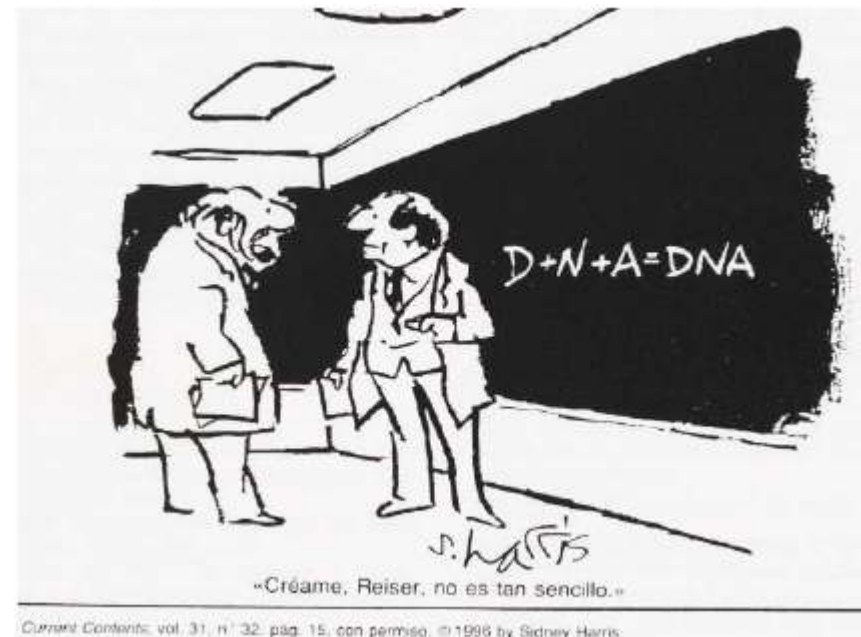
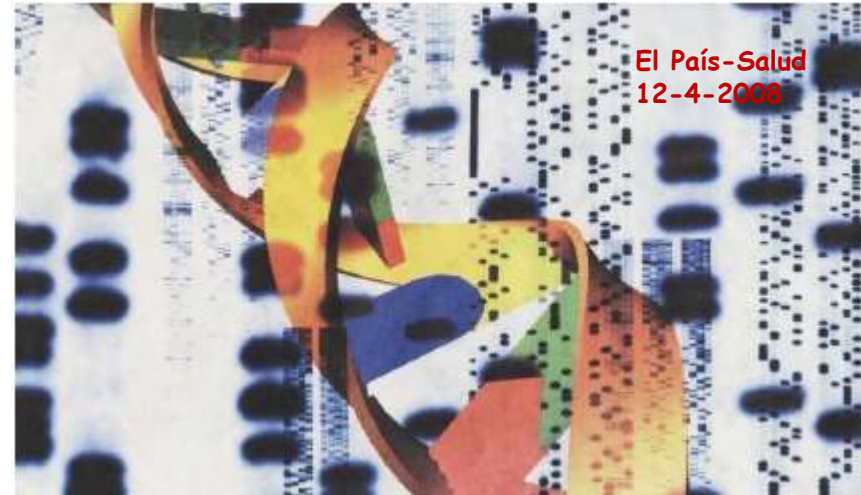


# Tema 29. MAPAS GÉNICOS Y GENOMA HUMANO

- \* Mapas genéticos y mapas físicos.
- \* Hibridación "in situ"
- \* Híbridos de células somáticas
- \* Investigación de genes candidatos
- \* Genoma humano
- \* Mapa de haplotipos



## MAPAS GÉNICOS: Mapas físicos y genéticos

Los mapas físicos localizan a los genes en los cromosomas midiendo la distancia entre sus loci en pares de bases. La longitud física de todo el genoma humano, referida al complemento cromosómico n (haploide), es de 3.200 Mb (millones de pares de bases). La realización de mapas físicos implica el uso de técnicas citogenéticas y moleculares.

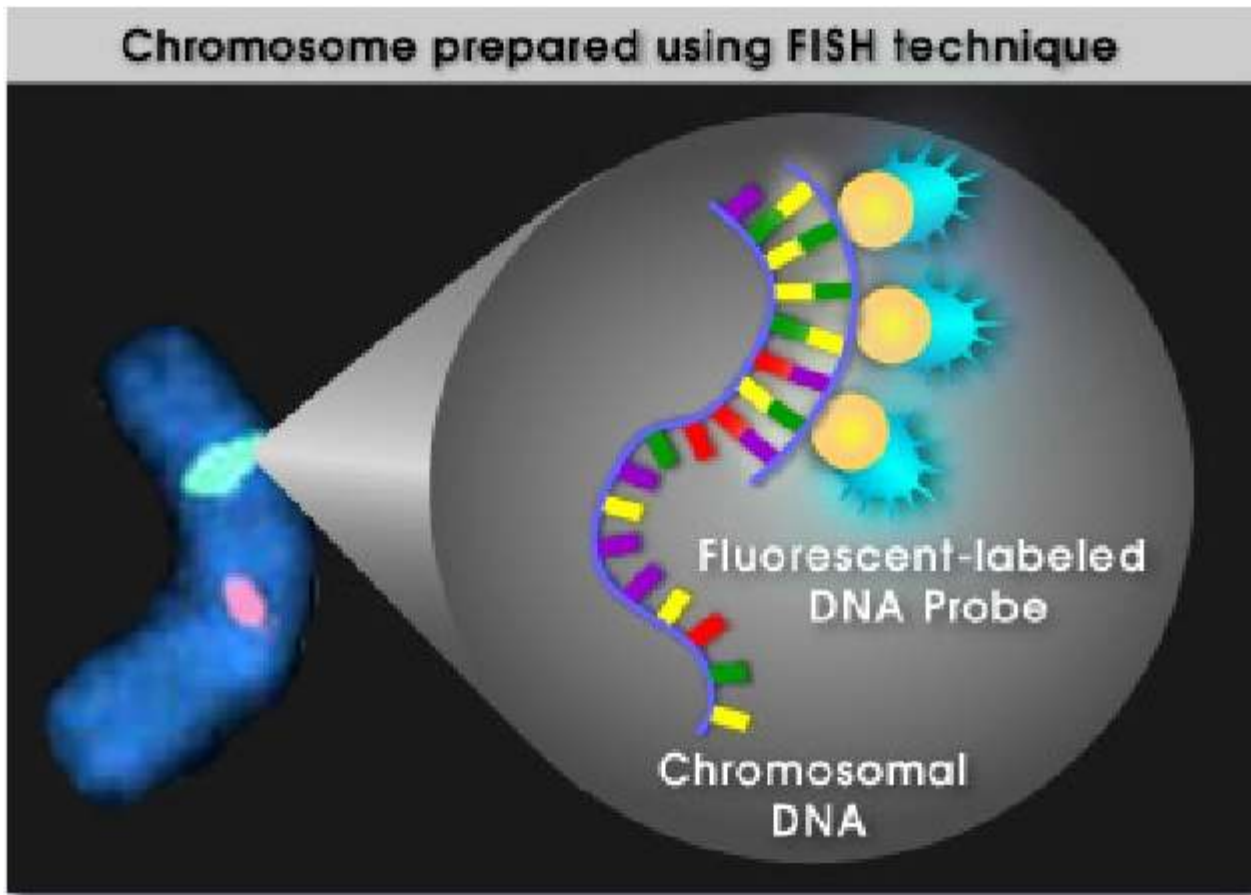
Los mapas genéticos estiman esa distancia por análisis estadístico de ligamiento, valorando en los "progenitores informativos" la frecuencia con la que los haplotipos familiares cambian de fase al formarse los gametos.

La longitud genética aproximada de los 23 (n) cromosomas humanos es de unos 3.615 cM. Es imprescindible que entre cada par de homólogos se dé al menos un sobrecruzamiento para que ambos cromosomas se separen correctamente en Anafase I. Dado que los sobrecruzamientos no se dan de forma homogénea (hay más en las meiosis femeninas y cerca de los telómeros) los cálculos de los mapas genéticos son sólo aproximados.

Utilizando numerosos marcadores se estima que como media 1,13 cM equivalen a 1 Mb.

## LOCALIZACIÓN DE GENES POR HIBRIDACIÓN "IN SITU"

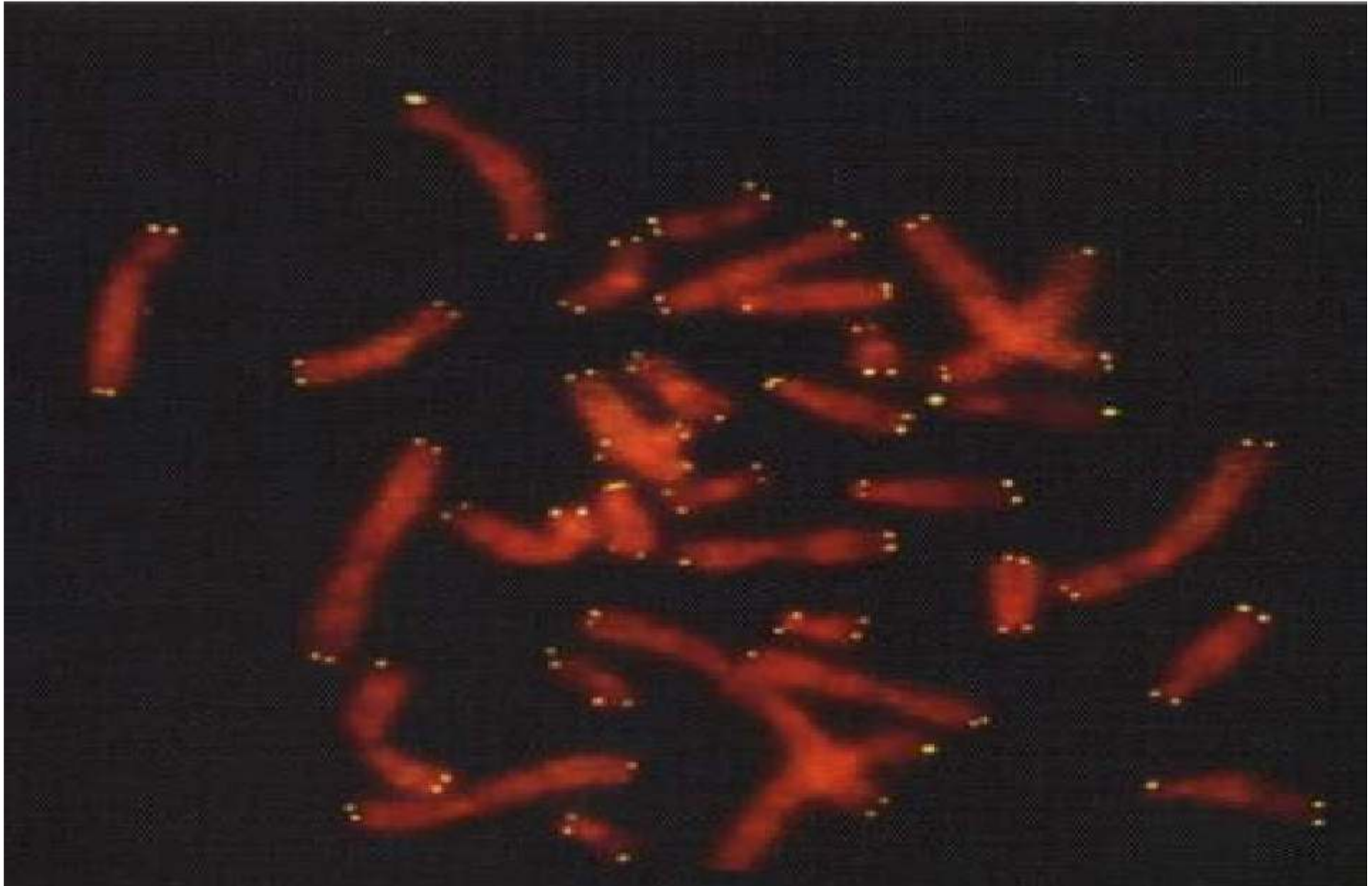
Una vez conocida la secuencia de un gen, puede clonarse y marcarse con fluorescencia parte de ella, para que hibride directamente en el locus del cromosoma metafásico correspondiente.



La señal se sitúa específicamente en el locus del fragmento del gen cuya secuencia es complementaria a la marcada (sonda).

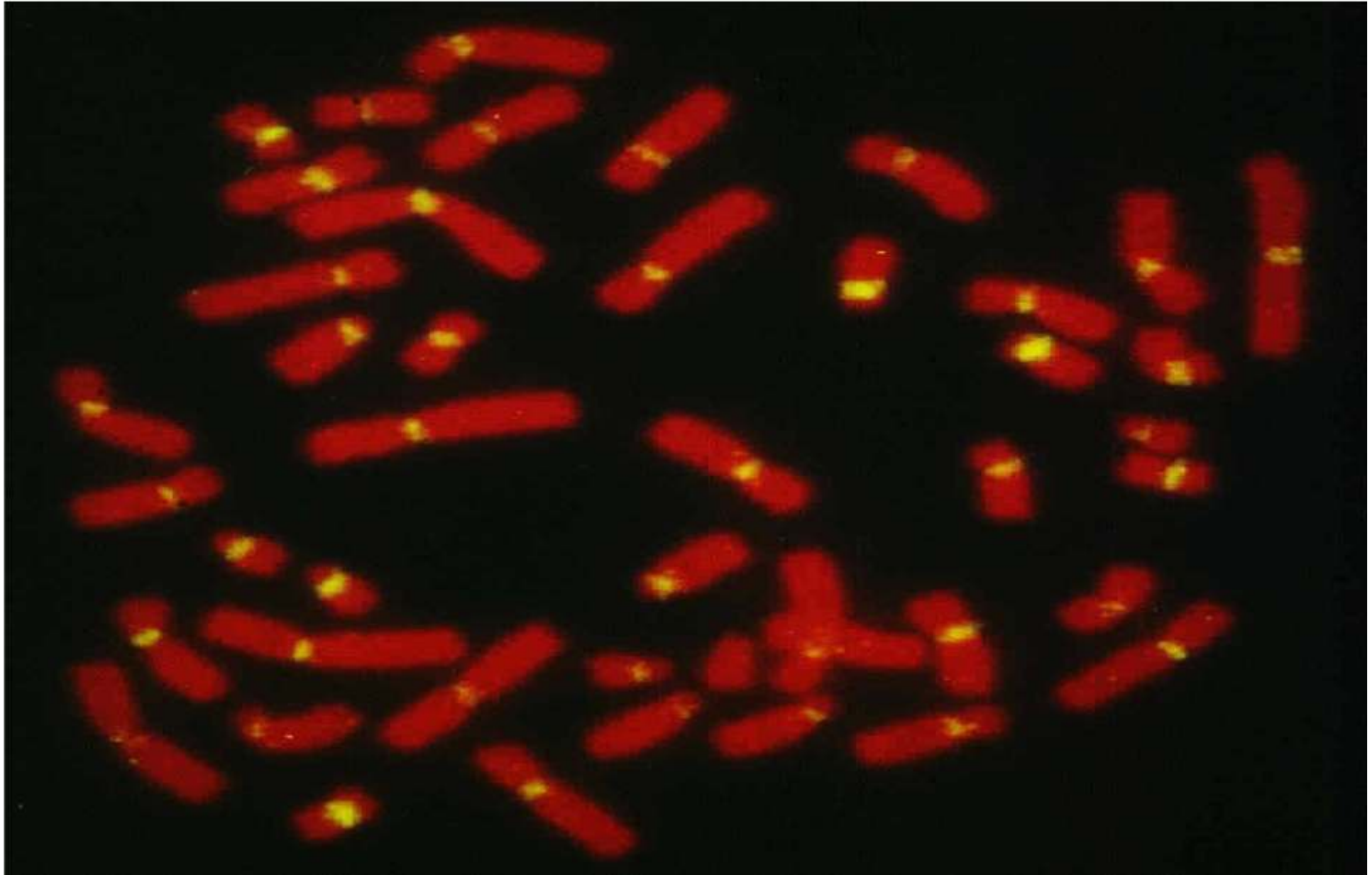
Pueden utilizarse varias sondas en la misma prueba, marcando así distintos loci.

FISH con una sonda para la secuencia telomérica TTAGGG





HIBRIDACIÓN "IN SITU": SONDA PARA REPETICIONES DE ADN  $\alpha$  SATÉLITE  
CENTROMÉRICO

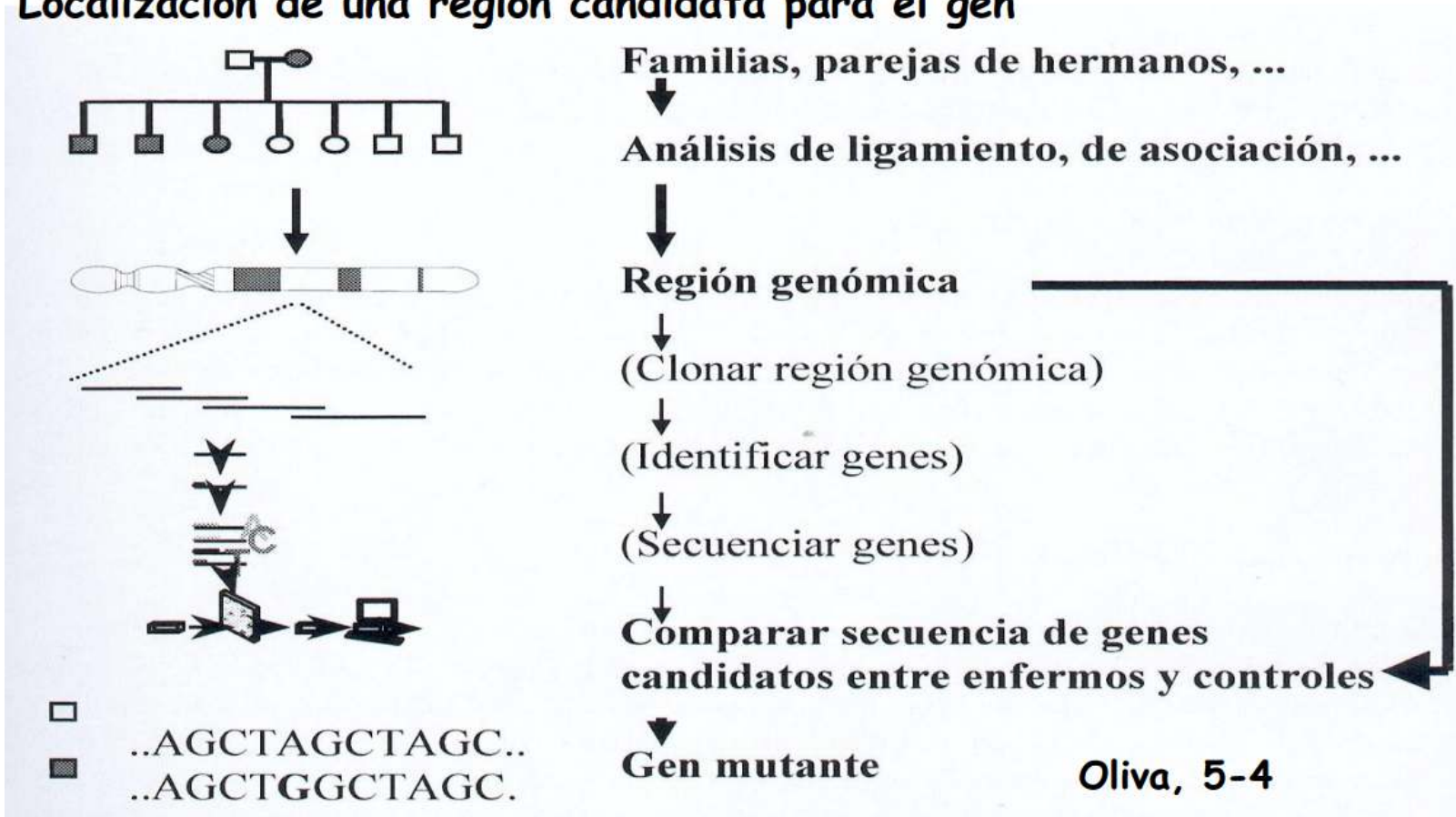


¿Que se hace si no se conoce la secuencia del gen?

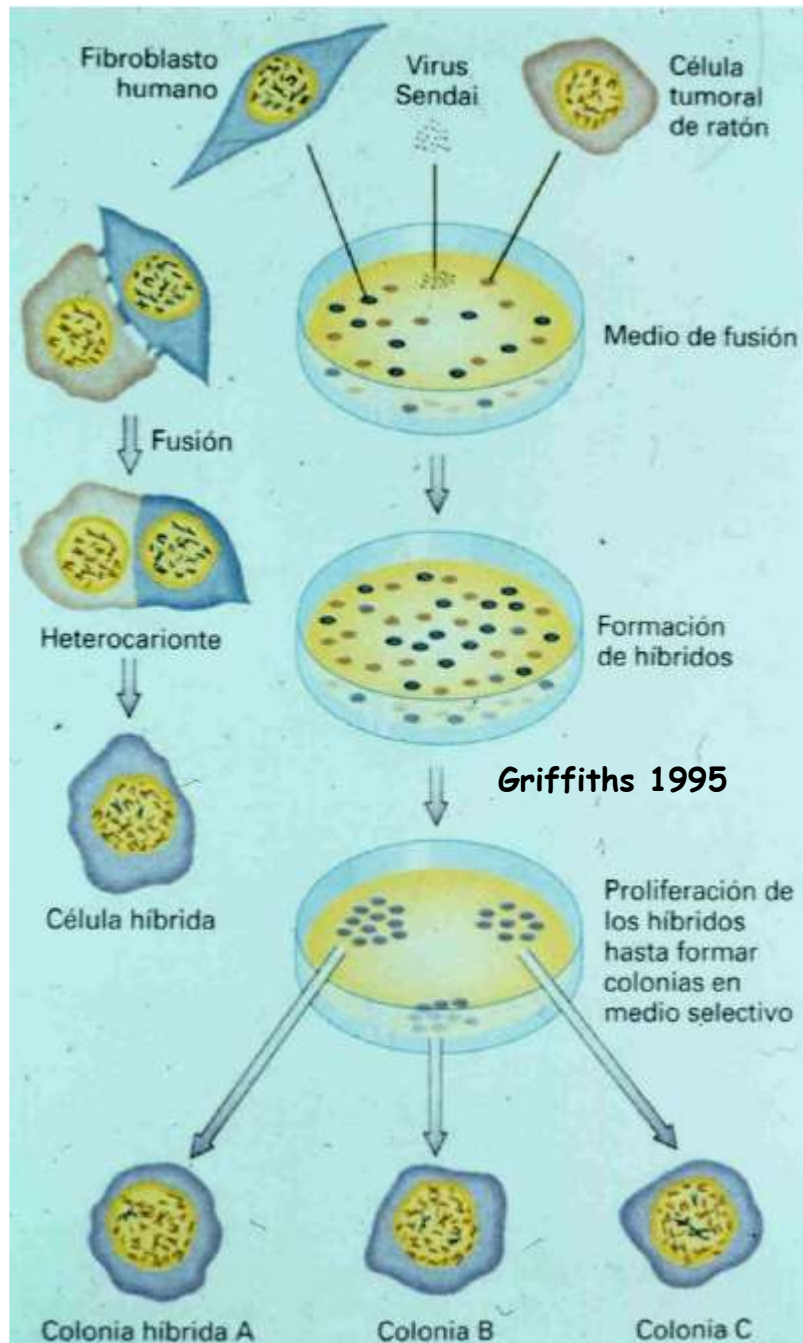
## Localización de genes por clonación posicional

Método indirecto de búsqueda de genes: solo se conoce el fenotipo  
Estudio de marcadores polimórficos en todos los miembros de la familia  
Análisis de ligamiento entre los marcadores y la manifestación del fenotipo

Localización de una región candidata para el gen



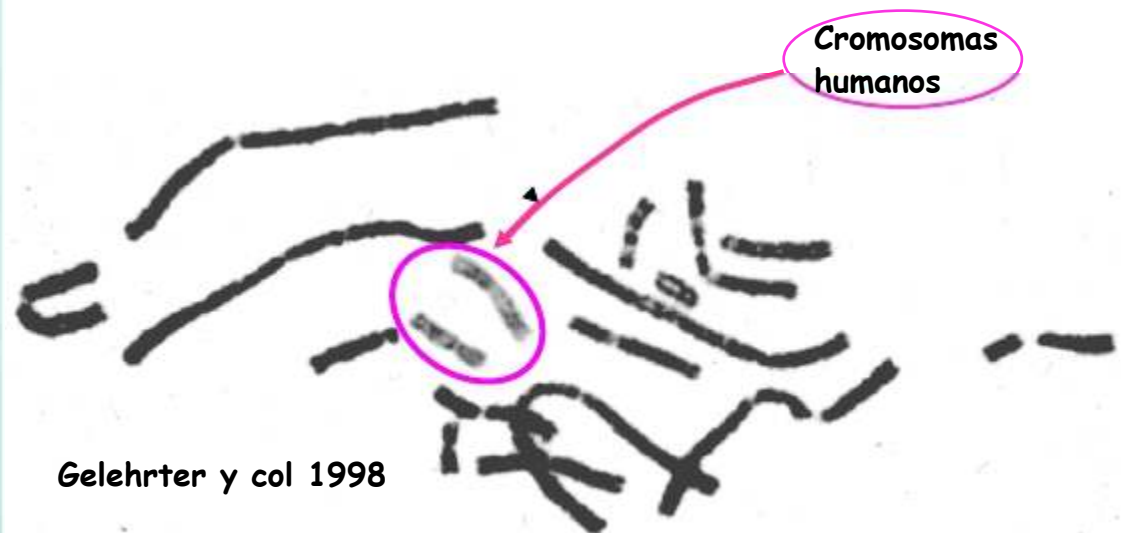
## MAPAS FÍSICOS CON CÉLULAS SOMÁTICAS HÍBRIDAS



Se hibridan células somáticas humanas y de ratón, perdiéndose la mayoría de los cromosomas humanos.

Después de varias divisiones, se obtienen líneas celulares con pocos cromosomas humanos, distintos a los del ratón, como se muestra en la imagen.

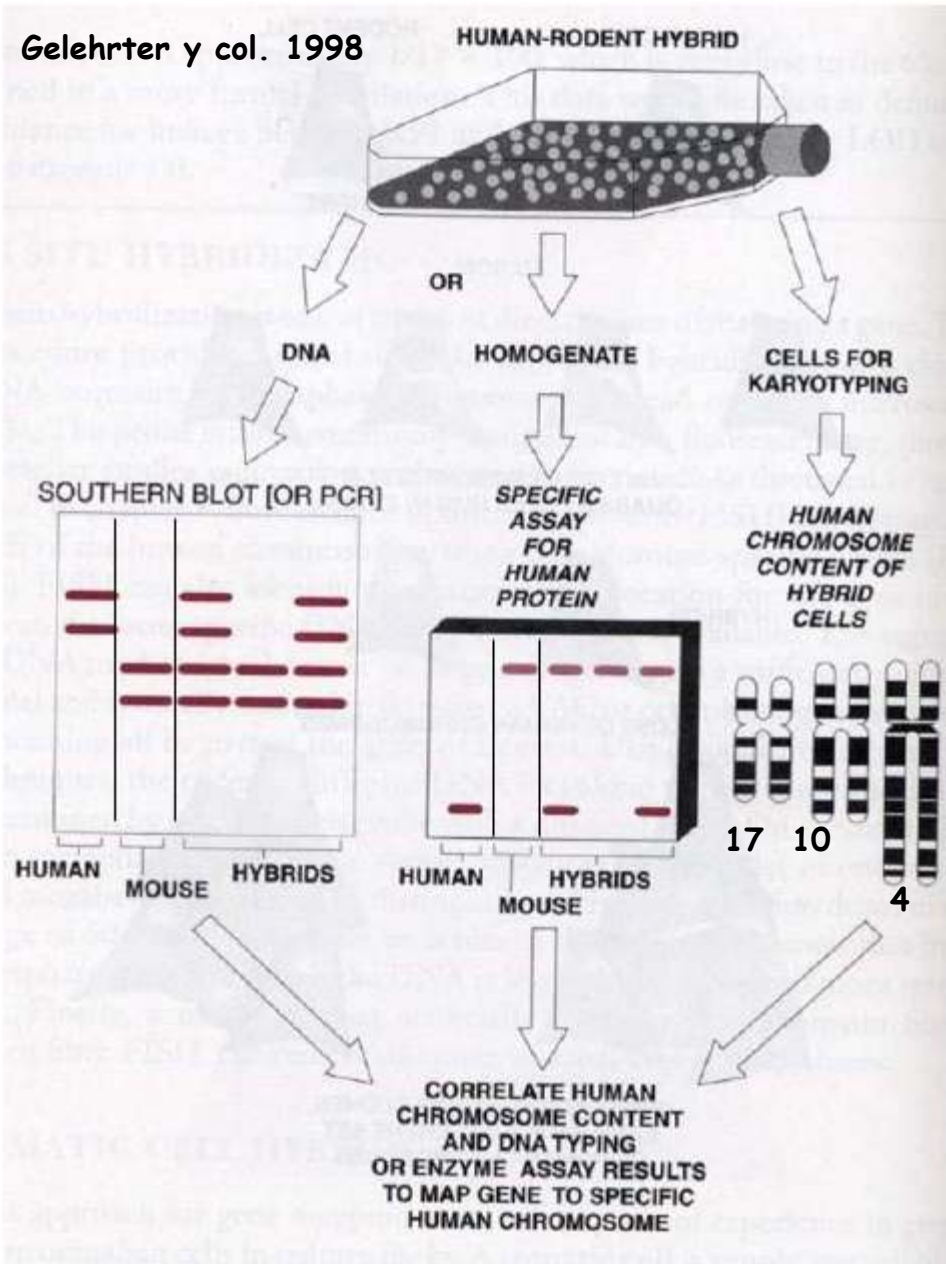
Los clones de las líneas celulares híbridas se utilizan para localizar genes humanos.





## LOCALIZACIÓN DE GENES EN LOS CLONES DE CÉLULAS SOMÁTICAS HÍBRIDAS

Gelehrter y col. 1998



En los clones híbridos se analiza directamente el ADN de los cromosomas humanos por Southern o por PCR y se estudian las enzimas humanas expresadas llegando a identificar el cromosoma en el que se encuentra el gen.

Por ejemplo, el oncogén MET humano se localizó en clones híbridos. Se expresaba sólo en los que incluían al cromosoma 7 humano (Nature 1984:311), señalando así su locus.

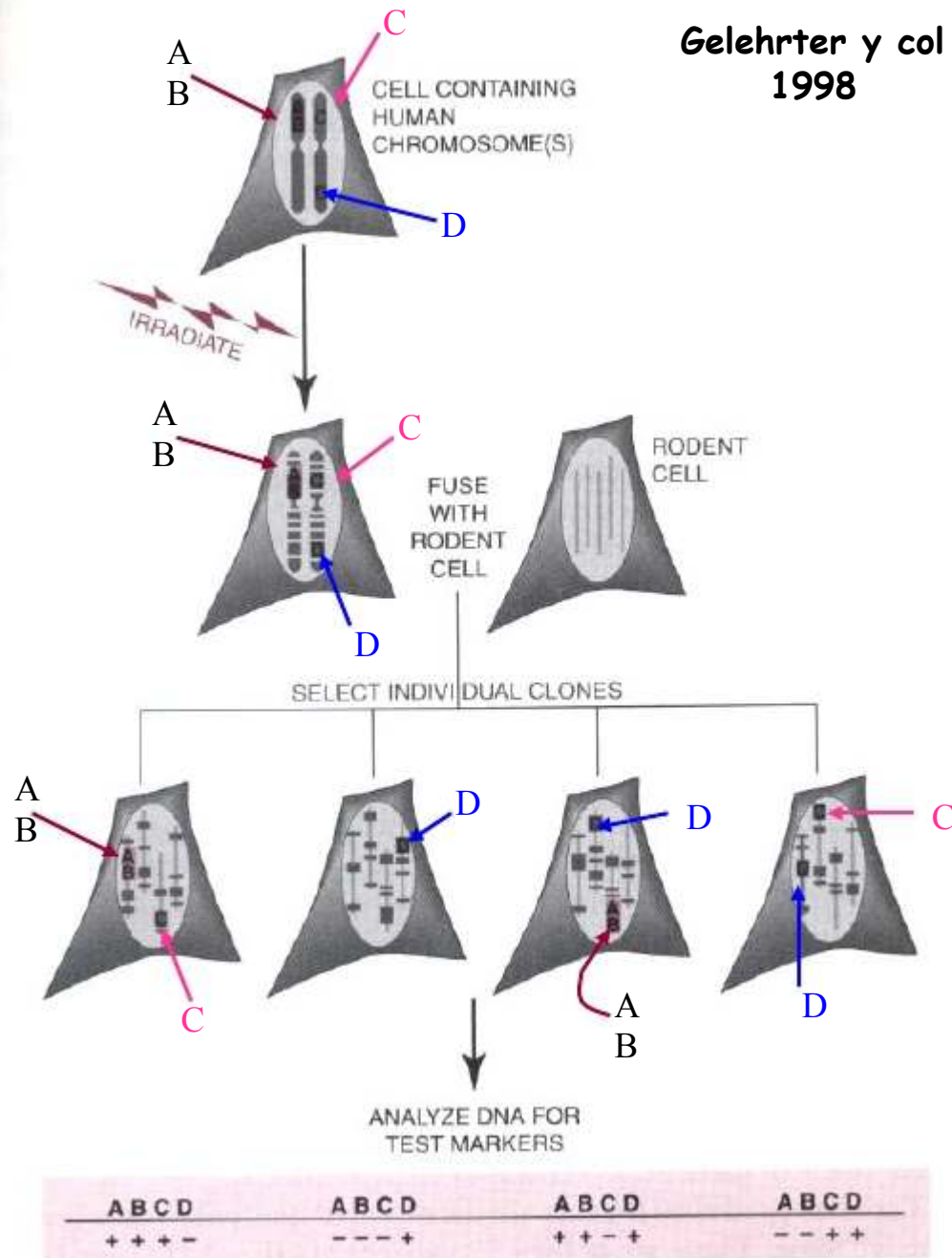
		HUMAN CHROMOSOMES																						Human MET			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X			
HYBRID	A																									-	
	B																										-
	C																										+
	D																										-
	E																										-
	F																										-
	G																										-
	H																										-
	I																										-
	J																										+
	K																										+
	L																										-

Gelehrter y col. 1998



## HÍBRIDOS DE CÉLULAS SOMÁTICAS IRRADIADAS

Gelehrter y col  
1998



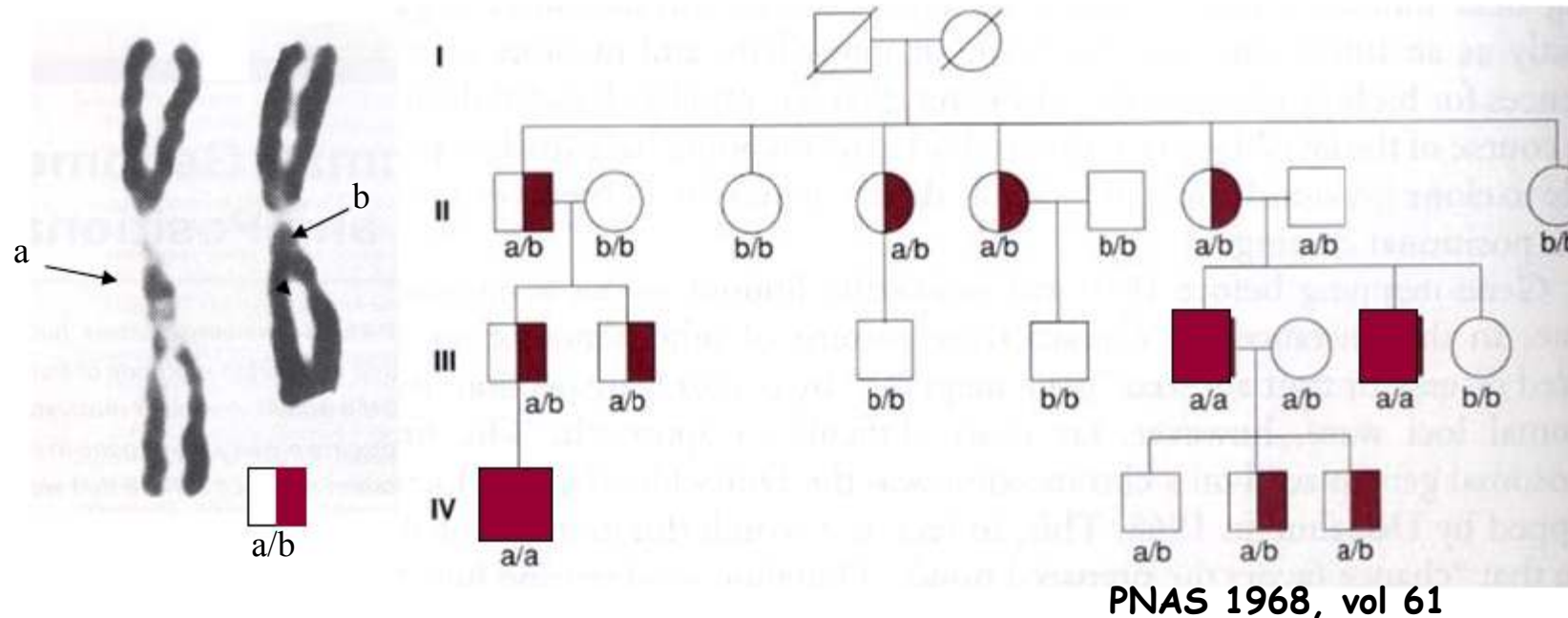
Sometiendo antes a las células humanas a rayos X, se obtienen clones híbridos con fragmentos de cromosomas humanos, que informan del orden de los genes de cada fragmento.

Cuanto más cerca estén dos loci más veces estarán (o faltarán) los dos en los clones híbridos.

La interpretación de estos datos es mucho más compleja que la de los híbridos no irradiados, requiriendo un potente análisis estadístico.

## LOCALIZACIÓN DE GENES POR CAMBIOS EN LA FORMA DE LOS CROMOSOMAS QUE SE VEN CON EL MICROSCOPIO ÓPTICO

Hasta 1950 sólo se localizaron genes del cromosoma X, gracias a su modo peculiar de herencia. El primer gen localizado en un autosoma fue el del grupo sanguíneo Duffy (1968, Donahue), que se encuentra en la región subcentromérica del cromosoma 1, con heterocromatina mucho menos condensada en presencia del alelo a (grupo Duffy).



Los familiares con este cromosoma 1 heteromórfico (variante no patológica fácil de distinguir al microscopio óptico) tenían todos ellos el grupo sanguíneo Duffy (a).

## OMIM Statistics for January 7, 2013

### Synopsis of the Human Gene Map: Number of Entries

	Autosomal	X-Linked	Y-Linked	Mitochondrial	Total
* Gene with known sequence	<a href="#">13370</a>	<a href="#">651</a>	<a href="#">48</a>	<a href="#">35</a>	<a href="#">14104</a>
+ Gene with known sequence and phenotype	<a href="#">124</a>	<a href="#">4</a>	0	<a href="#">2</a>	<a href="#">130</a>
# Phenotype description, molecular basis known	<a href="#">3371</a>	<a href="#">271</a>	<a href="#">4</a>	<a href="#">28</a>	<a href="#">3674</a>
% Mendelian phenotype or locus, molecular basis unknown	<a href="#">1627</a>	<a href="#">133</a>	<a href="#">5</a>	0	<a href="#">1765</a>
Other, mainly phenotypes with suspected mendelian basis	<a href="#">1765</a>	<a href="#">125</a>	<a href="#">2</a>	0	<a href="#">1892</a>
<b>Total</b>	<a href="#">20257</a>	<a href="#">1184</a>	<a href="#">59</a>	<a href="#">65</a>	<a href="#">21565</a>

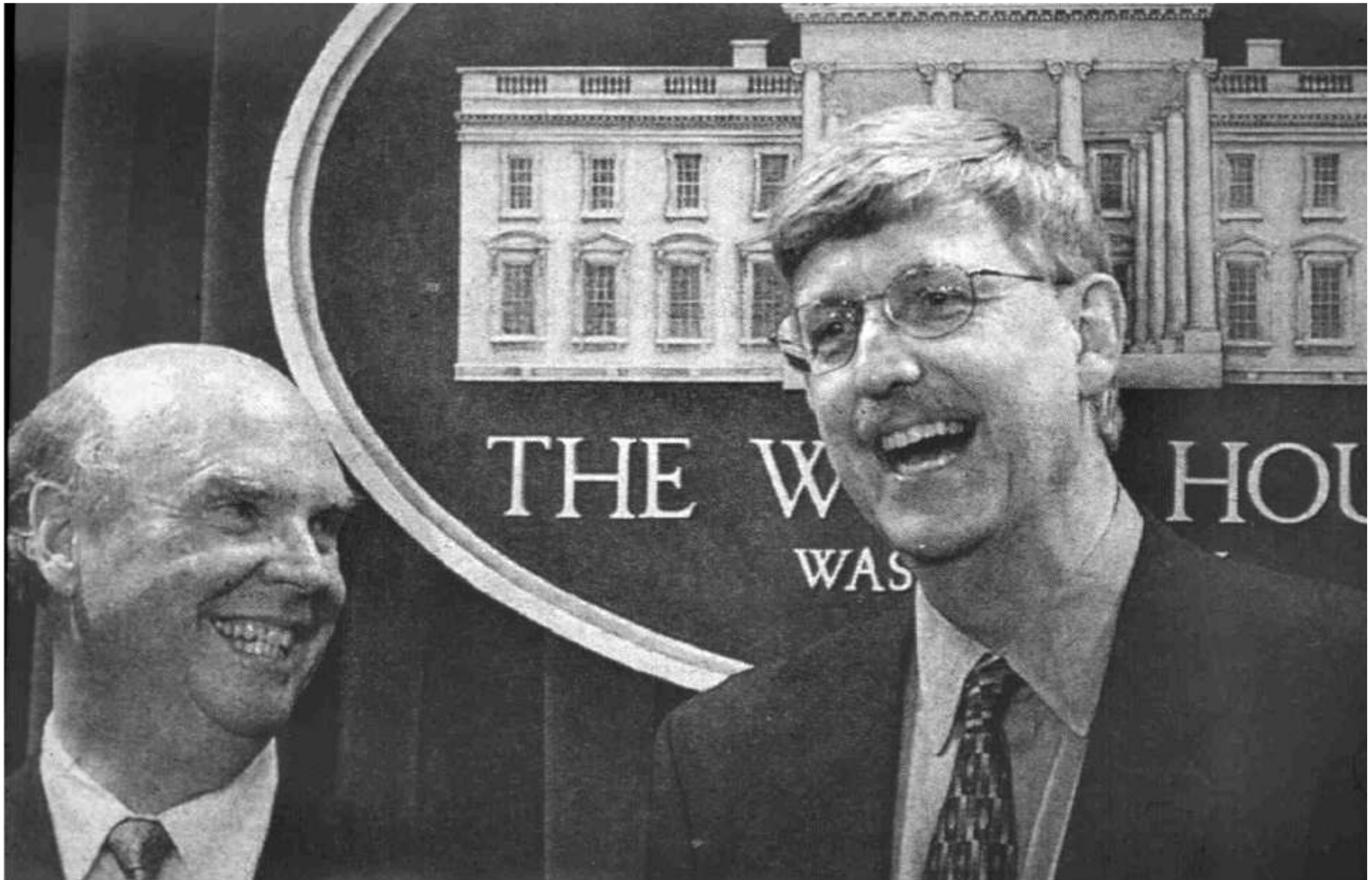
Chromosome	Loci
<a href="#">1</a>	1358
<a href="#">2</a>	870
<a href="#">3</a>	726
<a href="#">4</a>	537
<a href="#">5</a>	631
<a href="#">6</a>	815
<a href="#">7</a>	636
<a href="#">8</a>	486
Chr.	Loci
<a href="#">9</a>	518
<a href="#">10</a>	499
<a href="#">11</a>	841
<a href="#">12</a>	719
<a href="#">13</a>	255
<a href="#">14</a>	440
<a href="#">15</a>	414
<a href="#">16</a>	563
Chr.	Loci
<a href="#">17</a>	797
<a href="#">18</a>	200
<a href="#">19</a>	864
<a href="#">20</a>	349
<a href="#">21</a>	147
<a href="#">22</a>	332
<a href="#">X</a>	735
<a href="#">Y</a>	46
Total number of loci: <b>13778</b>	

**Tenemos cada vez más información sobre genes localizados en todos los cromosomas, pero aún ignoramos qué genes contribuyen a causar la mayoría de las enfermedades.**



# Proyecto Genoma Humano

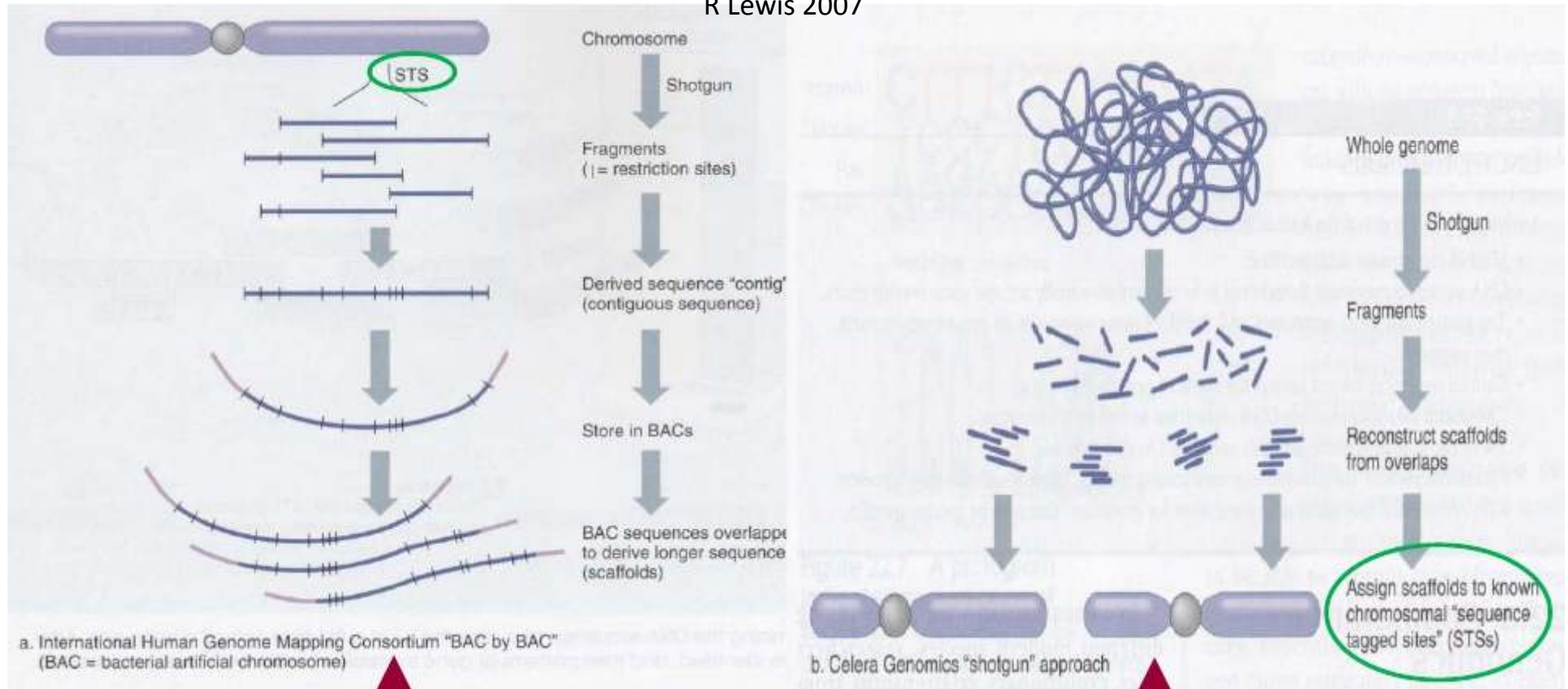
- 1990 Se inicia el Proyecto Genoma Humano.
- Se identifican numerosas **STS** (sequence-tagged site ) son secuencias cortas (200-500 pb) marcadoras, cada una de un lugar único del ADN -"etiquetas" características de la posición física que ocupan).
- Se decide ofrecer un banco de datos público y actualizado en la red.
- 1998 El Consorcio Público ofrece datos preliminares del 98% del genoma
- 2003 Se termina la secuenciación del ADN del Genoma Humano. Todas las secuencias que codifican proteínas están disponibles en micromatrices de ADN



**2003: Venter (Celera Genomics) y Collins (Consortio Internacional) presentan juntos el primer borrador de la secuencia completa del Genoma Humano, uno de los objetivos del Proyecto Genoma Humano (HUGO: Human Genome Organization).**

# DOS CAMINOS PARA SECUENCIAR EL GENOMA HUMANO

R Lewis 2007



## Consortio Internacional (público):

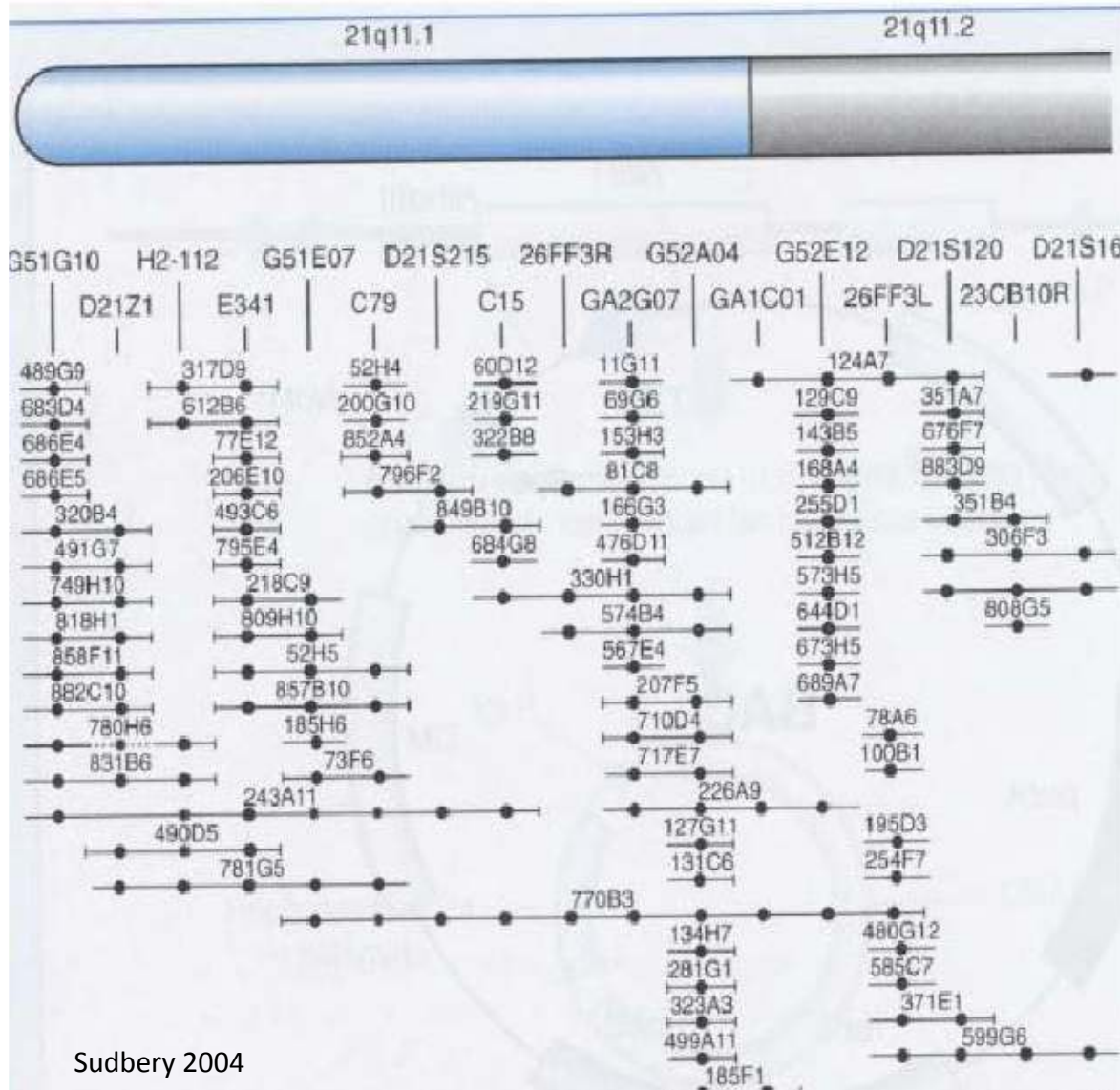
Utilizando STS de cada cromosoma reconstruyeron la secuencia de fragmentos largos por solapamiento de secuencias contiguas. La secuencia de los trozos pequeños se reconstruyó a partir de la de clones solapantes.

## Celera Genomic (iniciativa privada):

Partiendo de fragmentos pequeños de varias copias de un genoma entero, identificaron secuencias solapantes y las colocaron con ayuda de ordenadores. Se beneficiaron de muchos datos del Consortio público.



Los clones con secuencias solapadas (clones contiguos o cóntigos) se pueden ordenar ya que comparten **Sitios de Secuencias marcadoras (STS)**



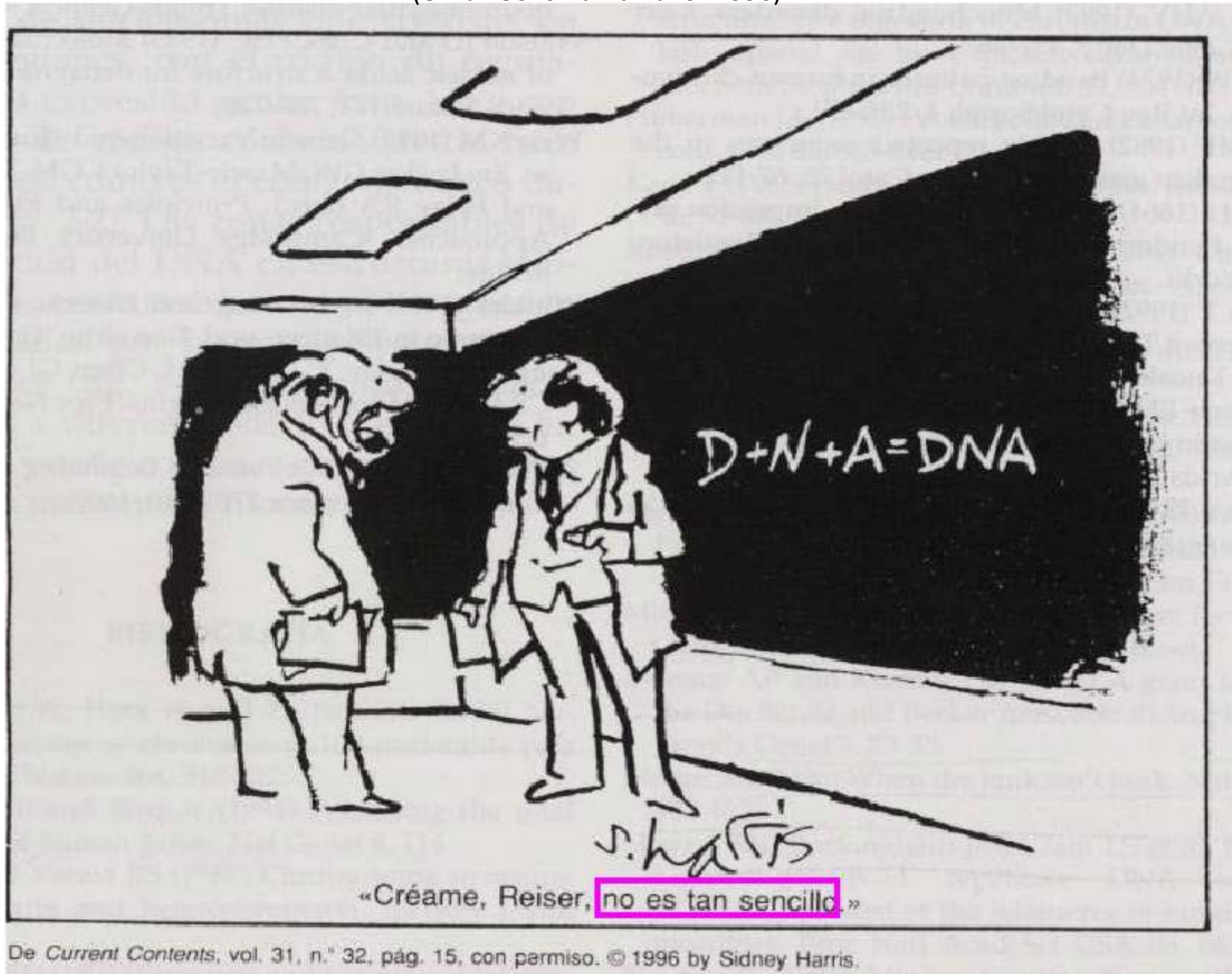
Sudbery 2004

Es posible la ordenación de la secuencia de un fragmento de un del cromosoma al ordenar sus cóntigos, gracias al solapamiento sucesivo de sus secuencias STS.

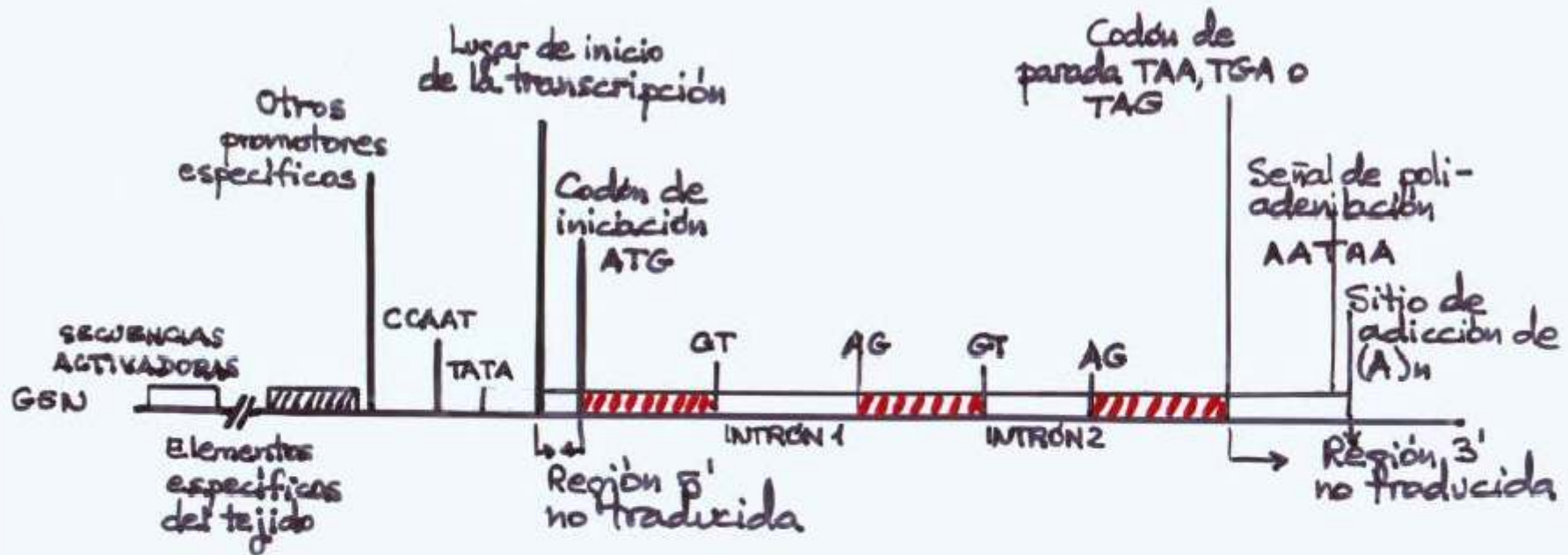
Los clones contiguos comparten **STS que están catalogados y disponibles en bases de datos muy completas (Genbank, en internet)**

## Pero ¿Qué significan las secuencias?...

(Oliva "Genoma Humano" 1996)



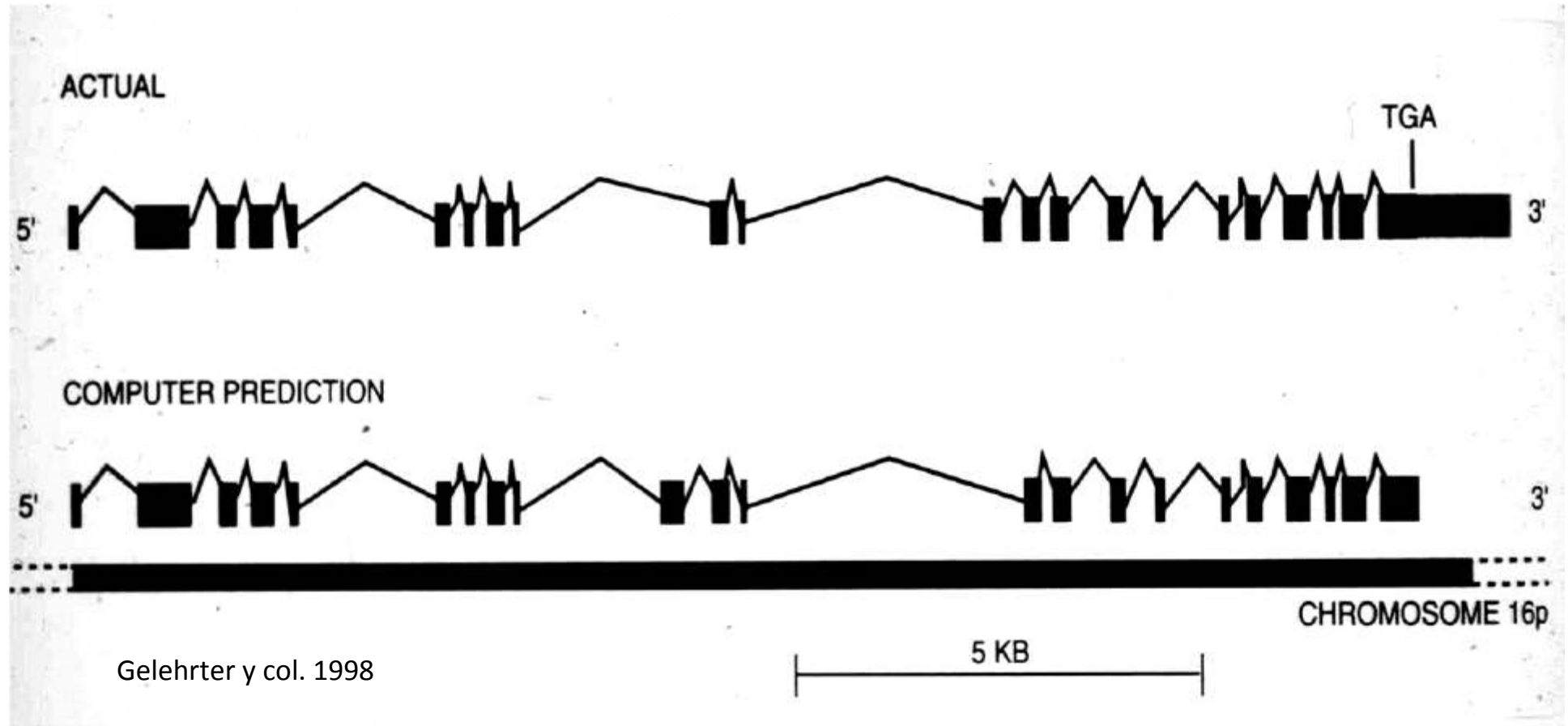
## ¿CÓMO SE RECONOCEN LAS SECUENCIAS UNA VEZ ORDENADAS?



Características generales de un gen  
(De Gelehrter y col 2da Ed. pag 62)

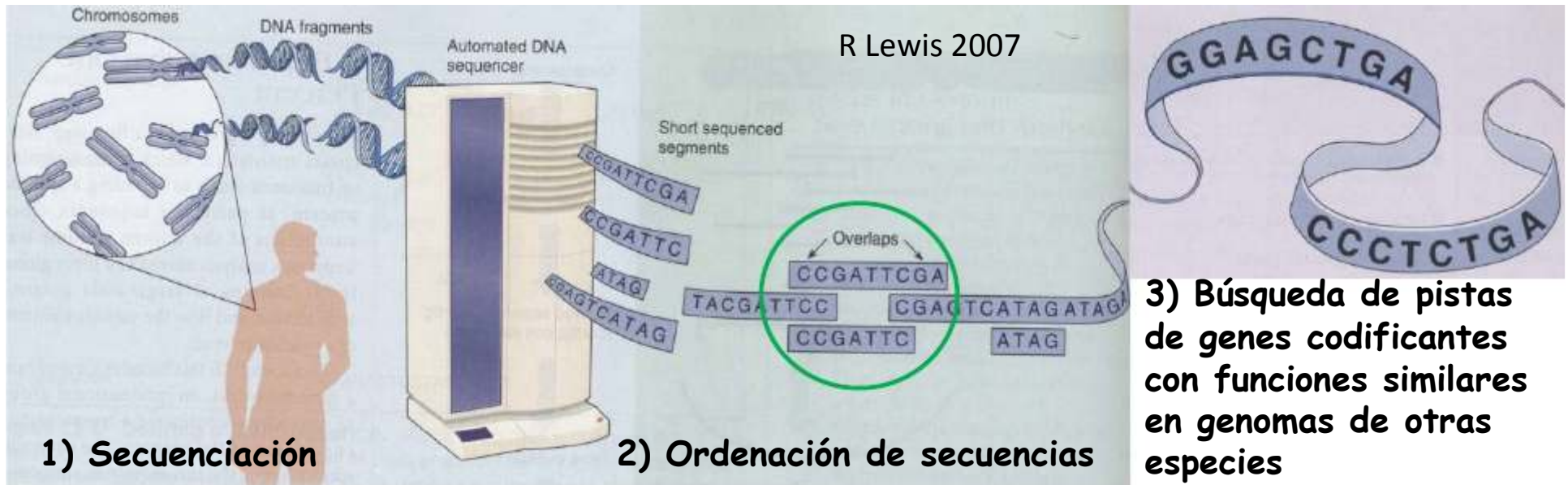


Potentes programas informáticos, como *GRAIL*, ayudan a predecir las partes de las secuencias que podrían ser codificantes

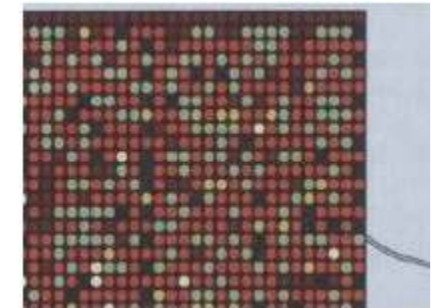


Fragmento del gen cuya mutación AD da lugar a poliquistosis renal, en el brazo corto del cromosoma 16

## LA SECUENCIACIÓN DEL GENOMA ES SÓLO EL PRIMER PASO



Una vez secuenciado el genoma humano, junto con el de otras especies, **se trata de identificar genes que codifiquen no sólo proteínas, sino también ARNs**, cuya función reguladora es de crucial importancia. El último y más apasionante de nuestros objetivos es descubrir los patrones de expresión de los genes en cada tipo de célula diferenciada y durante el desarrollo.



4) Perfiles de expresión génica y proteómica con micromatrices de ADNc a partir de ARNm de células de distintos tejidos

## Proyecto Genoma Humano: secuencia encontrada en 2003

CROMOSOMA 1



Oliva y col '04 Fig 5.3

CROMOSOMA 21

La secuencia completa de las 3.000 millones de bases ocuparía 200 libros de 1000 páginas.

Acceso gratuito en:

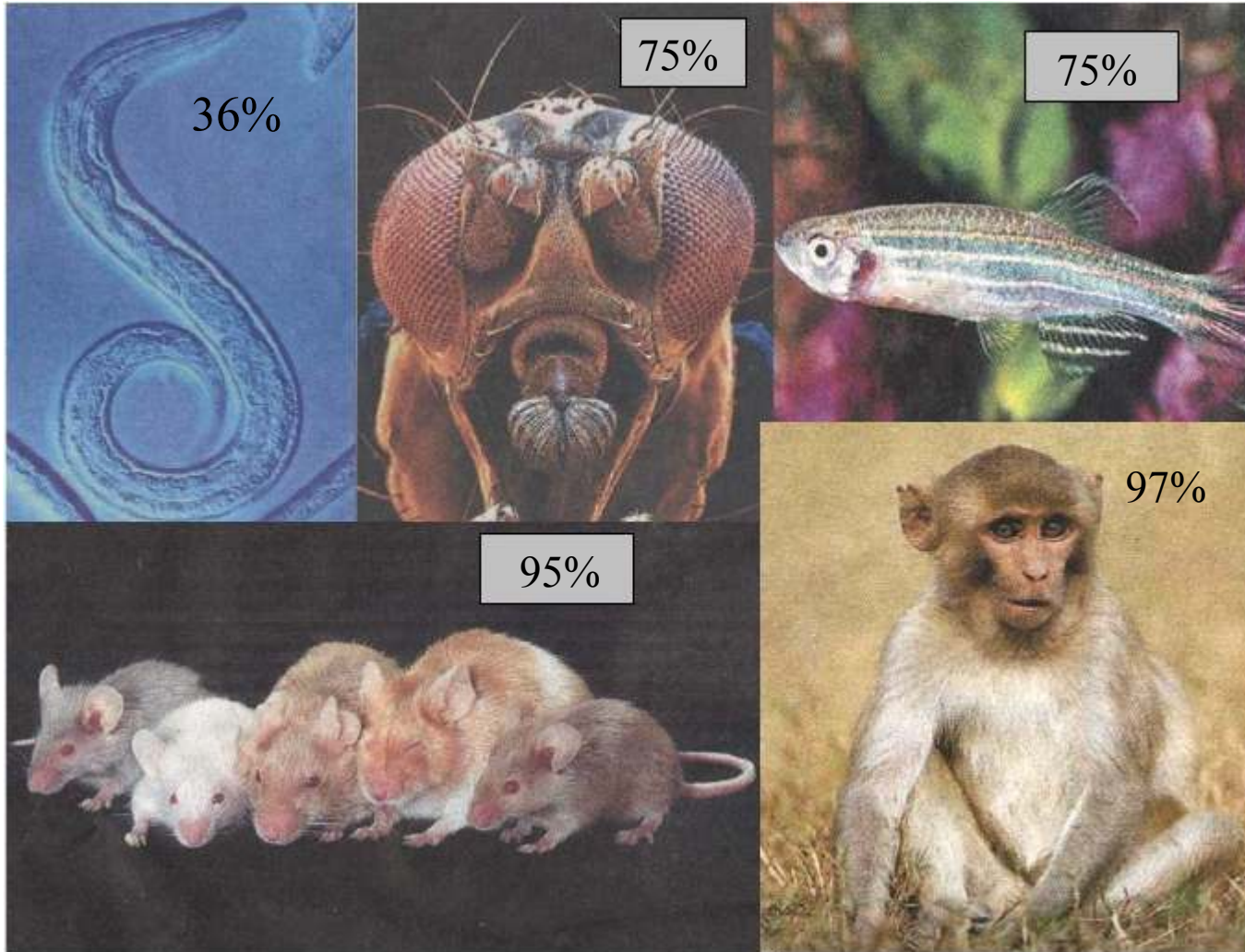
<http://genome.ucsc.edu>

- \* Se identificaron unos 35.000 genes codificantes
- \* Aproximadamente la mitad de las secuencias son repetitivas
- \* Muchos genes humanos proceden directamente de procariotas a través de transferencia evolutiva horizontal
- \* El concepto de "raza humana" carece de contenido científico, ya que las diferencias genómicas entre personas de distintas "razas" no lo sustentan

•Una proporción muy elevada de los genes se dedican a regular la propia expresión génica, y a controlar mecanismos esenciales para todos los seres vivos, como la división celular.



**La secuenciación del genoma de otras especies demuestra un elevado porcentaje de homología con nuestros genes**



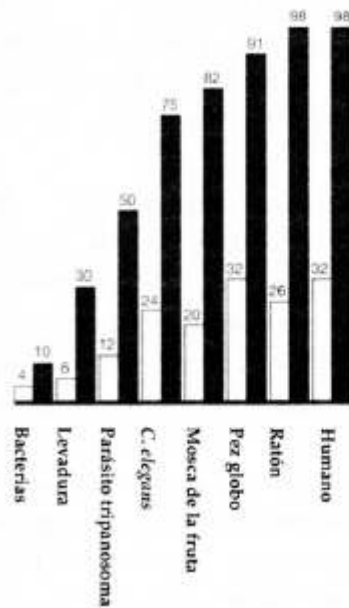
**Los genes que codifican proteínas en otras especies con secuencia homóloga a la de los nuestros, cumplen con frecuencia funciones muy similares en mecanismos básicos para todos.**

Más que los genes codificantes de proteínas, lo que nos diferencia es la **regulación génica**

La diferencia fundamental entre las especies radica en el **grado de complejidad de su regulación génica, mayor a medida que se asciende en la escala filogenética**. De modo que en las especies con desarrollo más complejo, la misma secuencia génica puede transcribirse en distintos ARNm que conducen a distintas isoformas de una misma proteína.

Por tanto, **en eucariotas el número de genes no se correlaciona con el tamaño del genoma ni con la complejidad del organismo.**

□ Parte del genoma que codifica proteínas, en millones de pares de bases  
■ Porcentaje del genoma que no codifica proteínas



De hecho, un % muy elevado del genoma de las especies con mayor complejidad no codifica para proteínas (98% del genoma humano).

El ARN no codificante participa en **cambios epigenéticos** fundamentales, como la inactivación del cromosoma X y el sello genómico.



Los **cambios epigenéticos** afectan a la estructura de la cromatina, que a su vez controla tanto la replicación como la transcripción del ADN. Cada vez son más los datos que demuestran la **importancia de dichos cambios en el funcionamiento del genoma y su contribución en enfermedades como el cáncer**. El CNIO colabora con un consorcio internacional en el Proyecto Epigenoma Humano, que se propone establecer modificaciones del genoma de todos los tipos celulares en personas sanas y enfermas.

**EL Proyecto ENCODE trata de identificar los elementos funcionales de la secuencia del Genoma Humano**

El estudio en profundidad del 1% del genoma humano llevado a cabo entre 2003 y 2007 en la primera fase del proyecto reveló que:

- \* **Se transcriben numerosas secuencias que antes se consideraban silenciosas.**  
(Hay más promotores de los inicialmente previstos)
- \* La estructura de la cromatina determina tanto la replicación como la transcripción del ADN (**control epigenético**).
- \* **El 5% del genoma humano está muy conservado en todos los mamíferos.**
- \* **Las partes funcionales del genoma pueden variar entre las especies, pudiendo representar reservas propias para la evolución de cada una de ellas.**



**El proyecto HapMap (Mapa de Haplotipos) investiga la variabilidad alélica funcional que es clave para entender el genoma humano**

**Es uno de los proyectos fundamentales derivados de la secuenciación del genoma humano (GH)**

**Su objetivo es la realización de un mapa de haplotipos del GH; describirá las variaciones habituales de la secuencia del ADN presentes en las personas estudiadas de diversas poblaciones**

**Constituye una fuente muy valiosa de información para conocer mejor la predisposición genética que subyace a muchas enfermedades.**

**www.hapmap.org**

**Tabla 9-4** Poblaciones estudiadas en el proyecto HapMap internacional

Ciudad/país	Ascendencia	Muestras analizadas
Ibadán (Nigeria)	Yoruba	30 tríos (adulto y ambos progenitores)
Tokio (Japón)	Japonesa	45 individuos no relacionados
Pekín (China)	China	45 individuos no relacionados
Estados Unidos	Europa del norte y del oeste	30 tríos (adulto y ambos progenitores)

**Emery, tabla 9-4**

## HACIA LOS TRATAMIENTOS PERSONALIZADOS



Klug y col. 2006

El "Proyecto de los Mil Genomas" se propuso analizar variantes del ADN de 1000 voluntarios anónimos de múltiples etnias, con herramientas como este GenChip de primera generación con 50.000 genes candidatos.

El objetivo del "Proyecto Genoma Personal" (PGP) de la Universidad de Harvard es la secuenciación de 100.000 genomas de voluntarios.

Obama propuso en Enero de 2015 que se inviertan unos 200 millones de \$ para crear bases de datos con información genética de hasta 1 millón de voluntarios de EEUU. Se espera que el análisis de los datos recogidos permita ajustar tratamientos para enfermedades como el cáncer, de manera que el organismo responda óptimamente ante ellos reduciéndose al mínimo sus efectos secundarios.

[http://elpais.com/elpais/2015/01/30/ciencia/1422650626\\_915373.html](http://elpais.com/elpais/2015/01/30/ciencia/1422650626_915373.html)



# NIH takes next steps in Obama's Precision Medicine plan

CHICAGO | BY JULIE STEENHUYSEN



U.S. President Barack Obama talks about investments to improve health and treat disease through precision medicine while in the East Room of the White House in Washington January 30, 2015.  
REUTERS/LARRY DOWNING

The National Institutes of Health on Thursday approved a blueprint for U.S. President Barack Obama's Precision Medicine Initiative and named an NIH insider as interim director of the project, which aims to enroll 1 million volunteers in the next three to four years.

NIH Director Francis Collins said he would "act immediately" on recommendations delivered to him on Thursday by a working group created to develop a framework for the study.

Among the goals of the study are to develop better estimates on individuals' risk for developing disease. These would be drawn from a range of metrics, looking not only at genetic factors but also the role of environmental exposures and their impact on genetic predispositions.

Collins estimates that based on current prices, it would cost up to \$2 billion to do whole genome sequencing on 1 million individuals now, compared with \$30 to \$40 each for genotyping.

