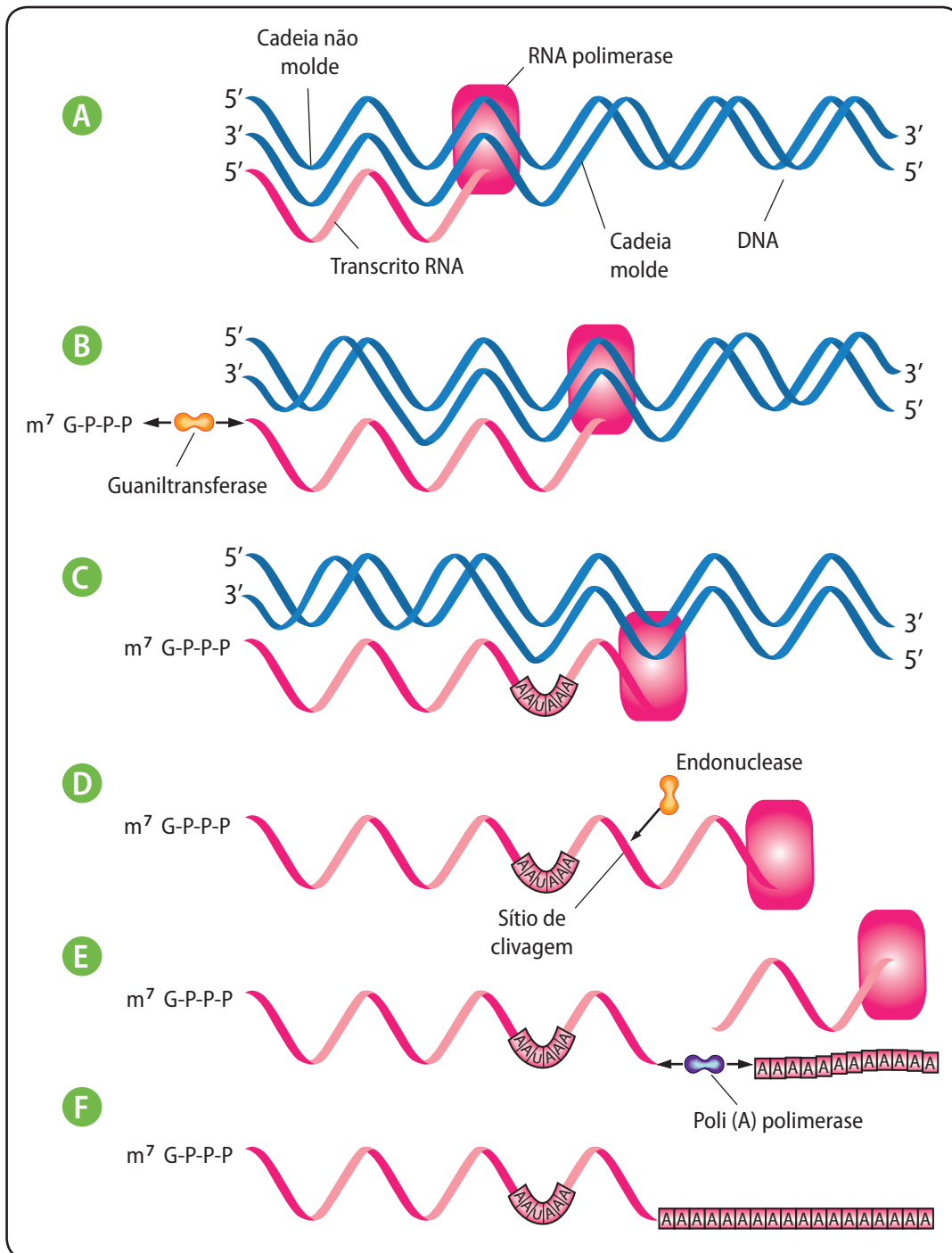


Figura 3.14 - Estrutura do RNA mensageiro maduro

Finalmente, na terminação 5' do mRNA de eucariotos a molécula apresenta um nucleotídeo guanina unido por uma ligação fosfodiéster não usual, com uma ligação 5'-5'. Grupos metil são adicionados a essa guanina e freqüentemente aos dois nucleotídeos adjacentes. Essa modificação é denominada **cap** e sua função não é compreendida (veja a Figura 3.14).

A estrutura dos mRNA nos cloroplastos assemelha-se aos mRNA procarióticos: não estão presentes o cap e a cauda poli A. A cauda poli A é encontrada nos mRNA de mitocôndrias, mas, ao contrário dos mRNA citoplasmáticos, a estrutura 5' cap não está presente.

### A) Processamento do RNA

O processamento do RNA é um dos assuntos mais excitantes da Biologia, uma vez que a detecção desse fenômeno foi totalmente inesperada.

Na realidade, muito pouco a respeito do processamento do mRNA era conhecido há poucos anos.

Enquanto nos procariotos e nos bacteriófagos de RNA o processamento do RNA transcrito é relativamente simples, o mesmo não ocorre nos eucariotos.

Os transcritos primários nos núcleos de eucariotos destinados a tornarem-se mRNA maduros são maiores e têm uma grande variabilidade de tamanho. Esse RNA, inicialmente denominado RNA heterogêneo nuclear, atormentou os biólogos por anos, desde que possui uma massa aproximadamente 10 vezes maior que o RNA maduro.

O **processamento** resulta na remoção de seqüências das extremidades e de seqüências internas do RNA. Vamos analisar um caso relativamente simples – o RNA mensageiro da cadeia  $\beta$  da globina de camundongo, proteína que é parte da hemoglobina. O mRNA que codifica a cadeia  $\beta$  foi isolado dos reticulócitos como um RNA 9S e a seguir hibridado com o DNA genômico, quando foram observadas, por meio de microscopia eletrônica, alças de cadeias simples ao lado de estruturas em fita dupla (veja a Figura 3.15). Essas alças são devido às seqüências intercaladas (íntrons) do DNA que não estão presentes no mRNA maduro.

*Nos organismos eucariotos, uma porção do DNA com a mensagem para a síntese de uma cadeia polipeptídica é, geralmente, interrompida por certas porções que não têm função codificante e, portanto, não aparecem representadas nas proteínas. Tais porções sem sentido são referidas como íntrons, enquanto as regiões com informação codificada são os éxons.*

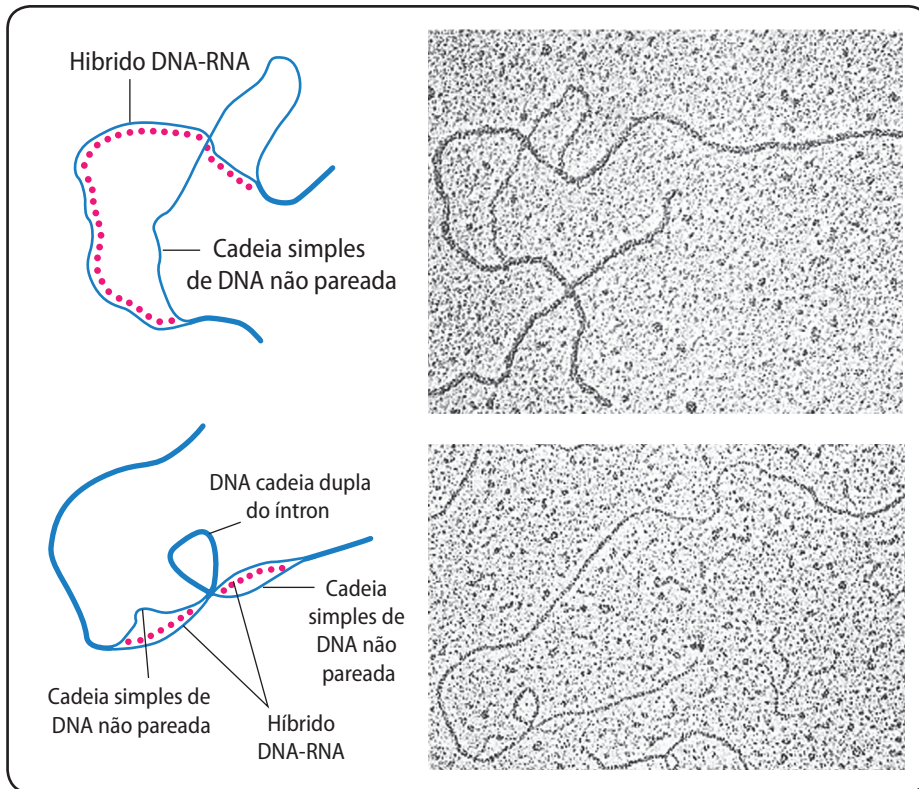


Figura 3.15 - Microfotografia eletrônica do híbrido DNA/RNA

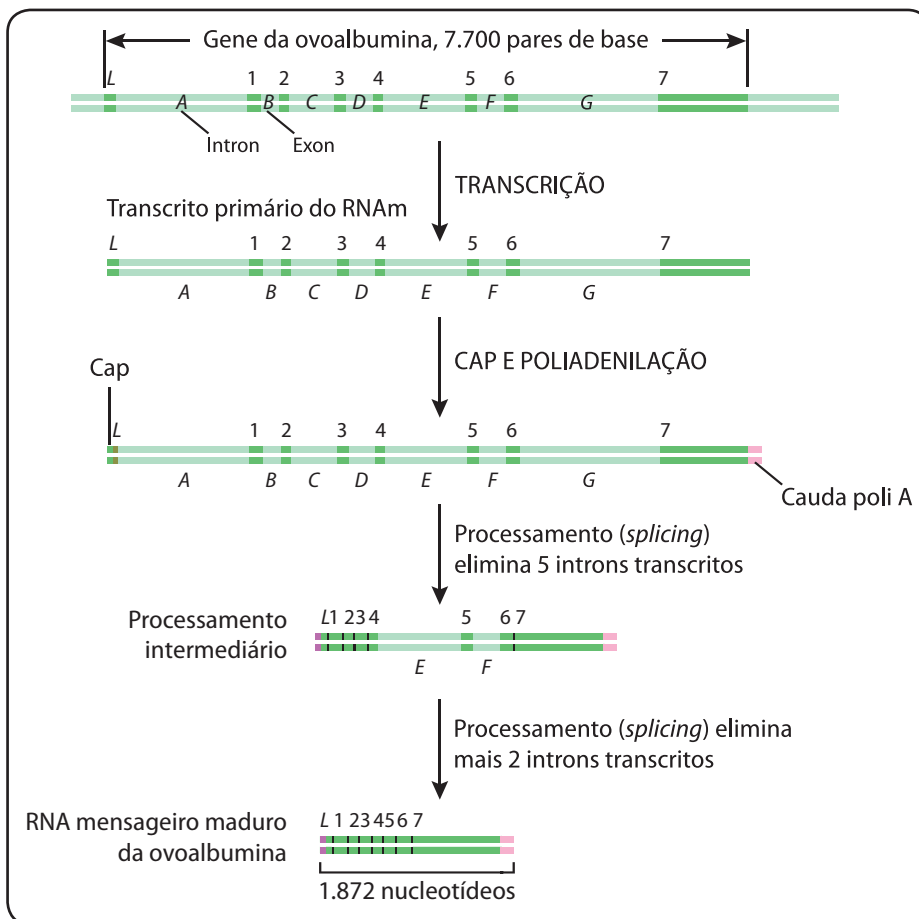


Figura 3.16 - Processamento do RNA da betaglobina



Uma dessas alças com 646 pares de bases está localizada entre os códons para os aminoácidos 30 e 31 do mRNA. Uma outra, com 646 pares de bases, está localizada entre os códons 104 e 105 (veja a Figura 3.16). Esses detalhes foram descobertos pela comparação da seqüência do DNA da  $\beta$  globina com a seqüência de aminoácidos da proteína.

Embora não previsto teoricamente, esse processo é bastante generalizado.

Apesar de os precursores do RNA mensageiro terem sido os primeiros em que as seqüências intercaladas foram identificadas, essas também são observadas nos genes dos rRNA e tRNA de eucariotos.

A questão que foi levantada inicialmente é se essas seqüências eram transcritas e depois removidas ou simplesmente ignoradas pelas RNA polimerases. O progresso na compreensão desse processo foi possível por meio da análise das pequenas seqüências intercalares do gene de tRNA acceptor da tirosina de levedura.

Há uma seqüência intercalada de 14 pares de bases próxima ao anticódon. A observação importante é que a seqüência de 14 bases estava presente no precursor, sendo, portanto, sintetizada durante a transcrição. A análise dos precursores isolados mostrou que as seqüências são removidas e as extremidades ligadas covalentemente numa seqüência de reações denominada *splicing*. As regiões que permanecem no RNA maduro são denominadas **éxons**, enquanto as regiões removidas são denominadas **íntrons**.

### B) Seleção do Local de Splicing

A questão central do conceito de splicing é a maneira pela qual o local é reconhecido e removido e a seqüência de RNA religada. No caso dos precursores do RNA mensageiro, a informação para a seleção do local de splicing provavelmente está localizada na seqüência primária dos locais de junção. A análise da estrutura primária de diversos genes com seqüências intercaladas levou à conclusão que existe uma seqüência de consenso na junção íntron-éxon (veja a Figura 3.17).

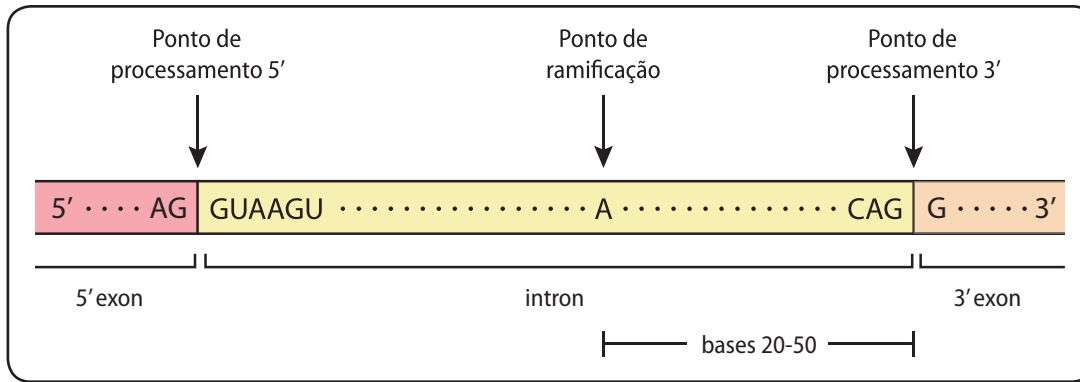


Figura 3.17 - Remoção dos íntrons

As seqüências intercalares são de alguma forma importantes para as células; as seqüências intercaladas foram encontradas em outros genes da globina  $\beta$  (de coelho e humanos) exatamente nas mesmas posições do gene do camundongo. As seqüências de bases dos íntrons de camundongos e coelho não são similares, exceto nas extremidades, mas sua posição nos genes é idêntica. A observação de seqüências divergentes nos íntrons é consistente com o

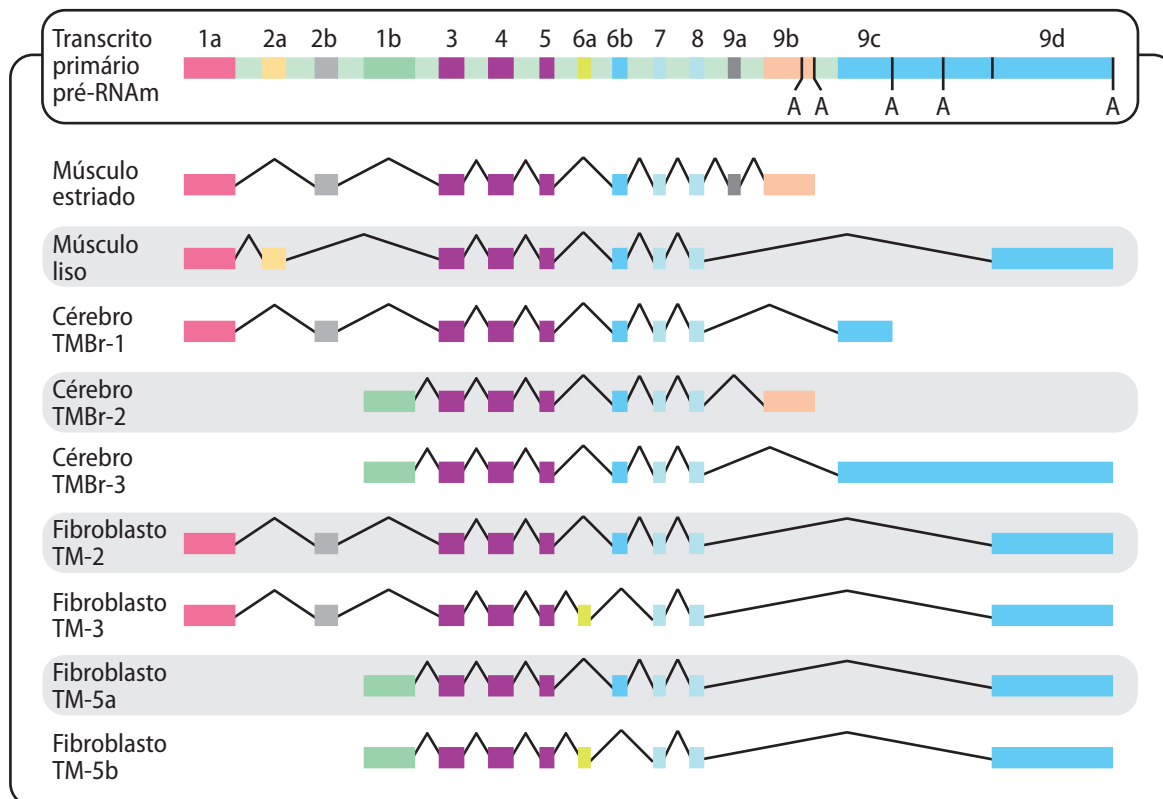


Figura 3.18 - Splicing alternativo do gene da actina humano

acúmulo de mutações durante a evolução. Quando se observam as características estruturais – como a localização dos íntrons nos genes –, que são conservadas pelas duas espécies durante a evolução, parece provável que a perda dessas características pode selecionar negativamente, especialmente se a distância evolutiva das duas espécies é grande.

Por outro lado, o processo de edição pode proporcionar uma diversidade adicional por meio da edição diferenciada de um mesmo gene, ou **splicing alternativo**, como verificado em alguns genes (veja a Figura 3.18).

### 3.7.3 RNA Ribossômico

O rRNA é o RNA celular majoritário, compreendendo 50-70% do RNA celular, sendo um RNA que apresenta grande estabilidade.

Os ribossomos constituem a maquinaria de síntese protéica das células. São formados por duas subunidades, diferentes em tamanho. Nos procariotos, a **subunidade maior** contém duas cadeias de RNA, uma grande com **23S** e outra menor com **5S**, e **34 moléculas de proteínas** diferentes; a subunidade pequena contém uma cadeia de RNA de **16S** e **21 moléculas protéicas**. Todas as cadeias de RNA são de cadeia simples (veja a Figura 3.19).

Os ribossomos dos procariotos são menores que os dos eucariotos. Os ribossomos das mitocôndrias e dos cloroplastos assemelham-se aos de procariotos – outra evidência consistente com a teoria endossimbiótica de origem das organelas eucarióticas.

A seqüência da molécula de RNA 5S de *E. coli* foi determinada por G. Brownlee e F. Sanger em 1966. A partir disso, diversas moléculas de RNA 5S de várias origens foram seqüenciadas, e os dados comparativos mostraram algumas características comuns dessa molécula:

- nos procariotos ocorrem quatro regiões helicoidais. Admita-se que são universais porque formam pares de bases mesmo quando a seqüência precisa dessas regiões não é conservada;
- em todos procariotos estudados, na posição 43-47 a seqüência é CGAAC. Esses nucleotídeos podem interagir com a seqüência

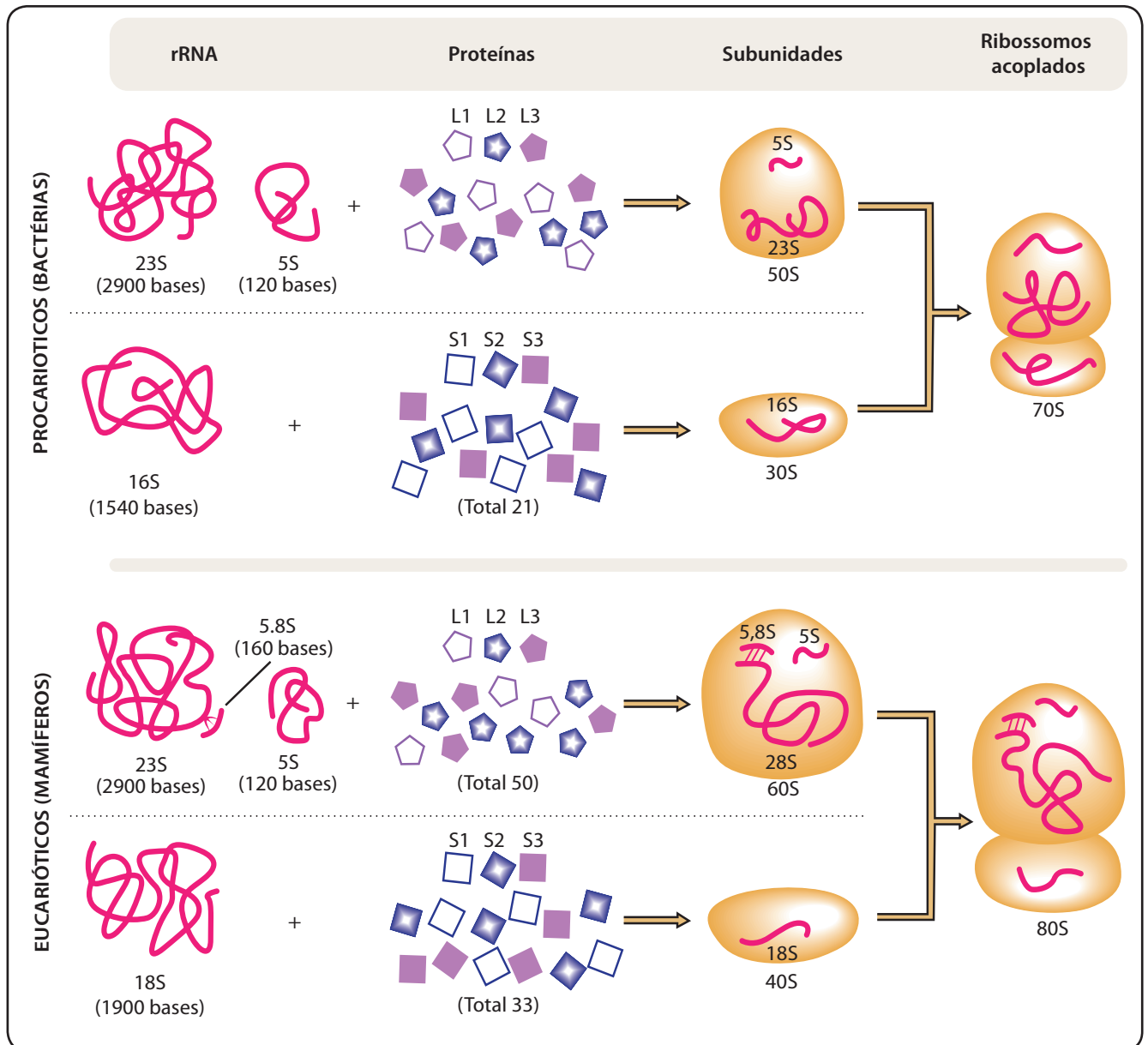


Figura 3.19 - Constituição dos ribossomos de procariotos e eucariotos

complementar comum GT $\psi$ CG na alça III do tRNA, podendo assim representar um sítio de interação entre o ribossomo e o tRNA;

- todos os procariotos apresentam uma seqüência altamente conservada (posições 67-80) protegida da hidrólise enzimática *in vitro* por três proteínas ribossômicas 50S. Essa região possivelmente representa um sítio de interação com outros componentes da subunidade 50S do ribossomo.

Da mesma forma, estudos comparativos da molécula de rRNA 16S da subunidade pequena do ribossomo, provenientes de diferentes organismos também mostram diversas regiões conservadas que possivelmente estão relacionadas a alguns aspectos importantes da estrutura e da função dos ribossomos. Um exemplo interessante é uma parte da seqüência 3' dessa molécula, conservada em todos os procariotos, denominada **seqüência de Shine e Delgarno**. As bases dessa região podem se parear com uma seqüência do mRNA viral conhecida como iniciadora. Isto significa que a região iniciadora pode ser reconhecida pelos ribossomos durante as etapas iniciais da síntese protéica. Esse pareamento de bases ocorre em vários tipos diferentes de moléculas mRNA, sugerindo que a extremidade 3' da molécula 16S do rRNA está envolvida no processo de iniciação da síntese de proteínas.

Quando se compara a seqüência da molécula de rRNA 23S da subunidade maior do ribossomo de *E. coli* com outros organismos, não restam dúvidas de que a seqüência dessa molécula possa estar relacionada a diversos aspectos da função do ribossomo.

Ainda se conhece muito pouco a respeito da estrutura secundária do rRNA (estrutura secundária significa a formação de alças entre nucleotídeos próximos aos outros). Visto que os ribossomos são partículas compactas, quase esféricas, é possível que o RNA esteja significativamente enovelado, isto é, que deve estar presente uma estrutura terciária, a qual seria conseqüência da compactação da estrutura de alças. Existem diversas evidências sugerindo uma estrutura terciária do rRNA. Essas evidências indicam que regiões da molécula 16S distantes na estrutura primária interagem com a mesma proteína no ribossomo. No entanto, ainda é necessário um número maior de dados experimentais para que seja possível identificar a estrutura tridimensional e as interações que estabilizam a estrutura terciária do rRNA.

Na *E. coli* existem sete genes separados para os rRNA. Em cada gene são encontradas seqüências para os três rRNA na ordem 16S, 23S e 5S. Esse RNA transcrito tem 30S, aproximadamente 6.000 bases e uma quantidade considerável de seqüências excedentes que são removidas durante a maturação. Esse processamento ocorre em geral concomitantemente à síntese.

Também ocorrem modificações nos nucleotídeos durante o processamento do transcrito primário, as quais resultam na presença de **bases modificadas** no RNA maduro.

Nos eucariotos o rRNA também é sintetizado a partir de moléculas precursoras maiores. A molécula precursora de 45S é clivada numa intermediária com 41S, e a seguir sucessivamente nos rRNA 28S, 18S e 5,8 S (veja a Figura 3.20). O rRNA 5,6 se origina de um precursor diferente. A clivagem de maturação ocorre preferencialmente em seqüências não metiladas. Não se sabe se as seqüências removidas durante o processamento no núcleo contêm tRNA ou outros RNAs com significado funcional.

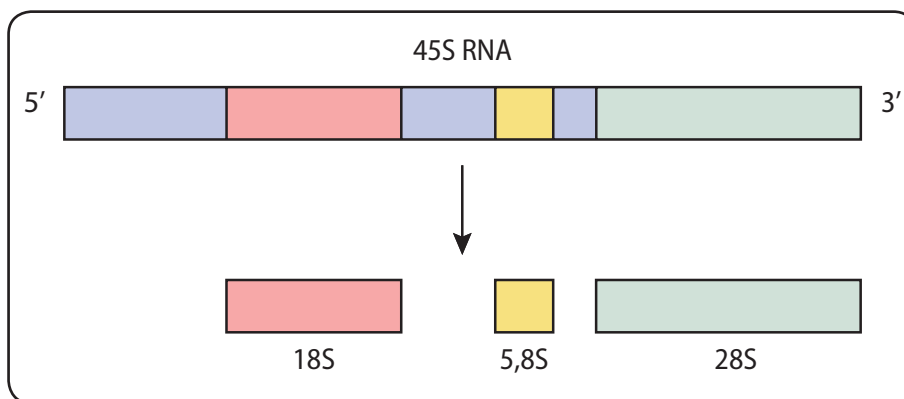


Figura 3.20 - Processamento do rRNA precursor de eucariotos

### Multiplicidade de Moldes para o rRNA

A necessidade de um grande número de ribossomos para a síntese protéica significa que um número igualmente grande de moléculas de rRNA devem ser sintetizadas. Se houver apenas uma cópia da seqüência de DNA específica para o rRNA (chamada rDNA), mesmo que diversas moléculas de RNA polimerase possam operar simultaneamente ao longo do molde, existem limites na velocidade em que o rRNA pode ser produzido. Esses limites estão abaixo do número necessário mesmo por células bacterianas, e a solução universal para os problemas é a multiplicação do número de genes de rRNA.

A hibridação do rRNA no genoma de *E. coli* fornece uma estimativa de sete cópias das seqüências complementares no DNA.

Em células animais a estimativa é de diversas centenas a alguns milhares, e em algumas plantas pode chegar a mais de 10 mil cópias dos genes de rRNA (veja a Tabela 3.2).

Tabela 3.2 - Número de genes de RNA em diferentes organismos			
Espécie	Genes 18S/28S	Genes 5S	Genes tRNA
<i>E. coli</i>	7	7	60
<i>S. cerevisiae</i>	140	140	250
<i>D. discoideum</i>	180	180	
<i>D. melanogaster</i> (X) (Y)	250 150	165	850
Humano	280	2.000	1.300
<i>X. laevis</i>	450	24.000	1.150

As cópias extras estão associadas em regiões especiais de um ou mais cromossomos, denominadas regiões organizadoras do nucléolo. Essas, por sua vez, estão associadas durante a síntese de rRNA dentro do nucléolo da célula (veja a Figura 3.21). Em cada organizador nucleolar, grupos de genes para rRNA estão arranjados em unidades repetidas em tandem com regiões não transcritas de tamanho variável entre eles, denominadas espaçadores.

Nem todas as cópias dos genes de rRNA estão localizadas nos cromossomos. Em alguns casos especializados há amplificação do número de genes – síntese de cópias extracromossômicas. Durante a oogênese em anfíbios a amplificação ocorre em quantidades extremas. O ovo em desenvolvimento contém ribossomos em número suficiente para toda a síntese protéica do período inicial de desenvolvimento embrionário, nenhum ribossomo novo é sintetizado nesse período. Durante a maturação do oócito de anfíbio há síntese de rRNA a partir de um número de cópias de genes de rRNA estimado em  $10^6$ .

O rRNA 5S dos eucariotos não é sintetizado no nucléolo.

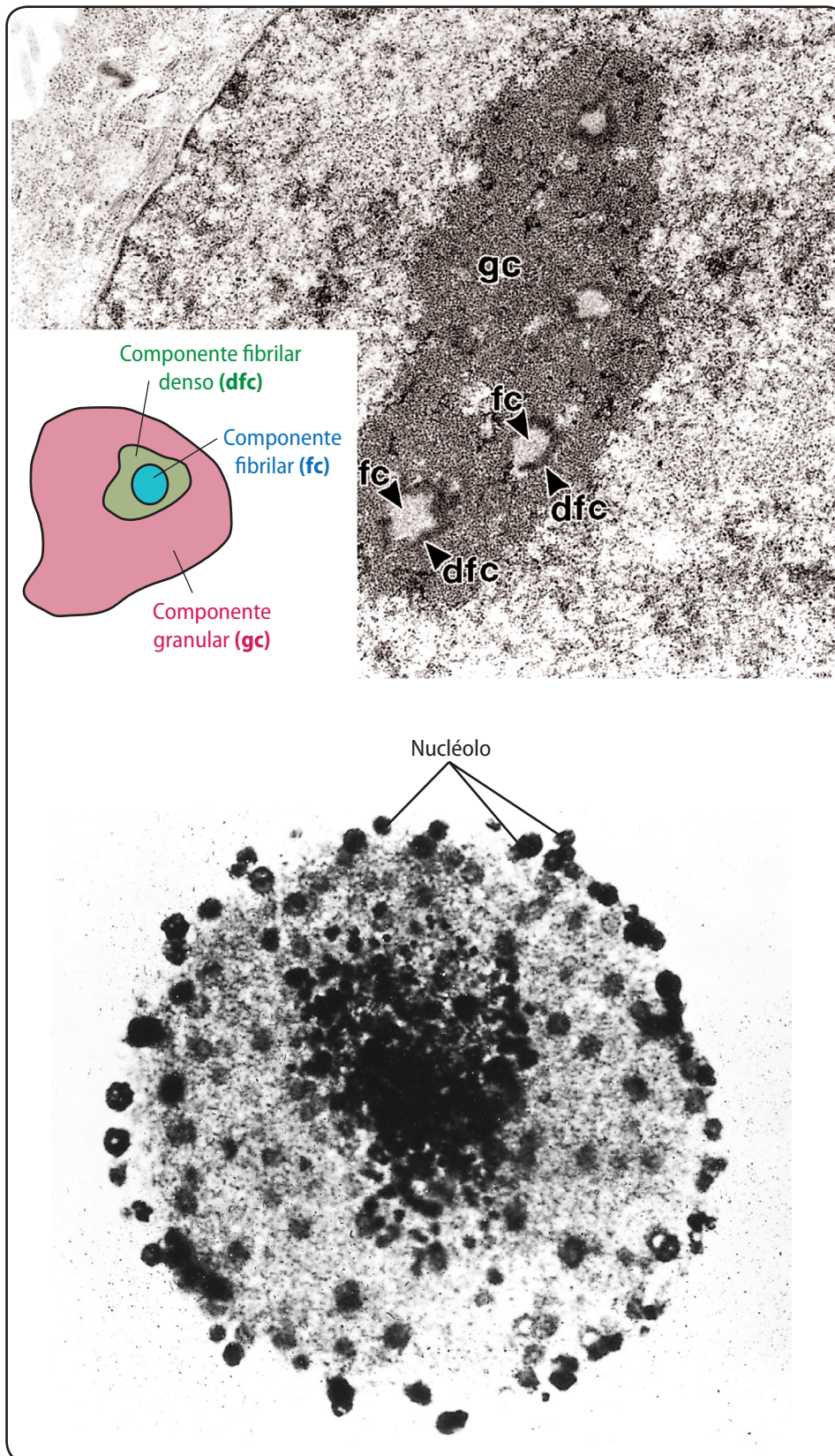


Figura 3.21 - Estrutura do nucléolo



### 3.7.4 RNA Transportador

Uma vez que ficou claro que o mRNA é o molde que fornece a informação para a síntese de proteínas, surgiu a questão de como os nucleotídeos no mRNA selecionam ou interagem especificamente com os aminoácidos adequados. F. Crick, baseado em argumentos teóricos, propôs que seria improvável que os nucleotídeos interagissem diretamente e de uma maneira específica com os aminoácidos. Assim, seria necessária uma molécula adaptadora relativamente pequena que poderia reconhecer o código individual de cada aminoácido. Também admitiu que o material ideal para essa adaptação poderia ser um ácido nucléico (presumivelmente RNA), o qual utilizaria o pareamento de bases para essa especificidade. Essa hipótese, agora confirmada, resolveu o problema da natureza da interação entre o adaptador e o mensageiro.

Logo após essa proposição, M. Hoagland purificou uma fração de RNA com as propriedades esperadas de molécula adaptadora. Quando um homogenado celular é centrifugado em alta velocidade (100.000g por 120 minutos), o RNA conjugado é sedimentado com os ribossomos e o sobrenadante contém RNA não ligado, o qual costuma ser denominado RNA solúvel. A maioria, e possivelmente todo, do RNA nesse sobrenadante tem a função de transferência – combina-se especificamente com aminoácidos ou peptídeos e interage especificamente com o códon do mRNA. Podemos distinguir os tRNA com base na sua afinidade pelos 20 aminoácidos conectados, assim, o tRNA para tirosina pode se conectar apenas à tirosina. A estrutura secundária do tRNA é crítica para a compreensão de sua função (veja a Figura 3.22).

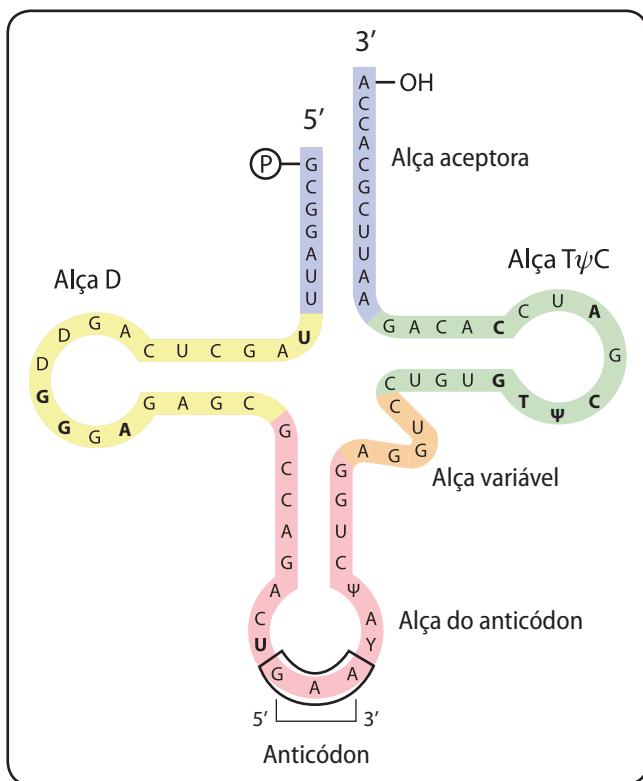


Figura 3.22 - Estrutura secundária do tRNA

Em 1965, R. Holley e colaboradores obtiveram a seqüência do tRNA da alanina em fungos. Foi a primeira seqüência completa de um ácido nucléico estabelecida.

Já havia sido verificado que o tRNA possui **bases não usuais**. Essas bases diferem das tradicionais A, C, G e U, possuindo um ou mais grupos metil substituídos em diversas posições na estrutura em anel (veja a Figura 3.23). Existem aproximadamente 50 dessas bases bizarras.

No artigo original, Holley e colaboradores determinaram diversas possibilidades de pareamento intramolecular, concluindo que uma estrutura em folha de trevo poderia maximizar essas possibilidades. As análises da estrutura primária de outras moléculas de tRNA citoplasmáticos confirmaram essa proposição, e para cada seqüência determinada subsequentemente o padrão de folha de trevo demonstrou uma estrutura com o máximo de possibilidades de pareamento de bases. Embora a seqüência possa variar consideravelmente de uma molécula para outra, o pareamento que forma a estrutura não varia.

Outra questão que Holley e colaboradores consideraram quando propuseram a estrutura em forma de folha de trevo foi a localização do anticódon. A extremidade CCA<sub>OH</sub> 3' da molécula já havia sido identificada como o sítio de conexão com o aminoácido, e

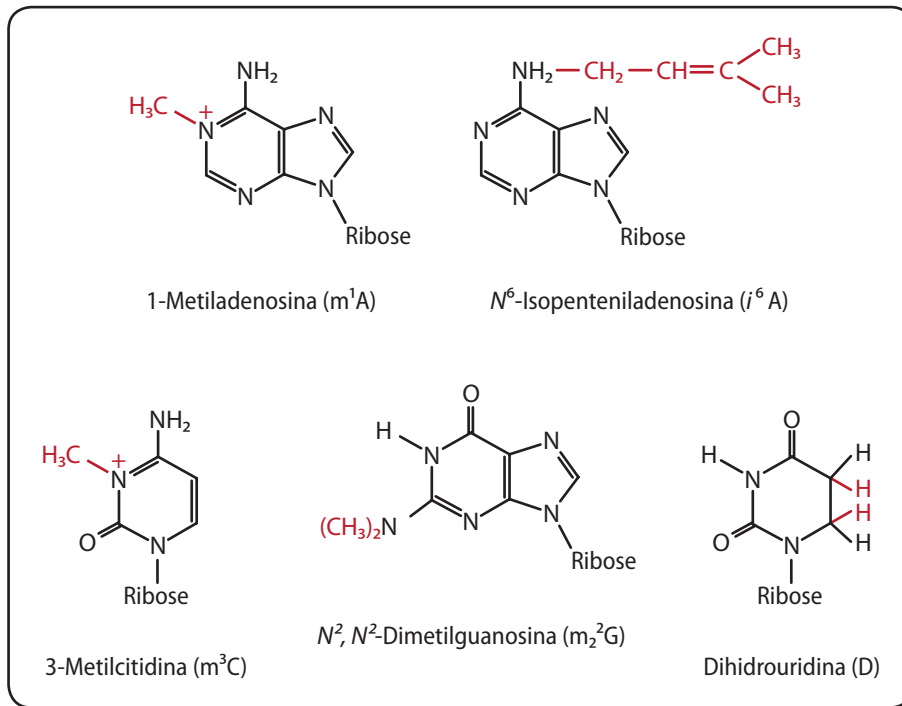


Figura 3.23 - Bases não usuais presentes na molécula de tRNA

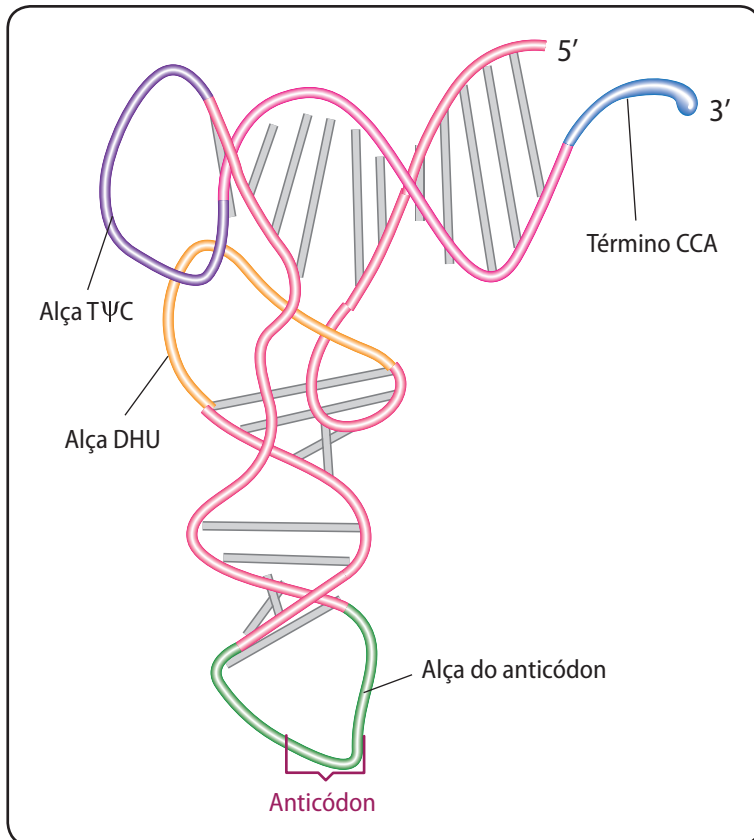


Figura 3.24 - Estrutura tridimensional do RNAt.

parecia lógico supor que o anticódon deveria estar localizado na extremidade oposta da folha de trevo. Nesta época, já eram conhecidos os quatro tripletes do mRNA que codificam a alanina: GCU, GCC, GCA e GCG. Portanto, o anticódon deveria ter a seqüência complementar 3'-CGX-5', e na seqüência do tRNA da alanina foi identificada a seqüência CGI na extremidade oposta da região de ligação ao aminoácido. Em virtude de a inosina (I) ser análoga à guanina, seria esperada sua capacidade de pareamento com C e, portanto, poderia ser previsto que o código no mRNA para esse tRNA da alanina seria GCC.

As moléculas de tRNA possuem diversas características comuns: o

sítio de ligação ao aminoácido, que é invariável; um braço extra variável; e as alças DHU e TΨC com sua composição de bases altamente conservadas. A relativa invariabilidade dessas duas alças sugeriu que teriam uma função importante no empacotamento da molécula numa estrutura terciária precisa, comum a todas as moléculas de tRNA citoplasmáticas. A presença de bases não usuais no tRNA levou à sugestão de que poderiam conferir ao RNA uma propriedade similar às proteínas na formação de uma estrutura tridimensional precisa (veja a Figura 3.24).

Estudos posteriores demonstraram a função dessa estrutura tridimensional:

1. a principal característica da estrutura tridimensional do tRNA é a sua forma em L, com a seqüência aceptora de aminoácidos localizada numa extremidade e o anticódon na outra;
2. as alças DHU e TΨC são empacotadas interagindo fortemente com cada outra e estabilizando a estrutura em forma de L;

3. a estrutura terciária da molécula é estabilizada por pontes de hidrogênio não usuais, ou seja, as pontes de hidrogênio ocorrem de uma maneira diferente da verificada no DNA. Na estrutura terciária as pontes estruturais envolvem algumas das bases conservadas e algumas das bases variáveis. Uma vez que a maioria das bases conservadas é modificada, ficou confirmada a proposição de que as modificações têm função no empacotamento do tRNA na estrutura tridimensional;
4. a alça do anticódon é mantida por poucas pontes de hidrogênio terciárias, sugerindo que pode ocorrer movimento do anticódon durante a síntese de proteínas.

### Síntese de RNA-RNA polimerases: Processo de transcrição

1. Ligação da RNA polimerase na seqüência de DNA do início da transcrição: **promotor**;
  - Polimerases de **procariotos** reconhecem o promotor e ligam-se diretamente;
  - Polimerases de **eucariotos** dependem de **fatores de transcrição**.
2. Separação (desnaturação) da dupla hélice de DNA pela **helicase**;
  - Polimerases de procariotos têm atividade de helicase;
  - Polimerases de eucariotos requerem fator de transcrição específico.
3. Síntese de RNA com NTPs;
4. Término.
  - Mecanismo distinto em procariotos e eucariotos.

**Tabela 3.3 - Diferentes moléculas de RNA presentes na célula**

Tipo de RNA	No. de moléculas diferentes	Tamanho	Distribuição
tRNA	80-100	75-90	P-E
rRNA 5S	1-2	120	P-E
rRNA 5,8S	1	155	E
rRNA 16S	1	1.600	P
rRNA 23S	1	3.200	P
rRNA 18S	1	1.900	E

Tipo de RNA	No. de moléculas diferentes	Tamanho	Distribuição
rRNA 28S	1	5.000	E
hnRNA	milhares	vários	E
mRNA	milhares	vários	P-E
scRNA	dezenas	90-330	P-E
snRNA	dezenas	58-220	E

Moléculas lineares não circulares

### 3.8 Tradução

A tradução é um dos processos mais conservados evolutivamente e um dos eventos de maior custo energético para a célula. O conteúdo protéico da célula é extremamente dinâmico, e as proteínas são continuamente sintetizadas e degradadas.

Nesse processo a informação contida na seqüência de nucleotídeos do mRNA é traduzida numa seqüência de aminoácidos.

A maquinaria celular para esse processo inclui o conjunto de tRNA, os ribossomos, e o mRNA; também são necessários aminoácidos e diversas proteínas específicas.

Enquanto os tRNA e os ribossomos são reciclados pela célula, podendo ser utilizados para a síntese de diferentes proteínas, o mRNA é específico para cada proteína, sendo continuamente produzido e degradado. O mesmo mRNA pode ser traduzido diversas vezes antes de ser degradado.

A tradução tem os mesmos princípios gerais em eucariotos e procariotos, no entanto, existem diferenças importantes. Nos procariotos não existe separação geográfica entre transcrição e tradução. Existem também diferenças estruturais nos ribossomos. Nos eucariotos a separação geográfica entre o núcleo e o citoplasma torna os dois processos independentes.

### 3.8.1 Ativação do tRNA

Antes de serem incorporados às proteínas, os aminoácidos se ligam ao tRNA. A especificidade da ligação do aminoácido ao tRNA não é óbvia, mas a acuracidade da síntese protéica é dependente dessa especificidade.

Na extremidade 3' do tRNA há uma seqüência 5'-CCA-3', em que é adicionado o aminoácido por meio de uma ligação entre o grupo carboxila do aminoácido e a hidroxila 2' ou 3' da adenosina na posição 3'.

Esse processo ocorre em duas etapas: inicialmente o aminoácido é ativado por uma molécula de ATP de forma a criar uma afinidade entre o aminoácido e o tRNA (veja a Figura 3.25). Nessa reação ocorre a hidrólise do ATP, sendo irreversível.

Esses aminoácidos adenilados ligam-se ao tRNA apropriado pela ação da tRNA sintetase específica (veja a Figura 3.25). Essas enzimas possuem dois sítios de reconhecimento, um deles reconhece o aminoácido e o outro o tRNA, desta forma são necessárias 20 tRNA sintetases, sendo possível que a mesma enzima reconheça diferentes tRNA, uma vez que algumas vezes diversos tRNA se ligam ao mesmo aminoácido. A estrutura tridimensional do tRNA é importante nesse processo.

Os tRNA aminoacilados descarregam o aminoácido diretamente no ribossomo durante a tradução.

Da mesma forma que a transcrição, a tradução ocorre em etapas subseqüentes:

### 3.8.2 Iniciação

A iniciação é a primeira etapa da tradução, representada pela ligação de uma molécula de RNA mensageiro à subunidade pequena do ribossomo. Essa ligação é mediada por um componente protéico denominado **fator de iniciação 3 ou IF3**. Há uma ligação transitória entre a subunidade 16S do RNA ribossômico e o RNA mensageiro. Essa complementaridade entre os dois RNAs é decorrente de uma seqüência de consenso denominada seqüência de **Shine-Dalgarno** 5'-AGGAGGU-3' (veja a Figura 3.26). Nos eu-

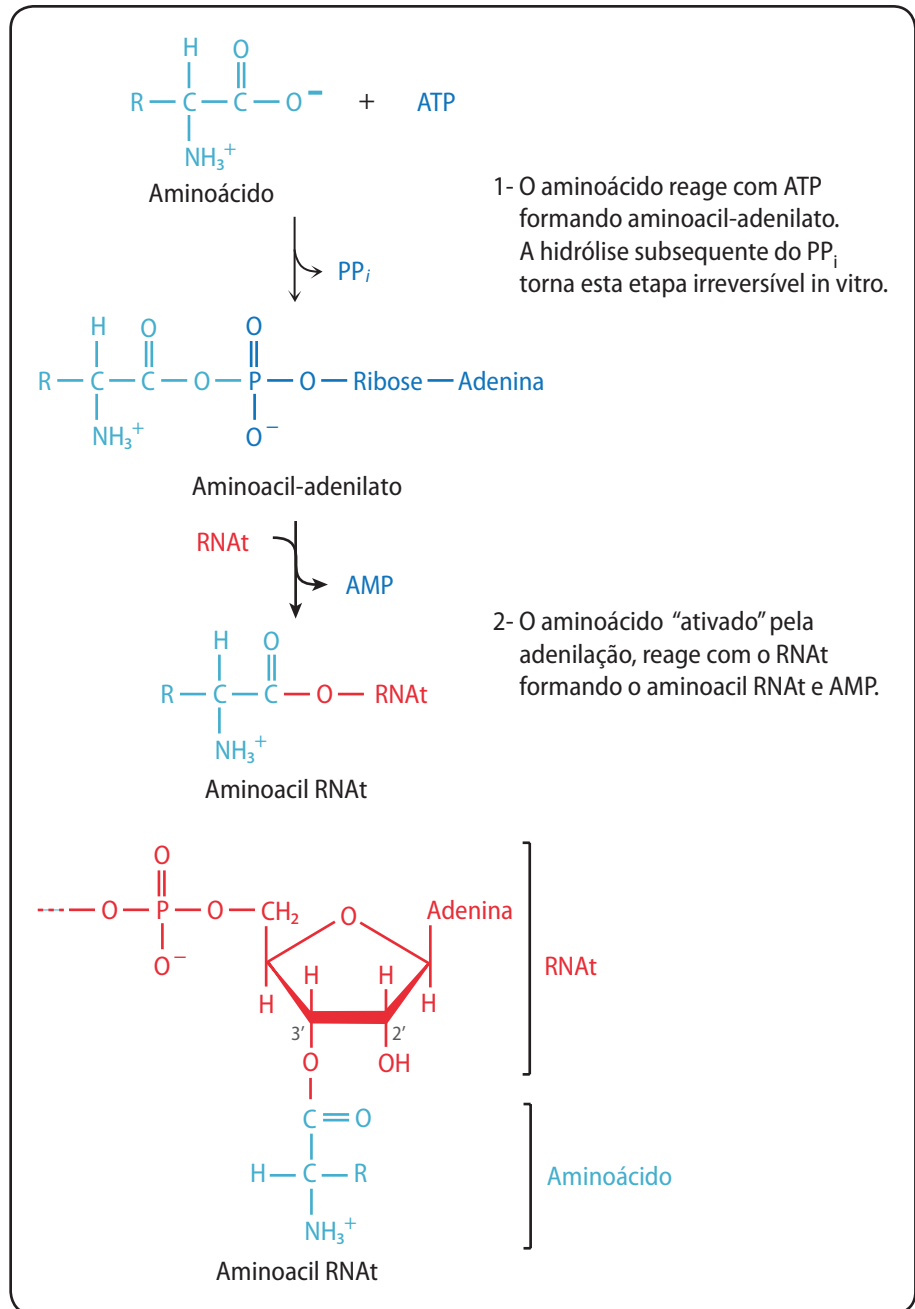


Figura 3.25 - Etapas da ativação do aminoácido

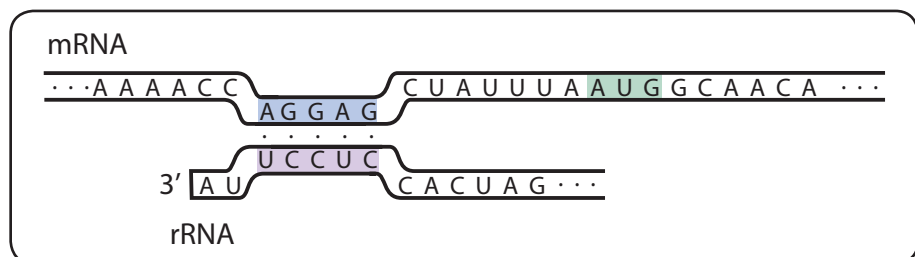


Figura 3.26 - Pareamento entre o mRNA e rRNA

cariotos também existe uma seqüência, denominada **Kozak**, que inclui o primeiro códon, mas a ligação do mRNA à subunidade pequena do ribossomo se dá pelo **CAP-5'**.

A seguir, nos procariotos, a subunidade pequena do ribossomo (30S) migra pela fita de mRNA até encontrar o códon de iniciação (AUG). Nesse momento o **fator de iniciação IF2** ativado pelo **GTP** conduz o primeiro tRNA, fMet-tRNA, ao complexo de iniciação numa região do ribossomo denominada **sítio P** (peptidil). Esse primeiro tRNA possui o aminoácido N-formilmetionina, uma modificação da metionina que bloqueia o grupo amino, desta forma a ligação peptídica só pode ocorrer na direção do grupo carboxil.

Uma vez que se estabelece a ligação entre o códon de iniciação e o anticódon do tRNA, os **fatores de iniciação IF1 e IF3** são liberados e a subunidade grande do ribossomo é acoplada ao conjunto (veja a Figura 3.27).

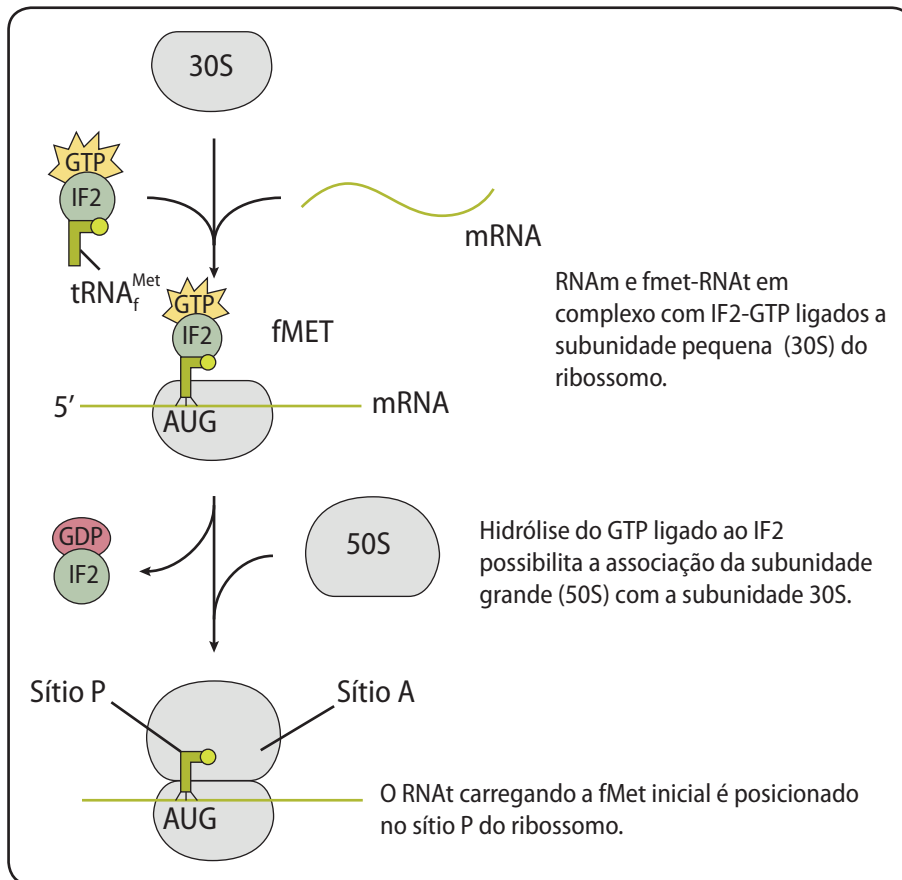


Figura 3.27 - Iniciação da tradução



O ribossomo assim organizado possui regiões específicas onde se encaixam simultaneamente dois tRNA, denominadas **sítio A** (aminoacil) e **sítio P** (peptidil), e o **sítio E** (exit) (veja a Figura 3.28).

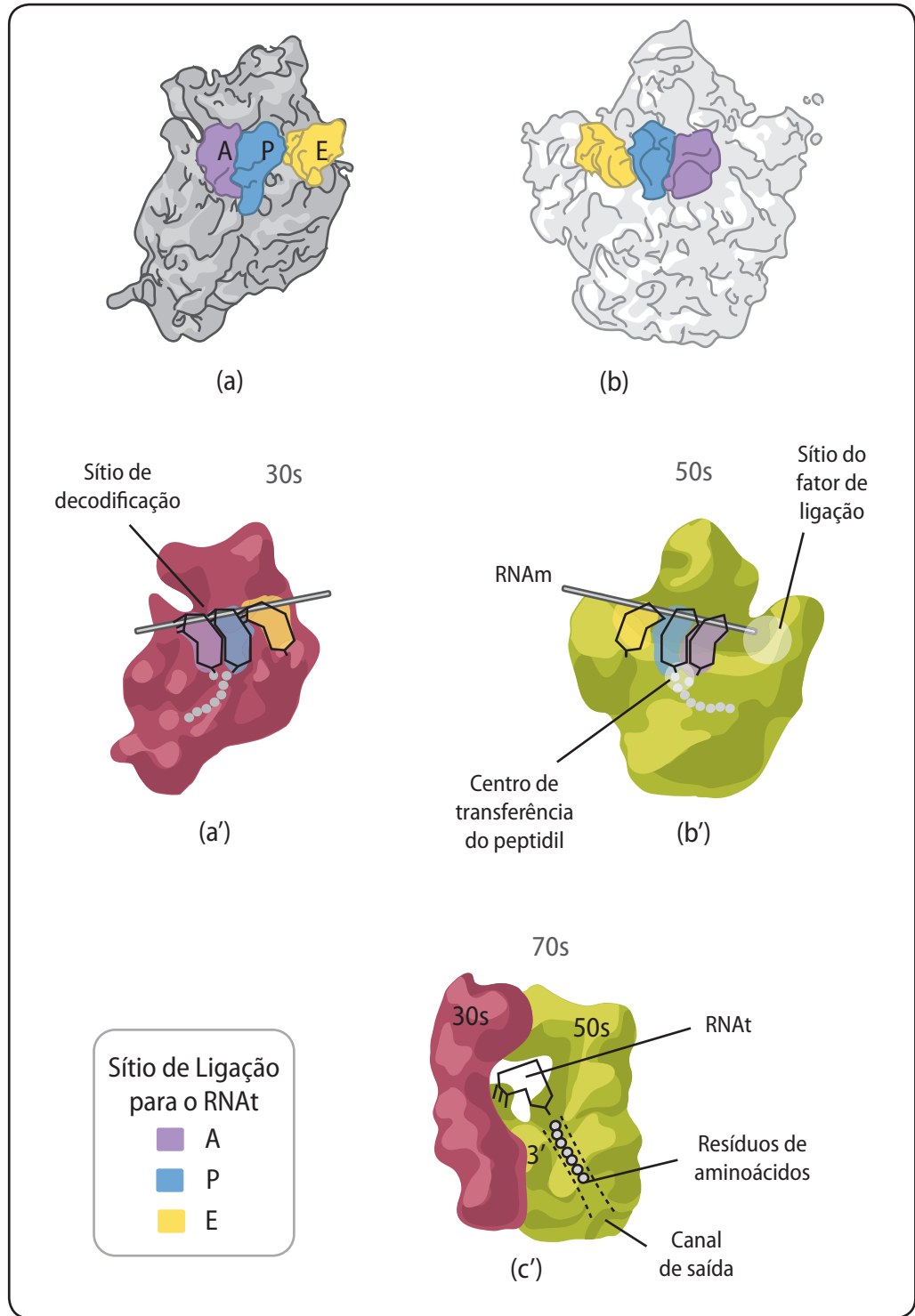


Figura 3.28 - Sítios funcionais do ribossomo

### 3.8.3 Alongamento

Após a montagem do ribossomo, inicia-se a etapa do alongamento. Os tRNA ligados aos aminoácidos, com o auxílio do **fator de alongamento EF-Tu** ativado pelo GTP, são transportados ao ribossomo, sendo posicionados no **sítio A** pela complementaridade **códon/anticódon**. Neste momento o fator de alongamento é liberado, sendo regenerado com outra molécula de GTP (veja a Figura 3.29). Neste momento, ocorre a **ligação peptídica** entre os dois aminoácidos (veja a Figura 3.30). A enzima que catalisa essa reação, denominada **peptidil transferase** é possivelmente uma combinação de diversas proteínas ribossômicas.

Inicia-se assim a **cadeia polipeptídica**, após a ligação dos aminoácidos, ocorre a movimentação do ribossomo no sentido  $5' \rightarrow 3'$ . Essa movimentação é conduzida pelo **fator de alongamento EF-G** (ativado pelo GTP), liberando o tRNA (cada ciclo requer a hidrólise de um GTP). Nessa movimentação ocorre a transferência do **peptidil-tRNA** para o sítio P e a **translocação** do ribossomo no mRNA em direção ao códon seguinte (veja a Figura 3.31).

O alongamento evolui numa velocidade de cerca de 20aa/segundo em bactérias (a  $37^\circ\text{C}$ ), enquanto numa célula humana o processo é mais lento, cerca de 2aa/seg.

### 3.8.4 Término

O alongamento prossegue até que ocorra o posicionamento do **códon de terminação** (UAA, UAG ou UGA) no sítio A. Não existem tRNA complementares a esses códons, que são reconhecidos pelos fatores de liberação. Os **fatores de liberação** entram no sítio A, alterando a atividade da peptidil transferase (adição de  $\text{H}_2\text{O}$  em vez de aminoácido), produzindo a clivagem do peptídeo do último tRNA (veja a Figura 3.32). São conhecidos dois fatores de liberação, RF1, que reconhece os códons UAA e UAG, e RF2, que reconhece os códons UAA e UGA.

O último evento é a dissociação do ribossomo (consumo de uma molécula de GTP). As subunidades do ribossomo continuam íntegras, ficando disponíveis para nova tradução, completando-se assim o **ciclo ribossômico**.

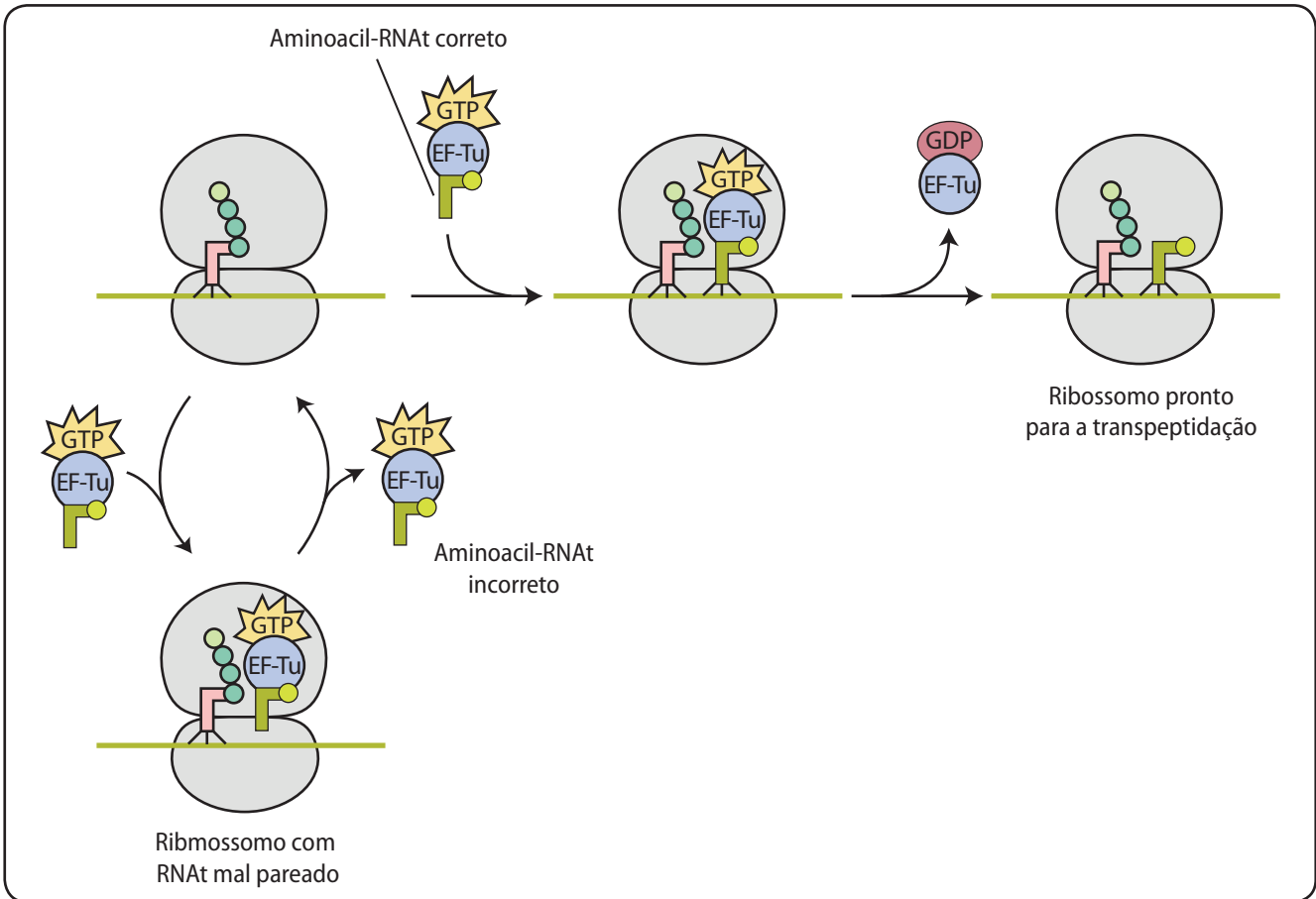


Figura 3.29 - Tradução – alongamento

**LIGAÇÃO PEPTÍDICA**

Os aminoácidos são ligados por uma ligação chamada ligação peptídica

**Ligação peptídica:** os quatro átomos em cada caixa sombreada formam uma unidade planar rígida. Não há rotação entre a ligação C-N.

$$\begin{array}{c} \text{H} \\ | \\ \text{H}-\text{N}-\text{C}-\text{C}=\text{O} \\ | \quad | \\ \text{H} \quad \text{R} \\ \text{OH} \end{array} + \begin{array}{c} \text{H} \\ | \\ \text{H}-\text{N}-\text{C}-\text{C}=\text{O} \\ | \quad | \\ \text{H} \quad \text{R} \\ \text{OH} \end{array} \xrightarrow{\text{H}_2\text{O}} \begin{array}{c} \text{H} \quad \text{H} \\ | \quad | \\ \text{H}-\text{N}-\text{C}-\text{C}-\text{N}-\text{C}-\text{C}=\text{O} \\ | \quad | \quad | \quad | \\ \text{H} \quad \text{R} \quad \text{H} \quad \text{R} \\ \text{OH} \end{array}$$

terminação amino (NH<sub>2</sub>)

terminação carboxil (COOH)

Proteínas são longos polímeros de aminoácidos ligados por ligações peptídicas, e são sempre anotados com a extremidade N na esquerda. A sequência deste tripeptídeo é histidina- cisteína- valina

Estas duas ligações simples permitem rotação, assim as longas cadeias de aminoácidos são muito flexíveis.

Figura 3.30 - Ligação peptídica

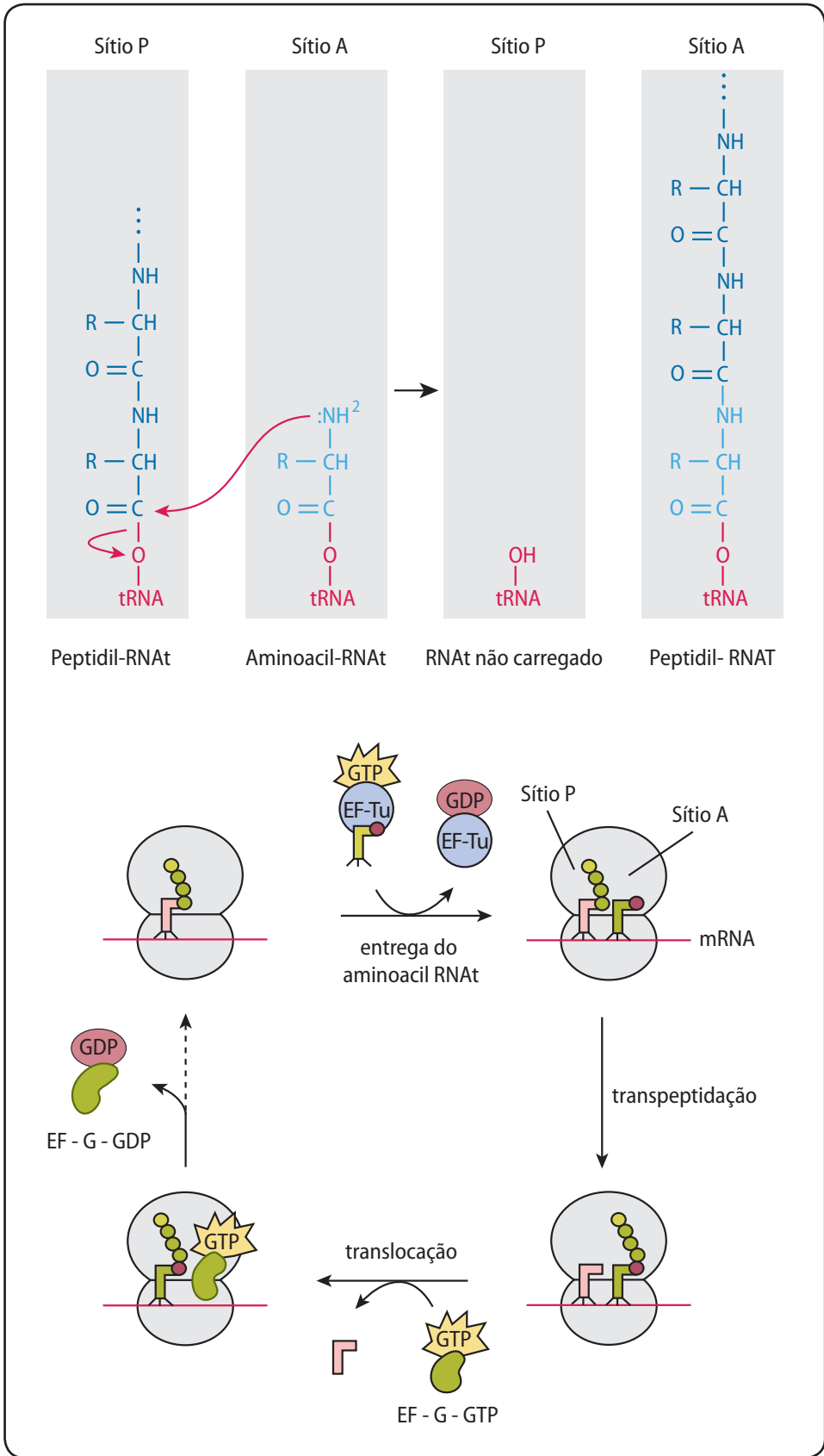


Figura 3.31 - Translocação do ribossomo

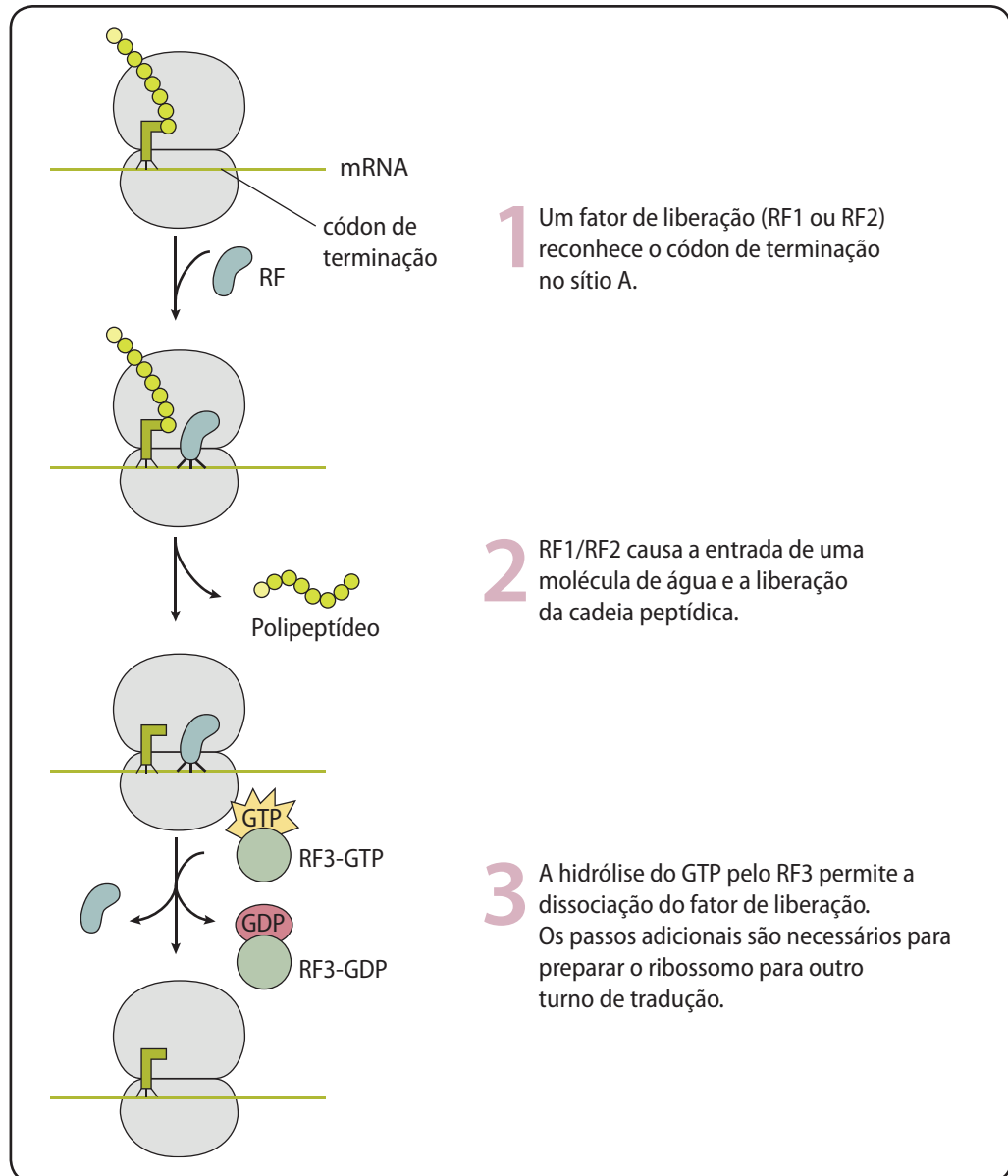


Figura 3.32 - Término da tradução

Verifica-se que o mesmo mRNA pode ser traduzido simultaneamente por diversos ribossomos sendo visualizados como estruturas denominadas **polissomos**, que evoluem progressivamente ao longo da molécula de RNA (veja a Figura 3.33), permitindo que várias cadeias polipeptídicas sejam sintetizadas a partir do mesmo mRNA.

Desta forma, pode-se constatar que o processo de tradução apresenta um **custo energético** para a célula proporcional ao nú-

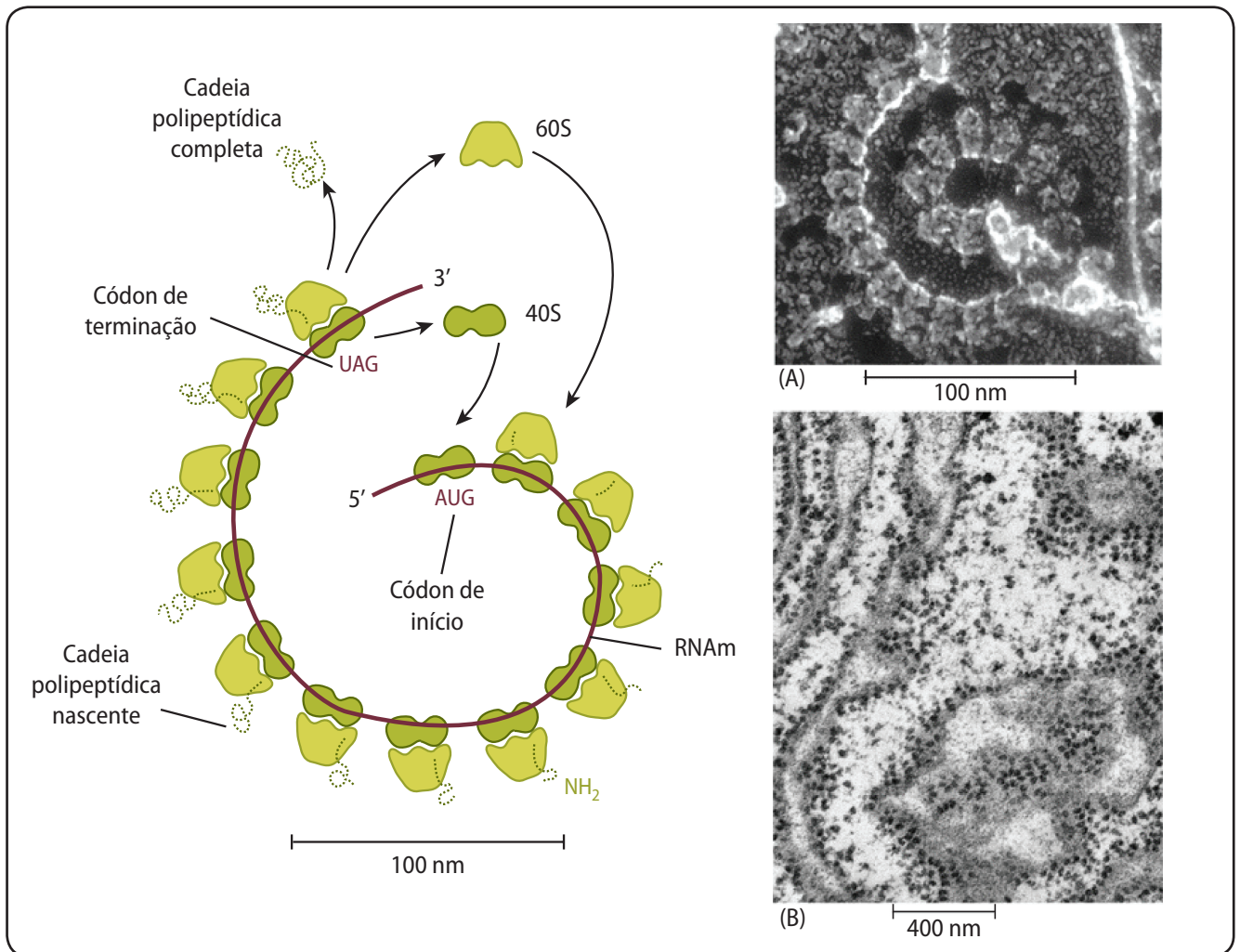


Figura 3.33 - Microfotografia eletrônica de uma fita de mRNA sendo traduzida simultaneamente por diversos ribossomos

mero de aminoácidos, representado pelas moléculas trifosfatadas, como sumariado na Tabela 3.4. Assim, parece natural que a célula busque alternativas para racionalizar o tipo e a quantidade de proteínas sintetizadas. Os mecanismos utilizados para esse fim atuam nas diferentes etapas do processo da informação, sendo o controle da transcrição o processo mais bem caracterizado.

**Tabela 3.4 - Consumo energético**

Ligação aminoácido – RNAt	1 ATP
IF2	1 GTP
EF-TU	1 GTP/aminoácido
EF-G ativado pelo GTP	1 GTP/aminoácido (movimentação do ribossomo)
Dissociação do ribossomo	1 GTP
<b>Para cada aminoácido, 1 ATP + 2 GTP</b>	

Após a tradução, inicia-se o **processamento do peptídeo**. Apenas a tradução é compreendida pela informação do RNA mensageiro, ainda não estão esclarecidos os eventos subsequentes que ocorrem no processamento do peptídeo, os quais influenciam a sua funcionalidade.

### 3.9 Diferenças entre Procariotos e Eucariotos

A tradução da informação é virtualmente similar em todos os seres vivos, de forma que a informação produzida em um organismo pode ser processada por outro organismo e interpretada fielmente. No entanto, já foram identificadas algumas diferenças na interpretação dos códons, como podemos conferir na Tabela 3.5.

**Tabela 3.5 - Diferenças na interpretação dos códons nas mitocôndrias de diferentes organismos**

Códon	Código "Universal"	Código Mitocondrial			
		Mamíferos	Drosófila	Fungos	Plantas
UGA	STOP	<i>Trp</i>	<i>Trp</i>	<i>Trp</i>	STOP
AUA	Ile	<i>Met</i>	<i>Met</i>	<i>Met</i>	Ile
CUA	Leu	Leu	Leu	<i>Thr</i>	Leu
AGA	Arg	<i>STOP</i>	<i>Ser</i>	Arg	Arg
AGG					

Códigos destacados indicam diferença com relação ao código "Universal"

Como o código genético possui 64 códons, seria previsível que houvesse 61 tipos de tRNA contendo o anticódon correspondente a cada códon. No entanto, o repertório de tRNA, em geral, é mais limitado. Nas bactérias, esse conjunto compreende 30 a 40 tRNA com diferentes anticódons, enquanto animais e plantas contam com cerca de 50 tipos de tRNA. O menor conjunto conhecido é encontrado nas mitocôndrias, que possuem apenas 22 tipos de tRNA. Essa diminuição não compromete a precisão da tradução, pois existem apenas 20 aminoácidos. Nos códons sinônimos o

mesmo tRNA pode ser utilizado, isto é possível porque o pareamento entre o códon e o anticódon apresenta uma curvatura, duas bases se pareiam normalmente e a terceira fica oscilante; ainda a presença de bases modificadas no anticódon do tRNA possibilita outras opções de pareamento.

Existem diferenças significativas na tradução em procariotos e eucariotos (veja a Tabela 3.6). Enquanto nos procariotos não há separação geográfica entre a transcrição e a tradução, nos eucariotos a transcrição ocorre no núcleo e a tradução, no citoplasma. Nos procariotos as moléculas de mRNA codificam mais de uma proteína (RNA policistrônico), enquanto nos eucariotos a maior parte de mRNA codifica uma única proteína (veja a Figura 3.34). Também os ribossomos apresentam diferenças estruturais entre procariotos e eucariotos. No entanto, as organelas citoplasmáticas, as mitocôndrias e os cloroplastos, que possuem seu próprio genoma, produzem também seus próprios ribossomos, similares aos procarióticos, coexistindo assim na mesma célula ambos os sistemas operacionais.

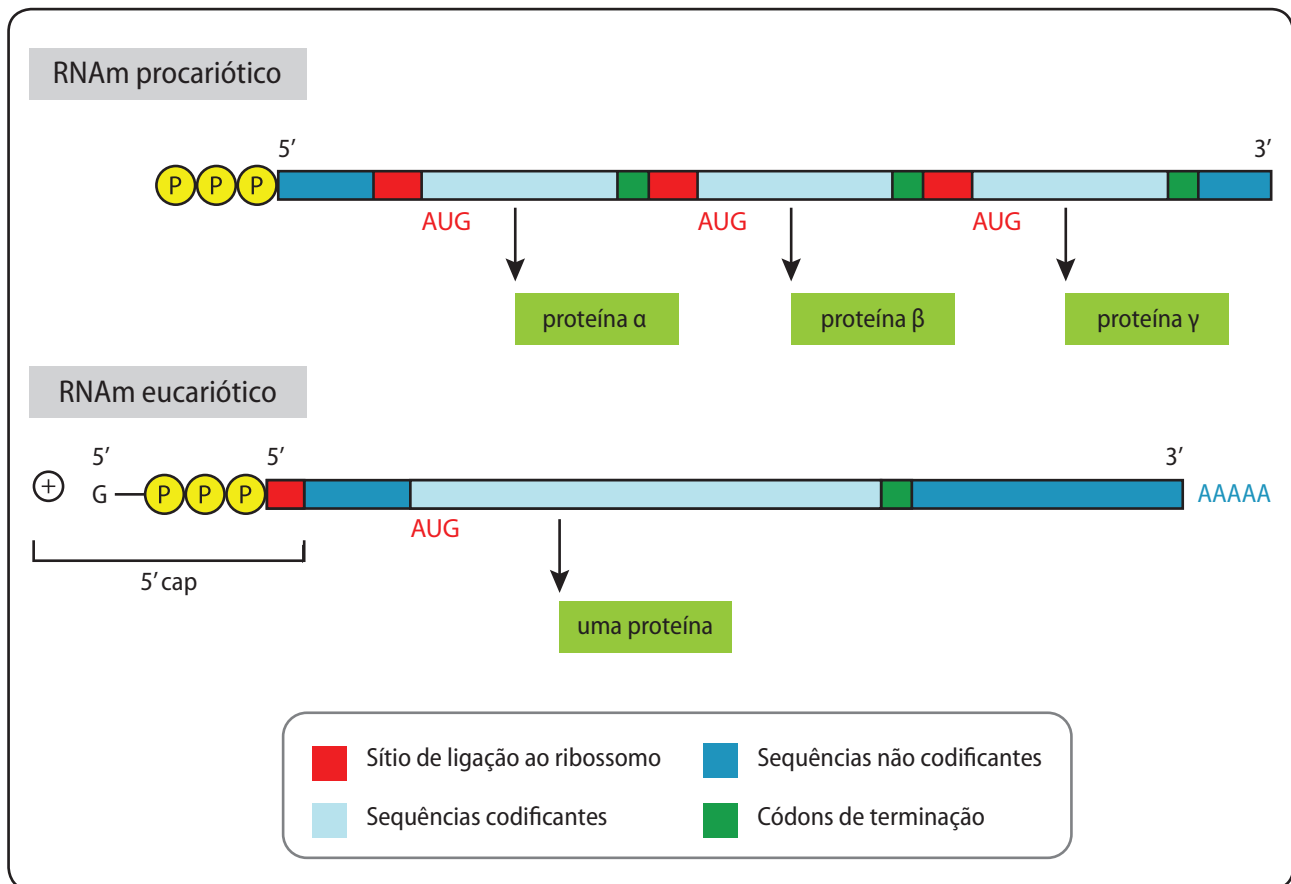


Figura 3.34 - Diferenças entre procariotos e eucariotos



**Tabela 3.6 - Principais diferenças entre a síntese de proteínas de eucariotos e procariotos**

PROCARIOTOS	EUCARIOTOS
não há separação física entre transcrição e tradução	transcrição ocorre no núcleo e tradução no citoplasma
início da tradução antes do término da transcrição	moléculas mais estáveis (cauda poli A)
mRNA policistrônico – pode conter até 20 proteínas	células especializadas – RNA específico abundante
vida média de 0,7 minutos	vida média 1-24h
moléculas instáveis	

### 3.10 Inibidores de Síntese de Proteínas

Foram reconhecidas diversas substâncias que atuam como **inibidoras da síntese de proteínas**, interferindo seletivamente na tradução ou na transcrição. Uma vez que os sistemas celulares são ligeiramente diferentes, essas substâncias atuam diferencialmente em eucariotos e procariotos.

Essas substâncias atuam em etapas específicas do processamento da informação. As substâncias que atuam nos **procariotos** são utilizadas como antibióticos no controle de infecções bacterianas (veja a Figura 3.35). Por exemplo:

- tetraciclina – ligação do tRNA ativado ao ribossomo;
- estreptomicina – bloqueio da iniciação e do alongamento;
- cloranfenicol – bloqueia peptidil transferase; e
- rifamicina – ligação da RNA polimerase (inibe mRNA).

Outras substâncias atuam tanto em **procariotos** quanto em **eucariotos** como a puromicina, que interrompe cadeia polipeptídica, e a actinomicina D, que bloqueia o movimento da RNA polimerase. E ainda existem substâncias seletivas para **eucariotos**, como a anisomicina, que induz o bloqueio da peptidil transferase, a  $\alpha$ amanitina, que bloqueia a síntese do mRNA pela ligação com a RNA polimerase II, e a ciclohexamida, que bloqueia a translocação dos ribossomos.

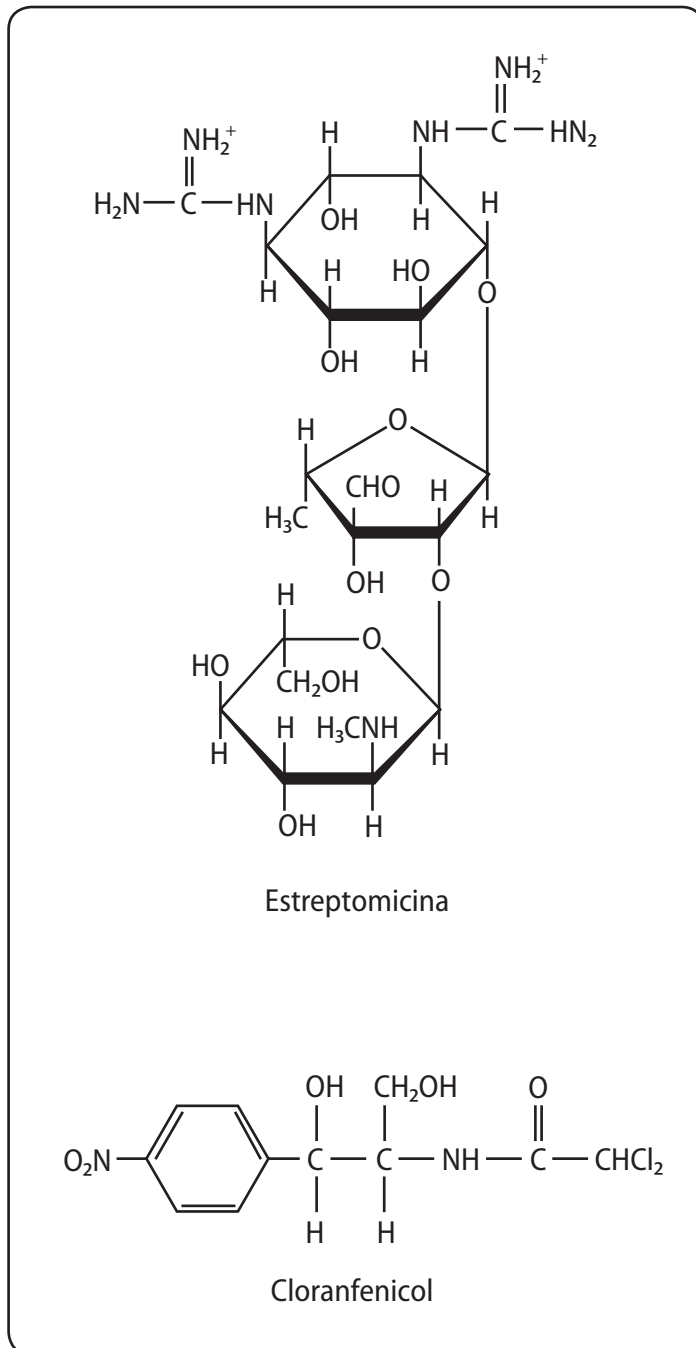


Figura 3.35 - Fórmula de compostos que atuam como inibidores de síntese de proteínas

As mitocôndrias das células eucarióticas possuem o seu próprio DNA e ribossomos, possuindo um sistema de síntese de proteínas independente, similar aos sistemas procarióticos, sendo também influenciadas pelos inibidores de síntese de proteína bacterianos.

## Resumo

O fluxo da informação compreende duas etapas distintas, a **transcrição**, em que as moléculas de DNA são sintetizadas a partir do molde de DNA, e a **tradução**, em que ocorre o alinhamento dos aminoácidos a partir da seqüência de nucleotídeos do RNA. Enquanto nos **procariotos** esses processos ocorrem simultaneamente, no mesmo ambiente, nos **eucariotos** há uma separação espacial e temporal desses processos.

A transcrição é conduzida pela enzima RNA polimerase, utilizando como molde uma das fitas do DNA e ribonucleotídeos trifosfatados e ocorrendo no sentido 5'-3'. O início da transcrição ocorre em uma seqüência de consenso denominada **promotor**, que é o sítio de ligação da RNA polimerase. O início da transcrição necessita de componentes adicionais da RNA polimerase, em procariotos é utilizada uma subunidade da enzima denominada subunidade sigma. Nos eucariotos há uma diversidade de fatores que atuam nesse processo que são denominados genericamente **fatores de transcrição**.

São reconhecidos três tipos principais de moléculas de RNA envolvidos na síntese de proteínas: o RNA mensageiro (mRNA), que possui uma seqüência de nucleotídeos característica que especifica a seqüência de aminoácidos no polipeptídeo; os RNAs ribossômicos (rRNA), que são parte estrutural e funcional dos ribossomos, organelas onde ocorre a síntese de proteínas; e os RNAs transportadores (tRNA), moléculas especializadas no transporte de aminoácidos para o ribossomo. O mRNA e tRNA possuem regiões complementares que possibilitam a orientação correta dos aminoácidos no ribossomo. São utilizadas seqüências de três nucleotídeos para codificar cada aminoácido, sendo possível a partir dos quatro nucleotídeos 64 combinações diferentes. Essas trincas são denominadas códons; como são utilizados apenas 20 aminoácidos para a construção dos polipeptídios, há redundância nos códons, sendo observados aminoácidos que são codificados por diversos códons. Há um códon específico para o início, que codifica a metionina, e três códons que não codificam nenhum aminoácido e sinalizam o término da tradução. A interpretação dos códons é

universal nos sistemas biológicos, o que permite que a informação de um organismo possa ser interpretada acuradamente por outro organismo; essa correspondência é denominada **código genético**. As trincas específicas do tRNA são denominadas anticódons, o conjunto de RNAs transportadores é variável, o menor sistema conhecido é o sistema mitocondrial, que tem 22 tipos de tRNA. Os códons de terminação não possuem tRNA correspondente.

Existem diferenças entre os ribossomos e os RNAs ribossômicos de eucariotos e procariotos, em geral os genes são redundantes e o produto da transcrição inicial pode conter diversas moléculas que são processadas após a transcrição, muitas vezes intercaladas com moléculas de tRNA. Os rRNA e tRNA podem possuir regiões de pareamento intramolecular, que conferem um significado estrutural adicional a essas moléculas, sendo também encontradas bases não usuais, originadas por modificações pós-transcrição.

Nos procariotos é comum uma mesma molécula de mRNA codificar diferentes polipeptídios. Nos eucariotos o produto da transcrição é significativamente maior do que o produto final. Esse transcrito inicial é processado, recebendo modificações nas extremidades 5' (CAP) e 3' (cauda poli A), mas o processo mais dramático é o splicing, que remove regiões intersticiais da molécula, sendo as regiões removidas denominadas íntrons e as regiões que estão presentes no mRNA maduro denominadas éxons. Existem diversos exemplos de splicing alternativo, de forma que uma mesma seqüência do DNA pode codificar polipeptídios diferentes.

A **tradução** ocorre no ribossomo, em que os três tipos de RNA (mRNA, rRNA e tRNA) atuam para construir a molécula polipeptídica de acordo com a seqüência de bases do mRNA. Esse processo envolve grande consumo de energia para ativação dos aminoácidos e deslocamento dos ribossomos.

Diversas substâncias podem interferir nas diferentes etapas do fluxo da informação. Uma vez que os sistemas procarióticos e eucarióticos são estruturalmente diferentes, sua ação é seletiva.

## Referência Comentada

### **A célula: uma abordagem molecular**

*Geoffrey M. Cooper, Robert E. Hausman*

Os capítulos 6 e 7 fazem uma abordagem adequada do conteúdo numa linguagem acessível e ricamente ilustrada.

COOPER, G. M.; HAUSMAN, R. E. **A célula: uma abordagem molecular**. 3. Ed. Porto Alegre: Artmed, 2007. Caps. 6–7.

## Referências Bibliográficas

ALBERTS, B. et al. **Molecular Biology of the Cell**. 4th. ed. New York: Garland Pub. Inc., 2002.

GRIFFITHS, a. j. f. et al. **Introdução à Genética**. 8. ed. Rio de Janeiro: Guanabara-Koogan, 2006.

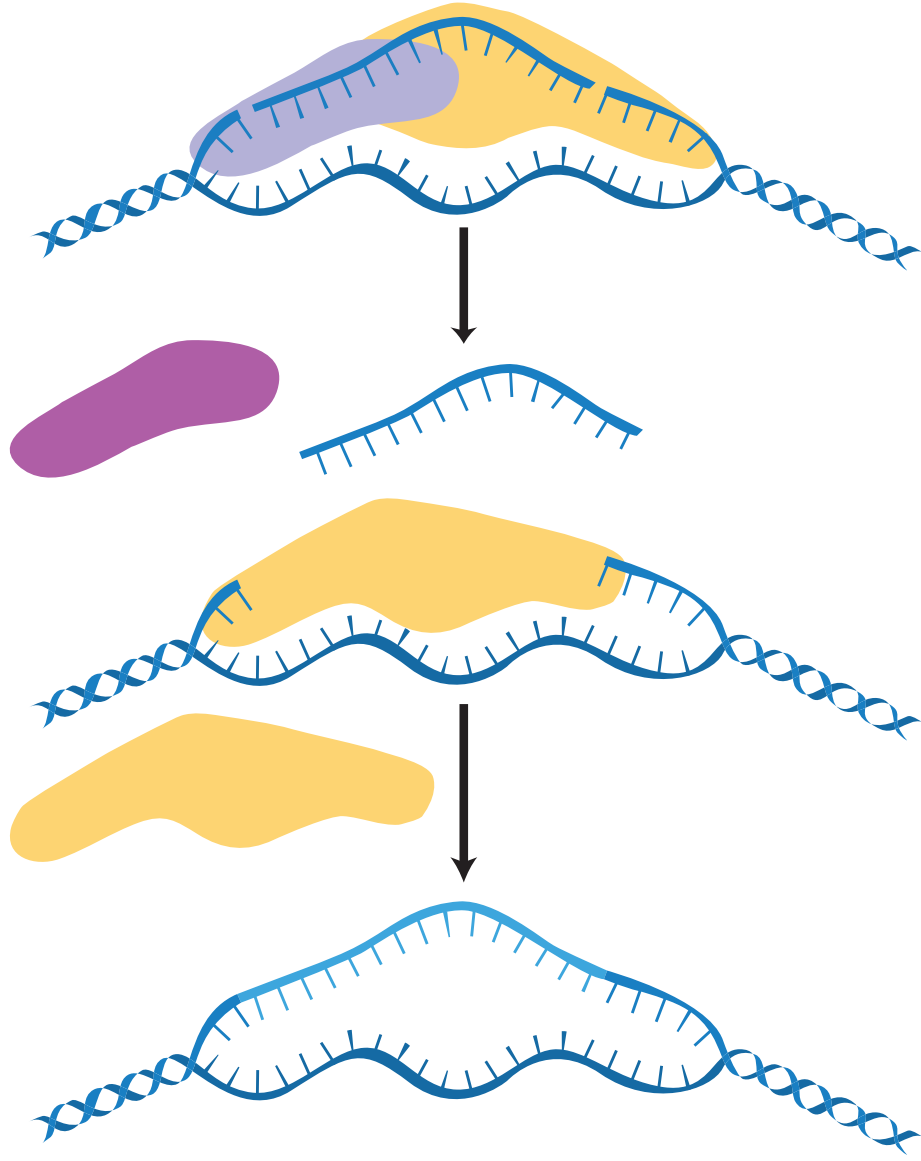
LEWIN, B. **Genes VII**. Porto Alegre: Artmed, 2001.

PURVES, W. K. et al. **Vida: a ciência da Biologia**. Célula e hereditariedade. 6. ed. Porto Alegre: Artmed, 2002. v. I.

STRACHAN, T.; READ, A.P. **Genética molecular humana**. 2. ed. Porto Alegre: Artmed, 2002.

WATSON, J. D. et al. **Biologia molecular do gene**. 5. ed. Porto Alegre: Artmed, 2006.

# CAPÍTULO 4



## Variabilidade Genética

*Neste capítulo vamos abordar a plasticidade do genoma e os processos que contribuem para o aumento da variabilidade genética e diversidade, buscando compreender os efeitos biológicos das alterações na seqüência de DNA. Também iremos esclarecer o efeito dos agentes mutagênicos no DNA e reconhecer a importância dos mecanismos celulares de reconhecimento e reparo das lesões do DNA.*

*Para finalizar, vamos abordar a contribuição da recombinação e da transposição para a diversidade genética.*





O material genético deve possibilitar plasticidade ao genoma, permitindo a variabilidade genética e, portanto, a capacidade de adaptação ao ambiente. Uma vez que as características da molécula do DNA e seu processo de replicação contribuem para a continuidade da informação, outros mecanismos atuam para gerar a indispensável variabilidade.

Quimicamente, o DNA é uma molécula reativa, sendo alvo de alterações espontâneas e induzidas por agentes físicos, químicos e biológicos que produzem alterações na molécula do DNA. Os sistemas celulares de reparo do DNA atuam para corrigir essas alterações, preservando a integridade do genoma. As lesões não corrigidas podem resultar em alterações permanentes na seqüência do DNA.

Os mecanismos de recombinação possibilitam o reagrupamento da informação, gerando novas combinações de seqüências codificantes e individualidade aos organismos. Recentemente foram caracterizados mecanismos adicionais de variabilidade genética como a transposição e a retrotransposição, evidenciando a amplitude dos recursos celulares disponíveis para gerar variabilidade genética. Também foi esclarecido o processo que resulta na transformação, fenômeno que permitiu demonstrar que o DNA é o material hereditário.

## 4.1 Mutação

O termo “mutação” é largamente utilizado em Biologia, compreendendo as alterações no material genético e o processo que causa essas alterações e designando ainda as alterações na sequência de nucleotídeos.

As mutações podem ocorrer espontaneamente, decorrentes de erros no processo de replicação do DNA ou decorrentes de lesões nas moléculas que constituem o DNA, induzidas por substâncias produzidas pelo metabolismo celular. Essas lesões podem ocorrer também por meio da exposição das células a agentes que lesam o DNA, os quais são denominados **agentes mutagênicos**.

Uma gama de agentes mutagênicos já foi identificada, o estilo de vida moderno acarreta maior exposição a esses agentes.

As mutações apresentam uma gama de efeitos biológicos, sendo uma fonte importante de variabilidade genética, produzindo o substrato para a evolução. A variabilidade genética possibilita a adaptação a novos ambientes.

Enquanto para a espécie a variabilidade é fundamental, para o organismo pode ser deletéria ou adaptativa, produzindo fenótipos diferentes do original. Os indivíduos com esses fenótipos são denominados mutantes.

O termo “mutação” engloba diferentes tipos de alteração no material genético. As euploidias e as aneuploidias representam alterações no número de cromossomos, enquanto nas poliploidias (euploidias) há alterações na quantidade de conjuntos cromossômicos, nas aneuploidias há alteração no número de cópias de um determinado cromossomo, estando associadas a alterações fenotípicas características. Esses fenômenos foram descobertos em vegetais por meio de análise microscópica. Nos vegetais, as alterações na ploidia estão associadas ao aumento no potencial nutritivo, a maior parte dos cereais e das frutas utilizados como alimentos são **poliplóides**. Algumas orquí-

### Agentes mutagênicos

- **físicos:** radiação ionizante e raios UVC, capazes de destruir as ligações químicas ente os nucleotídeos, e UVB, cujo espectro é absorvido pelo DNA. Os danos desses agentes são grandemente amplificados em presença de água e oxigênio;
- **químicos:** inúmeras substâncias ditas cancerígenas, que atuam danificando ligações químicas ou mesmo substituindo nucleotídeos normais por moléculas análogas;
- **biológicos:** ação de vírus e bactérias, que injetam parte de seu DNA na célula hospedeira, ocasionalmente, integrando-a a cadeia de DNA do hospedeiro. Também podem ocorrer mutações por falhas de ordem genética.

## Poliploidia

A **poliploidia** é muito comum em plantas, mas rara em animais. De fato, um aumento no número de conjuntos cromossômicos foi um fator importante na origem de algumas espécies de novas plantas. Tradicionalmente os poliplóides são classificados em autopoliplóides, originados pela duplicação de um mesmo genoma, e alopoliplóides, originados pela duplicação de genomas diferentes, normalmente após um evento de hibridação. Em torno de 40% das espécies cultivadas são poliplóides (SIMMONDS, 1980), como alfafa (*Medicago sativa*), algodão (*Gossypium hirsutum*), batata (*Solanum tuberosum*), batata-doce (*Ipomoea batatas*), café (*Coffea arabica*), cana-de-açúcar (*Saccharum officinarum*), fumo (*Nicotiana tabacum*), morango (*Fragaria ananassa*), trigo (*Triticum aestivum*), dentre outras.

deas poliplóides apresentam maior valor comercial, enquanto as aneuploidias foram associadas à exacerbação de uma característica pontual.

Podem ocorrer também alterações na estrutura dos cromossomos resultantes de rearranjos do DNA, acarretados por quebras de cadeia dupla, as quais podem ocorrer em um ou mais cromossomos, resultando em inversões, deleções, duplicações e translocações. Esses eventos podem comprometer genes localizados nas regiões das fraturas e também influenciar o funcionamento dos genes das regiões envolvidas. Esses rearranjos estão associados à evolução (veja a Figura 4.1).

Em humanos a primeira aneuploidia reconhecida foi a Síndrome de Down, seguida do reconhecimento de diversas outras alterações cromossômicas. As alterações cromossômicas são a forma mais prevalente de doença genética na população, estando associadas a malformações congênitas, retardo mental e do desenvolvimento neuropsicomotor, abortos recorrentes e alterações no desenvolvimento sexual. Nas neoplasias observa-se a presença dessas alterações, sendo que a caracterização das alterações cromossômicas vem se consolidando como uma ferramenta adicional para a classificação dos tumores.

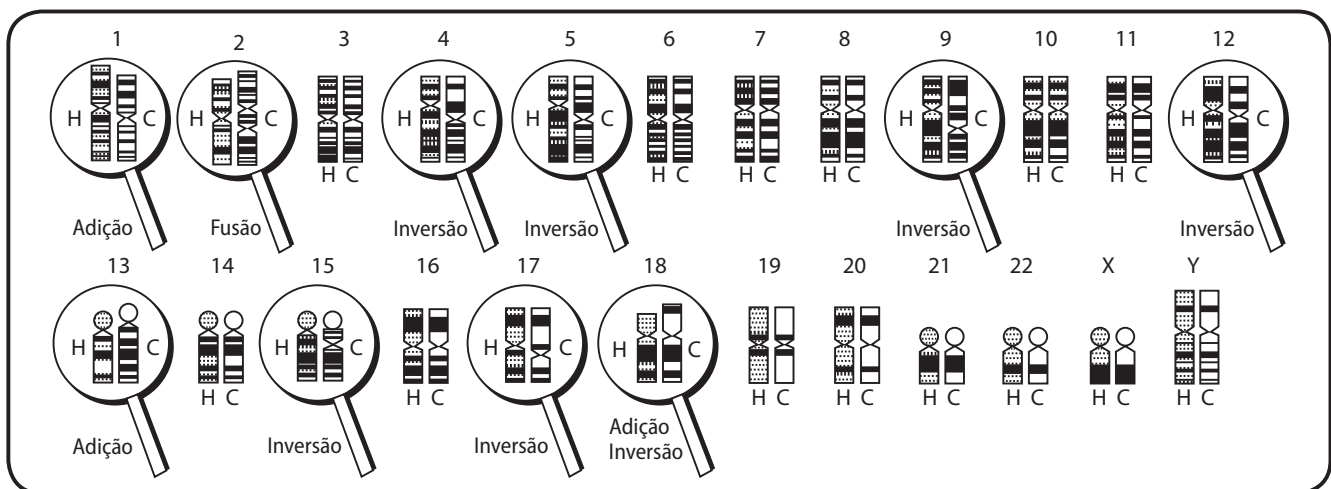


Figura 4.1 - Comparação entre os cromossomos humanos (H) e de chimpanzé (C)

Alterações mais discretas do DNA, que não podem ser visualizadas microscopicamente, acarretam alterações de genes individuais. Essas alterações incluem as mudanças de um único nucleotídeo, mas também podem compreender uma seqüência maior. As alterações de nucleotídeos uma vez ocorridas são transmitidas para as células-filhas, e no caso de ocorrerem em células germinativas, para as gerações subseqüentes.

A detecção de uma drosófila mutante, com olhos brancos, é um exemplo recorrente de organismo mutante (veja a Figura 4.2). Os mutantes também exercem uma influência poderosa na ficção, associada a atributos extraordinários de mocinhos e vilões, tendo uma grande repercussão no imaginário popular.

Um dos exemplos mais conhecidos de fenótipo mutante em humanos é o albinismo. Essa condição é ocasionada pela corrupção da informação de uma das enzimas da via biossintética da melanina. O albinismo não é um fenômeno exclusivamente humano, sendo verificado em diversas outras espécies.

Uma grande diversidade de condições hereditárias humanas já foi descrita e é compilada continuamente num banco de dados com acesso livre (OMIM – Online Mendelian Inheritance in Man). A identificação dos genes associados a algumas dessas condições oferece a oportunidade de compreender diversas nuances do funcionamento celular e do fluxo da informação. Por outro lado, ainda não existe tratamento disponível para a maioria dessas condições, acarretando aos portadores muitas vezes constrangimentos e intimidações, e apreensão nas famílias quanto à possibilidade de recorrência nos parentes. Não existem políticas públicas que busquem atenuar o impacto dessas condições e disponibilizar a atenção adequada a esses indivíduos. A maior parte dessas condições é rara na população e na maioria das vezes inexpressiva em termos estatísticos. Na maior parte das vezes, essas condições são decorrentes de mutações herdadas, que segregam na família pelas gerações. Eventualmente podem ocorrer novas mutações em células germinativas.



Figura 4.2 - Drosófilas com mutação para cor de olho

Por outro lado, as mutações em células somáticas não produzem um efeito sistêmico, mas apresentam maior impacto em termos de saúde pública, pois estão associadas a malformações congênitas e a doenças neoplásicas, ambas influenciadas pela exposição ambiental.

O feto em desenvolvimento é especialmente sensível aos efeitos dos agentes mutagênicos, de forma que durante a gestação deve ser redobrada a vigilância quanto à exposição. Em geral, quanto mais precoce a exposição maior o impacto no desenvolvimento embrionário. Diversas malformações são atribuídas à exposição intra-uterina a agentes mutagênicos.

A relação entre mutagênese e carcinogênese é antiga, no entanto, tornou-se mais evidente ao demonstrar-se que a incidência de câncer de pulmão estava associada ao tabagismo, sendo a associação de tumores com exposição ocupacional ou ambiental crescente. É necessário um esforço de informação e conscientização da população sobre a natureza desses agentes e os cuidados necessários para o seu manuseio e descarte. A poluição ambiental com **agentes mutagênicos** tem repercussão em todos os seres vivos do local, esse efeito é duradouro e acumulativo.

Em termos moleculares, o efeito das mutações em genes individuais é melhor compreendido do que os rearranjos e as alterações cromossômicas, uma vez que os detalhes do funcionamento panorâmico do genoma ainda são pouco esclarecidos.

Desta forma, as alterações em nucleotídeos individuais são classificadas em transições, quando envolvem a substituição de uma purina por outra purina, e de uma pirimidina por outra pirimidina, ou transversões, quando a troca é de uma purina por uma pirimidina, ou vice-versa. Também podem ocorrer deleções e adições de nucleotídeos.

O impacto dessas alterações é variável e classificado de acordo com o efeito na informação (veja a Figura 4.3). As mutações silenciosas consistem em trocas de nucleotídeo, que resultam num códon sinônimo, a proteína resultante possui a mesma seqüência de aminoácidos codificada pela seqüência original. Nas mutações sinônimas ou neutras ocorre a substituição por aminoácidos funcionalmente similares, as propriedades da proteína são preservadas.

No site <<http://monographs.iarc.fr/ENG/Classification/Listagentsalphorder.pdf>>, está disponível um documento que relaciona as substâncias reconhecidas como mutágenos.

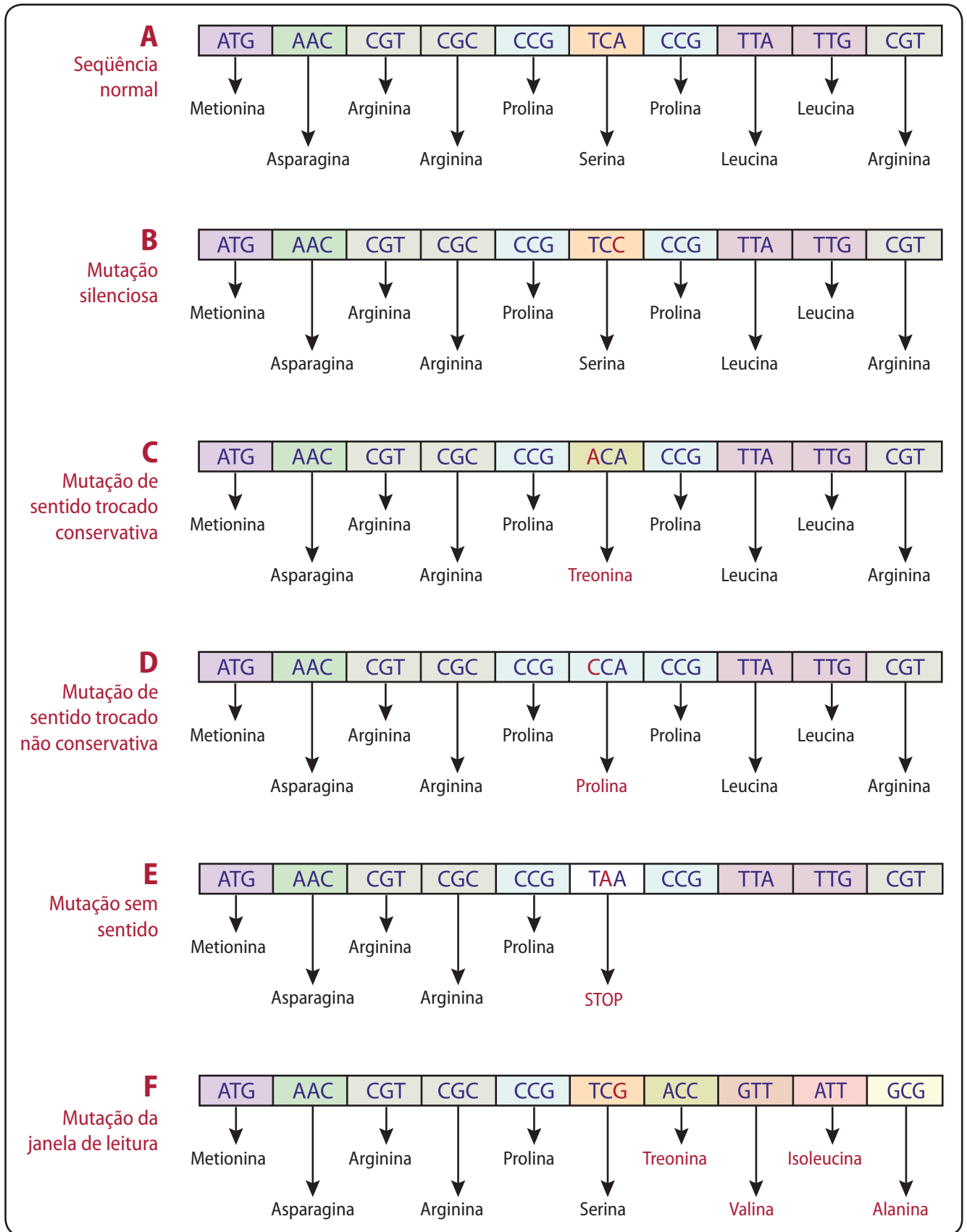


Figura 4.3 - Classificação das mutações de nucleotídeos individuais pelo efeito na tradução

Na mutação de sentido alterado, o aminoácido é substituído por outro funcionalmente diferente, alterando as propriedades e a funcionalidade da proteína. A mutação sem sentido resulta na substituição de um códon codificante por um códon de terminação, acarretando o término prematuro da tradução, e a síntese de uma proteína truncada.

As inserções e as deleções têm um impacto mais amplo; quando o número de nucleotídeos não é múltiplo de 3, ocorre a alteração em todos os nucleotídeos subsequentes, deslocando a janela de leitura a partir do ponto em que ocorreu a deleção ou a inserção.

Os processos que acarretam as mutações também podem produzir reversão das mutações, restabelecendo assim a seqüência original.

Outro tipo de alteração pontual consiste na expansão de seqüências de nucleotídeos, em que uma determinada seqüência de nucleotídeos pode expandir seu número de cópias. A expansão de seqüência de trinucleotídeos está associada a algumas doenças hereditárias humanas (veja a Figura 4.4).

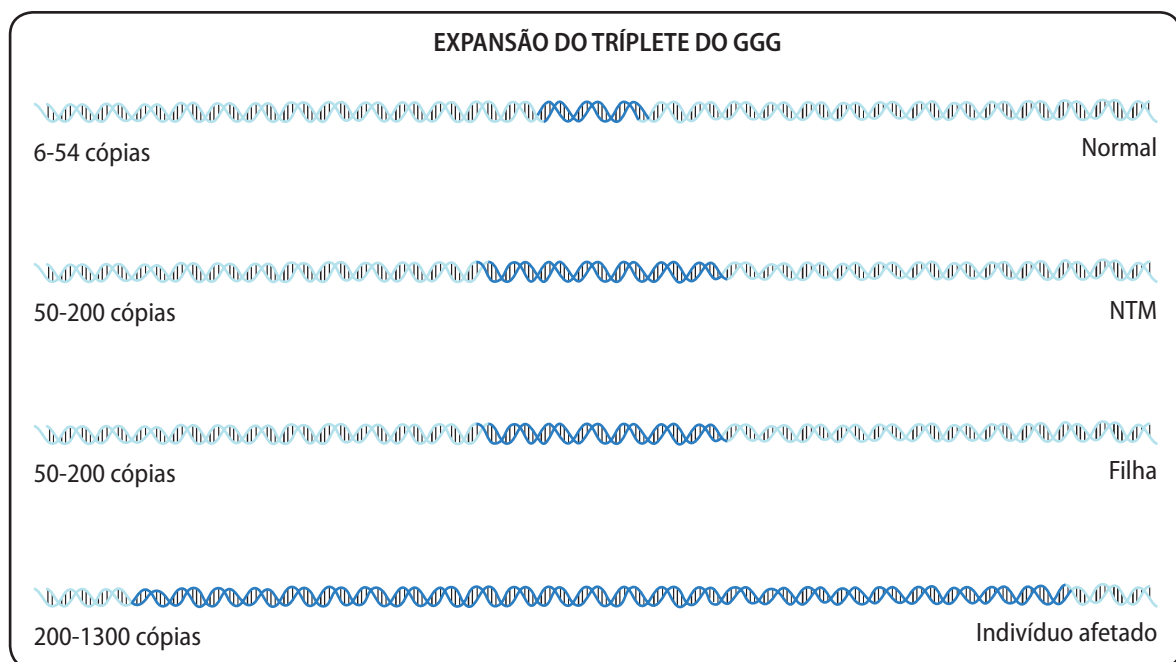


Figura 4.4 - Expansão de uma trinca CGG no gene FMR-1 observada na Síndrome do X Frágil

O impacto das alterações ainda depende da região do DNA que foi modificada, a mutação nas seqüências codificantes tem um impacto diferenciado das mutações que ocorrem em íntrons. Por ou-

tro lado, as mutações nos promotores podem exercer um impacto surpreendente.

Quando a frequência populacional de uma mutação é muito elevada ( $\geq 0,01$ ) recebe a denominação de polimorfismo.

As regiões não codificantes apresentam maior variabilidade que as regiões codificantes. Entre as regiões não codificantes destaca-se o DNA de seqüências repetitivas. Essas seqüências podem estar dispersas no genoma, como as seqüências denominadas LINES e SINES, e as seqüências com repetições em tandem (VNTR)

(veja a Figura 4.5). A denominação dessas seqüências reflete o tamanho da seqüência repetitiva. O DNA satélite compreende seqüências longas de DNA altamente repetitivas ( $>10^4$  cópias), de 100 a 2.000 pares de bases, que constituem uma fração importante do genoma dos eucariontes superiores, variando de 25% em algumas espécies até 50% em outras. Os minissatélites são representados por seqüência de DNA moderadamente repetitiva de 10 a 100 pares de bases. Os microsatélites (STR ou short tandem repeats) compreendem seqüências de DNA moderadamente repetitivas de 2 a 9 pb, altamente polimórficas (veja a Figura 4.6).

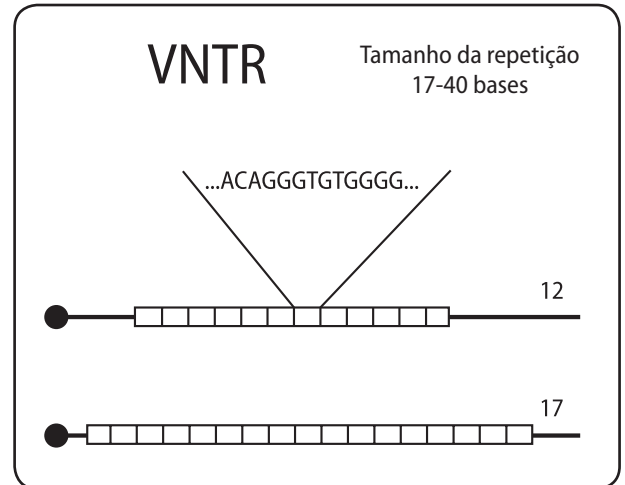


Figura 4.5 - Número variável de repetições em tandem (VNTR)

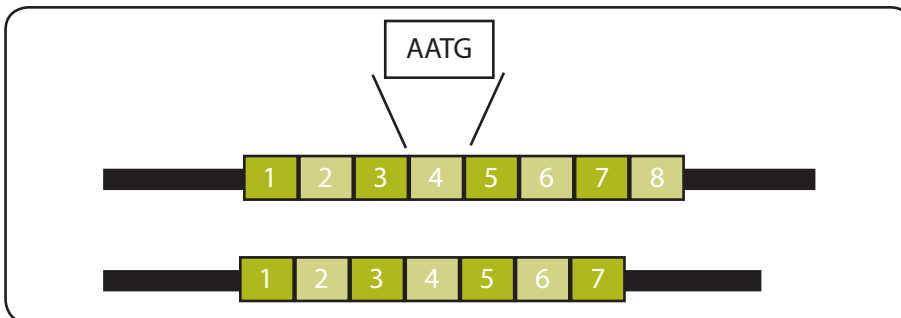


Figura 4.6 - Microsatélites ou pequenas repetições em tandem (STR)

### Mutações gênicas – regiões codificantes

- **mutação silenciosa:** código degenerado;
- **mutação com sentido alterado:** altera aminoácido;
- **mutação neutra:** troca de aminoácidos sem alteração da função do polipeptídeo;
- **mutação de janela de leitura:** deleção ou inserção de nucleotídeos em número diferente de 3.



## Nomenclatura para descrição das mutações

(Antonarakis et al., 1998)

### 1. Substituição de aminoácidos:

Código de uma letra: A, alanina; C, cisteína; D, ácido aspártico; E, ácido glutâmico; F, fenilalanina; G, glicina; H, histidina; I, isoleucina; K, lisina; M, metionina; N, asparagina; P, prolina; Q, glutamina; R, arginina; S, serina; T, treonina; V, valina; W, triptofano; Y, tirosina; X fim.

O código de três letras também é aceito.

- **R117H ou Arg117His** – substitui arginina 117 pela histidina (o códon iniciador metionina é o códon 1).
- **G542X or Gly542Stop** – glicina 542 substituída pelo códon de terminação.

### 2. Substituição de nucleotídeo:

O A do códon de iniciação ATG é +1; a base imediatamente precedente é -1. Não há zero. Número do nucleotídeo seguido pela alteração. Se necessário, usar g ou c para designar seqüências genômicas e de cDNA. Para seqüências intrônicas, quando ape-

nas a seqüência de cDNA é conhecida, especificar o número do íntron por IVSn ou pelo número da posição no éxon mais próxima.

- **1162G>A** – substitui guanina na posição 1162 pela adenina.
- **621+1G>T ou IVS4+1G>T** – substitui G por T na primeira base do íntron 4; éxon 4 termina no nt 621.

### 3. Deleções e inserções:

Use del para deleções e ins para inserções. Como nos itens anteriores, nas mudanças no DNA, primeiro vem a posição do nucleotídeo ou do intervalo, e nas mudanças em aminoácidos, o símbolo do aminoácido vêm primeiro.

- **F508del** – deleção fenilalanina 508.
- **6232-6236del** ou **6232-6236delATAAG** – deleção 5 nucleotídeos (os quais podem ser especificados) começando com nt 6232.
- **409-410insC** – inserção C entre nt 409 e 410.

A freqüência de mutação é influenciada por diversos fatores, diferindo em eucariotos e procariotos; em bactérias é da ordem de 1 mutação/ $10^8$  nucleotídeos, enquanto em humanos é de 1 mutação/ $10^9$  nucleotídeos. A acurácia da maquinaria de replicação do DNA e a eficiência do sistema de reparo contribuem decisivamente, no entanto, a exposição a agentes mutagênicos pode modificar essa freqüência. Considerando-se genes individuais, a freqüência de mutações está associada ao seu tamanho.

Os agentes mutagênicos atuam diferenciadamente no DNA produzindo lesões e alterando sua estrutura, de forma que alteram a capacidade de replicação e transcrição. As alterações que deformam a molécula de DNA são inadmissíveis para a célula, existindo uma maquinaria especializada no reparo dessas lesões.

Os erros também podem ser decorrentes da incorporação de bases tautoméricas durante a replicação do DNA (veja a Figura 4.7).

Após o término da replicação do DNA, é realizada uma checagem, rastreando-se os nucleotídeos malpareados; esse processo é denominado reparo de malpareamento (mismatch repair) (veja a Figura 4.8). No entanto, como esse processo consegue diferenciar

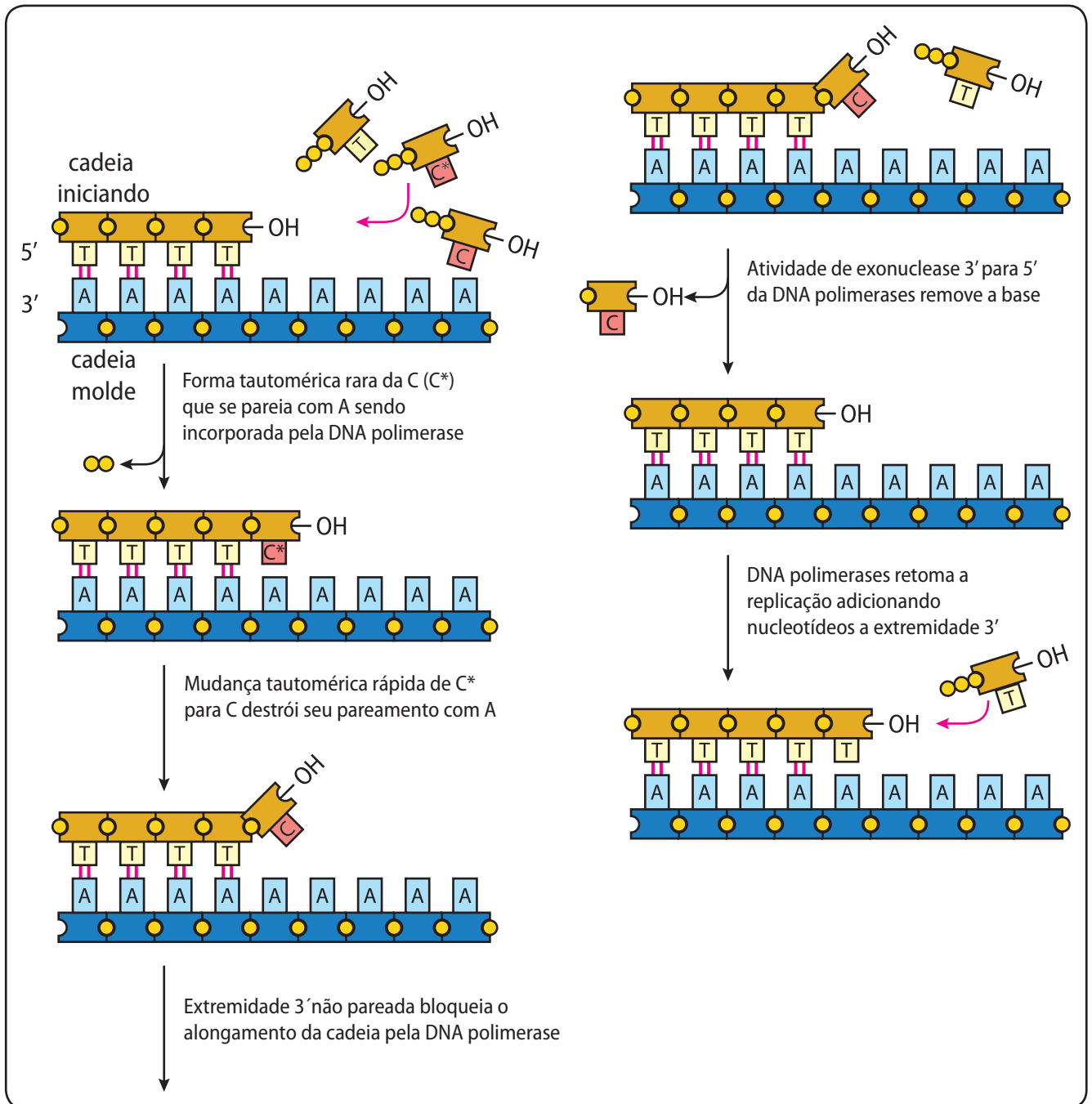


Figura 4.7 - Incorporação de base tautomérica e seu respectivo reparo

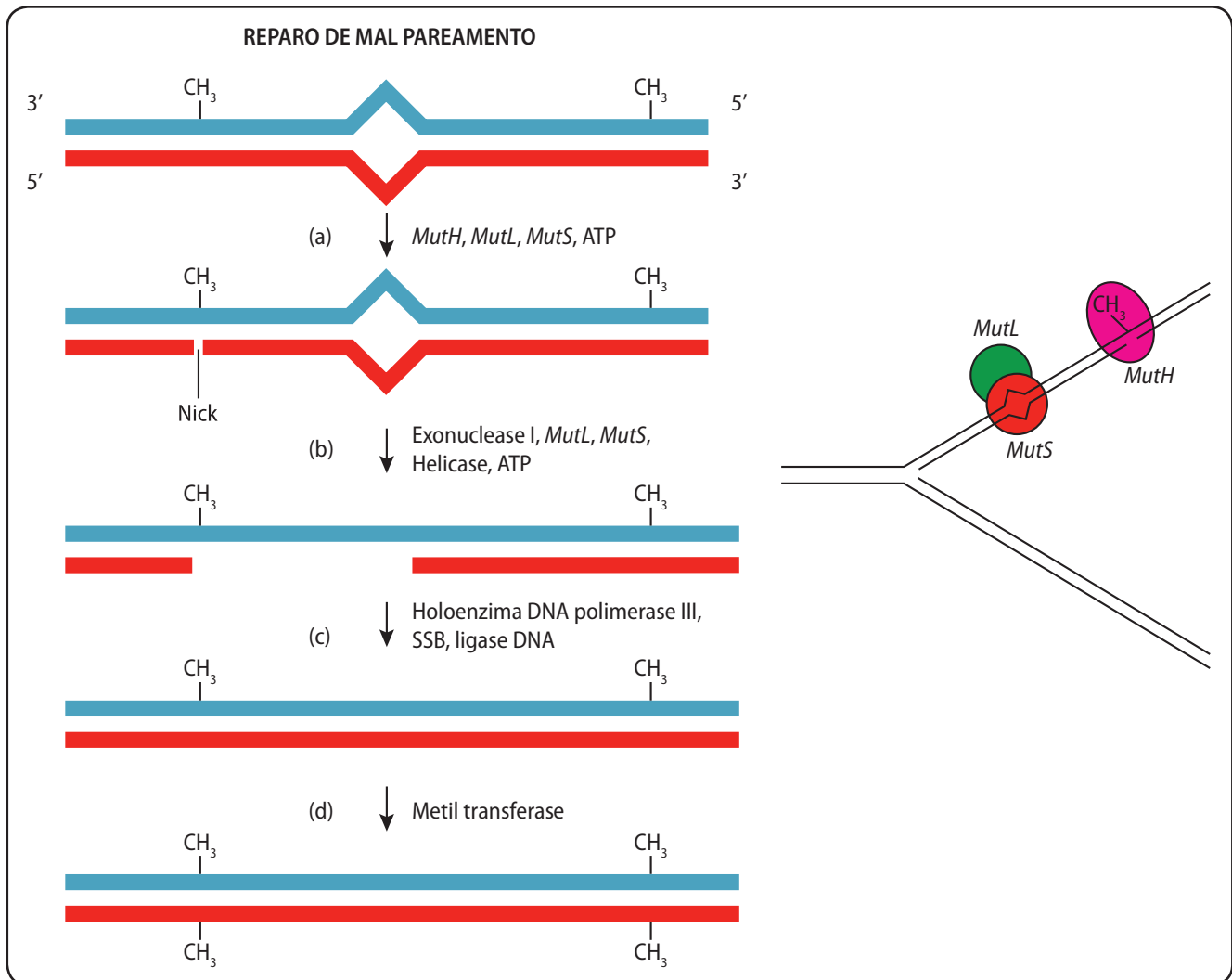


Figura 4.8 - Reconhecimento e reparo do malpareamento (*mismatch repair*)

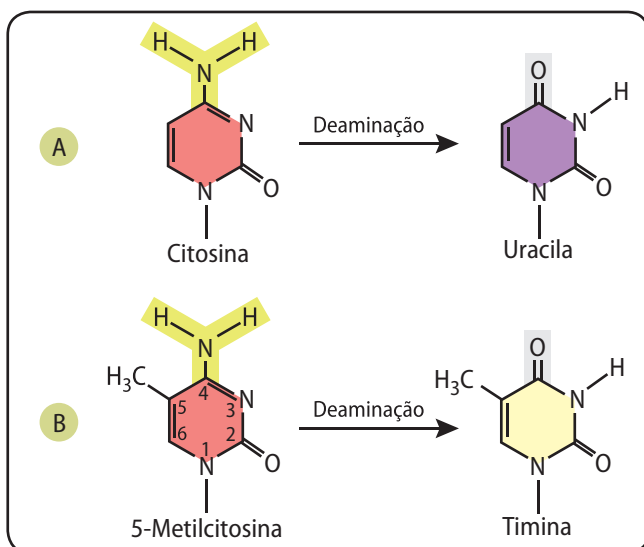


Figura 4.9 - Lesões nas bases que acarretam mudanças nas propriedades de pareamento

a cadeia correta e a cadeia alterada, ainda não é compreendido, admite-se que a maquinaria tenha capacidade de identificar a fita de DNA molde.

As lesões na molécula de DNA podem ainda modificar as bases, acarretando mudanças nas propriedades de pareamento, quando utilizadas como molde para replicação resultam em moléculas com seqüência diferente da original (veja a Figura 4.9). Esse

efeito também é observado quando na replicação são incorporadas as bases tautoméricas ou análogas das bases (veja a Figura 4.10).

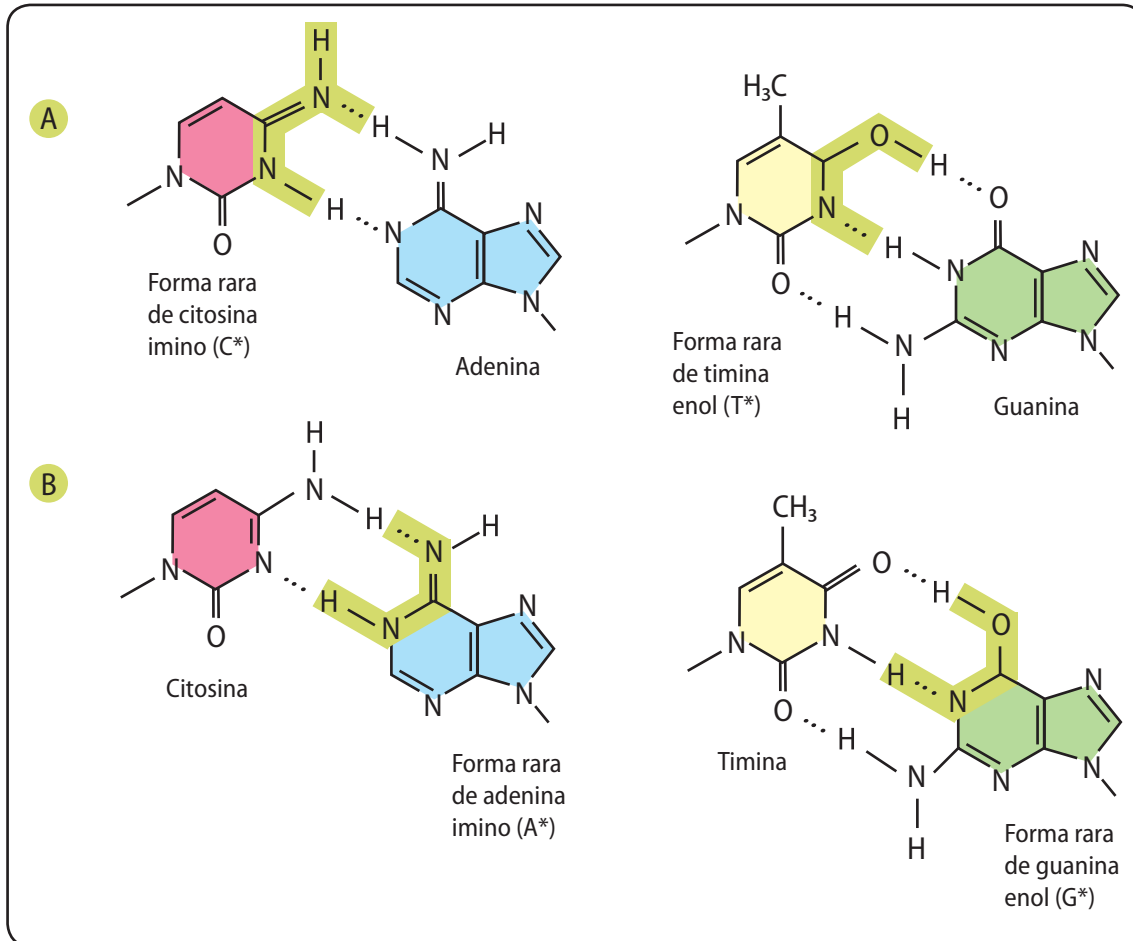


Figura 4.10 - Bases tautoméricas e seu pareamento

A disponibilidade dessas moléculas influencia a frequência dessas lesões. O efeito da incorporação de bases tautoméricas se reflete na replicação do DNA subsequente (veja a Figura 4.11).

As quebras nas ligações glicosídicas resultam em sítios abásicos (sem base) que deformam a molécula (veja a Figura 4.12).

As quebras nas ligações fosfodiéster resultam em quebras na cadeia de DNA que podem ser de fita simples ou fita dupla; essas lesões são críticas para o funcionamento do DNA, bloqueando a DNA e a RNA polimerases. Essas lesões estão associadas ao envelhecimento e ao câncer.

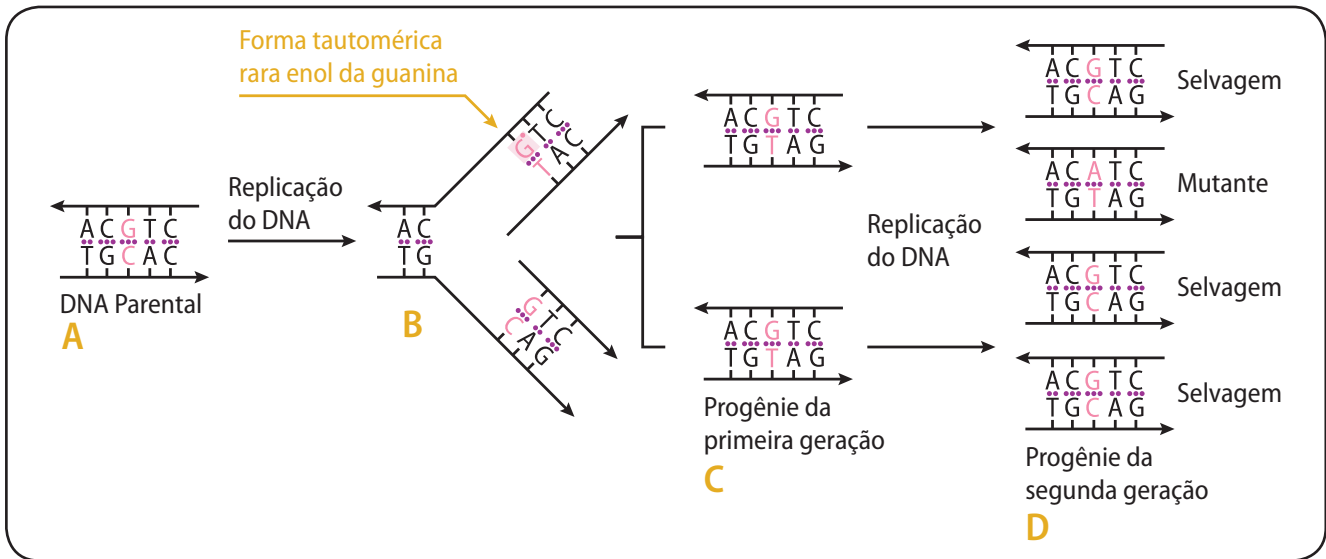


Figura 4.11 - Efeito da incorporação de base tautomérica

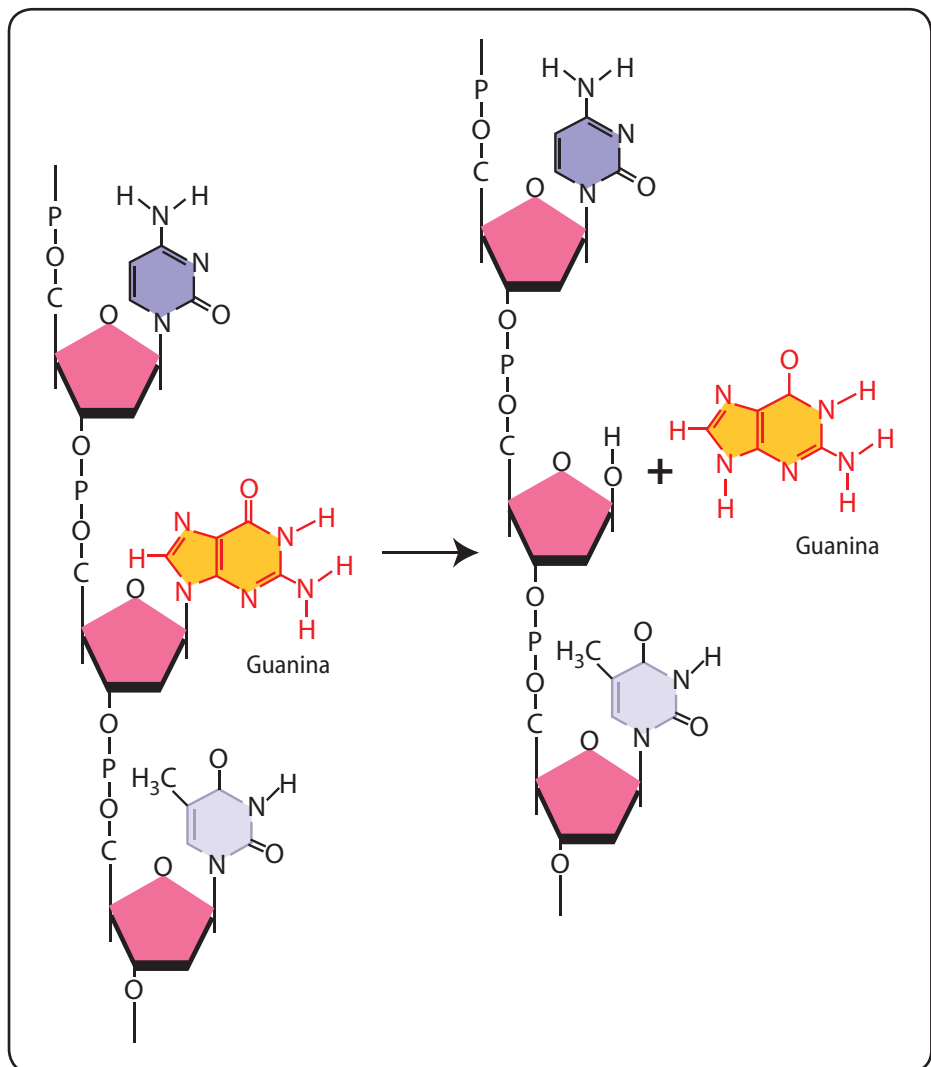
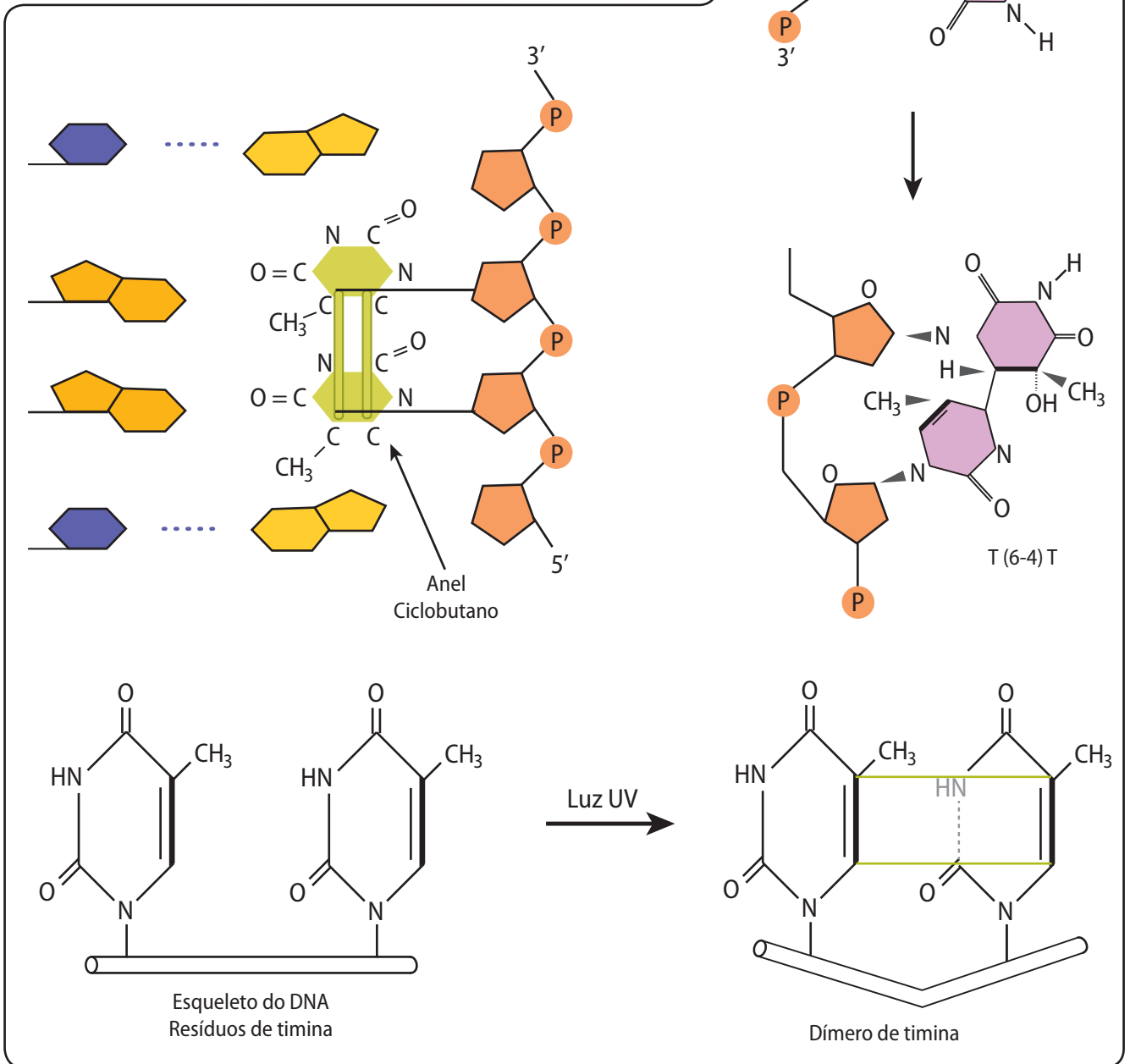


Figura 4.12 - Depurinação do DNA

A radiação ultravioleta produz a formação de dímeros entre pirimidinas adjacentes. Algumas moléculas se intercalam na molécula de DNA, alterando o seu funcionamento (veja a Figura 4.13).

Figura 4.13 - Dímeros de pirimidina formados pela ação da luz UV



Parte dessas alterações resulta da ação de radicais livres gerados pelo metabolismo celular normal e influenciados pela exposição a agentes mutagênicos como a radiação ionizante (veja a Figura 4.14A, B e C).

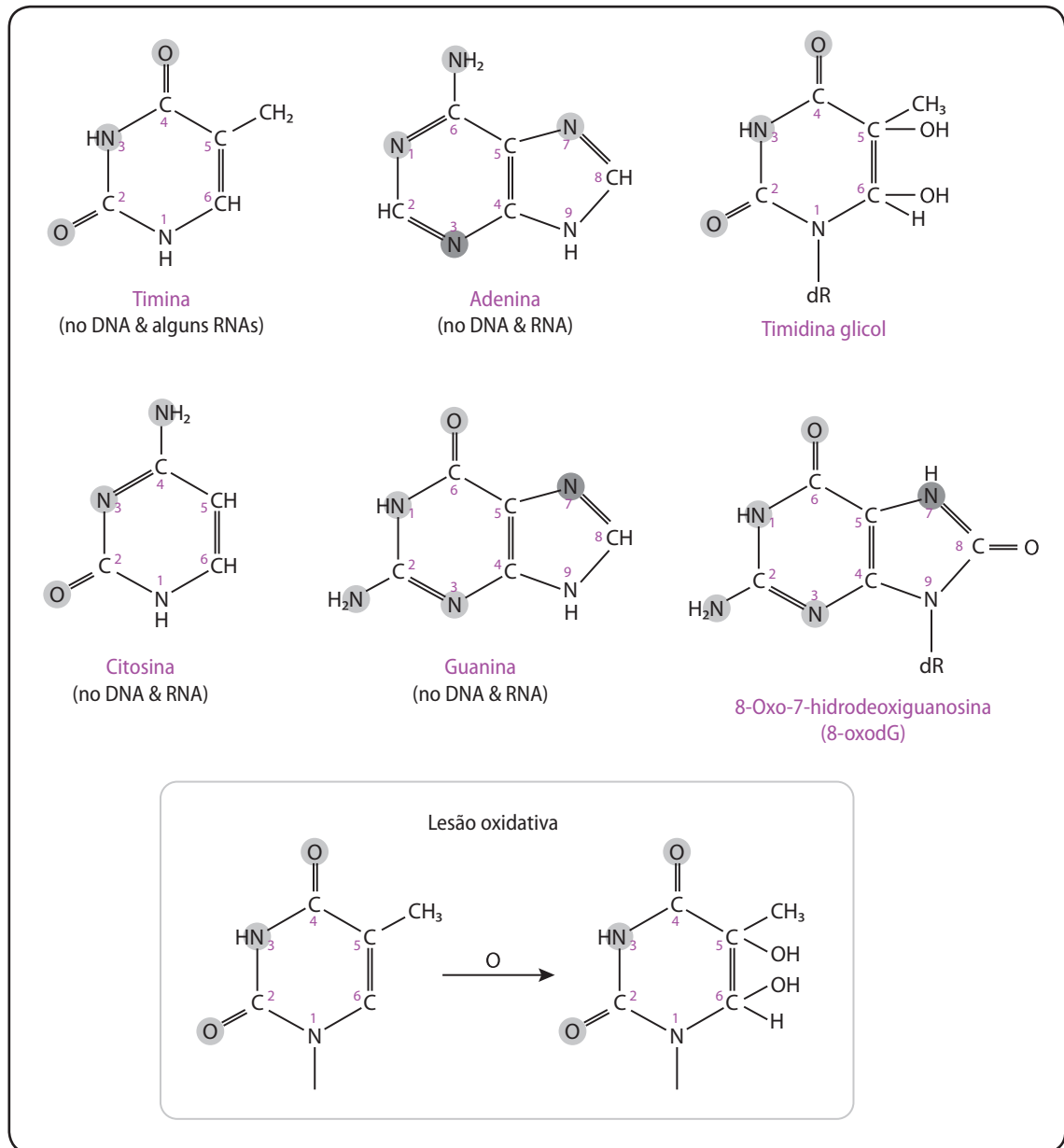


Figura 4.14A - Produtos de dano de DNA formados após o ataque de radicais de oxigênio.

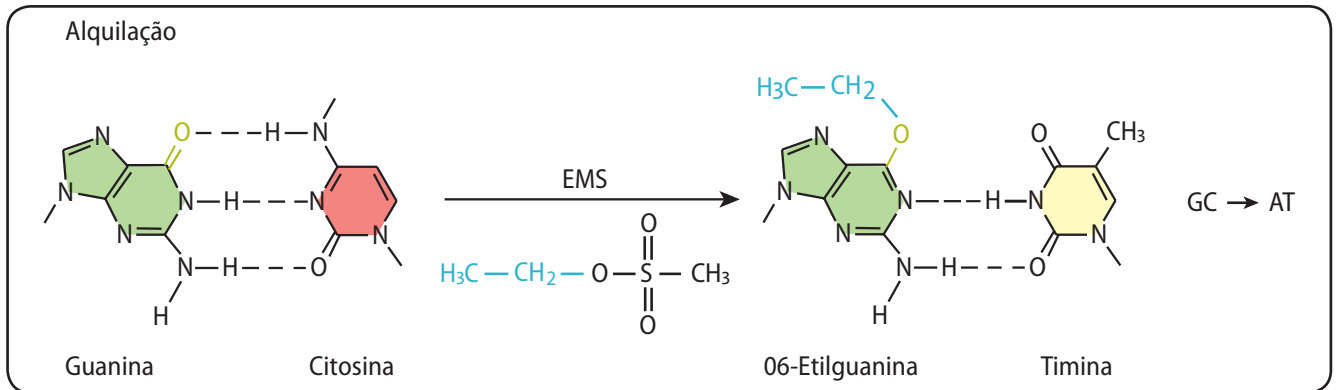


Figura 4.14B - Alquilação causada pelo etil-metano-sulfonato (EMS)

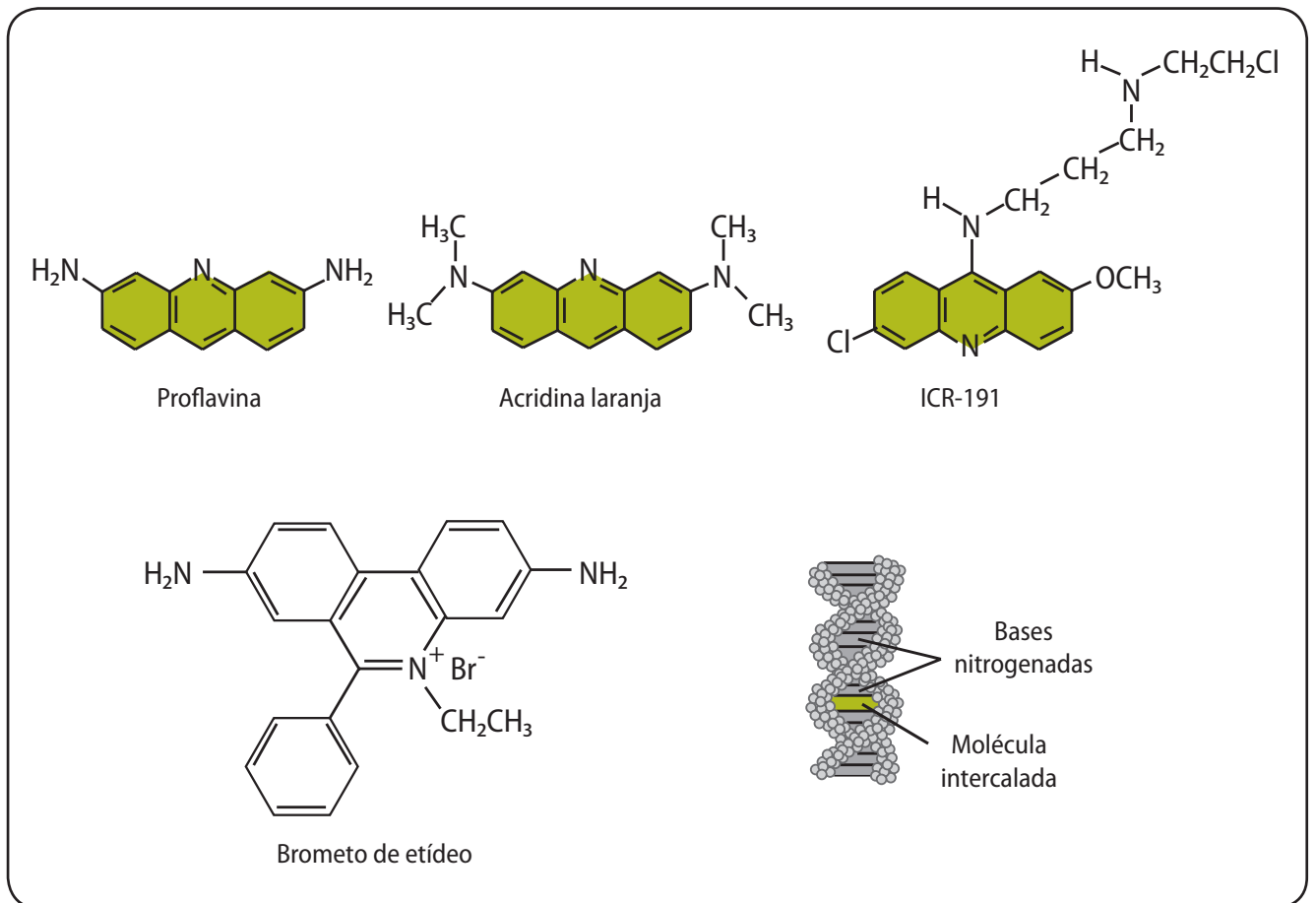


Figura 4.14C - Agentes que se intercalam na molécula de DNA (agentes intercalantes)



## 4.2 Reparo do DNA

Cada agente mutagênico atua de uma maneira específica no DNA, consistindo na assinatura do agente mutagênico, sendo diferenciados os alvos na mesma molécula de DNA (veja a Figura 4.15). A somatória dessas lesões pode ser crítica, de forma que a célula conta com um arsenal de recursos para o reparo dessas lesões (veja a Tabela 4.1). Quando a quantidade de lesões é superior à capacidade de reparo, a célula dispara um mecanismo de morte celular programada (apoptose), que impede a transmissão das lesões para as células-filhas.

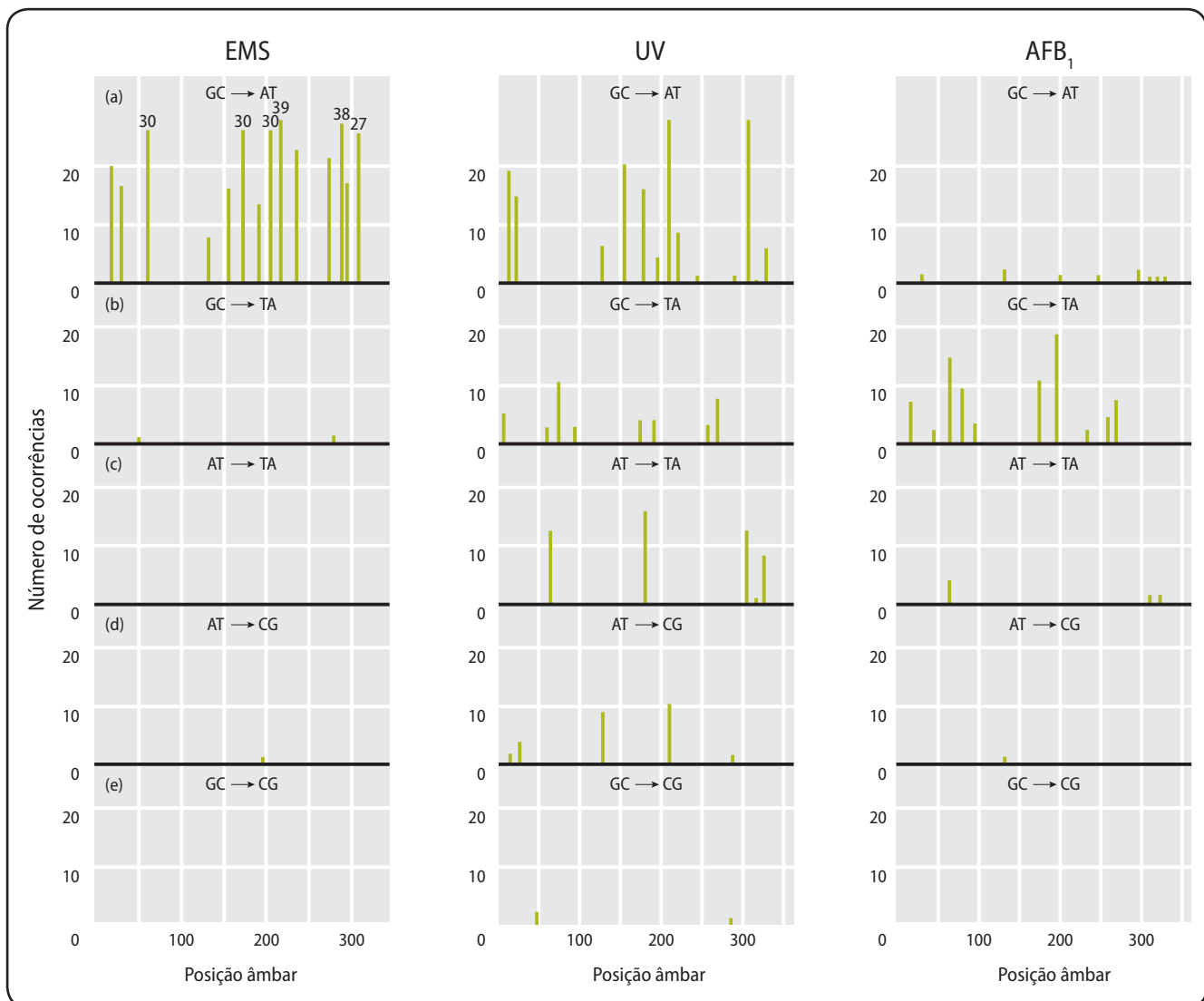
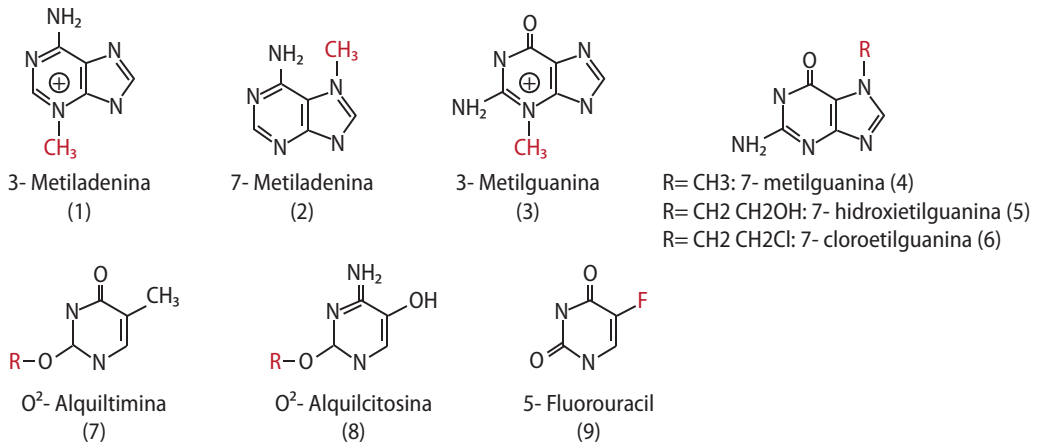


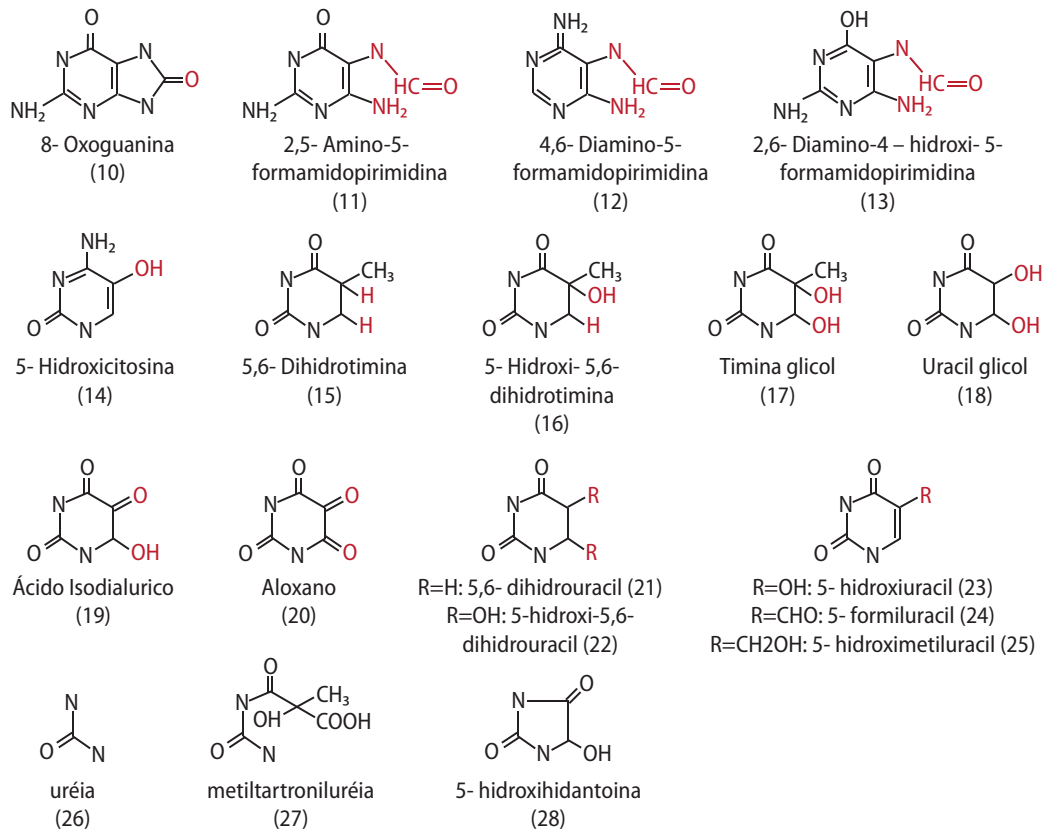
Figura 4.15 - Tipo e posição das mutações induzidas por diferentes agentes mutagênicos no mesmo gene.

*EMS* – etil-metano sulfonato; *UV* – ultra violeta; *AFB<sub>1</sub>* – aflatoxina

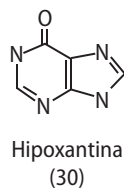
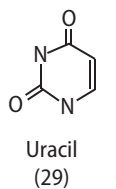
## Bases alquiladas e halogenadas



## Bases oxidadas e com anel fragmentado



## Bases deaminadas



## Outras

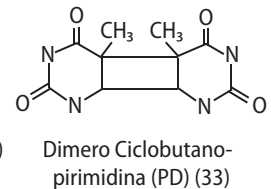
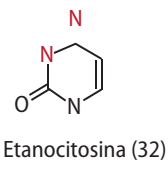
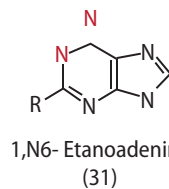


Tabela 4.1 - Principais lesões no DNA e enzimas associadas ao seu reparo

Enzima	Fonte/gene	Substrato relatado no DNA	Atividade de $\beta$ liase
Uracil DNA glicosilase UDGs	Viral	29	Não
	Bacteriana (UNG)	9,29,13	Não
	<i>S. cerevisiae</i> (UNG1)	29	Não
	Plantas	29	Não
	Humana (ung)	19, 20, 23	Não
DNA glicosilases de malpareamento G/T (U)	<i>M. thermoautotropicum</i>	G/G, A/G, T/C, U/C	Não
	Insetos	29	Não
	humana	29	Não
DNA glicosilases bases alquiladas	<i>E. coli</i> (tag)	3	Não
	<i>E. coli</i> (alkA)	8, 24, 25, 30-32 (G, A)	Não
	<i>S. cerevisiae</i> (MAG)	6, 30, 31	Não
	<i>S. pombe</i> (mag1)	1	?
	<i>A. thaliana</i> (MPG)	1	?
	Roedores/humanos (MPG)	6, 10, 30, 31	Não
DNA glicosilases 5- metilcitosinas	Galinha	T em G/T	Não
	Humano	5-MeC	Não
DNA glicosilases específicas para malpareamento de adenina	<i>E. coli</i> (mutY)	A em G/A, C/A	Sim/ não
	Bovina/humana (MYH)	A em 10?A, C/A	Sim?
DNA glicosilases removedoras de pirimidinas oxidadas	<i>E. coli</i> Endo III (nth)	14-18, 20-23, -28	Sim
	<i>S. cerevisiae</i> (NTG1)	11, 12, 13, 17	Sim
	<i>S. pombe</i> (nth)	17,	Sim
	Bovina/humana homóloga Endo III	17	Sim
Endo VIII	<i>E. coli</i>	15,17	Sim
Endo IX	<i>E. coli</i>	26	?
Hidroximetil DNA glicosilase	Camundongo	29	Não
	Bovina	25	Sim ?
Formiluracil DNA glicosilase	Humana	24	?
DNA glicosilases removedoras de purinas oxidadas	<i>E. coli</i> (fpg)	10-13	Sim
	<i>S. cerevisiae</i> (OGG1)	10 (oposta ou T)	Sim
	<i>S. cerevisiae</i> (OGG2)	10 (oposta ou A)	Sim
	<i>D. melanogaster</i> S3	10	Sim
DNA glicosilases dímero de pirimidina	T4	,12	Sim
	<i>M. luteus</i> (pdg)	33	Sim
	<i>N. mucosa</i>	33	Sim

Tabela 4.1 (cont.) - Principais lesões no DNA e enzimas associadas ao seu reparo

Outra opção é o sistema SOS, em que são recrutadas proteínas que se ligam à molécula de DNA lesada, possibilitando a síntese translesão, que, no entanto, apresenta pouca acurácia.

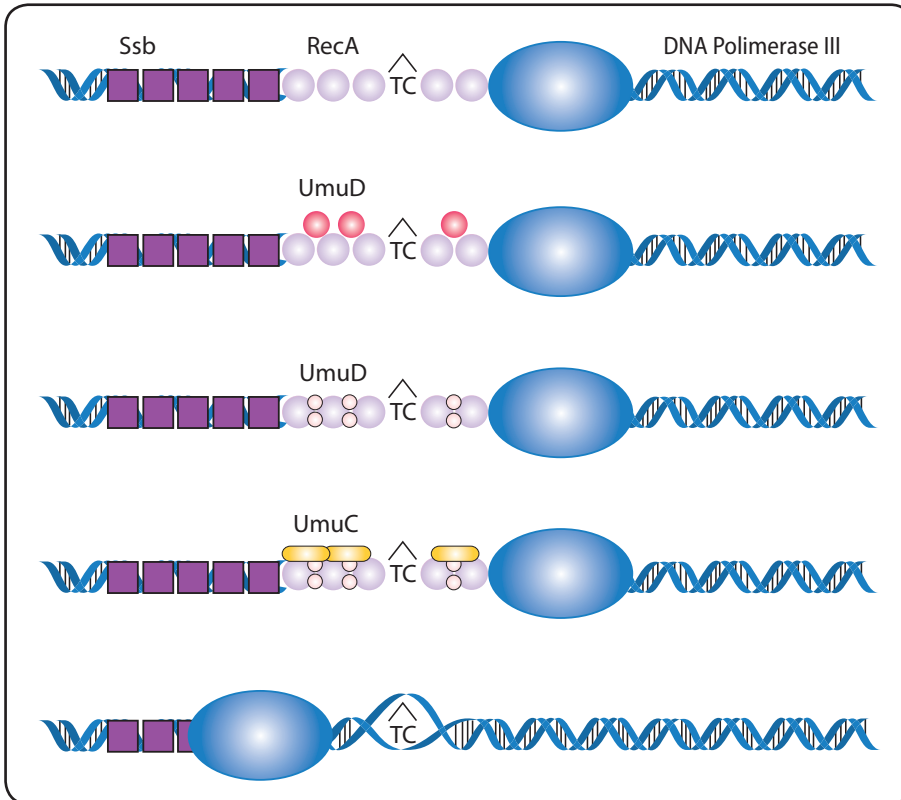


Figura 4.16 - Sistema SOS. A DNA polimerase III pára em uma lesão, gerando regiões unilamentares que atraem as proteínas Ssb e RecA, que formam filamentos. A presença dos filamentos RecA ajudam a ativar a célula para a produção de UmuD, que é clivada por RecA para produzir UmuD' e UmuC. A UmuC é recrutada para formar um complexo com UmuD permitindo que a polimerização pela DNA polimerase continue além da lesão bloqueadora

O reparo por fotorreativação é específico para os dímeros de pirimidina, as ligações entre as pirimidinas adjacentes são quebradas por uma enzima específica, a fotoliase, a atividade dessa enzima é dependente da luz branca. Essa enzima é amplamente conservada, estando presente em procaríotos e eucariotos (veja a Figura 4.17). É curioso que os mamíferos placentários não produzem essa enzima, não efetuando esse tipo de reparo de DNA.

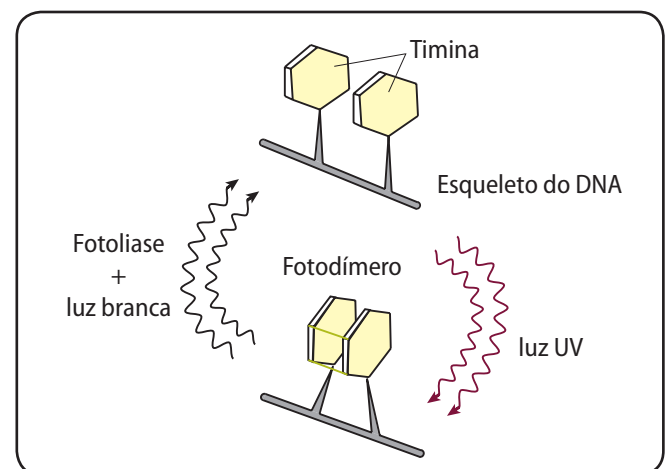


Figura 4.17 - Reparo do fotodímero induzido por UV

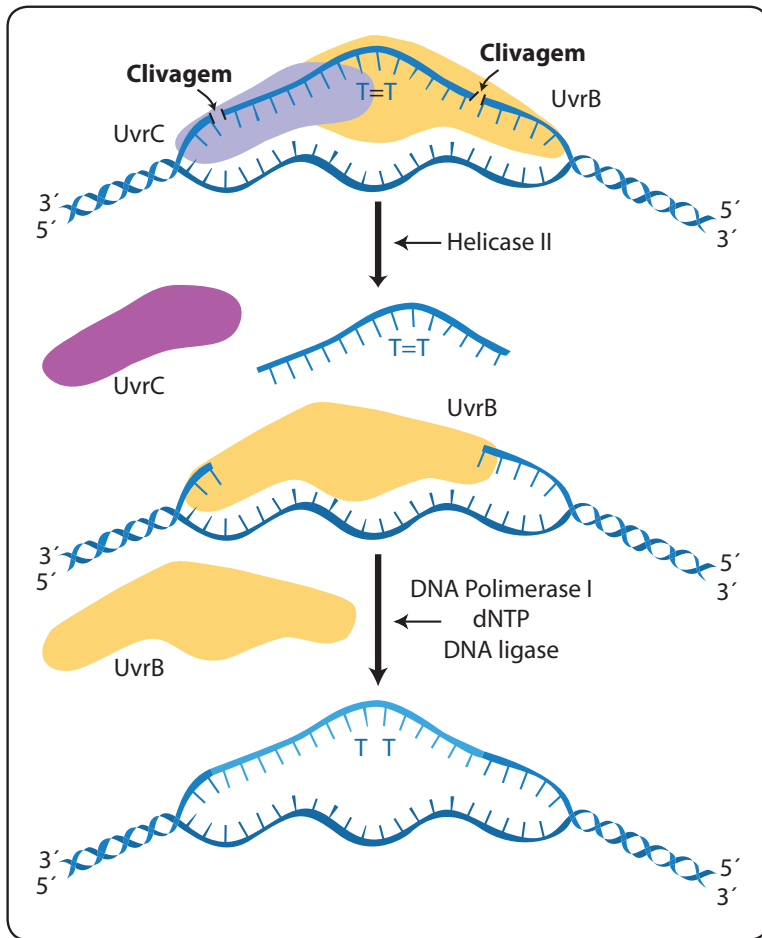


Figura 4.18 - Reparo de excisão de nucleotídeo para remoção de dímero de timidina



Figura 4.19 - Paciente com *Xeroderma pigmentosum*

As células humanas não realizam reparo por fotorreativação, os dímeros de pirimidina são reparados por outro processo denominado reparo de excisão de nucleotídeos. Este processo foi o primeiro mecanismo de reparo de DNA reconhecido. Em 1957, Philip Hanawalt, estudando os efeitos da radiação ionizante, verificou que, após serem submetidas à radiação, as células em cultura sintetizam pequenas quantidades de DNA, que foi denominada síntese de DNA não programada (UDS – Unscheduled DNA Synthesis), ou seja, que ocorre fora da fase S. Nessa ocasião propôs que poderia ocorrer um processo de reparo de DNA, revelando um tipo de atividade ainda insuspeitável. Posteriormente, foram identificadas diversas enzimas que atuam nesse processo, as quais

também são conservadas evolutivamente. O reconhecimento da lesão é feito por proteínas específicas que sinalizam a região do DNA lesada, uma endonuclease produz uma clivagem no trecho da fita de DNA com a lesão, a seguir uma DNA helicase separa o fragmento de DNA com a lesão, expondo a cadeia complementar, é então recrutada uma DNA polimerase específica que sintetiza um pequeno trecho da molécula de DNA, sendo o processo finalizado pela DNA ligase (veja a Figura 4.18).

Em humanos foi identificada uma condição de herança mendeliana associada com a inativação desse processo de reparo do DNA, o Xeroderma pigmentosum. Essa condição apresenta heterogeneidade genética, e os genes associados codificam proteínas indispensáveis para o reparo de DNA por excisão de nucleotídeos (veja a Figura 4.19).

Verifica-se que a eficiência do reparo é maior nas regiões transcritas, evidenciando que o bloqueio da RNA polimerase dispara o mecanismo de reparo de DNA (Figura 4.20). Em humanos está presente um sistema adicional que é o reparo genômico global, o qual ocorre em regiões transcritas e não transcritas (veja a Figura 4.21).

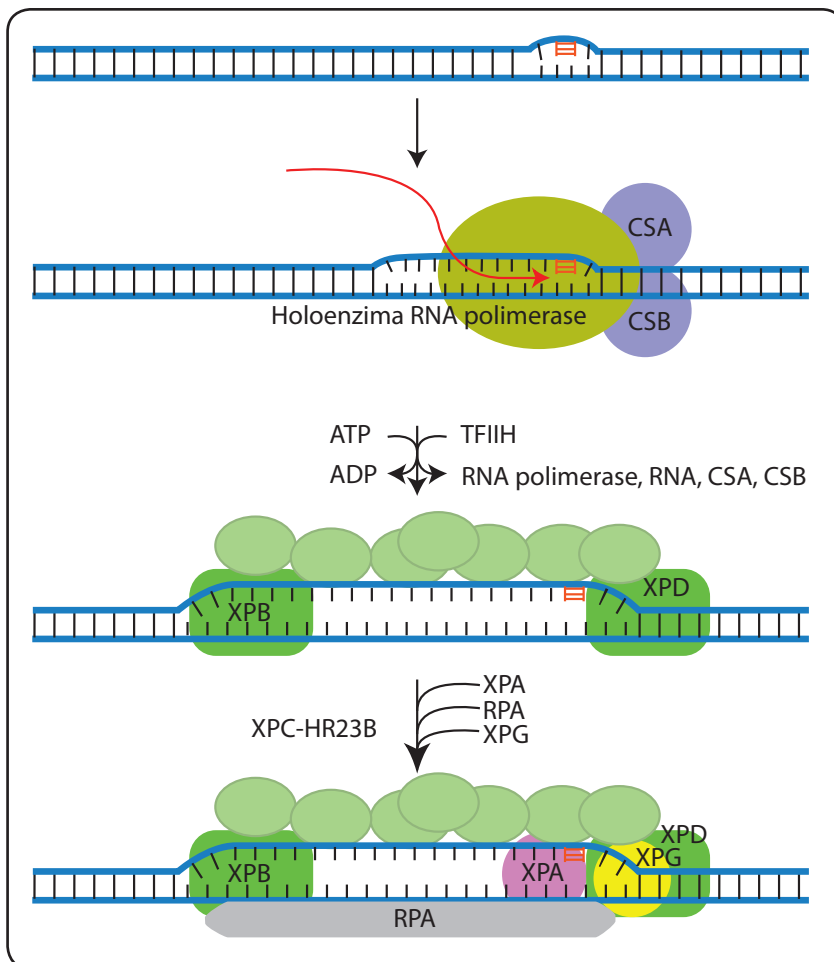


Figura 4.20 - Reparo de DNA acoplado a transcrição

Outro processo de reparo é o reparo por excisão de bases, uma enzima denominada DNA glicosilase produz uma clivagem na ligação glicosídica, removendo a base lesada, a seguir é recrutada uma AP endonuclease que produz uma quebra na ligação fosfodiéster, removendo o nucleotídeo. A seguir uma DNA polimerase recompõe a fita, processo finalizado pela DNA ligase (veja a Figura 4.22). Esse processo é utilizado para a remoção de bases alteradas.

Figura 4.21 - Reparo genômico global

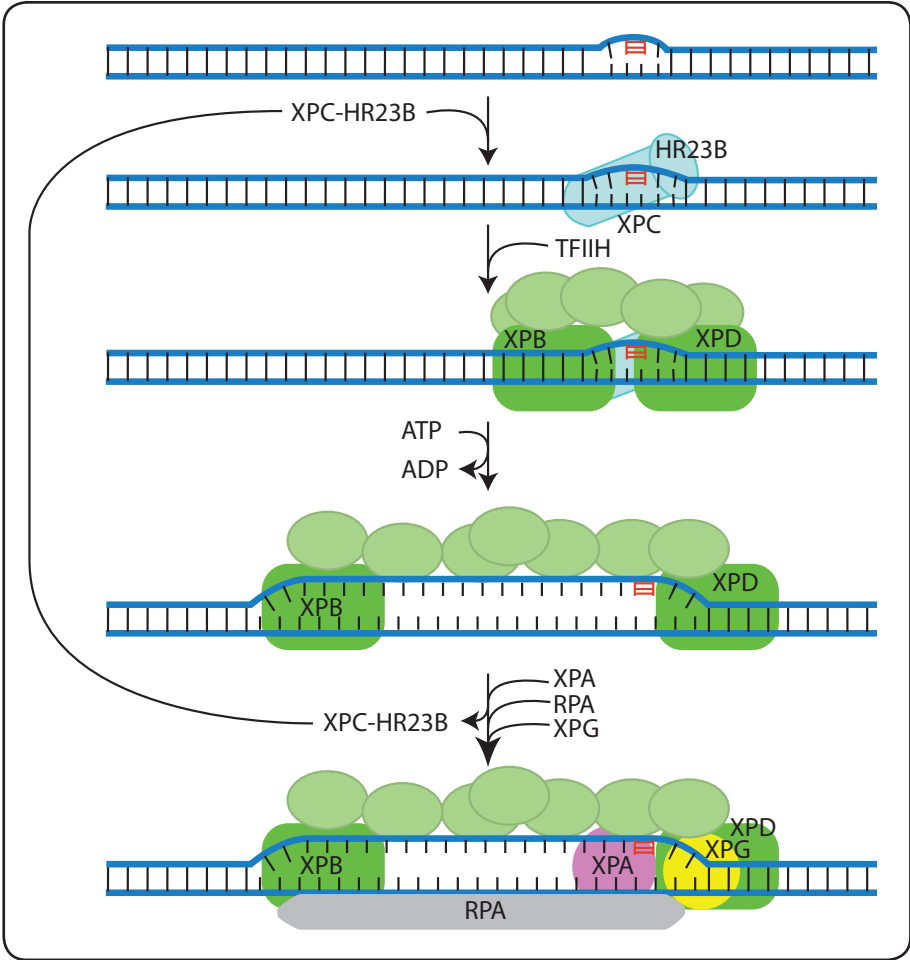
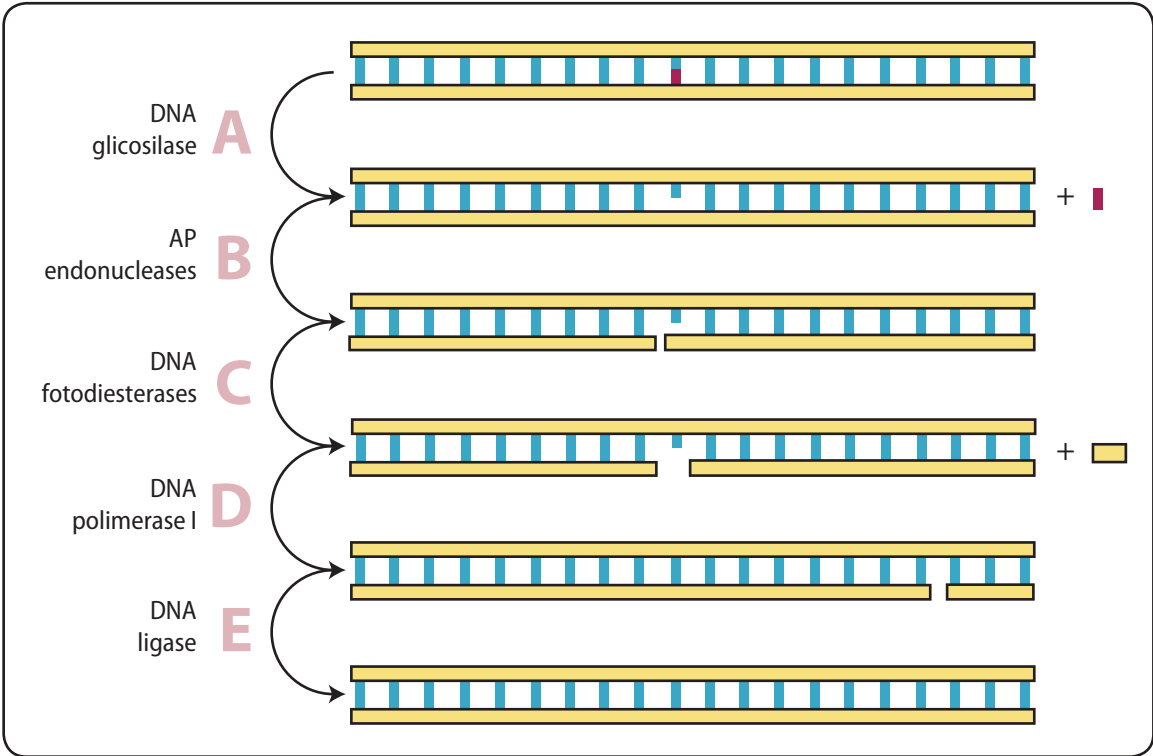


Figura 4.22 - Reparo por excisão de base



Em situações em que há grande quantidade de lesões, tornando inviável o reparo individual de cada uma, há a possibilidade de síntese de DNA translesão (bypass), esse processo convoca uma DNA polimerase especial, e embora seja possível a síntese do DNA, a acurácia é comprometida, sendo uma situação-limite (veja a Figura 4.23).

O reparo pós-replicação envolve a recombinação com o cromossomo homólogo e são utilizadas quebras de cadeia do DNA. Esse processo tem certa similaridade com o processo de recombinação que ocorre na meiose (veja as Figuras 4.24 e 4.25).

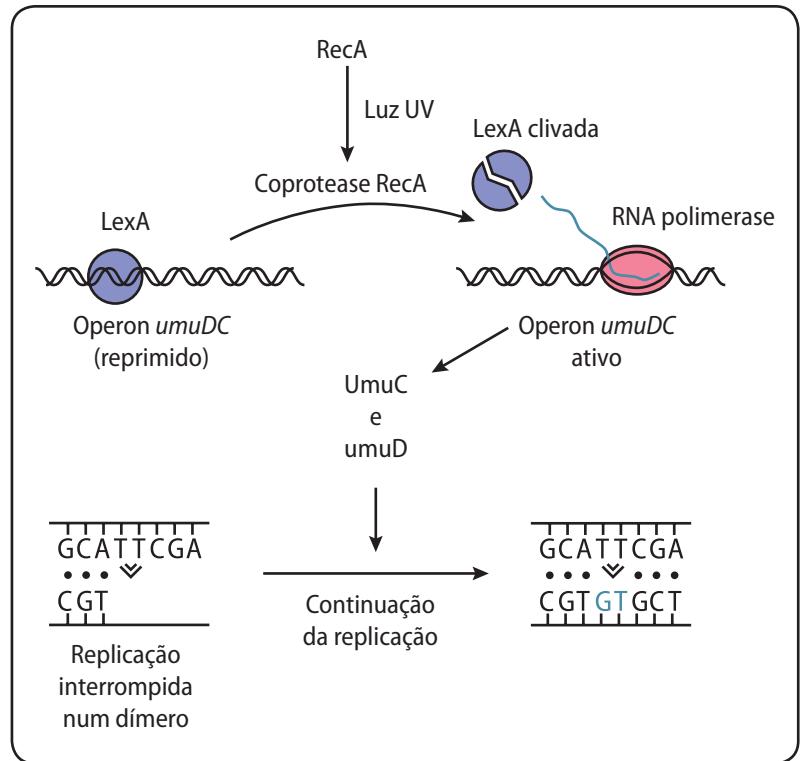


Figura 4.23 - Síntese translesão

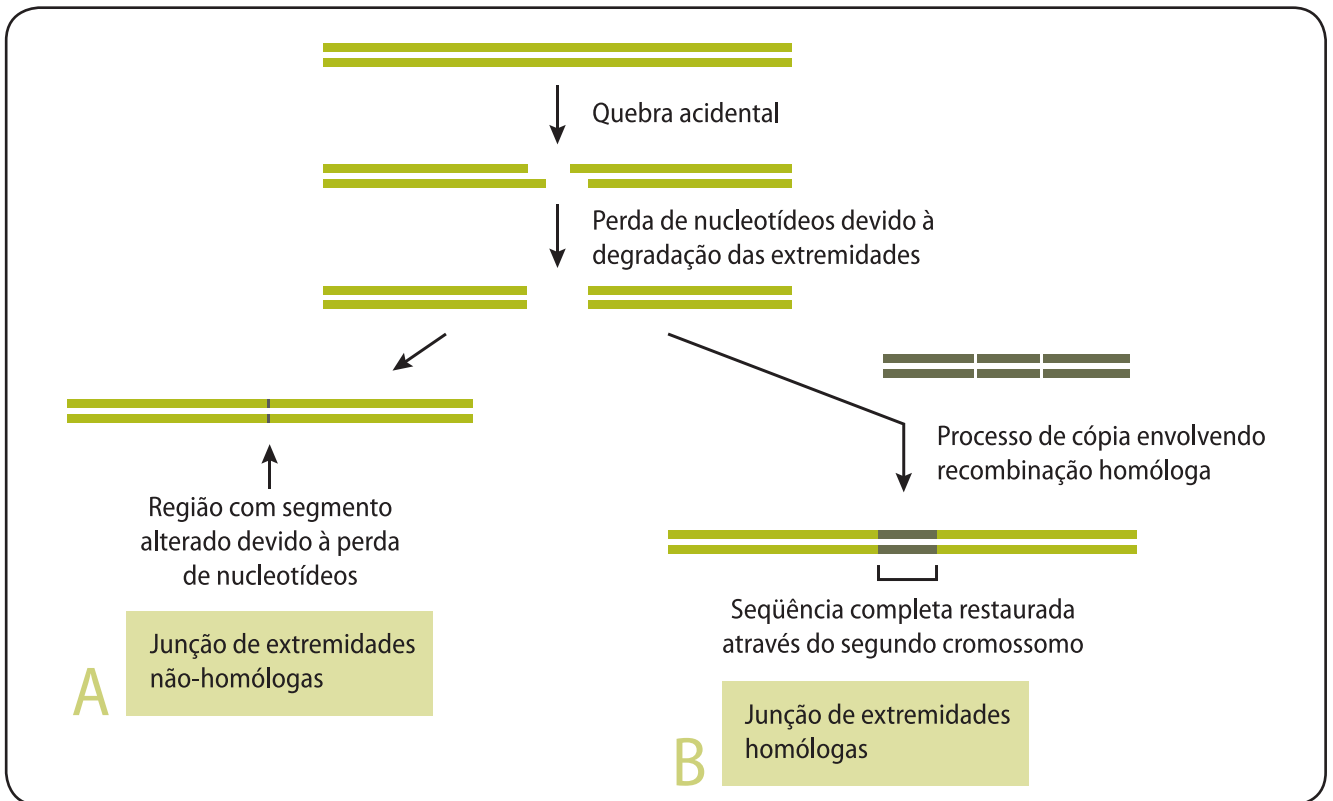


Figura 4.24 - Reparo de quebra de cadeia



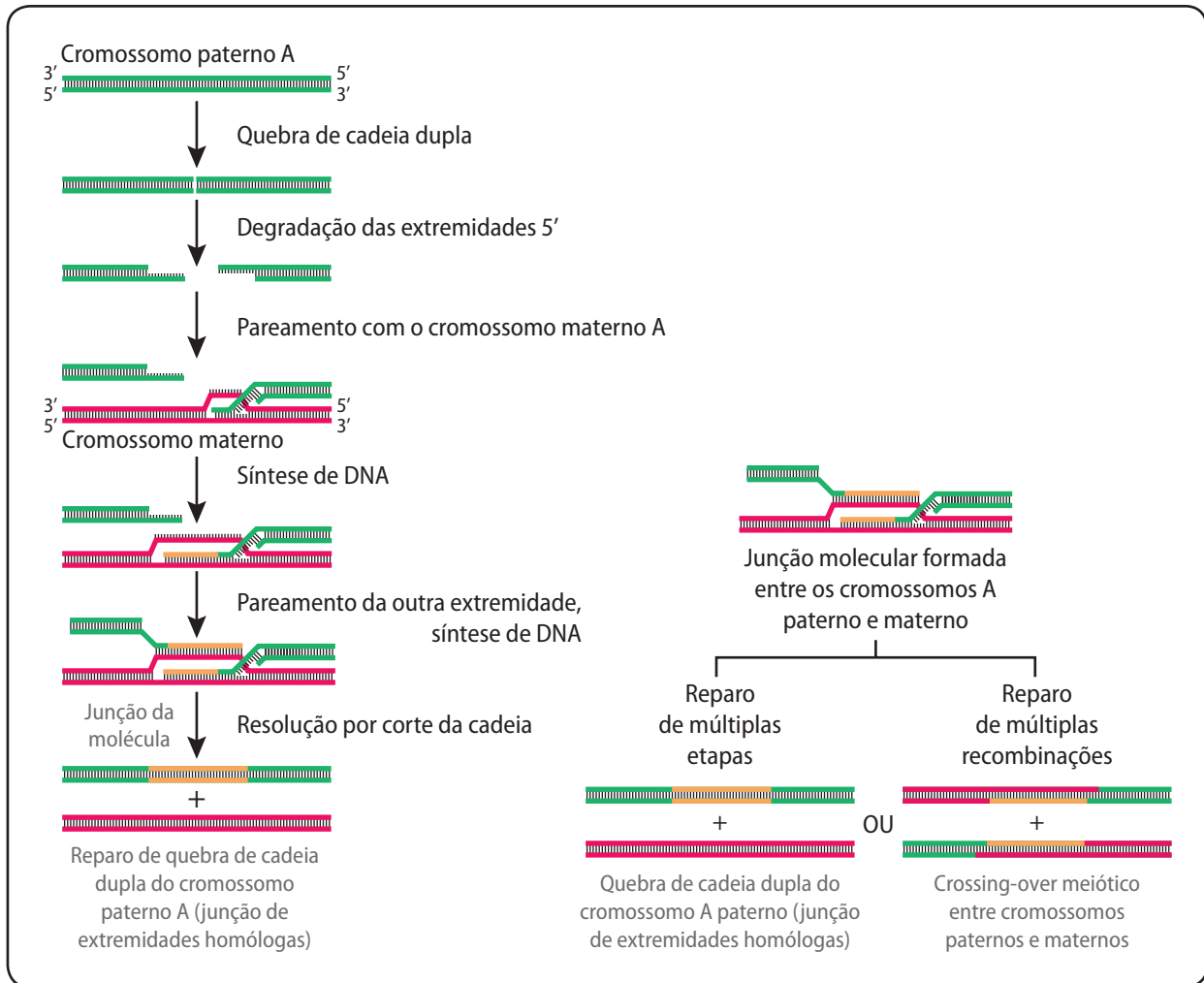


Figura 4.25 - Reparo por recombinação

Outra alternativa celular, quando estão presentes lesões acima da capacidade de reparo do DNA, é a morte celular ou a apoptose, nesse processo o DNA é fragmentado nos nucleossomos (~200 pb), inviabilizando a continuidade celular. A proteína p53 desempenha um papel fundamental nesse processo.

É curioso que parte dos genes identificados à susceptibilidade hereditária ao câncer atua no processo de reparo de DNA, demonstrando que a diminuição da capacidade de reparo de DNA possibilita o acúmulo de lesões críticas na célula, o que produz a transformação neoplásica.

Embora a continuidade da informação disponha de uma infraestrutura celular refinada, a variabilidade da informação é o condutor da evolução, possibilitando adaptação. Quanto maior a variabilidade maior a viabilidade da espécie.

### 4.3 Recombinação

A recombinação é um fenômeno celular normal. Entre moléculas que apresentam **homologia**, a **recombinação homóloga** ocorre durante a meiose entre **cromossomos homólogos**.

Também é reconhecida a recombinação sítio específica que ocorre entre fitas de DNA que tenham alguma região de homologia e está associada à integração de genomas virais no hospedeiro e na transposição de segmentos de DNA no genoma.

Esses eventos contribuem para o aumento da variabilidade genética e desempenham um papel fundamental na evolução.

### 4.4 Recombinação Homóloga ou Generalizada

Essa forma de recombinação ocorre no início da meiose, sendo evidenciada citologicamente como quiasmas, no início da prófase I.

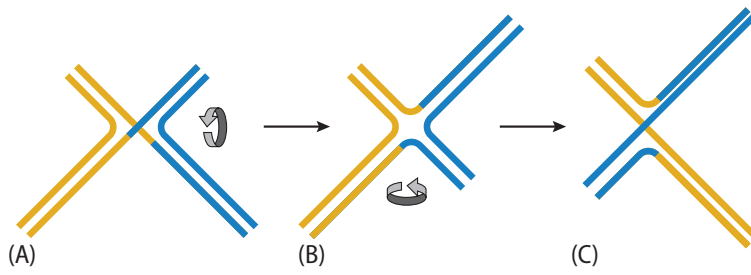
É um evento gerador de diversidade dos gametas e permite que mutações favoráveis ou desfavoráveis sejam separadas e testadas como unidades individuais em novas combinações.

Não há informação detalhada acerca dos eventos moleculares envolvidos na recombinação em células de eucariotos superiores.

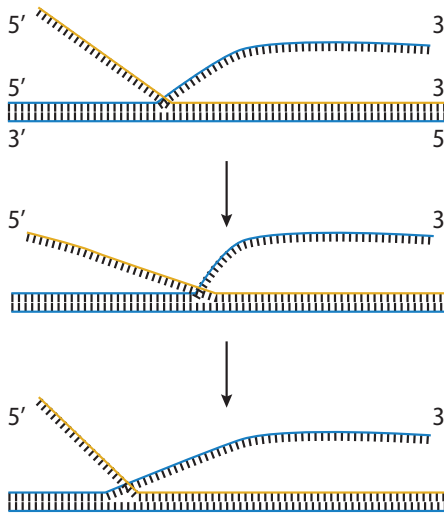
No início da meiose, os cromossomos estão duplicados, sendo constituídos de duas fitas duplas de DNA, com quatro fitas duplas nos cromossomos homólogos e duas fitas envolvidas no evento. Esse processo acarreta troca física do DNA, ocorrendo quebra e reunião nas moléculas (veja a Figura 4.26).

Além dos intrincados eventos moleculares que possibilitam a reunião de moléculas de origem diferente, outros detalhes ainda não são esclarecidos. A frequência é diferente no sexo feminino e no masculino, em algumas espécies ocorre apenas na meiose feminina. Também é influenciada pela estrutura do cromossomo, ocorrendo em menor frequência nas regiões de heterocromatina.

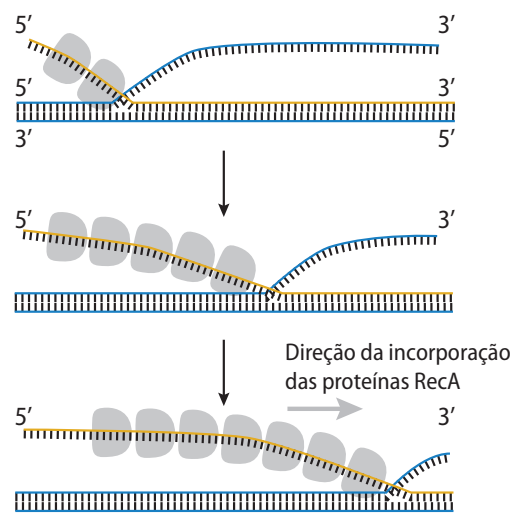
Os estudos de recombinação possibilitaram a elaboração dos primeiros mapas cromossômicos, onde quanto maior a proximi-



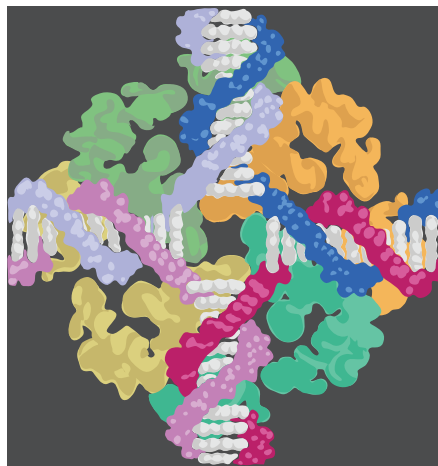
Junção Holliday e resolução das fitas



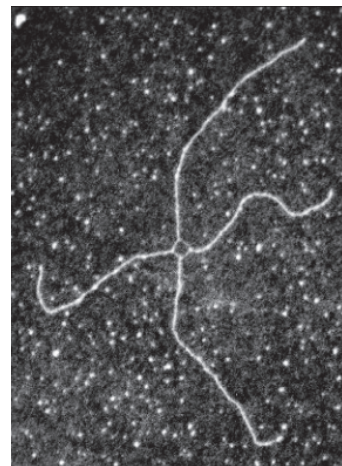
(A) Migração espontânea da ramificação



(B) Migração da ramificação direcionada por proteínas



Modelo proposto



Visualização ao microscópio eletrônico

Figura 4.26 - Recombinação homóloga

dade dos genes, menor a possibilidade de serem separados por um evento de recombinação. Essa metodologia ainda é utilizada para mapeamento de QTLs (Quantitative Trait Locus) no genoma.

## 4.5 Recombinação Sítio Específica

A recombinação sítio específica promove a integração de genomas virais no cromossomo bacteriano, esse fenômeno envolve seqüências específicas do DNA viral e do DNA bacteriano, com uma região de homologia. As enzimas envolvidas atuam somente sobre o par específico de alvo. Esse evento também possibilita a integração do plasmídio ao cromossomo bacteriano.

## 4.6 Transposição

Os experimentos realizados com milho na década de 1930 levaram Barbara McClintock a propor que deveria haver genes saltadores, explicação que seria plausível para o desvio das proporções mendelianas verificadas em alguns cruzamentos (veja a Figura 4.27).

Atualmente, constatou-se não apenas a pertinência dessa suposição, mas também a sua abrangência.

Esse processo permite que uma seqüência de DNA seja inserida em outra, sem a dependência de homologia entre elas, ocorre quebra e reunião das fitas de DNA envolvidas.

Os elementos de transposição ou transposons são seqüências discretas do genoma que têm capacidade de mobilização, a qual pode se efetivar através de DNA ou de RNA.

As regiões mobilizadas são denominadas seqüências de inserção (IS), são constituintes normais dos cromossomos. São delimitadas por repetições terminais invertidas que contêm uma região codificante para a enzima transposase relacionada aos eventos bioquímicos da transposição.

A seqüência é transposta para um sítio-alvo no DNA hospedeiro, também caracterizado por uma seqüência de DNA específica que

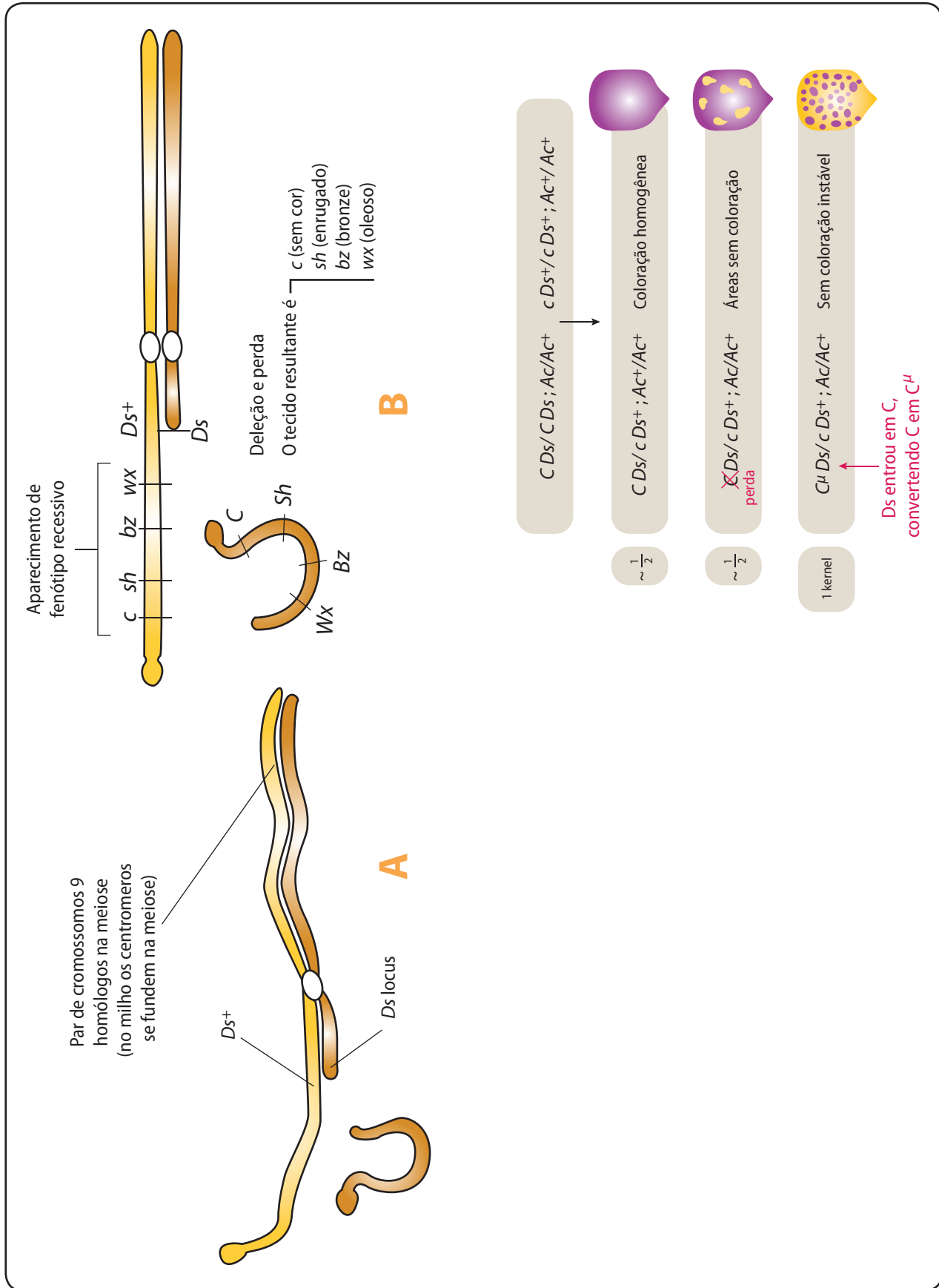


Figura 4.27 - Genes saltadores

é duplicada no sítio da inserção. A frequência de transposição é da ordem de aproximadamente  $10^{-3}$ - $10^{-4}$ /elemento/geração. A ocorrência de transposição é mais bem documentada em vegetais.

Os elementos controladores foram identificados no milho, pois quando inseridos próximos a genes que produzem efeitos visíveis sobre o fenótipo, uma vez que afetam a atividade dos genes adjacentes ao local de inserção. O evento da transposição pode causar a inativação do gene-alvo, um evento subsequente pode excisar o segmento transposto e o gene-alvo voltar à constituição original. Esse evento está associado à variegação.

Do ponto de vista estrutural, somente as extremidades são necessárias para que um transposon sirva de substrato para a transposição, podendo utilizar o gene da transposase de outro elemento presente na célula para efetivar a transposição. Os elementos autônomos são capazes de excisão e transposição, gerando alelos instáveis. Os elementos não autônomos são mais estáveis, não sofrem transposição independentemente e dependem de um membro autônomo da mesma família. Esses elementos perderam as seqüências internas, mas possuem as repetições invertidas nas extremidades (veja a Figura 4.28).

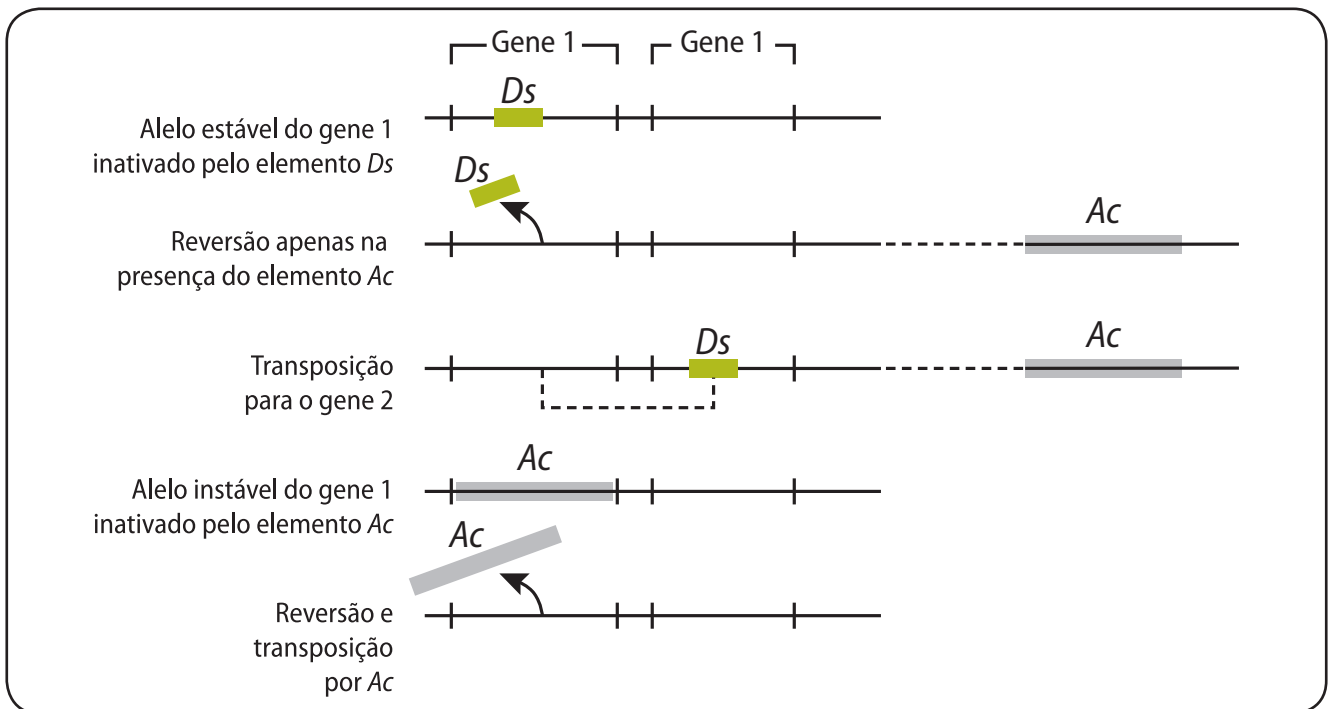


Figura 4.28 - Elementos de transposição

Também existem transposons que contêm genes adicionais, denominados transposons compostos. Esse tipo de transposição está associado ao aparecimento de resistência a antibióticos em bactérias (veja a Figura 4.29).

O mecanismo de transposição envolve a proteína transposase, sendo conhecidos transposons que se mobilizam por um mecanismo de corte e colagem (veja a Figura 4.30) e elementos com transposição conservativa, quando uma cópia do transposon é inserida em outro segmento do DNA (veja a Figura 4.31).

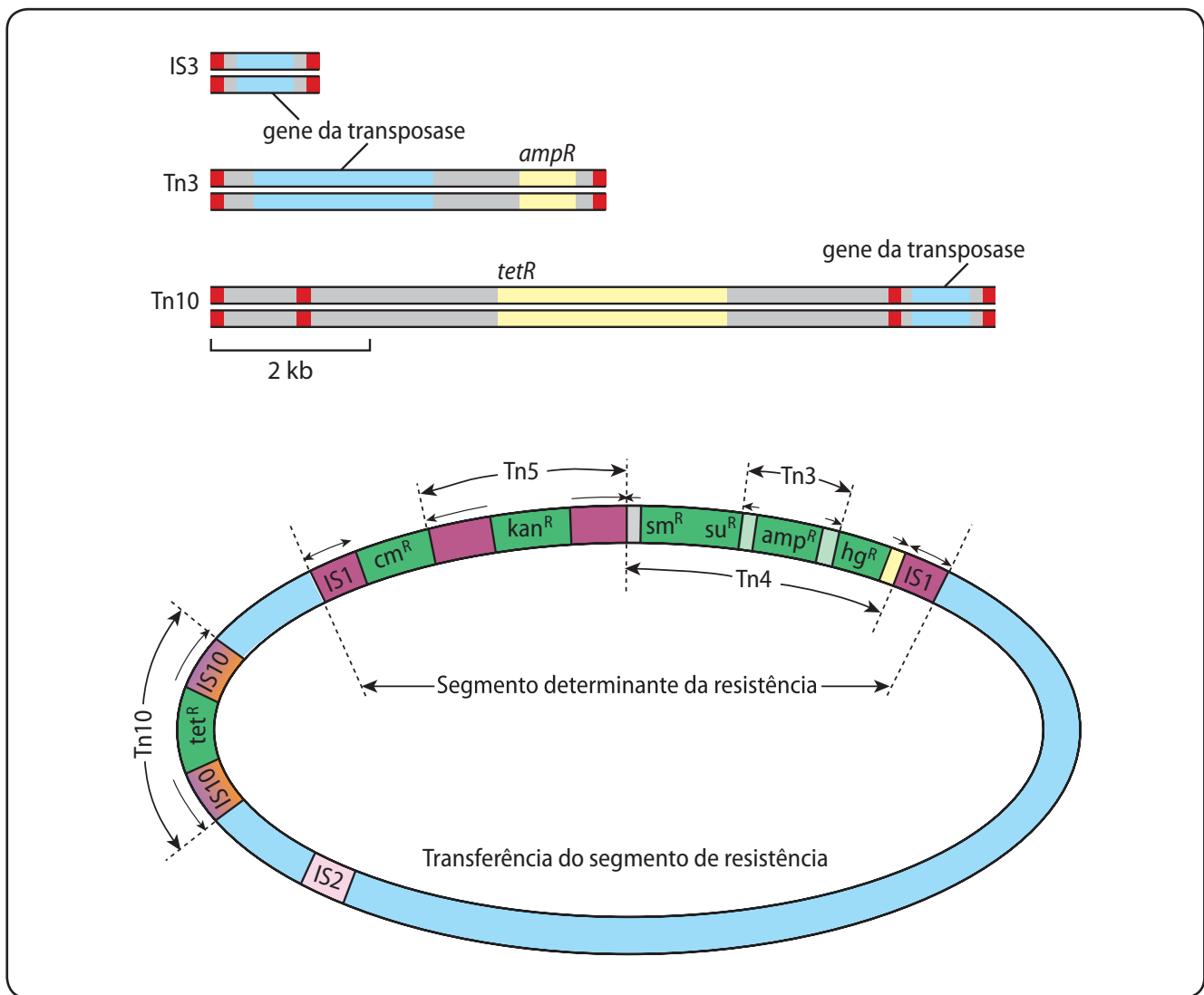


Figura 4.29 - Elementos de transposição compostos com genes de resistência a antibióticos

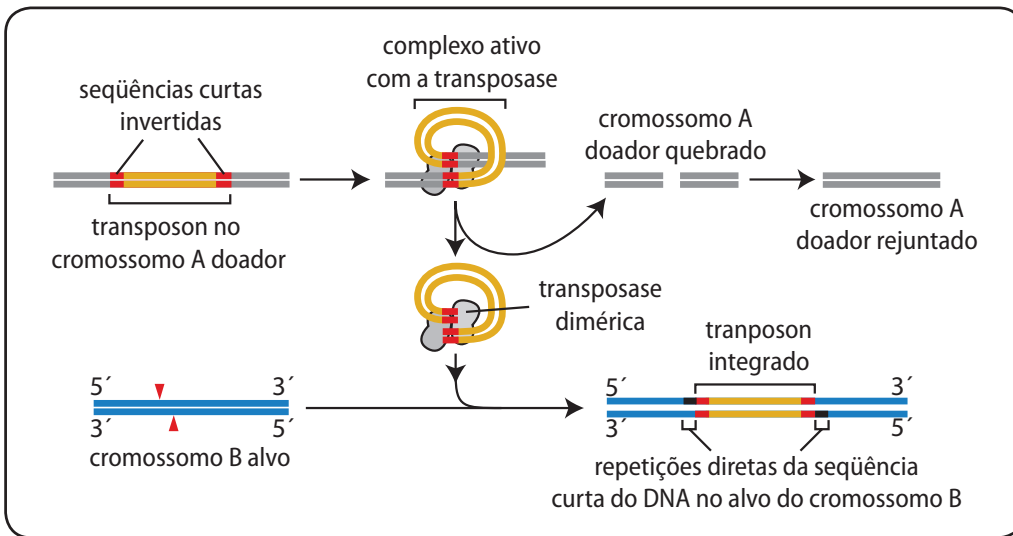


Figura 4.30 - Transposição entre dois cromossomos diferentes

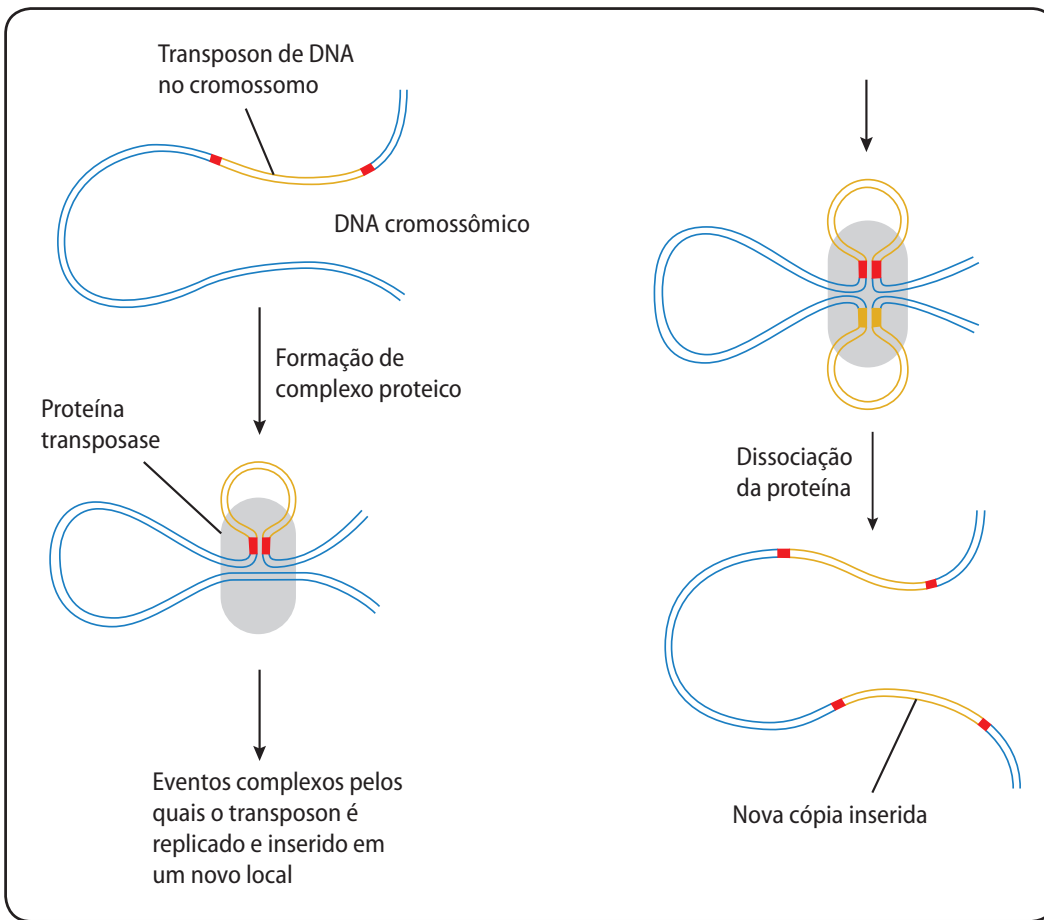


Figura 4.31 - Transposição conservativa no mesmo cromossomo



## 4.7 Retrovírus e Retrotransposons

Os retrovírus são vírus de RNA de fita simples que se replicam por intermediário de DNA de fita dupla. Esse DNA de fita dupla é inserido no genoma do hospedeiro, sendo então denominado pró-vírus num evento semelhante à transposição (veja a Figura 4.32). *A integração é uma etapa normal do ciclo vital e necessária para a transcrição.* Após a integração, o pró-vírus pode permanecer

no genoma da célula hospedeira e sendo replicado juntamente com o DNA genômico (ciclo lisogênico) ou então iniciar um processo de síntese de RNA viral e proteínas virais, utilizando a maquinaria da célula hospedeira, a qual resulta na produção de novos vírus e acarreta lise da célula hospedeira, liberando a progênie viral (veja a Figura 4.33).

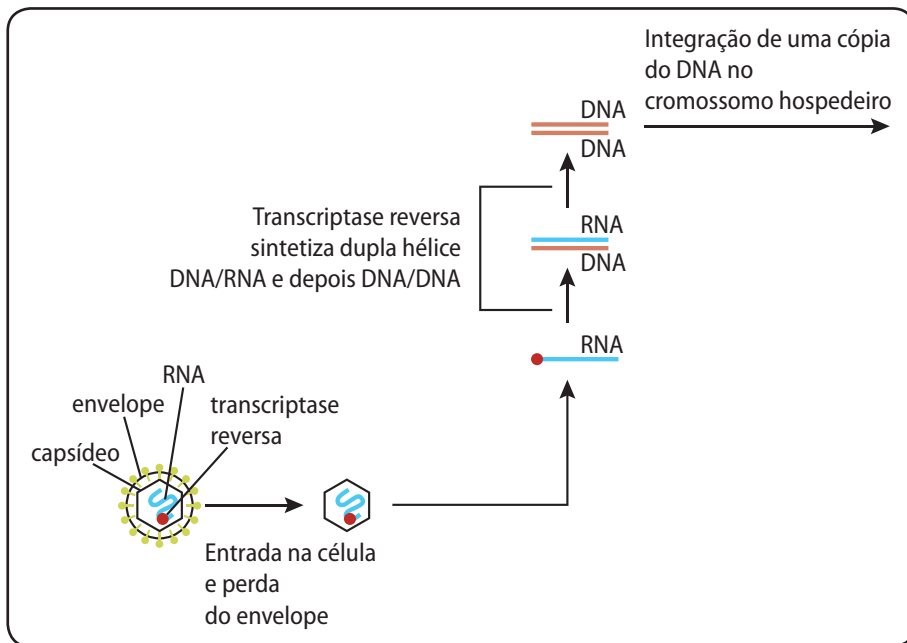


Figura 4.32 - Integração do DNA viral no cromossomo hospedeiro

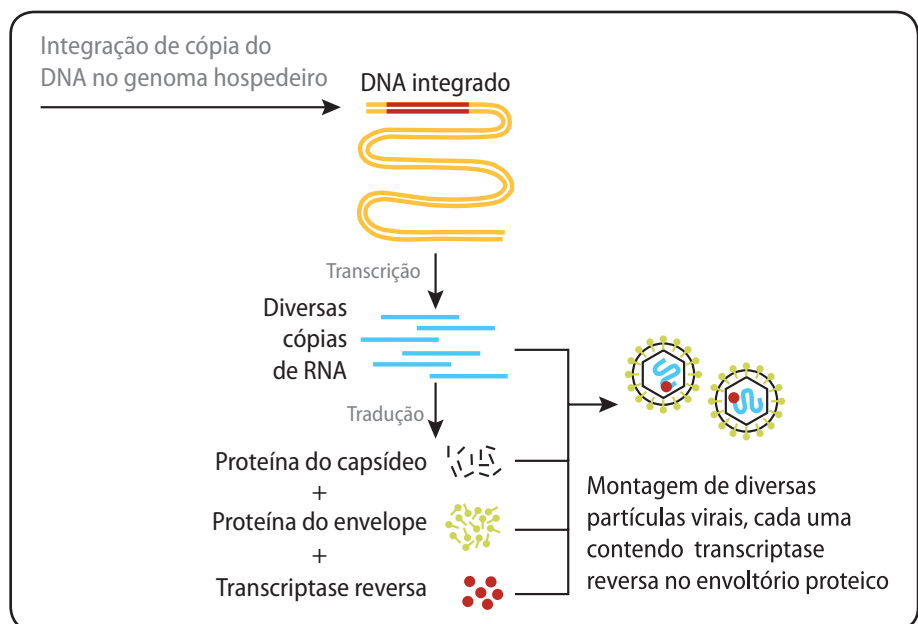
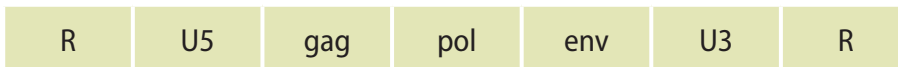


Figura 4.33 - Eventos na célula infectada por retrovírus

As seqüências retrovirais contém três ou quatro genes na seguinte ordem:

- gag – MA (matriz), CA (capsídeo principal), NC (capsídeo nuclear);
- pol – PR (protease), RT (transcriptase reversa), IN (integrase);
- env – SU (proteína de superfície), TM (transmembrana).

A extremidade do cromossomo viral é constituída por regiões terminais invertidas, similares às seqüências dos transposons, e a inserção ocorre em regiões específicas que também são duplicadas.



Os retrotransposons são transposons que se mobilizam via RNA, também dependentes da transcriptase reversa que produz a conversão do RNA em DNA. A principal diferença entre retrovírus e retrotransposons é de que os retrovírus têm capacidade de infectar outras células. A integração do retrovírus é bem documentada em **procariotos**, estando associada ao fenômeno da lisogenia. Os retrotransposons só podem se mobilizar na mesma célula (veja a Figura 4.34).

As diversas formas de transposição (veja a Tabela 4.2) contribuem expressivamente para a variabilidade genética, sendo encontrados na maioria dos genomas vestígios de transposição, também denominados DNA fósseis.

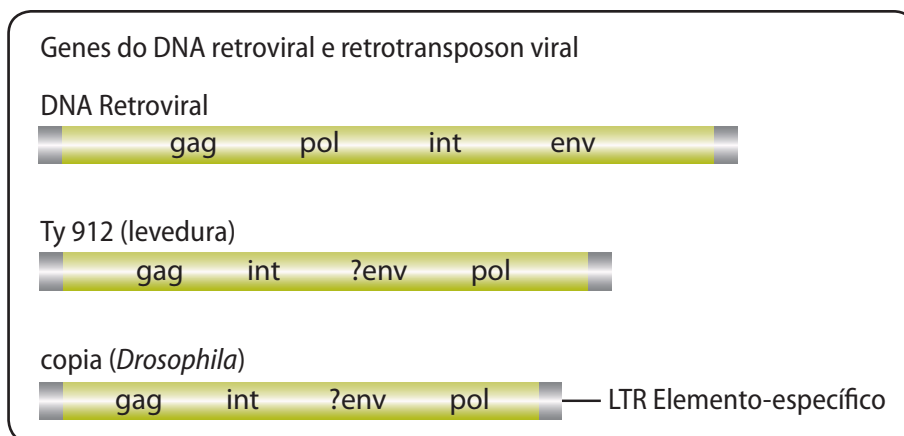





Figura 4.34 - Genes no DNA retroviral e retrotransposons virais

Tabela 4.2 - Principais tipos de elementos de transposição

Descrição de Classe e Estrutura	Gene	Modo de Movimentação	Exemplos
<b>DNA – Somente transposons</b>  Curtas repetições invertidas nas extremidades	Codificam transposases	Movem-se como DNA seguindo a via de replicação em cada fita	Elemento P (Drosófila) Ac-Ds (milho) Tn3 e IS1 ( <i>E.coli</i> ) Tam3 (boca de leão)
<b>Retrovírus – como retrotransposons</b>  Longas repetições terminais (LTRs) em seqüências repetidas nas extremidades	Codificam transcriptase reversa e se assemelham a retrovírus	Movem-se através de um RNA intermediário produzido por um promotor LTR	Copia (Drosófila) Ty1 (Fungo) THE-1 (Humano) Bs1 (milho)
<b>Retrotransposons não-retrovirais</b>  Poly A em 3' encerra o transcrito de RNA; a terminação 5' é geralmente truncada	Codificam transcriptase reversa	Movem-se através de um RNA intermediário que geralmente é produzido por um promotor próximo	Elemento F (Drosófila) L1 (Humano) Cin4 (milho)

Esses elementos apresentam seqüências de 1.000 a 12.000 pares de nucleotídeos; cada família contém muitos membros, com poucos deles exemplificados no quadro. Ainda com relação a transposons, existem determinados vírus capazes de se mover através de cromossomos celulares, sendo relatados nas duas primeiras classes de transposons.

## 4.8 Transformação (Exp. GRIFFITH)

A compreensão do fenômeno da transformação verificado por Griffith não envolve elementos de transposição, mas apenas a fita de DNA doadora, isto ocorre por um processo ativo com consumo de energia e depende de células competentes (fator de competência).

Inicialmente ocorre uma ligação reversível do DNA de dupla fita aos sítios receptores da membrana plasmática, com ligação irreversível do DNA doador, ocorrendo conversão do DNA doador

para DNA fita simples. A degradação de uma fita do DNA gera energia para impulsionar a outra fita do DNA pela membrana. No interior da célula ocorre a integração de todo ou parte do DNA fita simples no cromossomo recipiente. Para esse processo é necessária a homologia de aproximadamente 500 nucleotídeos entre as fitas, ocorrendo a inserção de até 20.000 nucleotídeos (1/200 do DNA genômico).

## Resumo

A plasticidade da molécula de DNA possibilita a variabilidade genética e a diversidade, que são fundamentais para o processo evolutivo.

O termo “mutação” refere-se tanto ao processo que gera alterações na seqüência do DNA, como as suas conseqüências fenotípicas.

As alterações podem envolver conjuntos cromossômicos (euploidias), cromossomos individuais (aneuploidias) ou alterações nas seqüências de nucleotídeos do DNA. As alterações cromossômicas têm impacto no funcionamento gênico, assim como alterações nas seqüências reguladoras.

As alterações nas seqüências codificantes podem acarretar mudanças na seqüência de aminoácidos dos peptídeos, enquanto as alterações nas seqüências sem atividade genética contribuem para a diversidade genética. É utilizada uma nomenclatura específica para designar as mutações.

Quando as mutações ocorrem em células germinativas, têm repercussão nas gerações subseqüentes. A diversidade gerada por essas alterações é fonte de variabilidade genética, sendo indispensável para a evolução dos seres vivos, possibilitando a adaptação a novos ambientes.

As alterações nas células somáticas se limitam ao indivíduo, estando associadas ao câncer e a doenças degenerativas.

As mutações podem ser decorrentes de processos celulares normais, como uma imprecisão na replicação, ou podem ser induzi-

das por uma gama de agentes físicos, químicos e biológicos. A exposição a esses agentes pode acarretar um aumento na frequência de mutações.

A maior parte dos agentes mutagênicos produz lesões características na molécula de DNA. Essas lesões são inadmissíveis, pois comprometem o funcionamento do DNA; as células dispõem de um elaborado aparato para a detecção e a correção dessas lesões. Os principais processos de reparo do DNA são a fotorreativação, o reparo por excisão de bases, o reparo por excisão de nucleotídeos e o reparo por recombinação. Esses processos ocorrem pela ação orquestrada de diversas proteínas que são conservadas evolutivamente. Algumas condições hereditárias raras envolvem a inativação desses processos por corrupção da informação do DNA, que codifica essas proteínas, acarretando hipermutabilidade das células. Quando o nível das lesões celulares está acima da capacidade do reparo, a célula aciona o processo de apoptose, que produz a fragmentação do DNA nos nucleossomos e a morte celular.

Um processo celular que contribui significativamente para a variabilidade genética é a recombinação do DNA, a recombinação homóloga ocorre durante a meiose envolvendo a quebra e a reunião das moléculas de DNA dos cromossomos homólogos. Alguns tipos de lesão de DNA, especialmente as quebras de cadeia dupla, são reparados por um processo análogo.

A recombinação sítio específica envolve segmentos de DNA com pequenas regiões de homologia. Esse processo é utilizado para a integração de genomas virais ao cromossomo bacteriano e também para a transposição. Esse fenômeno, postulado inicialmente por Barbara McClintock, envolve a mobilização intracelular de seqüências, os transposons, ou genes saltadores, utilizando uma variada gama de alternativas. A transposição está associada a efeitos fenotípicos diversos, desvios das proporções mendelianas e aparecimento de resistência a múltiplos antibióticos em bactérias. Estima-se que esses eventos sejam propulsores importantes da variabilidade genética e associados ao aumento do conteúdo do DNA.

## Referência Comentada

### Biologia molecular do gene

*James D. Watson, Michael Levine, Alexander Gann, Richard Losick, Tania A. Baker, Stephen P. Bell*

*Capítulos 9, 10 e 11:* Apresenta uma visão aprofundada e rica em detalhes, refletindo um conhecimento dinâmico numa linguagem acessível, com ilustrações que complementam o texto e facilitam a compreensão.

WATSON, J. D. et al. **Biologia molecular do gene**. 5. ed. Porto Alegre: Artmed, 2006.

## Referências Bibliográficas

ALBERTS, B. et al. **Molecular Biology of the Cell**. 4<sup>th</sup>. ed. New York: Garland Pub. Inc., 2002.

GRIFFITHS, A. J. F. et al. **Introdução à Genética**. 8. ed. Rio de Janeiro: Guanabara-Koogan, 2006.

LEWIN, B. **Genes VII**. Porto Alegre: Editora Artmed, 2001.

PURVES, W. K. et al. **Vida: a ciência da Biologia. Célula e hereditariedade**. 6. ed. Porto Alegre: Artmed, 2002. v. I.

STRACHAN, T.; READ, A. P. **Genética molecular humana**. 2. ed. Porto Alegre: Artmed, 2002.