

Organització dels genomes

GEN

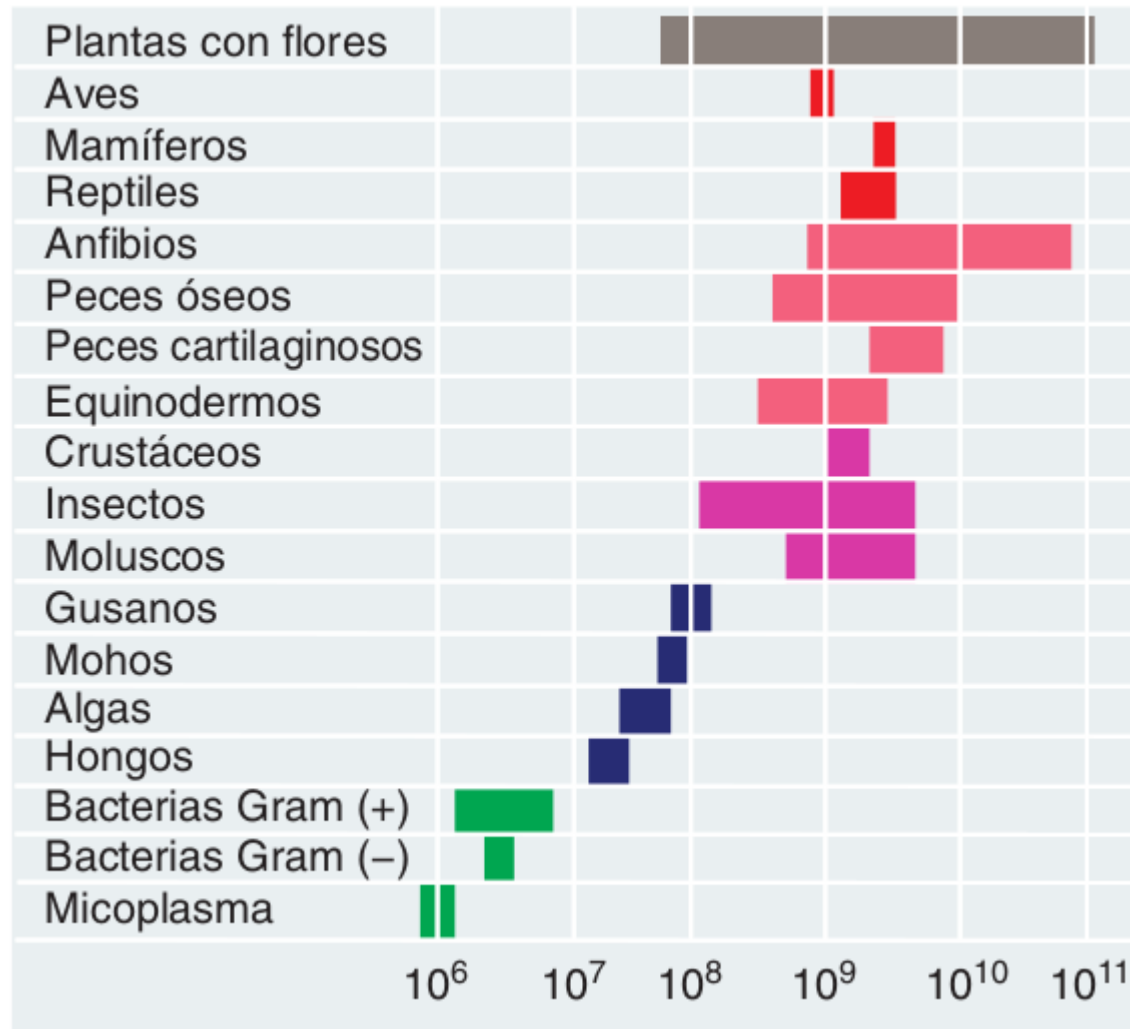
Un gen és un segment de DNA que a l'expressar-se dona lloc a un producte funcional que pot ser una proteïna o un RNA.

(amb aquesta definició de gen s'engloben tan els gens que informen sobre proteïnes estructurals i enzims com els gens que informen sobre proteïnes reguladores, com els que contenen informació pels RNA diferents dels mRNA)

GENOMA

El **genoma** és tota la informació genètica d'un organisme. Inclou no sols els gens, sinó també els segments no gènics del DNA.

- El genoma conté la informació necessària per a que una cèl·lula pugui existir i reproduir-se.
- La grandària del genoma generalment és més petita en procariotes que en eucariotes tant en nombre de parells de bases com en nombre de gens.



(milions de parells de bases per genoma haploide)

La quantitat total de DNA per genoma haploide és característica de cada espècie. Tal com s'aprecia a la figura, existeix una gran variació dintre del mateix grup taxonòmic (grups d'organismes classificats junts) i entre els diferents grups. El contingut en DNA s'incrementa amb la complexitat morfològica en els eucariotes inferiors, però varia considerablement dintre d'alguns grups d'eucariotes superiors.

DNA organismes procariotes vs eucariotes

Procariotes:

- El DNA conté una còpia de cada gen.
- Els gens no contenen introns.
- Gairebé no hi ha DNA que no es transcriu.
- Els gens es disposen un a continuació dels altres (existeix poc DNA espaiador).
- Més del 90% del genoma dels bacteris són gens.

Eucariotes:

- Tot i que tenen molt més DNA que els procariotes, només una petita part són gens que codifiquen per a proteïnes (gens codificants).
- Els gens contenen introns la qual cosa fa augmentar molt la seqüència codificant.
- Els gens estan separats entre si per DNA espaiador.
- Molts DNA eucariotes estan dominats per seqüències de DNA repetides.

GENOMA EUCARIOTA

Tipus de seqüències de nucleòtids

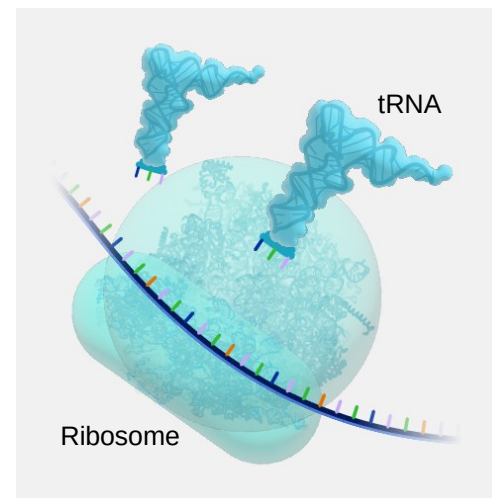
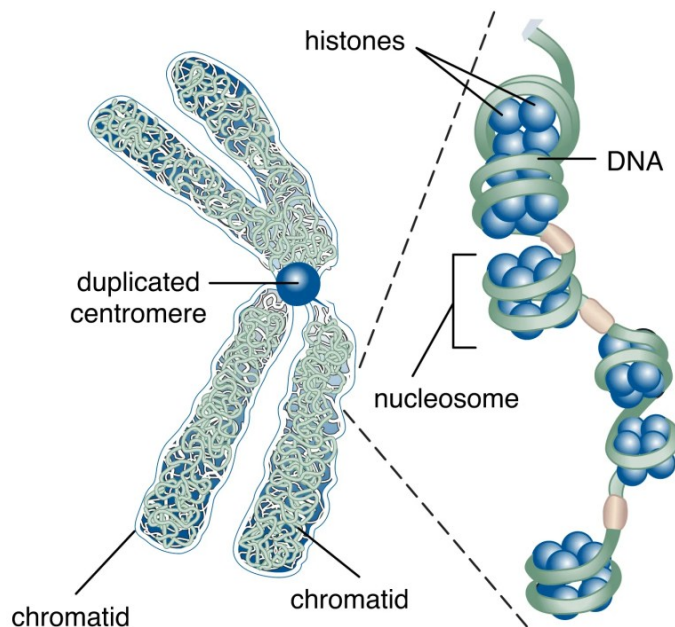
El DNA eucariota està format per tres tipus de seqüències:

- Seqüències de DNA altament repetitives.
- Seqüències de DNA moderadament repetitives.
- Seqüències de DNA de còpia única.

DNA moderadament repetitiu

Format per **seqüències llargues**, entre 300 i 6000 pb, repetides entre 10 i més de 1000 vegades. **No ocupen uns punts concrets** en els cromosomes.

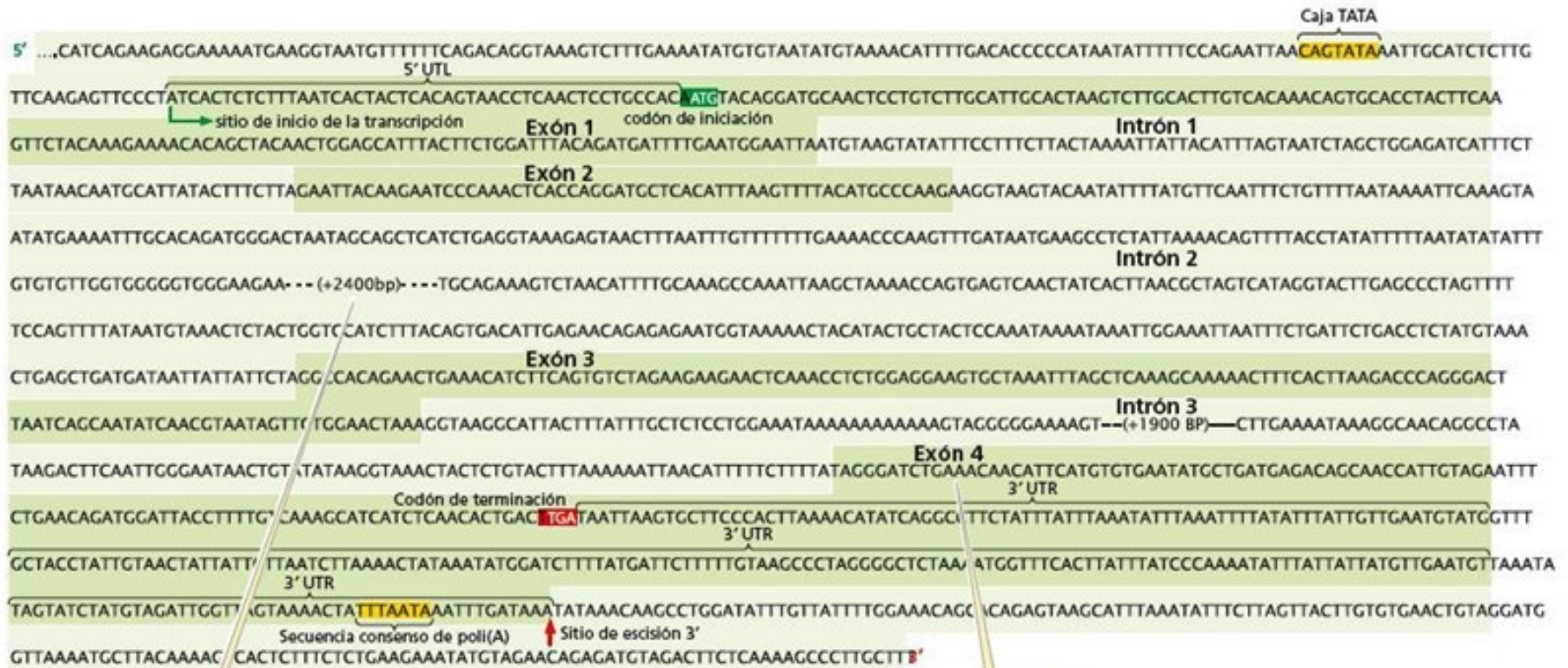
- **La major part són** seqüències que no es transcriuen, que es troben disperses pel genoma i de les que se'n desconeix la funció.
- **Una part són** seqüències que sí que es transcriuen i corresponen als gens que codifiquen per a les **histones** i als gens que porten la informació per als diferents RNA (excepte els mRNA) com per exemple els **RNA ribosòmics** i els **RNA de transferència**.



DNA de còpia única, simple o no repetitiu

Format per seqüències de nucleòtids que no es repeteixen (o ho fan molt poc).

- Inclou:
 - Seqüències de DNA codificants:
 - Gens que codifiquen per a proteïnes estructurals i enzims, anomenats **gens estructurals**.
 - Gens que codifiquen per a proteïnes que controlen l'expressió dels gens estructurals, anomenats **gens reguladors**.
 - Seqüències de DNA no codificants:
 - Seqüències intercalades dins dels gens, anomenades **introns**.
 - Seqüències que separen els gens, anomenades **DNA espaiador**.



Puede observar que los intrones no codificantes ocupan grandes partes de los genes, aun cuando no se enumeren de manera individual grandes números de bases.

Los exones codifican menos de 165 aminoácidos, una proteína pequeña.

Genética. Un enfoque conceptual. 5ª ed. B.A. Pierce
 © W.H. Freeman and Co.
 © Traducción Editorial Médica Panamericana.

Representació del gen que codifica per a la Interleucina 2 humana. Inclou la caixa TATA, el lloc d'inici de la transcripció, els codons d'inici i d'aturada, introns, exons, seqüències poliA...

Projecte GENOMA HUMÀ



- Objectiu: seqüenciar el genoma humà.
- Organitzat i dut a terme per un grup internacional d'investigadors de diferents universitats i instituts d'investigació.
- Comença: any 1990
- L'any 2003 ja es tenia la seqüència nucleotídica de la major part (92%) del DNA de cada cromosoma humà (dels 22 autosomes i de la parella de cromosomes sexuals). Faltava seqüenciar especialment certes regions altament repetitives.
- La millora de les tècniques de seqüenciació han permés actualment (2022) disposar de la seqüència completa de tot el genoma humà.

De què pot servir conèixer el genoma humà?

- El coneixement de la seqüència de nucleòtids de cada gen i de la seva forma d'expressar-se permet tenir una referència fonamental per a la teràpia gènica.
- Es podrà conèixer si s'és portador d'una malaltia genètica que no es manifesta en els primers anys de vida, i la possibilitat de que arribi a manifestar-se, i d'aquesta manera es podran prendre les precaucions necessàries.
- Les malalties que primer es poden beneficiar d'aquests coneixements són el càncer, la malaltia de l'Alzheimer, determinats tipus d'asma i algunes malalties neuronals, com ara l'esquizofrènia.

GENOMA HUMÀ: algunes dades

- Grandària aproximada: 3.200 milions de parells de bases ($3,2 \times 10^9$ pb) per genoma haploide.
- Nombre de gens: aproximadament 20.500
- Longitud mitjana d'un gen: 27.000 pb
- Nombre mitjà d'exons per gen: 9
- Longitud mitjana d'un intró: 3.365 pb
- La densitat gènica varia entre els cromosomes: els cromosomes 17, 19 i 22 tenen les densitats més elevades, mentre que els X, Y, 4, 13 i 18 les més baixes.

cromosoma	pares de bases	nucleótidos	genes	genes no codificantes	pseudogenes	densidad
1	248.956.422	497.912.844	2059	2090	1293	8
2	242.193.529	484.387.058	1300	1768	1079	5
3	198.295.559	396.591.118	1077	1283	800	5
4	190.214.555	380.429.110	753	1112	757	4
5	181.538.259	363.076.518	886	1293	738	5
6	170.805.979	341.611.958	1049	1150	831	6
7	159.345.973	318.691.946	1001	1067	910	6
8	145.138.636	290.277.272	686	1120	649	5
9	138.394.717	276.789.434	779	840	691	6
10	133.797.422	267.594.844	729	972	601	5
11	135.086.622	270.173.244	1320	1131	835	10
12	133.275.309	266.550.618	1034	1268	654	8
13	114.364.328	228.728.656	321	658	395	3
14	107.043.718	214.087.436	817	919	532	8
15	101.991.189	203.982.378	612	1032	532	6
16	90.338.345	180.676.690	859	1061	506	10
17	83.257.441	166.514.882	1185	1228	546	14
18	80.373.285	160.746.570	269	665	260	3
19	58.617.616	117.235.232	1472	915	527	25
20	64.444.167	128.888.334	545	639	264	8
21	46.709.983	93.419.966	234	425	188	5
22	50.818.468	101.636.936	495	538	343	10
X	156.040.895	312.081.790	852	665	891	5
Y	57.227.415	114.454.830	64	108	394	1
Mitocondria	16.569	33.138	13	24	0	785
No asignado	183.830.549	367.661.098	29	24	6	0
Total	3.272.116.950	6.544.233.900	20440	23995	15222	

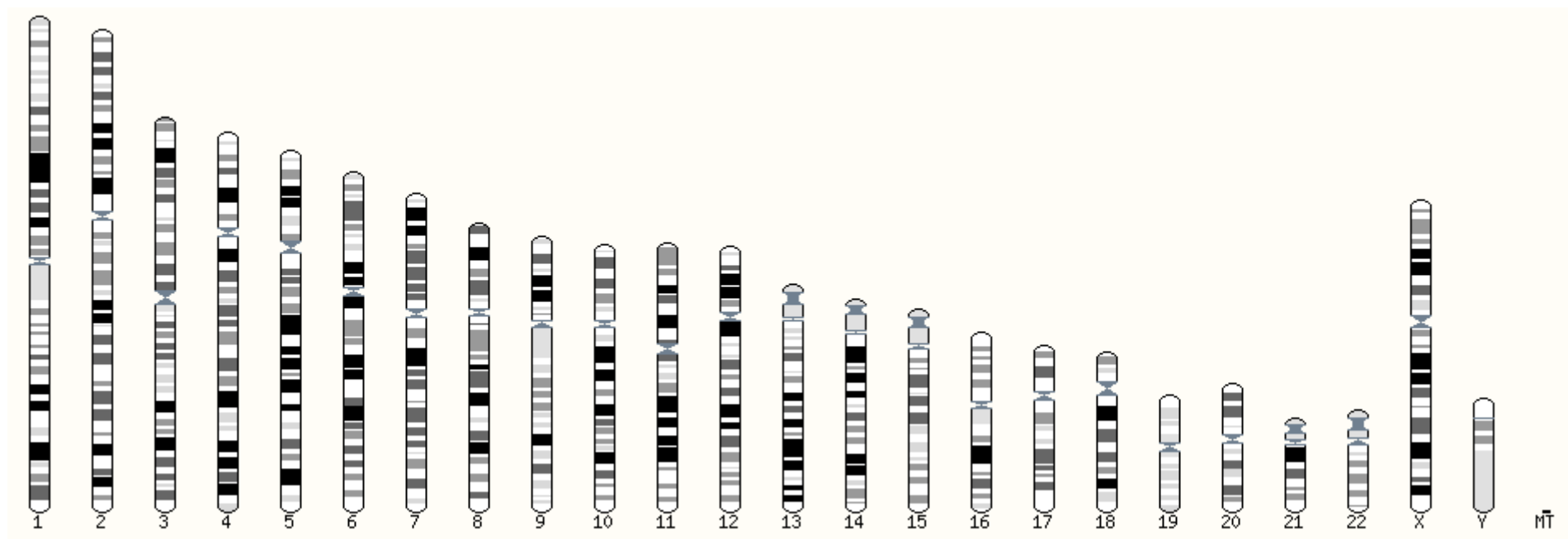
Nombre de gens (codifiquen per a proteïnes), gens no codificants (gens amb informació per als RNA sense comptar el mRNA) i pseudogens (gens que han quedat inactius al llarg de l'evolució per acumulació de nombroses mutacions) existents en cadascun dels cromosomes humans. La densitat fa referència al nombre de gens que te cada cromosoma per cada 1.000.000 de parells de bases.

El nostre genoma és immens!!!



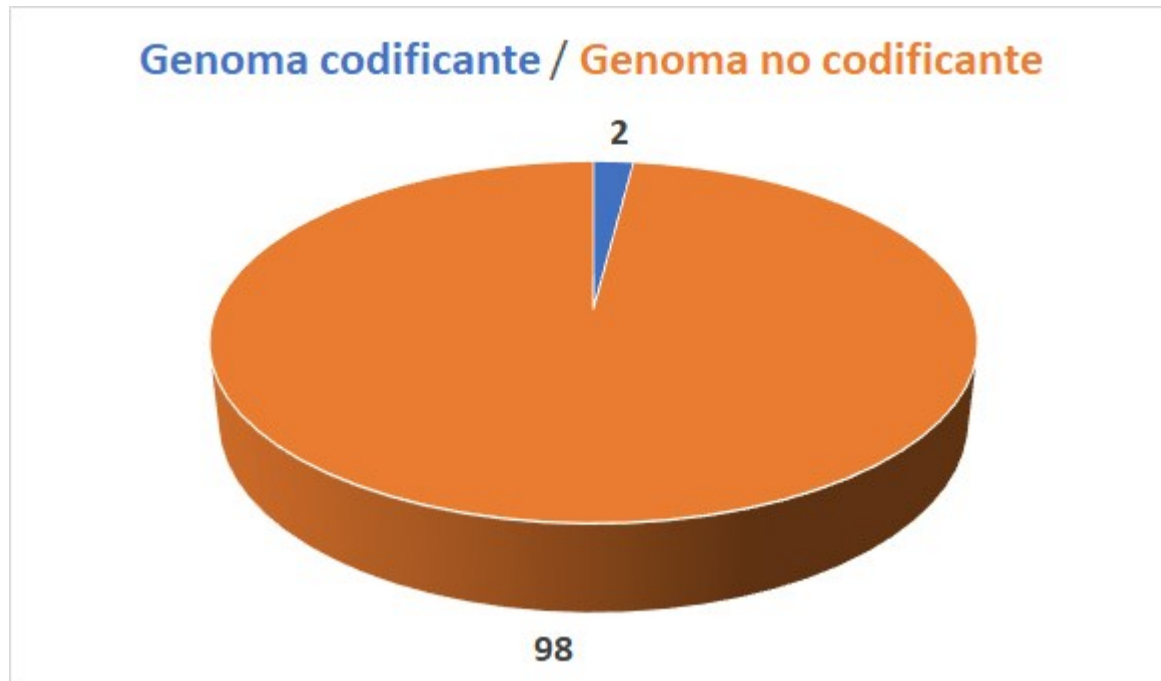
Si cadascun dels nucleòtids que formen el nostre genoma diploide fos una lletra d'un llibre de text, el genoma ocuparia unes 8,5 milions de pàgines. Amb una mitjana de 1.000 pàgines per llibre, tindríem més de 8.500 llibres.

Tota una biblioteca completa en cadascuna de les nostres cèl·lules !!



Representació gràfica del conjunt de cromosomes humans i del DNA mitocondrial

Font: www.ensembl.org/Homo_sapiens/Location/Genome



A los efectos de la cuantificación de genes en genomas el acuerdo actual, de nuevo arbitrario, es contar separadamente los genes codificantes (los que portan información que acaba traducándose en forma de proteínas) y los genes no codificantes (los que se transcriben a ARN y actúan como ARN, sin traducirse a proteínas). Adicionalmente, cuando nos referimos al número de genes de un organismo solemos referirnos a la primera de estas categorías, solamente teniendo en cuenta los genes codificantes.

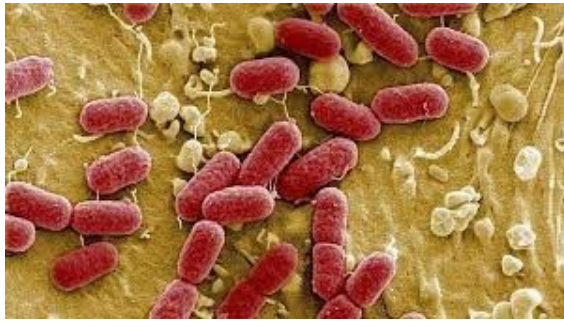
Si juntamos todos los segmentos de ADN que corresponden a genes codificantes, que se transcriben a ARN y se traducen a proteína, el resultado será, aproximadamente un 2% del tamaño total de nuestro genoma (de referencia). O sea, **de los 3.272.116.950 pares de bases que tiene el genoma de referencia** solamente un 2% correspondería a genoma codificante, es decir: 65.442.339 pares de bases.

¿Y que hay en el 98% restante que corresponde al genoma no-codificante? Pues muchos elementos reguladores, que determinan cuando y dónde se deberá expresar un gen, también están aquí todos los genes no codificantes (miRNA, lncRNA, rRNA y tRNA), diferentes familias de elementos móviles, **transposones** y **retrotransposones**, y múltiples familias también de ADN repetitivo (satélite). También encontramos los **pseudogenes**, generalmente correspondientes a secuencias parciales o totales de genes que han quedado inactivos a lo largo de la evolución al acumular numerosas mutaciones.

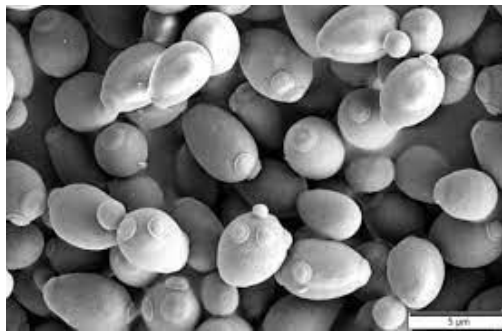
Grandària d'alguns genomes i nombre de gens

Organisme	Genoma haploide (parells de bases)	Nombre de gens	Densitat *
<i>Streptococcus pyogenes</i> (bacteri)	1.837.281	1.828	995
<i>Escherichia coli</i> (bacteri)	4.558.660	4.051	889
<i>Saccharomyces cerevisiae</i> (llevat)	12.157.105	6.600	543
<i>Drosophila melanogaster</i> (mosca de la fruita)	142.573.024	13.947	98
<i>Caenorhabditis elegans</i> (nematode)	103.022.290	20.191	196
<i>Homo sapiens</i> (humà)	3.272.116.950	20.440	6
<i>Mus musculus</i> (ratolí)	3.486.944.526	22.519	6
<i>Danio rerio</i> (peix zebra)	1.674.207.132	25.592	15
<i>Arabidopsis thaliana</i> (planta)	135.670.229	27.655	204
<i>Oryza sativa</i> (arròs)	411.710.190	40.745	99
<i>Triticum aestivum</i> (blat)	14.547.261.565	107.891	7

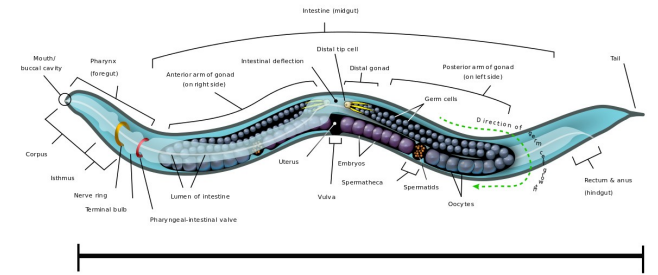
* La densitat es refereix al nombre de gens per cada 1.000.000 de parells de bases



Escherichia coli
5 Mb
4.051 gens



Saccharomyces cerevisiae
12 Mb
6.600 gens



~1 mm

Caenorhabditis elegans
103 Mb
20.191 gens



Arabidopsis thaliana
136 Mb
27.655 gens



Danio rerio
1.700 Mb
25.592 gens



Drosophila melanogaster
142 Mb
13.947 gens



Mus musculus
3.500 Mb
22.519 gens



Homo sapiens
3.200 Mb
20.440 gens

Grandària dels genomes vs complexitat

Els procariotes són els organismes que tenen menys gens i també els que tenen els genomes més petits. Amb els genomes tan reduïts també són els que tenen major densitat gènica, al voltant de 1.000 gens per milió de parells de bases.

Els genomes dels eucariotes pluricel·lulars són més grans que els dels eucariotes unicel·lulars com per exemple, el llevat.

Entre els eucariotes pluricel·lulars, no hi ha una relació directa entre la grandària del genoma i la complexitat dels organismes. Per exemple, el nemàtode *Caenorhabditis elegans* és més complex des del punt de vista estructural que la planta *Arabidopsis thaliana*, però té molta menys quantitat de DNA.

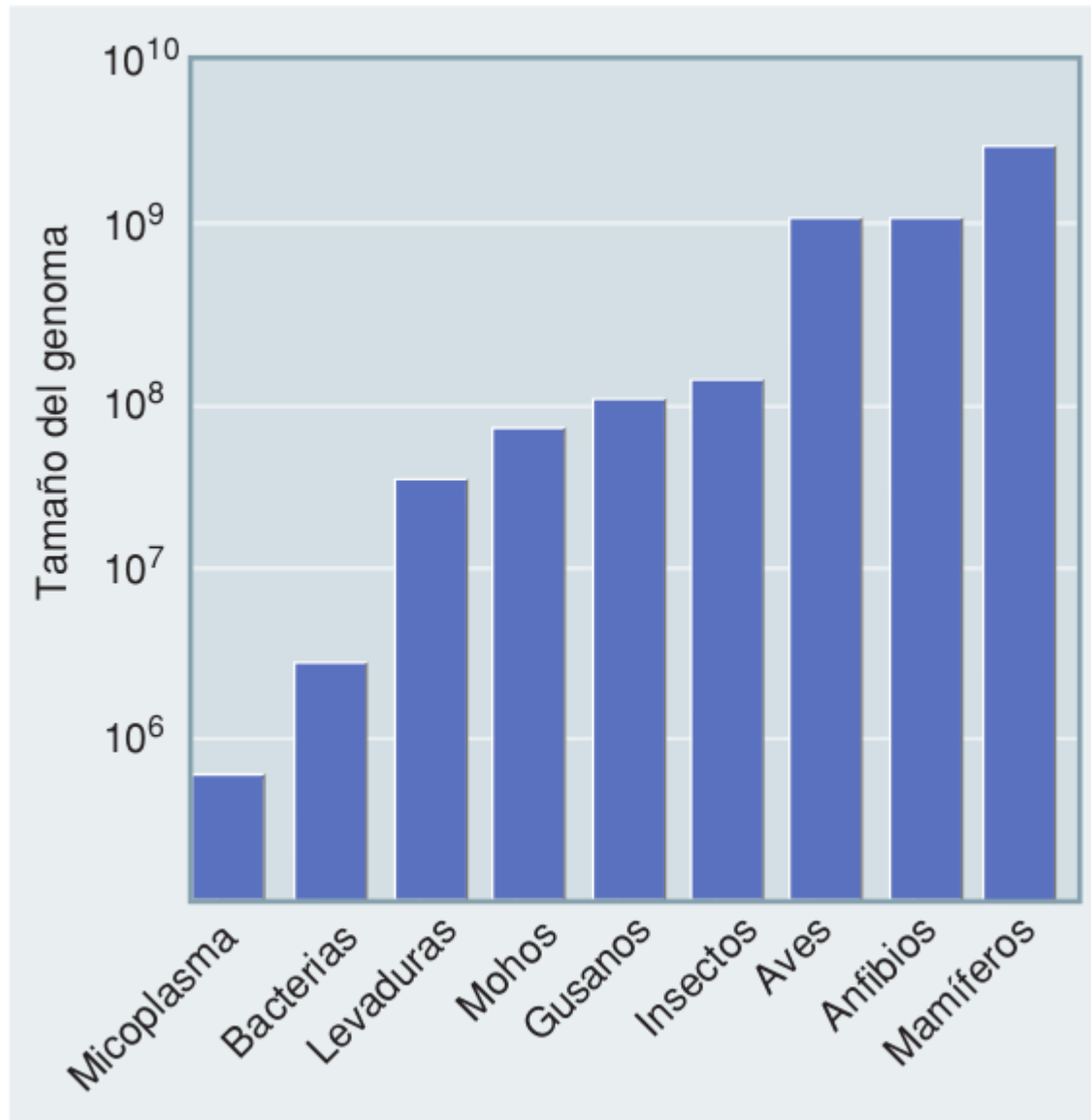
Entre els eucariotes no hi ha tampoc una correlació entre la grandària del genoma i el nombre de gens que conté, ni tampoc entre el nombre de gens i la complexitat fenotípica. Per exemple, els éssers humans tenen més gens que els invertebrats, però sols el doble que les mosques de la fruita i gairebé els mateixos que la planta *Arabidopsis thaliana*, tot i que l'ésser humà és molt més complex. El nemàtode *Caenorhabditis elegans* té més gens que *Drosophila melanogaster*, però és menys complex fenotípicament.

La gran complexitat de l'ésser humà no depèn tant del nombre de gens, sinó de les diferents maneres d'interactuar entre si, i per tant, de les seves múltiples formes d'expressió.

Existeixen múltiples estratègies que han anat apareguent al llarg de l'evolució, que han fet augmentar la complexitat estructural i funcional del genoma.

Amb 20.000 gens codificants penseu que fabriquem 20.000 proteïnes?

En realitat som capaços de fabricar moltíssimes més!!!



La grandària del genoma mínim que es troba en cada taxó augmenta des dels procarïotes fins els mamífers.

Seres modélicos

Entre la naturaleza y el laboratorio

[Proyectos genoma](#) [Blog](#)

Bacteria

Escherichia coli



Levadura

Saccharomyces cerevisiae



Planta

Arabidopsis thaliana



Gusano

Caenorhabditis elegans



Mosca

Drosophila melanogaster



Pez

Danio rerio



Ratón

Mus musculus

