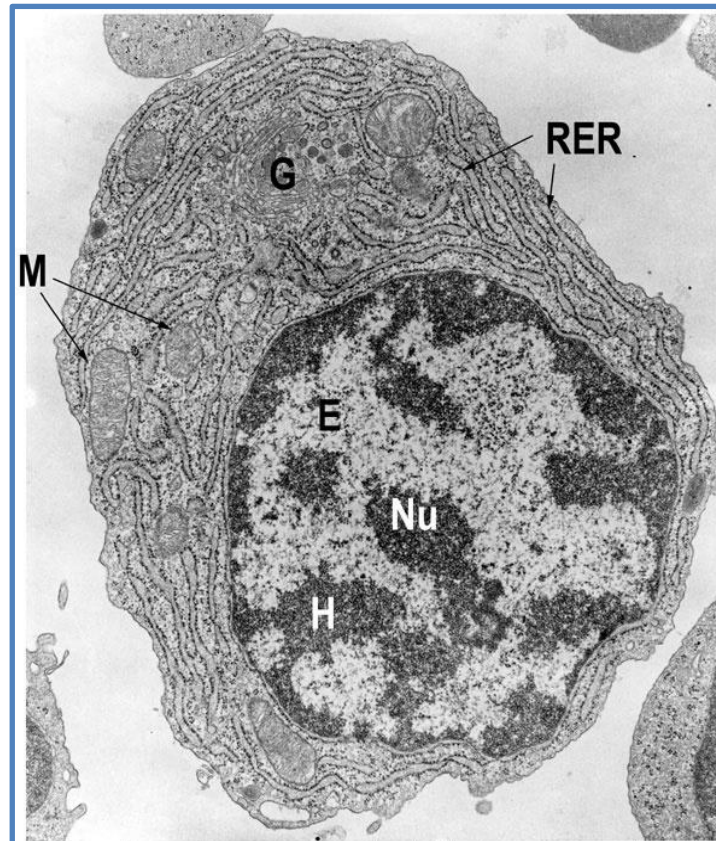


# **TEMA 1.-**

Organización del genoma humano.  
Genoma mitocondrial. Genoma nuclear.  
Tipos de secuencias.

El término **Genoma humano**, describe la información genética total (contenido de ADN) de las células humanas.

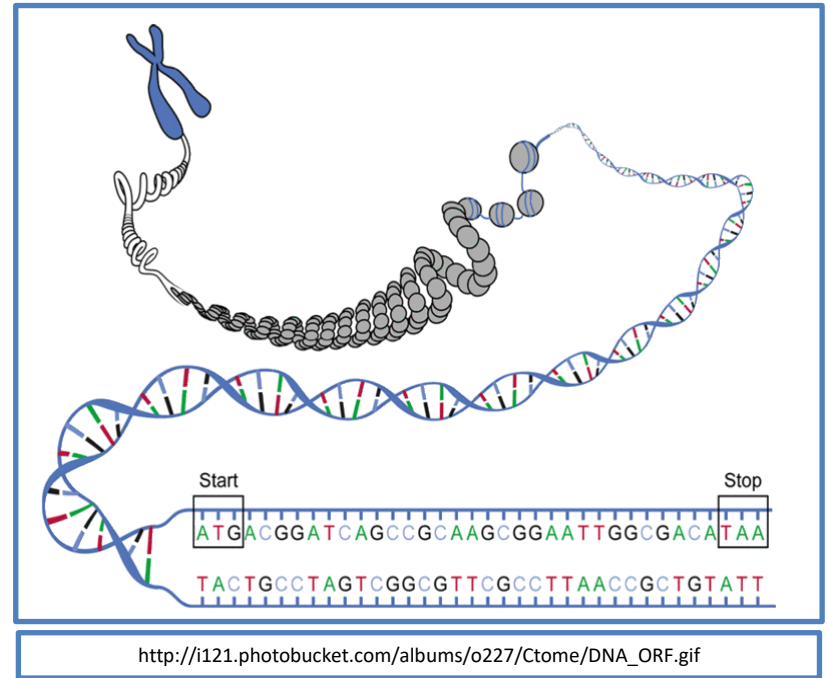


Nu-nucleus, E-euchromatin, H-heterochromatin, M-mitochondria, RER-rough endoplasmic reticulum, G-golgi complex

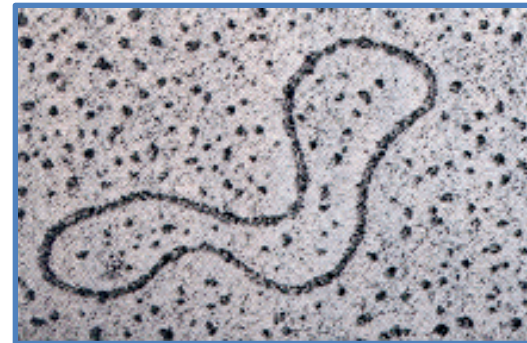
[http://missinglink.ucsf.edu/lm/IDS\\_101\\_histo\\_resource/images/cell\\_structur\\_e\\_lab\\_micrograph\\_B-labelled.jpg](http://missinglink.ucsf.edu/lm/IDS_101_histo_resource/images/cell_structur_e_lab_micrograph_B-labelled.jpg)

Incluye dos genomas:

- Un gran genoma nuclear ( $\approx 3.200.000.000$  pb), repartido en 24 moléculas de ADN duplexo y lineal (22 autosomas + cromosoma X + cromosoma Y)



- Un genoma mitocondrial (16.569 pb), formado por ADN duplexo y circular



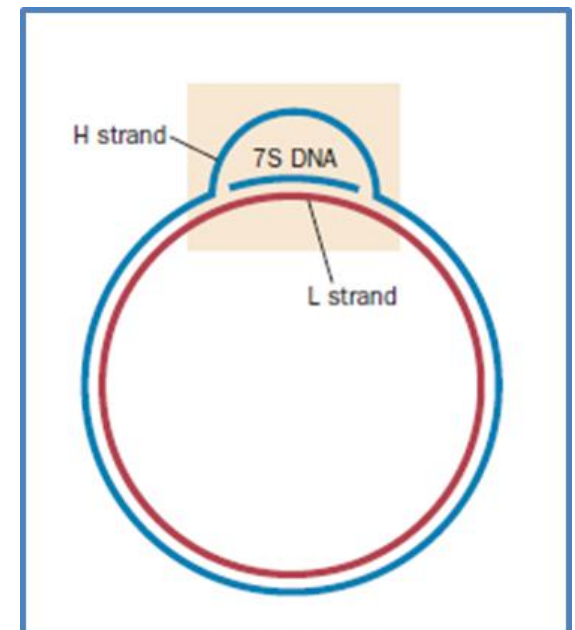
# Genoma mitocondrial

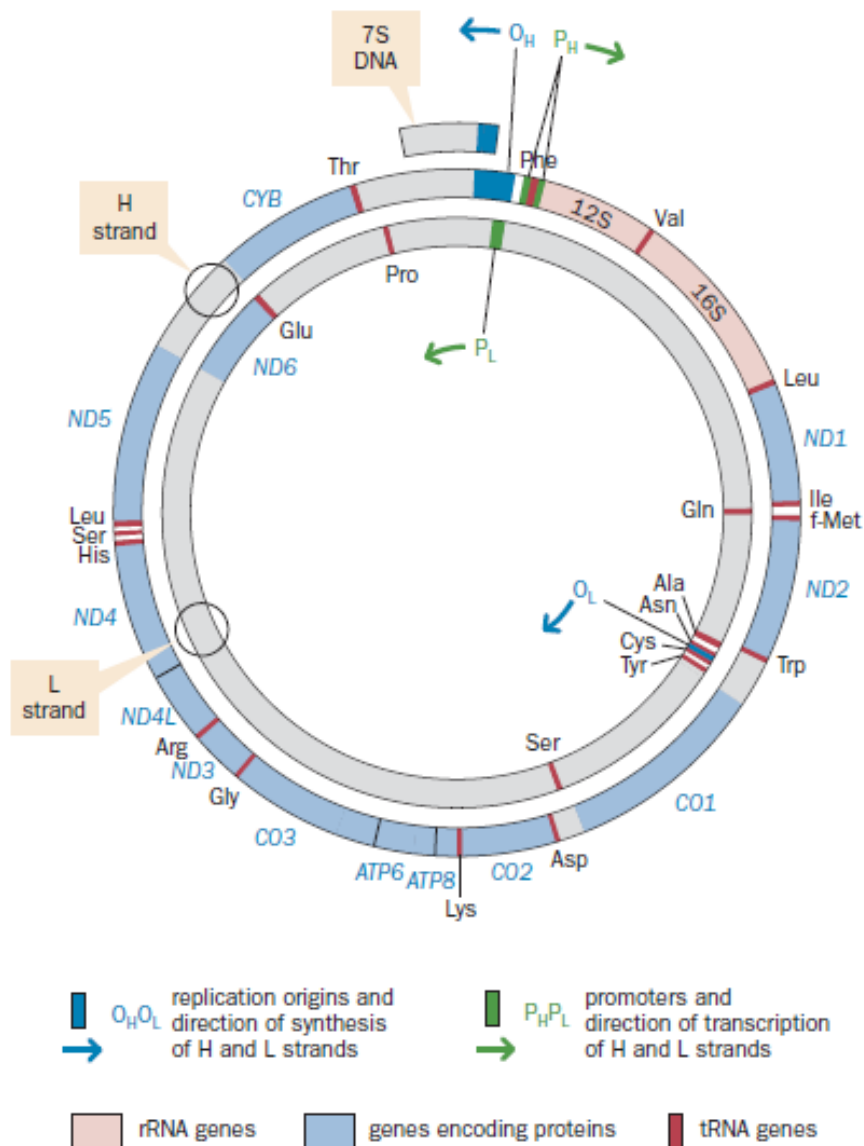
- ADN duplexo circular, 16.569 pb. Cada mitocondrial tiene 2-10 copias.
- Cada célula tiene muchas mitocondrias. Según el tipo celular el ADN mitocondrial supone entre el 0,5% y el 30% (en ovocitos) del ADN total de la célula.
- Secuenciado por Anderson et al. 1981
- % G-C= 44 %
- Cadenas H y L
- Lazo D
- Herencia materna

## Replicación

Unidireccional, comienza en el lazo D la de la cadena H.

A los 2/3 se encuentra el origen de replicación de la cadena L.





37 genes: 28 H + 9 L

- 22 ARNt
- 2 ARNr (12S, 16S)
- 13 ARNm

### Transcripción:

- Dos promotores PH y PL
- Maduración de los ARN
- Sin intrones

### Traducción:

Excepciones al código genético universal.

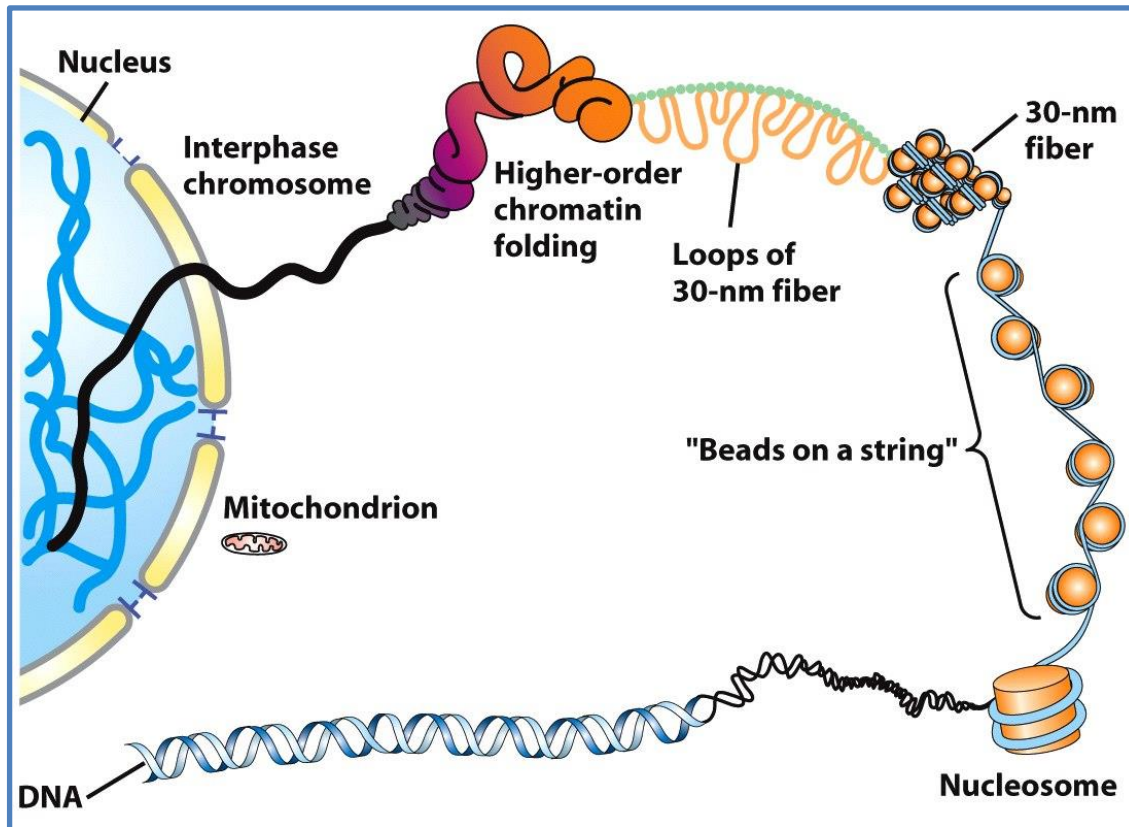
- AGA → codón de stop (??)
- AGG → codón de stop (??)
- AUA → codón para met
- UGA → codón para trp en lugar de stop.

[www.mitomat.org](http://www.mitomat.org) (Temperly, 2010)

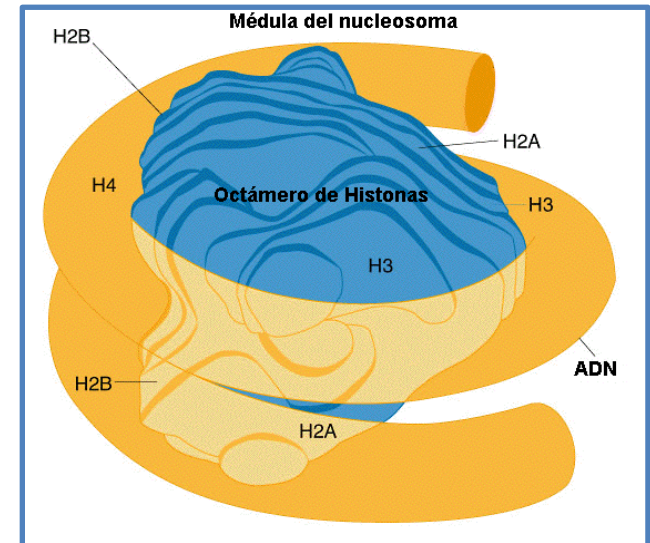
# Genoma nuclear

24 moléculas de ADN duplexo y lineal ( $\approx 3.1$  Gpb), asociado con proteínas principalmente de tipo histona para dar lugar a la fibra de cromatina.

- 22 cromosomas autosómicos
- Cromosoma X y cromosoma Y



2.9 Gpb de eucromatina  
0,2 Gpb de heterocromatina



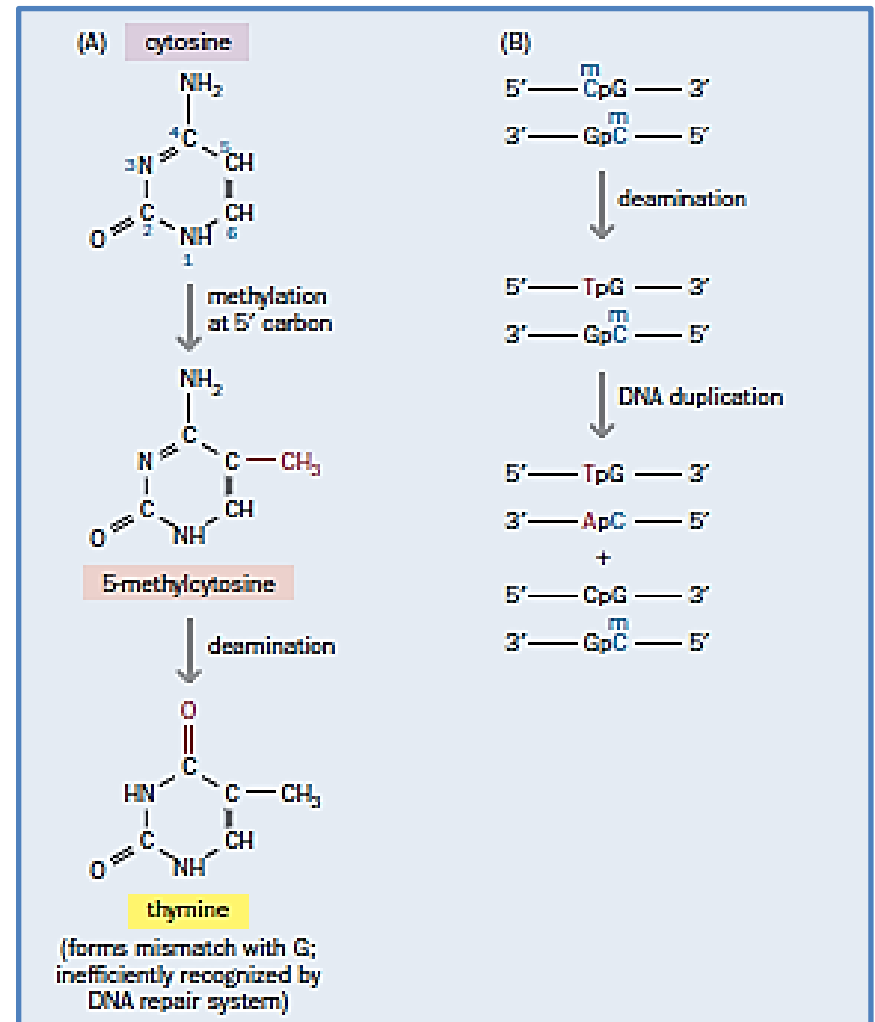


Composición de bases del ADN nuclear:

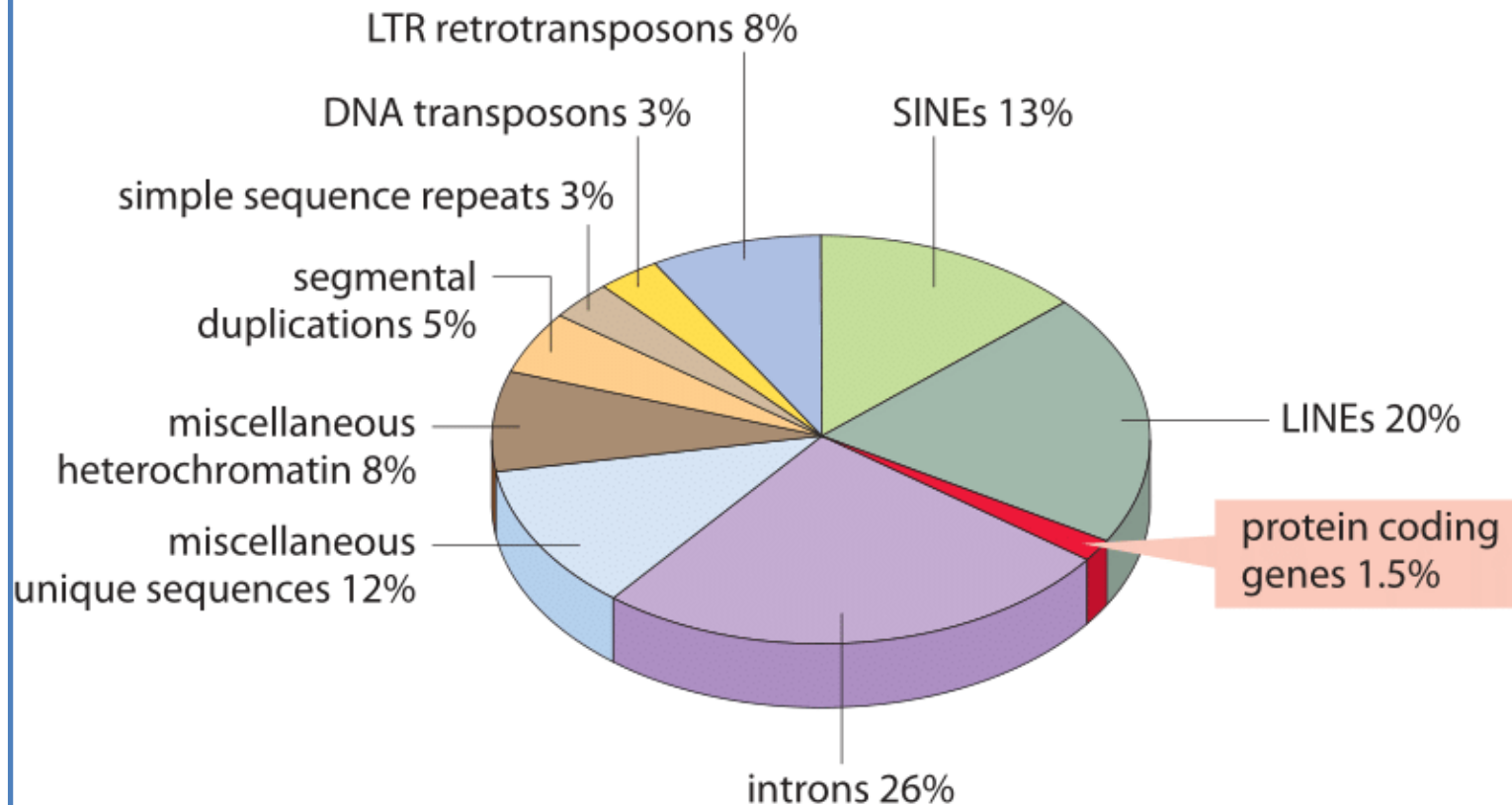
41 % (G+C), con variaciones intercromosómicas. Ej. Cromosomas 4 y 13 % (G+C)= 38%. Cromosoma 19 49 % (G+C).

Variaciones intracromosómicas del % de (G+C).

Variaciones en la proporción de ciertas combinaciones de dinucleótidos. Disminución del dinucleótido CpG, excepto en las islas CpG



## main components of the human genome



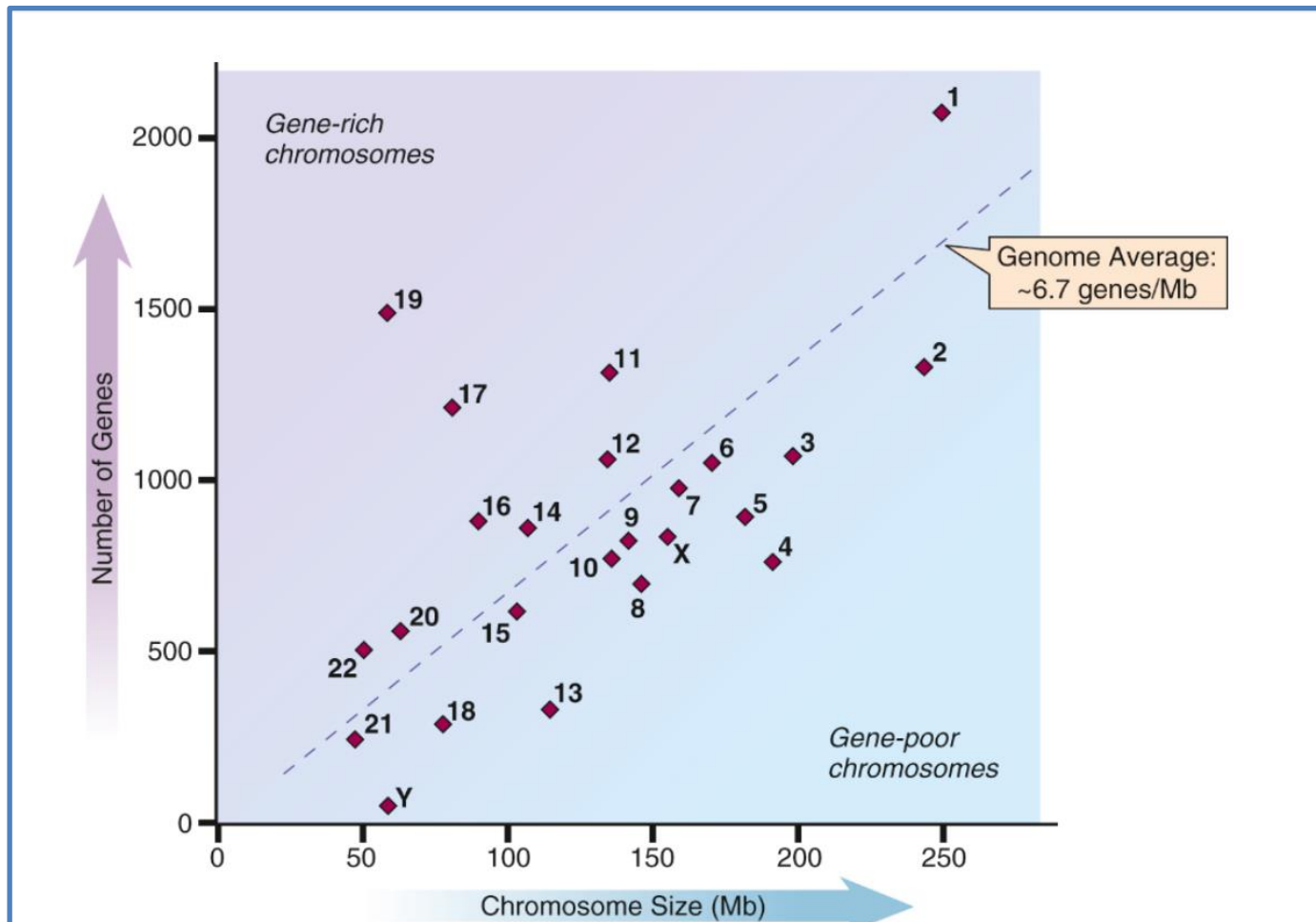
Se estima que el genoma nuclear contiene unos 22.000 genes que codifican proteínas, aunque en muchos casos, cada uno de ellos puede dar lugar a dos o más productos que en su conjunto constituyen el proteoma. Muchos genes tienen como productos finales moléculas de ARN son los ncRNA y se estima que puede haber entre 20.000 y 25.000 en todo el genoma con funciones variadas.

Nuevas revisiones estiman que el número de genes codificantes de proteínas sería unos 19.000.

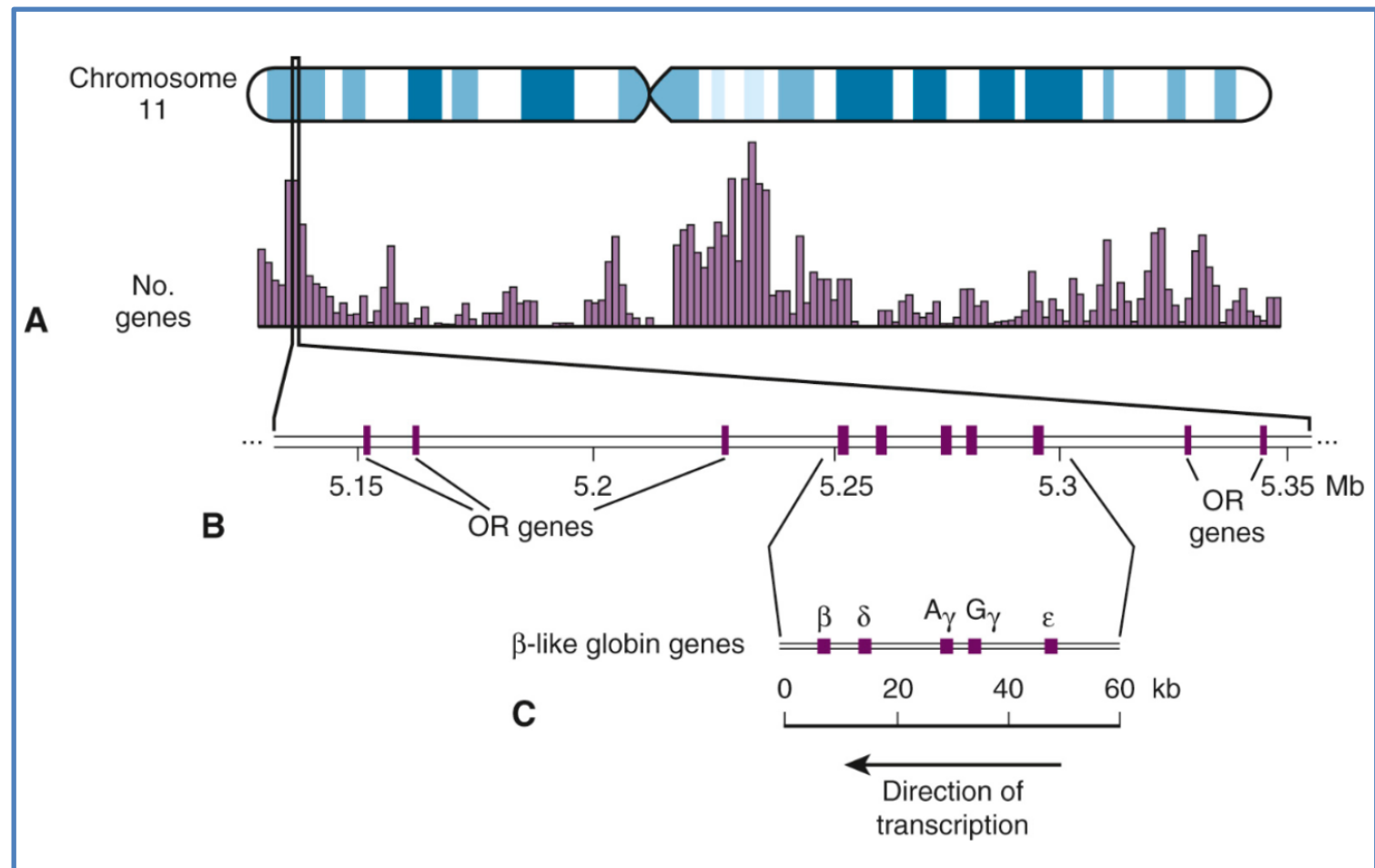
**7070–7084** *Nucleic Acids Research*, 2018, Vol. 46, No. 14 Published online 30 June 2018. **Loose ends: almost one in five human genes still have unresolved coding status**

Federico Abascal, David Juan, Irwin Jungreis, Laura Martinez, Maria Rigau, Jose Manuel Rodriguez, Jesus Vazquez and Michael L. Tress,





Dotted diagonal line corresponds to the average density of genes in the genome, approximately 6.7 protein-coding genes per megabase (Mb). Chromosomes that are relatively gene rich are above the diagonal and trend to the upper left. Chromosomes that are relatively gene poor are below the diagonal and trend to the lower right.



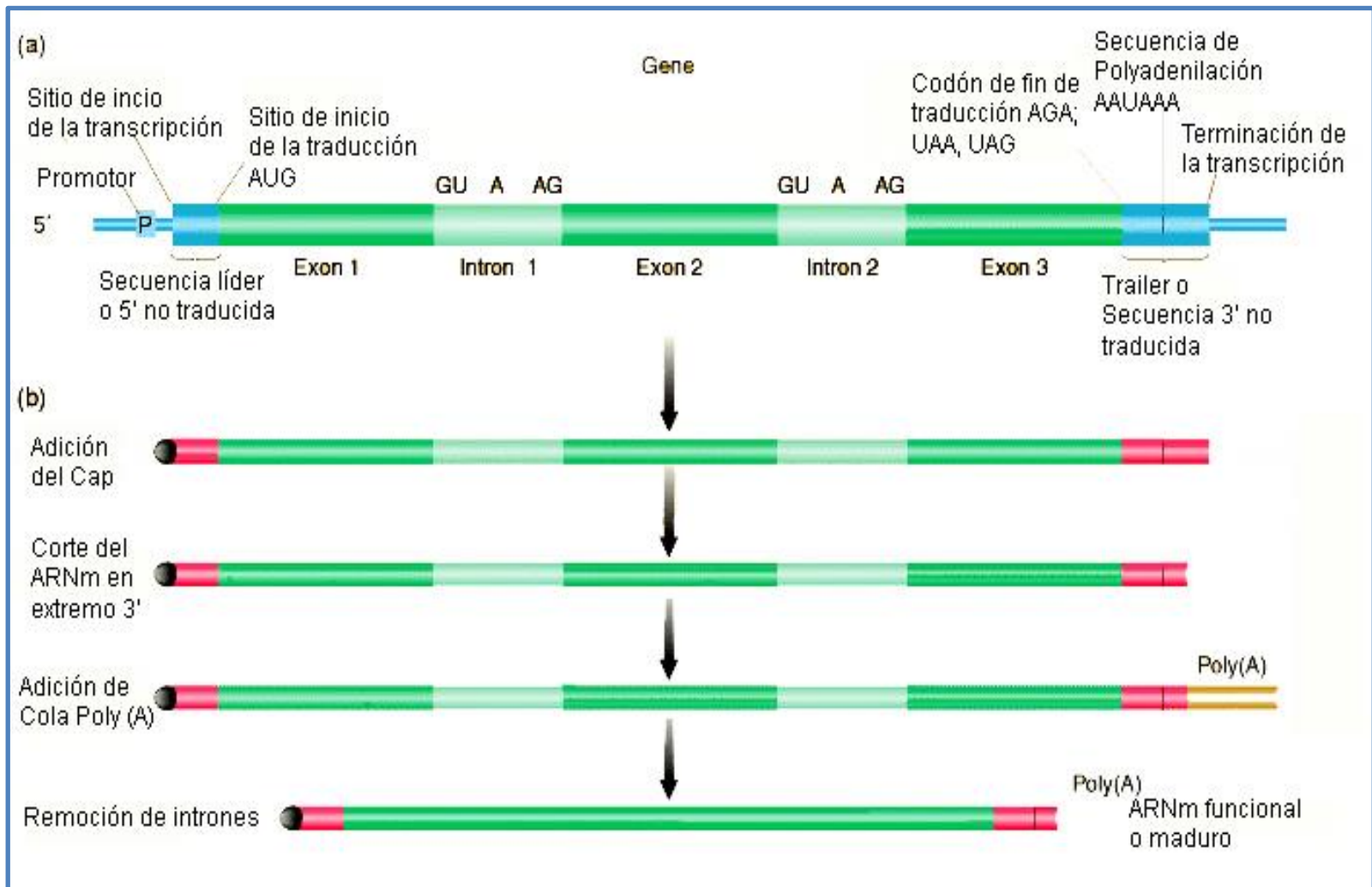
A, The distribution of genes is indicated along the chromosome and is high in two regions of the chromosome and low in other regions. B, An expanded region from 5.15 to 5.35 Mb (measured from the short-arm telomere), which contains 10 known protein-coding genes, five belonging to the olfactory receptor (OR) gene family and five belonging to the globin gene family. C, The five β-like globin genes expanded further.

## Genes que codifican para proteínas

Gran variación en tamaño y en organización interna (intrones-exones).

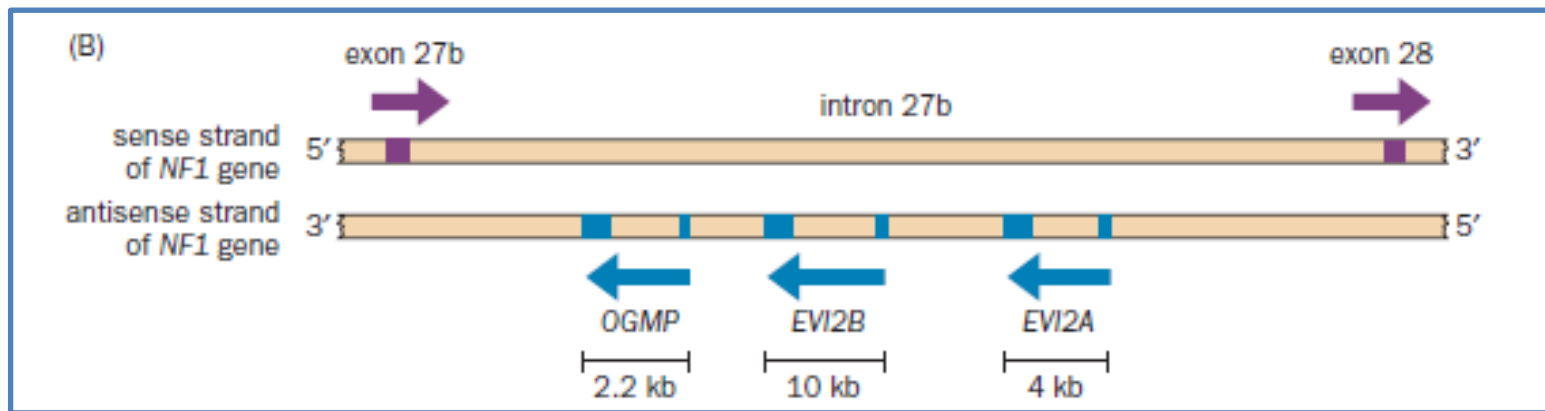
**TABLE 9.4 STRUCTURAL VARIATION IN SIZE AND ORGANIZATION OF HUMAN PROTEIN-C**

Human protein	Size of protein (no. of amino acids)	Size of gene (kb)	No. of exons	Coding DNA (%)
SRY	204	0.9	1	94
$\beta$ -Globin	146	1.6	3	38
p16	156	7.4	3	17
Serum albumin	609	18	14	12
Type VII collagen	2928	31	118	29
p53	393	39	10	6.0
Complement C3	1641	41	29	8.6
Apolipoprotein B	4563	45	29	31
Phenylalanine hydroxylase	452	90	26	3
Factor VIII	2351	186	26	3
Huntingtin	3144	189	67	8.0
RB1 retinoblastoma protein	928	198	27	2.4
CFTR (cystic fibrosis transmembrane receptor)	1480	250	27	2.4
Titin	34,350	283	363	40
Utrophin	3433	567	74	2.2
Dystrophin	3685	2400	79	0.6

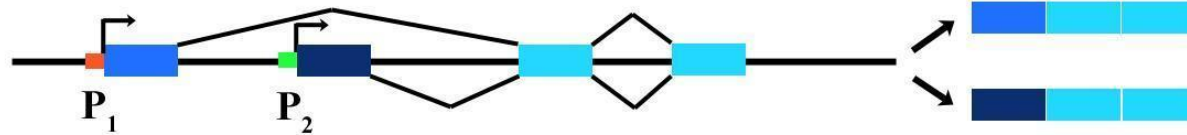


Aproximadamente el 9% de los genes que codifican proteínas se solapan parcialmente entre si, en la mayor parte de los casos (90%) porque se transcriben ambas cadenas.

En algunos casos, se han descrito la existencia de genes dentro de los intrones de otro gen, como por ejemplo en el gen NF1 (neurofibromatosis de tipo I), en el que hay que genes pequeños que se transcriben a partir de la otra cadena.



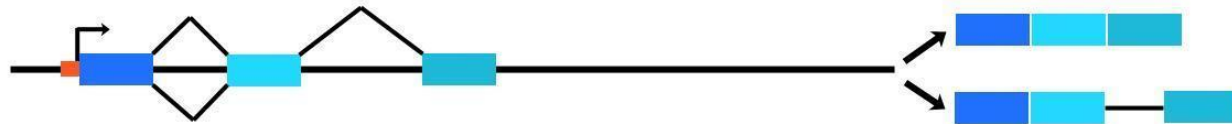
(a) Selección de promotores alternativos (Ej: Transcrito primario de miosina)



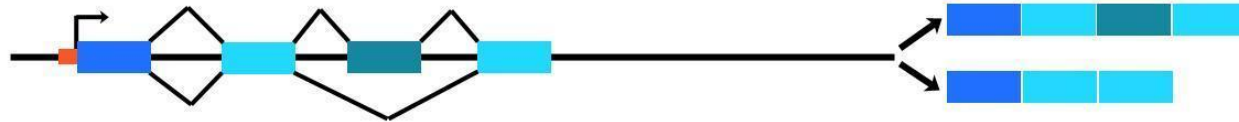
(b) Selección de sitios de poliadenilación alternativo (Ej: Transcrito primario de tropomiosina)

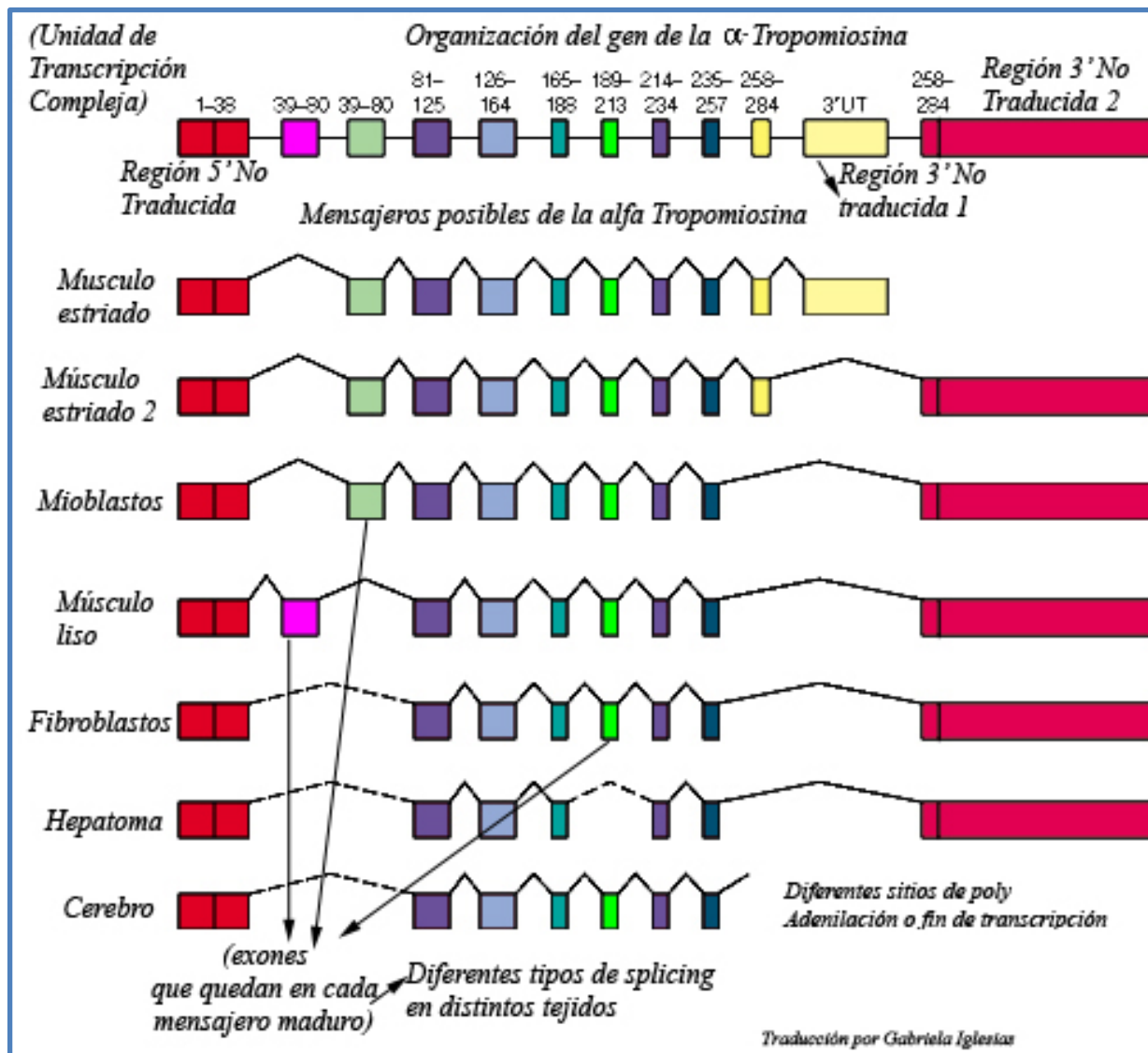


(c) Retención de intrones



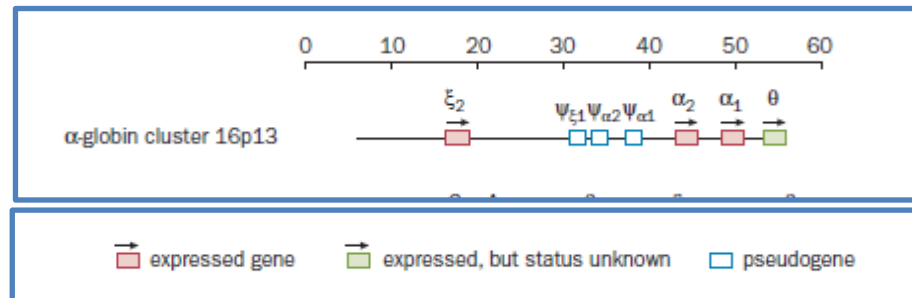
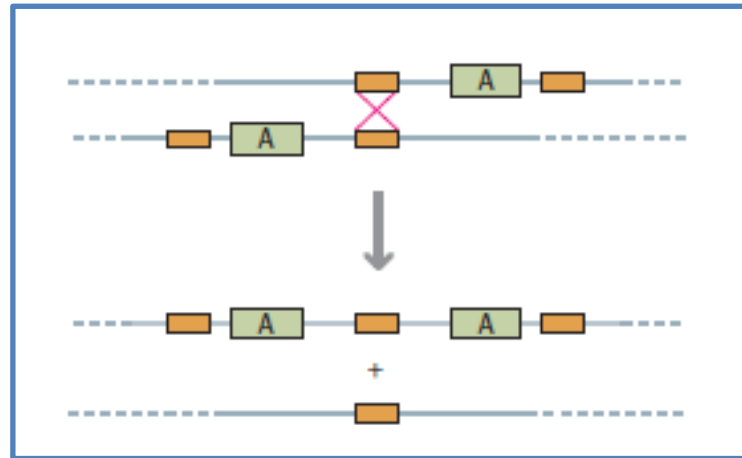
(d) Ayustación de exones (Ej: Transcrito primario de la troponina)





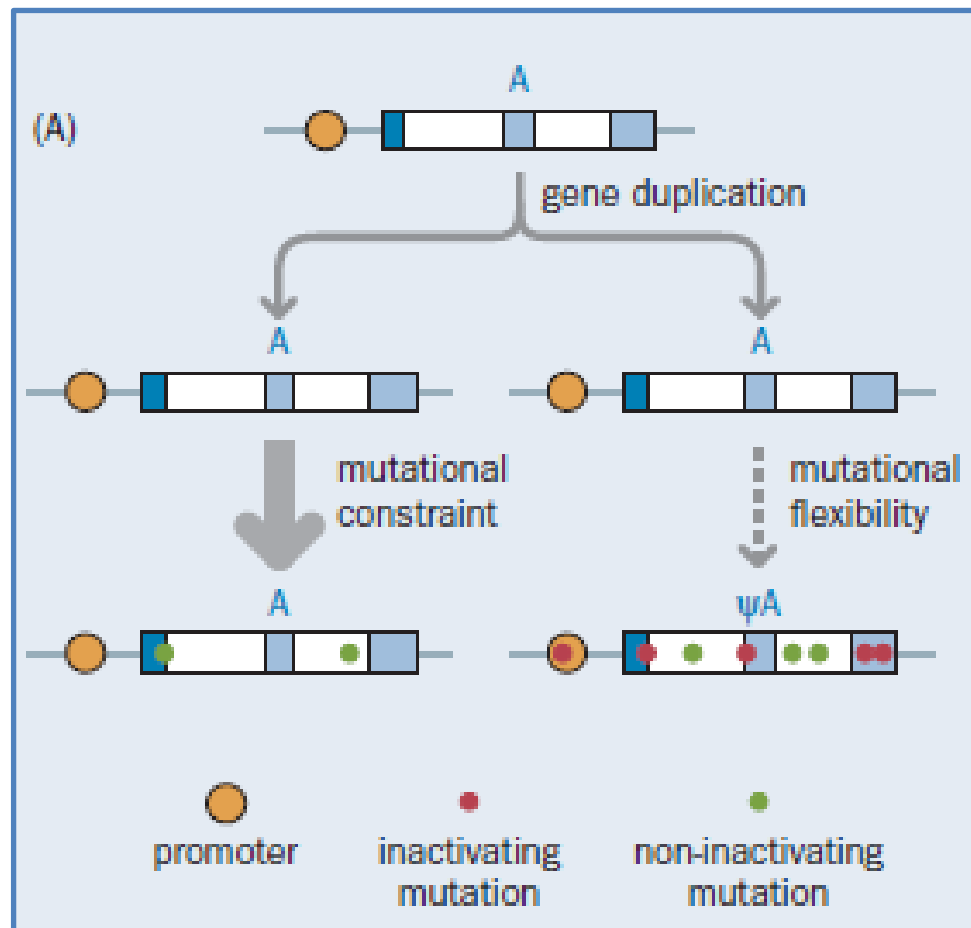


## Familias génicas: en clúster o dispersas

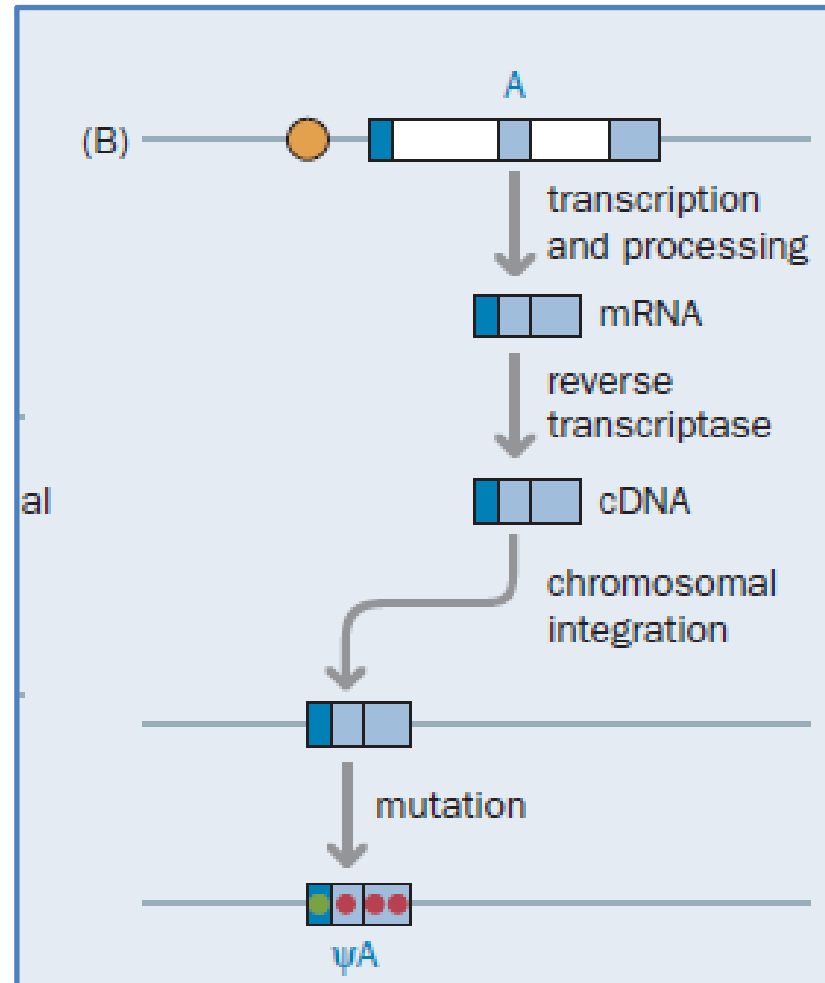


Ejemplo de familia génica con múltiples cluster son los genes que codifican receptores olfativos, que se encuentran localizados en diferentes cromosomas (900 genes en 25 localizaciones cromosómicas).

La duplicación génica también puede dar lugar a pseudogenes, con la secuencia completa del gen o con sólo algunas partes por ejemplo algún exón (fragmentos génicos).



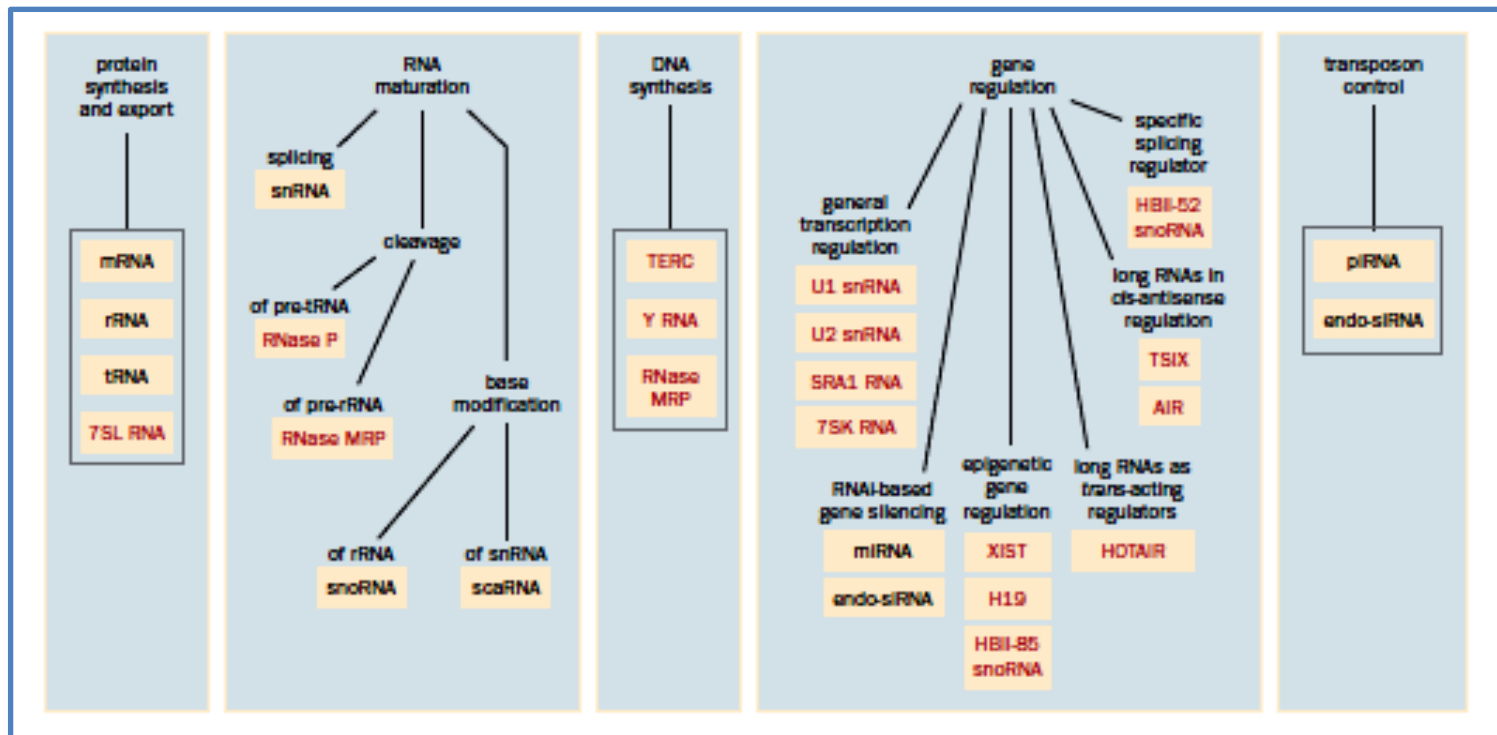
Pseudogenes procesados. Se trata de copias defectuosas de genes que contienen sólo las secuencias exónicas, careciendo por tanto de intrones y de la secuencia promotora. Se originan por retrotranscripción



## Genes que codifican para ARN

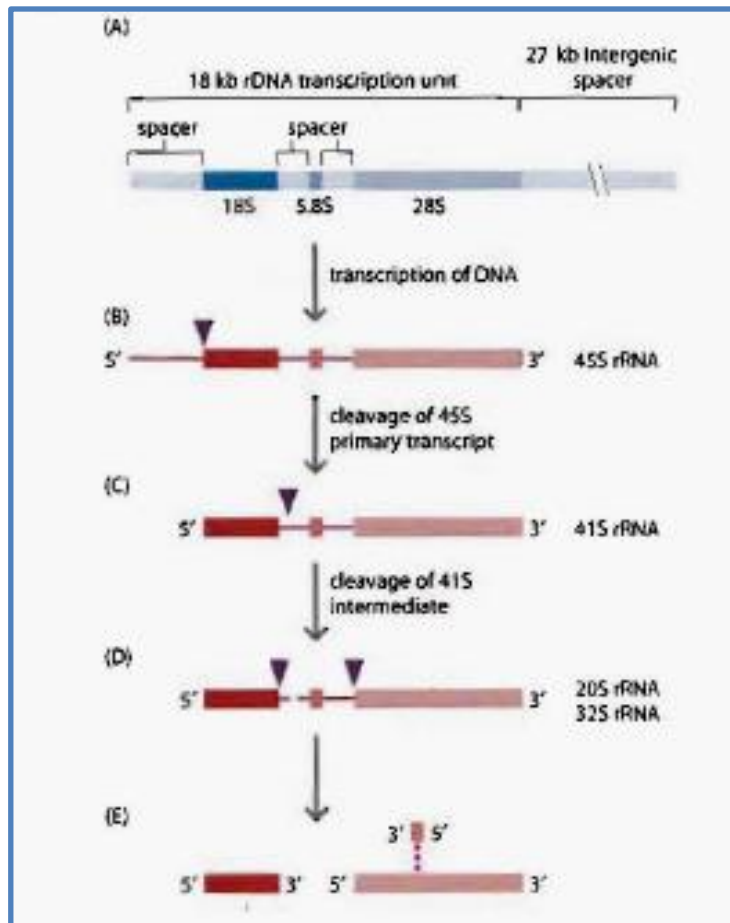
Hasta hace pocos años se consideró que los genes cuyo producto final eran moléculas de ARN no codificante (ncRNA) sólo participaban en la síntesis de proteínas: ARNr; ARNt...

Actualmente se ha visto que entre el 85 y el 90% del genoma correspondiente a zonas eucromáticas se transcribe al menos a partir de una de las cadenas de ADN.



Hay más de 1000 genes en el hombre que codifican para ARNr o ARNt, que están situados principalmente en clusters.

## ARN ribosómico



**Figure 1.22** The major rRNA species are synthesized by cleavage of a shared primary transcript. (A) In human cells, the 18S, 5.8S, and 28S rRNAs are encoded by a single transcription unit that is 13 kb long. It occurs within tandem repeat units of about 40 kb that also includes a roughly 27 kb non-transcribed (intergenic) spacer. (B) Transcription by RNA polymerase I produces a 13 kb primary transcript (45S rRNA) that then undergoes a complex series of post-transcriptional cleavages. (C–E) Ultimately, individual 18S, 28S, and 5.8S rRNA molecules are released. The 18S rRNA will form part of the small ribosomal subunit. The 5.8S rRNA binds to a complementary segment of the 28S rRNA; the resulting complex will form part of the large ribosomal subunit. The latter also contains 5S rRNA, which is encoded separately by dedicated genes transcribed by RNA polymerase III.

## **ARN de transferencia**

Hay 516 genes que codifican para ARNt, clasificados en 49 familias en función de su anticodón específico.

Están dispersos por los cromosomas, aunque con diferente densidad. Así 273 de estos genes están entre los cromosomas 6 y 1.

### **ARN nucleares pequeños: snRNA**

Muchos son ricos en Uridina y se denominan por ejemplo como U2 snRNA e intervienen en el procesamiento del ARNm, formando parte de los spliceosomas.

### **ARN pequeños nucleolares: snoRNA**

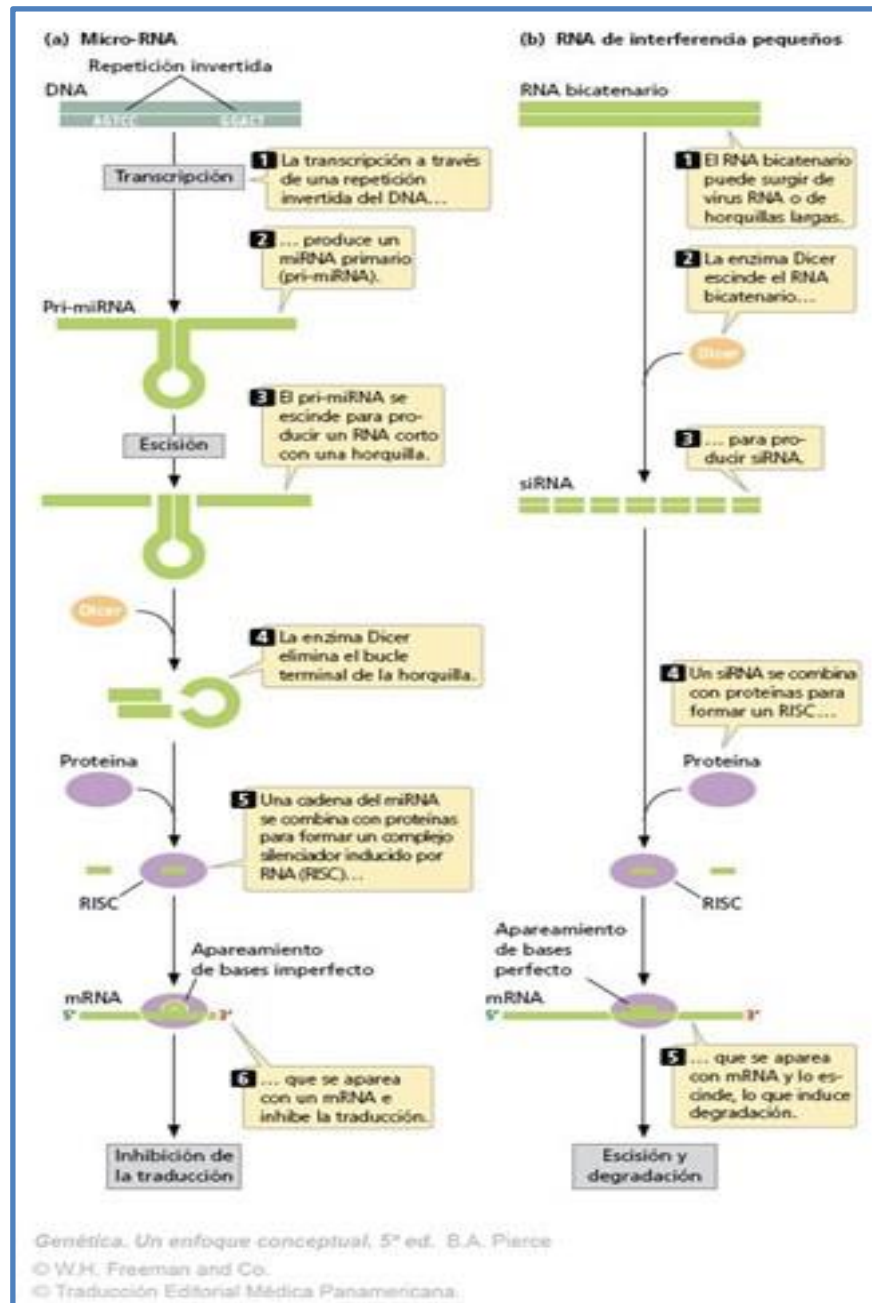
Intervienen en la modificación de los ARNr

### **ARN del cuerpo de Cajal pequeños**

Intervienen en la maduración de los ARNsn

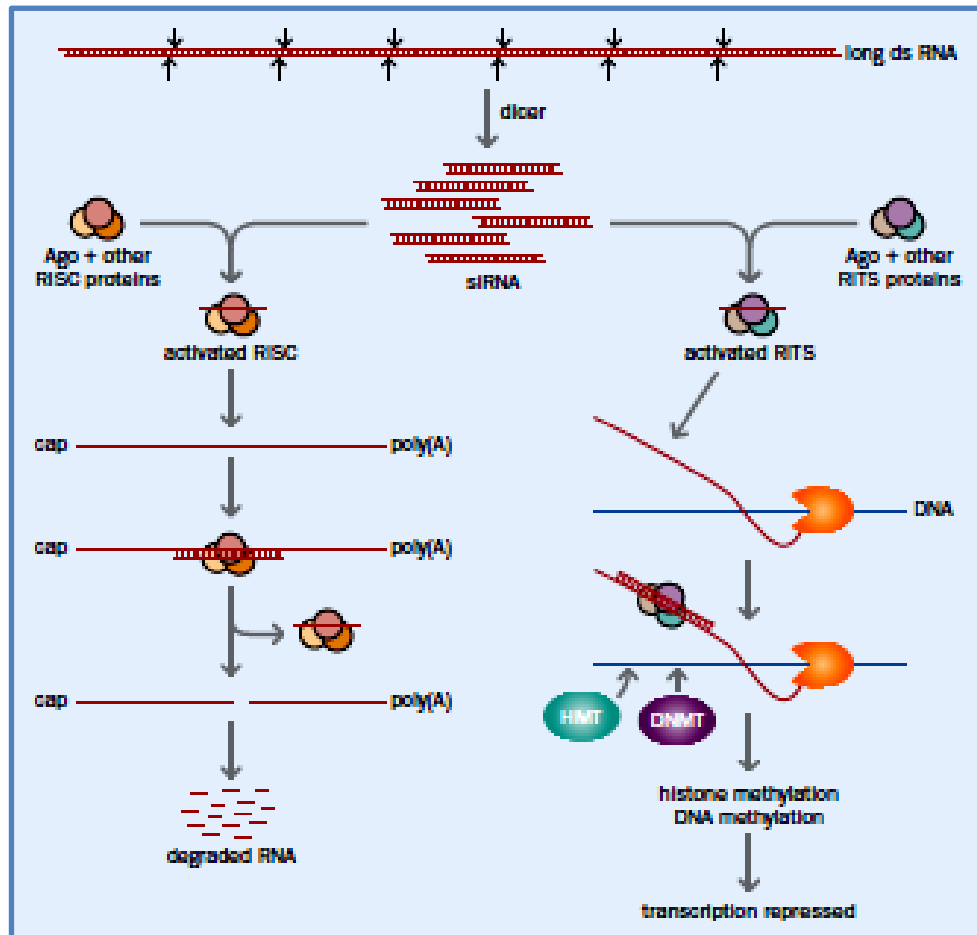
# Interferencia por ARN:

## miRNA y siRNA





## siRNA



**Figure 1 RNA interference.** Long double-stranded (ds) RNA is cleaved by cytoplasmic dicer to give siRNA. siRNA duplexes are bound by argonaute complexes that unwind the duplex and degrade one strand to give an activated complex with a single RNA strand. By base pairing with complementary RNA sequences, the siRNA guides argonaute complexes to recognize target sequences. Activated RISC complexes cleave any RNA strand that is complementary to their bound siRNA. The cleaved RNA is rapidly degraded. Activated RITS complexes use their siRNA to bind to any newly synthesized complementary RNA and then attract proteins, such as histone methyltransferases (HMT) and sometimes DNA methyltransferases (DNMT), that can modify the chromatin to repress transcription.

## ARN no codificantes largos (lncRNA): Xist

# ADN altamente repetido

## ADN altamente repetido en tándem

- Heterocromatina constitutiva formada principalmente por grandes bloques de un gran número de repeticiones en tándem de secuencias de ADN y situados en ciertas regiones subcromosómicas. Este tipo de ADN permanece condensado la mayor parte del ciclo celular.
- El ADN que forma la heterocromatina comprende unas 200 Mpb es decir un 6,5% del genoma. Está situado en las regiones centroméricas y también en los telómeros. La mayoría del cromosoma Y, y la mayor parte de los brazos cortos de los cromosoma acrocéntricos (13, 14, 15, 21 y 22). También en ciertas regiones próximas a los centrómeros de los cromosomas 1, 9, 16 y 19.
- El ADN que constituye la heterocromatina está formado por grandes ristras de secuencias de ADN altamente repetido conocido por ADN satélite (cientos de kilopares de bases).
- Las ristras más pequeñas se conocen como minisatélites (0,1 a 20 Kpb) y los agrupamiento aún más pequeños se denominan microsatélites (menos de 100-150 pb).

**TABLE 9.13 MAJOR CLASSES OF HIGH-COPY-NUMBER TANDEMLY REPEATED HUMAN DNA**

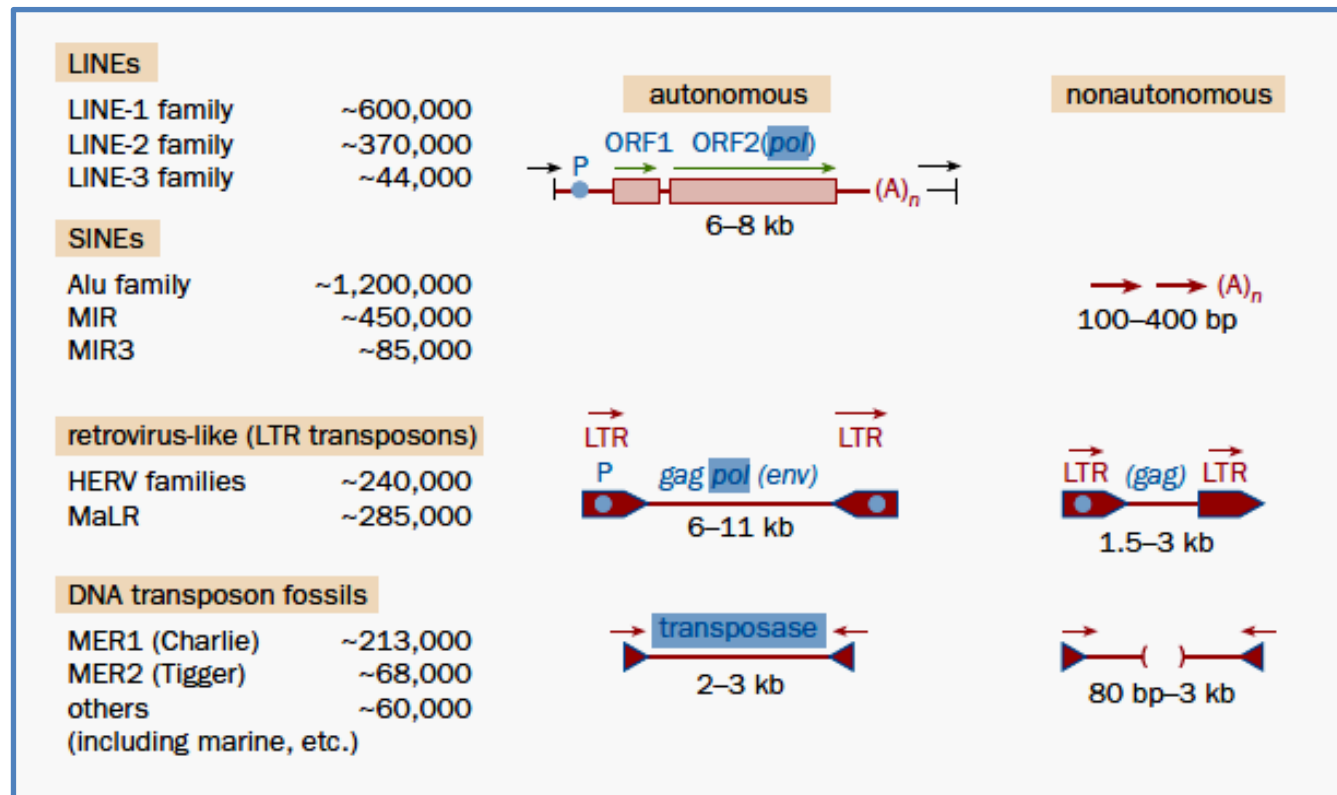
Class <sup>a</sup>	Total array size unit	Size or sequence of repeat unit	Major chromosomal location(s)
Satellite DNA <sup>b</sup>	often hundreds of kilobases		associated with heterochromatin
$\alpha$ (alphoid DNA)		171 bp	centromeric heterochromatin of all chromosomes
$\beta$ (Sau3A family)		68 bp	notably the centromeric heterochromatin of 1, 9, 13, 14, 15, 21, 22, and Y
Satellite 1		25–48 bp (AT-rich)	centromeric heterochromatin of most chromosomes and other heterochromatic regions
Satellite 2		diverged forms of ATTCC/GGAAT	most, possibly all, chromosomes
Satellite 3		ATTCC/GGAAT	13p, 14p, 15p, 21p, 22p, and heterochromatin on 1q, 9q, and Yq12
DYZ19		125 bp	~400 kb at Yq11
DYZ2		AT-rich	Yq12; higher periodicity of ~2470 bp
Minisatellite DNA	0.1–20 kb		at or close to telomeres of all chromosomes
Telomeric minisatellite		TTAGGG	all telomeres
Hypervariable minisatellites		9–64 bp	all chromosomes, associated with euchromatin, notably in sub-telomeric regions
Microsatellite DNA	< 100 bp	often 1–4 bp	widely dispersed throughout all chromosomes

<sup>a</sup>The distinction between satellite, minisatellite, and microsatellite is made on the basis of the total array length, not the size of the repeat unit.

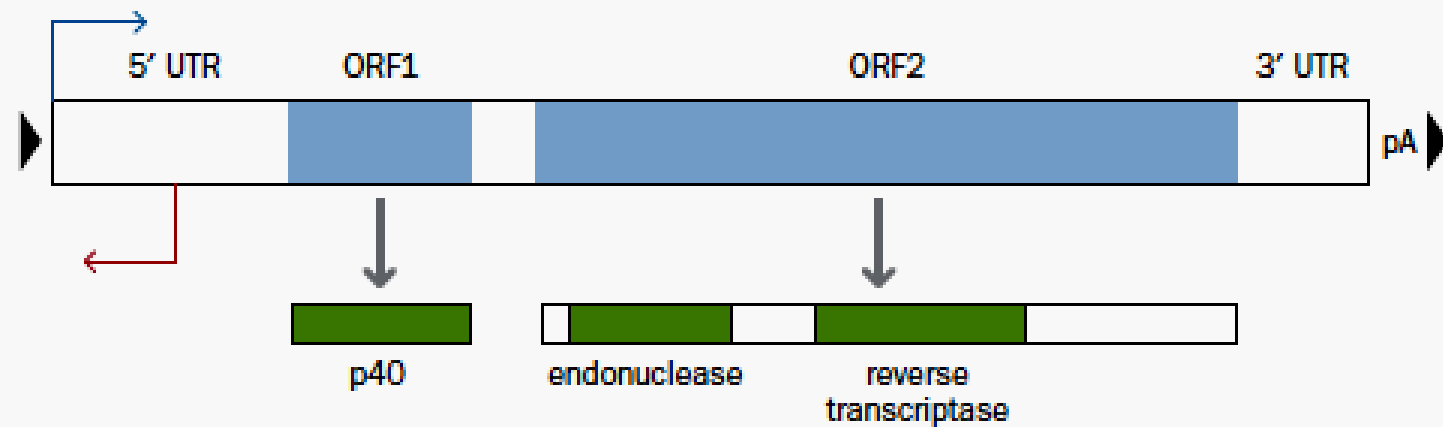
<sup>b</sup>Satellite DNA arrays that consist of simple repeat units often have base compositions that are radically different from the average 41% G+C (and so could be isolated by buoyant density gradient centrifugation, when they would be differentiated from the main DNA and appear as *satellite bands*—hence the name).

## ADN altamente repetido disperso

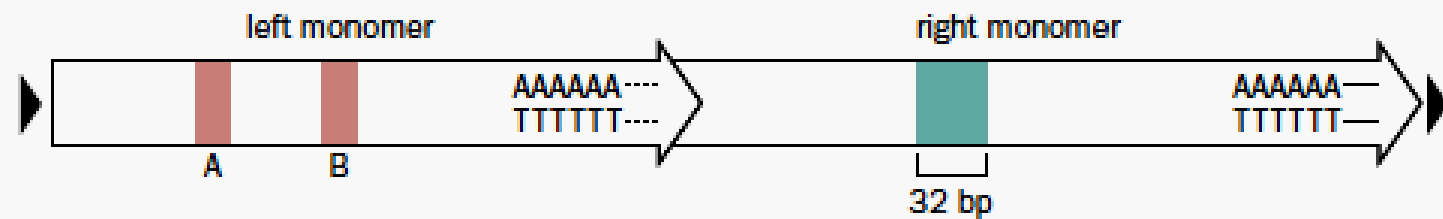
La mayoría deriva de transposones o también denominados elementos transponibles. Aproximadamente el 45 % del genoma pertenece a este tipo de elementos y puede ser más aun que es difícil de reconocer por los cambios que han sufrido a lo largo de la evolución



(A) LINE-1 repeat element



(B) Alu repeat element





**Figura 20-36.** Elementos repetidos encontrados en el gen humano (*HGO*) que cifra la dioxigenasa del ácido homogentísico, la enzima cuya falta causa la alcaptonuria.