Organização do Genoma Humano

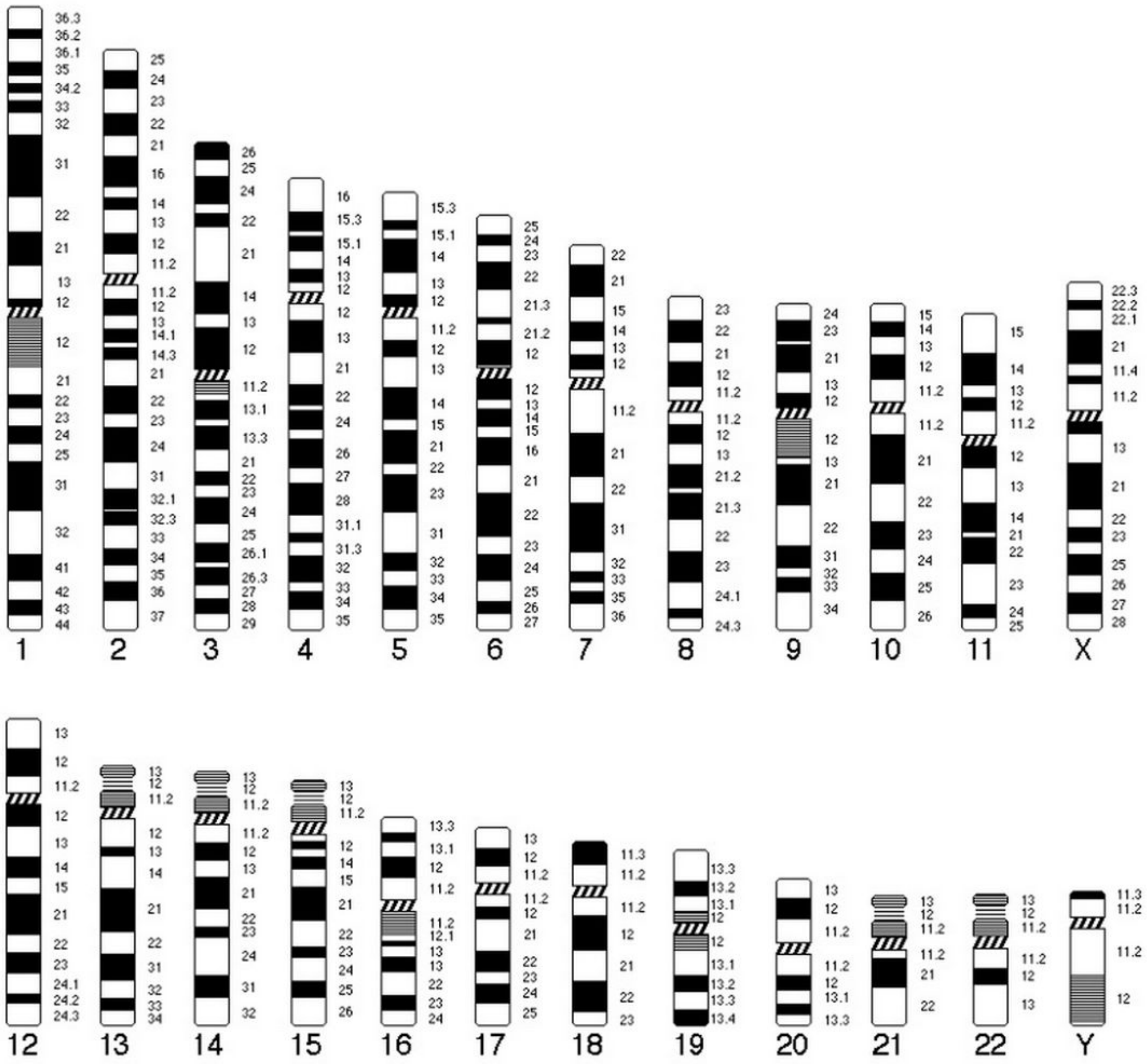
RIB0102 - Genética Molecular

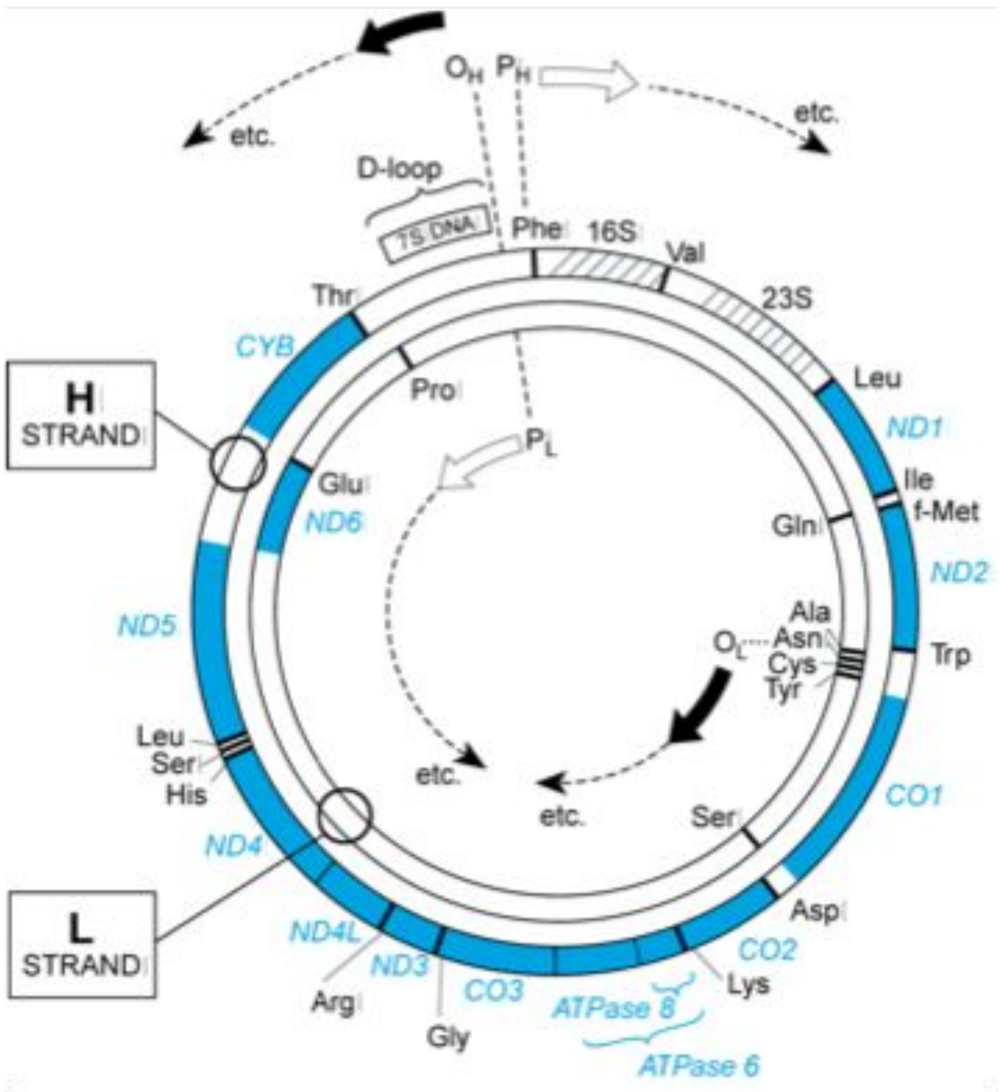
Wilson Araújo Silva Jr
FMRP/USP

Genoma Humano

Genoma nuclear

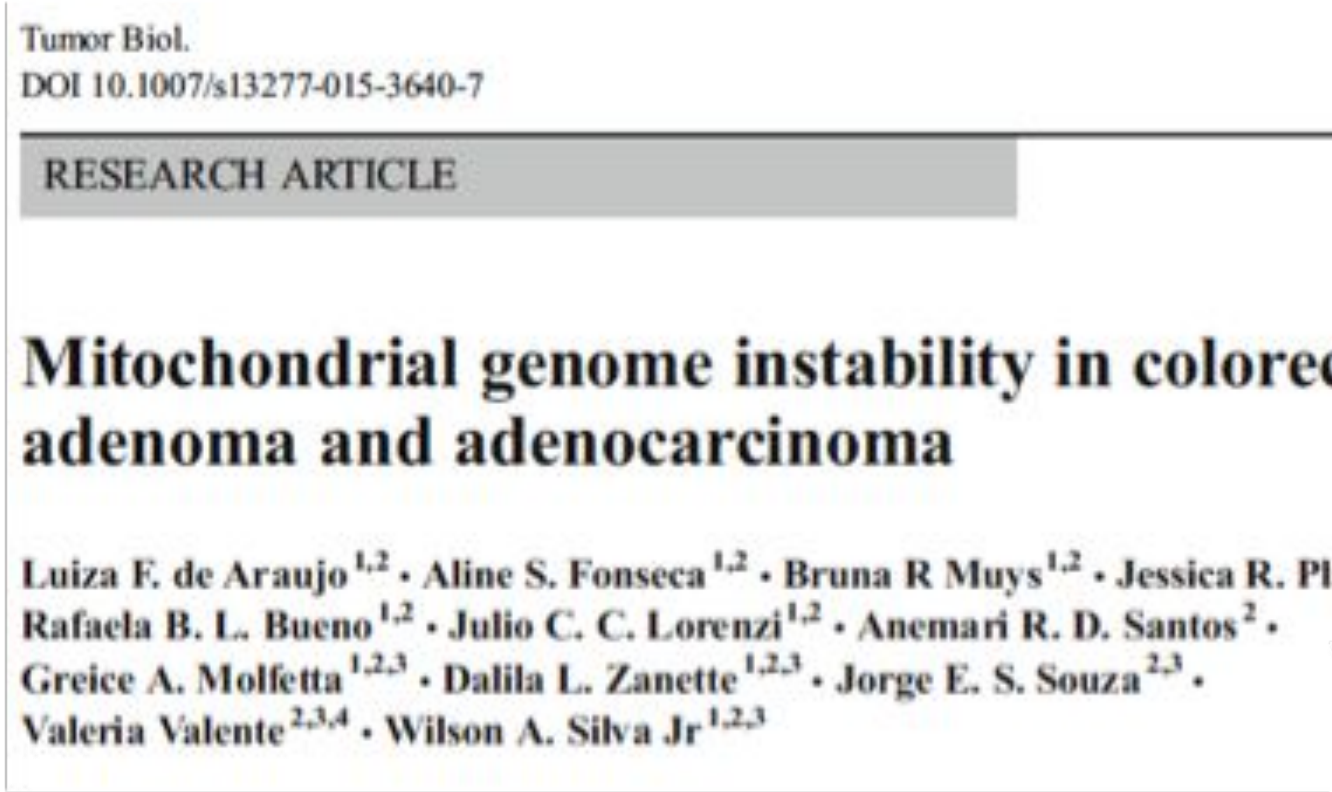
Genoma mitocondrial





Principais diferenças entre o genoma nuclear e mitocondrial

	Nuclear	Mitocondrial
Tamanho	3.300 Mb	16.6 kb
No. de moléculas de DNA	23 (em XX) ou 24 (em XY), linear	um molécula de DNA circular
No. total de mol. DNA/células	23: Cels. haploides, 46: Cels. Diploides	milhares genomas
Proteinas associadas	várias classes de histonas e proteínas não-histonas	milhares genomas por mitocondria
Número de genes	~ 25.000 - 30.000	37
Densidade gênica	~ 1/40 kb	1/0.45 kb
DNA repetitivo	grande porção	muito pouca
Transcrição	os genes são transcritos individualmente	transcrição contínua de muitos genes
Introns	encontrado em muitos genes	ausente
Percentagem de DNA codificador	~ 3%	~ 93%
Codon usage	?	?
Recombinação	pelo menos uma vez por cada par de homólogos	transcrição contínua de muitos genes
Herança	mendeliana no cromossomo X e autossomos, paternal no cromossomo Y	Exclusivamente materna

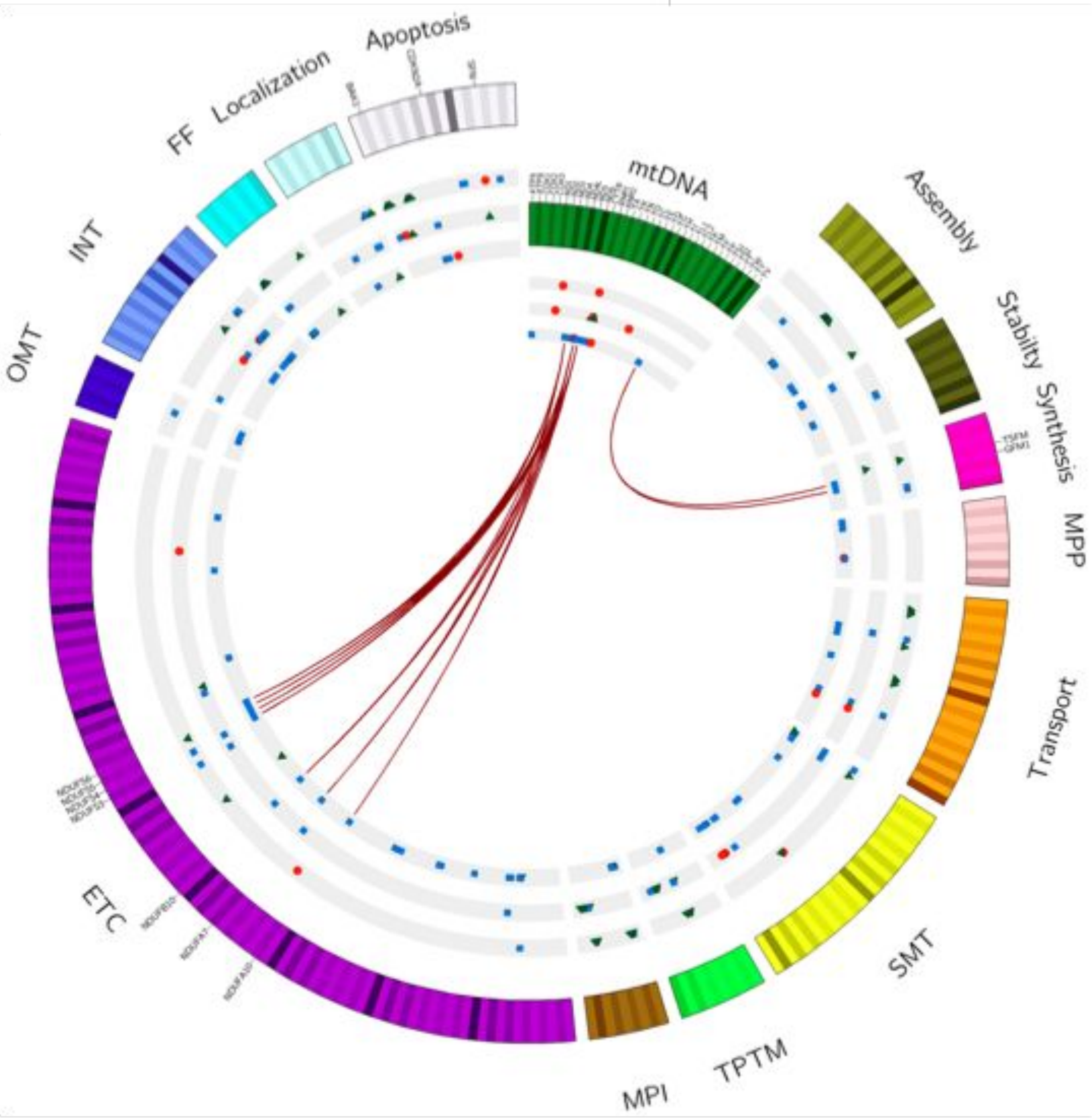


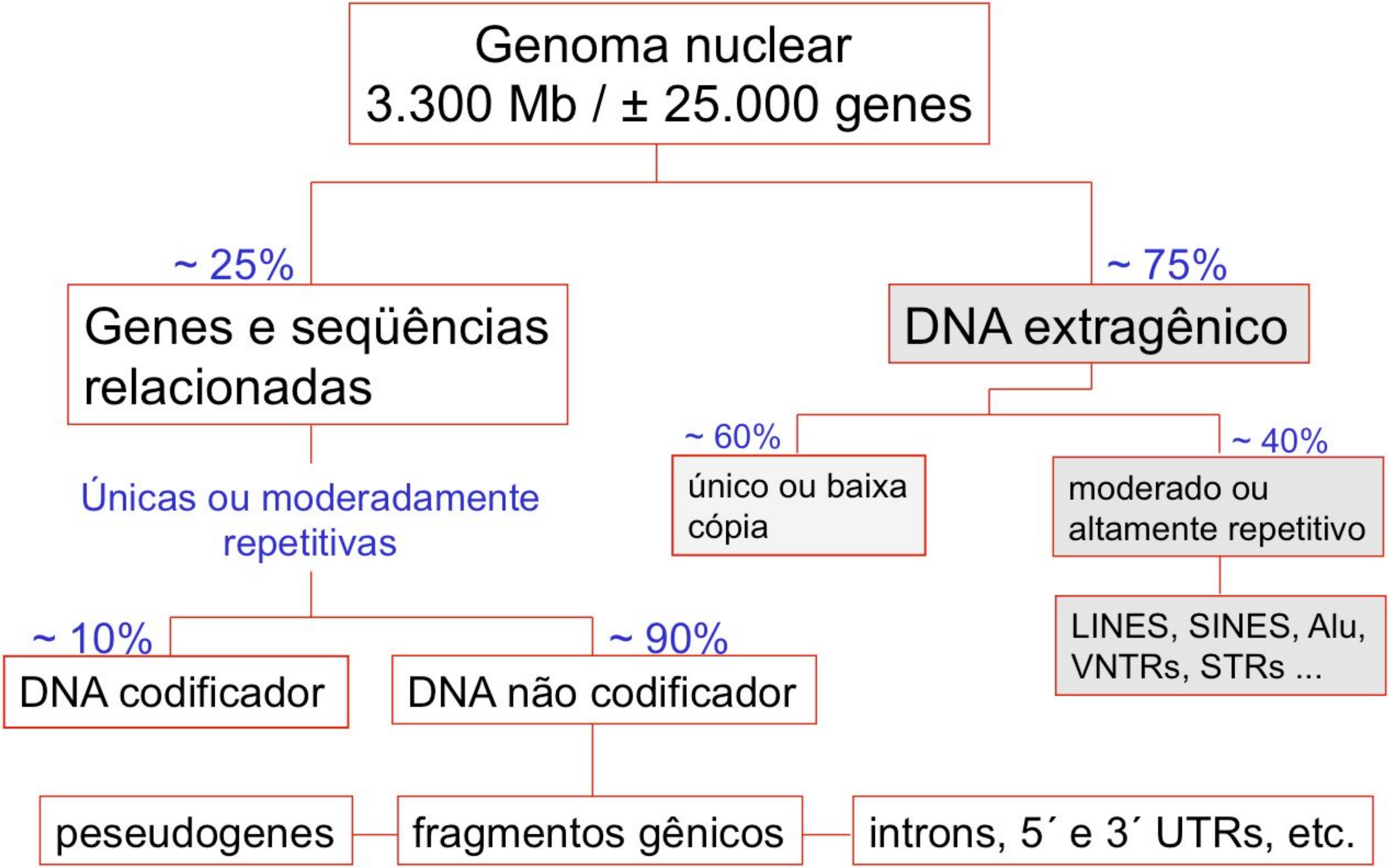
Tumor Biol.
DOI 10.1007/s13277-015-3640-7

RESEARCH ARTICLE

Mitochondrial genome instability in colorectal adenoma and adenocarcinoma

Luiza F. de Araujo[1,2] · Aline S. Fonseca[1,2] · Bruna R Muys[1,2] · Jessica R. Plaça[2] · Rafaela B. L. Bueno[1,2] · Julio C. C. Lorenzi[1,2] · Anemari R. D. Santos[2] · Greice A. Molfetta[1,2,3] · Dalila L. Zanette[1,2,3] · Jorge E. S. Souza[2,3] · Valeria Valente[2,3,4] · Wilson A. Silva Jr[1,2,3]





Genoma nuclear
3.300 Mb / ± 25.000 genes

~ 25%

Genes e seqüências relacionadas

Únicas ou moderadamente repetitivas

~ 10%

DNA codificador

~ 90%

DNA não codificador

peseudogenes

fragmentos gênicos

introns, 5´ e 3´ UTRs, etc.

~ 75%

DNA extragênico

~ 60%

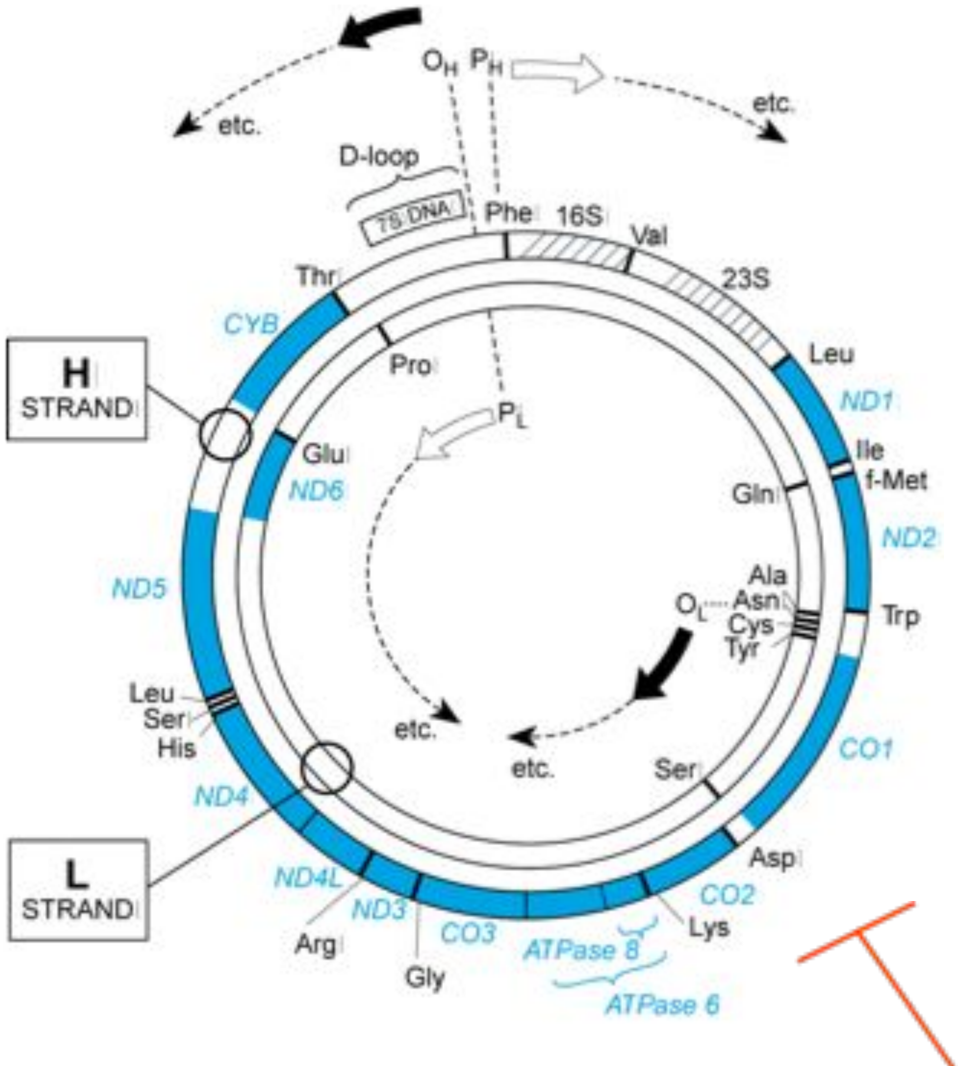
único ou baixa cópia

~ 40%

moderado ou altamente repetitivo

LINES, SINES, Alu, VNTRs, STRs ...

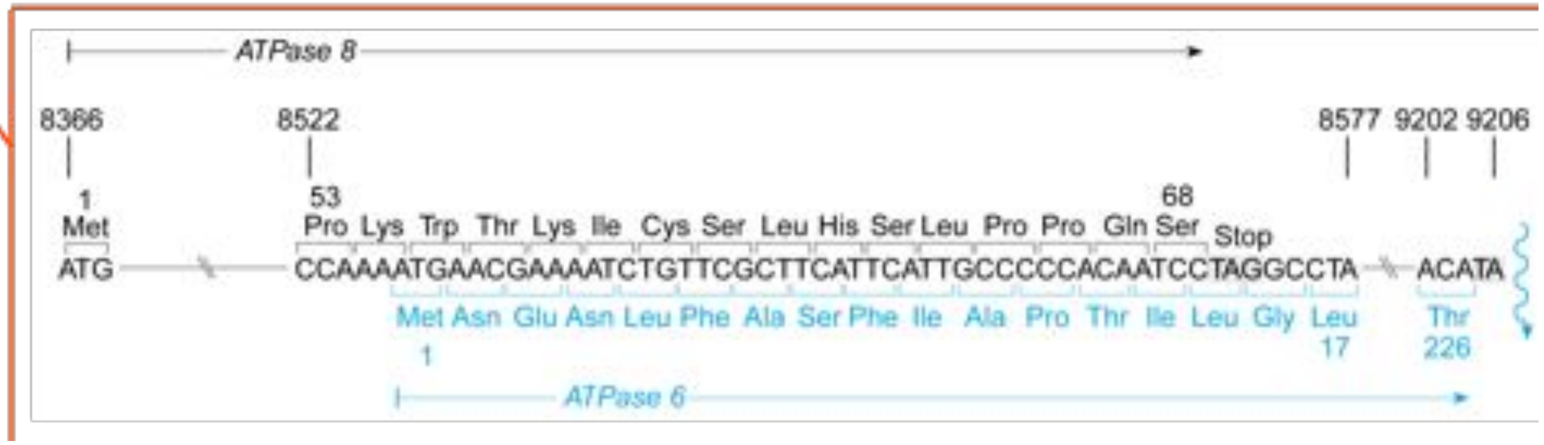
O DNA mitocondrial consiste em um pequeno duplex de DNA circular com alta densamente gênica



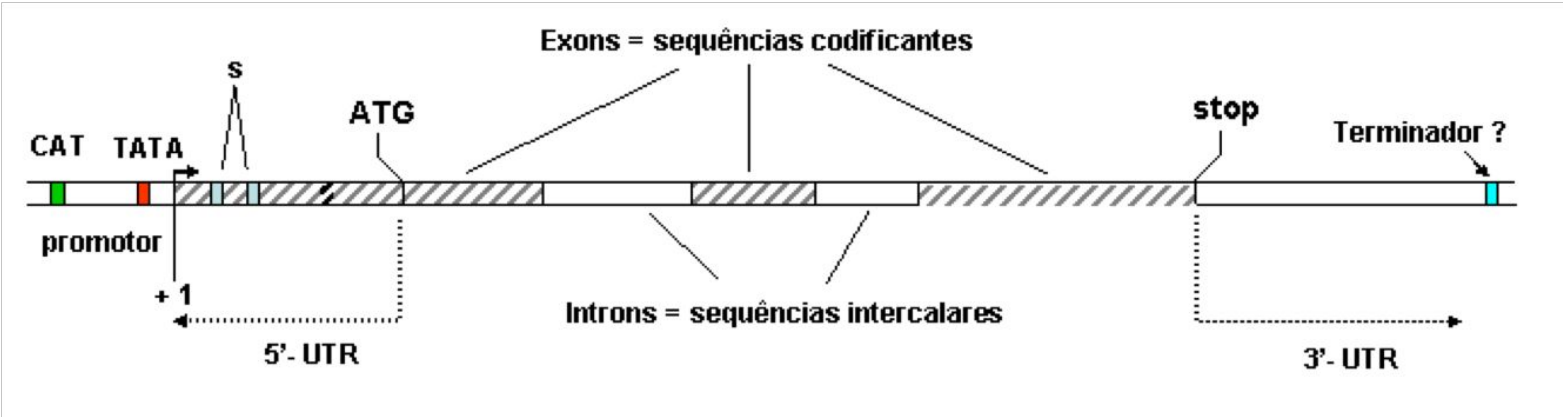
Herança Materna

Genes:

2 RNA ribossomal
22 tRNA
13 subunidades de enzimas da fosforilação oxidadtiva

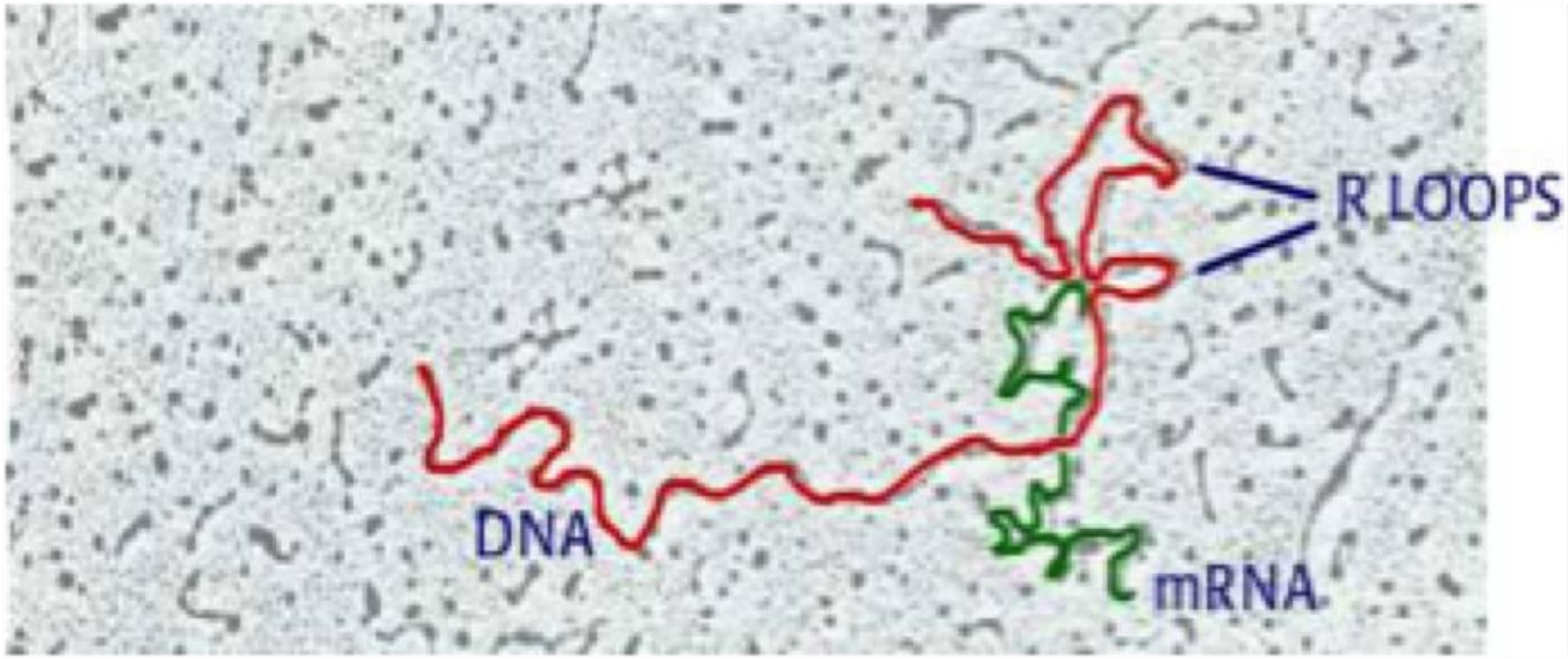


Estrutura de genes eucariotos



Phillip Sharp

Genes de eucariotos são interrompidos

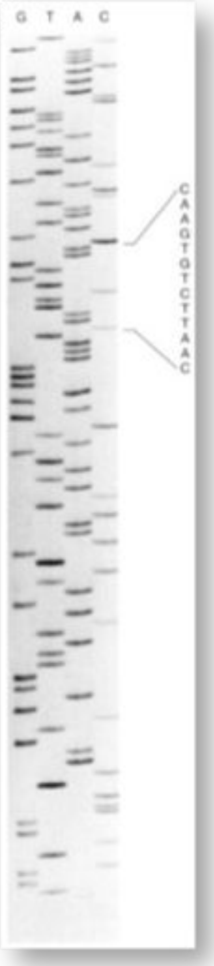


1975-1977

1987

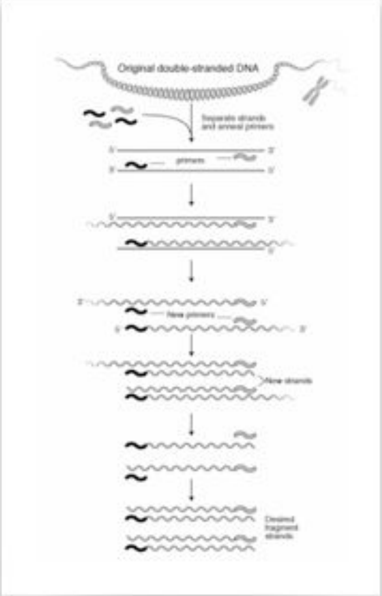
Frederick Sanger e Walter Gilbert

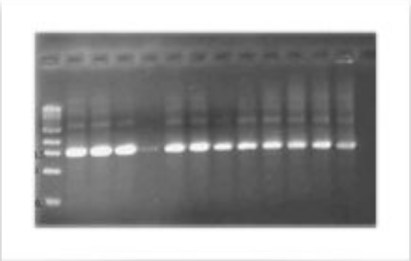
Sequenciamento do DNA



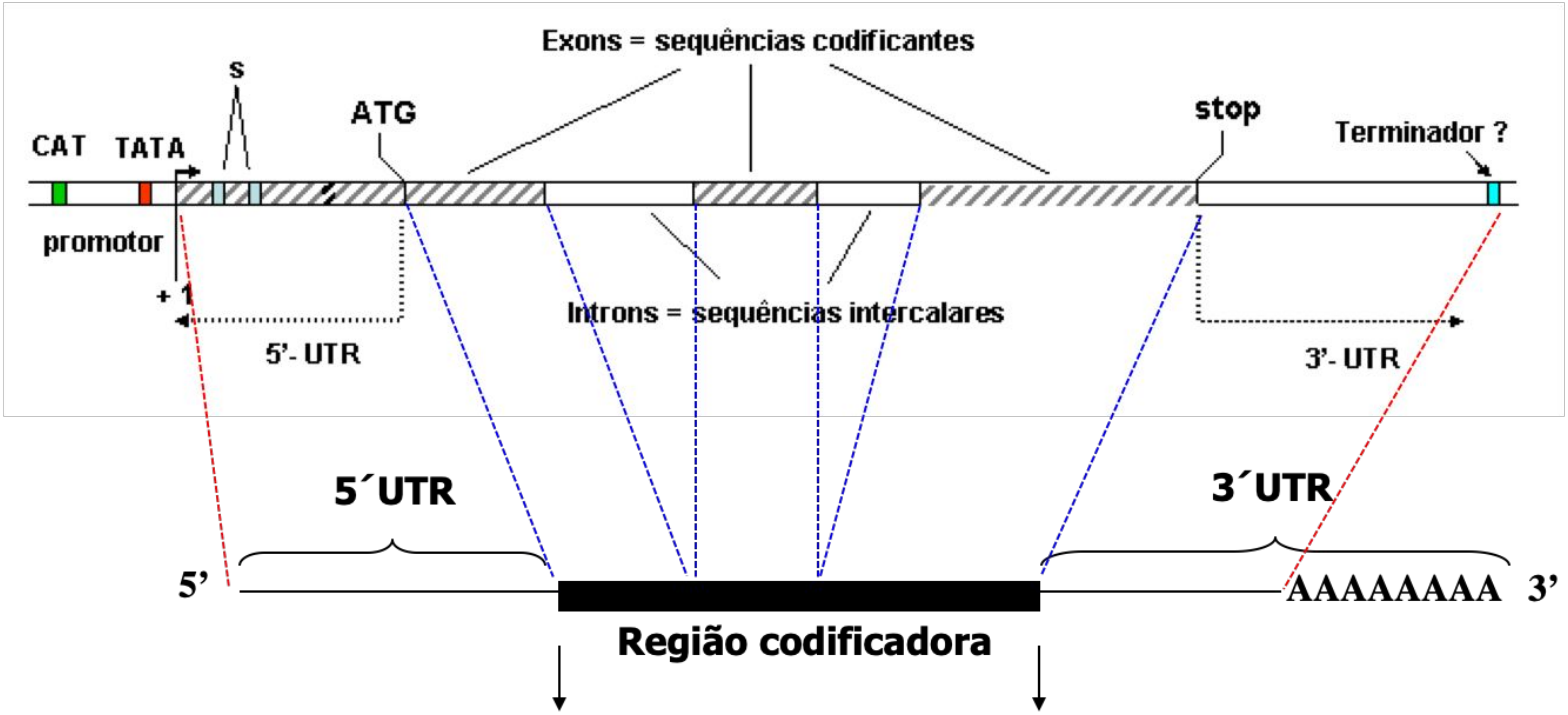
Kary Mullis

Reação em cadeia da polimerase (PCR)





Estrutura de genes eucariotos

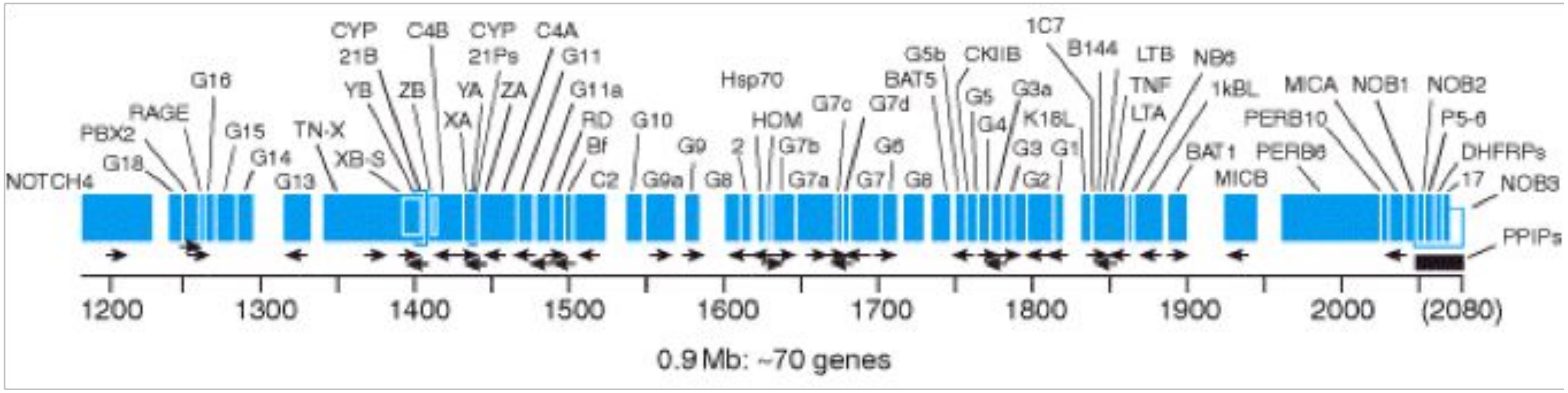


AUG -Met

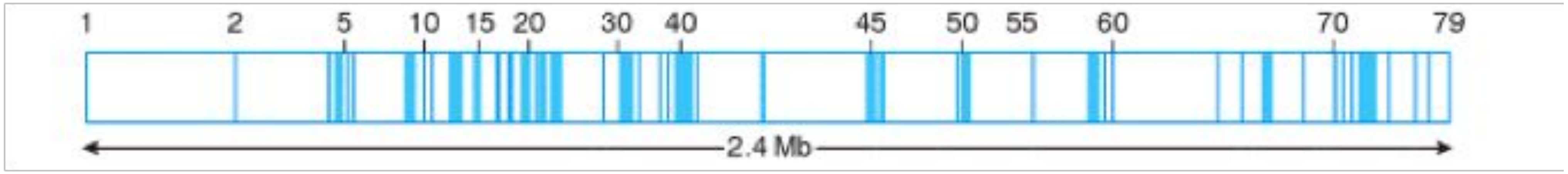
ACG -Thr
CUG -Leu
GUG -Val

TAA
UAG
UGA

HLA classe III



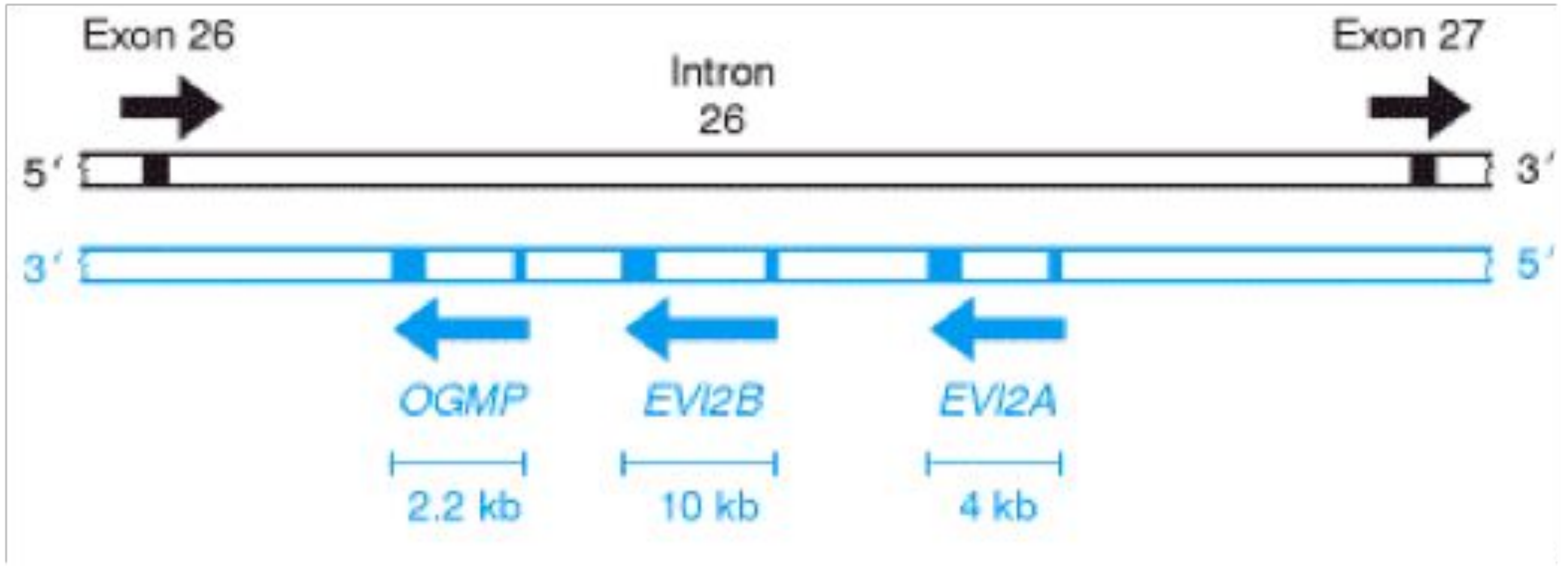
Distrofina



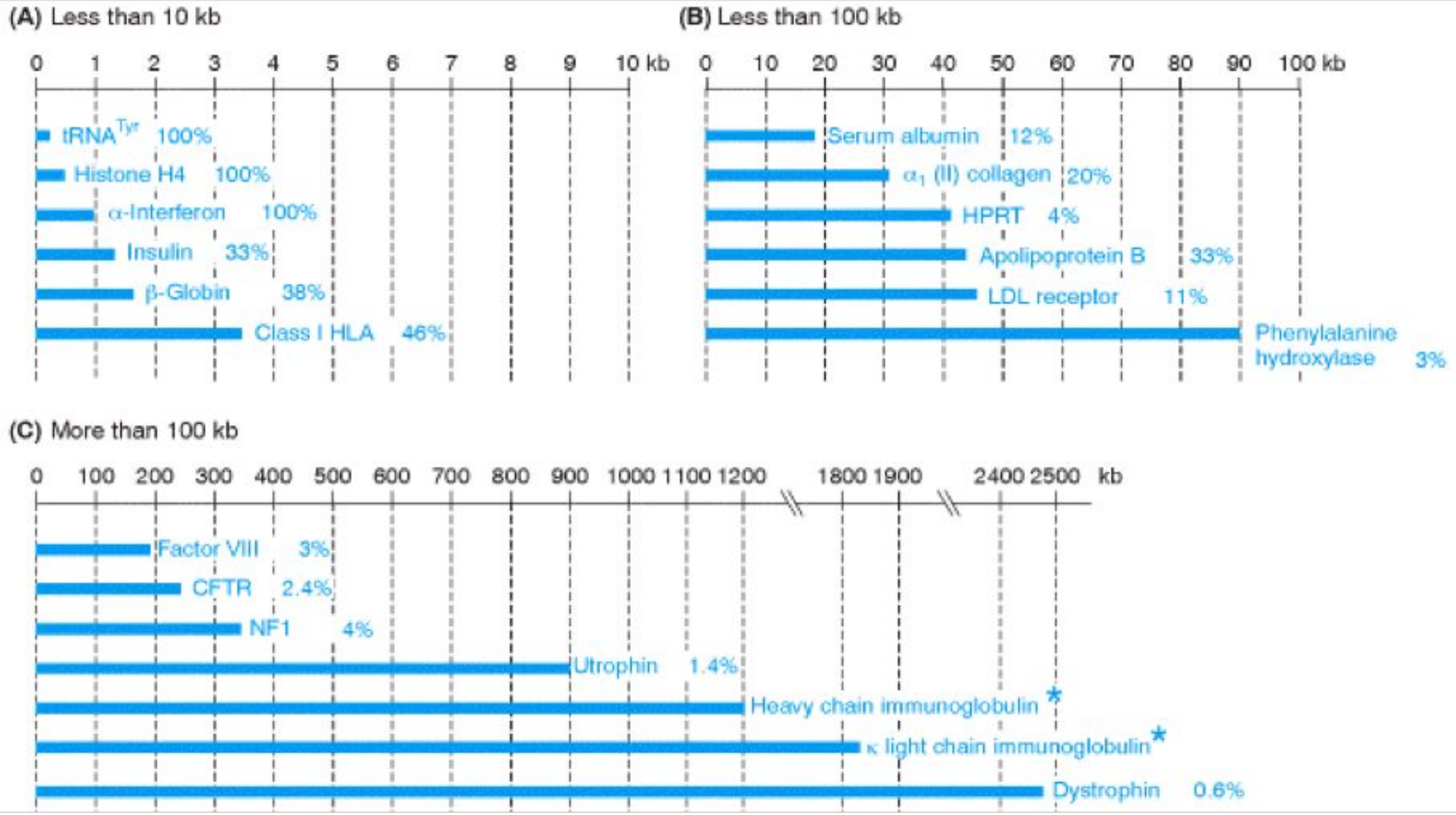
Neurofibronina 1

Referência, sense, +

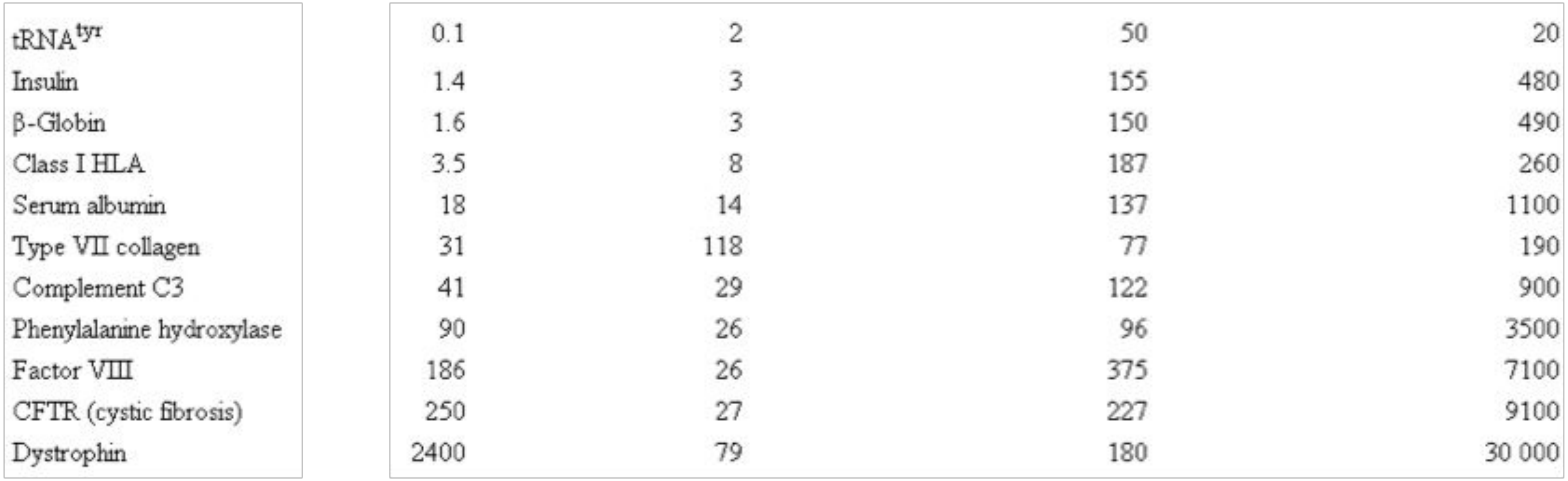
Complementar, anti-sense, -



Variação do tamanho dos genes humanos



Tamanho médio de exons e introns



Gene	Tamanho (kb)	Número de Exons	Tamanho Médio dos Exons	Tamanho Médio dos Introns
tRNA^{tyr}	0.1	2	50	20
Insulin	1.4	3	155	480
β-Globin	1.6	3	150	490
Class I HLA	3.5	8	187	260
Serum albumin	18	14	137	1100
Type VII collagen	31	118	77	190
Complement C3	41	29	122	900
Phenylalanine hydroxylase	90	26	96	3500
Factor VIII	186	26	375	7100
CFTR (cystic fibrosis)	250	27	227	9100
Dystrophin	2400	79	180	30 000

DNA repetitivo

Região não codificadora com repetição em tandem:

DNA-satélite, minissatélite e microssatélite.

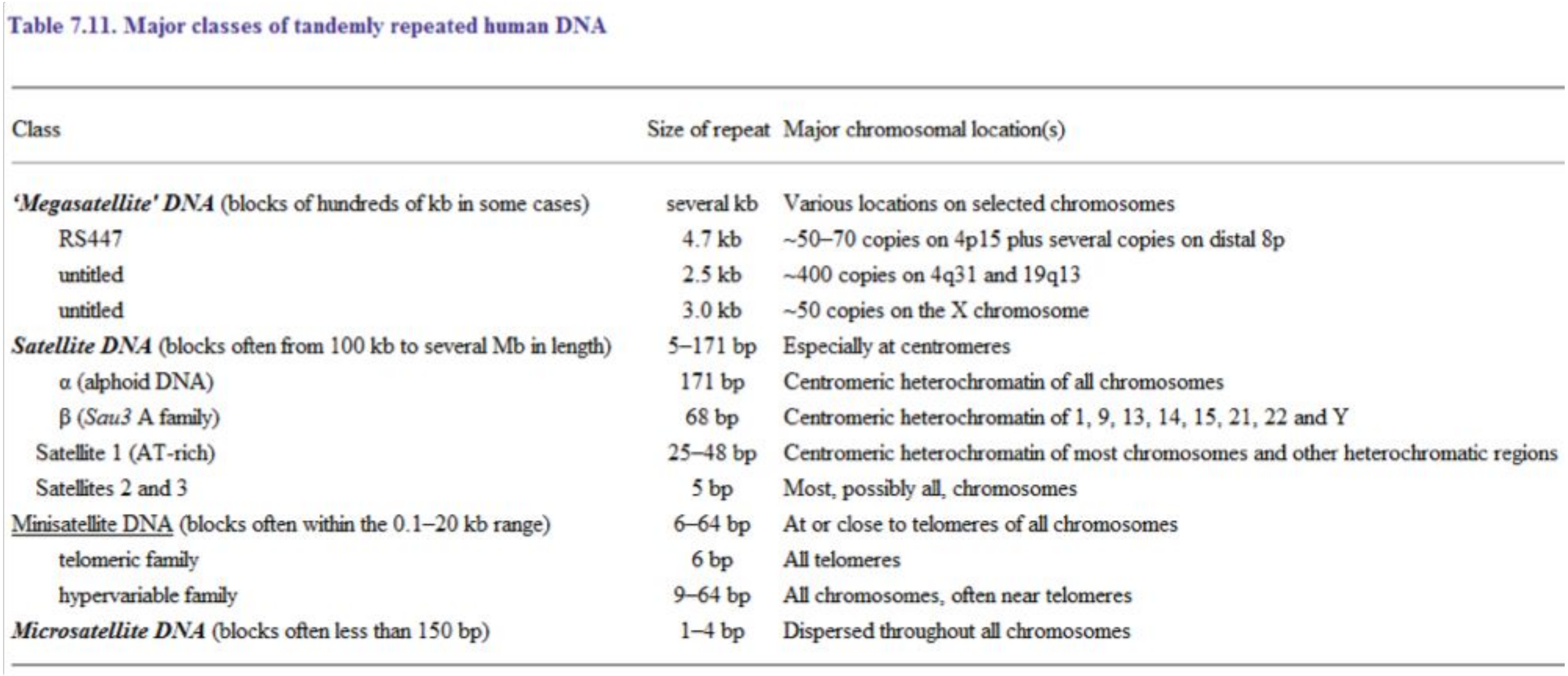


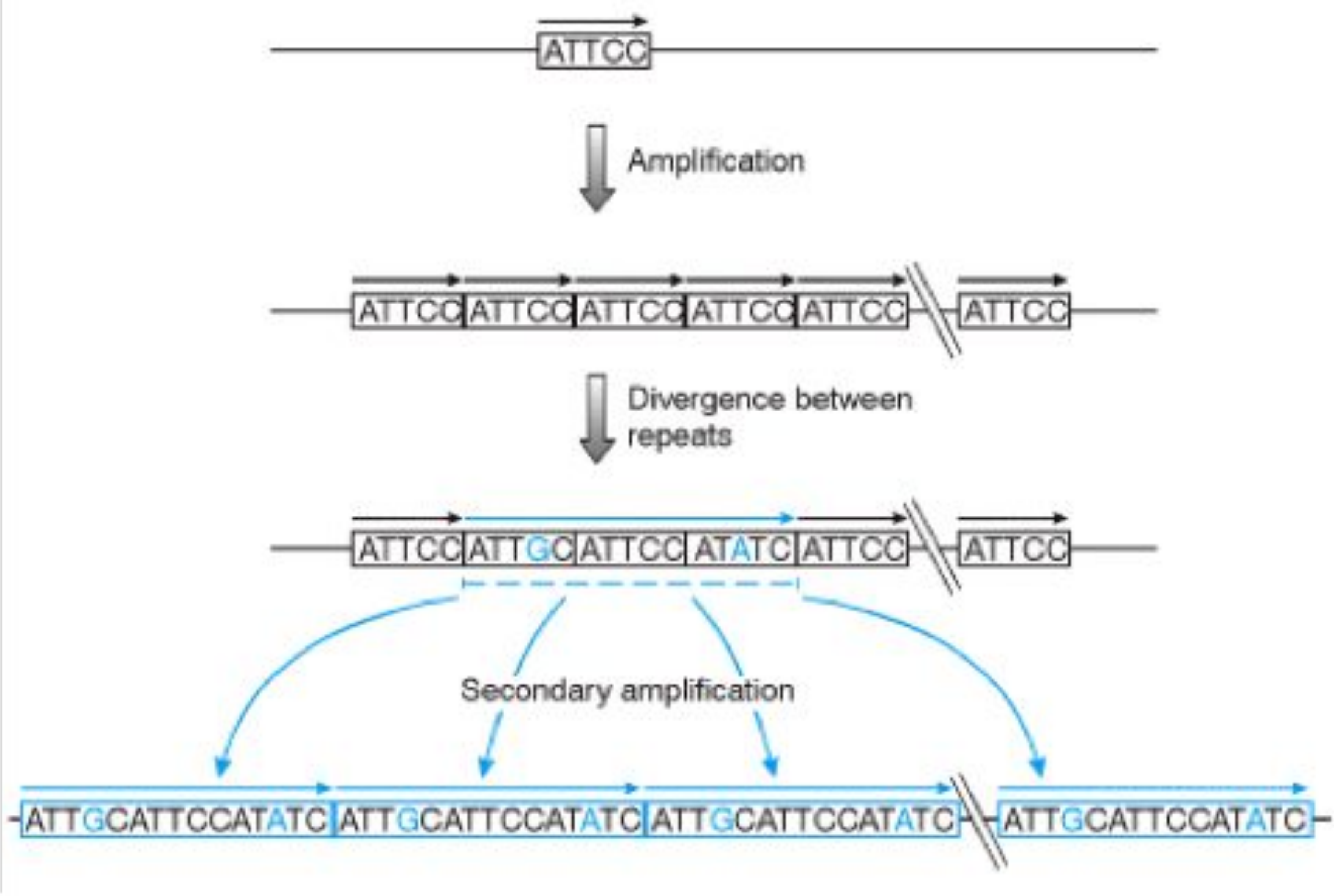
Table 7.11. Major classes of tandemly repeated human DNA

Class	Size of repeat	Major chromosomal location(s)
'Megasatellite' DNA (blocks of hundreds of kb in some cases)	several kb	Various locations on selected chromosomes
RS447	4.7 kb	~50–70 copies on 4p15 plus several copies on distal 8p
untitled	2.5 kb	~400 copies on 4q31 and 19q13
untitled	3.0 kb	~50 copies on the X chromosome
Satellite DNA (blocks often from 100 kb to several Mb in length)	5–171 bp	Especially at centromeres
α (alphoid DNA)	171 bp	Centromeric heterochromatin of all chromosomes
β (*Sau3* A family)	68 bp	Centromeric heterochromatin of 1, 9, 13, 14, 15, 21, 22 and Y
Satellite 1 (AT-rich)	25–48 bp	Centromeric heterochromatin of most chromosomes and other heterochromatic regions
Satellites 2 and 3	5 bp	Most, possibly all, chromosomes
Minisatellite DNA (blocks often within the 0.1–20 kb range)	6–64 bp	At or close to telomeres of all chromosomes
telomeric family	6 bp	All telomeres
hypervariable family	9–64 bp	All chromosomes, often near telomeres
Microsatellite DNA (blocks often less than 150 bp)	1–4 bp	Dispersed throughout all chromosomes

DNA repetitivo

Região não codificadora com repetição em tandem:

DNA-satélite, minissatélite e microssatélite.

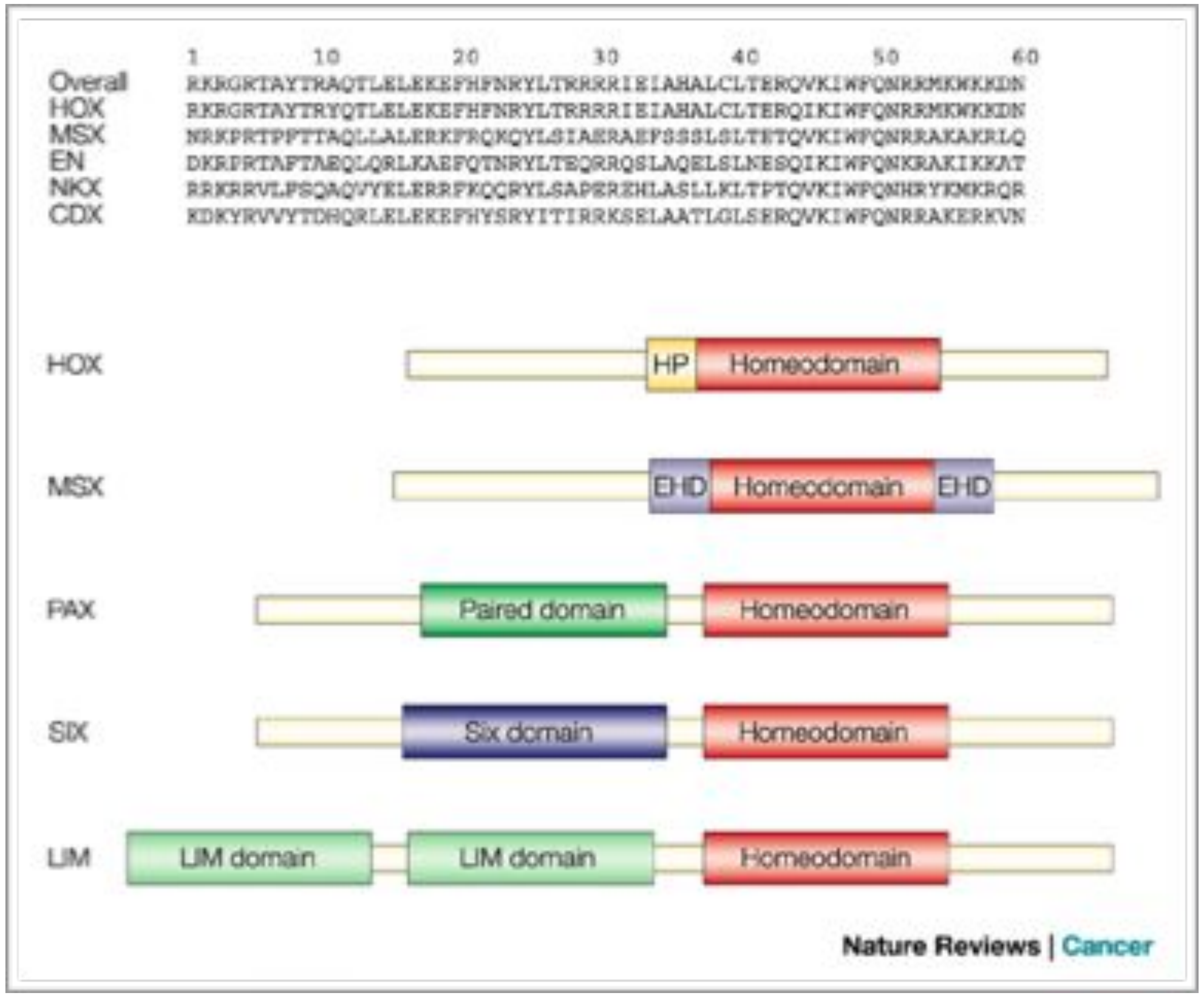


Famílias gênicas clássicas

Alto grau de homologia e identidade entre os membros: Histonas e RNAr

Famílias gênicas com domínios grandes e altamente conservados

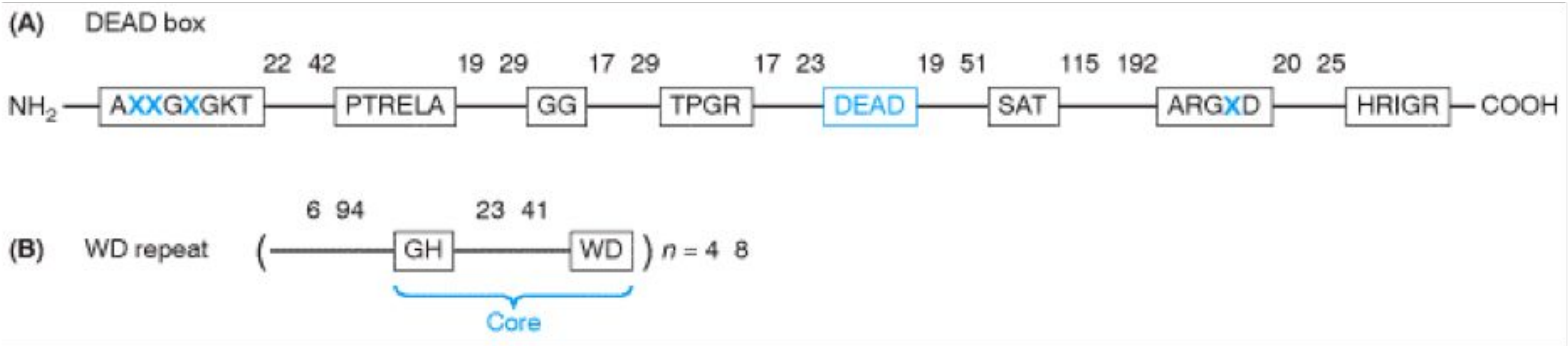
Homologia em regiões específicas



Nature Reviews | Cancer

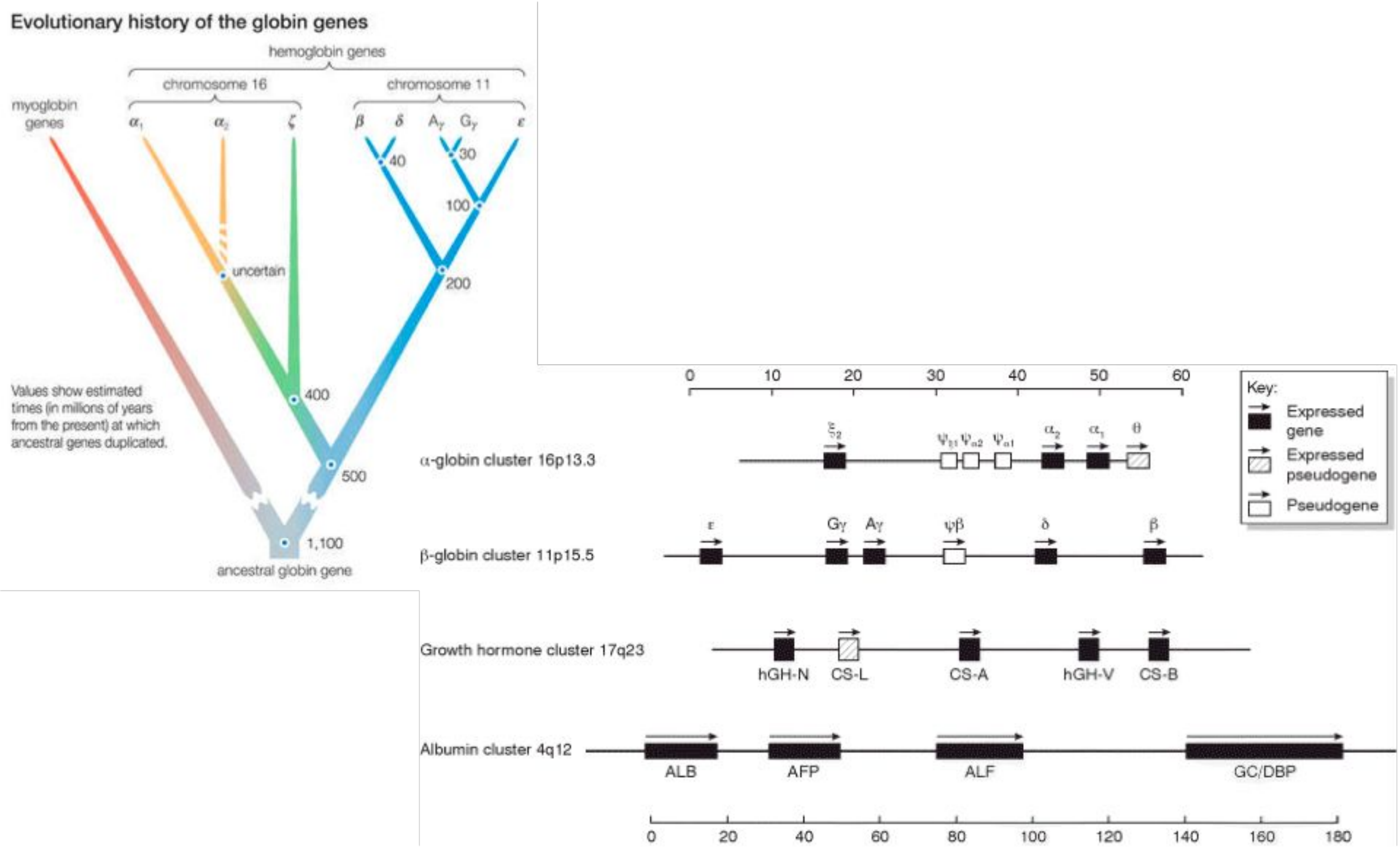
Famílias gênicas com motivos de aminoácidos conservados muito curtos

Homologia muito baixa, mas codificam produtos com função comum



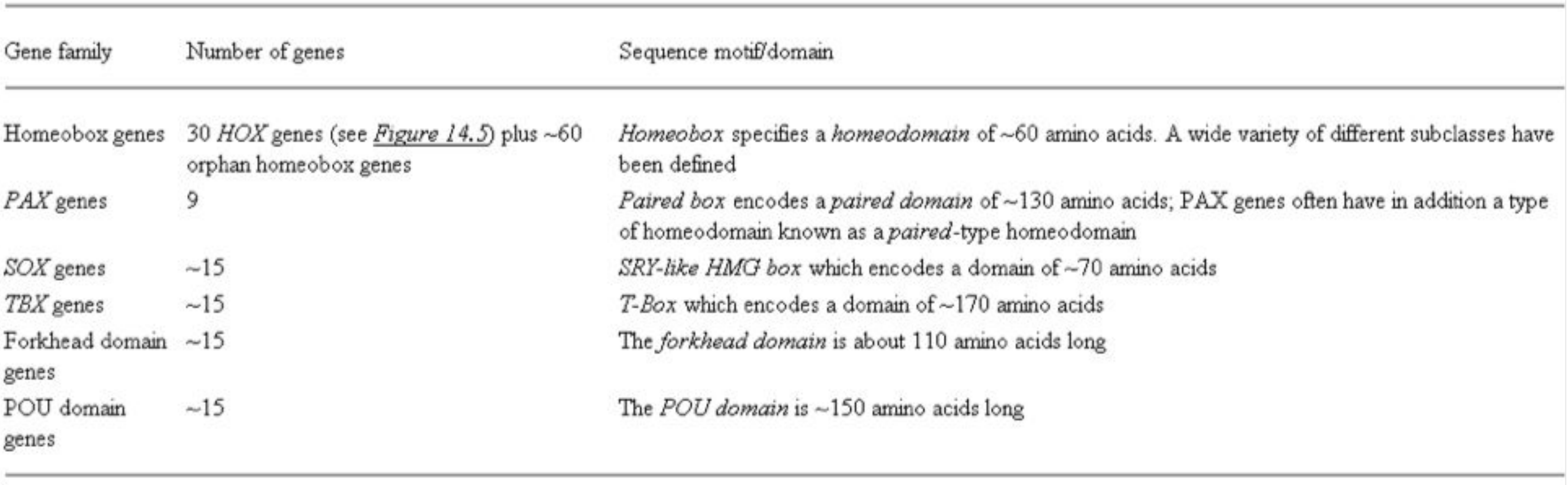
Famílias gênicas organizadas em agrupamentos únicos

Agrupamentos compostos: genes intercalados no agrupamento da família



Famílias gênicas dispersas

Originadas de antigas duplicações do genoma ou de eventos de duplicação gênica: PAX9, HOX, SOX, etc.



Gene family	Number of genes	Sequence motif/domain
Homeobox genes	30 *HOX* genes (see *Figure 14.5*) plus ~60 orphan homeobox genes	*Homeobox* specifies a *homeodomain* of ~60 amino acids. A wide variety of different subclasses have been defined
PAX genes	9	*Paired box* encodes a *paired domain* of ~130 amino acids; PAX genes often have in addition a type of homeodomain known as a *paired*-type homeodomain
SOX genes	~15	*SRY-like HMG box* which encodes a domain of ~70 amino acids
TBX genes	~15	*T-Box* which encodes a domain of ~170 amino acids
Forkhead domain genes	~15	The *forkhead domain* is about 110 amino acids long
POU domain genes	~15	The *POU domain* is ~150 amino acids long

Famílias gênicas dispersas

Originadas predominantemente por eventos de retrotransposição:

Elementos Transponíveis

Transposição Mediada por RNA
(Frequente)

Retrotransposons

Incapaz de codificar Transcriptase Reversa

Família não viral

↓

Pseudogenes processados

↓

SINE-1 (Alu family)

Capaz de codificar Transcriptase Reversa

Família viral

Retrovírus endógenos ou elementos semelhantes a retrovirus

↓

Famílias HERV/RTLV e THE-1

Retrotransposons

↓

Sem LTRs

LINE-1 (Kpn family)

Transposição Mediada por DNA
(Rara)

Transposons

Famílias gênicas dispersas

Originadas predominantemente por eventos de retrotransposição:

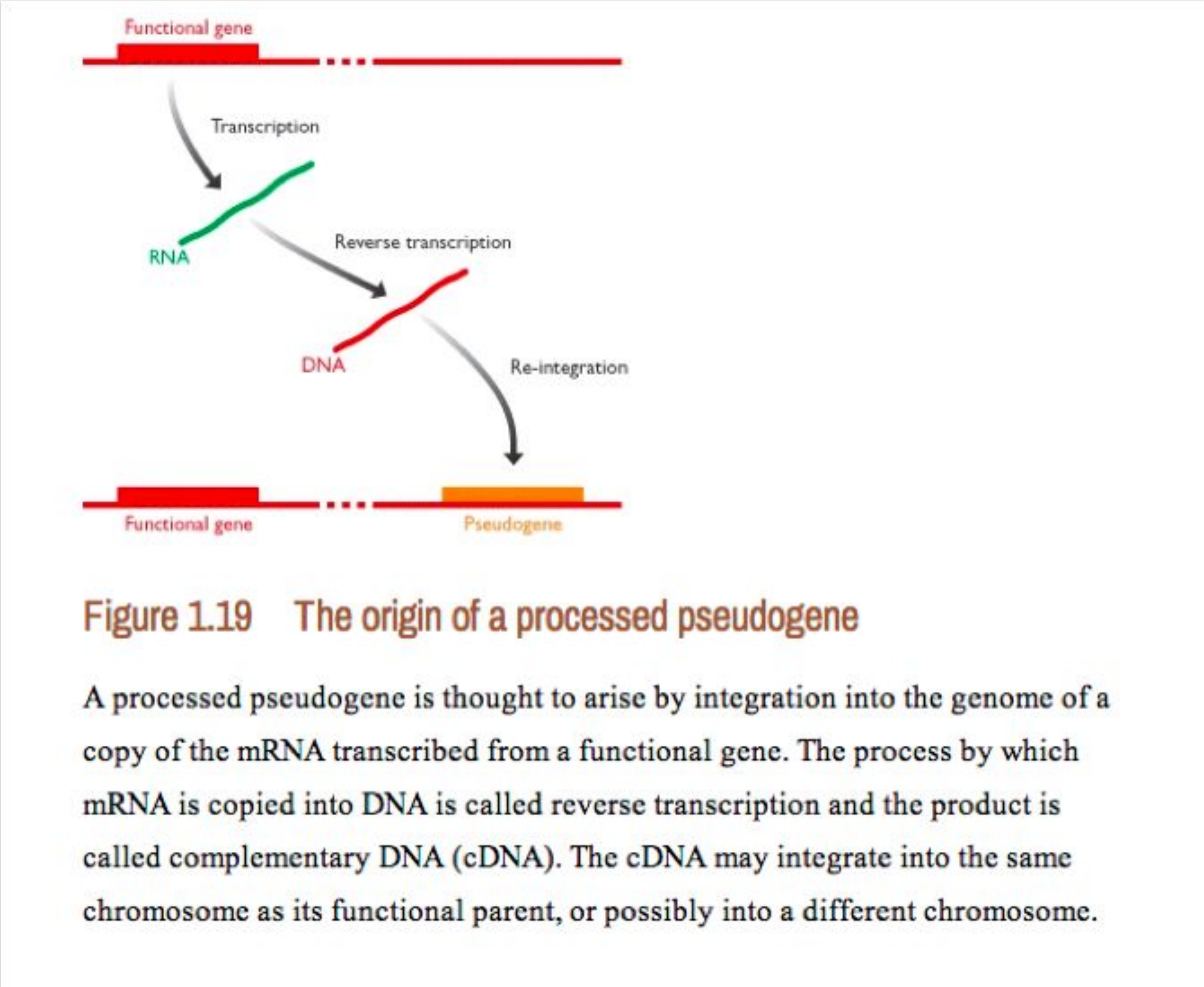


Figure 1.19 The origin of a processed pseudogene

A processed pseudogene is thought to arise by integration into the genome of a copy of the mRNA transcribed from a functional gene. The process by which mRNA is copied into DNA is called reverse transcription and the product is called complementary DNA (cDNA). The cDNA may integrate into the same chromosome as its functional parent, or possibly into a different chromosome.

Famílias gênicas dispersas

Originadas predominantemente por eventos de retrotransposição:

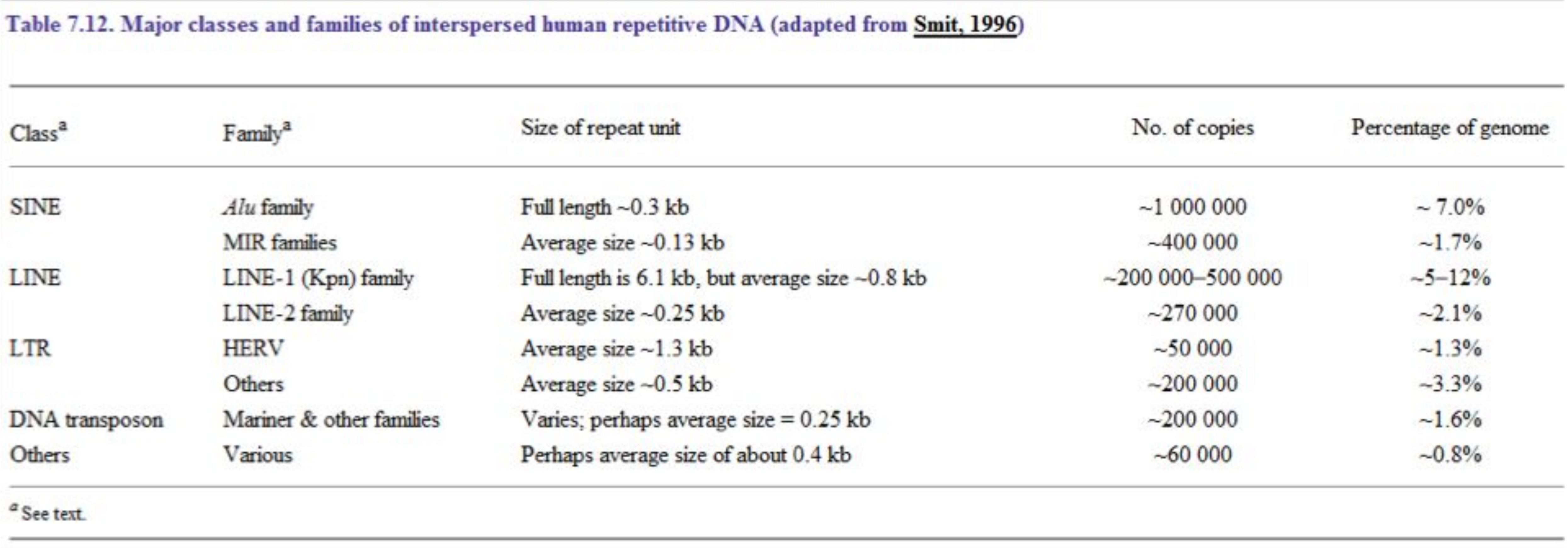


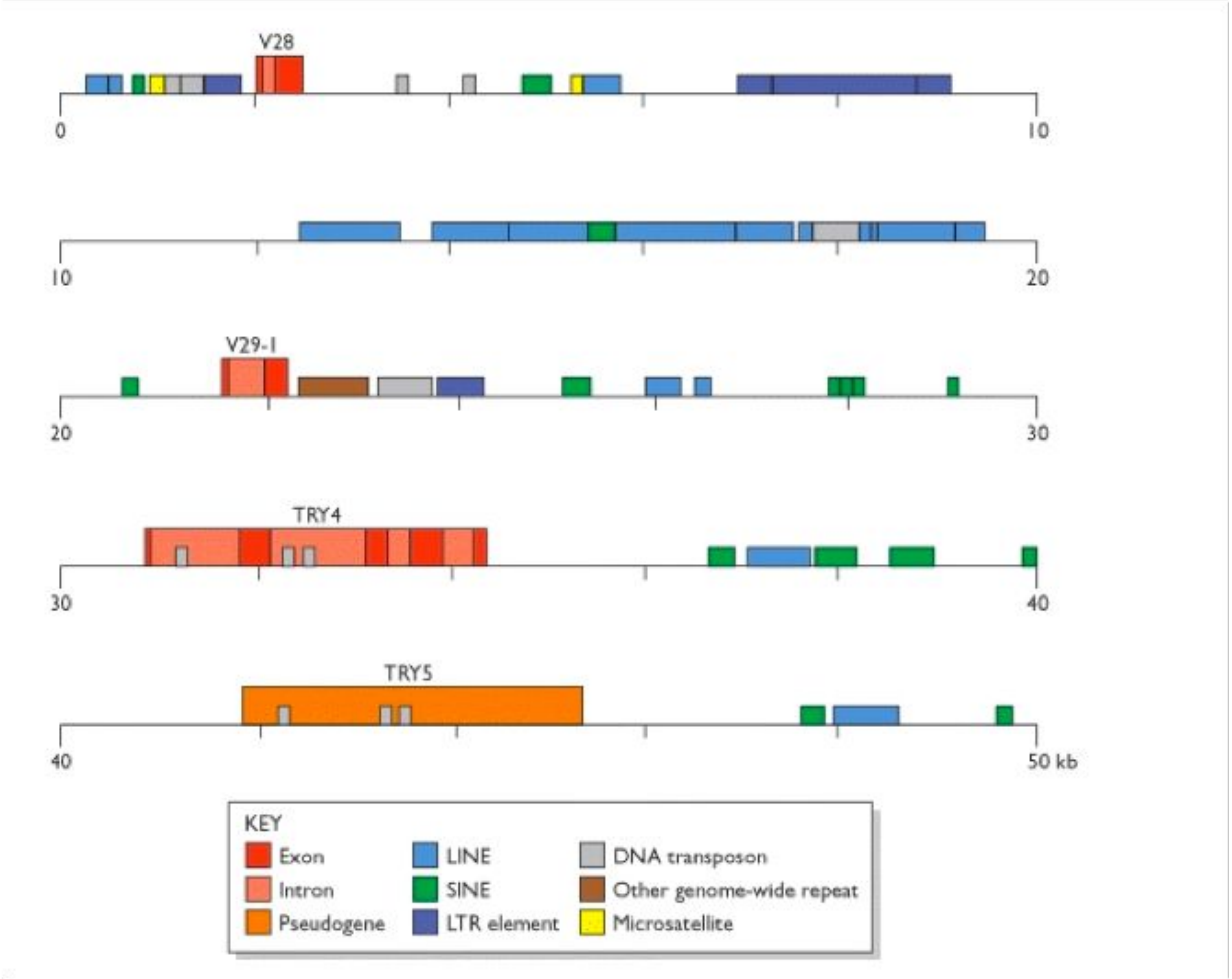
Table 7.12. Major classes and families of interspersed human repetitive DNA (adapted from Smit, 1996)

Class[a]	Family[a]	Size of repeat unit	No. of copies	Percentage of genome
SINE	*Alu* family	Full length ~0.3 kb	~1 000 000	~ 7.0%
	MIR families	Average size ~0.13 kb	~400 000	~1.7%
LINE	LINE-1 (Kpn) family	Full length is 6.1 kb, but average size ~0.8 kb	~200 000–500 000	~5–12%
	LINE-2 family	Average size ~0.25 kb	~270 000	~2.1%
LTR	HERV	Average size ~1.3 kb	~50 000	~1.3%
	Others	Average size ~0.5 kb	~200 000	~3.3%
DNA transposon	Mariner & other families	Varies; perhaps average size = 0.25 kb	~200 000	~1.6%
Others	Various	Perhaps average size of about 0.4 kb	~60 000	~0.8%

[a] See text.

Famílias gênicas dispersas

Originadas predominantemente por eventos de retrotransposição:



Famílias gênicas dispersas

Originadas predominantemente por eventos de retrotransposição: