

ESTRUTURA DO GENE E TRANSCRIÇÃO GÊNICA

QBQ 0104 Bioquímica de macromoléculas – Fisioterapia USP

O DNA e o genoma

- O conteúdo de DNA de uma célula é chamado de genoma
- Bactérias, fungos, plantas e animais, têm genomas de tamanhos diferentes
- Organismos da mesma espécie têm o mesmo genoma e o mesmo número de cromossomos
- O genoma humano tem 3 bilhões de nucleotídeos (pares de bases) distribuídos em 23 cromossomos
- Nós temos duas cópias de cada cromossomo, uma do pai e outra da mãe

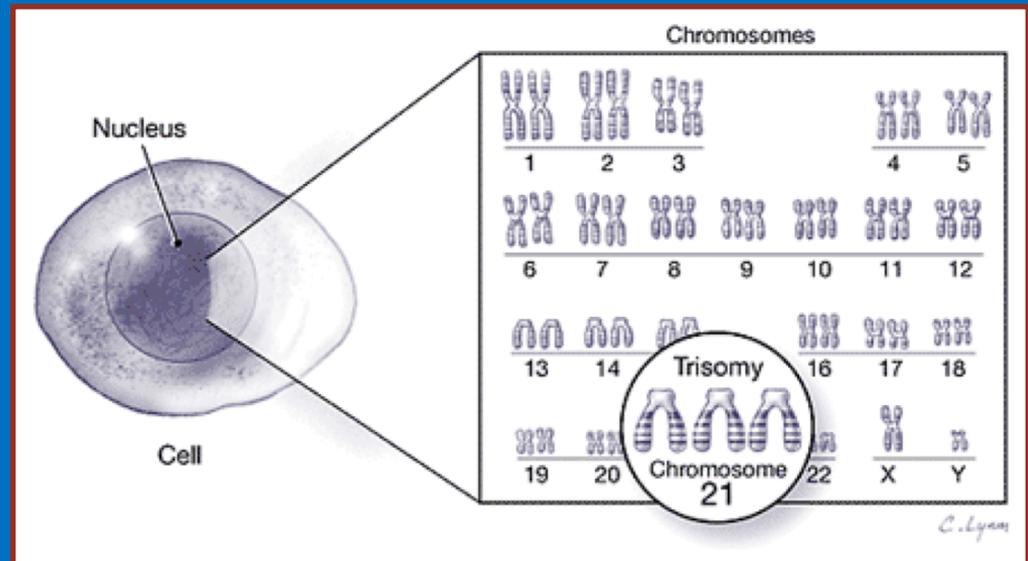


TABLE 24-2 DNA, Gene, and Chromosome Content in Some Genomes

	Total DNA (bp)	Number of chromosomes*	Approximate number of genes
<i>Escherichia coli</i> K12 (bacterium)	4,639,675	1	4,435
<i>Saccharomyces cerevisiae</i> (yeast)	12,080,000	16 [†]	5,860
<i>Caenorhabditis elegans</i> (nematode)	90,269,800	12 [‡]	23,000
<i>Arabidopsis thaliana</i> (plant)	119,186,200	10	33,000
<i>Drosophila melanogaster</i> (fruit fly)	120,367,260	18	20,000
<i>Oryza sativa</i> (rice)	480,000,000	24	57,000
<i>Mus musculus</i> (mouse)	2,634,266,500	40	27,000
<i>Homo sapiens</i> (human)	3,070,128,600	46	29,000

Note: This information is constantly being refined. For the most current information, consult the websites for the individual genome projects.

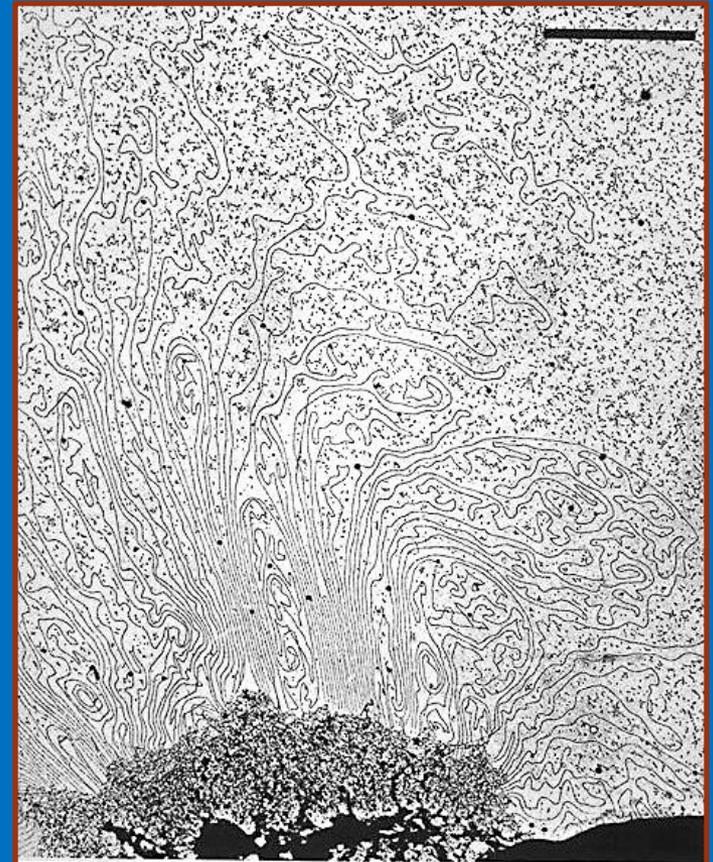
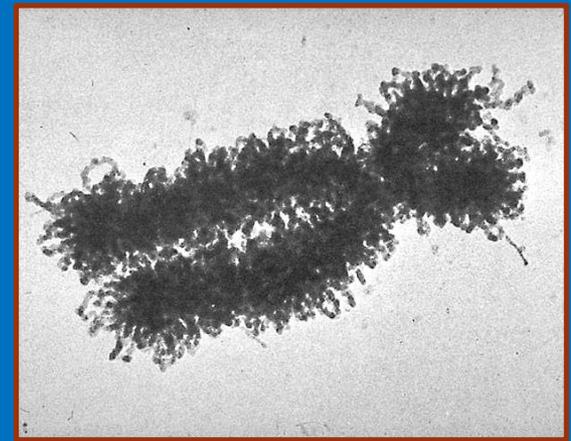
*The diploid chromosome number is given for all eukaryotes except yeast.

[†]Haploid chromosome number. Wild yeast strains generally have eight (octoploid) or more sets of these chromosomes.

[‡]Number for females, with two X chromosomes. Males have an X but no Y, thus 11 chromosomes in all.

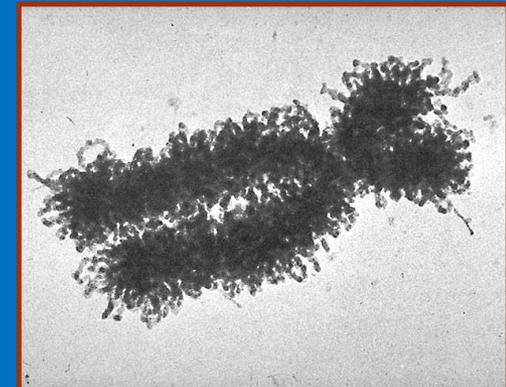
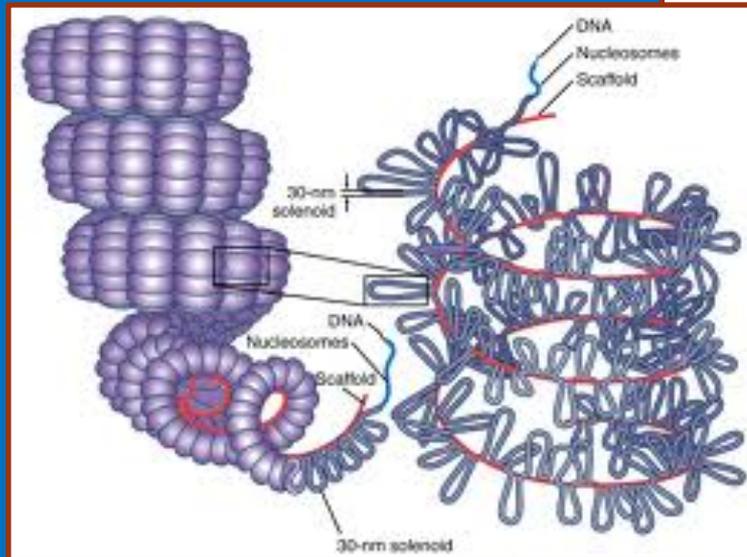
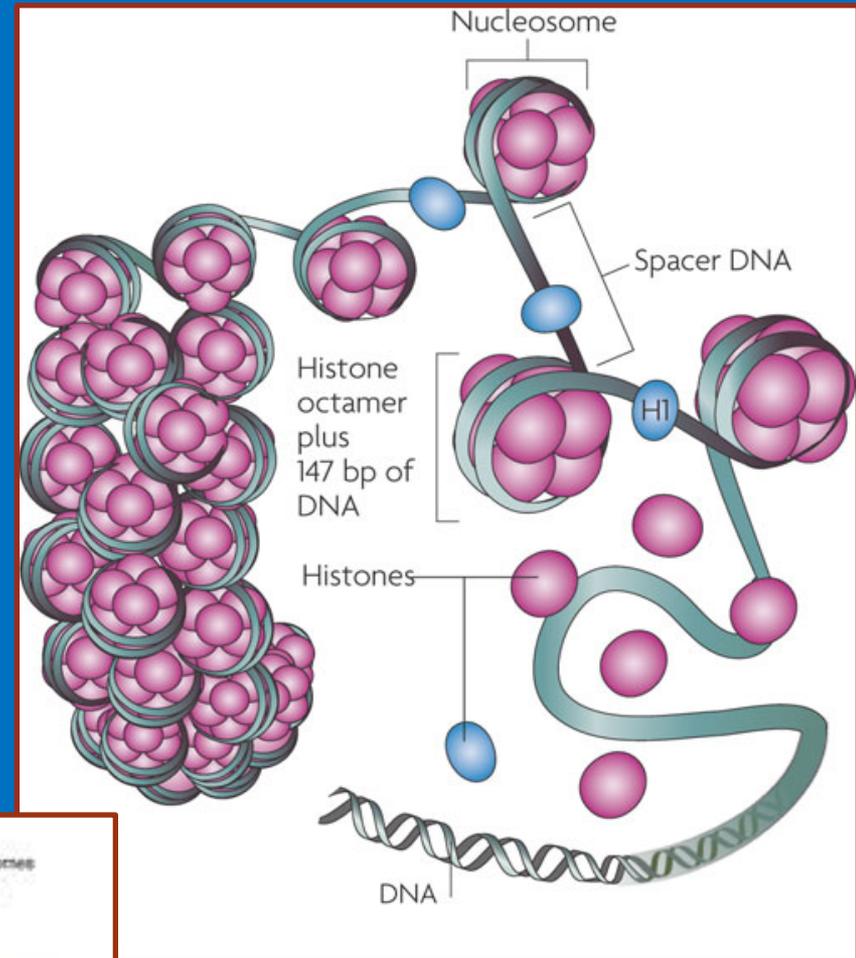
Cromossomos e o DNA

- Em média, o DNA de uma única célula humana, se desenovelado, tem ~5cm de comprimento.
- Você tem ~10 trilhões de células; se cada molécula de DNA fosse amarrada umas as outras, você obteria um filamento com 1.190 milhões de Km. Suficiente para ir e voltar da lua 1.500!
- Como um fita de DNA tão longa pode caber dentro de uma célula???



Cromossomos e o DNA

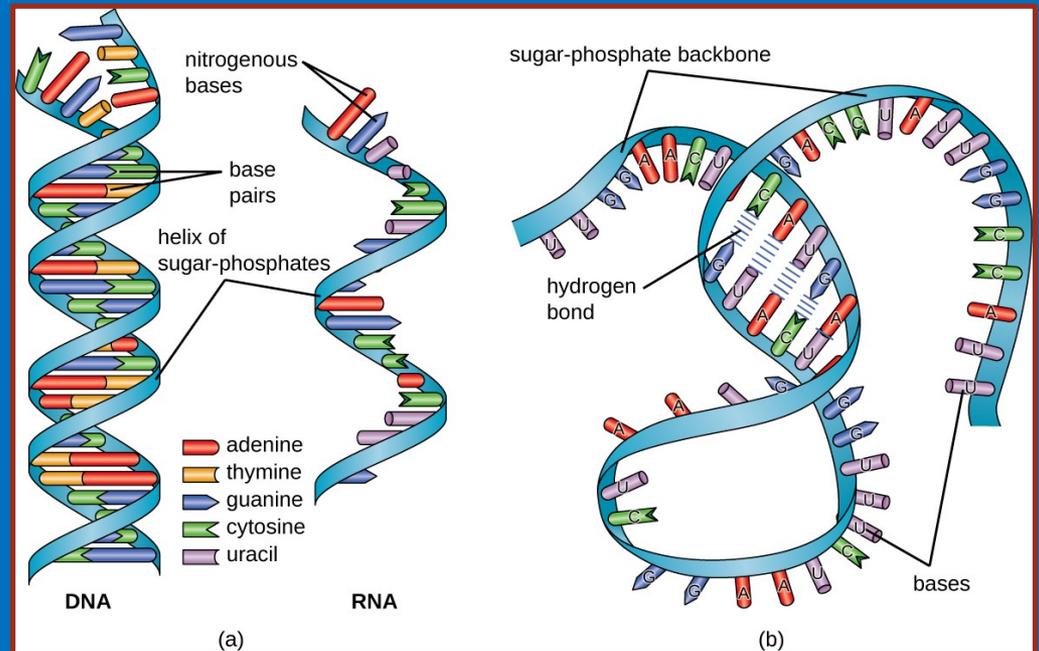
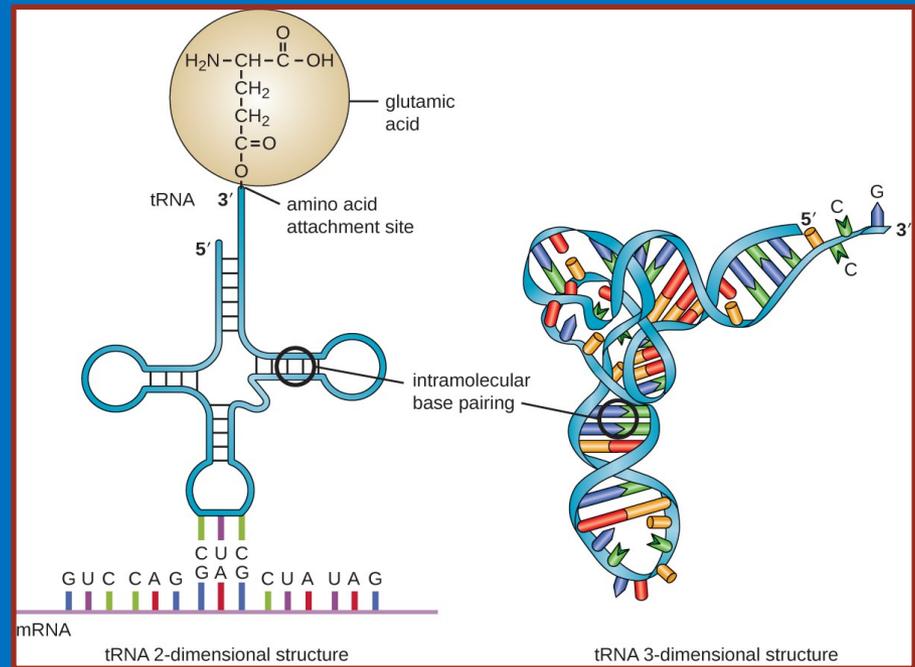
- Isto é possível, porque dentro da célula o DNA está empacotado
- Proteínas especiais (histonas) se ligam ao DNA e sua fita é "enrolada" em torno destas proteínas
- Estas unidades são chamadas de nucleossomos
- É desta forma, compactada, que o DNA cabe dentro da célula



EoRNA?

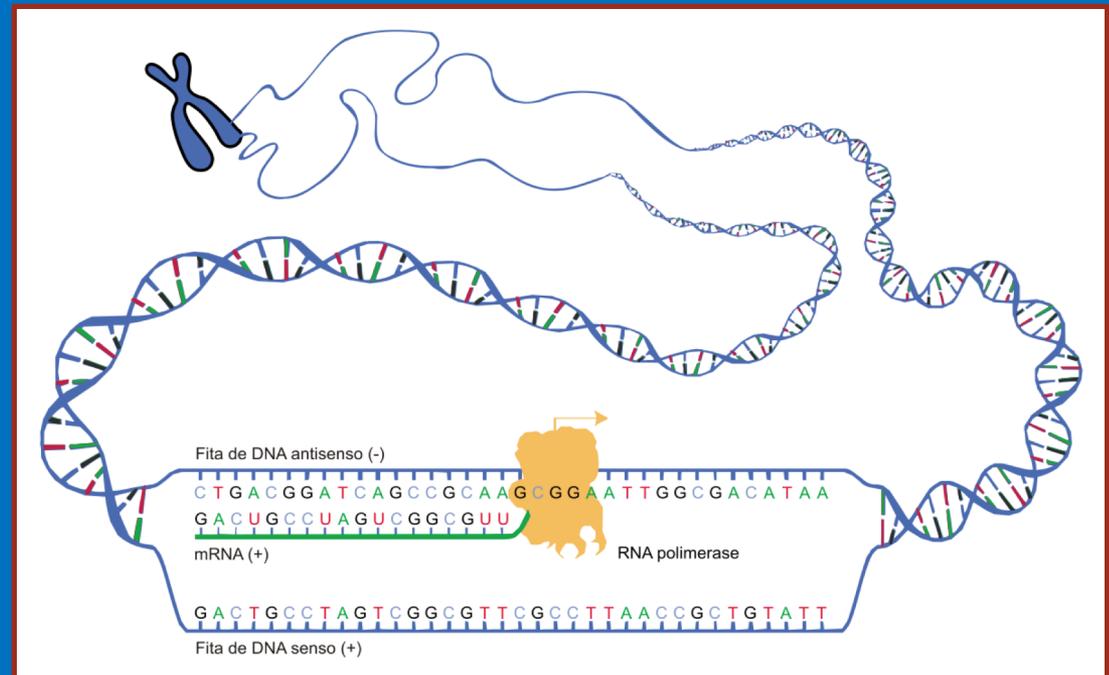
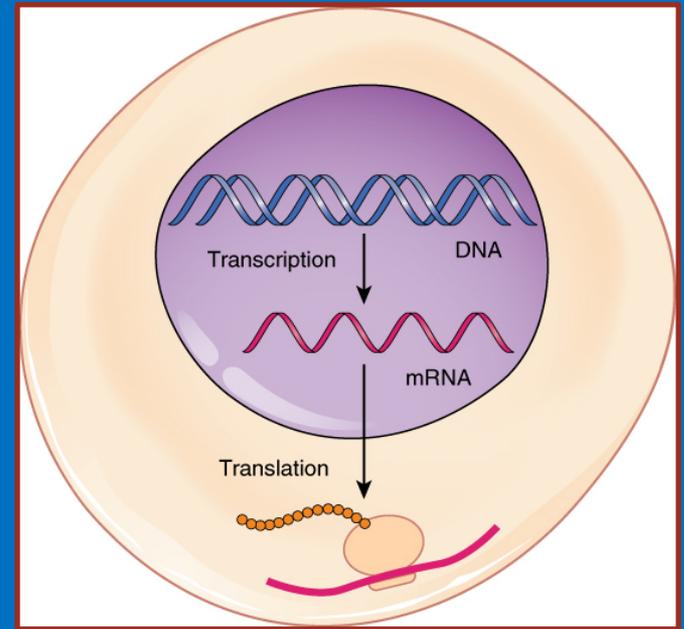
ORNA

- O RNA não forma uma dupla hélice
- Porém, moléculas de RNA podem assumir estruturas secundárias
- Por exemplo, o ribossomo, importante para a síntese de proteínas, é formado por RNA



Transcrição e a RNA polimerase

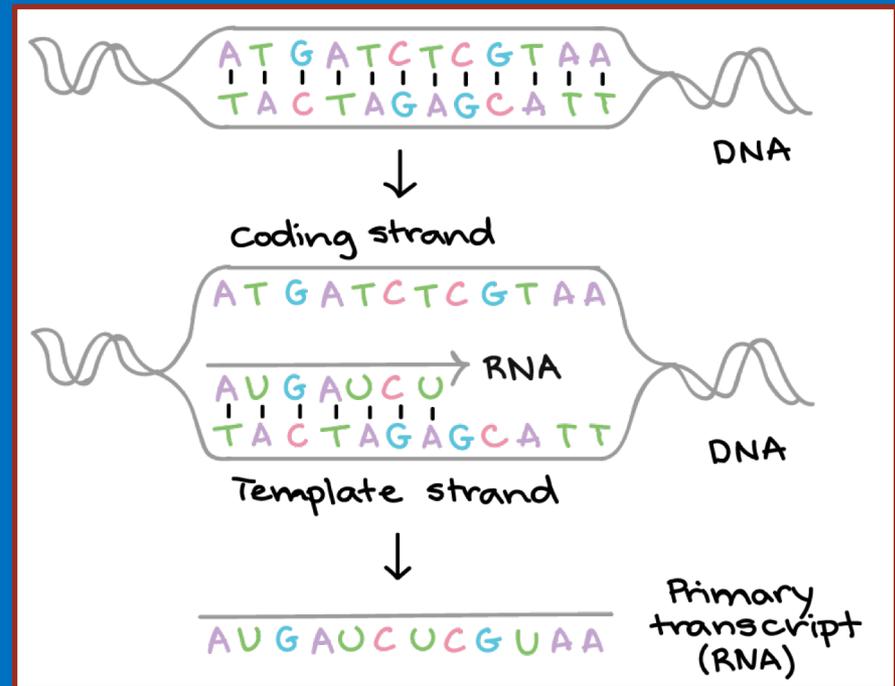
- Uma das funções mais importantes do RNA é transmitir a informação contida no DNA para a síntese de proteínas
- Isto é feito através do RNA mensageiro (mRNA)
- E pela enzima RNA polimerase
- Esta enzima, lê a informação contida no DNA e a converte numa molécula de RNA



A síntese do RNA é feita com base na sequência do DNA

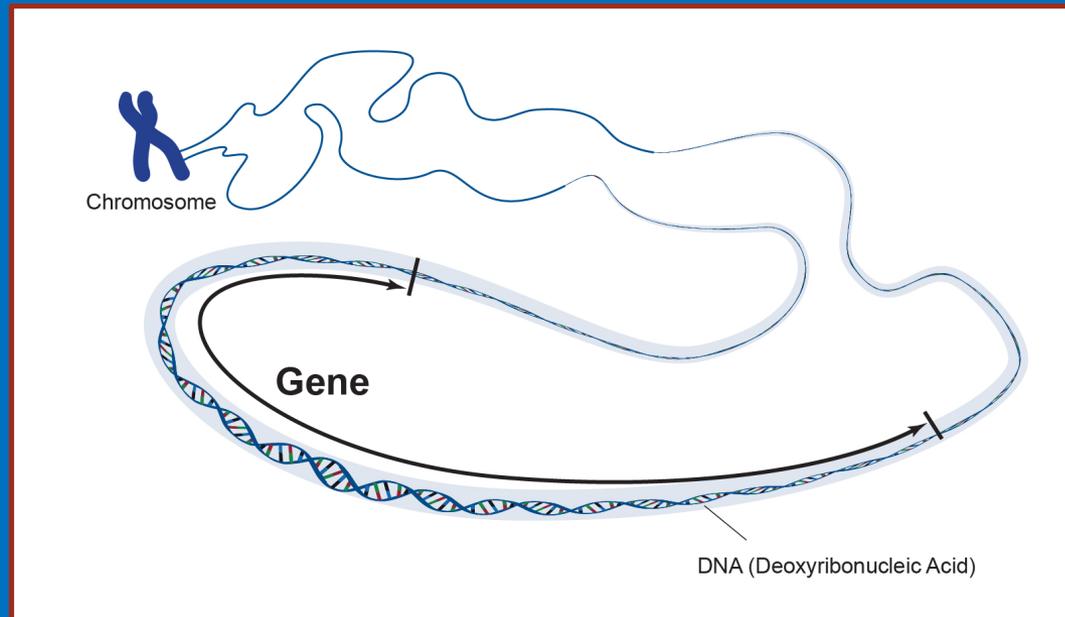
- A síntese da molécula de RNA é feita de acordo com a sequência do DNA
- Para os nucleotídeos C, G e T, o RNA utiliza as mesmas bases do DNA (G, C e A)
- Porém, quando a RNA polimerase encontra um A, ela incorpora um U (uracila)

	DNA	RNA
Pentose sugar	Deoxyribose	Ribose
Base Composition	Adenine (A) Guanine (G) Cytosine (C) Thymine (T)	Adenine (A) Guanine (G) Cytosine (C) Uracil (U)
Number of strands	Double stranded (forms a double helix)	Single stranded



O que é um gene?

- Nem todo o DNA é utilizado pela célula para produzir RNA ou proteína. Acima de tudo, um gene é uma região discreta do DNA (figura abaixo)
- É uma região, unidade, do DNA que quando transferida (hereditariamente) para o(a) filho(a), determina uma característica biológica
- É uma região, unidade, do DNA responsável pela produção de um RNA e/ou proteína
- De forma geral um gene tem uma região regulatória/promotora e uma região contendo a informação. A região regulatória permite controlar quando o gene é expresso



Qual o significado do tamanho de um genoma?

- O conjunto de todo o DNA de um organismo é o seu genoma
- Humanos tem um genoma de 3×10^9 pb (pares de bases) (genoma haploide, ou seja, contando apenas uma cópia de cada cromossomo)
- Uma bactéria *Escherichia coli*, encontrada no nosso intestino, tem um genoma bem menor, de 5×10^6 pb
- Porém, apesar da grande diferença no tamanho do genoma, apenas uma parte do genoma contem genes
- Temos mais ou menos 30.000 genes, que ocupam apenas 1% do nosso genoma. O restante, são regiões estruturais e regulatórias
- Já a bactéria *Escherichia coli* utiliza uma percentagem maior do seu genoma com genes: mais de 90%!

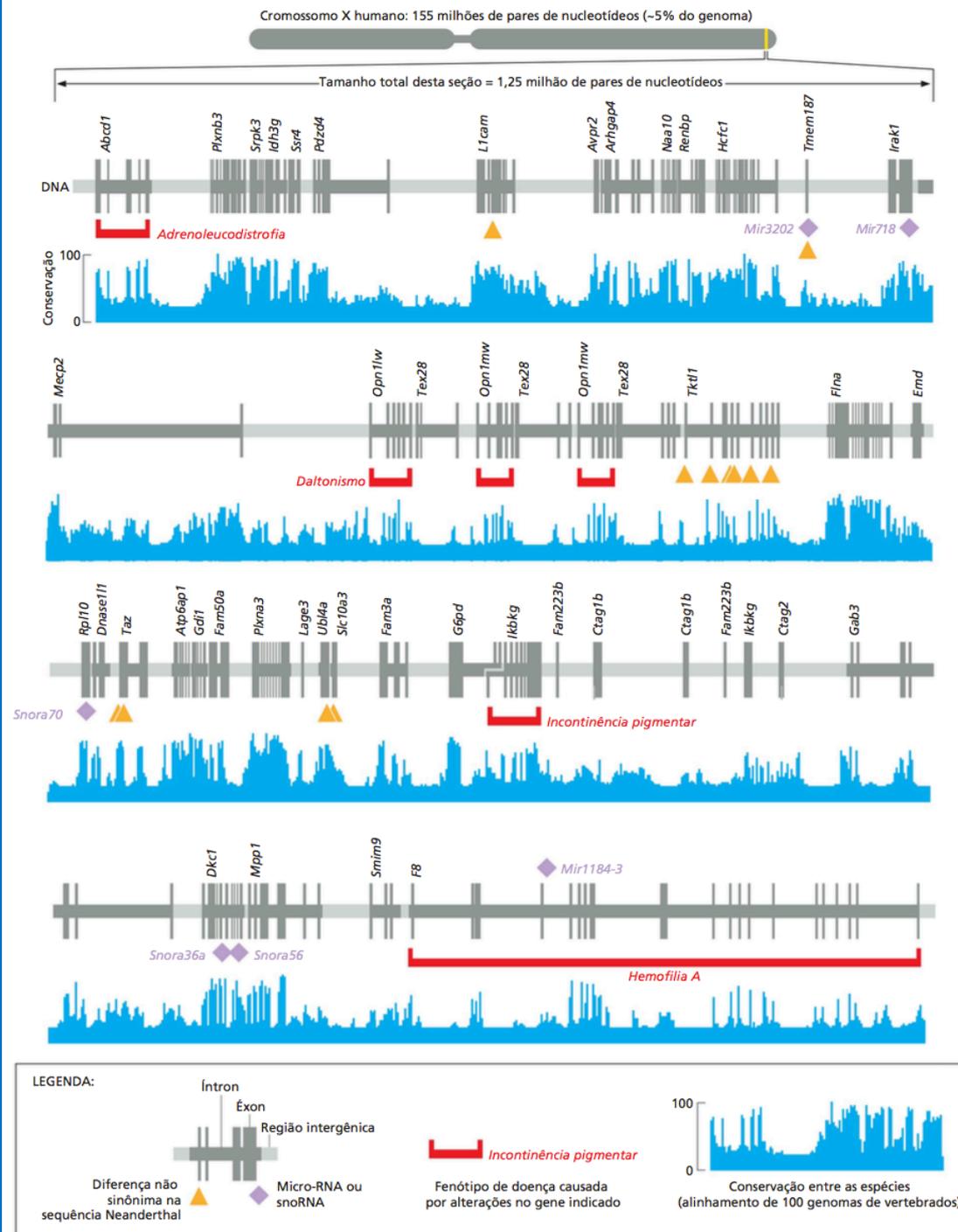
TABELA 1-2 Alguns organismos-modelo e seus genomas

Organismo	Tamanho do genoma* (pares de nucleotídeos)	Número aproximado de genes
<i>Escherichia coli</i> (bactéria)	$4,6 \times 10^6$	4.300
<i>Saccharomyces cerevisiae</i> (levedura)	13×10^6	6.600
<i>Caenorhabditis elegans</i> (verme cilíndrico)	130×10^6	21.000
<i>Arabidopsis thaliana</i> (planta)	220×10^6	29.000
<i>Drosophila melanogaster</i> (mosca-das-frutas)	200×10^6	15.000
<i>Danio rerio</i> (peixe-zebra)	1.400×10^6	32.000
<i>Mus musculus</i> (camundongo)	2.800×10^6	30.000
<i>Homo sapiens</i> (humano)	3.200×10^6	30.000

*O tamanho do genoma inclui uma estimativa para a quantidade de seqüências de DNA altamente repetidas que não estão nos bancos de dados de genomas.

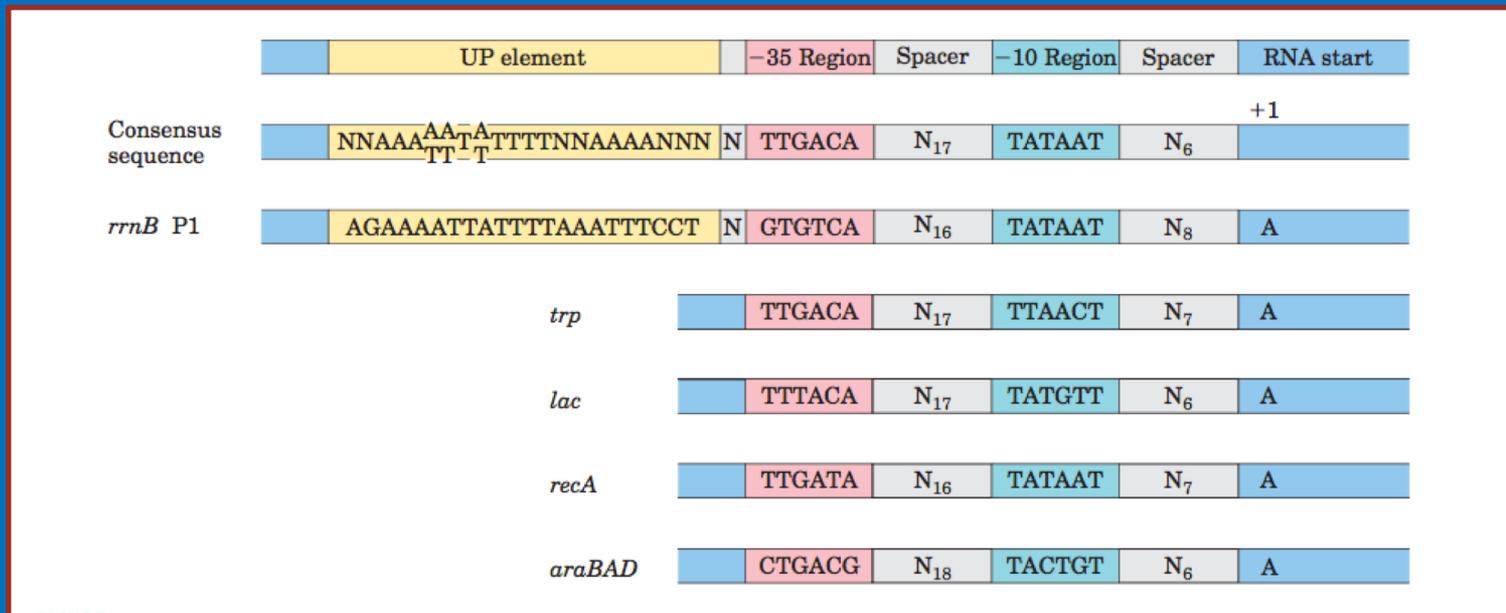
O que é um gene

- Ao lado está a representação de um pequena região do cromossomo X humano
- Estão indicados os genes, que geralmente recebem abreviações do tipo *Abcd1*, *Plxnb3* etc
- Em azul, está indicada a conservação do DNA entre espécies. Isto é, se a sequência do DNA nesta região é mais parecida ou diferente da de outros animais (espécies)
- Note que um gene humano é composto de exons e introns (vamos discutir isto logo adiante)
- Entre um gene e outros, temos as regiões intergênicas (cinza claro)
- Em vermelho, está indicados genes importante para algumas doenças



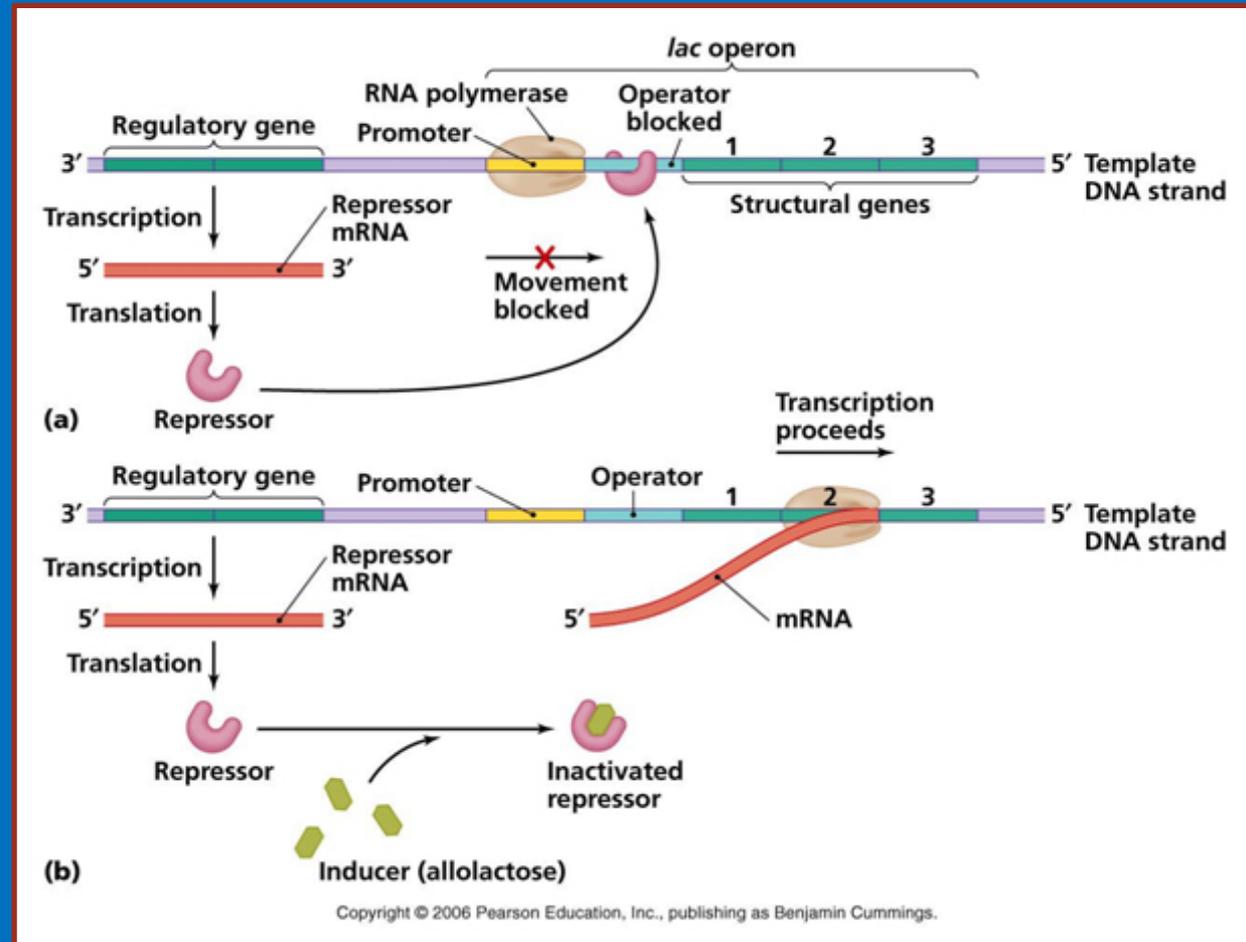
Os fatores de transcrição e a regulação gênica

- Bactérias utilizam fatores de transcrição do tipo **sigma** para controlar a expressão de vários genes. Abaixo, está ilustrado um promotor típico para o fator sigma-70
- Todos os promotores apresentam regiões com seqüências do DNA conservadas (p.ex. TATA-box) que se encontram em posições específicas. Por exemplo, o TATA-box (região rica em T e A) localiza-se a -10 pares de bases antes do início da transcrição
- Estas regiões são importantes para que o fator de transcrição se ligue e recrute a RNA polimerase



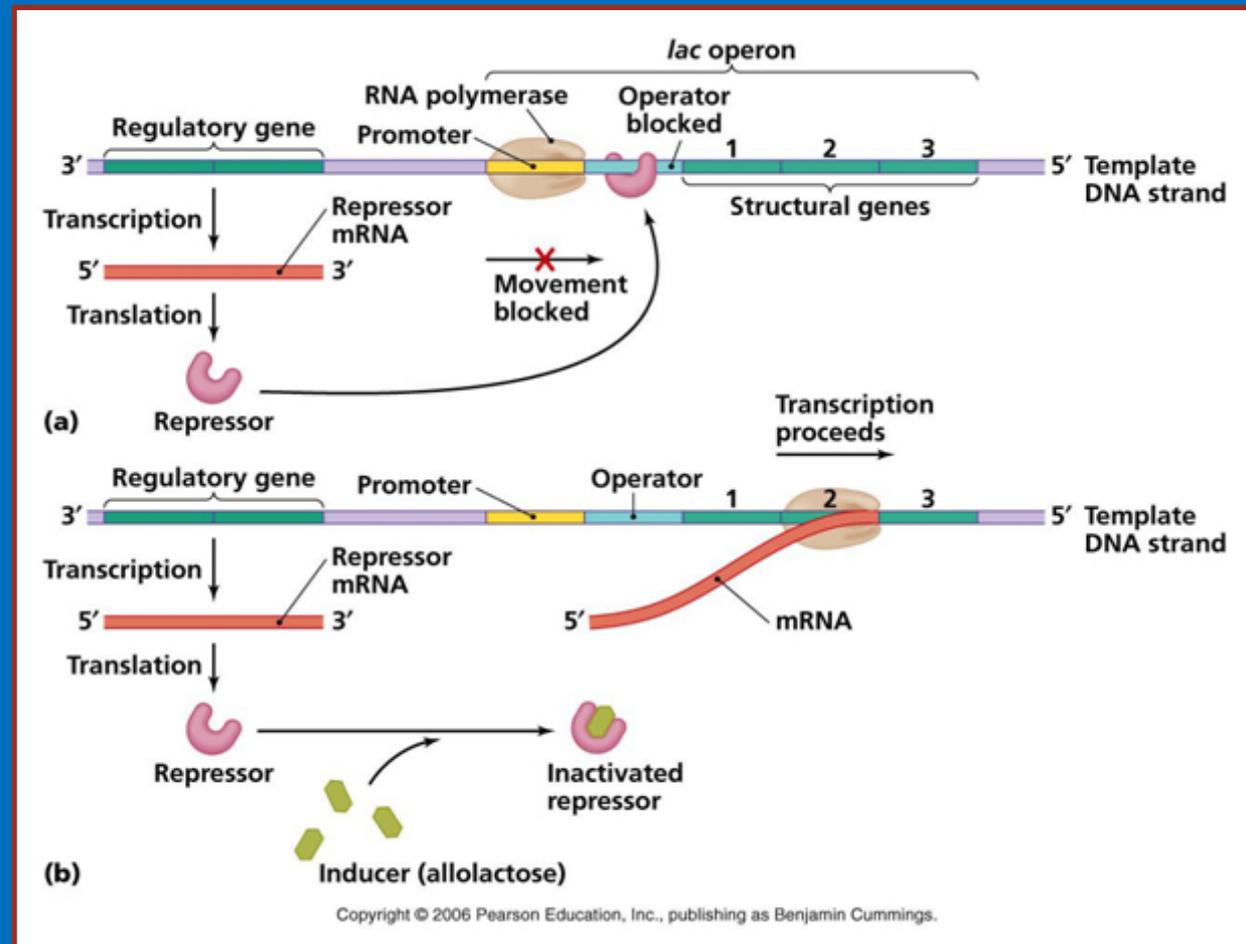
O início da transcrição em procariotos

- Mas como controlar a expressão de um gene?
- O operon lac é um bom exemplo do controle da expressão gênica em bactérias
- Quando não há o açúcar lactose no meio, a bactéria não precisa produzir as enzimas que metabolizam este açúcar
- Por isso, os genes (representados como 1, 2 e 3 – verde) não precisam ser transcritos em mRNA



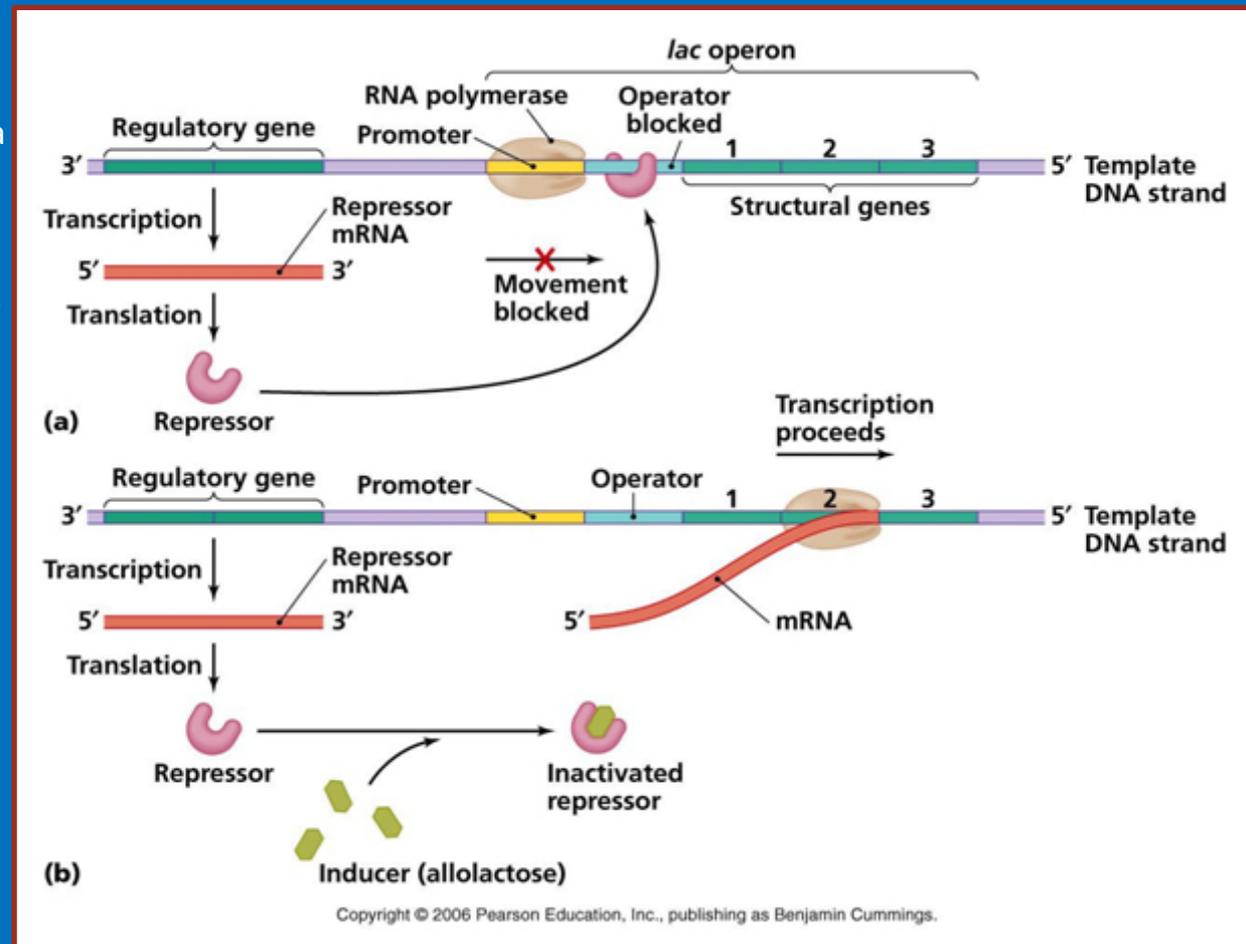
O início da transcrição em procariotos

- Neste caso, a proteína repressora, codificada pelo gene regulador, liga-se ao promotor e impede que a RNA polimerase leia o DNA e produza o mRNA
- Porém, na presença de lactose, esta se liga a proteína repressora, impedido que se ligue ao promotor
- Note que o gene da proteína repressora é CONSTITUTIVO. Isto é, este gene está sempre ligado, produzindo mRNA e a proteína repressora



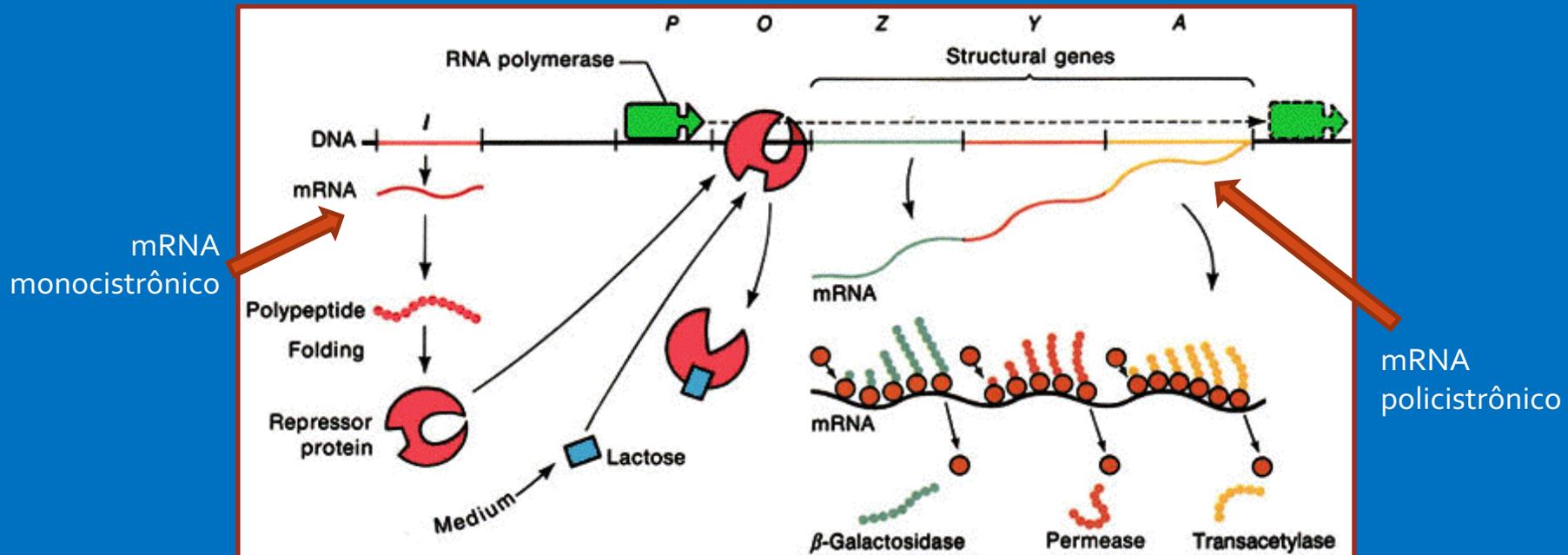
O início da transcrição em procariotos

- Já os genes 1, 2 e 3 do operon lac, são induzidos pela presença da lactose
- Ou seja, sem lactose no meio, o mRNA não é produzido
- O termo operon se deve pelo fato das bactérias produzirem um único mRNA para mais de uma proteína
- Em organismos eucariotos, isto é diferente.



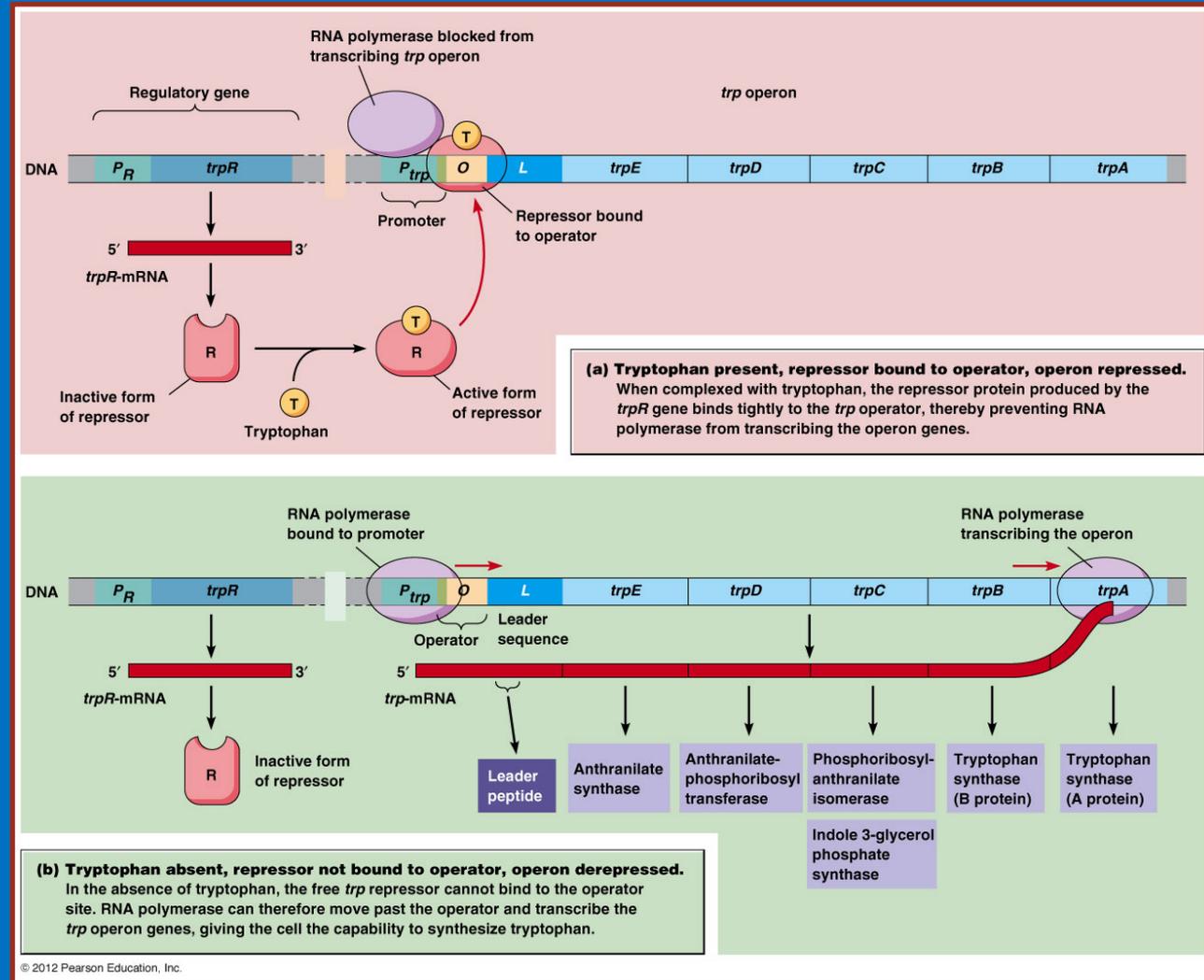
O mRNA de procariotos pode ser policistrônico

- O operon Lac ilustra outra característica importante do muitos RNA mensageiros (mRNA) de organismos procariotos (bactérias): eles podem conter mais de um gene
- Desta forma, bactérias "economizam" e utilizam o mesmo promotor para expressar mais de um gene, e produzir mais de uma proteína. Geralmente, estes genes estão correlacionados na função e importância para a célula
- Portanto, dizemos que o mRNA é monocistrônico, quando ele carrega a informação para produzir apenas uma proteína; e policistrônico, quando ele contém informação para produzir mais de uma proteína



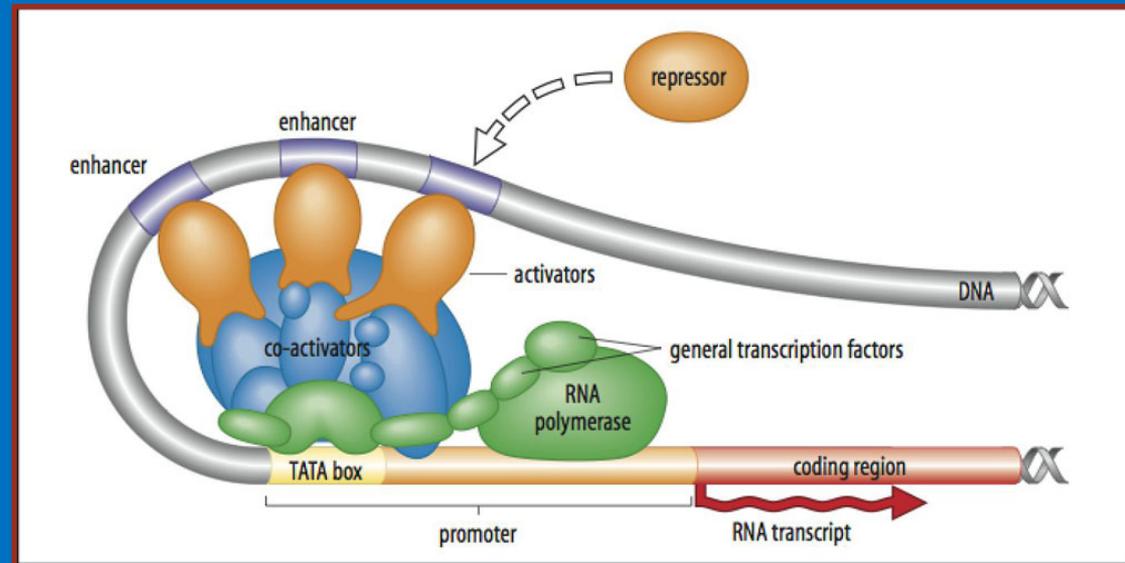
Os fatores de transcrição e a regulação gênica

- No caso dos genes para a síntese do aminoácido triptofano, ocorre o oposto
- Quando há um excesso deste aminoácido, o triptofano se liga a proteína repressora
- Isto faz com que o repressor fique ativo e se ligue ao DNA, inibindo a transcrição
- Quando acaba o triptofano, a proteína repressora é inativada, ativando a expressão gênica



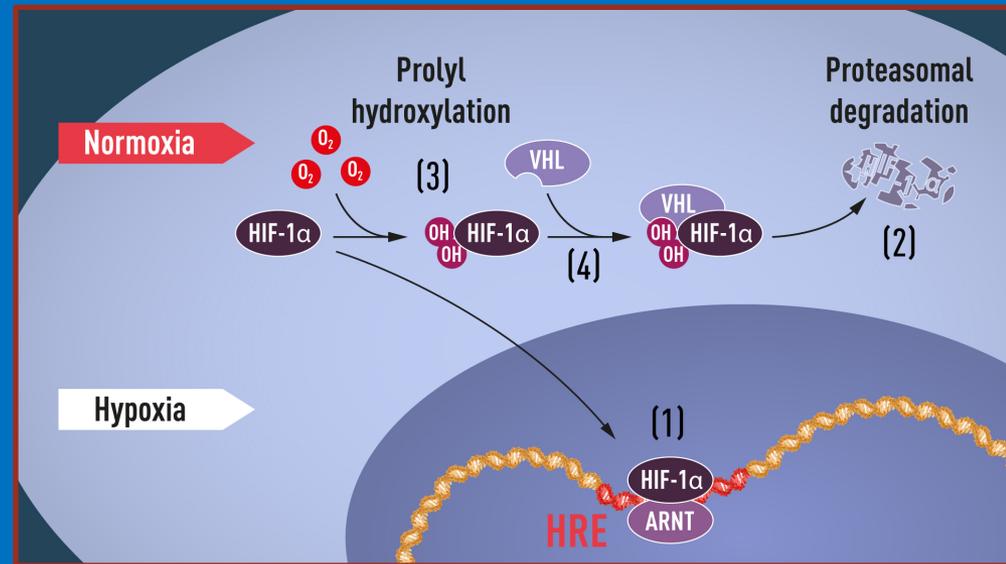
Os fatores de transcrição e a regulação gênica em eucariotos

- Da mesma forma que organismos procaríotos, os organismos eucariotos não precisam utilizar todos os genes ao mesmo tempo, nem em todas as células ou tecidos
- Porém, a regulação gênica em organismos eucariotos é bem mais complexa
- Uma diferença importante é que temos um promotor para cada gene
- Nós (e todos os organismos eucariotos) não produzimos mRNA policistrônico



Regulação gênica em eucariotos: resposta ao oxigênio

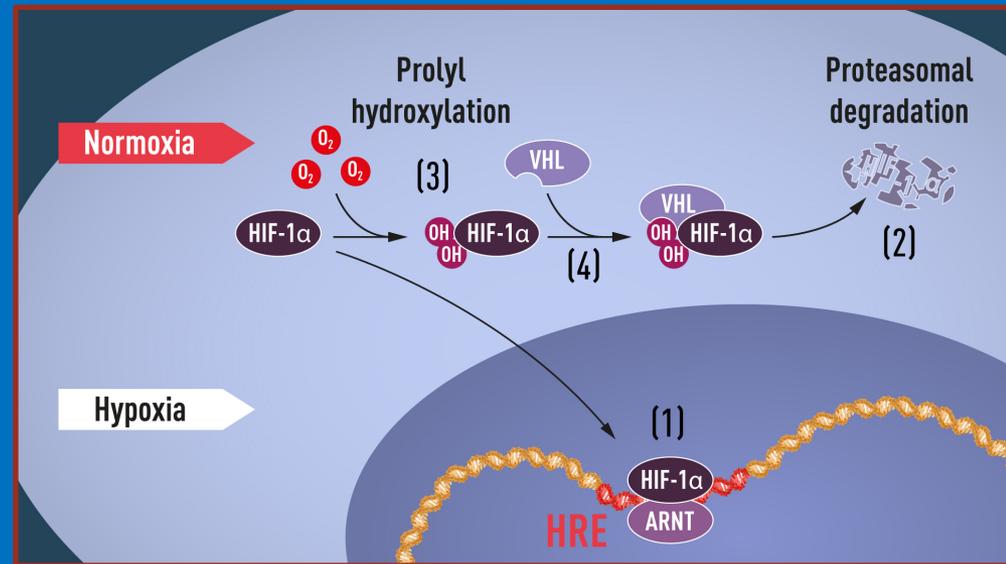
- A regulação gênica em eucariotos é bem mais complexa
- Vamos ver dois exemplos
- Um deles é a resposta a oxigênio, mecanismo que foi desvendado, em boa parte, por 3 pesquisadores
- Eles receberam o prêmio Nobel em Medicina e Fisiologia em 2019 pela suas contribuições



www.nobelprize.org/prizes/medicine/2019/advanced-information/

Regulação gênica em eucariotos: resposta ao oxigênio

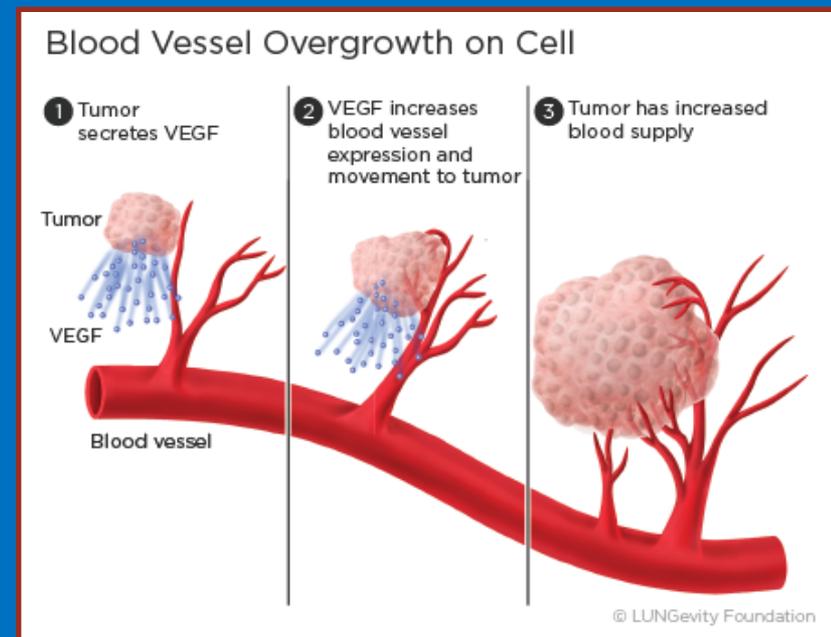
- Nossas células produzem, constantemente (constitutivamente), uma proteína chamada HIF (hypoxia inducible fator)
- Porém, o HIF é hidroxilado por uma enzima chamada PHD (prolina hidroxilase), e isto faz com que o HIF seja imediatamente degradado
- Porém, se faltar oxigênio (hipóxia), não haverá substrato (O_2) para a PHD hidroxilar o HIF
- Desta forma, ele não é degradado e consegue entrar no núcleo da célula
- Como ele é um fator de transcrição, ele se liga a vários genes relacionados com a falta de oxigênio, ativando-os
- Ou seja, a célula começa a produzir mRNA e proteínas importantes para combater a



www.nobelprize.org/prizes/medicine/2019/advanced-information/

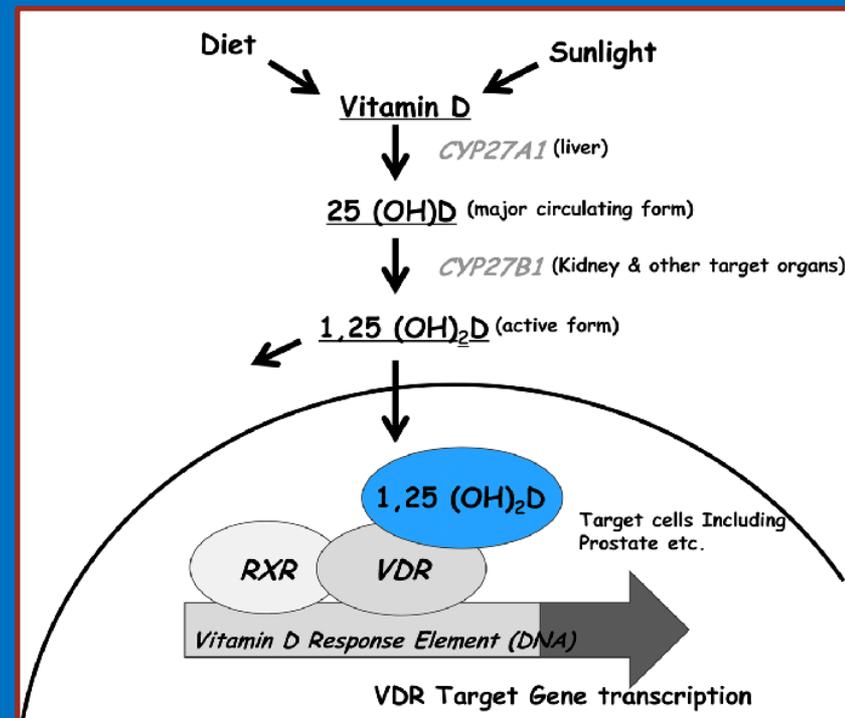
Transcrição gênica e angiogênese

- Entre estes genes, estão enzimas do metabolismo, por exemplo, durante um exercício físico intenso (correr uma maratona)
- E até mesmo, em casos de hipóxia muito forte, a indução da angiogênese (formação de vasos sanguíneos)
- Isto permite que mais oxigênio seja levado para os tecidos com hipóxia
- Este mesmo mecanismo é utilizado pelo câncer
- Quando um tumor cresce, a falta de oxigênio induz a formação de vasos sanguíneos para alimentar o tumor
- Por isso, dizemos que os tumores são dependentes da angiogênese



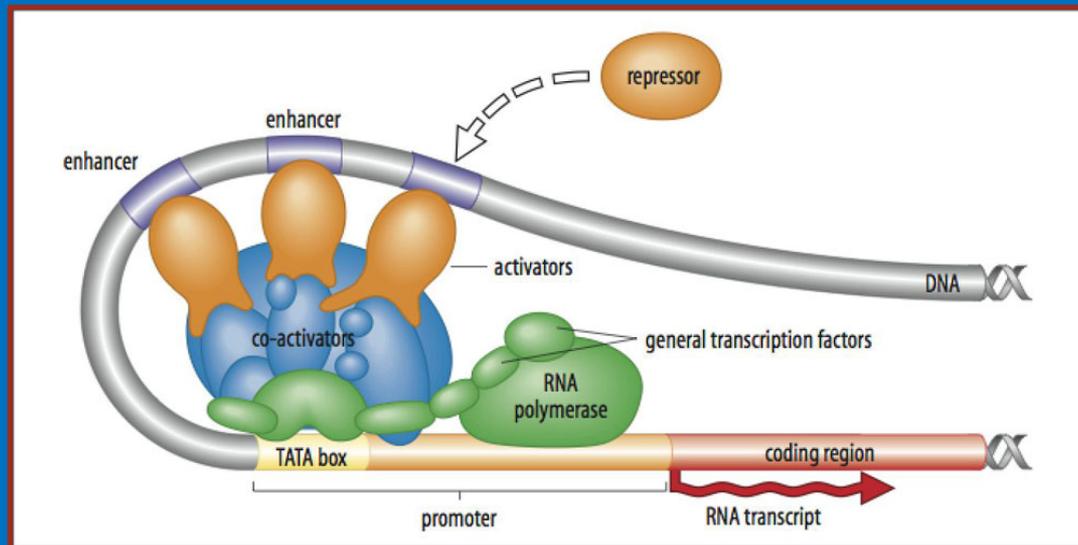
Regulação gênica em eucariotos: vitamina D

- Outro exemplo interessante é o caso da vitamina D
- A vitamina D é um importante ativador da expressão gênica
- A vitamina D é metabolizada no organismo e convertida em 1-alfa, 25-hidroxi-vitamina D₃
- O 1,25(OH)₂D é a forma ativa da vit. D
- O 1,25(OH)₂D entra no núcleo e liga-se um fator de transcrição chamado de "VDR" (vitamin D receptor)
- O complexo 1,25(OH)₂D/VDR liga-se ao promotor (chamado de "vitamin D responsive elements), recruta a RNA polimerase e inicia a transcrição gênica
- Em resumo, genes que tiverem um promotor de vitamina D, irão ser ativados na presença desta vitamina



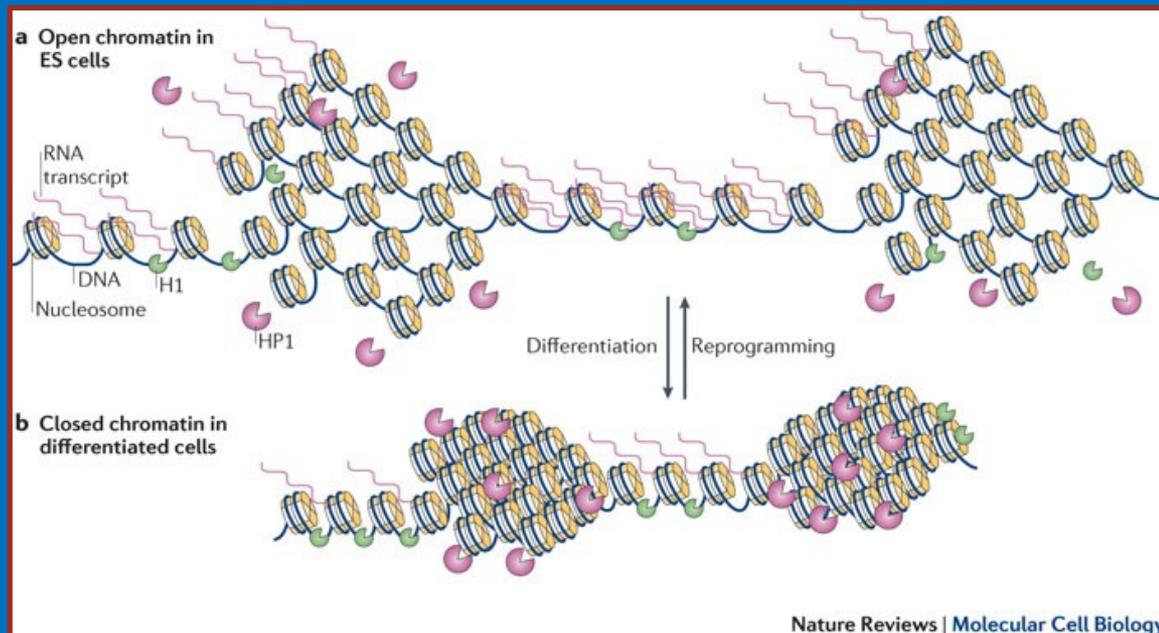
A transcrição gênica em eucariotos é bastante complexa

- Por isso, diferentes fatores de transcrição são produzidos, ligando e desligando genes específicos
- Quando fatores de transcrição se ligam ao DNA, eles atraem a RNA polimerase, iniciando a transcrição
- Porém, como vocês podem ver na imagem abaixo, existem outros elementos, como proteínas repressoras que podem inibir a transcrição;
- Certas regiões do DNA, próximas ao promotor podem aumentar a expressão gênica (enhancers)
- Porém, não iremos nos preocupar com estes detalhes neste curso



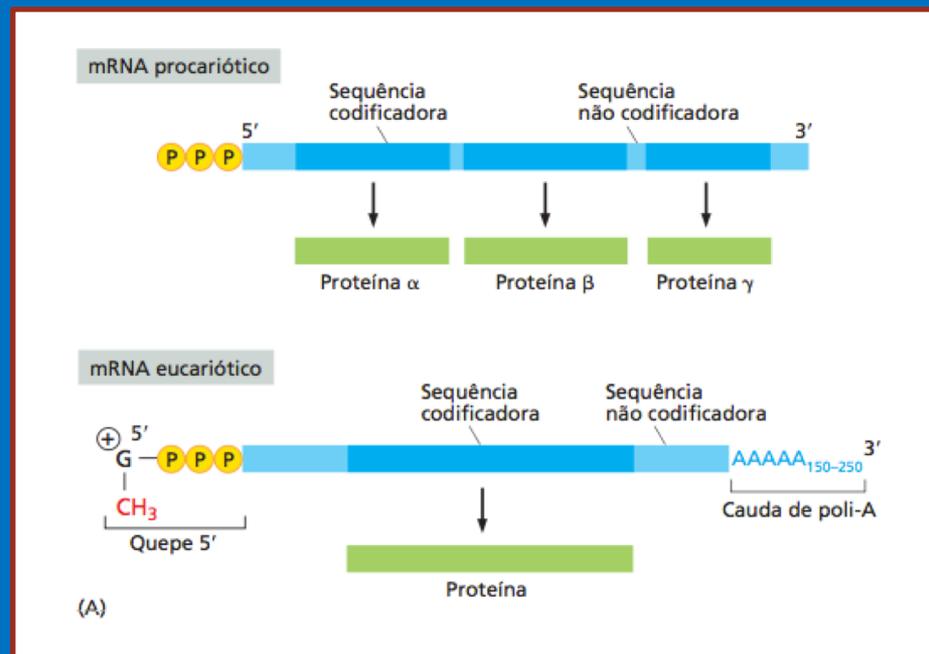
Cromatina: eucromatina e heterocromatina

- Outro aspecto importante que regula a transcrição gênica é o estado da cromatina
- Quando o DNA está todo empacotado ele é chamado de heterocromatina
- Genes dentro da heterocromatina não são acessíveis aos fatores de transcrição e à RNA polimerase. Por isso, não são transcritos
- Já na eucromatina, os nucleossomos estão mais afastados, permitindo o acesso dos fatores de transcrição ao promotor, assim como da RNA polimerase



O mRNA

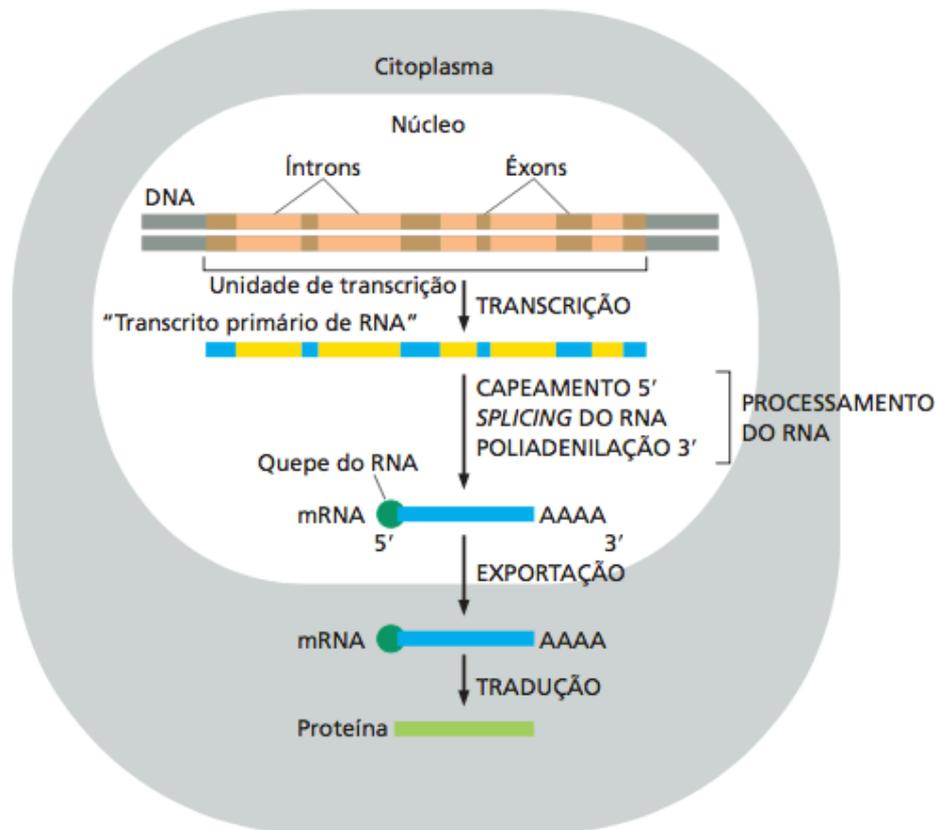
- Como mencionamos, diferentemente das bactérias, nos produzimos um mRNA para cada gene
- Bactérias produzem mRNA policistrônicos, contendo a informação para produzir mais de uma proteína
- Nosso mRNA contém a informação para produzir apenas uma proteína
- Ele é também modificado com a adição de uma cauda poli-A (contendo uma série de adeninas) e quepe de 5-metil-guanina, uma base modificada. Todas estas adições são importantes para indicar que este é um mRNA maduro, pronto para ser utilizado para a síntese proteica



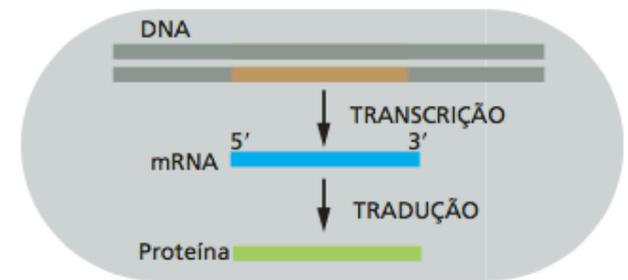
Introns e exons

- A adição da cauda poli-A e do quepe de metil-guanosina indicam que o RNA foi processado e os introns removidos
- O que são introns e exons?

(A) EUCARIOTOS



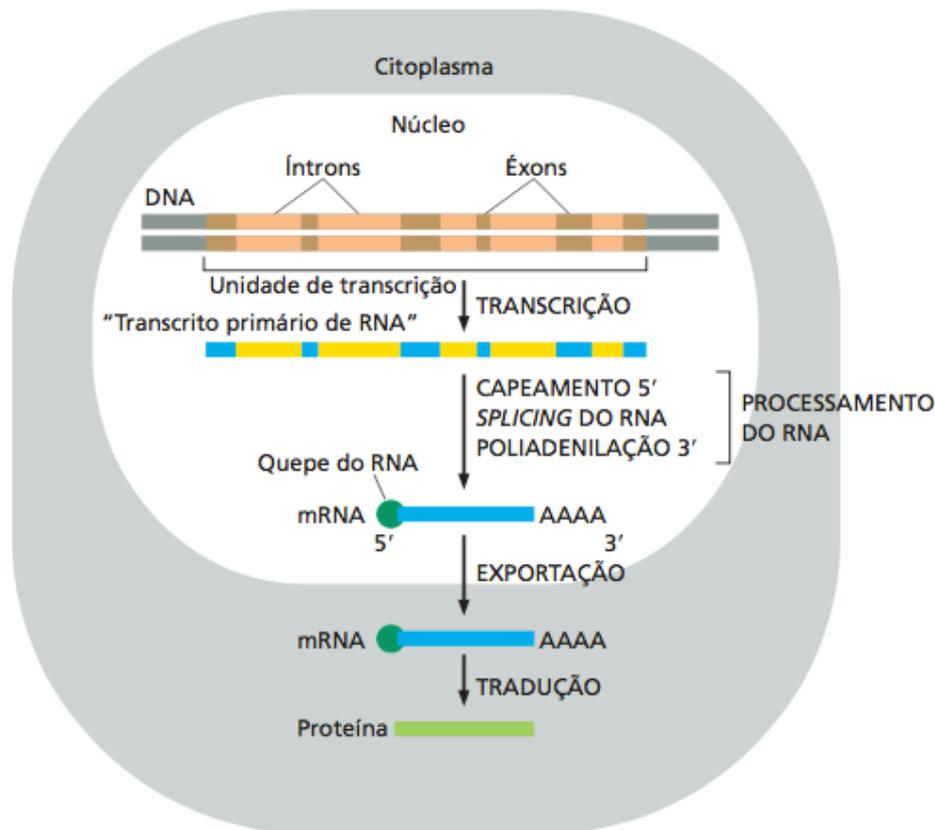
(B) PROCARIOTOS



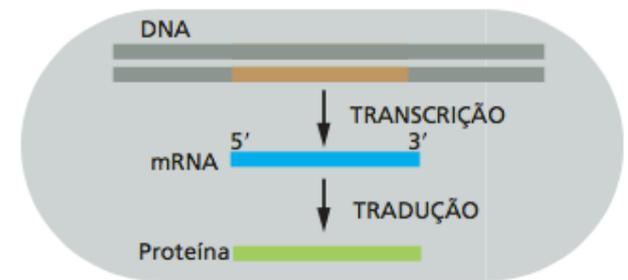
Introns e exons

- Diferentemente do gene de procaríotos, os genes de eucariotos contém sequências que não são utilizadas para produzir proteína: são os introns
- Apenas os exons contém informação genética utilizada para produzir uma proteína
- Por isso, os introns precisam ser removidos

(A) EUCARIOTOS



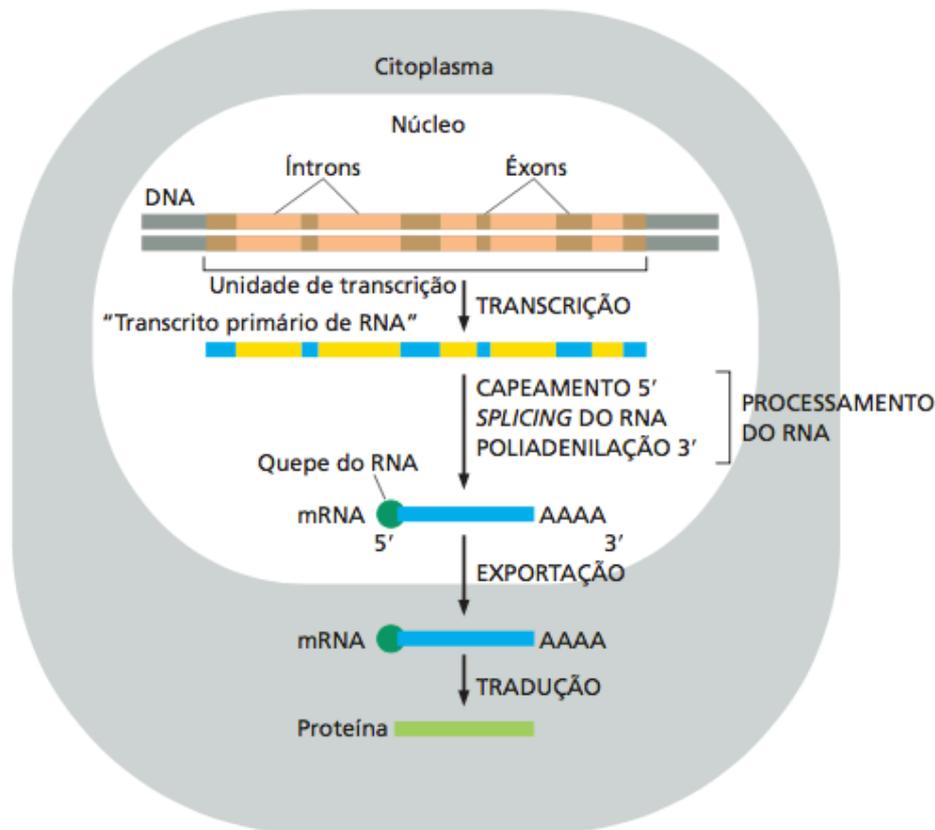
(B) PROCARIOTOS



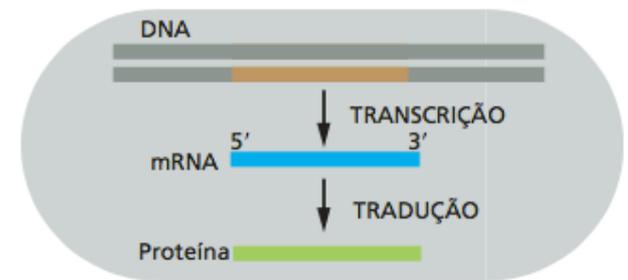
Introns e exons

- Este processo para remover os introns é chamado de "splicing", ou processamento do RNA
- Mas por que temos introns? Por que este desperdício de DNA?

(A) EUCARIOTOS



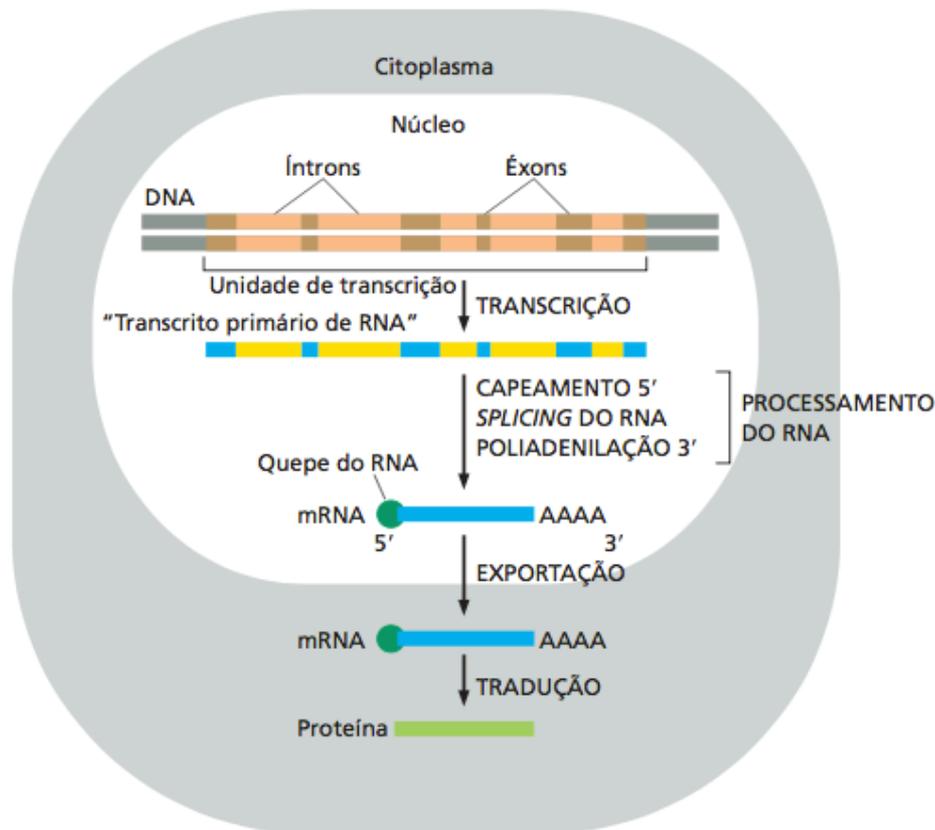
(B) PROCARIOTOS



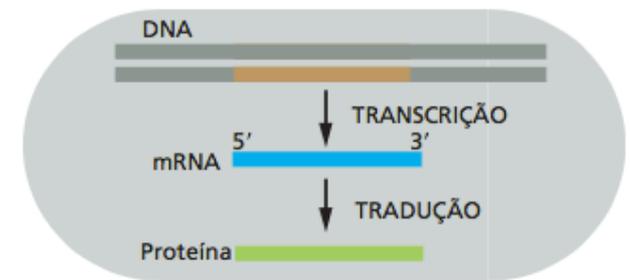
Introns e exons

- Este processo para remover os introns é chamado de "splicing", ou processamento do RNA
- O splicing é realizado pela splicessomo, um conjunto de proteína que reconhece o intron e o remove

(A) EUCARIOTOS

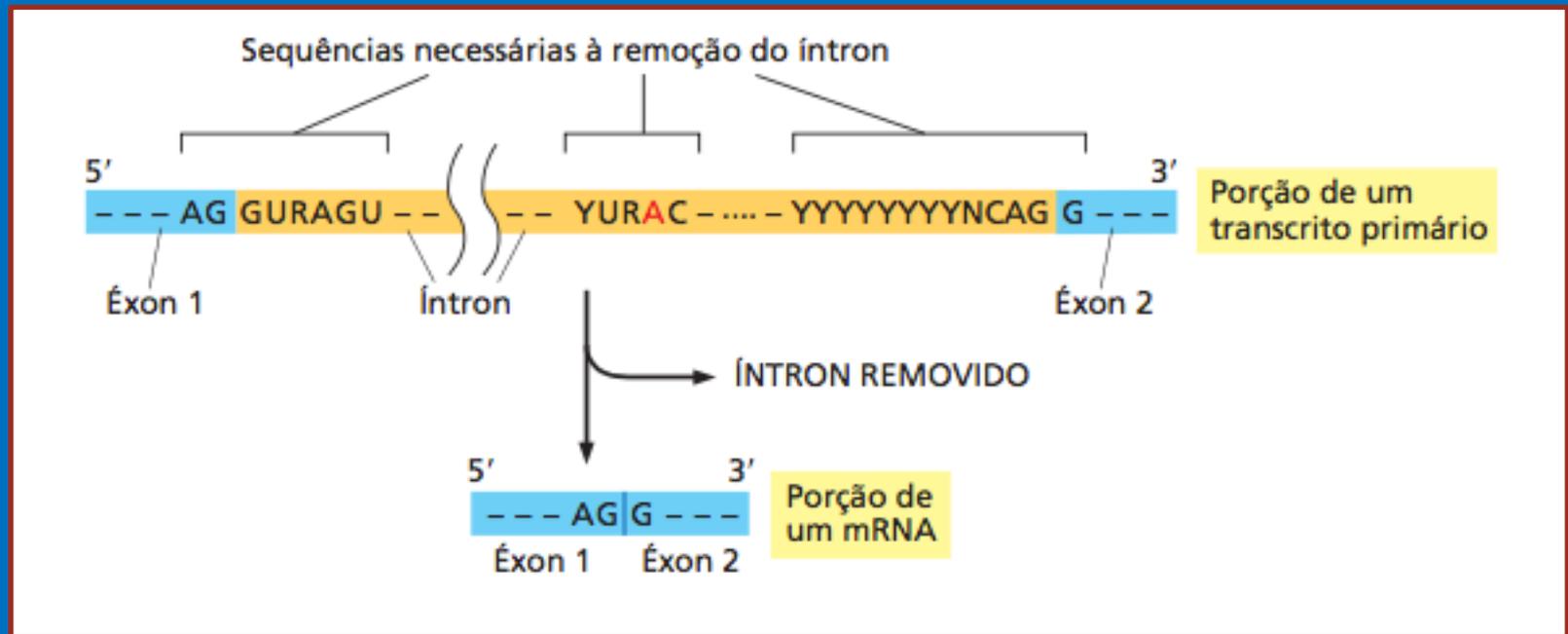


(B) PROCARIOTOS



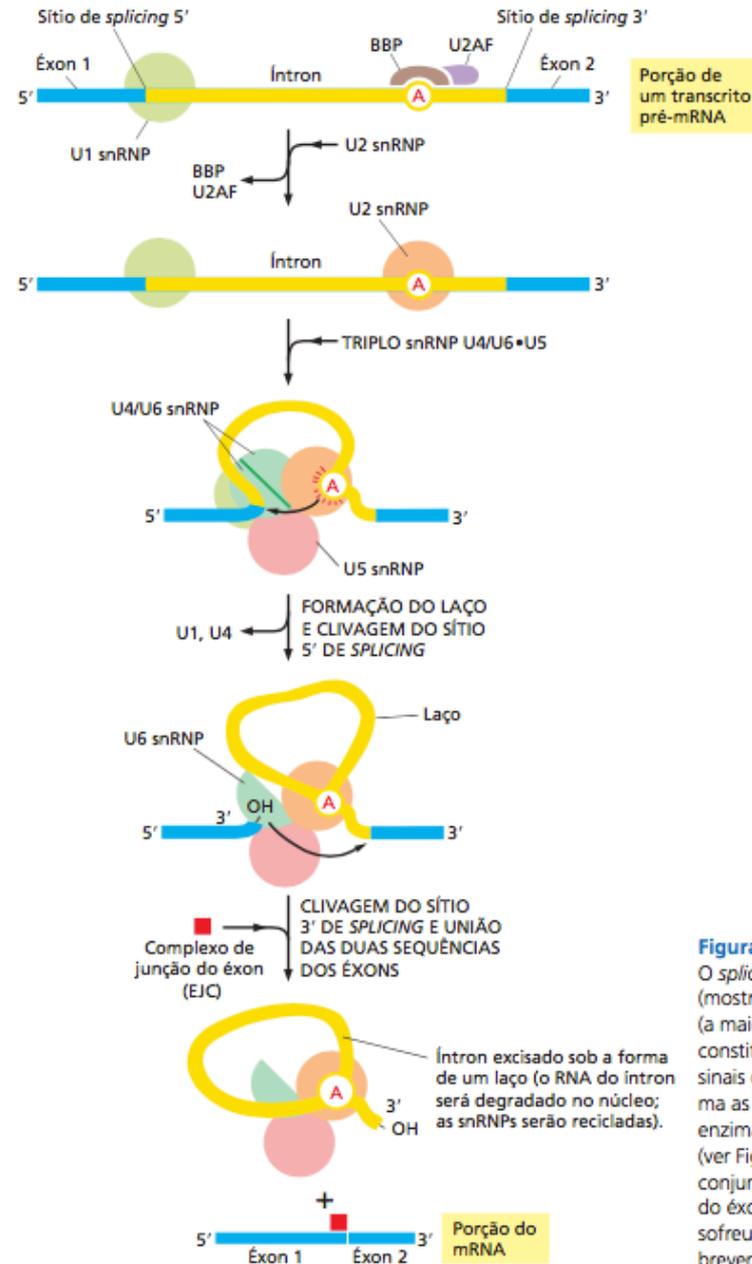
Introns e exons

- Sequências de DNA específicas permitem que o spliceossomo reconheça o introns (figura abaixo; R = purina [A ou G] e Y = representa uma pirimidina [C ou T])



Introns e exons

- O spliceossomo é composto por proteínas e pequenos RNAs guias (U₁, U₂, etc)
- Eles ajudam as proteínas a reconhecer as sequências do intron e fazer a remoção na região correta
- Por isso, são chamados de RNA guia, ou seja, indicam para o complexo de splicing onde se ligar e remover o intron



A snRNP U1 pareia com a junção 5' de *splicing* (ver Figura 6-29) e a BBP (proteína de ligação ao ponto de forquilha) e o U2AF (fator auxiliar de U2) reconhecem o sítio do ponto de forquilha.

A snRNP U2 desloca BBP e U2AF e forma pares de bases com a sequência consenso do sítio do ponto de forquilha.

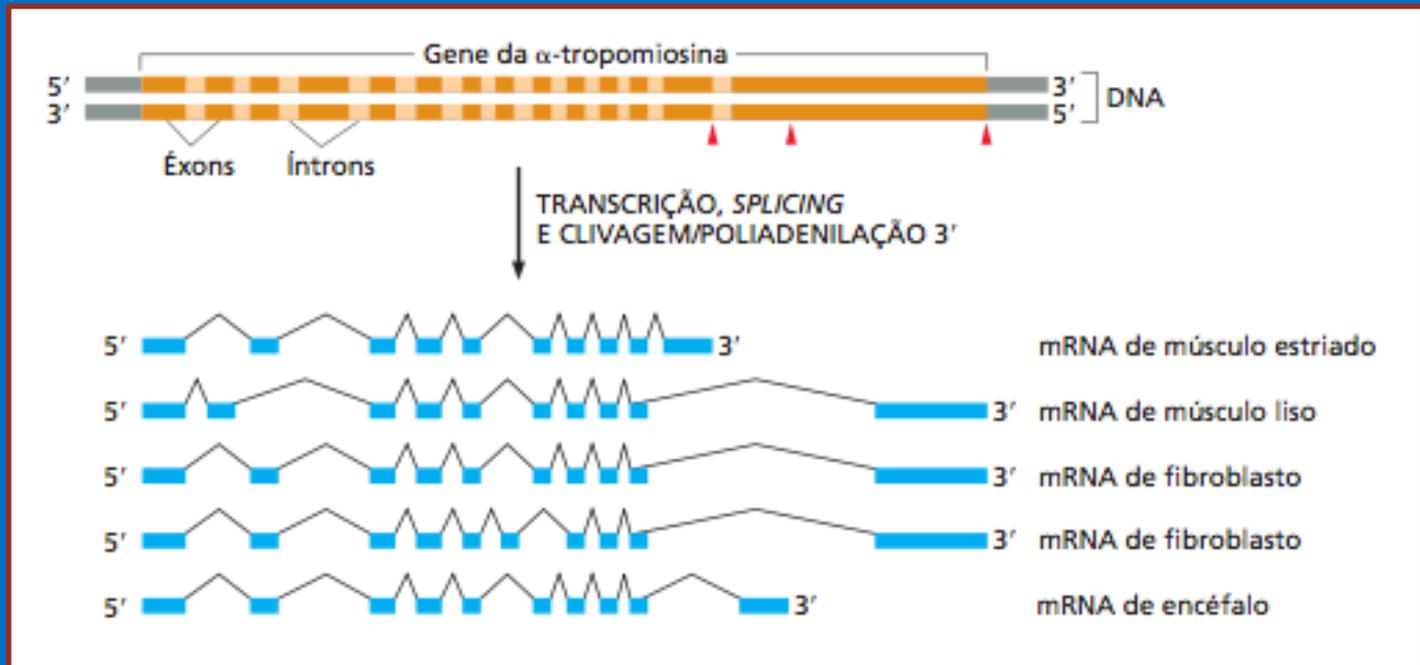
A tripla snRNP U4/U6•U5 entra na reação. Na tripla snRNP, os snRNAs U4 e U6 são mantidos em firme união pelas interações de pares de bases. Os rearranjos subsequentes quebram os pareamentos de U4/U6 permitindo que U6 desloque U1 na junção 5' de *splicing* (ver Figura 6-29). Isso cria um sítio ativo que catalisa a primeira reação fosforiltransferase.

Rearranjos RNA-RNA adicionais criam o sítio ativo para a segunda reação fosforiltransferase, a qual, então, completa o *splicing* (ver Figura 6-25A).

Figura 6-28 Mecanismo de *splicing* do pré-mRNA. O *splicing* do RNA é catalisado por um arranjo de snRNPs (mostrado como círculos coloridos) e de outras proteínas (a maioria das quais não é mostrada) que, em conjunto, constituem o spliceossomo. O spliceossomo reconhece os sinais de *splicing* em uma molécula de pré-mRNA, aproxima as duas extremidades dos introns e fornece a atividade enzimática para as duas etapas necessárias da reação (ver Figura 6-25A e Animação 6.5). Como indicado, um conjunto de proteínas do chamado complexo de junção do éxon (EJC) permanece sobre a molécula de mRNA que sofreu *splicing*; o seu papel subsequente será discutido brevemente.

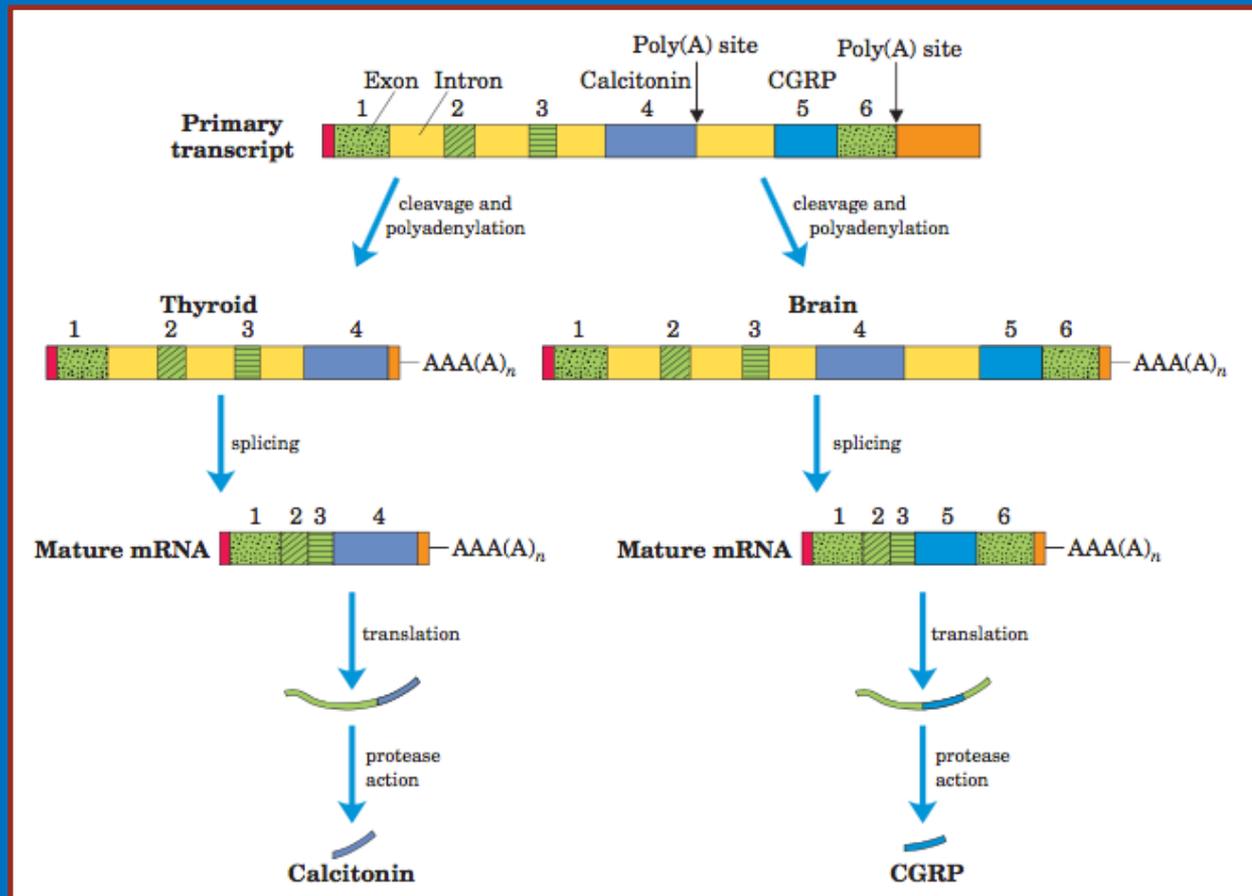
Mas, por que temos introns e exons?

- Por que nossos genes têm introns e exons? Parece um desperdício de DNA...
- Na verdade, isto permite que um mesmo gene possa codificar proteínas "diferentes"
- Durante o processamento de splicing, exons podem ser removidos também: é o chamado "splicing alternativo" onde exons podem ou não ser incluídos no mRNA final
- Desta forma, um mesmo gene pode produzir diferentes mRNAs
- Notem no exemplo abaixo, que o mesmo genes (da tropomiosina alfa, pode produzir tropomiosinas "diferentes" no músculo, parênquima (fibroblasto) e no cérebro



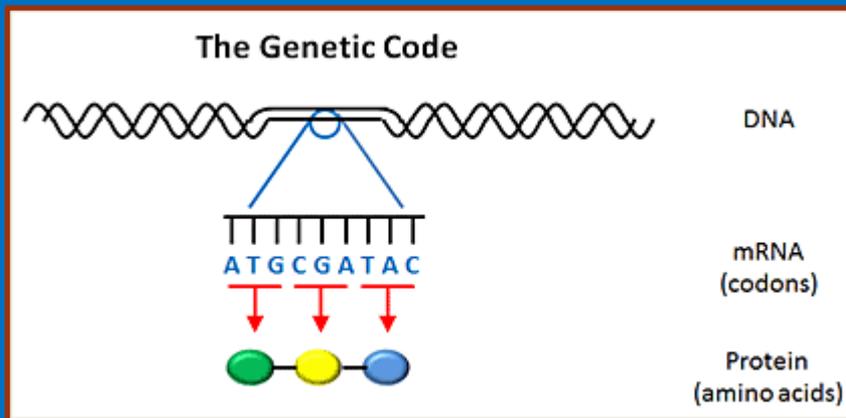
Mas, por que temos introns e exons?

- Por exemplo, o gene da calcitonina, pode produzir dois hormônios por splicing alternativo
- Na glândula tireoide, produz-se a calcitonina. Já no cérebro, é produzido o CGRP (peptídeo relacionado a calcitonina). Tudo a partir do mesmo gene



DNA e a informação genética

- O RNA carrega a informação genética para que a célula sintetize as proteínas correspondentes.
- Cada três bases correspondem a um codon, que correspondem a um aminoácido.

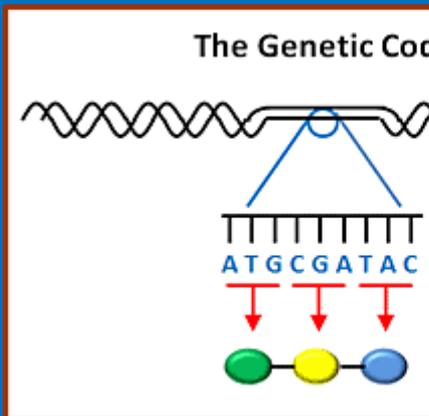


		Second letter				Third letter
		U	C	A	G	
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA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

Mas iremos discutir isto com mais calma na próxima aula

- O RNA carrega a inform
- Cada três bases corres

respondentes.

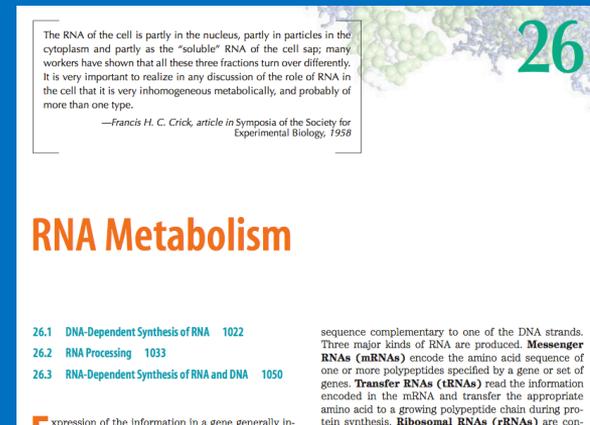
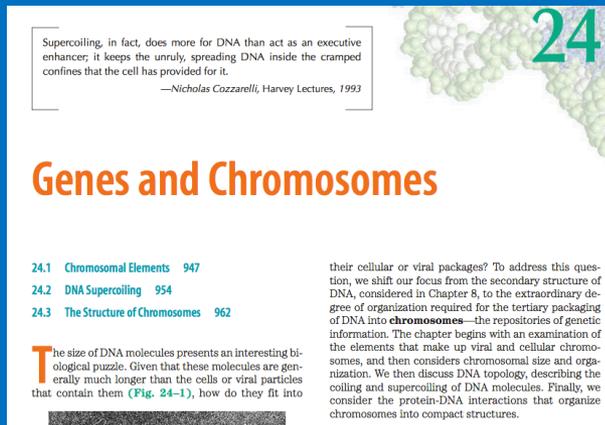


		cond letter		
		A	G	
r	UAU } Tyr	UGU } Cys	U C A G	
	UAC } Tyr	UGC } Cys		
	UAA Stop	UGA Stop		
	UAG Stop	UGG Trp		
o	CAU } His	CGU } Arg	U C A G	
	CAC } His	CGC } Arg		
	CAA } Gln	CGA } Arg		
	CAG } Gln	CGG } Arg		
r	AAU } Asn	AGU } Ser	U C A G	
	AAC } Asn	AGC } Ser		
	AAA } Lys	AGA } Arg		
	AAG } Lys	AGG } Arg		
G	GUC } Val	GCC } Ala	U C A G	
	GUA } Val	GCA } Ala		
	GUG } Val	GCG } Ala		
		GAA } Glu		GGA } Gly
	GAG } Glu	GGG } Gly		

Third letter

Bibliografia

- Lehninger – Capítulo 24 – Genes e cromossomos e Capítulo 26 – Metabolismo RNA (5 ed.)
- Alberts (Biologia Molecular da Célula) – Capítulo 6 – Como as células leem o genoma (6 ed.)



Como as células leem o genoma: do DNA à proteína

CAPÍTULO
6

NESTE CAPÍTULO

- DO DNA AO RNA
- DO RNA À PROTEÍNA
- O MUNDO DE RNA E A ORIGEM DA VIDA

Desde a descoberta da estrutura do DNA no início de 1950, os progressos em biologia celular e molecular têm sido surpreendentes. Atualmente, são conhecidas as sequências genômicas completas de milhares de organismos diferentes, o que nos revelou detalhes fascinantes da sua bioquímica, bem como indícios importantes sobre a forma como esses organismos evoluíram. Também foram determinadas as sequências completas do genoma de milhares de seres humanos individuais, bem como de alguns de nossos parentes agora extintos, como os neandertais. O conhecimento da quantidade máxima de informação necessária para que se produza um organismo complexo, como um ser humano, nos permite identificar os limites das características bioquímicas e estruturais das células e deixa claro que a biologia não é infinitamente complexa.

Como discutido no Capítulo 1, o DNA genômico não controla a síntese de proteínas diretamente, mas utiliza o RNA como intermediário. Quando a célula requer uma proteína específica, a sequência de nucleotídeos da região apropriada de uma molécula de DNA ex-