

INTRODUÇÃO:

BIOLOGIA MOLECULAR
visão geral

2017

Moléculas da hereditariedade

PROTEÍNAS

centenas de diferentes polímeros formados pelos mesmos 20 aminoácidos, com seqüências distintas.

Como a célula obtém a informação: qual seqüência proteica sintetizar?



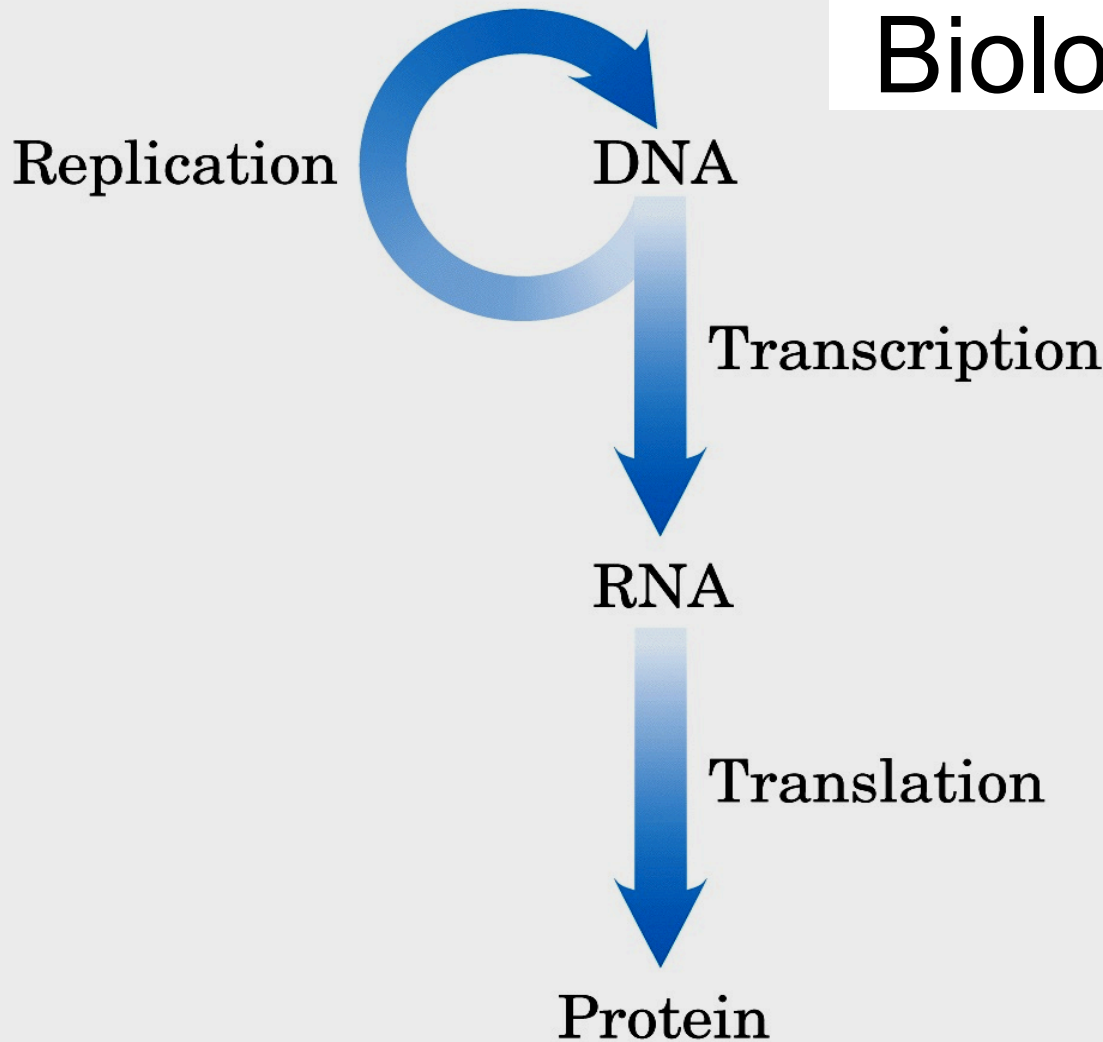
CROMOSSOMOS (genes)

Ácidos nucleicos + proteínas



DNA (O. Avery)
Um gene — uma proteína
(G. Beadley, E. Tatum)
(década de 1940)

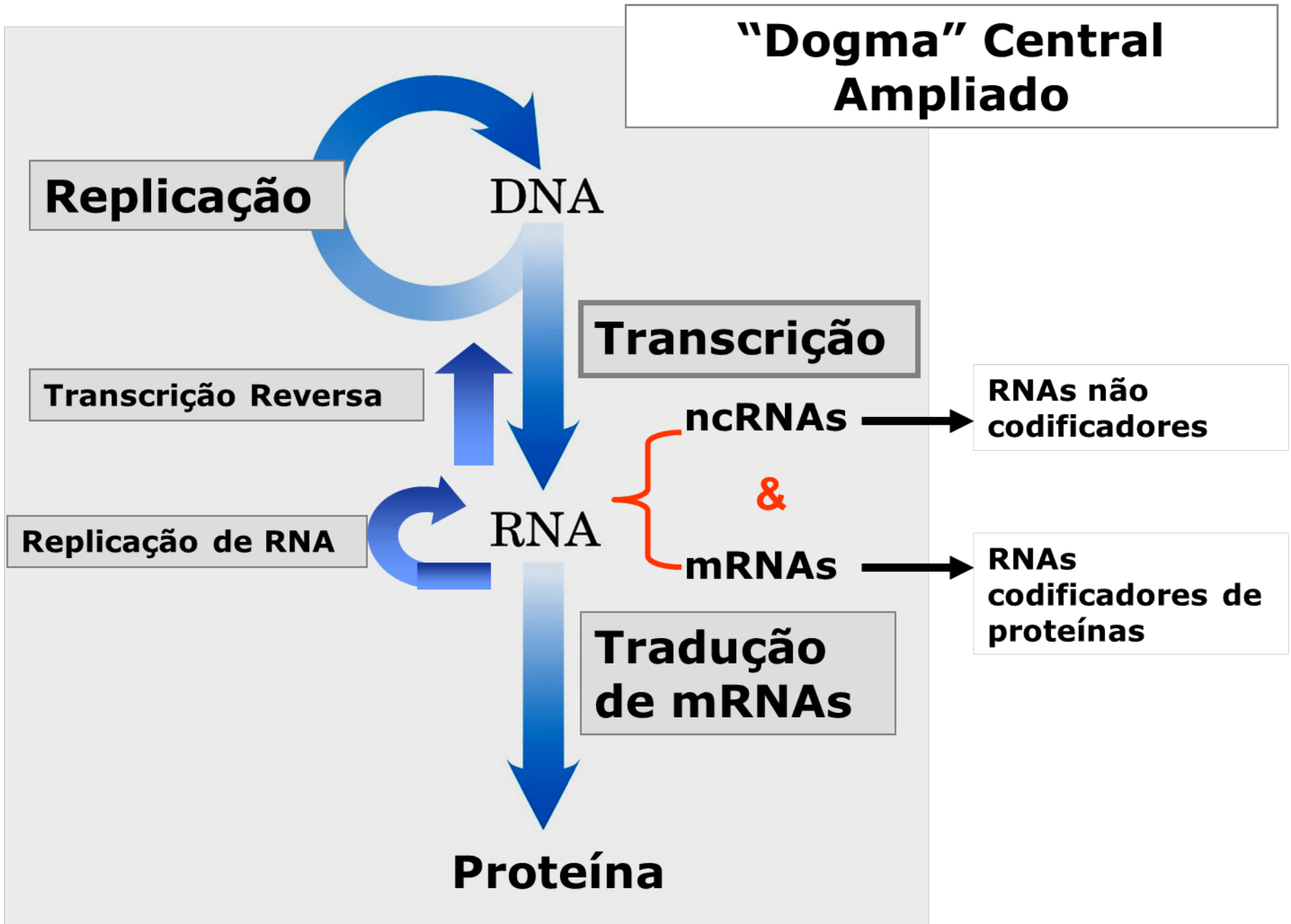
Dogma Central da Biologia Molecular



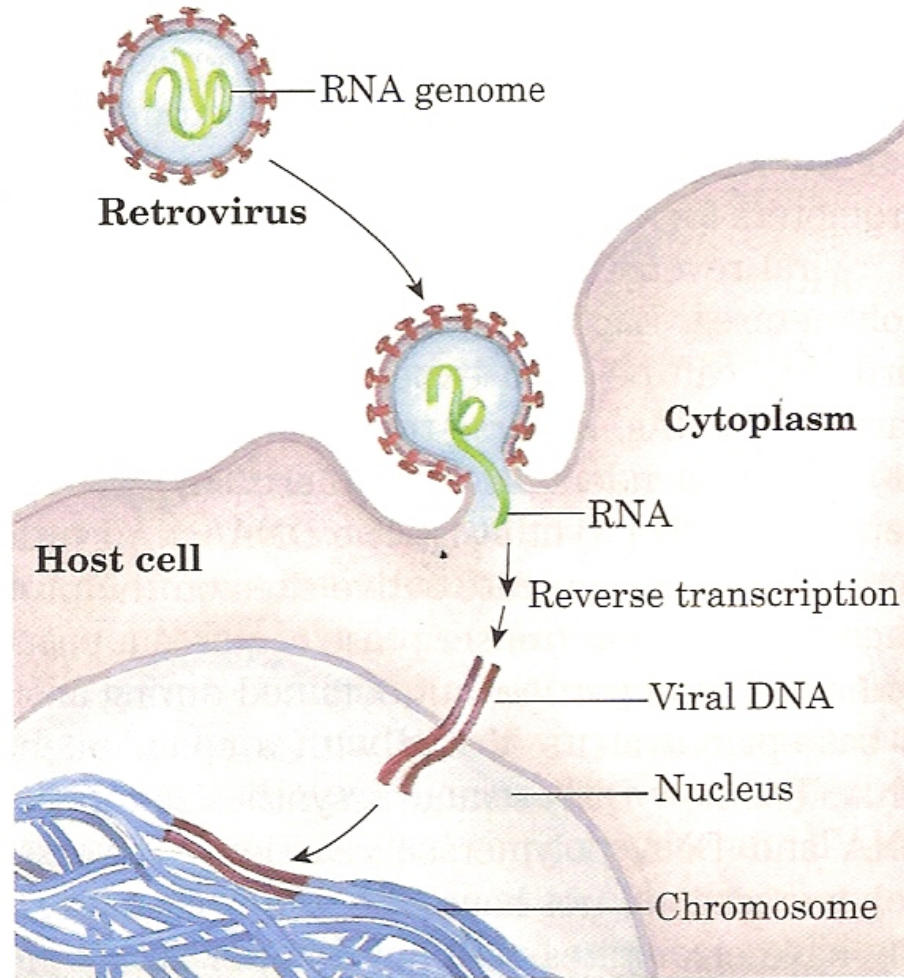
Como a informação do DNA é transferida para a célula ou organismo?

Francis Crick, 1956

"Dogma" Central Ampliado



Retrovirus



Transcriptase reversa: Temin 1962, 1970; Baltimore, 1970

Processo para síntese de novas cadeia de DNA, catalisado pelas DNAs polimerases

Replicação

DNA

Transcrição Reversa

Transcrição

Replicação de RNA

RNA

ncRNAs

&

mRNAs

Tradução de mRNAs

Proteína

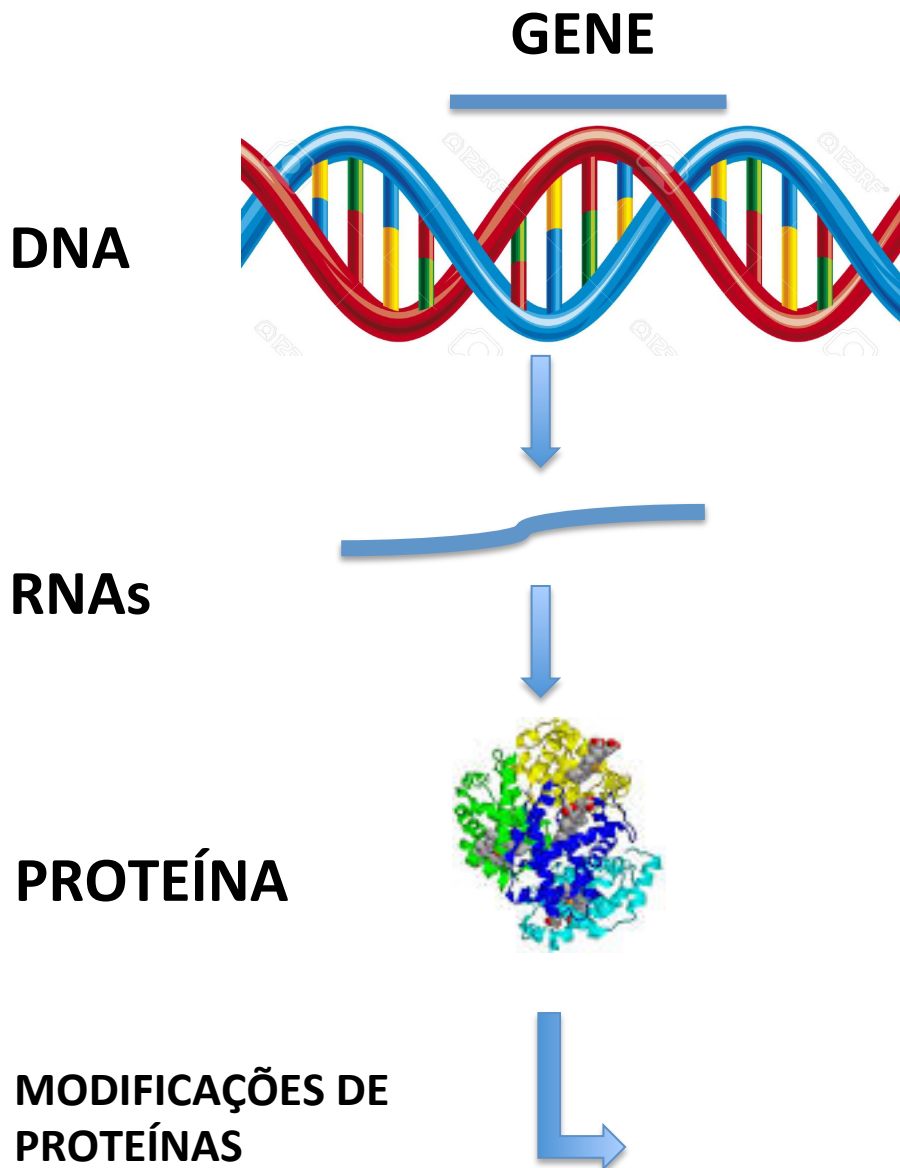
Processo para síntese de todos os RNAs da célula

Catalisado pelas RNAs polimerases

O conjunto de genes expressos é variável

Processo para síntese das proteínas da célula

Ocorre nos ribossomos



NATUREZA DO POLÍMERO

SEQUÊNCIA DE NUCLEOTÍDIOS

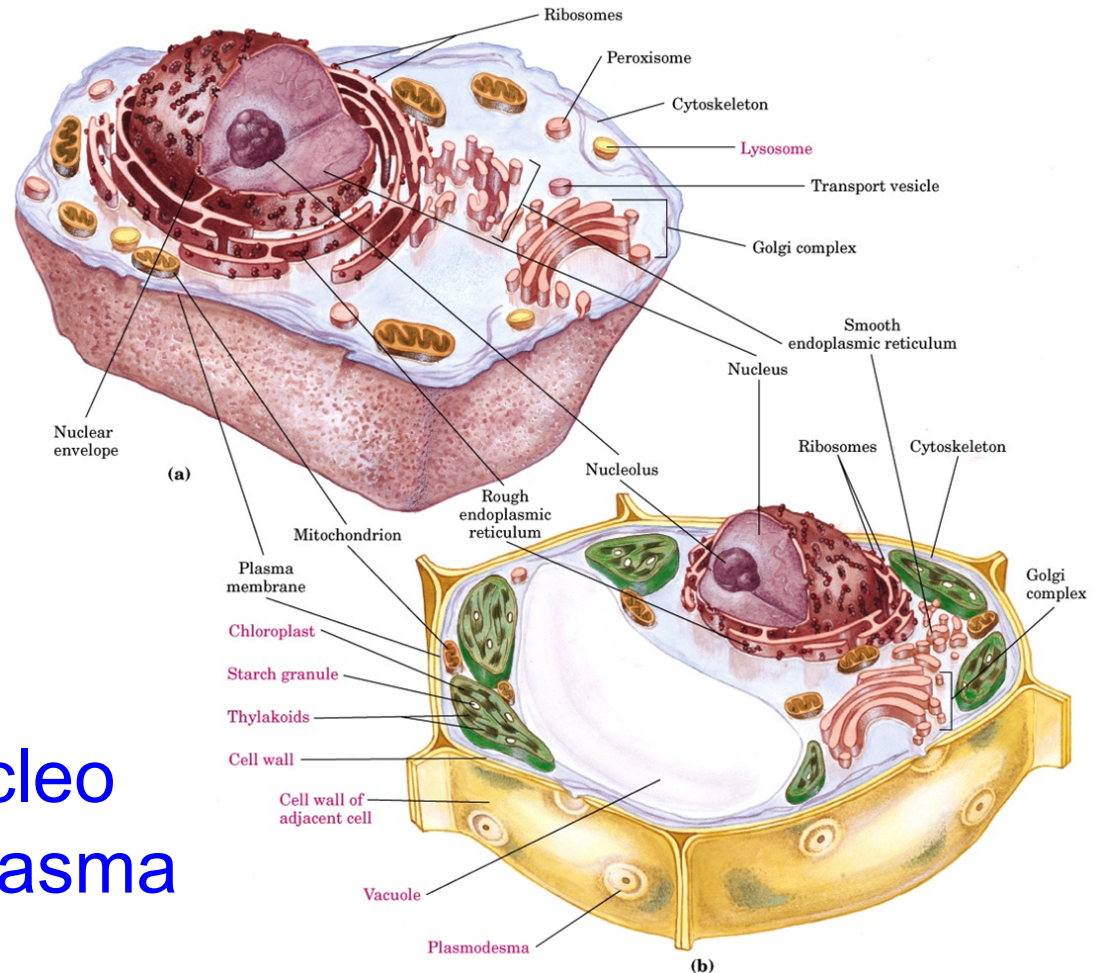
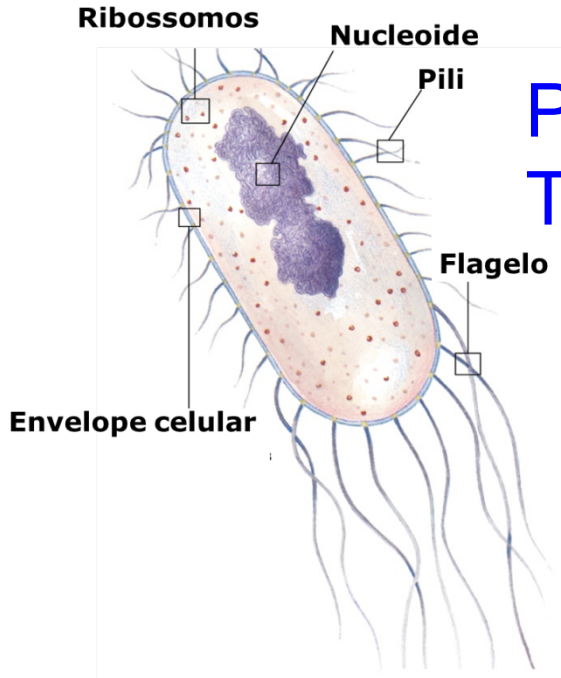
SEQUÊNCIA DE NUCLEOTÍDIOS

Adaptador
de linguagem:
tRNA

SEQUÊNCIA DE AMINOÁCIDOS

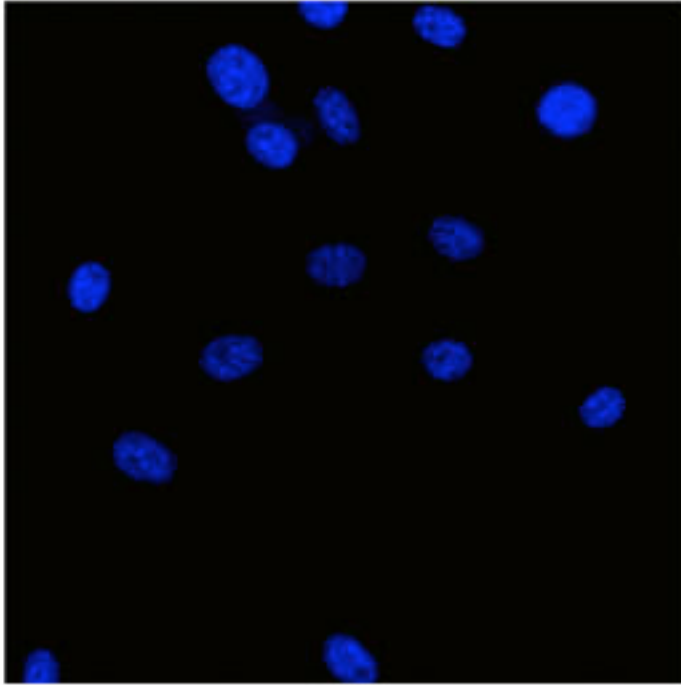
FOSFATO, AÇÚCAR, ACETILA,
ÁCIDOS GRAXOS, FERRO, ETC

Procariotos: Transcrição e tradução acopladas

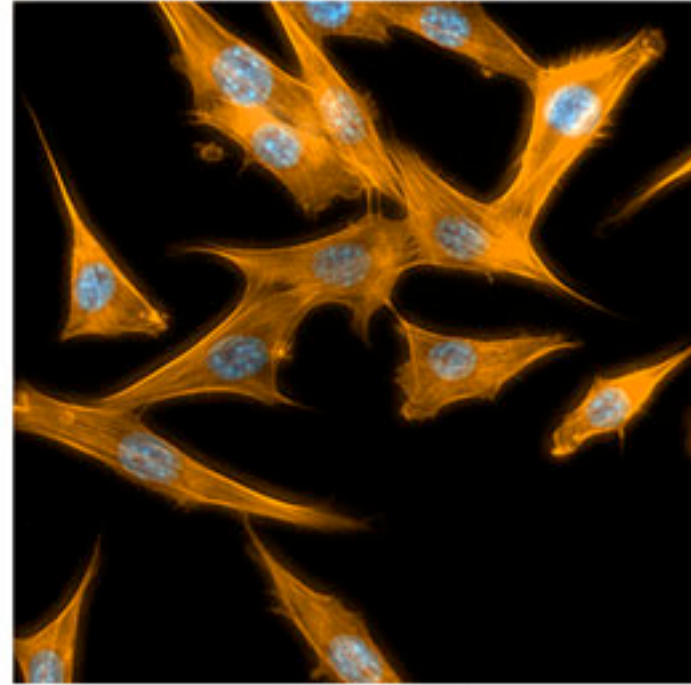


Eucariotos:
Transcrição no núcleo
Tradução no citoplasma

Visualização do DNA na célula- exemplo



DAPI

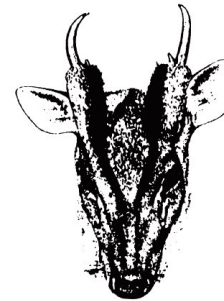
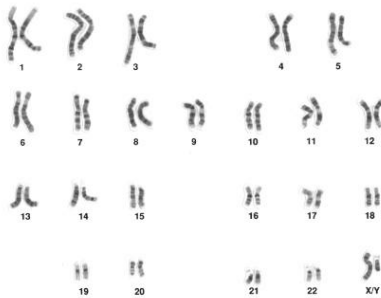


DAPI + Phalloidin

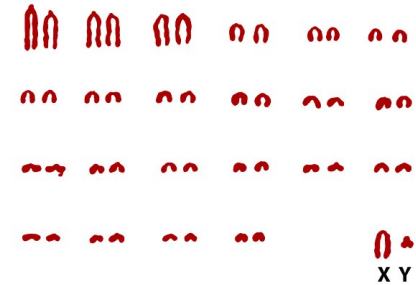
Moléculas da hereditariedade

- Número de cromossomos varia entre as espécies, assim como o formato.

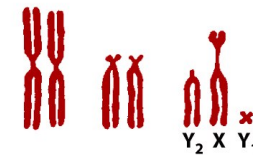
Humanos: 46 cromossomos em células somáticas



Chinese muntjac



Indian muntjac



Gene

- **Gene**- sequência de DNA que inclui todas as sequências de nucleotídeos que levam à síntese de um produto (mRNA para a síntese de proteína ou outros tipos de RNAs). Inclue as regiões codificantes e as reguladoras.
- Genes estão orientados de 5' para 3' (fita codificadora), cuja sequência estará representada no mRNA. A outra fita de DNA, complementar à codificadora é a fita molde para a síntese do RNA.
- O RNA portanto tem a mesma orientação e sequência da fita de DNA codificadora

Gene

- Sequências codificantes podem estar interrompidas por sequências sem função de codificadora.
- **Exons**- codificantes- **expressed sequences**
- **Introns**- não codificantes- **intervening sequences**.
- Alguns organismos não apresentam introns (incluindo alguns eucariotos)
- Genes podem ter cópia simples ou múltiplas no genoma

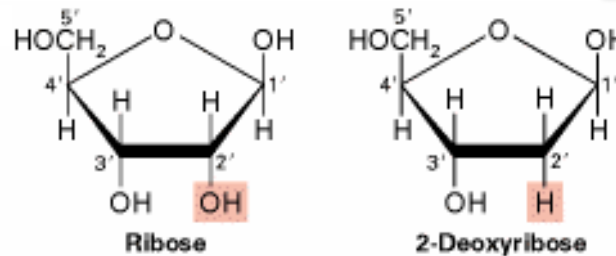
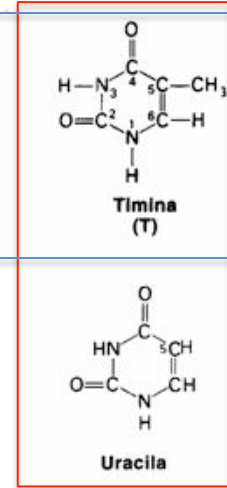
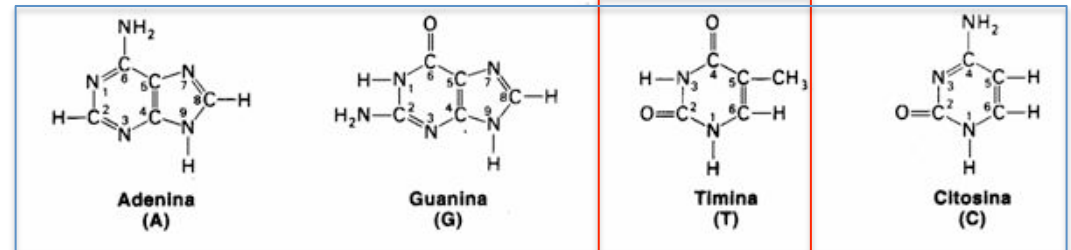
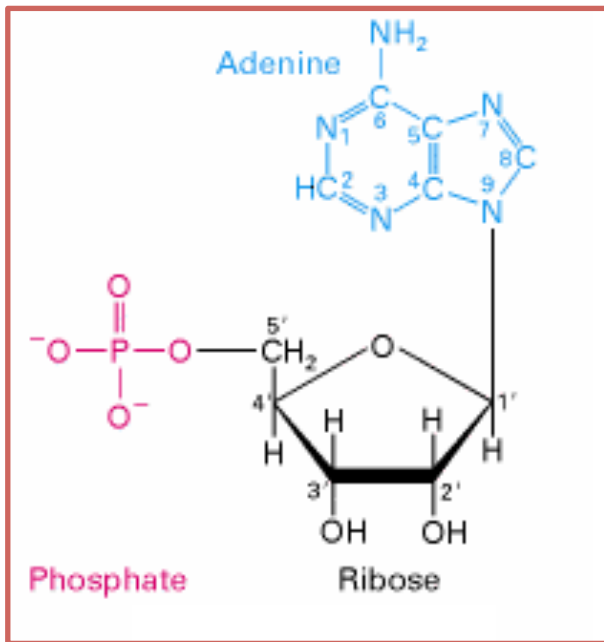
Genomes vary greatly in size

Genome	Gene Number	Base Pairs
Organisms		
Plants	<50,000	<10 ¹¹
Mammals	30,000	~3 x 10 ⁹
Worms	14,000	~10 ⁸
Flies	12,000	1.6 x 10 ⁸
Fungi	6,000	1.3 x 10 ⁷
Bacteria	2-4,000	<10 ⁷
Mycoplasma	500	<10 ⁶
dsDNA Viruses		
Vaccinia	<300	187,000
Papova (SV40)	~6	5,226
Phage T4	~200	165,000
ssDNA Viruses		
Parvovirus	5	5,000
Phage φX174	11	5,387
dsRNA Viruses		
Reovirus	22	23,000
ssRNA Viruses		
Coronavirus	7	20,000
Influenza	12	13,500
TMV	4	6,400
Phage MS2	4	3,569
STNV	1	1,300

Componentes de ácidos nucleicos (DNA e RNA)

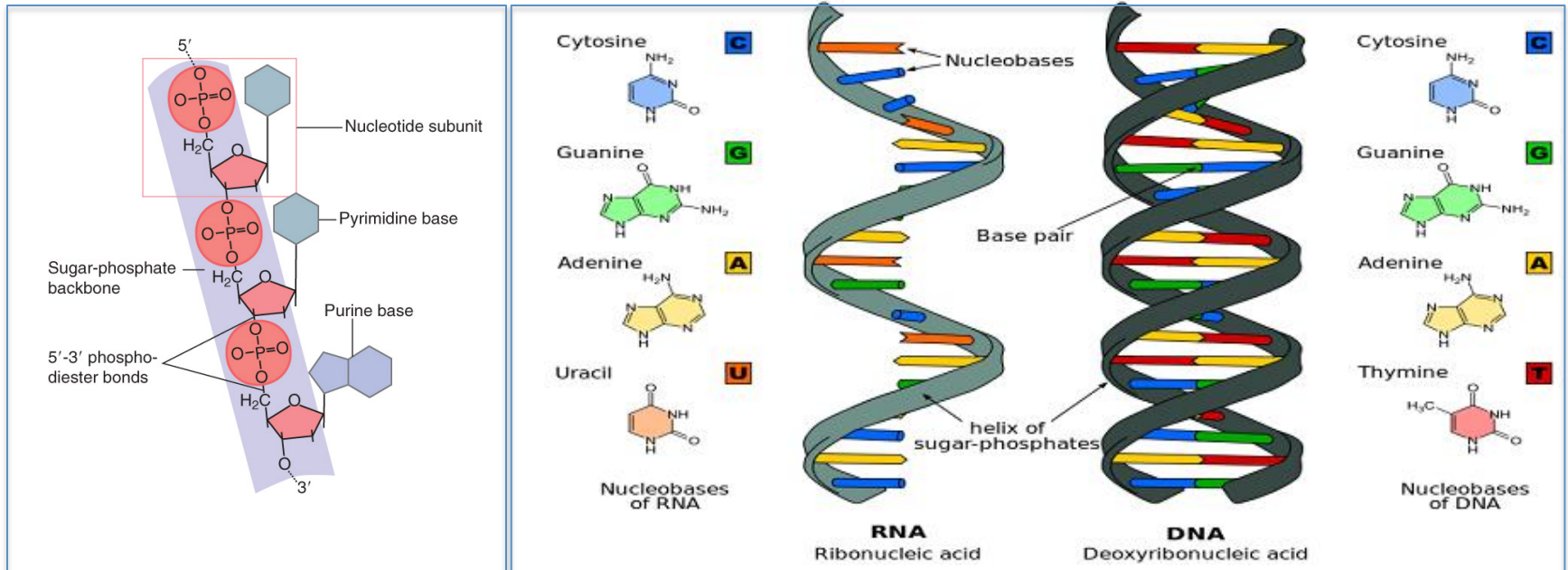
purinas

pirimidinas

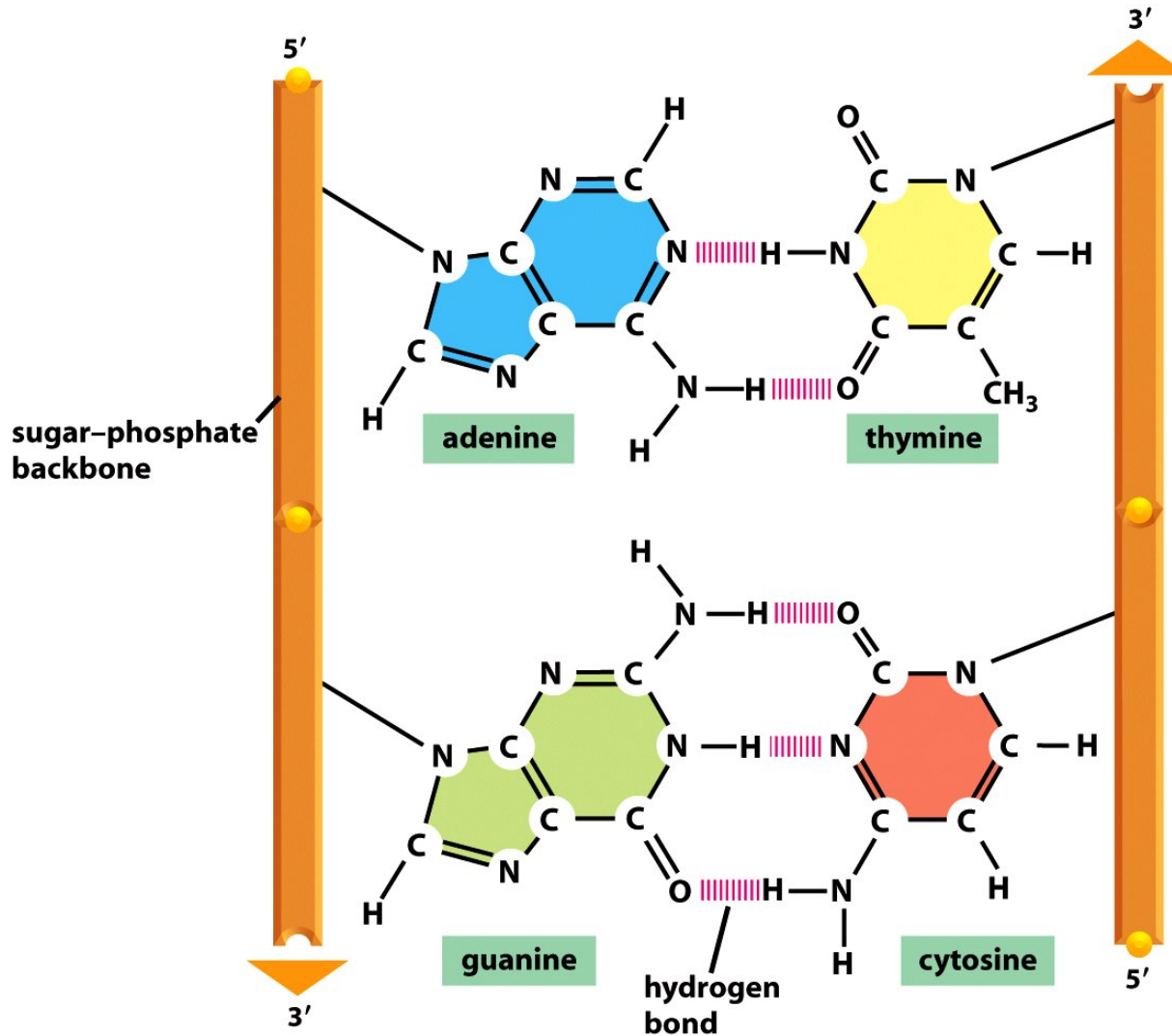


nucleotídio

Estructura de ácidos nucleicos



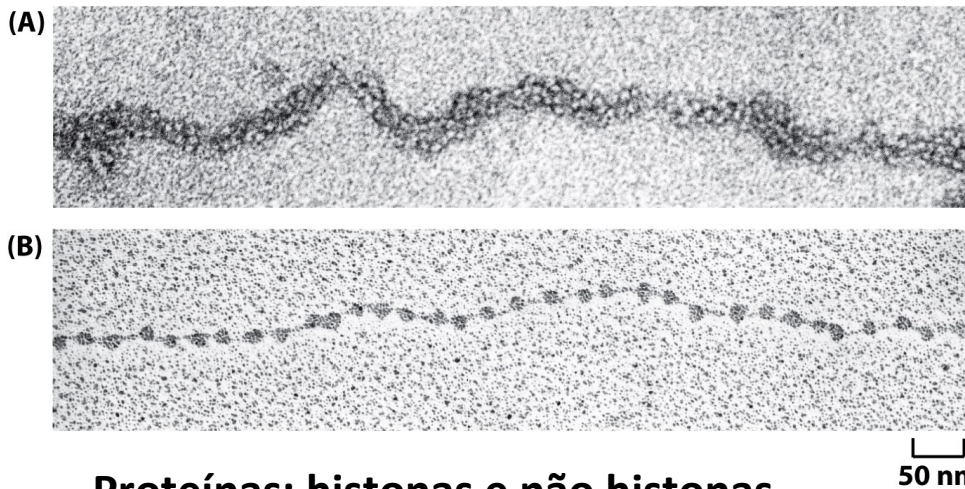
DNA (dupla fita)



Nota:
Princípio da interação DNA-RNA-
Pareamento de bases

Moléculas da hereditariedade

- **Cromossomos**- composição determinada- ácido nucleico e proteínas (histonas e não histonas). Histonas-proteínas básicas, abundantes



Proteínas: histonas e não histonas
(~60 milhões de moléculas de cada tipo de histona/núcleo)

DNA humano “linearizado”= 1m

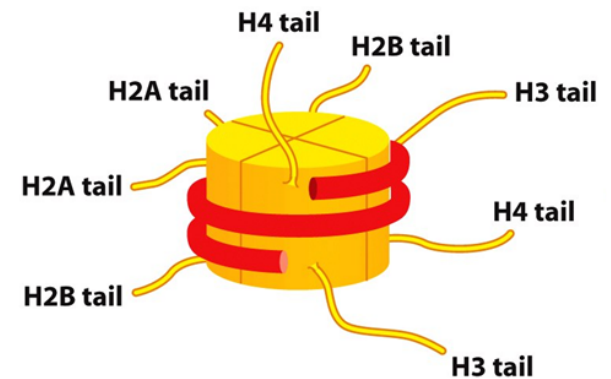


Figure 4-33a Molecular Biology of the Cell 5/e (© Garland Science 2008)

Moléculas da hereditariedade

- **Maioria dos cromossomos de bactéria-circulares. Lineares são quase exceções.**

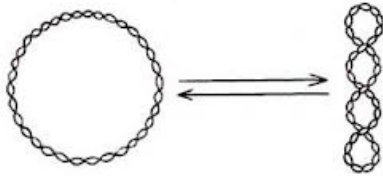
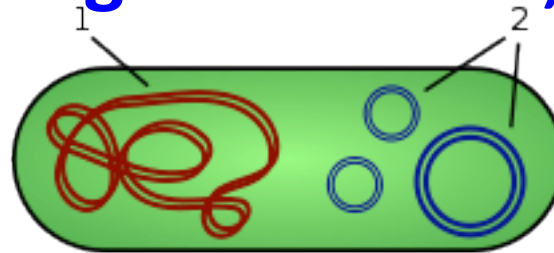
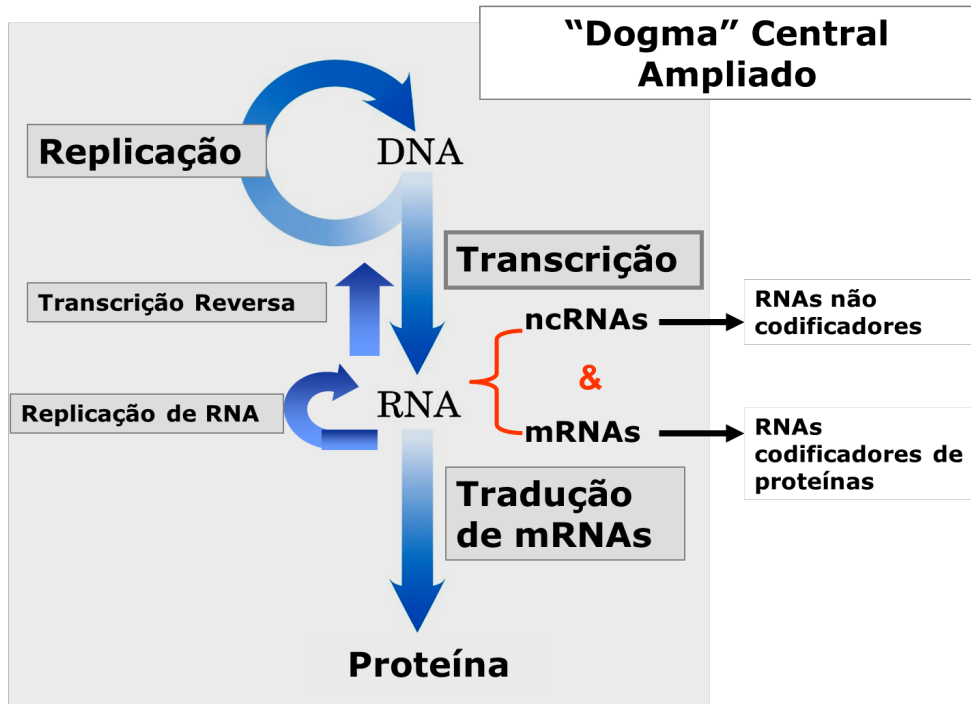


Fig. 3.9. Super coiling

- **Procariotos podem conter plasmídios (DNA extracromossomal, em geral circular; variam de 1 a inúmeros).**



- **Eucariotos- plasmídios em levedo**



Síntese de DNA

A replicação está baseada no pareamento das bases da dupla hélice do DNA.

A estrutura do DNA contém a informação necessária para perpetuar sua sequência de bases

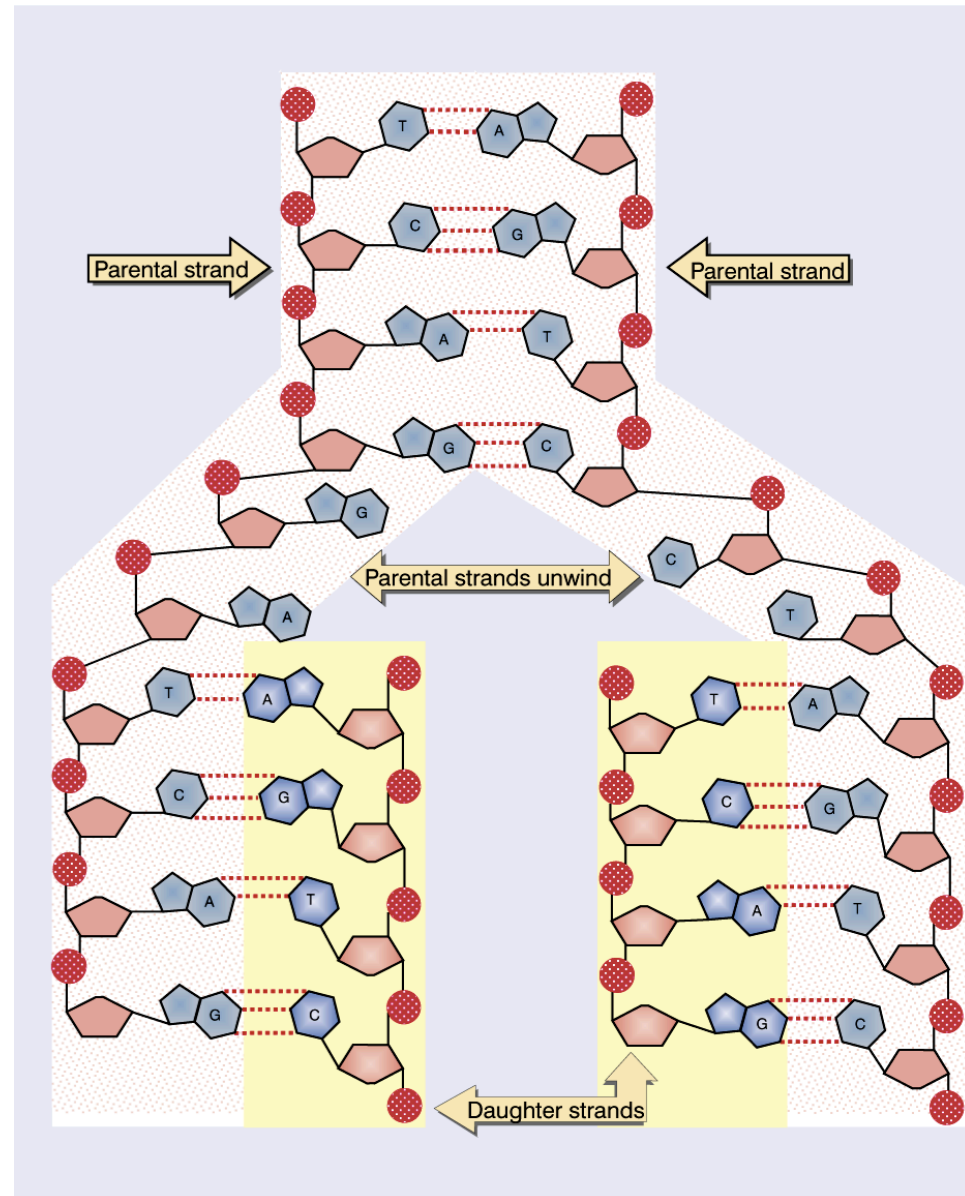
Humanos- cerca de 3 bilhões de pares de bases/ célula a serem duplicadas em cada ciclo celular.

Humanos- ~ 1 trilhão de células

James D. Watson e Francis Crick

Estrutura do DNA – 1953

Replicação semiconservativa - 1958



Etapas da replicação

(uma vez por ciclo celular)

- **Iniciação** (origem de replicação- replicon-1 na maioria dos procariotos; até 3 em Archaea; ~400 em mamíferos). Bactérias: origem DNA metilado.
- **Elongação**- forquilha de replicação. As duas fitas são sintetizadas ao mesmo tempo. A velocidade da forquilha de replicação bacteriana é 500.000pb/min; (~4.600.000pb) e 2.000 pb/min em mamíferos
- **Término**

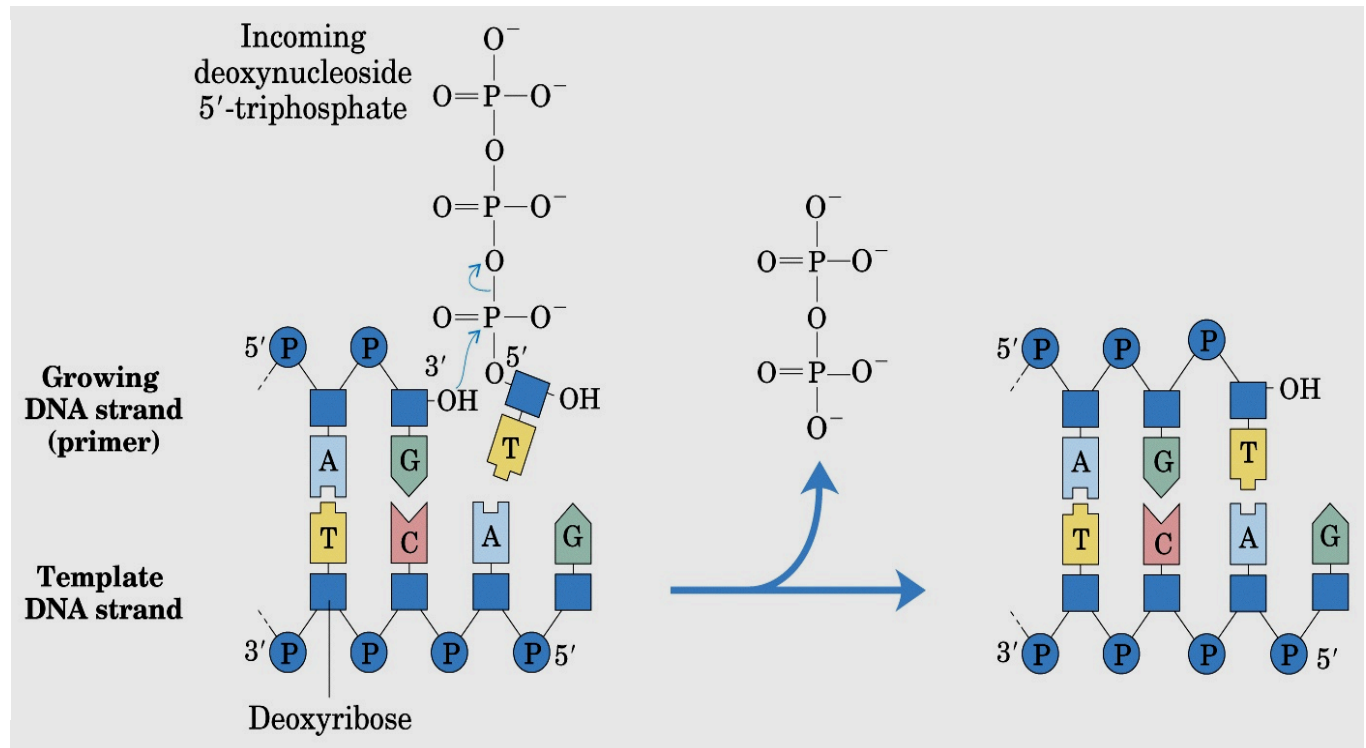
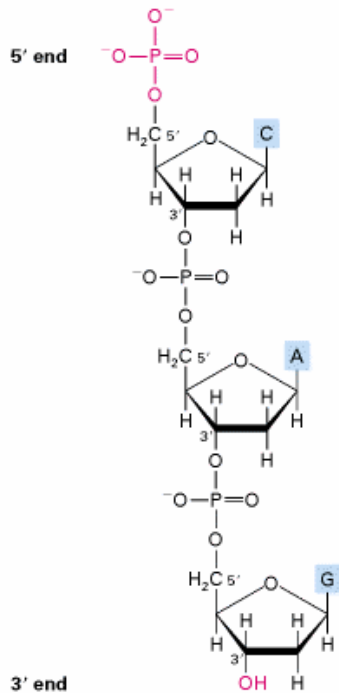
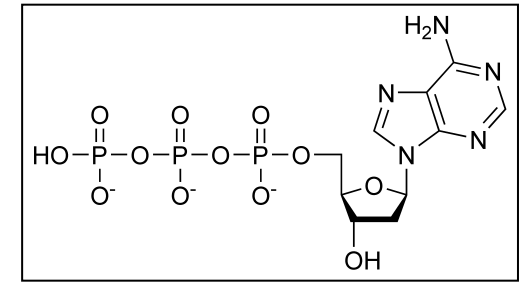
Enzimas: DNA polimerases

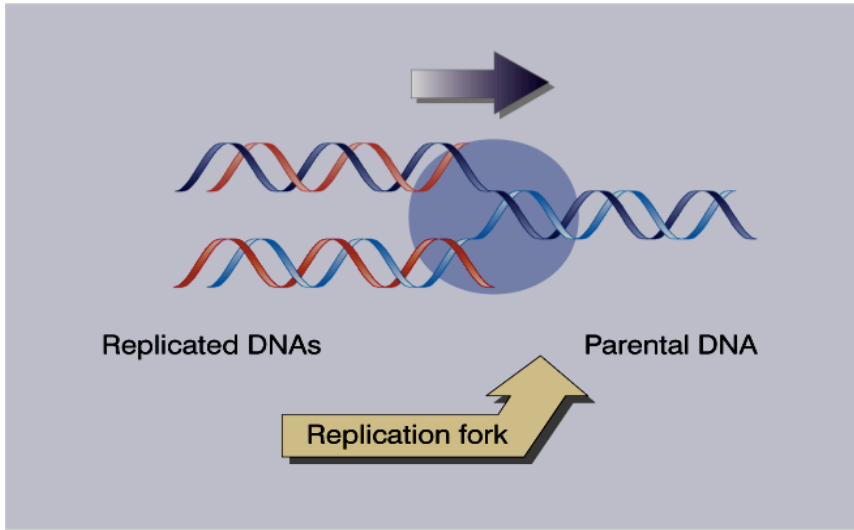
Substrato da DNA polimerase:

desoxiribonucleotídios
(dATP, dCTP, dTTP, dGTP)

Fita molde

Cadeia pré-existente (*primer*)- RNA primase
(ATP, CTP, UTP, GTP). RNA removido

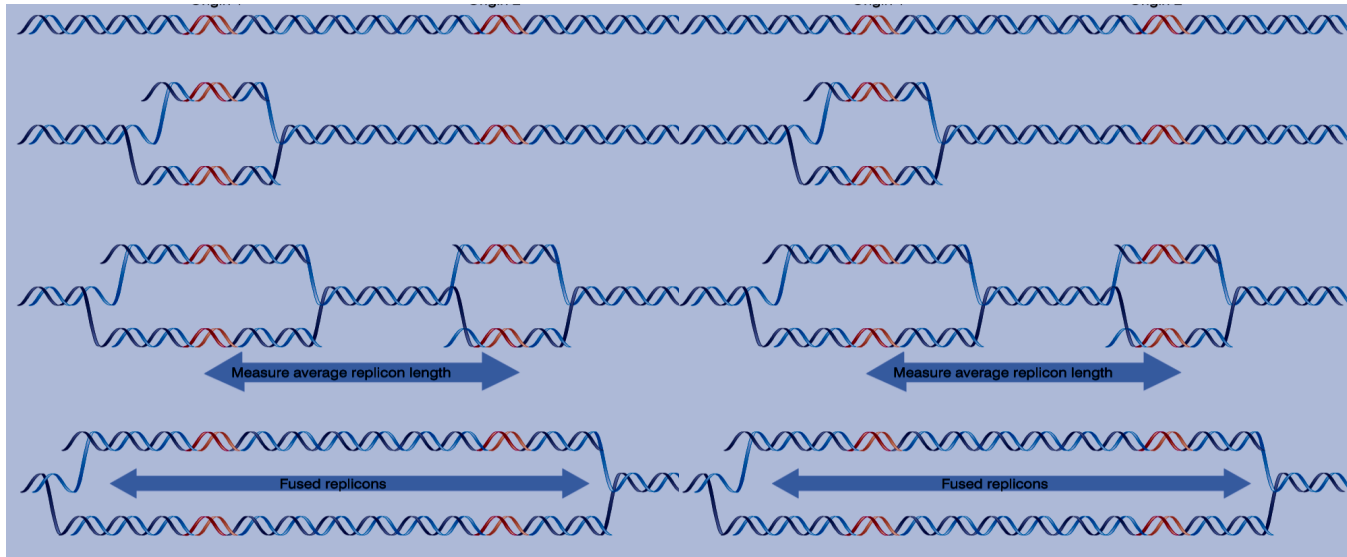




Forquilha de replicação

- Replicação é complexa-proteínas se arranjam como fábricas de síntese (replissomos).
- Replissomo- polimerases, helicases, primases, etc
- Bactérias- replissomos ligados à membrana
- A replicação é um processo extremamente fiel. As DNA-polimerases têm atividade revisora.
- 1 erro a cada 10^9 - 10^{10} em *E. coli*
- Mutações

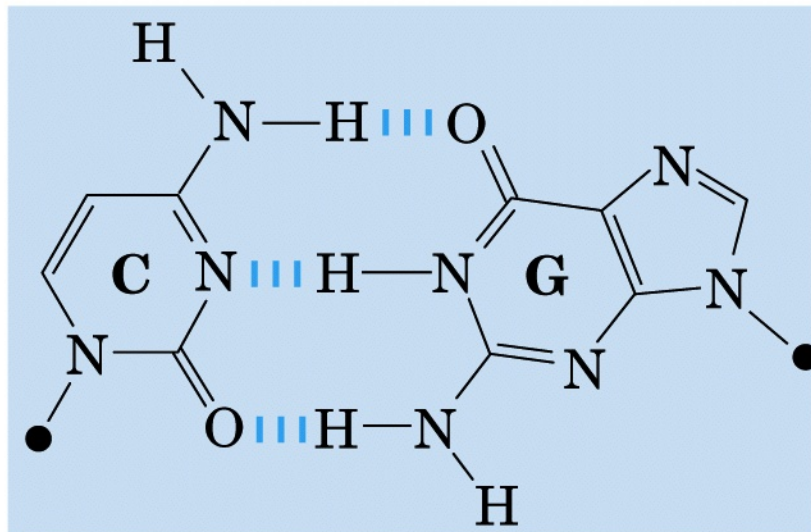
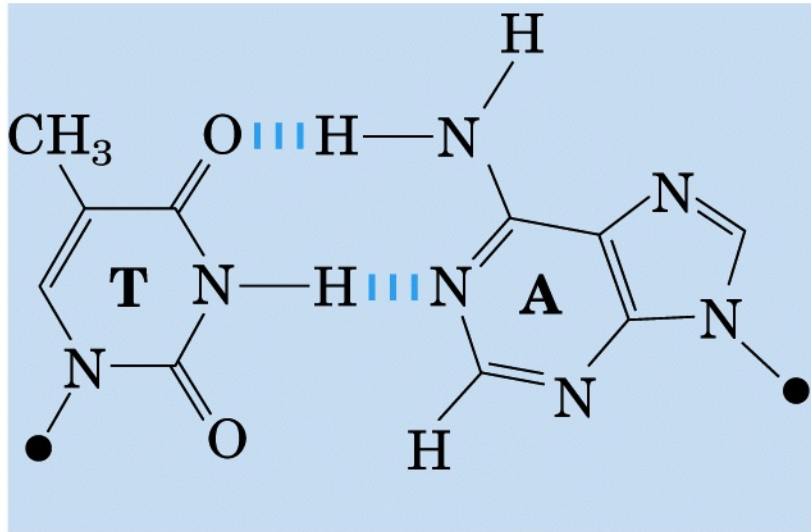
Genoma eucariótico: vários replicons



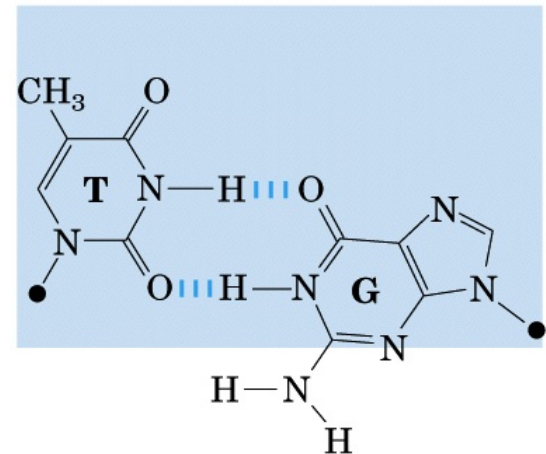
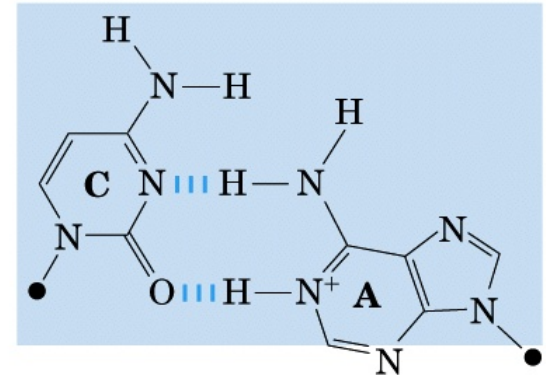
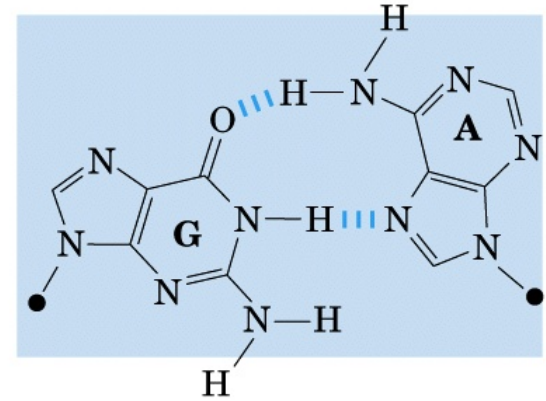
- A velocidade da forquilha de replicação eucariótica é 2.000pb/min (~1/20 de E. coli)
- Os replicons eucarióticos tem 40-100 kb e são iniciados em tempos diferentes. ~400 replicons; 30.000-300.000 pb afastados.
- se fosse 1 replicon- 500 hs para duplicar o DNA
- Fase S demora ~ 6hrs em uma célula somática

Pareamento de bases

correto



incorreto



Células eucarióticas tem muito mais DNA que células procarióticas

Drosófila – 25 x mais que *E.coli*

Célula humana – 600 x mais que *E.coli*

Genoma eucariótico tem alta proporção de DNA não codificador

~2500 genes/mm DNA de *E.coli*

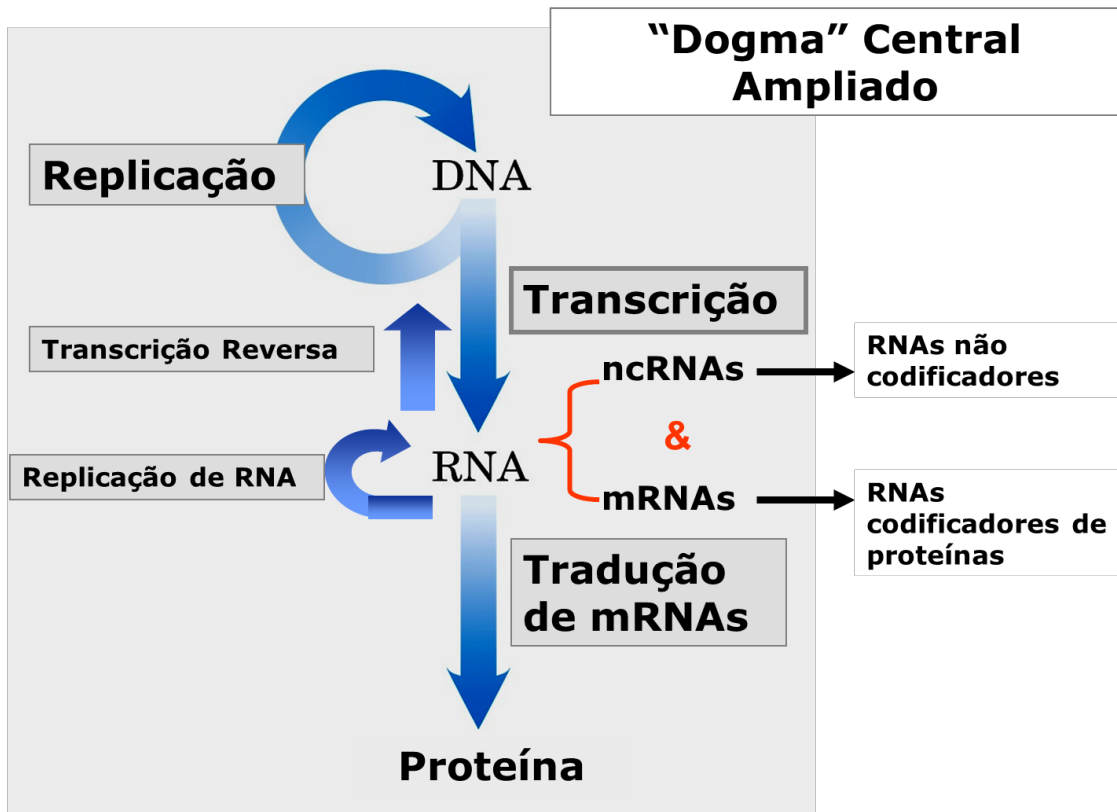
~50 genes/mm DNA humano

Introns

Sequências intergênicas

DNA repetitivo

- **Somente uma fração dos genes é expressa em um determinado tempo**
- **Alguns produtos gênicos são mais (fatores de alongação de proteínas em bactéria) ou menos (enzimas de reparo) abundantes**
- **Necessidade de determinados produtos varia com a condição fisiológica da célula**
- **Algumas proteínas são órgão específicas**
- **Genes constantemente expressos- constitutivos (housekeeping genes- enzimas do metabolismo central)**
- **Genes que respondem a sinais moleculares- indução ou repressão**
- **Regulação- feita por interação de ácidos nucleicos- proteína**



Síntese de RNA

Principais Tipos de RNA

RNA mensageiro (mRNA): contém a informação genética para a sequência de aminoácidos das proteínas

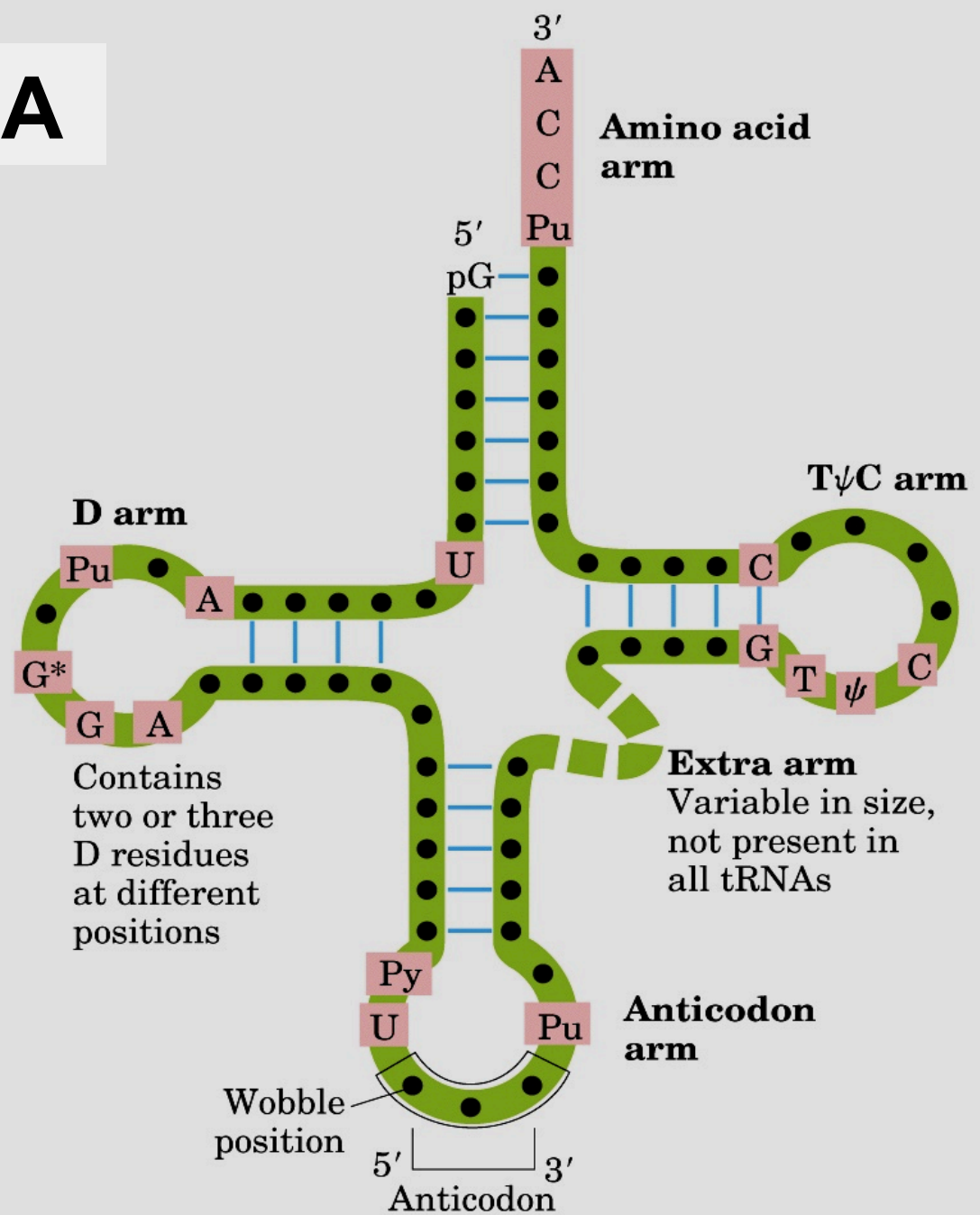
RNA transportador (tRNA): identifica e transporta os aminoácidos até o ribossomo

RNA ribossômico (rRNA): constituinte dos ribossomos

tRNA

-Estrutura secundária com grampos e alças formando um trevo

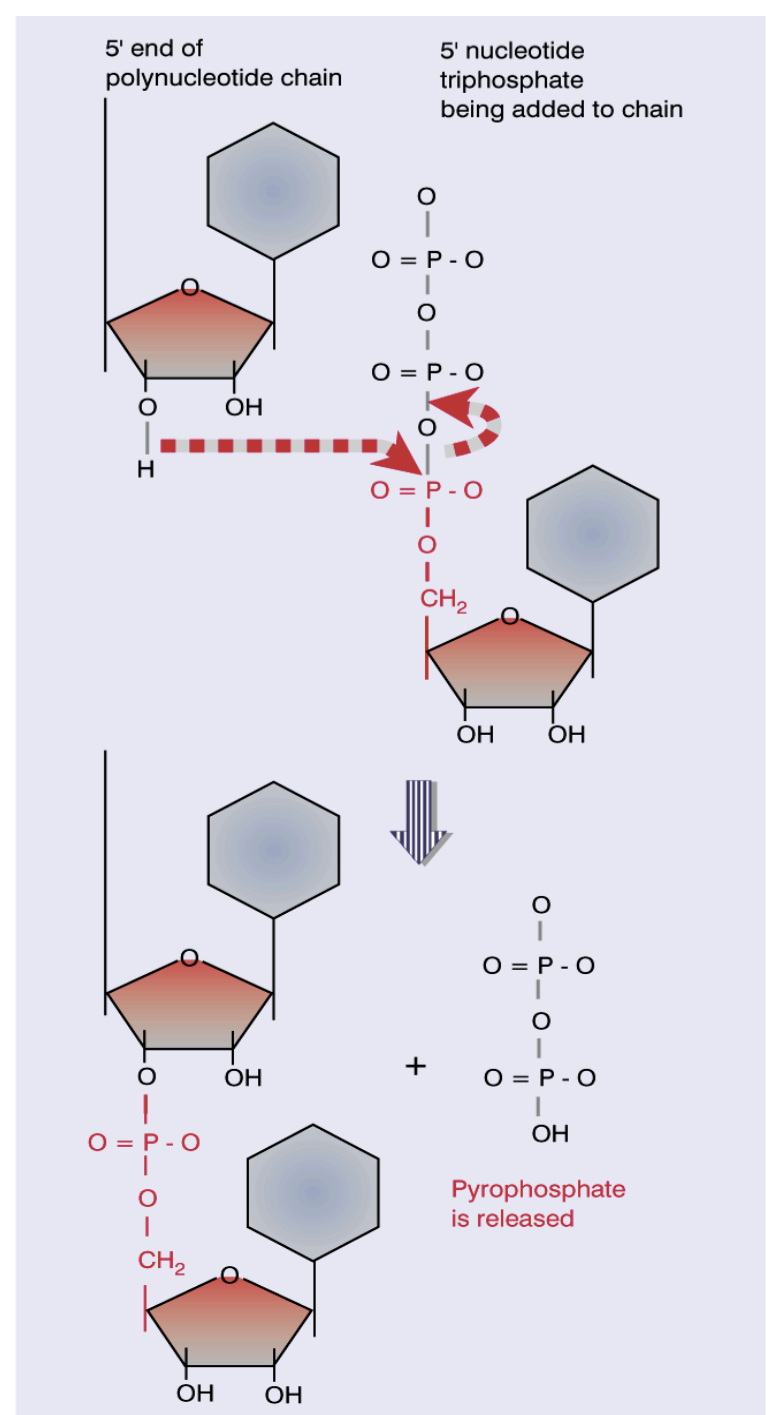
-Alto número de bases modificadas depois da sua transcrição



Síntese de RNA: formação da ligação fosfodiéster envolve o ataque nucleofílico do 3' OH da cadeia crescente no fosfato do carbono 5' do ribonucleosídeo trifosfatado que será incorporado

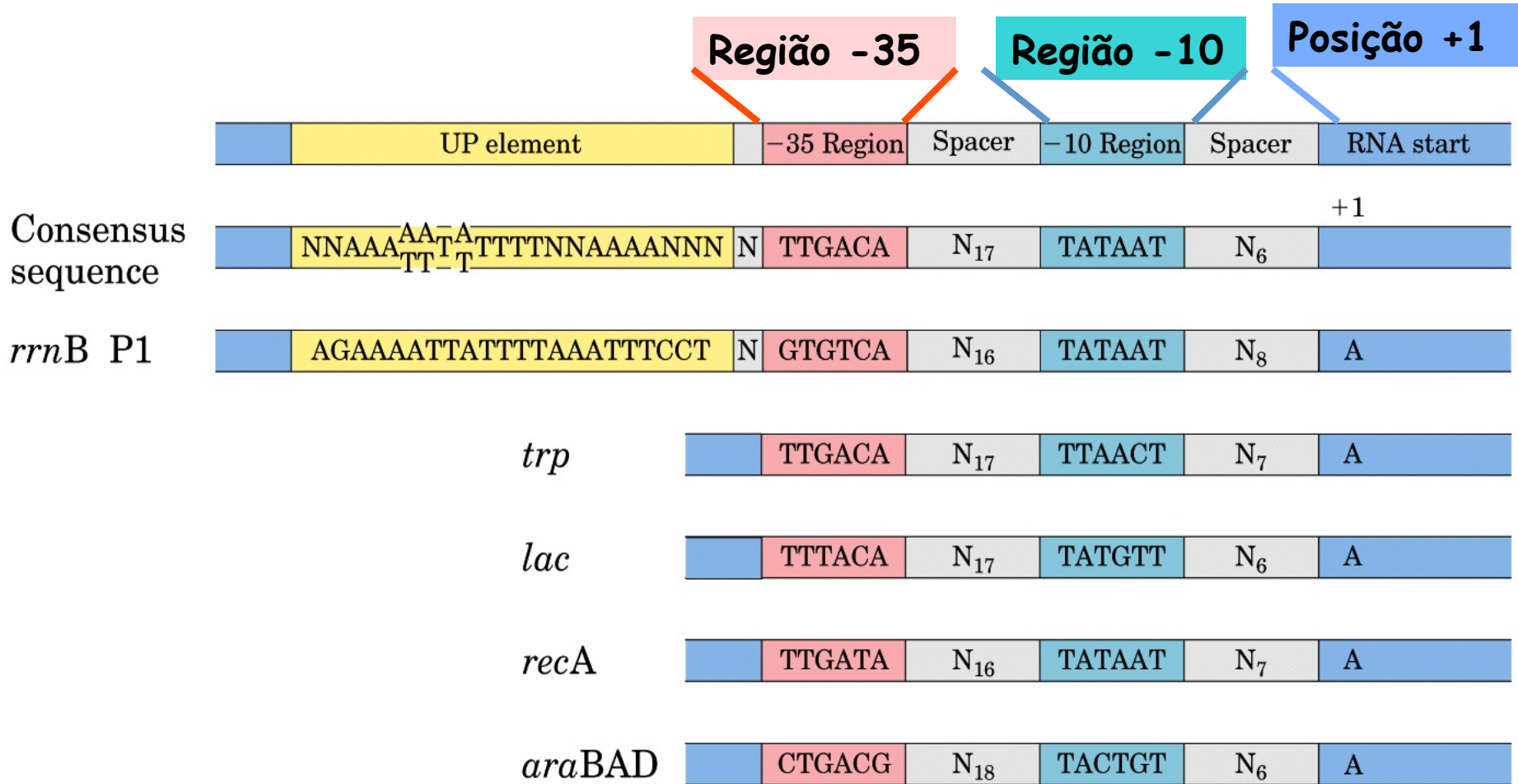
A RNA polimerase não requer um iniciador (primer) para iniciar a síntese do RNA

A RNA polimerase não tem mecanismo de reparo

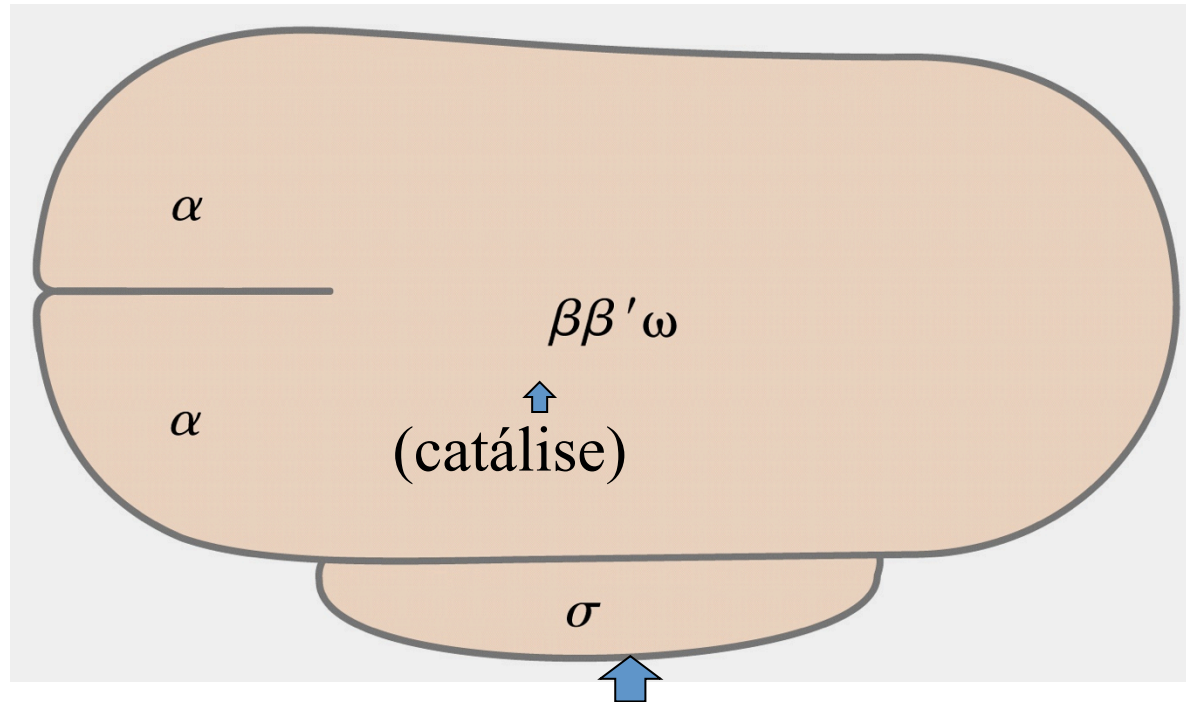




Promotores de genes bacterianos



RNA polimerase de *E.coli*

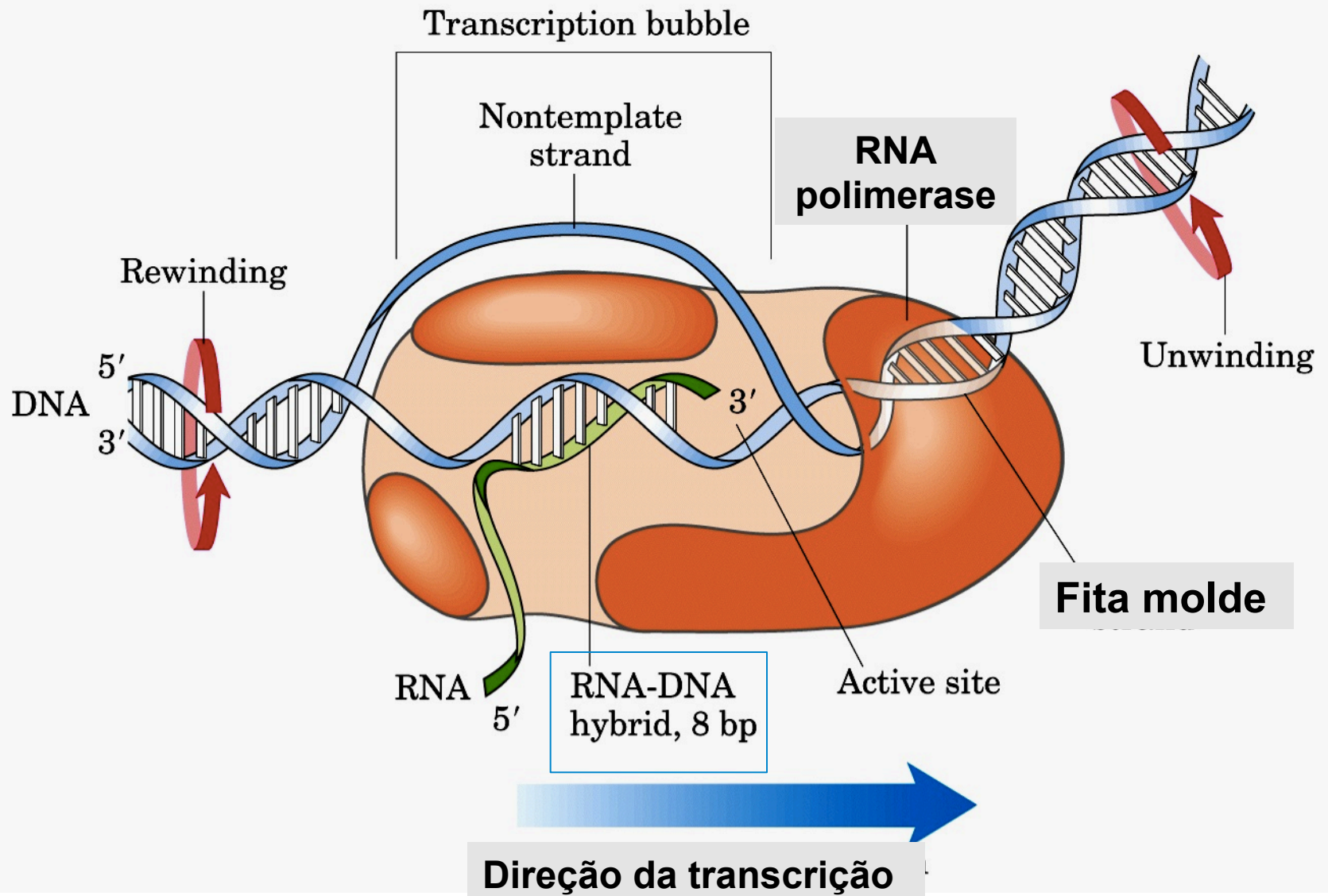


50-90 nucleotídios/seg
1 erro a cada $10^4 - 10^5$

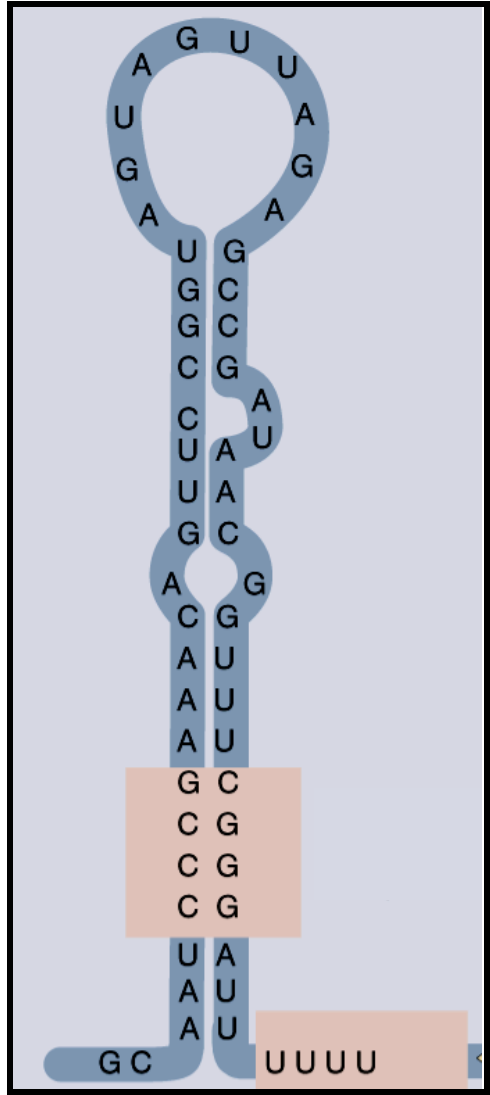
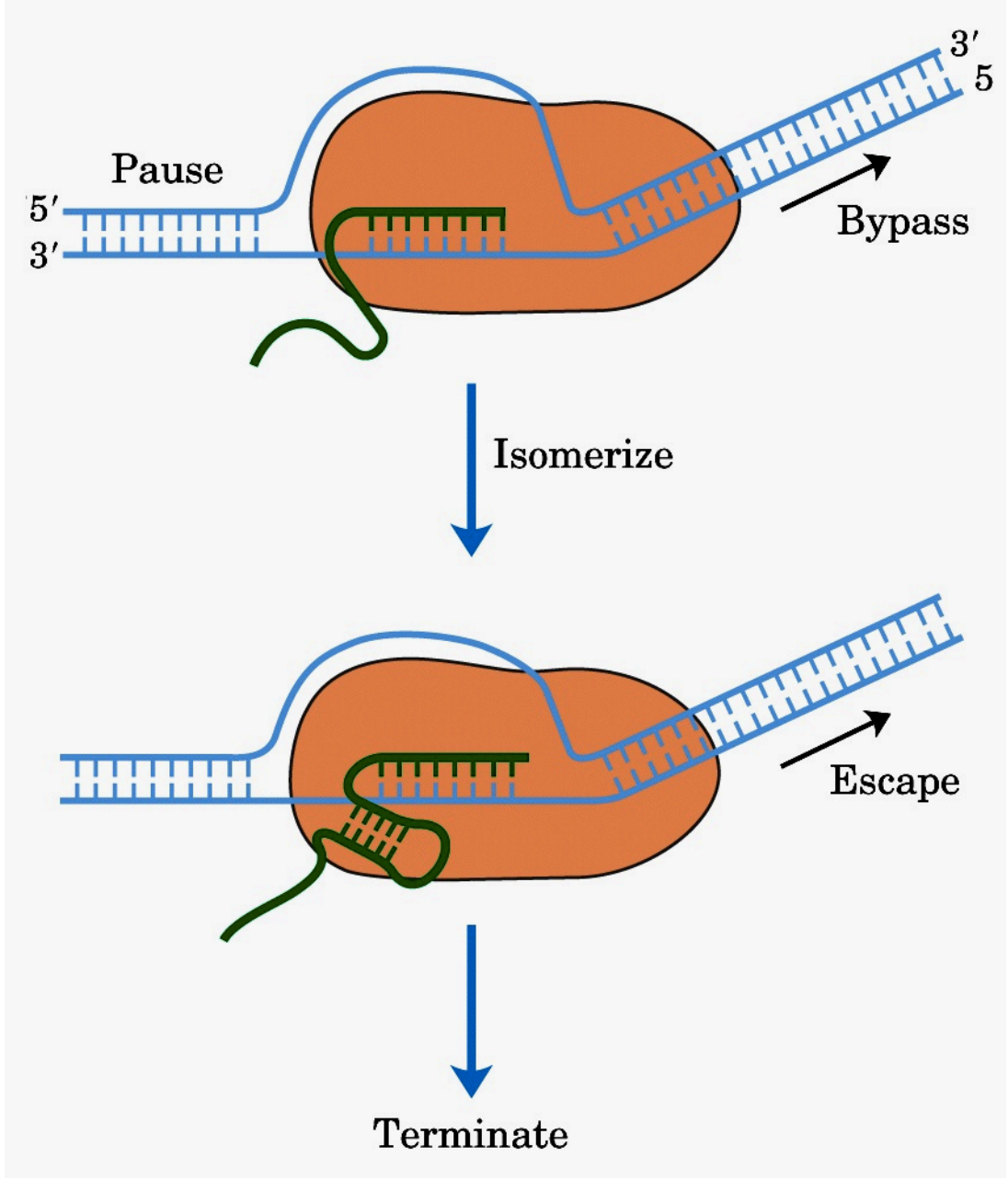
Reconhecimento de promotor
Diferentes Sigmas
Ausente no alongamento

Duplicação- 1 erro a cada $10^9 - 10^{10}$

Bolha- ~17 pares de bases



Terminação da transcrição através da formação do grampo de terminação



Terminação da transcrição dependente da proteína Rho

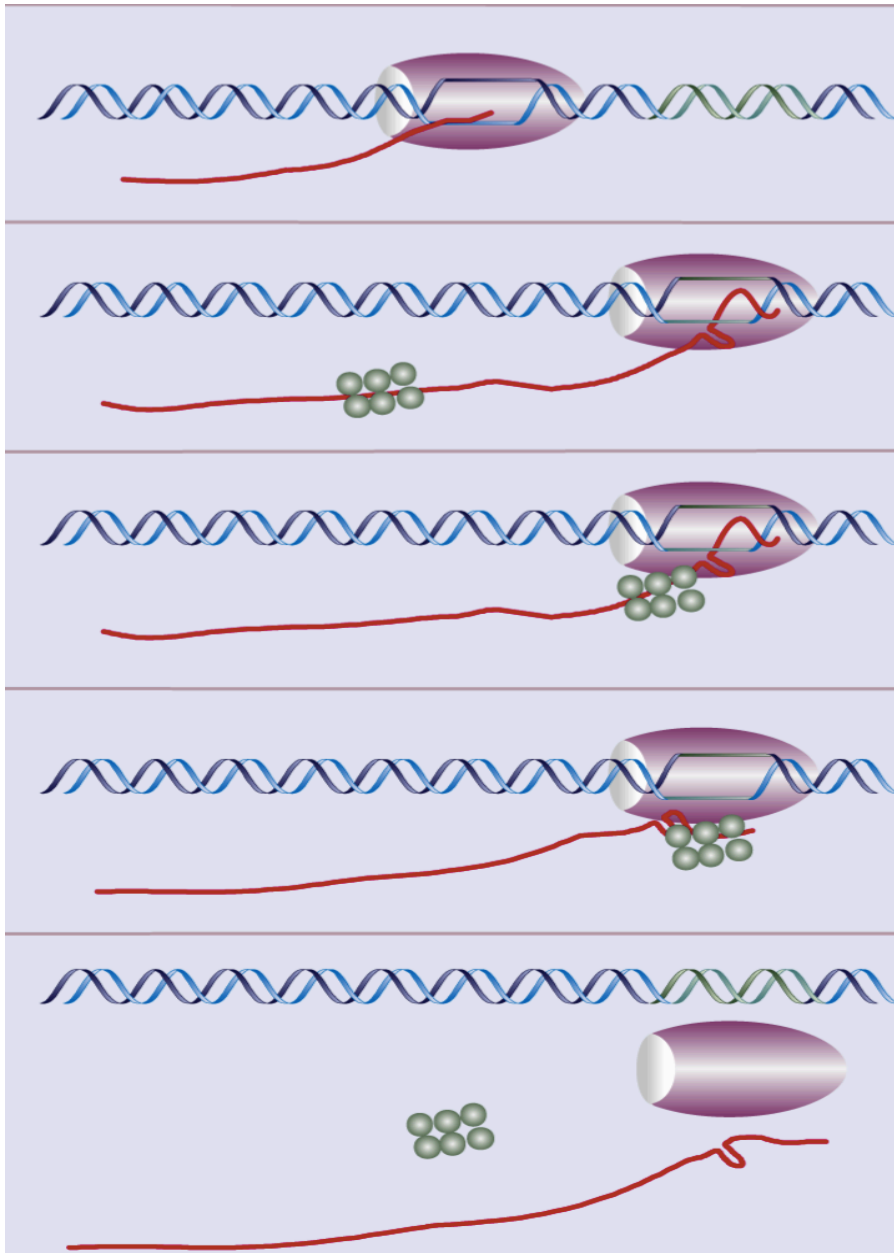
Rho (46kDa, ativa como hexâmero) move-se ao longo do RNA acompanhando a RNA polimerase

Com a pausa da RNA polimerase na seqüência de terminação, Rho alcança a enzima

Rho desenrola o híbrido RNA-DNA

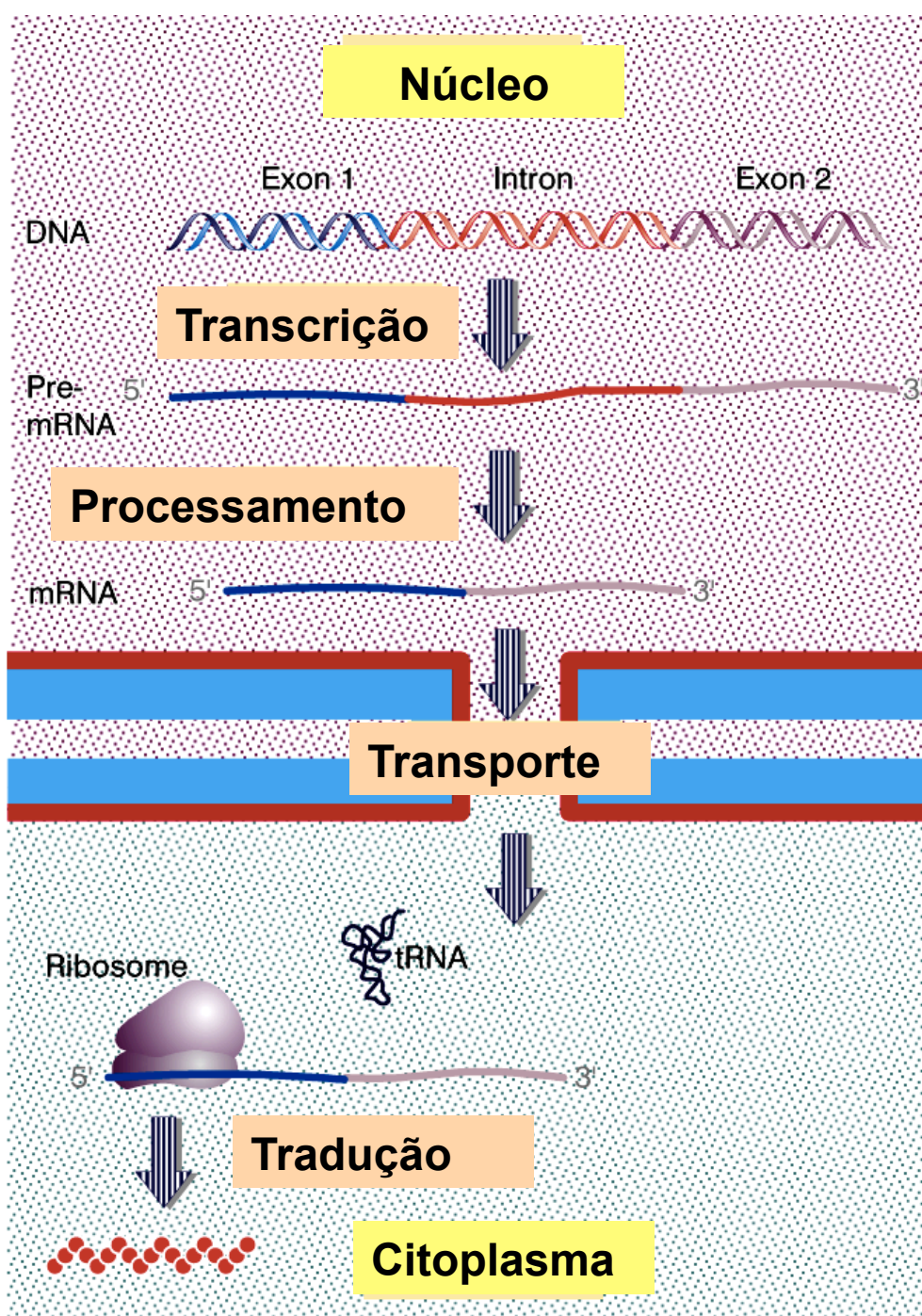
Seqüência de Terminação dependente de Rho: rica em C

AUCGCUACCUCAUAUCCGCACCUCUCAAACGCUACCUCGACCAGAAAGGCGUCUCUU



Proteínas necessárias para transcrição a partir de promotores reconhecidos pela RNA Pol II

Transcription factor	Number of subunits	Subunit M_r	Functions
Initiation			
RNA polymerase II	12	10,000–220,000	Catalyzes RNA synthesis
TBP (TATA-binding protein)	1	38,000	Specifically recognizes the TATA box
TFIIA	3	12,000, 19,000, 35,000	Stabilizes binding of TFIIIB and TBP to the promoter
TFIIIB	1	35,000	Binds to TBP; recruits RNA polymerase–TFIIF complex
TFIID	12	15,000–250,000	Interacts with positive and negative regulatory proteins
TFIIE	2	34,000, 57,000	Recruits TFIIH; ATPase and helicase activities
TFIIF	2	30,000, 74,000	Binds tightly to RNA polymerase II; binds to TFIIIB and prevents binding of RNA polymerase to nonspecific DNA sequences
TFIIH	12	35,000–89,000	Unwinds DNA at promoter; phosphorylates RNA polymerase; recruits nucleotide-excision repair complex
Elongation*			
ELL [†]	1	80,000	
P-TEFb	2	43,000, 124,000	
SII (TFIIS)	1	38,000	
Elongin (SIII)	3	15,000, 18,000, 110,000	



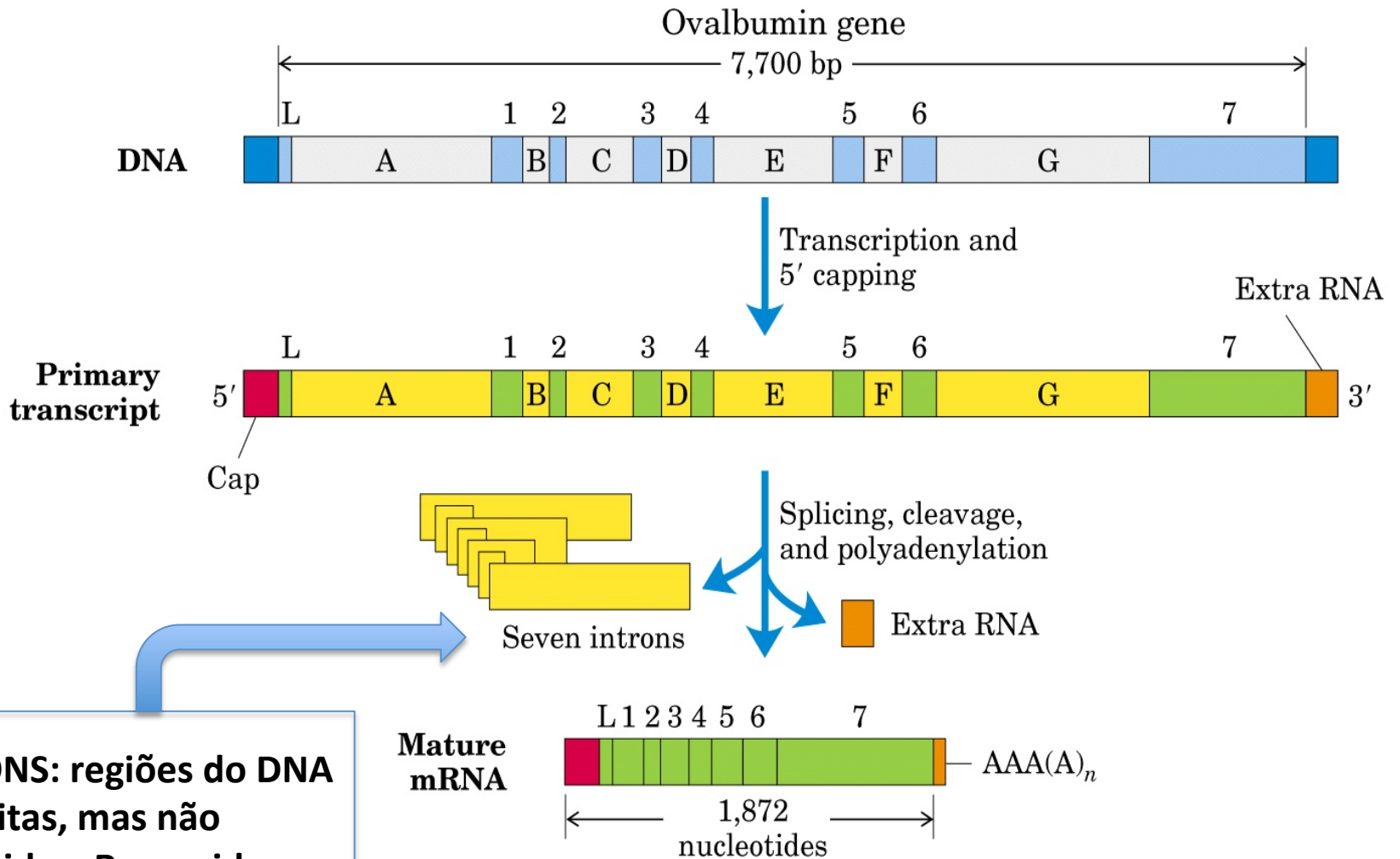
Etapas envolvidas na expressão de genes em eucariotos

Eucariotos

- Há processamento de:
 - mRNA
 - tRNA
 - rRNA

Processamento de mRNA de eucariotos

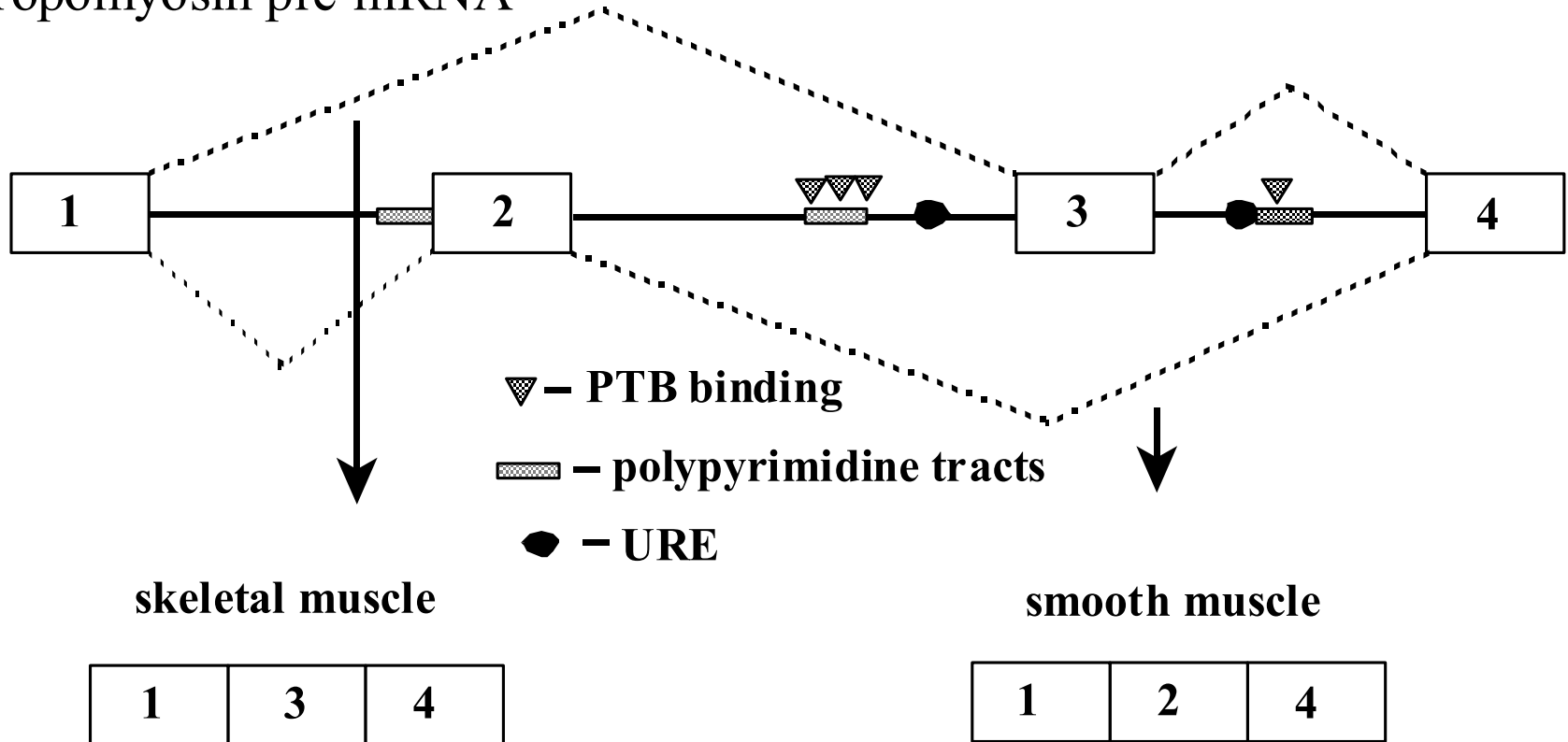
- **Adição de 5' CAP**- proteção a exonucleases e ligação a ribossomo
- **Retirada de regiões não codificantes- “splicing”**. Introns (50- 20000 nucleotídios): presentes em mRNA, rRNA, tRNA de genes nucleares, mitocondriais e de cloroplastos; raros em bactéria.
- **Adição de cauda de poliA (3')**- poliadenilato polimerase



INTRONS: regiões do DNA transcritas, mas não traduzidos. Removidos no processamento do mRNA

Alternative splicing of a tropomyosin

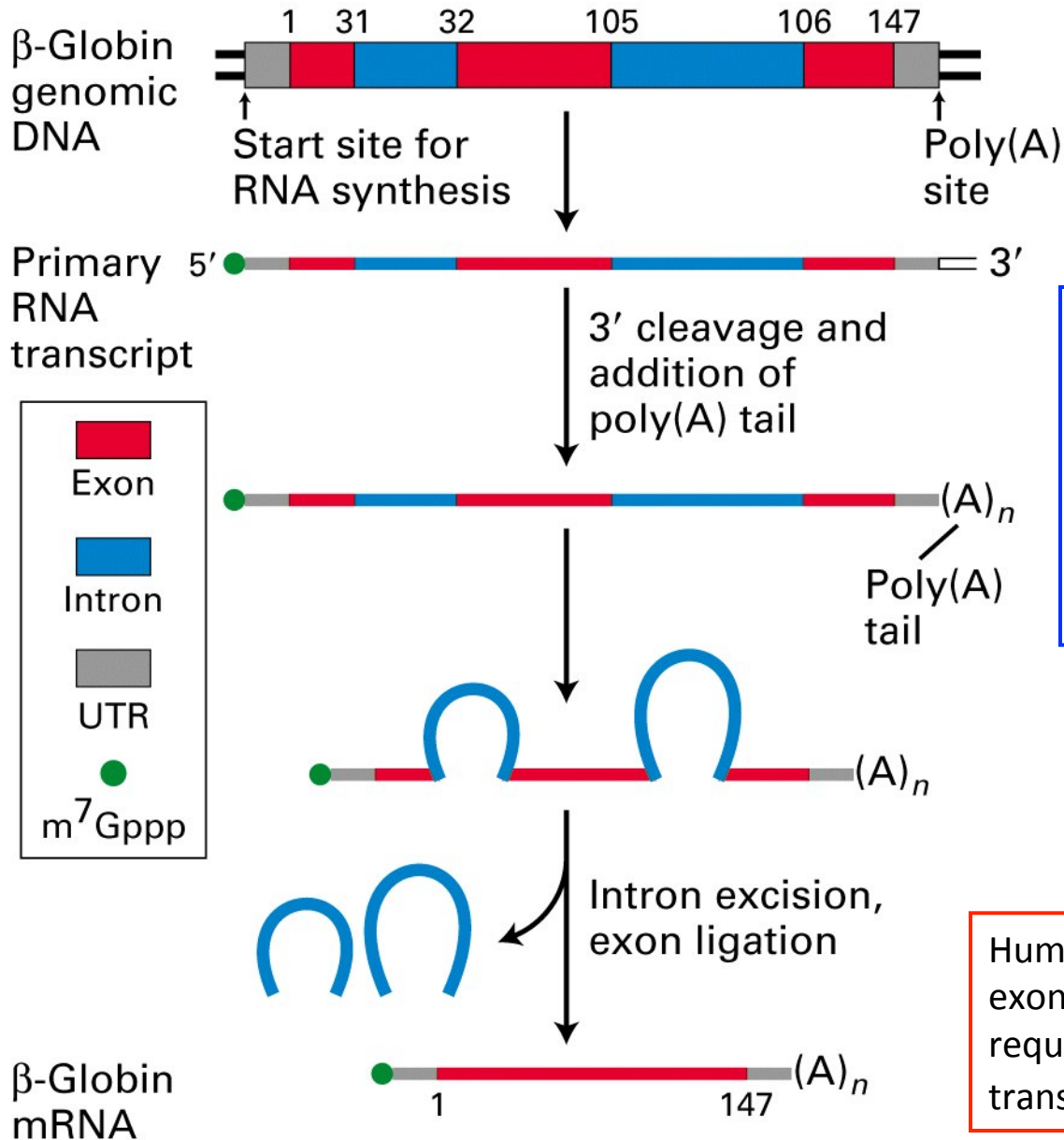
α -tropomyosin pre-mRNA



There are 3 forms of polypyrimiding tract binding protein (PTB) PTB1, PTB2 and PTB4.

Binding of PTB4 to the polypyrimidine suppresses splicing while binding of PTB1 promotes splicing. In smooth muscle exon 3 of α -tropomyosin is not present. Thus, PTB4 is expressed in smooth muscle while PTB1 is not.

Processing of eukaryotic pre-mRNA



For primary transcripts containing multiple exons and introns, splicing occurs before transcription of the gene is complete--co-transcriptional splicing.

Human dystrophin gene has 79 exons, spans over 2,300-Kb and requires over 16 hours to be transcribed!

Regulação da Expressão Gênica

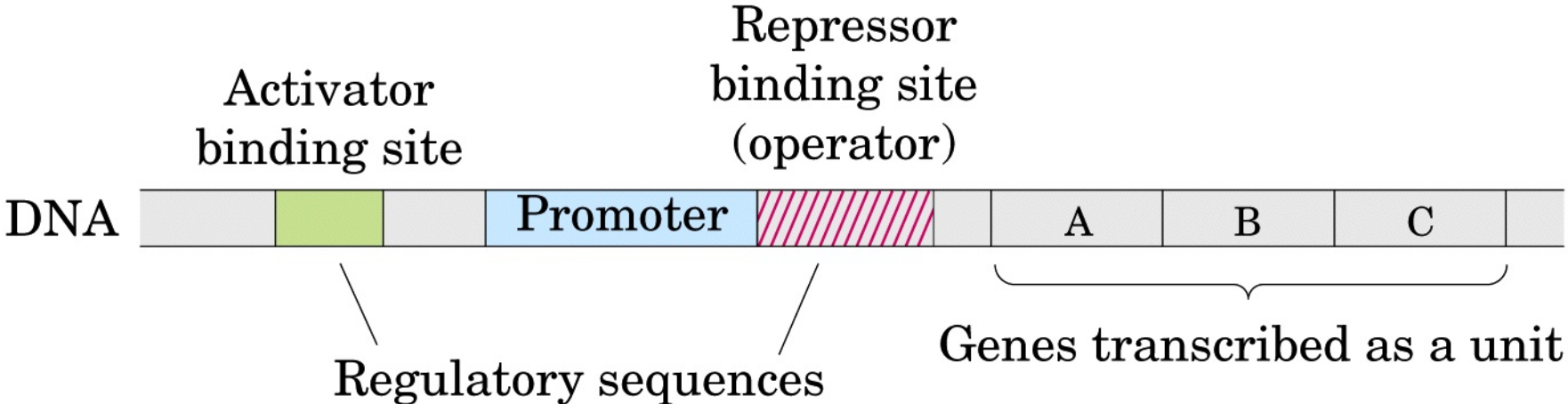
Procaríotos:

- Resposta direta a variações nas condições nutricionais (genes ativados e reprimidos)
- Transcrição pode ser acoplada com a tradução (simultânea)

Eucariotos multicelulares:

- Limitação na resposta direta às variações nas condições externas (células estão organizadas em tecidos e órgãos)
- Transcrição ocorre em compartimento distinto da tradução eliminando a possibilidade de acoplamento
- Regulação dependente de desenvolvimento e tecido

Procarotos



CRP- catabolic repression protein
CAP- catabolic gene activator protein

Desmetilação do DNA aparentemente está associada à ativação da cromatina



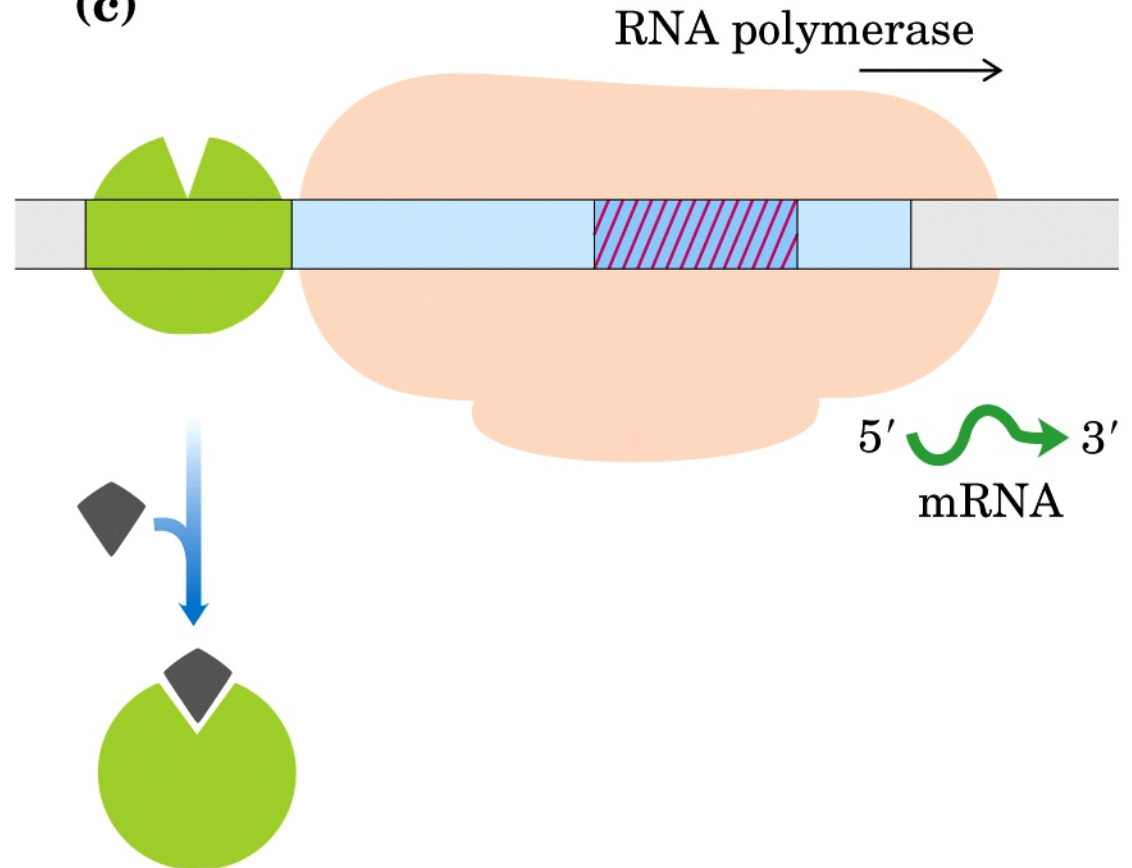
Estado de metilação do genes varia em diferentes tecidos e está correlacionado com a ativação da expressão

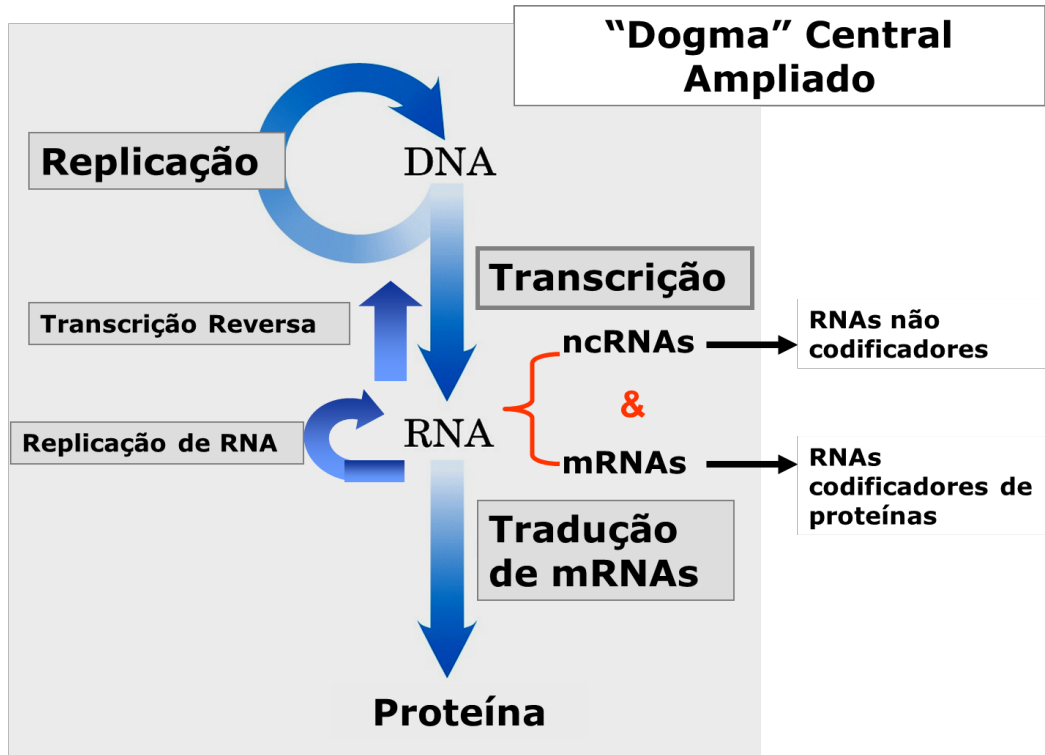
Exemplo de regulação de transcrição

Positive regulation
(bound activator facilitates transcription)

(c)

Molecular signal
(◆) causes *dissociation*
of regulatory protein
from DNA





Síntese de proteínas

Síntese proteica

- **~300 moléculas envolvidas**
- **~90% da energia gasta nos processos biossintéticos**
- **Bactérias: ~35% do peso: 20.000 ribossomos, 100.000 proteínas (fatores e enzimas), 200.000 tRNAs. Tradução acoplada à transcrição.**
- **Processo rápido: ~20 resíduos/seg**
- **Erro: 1 a cada 10.000 aa adicionados**

Código Genético: de RNA a proteína

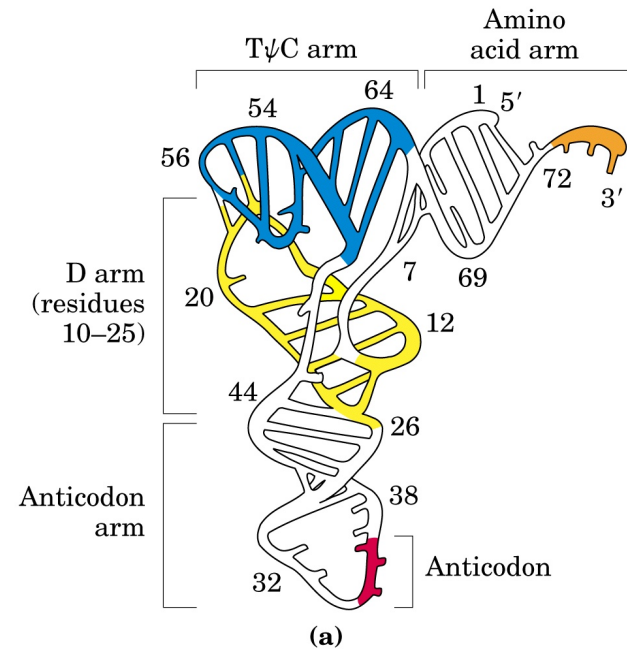
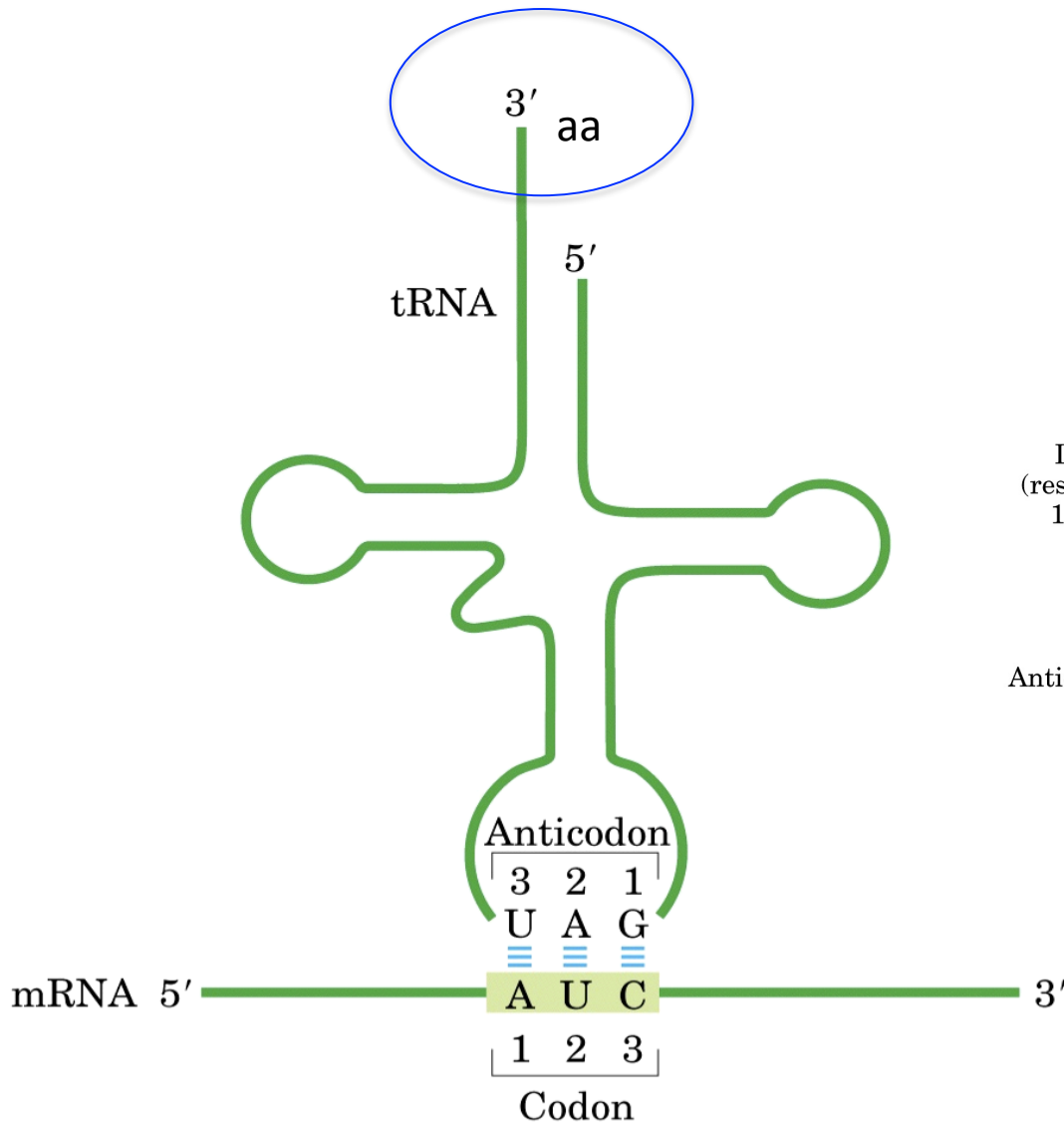
Alfabeto do DNA-
4 letras
Proteínas- 20 letras
3 nucleotídios- tripletes
4³- 64 trinucleotídios
possíveis

Tripletes= codons

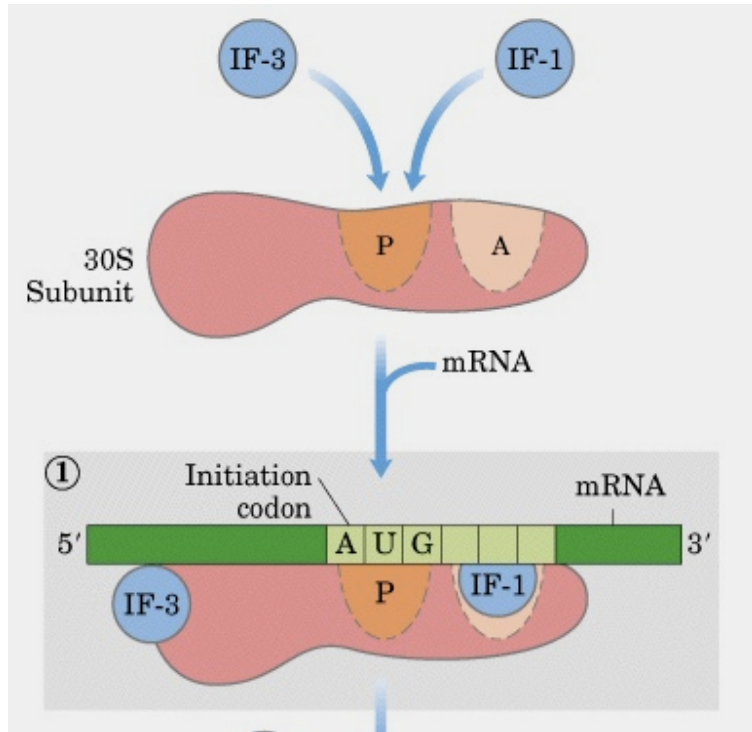
		Second letter				
		U	C	A	G	
First letter	U	UUU Phenylalanine UUC Phenylalanine UUA Leucine UUG Leucine	UCU Serine UCC Serine UCA Serine UCG Serine	UAU Tyrosine UAC Tyrosine UAA Stop codon UAG Stop codon	UGU Cysteine UGC Cysteine UGA Stop codon UGG Tryptophan	U C A G
	C	CUU Leucine CUC Leucine CUA Leucine CUG Leucine	CCU Proline CCC Proline CCA Proline CCG Proline	CAU Histidine CAC Histidine CAA Glutamine CAG Glutamine	CGU Arginine CGC Arginine CGA Arginine CGG Arginine	U C A G
	A	AUU Isoleucine AUC Isoleucine AUA Isoleucine AUG Methionine; start codon	ACU Threonine ACC Threonine ACA Threonine ACG Threonine	AAU Asparagine AAC Asparagine AAA Lysine AAG Lysine	AGU Serine AGC Serine AGA Arginine AGG Arginine	U C A G
	G	GUU Valine GUC Valine GUA Valine GUG Valine	GCU Alanine GCC Alanine GCA Alanine GCG Alanine	GAU Aspartic acid GAC Aspartic acid GAA Glutamic acid GAG Glutamic acid	GGU Glycine GGC Glycine GGA Glycine GGG Glycine	U C A G

Códon de iniciação- AUG
Códons de terminação: UAA, UAG, UGA

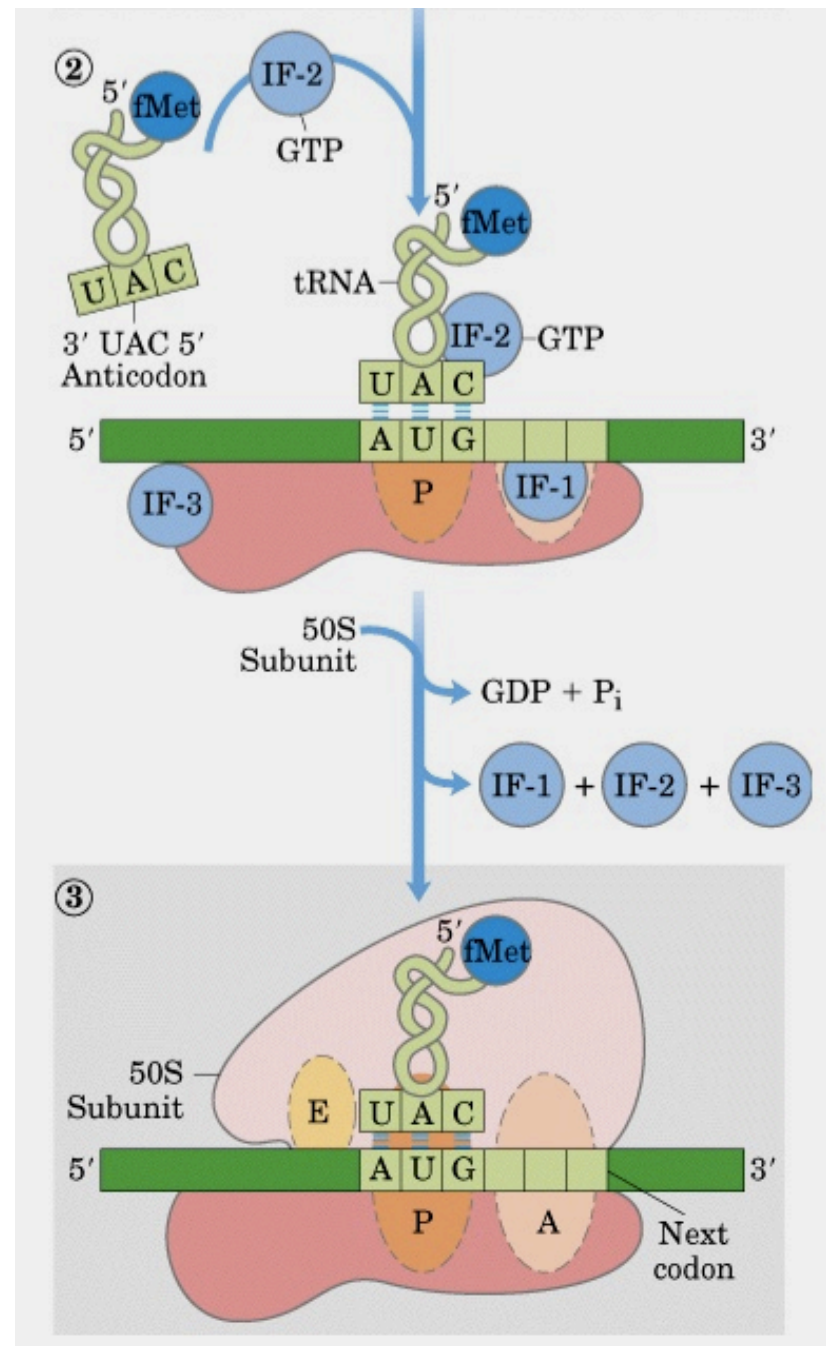
tRNA como adaptador na síntese proteica: reconhecimento codon-anticodon



2. Iniciação (bactéria)



Ribossomo bacteriano (16S) reconhece uma sequência no mRNA- Shine Delgarno



Esquema de síntese proteica

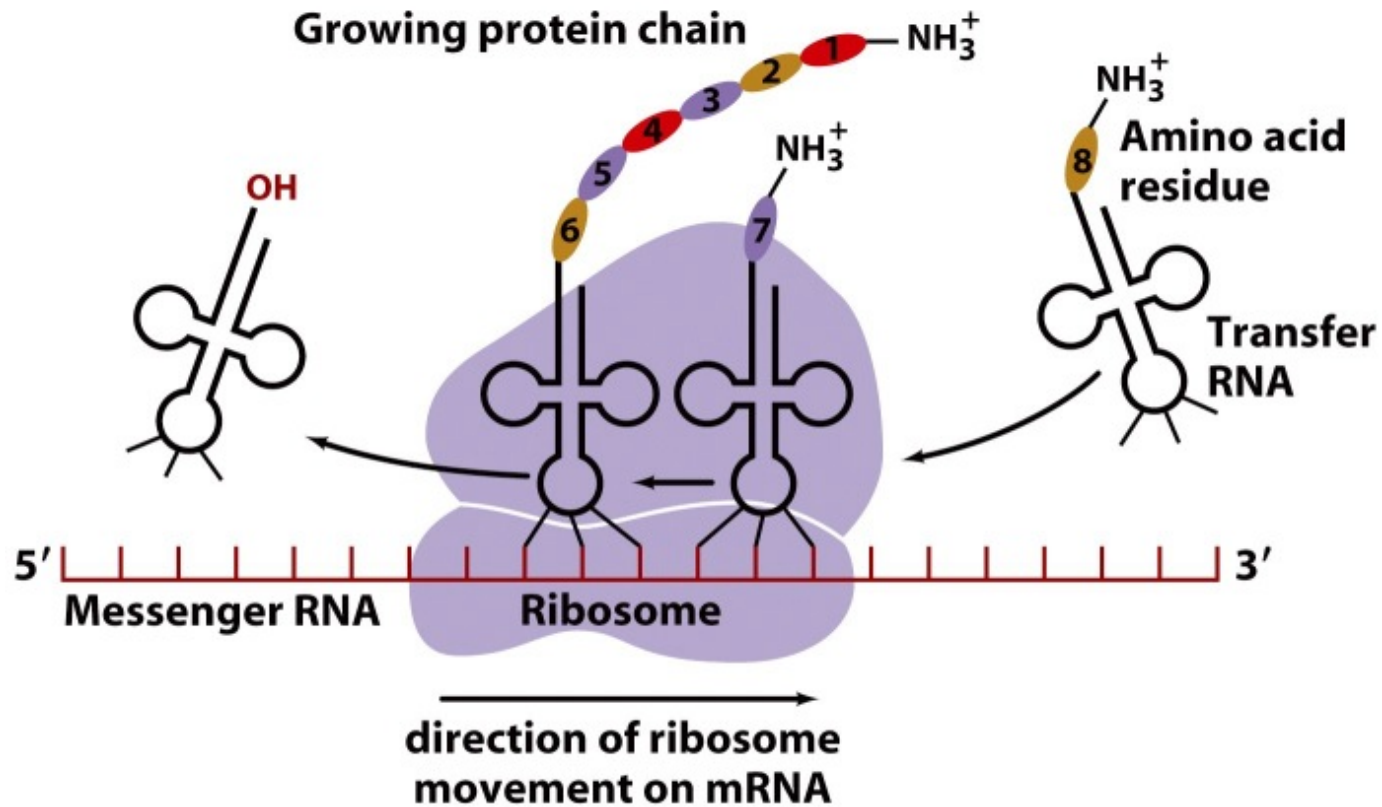
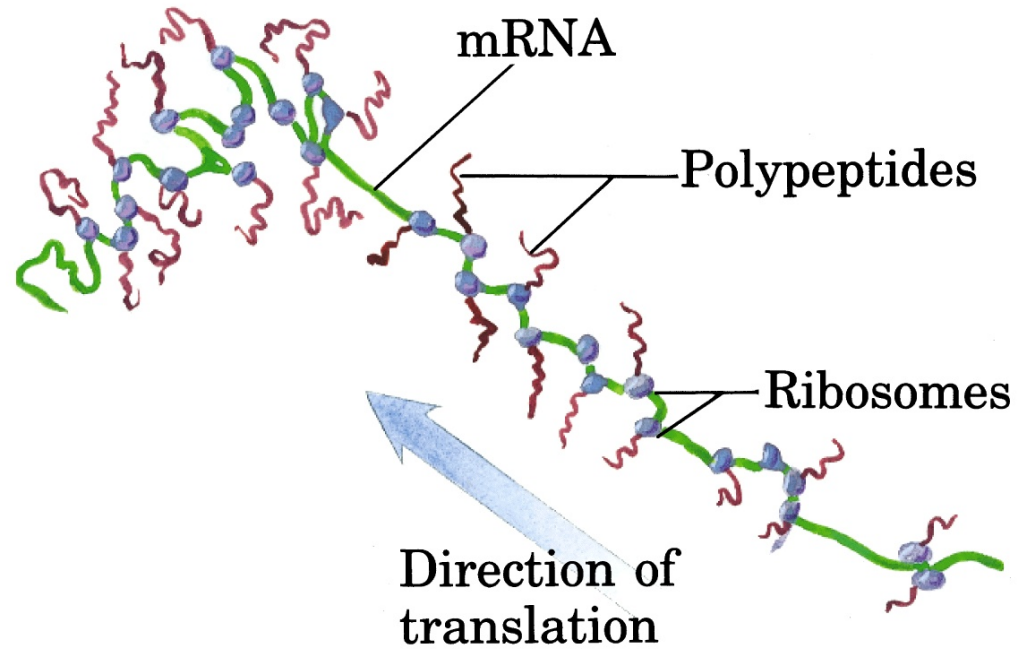
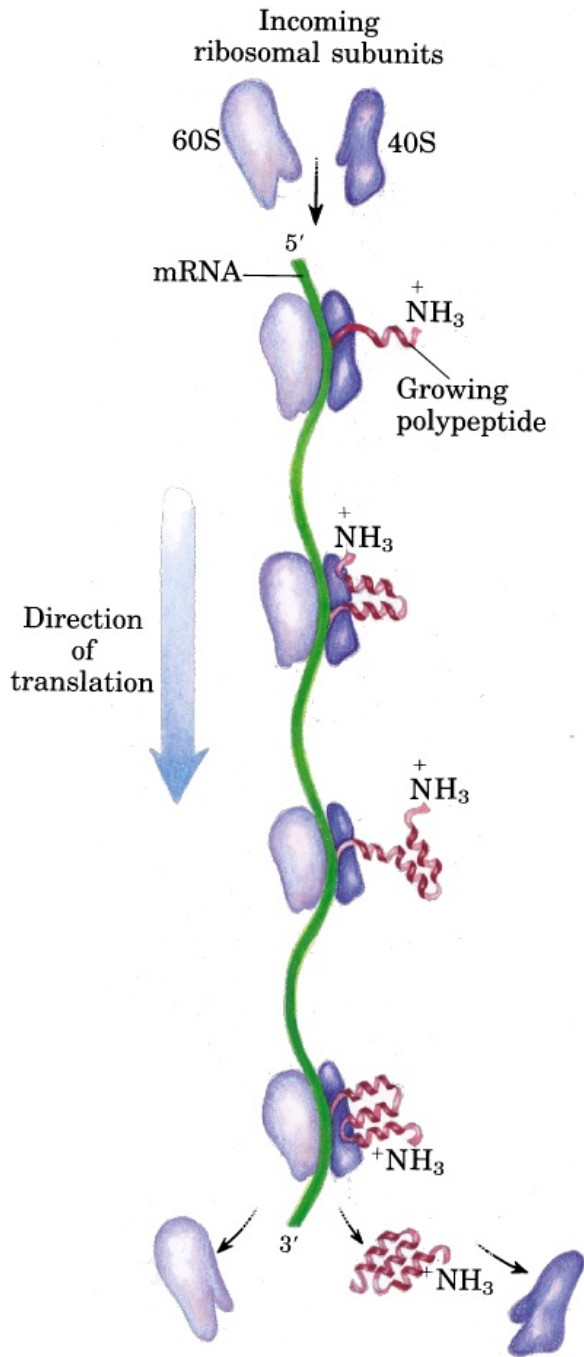


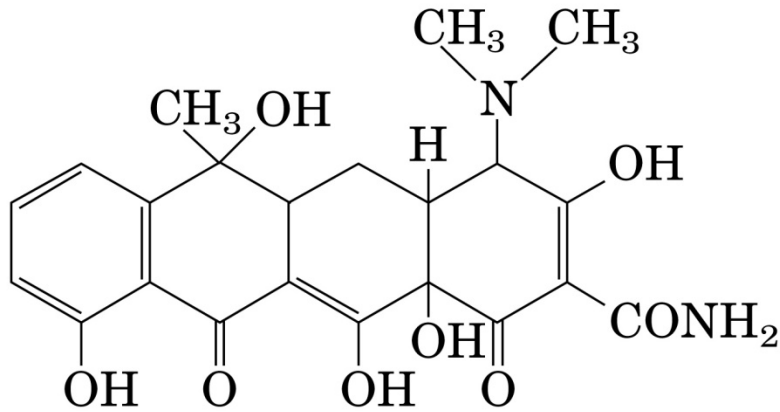
Figure 3-14 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

Polissomo



Polissomos- 10 a 100 ribossomos

Inibição da síntese proteica por antibióticos e toxinas

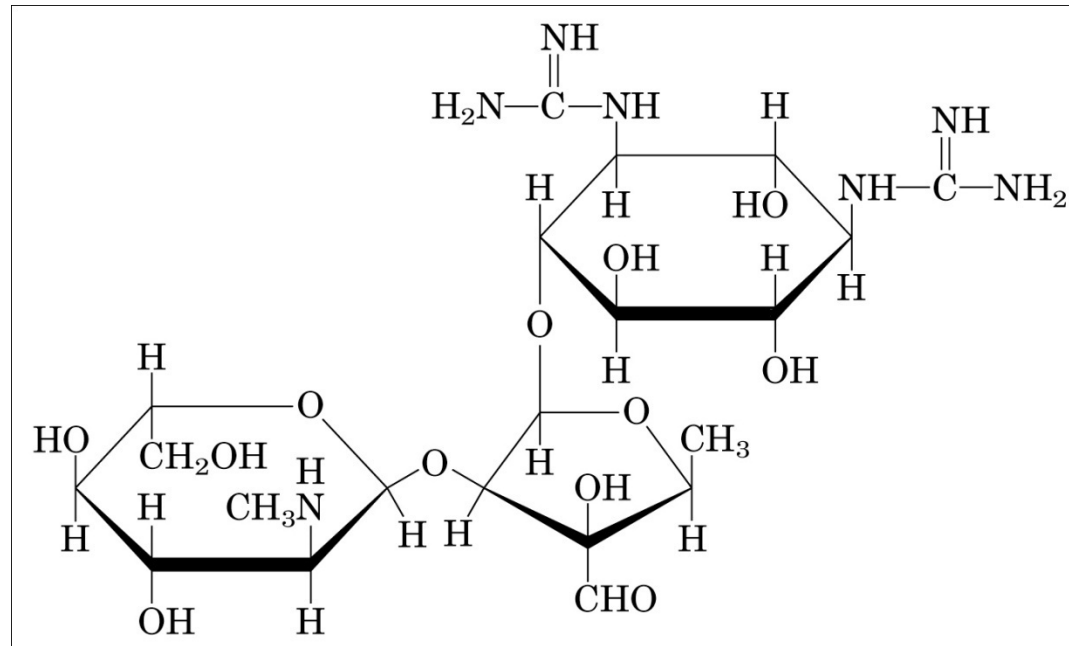


Tetraciclina

Bloqueia o sítio A do ribossomo bacteriano e inibe associação do aminoacil-tRNA

Estreptomicina

Causa leitura incorreta dos códons e inibe iniciação



**20 aminoácidos
diferentes**

**Modificações covalentes
nos aminoácidos
(modificações pós-
traducionais):**

Ex:

Metilação

Acetilação

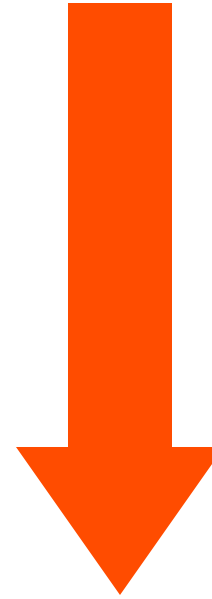
Hidroxilação

Glicosilação

Fosforilação

Grupos prostéticos

Acilação



**~200 aminoácidos
diferentes**

(modificados covalentemente)

genoma

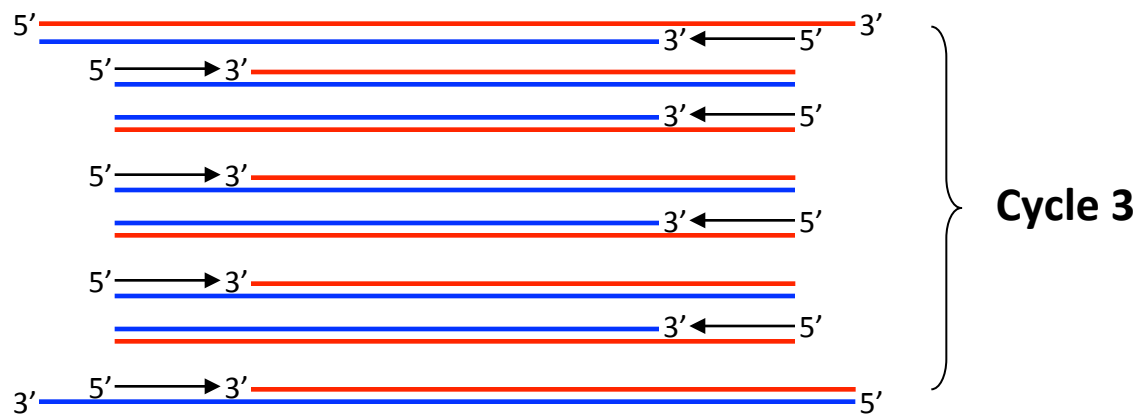
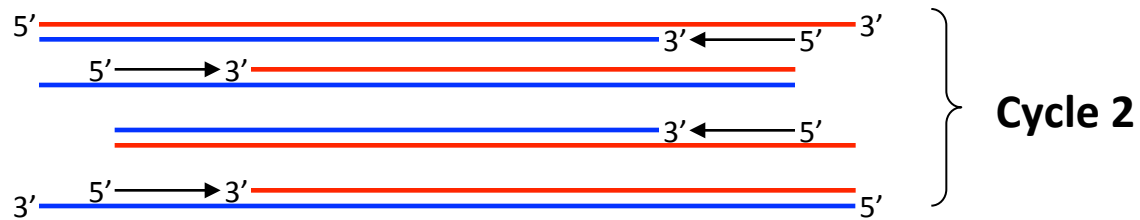
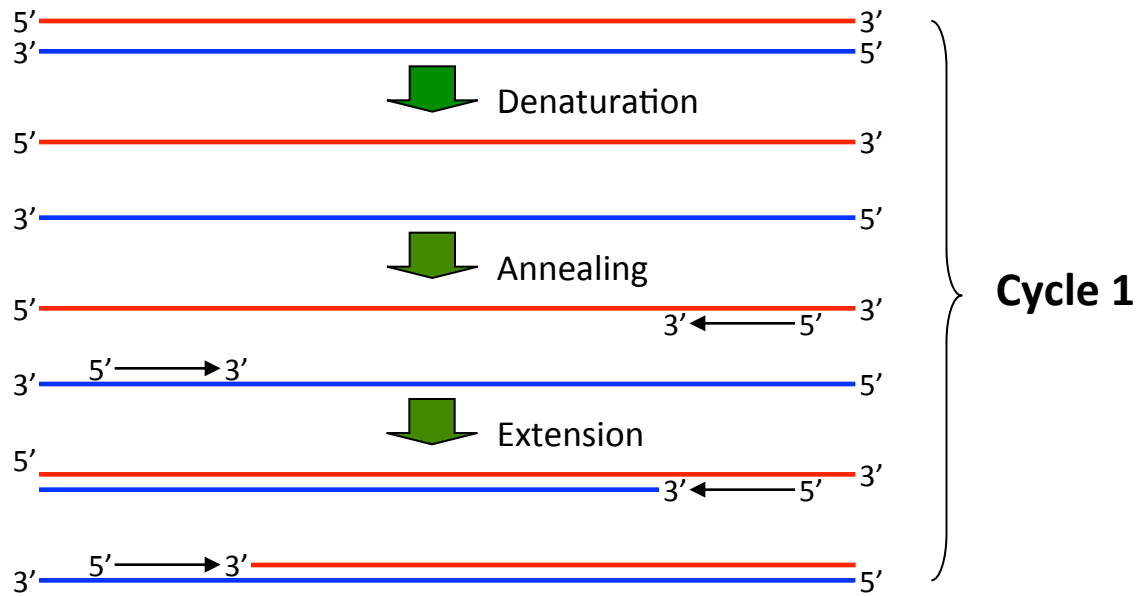
- Diferentes métodos de sequenciamento do DNA (primeiros em 1977- Sanger). Fragmentos de DNA sequenciados com superposição
- Assembly
- Anotação- isto é que sequência de DNA corresponde a uma proteína (portanto com base nas proteínas já sequenciadas em algum organismo)
- Muitas sequências ainda classificadas como hipotéticas
- Primeiro genoma sequenciado- 1977 (fago)
- Primeiro genoma humano- 2001
- Bioinformática
- Banco de dados (Kregg, genbank, etc)

Sequências no DNA não traduzidas genoma humano

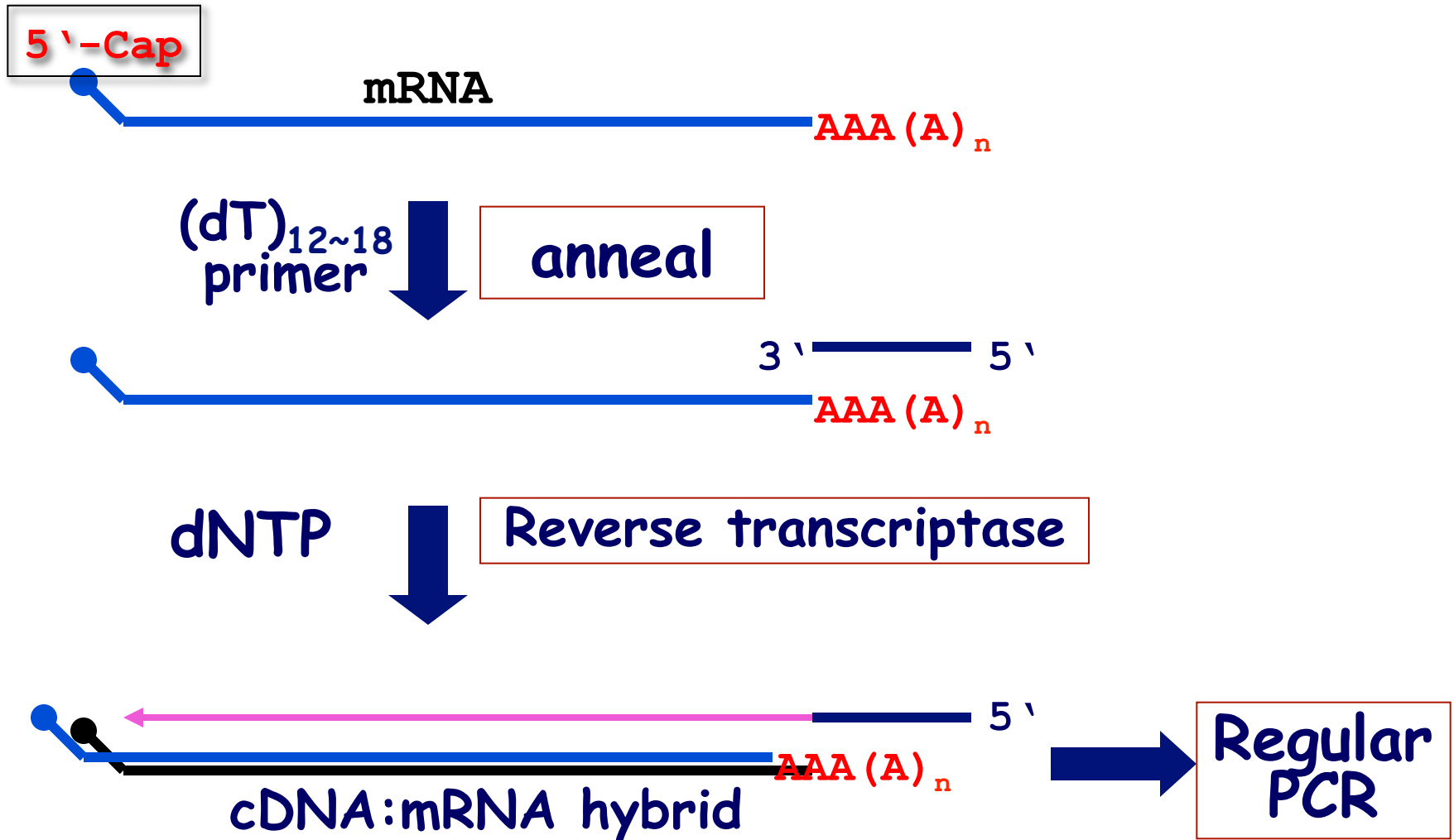
- 3% - Exons (codificantes)
- 10% - sequências conservadas na evolução. Função?
- 99% do genoma de mamífero é transcrito- além do mRNA, pequenos e longos RNA não codificantes, introns, enhancer, telômeros, etc
- Mutações nas regiões não codificantes- *single nucleotide polymorphism (SNP)*- marcadores de patologias (~100 associadas a estados patológicos; pessoa saudável: ~25.000 SNPs; 3.500.000 SNPs identificados

Polymerase chain reaction (PCR)

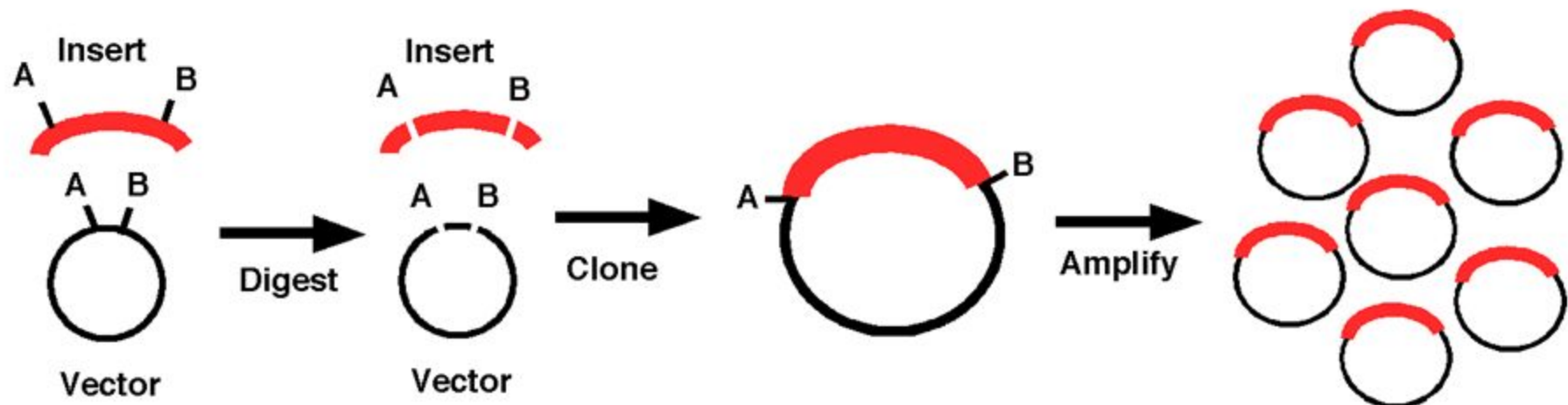
- Amplificar segmentos de ac. Nucleicos-
- DNA
- RNA (RT-PCR)



RT-PCR



Cloning Scheme

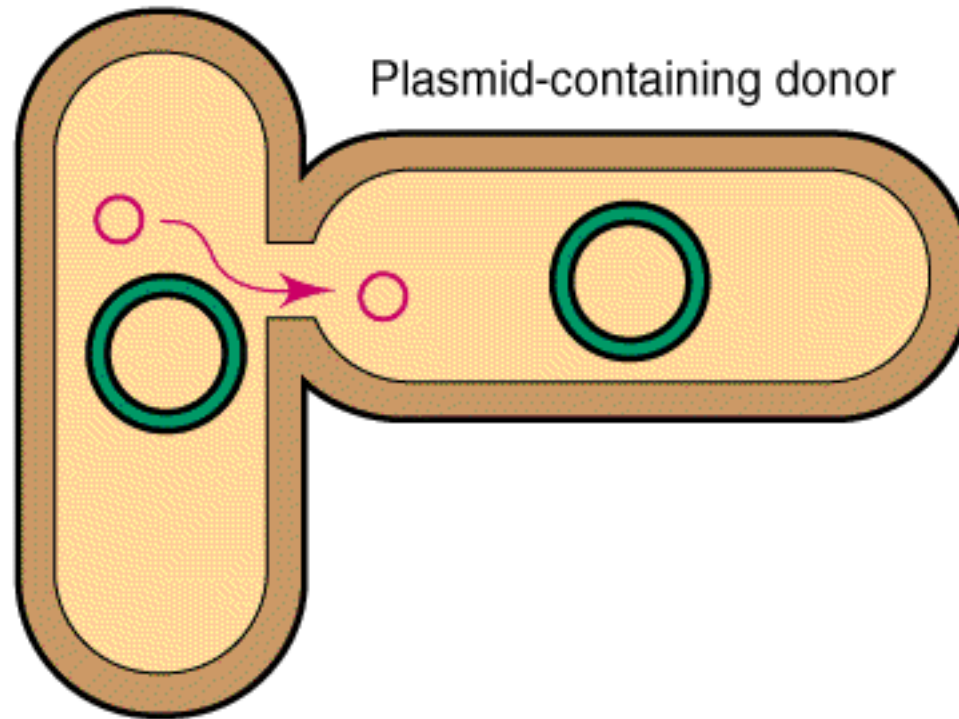


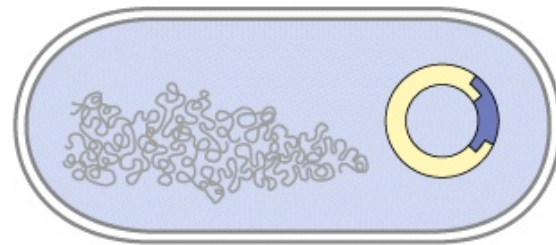
Digest

Ligate

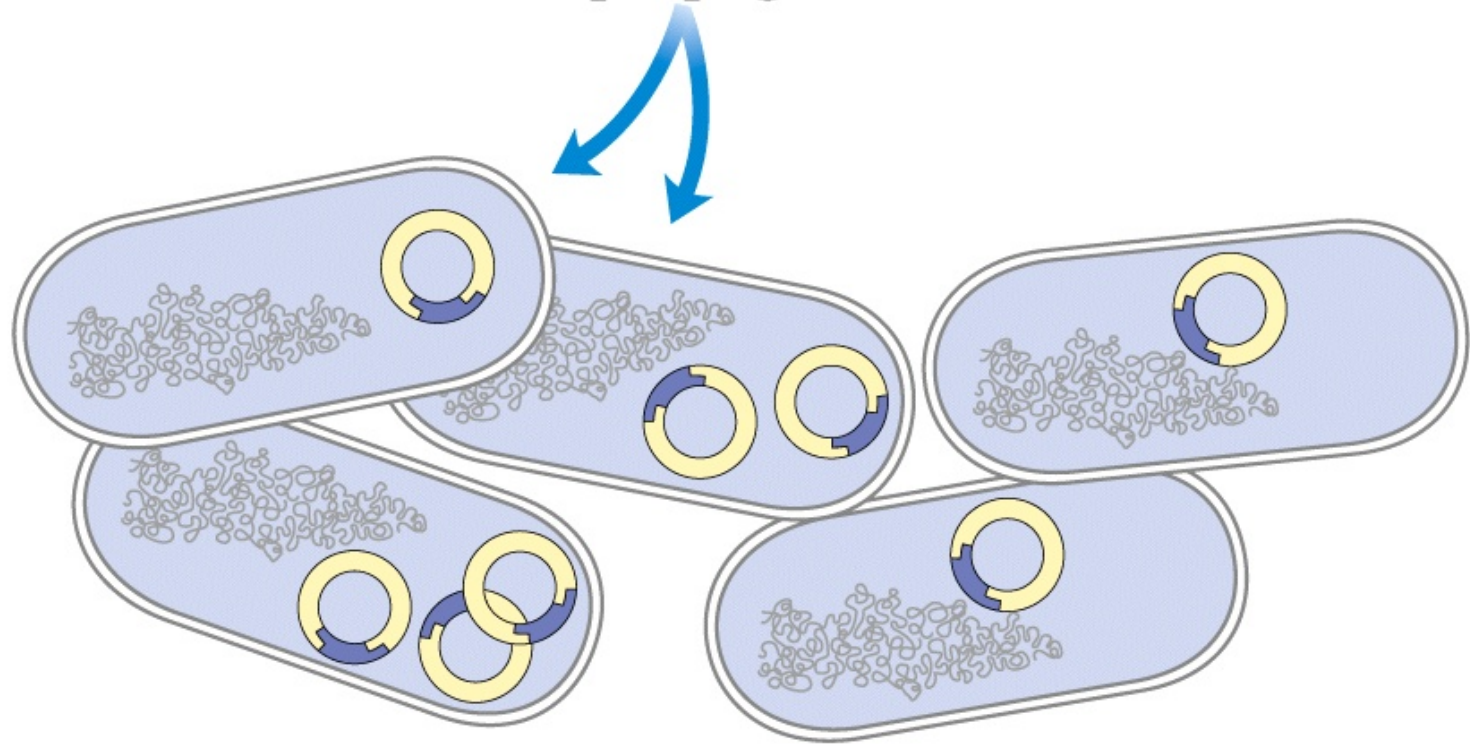
Amplify and Prep

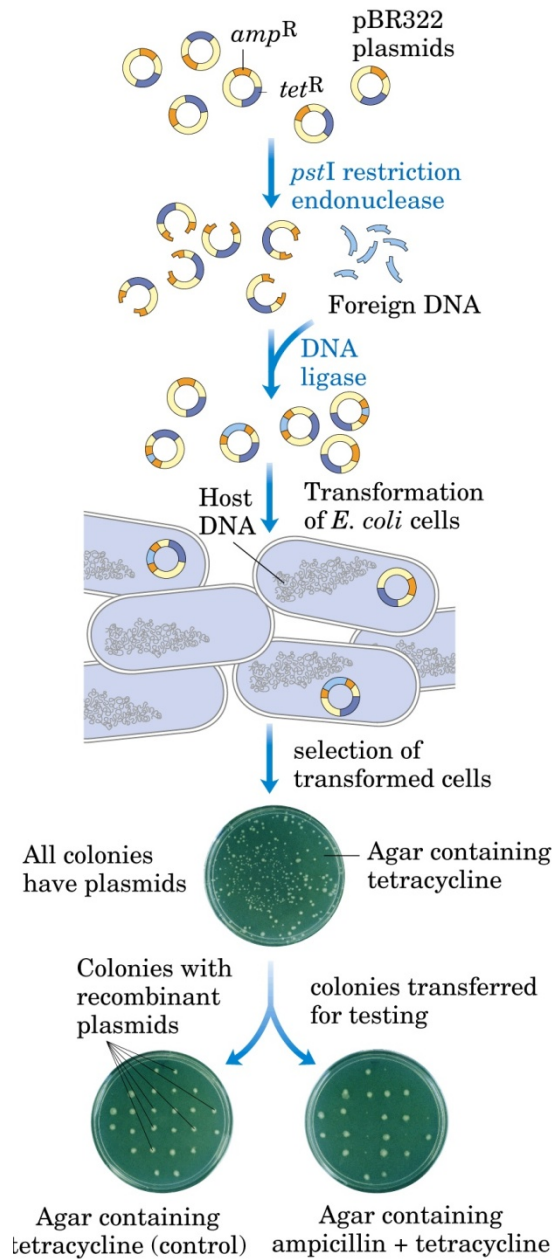
Conjugação: transferência de plasmídeo





cell propagation





Informações adicionais

Fluxo da Informação Gênica

DNA 5' — A-G-A-G-G-T-G-C-T — 3'

3' — T-C-T-C-C-A-C-G-A — 5'

↓ **Transcrição**

mRNA 5' — A-G-A-G-G-U-G-C-U — 3'

tRNAs 3' U-C-U C-C-A C-G-A 5'

┌───┐ ┌───┐ ┌───┐
└───┘ └───┘ └───┘
Arginine Glycine Alanine

↓ **Tradução**

Protein —Arginine—Glycine—Alanine—

RT-PCR Cloning of a Partial cDNA for the Porcine Arylhydrocarbon Receptor (AhR)

Ld 1 2 3 4 nt nt ntq Ld



Ld = 100 bp ladder; nt = No template; Ntq = No Taq
1 = Adrenal; 2 = Pancreas; 3 = Uterus; 4 = Lung
Arrow head indicates 200 bp

RNAs

Tipo	Tamanho	Função
tRNA	Pequeno	Transporte de aa para o local de síntese
rRNA	Diversos	Forma os ribossomos, juntamente com proteínas
mRNA	Diversos	Determina a sequência de aa na proteína
snRNA	Pequeno	Processa o mRNA inicial nos eucariotos
miRNA	Pequeno	Afeta a expressão gênica (crescimento, desenvolvimento)
siRNA	pequeno	Afeta a expressão gênica. Cientistas utilizam para bloquear a expressão do gene de interesse

GENOMA

TRANSCRIPTOMA

PROTEOMA

Genes
(2×10^4)

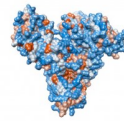
transcrição

mRNAs
($>10^6$)

tradução

Proteínas
($>10^6$)

Proteínas
Modificadas
($>10^7$)



FENÓTIPO

Reações enzimáticas

METABOLOMA

(3×10^3 metabolitos)

telômeros

- Sequências na ponta dos cromossomos lineares
- Telômeros de levedo- ~100 pares de bases repetidas, sem muita precisão- 5'(TxGy)_n; 3'(AxCy)_n- x e y ~1-4; de mamífero-- ~1500 pb
- A cada mitose- perde cerca de 50- 200 nucleotídeos/ divisão – a DNA polimerase não consegue refazer os fragmentos na ponta do cromossomo linear (correspondente ao *primer*). Diminuição do telômero- associado com envelhecimento- sinal limita o no. de divisões (também contestado)
- Células imortais (incluindo cancer)- Enzima telomerase: composta por RNA e proteína. RNA tem cópia da sequência repetida do telômero- funciona como template para a síntese de TxGy de uma fita- funciona portanto como transcriptase reversa. Síntese na direção 5' - 3'

Telômeros na ponta dos cromossomos

