

XENOMA. PROXECTO XENOMA



Carmen Cid Manzano

I.E.S. Otero Pedrayo. Ourense.

Departamento Bioloxía e Xeoloxía.

O xenoma é...

o conxunto completo de información xenética dun organismo.

- Herdase de xeración en xeración
- Codificado en secuencias de ácidos nucleicos
- Organizado en cromosomas (cada cromosoma ten unha molécula de ADN).



“Os antigos, para coñecer o seu destino, consultaban ós deuses. Nós, modernos, aprendemos dos xenes”.

*F. Jacob
Premio Nóbel de Medicina 1965*

Os xenomas diferéncianse polo seu tamaño e complexidade

-En virus

- Xenoma de DNA ou RNA, de cadea simple ou dobre, circular ou lineal.

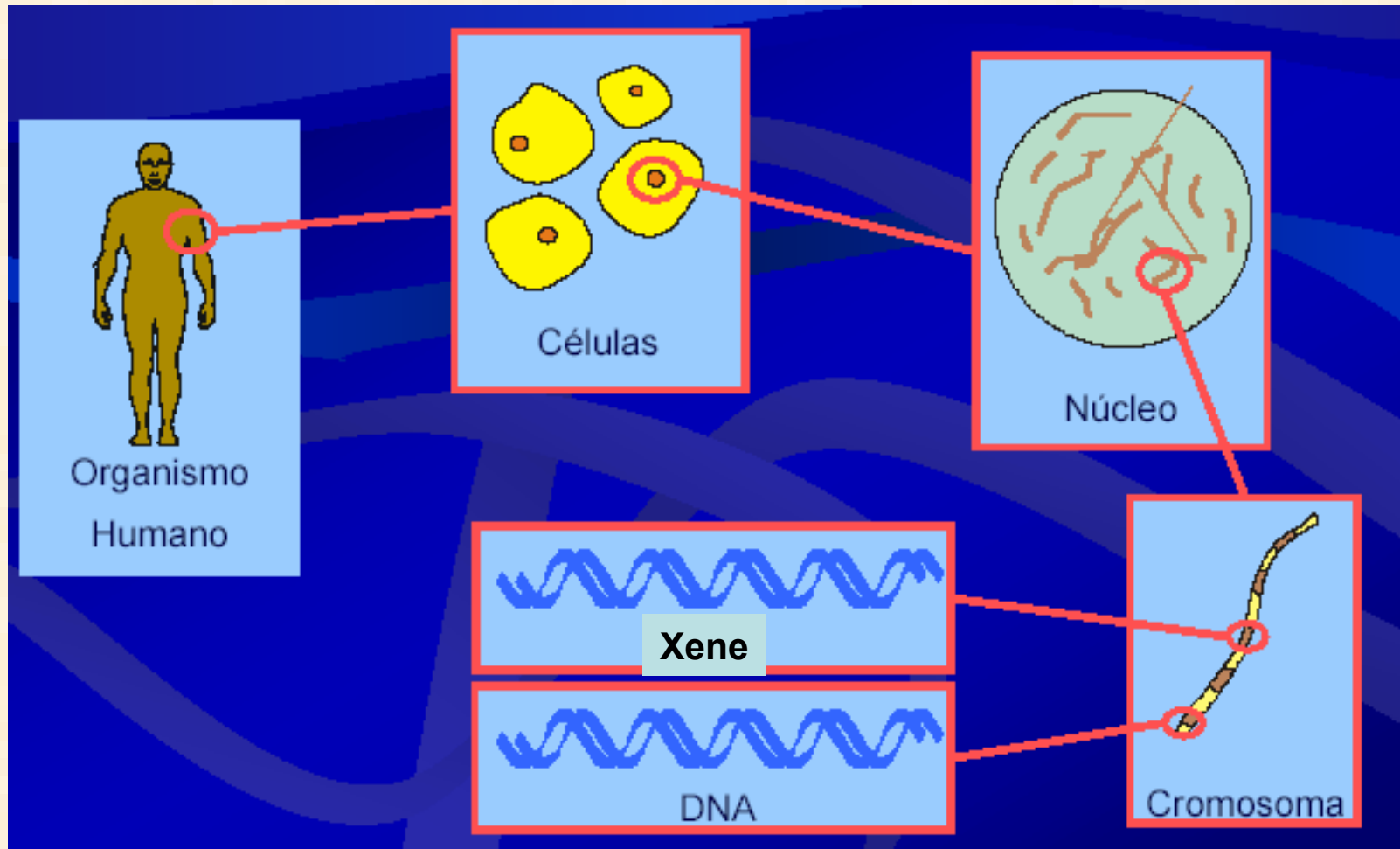
- En procariotas

- Un só cromosoma de DNA circular de cadea dobre
- DNA extracromosómico circular de cadea dobre (plásmidos)

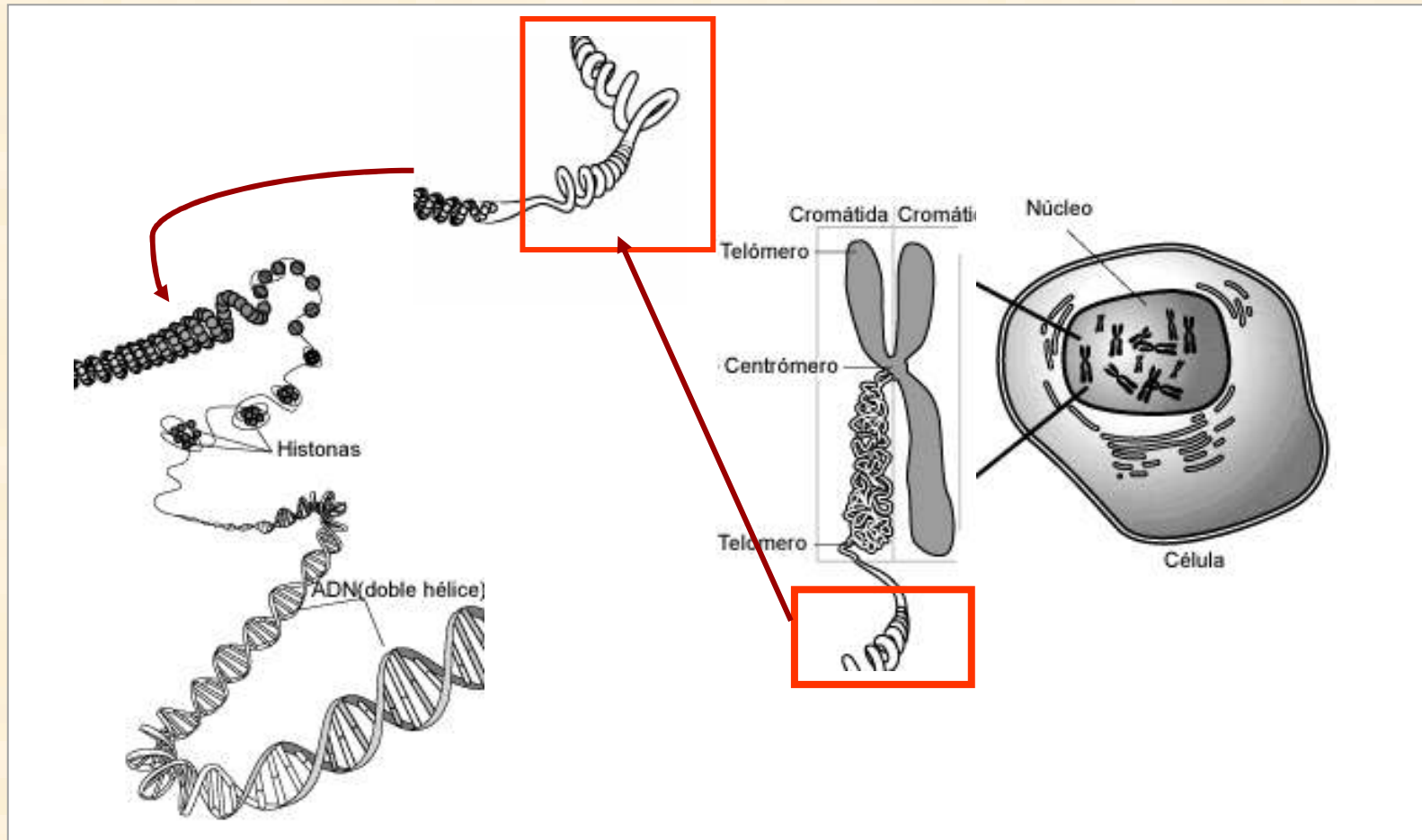
-En eucariotas

- Múltiples cromosomas de DNA lineal de cadea dobre
- Xenoma diploides, excepto en gametos (haploide)
- Gran cantidade de DNA non codificador
- As mitocondrias e cloroplastos teñen xenoma propio (similar o xenoma de procariotas)

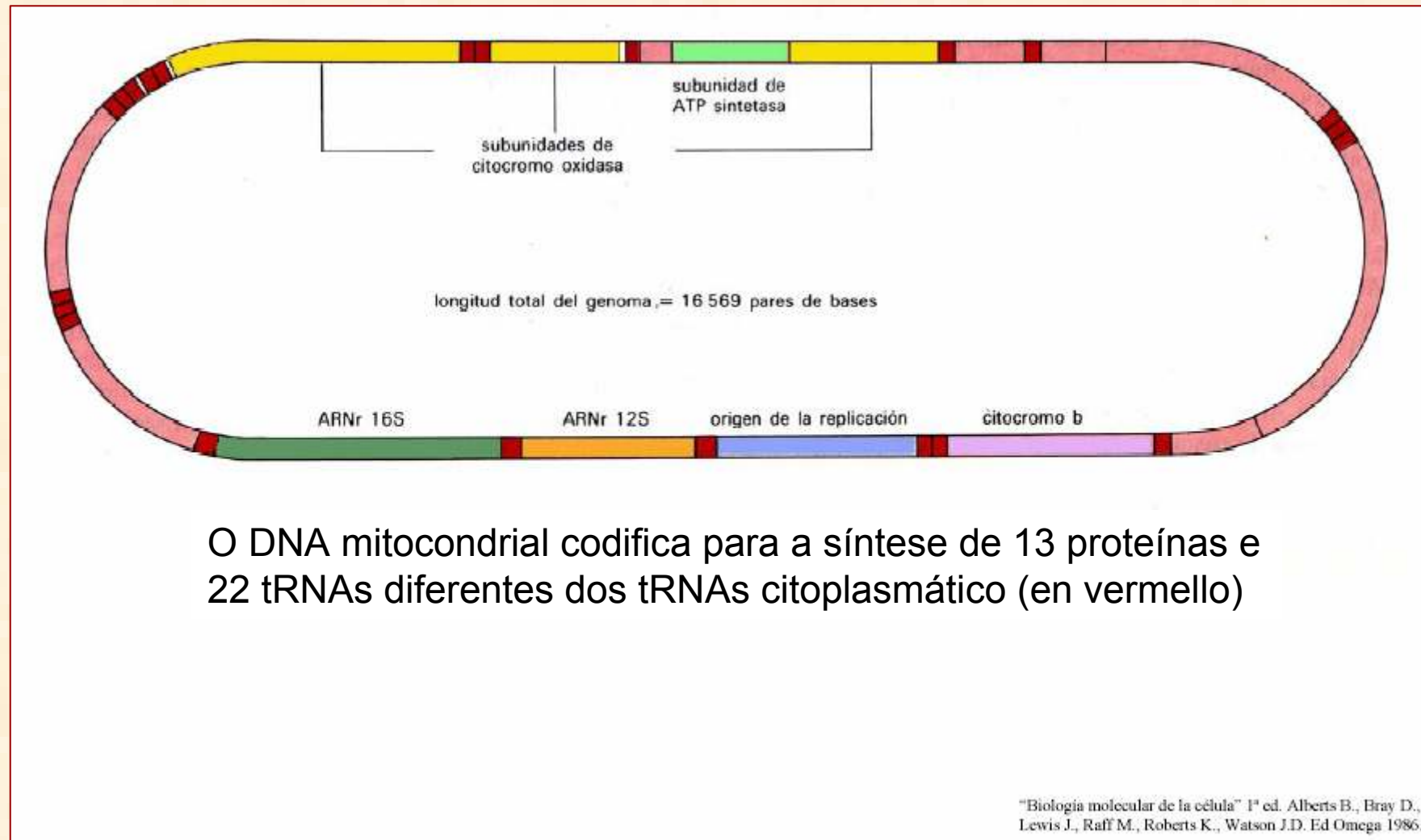
O corpo humano contém 50 bilhões de células



Xenoma eucariota nuclear



Xenoma mitocondrial



O termo **xenoma** úsase para referirse a todos os **alelos** que posúe un organismo (ou unha poboación, especie, ou un grupo taxonómico).

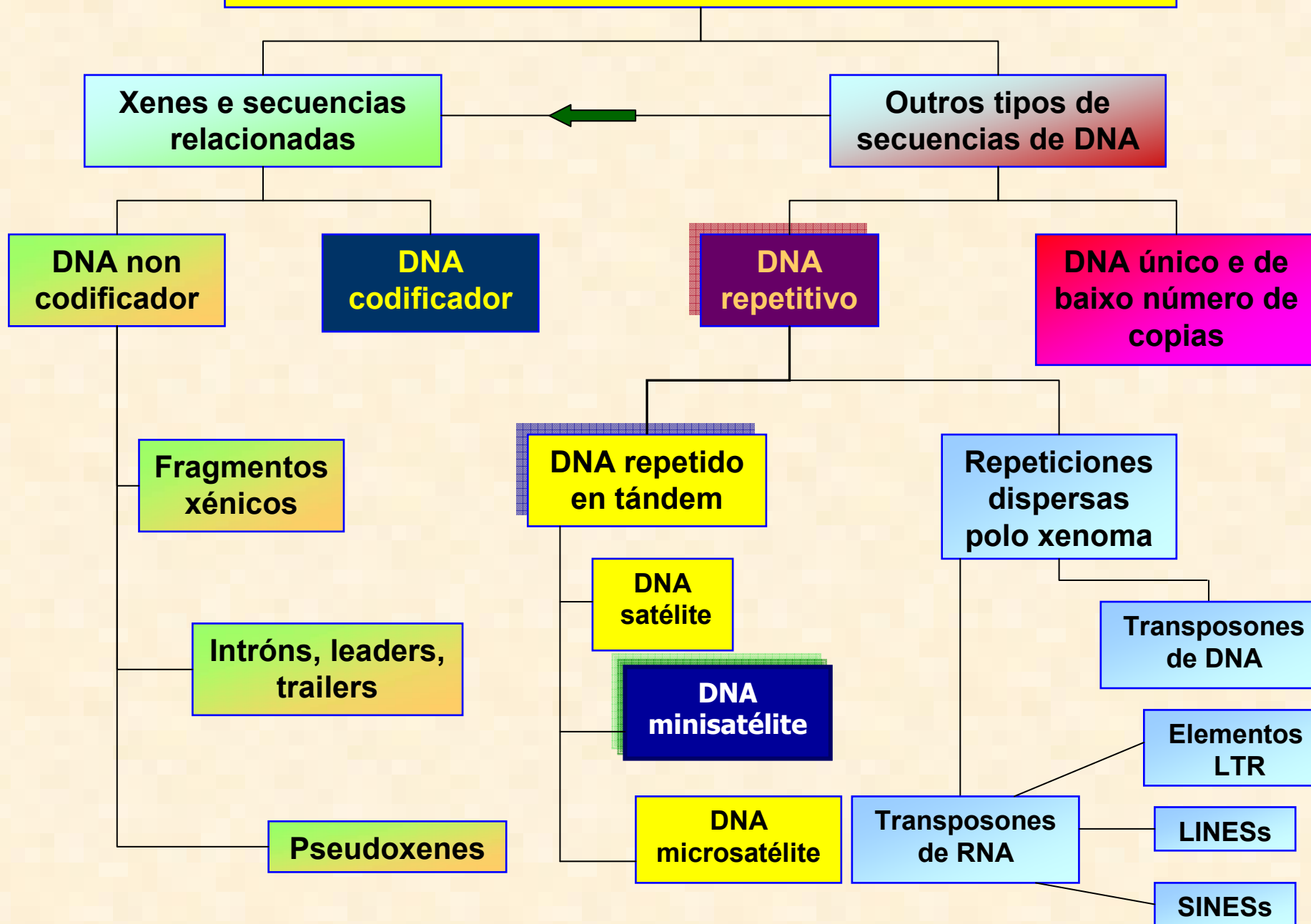
Se ben a cantidade de ADN dunha célula **diploide** é constante para unha especie, existen grandes diferencias entre as mesmas. *Homo sapiens* ten 3.5×10^9 pares de bases e *Drosophila* ten 1.5×10^8 , por xenoma haploide.

O tamaño do xenoma en eucariotas defínese como o valor C ou cantidade de DNA por xenoma haploide, tal como o que existe no núcleo dun espermatozoide. Denomínase C, por constante ou característico, para indicar o feito de que o tamaño é practicamente constante dentro dunha especie.

Paradoxa do valor C:

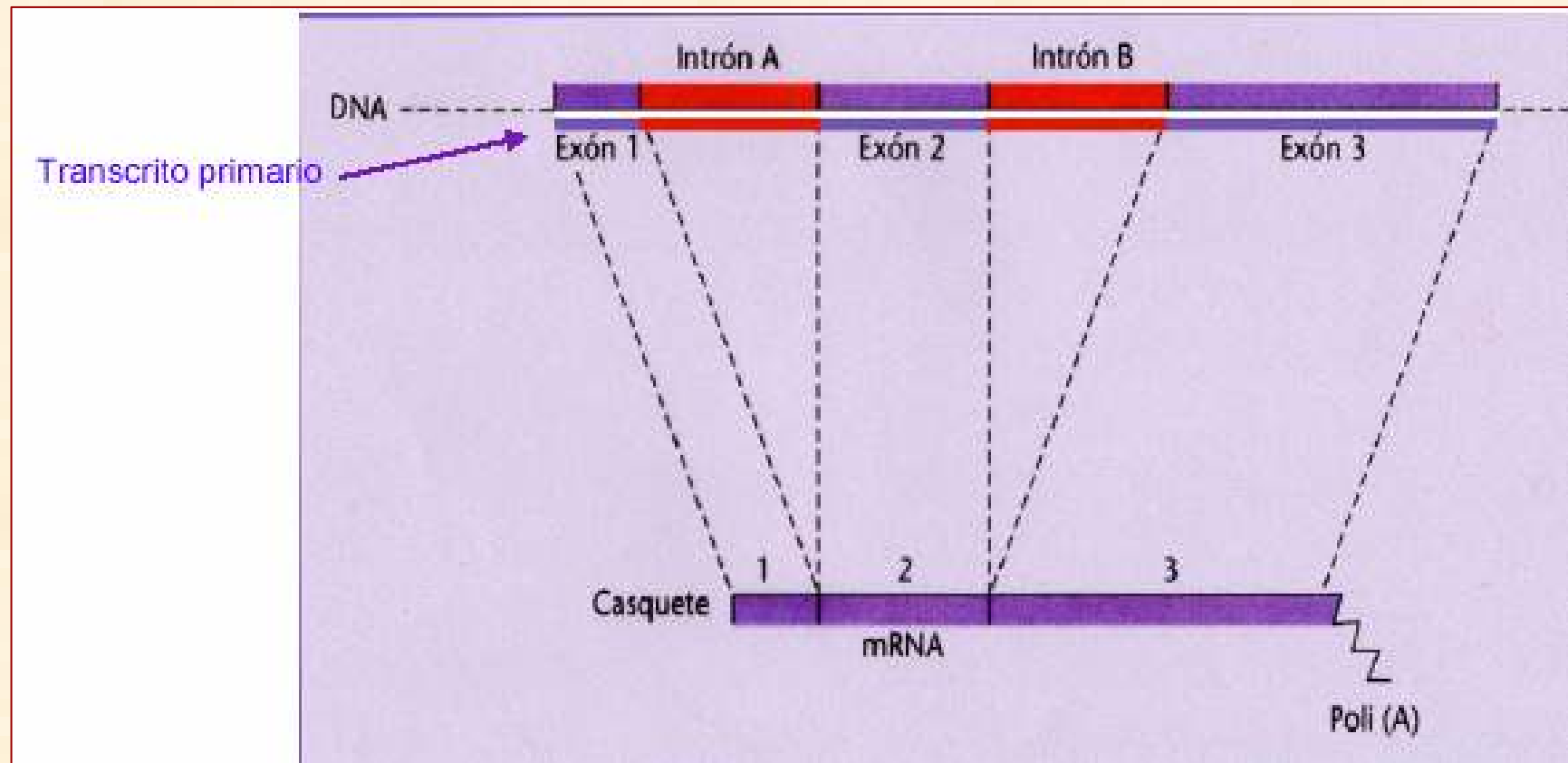
Non hai unha total correspondencia entre o valor C e a cantidade de información xenética esperada segundo a complexidade do organismo.

Clases de Secuencias do XENOMA



DNA codificador

Exóns: 2% ADN eucariota que codifica proteínas e secuencias reguladoras.



DNA non codificador

Algunhas secuencias xogan un importante papel na expresión xénica.

- **Intróns:** Están formados por unha secuencia sinxela de bases que se dispoñen entre rexións codificantes do xenoma (exóns). Deles poden derivar os ARN de interferencia.
- **Fragmentos xénicos** son trozos de xenes incompletos e polo tanto presumiblemente inviables
- **Leaders:** secuencias que se transcriben a nucleótidos de ARNm, que se estende desde o extremo 5' hasta o codón de iniciación e que, polo tanto, non se traduce.
- **Trailers:** secuencias que se transcriben a nucleótidos de ARNm, exténdense desde o codón de terminación hasta o extremo 3' e que, polo consiguiente, non se traducen.
- **Pseudoxenes:** trátase de xenes completos ou sen intróns pero que non se expresan, xa sexa pola falla de promotor, perda do codón de inicio, etc..

Outros tipos de secuencias de DNA

DNA único e de baixo número de copias

- É o DNA interxénico non repetitivo

DNA repetitivo

DNA repetido en tándem

- ADN con motivos repetidos en tándem (satélite, minisatélites e microsatélites, segundo o tamaño da unidade que se repite): secuencias de nucleótidos colocadas unhas detrás de outras (*en tándem*) e nas que existe unha secuencia común relativamente curta que se repite unha tras outra nun fragmento de ADN.
- ATCGG ATCGG ATCGG ATCGG ATCGG -----

DNA repetitivo

Repeticiones
dispersas
polo xenoma

ADN con motivos repetitivos dispersos: trátase dunha secuencia que aparece repetida, pero non unhas despois doutras senón en diferentes posicións no xenoma.

□ **Transposóns de DNA:** repeticións dispersas que poden ser duplicadas e insertarse noutro lugar do xenoma e non teñen intermediarios de RNA. Moito menos abundantes que os transposóns de RNA nos eucariótas.

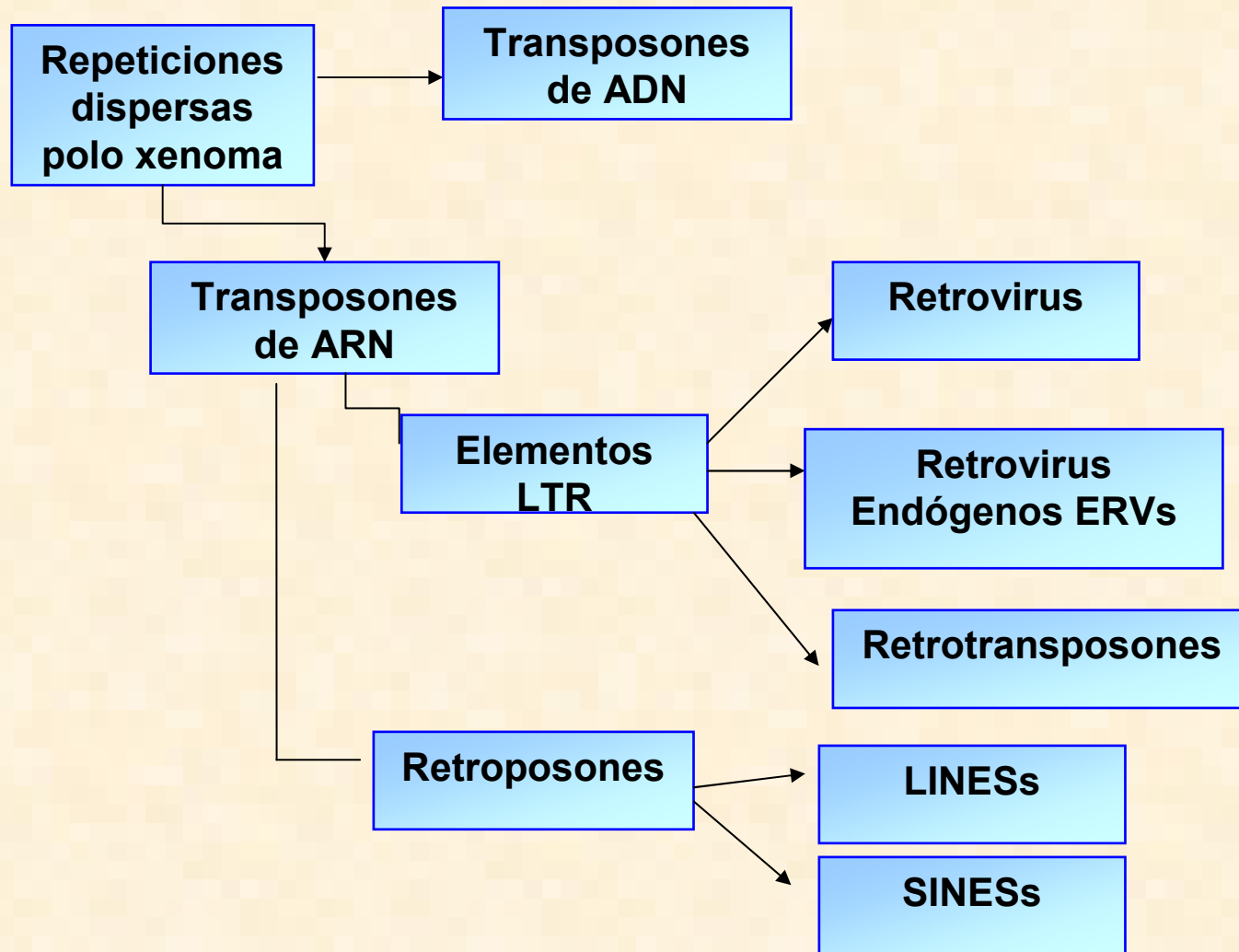
□ **Transposóns de RNA ou retroelementos:** *repeticións dispersas derivadas dun RNA intermediario que se retrotranscribe e insértase noutro lugar do xenoma.*

Elementos LTR (long terminal repeat): caracterízanse pola presenza de longas repeticións terminais (LTR) nos extremos. Existen varios tipos:

- Retrovirus: Virus de RNA en fase lisoxénica e que poden xerar novos virus.
- Retrovirus endóxenos (ERVs) e elementos tipo retrovirus (RTVLs). Similares en estrutura ó grupo anterior pero que a maioría perderon a capacidade para formar novos virus.
- Retrotransposóns: De estrutura similar ós ERVs, son especialmente abundantes en invertebrados e plantas.

Retroposóns: diferénciase basicamente do grupo anterior na ausencia de repeticións terminais longas nos seus extremos.

- **LINEs** (*long interspersed elements*), longos elementos dispersos. Posúen un xene tipo transcriptasa inversa probablemente asociado a súa transposición.
- **SINEs** (*short interspersed elements*) elementos curtos disperso no xenoma. Carecen de transcriptasa inversa propia e utilízan a de outros transposóns para a súa transcripción. O exemplo máis destacable é a familia Alu con un número de copias por xenoma humano próximo ó millón.



DNA altamente repetitivo

Satellite	Predominant Sequence	Total Length	Genome Proportion
I	ACAACT TGTTGA	1.1×10^7	25%
II	ATAACT TATTTGA	3.6×10^6	8%
III	ACAAATT TGTTTAA	3.6×10^6	8%
Cryptic	AATATAG TTATATC		



5' ATAAACTATAAACTATAAACT...
3' TATTTGATATTTGATATTTGA...

Telómeros



Telómeros, atópanse nos extremos dos cromosomas. Nos ser humano esta secuencia é TTAGGG e repítese unhas 2500 veces.

O **ADN codificante** está formado por secuencias altamente conservadas con moi poucas variacións interindividuais e interxeneracionais, xa que do contrario poderíanse ver afectadas funcións básicas para a vida das persoas. Os mínimos cambios que ten lugar, cando son viables, aumentan o polimorfismo de proteínas e encimas e poden ter efectos negativos.



Anemia falciforme



Glóbulos vermellos normais

A anemia falciforme débese á presenza nas persoas que a padecen dunha forma de hemoglobina na que se produciu un cambio no aminoácido 6 da cadea de globina Beta (o ácido glutámico substituíuse por valina). Debido a isto os glóbulos vermellos adoptan forma de foz.

O **ADN non codificante** de tipo repetitivo en tándem, presenta maior variabilidade duns individuos a outros, xa que algunhas destas secuencias afectan en menor medida á fisioloxía do individuo. As variacións debidas a cambios de bases no ADN durante o proceso de replicación do mesmo ou de intercambio de ADN (recombinación) durante a formación das células xerminais (meiose), fan que se modifiquen o número de repeticións ou a orde das bases dun determinado fragmento repetitivo, podendo producirse nun *locus* sinxelo ou en múltiples *loci*, sendo este a orixe da variación que fai que non haxa dous persoas, a excepción dos xemelgos univitelinos, que teñan a mesma secuencia do ADN.

Single Nucleotide Polymorphism: SNP

Os lugares da secuencia de ADN onde os individuos difiren nunha soa base coñecese como Polimorfismos dunha soa base (***Single Nucleotide Polymorphism: SNP***).

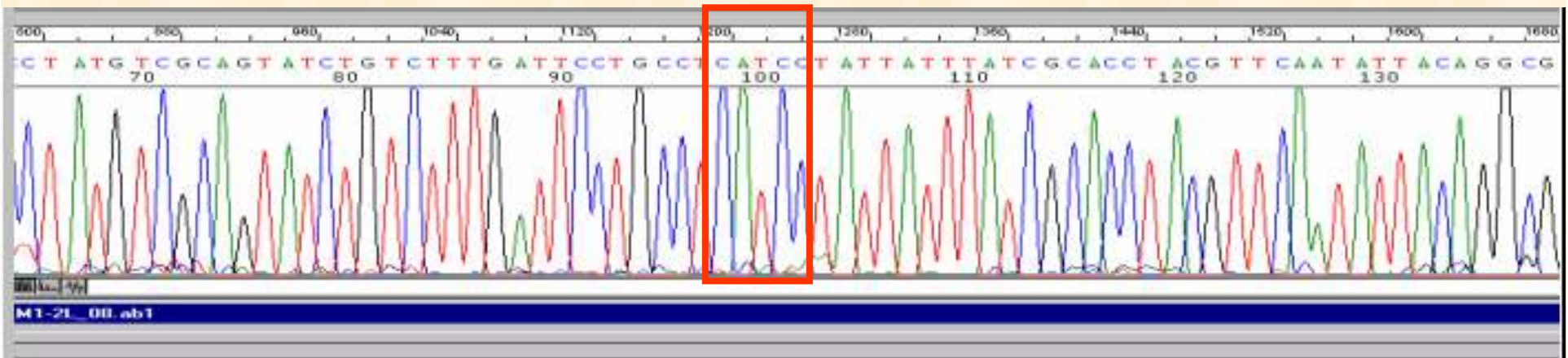
ATCCCATCGGG**T**ATCATAACGGG



ATCCCATCGGG**A**ATCATAACGGG

Os SNP utilízanse para detectar pegadas xenéticas ou localizar xenes implicados en enfermidades xenéticas.

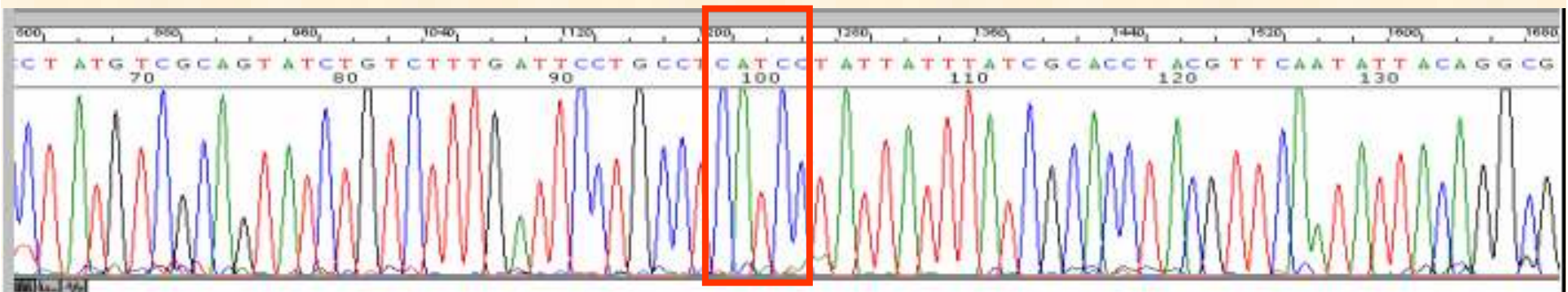
PELOS ATOPADOS NA MAN DA VÍCTIMA



INDIVIDUO 1



INDIVIDUO 2



A pegada xenética

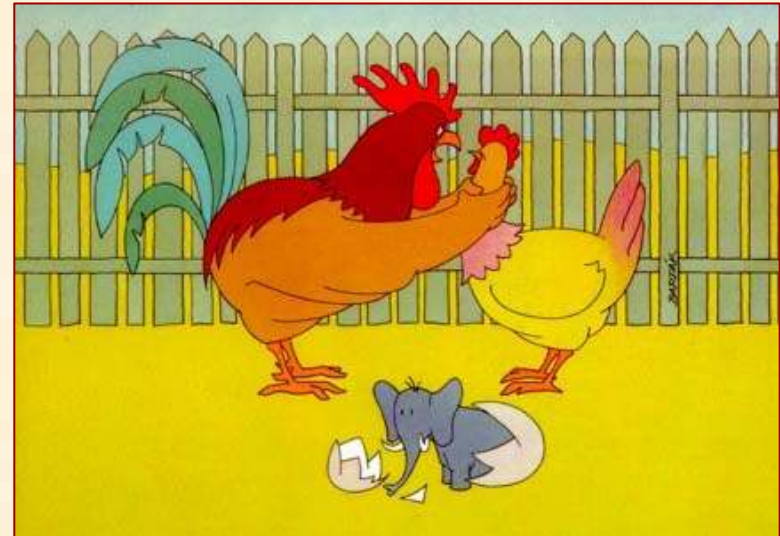
Medicina forense :

Denomínase: **Proba de ADN.**

Consiste: en comparar rexións de material xenético non codificante de dúas mostras.

Realízase: dun resto de saliva nun cigarro (contén células da mucosa bucal), unha mancha de sangue (leucocitos), restos de esperma (espermatozoides) ou un pelo (células do folículo piloso), etc.

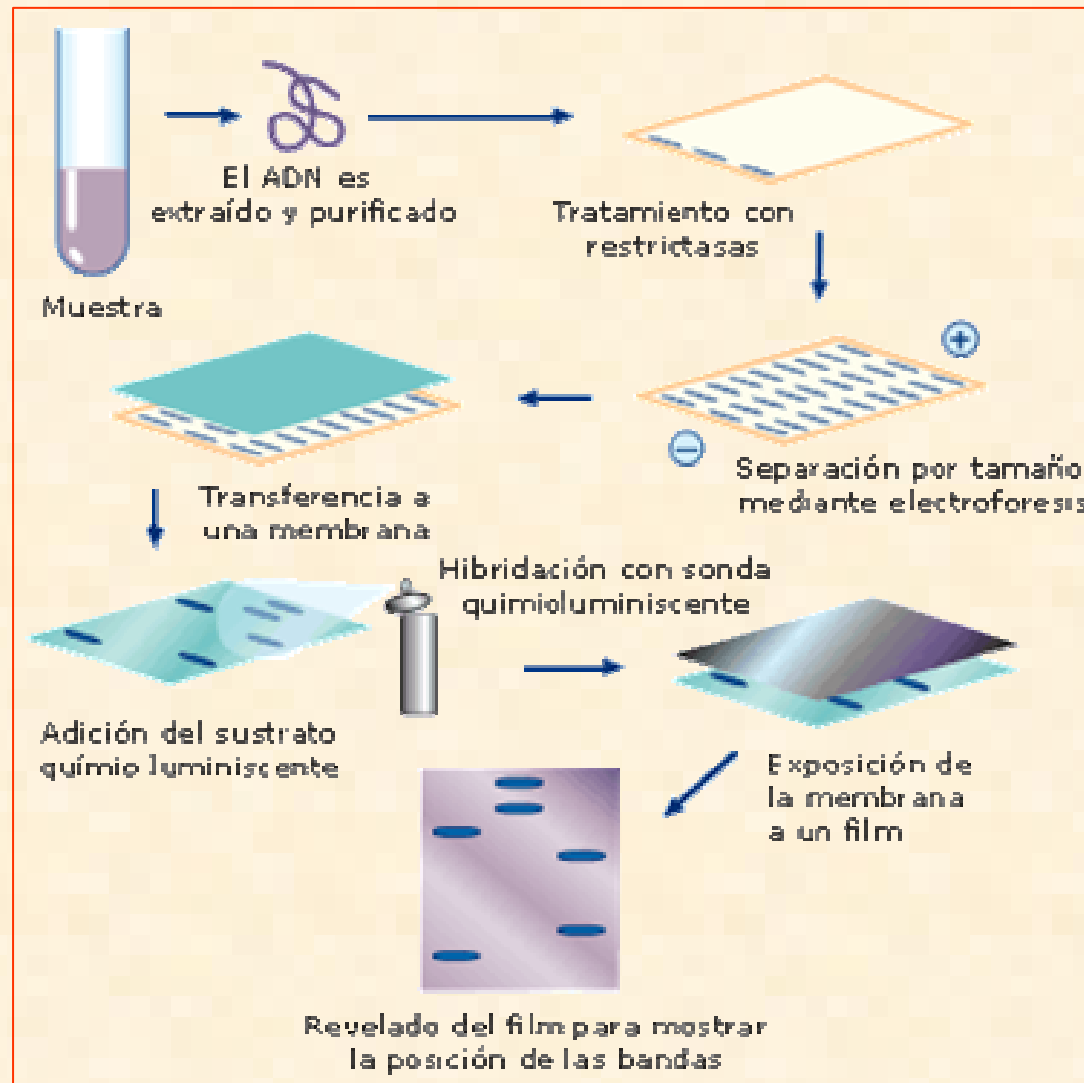
A pegada xenética tamén pode utilizarse para realizar **probas de paternidade** e para **identificar a persoas desaparecidas.**



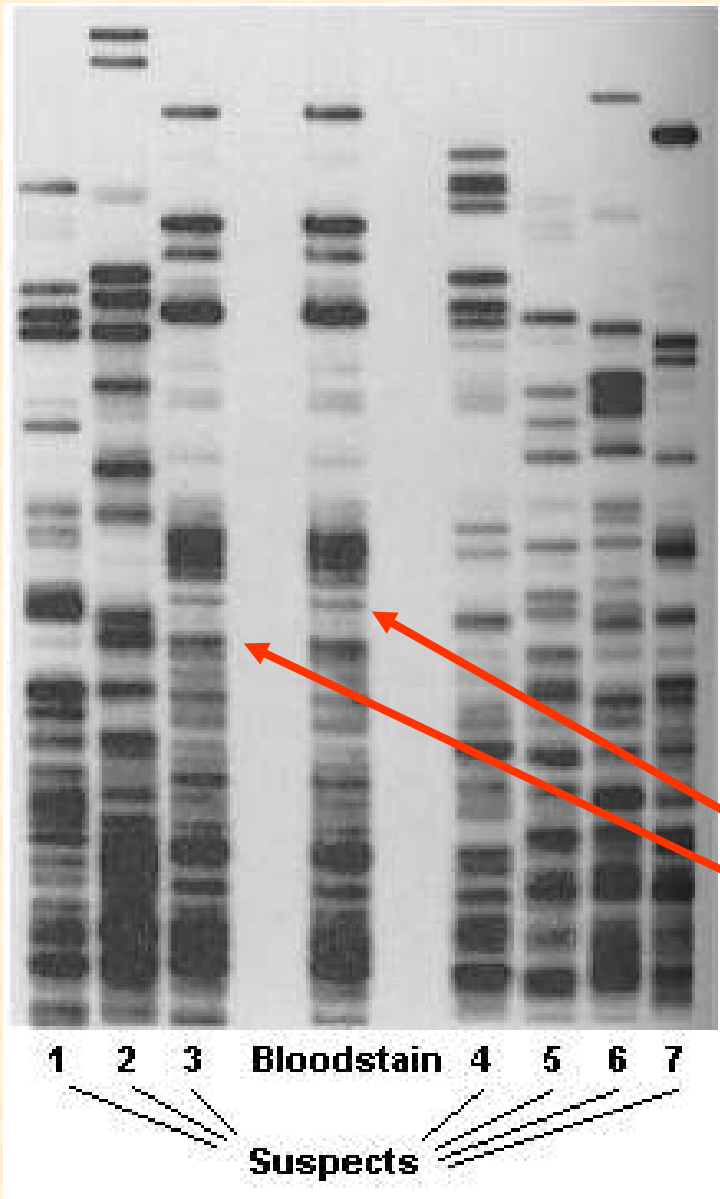


As características xerais do **ADN non codificante** faino especialmente útil para a súa aplicación en Medicina Forense. Así por exemplo os microsateles utilízanse para establecer consanguinidade ou parentesco e serve como pegada biolóxica para a identificación de persoas na escena dun crime ou outro acto.

A pegada xenética



Técnica utilizada na proba del ADN



¿Quen é o asasino?

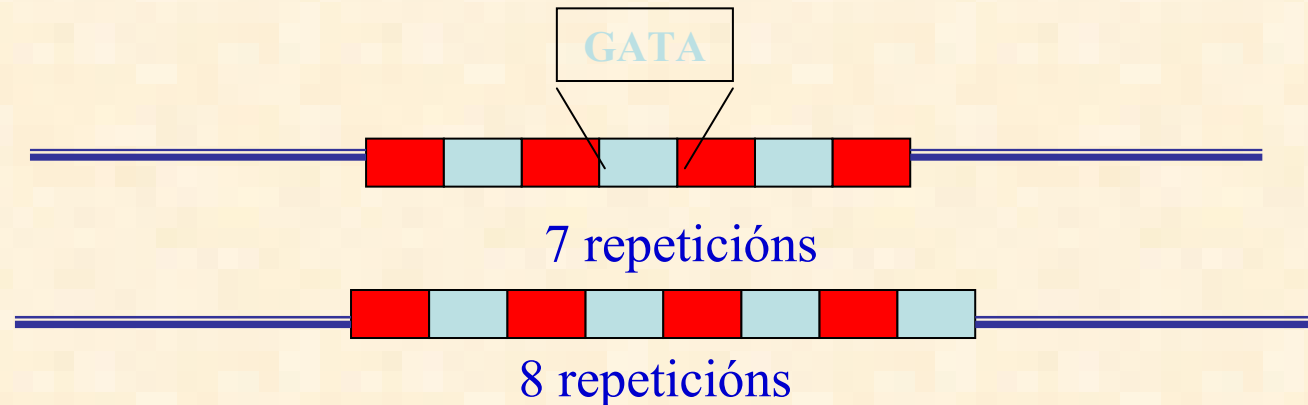
No centro o ADN do sangue recollida no lugar do crime.

A ambos lados mostras de ADN de 7 sospeitosos.

Comparativamente similar ó indicio primario será a mostra nº 3

Repeticións curtas en tandem

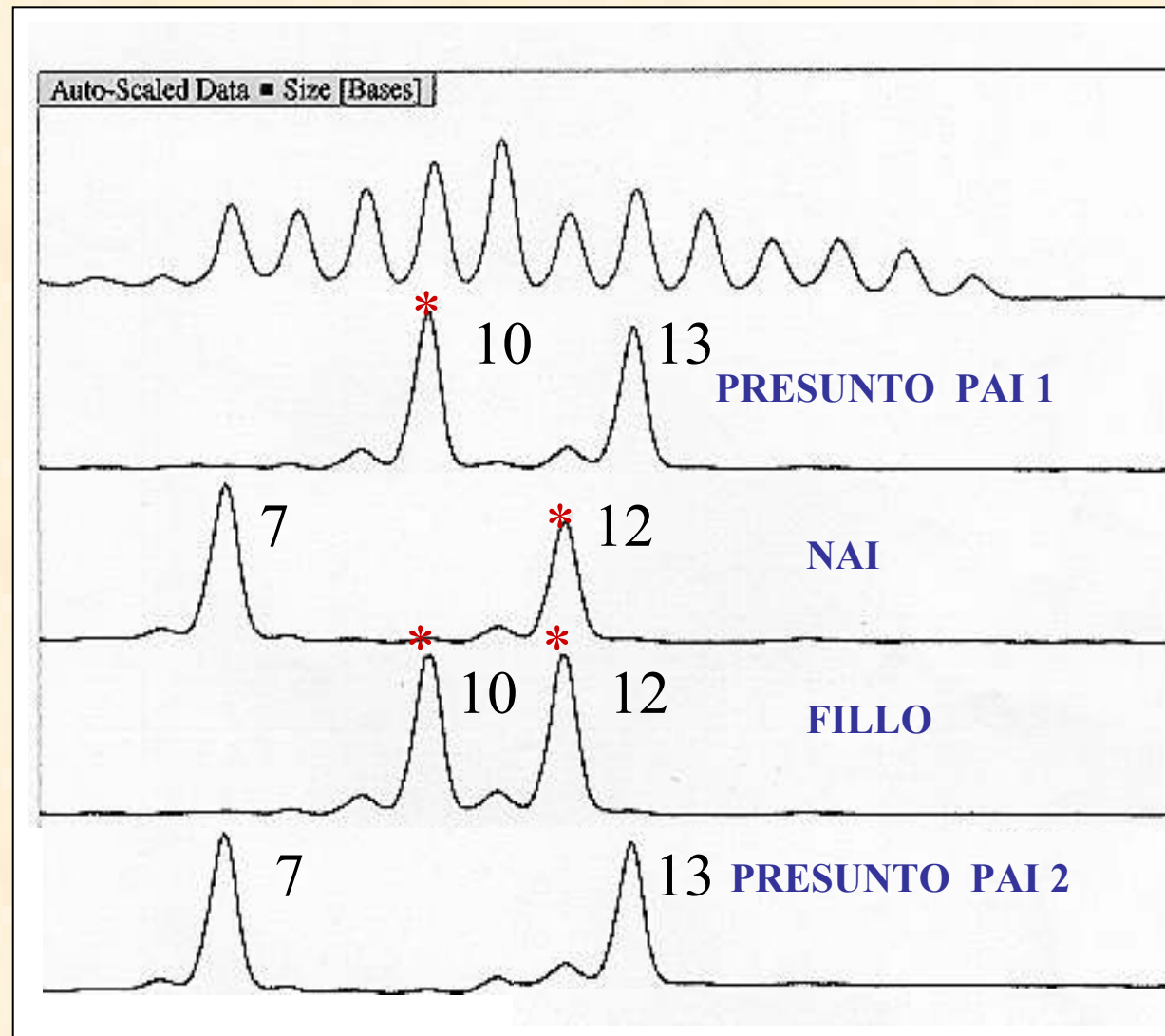
Short Tandem Repeats (STRs)

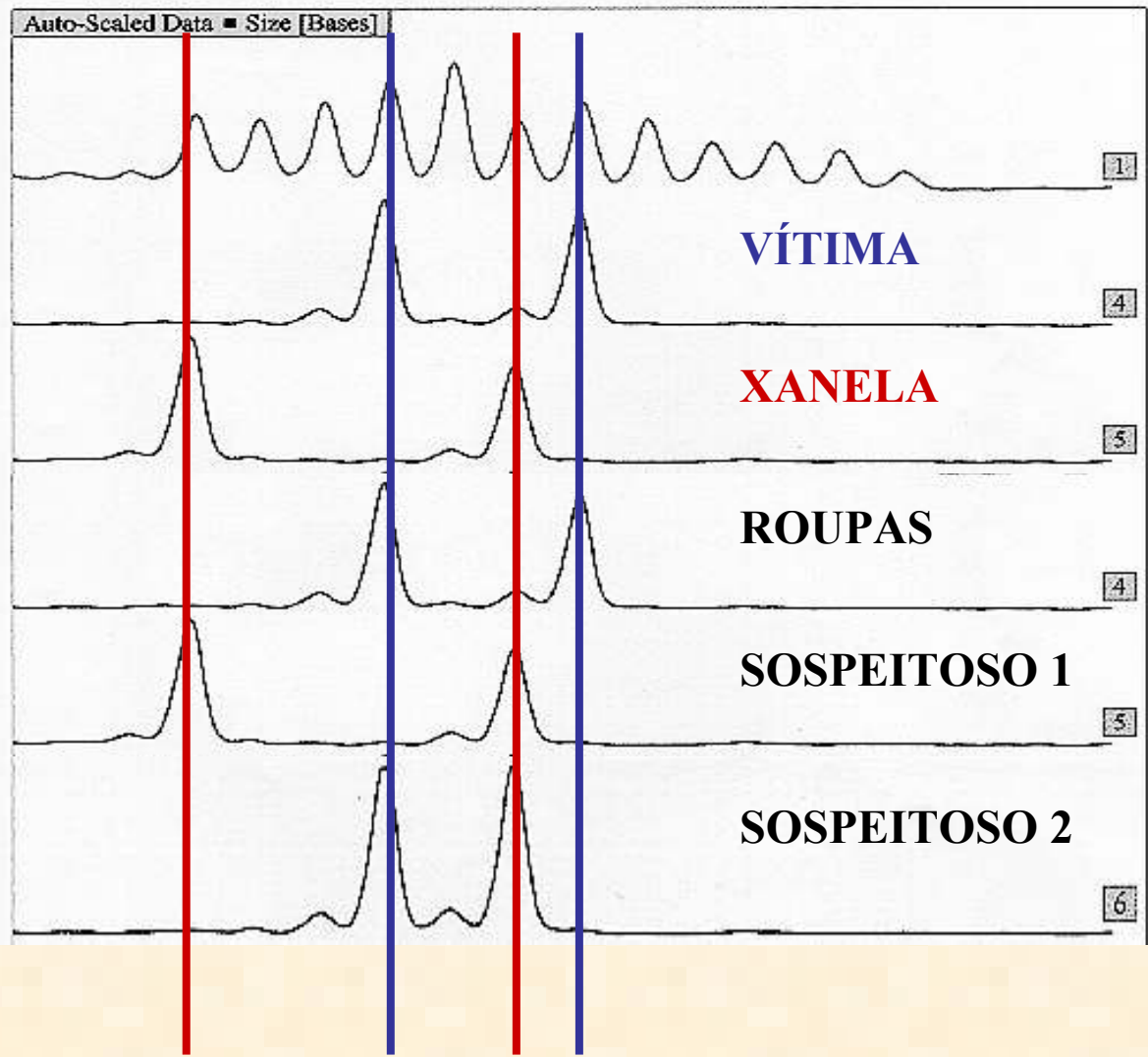


As STR son secuencias curtas de ADN, normalmente cunha lonxitude de 2 a 5 pares de bases, que se repiten moitas veces de forma consecutiva (gatagatagatagata representaría 4 copias dispostas cabeza con cola do tetrámero "gata").

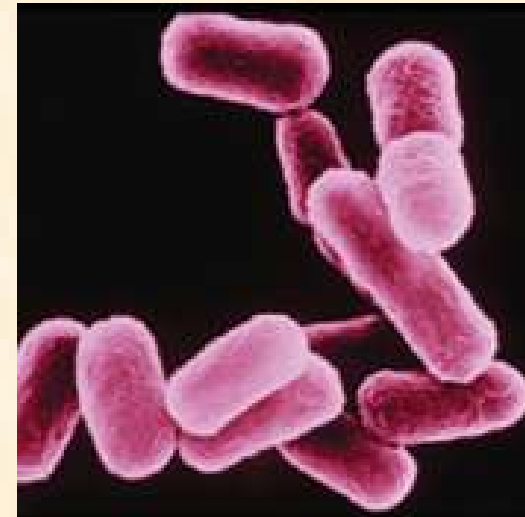
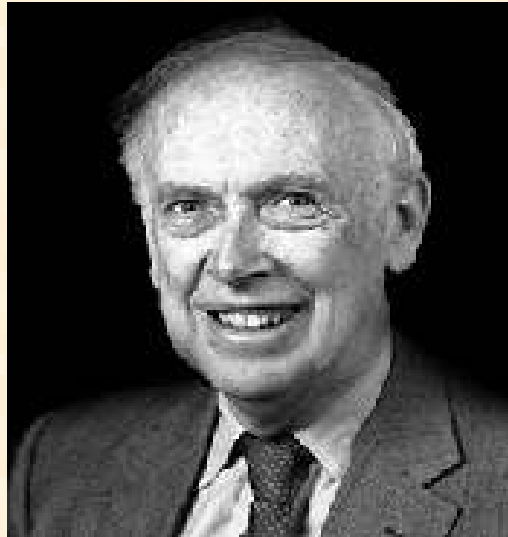
Os polimorfismos en STR (microsatelites) débense ó distinto número de copias do elemento repetido que pode aparecer nunha poboación de individuos.

Los STRs son moi útiles para comparar ADNs porque se necesita moi pequena cantidade de mostra.





PROXECTO XENOMA HUMANO



Abril 1990

O NIH e o DOE acordan un plan para completar mapas xenéticos, a secuencia completa dun mapa xenético e ademais secuenciar organismos como: *Mycoplasma capricolum*, *Escherichia coli* (na figura), *Caenorhabditis elegans*, *Saccharomyces cerevisiae* e o *Xenoma humano*

Objetivos do proxecto genoma humano



1.- Localizar os xenes principalmente os implicados en enfermidades hereditarias



2.Elaborar un mapa de ligamento co fin de colocar de forma ordenada os xenes.



3.- Secuenciar cada xene, é dicir, averiguar a secuencia de nucleótidos.



4.- Determinar a función dos xenes

Estamos aquí!!!!

Historia do Proxecto Xenoma Humano

- 1983 Ubícase o xene para unha enfermidade humana: A enfermidade de Huntington (no cromosoma 4).
- En 1985 iniciouse unha xigantesca empresa dirixida a describir a totalidade da *información xenética contida nos 23 pares de cromosomas humanos mediante a secuenciación do ADN.*
- El 01-10-1990 comeza o traballo de laboratorio, a duración estimouse en 15 anos.
- Desígnase inicialmente a James Watson a cargo do Proxecto.
- 1995 pode lerse o xenoma de *Haemophilus influenzae* e *Sacharomyces cereviseae*.
- 1999 tense o primeiro cromosoma secuenciado.

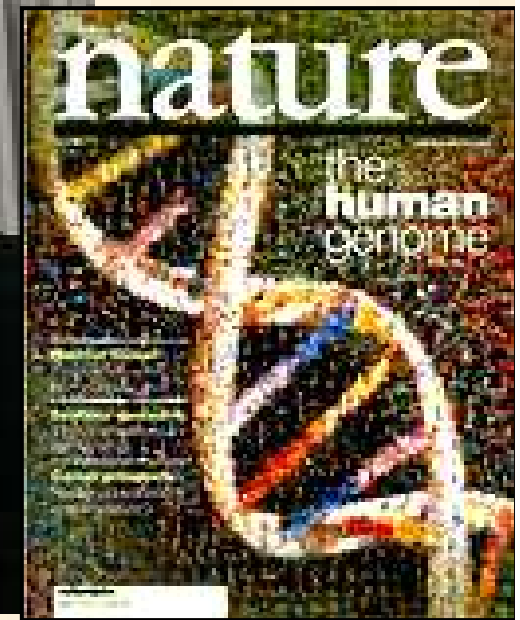
Francis Collins é o director do Instituto Nacional para a Investigación do Xenoma Humano.



Dous adversarios participaron na carreira por secuenciar o material xenético humano. O consorcio público Proxecto Xenoma Humano (PGH) e a empresa privada PE Celera Genomics.

2001

O consorcio HGP publica un borrador de traballo en *Nature* (15 Febreiro), e Celera en *Science* (16 Febreiro)



2003: Primeira fase considerada concluída. Termínase a secuencia completa do Xenoma Humano

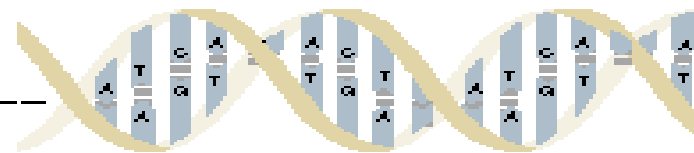
Cómo se lee el ADN

La totalidad de los genes del ser humano (genoma) está repartida en **46 cromosomas** agrupados en 23 pares. El número total de genes humanos se estima en unos 100.000 y las secuencias de bases, en 3.000 millones.

Realmente ó redor de 30.000 xenes

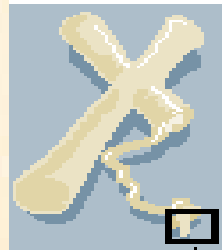
La doble hélice

Cada cromosoma contiene una larguísima molécula, la doble hélice del ADN, que es el sustrato de la información genética. Consiste en dos largas hileras. Cada hilera soporta las letras de la información genética, llamadas bases.



Son de cuatro tipos: **A, G, C y T**. En cada hilera las bases pueden disponerse en cualquier orden y ese orden es el contenido de la información. La secuenciación consiste en determinar ese orden exacto.

La técnica del consorcio público



- 1 Los 23 pares de cromosomas humanos se dividen y estudian uno a uno para poder ser analizados.
- 2 Se dividen los cromosomas en regiones para ir afinando mapas de una escala cada vez más precisa.

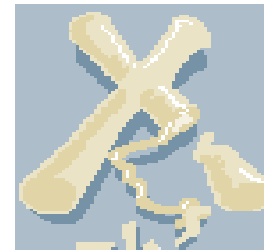
El sistema de estudio es más lento pero más ordenado.

Mapa



La secuenciación es el último paso

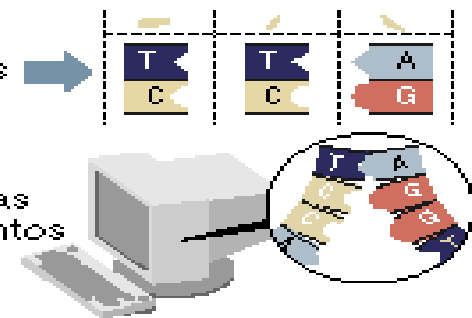
La técnica de Celera (Shotgun)



- 1 Los 23 pares de cromosomas humanos se parten en millones de fragmentos al azar.
- 2 Cada fragmento resultante se secuencia para su posterior ordenamiento del mapa genético.

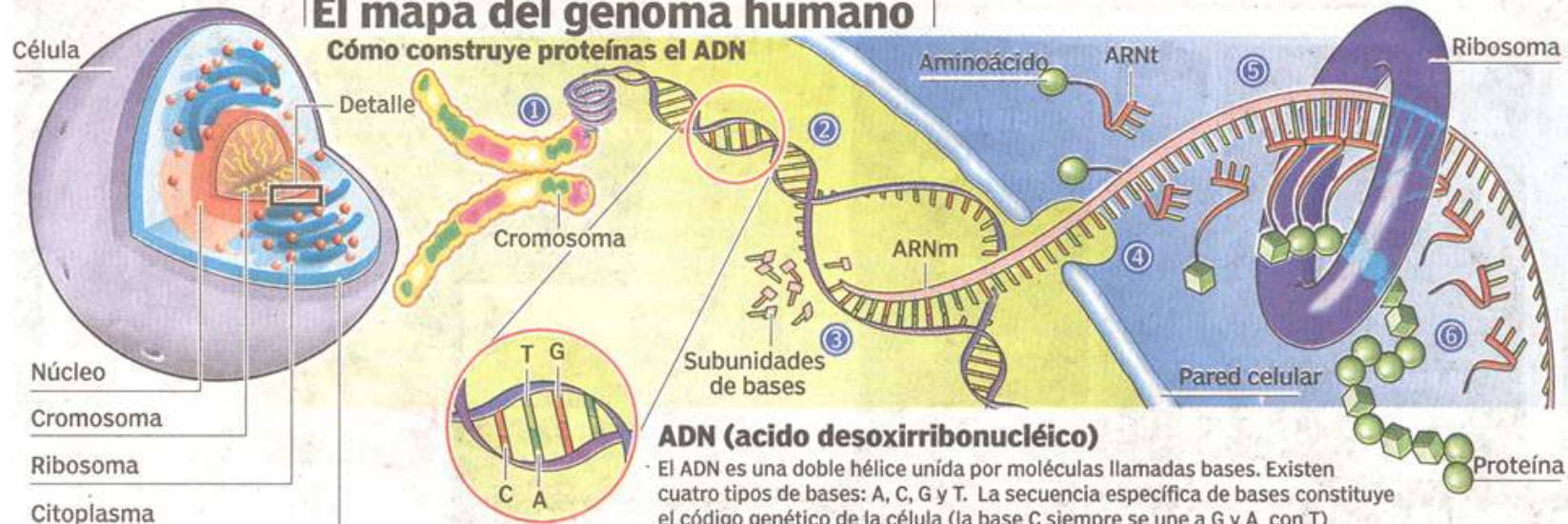
Fragmentos

- 3 Los programas informáticos intentan ensamblar las secuencias gracias a los solapamientos entre fragmentos.



El mapa del genoma humano

Cómo construye proteínas el ADN



- ① **Cromosoma**
Un racimo de moléculas enrolladas en una fibra de ADN
- ② La fibra de ADN se desenrolla para autocopiarse
- ③ Las bases se emparejan formando una cadena llamada ARN mensajero (ARNm)
- ④ El ARNm atraviesa la pared del núcleo y se une a uno o varios ribosomas (moléculas que construyen proteínas)
- ⑤ Los aminoácidos se unen a las hebras de ARN de transferencia (ARNt) adherido al ARNm en orden estricto de bases
- ⑥ Tras pasar por el ribosoma, el ARNt se disgrega, dejando los aminoácidos unidos en una cadena que constituye la proteína

La carrera del genoma humano

1866	1871	1951	1953	1960	1977	1975-79	1986	1995	1999	2003
Los genes son descubiertos	Descubren los ácidos nucleicos	Primera secuencia de proteína	Estructura del ADN	Explicación del código genético	Primera secuenciación de ADN	Primeros genes humanos aislados	Automatización de la secuenciación del ADN	Primer genoma	Primer cromosoma humano	Mapa del genoma humano completado

Fuente: Nature, The Guinness Encyclopedia of the Human Being

REUTERS/LA VOZ

¿Que sabemos?

- Que temos uns 3 mil millóns de pares de bases (A,C,T, G)
- Que un xene ten de media 3000 bases pero varía moito o seu tamaño
- Que o número de xenes total que temos é ó redor de 30.000
- Que os seres humanos somos un 99,9% idénticos
- Que só o 2% do xenoma codifica proteínas. A gran maioría do ADN humano non ten unha función coñecida
- Que a maioría do ADN é non codificante (mal chamado lixo)
- Que un 50% do ADN son motivos que se repiten (moito máis que ningunha outra especie)

Variación no Xenoma Humano



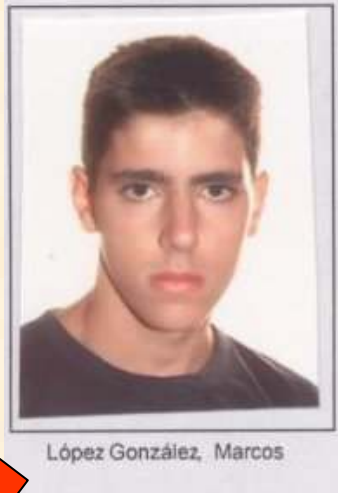
99.9% idénticos

Non se atoparon diferencias significativas nos xenomas de humanos pertencentes a distintas razas. O que deixa sen base científica as elucubracións concernientes á “superioridade” dunha raza sobre outra.





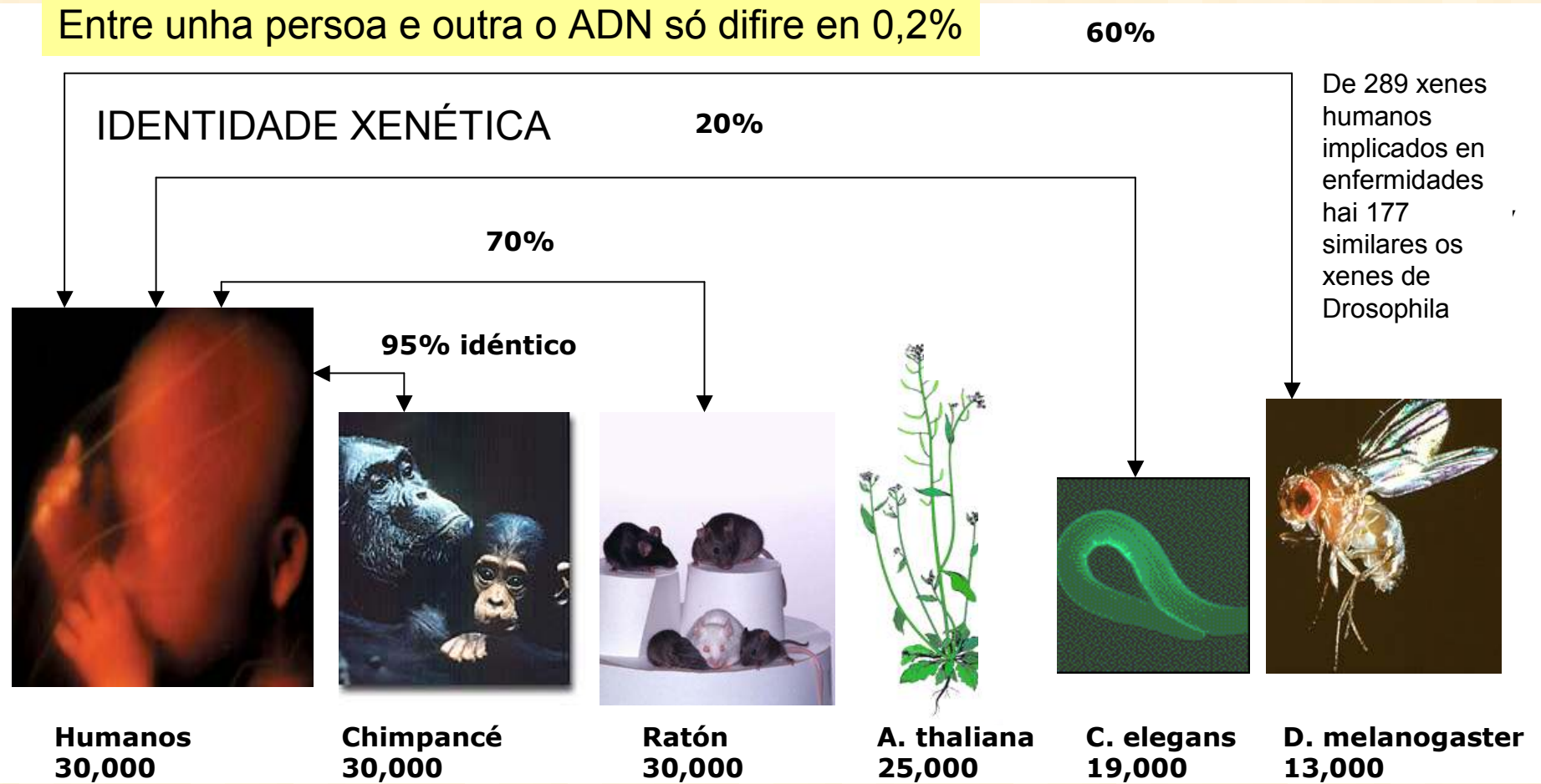
99.9% idéntica



99.9% idéntico

**A secuencia do
xenoma humano
é única para cada
individuo**

Entre unha persoa e outra o ADN só difire en 0,2%



¿Que non sabemos?

- O número exacto dos xenes.
- A función da maioría.
- O xeito de como interaccionan.
- A maioría dos xenes implicados en enfermidades comúns.
- Moitas das funcións do ADN non codificante.
- A regulación exacta da expresión xénica, síntese de proteínas e eventos postrasduccionais.
- O proteoma (contido total de proteínas, a súa función e interrelacións).

SE ME PERDIÓ MI DNI.
¿ PUEDO VOTAR CON EL
ADN?

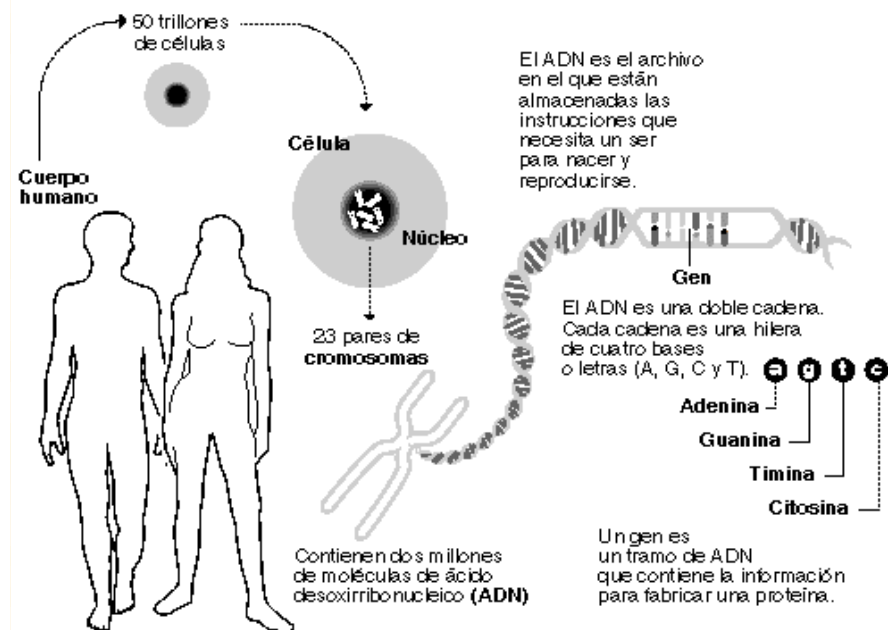
MESA 4



Landru!

El secreto de la vida

DEL CROMOSOMA A LOS GENES



FASES DEL PROYECTO GENOMA HUMANO

1 SECUENCIAR

Identificar las letras químicas que componen cada gen humano.



2 IDENTIFICAR

Tras descifrar cuál es el orden correcto de cada letra, se lee la información de cada gen, y una vez identificado, se averigua su función.



3.200 millones de pares de bases forman genes, repartidos entre los 23 pares de cromosomas

CIFRAS Y CLAVES DEL GENOMA

- La secuencia está completa en un **88%**.
- Más del **90%** de la secuencia es exacta en un **99,99%**.
- Menos del **5%** de la secuencia contiene genes portadores de instrucciones para hacer proteínas.
- En ADN, el genoma humano es 200 veces menor que el de la ameba.
- Al menos 223 de nuestros genes parecen proceder de las bacterias.

CUÁNTOS GENES TENEMOS

Según el consorcio público: **31.780 Genes**

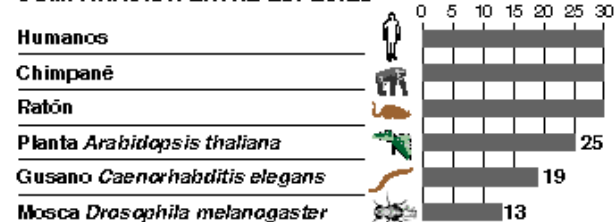
(22.000 identificados, el resto supuestos)

Según Celera genomis: **38.500 genes**

(26.588 identificados)

En el proyecto de secuenciación pública han trabajado 20 grupos de seis países, y ha costado 55.000 millones de pesetas.

COMPARACIÓN ENTRE ESPECIES



Si todo el ADN del cuerpo humano fuera puesto en fila haría 600 veces el trayecto entre la Tierra y el Sol.

La información llenaría una pila de libros de 61 metros de altura o 200 guías telefónicas de 500 páginas cada una.

Las cuatro letras del ADN llevan las instrucciones para hacer todos los organismos. Cada bloque de tres letras corresponde a un aminoácido.

Hay 20 aminoácidos distintos que se usan en cualquier combinación para producir proteínas tan diferentes como la creatina para el pelo o la hemoglobina de la sangre.

La mayor parte del ADN (más de un 95%) en el Genoma Humano no tiene función conocida y se considera ADN basura.

El ADN humano es al menos en un 98% idéntico al de los chimpancés y otros primates, los parientes más cercanos del hombre.

Entre una persona y otra el ADN difiere sólo en un 0,2%.





*Departamento Bioloxía e Xeoloxía
I.E.S. Otero Pedrayo. Ourense.*