



Universidade de São Paulo
Instituto de Química

QBQ0317 – 2020

Aula 4

Estrutura de Genes e Genomas

O que é gene?

1. *Pedaço de DNA que contém uma informação para uma característica específica*
2. *É uma sequência de bases nitrogenadas do DNA com localização (locus) específica no cromossomo, e que contém informação para a produção de proteínas*
3. *É a região com bases específicas que codifica e produz proteínas*
4. *É a unidade responsável pelo armazenamento da informação sobre uma proteína. As proteínas em conjunto determinam as características de um indivíduo.*
5. *Gene é o material genético contido no DNA*
6. *É uma proteína que carrega a informação genética*
7. *Gene é toda informação contida no DNA*
8. *Gene é um trecho do DNA responsável pela realização de um conjunto de funções/fenótipos*

Algumas das respostas da avaliação prévia

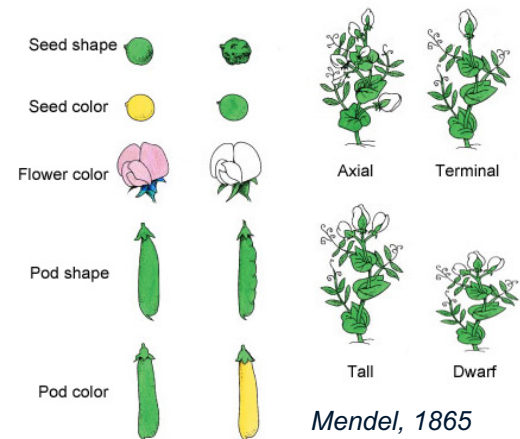
O que é gene?

1. Pedaco de **DNA** que contém uma **informação** para uma **característica** específica
2. É uma **sequência** de bases nitrogenadas do **DNA** com **localização** (locus) específica no cromossomo, e que contém **informação** para a produção de **proteínas**
3. É a **região** com bases específicas que codifica e produz proteínas
4. É a unidade responsável pelo armazenamento da **informação** sobre uma **proteína**. As **proteínas** em conjunto determinam as **características** de um indivíduo.
5. Gene é o material genético contido no **DNA**
6. É uma **proteína** que carrega a **informação** genética
7. Gene é toda **informação** contida no **DNA**
8. Gene é um **trecho** do **DNA** responsável pela realização de um conjunto de **funções/fenótipos**

Algumas das respostas da avaliação prévia

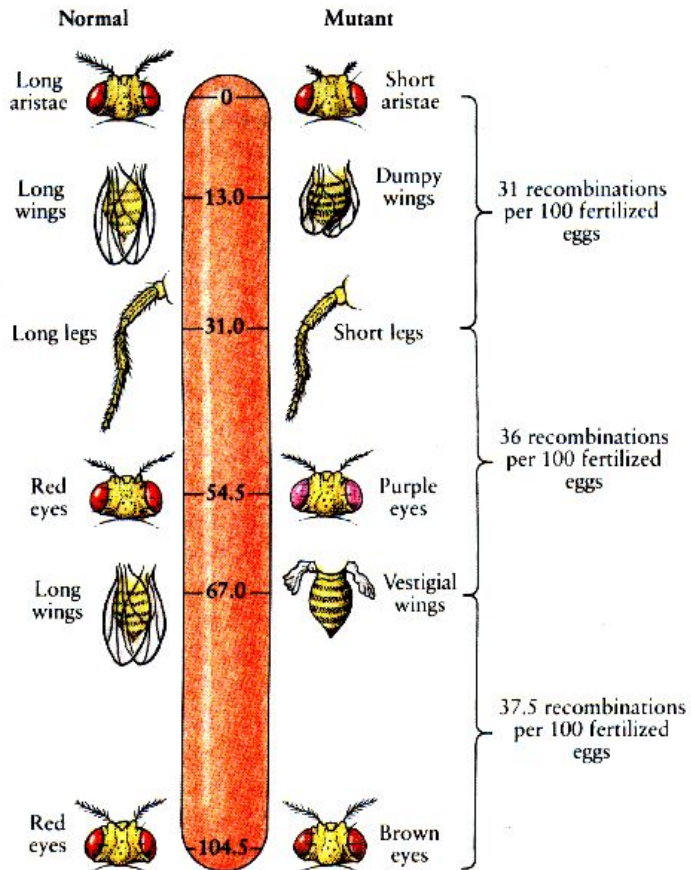
O que é um gene?

- Herdado entre gerações:
 - cada descendente tem uma cópia física do material (**genótipo**)
- Informação sobre estrutura, função e outros atributos biológicos de quem o carrega (confere um **fenótipo**)



1909 - Wilhelm Johannsen (botânico dinamarquês)

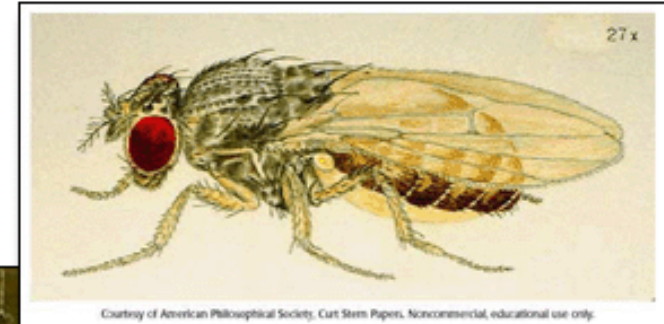
Genes estão arranjados de maneira linear nos cromossomos



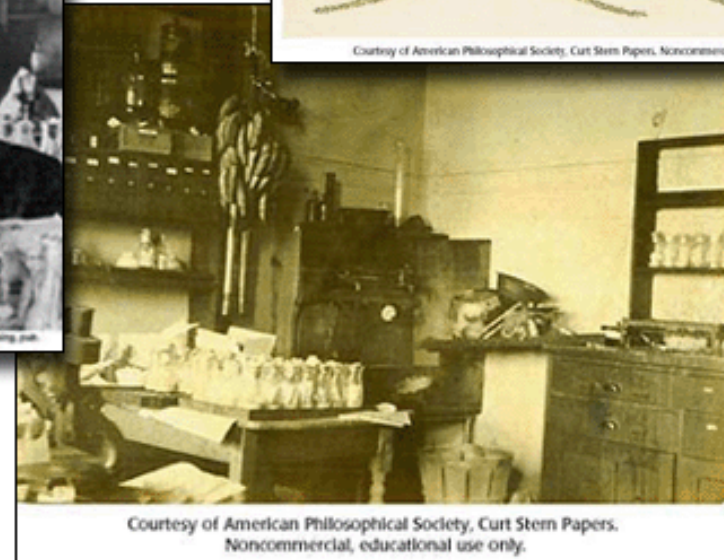
Portrait of Morgan at Work



Early Fly drawing



Courtesy of American Philosophical Society, Curt Stern Papers. Noncommercial, educational use only.



Courtesy of American Philosophical Society, Curt Stern Papers. Noncommercial, educational use only.

Photo of Morgan's fly room at Columbia University

1941 – Thomas Morgan

O que é um gene do ponto de vista molecular?

Estudos que começaram a relacionar a Bioquímica com a Genética



Bases para o início da Biologia Molecular

O que é um gene do ponto de vista molecular?

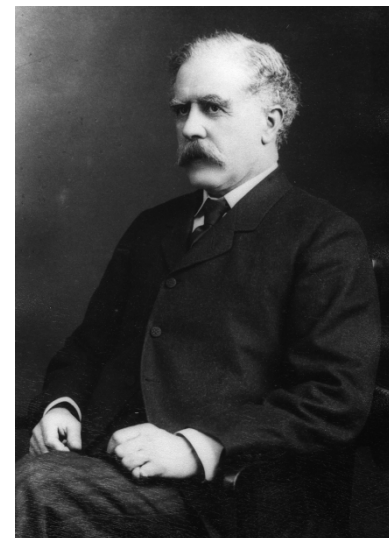
- *Archibald Garrod, 1902*
- *Beadle and Tatum, 1941*
 - http://www.dnalc.org/resources/nobel/beadle_tatum.html

O conceito de gene foi e ainda está evoluindo com novos experimentos e conhecimentos

O que é um gene do ponto de vista molecular?

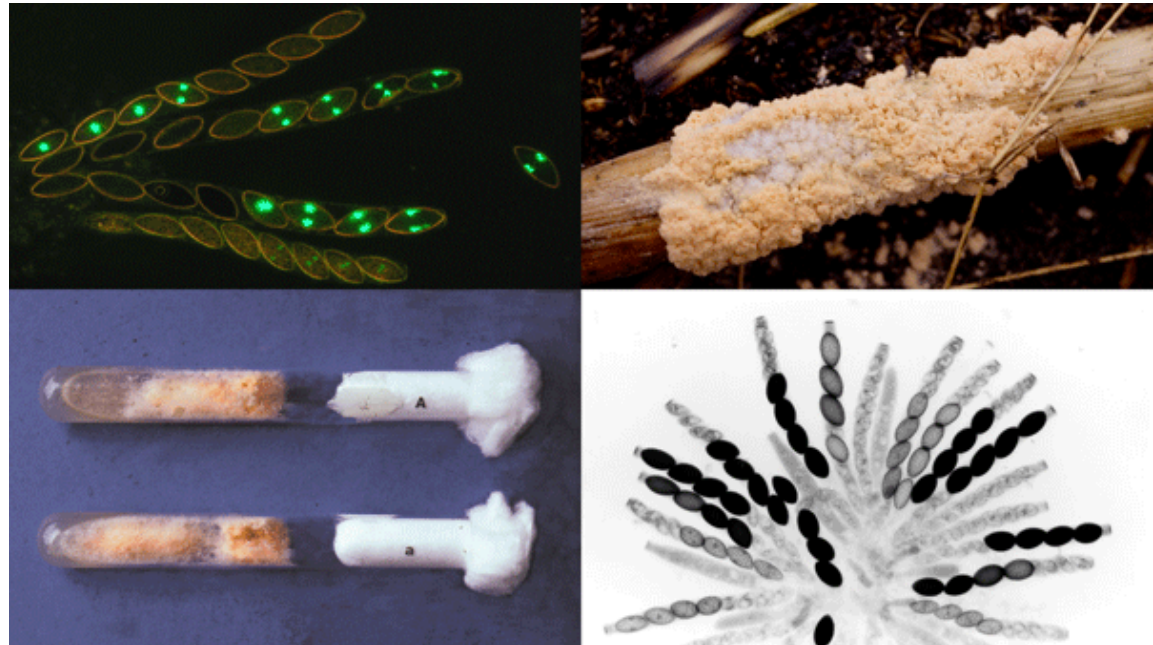
- Erros inatos do metabolismo
 - *Archibald Garrod, 1902*
 - alcaptonúria como uma doença genética recessiva
 - Urina de cor escura
 - Mutações no gene HGD (homogentisato 1,2-dioxigenase)
 - Metabolismo da tirosina

Owing, as I believe, to their chemical individuality different human beings differ widely in their liability to individual maladies, and to some extent in the signs and symptoms which they exhibit.

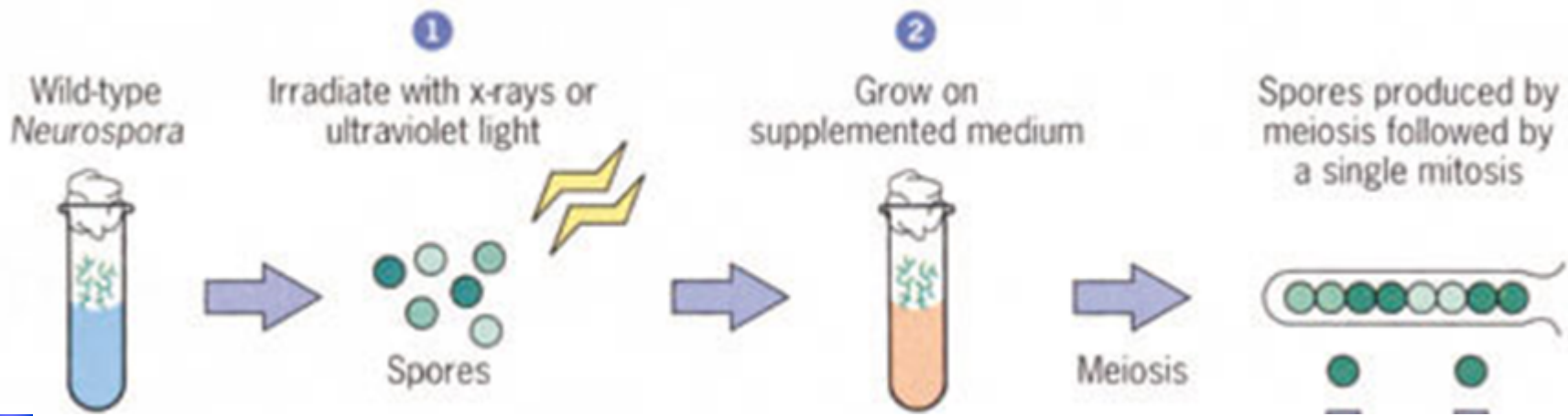


Neurospora crassa

- Bolor do pão
- Esporos haplóides
- “Genética Bioquímica”
 - Beadle and Tatum: um gene, uma enzima



Beadle e Tatum - mutantes de *Neurospora*



Meio mínimo:
Contém apenas
açúcar e sais

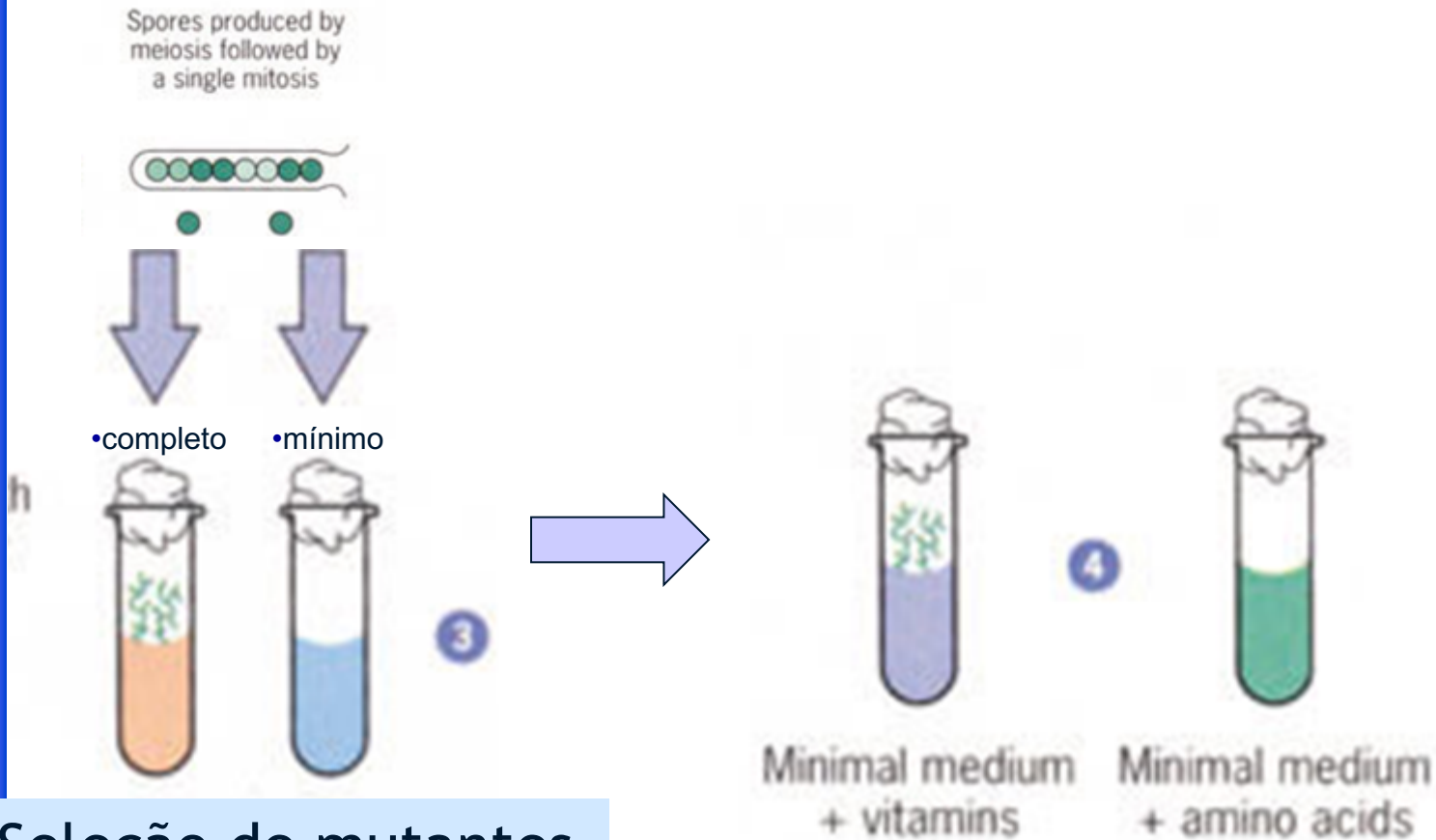
Neurospora tem que
produzir todos os
aminoácidos e vitaminas
para crescer

SELVAGEM

Meio completo:
Contém todos os
aminoácidos e
vitaminas

MUTANTES
(auxotróficos)
também crescem

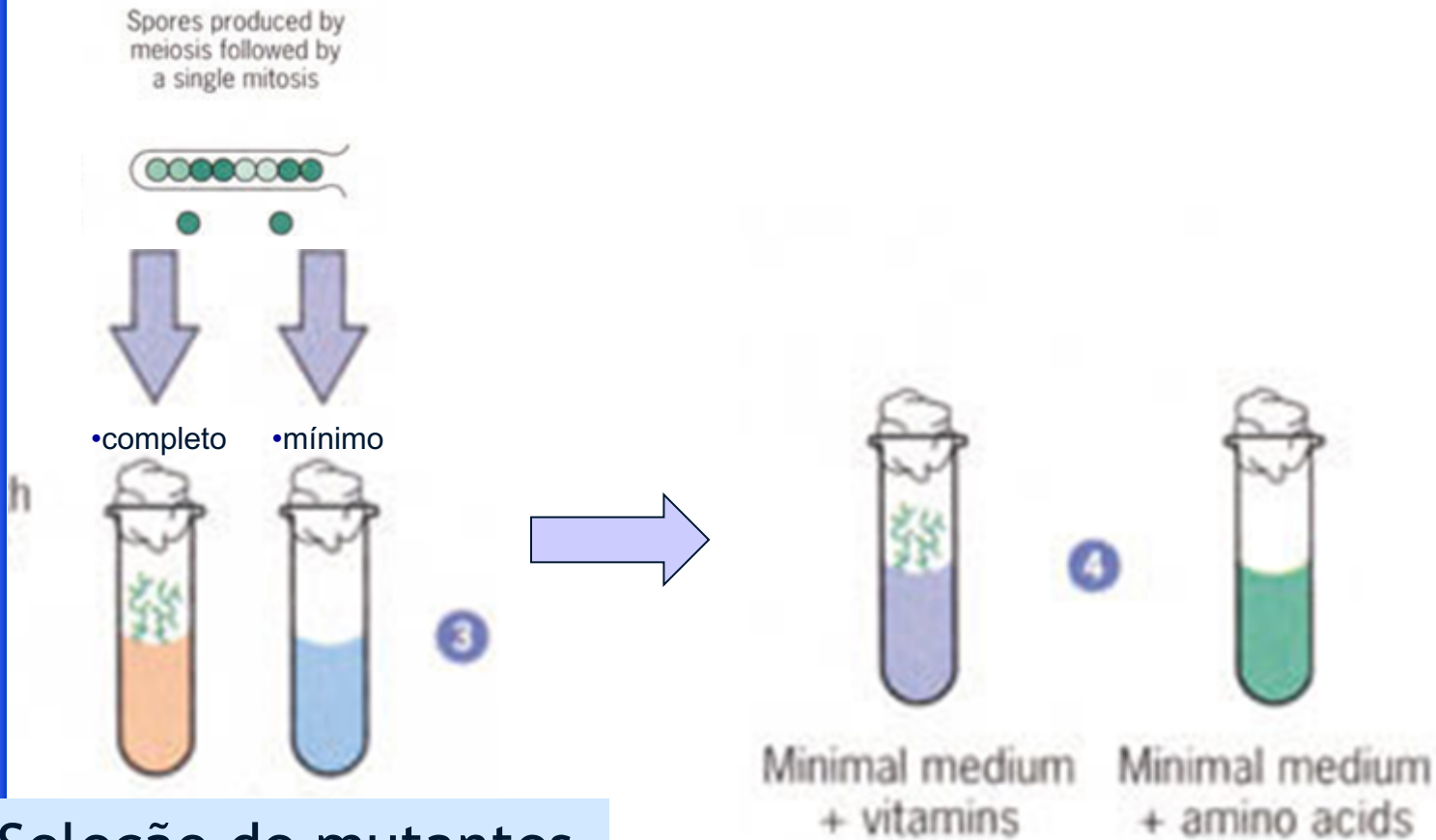
Beadle e Tatum - mutantes de *Neurospora*



Seleção de mutantes que **não** crescem em meio mínimo (crescem no meio completo)

Crescimento apenas em meio mínimo + vitaminas

Beadle e Tatum - mutantes de *Neurospora*



Seleção de mutantes que **não** crescem em meio mínimo (crescem no meio completo)

Crescimento apenas em meio mínimo + vitaminas

Beadle e Tatum - mutantes de *Neurospora*



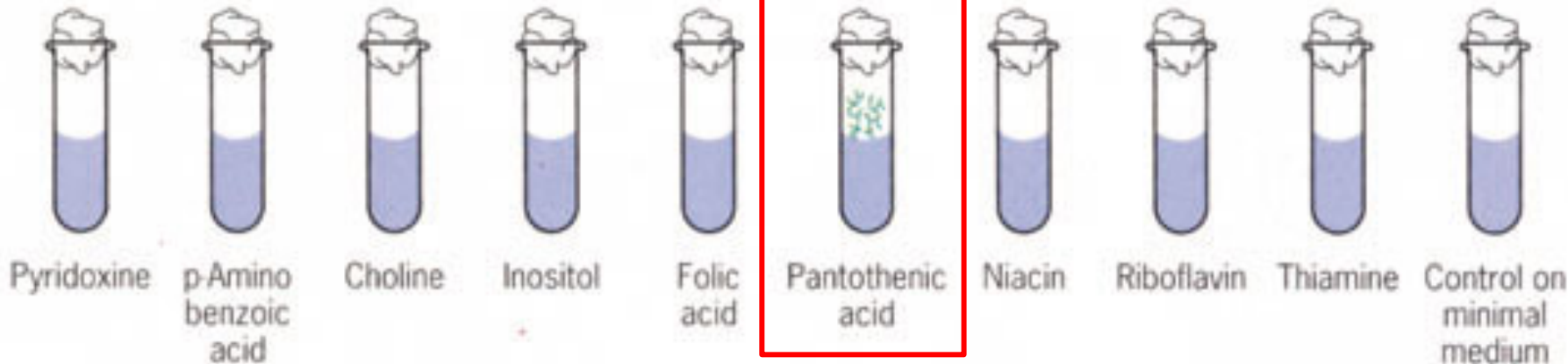
Deficiência na síntese de qual vitamina?

Minimal medium + vitamins

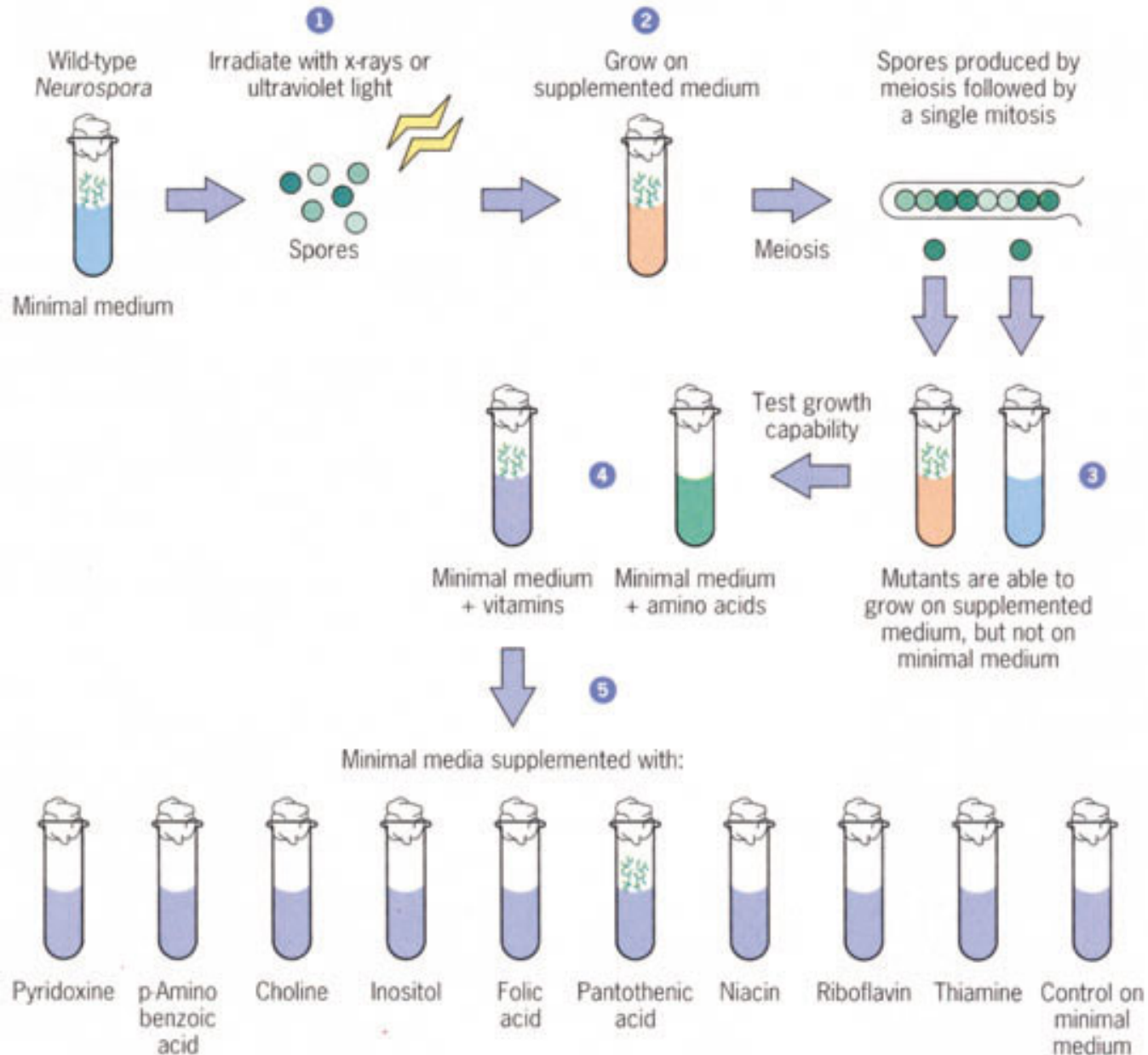


O mutante não sintetiza ácido pantotênico

Minimal media supplemented with:



Beadle e Tatum - mutantes de *Neurospora*



Beadle e Tatum & Srb e Horowitz genes/enzimas ordenados em vias

Seleção de mutantes deficientes na síntese do
aminoácido arginina

mutante A

mutante B

mutante C

selvagem



Minimal
medium





Mutantes não crescem



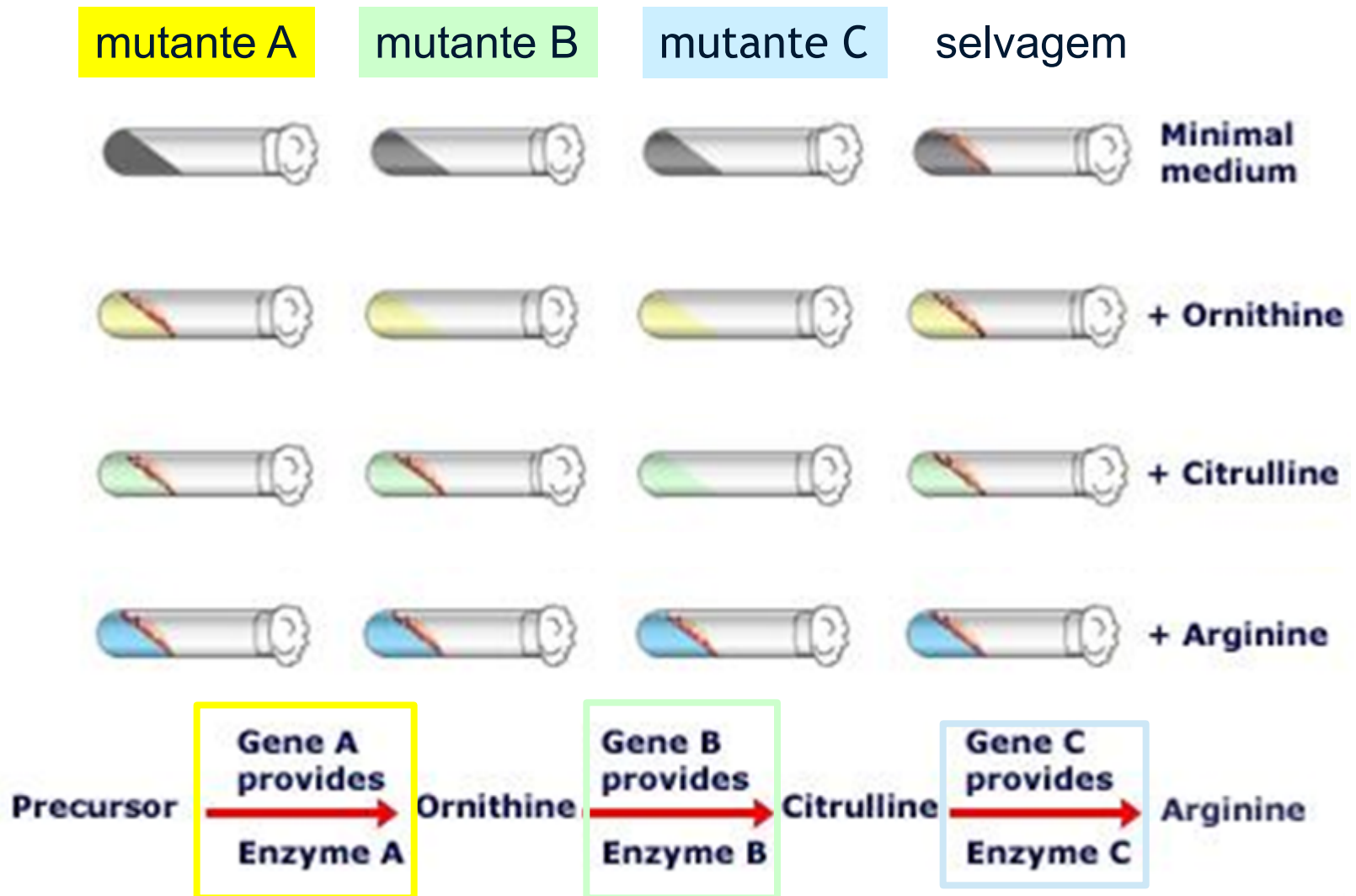
+ Arginine

Todos crescem

Beadle e Tatum & Srb e Horowitz genes/enzimas ordenados em vias

mutante A	mutante B	mutante C	selvagem	
				Minimal medium
				+ Ornithine
				+ Arginine

Beadle e Tatum & Srb e Horowitz genes/enzimas ordenados em vias



O que é um gene do ponto de vista molecular?

- Um gene, uma enzima
- Um gene, uma proteína
- Um gene, um polipeptídeo
- Um gene, um RNA



Evolução
do conceito

- **Um gene é um segmento de DNA que contém a informação para um produto final, que pode ser RNA ou proteína**

Um gene corresponde a um transcrito de RNA

Região regulatória



Região não traduzida 5'

5'UTR

5'



3'UTR

Região não traduzida 3'

3' RNA

Length of RNA defines region of gene

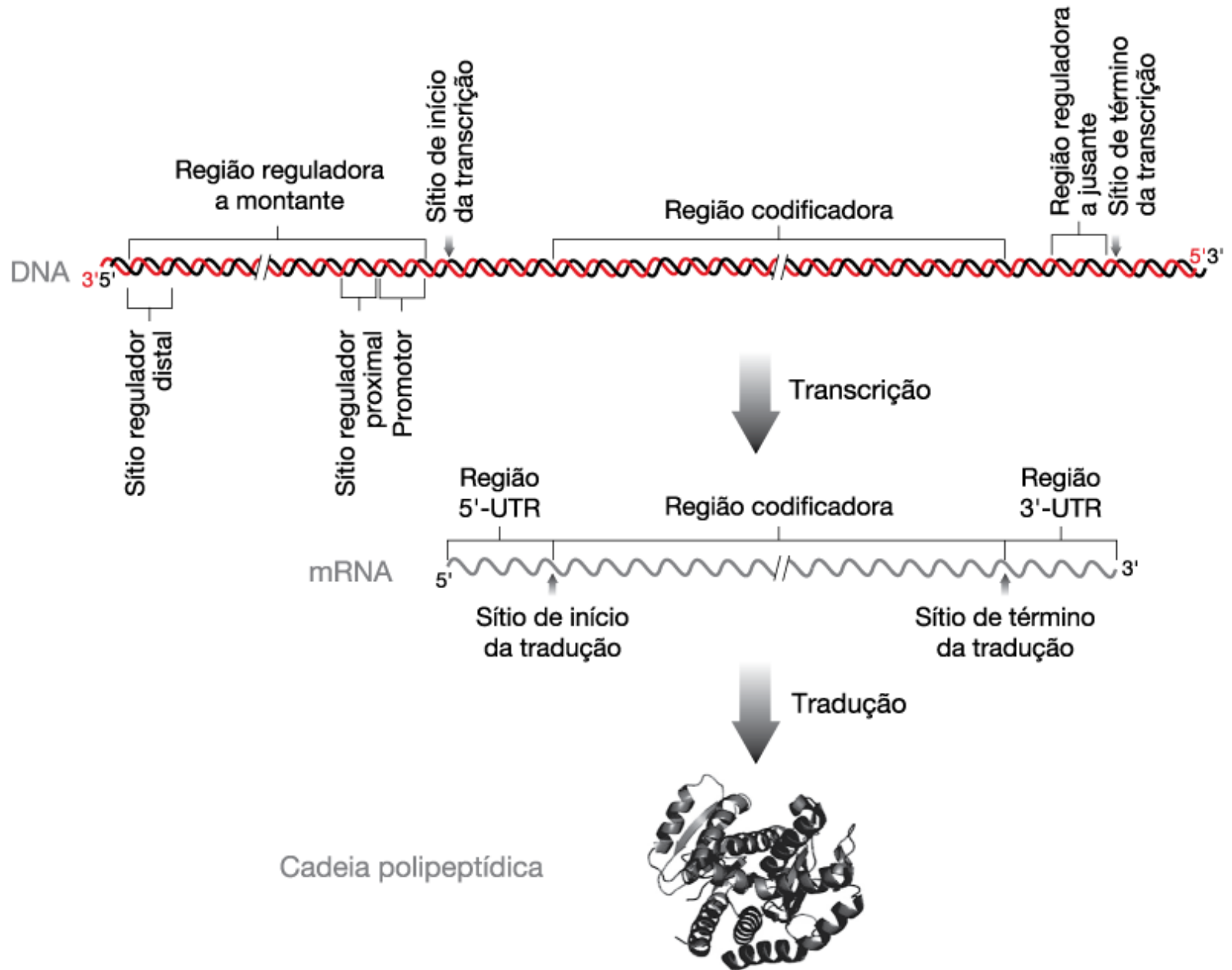


Protein defines coding region

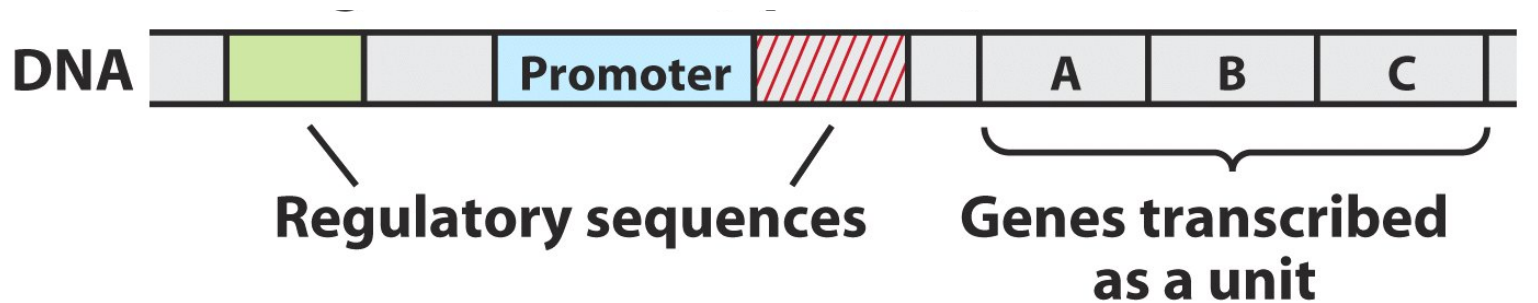
Do que é composto um gene?

- Sequência codificadora
 - informação do produto final (RNA ou proteína)
- Introns e exons (em eucariotos)
- Sequências regulatórias
 - indicam quando, onde e como o gene deve ser expresso

Gene Procariótico

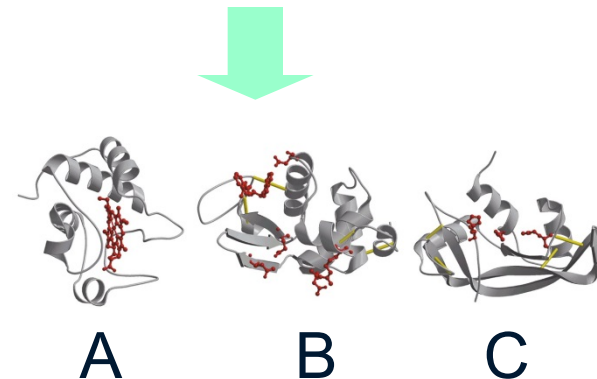


Operon Bacteriano

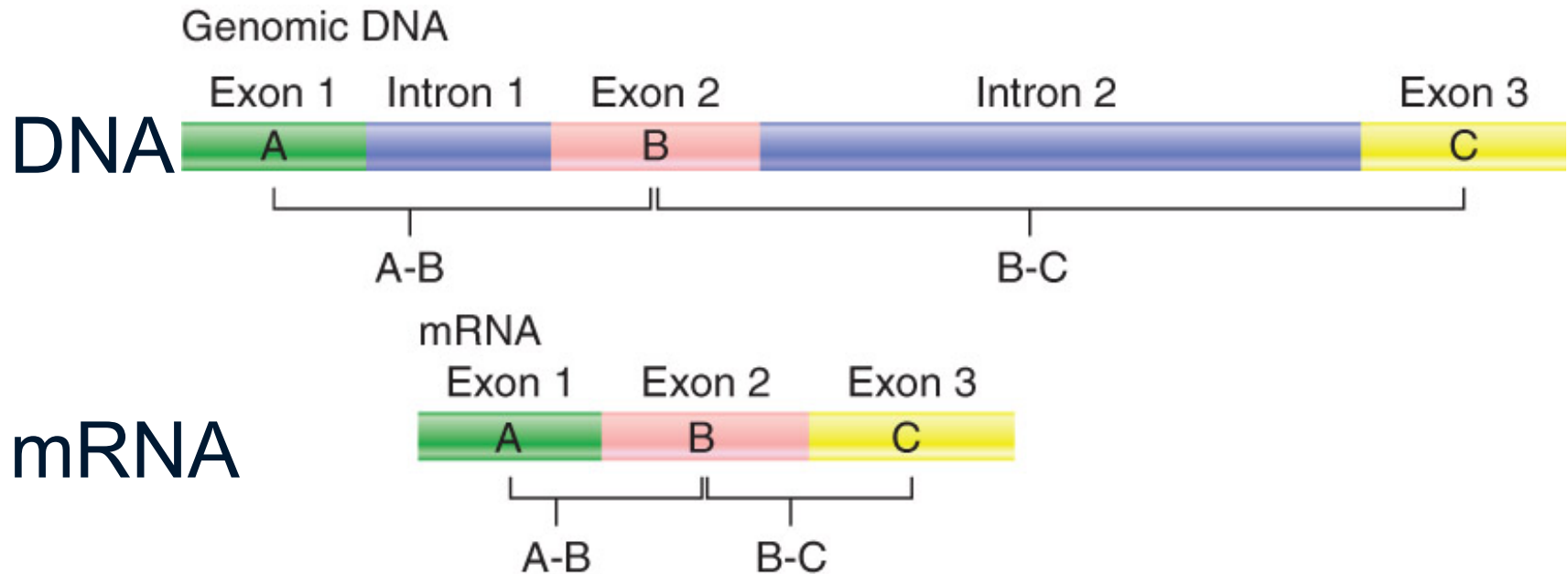


mRNA policistronico 5' ————— 3'

- Mais de uma proteína codificada no mesmo mRNA
 - nesse caso, um mRNA corresponde a mais de um gene)

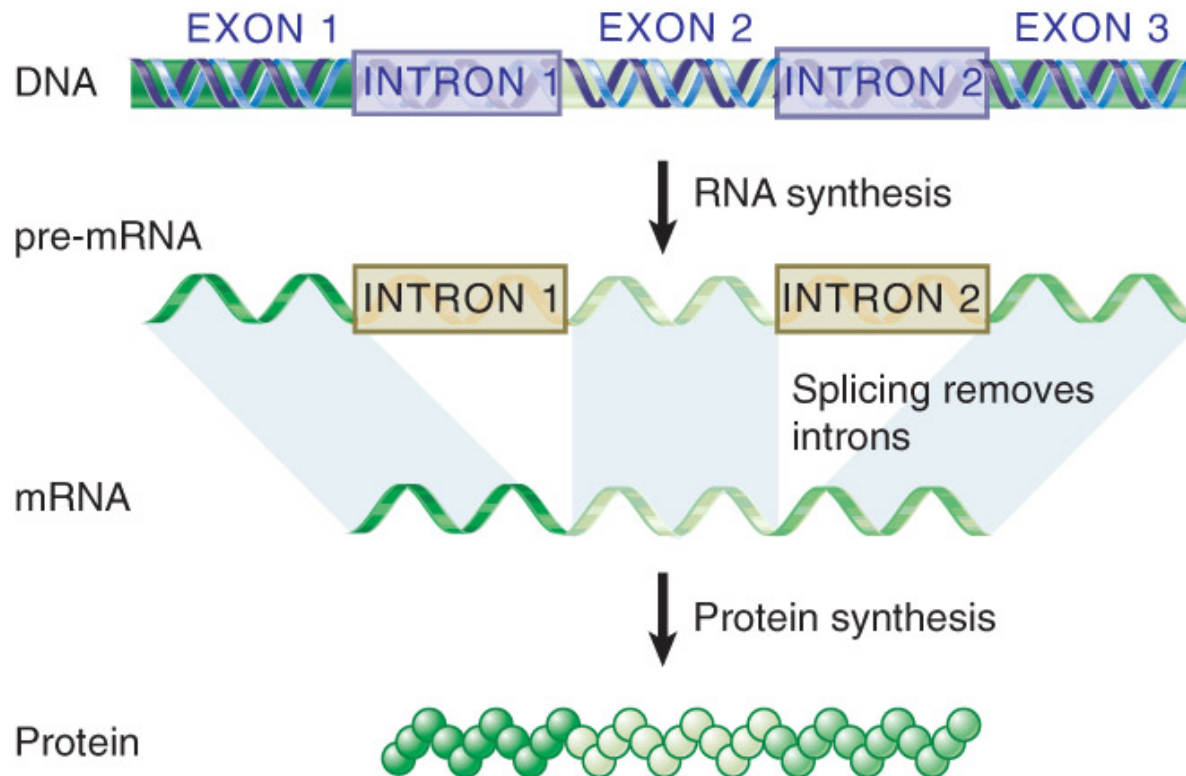


Eucariotos possuem genes interrompidos por introns



- **Exons:** regiões codificadoras
- **Introns:**
 - presentes no DNA, mas não no mRNA
 - removidos do RNA precursor (*splicing*)

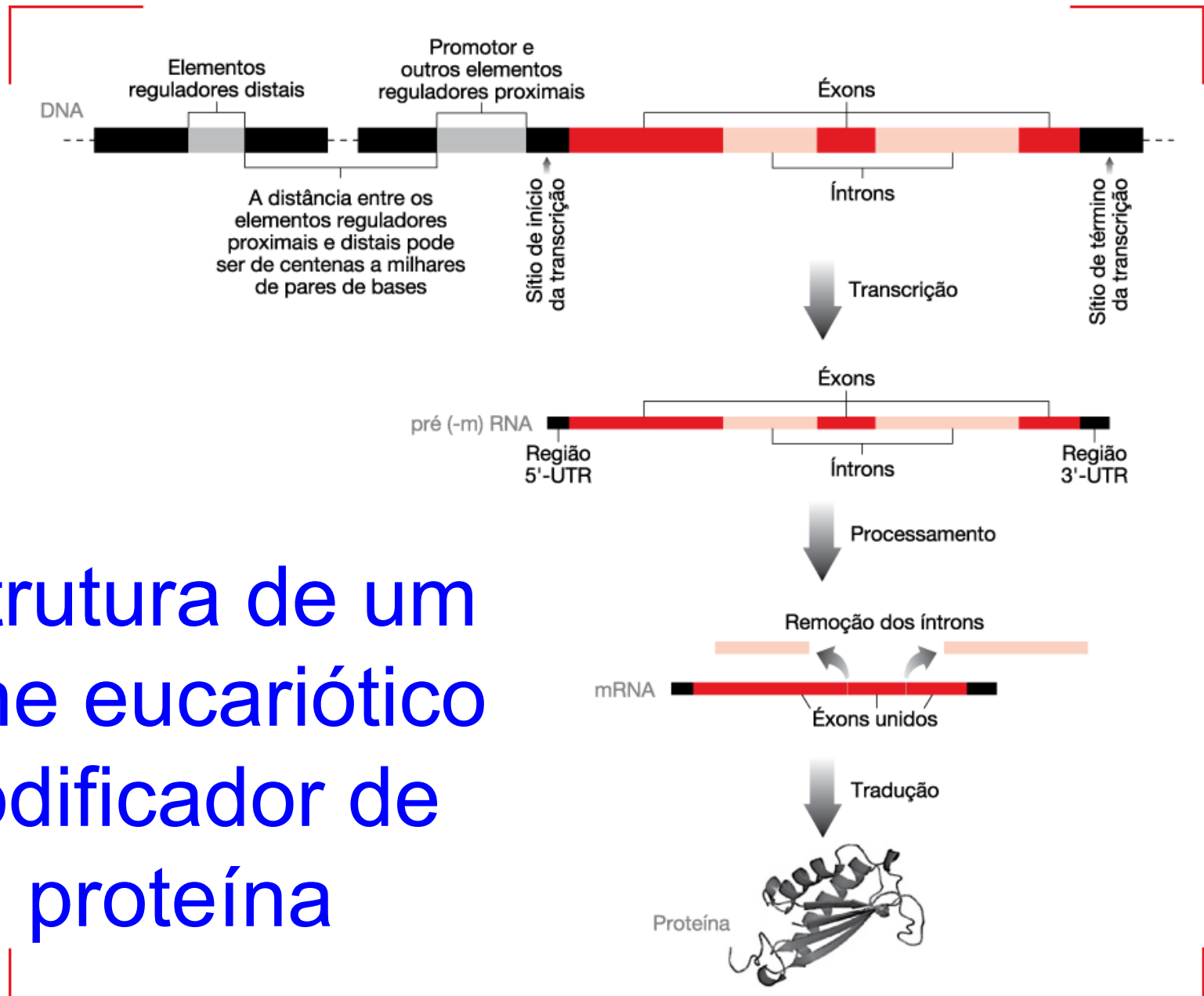
Genes eucarióticos contêm introns



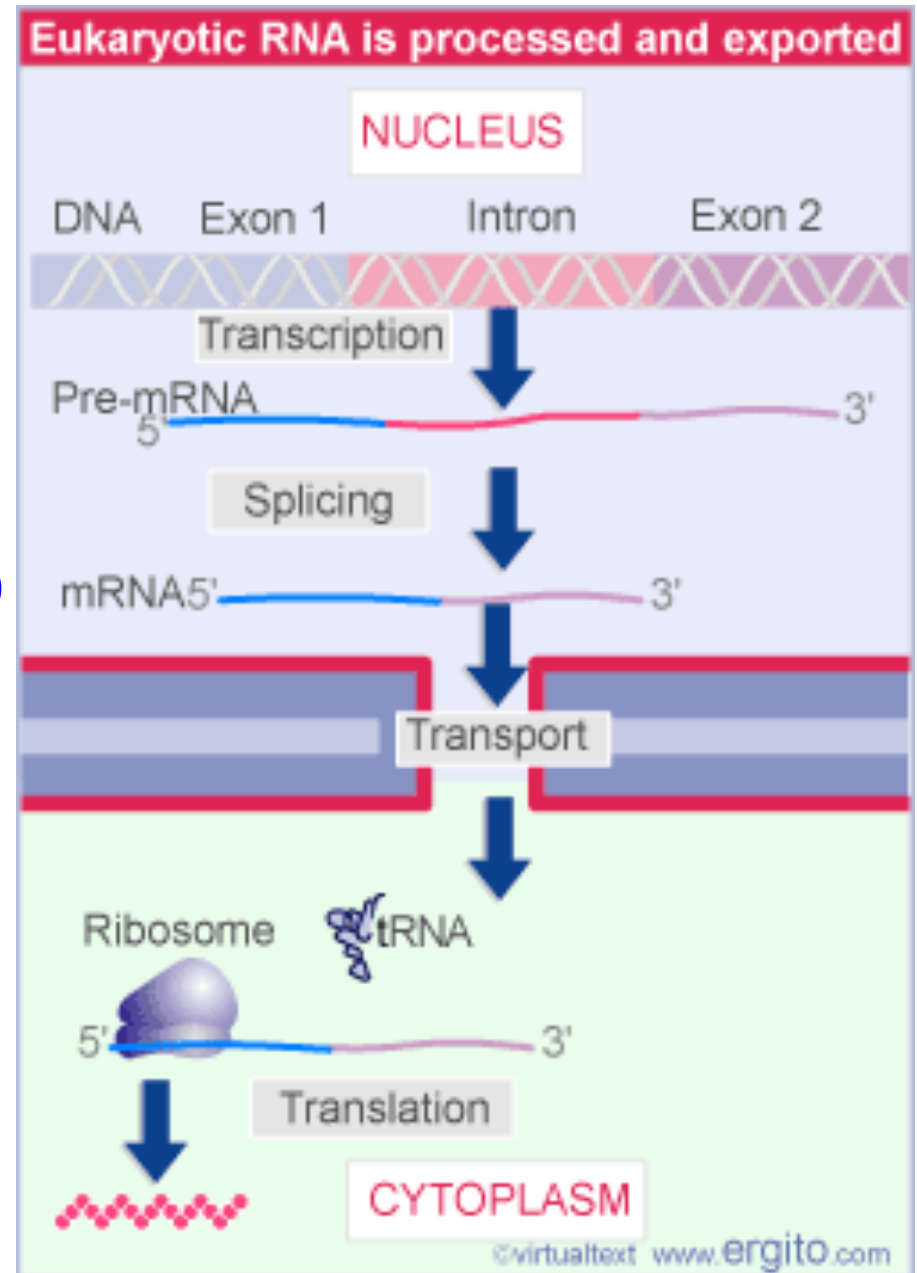
Length of precursor RNA (not mRNA) defines region of gene

Individual coding regions are separated in gene

Estrutura de um gene eucariótico codificador de proteína



Genes de eucariotos:
mRNA é processado
no núcleo
antes de ser traduzido



Quantos genes fazem um organismo?

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	~20.000

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.

Table 24-2

Lehninger Principles of Biochemistry, Fifth Edition

© 2008 W. H. Freeman and Company

O que é o genoma?

1. *Genoma é o conjunto de genes*
2. *É o conjunto do material genético de uma espécie, a sequência*
3. *O conjunto do material genético, todos os cromossomos*
4. *Conjunto de todos os genes de um organismo*
5. *Genoma é toda a informação genética de uma espécie*
6. *Código genético de cada espécie*
7. *É o conjunto que reúne todas as informações do DNA humano*
8. *É o conjunto de genes, a junção das características que eles expressam*
9. *Genoma é o código genético que determina uma espécie*
10. *É o sequenciamento de todo material genético de uma espécie*
11. *Genoma é o conjunto de informações contido no código genético*
12. *Agora não sei diferenciar código genético de genoma*

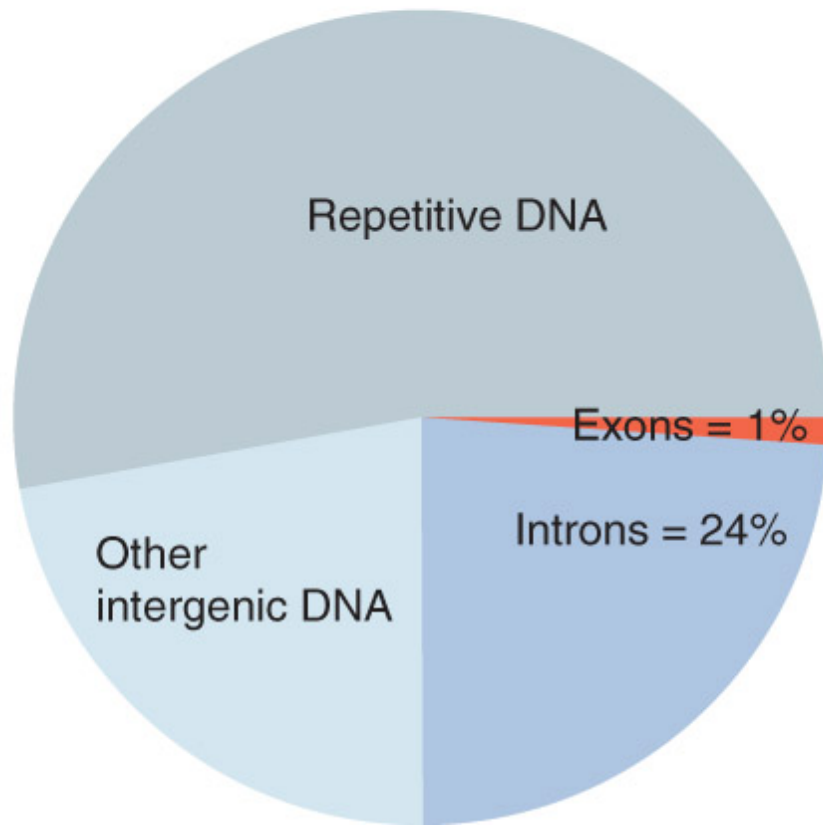
Algumas das respostas da avaliação prévia

O que é o genoma?

- Conjunto de toda a informação genética do organismo
 - Cromossomos
 - Plasmídeos
 - DNA das mitocôndrias e cloroplastos
 - Inclui genes, regiões não codificadoras de proteínas, regiões regulatórias, regiões intergênicas

Não é sinônimo de “código genético”!

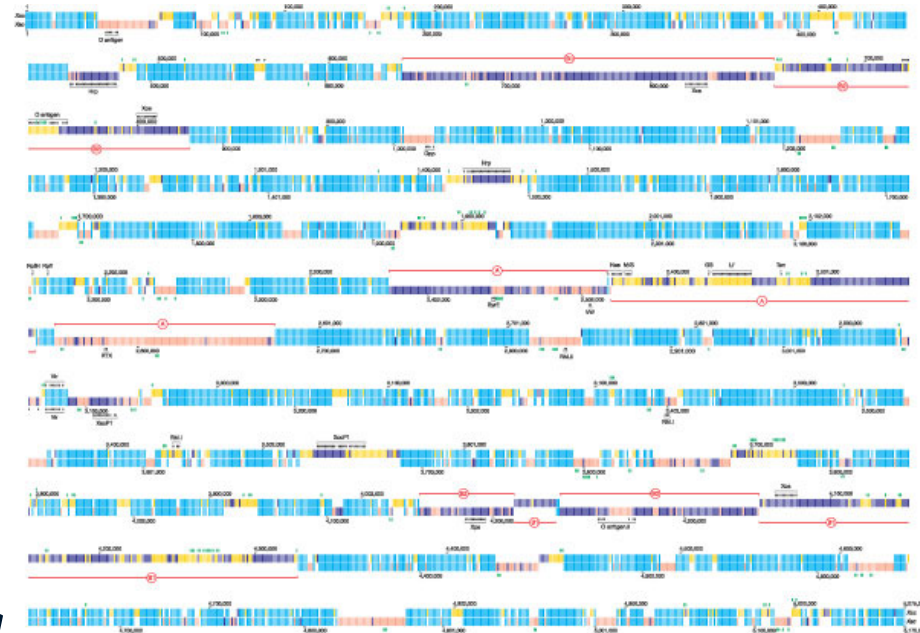
Exons representam apenas 1% do genoma humano!



- Muitos genes ainda não têm função conhecida
- O que já foi considerado “DNA lixo” pode ter funções regulatórias importantes
- O genoma não é apenas o “conjunto de genes”

Genoma bacteriano

- Genes com 1kb (media)
- Regiões regulatórias e intergênicas pequenas
- Ausência de introns
- Presença de operons
- poucas sequências repetidas
- Alta densidade gênica



Xylella fastidiosa

Sequenciamento de genomas

- Permite conhecer toda a sequência de bases do DNA
- Ferramentas de predição de genes/ORFs
- Comparações com genes de função conhecida
- Comparações com outros organismos
- Evolução
- Predisposição a doenças ou sensibilidade a tratamentos clínicos

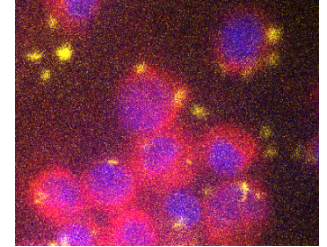
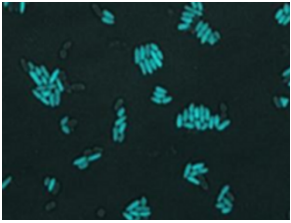
- Ainda há mais perguntas que respostas!
 - *Aulas sobre Sequenciamento e Genômica*

Como relacionar a sequência à função de um gene?

- Ao longo da evolução, genes que codificam processos similares se mantêm, mas sofrem variações e duplicações
- Algoritmos que comparam as sequências de genes ou de seus produtos preditos entre organismos diferentes
- Valores de identidade e similaridade definem homologia

Quais as diferenças e
semelhanças entre os genes e os
genomas de bactérias e
eucariotos?

Bactérias x Eucariotos



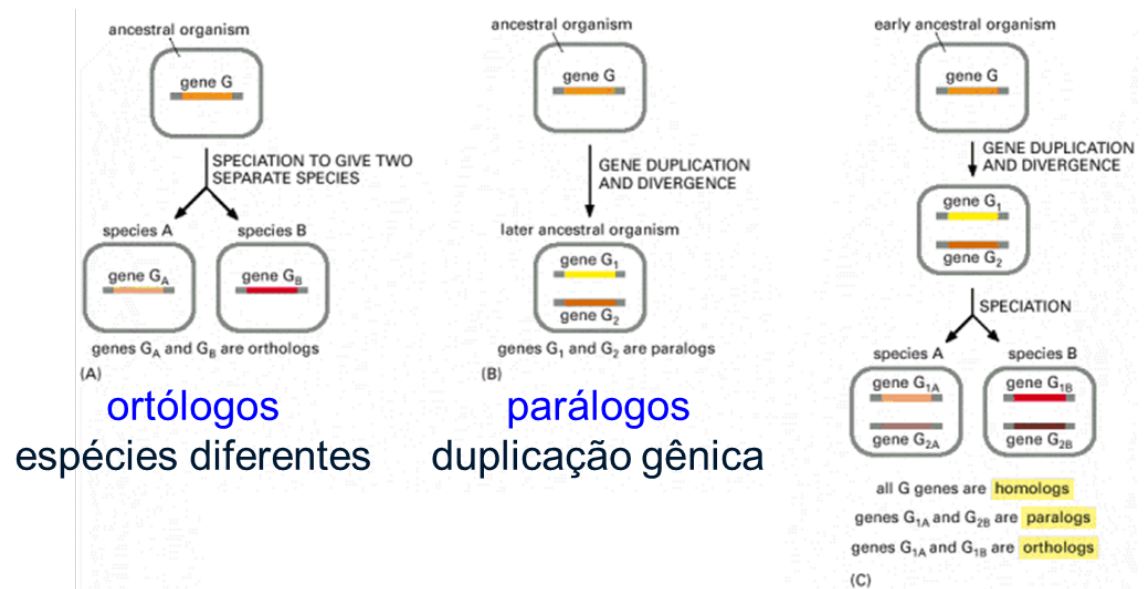
- 1 cromossomo circular
 - Genes em operons
 - ~1kb/gene
 - Alta densidade gênica
 - Regiões regulatórias curtas
 - Poucas e pequenas regiões intergênicas
 - Poucas sequências repetidas
- Vários cromossomos
 - Cada gene tem sua região regulatória
 - Genes com tamanhos muito variáveis
 - Genes interrompidos por introns
 - Regiões regulatórias longas e distantes
 - Regiões intergênicas longas
 - Muitas sequências repetidas

Material extra

- Estrutura de genes e genomas

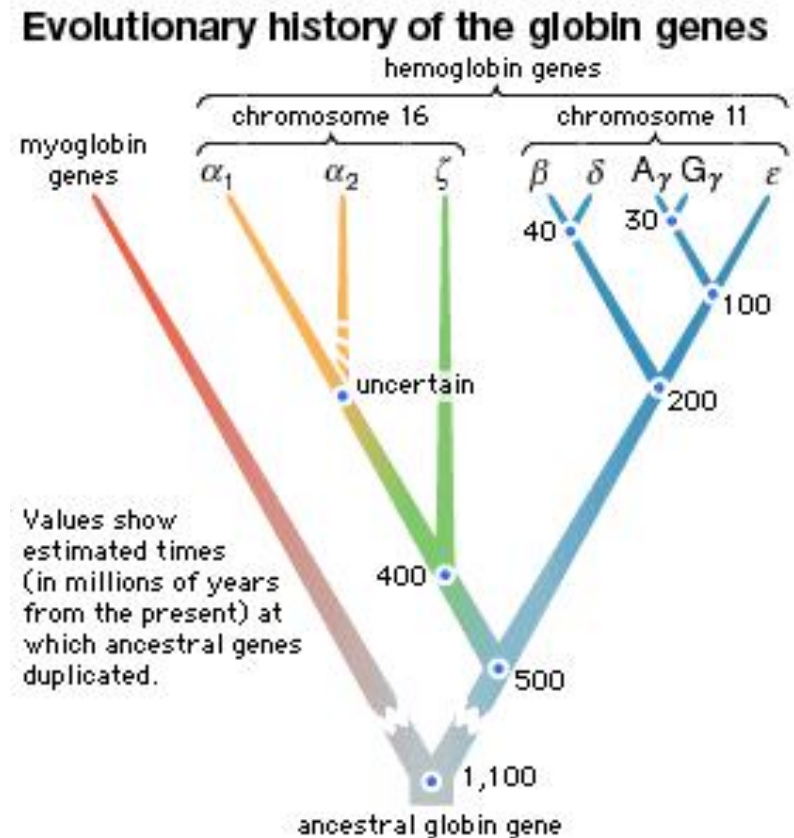
Homologia x Identidade/Similaridade

- **Homologia:** supõe um ancestral comum
 - qualitativo
- **Identidade:** o quanto duas sequências são idênticas
- **Similaridade:** o quanto suas sequências de proteínas são similares em termos de aminoácidos de mesma classificação



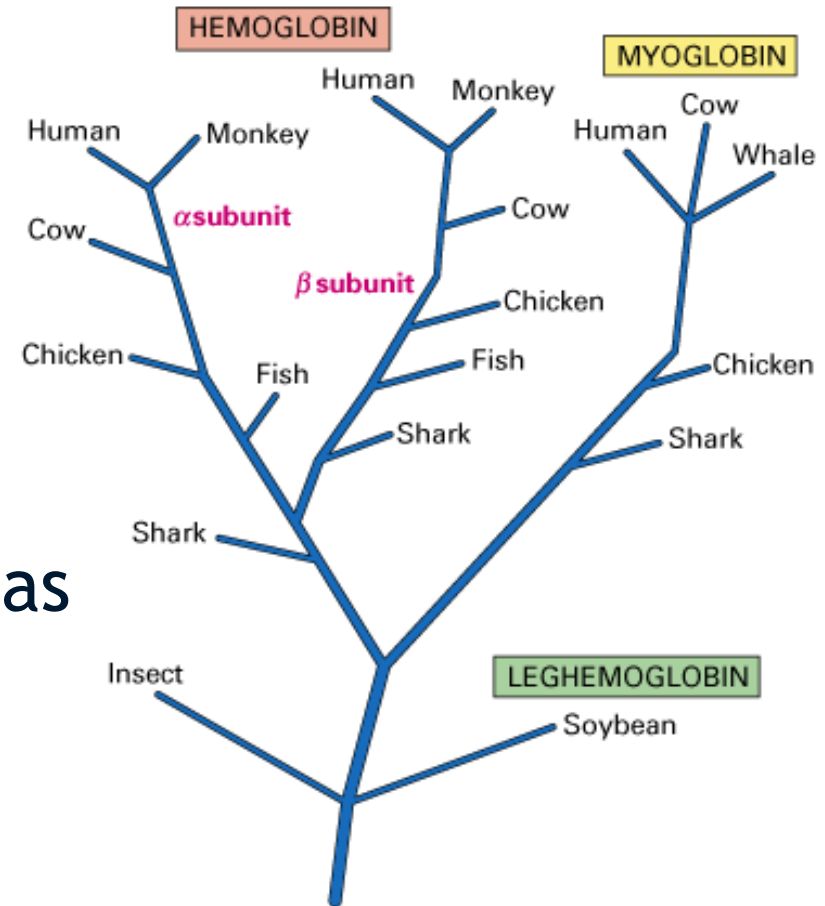
Famílias multigênicas

- Resultado de eventos de duplicação e diversificação (genes parálogos)
- conjunto de genes que codificam variantes de uma proteína
 - Ex: hemoglobinas, histonas, imunoglobulinas
- Padrões diversos de expressão
 - Local
 - Temporal



Homologia x Similaridade

- **Homologia:** supõe um ancestral comum
 - qualitativo
- **Identidade:** o quanto duas sequências são idênticas entre si
 - quantitativo



Organização e conteúdo do genoma humano

3200 Mb ($3,2 \times 10^9$ pb)

**DNA intergênico
~75%**

**RNAs
regulatórios
Microssatélites**

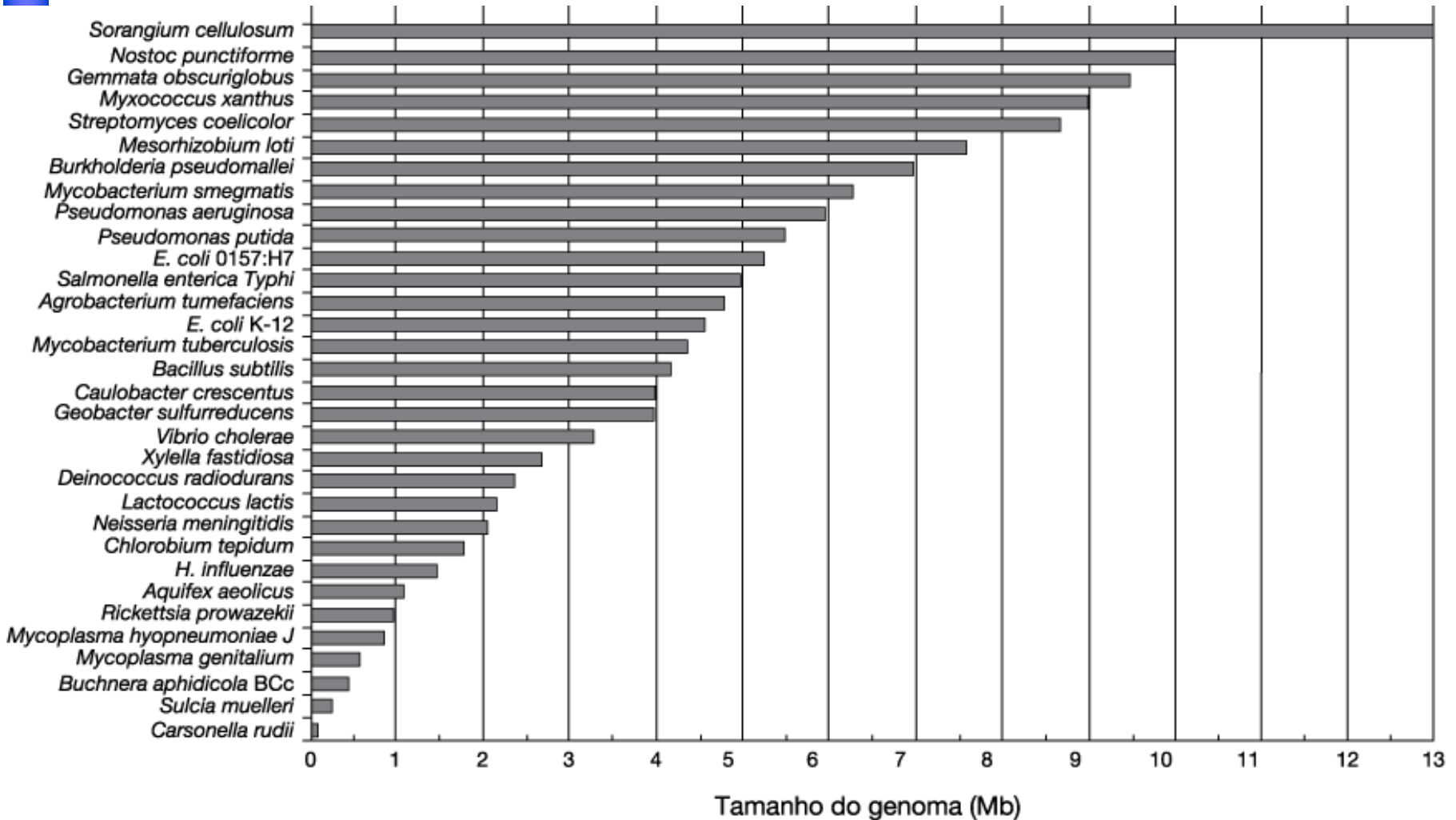
Repetições

**Genes e sequências
relacionadas a genes
codificantes de proteínas
(~25000 genes \rightarrow ~25% do
genoma)**

**Exons
(~1% do genoma)**

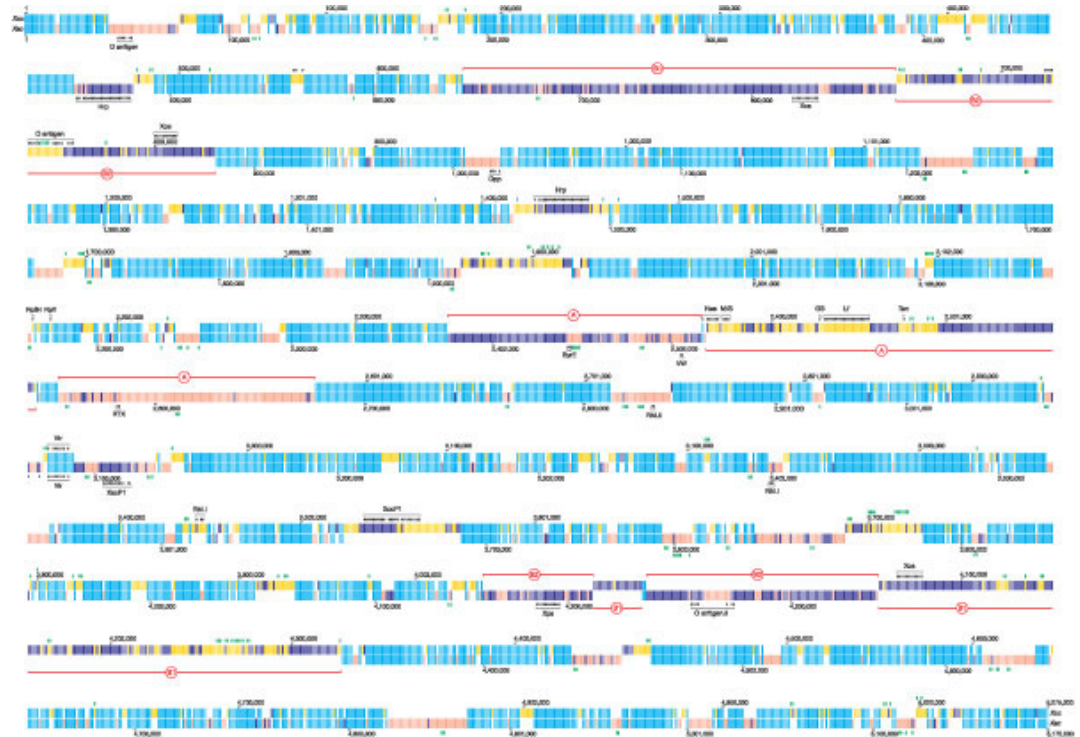
**Sequências
relacionadas a
estes genes:
Introns
UTRs
pseudogenes**

Tamanho dos genomas bacterianos



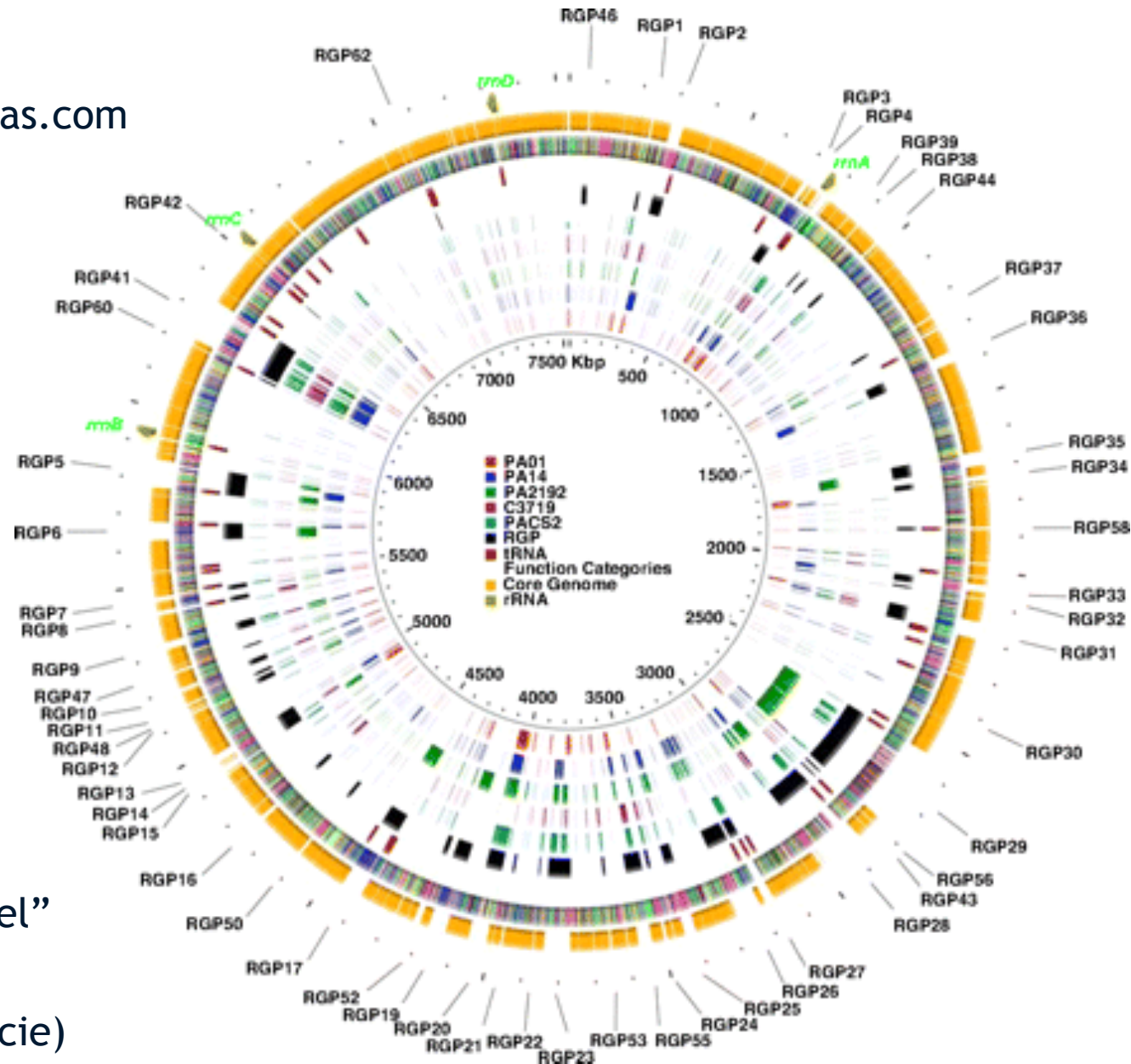
Xyلةla fastidiosa

- Bactéria que causa doença em frutas cítricas e outras plantas (fitopatógeno)
- Projetos Fapesp
- Envolveram e capacitaram muitos laboratórios de São Paulo para a Biologia Molecular



Exemplo de genomas: *Pseudomonas*

www.pseudomonas.com



Genoma central

+

Genoma “dispensável”

=

Pangenoma (da espécie)

Transferência horizontal de genes

- Parte do genoma de uma espécie é incorporada por outra
- Procariotos e eucariotos
- (transferência vertical: de “pais/mães” para “filhos/filhas”)

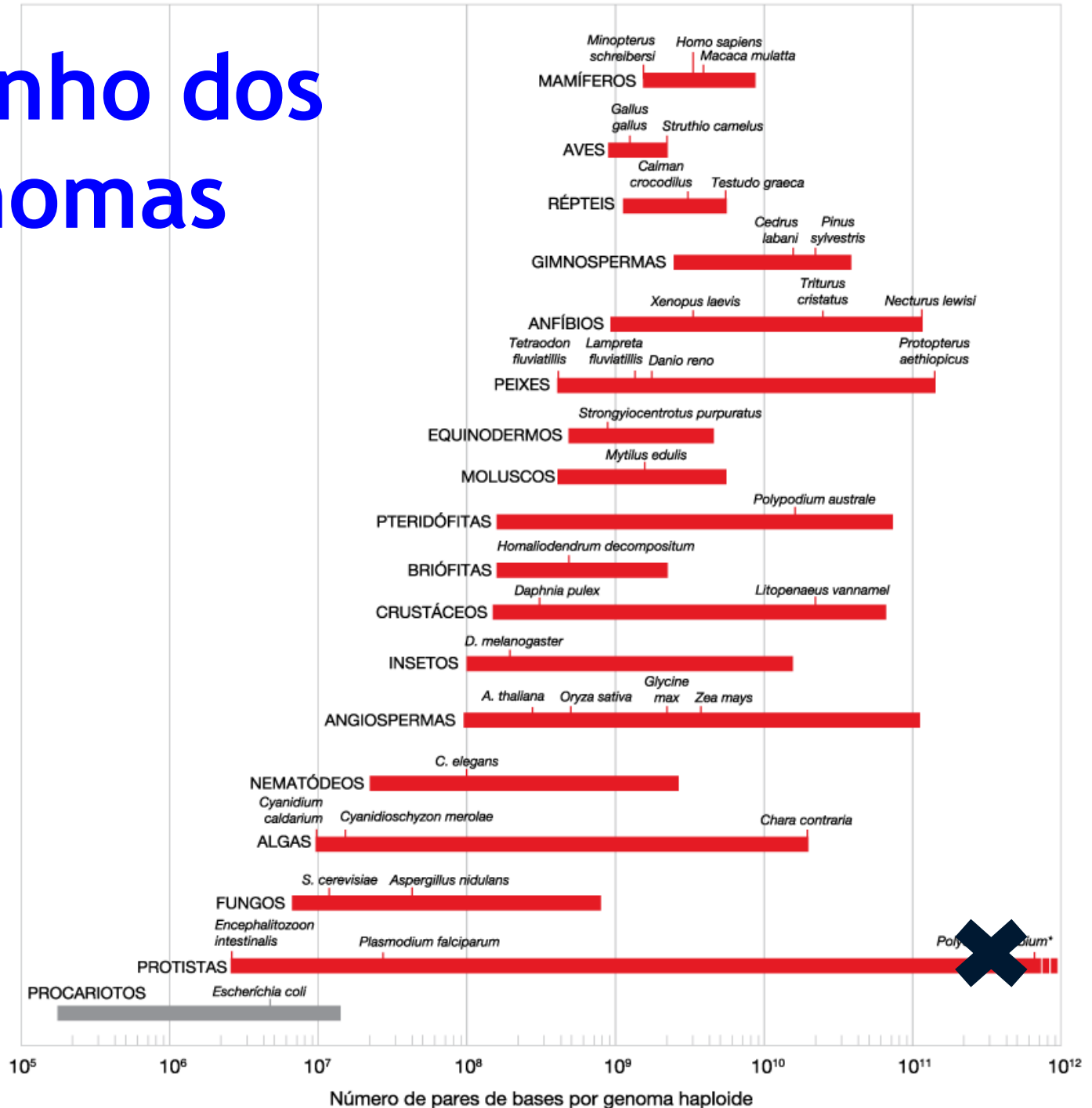
Sequenciamento de genomas

937 espécies eucarióticas (2014)

Qual é o número atual?

	TOTAL NUMBER OF SPECIES (ESTIMATED)	NUMBER OF SPECIES IDENTIFIED/ DESCRIBED (APPROXIMATE)	NUMBER OF SPECIES WITH COMPLETE GENOMES SEQUENCES* (ESTIMATED)	REFERENCES
BACTERIA, ARCHAEA	100,000 to 10 million	12,000 (460 cultured Archaea)	17,420 bacteria, 362 Archaea genomes	GOLD database; <i>Nucleic Acids Research</i> (January 2012); World Data Centre for Microorganisms
FUNGI	1.5 million	100,000	356	JGI
INSECTS	10 million	1 million	98	<i>Pest Management Science</i> (May 2007); NCBI; R.G. Foottit and P.H. Adler, <i>Insect Biodiversity: Science and Society</i> (2009)
PLANTS	435,000 (land plants and green algae)	300,000	150	Botanic Gardens Conservation International; The Plant List; W.S. Judd et al., <i>Plant systematics—a phylogenetic approach</i> (2008); D. Bramwell, <i>Plant Talk</i> , 28:32-4 (2002)
TERRESTRIAL VERTEBRATES, FISH	80,500 (5,500 mammalian)	62,345 (5,487 mammalian)	235 (80 mammalian)	Genome 10K; Zoological Society of London (2012); <i>Science</i> (December 2010); Australian Biodiversity Information Services (2009)
MARINE INVERTEBRATES	6.5 million	1.3 million	60	<i>Journal of Heredity</i> (May 2014); <i>Current Biology</i> (December 2012); Zoological Society of London (2012); <i>Oceanis</i> 19:5-24 (1993)
OTHER INVERTEBRATES	1 million nematode, several thousand <i>Drosophila</i>	23,000 nematode, 1,300 <i>Drosophila</i>	17 nematode, 21 <i>Drosophila</i>	M. Blaxter et al., "The evolution of parasitism in Nematoda," <i>Parasitology</i> (in press); Texas A&M University; <i>Molecular Biology and Evolution</i> (1995)

Tamanho dos genomas



Tipos de sequências no genoma humano

