

FELIPE GARCÍA V. - MARTHA C. DOMÍNGUEZ

# Las Hélices Paralelas

Una Visión Crítica  
de la Era Genómica y Postgenómica

Colección Salud



Programa  Editorial

FELIPE GARCÍA VALLEJO  
MARTHA CECILIA DOMÍNGUEZ

# Las Hélices Paralelas

Una Visión Crítica  
de la Era Genómica y Postgenómica



Colección Salud

El proyecto del Genoma Humano es, sin duda, en el área de la biología, la mayor empresa de investigación que a través de los tiempos ha emprendido el hombre. Los dos autores han entendido que esta es una poderosísima herramienta para comprender los diferentes fenómenos biológicos y los secretos de la evolución de las especies. Este libro es un intento brillante por actualizar toda la discusión que se adelanta en torno al Genoma Humano y las perspectivas que ella abre.



Universidad  
del Valle

Programa  Editorial

FELIPE GARCÍA VALLEJO  
MARTHA CECILIA DOMÍNGUEZ

# Las Hélices Paralelas

Una Visión Crítica  
de la Era Genómica y Postgenómica



Colección Salud



**Universidad del Valle**

**Programa Editorial**

Título: Las hélices paralelas: una Visión Crítica de la Era Genómica y Postgenómica

Autores: Felipe García Vallejo, Martha Cecilia Domínguez

ISBN: 978-958-670-251-5

ISBN-PDF: 978-958-5164-63-5

DOI: 10.25100/peu.537

Colección: Salud

**Primera Edición Impresa abril 2003**

Rector de la Universidad del Valle: Édgar Varela Barrios

Vicerrector de Investigaciones: Héctor Cadavid Ramírez

Director del Programa Editorial: Omar J. Díaz Saldaña

© Universidad del Valle

© Felipe García Vallejo, Martha Cecilia Domínguez

Diseño de carátula: Henry Naranjo

Este libro, o parte de él, no puede ser reproducido por ningún medio sin autorización escrita de la Universidad del Valle.

El contenido de esta obra corresponde al derecho de expresión del autor y no compromete el pensamiento institucional de la Universidad del Valle, ni genera responsabilidad frente a terceros.

El autor es el responsable del respeto a los derechos de autor y del material contenido en la publicación, razón por la cual la Universidad no puede asumir ninguna responsabilidad en caso de omisiones o errores.

Cali, Colombia, diciembre de 2020

# TABLA DE CONTENIDO

Prólogo.....	9
Prefacio .....	11
<b>Capítulo 1</b>	
Dos “ <i>ISMOS</i> ” frente al genoma humano .....	15
<b>Capítulo 2</b>	
Información, genoma y sociedad .....	21
<b>Capítulo 3</b>	
Los orígenes del proyecto: De la biología molecular al genoma.....	29
<b>Capítulo 4</b>	
La agenda de trabajo hacia el proyecto genoma humano .....	41
<b>Capítulo 5</b>	
Aprendiendo a entender el genoma .....	51
<b>Capítulo 6</b>	
Como secuenciar y cartografiar el genoma .....	63
<b>Capítulo 7</b>	
Los otros proyectos genoma .....	73
<b>Capítulo 8</b>	
¿Qué sigue en la era postgenómica? .....	87

<b>Capítulo 9</b>	
Hacia una patología genómica .....	95
<b>Capítulo 10</b>	
La genómica funcional: La sinfonía de los genes. ....	111
<b>Capítulo 11</b>	
La terapia génica. ....	121
<b>Capítulo 12</b>	
La farmacogenómica .....	131
<b>Capítulo 13</b>	
El proyecto proteoma humano .....	139
<b>Capítulo 14</b>	
El negocio de la genómica .....	143
<b>Capítulo 15</b>	
El genoma humano frente al futuro de la humanidad .....	149
<b>Glosario</b> .....	159
<b>Bibliografía</b> .....	173
<b>Motores de búsqueda de páginas en la Web</b> .....	181



Universidad  
del Valle

---

**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

# PRÓLOGO

El proyecto del Genoma Humano es, sin duda, en el área de la biología, la mayor empresa de investigación que a través de los tiempos ha emprendido el hombre. Su verdadero alcance sólo lo podremos constatar en el curso de algunos años, una vez se hayan perfeccionado técnicas que permitan una amplia aplicación y se terminen de entender algunos fenómenos que hasta ahora han sido elusivos a la curiosidad y al trajinar de los científicos.

Felipe García, uno de los científicos más versados en el campo de la Biología Molecular en Colombia, desde hace años entendió que ésta era una poderosísima herramienta para comprender los diferentes fenómenos biológicos y los secretos de la evolución de las especies. Dentro de ese espíritu inquisitivo y crítico, ha incursionado en el estudio del genoma de los virus, bacterias, aves, delfines y algunos genes humanos. Y lo que para algunos podría haber parecido como dispersión en lo científico, tiene una natural y necesaria convergencia, pues tal como magistralmente él lo describe en este libro, todos los seres vivos están hechos de los mismos materiales, que gracias a la evolución han dado origen a organismos y estructuras cada vez más complejos y dotados de nuevas funciones y capacidades.

En todas esas empresas intelectuales, Felipe ha tenido en Martha Domínguez a su cómplice, compañera de trabajo, interlocutor y crítico más severo y como fruto de ese compartir esforzado y riguroso, los autores presentan este libro, “LAS HÉLICES PARALELAS”, centrado en el planteamiento de que en la vida todo está interrelacionado, por lo cual a través de sus páginas nos llevan de la mano desde los comienzos de la Genética hasta las disquisiciones y motivaciones que dieron origen al proyecto Genoma Humano y nos muestran, también, cómo este proyecto no nació de la noche a la mañana sino que es un continuo en el pensamiento y en la investigación que va concatenando hallazgos, en forma tal que los conocimientos obtenidos en guisantes y moscas de la fruta y luego en virus, levaduras y mamíferos han ido dando respuestas a la forma en que está codificado el secreto de la vida desde los especímenes más pequeños hasta llegar a los mamíferos y el hombre.

Describen luego la forma en que se planea lo que a todas luces parecía fantástico por su magnitud, costo y duración y narran muy amablemente las motivaciones, contradicciones e

intereses, a veces solidarios y en ocasiones encontrados, entre quienes participaron en dicha empresa, evidenciando que ni aún la ciencia está exenta de rivalidades, pero que en este caso, afortunadamente, contribuyeron a acelerar y a abaratar el costo general del Proyecto.

El capítulo cinco nos enseñan cómo estudiar los genes, los cromosomas y el genoma. Partiendo de esos presupuestos, los autores describen el porqué ha sido necesario ir secuenciando paralelamente con el genoma humano, otros genomas de animales y plantas, desde los más simples hasta los más complejos, con el fin de entender los procesos evolutivos que han llevado a la aparición del *homo sapiens* y al origen y dispersión de la enfermedad genética.

En los últimos capítulos el lector podrá darse cuenta de lo importante que son los resultados del Proyecto del Genoma Humano para un buen número de enfermedades genéticas cuyo diagnóstico y en algunos casos sus manejos se han beneficiado del conocimiento adquirido a través del estudio de los genes, así como un detallado análisis de las cosas que aún están por comprender y dominar para poder poner en práctica todas las enseñanzas que de él se derivan.

En este libro, Felipe García y Martha Domínguez, quienes han recorrido juntos el camino de la vida y de la ciencia, y cuya unión de genes y de amor ha dado vida a dos hijos, han sumado una vez más su inteligencia, su capacidad de trabajo y su dedicación y entusiasmo por la ciencia para adentrarnos en el fascinante mundo de la genómica en forma amena, coherente, rigurosa en lo científico, pero didáctica y ágil por lo fácil de comprender para el profano.

Gracias a lo anterior, los autores nos transmiten y contagian su fascinación y pasión por el tema, no sólo a versados en la Biología, sino también a todos aquellos que quieran entender desde el ángulo de las ciencias sociales los códigos en que está escrita la vida, o a quienes conociendo un poco más del tema quieran aproximarse a los secretos del genoma y sus potencialidades en el conocimiento, tratamiento de las enfermedades y las implicaciones que estos hallazgos, descubrimientos tiene en nuestra vida y en el destino de la humanidad.

Los invito con entusiasmo a degustar, como lo he hecho yo, estas páginas motivadoras y preñadas de fervor y de conocimientos que por la importancia del tema y el tratamiento que le dan, abren las puertas al saber, pero también a la expectación y la esperanza, útil para todos.

*Luis Alejandro Barrera Ph.D.*

*Director Instituto Errores Innatos del Metabolismo.*

*Pontificia Universidad Javeriana. Santa Fe de Bogotá*

# PREFACIO

Desde hace más de una década, un grupo de hombres ejecuta un proyecto de investigación diferente a cualquier otro intentado hasta el momento. Tratar de secuenciar tres mil millones de nucleótidos (componentes básicos de la molécula de ADN) y localizar e interpretar los casi 40.000 paquetes de información biológica representada por los genes. Este proyecto, en última instancia, cambiará no sólo nuestras vidas sino la concepción de nosotros mismos. Además tendrá un tremendo impacto en la medicina y en nuestro conocimiento de la biología humana. Su culminación llevará, en un futuro no muy lejano, a poder controlar muchas de las enfermedades que afligen al hombre y nos conducirá, entre otras cosas, a intervenir sobre el envejecimiento y la muerte. Sin embargo, sus efectos ya plantean serios dilemas éticos, legales, filosóficos, humanísticos y sociales.

Esta moderna caja de Pandora se conoce como el Proyecto Genoma Humano. Los que trabajamos en el tema abreviamos su nombre mediante tres letras: PGH; tres letras que encierran la esperanza para algunos, el conocimiento de la dura realidad para otros y la incertidumbre para todos. Sus objetivos finales son secuenciar, localizar y caracterizar todos y cada uno de los genes que conforman nuestra especie. Una vez finalizado, dispondremos de una fuente de conocimientos prácticamente inagotable para tener una visión diferente del hombre como individuo y como especie, nos estaremos planteando un nuevo paradigma cognitivo, el hombre genómico. Sabremos cuáles de nuestras características son heredables y cuáles dependen del ambiente. Podremos conocer también los defectos genéticos que portamos de forma latente y que podríamos legar a nuestros hijos. También podremos llegar a conocer, a veces con varias décadas de adelanto, si nosotros mismos vamos a padecer alguna enfermedad cuyo proceso aún no ha comenzado.

En las siguientes páginas trataremos de aproximarlos al PGH y sus consecuencias posteriores, la postgenómica; algunas veces, habrá grandes rodeos partiendo de terreno conocido, intentando introducir aquellos descubrimientos y hechos que resultaron cruciales en una cadena de acontecimientos que condujo, prácticamente, de forma inevitable, al Proyecto Genoma Humano. Trataremos de narrar cómo la biología molecular alcanzó un nivel tal



que hizo posible plantear una iniciativa tan ambiciosa. Otras veces, haremos un enfoque de manera más directa, proporcionando datos concretos sobre metas que ya han sido alcanzadas.

También es el cometido de este trabajo tratar de enumerar los alcances y consecuencias que han tenido y tendrán, inexorablemente, el PGH y sus derivaciones para nuestra sociedad. Puede parecer alarmante y el último de nuestros deseos sería crear en el lector una sensación de preocupación, de que existe una conspiración contra la humanidad; sin embargo, es importante reflexionar que, mientras millones de personas permanecen ajenas e inconscientes, grupos de científicos trabajan en la sombra para cambiar sus vidas, lo sepan o no, lo quieran o no. Esto no es nada nuevo y así es como ha funcionado siempre la ciencia. Sin embargo, en los últimos años la conciencia ciudadana parece haberse despertado en los países desarrollados y grupos de no especializados se han interesado por el trabajo que las instituciones científicas llevan a cabo, participando de manera activa, para apoyar en algunos casos, censurar algunas veces e impedir, en contadas ocasiones, potenciales excesos. En alguna que otra ocasión, los grupos de investigadores se han autoimpuesto moratorias para no realizar determinadas investigaciones. De cualquier manera se nota una tendencia a que la conciencia colectiva de la sociedad opine sobre los trabajos de la ciencia y sus repercusiones en nuestras sociedades.

A pesar del estricto control que toda la sociedad y la comunidad científica en particular puedan ejercer sobre el Proyecto Genoma Humano, la mayoría de la gente que oye hablar del mismo por primera vez no puede evitar, como su primera impresión, un sentimiento de inquietud, un primer acto reflejo, que casi siempre es un escalofrío que recorre el cuerpo.

Todo conocimiento tiene un aspecto ético-moral y otro oscuro. Es la tarea de la humanidad decidir que el lado que contemplemos sea siempre el ético y moral. En algunas ocasiones, el hombre ha conocido el reverso oscuro de la ciencia; como los miles de personas que estaban en Hiroshima y Nagasaki en 1945 lo saben muy bien, también los descendientes de los judíos alemanes que los nazis utilizaron, entre otras cosas, además de conejillos de indias, como combustible para calderas, los que pierden la vida por una bala salida del arma de un delincuente y los que mueren cada semana en accidentes de automóvil en nuestras carreteras.

El Proyecto Genoma Humano cambiará nuestras vidas, siendo casi seguro que el cambio es para mejorar la calidad de vida de una mayoría de la población, en diversos grados. El PGH tiene muchas connotaciones positivas; la mayoría de sus razones, sus justificaciones y sus consecuencias se englobarían dentro de la categoría de «buenas», al menos para la generalidad de las personas. Sin embargo, no hay que olvidar que también tiene, y nadie puede negarlo, posibilidades de ser empleado en favor de unos pocos y en perjuicio de la mayoría entre los que se incluyen las naciones menos desarrolladas de nuestro planeta. Estas son las connotaciones negativas del PGH, los horrores encerrados en el escaparate de los genes. Creemos que vale la pena el esfuerzo. No podemos renunciar a unos beneficios jamás soñados por el ser humano más optimista sólo porque existan unas posibilidades mínimas de hacer un mal uso de los conocimientos que podremos obtener.

Nuestro trabajo ha sido recopilar y comentar de manera crítica, con base en nuestra experiencia como científicos y biólogos moleculares, los diferentes aspectos del proyecto genoma humano y sus implicaciones en la era postgenómica. Aunque necesariamente mucho del lenguaje utilizado implica un cierto nivel de comprensión biológica, nos hemos dado a la tarea de ponerlo, en la medida de lo posible, para que todo el mundo pueda entender tanto la parte metodológica como la conceptual. Sin embargo, en algunos capítulos, necesariamente debemos introducir términos y conceptos especializados que intentaremos explicar de la manera más sencilla posible. Hemos ido recopilando un volumen grande de bibliografía escrita sobre la genética y sobre el PGH, tanto en papel como en Internet. En realidad, nuestro trabajo se ha limitado a tratar de dar una visión general y crítica del rompecabezas molecular de nuestra especie.

Después de un arduo proceso de pensamiento le hemos dado al libro el título de “Las Hélices Paralelas” puesto que la información contenida en la doble hélice del ADN implica dos grandes paradigmas que la humanidad debe replantear en este siglo XXI. Uno de ellos, el científico o positivista, que ha comenzado a transformar el conocimiento completo del hombre biológico, mientras que el otro, el paradigma postmodernista desde el enfoque de la realidad cultural y humanística; de acuerdo con nuestra percepción, estos dos enfoques son actualmente paralelos, aunque ambos surgen de la misma esencia de la doble hélice.

Esperamos cumplir los objetivos que nos hemos trazado y que usted, lector, pueda tener después de leer este libro, una percepción más integral de lo que representará para la humanidad el haber abierto el escaparate de los genes contenidos en nuestro genoma o destapado la moderna caja de Pandora, o como desee llamarlo.

**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

## **DOS “ISMOS” FRENTE AL GENOMA HUMANO**

*“El hombre blanco dibujó un pequeño círculo en la arena y dijo al piel roja: “Esto es lo que sabe el indio”; entonces trazando otro círculo mas grande que rodeaba al anterior exclamó: “Esto es lo que sabe el hombre blanco”. El indio tomó la rama, trazó rapidamente un inmenso anillo encerrando los dos círculos y dijo: “Aquí es donde el hombre blanco y el piel roja no saben nada”.*

**CARL SANDBURG**  
**(“The people, yes”) La gente, si 1936.**

### **LA EVOLUCIÓN DEL SIMPLISMO INTERPRETATIVO**

En el siglo III antes de Cristo, Arquímedes, descubrió la ley fundamental de la estática, o ley de la palanca. Es una de las primeras leyes físicas, expresables en términos matemáticos conocidas por el hombre. El sabio de Estagira aprovechó su descubrimiento para impresionar al rey Hierón de Siracusa, del cual era pariente y consejero; moviendo, él solo, sin esfuerzo aparente, mediante un sistema de palancas y poleas, un pesado mercante de tres palos repleto de carga que había sido transportado hasta tierra por un grupo de hombres, no sin gran esfuerzo. Eufórico por la sencillez y potencia de su descubrimiento, se cuenta que dijo, su famosa frase: «*Dadme un punto de apoyo y moveré el Mundo*». Sin duda, pensaba que la física estaba muy desarrollada por aquel entonces y en cuanto se descubrieran dos o tres leyes más, se tendría un conocimiento absoluto del funcionamiento del Universo. Pasaron dos mil años y la física encontró muchas y muy variadas leyes para explicar todo

tipo de fenómenos. No todas las leyes físicas eran tan sencillas como la ley de la palanca de Arquímedes; sin embargo, la observación y los desarrollos matemáticos posibilitaron un conocimiento bastante útil del sistema del Mundo. Fue el físico y matemático inglés Isaac Newton que, prácticamente, inventó toda la física en unos tres años de trabajo tras los cuáles dedicó su tiempo a ajusticiar falsificadores y malversadores de fondos, como Director de la Real Casa de la Moneda. De cualquier modo, las leyes de Newton parecían bastarse por sí solas para explicar la casi totalidad de los fenómenos físicos. El ambiente de euforia entre la comunidad científica era general. El Marqués Pierre Simón de Laplace, otro físico eminente, quien fue uno de los máximos exponentes de esta euforia; hacia finales del siglo XVIII, afirmó que si pudiéramos conocer las condiciones iniciales, aplicando sencillamente las leyes de Newton, podríamos calcular perfectamente todos los estados pasados, presentes y futuros del Universo. Así, el conocimiento del mundo parecía total y absoluto. ¿No parece esta afirmación tan presuntuosa como la de Arquímedes: «*Dadme unas condiciones iniciales y os diré el futuro del Universo*»?

### **SE DESVANECE LA EUFORIA DE LAPLACE**

Sin embargo, a finales del siglo XIX comenzaron a aparecer fenómenos físicos que no podían ser descritos mediante las leyes de Newton. Los físicos tuvieron que introducir nuevas leyes más complicadas; tanto, que a veces ni ellos mismos las entendían. En este nuevo marco surgieron dos grandes teorías, la teoría de la relatividad la cual no era más que un largo rodeo para explicar por qué los cuerpos a altas velocidades no obedecían las leyes de Newton y la mecánica cuántica que introdujo un concepto desconocido hasta entonces, la incertidumbre; así, en los niveles más profundos de la materia subyace una indeterminación, una aleatoriedad que implica, en la práctica, que nunca llegaremos a conocer perfectamente el estado de un sistema.

En la actualidad, todos los físicos opinan que la euforia de Laplace era desmesurada y ya no están tan seguros de poder llegar algún día a conocer, absolutamente, el Universo. Incluso la teoría de la relatividad, que no deja de ser determinista, está siendo revisada hoy día, introduciendo el componente de azar y caos del que carece en su formulación actual. Incluso en las matemáticas, las ciencias exactas, cuyo desarrollo ha sido paralelo al de la física, la incertidumbre comienza a ser un elemento determinante. Durante el siglo XIX se establecieron definitivamente las leyes de la lógica y del álgebra; parecía que los sistemas formales matemáticos habían alcanzado su esplendor. A partir de unos pocos postulados se podían deducir y demostrar, de una forma sencilla, todas las afirmaciones matemáticas posibles. Se habían acabado las controversias sobre las demostraciones de enunciados complicados. Sin embargo, a comienzos del siglo XX, un lógico llamado Kurt Gödel demostró más allá de toda duda que no podía existir un sistema lógico completo. Con cualquier conjunto de postulados de partida, siempre existirían afirmaciones que no podrían ser ni demostradas ni revocadas basándose en dichos postulados. Por tanto, siempre deberíamos aceptar, como acto

de fe, determinadas proposiciones matemáticas. El teorema de la incompleción de Gödel, como se conoce a este enunciado, tiene en jaque desde entonces a los matemáticos. Ellos ya no están seguros de nada y han aparecido corrientes, como la lógica difusa que tratan las matemáticas de una forma más empírica y semialeatoria que nunca.

Parece ser un devenir general en todas las ciencias. Una vez que se van desarrollando, vamos conociendo más y más leyes sencillas que permiten explicar los fenómenos observados. Inevitablemente pensamos que lo conocemos todo y que podemos deducir el comportamiento real de las cosas a partir de un conjunto pequeño de leyes simples. Llegamos entonces al reduccionismo y al determinismo: teniendo diez o doce leyes fundamentales, una visión reducida del mundo en toda su complejidad, se puede determinar perfectamente como va a evolucionar un sistema real; somos omniscientes y casi omnipotentes.

Tras la etapa de euforia omnisciente, generalmente el progreso de una ciencia, conlleva descubrimientos que imponen límites al conocimiento, produciendo una modelación de la realidad que obliga a plantear nuevos paradigmas. También, los mismos conocimientos permiten avances casi inimaginables hasta ese entonces; una vez vencido ese paradigma determinista, la nueva generación de científicos “liberados” puede evolucionar hasta descubrir la fusión nuclear, los agujeros negros, la teoría del caos, la lógica de redes neuronales, la microelectrónica o la nanobiología.

### ***DETERMINISMO Y REDUCCIONISMO VS EL GENOMA HUMANO***

En lo que va corrido del siglo XXI, la biología es la ciencia de moda. Sin embargo en la segunda mitad del siglo XX esta disciplina científica evolucionó más rápida y más espectacularmente que ninguna otra rama del conocimiento científico. De la simple observación y clasificación de los fenómenos biológicos hemos pasado a una comprensión a nivel molecular de las leyes que gobiernan los organismos vivos, inaugurándose la nueva etapa de la «*biología molecular*». Disponemos de diez o doce leyes que parecen explicar todos los fenómenos que regulan nuestros cuerpos y el de todos los seres vivos de este planeta y eso que aún no conocemos ningún ser vivo fuera del nuestro. Sabemos que la información genética se almacena en los ácidos nucleicos y se ejecuta en forma de proteínas. Conocemos la manera en que algunos de los genes se regulan y se expresan de una forma casi perfectamente predecible, basta con leer el libro de divulgación: «*El azar y la necesidad*» escrito por Jacques Monod para darse cuenta de la euforia que embargó a este premio Nóbel francés en la década de los setenta, del siglo pasado. En fin, parece que estamos llegando a una comprensión del funcionamiento de la vida, al menos tan buena como la que establecen las leyes de Newton para el funcionamiento de la materia o la ley de la palanca para el funcionamiento de los objetos estáticos.

¿Acaso no estamos los biólogos en la etapa de Laplace?. Sin embargo, creemos que todavía falta el Gödel que imponga límites a la biología y el Heisenberg que nos enseñe

que algunos de los fenómenos de la vida poseen una incertidumbre intrínseca a niveles profundos, que nunca podremos superar. Existe actualmente una tendencia a cambiar el paradigma reduccionista de la biología por uno que incluya la noción de integración holística con un componente de incertidumbre que obliga a redefinir el concepto de complejidad generando un cambio en el enfoque sobre el hombre. Los sistemas irreversibles y metaestables comienzan a explicar como nos diferenciamos a partir de una masa informe de células hasta tener una serie de estructuras de alta complejidad que funcionan e interactúan entre si. Sin embargo, es posible que hacia el año 2003, fecha en que está prevista la terminación de la secuenciación completa del genoma humano, aparecerán declaraciones de algún biólogo que sonarán, algo así como: «*Ya hemos terminado: dadme el genoma de un individuo y os diré todo sobre éste*». Creemos aún en el determinismo biológico. Somos víctimas de nuestros propios conocimientos y de una visión reduccionista de las cosas. Aunque algunos científicos poseen más sentido común que otros, la mayoría de ellos, así como prácticamente la totalidad de la «gente de la calle», al ser interrogados reconocerán sin duda alguna que «*los genes determinan nuestra biología y nuestro comportamiento*». Algunos reconocerán que «*el ambiente también tiene algo que ver, pero en mucha menor medida que los genes*».

### **CONOCIENDO LOS GENES TENDREMOS UN CONOCIMIENTO ABSOLUTO DEL HOMBRE**

Esta afirmación, por demás preocupante, es la principal justificación epistemológica del Proyecto Genoma Humano. Se basa en la falsa esperanza del determinismo genético. En nuestra opinión, aunque conozcamos los genes humanos, estaremos muy distantes de conocer la esencia biológica del hombre, sin embargo habremos dado un paso gigante hacia el establecimiento de nuevos paradigmas frente al concepto de hombre como hombre genómico.

También el determinismo genético es el que ha generado muchos de los problemas éticos y legales que pueden aparecer como consecuencias del Proyecto Genoma. La repetida casuística sobre compañías de seguros que se niegan a asegurar a portadores de posibles genes defectuosos, se entiende, en gran medida, si dejamos a los genes la mayor parte de la responsabilidad en el proceso de formación del hombre. Qué duda cabe de que existen algunos genes que, hasta ahora, se han mostrado completamente deterministas. Estamos acostumbrados a que nos digan que si portamos el gen de la hemofilia, inevitablemente padeceremos esta enfermedad, o que si somos homocigotos para el gen defectuoso de la hemoglobina, sufriremos anemia falciforme. Bajo el enfoque reduccionista y simplista no entendemos que portar uno de los genes que pueden provocar cáncer no significa necesariamente que padeceremos esta cruel enfermedad. No se nos ha enseñado que existe un amplio abanico de posibilidades y un gran espectro de severidad en la mayoría de las enfermedades poligénicas y que, posiblemente, el ambiente y las condiciones de vida o incluso quizás los aspectos psicológicos, pueden influir en el desarrollo de la enfermedad. Corremos el riesgo de que, si nos dicen que tenemos ciertas

posibilidades de padecer un cáncer de hígado a los cuarenta, acabemos suicidándonos a los treinta y nueve. Adelantémonos a nuestro tiempo y aprendamos a no ser reduccionistas. Sólo así podremos superar los problemas en los que nos veremos inmersos por causa del Proyecto Genoma Humano. No podemos discriminar a nadie por sus genes y tengamos la humildad de reconocer que los genes no son todo en la vida y que tenemos opciones de vivir, a pesar de que nos hayan detectado un gen que nos confiera propensión a la calvicie o a la miopía.



**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

## INFORMACIÓN, GENOMA Y SOCIEDAD

*“Junto a una posible limitación de lo que el ser humano puede saber, cabe también hacernos la pregunta de lo que debe saber. Dicho de otro modo: ¿Hay datos que conduzcan a un acontecimiento que sería preferible que los seres humanos no adquirieran?.....”*

**FRANÇOIS JACOB. “LA SOURIS, LA MOUCHE ET L’HOMME”  
(El ratón, la mosca y el hombre). 1997.**

SETI (“*Search for ExtraTerrestrial Intelligence*”) es la sigla de una red integrada por científicos de diferentes especialidades que exploran permanentemente la posibilidad de existencia de otras inteligencias en el universo. Una manera de poder algún día llegar a tener, una evidencia preliminar de la existencia de sociedades inteligentes en el universo, es la búsqueda de emisiones en la frecuencia de los 300 Hz. Este rango de emisión del espectro electromagnético no es emitido por ninguna de las fuentes emisoras en el universo incluyendo a los agujeros negros y los infiernos nucleares de las estrellas. Se considera pues que la emisión en la franja de 300 Hz sólo puede provenir de sociedades que han desarrollado la tecnología suficiente como para enviar mensajes cifrados al espacio. SETI representa la esperanza de la sociedad humana por no considerarse la única en el universo además de la necesidad de compartir información en diferentes grados de desarrollo tecnológico.

Si el párrafo anterior hubiese sido escrito hace 20 años, haría parte de una de las muchas obras de ciencia ficción que han escrito sobre el tema, si no fuese porque SETI existe como programa oficial sustentado parcialmente por la Fundación Nacional de Ciencia (“National Science Foundation”) de los Estados Unidos. Muchos hechos establecidos nos obligan, en el inicio del tercer milenio, a retomar el debate sobre la existencia de sociedades galácticas posiblemente iguales o tecnológicamente mejor desarrolladas que la nuestra. Si bien es cierto que para identificar vida extraterrestre se requiere mucho más que sólo encontrar una señal de banda estrecha, la esperanza es que dicha señal sería la señal portadora dentro de la cual se encontrarían mensajes codificados que podrían ser decodificados utilizando instrumentos

cuya tecnología posiblemente no hemos desarrollado todavía. Se asume que la información codificada lanzada al espacio en forma de ondas electromagnéticas, equivaldría a la existencia de vida inteligente extraterrestre. En este punto el lector se preguntará ¿Qué tiene que ver la postgenómica con el proyecto SETI y con el concepto de sociedades galácticas? Pues bien una respuesta muy sencilla es que la base de toda sociedad organizada es la información y el genoma humano contiene información.

### **INFORMACIÓN Y GENOMA UNA ALIANZA ESTRATÉGICA**

A partir de su consolidación como especie biológica, el hombre inició el lento pero continuo camino hacia el establecimiento de una estructura social, la cual evolucionó a la par del desarrollo cerebral y de la selección rigurosa de sociedades que paulatinamente distribuyeron y canalizaron la información como fuente de supervivencia y de desarrollo cultural.

Muchos de los textos de biología definen la vida en términos de una organización funcional que se mantiene a base de reciclamiento de la energía, sin embargo no enfocan el concepto de vida de forma integral. Siempre se ha asociado a la vida el concepto de metabolismo como un proceso del cual todo organismo prolifera y puebla un determinado nicho ecológico. A pesar de que esto es cierto, es tal vez igual o más importante considerar que la información es un atributo inherente de la vida. Los genomas de los organismos contienen una información biológica que ha sido moldeada por millones de años de evolución pero que no siempre debe estar asociada a un metabolismo y ni siquiera a la serie de estructuras organizadas que forman la célula. Si esto es cierto los virus y otros complejos moleculares deben ser considerados como estructuras vivas, mas aún cada una de las moléculas que forman los organismos que contienen información. La información desde este punto de vista es una frontera que la biología actual debe incluir dentro de sus nuevos paradigmas.

La biósfera está plagada de información, cuando un lobo aúlla, está enviando mensajes que en muchos casos tienen que ver con alertar, identificar o buscar su pareja para procrear. La secreción, por muchas especies de animales, de sustancias como las feromonas constituye un claro lenguaje químico que contiene miles de unidades de información que se traducen en acciones que van desde escapar hasta atraer compañeros sexuales. Internamente en los organismos multicelulares se establecen redes de información química cuyo efecto es el activar la expresión de genes o liberar un neurotransmisor para que se desencadene un impulso nervioso que medie una contracción muscular. En cada uno de los ejemplos expuestos anteriormente existe un mensaje codificado. Éste tipo de comunicaciones son la esencia de la vida.

Una de las adaptaciones más espectaculares de la especie humana es la capacidad de desarrollar continuamente y adoptar maneras nuevas y mejoradas de intercambio de información. Ninguna otra especie se acerca a nuestras capacidades de lenguaje y escritura. La moderna tecnología de la información facilita el almacenamiento, procesamiento y transporte de información. Los electrones casi ingravidos y ondas electromagnéticas viajan a velocidades

cercanas a la de la luz (aproximadamente 300.000 Km/segundo), éste avance en la información ha ocurrido tan solo en unos cuantos cientos de años, sin embargo en muy poco tiempo estos conductos de información transformaron la sociedad humana puesto que permitieron avances fundamentales en la Internet y otras redes de comunicación global hasta el grado de desarrollar redes muy complejas que actúan como inteligencia colectiva.

Si hacemos un análisis evolutivo podemos darnos cuenta que la información fluye a través de todos los sistemas biológicos y que muchos de los desarrollos de computación e informática son imitaciones de los procesos de información que se suceden en los organismos. Los sentidos son la entrada a una red de información interna. Los oídos, los ojos, la nariz y la boca convierten mensajes de tacto, sonido, olor, gusto o luz en patrones de impulsos nerviosos. Ésta información es transferida en forma de flujos de iones que se mueven a lo largo de la membrana de las células nerviosas. Estas ondas despolarizantes denominadas potencial de acción, conducen información hacia y desde el cerebro. A pesar de que la velocidad de transmisión de estas señales parece lenta, aproximadamente unos 100 Km /hora, si la comparamos con las velocidades de movimiento de electrones y de fotones, se ha convertido, desde su aparición evolutiva hace unos 500 millones de años, en una adaptación revolucionaria para la transmisión de información biológica interna. El sistema nervioso se desarrolló a partir de esta adaptación generando circuitos de información cada vez más complejos para la transmisión y procesamiento de información que han permitido moldear las estructuras sociales a lo largo de la evolución humana. Así el cerebro procesa información, la clasifica, la cualifica y la almacena de manera dinámica en un proceso que denominamos memoria. La memoria, es pues, la base de toda la cultura humana desde el paleolítico hasta nuestras sociedades postmodernas.

Sin embargo a pesar de que nos parece espectacular el desarrollo de circuitos neuronales complejos, existe una red de información que se yuxtapone al sistema nervioso. Esta red es inclusive más antigua que el desarrollo de los organismos multicelulares y hace parte constitutiva de todas las formas de vida existentes en nuestro planeta. Todas las formas vivas están compuestas por moléculas cuyas características químicas son muy variadas. Poseemos millares de proteínas, lípidos, carbohidratos y dos tipos de ácidos nucleicos, moléculas que son comunes a todos los organismos. En cada molécula biológica hay un contenido de información cuya memoria evolutiva es cualitativamente más complejo que la que se procesa en el sistema nervioso; mas aún, el sistema nervioso no procesaría información si no existiesen estas moléculas.

Merece especial atención aquella información contenida en las secuencia de nucleótidos de los ácidos nucleicos. Desde la demostración que la información biológica está contenida en segmentos lineales de nucleótidos, el gen ha sido la piedra roseta de la sociedad moderna. En 1943 Erwin Schrödinger, el fundador de la mecánica cuántica una teoría que representó una de las mayores revoluciones científicas de la humanidad, especuló que los genes eran "*cierta clase de guión codificado*". Con ésto Schrödinger quiso decir, aun antes de postularse

el modelo de la doble hélice del ADN, que la información biológica está cifrada en un formato químico. Este formato químico representa el genoma en toda su extensión y es la base no solamente de la existencia de la vida en la tierra si no de su continuo cambio evolutivo.

Un aspecto esencial del genoma como información es que no es importante el cómo esté codificado, lo que es fundamental es cómo pueda recibirse y decodificarse el mensaje, en últimas cómo se construyen las proteínas que permiten la vida. Mediante un sistema de desciframiento denominado código genético, en grupos de tres nucleótidos, denominados codones, se descifra la información contenida en la secuencia de codones del ARN mensajero a secuencias específicas de aminoácidos que hacen parte de las proteínas. No es sorprendente pues, que el proceso de evolución por selección natural, impuso fuertes restricciones para el desciframiento del mensaje, por esta razón el código genético, es la principal adaptación evolutiva para la transferencia de la información biológica.

Uno de los conceptos básicos con relación a la evolución del genoma se refiere al aumento en el contenido del formato químico, el ADN, sin que haya un aumento proporcional de la información biológica. Este hecho, conocido como la paradoja del valor C, ha sido uno de los retos epistemológicos de la biología molecular. En las bacterias más del 98% de su genoma, que tiene un tamaño de aproximadamente 3 millones de nucleótidos, codifica por proteínas mientras que en los mamíferos, cuyo genoma tiene un rango de tamaño de 1000 a 4000 millones de nucleótidos, solo el 5 al 2% lo hace; lo anterior quiere decir en términos generales que no ha ocurrido un cambio sustancial de la información biológica a lo largo de la evolución, si esto es así la pregunta lógica es qué sucedió. La respuesta a este interrogante está en un tipo muy curioso de ADN, el denominado ADN de relleno (“junk DNA”); éste no es más que un tipo de ADN compuesto por secuencias altamente repetidas o redundantes que carecen de significado con relación a la información que especifica proteínas. Podríamos decir que la evolución de la información que especifica proteínas en los genomas ha sido muy poca mientras que la redundancia, o el gagueo genético, ha sido espectacularmente alta. El hecho anterior plantea una serie de problemas a la hora de considerar la calidad de la información.

En los párrafos anteriores hemos expuesto de manera muy general el problema de la información en el contexto de la evolución de los genomas, sin embargo no hemos planteado cómo se ha resuelto este dilema biológico. El inicio de la aproximación a una solución ocurrió a finales de la década de los cuarenta del siglo pasado cuando Claude Shannon, un matemático que trabajaba en los laboratorios de investigación de la compañía de teléfonos Bell, publicó en el *Bell System Technical Journal* un trabajo titulado “Una teoría matemática de la comunicación” según Shannon el problema principal de la comunicación es reproducir en un punto, de forma exacta o aproximada, un mensaje seleccionado en otro punto. La importancia del trabajo de Shannon es que consideró el contenido de información en términos de su calidad de ser reciente. Una comunicación no contiene información si vuelve a declarar lo que se sabe o lo que ya había sido anticipado. Así la repetición de un mensaje predecible no contiene información adicional. Un libro entero puede estar escrito con las palabras “la evolución de los genomas es redundante”

repetidas 100.000 veces, esto sería aburrido y monótono al lector. La misma información puede transmitirse en una sola repetición de la oración. Esta aproximación nos lleva de nuevo al genoma humano; ya habíamos dicho que sólo del 2 al 5% era informacional en términos de proteínas, o sea una sola repetición de cada una de los genes codificados por el ADN, mientras que más del 95% es repetido y representaría una información tediosa y monótona para la vida. Sin embargo el ADN repetido hace parte de los genomas de los mamíferos incluidos nosotros.

Si analizamos la siguiente secuencia de nucleótidos de un trecho de una molécula de ADN, (ATAATAATAATAATAATAATA) que equivale a una continua repetición de nucleótidos de Adenina y Timina, rápidamente nos podemos dar cuenta que la unidad ATA se repite ocho veces, o sea  $(ATA)_8$ , en este sentido la información reciente sería ATA, el resto es ruido de fondo. Pero si consideramos la siguiente secuencia (ATATATATATATATATATATATAT) tendremos que la información reciente es solamente AT que tiene un ruido de doce repeticiones AT que en términos de información reciente, no cuentan, sin embargo hacen parte estructural del genoma. Situaciones como las que hemos descrito, se encuentran de manera muy frecuente en el genoma humano. ¿Será que el genoma humano es un libro repetido, aburrido y monótono para leerlo?, ¿Por qué tanta repetición?, ¿Qué coste evolutivo tiene la repetición? Estas son preguntas a las que nos enfrentamos los que trabajamos en este campo y cuyas respuesta están todavía por conocerse.

Para capturar la idea de información reciente, Shannon consideró un mensaje como una secuencia de símbolos (podía ser 0 y 1) y desarrolló una manera matemática para medir la posibilidad de que un símbolo no fuese anticipado. Si había N posibilidades de elección de símbolos cada uno con una posibilidad de que ocurriera, entonces la posibilidad de que cualquier símbolo ocurriera en un determinado evento, sería de 1 dividido entre N ( $P = 1/N$ ). La probabilidad de que dos eventos ocurrieran simultáneamente sería entonces ( $P_{1,2} = P_1 \times P_2$ ). Para hacer que la información aumente conforme disminuye la probabilidad, Shannon hizo que la información (I) fuese proporcional a  $1/P$ . Si fuésemos a calcular muchas informaciones simultaneas, o en el mismo formato, por ejemplo el ADN, podemos expresarlas como  $I_{total} = I_1 + I_2 + I_3 + \dots$  o en términos reales  $Log(1/P_{1,2}) = Log(1/P_1) + Log(1/P_2)$ . Con base en esta aproximación, Shannon definió la unidad de información reciente más pequeña como un *bit*. En estos términos la información repetida dentro del ADN adquiere un nuevo contexto, el cual muestra sutilezas evolutivas que hacen del ADN chatarra especialmente importante para que se pueda mantener la información reciente o semántica.

Las matemáticas de Shannon y el trabajo subsiguiente de Warren Weaver, Andrey Kolmogorov y otros hicieron posible que la tecnología de la información pudiese desarrollarse hasta el punto actual en que es posible transitar por las autopistas de información y poder obtener información reciente y en tiempo real de las muchas bases de datos que están disponibles. En este sentido el trabajo de Shannon revolucionó no solamente el concepto de información, si no, que determinó el cambio de la sociedad moderna a la sociedad postmoderna que impuso la globalización como punto de contacto no solo social si no biológico.

## **GENOMA Y ENTROPÍA**

El genoma, como formato químico, es mantenido y replicado dentro de estructuras biológicas como un proceso organizado que toma precursores simples, los desoxirribonucleótidos trifosfatos o dNTPs y los va agregando de manera dirigida para formar una nueva molécula de ADN. Lo anterior implica un desarrollo evolutivo para crear orden a partir del desorden químico. El físico Ludwig Boltzmann determinó en el siglo XIX que la entropía, una medida del desorden termodinámico, de una estructura cualquiera que tuviese muchos componentes, se relaciona con el número de disposiciones (subsistemas) que dichos componentes pudieran tener y la probabilidad de que ocurriera la disposición tal como está (el sistema *in toto*). En otras palabras un montículo de arena tendría un elevado grado de entropía mientras que un castillo de arena tendría un bajo grado de entropía. O de modo similar, una hilera al azar de aminoácidos que no asume una organización particular tendría un elevado grado de entropía en comparación con una secuencia plegada de aminoácidos que toma su forma por fuerzas evolutivas con el fin de poder cumplir una función o llevar un mensaje específico.

La baja entropía equivale a un alto contenido informacional, así la única diferencia entre la ecuación de entropía de Boltzmann y la ecuación de la información de Shannon, es que la primera tiene signo negativo y las probabilidades se multiplican por una constante termodinámica, la constante de Boltzmann (se expresa en calorías/grado centígrado). La correspondencia entre entropía e información es un aspecto intrínseco de todo organismo y ha obligado a considerar la vida en términos probabilísticos, los análisis probabilísticos son la base para la búsqueda de vida interna inteligente terrestre, las señales ocultas dentro del ADN, las proteínas y otras moléculas de la vida.

## **LAS BASES DE DATOS**

Una de las características culturales del hombre es su capacidad de recolectar información, procesarla y difundirla. Desde las pinturas rupestres de las cuevas de Altamira, en las que se informa la caza en grupos, pasando por los legados de los jeroglíficos de las culturas Egipcia y Azteca, la sistematización del conocimiento de la época consignada en la biblioteca de Alejandría, hasta la actualidad se han construido y manejado bases de datos. La memoria social es la base del desarrollo de la sistematización de la información que ha permitido los grandes cambios sociales en la historia de la humanidad.

Para que los descubrimientos e invenciones tengan un impacto duradero sobre la civilización, deben ser recordados y diseminados entre quienes sean capaces de actuar sobre ellos. La base del desarrollo de la ciencia occidental se debió en gran parte a la existencia de una estructura de información que, representada en revistas científicas, permitió la difusión y expansión de los conocimientos científicos. Actualmente gran cantidad de pensamientos innovadores, hallazgos experimentales y trabajos de creación son capturados y diseminados de

manera instantánea mediante computadores en red. Las bibliotecas de información biológica ahora son parte integral para la investigación biológica.

En 1979 un grupo de biólogos y matemáticos se reunió en la Universidad Rockefeller en Nueva York y propusieron la creación de un banco de datos para almacenar secuencias de ADN, nació la primera base de datos de secuencias de ADN. A medida que se desarrollaba la biología molecular y las metodologías de secuenciación se hicieron más fáciles, aumentó el volumen de secuencias de ADN de muchos organismos. Era necesario no solamente almacenar, si no, dar la oportunidad de comparar, mediante alguna herramienta matemática, los parecidos entre secuencias de ADN de diferentes organismos, mas aún, era importante compartir toda esta información con otros investigadores, en suma crear una red integrada que permitiera en tiempo real obtener información a escala mundial de secuencias de ADN. La importancia de este fenómeno motivó a muchos científicos a dedicarse al desarrollo de herramientas de computación que pudiesen satisfacer las necesidades científicas de los investigadores, nació una nueva disciplina científica, la bioinformática. Ésta se consolidó como un arma muy poderosa que abonó el camino para emprender la secuenciación del genoma humano y del de otros organismos. Hacia finales de los 80 del siglo XX, el desarrollo de la tecnología de la secuenciación y de las herramientas de bioinformática permitió la iniciación de la genómica.

Como una respuesta a la iniciativa de la Universidad Rockefeller, en 1981 el laboratorio Europeo de biología molecular (EMBL) con sede en Heidelberg, estableció la biblioteca de datos del EMBL y en 1982, después de una ardua discusión, el Instituto Nacional de Salud de los Estados Unidos, apoyó una iniciativa de Michael Waterman y Temple Smith del Departamento de Energía (DOE) del laboratorio nacional de los Álamos en Nuevo México, para establecer el GeneBank en los Álamos con un presupuesto de 3.5 millones de dólares por cinco años.

El GeneBank no es sólo un depósito de secuencias de ADN, es a la vez un instrumento de información relacionada con las secuencias de ADN consignadas en el banco de datos. La posibilidad de difundir de manera dinámica las secuencias de ADN permitió el establecimiento de una nueva división de la biblioteca nacional de medicina (NLM), el centro nacional para la información de la biotecnología (NCBI) para que se encargara de la difusión del GeneBank, actualmente esta división tiene a MEDLINE que es un conector de información no solo de secuencias de ADN si no de la literatura concerniente a ellas. Basta entrar a la dirección [www.nlm.nih.gov](http://www.nlm.nih.gov) y tener a disposición una gran cantidad de información referente a genes, organización de los genes y bancos de datos de secuencias. Hoy en día gran parte de los recursos del NCBI y la NLM son de dominio público, existiendo algunas restricciones en el uso de programas de computador para el manejo de las secuencias de ADN y de proteínas. Actualmente existen en el GeneBank unos 7 millones de secuencias que pertenecen a más de 50.000 especies de organismos incluyendo la totalidad de las secuencias de ADN humano conocidas. Los esfuerzos por sistematizar la información han sido a escala mundial, Países como Japón y Suiza entre otros han desarrollado y consolidado bancos de datos de secuencias



de proteínas, y de ADN. Los bancos de datos son la memoria de la sociedad postmoderna además serán el punto de apoyo del cambio biotecnológico global que nos depara el futuro cercano con base en los resultados del proyecto genoma humano (PGH).

Además del GenBank público que tiene secuencias de más de 7 mil millones de unidades de ADN, Celera Genomics, una empresa privada, dice que tiene 50 terabytes de datos almacenados, lo que equivale a 80.000 discos compactos. Las secuencias brutas son secuencias de las 4 letras - A, T, C y G - lo que corresponde a 3 mil millones de bases presentes en el genoma humano. Es imposible acceder a estos datos si no se tiene un software especial que le dé sentido a estas secuencias. Se están desarrollando y poniendo en el dominio público software, pero el acceso a las bases de datos de las empresas privadas está disponible solo después de pagar una suscripción. Incyte, una empresa genómica privada, lanzó un programa genómico en marzo que permite a los investigadores ordenar secuencias de más de 100.000 genes “on-line”. Entre los subscriptores a las bases de datos se incluyen las empresas farmacéuticas gigantes tales como Pfizer, Bayer y Eli Lilly. La suscripción a los datos de genes de Celera, *gene notes*, costará entre 5 a 15 millones de dólares y para académicos, entre 2000 a 15000 dólares al año.

La sociedad postmoderna basa su existencia en el conocimiento y la manera de transferir y utilizar la información generada por este conocimiento. Si toda la información proveniente de un sistema vivo pudiese capturarse y digitalizarse sin cambiar su naturaleza fundamental, estaríamos a las puertas de un gran paso evolutivo de la información, la creación de la vida electrónica *o in silico*. Aunque este tema actualmente parece ciencia ficción, ya existen esfuerzos pioneros por desarrollar sistemas de inteligencia artificial cada vez mas complejos con el objetivo de poder en un futuro no muy lejano, reinventar el proceso de la vida pero sin átomos de carbono, una vida que utilizará el silicio y otros elementos comunes en los actuales computadores. En el momento en que esto se llegue a lograr, estaremos frente a un cambio crucial de la sociedad y de la especie humana.

# **DE LA BIOLOGÍA MOLECULAR AL GENOMA**

*“ ...Por si mismos los datos genéticos son mera sintaxis. La sorprendente utilidad de la información genética codificada deriva del hecho de que los aminoácidos la comprenden ”*

**PAUL DAVIES “THE FIFTH MIRACLE”  
(El quinto milagro). 1999.**

## **LA MOLÉCULA DE LA GUERRA FUE EL INICIO DE LA HISTORIA**

Si tuviéramos que establecer los orígenes del proyecto genoma humano, una historia que cambió el destino de la humanidad, posiblemente lo fijaríamos en 1869 cuando el químico suizo Johann Friedrich Miescher analizó muestras de células blancas de la sangre que había recogido de unos vendajes manchados de pus en un hospital local durante la guerra franco-prusiana. Meischer rompió las membranas de estos glóbulos blancos utilizando enzimas digestivas, obteniendo así muestras casi puras del material contenido en el núcleo celular, del cual extrajo altas concentraciones de una sustancia química, rica en fósforo, que denominó «nucleína». Meischer conjeturó que la síntesis de nucleína podría ser un modo de la célula para almacenar fósforo o quizá tuviese «algo que ver con la herencia». No sabemos cómo ni por qué se le ocurrió al químico suizo este alarde de genialidad, ya que es difícil encontrar una relación *a priori* entre un material extraído de los glóbulos blancos y algo tan esotérico como era la herencia en 1869.

En 1889 otros químicos se dedicaron a purificar aún más la nucleína, eliminando las últimas trazas de proteína, obteniendo así una sustancia gomosa con carácter levemente ácido; el ADN había sido aislado, purificado, secado y guardado en una botella con la etiqueta «Ácido nucléico». El frasco que contenía un extraño polvo blanco fue depositado en un estante del

laboratorio. Sin embargo transcurrirían casi sesenta años antes de que se revelase que era, de hecho, un frasco de genes.

La idea de gen no era nueva en tiempos de Miescher. Cualquiera puede ver que los hijos suelen parecerse a sus padres, heredando determinadas características de cada uno de sus dos progenitores. Los antiguos griegos habían pensado que los hijos eran concebidos por «coagulación» del esperma y del fluido menstrual. La capacidad de estos fluidos para transmitir características de sus padres se debía, según Hipócrates, al hecho de que contenían los «gonos» o «semillas» aportadas por todas las partes del cuerpo, que se mezclaban para producir al niño. Igualmente Aristóteles sugirió que ciertas partículas invisibles o «pangenes» surgían del cuerpo y se unían para formar los fluidos reproductores. Esta teoría de la pangénesis fue utilizada por Lamarck y Darwin veintiún siglos después y sirvió como explicación sorprendentemente buena para la idea de Lamarck de que las características adquiridas por un individuo pueden ser transferidas a su descendencia. Lo que parece increíble es el hecho de que Aristóteles e Hipócrates, sobre las mismas bases, no desarrollaran la idea de selección natural.

Sin embargo, a pesar de que sus resultados eran tan evidentes que todo el mundo podía verlos, las «unidades de herencia» no fueron reducidas a nada específico hasta una fecha tan sorprendentemente reciente como 1944. Las «gémulas» darwinianas, sus partículas invisibles imaginarias que fluían de todas las partes del cuerpo hasta los órganos reproductores, habían dominado la idea de herencia por muchos años. Se creía que estas unidades servían de vehículos de transmisión y directores del desarrollo y se supuso que eran autorreproductoras y circulaban por todo el organismo. Herbert Spencer prefirió llamarlas «unidades fisiológicas» y Karl von Naegeli «ideoplasma». Distintos nombres para la misma idea imaginaria, desarrollada sin evidencia experimental alguna, más de veinte siglos atrás.

### ***LAS ARVEJAS DE UN MONASTERIO DIERON LA CLAVE PARA LA HERENCIA***

La evidencia experimental que sustentó la teoría de la evolución propuesta por Charles Darwin, fue minuciosamente estudiada y maravillosamente documentada por un fraile agustino de Brün, Austria (actualmente Brno en la República Checa). Se trata de Gregor Mendel, cuyo visionario trabajo fue ignorado por la ciencia durante sesenta años. Johann Mendel nació en 1822, de padres campesinos, en el pequeño pueblo austriaco de Heitzendorf, localidad que gozaba de una gran tradición en suministrar diestros jardineros a los opulentos terratenientes de la zona. Ingresó a la edad de veintiún años en la Abadía de St. Thomas, tomando el nombre de Gregor. Como los agustinos proveían de maestros a la escuela secundaria local, Mendel fue enviado a la Universidad de Viena donde cursó estudios de ciencias y matemáticas. Afortunadamente, fue introducido en la recién nacida ciencia de la estadística, apoyo imprescindible para sus investigaciones posteriores. A su regreso al monasterio, el Abad de St. Thomas, F.C. Napp, había establecido un laborioso programa de experimentación con plantas en los jardines del monasterio, que pronto se convirtieron en el dominio de Mendel.

Podemos considerar a Mendel como el primer biólogo matemático de la historia. A diferencia de cualquiera de sus predecesores o contemporáneos, intentó disponer de una forma estadísticamente precisa sus resultados. Sus experimentos metódicos y meticulosos eran muy distintos a cualquier cosa que se hubiera hecho antes en biología. Mendel reunió 34 variedades de plantas de arvejas de viveros de toda Europa y durante varios años afinó cuidadosamente sus selecciones hasta que obtuvo verdaderas plantas reproductoras que diferían entre sí en siete pares de rasgos, incluyendo la forma de la semilla y las vainas, el color de la semilla, la longitud del tallo y la posición de las flores.

En un cruce entre plantas surgidas de semillas verdes y otras surgidas de semillas amarillas, los descendientes no mostraron una mezcla de esos colores, sino que sus semillas eran todas amarillas. El rasgo verde había desaparecido aparentemente. Entonces Mendel cruzó entonces miembros de esta primera generación de semillas amarillas entre sí y observó sorprendido, como algunos descendientes de esta primera generación volvían a mostrar el rasgo verde. La relación entre plantas amarillas y verdes en esta segunda generación era claramente de tres plantas amarillas por cada planta verde. Mendel sabía que el análisis estadístico era imposible sin un número adecuado de plantas. En el experimento descrito, leemos en los registros de Mendel que había cultivado y comparado 8.023 híbridos, de los cuáles 6.022 tenían semillas amarillas y 2.001 eran verdes.

Mendel había encontrado, descrito y dado nombre a los atributos biológicos de las arvejas transmisibles mediante herencia: al rasgo amarillo, lo denominó «dominante», al verde, que desaparecía cuando estaba en presencia del amarillo, lo denominó «recesivo». Con base en estos análisis propuso una teoría compatible con las proporciones relativas que había encontrado, según la cuál cada característica de un individuo estaba determinada por dos factores o «elementos», uno heredado del padre y otro de la madre. Únicamente cuando los dos factores eran de la forma recesiva se podía desarrollar la semilla de color verde.

El siguiente paso de Mendel fue considerar la herencia conjunta de dos características. Para ello cruzó plantas con semillas amarillas lisas con otras que tenían semillas verdes arrugadas. Las semillas de la primera generación eran todas amarillas lisas (los dos rasgos dominantes). Sin embargo en la segunda generación se produjeron semillas amarillas lisas, verdes lisas, amarillas arrugadas y verdes arrugadas en una proporción de 9:3:3:1, como preveía su teoría.

En 1865, Mendel presentó los resultados de ocho años de investigación en una reunión de la Sociedad de Ciencia Natural de Brün a una audiencia de entusiastas científicos locales. Las actas de la reunión que, sorprendentemente, aún se conservan, registran que no se hizo ni una sola pregunta. Los oyentes rápidamente se enfrascaron en una discusión sobre el tema caliente del día: «*El Origen de las Especies*», de Darwin, que había sido publicado seis años antes. El artículo de Mendel se publicó en la Revista de la Sociedad al año siguiente y una reseña de una página apareció en una enciclopedia alemana de cultivo de plantas. Le siguió, en palabras de L.C. Dunn, «uno de los más extraños silencios en la historia de la Biología».

## **LA DIVISIÓN CELULAR, UN DIVERTIMIENTO NECESARIO PARA LA HERENCIA.**

Para los microscopistas de la segunda mitad del siglo XIX, estaba ya claro que la materia viva estaba formada por células individuales. Ya en 1838, Mathhias Jakob Schleiden había afirmado que «*todas las plantas son agregados de seres separados, independientes e individuales, esto es, las propias células*». El botánico y amigo íntimo de Darwin, Robert Brown, había sido, en 1831, el primero en observar «en cada célula... una única aureola circular, o núcleo de la célula, como puede denominársela». La posición central del núcleo en la célula llevó a los investigadores a pensar que su función era importante. Por una vez, la evidencia posterior iba a estar a favor de las teorías científicas emitidas con base en la pura ideología, demostrando que el núcleo era, realmente, parte fundamental de la célula y el lugar donde residen los genes que controlan el funcionamiento de la misma.

En 1873, Friedrich Schneider, otro microscopista, había observado un fenómeno realmente interesante, estudiando las células del diminuto gusano plano transparente *Mesostomium*, visualizó una serie de etapas durante la división de algunas de las células del gusano; notó que algunas de estas etapas involucraban un movimiento y separación de ciertos cuerpos filamentosos en el interior del núcleo que, en estado normal, suele tener un aspecto homogéneo. El mismo fenómeno fue descrito dos años más tarde por Edward Strasburger, en células de embriones de coníferas en desarrollo. En 1879, Walter Flemming confirmó la universalidad del fenómeno, que denominó «mitosis», aparentemente en todas las células en división. Informó que, a medida que las etapas de división progresaban, lo que parecía ser una madeja continua de hebras entrelazadas (la cromatina) en el área nuclear, sufría una división longitudinal y luego se separaba en fragmentos que migran hacia las dos células resultantes. En sus estudios, Flemming fue el primer hombre que vio los cromosomas humanos.

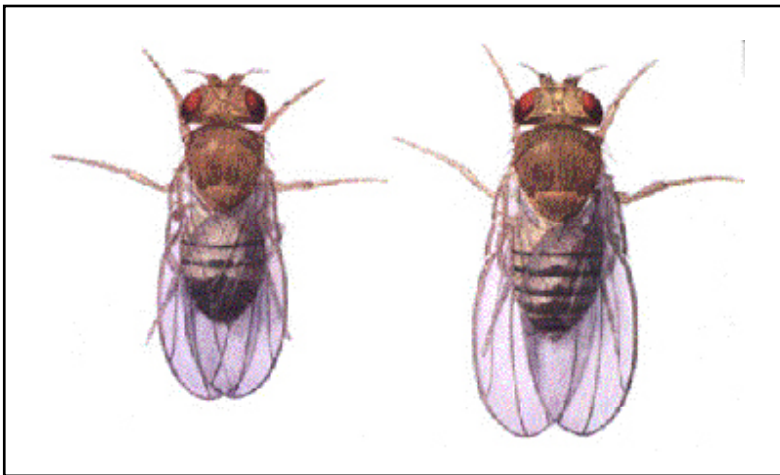
En 1883, Wilhelm Roux estudiando el desarrollo de los huevos de rana, interpretó su observación del núcleo durante la mitosis como un conjunto de «partículas hereditarias». Éstas se repartían entre las células hijas, como grupos de partículas con cualidades diferentes que permitían a cada célula diferenciarse de su hermana. Estaba equivocado acerca de las diferencias en las dotaciones de las partículas hereditarias, que hoy sabemos que son básicamente idénticas (excluyendo mutaciones), pero logró atraer una mayor atención sobre las enigmáticas hebras coloreadas. Porque se teñían fácilmente, en 1888 Wilhelm van Waldeyer los llamaría «cromosomas», lo que significa «cuerpos coloreados».

## **LAS MOSCAS DE LA FRUTA, LAS DROSOPHILAS, REVELAN SUS SECRETOS**

A partir de los primeros años del siglo XX, el liderazgo científico e intelectual correspondió a los Estados Unidos, cargo que continúa ocupando este país en la actualidad. Uno de los principales artífices del desarrollo de las ciencias de la vida en los Estados Unidos fue el embriólogo y genetista Thomas Hunt Morgan. El había llegado a la Universidad de Columbia

en Nueva York, en 1902. Durante mucho tiempo se había sentido incómodo con el hecho de que la relación entre factores hereditarios y cromosomas había sido básicamente inferencial y demandaba una demostración experimental. También estaba interesado en reemplazar la noción «nebulosa» de variación de Darwin por el concepto más preciso y reciente de «mutación», tal como lo había propuesto Hugo de Vries, cuyo laboratorio había visitado en 1903. La idea Darwiniana de variación conlleva continuidad y cambios graduales a lo largo de las sucesivas generaciones; por el contrario, las mutaciones son repentinas, imprevisibles y producen un cambio discontinuo, resultando en un organismo con una característica radicalmente distinta de sus antecesores. Morgan encontró un colaborador invaluable en Edmond B. Wilson, que había comenzado a impartir en Columbia un novedoso curso sobre la herencia, cromosomas, variación y evolución. Entre 1910 y 1925, un pequeño laboratorio de Columbia, habitado por varias generaciones de estudiantes y varios miles de moscas de la fruta, se convirtió en el centro gestor de investigación mundial de la nueva ciencia de la genética.

Las diminutas moscas de la fruta (*Drosophila melanogaster*) figuran, en la actualidad, entre los organismos vivos más estudiados (figura 3.1) y su nombre científico es posiblemente el segundo en popularidad después del *Homo sapiens*.



**Figura 3.1.** Aspectos morfológicos de la mosca de la fruta (*Drosophila melanogaster*). Izquierda macho, derecha hembra.

En 1910, un amigo entomólogo, Frank Lutz sugirió a Morgan que la *Drosophila* podía resultar útil como herramienta experimental. Con un poco de comida y calor, la mosca cumpliría en producir una nueva generación cada dos semanas. Sólo ocho horas después del nacimiento, las diminutas moscas comenzarían a aparearse. Un poco de éter para anestasiarlas, un poco de destreza y una lupa podrían servir para separar los machos de las hembras no apareadas

y realizar fecundaciones controladas de moscas con características específicas, tales como la forma de las alas o el color de los ojos. Quizás tras algunas generaciones aparecerían mutaciones, que se podrían llegar a relacionar con alguno de los cuatro cromosomas que tiene cada célula de la mosca.

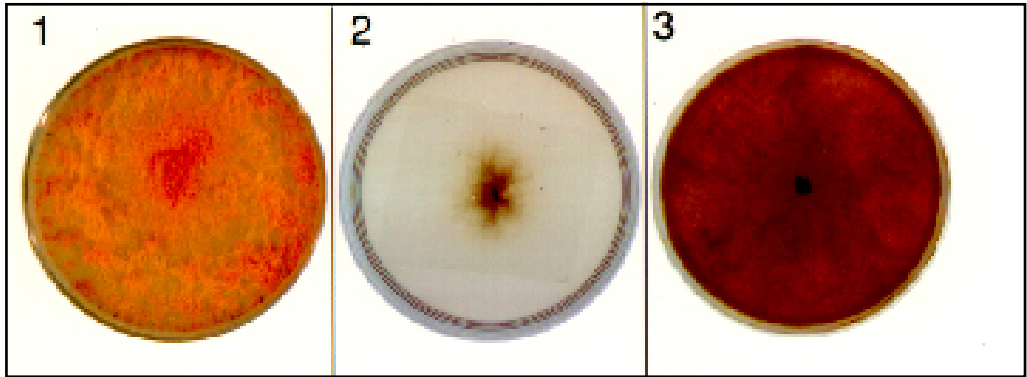
Durante quince años, Morgan y sus colaboradores repitieron y mejoraron los experimentos de Mendel; esta vez con animales y con mucho más conocimiento acumulado. Prácticamente todas las bases de la genética se deben a los resultados de este grupo de investigación: los cromosomas sexuales X e Y, la aparición de mutaciones esporádicas que se transmiten a las generaciones sucesivas, la herencia ligada al sexo, los mapas genéticos (o localización de los distintos genes en cada uno de los cromosomas), la recombinación o sobrecruzamiento entre cromosomas, a partir de la cuál se puede estimar la «distancia genética» entre dos genes, según la probabilidad que tienen de heredarse de forma conjunta.

En 1933, Morgan recibió el premio Nobel en reconocimiento a su exhaustivo trabajo. Se había publicado su libro *«El mecanismo de la herencia mendeliana»*, en colaboración con Sturtevant, Bridges y Muller. En cierto sentido, la búsqueda de los genes había terminado. A partir de ese momento el gen había dejado de ser un mero ejercicio intelectual para convertirse en algo tangible, físicamente localizado en una porción de cromosoma dentro del núcleo celular, cuya herencia seguía unas cuantas leyes extraordinariamente precisas.

A pesar de estos avances en el conocimiento de la mecánica de la herencia, la búsqueda acababa de comenzar, puesto que perduraban sin responder los mayores misterios. ¿Qué era un gen? ¿Cómo su presencia era capaz de provocar que una mosca tuviera los ojos blancos en lugar de rojos o afectar a la forma de la vaina de las arvejas? Las respuestas a estas y otras preguntas sólo pudieron conocerse cuando los biólogos ampliaron su visión de lo biológico para introducir en la biología los conceptos de la química, la física, la matemática y, en los últimos tiempos, la informática.

### **LA QUÍMICA SE INVOLUCRA CON LA HERENCIA**

Los siguientes avances en genética vinieron también de la mano de Thomas Morgan, que había dejado la Universidad de Columbia en 1927, para fundar la sección de biología del Instituto de Tecnología de California, conocido en todo el mundo sencillamente como CalTech. George Beadle y Edward L. Tatum (premios Nobel en 1958) fueron alumnos de Morgan y Sturtevant en CalTech, representando posiblemente la primera generación de «genetistas moleculares». Habían cambiado la *Drosophila* por el hongo rojo *Neurospora crassa* (figura 3.2), cuya biología resultaba, sin lugar a dudas más sencilla.



**Figura 3.2.** Aspectos de tres tipos de mutantes nutricionales del hongo rojo *Neurospora crassa*.

Estudiando ciertas mutaciones nutricionales, o sea aquellas *Neurosporas* que necesitaban de algún nutriente, presente en el medio de cultivo, para vivir, Beadle y Tatum llegaron a la conclusión de que cada mutación afectaba la capacidad del hongo para llevar a cabo una determinada reacción bioquímica. Por ejemplo, uno de los mutantes era incapaz de sintetizar vitamina B6 (o piridoxina), molécula que la cepa normal del hongo fabricaba por sí sola sin ningún problema.

En aquella época era perfectamente conocido que las distintas rutas metabólicas que permitían a los seres vivos la síntesis y la ruptura de compuestos químicos estaban controladas por una serie de moléculas conocidas como «enzimas». Con base en ésta premisa, Beadle y Tatum postularon que la mutación de *Neurospora* afectaba a una de las enzimas implicadas en la síntesis de piridoxina, un precursor de la vitamina B6, impidiendo la fabricación de ésta. Era el primer paso hacia la comprensión de la función de los genes: si el gen mutante impedía la fabricación de una determinada enzima; entonces, la función normal del gen era guiar la fabricación de la enzima. La hipótesis de Beadle y Tatum se conoce como la hipótesis «un gen, una enzima». Ya que, entre 1926 y 1930, James Sumner y John Northrop habían demostrado que todas las enzimas conocidas eran proteínas, fue fácil el cambio a la hipótesis: «un gen, una proteína». Con este nuevo enfoque químico, se dieron los primeros pasos para elucidar el mecanismo de acción de los genes, pero aún quedaba lo más interesante, ¿cuál era la naturaleza química del propio gen? Se sabía que los cromosomas eran una mezcla de ácido nucléico y proteínas. Para los bioquímicos de los años 30, el ácido nucléico era una molécula repetitiva y aburrida, formada sólo por la unión de cuatro tipos distintos de pequeñas unidades. Parecía difícil, por tanto, que el ácido nucléico o ADN pudiera englobar la compleja cantidad de información necesaria que debían contener los genes. Por el contrario, las proteínas eran macromoléculas apasionantes, formadas por una complicada secuencia de al menos 20 componentes distintos, cuyas propiedades químicas eran muy diferentes y permitían abrir un amplio abanico de posibilidades. La mayoría de los bioquímicos, por tanto, se inclinó



por la hipótesis de que los genes eran proteínas, y el ácido nucléico servía únicamente como soporte o «pegamento» para unir las distintas proteínas que formaban los diferentes genes.

Sin embargo, los experimentos de E. Griffith en 1928 y de Oswald Avery, Colin McLeod y MacLyn McCarty, en 1944 demostraron lo contrario. Avery y sus colaboradores marcaron radiactivamente una cepa patógena de la bacteria *Pneumococcus*, que produce la pulmonía tanto en hombres como en ratones. Hicieron dos experimentos: en uno de ellos usaron un isótopo radiactivo del fósforo (que sólo se encuentra en el ácido nucléico, pero no en las proteínas) y en el otro, utilizaron azufre radiactivo (que se encuentra en las proteínas, pero no en el ácido nucléico). Pusieron en contacto estas bacterias radiactivas, previamente muertas por calor, con otra cepa de *Pneumococcus* que no era patógena. Tras comprobar que la cepa inocua se había transformado en patógena, demostraron que la información que había pasado de una bacteria a la otra estaba en forma de ácido nucléico, y no de proteína. El material genético, o «principio transformador» era el ácido nucléico.

### **LOS VIRUS SE CONVIERTEN EN LOS OBJETOS DE INVESTIGACIÓN DE LA HERENCIA**

En Nueva York cerca del temido barrio del Bronx, pero en un apacible lugar a orillas del Atlántico, se encuentran los Laboratorios de Cold Spring Harbor. En los mismos, una serie de personas habían formado un grupo para investigar ciertos virus que parasitan a las bacterias, los denominados bacteriófagos. El grupo que se conoció como «Grupo de los Fagos» incluía nombres que fueron galardonados con el premio Nobel, como Max Delbrück, Salvatore Luria, Alfred Hershey, Emily Ellis y Martha Chase. Cada uno por sí solo hubiera marcado un hito en la historia de la biología, todos ellos juntos, revolucionaron la historia de la humanidad, puesto que sentaron las bases de la «biología molecular» y sus derivaciones, la «ingeniería genética» y la «biotecnología moderna».

Los virus que formaban parte del trabajo diario del Grupo de los Fagos sentían una especial debilidad por infectar a una bacteria insignificante la *Escherichia coli*. Este organismo es un verdadero ganador, existen más *E. coli* en el intestino de cualquier persona que personas en todo el Planeta. Las células de *E. coli* son diminutos bastoncillos de dos micras de longitud y media micra de ancho. Es un organismo muy práctico para los estudios genéticos puesto que es unicelular, se multiplica con rapidez, dando una nueva generación aproximadamente cada treinta minutos y puede prosperar en soluciones acuosas con un poco de glucosa y sales minerales. El ácido nucléico bacteriano está organizado en un único cromosoma circular, que se encuentra libre en el interior de la célula, ya que las bacterias no poseen núcleo. Este único cromosoma, una sola molécula de ácido nucléico o ADN, constituye el genoma de *E. coli* (aunque, normalmente, las distintas células de *E. coli* pueden poseer también una serie de minicromosomas, denominados plásmidos).

Durante más de tres décadas, desde 1940 hasta 1973, las bacterias y los fagos fueron los únicos sistemas adecuados para estudiar con cierto detalle la organización y el funcionamiento

a nivel molecular de los genes, así como las consecuencias de la introducción de nuevo material genético en la célula. Lo cierto es, que casi todos los avances de la genética molecular se produjeron como resultado de los estudios de las bacterias y sus fagos. La relativa facilidad con la que se pueden elaborar mapas genéticos bacterianos contrasta con las enormes dificultades con las que se encontraron los primeros investigadores que pretendieron cartografiar el genoma humano. El genoma de la bacteria ya ha sido completamente secuenciado y el mapa genético de *E. coli* incluye en la actualidad más de mil genes.

Pero el más importante descubrimiento generado del estudio de bacterias y virus fue el fantástico arsenal de «enzimas para biología molecular», que constituyen las herramientas de trabajo del ingeniero genético: enzimas que cortan el ADN en sitios determinados y específicos (enzimas o endonucleasas de restricción), enzimas que lo vuelven a unir (ligasas), enzimas que fabrican millones de copias a partir de una molécula de ADN (polimerasas), enzimas que transcriben el ADN hasta ARN (transcriptasas) o que hacen el proceso inverso (retrotranscriptasas o transcriptasas inversas). En fin, para cada necesidad que pueda tener el biólogo molecular en su trabajo diario, siempre encuentra la herramienta adecuada, molecularmente precisa. Sólo es cuestión de decidir para qué utilizarlas. Estas enzimas constituyen la base de la ingeniería genética. Las endonucleasas de restricción han hecho posible que se puedan fabricar toneladas de insulina humana en reactores industriales que contienen bacterias, que las mareas negras puedan ser disueltas mediante bacterias que eliminan el petróleo y después se suicidan, que el SIDA/VIH sea una enfermedad cada vez más llevadera y sin duda alguna, harán posible que algún día se encuentre la cura contra muchos tipos de cáncer que afligen a los seres humanos. Estas enzimas, en fin, han hecho posible el desarrollo del Proyecto Genoma Humano y de todas sus consecuencias.

### **LAS ESPIRALES DE LA VIDA**

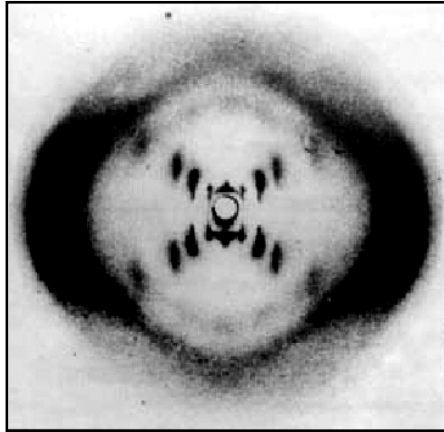
James Dewey Watson, quien fue un «niño prodigio» había ingresado en la Universidad de Chicago a la edad de quince años. Obtuvo su título en 1946 y se quedó un año más para tomar clases de zoología. En 1947 llegó a la Escuela de graduados de la Universidad de Indiana, donde se encontraba Herman Muller, quien había sido recientemente galardonado con el Nobel por su trabajo sobre las mutaciones inducidas por los rayos X. Watson fue influenciado por Delbrück y Luria, del Grupo de los Fagos. En este sentido, Luria le convenció para realizar un proyecto de investigación sobre los efectos de los rayos X sobre los fagos y en mayo de 1950, a la edad de 22 años, Watson completó su doctorado. Después de unas seis semanas en Cold Spring Harbor partió hacia Europa.

En este continente, Watson conoció en un congreso a Maurice Wilkins, quien estaba intentando discernir la estructura molecular del ADN en el King's College de Londres, utilizando la técnica de difracción de rayos X que había sido rutinariamente empleada desde principios de siglo para averiguar la estructura interna de los cristales de los compuestos inorgánicos.

Watson fue incapaz de conseguir una invitación para ir a trabajar con Wilkins. En cambio, gracias a los esfuerzos de Luria, consiguió unirse al grupo de Max Perutz en el laboratorio Cavendish, en la Universidad de Cambridge, cerca del King's College. El grupo de Perutz había conseguido recientemente discernir la primera estructura tridimensional de una proteína, la hemoglobina, mediante difracción de rayos X. Además, el laboratorio Cavendish estaba por aquella época dirigido por Sir Lawrence Bragg, uno de los inventores de la cristalografía de rayos X.

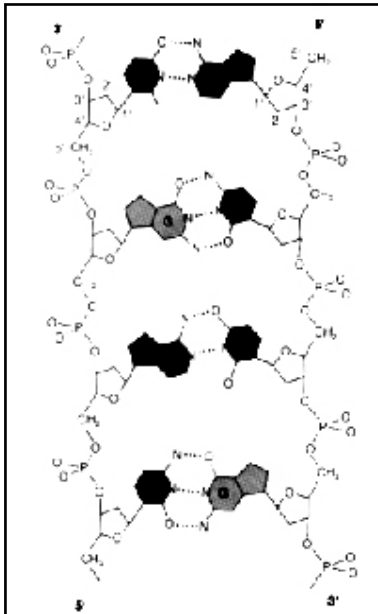
Durante su primer día en Cambridge, Watson conoció al físico británico Francis Crick y los dos se entendieron inmediatamente. En esa época, Francis Crick tenía 35 años y era aún un estudiante graduado. Pronto estuvieron de acuerdo en colaborar en base a que Crick uniese sus conocimientos sobre cristalografía de rayos X y de física a los conocimientos de Watson en genética. No sería justo, aunque numerosos historiadores de la ciencia caen en el error, contar la historia del descubrimiento de la doble hélice del ADN sin nombrar a Rosalind Franklin. Franklin estaba trabajando en el King's College con Maurice Wilkins; había realizado ya algunos análisis de difracción del ADN, pero había acordado con el director de la sección de biofísica del College, John Randall, emplear su tiempo en otro tipo de experimentos. Rosalind Franklin era una mujer de treinta años, brillante, analítica e independiente, que tuvo que sufrir la atmósfera de club masculino del King's College. Nunca congenió con Maurice Wilkins, que se empeñaba en tratarla como asistente y no como colega. En su best-seller de 1968, *«La doble hélice»*, Watson se refiere a Franklin como «Rosy» y se pregunta «cómo sería si se quitase las gafas e hiciese algo distinto con su cabello». Francis Crick, en su libro *«¡Qué loco propósito!»*, admite que en el College había restricciones irritantes para Rosalind, por ejemplo, no se le permitía tomar café en una de las salas reservada sólo para los hombres, pero considera que «éstas eran relativamente triviales o así lo parecían en la época».

Lo cierto es que Watson y Crick estaban impedidos para realizar experimentos con el ADN, por lo que dedicaban su tiempo a la realización de modelos teóricos de hojalata. Una mañana llegó un manuscrito de Linus Pauling desde Estados Unidos, en el cual detallaba sus conclusiones acerca de la estructura del ADN a su hijo Peter, que por casualidades del destino compartía una pequeña oficina con Crick y Watson. Pauling, posiblemente el químico más eminente de su época, representaba un duro competidor para los dos jóvenes de Cambridge que ni siquiera tenían acceso al laboratorio para hacerle pruebas al ADN. Watson y Crick leyeron el manuscrito y Watson corrió al King's College con las noticias de Pauling. Allí se encontró en un pasillo a Rosalind Franklin, quien se enojó, porque el manuscrito que Watson llevaba en la mano contenía información que ella hacía un mes había solicitado infructuosamente al laboratorio de Pauling. Wilkins llegó en ese momento y prácticamente empujó a Watson a su despacho. Wilkins, eufórico, mostró a Watson un par de excelentes fotografías de difracción que Rosalind Franklin había tomado al ADN (figura 3.3). Watson escribiría en su libro «En el instante en que vi la imagen, mi boca se abrió y mi pulso comenzó a acelerarse.»



**Figura 3.3.** Patrón de difracción de Rayos X de una fibra de ADN de la bacteria *E. coli* obtenida por Rosalind Franklin. Tomada de archivo público de Internet.

Franklin había deducido, mediante cálculos precisos, que las bases nitrogenadas que entraban a formar parte de la composición del ADN debían estar hacia adentro de una estructura helicoidal, con el esqueleto de azúcar-fosfato en su exterior. Además, había calculado varios parámetros de la hélice, como la distancia o período de repetición. Con estos nuevos datos, Watson volvió a su laboratorio y entre él y Crick, en un mes llegaron a su modelo teórico para la estructura del ADN «Era demasiado hermoso para no ser verdadero» (figura 3.4).



**Figura 3.4.** Modelo de la doble hélice de la molécula de ADN propuesto por James Watson y Francis Crick en 1953. De acuerdo con el modelo, el esqueleto azúcar-fosfato de las dos cadenas polinucleotídicas soporta las bases nitrogenadas las cuales son complementarias en cada cadena. Así la Adenina de una cadena se aparea mediante dos enlaces de hidrógeno con la Timina de la banda complementaria. Lo mismo es válido para el apareamiento entre la Guanina y la Citosina que se realiza mediante tres enlaces de hidrógeno.

La doble hélice de Watson y Crick permitía atar diversos cabos en la estructura y funcionamiento de los genes. El hecho de que cada base nitrogenada se aparee específicamente con otra base de la cadena opuesta explica satisfactoriamente la replicación del ADN. En efecto, durante la replicación de los genes previa a la división celular, la doble hélice se abre en sus dos hebras individuales, cada una de las cuales contiene exactamente la misma información. Entonces, la enzima ADN-polimerasa es capaz de sintetizar, tomando como molde las dos hebras separadas, otras dos hebras exactamente complementarias a las iniciales. Se explicaba así la aparición de los dos juegos de cromosomas durante la mitosis, cada uno de ellos conteniendo prácticamente la misma información, salvo los errores de copiado (mutaciones) y la herencia de los caracteres de los padres a los hijos.

El modelo de la doble hélice del ADN de Watson y Crick ha sido, quizás merecidamente, el hallazgo científico más profusamente representado en los más diversos materiales y medios, como emblema de los logros de la ciencia del siglo XX. A la entrada de la Facultad de Ciencias de la Universidad de Upsala, en Suecia, se encuentra, en bronce, un modelo algo extraño. Representa el ADN con el esqueleto de azúcar-fosfato en el interior y las bases nitrogenadas como protuberancias dirigidas hacia el exterior de la molécula. Se trata del modelo que Linus Pauling describía a su hijo en el famoso manuscrito que Watson enseñó a Wilkins. Pauling había ideado un modelo de ADN completamente equivocado, exactamente al contrario de cómo es la molécula en realidad. No hay que sentir lástima de Pauling por haber perdido su oportunidad de recibir el premio Nóbel de medicina y fisiología; el químico de Oregón fue galardonado a lo largo de su vida con otros dos: el de química, por sus estudios sobre la naturaleza del enlace químico y el de la paz, por haber promovido y participado activamente en una asociación de científicos en contra de las armas nucleares.

Rosalind Franklin no sería elegida para compartir la gloria de Watson, Crick y Wilkins, quienes recibieron el premio Nóbel en 1962 (Watson contaba 34 años), ni siquiera de forma póstuma. La voluntariosa investigadora murió de cáncer en 1958, a la edad de 37 años. Según Watson, «continuó trabajando al más alto nivel hasta pocas semanas antes de su muerte».

## **LA AGENDA DE TRABAJO HACIA EL PROYECTO GENOMA HUMANO**

*“Los cuerpos pueden ser colonias de genes, pero en cuanto a su comportamiento se refiere, han adquirido indudablemente, una individualidad propia. Un animal se mueve como un conjunto coordinado, una unidad.”*

**RICHARD DAWKINS.**

**“The Selfish Gene” (El gen egoista). 1982.**

### **LOS CIENTÍFICOS DISCUTEN**

Los primeros pasos del Proyecto Genoma Humano se dieron en los Estados Unidos de América. Por esta razón, conviene describir con cierto detalle la génesis del PGH en dicho país. Sin embargo, el primer acto de esta historia se va a desarrollar fuera de este País, de hecho, tampoco va a tener lugar en la vieja Europa. Nos iremos por un momento algo más lejos, hasta una ciudad tristemente célebre, cuyos habitantes seguramente hubieran preferido no haber llegado nunca a ser tan conocidos. Se trata de Hiroshima en el Japón.

Cuarenta años más tarde de que 71.000 personas murieran en unos instantes, en marzo de 1984, se llevó a cabo una conferencia científica en la ahora reconstruida Hiroshima, cuyo eje central eran los efectos que la radiación es capaz de causar sobre la genética humana. En la conferencia, los participantes destacaron la necesidad de usar herramientas analíticas moleculares del ADN para poder detectar directamente las mutaciones que los descendientes de las víctimas de la masacre pudiesen haber heredado.

El interés por estas cuestiones planteadas en la conferencia condujo a muchos de los estadounidenses allí presentes, posiblemente sintiéndose culpables de parte del mal causado, a organizar otra conferencia sobre el mismo tema nueve meses después, en Alta, Utah. Esta

vez, la conferencia estaba apadrinada por el Departamento de Energía (DOE) de los Estados Unidos y por la Comisión Internacional para la Protección contra Mutágenos y Carcinógenos Ambientales. Los 19 científicos expertos en biología molecular que participaron, llegaron a la conclusión de que sólo un esfuerzo conjunto, específicamente diseñado para mejorar la tecnología en varios órdenes de magnitud, podría servir para detectar las mutaciones de las víctimas de Hiroshima, estimadas en más de 30 mutaciones por genoma individual y por generación. Las ideas e interacciones generadas por la reunión de investigadores, muchos de los cuales jamás se habían visto antes de la conferencia, finalmente se aglutinarían en el Proyecto Genoma Humano.

A su vuelta a Washington desde Utah, Michael Gough redactó un informe para la Oficina de Análisis de Tecnologías (OTA) encargada de asesorar al Congreso sobre la conveniencia de financiar o no los futuros avances científicos. Charles DeLisi, que era en ese entonces director de la Oficina de Investigación Sanitaria y Ambiental (OHER) leyó un borrador inicial del trabajo de Gough, que apareció en 1986. Mientras lo leía, DeLisi tuvo por primera vez la idea de un Proyecto Genoma Humano, como puede leerse en su artículo, publicado en el número de septiembre/octubre de 1988 de *American Scientist*. En ese artículo, DeLisi señaló que las propuestas de secuenciar el genoma humano se remontaban a 1985, en que, durante un taller sobre el tema organizado por Robert Sinsheimer, biólogo molecular que era rector de la Universidad del Sur de California en Santa Cruz; la idea de Sinsheimer había surgido a partir de un proyecto millonario fallido para la construcción de un gigantesco telescopio astronómico, el cual jamás se llegó a construir. Sinsheimer se encontró entonces con un superávit de 36 millones de dólares y se le ocurrió la idea de construir un Instituto para la Secuenciación del Genoma Humano en su universidad.

En marzo de 1986, el DOE organizó una importante reunión científica en Santa Fe, Nuevo México, para discutir las ideas de DeLisi. Los participantes respaldaron con entusiasmo la idea de secuenciar el genoma humano y el DOE aceptó la iniciativa. La discusión en Santa Fe giró en torno a cómo podría lograrse la secuenciación; no, si debía hacerse. El premio Nobel Walter Gilbert propuso la creación de un Instituto del Genoma Humano dedicado a la secuenciación a gran escala. Sugirió que los esfuerzos iniciales se dedicasen a secuenciar genes conocidos en regiones de importancia. La mayoría, sin embargo, opinaba que aunque se necesitaba alguna forma de control central, el trabajo debía ser distribuido en los laboratorios de toda la nación y quizás de forma internacional. Esto estimularía el surgimiento de un amplio espectro de soluciones creativas para la obvia necesidad de producir mejoras importantes en la tecnología de secuenciamiento.

La biología molecular nunca había sido el campo principal de los programas de la DOE. DeLisi quería cambiar eso. La reunión de Santa Fe excedió sus expectativas, culminando con la impresión de que “el objetivo era meritorio, obtenible y sería un logro sobresaliente de la biología moderna”. Sin embargo, el PGH también tenía sus detractores; en 1990, en un artículo escrito por Walter Davis, genetista molecular de Harvard, y aprobado por 22 miembros del departamento del autor, se emitió un juicio realmente duro: “El PGH fue formulado por un administrador políticamente astuto en la DOE, convencido de que las poderosas herramientas

de la biología molecular hacían que fuese apropiado introducir la ‘gran ciencia’, administrada centralizadamente, en la investigación biomédica. La idea desarrolló rápidamente un fuerte atractivo político. Davis prosigue sugiriendo la necesidad de reevaluar el proyecto completo. Robert Weinberg, del Instituto Whitehead, declaró en 1986: “Me sorprende que adultos responsables se presten a hablar de esto en público. No tiene sentido”. James Walsh, de la Universidad de Arizona, escribió en la revista Nature: “Secuenciar el genoma humano resultaría más o menos tan útil como traducir las obras completas de Shakespeare a la escritura cuneiforme, aunque no tan sencillo ni tan fácil de interpretar”.

En mayo de 1986, tres meses después del eufórico congreso de Santa Fe, se celebró otro muy diferente en Cold Spring Harbor, sobre “La biología molecular del *Homo sapiens*”. James Watson había sido director del laboratorio desde 1968, despertándolo de un cierto período de latencia en el que se había visto sumido tras los éxitos iniciales del grupo de los Fagos. El conjunto de biólogos moleculares reunidos en Cold Spring Harbor, era bien distinto del reunido en Santa Fe bajo los auspicios del DOE. La mayoría de los presentes poseía una amplia tradición de trabajo en las ciencias biomédicas, bajo el patrocinio de los Institutos Nacionales de la Salud (NIH). Muchos estaban en contra de que el DOE, con su tradición de proyectos millonarios en el campo de las ciencias físicas invadiera su terreno, precisamente con un proyecto que representaba la culminación de la ciencia de la biología molecular. A pesar de que algunos, como Gilbert, preferían dejar el proyecto al DOE para que no interfiriera con los siempre ajustados presupuestos de los NIH, la mayoría estaban de acuerdo con Watson en que los NIH debían tomar parte activa en el proyecto, siempre que el Congreso asignase fondos especiales para no entrar en conflicto con los presupuestos de las investigaciones que estaban en marcha en los NIH.

El bioquímico y premio Nobel Paul Berg, de Stanford, trató de que los participantes dejasen de lado la cuestión económica y tuvieran en cuenta el valor intrínseco de un proyecto de secuenciamiento del genoma. Afirmó “¿Vale su costo, no en términos de dólares, sino en términos de su impacto sobre el resto de la ciencia biológica?”. A pesar del consejo, muchos investigadores jóvenes reaccionaron con hostilidad directa. Creían que un proyecto megamillonario inevitablemente desviaría el dinero de los subsidios de investigación a investigadores aislados, retardando el ritmo de la investigación biológica y médica de alta calidad, preocupación que años más tarde se tornó realidad. En el campo de la biología no existía tradición de grandes proyectos, como los que emprendían normalmente los físicos y los astrónomos del DOE. Watson y Crick habían desentrañado la estructura del ADN en un minúsculo despacho de Cambridge y, a finales de los cincuenta, Crick había tenido que trabajar en un cobertizo para bicicletas habilitado en el patio trasero del laboratorio.

A medida que crecía la controversia, surgió otro punto de vista que acabaría dominando la evolución del proyecto. Maxine Singer, del Instituto Nacional del Cáncer, había insistido en que realizar solamente el secuenciamiento no tenía sentido; podía obtenerse información mucho más útil si se llevaban otras líneas paralelas de investigación, como la genética y la



bioquímica, que eran el terreno natural de los científicos de los NIH. “No sabremos si valía la pena como proyecto hasta que no lo hayamos hecho”, dijo, “pero sabemos que cuando hacemos secuenciamiento y bioquímica juntos aprendemos mucho”.

El paso decisivo lo dió el Consejo Nacional de Investigación (NRC), posiblemente el órgano que puede erigirse en portavoz de la comunidad científica estadounidense con más autoridad que ningún otro. En Agosto de 1986, el Consejo convocó una reunión en Woods Hole, Massachusetts, para debatir específicamente la posibilidad de llevar a cabo el proyecto y qué agencia debía encargarse de dirigirlo. Según James Watson, “poco después de comenzadas las deliberaciones del NRC, se hizo evidente que, en la sala de reunión, el proyecto no era objetado en sí mismo, ¿quién podía estar en contra de lograr los mapas moleculares físicos y genéticos del ADN humano de mayor resolución, que serían necesarios para comenzar?”. El NRC respaldaba los anteriores cálculos sobre el coste de la empresa, sugiriendo un presupuesto anual de 2000 millones de dólares durante 15 años. Insistía, además, en la importancia de estudiar los genomas de otros organismos, además del humano, para poder interpretar biológicamente los datos de este último. El comité del Consejo declaró lo siguiente: *“Para obtener los mayores beneficios de una secuencia del genoma humano, será necesario disponer de una amplísima base de datos sobre las secuencias de ADN del ratón (cuyo genoma es del mismo tamaño que el humano) y de organismos más simples, con genomas mucho más pequeños, como las bacterias, las levaduras, la Drosophila melanogaster y el Caenorhabditis elegans (un gusano nematodo)... Así pues, para que este proyecto tenga éxito, no debe limitarse al genoma humano, sino que debe incluir un análisis intensivo de secuencias de genomas de otras especies escogidas.”* Pero dicho análisis significaba que el trabajo y el presupuesto se duplicarían con creces. Watson se atribuyó el mérito de la ampliación del proyecto: “Esta ha sido mi contribución más importante al proyecto en conjunto”, declaró al New Scientist.

### **¿QUIÉN LE PONE EL CASCABEL AL GATO?**

¿Quién estaría a cargo de un proyecto de tal magnitud y de futuro tan incierto? El comité sugirió que una sola agencia federal estuviese a cargo de organizar y coordinar el esfuerzo nacional en el estudio de los genomas. No identificaron ninguna agencia particular pero en privado, sin embargo, muchos creían que la agencia rectora debía ser la NIH, debido a su registro de investigaciones biomédicas intra y extramuros”. Ciertamente, el entusiasmo inicial por el secuenciamiento a escala completa del genoma humano, había evolucionado hasta convertirse en la meta, más modesta y realista, del cartografiado cromosómico. La secuencia completa de los más de 3.000 millones de bases se había convertido en una “meta subsidiaria”. En el plan quinquenal del DOE/NIH, fechado en Febrero de 1990, se lee: “La comprensión de nuestra herencia genética, el Proyecto Genoma Humano, los primeros cinco años: 1991-1995”. Las metas a realizar en dicho período de tiempo eran las siguientes:

1. Realizar un mapa genético humano completamente conectado con marcadores espaciados en promedio de dos a cinco centimorgans. Identificar cada marcador por un STS (“sequenced tagged sites” o lugar etiquetado mediante secuencia).
2. Reunir mapas STS de todos los cromosomas humanos con la meta de tener marcadores espaciados a intervalos de aproximadamente cien mil pares de bases.
3. Generar conjuntos superpuestos de ADN clonado o marcadores cercanamente espaciados y ordenados no ambiguamente con continuidad, sobre longitudes de dos millones de pares de bases para partes grandes del genoma humano.

El centimorgan (cM) es la unidad de distancia genética para mapas de ligamiento, llamada así en honor a Thomas Hunt Morgan. Los cromosomas humanos poseen unas longitudes comprendidas entre 58 y 293 cM, excepto el pequeño cromosoma Y, cuya longitud no puede medirse en cM, ya que posee muy pocos genes que tomar como referencia. No existe una equivalencia exacta entre la distancia genética en cM y el número de pares de bases, pero por término medio, 1 cM equivale a un millón de pares de bases (una Megabase, o 1 Mb).

Los STSs (abreviatura de “Sequence Tagged Site”), que se puede traducir como “lugar etiquetado mediante secuencia”, consisten en secuencias cortas de 200 a 500 pares de bases, elegidas de forma que aparezcan una única vez en todo el genoma y que pueden ser tomadas como referencia para fijar la posición de un determinado gen. Otras entidades relacionadas son los ESTs (Expressed Sequence Tags) o “lugares etiquetados expresados”; equivalen a los STS, pero en lugar de ser fragmentos del genoma, corresponden a fragmentos del ARN mensajero, es decir, a secuencias concretas de proteínas que se han expresado en un determinado tejido.

Los informes del NRC y de la OTA se publicaron en 1988. En Febrero del mismo año, el entonces director de los NIH, James Wyngaarden, convocó otra reunión en Reston, Virginia. La reunión de Reston transformó la actitud de los NIH y estableció objetivos concretos y detallados para el programa. Wyngaarden propuso a Watson asumir la responsabilidad de dirigir la investigación del genoma en el NIH, cargo que Watson aceptó encantado. Su nombramiento, según Norton Zinder, supuso “un salto cuantitativo en la credibilidad del programa”. El PGH había emprendido su marcha en Estados Unidos y con los NIH a la cabeza, en lugar del DOE.

### **EL CONSORCIO HUGO**

Otros proyectos de secuenciar el genoma humano fueron emprendidos en diferentes países. Principalmente Gran Bretaña, Francia y Japón. Con sus recursos limitados, los científicos europeos no aspiraban a competir con sus colegas norteamericanos. En lugar de ello, centraron sus objetivos en el estudio de los ADNc, obtenidos a partir de los ARNs mensajeros expresados en distintos tejidos celulares. Además, los británicos decidieron secuenciar también el genoma completo del *Caenorhabditis elegans* un gusano nemátodo de vida libre. Hace más de 20 años, el británico de origen sudafricano Sydney Brenner sugirió que este pequeño

gusano (1 mm de longitud) sería el organismo más apropiado para estudiar el desarrollo del huevo hasta transformarse en animal adulto, ya que su cuerpo posee únicamente 945 células, casi completamente transparentes.

Francia tiene dos instituciones que están compitiendo fuertemente con el programa británico. En 1984 se fundó el Centro de Estudios del Polimorfismo Humano (CEPH), dedicado al estudio de las enfermedades genéticas que estaba financiado en gran parte por asociaciones de familiares de enfermos. Casi todo el mapa genético humano que existe en la actualidad se ha construido utilizando el material proporcionado por el CEPH, que almacena material de más de sesenta familias de Francia, Pennsylvania, Utah y Venezuela. Otro importante centro francés es el Genethon, inaugurado en 1990 en Evry (París) y que ha colaborado intensamente en la realización de los mapas genéticos. En el resto de Europa, Dinamarca, Italia y Alemania han emprendido programas propios de estudio del genoma, aunque nunca tan bien financiados y coordinados como el británico y el francés.

En Japón una multitud de ministerios, agencias gubernamentales y empresas privadas participan en la financiación, como el Ministerio de Educación, Ciencia y Cultura, la Agencia de Ciencia y Tecnología, el Ministerio de Sanidad y Bienestar y el Ministerio de Industria y Comercio Internacional. No siempre está claro si existe alguna coordinación entre ellos. En Japón existe un considerable interés por el genoma de la planta del arroz, que se investiga con fondos aportados por el Ministerio de Agricultura, Silvicultura y Pesca.

Cuando comenzó a crecer el interés internacional por el proyecto genoma, se hizo evidente la necesidad de un foro internacional. En 1988, durante una reunión celebrada en Cold Spring Harbor, los investigadores decidieron formar la Organización del Genoma Humano (HUGO), encargada de coordinar los trabajos internacionales, procurando evitar las repeticiones y solapamientos. Su primer director fue el genetista norteamericano Victor McCusick, al que sucedió el británico Walter Bodmer. Su sede oficial se encuentra en Ginebra, HUGO es la Organización de las Naciones Unidas para el estudio del genoma. Nunca una comparación fue tan acertada, ya que la HUGO, al no disponer de fondos propios para financiar la investigación, se ha topado con dificultades de aceptación, pareciendo estar condenada a la impotencia de dar consejos que nadie se siente obligado a aceptar.

### **GENOMA, GENES Y DINERO**

El PGH, siguió su curso. Sin embargo, tras unos comienzos tormentosos que finalizaron en un consenso mayoritario, la continuación del proyecto no fue precisamente un camino fácil. A mediados de 1991, los planes de cooperación internacional sufrieron un nuevo y duro golpe, al anunciarse que los NIH pretendían patentar 337 secuencias de ADN descubiertas por sus investigadores (principalmente por Craig T. Venter). Para agravar la situación, las secuencias que se pretendían patentar correspondían a trozos de genes que se manifestaban

en las células del cerebro, de modo que, de aceptarse la solicitud, el gobierno de Estados Unidos no sólo sería propietario de parte del genoma humano, sino que también habría logrado patentar parte del cerebro humano. Para colmo, nadie sabía a qué genes correspondían las secuencias ni cuál era su función. Correspondían a ESTs aisladas por retrotranscripción del ARN expresado en las células cerebrales.

El tema provocó una gran controversia en todo el Mundo, sobre todo cuando se supo que los NIH habían actuado sin contar con James Watson, director del Centro Nacional para la Investigación del Genoma Humano. Watson se opuso rotundamente a la idea, pues comprendió que dificultaría el libre flujo de la información científica y la colaboración internacional, retrasando así el progreso del PGH. Los franceses consideraron que los genes humanos no debían ser patentables. Los británicos, como siempre, buscaron un compromiso; su propuesta era que las secuencias génicas sin utilidad conocida, como las que Venter pretendía patentar, no debían poder patentarse, pero que las secuencias con función conocida (por ejemplo, el gen mutante de la fibrosis quística) sí podían ser susceptibles de propiedad intelectual. El gobierno británico añadía una cláusula de advertencia: aunque se oponía a la idea de patentar secuencias sin utilidad conocida, patentaría todas las que tenía si los norteamericanos persistían en su actitud. Al final, parece ser que la propuesta británica ha sentado las bases de la política de patentes en todo el Mundo. Craig T. Venter no pudo patentar sus ESTs cerebrales y dejó los NIH para fundar una empresa privada, *Celera Inc.* que ha contribuido activamente desde entonces a la secuenciación de genomas bacterianos comercialmente explotables. Además firmó una alianza comercial con la multinacional farmacéutica SmithKline-Beecham. En la actualidad poseen más de 170.000 ESTs de diversos tejidos humanos, cuya secuencia guardan celosamente. Esta curiosa situación, provocó que la multinacional Merck entrara en el juego, comenzando a secuenciar ESTs a gran escala. Sin embargo, a diferencia del tándem TIGR - SmithKline-Beecham, Merck ofrece gratuitamente su información; ya ha puesto más de 350.000 secuencias de ESTs a disposición del público. Finalmente *Celera Inc.* en el 2001 y a la par con el Instituto del Genoma Humano (GHI), logró secuenciar casi todo el genoma humano, gracias a una estrategia novedosa de producir clones con secuencias humanas producto de restricción con múltiples endonucleasas de restricción, estrategia conocida como tiro de escopeta (“shot gun”).

Otro obstáculo en el desarrollo del PGH ocurrió en 1992, cuando una empresa privada estadounidense intentó contratar a John Sulston y Bob Waterston, los directores de los proyectos de secuenciación del *Caenorhabditis* en Cambridge y Saint Louis. Ambos habían desarrollado nuevos métodos automatizados de secuenciación y tratamiento informatizado de los datos, investigaciones que habían sido financiadas con dinero público. La empresa pretendía que Sulston y Waterston aportaran toda su experiencia para ofrecer un servicio contratado de secuenciación. Cuando se emprendiera la tarea de la secuenciación a gran escala, los laboratorios se encontrarían con que las patentes de los métodos de secuenciación pertenecerían a la empresa y los gobiernos tendrían que pagar regalías por algo cuya idea

había sido desarrollada previamente, cuando ambos investigadores trabajaban con dinero público. De nuevo, Watson se opuso enérgicamente, esta vez con éxito. Tanto Sulston como Waterston permanecieron en sus puestos.

Estos incidentes enfrentaron a Watson con diversos sectores de la burocracia norteamericana y en Abril de 1992, el NIH emprendió una investigación ética sobre la supuesta incompatibilidad entre el cargo de Watson como director del Centro Nacional para la Investigación del Genoma Humano y su participación en empresas privadas de Biotecnología. En realidad, nada hacía sospechar que Watson hubiese aprovechado su posición de poder en beneficio propio, por el contrario, la opinión generalizada es que fue víctima de un choque de personalidades y de política interna de los NIH. El mismo mes en que se abrió la investigación, Watson anunció que el desempeño simultáneo de los dos cargos, como director del Centro Nacional para la Investigación del Genoma y del laboratorio de Cold Spring Harbor se había convertido en una carga muy pesada, renunció a la dirección del primero y volvió a su viejo despacho de Cold Spring Harbor, en donde es director desde 1968. Los NIH se apresuraron a nombrar un director interino entre los miembros de su personal, hasta que Francis Collins, el codescubridor del gen de la fibrosis quística, aceptó el cargo.

En la actualidad, el Proyecto Genoma Humano sigue su curso. En 1996 se completó el mapa de marcadores genéticos, a partir del cuál se puede comenzar la secuenciación masiva de todo el genoma; así mismo, los avances y bajos costos de las tecnologías de secuenciación y de procesamiento de datos permiten pronosticar que era posible completar la secuenciación de los más de 3.000 millones de pares de bases para el 2003. Sin embargo a principios del año 2002 tanto el NIH como Celera Inc. cuyo presidente es Craig Venter, anunciaron conjuntamente haber secuenciado el 99% del genoma humano.

Las empresas privadas se han mostrado cada vez más interesadas por los resultados del proyecto y el negocio de la secuenciación. Esto permite que la financiación del PGH esté asegurada, hasta su cabal término. Pero la participación, de quizá cientos, de empresas comerciales creará un terreno abonado para los conflictos de intereses en lo referente al desarrollo de tecnologías y en la propiedad de los resultados. El sucesor de Watson no sólo tendrá que enfrentarse a la ya de por sí difícil tarea de dirigir un nutrido colectivo científico repartido por todo el Mundo, sino que tendrá que desarrollar un nuevo sistema de relaciones con la empresa privada.

### ***COMPARAR ES ÚTIL PUES NOS ACLARA MUCHO.***

Parece que resulta inevitable, para cualquiera que escriba acerca del Proyecto Genoma Humano, realizar comparaciones del mismo con otros proyectos de la “gran ciencia” americana, representada principalmente por el DOE y la NASA. Aunque reconozcamos que el título que abre este apartado representa una gran verdad, en realidad, debemos ser escépticos a las comparaciones que se han hecho. Básicamente, los detractores del PGH tratan de localizar

los puntos en común del mismo con otros proyectos multimillonarios de dudosa fama, el más importante es, sin duda, el Proyecto Manhattan, que permitió en los primeros años cuarenta el desarrollo de la bomba atómica. Sin embargo, en la mayoría de los casos, el PGH sale bien librado del trance en el que se le ve sometido, ya que es, básicamente, más barato, más científico y con mayores beneficios en potencia que ninguno de los grandes proyectos hasta ahora llevados a cabo por la humanidad.

Irónicamente el PGH tiene bastantes puntos de contacto, al menos en sus orígenes, con el Proyecto Manhattan. Cabe esperar que las reuniones de Hiroshima y Alta nunca se hubieran llevado a cabo si los Estados Unidos jamás hubieran tirado dos bombas atómicas en 1945 sobre dos ciudades japonesas. Los dos centros del DOE que más directamente han estado implicados en el PGH, los laboratorios de Los Álamos y de Oak Ridge, fueron inicialmente financiados por el secretísimo Proyecto Manhattan. Otra coincidencia, cuando James Watson era un precoz estudiante en la Universidad de Chicago, en 1942, dicha universidad era la sede de la primera pila atómica del mundo que, bajo la dirección de Enrico Fermi, había sido instalada bajo las gradas de un estadio de fútbol inutilizado, de forma tan secreta que el propio rector de la universidad lo ignoraba. Entre el grupo de físicos que estaban presentes el día que se puso en marcha la pila, se encontraba el húngaro Leo Szilard, quien cinco años después se uniría al Grupo de los Fagos de Cold Spring Harbor, colaborando a fundar las bases de la biología molecular.

Más caústica es la visión de Robert Wright, publicada en *The New Republic*: “A mediados de los ochenta, la misión ostensible del DOE de fabricar armas nucleares y confeccionar la política energética se estaba hundiendo en la agenda de la nación. Alguien tuvo la idea de pedir al Congreso montones de dinero para hacer el Proyecto Genoma Humano”.

El Proyecto Manhattan, el Proyecto Apolo, el Superacelerador Superconductor (SSC) o el Proyecto de Estación Espacial Alfa de la NASA, todos han sido, en tanto que pertenecientes a la gran ciencia, usados como parangones del PGH. Algunos de los iniciadores y líderes del PGH pensaron que ligarlo a estas empresas masivas, fuertemente financiadas por el gobierno estadounidense y apoyadas por el contribuyente como chauvinistas triunfos de la técnica estadounidense y su liderazgo internacional, contribuiría a aumentar la aceptación del PGH y dotarlo de un apoyo económico continuado, particularmente por el Congreso. Sus detractores usan las analogías como demostraciones de que el PGH es excesivamente costoso y con gran parte de contenido “maléfico”, denominándolo “superciencia supercara” e “investigación rococó” (*New York Times*) y recordando que el Proyecto Manhattan, el Proyecto Apolo, el SSC y la Estación Espacial fueron inicialmente diseñados, en grado variable, con propósitos militares.

Algunos escapan a las contradicciones inherentes a las analogías negando su validez. Renato Dulbecco sostiene que el rótulo de “gran ciencia” es inapropiado para el PGH. Charles Cantor argumenta que a diferencia de otros tipos de gran ciencia, el PGH tiene el éxito garantizado en sentido amplio no es investigación básica por el contrario es ingeniería para

proveer una herramienta que será usada por las personas que hagan investigación básica. En su ponencia en la conferencia Genoma Humano II, James Watson rechazó las críticas contra el PGH diciendo “bueno, el genoma humano es grande... nunca pensé que ser grande fuese necesariamente malo”, continuó asegurando que “no estamos pensando construir una ciudad ADN secreta en algún lugar”. Después, en un tono más serio, declaró: “queremos que sea lo suficientemente grande como para que el costo pueda ser bajado a un nivel razonable y queremos hacerlo en un período razonable de tiempo... no debiéramos estar peleando entre nosotros”. Si comparamos costos de inversión, el PGH es ciencia pequeña en comparación con otros proyectos, es ciencia grande para los biólogos que son empresarios infantiles. El PGH, con su costo total estimado de 3.000 millones de dólares en quince años, se queda minúsculo al compararlo con el Superacelerador Superconductor, un túnel de casi 90 Km de largo que se planea construir en el desierto de Texas para investigar las partículas elementales a un costo de 8.000 millones. Con la Estación Espacial Alfa que costará unos 40.000 millones para su construcción y quizás otros 80.000 millones para que sea operativa. Sin embargo una gran diferencia entre estos proyectos y el del genoma humano es que si el SSC y la Estación Alfa fuesen interrumpidos a mitad de camino, no tendríamos esencialmente ninguna información, por el contrario si se interrumpiese el PGH a la mitad todavía nos quedaría una gran cantidad de información. En este sentido el PGH ya ha empezado a dar sus frutos, como veremos en los capítulos siguientes.

Si comparamos el PGH con otros proyectos menos conocidos, como los estrictamente militares, nos quedamos boquiabiertos. Construir un submarino Trident cuesta 1.400 millones de dólares. El programa del bombardero Stealth está estimado en 68.000 millones. La financiación de la Iniciativa de Defensa Estratégica entre los años fiscales de 1988 y 1992 fue de 38.000 millones. Varios premios Nóbel argumentan que seis meses de los gastos en el mantenimiento de cabezas nucleares en los Estados Unidos serían suficientes para financiar los quince años proyectados del PGH, cuyo propósito es comprender más profundamente la propia vida.

Hemos relatado, desde el mismo comienzo, una larga historia. La historia de un siglo y medio de adelantos que nos han conducido a estar más cerca que nunca de comprendernos a nosotros mismos. El potencial para aprender acerca del funcionamiento de los sistemas vivientes, unido a las posibilidades inmediatas de diagnosticar y quizás de prevenir y curar las enfermedades genéticas, hace impracticable cualquier idea de retroceder. Debemos mirar hacia adelante y pensar en las implicaciones futuras. Antes de eso, quizás convendría echar una mirada a nuestros conocimientos presentes.

## **APRENDIENDO A ENTENDER EL GENOMA**

*“He escrito la vida como un acuerdo firmado entre ácidos nucleicos y proteínas. Sin embargo, estas moléculas habitan en dominios químicos muy diferentes; en realidad apenas se hablan”*

**PAUL DAVIES. “THE FIFTH MIRACLE”  
(El quinto milagro).1999.**

### ***EL MENSAJE DE LOS GENE SE DESCIFRA: EL CÓDIGO GENÉTICO.***

A comienzos de la década de los 50 del siglo XX no estaba, ni mucho menos, claro cómo el control genético aparecía de algún modo involucrado en la producción de proteínas. Ya se había demostrado que el principio transformador y portador básico de la herencia era el ADN, pero su funcionamiento era aún muy oscuro. Incluso, en 1953, Alfred Hershey, que había corroborado en los fagos y en *E. coli* los experimentos de Avery, McLeod y McCarty, sostenía que el ADN no era un determinante único de la especificidad genética.

Casi todas las evidencias indicaban que el citoplasma, la zona de la célula fuera del núcleo, era el lugar donde se sintetizaban las proteínas. Esto requería la existencia de un intermediario entre el ADN, que sólo podía localizarse en el núcleo y los minúsculos orgánulos donde se fabrican las proteínas, denominados «ribosomas». Un candidato importante para este «mensajero» era otro ácido nucléico, el ARN, presente en cantidades muy bajas en las células.

Durante los meses en que estuvo preocupado por el modelo del ADN, Watson había pegado en la pared un folio donde había escrito la frase: «ADN a ARN a proteína». Crick se



dirigió en 1957 al Simposio de la Sociedad de Biología Experimental, que ese año se reunía en Cambridge, llamando a la frase escrita en el papel de su laboratorio el «Dogma Central». Según Horace Judson en su libro «El octavo día de la creación», la elección de la palabra «Dogma» se basaba en un ligero desconocimiento de Crick del significado del término, cuyo sentido real es «una idea para la cuál no existe demostración razonable». Crick le dijo a Judson: «Yo no sabía qué significaba exactamente dogma... era sólo un término para atraer la atención». En el mismo simposio, Crick hizo pública la «hipótesis de la secuencia», según la cuál el orden en que aparecen los aminoácidos en cada proteína está determinado por un código simple, contenido en la secuencia de ADN del gen que «codifica» dicha proteína. En los dos años siguientes a la propuesta de Crick, reinó la confusión en los intentos por definir los detalles del código genético; es decir, el lenguaje en el que están escritas las instrucciones para determinar la secuencia de las proteínas. El lenguaje del ADN estaba compuesto sólo por cuatro letras, cada uno de los cuatro nucleótidos, que debían traducirse a un lenguaje de veinte letras distintas, los veinte distintos tipos de aminoácidos que forman parte de las proteínas. Era pues evidente que, por combinaciones de dos de las cuatro letras, sólo se podían formar 16 palabras distintas, insuficientes para definir todos los aminoácidos. Con tres letras, en cambio, existían 64 combinaciones, suficientes para codificar los diferentes aminoácidos. Pero, ¿funcionaba así la naturaleza? y ¿qué tripletes específicos de nucleótidos en el ADN codificaban cada aminoácido?

La elucidación del código genético debe gran parte de su éxito inusitadamente rápido a Marshall Warren Nirenberg y Johann Mathei, que trabajaban en los Institutos Nacionales de la Salud (NIH) cerca de Washington D.C. Utilizando un sistema de síntesis protéica libre de células, que había sido recientemente publicado por Paul Zamecnik, descubrieron que cuando se añadía a su sistema ribonucleasa, una enzima que destruye el ARN pero no el ADN, la síntesis de proteínas se detenía inmediatamente; por el contrario, la adición de una desoxirribonucleasa, que destruía sólo el ADN, la síntesis de proteínas era normal hasta después de unos 30 minutos. Estos resultados confirmaban experimentalmente el «Dogma Central», en que el ARN tenía el papel de mensajero entre el ADN y las proteínas era una propuesta correcta. El 22 de Mayo de 1961, Mathei decidió probar un ARN que sólo contenía uracilo en la secuencia (un poliU), en su sistema libre de células. Comprobó que la proteína que se fabricaba entonces estaba compuesta únicamente de unidades de fenilalanina. Este hecho inició el desciframiento del código genético y el primer triplete de nucleótidos UUU era la señal para añadir una fenilalanina a una cadena de proteínas en crecimiento. Entretanto, Hans Gobind Khorana, en la Universidad de Wisconsin, había perfeccionado la bioquímica necesaria para hacer largos trechos de ARN con secuencias repetidas de tres nucleótidos. Sólo bastaba ir probando con los ARNs de Khorana para llegar a conocer qué tripletes (o «codones») de nucleótidos codificaban cada aminoácido de las proteínas. Además, se dedujo qué codones permitían comenzar la síntesis de una cadena protéica y qué codones marcaban el punto final de la misma. El código genético se

había descifrado por completo en tan sólo cuatro años. Nirenberg y Khorana recibieron el premio Nobel en 1968 por sus contribuciones que ayudaron a descifrar completamente el lenguaje mediante el cual se plasma la vida.

A partir de este momento, el genoma de cualquier organismo pudo ser comprendido de un modo tan detallado como nunca se había imaginado hasta entonces. El genoma era un diccionario de palabras en código, que traducido, determinaba cuales proteínas podía fabricar cada organismo. Era el centro de control de la célula. Las secuencias de aminoácidos de algunas proteínas y la de bases nitrogenadas de algunas moléculas de ARN habían sido concretamente analizadas por los científicos. Si pudiese descubrirse la secuencia de bases nitrogenadas del ADN, esto es, «secuenciarlo», se conocería el conjunto completo de genes de un organismo.

La molécula promedio de ADN en una célula humana es de unos 3300 millones de pares de nucleótidos. A pesar de los increíbles avances que habían conducido a la comprensión de la naturaleza del gen, cualquier intento por determinar la secuencia completa de nucleótidos de cualquier genoma completo, y mucho más del humano, era algo imposible en ese entonces; sin embargo, una década más tarde, ya no sería tan impensable. La biología molecular abriría un nuevo capítulo de la ciencia, el gen se transformaría en una entidad que podía ser aislada, reproducida en el laboratorio y llevada de un organismo a otro. Había comenzado la era de la ingeniería genética.

### **UNA QUIMERA BACTERIANA ABRIÓ EL CAMINO HACIA EL PROYECTO GENOMA HUMANO.**

En 1966, algunos investigadores en biología molecular habían dejado la genética molecular por una disciplina naciente y enigmáticamente apasionante la neurobiología. Entre ellos, Francis Crick, que se había convertido una década antes en el líder intelectual de toda una generación de científicos. Crick creía que las bases de la biología molecular habían sido bien establecidas y sólo restaba «llenar los muchos detalles». Desafortunadamente no previó que una verdadera avalancha de descubrimientos pronto inundaría a la comunidad científica e iniciaría un período de excitación y logros que ha tenido pocos paralelos en la historia de la investigación biológica.

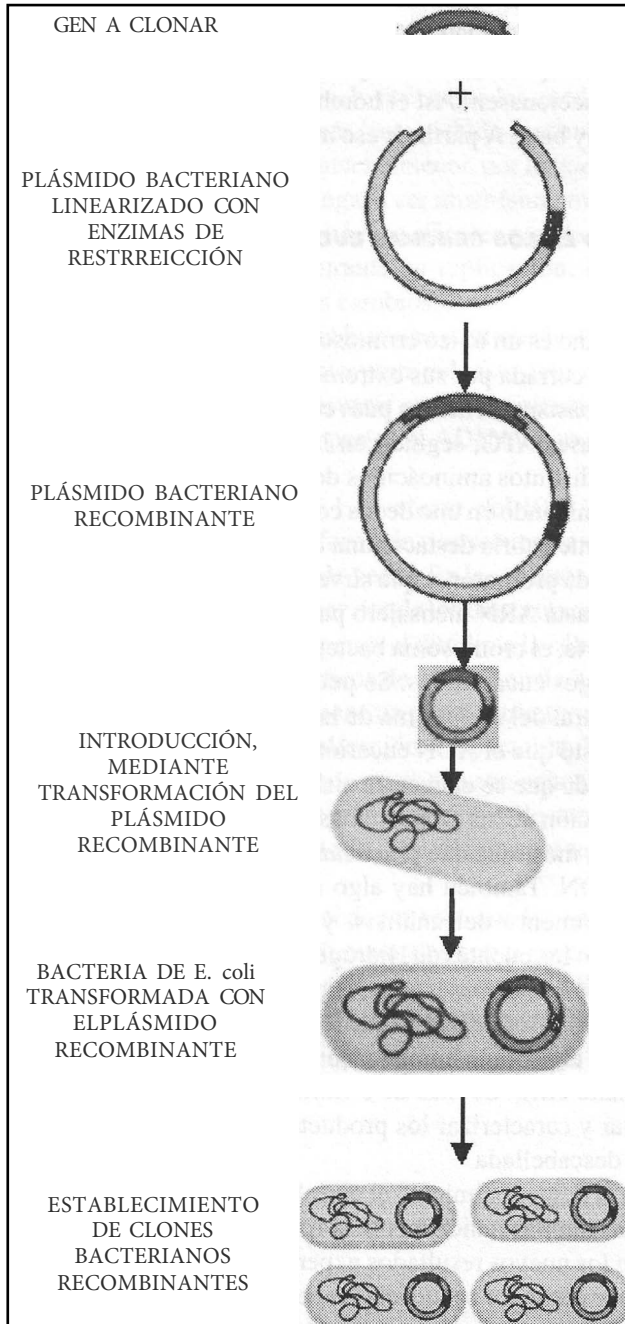
Todo comenzó cuando los biólogos echaron mano de la completa caja de herramientas enzimáticas que había proporcionado el estudio de la bioquímica bacteriana y viral. Hoy en día, varios cientos de estas enzimas están rutinariamente disponibles en los laboratorios; se puede simplemente examinar un catálogo de suministros de biología molecular, llamar por teléfono y pedir cualquiera de ellas.

La primera de las enzimas mágicas fue aislada en 1970 por Hamilton Smith, en la Universidad Johns Hopkins, a partir de la bacteria *Haemophilus influenzae*, esta bacteria ha sido culpada durante años de producir la gripe, hecho al cual debe su nombre latino; hoy en día sabemos que esta frecuente enfermedad es en realidad causada por el virus de la influenza. Smith y sus colaboradores comprobaron que la enzima, a la que denominaron Hind II, por

ser la segunda enzima capaz de cortar el ADN que se localizaba en *H. influenzae*, era muy distinta de su predecesora, pues sólo cortaba el ADN cuando reconocía una determinada secuencia de nucleótidos y nunca en otro lugar. Mas sorprendente era la enzima EcoR I, aislada poco después por Robert Yoshimori, en la Universidad de California en San Francisco, a partir de una cepa de la familiar *E. coli*. Esta enzima introdujo rupturas que cortaban las dos hélices del ADN en sitios distintos, separados por cuatro nucleótidos, por lo que dejaban un extremo sobresaliente en los trozos así obtenidos. En 1972, Janet Mertz y Ronald E. Davis, de la Universidad de Stanford, informaron que el ADN cortado con EcoR I podía volver a aparearse entre sí y era unido permanentemente por la enzima ADN-ligasa, aislada del fago T4.

En 1973, A.C.Y. Chang y Stanley Cohen, de la Universidad de Stanford y Herbert Boyer y Robert Helling, de Universidad de California en San Francisco, informaron a la comunidad científica, de la primera unión de dos moléculas biológicamente funcionales de dos organismos diferentes. Habían tomado el ADN de una cepa de *E. coli*, lo habían cortado con EcoR I y lo habían unido con ADN de otra cepa distinta de la misma bacteria. Después de transformar la segunda cepa con la nueva molécula así formada, el ADN resultaba completamente funcional. Los investigadores llamaron a su molécula «ADN quimera», en honor del ser mitológico con cabeza de león, cuerpo de cabra y cola de serpiente. Hoy día, se prefiere el término más técnico y con menos simbolismo de ADN recombinante o, sencillamente, ADNr (figura 5.1). No parecía muy espectacular el intercambio entre dos *E. coli*; al fin y al cabo, estas bacterias también intercambian de vez en cuando su ADN de forma natural, mediante un proceso denominado «conjugación». Más impresionantes resultaron los siguientes experimentos de Cohen y Boyer: primero, introdujeron ADN de la peligrosa bacteria *Staphylococcus* en la *E. coli*; no contentos con eso, dieron un salto cualitativo importante, al cortar ADN de células eucariotas, procedentes del sapo con garras africano, *Xenopus laevis*, empalmarlo con ADN de *E. coli* e introducirlo en la bacteria. El gen del sapo codificaba un ARN ribosómico eucariótico. Al introducirlo en las células bacterianas, éstas obedientemente se ponían a sintetizar el ARN del sapo como si fuera propio. La mitología había cobrado vida. Una quimera animal-bacteria había sido creada en el laboratorio. Las barreras naturales entre los reinos vivientes habían sido traspasadas.

En realidad, el experimento del sapo no tenía utilidad práctica alguna, aunque su nombre científico, *Xenopus* o «extraños pies» resultaba suficientemente exótico para llamar la atención. Las implicaciones prácticas eran realmente impresionantes, ¿Se podría insertar el gen como el de la insulina humana en *E. coli* para poder fabricarla en grandes cantidades? ¿Podría hacerse con la somatotropina, u hormona del crecimiento, y poder curar así el enanismo humano? Sin embargo, no todas las aplicaciones eran positivas, ¿Qué pasaría si un gen peligroso, como la toxina botulínica, fuese introducido en una bacteria que pudiese vivir en el intestino humano y ésta se liberase en las cañerías de agua potable de un país enemigo?



**Figura 5.1.** Esquema de un experimento de ADN recombinante en el que se introduce un gene humano dentro de un plásmido vector. Los plásmidos resultantes se denominan recombinantes y pueden transformar más de varios genes bacterianos.

Es conveniente indicar que todas las enzimas utilizadas en los procedimientos de la ingeniería genética, así como los ADNs de partida, existían de forma natural en los organismos vivos, lo único artificial que había en el procedimiento era ponerlos juntos en un tubo de ensayo y hacer que reaccionasen. Así el hombre había liberado fuerzas de la naturaleza que aún no comprendía muy bien. A partir de ese momento, tendría que aprender a manejarlas y a convivir con ellas.

**ALGO INESPERADO EN LOS GENOMAS EUCARIÓTICOS CREÓ EL DESCONCIERTO: SOMOS MUY COMPLEJOS.**

El genoma bacteriano es un único cromosoma circular que contiene una sola molécula de ADN de doble cadena cerrada por sus extremos. Los genes bacterianos (así como los de los virus) son, así mismo, bastante sencillos pues consisten en un codón de iniciación formado por la secuencia de tres bases ATG, seguido en la molécula por el conjunto de pares de bases correspondiente a los distintos aminoácidos de la proteína codificada por el gen (3 bases por cada aminoácido), terminando en uno de los codones de terminación, TAA, TGA o TAG. Eso es casi todo. Únicamente cabría destacar una determinada secuencia delante de la secuencia codificante, denominada promotor, y que sirve para que la enzima ARN-polimerasa se una al ADN y lo transcriba hasta ARN mensajero para que pueda ser expresada la proteína correspondiente. Por otra parte, el cromosoma bacteriano tiene asociada muy poca proteína, al contrario que sus homólogos eucarióticos. Se puede decir que es ADN prácticamente desnudo, situado en la zona central del citoplasma de la célula bacteriana.

Nadie había supuesto que el ADN eucariótico fuera muy diferente al bacteriano, excepto las salvedades obvias de que se encuentra aislado del resto de la célula por la envoltura nuclear y que la composición de los cromosomas eucarióticos, que se podían analizar por métodos químicos clásicos, mostraba que prácticamente el 60% de su peso eran proteínas, siendo el 40% restante el ADN. También hay algo de ARN, es el ARN mensajero que se estaba sintetizando en el momento del análisis, y que permanece unido al ADN por fuerzas intermoleculares, como los puentes de Hidrógeno. Puesto que el cromosoma de *E. coli* contiene unos 3 millones de pares de bases y se suponía que tenía unos 2000 genes, con una media de un gen por cada 1000 pares de bases (que formaría una proteína media de 300 aminoácidos de longitud), entonces el genoma humano, con más de 3000 millones de pares de bases, contendría la escalofriante cifra de más de 3 millones de genes. La idea de llegar algún día a conseguir secuenciar y caracterizar los productos proteicos de semejante cantidad de genes era, sencillamente, descabellada.

Pero no era sino el desconocimiento el que abrumaba a los científicos. Las nuevas investigaciones llevarían a cambiar radicalmente la forma de pensar de los biólogos. Básicamente, lo que se obtenía de los nuevos resultados experimentales era que el ADN eucariótico estaba formado en su mayor parte por secuencias repetidas sin capacidad codificadora; es decir, la mayor parte del ADN humano estaba formado por ADN, supuestamente de relleno, cuya función solo se comenzó a entender muy entrada la última década del siglo XX.

Los genomas bacterianos tienden a ser extremadamente compactos. Podemos postular que esto es la consecuencia lógica de miles de billones de generaciones de bacterias, evolu-

cionando rápidamente en un medio cambiante. Así una posible interpretación a este hecho es una adaptación a un medio ambiente cambiante, con el tiempo, se adquiere la tendencia a perder el «equipaje inservible» que pudiera haber llevado en sus cromosomas, de modo que actualmente los genes funcionales son, prácticamente, lo único que les queda.

En los eucarióticos la evolución es mucho más lenta, debido fundamentalmente a dos razones: una es que su reproducción tiene un ritmo mucho menor, por lo que el tiempo que transcurre entre una generación y la siguiente puede llegar a ser muchísimo más largo; la otra, es que las células eucariotas poseen mecanismos mucho más eficaces para la reparación de los errores de copia que se producen en el ADN durante su replicación. Por ello, los genomas eucarióticos presentan mucha más inercia a los cambios.

Parece que sólo de un 2 a un 5 % del genoma humano contiene instrucciones para fabricar proteínas (se incluyen los intrones). Seguramente, parte del resto debe de tener alguna función reguladora de la actividad de los genes o intervenir en la organización de la estructura del cromosoma o en su replicación. Pero la mayor parte del ADN humano, se estima que más de un 90%, no parece tener función alguna.

Con base en lo que se muestra en la tabla 5.1, en efecto, del 10 al 25% del total del ADN humano y de otros eucariotas superiores está formado por secuencias cortas de cinco a diez pares de bases, que se repiten en tándem miles de veces. En los cromosomas, estas «repeticiones cortas» se localizan en los centrómeros, lugar donde los dos cromosomas hermanos continúan unidos tras la replicación, hasta que se separan al dividirse la célula. Otros tres millones de secuencias representan variaciones polimórficas de un solo nucleótido, son los denominados SNPs. Aparentemente, este área del ADN tiene un papel estructural, más que genético, en la célula y, de algún modo, opera durante la replicación y separación. También se encuentran repeticiones cortas en los extremos de los cromosomas, que son conocidos como telómeros.

Otra gran parte del genoma está compuesto por repeticiones de secuencias más largas, las llamadas «repeticiones largas», repartidas por todo el genoma, cuya función, si es que existe, permanece aún desconocida.

Por último, otros componentes de este «ADN de relleno» son los «pseudogenes» (del griego «falsos genes»). Son secuencias que fueron funcionales como genes en estadios evolutivos anteriores, pero que han perdido su capacidad codificadora debido a diversos tipos de mutaciones. Estos pseudogenes pueden ser importantísimos a la hora de llevar a cabo una investigación de la historia evolutiva del hombre.

**Tabla 5.1.** Descripción de las secuencias del genoma de los humanos con base en su grado de repetición

Tipo de secuencia	Denominación	Características
De alta repetición	DNA satélite	10 <sup>6</sup> copias. Ejemplos: Alfoide (centrómeros ; unidad de repetición 171 bp) DYZ1 y DYZ2 (heterocromatina del Y) Satélite (unidad de repetición 68 bp)
	Retroposones	LINES, SINES, Pseudogenes procesados (10 a 10 <sup>4</sup> copias o mas)
De moderada repetición	Transposones	Transposón THE (10 a 10 <sup>2</sup> copias)
	Familias de genes	rRNA de 18 y 28 s; tRNAs
	DNA telomérico	(TTAGGG) <sub>n</sub> 10 <sup>4</sup> copias
	DNA minisatélite	Altamente polimórfico 10 <sup>2</sup> a 10 <sup>3</sup> copias
De secuencia única	DNA microsátélite	(CA) <sub>n</sub> ; (TG) <sub>n</sub> ; 10 <sup>5</sup> copias altamente polimórfico
	Genes que codifican proteínas	Copia única
	DNA intergénico	Función desconocida

**SORPRESA, EL GEN EUCARIÓTICO ESTÁ DIVIDIDO**

Además de este ADN de relleno, en 1977, se hizo un descubrimiento inesperado; los patrones de fragmentos producidos por enzimas de restricción en el ADN cromosómico revelaron que había regiones dispersas dentro de los genes que no eran parte del ARN mensajero, y de esta manera no se expresaban en proteínas. Los genes eucarióticos solían estar interrumpidos por una o varias secuencias que no parecían tener función alguna, al contrario de los genes bacterianos, que eran siempre ininterrumpidos. Walter Gilbert, de la Universidad de Harvard, quien había ideado el primer método útil para la secuenciación del ADN, llamó «exones» a las secuencias del gen que se expresan en proteínas e «intrones» a aquellas secuencias intermedias silenciosas que no se convierten en aminoácidos. Los intrones aumentan extraordinariamente la longitud de los genes eucarióticos, ya que la longitud

total de los intrones incluidos en un gen puede ser muchísimo mayor que la longitud total de las secuencias codificadoras. Algunos genes humanos son realmente gigantes, como el de la distrofina, la proteína que resulta defectuosa en los enfermos afectados de distrofia muscular de Duchenne, que puede extenderse a lo largo de dos millones de pares de bases y tener más de sesenta intrones. Si todos los genes humanos fueran tan largos como el de la distrofina, aún sin considerar la existencia del ADN de relleno, los tres mil millones de bases del genoma humano sólo codificarían 1500 genes, un número menor que el número de genes que posee la bacteria más sencilla, ¿Sería esto cierto?

### ***EL ASTUTO INTRÓN***

Cuando el ADN es copiado hasta ARN, los intrones son también copiados, de modo que el ARN sintetizado directamente del ADN (el llamado «transcrito primario») de la distrofina posee 2 millones de nucleótidos. Una vez sintetizado, el transcrito primario madura en el núcleo durante un proceso llamado de «corte y empalme» (traducción del término inglés intraducible «splicing»), durante el cuál son eliminados los intrones y se añade una cola de longitud variable de nucleótidos de adenosina (la cola de poliA), antes de que el ARN mensajero, ya maduro, salga del núcleo al citoplasma para servir de patrón en la síntesis de proteínas. Esto deja al ARN mensajero con el código completo e ininterrumpido del gen, excluidos los intrones.

Este fenómeno apunta a dos problemas relacionados con la secuenciación y el clonado de genes (se llama así al proceso de introducir un gen exógeno en una bacteria para que lo exprese). El primero es que no se puede deducir directamente la secuencia exacta de la proteína que se expresará a partir de la secuencia de ADN genómico (aunque existen ciertas secuencias encontradas repetidamente en algunos intrones, que han sido identificadas como las zonas que limitan el intrón y que son eliminadas durante el splicing, los intrones siempre parecen comenzar por el par de bases GT y acabar en AG). El otro problema es que un ADN recombinante que posea ADN genómico con intrones no expresará la proteína correcta al ser introducido en una bacteria, ya que éstas no poseen mecanismos para la eliminación de los intrones. Se puede obviar este problema utilizando para el clonado del gen, no ADN genómico, sino una copia en ADN del ARN mensajero correspondiente, que no posee intrones. La copia del ARN mensajero se puede conseguir utilizando la enzima «transcriptasa reversa» o «retrotranscriptasa», aislada de algunos retrovirus. Esta enzima permite copiar una cadena de ARN en su complementaria en ADN. El ADN así obtenido se conoce como ADN copia, ADN complementario o, sencillamente, ADNc.

Con estos nuevos conocimientos, podemos realizar una nueva estimación del número de genes que posee el genoma humano. En la actualidad se estima entre 30.000 y 45.000 genes, un número adecuadamente pequeño como para poder pensar en su secuenciación. Después de haber leído los párrafos anteriores, no debe extrañarnos semejante falta de precisión; sencillamente, no conocemos cuál es el tamaño medio de un gen humano. Acerca de esta falta de conocimientos, hay que tener en cuenta que, en fecha tan reciente como 1956, ni siquiera se sabía el número exacto de cromosomas que componían el genoma humano. La mayoría de los biólogos suponían que eran 48, hasta que las observaciones cuidadosas de Jo Hin Tjio y Albert Levan demostraron que el número exacto era 46.



## REPARTIENDO EL GENOMA

Las modernas técnicas de la ingeniería genética proporcionan una herramienta indispensable para el estudio de los genomas. A nadie se le ocurre ponerse a estudiar directamente un cromosoma humano completo; sencillamente, son demasiado largos como para poder ser manejados adecuadamente por nuestra caja de herramientas enzimáticas. Si pudiéramos a cortar uno de nuestros cromosomas con EcoR I, obtendríamos miles de fragmentos de todos los tamaños.

En lugar de ello, los biólogos moleculares hacemos uso de la vieja estrategia de «divide y vencerás». Ninguno de los estudios actuales que se están llevando a cabo sobre el genoma humano ha sido realizado sobre cromosomas humanos intactos. Por el contrario, se llevan a cabo sobre «genotecas» o «bibliotecas de genes», construidas mediante la técnica que detallaremos a continuación.

Las primeras genotecas fueron construidas utilizando como huéspedes bacterias, en concreto, nuestra incansable colaboradora *E. coli*. La idea consiste en tomar todo el genoma humano (o de cualquier otro organismo), cortarlo en cientos de miles de trozos con una enzima de restricción que produzca extremos cohesivos y ligar los fragmentos de ADN así obtenidos en un vector genético adecuado (los bacteriófagos han demostrado sobradamente su capacidad), esto fue lo que hicieron los científicos de Celera Inc. cuyo presidente es Craig Venter. Con los vectores recombinantes así obtenidos, infectamos un cultivo de bacterias, cada una de las cuales será transformada por uno o varios bacteriófagos que depositarán en su interior uno o varios fragmentos del ADN humano que portan. De esta manera se pueden construir genotecas de ADN genómico o de ADNc. La técnica en este último caso es semejante, excepto que el material de partida es el conjunto de ARN mensajero que se está expresando en un determinado tejido celular. De este último se obtienen genotecas sin intrones ya listas para ser empleadas en la expresión heteróloga de las proteínas humanas.

El problema, si tenemos en mente el estudio de un gen o de una proteína en concreto, es saber cuál de los cientos de miles de bacterias ha incorporado el gen que nos interesa. Existen diversos métodos de «screening» o tamizado, el más extendido de los cuales consiste en rastrear la genoteca con una sonda de ARN marcada radiactivamente cuya secuencia sea análoga a la del gen que buscamos (por ejemplo, utilizando la secuencia de un gen conocido de un organismo relacionado, o la secuencia de la proteína, obtenida mediante técnicas de secuenciación de proteínas). La sonda de ARN se «hibridizará» con el ADN del gen que buscamos, y la colonia de bacterias producida a partir de la bacteria que inicialmente contenía el gen emitirá radiactividad, que podremos detectar mediante una película fotográfica. Si no nos interesa ningún gen concreto, sino secuenciar todo el genoma, podemos entonces coger la bacteria que más rabia nos dé y ponernos a secuenciar la secuencia que transporta. Si hacemos esto con un número suficiente de bacterias, llegaremos a disponer de numerosos trozos de secuencia de ADN humano que, con un poco de suerte, se superpondrán entre sí. A dichos trozos se les denomina «contigs» y con ellos, en teoría, se podría resolver, como un rompecabezas, todo la secuencia del genoma.

El problema de las genotecas en bacterias es que los trozos que pueden almacenar los fagos y otros vectores bacterianos suelen ser de tamaño corto (hasta 20 kilobases). Para resolver

este inconveniente se han desarrollado otros tipos de genotecas. Las más utilizadas en la actualidad para el estudio del genoma humano son las genotecas que utilizan como huésped a la levadura *Saccharomyces cerevisiae*. Las levaduras son células eucariotas, como los humanos, pero extremadamente sencillas y que, al igual que las bacterias, pueden crecer fácil y rápidamente en medios de cultivo líquidos. Los vectores adecuados para las levaduras se denominan YACs, que es la abreviatura de «cromosoma artificial de levadura», desarrollados en 1987 por David T. Burke, George F. Carle y Maynard V. Olson, en la Universidad de Washington. Pueden contener fragmentos de ADN mucho mayores, los modernos «MegaYACs» pueden albergar fragmentos de un millón de bases. Basta con 3000 de estos MegaYACs para completar una genoteca humana.

A mediados de los ochenta, los biólogos disponían de técnicas apropiadas, lo suficientemente potentes como para pensar en realizar la hercúlea tarea de cartografiar las posiciones de los genes en el genoma humano como paso previo para descifrar el conjunto de instrucciones genéticas esparcidas aquí y allá en la secuencia de pares de bases de los cromosomas. Los posibles beneficios de semejante esfuerzo eran tan abrumadoramente numerosos como para que los científicos pusieran todo su empeño en llevarlo a cabo.

**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

# **CÓMO SECUENCIAR Y CARTOGRAFIAR EL GENOMA**

*“El proyecto genoma no es tan solo un esfuerzo aislado por parte de los biólogos moleculares. Es un desarrollo natural de temas de la biología como un todo. En el mas simple sentido, la idea de secuenciar el genoma humano es un esfuerzo por definir todos los genes que constituyen un humano”*

**WALTER GILBERT. “A VISION OF THE GRAIL”.  
(Una vsion del grial). 1992**

Hacia mediados de la década de los años 80 del siglo pasado, la metodología del ADN recombinante y sus técnicas asociadas que incluían vectores de clonación, enzimas de restricción, transformación artificial de células procarióticas y eucarióticas, bibliotecas de genes, sondas moleculares, secuenciación, genética inversa, PCR, etc. habían alcanzado una madurez suficiente como para que se planteara la pertinencia y viabilidad de un proyecto coordinado de caracterización detallada que fuese hasta nivel de secuencia de nucleótidos del genoma humano y del genoma de una serie de organismos modelo.

## **¿POR QUÉ CARTOGRAFIAR Y SECUENCIAR GENOMAS?**

Uno de los paradigmas de la biología moderna es encontrar respuestas integrales a los procesos de la vida. Puesto que el ADN es la molécula que en nuestro planeta contiene la información biológica para el funcionamiento y evolución de los organismos, la tendencia natural ha sido terminar buscando explicaciones al nivel de ADN. Esperamos que el conocimiento de los procesos moleculares de la clave de muchos fenómenos que hoy entendemos a niveles menos profundos ya descritos por otras ciencias biológicas como la fisiología, la biología celular y la bioquímica entre otras.

Ha llegado un momento en que se plantea que abordar el estudio detallado de los genomas de los organismos es mucho menos costoso y más interesante intelectualmente, logrando el conocimiento detallado de la secuencia. Los Proyectos Genoma se han constituido en un punto de arranque para generar nuevos descubrimientos en las ciencias biomédicas. Con los datos de secuencias habrá que trabajar para dar respuestas a cuestiones de expresión de genes, de regulación genética, de interacción de las células con sus entornos. Además, la secuenciación de genomas de plantas y animales domésticos conducirá a nuevos avances en la mejora agronómica y ganadera. Para comprender la evolución será cada vez más esencial el disponer de datos de secuencias. La bioinformática permite comparar genes y genomas completos, lo que junto con otros datos biológicos y paleontológicos, está dando una nueva visión de la evolución de la vida en nuestro planeta.

### ***CARTOGRAFÍAS GENÓMICAS***

Tal como hemos dicho en capítulos anteriores, los tres objetivos del proyecto genoma humano son secuenciar, caracterizar y localizar la información física y genética del genoma. El PGH ha utilizado dos tipos de estrategias para cartografiar el genoma, aunque en última instancia los mapas obtenidos por distintos métodos han de ser correlacionados e integrados. Básicamente la localización de secuencias de nucleótidos utiliza dos estrategias claves; la cartografía genética de ligamiento y la cartografía física.

### ***MAPEO GENÉTICO POR LIGAMIENTO***

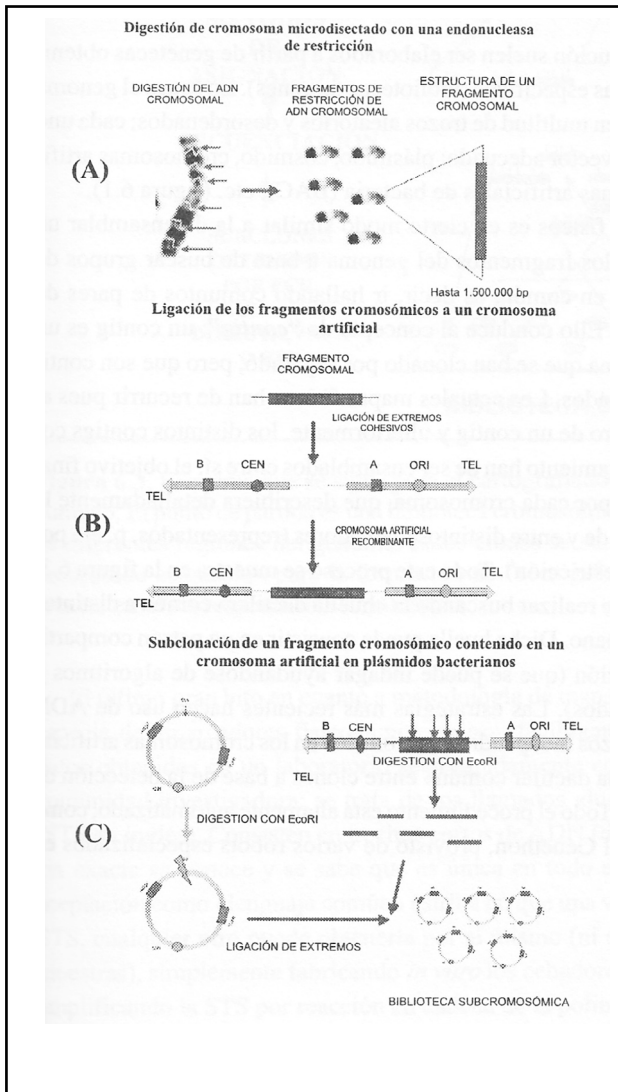
Este tipo de cartografía genética se basa en el cálculo de la frecuencia a la que se co-heredan formas alternativas (alelos) de dos loci genéticos que están ligados y forman parte de un mismo cromosoma. Hasta el advenimiento de las técnicas moleculares, los mapas genéticos de ligamiento en humanos eran bastante rudimentarios, ya que en su elaboración no puede intervenir (por obvios motivos éticos) la experimentación de laboratorio que se usa en animales y porque los datos habían de basarse casi exclusivamente en la comparación de fenotipos normales y los mutantes correspondientes a determinadas enfermedades genéticas y en el recurso a análisis de familias, a ser posible con registros de varias generaciones y con gran número de individuos.

La revolución de la cartografía genética de ligamiento sobrevino cuando a finales de los años 70 se recurrió al análisis molecular de zonas de ADN no codificadoras y que son muy polimórficas. Tal como hemos descrito anteriormente, existen varios tipos de secuencias (algunas de ellas de naturaleza repetitiva, como los VNTR, los microsatélites, etc.), dispersas por el genoma, cada una de ellas con varios alelos distribuidas diferencialmente en la población. Entre las ventajas de los microsatélites se cuentan: contenido informativo muy alto, con lo que los análisis estadísticos mejoran en fiabilidad; distribución abundante y relativamente

uniforme por todo el genoma; además que se pueden experimentar fácilmente mediante PCR. De otra manera estos loci sirven en genética clínica como marcadores útiles para localizar genes relacionados con enfermedades.

Los polimorfismos moleculares han permitido que en la actualidad el PGH haya generado detallados mapas genéticos del genoma humano a un nivel de resolución en torno a 1 centimorgan (cM) o incluso menos. Esto ya se logró en 1994, un año antes de lo previsto y en buena parte con resoluciones mejores (0.7 cM).

**CARTOGRAFÍA FÍSICA E INTEGRACIÓN DE LOS MAPAS.**



Los mapas físicos tienen como objetivo especificar distancias físicas mensurables en pares de bases (pb) o alguno de sus múltiplos. Obviamente, el mapa físico de mayor detalle es la propia secuencia del genoma. Pero antes de llegar a obtenerla, hubo que

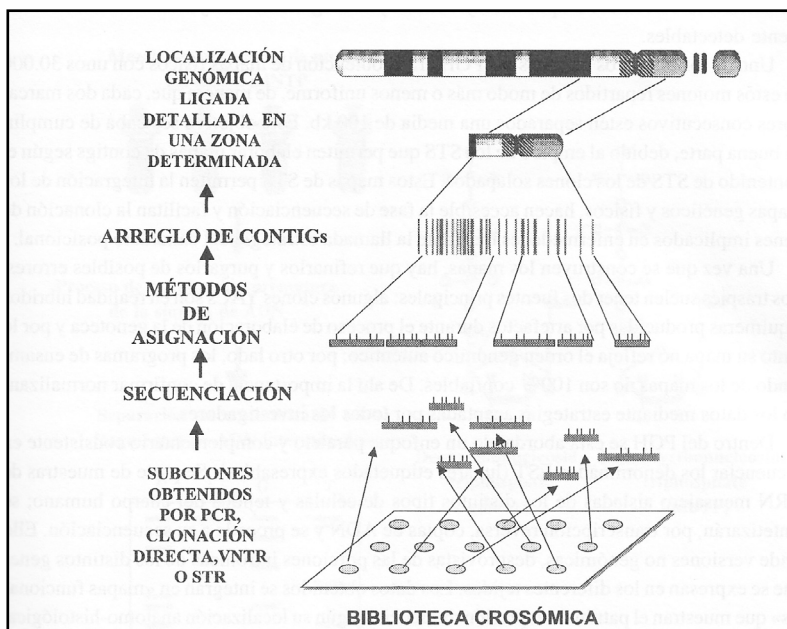
**Figura 6.1.** Procesos que se realizan para la construcción de una biblioteca que contenga subfragmentos de un cromosoma. Inicialmente se debe fragmentar el cromosoma microextraído de una célula, mediante el tratamiento con endonucleasas de restricción, posteriormente fragmentos cromosómicos de hasta 150.000 pares de bases (pb) se clonan en cromosomas artificiales. Una vez identificado el correspondiente clon, éste se somete de nuevo a tratamiento con enzimas de restricción para poder tener una colección de muchos fragmentos cuyo tamaño no es mayor a 10.000 pares de bases. Éstos son luego clonados en plásmidos, vectores apropiados, con lo que se obtiene una biblioteca subcromosómica.

elaborar mapas físicos partiendo de resoluciones bajas y avanzando hacia las resoluciones cada vez mayores. En cierta manera, los mapas físicos de menor resolución son los propios cariotipos: la visualización microscópica de la dotación cromosómica haploide humana teñida con colorante de Giemsa nos muestra un patrón alternante de bandas claras y oscuras, en el que cada banda tiene una media de unos 7 millones de pares de bases. Si bien los métodos citogenéticos tienen sus limitaciones, no hay que olvidar que actualmente existen novedosas herramientas de citogenética molecular (como la hibridización *in situ* con sondas fluorescentes *in situ* o FISH, la “pintura de cromosomas», etc.) que permiten un mayor detalle y que, unidas a otras técnicas, aumentan el arsenal de enfoques para el estudio de los genomas, de su dinámica y de sus alteraciones.

Los mapas físicos de mayor resolución suelen ser elaborados a partir de genotecas obtenidas por fragmentación de cromosomas específicos (bibliotecas de genes). En éstas, el genoma a estudiar se encuentra fragmentado en multitud de trozos aleatorios y desordenados; cada uno de ellos, clonado por separado en un vector adecuado: plásmido, cósmido, cromosomas artificiales de levadura (YAC), cromosomas artificiales de bacteria (BAC), etc. (figura 6.1).

La idea para elaborar los mapas físicos es en cierto modo similar a la de ensamblar un rompecabezas: consiste en ordenar los fragmentos del genoma a base de buscar grupos de fragmentos que tienen alguna zona en común, es decir, ir hallando conjuntos de pares de fragmentos parcialmente solapados. Ello conduce al concepto de “*contig*”: un contig es un conjunto de fragmentos de un genoma que se han clonado por separado, pero que son contiguos y que están parcialmente solapados. Los actuales mapas físicos han de recurrir pues al ensamblaje de esos fragmentos dentro de un contig y ulteriormente, los distintos contigs correspondientes al mismo grupo de ligamiento han de ser ensamblados entre sí: el objetivo final (ideal) sería obtener un gran contig por cada cromosoma, que describiera detalladamente la posición y distancia física (en bases) de y entre distintos marcadores (representados, p. ej., por secuencias blanco para enzimas de restricción). Todo este proceso se muestra en la figura 6.3.

La cartografía de contigs se puede realizar buscando la «huella dactilar» común a distintos clones de una genoteca de ADN humano. Dicha huella puede consistir en un patrón compartido de dianas de enzimas de restricción (que se puede indagar ayudándose de algoritmos y programas computacionales adecuados). Las estrategias más recientes hacen uso de ADN humano en forma de unos 20.000 trozos independientes clonados en los cromosomas artificiales de levadura y buscando la «huella dactilar común» entre clones a base de la detección de determinadas secuencias repetitivas. Todo el procedimiento está altamente automatizado, como en el famoso laboratorio francés del Génethon, provisto de varios robots especializados en procesar y analizar las muestras.



**Figura 6.3.** Esquema que se sigue para el cartografiado físico de las secuencias del genoma humano. El punto de partida es una biblioteca cromosómica que contenga fragmentos clonados de diferentes regiones del genoma. Estos clones se someten a secuenciación y se inicia el proceso de asignación para la formación de “contigs” que luego permiten su localización precisa en el correspondiente cromosoma.

El último gran hito en cuanto a metodología de mapas físicos ha sido el desarrollo de una especie de «marcadores físicos universales», fácilmente generables, que permiten que los datos obtenidos en un laboratorio sean rápidamente compartidos y asimilados por toda la comunidad investigadora: se trata de los llamados «lugares etiquetados por su secuencia» (STS en inglés). Consisten en trechos cortos de ADN (de entre 100 y 1000 pb) cuya secuencia exacta se conoce y se sabe que es única en todo el genoma. Su facilidad de uso y su aceptación como «lenguaje común» estriba en que una vez que un investigador descubre una STS, cualquier otro puede obtenerla por sí mismo (ni siquiera hace falta el envío físico de muestras), simplemente fabricando *in vitro* los cebadores correspondientes a sus extremos y amplificando la STS por reacción en cadena de la polimerasa (PCR). Los



STS definen puntos concretos únicos del mapa físico y constituyen magníficos «mojones» o señales fácilmente detectables.

Uno de los objetivos iniciales del PGH era la obtención de mapas físicos con unos 30.000 de estos mojones repartidos de modo más o menos uniforme, de manera que, cada dos marcadores consecutivos estén separados una media de 100 kb. Este objetivo se acaba de cumplir, en buena parte, debido al empleo de los STS que permiten elaborar mapas de contigs según el contenido de STS de los clones solapados. Estos mapas de STS permiten la integración de los mapas genéticos y físicos, hacen accesible la fase de secuenciación y facilitan la clonación de genes implicados en enfermedades mediante la llamada estrategia del candidato posicional.

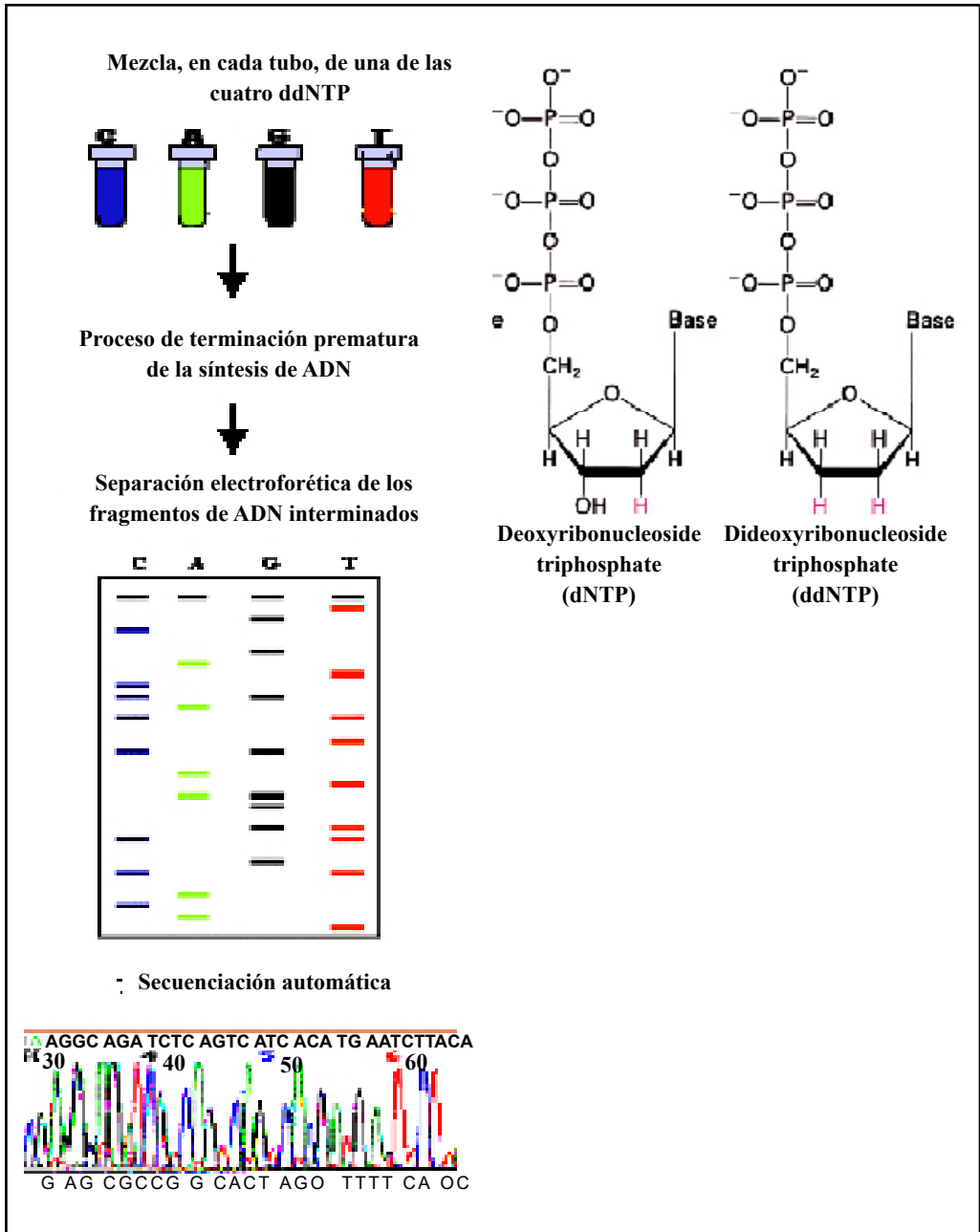
Una vez que se construyen los mapas, hay que refinarlos y purgarlos de posibles errores. Los traspies suelen tener dos fuentes principales: algunos clones YACs son en realidad híbridos o quimeras producidas por artefactos durante el proceso de elaboración de la genoteca y por lo tanto su mapa no refleja el orden genómico auténtico; por otro lado, los programas de ensamblado de los mapas no son 100% confiables. De ahí la importancia de confirmar normalizando los datos mediante estrategias aceptadas por todos los investigadores.

Dentro del PGH se está abordando un enfoque paralelo y complementario consistente en secuenciar los denominados EST (lugares etiquetados expresables). Se parte de muestras de ARN mensajero aisladas de los distintos tipos de células y tejidos del cuerpo humano; se sintetizarán, por transcripción reversa, copias de ADN y se procede a su secuenciación. Ello rinde versiones no genómicas, desprovistas de las porciones intrónicas de los distintos genes que se expresan en los diferentes tejidos. Los datos obtenidos se integran en «mapas funcionales» que muestran el patrón de expresión diferencial según su localización anátomo-histológica.

### **LA SECUENCIACIÓN DEL ADN**

Secuenciar es identificar uno por uno linealmente los distintos nucleótidos que componen el genoma de un organismo. Esta asignación lineal de nucleótidos es la forma más fina de hacer un cartografiado físico y es por su sencillez un enfoque experimental y tecnológico utilizable en todos los proyectos genoma.

La secuenciación del orden de nucleótidos de un ADN se lleva a cabo mediante dos métodos: el de Maxam y Gilbert, que consiste en introducir modificaciones químicas en algunas de las bases nitrogenadas que conforman el ADN a secuenciar. Posteriormente se pueden diferenciar mediante métodos fisicoquímicos. Debido a que el proceso es muy largo y tedioso, este método es actualmente muy poco utilizado. El otro método es el de Sanger, desarrollado por Fred Sanger de la Universidad de Stanford que se ha denominado de terminación precoz de la síntesis de cadenas de ADN y utiliza para ello una serie de nucleótidos modificados denominados didesoxiribonucleósidos -5'-trifosfatos (ddNTP). El proceso es bastante simple y como muestra la figura 6.4, genera una colección de fragmentos de ADN



**Figura 6.4.** Protocolo seguido para la secuenciación automática de ADN mediante el método de Sanger. A la derecha se muestran las estructuras químicas de un dNTP y de un ddNTP.

Cada ddNTP se marca con un colorante fluorescente diferente. Ello permite hacer la electroforesis en el mismo carril del gel. Las bandas de ADN son detectadas por su fluorescencia según pasan delante del detector. Si el detector lo hacemos mover en horizontal, podrá leer varias secuenciaciones al mismo tiempo. Los datos pasan a un sistema computarizado.

### **NUEVOS MÉTODOS DE SECUENCIACIÓN**

#### **MICROSCOPIAS DE EFECTO TÚNEL (STM) Y DE FUERZA ATÓMICA (AFM)**

En la microscopía de barrido (*scanning*) de efecto túnel (STM). Una fina sonda se mantiene muy cerca del objeto (en este caso ADN), por medio de un sistema de control basado en la detección de una minúscula corriente inducida por el efecto túnel entre la punta de la sonda y el ADN. Una técnica muy parecida, es la microscopía de fuerza atómica (AFM), en la que el control se debe a la medida de las fuerzas de Van Der Waals (fuerzas atractivas del núcleo atómico) entre la sonda y la muestra. En cualquiera de los dos casos, la punta se mueve a lo largo del objeto, de modo que sus desplazamientos en vertical se miden y registran, generando una imagen de la superficie del objeto. Aunque se han obtenido imágenes del esqueleto azúcar-fosfato de ADN de cadena sencilla y de cadena doble, está por verse si se pueden «ver» las bases nitrogenadas. Si es así, la técnica podría secuenciar del orden de 1 Mb cada día.

#### **SECUENCIACIÓN POR HIBRIDACIÓN: “CHIP” DE HIBRIDACIÓN**

La secuenciación por hibridación en chips con oligonucleótidos. Se basa en sintetizar distintas sondas de oligonucleótidos y unir las en disposiciones ordenadas (“*arrays*”) a una fina placa de nylon o vidrio. Este chip se prueba frente a un ADN marcado fluorescentemente de modo que el patrón y cantidad de fluorescencia suministra información sobre la secuencia del ADN en cuestión. La última generación de este enfoque es la combinación de técnicas fotolitográficas (como la de los chips de silicio para computadores) con síntesis química en fase sólida, que logra chips con ordenaciones de decenas e incluso centenares de miles de oligos distintos; estos chips pueden usarse para identificar secuencias marcadas fluorescentemente en cuestión de minutos, utilizando un microscopio confocal de fluorescencia, totalmente automatizado, que registra los datos.

A modo de estudio piloto sobre sus posibilidades, la empresa, Affimetrix ha logrado resecuenciar por este método las 16 kb de ADN mitocondrial humano, con un dispositivo formado por 135.000 oligonucleótidos. Con la tecnología actual se puede llegar a sintetizar en un día 400.000 oligonucleótidos de 20 bases cada uno, dispuestos en un chip de 1,6 cm<sup>2</sup>. Pero el objetivo final es lograr un chip con los 4 millones de sondas necesarias para secuenciar todo el genoma humano en una sola hibridación.

## **ESTRATEGIAS DE SECUENCIACIÓN DEL GENOMA HUMANO.**

### **SECUENCIACIÓN GENÓMICA**

Hasta hace muy poco, la secuenciación genómica dependía de la previa disponibilidad de mapas físicos detallados, para que los fragmentos secuenciados pudieran ser ensamblados correctamente. Este enfoque requiere el manejo de una gran cantidad de datos y no es totalmente automatizable. Resumamos brevemente cuál es la estrategia habitual para secuenciar genomas:

El ADN genómico de un cromosoma determinado, se corta aleatoriamente en fragmentos, que se separan en distintas clases de tamaños, y se insertan en distintos tipos de vectores, cada uno con una capacidad media diferente (por ejemplo, los YACs portan insertos entre 100 y 2000 kb, los cósmidos unas 40 kb). Se preparan mapas físicos de baja resolución a base de clones solapados de insertos de YACs.

Se construyen mapas físicos de alta resolución, preparados para la secuenciación, por medio de la subclonación en cósmidos de fragmentos aleatorios de insertos de YACs.

Finalmente, se escoge un conjunto de clones de cósmidos con solapamientos mínimos, sus insertos se rompen aleatoriamente y se subclonan en vectores para la secuenciación (derivados del fago M13, plásmidos de la serie pUC, etc.).

Por cada cósmido hay que secuenciar unos 800 clones de fago M13, con un tamaño medio de inserto de 400 pb, y la secuencia de 40 kb del inserto del cósmido original se ensambla computacionalmente. Una variación de este método es la secuenciación aleatoria (“*shotgun*”) que parte del DNA completo el cual es digerido en muchos fragmentos los cuales se clonan y se secuencian, este enfoque es tedioso puesto que obliga a secuenciar muchas veces, (de 8 a 10), el mismo segmento de secuencia; lo que tiene de ventaja es que asegura una mayor confiabilidad de los datos obtenidos. Sin embargo, como ya dijimos, presenta varios inconvenientes: hay que disponer previamente de buenos mapas; los YACs son inestables y muchos de ellos son quimeras, artefactos de clonación que no reflejan el orden genómico (identificarlos y descartarlos es una tarea que lleva tiempo y esfuerzo); y como ya sabemos, esta metodología no se puede automatizar totalmente. Un tema importante es tener tasas de errores lo más bajas posibles: del orden de 0.02-0.2%.

### **SECUENCIACIÓN DE ADN COMPLEMENTARIO (ADNc).**

No es una alternativa sino un complemento a la secuenciación genómica. No es el gen lo que estamos secuenciando sino la «retrocopia» de su ARNm, obtenida por transcripción reversa, que está desprovisto de intrones y de secuencias reguladoras no traducidas. Si bien la secuenciación de ADNc no nos da información de la estructura del gen, sí nos la puede dar sobre su expresión: en qué tejidos se expresa, bajo qué condiciones, etc., lo que permite iniciar mapas funcionales del genoma humano.

## NUEVAS ESTRATEGIAS DE SECUENCIACIÓN GENÓMICA.

Recientemente se ha propuesto una estrategia de secuenciación genómica que no depende de mapas físicos previos y que en lugar de YACs emplea otro tipo de vectores; los cromosomas artificiales de bacteria (BACs), que aunque permiten insertos de entre 100 y 350 kb, tienen la gran ventaja de aceptar insertos genómicos con gran fidelidad (sin quimerismos, ni apenas reordenaciones del material insertado). La estrategia consiste en lo siguiente:

Se crea una genoteca humana de BACs, con un tamaño medio de 150 kb, y una redundancia de 15 veces, de modo que consta de unos 300.000 clones. Los clones de la genoteca se reparten en pocillos de placas de microtitulación. Se secuencian unas 500 bases en cada uno de los dos extremos del inserto de cada clon. Esto significa que las 600.000 secuencias de los extremos de los clones están repartidas a razón de una cada 5 kb de genoma, y constituyen el 10% de la secuencia genómica.

A estas secuencias se las denomina STCs, acrónimo inglés de «conectores» etiquetados por su secuencia», ya que permiten que por término medio cada clon BAC pueda ser conectado con otros 30 (un inserto típico de 150 kb dividido por 5 kb, estará representado en otros 30 BACs). Las secuencias de las STCs se pondrían enseguida en Internet, a disposición de la comunidad investigadora. Con una enzima de restricción se toma la «huella dactilar» de cada clon BAC.

Un clon BAC que actúa de semilla se secuenciará por cualquier método, y se compara con la base de datos de STCs para identificar los clones solapados. Se secuencian los dos clones BAC que muestren consistencia interna entre las huellas dactilares y solapamiento mínimo en ambos extremos con respecto al BAC semilla.

El proceso se va haciendo reiterativo a ambos lados del BAC semilla (lo que podríamos llamar «paseo genómico con BACs»). De esta manera, el genoma humano completo se podría secuenciar con sólo 20.000 clones BAC.

Aunque la estrategia es muy atractiva por su aparente rapidez y menor costo, tiene varios inconvenientes, entre ellos el que obliga a mantener uno o varios centros depositarios de las placas con los clones BACs y al envío de clones entre laboratorios. Dado que el PGH ha funcionado muy bien hasta ahora de modo descentralizado, es posible que haya reticencias en grupos que vean amenazada la actual cuasi-democracia investigadora. No se puede olvidar quién hace la propuesta: el director de uno de los centros privados más «agresivos» en investigación genómica, Carig Venter.

## **LOS OTROS PROYECTOS GENOMA**

*“... Obviamente, la vida si se inició: nuestra existencia lo demuestra. Pero ¿tenía que iniciarse? En otras palabras, ¿era inevitable la emergencia de la vida a partir de un caldo químico o cualquier otra cosa, contando con millones de años?”.*

**PAUL DAVIES. “THE FIFTH MIRACLE”  
(El quinto milagro). 1999.**

### **SECUENCIAR UN GENOMA NO ES SUFICIENTE**

Casi desde el mismo comienzo de la idea de secuenciar el genoma humano, la mayoría de los biólogos estuvieron de acuerdo en que sería mucho más útil tener datos sobre genomas de otros organismos relacionados. La utilidad es clara, puesto que resulta éticamente imposible realizar experimentos de genética humana. A nadie se le ocurriría modificar deliberadamente un par de bases del ADN humano, e implantar el óvulo así fecundado en el útero de una mujer, sólo para ver qué ocurre. Sin embargo los experimentos genéticos con bacterias, levaduras o ratones, por ejemplo, no plantean este tipo de problemas. Esta consideración, unida al hecho de que la gran mayoría de los genes presentes en el genoma humano se espera que tengan su contrapartida en los genomas de otras especies filogenéticamente relacionadas, explica por qué puede valer la pena invertir semejante cantidad de esfuerzo y dinero en la secuenciación de otros genomas, algunos de ellos, tan grandes como el nuestro.

Uno de los problemas con los que nos encontraremos cuando hayamos secuenciado el genoma humano, será que tendremos miles de secuencias que posiblemente codifiquen proteínas de las cuales sólo podremos conocer su secuencia de aminoácidos; a partir de la secuenciación del ADN, no podremos saber nada directamente sobre su bioquímica, su estructura o su función. En muchas ocasiones, será posible deducir indirectamente algunos de

estos rasgos, por comparación con otras proteínas conocidas en otros organismos diferentes, con las cuales la secuencia de la proteína en cuestión presente homología.

Es por ello que, en la actualidad, se están llevando a cabo o ya se han completado, proyectos para secuenciar los genomas completos de algunos organismos escogidos. Se trata de organismos que podrían resultar válidos como modelos de la genética humana, como el ratón; o que presentan potencial como fuente de alimentación, como el arroz; o que pueden tener importancia industrial o médica, como los genomas de algunas bacterias o protozoos parásitos. Por último, otros organismos cuyos genomas están siendo secuenciados o ya han finalizado son los de la planta *Arabidopsis thaliana*, el nemátodo *Caenorhabditis elegans* o la mosca de la fruta *Drosophila melanogaster*.

El estudio de los genomas de estos organismos no humanos nos permitirá obtener avances sin precedentes en la historia de la biología evolutiva. En este sentido ya se han producido las primeras sorpresas; una de las cuáles, ha sido la caracterización genómica de las llamadas «arqueobacterias», organismos procarióticos con características bioquímicas «excéntricas», consideradas, hasta hace muy poco tiempo, como muy primitivas y evolutivamente anteriores a las bacterias verdaderas o «eubacterias», están genéticamente más relacionadas con las células eucariotas que con las eubacterias. Esto ha cambiado la forma del árbol evolutivo principal; como veremos más adelante.

### **LA GENÓMICA EVOLUTIVA ES UNA FUENTE INAGOTABLE DE SORPRESAS**

En 1985, se descubrió que hace cinco a diez millones de años, los gatos adquirieron un virogén o gen viral del papión (un mono africano bastante agresivo) en algún lugar a orillas del mediterráneo. La historia más posible fue que un retrovirus infectó a un papión, incorporándose a uno de los genes del simio. Posteriormente, los descendientes de este retrovirus llegaron de algún modo a parasitar a un único gato o a un pequeño número de gatos que pasaban por allí, transfiriendo el gen del simio al genoma del felino. Toda esta secuencia de acontecimientos ha podido ser reconstruida mediante comparaciones entre algunos de los genes de ambos animales. Del mismo modo, se sabe que existen en ciertos genes de peces que fueron de bacterias y también, en ciertas bacterias cuyos genes que fueron de peces. Los procedimientos «artificiales» de la ingeniería genética para fabricar quimeras no parecen ahora tan nuevos; vemos que la evolución los había inventado ya hace millones de años. Es prácticamente seguro que la investigación conjunta de los genomas de varios organismos, aportará muchas otras sorpresas semejantes.

**QUIENES HAN ADQUIRIDO EL STATUS DE SECUENCIADOS.**

La secuenciación de los genomas de microorganismos abrió las puertas para poder tener instrumentos de conocimiento que pudiesen ofrecer explicaciones a muchas de las enfermedades infecciosas emergentes tanto bacterianas, virales como de otros patógenos humanos. Descontando la secuenciación de más de 600 genomas virales, desde 1995, año en que aparece publicada la secuencia completa del genoma de *Haemophilus influenzae*, hasta el momento, se han secuenciado completamente los genomas de mas de 74 microorganismos; muchos de ellos, patógenos de humanos. En la tabla 7.1 se presenta un resumen de los principales organismos cuyo genoma se ha secuenciado o está en proceso de ser completado.

**Tabla.7.1.** Resumen de los principales organismos cuyo genoma ya ha sido secuenciado o está en proceso de estarlo muy pronto.

> Organismo Procariotes	> Tamaño del genoma (*)	Año de > secuenciación
<i>Haemophilus influenzae</i>	1.830.137 pb	1995
<i>Mycoplasma genitalium</i>	580.070 pb	1995
<i>Methanococcus jannaschii</i>	1.660.000 pb	1996
<i>Helicobacter pylori</i>	1.667.867 pb	1997
<i>Escherichia coli</i>	4.639.221 pb	1997
<i>Bacillus subtilis</i>	4.214.810 pb	1997
<i>Treponema pallidum</i>	1.138.006 pb	1998
<i>Rickettsia prowazekii</i>	1.111 .523 pb	1998
<b>Eucariotes</b>		
<i>Saccharomyces cerevisiae</i>	12.052.000 pb (completo)	1996
<i>Caenorhabditis elegans</i>	97.000.000 pb (completo)	1998
<i>Arabidopsis thaliana</i>	140. 000.000 pb (completo)	2002
<i>Drosophila melanogaster</i>	180.000.000 pb (completo)	2000
<i>Homo sapiens</i>	> 99% genoma total	2001



Dentro de los microbios más importantes asociados con enfermedades en humanos cuyos genomas han sido secuenciados se incluyen *Helicobacter pylori*, cuya infección representa un gran riesgo para la adquisición de un cáncer del estómago denominado carcinoma gástrico; *Borrelia burgdorferi* la bacteria causante de la borreliosis o enfermedad de Lyme; el causante de la sífilis, la bacteria *Treponema pallidum*; *Mycobacterium tuberculosis* bacteria causante de la tuberculosis; *Rickettsia prowazekii*; varias especies del genero *Chlamidya*; *Mycoplasma genitalium*; *Escherichia coli*; además de otros que han sido empleados como modelos de investigación que incluyen bacterias como *Bacillus subtilis*; dos especies de bacterias termofílicas que habitan ambientes extremos marinos *Aquifex aeolicus* y *Thermotoga marítima*; cinco géneros representativos del dominio Archae (Arqueones) y el primer organismo eucariótico en ser secuenciado completamente, la levadura *Saccaromyces cerevisiae*. Recientemente la publicación parcial de la secuencia del genoma del *Plasmodium falciparum* ha abierto un panorama esperanzador para el combate de la malaria.

### **MICOPLASMA, ERES EL MÁS SIMPLE PROCARIÓTICO**

Los Micoplasmas constituyen un grupo de microorganismos que carecen de pared celular; poseen, posiblemente, los genomas de menor tamaño de todos los procariotes e incluyen tanto especies patógenas como otras que forman parte de la flora humana normal. *Mycoplasma genitalium* posee el genoma más pequeño (600 kb) de todos los organismos celulares conocidos que son capaces de reproducirse independientemente. Esto explica por qué ha sido uno de los primeros genomas de organismos no virales que han sido completamente secuenciados. Su pariente cercano, *Mycoplasma pneumoniae*, que es el causante de la neumonía atípica, posee un genoma de similar tamaño. Ambos microorganismos están morfológica y serológicamente relacionados. El éxito de su parasitismo parece estar relacionado con el producto de uno de sus genes, la proteína denominada adhesina, que sirve para adherirse a las células humanas durante la infección. Ésta y otras proteínas relacionadas son necesarias para que se produzca la adhesión y se desarrolle la enfermedad.

La historia natural de *M. genitalium* no está tan bien caracterizada como la de *M. pneumoniae*. El ADN de *M. genitalium* ha sido hallado en tejidos pulmonares de adultos hospitalizados y en muestras de pacientes con infecciones urogenitales. Se cree que *M. genitalium* puede ser la causa de la uretritis (inflamación de la uretra) no gonocócica y que el tracto urogenital puede ser la zona primaria de infección. La presencia del microorganismo es más común en personas homosexuales que heterosexuales.

El genoma de *M. genitalium* fue completamente elucidado en 1995. Su cromosoma circular posee 580.073 pares de bases. El contenido en las bases nitrogenadas G y C de los Micoplasmas es considerablemente menor a la media de otros organismos, variando entre el 27 y el 37%. *M. genitalium* tiene un contenido en G y C del 32%. El análisis del genoma

predice sorprendentemente la pequeña cantidad de 470 regiones codificadoras. 374 de estas regiones corresponden a genes cuyos productos génicos han sido identificados por comparación con otras secuencias de las bases de datos, como proteínas implicadas en la replicación del ADN, la transcripción y la traducción, la reparación del ADN, el transporte celular y el metabolismo energético. 96 regiones codificantes no parecen tener homologías con las secuencias conocidas en las bases de datos. Una característica interesante de la genética de *M. genitalium* es que el codón UGA, que normalmente representa una señal de parada, codifica en este organismo por el aminoácido triptófano.

Cabe esperar que la investigación en sus genomas nos provea de técnicas de tratamiento específico contra los micoplasmas. La lucha por bloquear la proteína adhesina u otras proteínas relacionadas con la adhesión parece ser muy esperanzadora.

### **UNO DE LOS CAUSANTES DE LA GRIPE, *HEAMOPHILUS INFLUENZAE*, ES SECUENCIADO**

*Haemophilus influenzae* es una bacteria gram-negativa que fue aislada durante la epidemia de gripe de 1890, siendo culpada inicialmente de provocar esta enfermedad, cuyo agente ahora sabemos que se trata de un virus. *H. influenzae* es única entre las bacterias anaeróbicas facultativas (se denominan así los microorganismos que pueden crecer en presencia de oxígeno o en su ausencia), ya que depende estrictamente del aporte externo del grupo hemo (un compuesto químico que entra a formar parte, como cofactor, de la hemoglobina y de otras proteínas relacionadas con la cadena respiratoria) para poder crecer de forma aeróbica. Existen seis tipos capsulares distintos de *H. influenzae*, denominados desde la a hasta la f. La cepa no patógena Rd fue, la que finalmente, fue completamente secuenciada. Existen otras cepas de *H. influenzae* que son parte de la flora normal del tracto respiratorio superior. El 95% de las infecciones sistémicas en la infancia son causadas por el serotipo b de esta bacteria. Esto incluye los típicos resfriados, pero también meningitis, sepsis, epiglotitis, neumonía y otitis media. La meningitis bacteriana y la epiglotitis debidas a *H. influenzae* son enfermedades muy graves, con una mortalidad del 5 al 25% de los afectados. Además, deja graves secuelas en el 35% de los afectados de meningitis. Estas estadísticas hacen del estudio de *H. influenzae* un capítulo importante de la lucha contra las enfermedades infantiles.

*H. influenzae* está desarrollando, cada vez más, resistencia a los antibióticos más comunes. El primer hallazgo de una cepa resistente a ampicilina data de 1984. La investigación farmacológica actual se centra en el desarrollo de nuevos antibióticos dirigidos específicamente contra este microorganismo. Se ha hallado además una asociación bastante clara entre la infección por *H. influenzae* y la infección por el virus de la inmunodeficiencia adquirida en humanos (SIDA) el VIH-1.

La única diferencia entre la cepa no infecciosa Rd y las terribles cepas b de *H. influenzae* consiste en la presencia, en esta última, de una serie de ocho genes en tándem que codifican proteínas fimbriales. Las fimbrias son factores de colonización que permiten la adherencia de las bacterias a las células humanas.

El genoma completo de *H. influenzae* fue el primero correspondiente a un organismo autónomo que se terminó, en 1995. El cromosoma circular de este microorganismo posee 1.830.140 pares de bases; el contenido total en G y C es de un 38%. Los autores identificaron 1743 genes potencialmente codificadores de proteínas. 1007 de ellos concordaban con proteínas homólogas en las bases de datos; 347 se parecían a proteínas hipotéticas (proteínas para las cuáles se conoce el gen, pero cuya presencia real no ha sido confirmada) y 389 no parecían tener homólogos conocidos. Recientemente, el genoma de *H. influenzae* ha sido reanalizado, encontrando un nuevo conjunto de genes homólogos a las nuevas entradas de las bases de datos.

#### **LA VIDA ARCAICA MUESTRA SUS SECRETOS GENÓMICOS.**

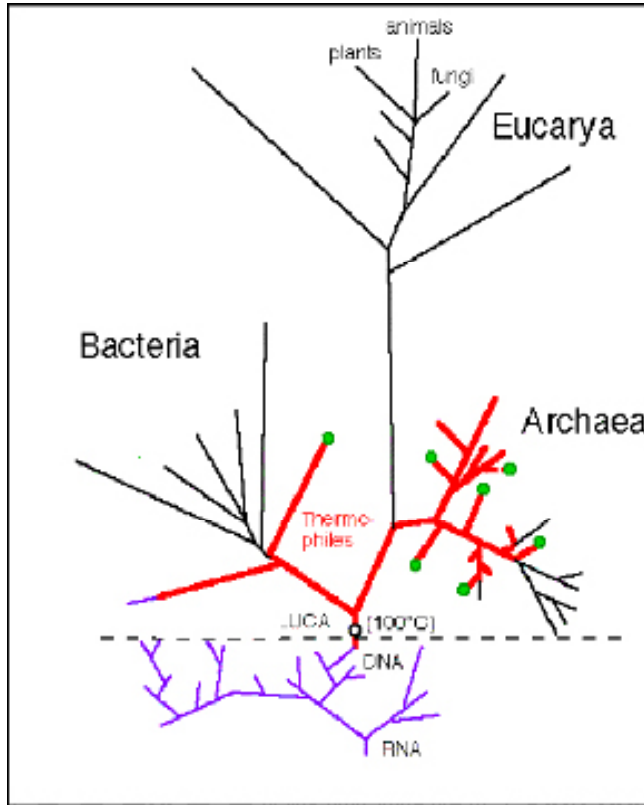
En 1983, una muestra de sedimentos tomada de una fumarola volcánica submarina, recogida por un batiscafo que exploraba el fondo de la costa de la península de Baja California en México, mostró la presencia de un nuevo organismo unicelular, que fue denominado *Methanococcus jannaschii*, en honor del microbiólogo marino Holger Jannasch. El estudio de esta forma de vida, previamente desconocida, ha revelado varios hechos notables relacionados con una de las discusiones más antiguas de la biología; la clasificación de todas las formas de vida. Varias características de *Methanococcus jannaschii* hacen que se incluya a este organismo en la categoría de las Arqueobacterias, un grupo aparte de seres, propuesto inicialmente en 1977 por Carl Woese. Las otras dos grandes ramas del árbol de la vida son los Procariotas, que carecen de núcleo e incluyen las bacterias tradicionales o «Eubacterias» y los Eucariotas, que poseen núcleo e incluyen protozoos, hongos, algas, plantas y animales. Las Arqueobacterias, o Arqueones, poseen características a medio camino entre los procariotas y los organismos superiores; por ejemplo, carecen de núcleo, pero también tienen varios caracteres eucarióticos. *M. jannaschii* muestra una gran variedad de características inusuales: como metanógeno, es capaz de producir metano (gas natural), es un microorganismo termofílico, capaz de vivir a temperaturas entre 48 y

94 °C, es un anaeróbico estricto (muere cuando se sitúa en presencia de oxígeno) y un autótrofo (obtiene toda su energía de fuentes inorgánicas). Este organismo vive a presiones superiores a las 200 atmósferas, usa hidrógeno y dióxido de carbono como sustratos para su crecimiento y la formación de metano y es capaz de fijar nitrógeno. Otra característica sobresaliente (en el sentido literal de la palabra) de *M. jannaschii* es la presencia de dos con-

juntos de flagelos en una formación parecida a un sacacorchos, que se encuentran insertados en un mismo punto de la superficie celular. El genoma de *M. jannaschii* fue el cuarto cuya secuencia fue completamente analizada, en 1996.

Con esta información, los biólogos han sido capaces de establecer comparaciones entre las tres ramas principales del árbol de la vida, sus dotaciones genéticas y su bioquímica. Estos estudios también han sido importantes para comprender el origen de la vida. Por ejemplo, alrededor del 60% de los genes de *M. jannaschii* no presentan homología con ninguno de los genes procariotas o eucariotas que se encuentran en las bases de datos. El 40% que sí la presentan, muestran parecidos tanto con bacterias como con eucariotas. Las secuencias relacionadas con genes bacterianos son las que regulan la producción de energía, la división celular y el metabolismo, mientras que los genes implicados en la transcripción, la traducción y la replicación del ADN son similares a sus homólogos eucarióticos. Además, el genoma de *M. jannaschii* consiste en tres elementos diferenciados; el cromosoma circular principal y dos elementos extracromosómicos circulares, uno mayor y otro menor. El cromosoma posee 1.664.976 pares de bases; contenido en G y C del 31.4%; los elementos extracromosómicos, 58.407 pb (28.2% de G y C) y 16.550 pb (28.8% de G y C) respectivamente. Existen un total de 1738 regiones con potencial codificador, 1682 en el cromosoma, 44 en el elemento extracromosómico mayor y 12 en el menor. La función de los elementos extracromosómicos es aún desconocida y los genes que estos portan pertenecen al 60% de los genes sin homología conocida. La secuenciación ha sido llevada a cabo mediante la colaboración de organismos públicos (Centro Nacional para Recursos Genómicos, Univ. de Illinois y Univ. Johns Hopkins) y privados (TIGR).

El estudio del genoma de *M. jannaschii* nos enseña varias cosas sobre biología evolutiva. Los tres grupos de organismos, arqueones, bacterias y eucariotas parecen compartir un antecesor común, habiendo divergido hace unos 3.000 millones de años cuyo nombre abreviado es LUCA (“Last Universal Common Ancestor”) (figura 8.1). Los Archae o Arqueones y las bacterias comparten características estructurales y de organización tales como los genomas circulares y los genes organizados en operones. Las arqueobacterias son también similares a las bacterias en su metabolismo, pareciendo que sus rutas bioquímicas se derivan del antecesor común. Los arqueones muestran similitudes con los eucariotas en cuanto al procesamiento de la información celular y los sistemas de secreción de sustancias al medio. El refinamiento de los sistemas de transcripción y traducción parecen indicar que los arqueones y evolución de la vida en nuestro planeta.



**Figura 7.1.** El árbol más probable de la evolución de la vida en el planeta Tierra. LUCA que se toma como el primer ancestro universal fue considerado como un microorganismo extremófilo, pues se asume vivió en condiciones de alta temperatura, bajo pH y ausencia completa de Oxígeno.

Los análisis comparativos de los genomas de microorganismos, han permitido completar la historia que los eucariotas comparten rutas bioquímicas comunes e independientes de su linaje bacteriano. Se piensa que el estudio de los genes bacterianos y arqueobacterianos ayudará a dar una explicación de la existencia de la vida en nuestro planeta.

### **¿CON 6000 GENES ES SUFICIENTE PARA UN EUCARIÓTICO COMO LA LEVADURA?**

Definitivamente gran parte de los desarrollos de la genética del siglo pasado, tuvieron como punto de apoyo tres animales modelo que en su conjunto nos llevaron a descifrar una buena parte de los secretos guardados del pasado evolutivo de nuestro planeta. Sin lugar a dudas después de la bacteria *E. coli* y de la levadura *S. cerevisiae*, que son unicelulares, el gusano nemátodo de vida libre *Chaenorabditis elegans*, la mosca de la fruta (*D. melano-*

*gaster*) y el ratón (*Mus musculus*) son los principales modelos animales utilizados en una buena parte de la experimentación en la genómica.

En abril de 1996, fue publicado el genoma completo la levadura *Saccharomyces cerevisiae*, un hongo unicelular, que tan obedientemente nos ayuda a fabricar la cerveza desde tiempo inmemorial. Consta de 12.068 kilobases, repartidas en 16 cromosomas. En su dotación genética se han encontrado 5.885 genes con capacidad potencial para codificar productos protéicos, junto con 140 genes que codifican ARN ribosómico (que forma parte estructural de los ribosomas), 40 genes para ARNs nucleares pequeños (implicados en el “splicing” y en otras funciones variadas en el núcleo celular) y 275 genes para ARNs que participan activamente en la síntesis de proteínas.

La posición de *S. cerevisiae* como organismo modelo se debe a sus ventajas intrínsecas como sistema experimental. Es un organismo unicelular que puede (a diferencia de la gran mayoría de eucariotas) crecer en un medio artificial de composición definida, lo que proporciona un control total sobre su ambiente físico y químico. Las levaduras poseen un ciclo vital que es ideal para los análisis de genética clásica lo cual ha permitido la construcción de un mapa genético detallado de sus dieciséis cromosomas. Más aún, han sido desarrolladas técnicas que permiten la sustitución de cualquiera de sus 6.000 genes por un alelo mutante, o su completa delección del genoma. La combinación de un genoma de tamaño relativamente pequeño, repartido entre un número relativamente grande de cromosomas, permitió dividir fácilmente el trabajo entre los diferentes grupos internacionales de investigación que participaron en el proyecto.

La característica fundamental del genoma de *S. cerevisiae* es su compactamiento genético, encontrándose un gen por cada 2 kb de genoma, en comparación con los genomas de *C. elegans* (un gen cada 6 kb) y humano (un gen cada 30 kb). Incluso es un genoma compacto comparado con otras levaduras y hongos, como la *Schizosaccharomyces pombe*, que contiene un gen cada 2.3 kb. Aproximadamente el 70% de la longitud total del ADN de *S. cerevisiae* se traduce a proteínas. Esto se debe a la notable ausencia de intrones en las secuencias de sus genes; únicamente un 4% de los genes codificantes poseen intrones (la mayoría de ellos son los que codifican proteínas ribosómicas, que suelen contener un único intrón al comienzo de sus secuencia), frente al 40% de los genes de *S. pombe*, que poseen intrones. Incluso se ha sugerido que la mayoría de los genes actuales de *S. cerevisiae* son, en realidad, ADNc procedente del ARN mensajero, generado por la acción de retrotranscriptasas especificadas mediante los retrotransposones (elementos Ty) que se encuentran en número elevado en el genoma de la levadura.

Con la secuenciación completa del genoma de *S. cerevisiae*, se conoció por primera vez el proteoma completo de un organismo eucariótico. El término proteoma ha sido acuñado para definir la dotación completa de proteínas de que dispone un organismo para desarrollar sus tareas vitales. El análisis del proteoma de *S. cerevisiae* permite clasificar un 50% de las proteínas teóricamente codificadas, con base a sus homologías con otras proteínas de

las bases de datos. Estas proteínas de función potencial conocida se pueden clasificar de la siguiente forma: un 11% del proteoma se dedica al metabolismo, el 3% a la producción y almacenamiento de energía, el 3% a la replicación, reparación y recombinación del ADN, el 7% a la transcripción y el 6% a la traducción. Un total de 430 proteínas están implicadas en el tráfico intracelular o traslocación de proteínas de un compartimento celular a otro y 250 proteínas poseen funciones estructurales. Se han identificado cerca de 200 factores de transcripción, así como 250 transportadores primarios y secundarios. Estas estadísticas se refieren únicamente a proteínas cuyos homólogos son conocidos.

Durante cierto tiempo, ha sido un artículo de fe, la creencia de que comprender el genoma de la levadura era un hecho crucial para la comprensión del genoma humano. Esto ha sido demostrado ahora, puesto que más de la mitad de las proteínas cuyas mutaciones causan enfermedades genéticas en el hombre poseen homología en su secuencia con proteínas de la levadura. Aunque es evidente que el genoma humano codifica por muchas más proteínas, comienza a evidenciarse que muchas proteínas de la levadura poseen homólogos en las células humanas. Esto significa que dichas proteínas humanas podrán ser clasificadas con base en su equivalencia funcional o estructural con aquellas del proteoma de la levadura.

Ahora que se conoce el genoma completo de *S. cerevisiae*, los genomas de otras levaduras de interés industrial o médico están al alcance de nuestro conocimiento. La secuenciación completa del genoma de estas especies puede ser innecesaria, si se confirma la sospecha de que numerosas especies de levaduras y hongos muestran un alto grado de sintenia (conservación del orden de los genes). Por ejemplo, estudios recientes sobre *Ashbya gossypii* (un hongo filamentoso que parasita las plantas de algodón) demuestran que la mayoría de sus genes muestran homología con los de *S. cerevisiae*, y que al menos la cuarta parte de los clones de este hongo, que se encuentran en los bancos de datos, poseen grupos de genes en el mismo orden y orientación relativa que los de sus equivalentes en *S. cerevisiae*. Esto proporciona considerables esperanzas sobre el rápido análisis de un gran número de organismos relacionados, tomando el genoma de *S. cerevisiae* como paradigma.

### **EL GENOMA DE CHAENORABDITIS EL PRIMER MULTICELULAR SECUENCIADO**

En animales, los primeros genomas en ser terminados han sido genomas pequeños de organismos que fueron escogidos por su interés como modelos biológicos. Actualmente se tiene la secuencia completa del genoma de este nemátodo de vida libre, el cual ha sido muy utilizado en estudios sobre la diferenciación celular y la embriogénesis animal. Fue el primer genoma de un organismos multicelular en ser secuenciado completamente como resultado de un esfuerzo cooperativo entre el Centro de Secuenciación Genómica de la Universidad de Saint Louis, el Centro Sanger en Hinxton Inglaterra con la ayuda del NIH y el NRC de los Estados Unidos. El genoma de *C. elegans* tiene un tamaño aproximado de 97.000.000 bp el cual codifica por unas 19.000 proteínas.

El proyecto genoma *Chaenorabditis elegans* se inició hacia finales de los años 80 del siglo pasado cuando investigadores de la Universidad de Saint Louis construyeron un primer mapa de ligamiento con base en recombinantes derivados de cromosomas artificiales de levaduras; esta primera aproximación demostró que a pesar de tener clonado un porción apreciable del genoma, todavía un 20% del mismo no había sido clonado. Hacia 1990 se había construido un mapa físico que consistía en aproximadamente unas 20 contigs. El esfuerzo por clonar y secuenciar el genoma de este nemátodo se vió recompensado cuando en 1996 se publicó la secuencia de más del 95% de su genoma conteniendo muy pocos espacios no secuenciados.

El análisis inicial del genoma reveló la existencia de 19.099 genes, cada uno de ellos teniendo en promedio cinco intrones. Un aspecto muy interesante de la distribución de genes fue que la mayor densidad génica está localizada en los brazos de los cromosomas. Las secuencias exónicas cubren un 27% del genoma total, siendo que el 42% de los productos protéicos predichos contiene secuencias homólogas con organismos de otros grupos taxonómicos, esto es un potencial importante para efectuar análisis genómicos comparativos. Se observó también que las secuencias repetidas en tandem son más abundantes en los autosomas que en los cromosomas sexuales. Los cromosomas tienen un contenido GC de 36%. Muchos de los genes de este organismo ya han sido asignados a los distintos cromosomas con lo cual se tiene ya un mapa físico muy completo.

### **EL GENOMA DE LA DROSOPHILA, LA MULTICELULARIDAD VS INFORMACIÓN GENÉTICA**

La secuencia del genoma de *Drosophila* fue publicada en el número del 24 de Marzo de 2000 en la revista *Science*. Los investigadores informan que han secuenciado entre el 97 y el 98 por ciento del genoma y quizás, el 99 por ciento de los 14.000 genes estimados. Los datos de la secuencia están disponibles para los científicos de todo el mundo a través del Genebank, base de datos de secuencias genéticas de los Institutos Nacionales de la Salud.

Con el correr de los años, *Drosophila* ha sido uno de los sistemas modelo más importantes para la investigación en genética. Cuando se aísla el homólogo de *Drosophila* de un gen mamífero importante pero mal entendido, se puede aplicar a su caracterización todo el arsenal de técnicas genéticas del sistema *Drosophila*.

El proyecto de secuenciación del genoma de la *Drosophila* se inició en 1991 como un proyecto cooperativo de Gerald Rubin de la Universidad de California en Berkeley y Allan Spradling de la Institución Carnegie de Nueva York. En Mayo de 1998, el Proyecto Genoma de la *Drosophila* de Berkeley con fondos del NIH había terminado de secuenciar el 20 por ciento. En ese año Craig Venter propuso que su recién formada compañía, Celera, secuenciaría el genoma de *Drosophila* gratis, utilizando una polémica técnica conocida como de fragmentación exhaustiva del genoma («whole genome shotgunning»). La técnica requiere que el ADN de *Drosophila* con los extremos cohesivos se reparta en tres millones de clones al



azar. Estos clones son, entonces, secuenciados automáticamente. Aunque en este caso se dispone solamente de la secuencia de los clones, es importante destacar que espera un arduo trabajo para ensamblar el rompecabezas del posicionamiento de los genes de este organismo.

Venter, después de su salida del NIH, había fundado la compañía Celera con el respaldo de PE Corporation (antes conocida como Perkin Elmer Corporation), que produce las máquinas de secuenciación de ADN, para la aventura comercial de secuenciar el genoma humano para el 2001, varios años antes de la fecha proyectada para la terminación por el Proyecto Genoma Humano internacional. A pesar de que Venter prometía que los datos estarían a disposición de los investigadores, también apostaba a que Celera podría hacer dinero licenciando a la industria farmacéutica para que mire, anticipadamente, los datos de la secuenciación.

El genoma de *Drosophila* terminado, ya parece ser notablemente revelador. Su secuencia completa de  $1.4 \times 10^8$  pb codifica por aproximadamente unos 14.000 genes. De los 289 defectos genéticos que se sabe causan enfermedades en seres humanos, se han encontrado homólogos en *Drosophila* para el 60 %, y para el 70 % de los genes implicados en muchos cánceres humanos. Entre los genes que se han identificado ya están los homólogos en *Drosophila* de los genes implicados en la enfermedad de Parkinson, y el homólogo de *Drosophila*, buscado por mucho tiempo, del gen supresor tumoral p53, que está involucrado en un sinnúmero de cánceres humanos.

La mayor sorpresa que surgió del proyecto de secuenciación de *Drosophila*, es que las moscas tienen solamente el doble del número de genes de las levaduras. Como ya se explicó, una levadura es un hongo unicelular, en contraste con las moscas que son multicelulares; sin embargo, las *Drosophilas* sólo necesitan el doble de los genes para ensamblar funcionalmente un animal que puede volar sin estrellarse contra las paredes, que tiene tejidos, nervios, músculos, memoria y otras clases de comportamientos complicados como ritmos circadianos. El mensaje que nos deja es que la mayor complejidad en animales como moscas y los humanos surge sin necesidad de agregar muchas nuevas piezas; se pueden construir con las mismas partes, sólo agregando algunos genes cuya función sería muy especializada.

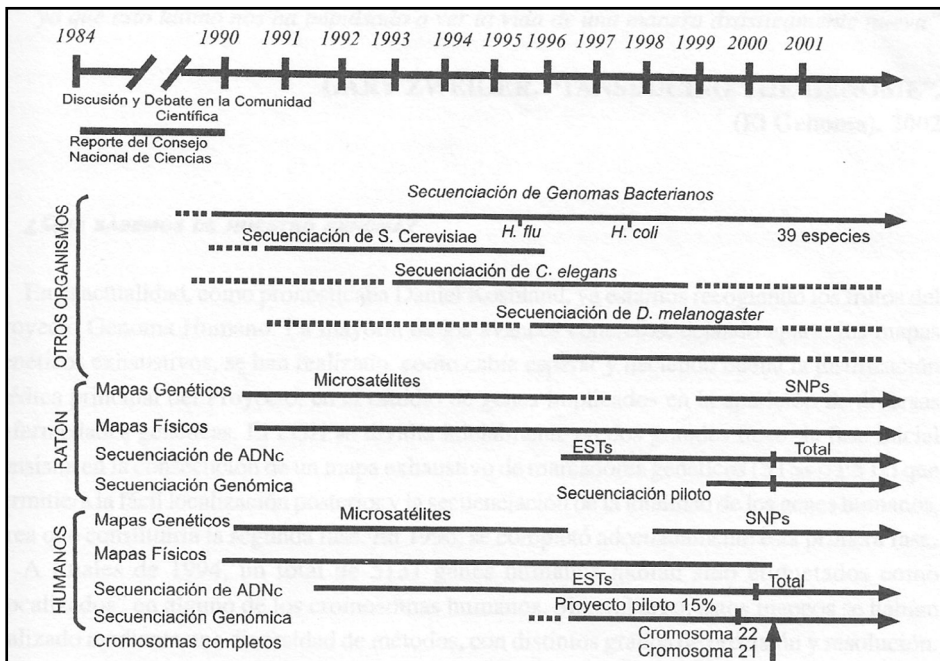
### **EL CONSORCIO GENOMA RATÓN**

Un paso fundamental en la genómica comparativa fue la iniciación de la secuenciación del genoma de un mamífero que sirviera de punto de comparación con el genoma humano. La elección era obvia, el ratón de laboratorio del cual se habían hecho innumerables estudios en genética ofrecía una serie de ventajas y coincidencias con el genoma humano. La investigación del genoma del ratón ha corrido de manera paralela a la desarrollada en el genoma humano por lo que se considera de fundamental importancia la terminación de la secuenciación del genoma del ratón para conocer secretos que nuestro genoma todavía se niega a revelar.

El estudio del genoma del ratón se ha hecho mediante la conformación de un consorcio integrado por los Institutos Nacionales de Salud de los Estados Unidos, la Wellcome Trust y tres compañías transnacionales que son Smith-Kline Beecham, El Instituto de Investigación Genómica de Merck

y Affymetrix Inc., este consorcio aportó US\$5.8 millones en 1998 para acelerar la secuenciación del genoma del ratón que se realizaba en el Instituto Sanger en Inglaterra y en dos Laboratorios de secuenciación de ADN en los Estados Unidos. Como resultado de esta inversión en Febrero del 2001 se obtuvo el primer borrador que incluyó más del 80% de las secuencias clonadas y hacia Diciembre del 2002 se espera tenerlo completamente secuenciado. Muchas y notables semejanzas con el genoma humano han hecho posible obtener un conocimiento valioso sobre la evolución genómica además de darnos un modelo de estudio que es muy parecido al nuestro, con lo cual se espera avanzar en la genómica funcional y la proteómica, para resolver problemas asociados con la salud humana. Se estima que de los 33.924 genes que codifican por proteínas, el 95% tengan un equivalente humano. Tal vez donde se ve una mayor aplicación del genoma del ratón es en la genómica funcional comparativa que permitirá contestar preguntas importantes con respecto a la embriogénesis, la proteómica, el metaboloma y finalmente el transcriptoma.

En la figura 7.2 hemos intentado resumir los principales hitos del desarrollo de los diferentes proyectos genoma, hemos hecho un énfasis especial por mostrar cronológicamente los logros más importantes comparándolos con los del proyecto genoma humano. Esperamos que esta figura les ayude a entender la posición cronológica de la investigación genómica y postgenómica.



**Figura 7.2.** Principales hitos en el desarrollo de los distintos proyectos genoma y los logros que se han obtenido y se espera obtener.

**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

## **¿QUÉ SIGUE EN LA ERA POSTGENÓMICA?**

*“¿Representa el planteamiento de información intensiva un nuevo y revolucionario marco de referencia para la comprensión del cáncer y otros fenómenos biológicos? En absoluto, ya que esto último nos ha impulsado a ver la vida de una manera drásticamente nueva”*

**GARY ZWEIGER. “TANSDUCING THE GENOME”.  
(El Genoma). 2002**

### **¿QUÉ SABEMOS DE NUESTRO GENOMA?**

En la actualidad, como pronosticaba Daniel Koshland, ya estamos recogiendo los frutos del Proyecto Genoma Humano. La mayoría de los avances concretos, dejando aparte los mapas genéticos exhaustivos, se han realizado, como cabía esperar y haciendo buena la justificación médica principal del Proyecto, en el estudio de genes implicados en la aparición de diversas enfermedades genéticas. El PGH se dividía inicialmente en dos grandes fases: la fase inicial consistía en la consecución de un mapa exhaustivo de marcadores genéticos (STSs o ESTs) que permitiera la fácil localización posterior y la secuenciación de la totalidad de los genes humanos, tarea que constituiría la segunda fase. En 1996, se completó adecuadamente esta primera fase.

A finales de 1994, un total de 5131 genes humanos habían sido etiquetados como “localizados” en alguno de los cromosomas humanos. Sin embargo, estos mapeos se habían realizado mediante una diversidad de métodos, con distintos grados de precisión y resolución. A comienzos de 1995, el número de ESTs almacenadas en los bancos de datos genéticos, experimentó un considerable aumento, principalmente debido a los datos aportados, entre

otros, por la multinacional Merck, en colaboración con la Universidad de Washington. En la actualidad, año 2002, más de 950.000 secuencias correspondientes a ESTs, o fragmentos de genes, se almacenan en el GenBank; se estima que la mitad de los genes humanos estén representados en dichas secuencias. Semejante conjunto de secuencias marcadoras ha permitido a un consorcio de investigadores de todo el mundo la realización de un nuevo mapa genético, que cumple, con creces, los requisitos establecidos en las directivas del DOE/NIH.

La revista Science publicó en su número del 25 de Octubre de 1996, una excitante revisión de algunos de los genes hasta ahora descubiertos, así como un mapa de densidad génica de cada uno de los cromosomas. El artículo estaba firmado por 104 investigadores de 18 laboratorios en cinco países. Los mapas de densidad génica tratan de representar gráficamente el número de genes por unidad de longitud localizados en las distintas zonas de cada cromosoma, ante la imposibilidad absoluta de representar individualmente, en un dibujo sencillo y de tamaño lo suficientemente pequeño como para resultar manejable, todos los genes hasta ahora localizados o “mapeados”, que ascienden a más de 16.000. Debemos dejar en claro que no son mapas definitivos y, por lo tanto, una acumulación de genes en una determinada zona no quiere decir que la misma sea especialmente rica en genes, sino que podría deberse a que ha sido mejor estudiada hasta el momento. De todas formas, llama la atención la distinta distribución de genes a lo largo de todo el genoma, existiendo zonas con muy alta concentración, mientras que otras están prácticamente vacías.

### ***LA QUIMERA DEL PROTEOMA.***

Secuenciar el genoma humano no ha sido más que un paso en el camino del objetivo final de tanta inversión, que ha sido conocer lo más exactamente posible el funcionamiento del organismo. Sin embargo, el genoma no es más que el libro de instrucciones generales; pero quienes realizan el trabajo de verdad son las proteínas. El conjunto de todas las proteínas que intervienen en los procesos biológicos de una especie es lo que se conoce como su proteoma y el objetivo que se plantea la proteómica ahora es llegar a determinar la composición, estructura y funciones de todas y cada una de ellas. Los laboratorios que están en primera línea de la investigación se marcan ya nuevas metas y retos que superar. Uno de los retos más apasionantes a los que puede enfrentarse un equipo de científicos es descifrar el secreto de las proteínas humanas. Pasar del genoma al proteoma es una secuencia lógica en la investigación, pues el trabajo básico para descifrar el genoma ha sido determinar el orden en que se encajan las cuatro bases nitrogenadas de la doble hélice de ADN (guanina, adenina, tirosina, citosina). Las proteínas guardan algunos de los enigmas más insondables de la biología. Su forma tridimensional es la que confiere a cada familia de proteínas unas propiedades biológicas particulares. La proteómica y la genómica funcional son nuevas disciplinas dedicadas al estudio de las proteínas que está en plena fase de crecimiento y desarrollo. Las proteínas guardan la clave de numerosas funciones fisiológicas y vitales. Intervienen en los procesos

digestivos, en la transmisión o generación de energía, en la transmisión de impulsos nerviosos, en la defensa inmunitaria, en el transporte de moléculas, etc. Una anomalía en un gen puede permitir un defecto en una proteína o su actividad defectuosa. Si una proteína sufre una deformación, no puede cumplir su función esencial, que es la de encajar con otra proteína.

Enfermedades genéticas como la Anemia Falciforme se deben a un defecto proteico. Otras como el Alzheimer, son el producto de proteínas deformadas. Una infección que se resiste a los antibióticos, una alergia a medicamentos, las responsables son las proteínas. Para solucionar enfermedades de este tipo será en muchos casos, mucho más ventajoso, que las modificaciones genéticas, reemplazar la proteína deficiente o bloquear su actividad.

La transnacional IBM, que ve enormes posibilidades para la informática en todo lo relacionado con la secuenciación humana y manejo de bases de dato genéticos, va a construir para el 2004 un superordenador llamado BLUEGENE, éste será capaz de organizar los movimientos de decenas de miles de átomos en el interior de una proteína en fase de acoplamiento con su gemela, un problema que tardaría en calcularse 300 años usando la tecnología actual.

**LA TABLA PERIÓDICA DEL GENOMA, LOS CROMOSOMAS HUMANOS**

Hacer sinopsis de datos a veces no es fácil, principalmente cuando existe una información incompleta y muy fragmentaria sobre la distribución cromosómica de los genes humanos. Sin embargo a pesar de esto, nos hemos propuesto la tarea de incluir en una tabla los principales rasgos tanto estructurales como informacionales de cada uno de los 23 cromosomas humanos. Esperamos que esta tarea represente una ayuda para poder enfocar de manera más global los impactos del Proyecto Genoma Humano (tabla 8.1)

**Tabla. 8.1.** Principales características de la eucromatina de los 23 cromosomas humanos hasta Mayo del 2002.

<b>Cromosoma</b>	<b>Tamaño de la eucromatina (Kbp)</b>	<b>Número de Contigs secuenciadas</b>	<b>Algunas enfermedades asociadas a mutaciones en genes ligados</b>
1	263.000	440	Síndrome de Chediack-Higashi, Enfermedad de Gaucher, Síndrome de Usher tipo 2
2	255.000	290	Síndrome de Alston, Distrofia tibio-muscular, Holoprosencefalia
3	214.000	351	Alkaptonuria, Síndrome de Von Hippel-Lindau, Leucemia mielode aguda
4	203.000	417	Enfermedad de Huntington, Periodontitis juvenil, Síndrome de Wolfram

5	194.000	396	Cri du Chat, Cancer colorectal, Displasia distrófica, Síndrome de Treacher Collins
6	183.000	36	Enfermedad celíaca, Hemocromatosis, Atrofia espinocerebelar 1,
7	171.000	91	Fibrosis quística, Síndrome de Pallister-Hall, Síndrome de Pender.
8	155.000	257	Síndrome Cohen, Síndrome Langer-Giedon,
9	145.000	103	Disautonomía, Ataxia de Friedreich, Anemia de Fanconi tipo C.
10	144.000	155	Enfermedad de Cowden, Enfermedad de Wolman, Síndrome de Jackson-Weiss
11	144.000	145	Ataxia-telangiectásica, Tumor de Wilms tipo 1, Síndrome QT 1
12	143.000	297	Fenilcetonuria, Enfermedad de Darier,
13	98.000	20	Retinoblastoma, Enfermedad de Wilson, Síndrome de Rieger tipo 2
14	93.000	36	Enfermedad de Alzheimer, Enfermedad de Graves, Paraplejía espástica 3A
15	89.000	121	Esclerosis lateral amiotrófica, Síndrome de Bloom, Enfermedad de Tay Sachs
16	98.000	164	Anemia de Fanconi tipo A, Enfermedad del riñón poliquístico,
17	92.000	136	Cáncer de seno 1, Enfermedad de Canavan, TP53 gen supresor de tumores
18	85.000	117	Cáncer colorectal, Enfermedad de Niemann-Pick tipo C, Síndrome de Tourette
19	67.000	127	Leucemia aguda de células T, Distrofia miotónica
20	72.000	7	Síndrome de Inmunodeficiencia combinada deficiencia en ADA
21	34.000	5	Síndrome de Down, Síndrome de Usher tipo 1E
22	34.491	12	Neurofibromatosis tipo 2, Sarcoma de Ewing, Síndrome del ojo de gato
X	164.000	170	Distrofia muscular tipo Duchenne, Adrenol eucodistrofia
y	35.000	6	Disgenesia Gonadal (SYR)

Toda la información consignada es el punto de partida para entender de manera más integral la existencia de una patología genómica.

### **¿QUÉ SIGUE EN LA ERA POSTGENÓMICA?**

Según Eric S. Lander, del Instituto Tecnológico de Massachussets (M.I.T), el Proyecto Genoma Humano tiene como fin último producir la tabla periódica de los genes. Esta tabla no constará de 108 elementos, sino de unos 45.000 genes. No será un rectángulo con los elementos en perfecto orden de filas y columnas, sino una estructura en árbol mostrando afinidades evolutivas y funcionales entre los genes humanos. La biología molecular ha tendido siempre a examinar los genes individualmente; ahora llega la hora de tomar una perspectiva global del funcionamiento celular, examinando varios genes a la vez. En algunos casos, considerando los casi 45.000 genes y sus productos de un modo simultáneo. Lander propone diez metas a conseguir en los próximos años. Se trata de metas básicamente tecnológicas y de infraestructura más que de objetivos científicos en sí. El autor sostiene que la forma de planificar la ciencia ha de ser centrándonos en el desarrollo de nuevos métodos, ya que en pocos casos pueden predecirse los resultados de los experimentos que se llevarán a cabo utilizando estas nuevas tecnologías. Lander divide sus diez objetivos en cuatro apartados: ADN, ARN, proteínas y sociedad. Veamos a continuación cuáles son dichos objetivos a conseguir en la era post-genómica.

#### **ADN**

##### ***1. RE-SECUENCIACIÓN DE RUTINA DE EXTENSAS REGIONES DE LOS GENOMAS HUMANO Y DE RATÓN.***

Sólo estudiando extensas regiones cromosómicas (de muchos Megabases) podremos localizar regiones de los cromosomas en las que existan genes que confieran predisposición a diversas enfermedades poligénicas, como la diabetes, la hipertensión o la esquizofrenia. Los métodos para re-secuenciar estas largas regiones está todavía en desarrollo.

##### ***2. IDENTIFICACIÓN SISTEMÁTICA DE TODAS LAS VARIANTES DE LOS GENES HUMANOS.***

Identificando los polimorfismos, existentes en todos los genes humanos, podremos llevar a cabo un estudio exhaustivo de las enfermedades y de las mutaciones que las causan.

##### ***3. SECUENCIAMIENTO RÁPIDO DE NOVO DE OTROS ORGANISMOS.***

La secuenciación comparativa proporciona la clave del estudio de la evolución y de la comprensión de cómo la naturaleza es capaz de generar tal diversidad de estructuras y funciones manteniendo, sin embargo, una gran economía en los genes.



## **ARN**

### **4. SEGUIMIENTO SIMULTÁNEO DE LA EXPRESIÓN DE NUMEROSOS GENES.**

Los niveles de ARN mensajero reflejan significativamente el estado de la célula, definiendo el tipo celular, las fases de su desarrollo y las respuestas a los estímulos externos. Para realizar un seguimiento de estos niveles de ARNm, es necesario el desarrollo de técnicas que permitan la medición simultánea de la concentración de estos en la célula.

### **5. HERRAMIENTAS GENÉRICAS PARA MANIPULAR LA RED CELULAR.**

Para llegar a comprender las funciones biológicas, no basta con medir los niveles de expresión de los genes, sino que resulta necesario aprender a manipular e impedir esta expresión. Las futuras investigaciones requerirán de un arsenal de métodos genéricos para manipular la expresión de los genes.

## **PROTEÍNAS**

### **6. SEGUIMIENTO DE LA CONCENTRACIÓN Y ESTADO DE MODIFICACIÓN DE TODAS LAS PROTEÍNAS.**

El caso del seguimiento global de los ARNm se aplica en el mismo grado a las proteínas, con la dificultad añadida de que será necesario controlar las modificaciones post-traduccionales que juegan un papel crucial en las funciones celulares.

### **7. CATÁLOGOS SISTEMÁTICOS DE INTERACCIONES ENTRE PROTEÍNAS.**

Una vez que se conozcan todas las proteínas humanas, podrá ser posible desarrollar un catálogo de interacciones entre las mismas, indispensable para poder estudiar la bioquímica celular. Esto ya ha sido realizado para el fago T7, que posee 55 genes. Finalmente, deben ser explorados otros métodos, distintos de las interacciones físicas entre proteínas, para discernir las rutas bioquímicas.

### **8. IDENTIFICACIÓN DE TODAS LAS FORMAS BÁSICAS DE LAS PROTEÍNAS.**

Los bioquímicos creen que será posible reunir un catálogo de formas tridimensionales básicas de las proteínas, donde se incluyan todas las proteínas, una vez conocida su secuencia. No sabemos si este catálogo se completará fácilmente, ni incluso si será posible su realización.

## **SOCIEDAD**

### **9. INCREMENTO DE LA ATENCIÓN A LOS ASPECTOS ÉTICOS, LEGALES Y SOCIALES.**

A medida que los análisis genéticos aumenten su potencia y disminuyan su coste, el potencial para la generación de aplicaciones intrusivas aumentará de forma exponencial. Los esfuerzos en los estudios ELSI requerirán de una aguda visión científica para anticipar los problemas y proponer soluciones.

### **10. ÉNFASIS EN LA EDUCACIÓN PÚBLICA.**

Los ciudadanos se enfrentarán a la posibilidad de obtener visiones globales de sus propios genomas y de interpretar semejante información. En la actualidad, el público en general posee una rudimentaria visión de la genética y carece de los medios para aprender más. Deberá realizarse un gran esfuerzo para educar a los escolares, al público en general, a los médicos y a los consejeros genéticos.

En resumen, los métodos científicos y tecnológicos avanzan de modo inexorable. La biología del siglo XXI promete ser mucho más excitante y poderosa que la genética del siglo XX. En general, no podemos impedir este avance, ni sería razonable oponernos al mismo, dado el gran potencial de mejora para toda la sociedad que es parte inherente de la era postgenómica. La única posibilidad es la educación, la posibilidad de que los maravillosos conocimientos de que dispondremos no sean únicamente patrimonio de un escaso grupo de científicos, sino que formen parte del acervo cultural de la humanidad, para que cada individuo sea capaz de aplicarlos libremente en su propio provecho. Debemos familiarizarnos con determinados conceptos y con las posibilidades que se abrirán ante nosotros, pero también con las posibilidades de manipulación que se tendrán para determinados grupos de individuos cuyas motivaciones pueden no ser éticamente lícitas. Quizás sea difícil al principio, pero somos optimistas en este aspecto. Confiamos plenamente en que la capacidad racionalizadora del hombre será suficiente para hacer frente a todas las posibilidades negativas. Ya lo hemos hecho una vez, con el poder, más mortífero si cabe, de las armas nucleares. ¿Por qué no íbamos a poder hacer frente a las manipulaciones de la genómica?

**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

## **HACIA UNA PATOLOGÍA GENÓMICA**

*“La genética mendeliana es tan apropiada para comprender la herencia del mundo real como lo es la geometría euclidiana para comprender la forma de un roble”*

**MATT RIDLEY.**  
**“Genoma”. 1999**

La principal justificación del PGH y la que, sin duda, más contribuyó a la decisión de acometer tan enorme tarea, era la posibilidad de conocer los genes relacionados con las distintas enfermedades de origen genético que afligen al hombre, para encontrar eventualmente un tratamiento o una cura para las mismas. Existen más de 4000 enfermedades «de un sólo gen», cuya herencia sigue las leyes de la genética mendeliana y puede ser seguida mediante estudios de familias afectadas, estas enfermedades son las que se han definido de manera poco precisa como monogénicas. El número de trastornos «poligénicos», cuya herencia se debe a defectos acumulados al azar en varios genes, puede llegar a doblar el anterior, aunque no se conoce con precisión si serán todas las enfermedades incluyendo aquellas definidas como monogénicas. Si aceptamos que existen enfermedades monogénicas como tales, éstas se pueden dividir en enfermedades recesivas, enfermedades dominantes y enfermedades ligadas al sexo.

Las enfermedades recesivas sólo se manifiestan si un hijo cuyos padres son ambos «portadores sanos» hereda los dos alelos recesivos del gen anómalo. Una enfermedad recesiva puede mantenerse latente en sucesivas generaciones de una familia, manifestándose solamente cuando uno de sus miembros tiene descendencia con otro portador del gen defectuoso. Por ello, las enfermedades recesivas son especialmente frecuentes entre grupos culturales o étnicos que mantienen un alto grado de endogamia. Es el caso de los judíos Azkenazíes, de los Amish o de las familias reales europeas. Un hijo de padres portadores posee sólo un 25% de probabilidad de padecer la enfermedad, pero también un 50% de ser un portador sano.

La mayoría de las enfermedades genéticas más extendidas son de este tipo, como la anemia falciforme o la fibrosis quística.

En las enfermedades dominantes, basta con que el hijo herede un sólo alelo del gen defectuoso para que sufra la enfermedad, de modo que es suficiente que uno de los padres sea portador (y por tanto, sufra la enfermedad) para que el hijo posea un 50% de probabilidades de padecerla. Las enfermedades dominantes que sean especialmente graves, manifestándose pronto durante la infancia, tienden a desaparecer rápidamente, ya que sus portadores no pueden llegar a reproducirse. Sin embargo, las enfermedades que aparecen de forma tardía, como la enfermedad de Huntington, pueden llegar a extenderse por la humanidad, ya que sus víctimas pueden llegar a reproducirse, transmitiéndola al 50% de sus descendientes. El caso de las enfermedades ligadas al sexo es algo más complejo. Se trata de enfermedades recesivas monogénicas cuyo gen está localizado en el cromosoma X. Si el descendiente es una mujer, como posee dos cromosomas X similares, es muy difícil que llegue a manifestar la enfermedad, porque tendría que poseer los dos alelos recesivos. Sin embargo, si el descendiente es un varón, como sólo posee un cromosoma X, bastaría con que heredara el alelo recesivo de su madre para que sufriera la enfermedad. Por ello, si una mujer es portadora, sus hijos varones poseen un 50% de probabilidades de padecer la enfermedad, aún cuando sus hijas nunca la padezcan (si el padre no porta la enfermedad). Es por ello que estas enfermedades aparecen con muchísima mayor frecuencia en varones que en mujeres y son las más fáciles de localizar en un cromosoma (el X). Son enfermedades con herencia ligada al sexo la hemofilia, el daltonismo y la distrofia muscular de Duchenne.

A continuación, incluimos a manera de ejemplo, una lista de algunas enfermedades genéticas y genes asociados a ellas, así como otros genes relacionados. No están ordenadas por orden de gravedad, ni por orden alfabético. Al contrario, hemos seguido un orden que no podría haberse seguido hace diez años, antes de que se iniciara el PGH, las hemos ordenado según el número de cromosoma donde se encuentra el gen que las produce.

### ***ENFERMEDAD DE GAUCHER.***

La enfermedad de Gaucher, llamada así en honor del médico francés Phillippe Gaucher (1854-1918), es producida por una mutación recesiva en el gen que codifica la proteína glucocerebrosidasa, que se localiza en el cromosoma 1. Esta enzima es la que normalmente metaboliza un tipo de lípidos llamados glucocerebrósidos. En los enfermos de Gaucher, estos lípidos no pueden ser descompuestos y se acumulan principalmente en el hígado, en el bazo y en la médula ósea. Los síntomas de la enfermedad de Gaucher incluyen fuertes dolores, fatiga, ictericia, daños óseos, anemia y muerte. Esta enfermedad es el trastorno genético más común entre los judíos de ascendencia europea oriental (Askenazíes), con uno de cada 450 nacimientos; aunque puede afectar a individuos de cualquier grupo étnico. En el conjunto de la población, afecta a uno de cada 100.000 nacimientos. En 1991, se hizo posible la primera

terapia efectiva contra esta enfermedad, inyectando la enzima sintetizada en *E. coli* en el torrente sanguíneo de los enfermos; este procedimiento denominado terapia de reemplazo enzimático es una luz de esperanza para muchas otras enfermedades génicas. El tratamiento es similar a la diálisis de los pacientes enfermos de riñón, durando unas dos horas y debiendo repetirse cada dos semanas. La terapia, cuyo desarrollo ha sido posibilitado directamente por el PGH, detiene el avance de los síntomas y, en muchos casos, los revierte, permitiendo a los pacientes mejorar sustancialmente su calidad de vida.

### **ENFERMEDAD DE ALZHEIMER.**

La enfermedad de Alzheimer, denominada así en honor del neurólogo alemán Alois Alzheimer (1864-1915), es una enfermedad degenerativa que destruye el cerebro, haciendo que los enfermos pierdan la memoria y el juicio e impidiendo normalmente que se valgan por sí solos. En algunos casos, puede comenzar en la mediana edad, pero su probabilidad de incidencia aumenta de forma espectacular con la edad del paciente. La enfermedad de Alzheimer aflige al 10% de las personas con más de 65 años y a casi el 50% de las personas con más de 80 años. La muerte ocurre normalmente dentro de los 5 a 10 años del inicio de los síntomas. En los países desarrollados, el Alzheimer es la cuarta causa principal de mortandad en los adultos, tras las enfermedades cardiovasculares, los accidentes de tráfico y el cáncer. Además, los costos que produce el Alzheimer a los sistemas públicos de salud son realmente elevados y se espera que el número de pacientes con Alzheimer se incremente dramáticamente en las próximas décadas, a medida que se incrementa la esperanza media de vida en los países desarrollados.

No existe un método seguro para su diagnóstico, excepto la autopsia. La enfermedad de Alzheimer puede ser de origen genético en un 20% de los casos documentados y debida a otros factores diversos en el 80% restante. Además, se han localizado varios marcadores para el Alzheimer de origen genético en los cromosomas 1, 14, 19 y 21, lo que ha llevado a diferentes tipos de Alzheimer, desde el tipo I hasta el IV. Es interesante que una trisomía en el cromosoma 21 (la presencia de un cromosoma 21 extra en cada célula) es la causante del síndrome de Down, la forma más común de retraso mental. Las personas con síndrome de Down, si viven lo suficiente, casi siempre suelen desarrollar la enfermedad de Alzheimer.

### **CÁNCER DE COLON.**

El cáncer ocurre cuando las células comienzan a dividirse sin orden ni control. En realidad, se denomina cáncer a un conjunto de más de 200 enfermedades bastante diferentes entre sí. El cáncer de colon es el segundo más frecuente, tras el cáncer de pulmón. En 1996, se diagnosticaron 94.500 nuevos casos, sólo en Estados Unidos y más de 50.000 personas mueren cada año debido a esta enfermedad. Los síntomas de esta enfermedad incluyen hemorragias

rectales, presencia de sangre en las heces y cambios en la frecuencia de visitas al servicio. El exámen rectal digital, el análisis de las heces y la sigmoidoscopia, análisis exploratorio del recto, permiten detectar precozmente el cáncer de colon, antes de que se produzcan los síntomas. Si se detecta a tiempo, la posibilidad de supervivencia en los siguientes 5 años es del 91%.

Existen varios factores de riesgo para el cáncer de colon. Los antecedentes familiares y la propensión a enfermedades que provocan inflamaciones intestinales son los más importantes. También podrían influir la vida sedentaria y la dieta rica en grasas y pobre en fibra. Estudios recientes demuestran que la terapia con estrógenos y los anti-inflamatorios no esteroideos (AINEs), como la aspirina, pueden reducir el riesgo de padecer cáncer de colon.

Recientemente, se ha descubierto un gen mutante en el cromosoma 2 que aparece con frecuencia en determinados cánceres de colon, que ha sido llamado MSH2. El producto del gen es una enzima que presenta homología con otras enzimas bacterianas que intervienen en los mecanismos de reparación del ADN, por lo que su carencia o mal funcionamiento podrían explicar el desarrollo incontrolado de las células cancerosas. Los tratamientos utilizados en la actualidad incluyen la cirugía, la quimioterapia, la radioterapia y la terapia biológica.

### **CÁNCER DE PULMÓN.**

El cáncer de pulmón es la forma más común de cáncer en los países desarrollados, igualmente en hombres que en mujeres. Los estadounidenses poseen la más alta tasa de cáncer de pulmón en el Mundo. En 1996 se diagnosticaron 177.000 nuevos casos en los Estados Unidos y 158.700 personas murieron. La probabilidad de sobrevivir cinco años tras el diagnóstico es únicamente de un desesperanzador 13%.

Desde 1940, el incremento en la mortalidad por cáncer de pulmón, según el sexo, ha seguido la pauta marcada por la adicción al tabaco, con un retraso de 20 años. También se ha demostrado que dejar de fumar reduce la probabilidad de padecer cáncer de pulmón, a cualquier edad. Aunque el tabaquismo es, con diferencia, la principal causa del cáncer de pulmón, no es la única. La exposición a los humos exhalados por otros fumadores incrementa el riesgo de padecer cáncer de pulmón entre los «fumadores pasivos». También la exposición a otros carcinógenos, como el amianto, principalmente en el lugar de trabajo, puede favorecer su aparición y los trabajadores (especialmente los fumadores) que están expuestos a altas dosis de radón, un gas radiactivo, presentan altas tasas de incidencia de esta enfermedad.

El cáncer de pulmón no suele causar síntomas externos durante sus primeras etapas de desarrollo. Sin embargo, con el paso del tiempo, aparece una tos persistente, un dolor constante en el pecho, jadeos, ataques repetidos de neumonía y bronquitis, aparición de sangre al toser y ronquera. Como todos los cánceres, el de pulmón puede causar fatiga, pérdida de apetito y de peso. El cáncer de pulmón normalmente se expande a los nódulos linfáticos y a otros tejidos del pecho (incluyendo el otro pulmón). A menudo se puede extender a otras

partes del cuerpo, como los huesos, cerebro o hígado. Si esta expansión (metástasis) ocurre, aparecen dolores de cabeza y fracturas óseas.

El tratamiento del cáncer de pulmón depende de numerosos factores, incluyendo el tipo de cáncer, su tamaño y localización y el avance de la enfermedad. Usualmente se recurre a la cirugía, la radioterapia y la quimioterapia; en la mayoría de los casos, con escasos resultados positivos.

Los científicos han localizado un gen en el cromosoma 3, llamado SCLC1, que predispone a padecer cáncer de pulmón. Sin embargo, es esta una enfermedad que depende del ambiente en la inmensa mayoría de los casos.

### ***ENFERMEDAD DE HUNTINGTON.***

La enfermedad de Huntington (HD), llamada así en honor del médico estadounidense George Sumner Huntington (1850-1916), es una enfermedad degenerativa del cerebro, que conduce a un deterioro mental que termina en demencia. Normalmente comienza a aparecer entre los 30 y los 50 años, aunque se han diagnosticado casos tan pronto como a los 2 años y tan tarde como a los 70. Es obvio que posee un amplio espectro de severidad.

Los síntomas incluyen cambios en la personalidad y en el estado de ánimo, depresión y pérdida gradual del control sobre los movimientos voluntarios, causando espasmos primero y grandes movimientos al azar posteriormente. Estos movimientos le valieron el nombre antiguo, no exento de un toque de humor negro, de «*corea de Huntington*»; corea significa «movimiento de baile». Trágicamente, las personas con la enfermedad mantienen una inteligencia social, una conciencia de que ya no pueden controlar sus facultades físicas y son incapaces de comunicar sus necesidades y sentimientos a los demás.

La enfermedad de Huntington afecta a todas las edades y grupos étnicos y a ambos sexos por igual. A su crueldad se añade el hecho de ser una enfermedad provocada por un gen dominante. Es decir, si uno de los padres posee el gen (y por lo tanto, ha desarrollado la enfermedad), los hijos poseen un 50% de padecerla. La HD no se salta generaciones. Si no se hereda el gen, no se puede transmitir a la descendencia. Del mismo modo, si se hereda el gen, inevitablemente se padecerá la enfermedad, más tarde o más temprano.

En 1993, tras una búsqueda de diez años, los científicos aislaron el gen que provoca la EH, localizado en el cromosoma 4. Ahora los esfuerzos se centran en averiguar la función exacta del gen en su estado normal y como el gen mutante puede causar tantos daños. Además, se trata de buscar las razones que hacen que la enfermedad se manifieste de forma tardía, para poder utilizarlas en nuestro provecho. Muchas líneas de investigación están en marcha con el fin de encontrar un tratamiento y una cura.

Con el descubrimiento del nuevo gen, además, se desarrolló una nueva prueba genética para detectar si había sido heredado por un individuo «en riesgo» o no. El nuevo test únicamente requiere de un sencillo análisis de sangre del «paciente», a diferencia de la prueba



antigua, que requería muestras de varios miembros de su familia. Esta prueba, inevitablemente, plantea a las personas en riesgo el dilema de conocer previamente, a veces con varias décadas de adelanto, si van a padecer una enfermedad terrible y mortal. También plantea la posibilidad de conocer, durante las primeras semanas de embarazo, si un futuro hijo de una persona afectada posee el gen. Así mismo, se plantea la duda razonable de si es ético o no abortar a una persona que quizás pueda llevar una vida normal y completamente sana durante al menos treinta años, con posibilidad de que ésta se alargue hasta los sesenta o los setenta. Una de las personas famosas afectadas de la enfermedad fue el conocido cantante folk estadounidense de los años sesenta, Woody Guthrie. Además, se da la ¿coincidencia? de que una de las investigadoras que más han contribuido al estudio del gen HD, Nancy Wexler, una psicóloga clínica estadounidense, perdió a su madre a causa de esta enfermedad. La doctora Wexler ha expresado públicamente su deseo de no someterse a la prueba genética. Prefiere la misericordia de la incertidumbre a conocer antes de tiempo una amarga y dura verdad.

### ***DISPLASIA DISTRÓFICA.***

La displasia distrófica es una anomalía del crecimiento caracterizada por huesos curvados, extremidades cortas y deformaciones en las articulaciones, los pies y las manos. Se ha descubierto un gen, en el cromosoma 5, asociado a esta enfermedad. Ha recibido el nombre de DTD. Su función permanece aún desconocida.

### ***DIABETES JUVENIL.***

La diabetes es una enfermedad metabólica crónica que impide que las células pancreáticas sinteticen insulina, la hormona que controla el nivel de glucosa en la sangre, controlando la conversión de glucosa a energía. La diabetes juvenil, también llamada diabetes de tipo I o diabetes insulino-dependiente, es la forma más severa; puede aparecer a cualquier edad, aunque suele diagnosticarse desde la infancia hasta los treinta años; obliga a los pacientes a inyectarse insulina varias veces al día, para poder permanecer vivos. Aunque un tratamiento cuidadoso, tanto por parte del médico como del paciente, puede incrementar enormemente la calidad de vida de los diabéticos, esta enfermedad sigue siendo una de las causas de muerte por enfermedad más frecuentes. Numerosas enfermedades están asociadas a la diabetes: las enfermedades cardiovasculares causan el 75% de las muertes en los diabéticos, la hipertensión y el infarto son hasta cuatro veces más frecuentes en diabéticos que en personas sanas, las cataratas suelen provocarles ceguera y anualmente es necesario amputar más de 50.000 miembros, sólo en los Estados Unidos, debido a necrosis diabética.

La diabetes juvenil posee una causa genética, aunque es una enfermedad poligénica de herencia complicada. Un gen, llamado IDDM1 localizado en el cromosoma 6, ha sido asociado con esta enfermedad. Su aparición tiene una causa inmunológica: los linfocitos T, por alguna

razón aún no bien comprendida, atacan y destruyen a las células productoras de insulina del páncreas. La terapia génica está siendo investigada, aunque aún no se han realizado pruebas en seres humanos.

### ***FIBROSIS CÍSTICA.***

La fibrosis cística es la enfermedad genética mortal más común entre los individuos caucásicos, afectando a uno de cada 1800 caucásicos (también afecta a uno de cada 17.000 descendientes de africanos). Sus síntomas aparecen en la infancia y se caracterizan por una infección crónica del pulmón, una función pancreática anormal y un alto contenido en sal en la transpiración. Esto se debe a un defecto en el gen que codifica una proteína que actúa como canal de cloruro, permitiendo la entrada de este ión en las células. Como consecuencia de esto, las células producen un mucus grueso que obstruye los pasajes respiratorios, sirviendo como medio de cultivo ideal para las bacterias patógenas. Además, se obstruyen los conductos del páncreas y el hígado, impidiendo a las enzimas digestivas realizar su función adecuadamente. Los niños afectados de fibrosis cística suelen sucumbir por las infecciones respiratorias. Después de siete años de búsqueda intensiva, en 1985, un equipo de investigadores canadienses y estadounidenses localizó el gen de la fibrosis quística en el cromosoma 7. Se trata de un gen recesivo que es portado por uno de cada veinte blancos, que no manifiestan la enfermedad. Si dos portadores sanos tienen un hijo, éste tiene un 25% de posibilidades de padecer la enfermedad y un 50% de ser portador sano. En 1989 se consiguió aislar el gen y comenzó su secuenciamiento.

En 1990, los científicos consiguieron que cultivos de células defectuosas fueran transformadas con el gen sano y recuperaran sus niveles normales de cloruro. La primera terapia génica a nivel experimental fue administrada a un paciente en 1993, con resultados no concluyentes. En la actualidad, se están estudiando numerosas líneas de investigación en varios laboratorios de todo el mundo, incluyendo la posibilidad de suministrar terapia génica de forma tan fácil como mediante un inhalador nasal. Si esto tiene éxito, significaría el fin del sufrimiento de cientos de miles de personas.

### ***OBESIDAD.***

La obesidad es un exceso de grasa en el cuerpo que frecuentemente conlleva una serie de problemas médicos asociados. Médicamente, se considera que las mujeres con más del 25% de su peso en grasa y los hombres con más del 30% son obesos. Se puede tener sobrepeso sin ser obeso; por ejemplo, un culturista con una gran masa muscular; sin embargo, a efectos prácticos, la mayoría de la gente que posee sobrepeso, padece también de obesidad. Una de cada cuatro personas en los países desarrollados está en riesgo para su salud, por ser obesos. En efecto, la obesidad es un factor de riesgo reconocido para algunas enfermedades

crónicas, como las cardiovasculares, la diabetes, la hipertensión y algunas formas de cáncer. También podría contribuir negativamente en los cálculos biliares, la osteoartritis y los problemas respiratorios, como la apnea. La evidencia sugiere que la obesidad puede tener varias causas: genética, ambiental, psicológica y otros factores pueden tener parte de la culpa. El gen mutante de la obesidad (Ob), que ha sido encontrado en los ratones, provee a los científicos de un buen modelo para estudiar la obesidad humana. La obesidad tiende a manifestarse en familias, por lo que podría tener una causa genética; sin embargo, esto podría resultar engañoso, ya que los miembros de una familia no sólo comparten sus genes, sino también sus costumbres y su dieta. Separar las causas genéticas de las ambientales supone a veces un desafío. En la actualidad, un número cada vez mayor de evidencias tienden a señalar la importancia del factor genético. En un estudio sobre adultos que fueron adoptados cuando eran bebés, los pesos de los sujetos adultos se correlacionaban mejor con los de sus padres biológicos que con los de sus padres adoptivos. El ambiente, que depende de la familia adoptiva, posee aparentemente menor influencia que los genes.

Los científicos han sido capaces de determinar la proteína mutada en los ratones obesos. Se trata de un polipéptido denominado leptina. La misma proteína se encuentra en el hombre. Está codificada por un gen localizado en el cromosoma 7. Queda por demostrar su relación con algunos casos de obesidad patológica.

Uno no puede cambiar sus genes, pero puede cambiar su dieta y su nivel de actividad. Algunas personas han sido capaces de perder una cantidad considerable de peso, siguiendo una dieta más nutritiva con menos grasa, aprendiendo a reconocer las características ambientales que estimulan su apetito (como los olores tentadores) y aumentando su actividad física. En algunos casos extremos, puede ser recomendable la cirugía gastrointestinal, que puede tener resultados espectaculares.

### ***LINFOMA DE BURKITT***

El linfoma de Burkitt es un tipo de cáncer poco frecuente, que afecta predominantemente a niños de África Central, aunque ha sido descrito también en otras muchas zonas. Se llama así en honor de Denis Parsons Burkitt, cirujano inglés nacido en 1911, que ejerció en Uganda y describió la enfermedad. Se suele manifestar como un gran tumor en la mandíbula, que rápidamente avanza hasta invadir las órbitas oculares. También suelen aparecer tumores en la cavidad abdominal y en el sistema nervioso central. El tratamiento, que se realiza mediante radioterapia y quimioterapia, es poco efectivo. Este cáncer se ha relacionado con una translocación de una parte del cromosoma 8, que afecta a un oncogen, que ha sido denominado c-myc.

### **NEOPLASIA ENDOCRINA MÚLTIPLE**

La neoplasia endocrina múltiple (MEN) es, en realidad, un grupo de enfermedades poco frecuentes, caracterizadas por defectos genéticos que conducen a la hiperplasia (multiplicación anormal o incremento anormal en el número de células normales de un tejido) y la hiperfunción (sobrefuncionamiento) de dos o más componentes del sistema endocrino. Las glándulas endocrinas son todas aquellas que segregan hormonas a la corriente sanguínea. Las hormonas son potentes transmisores químicos que viajan a través de la sangre, controlando y dirigiendo el funcionamiento de los distintos órganos. Normalmente, la cantidad de hormonas en la sangre en cada momento está cuidadosamente ajustada según las necesidades del cuerpo. Cuando una persona padece de MEN, ciertas glándulas, como las paratiroides, el páncreas o la pituitaria, se hacen hiperactivas, produciéndose un imbalance hormonal. Los resultados pueden ser variados: acumulación de calcio en el torrente sanguíneo (lo que conlleva la aparición de cálculos y daños renales), fatiga, debilidad, dolor muscular y esquelético, estreñimiento, indigestión o delgadez de los huesos.

Todas las formas de MEN se transmiten como enfermedades autosómicas dominantes. La de tipo 2A se ha localizado en un gen del cromosoma 10. Se caracteriza por la aparición de tumores en la glándula pituitaria, el tiroides, las paratiroides y las suprarrenales.

### **EL ONCOGÉN RAS DE HARVEY**

Este oncogén, localizado en el cromosoma 11, aparece mutado en una multitud de cánceres humanos de diversos tipos. Su producto protéico es un polipéptido denominado RAS, cuya función permanece aún desconocida.

### **FENILCETONURIA**

La fenilcetonuria es una rara enfermedad metabólica causada por una deficiencia en la enzima fenilalanina-hidroxilasa, que se sintetiza en el hígado de las personas sanas; la fenilalanina-hidroxilasa (PAH) cataliza la conversión de fenilalanina en tirosina. La deficiencia de esta enzima impide que los enfermos metabolicen correctamente el aminoácido esencial fenilalanina, que se acumula en los tejidos corporales, apareciendo los productos de oxidación del mismo en la orina. La fenilcetonuria es una enfermedad progresiva que produce retraso mental si no es tratada desde muy temprano. Otros síntomas pueden ser erupciones cutáneas, ataques de nervios, comportamiento irritable y un olor corporal acre.

Afortunadamente, el tratamiento es bien sencillo, si se detecta la enfermedad a tiempo. Basta con dar al enfermo una dieta especial, pobre en fenilalanina, durante las primeras semanas de vida y mantener una dieta equilibrada durante toda su vida. La fenilcetonuria se puede detectar inmediatamente después del nacimiento mediante un sencillo análisis de

sangre, obligatorio en la legislación de numerosos países, por lo que, en realidad, sus efectos en raras ocasiones llegan a manifestarse, siendo una de las enfermedades genéticas más «benignas» y una de las pocas con tratamiento efectivo. La dieta de los fenilcetonúricos ha de ser necesariamente pobre en proteínas, requiriéndose un control médico continuo de la misma. El edulcorante artificial aspartame, que se ha puesto de moda en los últimos años en bebidas dietéticas y chicles sin azúcar, es un dipéptido formado por una fenilalanina unida a un ácido aspártico modificado, por lo que no puede ser consumido por los fenilcetonúricos. La fenilcetonuria es una enfermedad monogénica localizada en el cromosoma 12.

### **ENFERMEDAD DE WILSON**

La enfermedad de Wilson, denominada así por el neurólogo inglés Samuel Alexander Kinnier Wilson (1877-1937), es una extraña enfermedad genética caracterizada por la acumulación de cobre en diversos tejidos corporales, especialmente en el hígado, el cerebro y la córnea. Esta acumulación anómala se atribuye a un defecto en el metabolismo del cobre. La enfermedad tiene herencia autosómica recesiva y el gen responsable, denominado ATP7B, se localiza en el cromosoma 13.

La enfermedad de Wilson se caracteriza por cirrosis hepática y cambios degenerativos en el cerebro. Los trastornos hepáticos son los síntomas más frecuentes en los niños afectados, mientras que los trastornos cerebrales son más frecuentes en los adultos jóvenes. En el ojo, la característica principal es un anillo pigmentado, denominado anillo de Kayser-Fleischer, en el margen externo de la córnea.

### **SÍNDROME DE MARFÁN**

El síndrome de Marfán, llamado así por el pediatra francés Antonin Bernard Jean Marfán (1858-1942), es una enfermedad congénita del tejido conjuntivo, que afecta a numerosos órganos y sistemas, incluyendo el esqueleto, los pulmones, los ojos, el corazón y los vasos sanguíneos. La enfermedad se caracteriza por un crecimiento anormal de las extremidades (especialmente de los dedos), una dislocación parcial del cristalino (en el 50% de los pacientes), anomalías cardiovasculares (la arteria aorta suele ser más ancha y más frágil que en las personas normales) y otras deformaciones.

La enfermedad afecta a hombres y mujeres de todos los grupos étnicos. Se estima que 40.000 personas en Estados Unidos la padecen. El síndrome de Marfan es difícil de diagnosticar, ya que no existe una prueba específica. Además, las características de la enfermedad varían ampliamente de un individuo a otro; en algunos casos resulta bastante benigna. Algunos médicos actuales piensan que Abraham Lincoln padeció esta enfermedad.

El síndrome de Marfán es una enfermedad autosómica dominante, como la Enfermedad de Huntington, por lo que los descendientes de personas afectadas poseen el 50% de po-

sibilidades de padecerla. Recientemente, han sido asociadas al síndrome de Marfan distintas mutaciones en el gen de la fibrilina, una proteína esencial del tejido conjuntivo, cuyo gen se localiza en el cromosoma 15. Esto ofrece grandes posibilidades respecto al desarrollo de una prueba de diagnóstico y quizás, un tratamiento y una cura.

### ***ENFERMEDAD DEL RIÑÓN POLIQUÍSTICO.***

Los enfermos afectados de la enfermedad poliquística del riñón (PKD) se caracterizan por presentar numerosos y grandes quistes en uno o en ambos riñones. Los pacientes generalmente suelen morir por fallo renal o debido a las consecuencias de éste, como la hipertensión (elevada presión arterial). Un gen asociado a esta rara enfermedad ha sido localizado en el cromosoma 16 y ha sido denominado PKD1. Se espera localizar otros genes implicados en la aparición de este trastorno, presumiblemente poligénico.

### ***SUPRESOR DE TUMORES TP53.***

TP53 es un gen, localizado en el cromosoma 17, cuyo producto protéico se engloba en el grupo de los «supresores de tumores». Son estas proteínas cuya ausencia o mal funcionamiento provocan, normalmente, que las células crezcan de forma incontrolada provocando un tumor o cáncer. El TP53 es uno de los supresores más importantes. Cuando está mutado o delecionado, el individuo puede desarrollar una gran variedad de tumores de distinta naturaleza.

### ***CÁNCER DE PÁNCREAS.***

El páncreas es una gran glándula de forma alargada situada tras el estómago. Produce y segrega directamente en el torrente sanguíneo la hormona insulina, junto con otras sustancias que juegan un papel crucial para la digestión de las proteínas y otras operaciones corporales. El cáncer de páncreas es bastante frecuente. Se diagnostican cada año 27.000 nuevos casos en Estados Unidos, donde afecta con mayor frecuencia a individuos de raza negra que a caucasianos. Poco se sabe acerca del cáncer de páncreas o de como prevenirlo. El riesgo se incrementa a partir de la edad de 50 años, ocurriendo la mayoría de los casos entre los 65 y los 79. Fumar es un factor de riesgo, la incidencia entre fumadores es el doble que entre no fumadores. Algunos estudios han asociado este cáncer con otros problemas de salud, como pancreatitis crónica, cirrosis o diabetes. En países cuya dieta es rica en grasas, la incidencia del cáncer de páncreas es mayor.

El cáncer de páncreas generalmente cursa sin síntomas externos hasta que se encuentra en estado avanzado, y se considera una enfermedad «silenciosa». Por el momento, sólo la biopsia interna (que requiere una operación quirúrgica) permite un diagnóstico preciso. Debido a la falta de síntomas externos, la necesidad de realizar una biopsia no resulta obvia hasta que la

enfermedad está avanzada y es, probablemente, demasiado tarde. Los investigadores están intentando encontrar nuevas formas de diagnóstico, antes de que se puedan observar los síntomas. Las imágenes mediante ultrasonidos (ecografía) y la tomografía computarizada se están investigando.

La cirugía, la radioterapia y la quimioterapia podrían ser tratamientos contra esta cruel enfermedad, pero usualmente no producen resultados positivos, debido al diagnóstico tardío. Para el conjunto de todas las formas de cáncer de páncreas, la probabilidad de sobrevivir un año tras el diagnóstico es del 20% y de hacerlo tras cinco años, tan sólo del 4%.

Los científicos han descubierto un gen, llamado DPC4, localizado en el cromosoma 18, cuyo producto protéico ha recibido el nombre de «supresor del carcinoma pancreático». Las mutaciones en este gen provocan que el cáncer de páncreas se manifieste de forma especialmente agresiva e invada otros tejidos circundantes.

### ***DISTROFIA MIOTÓNICA***

La distrofia miotónica es una enfermedad hereditaria de los músculos. La «miotonía» es el término médico que designa el hecho de que los músculos se contraigan fácilmente, pero difícilmente vuelvan a su posición relajada. Debido a este hecho, los músculos se debilitan progresivamente y se inutilizan. Además, la distrofia miotónica también puede causar deficiencia mental y pérdida del cabello, así como cataratas. Los síntomas usualmente comienzan en la adultez temprana, pero pueden ocurrir a cualquier edad, existiendo un amplio espectro de severidad.

La distrofia miotónica está asociada con una secuencia repetida de nucleótidos en el cromosoma 19. Una de las características poco usuales de esta enfermedad es que su severidad se va amplificando con el paso de las generaciones, debido a que aparece cada vez un mayor número de repeticiones de la secuencia que produce la enfermedad.

### ***INMUNODEFICIENCIA SEVERA COMBINADA***

El término «inmunodeficiencia» designa cualquier trastorno del sistema inmunitario que provoque una respuesta inmune deficiente, lo cual hace que cualquier enfermedad oportunista pueda atacar al paciente, sin que su cuerpo pueda hacer nada por evitarlo. La palabra se ha hecho tristemente célebre por formar parte del nombre del SIDA. (Síndrome de Inmunodeficiencia Adquirida), enfermedad viral que, como todos sabemos, ataca principalmente a las células inmunitarias. El término «adquirida» denota que la enfermedad es infecciosa, no de origen genético. La Inmunodeficiencia Severa Combinada (IDSC), por el contrario, es una enfermedad monogénica bien caracterizada. Se manifiesta por la incapacidad del sistema inmunitario para sintetizar anticuerpos que permitan neutralizar cualquier infección. Los niños afectados de IDSC pueden enfermar fácilmente si entran en contacto con cualquier

patógeno, no pueden recibir transfusiones de sangre y las vacunas convencionales pueden provocarles infecciones fatales. La única terapia conocida hasta hace muy poco para esta enfermedad consistía en hacer vivir al niño en un ambiente estéril protegido por una cortina de plástico, por lo que los niños con IDSC son conocidos popularmente como «*niños burbuja*». Es obvio que no pueden abandonar el hospital y su calidad de vida es absolutamente pobre, convirtiéndose su existencia en una amarga espera, hasta que aparece el germen oportunista que les provoca la muerte.

Los científicos siguieron la pista de esta enfermedad hasta encontrar al culpable: se trata de un defecto en la enzima adenosina-desaminasa (ADA), una enzima clave para la síntesis de precursores de los anticuerpos. En la actualidad se usa rutinariamente la inyección de ADA producida sintéticamente por ingeniería genética en *E. coli*, lo que ha permitido alargar la esperanza de vida de los pacientes, que anteriormente solían morir antes de cumplir el primer año. Además, la deficiencia de ADA, como se suele conocer en la actualidad a la IDSC, fue la primera enfermedad en ser tratada con terapia génica, con resultados bastante prometedores, como veremos más adelante en este trabajo. El gen que codifica la ADA se localiza en el cromosoma 20.

### **ESCLEROSIS LATERAL AMIOTRÓFICA**

La esclerosis lateral amiotrófica (ALS), más conocida por los estadounidenses como «*enfermedad de Lou Gehrig*», por un famoso jugador de béisbol de los New York Yankees que la padeció, posiblemente será más conocida en el futuro, al menos en Europa, como la «*enfermedad de Hawking*», ya que es la enfermedad que padece el mundialmente conocido físico teórico inglés Stephen W. Hawking. La ALS es un trastorno neurológico degenerativo que provoca la progresiva inutilización de las neuronas motoras de la médula espinal y del cerebro. El término «*amiotrófica*» proviene del griego y significa «*sin movimiento muscular*»; «*lateral*» proviene de la parte de la médula espinal de donde parten los nervios motores; a medida que las células de la médula se van degenerando, se produce un endurecimiento o «*esclerosis*» del tejido de dicha zona.

La ALS es una de las enfermedades degenerativas más crueles. A medida que avanza, se va produciendo la pérdida de la motilidad de las extremidades y posteriormente, de los músculos implicados en el habla, la alimentación y la respiración. La mente, sin embargo, permanece intacta y en completo funcionamiento, a pesar de la inmovilidad del cuerpo. Los pacientes, en la última etapa de la enfermedad, han de ser alimentados por vía intravenosa y provistos de respiración asistida. En este estado, parecido a un «*coma lúcido*», pueden permanecer durante años.

Al comienzo de la enfermedad, los síntomas son tan suaves que suelen permanecer inadvertidos. Incluyen calambres y tirones musculares, especialmente en las manos y los pies, dificultad en el uso de las extremidades y en proyectar la voz y, en etapas más avanzadas,



dificultad para respirar y tragar. Más de 5.000 personas son diagnosticadas de ALS cada año en los Estados Unidos, la mayoría entre los 40 y 70 años de edad. Se calcula que más de 300.000 de los americanos vivos actualmente, morirán de ALS. Además, el coste para las familias afectadas puede llegar a ser enorme. En las últimas etapas de la enfermedad, puede llegar a más de 20 millones de pesetas por año.

Sin embargo, existe una esperanzadora investigación en marcha. En 1991, un equipo de científicos localizó un gen ligado a la ALS en el cromosoma 21. En 1993, estaba aislado y secuenciado. Fue llamado SOD1 y presentaba una mutación recesiva. No se descarta que la ALS sea poligénica y aparezcan más genes ligados a ella, quizás en otros cromosomas. La forma más común de ALS es «esporádica»; es decir, sin herencia genética directa.

### **NEUROFIBROMATOSIS**

La neurofibromatosis, llamada también enfermedad de von Recklinghausen, es conocida por el público como la presunta causa del desfiguramiento grotesco de John Merrick, más conocido como «el hombre elefante», a finales del siglo XIX. Su tragedia fue magistralmente llevada al teatro y al cine, aunque es más que probable que el diagnóstico de la época fuera equivocado y que Merrick sufriera, en realidad, otro trastorno de síntomas similares, pero mucho menos frecuente, conocido como síndrome de Proteo.

La neurofibromatosis se caracteriza por manchas marrón claro y bultos en la piel causados por tumores neuronales subyacentes que suelen causar pérdida de la audición y la visión, y otras dificultades neurológicas graves. No hay ningún tratamiento posible más que la cirugía, para extirpar los tumores que, inevitablemente, vuelven a aparecer. En Mayo de 1987, dos grupos de investigadores, en la Universidad de Utah y en el Hospital General de Massachusetts, localizaron el gen, un alelo dominante, sobre el cromosoma 17. Sin embargo, tan pronto como lo localizaron, se dieron cuenta de que la enfermedad no era común entre los desórdenes genéticos. Uno de los pacientes presentaba un intercambio entre el cromosoma 17 y el 22. El gen localizado en el cromosoma 17 era diferente de cualquier otro gen humano descubierto hasta entonces. Tenía entre 500 y 2.000 kb de longitud y había por lo menos otros tres genes anidados dentro de él, orientados en dirección opuesta. Descubrimientos como éste muestran que el genoma puede ser mucho más complejo de lo que se pensaba con anterioridad. Hoy en día se consideran dos tipos de neurofibromatosis, la de tipo 1 se localiza en el cromosoma 17 y la de tipo 2, en el 22, siendo ésta más benigna que la primera. Ambos genes parecen codificar proteínas «supresoras de tumores». Esto quiere decir que el gen sano actúa inhibiendo la formación de tumores en las células del paciente. Cuando se produce una mutación, el gen se hace inservible y los tumores pueden desarrollarse a sus anchas. Se piensa que el gen de la neurofibromatosis, así como otros supresores de tumores, podrían actuar frente a otros tumores también, por lo que las investigaciones para su uso terapéutico frente al cáncer están en marcha.

### ***DISTROFIA MUSCULAR DE DUCHENNE***

La distrofia muscular de Duchenne (DMD), llamada así por el neurólogo francés Guillaume Benjamin Amand Duchenne (1806-1875), que la describió, es la distrofia más frecuente entre los varones (1 de cada 3500 niños varones). Se caracteriza por el crecimiento anormal (pseudohipertrofia) de los músculos. El término «distrofia» se usa para designar cualquier enfermedad que produzca debilitamiento, degeneración o desarrollo anormal de los músculos. La DMD se caracteriza por un rápido avance de la degeneración de los músculos, que hace que el paciente muera hacia los veinte años, por paro cardíaco o pulmonar. Está ligada al sexo, por lo que su frecuencia de aparición es mucho mayor en varones y su gen se localiza en el cromosoma X. En 1986, un equipo de genetistas de la Escuela Médica de Harvard, dirigido por Louis Kunkel, consiguió localizar el gen que produce la DMD. Es un gen enorme, que se extiende a lo largo de dos millones de bases y contiene al menos 60 exones. Este gen codifica una proteína hasta entonces desconocida y cuya función está siendo extensivamente estudiada, que ha recibido el nombre de «distrofina». Es inusualmente grande, teniendo diez veces el tamaño de una proteína media. La colaboración internacional ha sido fundamental para su localización. Un sólo artículo científico sobre el hallazgo reunió los nombres de 77 autores de 24 instituciones de investigación de 8 países. Ahora, las investigaciones se centran en averiguar el papel que desempeña la distrofina en el músculo y en desarrollar procedimientos para inyectar genes que codifiquen la proteína normal, directamente en los músculos. Algunos experimentos han demostrado que las inyecciones de genes en los músculos de animales estimularon la producción de la proteína por parte de las células musculares durante varios meses.

### ***FACTOR DETERMINANTE DE LOS TESTÍCULOS.***

El cromosoma Y está prácticamente vacío, en comparación con su compañero de fatigas, el X. Esto tiene su explicación, ya que, en realidad, el hombre y la mujer han de poseer prácticamente las mismas proteínas. Si una proteína estuviera codificada por el cromosoma Y, únicamente aparecería en el hombre y no en la mujer, que no posee este cromosoma; por el contrario, las proteínas codificadas por el cromosoma X aparecen en ambos sexos, ya que el hombre es XY. Uno de los pocos genes que han podido localizarse en el cromosoma Y es el llamado factor determinante de los testículos, o TDF. Se trata de un «factor de transcripción». Se denominan así las proteínas que son capaces de unirse al ADN, modulando (favoreciendo o empeorando) la transcripción del mismo a ARN mensajero y por tanto, controlando la síntesis de las proteínas correspondientes. El TDF es capaz de unirse a determinados genes (no necesariamente en el cromosoma Y), cuyos productos protéicos son necesarios para la formación de los testículos durante el crecimiento fetal, determinando así directamente el sexo del futuro individuo.

**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

# **LA GENÓMICA FUNCIONAL: LA SINFONÍA DE LOS GENOMAS**

*“Nuestra imaginación despliega sin cesar ante nosotros la imagen renovada de lo que está en trance de ser posible, de lo que es posible. Al amparo de de esta imagen confrontamos permanentemente nuestros temores con nuestras esperanzas”*

**FRANÇOIS JACOB. “LA SOURIS, LA MOUCHE ET L’HOMME”  
(El ratón, la mosca y el hombre). 1997.**

Es importante que hagamos énfasis en que el genoma de cualquier organismo es el contenido total de su información biológica empacada en un juego completo de cromosomas. En otras palabras, el genoma es el material portador de la herencia, o sea, es el que determina la transferencia de los caracteres hereditarios del progenitor a su progenie.

La disponibilidad de las secuencias completas de los genomas tanto de humanos como de otros organismos supone un cambio estratégico en el abordaje de los problemas biológicos. Con la casi totalidad de los genes en la mano, la aproximación reduccionista convencional de escudriñar gen por gen puede complementarse por una aproximación más global o integradora para considerar todos los genes a la vez. -Sobre la base de que son las intrincadas redes de interacciones entre proteínas (y moléculas de RNA) las que regulan los procesos biológicos, más que proteínas aisladas actuando en solitario-; tales aproximaciones, ayudan a comprender los mecanismos moleculares involucrados en las enfermedades humanas. Los proyectos de secuenciación de genomas se asocian, cada vez con mayor frecuencia, con el reto de comprender la función de los genes identificados en proyectos de secuenciación a gran escala. La “era postgenómica” ha sido declarada oficialmente abierta y ello incluye la descripción tanto de cada una de las proteínas expresadas por sus genes respectivos (proteómica), como del conjunto interactivo de las proteínas involucradas en un proceso o módulo biológico

en particular (proteómica funcional). La complejidad inherente del tratamiento simultáneo de decenas o de cientos de proteínas para formular problemas biológicos integrados exige disponer de mapas informativos de complejidad funcional creciente.

Los organismos pueden concebirse como sistemas de módulos moleculares, cada uno de ellos, responsable de una determinada función biológica. La función de los módulos puede ser básica como la síntesis de una proteína, la regulación del ciclo celular o la maquinaria de reparación del DNA; o sofisticada como las memorias a largo plazo e inmunológica. Uno de los principales retos de la era postgenómica es que la mayoría de los genes identificados en las secuencias genómicas completas no ha podido ser adjudicada a módulo funcional alguno. Ello conlleva a la pregunta de si cada módulo estudiado incluye más genes de los supuestamente involucrados o si existen más módulos que deban ser descubiertos.

### ***La Bioinformática, una nueva aplicación del genoma.***

En cualquier caso, la estrategia de partida asume que los genes copartícipes de un determinado módulo funcional deben compartir propiedades comunes (interactoma). Así, los mapas de transcriptomas consistirán en un conjunto de “grupos de expresión” o de genes coregulados, mientras que el interactoma contempla redes de proteínas interactivas (grupos de interacción). Por su parte y especialmente en relación con el genoma, una herramienta revalorizada es la técnica del noqueo génico –silenciamiento de un gen determinado– en organismos transgénicos o mediante la utilización de RNAs de interferencia (RNA complementario de una secuencia de DNA a la que se une y anula).

Una de las principales limitaciones de los mapas funcionales es que sólo ofrecen una aproximación grosera a los genes pertenecientes a un módulo; hoy sólo son capaces de generar conclusiones sobre la función de una determinada proteína. No cabe duda de que las restricciones inherentes a los mapas funcionales se irán disipando conforme más tipos de mapas vayan integrando conjuntos solapantes de características funcionales. De este modo, uno de los retos más evidentes de la era postgenómica es aprender cómo puede integrarse la información funcional aportada por los distintos tipos de mapas; una empresa que sólo se resuelve mediante la colaboración computacional. La bioinformática se asienta, cada vez con mayor solidez, como la herramienta clave de la empresa biológica del futuro. Otro reto es aprender cómo la información funcional de los diferentes mapas puede integrarse en un atlas biológico. Generaciones de bioinformáticos se ocuparán de desarrollar la infraestructura de los mapas funcionales y la arquitectura de los atlas biológicos. Un tema importante es la organización de la información vertical y horizontalmente. En el contexto vertical, cada característica conocida de genes y de proteínas específicas, podrá encontrarse en un documento individual; aunque ésta aproximación es útil en proyectos “un-gen-cada-vez” tiene poca utilidad en programas integradores más ambiciosos, para los que el formato horizontal

será más apropiado. En otras palabras, la organización de mapas en grupos y el subsiguiente solapamiento entre grupos de diferentes mapas exigirá el tratamiento de bases de datos horizontalmente; ello abordando, a la vez, todos los genes Y.

### ***Los chips de ADN.***

Un chip de DNA es un microarreglo de múltiples secuencias de DNA que potencialmente sirven de sondas para detectar, mediante hibridización de ácidos nucleicos, los distintos ADNc que se han sintetizado *in vitro* a partir de poblaciones de ARNm extraídas de una célula. Como su nombre lo infiere este arreglo no ocupa más de unos cuantos centímetros y permite colocar de manera individual aproximadamente 30.000 tipos diferentes de sondas incluyendo ADNc y EST (secuencias expresadas con marcas), que pueden identificar un número equivalente de ARN mensajeros o secuencias de nucleótidos expresadas en una célula. Lo que se obtiene con ésta tecnología es un perfil individual de la expresión de los genes, la cual es una radiografía del funcionamiento del genoma.

Actualmente existen dos formatos de microarreglos de ADN que se diferencian entre sí en el tipo de ácido nucleico empleado como blanco de hibridización.

Formato I (Microarreglo): Contiene sondas de ADNc (500 a 5000 bases en extensión) que se inmovilizan sobre una superficie sólida (vidrio) utilizando un robot aplicador bien separadamente o en una mezcla se pueden incluir entre 300 a 6000 blancos de hibridización.

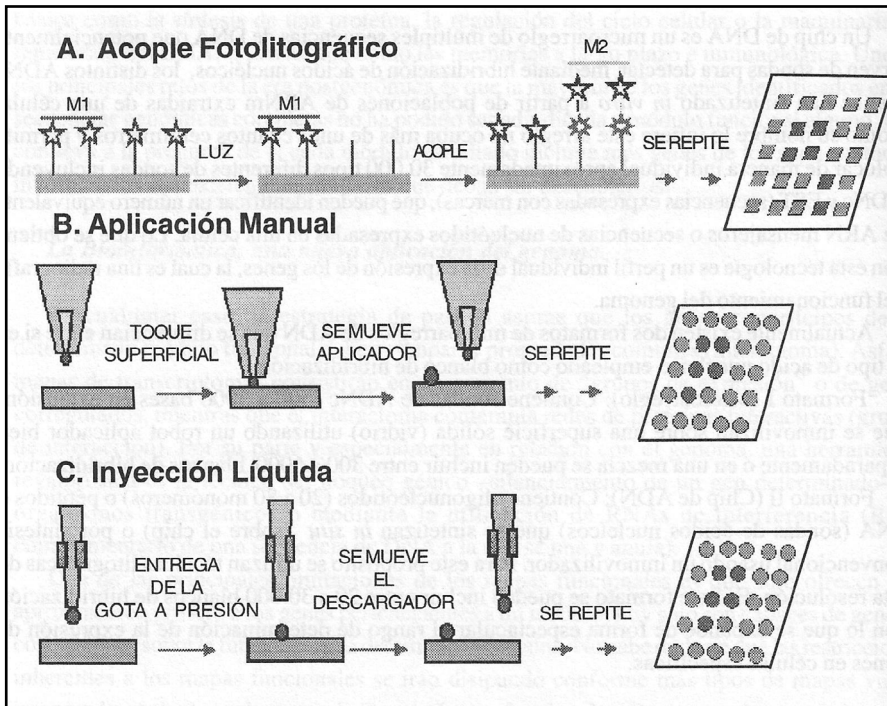
Formato II (Chip de ADN): Contiene oligonucleótidos (20 a 80 monómeros) o péptidos o PNA (sondas de ácidos nucleicos) que se sintetizan *in situ* (sobre el chip) o por síntesis convencional usando un inmovilizador. Para este propósito se utilizan técnicas litográficas de alta resolución. En este formato se pueden incluir entre 20 a 30.000 blancos de hibridización con lo que se expande de forma espectacular el rango de determinación de la expresión de genes en células específicas.

### ***Microarreglos de ADN de alta densidad.***

La clave que caracteriza a todos los microarreglos de ADN es que moléculas de ácidos nucleicos que están marcadas, hibridizan en solución con alta especificidad a secuencias inmovilizadas en un sustrato sólido. Así se facilita la medición cuantitativa en paralelo de muchas secuencias de ADN presentes en una mezcla compleja.

Aunque se han desarrollado muchos métodos de microarreglos de ADN, actualmente existen dos que han sido perfeccionados y son los de uso común en la actualidad. En uno de ellos, los microarreglos de ADN se construyen mediante unión física de fragmentos de ADN procedentes de clones procedentes de bibliotecas génicas o a partir de productos de la reacción en cadena de la polimerasa. Utilizando un robot capaz de arreglar e imprimir capilarmente secuencias de ADN se pueden fijar hasta 23.000 elementos en una placa de

microscópio. Por el otro método, los arreglos se construyen sintetizando *in situ* oligonucleótidos de cadena única empleando técnicas fotolitográficas (figura 9.1). El revelado de las secuencias hibridizadas se efectúa generalmente mediante la emisión fluorescente de los híbridos y permite su cuantificación en un citodensitómetro



**Figura 9. 1.** Diferentes métodos para la construcción de microarreglos de ADN con los que es posible determinar la expresión a gran escala de los genes que se transcriben en una célula. Esta tecnología ha permitido obtener una nueva visión del genoma.

Las ventajas del primer método son su bajo costo y su flexibilidad; además, de que no se requiere la secuencia primaria para imprimir el elementos de ADN. En el segundo método se pueden barrer hasta 28.000 características en una placa de 1.28 x 1.28 cm aumentado su poder de resolución. Aunque todavía existen algunos inconvenientes tecnológicos, los microarreglos de ADN ya son una herramienta poderosa aplicada en la genómica funcional.

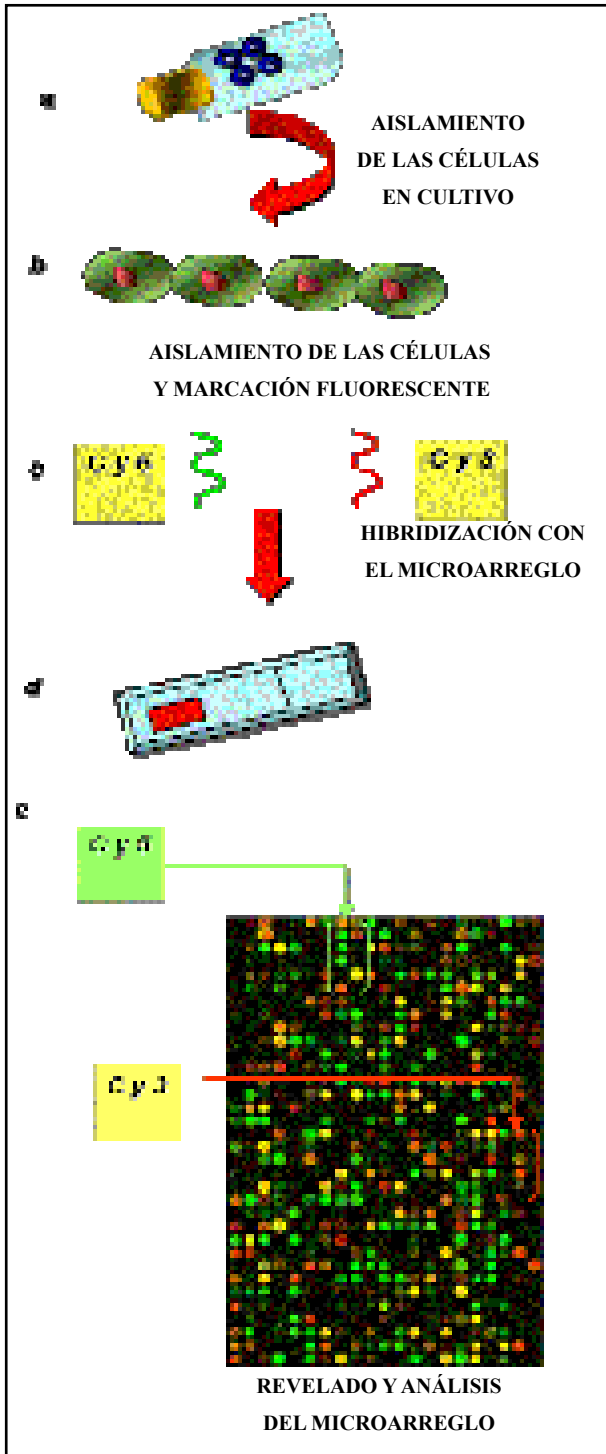
## **SEÑOR GENOMA TOQUE SU SINFONÍA**

Toda nuestra información genética está codificada en el genoma; así, de manera funcional, un gene es transcrito a un ARNm el cual es traducido a la correspondiente proteína con la cual se realizará una determinada función biológica. Lo anterior fue una de las propuestas más fijas en la biología molecular sin embargo, con algunas modificaciones introducidas, esto es lo que esencialmente ocurre cuando un gen se va a expresar. Hemos visto que el gene no es solamente una secuencia de nucleótidos que codifica por una secuencia definida de aminoácidos de una proteína, sino que tiene una serie de elementos reguladores que expresan otros sutiles atributos del mismo. ¿En qué célula se expresará?, ¿Bajo qué condiciones de control celular se expresará?, ¿Por qué si se transcribe en el cerebro no lo hace en el músculo?, ¿Por qué en diferentes tejidos la transcripción y maduración es distinta?. Además la existencia variaciones en las secuencias de nucleótidos de un mismo gene entre diferentes individuos nos lleva a postular otra gran serie de preguntas: ¿Cómo la secuencia exacta del genoma humano difiere entre los distintos individuos?, ¿Qué son las diferencias que resultan en la predisposición genética a adquirir ciertas enfermedades?, ¿Cuál es el papel específico de una proteína expresada por el genoma de un patógeno en el contexto de la expresión de genes de una célula humana?. ¿Cuánto se parecen las proteínas expresadas por los diferentes genomas? y ¿Cómo todas estas proteínas conforman de manera integrada la vida?. Todas estas preguntas conforman el paradigma de una novedosa rama del genoma humano, la proteómica.

La proteómica es tal vez una de las derivaciones del proyecto genoma humano que más ha sido debatida, no sólo por su complejidad sino por sus consecuencias tanto en la medicina como en la economía mundial. Es por esta razón que merece que describamos de alguna forma como se estudia y que se espera que pueda llegar a resolver. En la figura 9.2 se presenta un protocolo típico para poder determinar la sinfonía de la expresión de los genes de un genoma, este es tal vez el primer paso dentro de la proteómica.

Lo que se determina en un microarreglo de ADN son todos aquellos genes que están activos en una célula y que potencialmente serán traducidos en proteínas, desafortunadamente hasta el momento solamente a unos 20.000 de los casi 40.000 genes humanos se les ha podido determinar sus perfiles de expresión, o sea que más del 50% de los genes de nuestro genoma todavía no se han revelado con lo que su potencialidad funcional tampoco ha sido descrita. En términos prácticos la sinfonía del genoma humano es todavía una sinfonía inconclusa.





**Figura 9.2.** Protocolo general para determinar la expresión de los genes de un genoma mediante la utilización de microarreglos de ADN. (a) Se aíslan las células de un tejido y se les somete a condiciones especiales que permitan su crecimiento, es posible entonces introducir efectos que cambien la expresión de genes con respecto al ambiente no alterado. (b) De estas células se extrae los ARNm transcritos los cuales son convertidos en ADNc mediante la acción de la transcriptasa reversa de un retrovirus. (c) Estos ADNc son marcados mediante la unión de sustancias químicas que fluorescen y permiten reconocerlos mediante la emisión de un determinado color. (d) Una vez que se han marcado con diferentes fluorógenos, la población de ADNc es introducida en el microarreglo para hibridarla con los distintos oligonucleótidos blanco previamente inmovilizados en el micorarreglo. Una vez que se ha efectuado la reacción de hibridización, son revelados e interpretados y cuantificados con instrumentos de detección adecuados.

## **LA GENÓMICA FUNCIONAL ES UNA REALIDAD**

Aunque la individualización genómica permite enfocar la potencialidad del genoma de un individuo frente a un ambiente predeterminado endógenamente, la pregunta que surge es ¿cómo estudiar la compleja trama de interacciones a nivel génico que permiten la expresión del genoma?. En la última década del siglo pasado, el desarrollo de la nanotecnología permitió la creación y perfeccionamiento del famoso chip de ADN (“DNA chip”) como un instrumento que permitió diseccionar de manera muy fina la expresión, en forma de mRNAs de los genes de una célula, dando como resultado un perfil genómico (figura 9.2).

Trabajos recientes han mostrado que la variación en las condiciones nutricionales conlleva a una expresión global diferencial de los genes de una célula (cambio en el perfil genómico). Se ha demostrado que en el hígado de individuos sometidos a restricción calórica (RC), una disminución importante del suministro de grasa y carbohidratos en la dieta, se observa una expresión de genes que es opuesta a aquellos que se expresan en el envejecimiento. El grupo de genómica nutricional del Departamento de Bioquímica de la Universidad de California liderado por el Dr. Stephen Spindler ha demostrado que durante el envejecimiento de la célula hepática de ratones se expresan genes asociados con la inflamación, el stress celular, una disminuida capacidad para la apoptosis, el metabolismo xenobiótico, el ciclo celular normal y la replicación del ADN. En contraste la restricción calórica de largo plazo LT-RC (4 semanas) contrarresta la mayoría de estos efectos de expresión de genes produciendo un aplazamiento del envejecimiento. Aunque todavía no se conocen muchas de las respuestas génicas a las diferentes dietas en los humanos, la utilización de microarreglos de ADN es una herramienta poderosa para entender más detalladamente la expresión global de genes celulares en diferentes condiciones nutricionales. Pero lo más importante es que la genómica nutricional permitirá cruzar la información genómica individual con la alimentación y los componentes de los alimentos, de modo que el efecto sea muy positivo para la salud del individuo. La idea es que los alimentos riesgosos, como las grasas saturadas, podrían reemplazarse con otros potencialmente menos agresivos, que provoquen la descomposición del colesterol. Por ejemplo, comer salvado de avena en lugar de rodajas de queso. No se necesitan millones de dólares para llegar a esta conclusión, pero la ciencia de la nutrición todavía ignora cómo reacciona el organismo en el nivel molecular. La genómica nutricional será la encargada de descubrirlo. El perfil genómico funcional individual puede ayudar a mejorar la nutrición y la salud.

La genómica y la bioinformática tienen un gran potencial para identificar genes que causen enfermedades, lo cual se realiza mediante investigación de las bases de datos del genoma humano. La comparación de un genotipo individual con bases de datos genómica permitirá la prescripción de drogas individualizadas genotipo dependientes. El mismo enfoque se puede utilizar para el estudio de los metabolitos humanos, con lo cual se pueden, en un futuro no muy lejano, encontrar soluciones personalizadas para mejorar la salud nutricional además de permitir una planificación racional de la producción de alimentos para las siguientes generaciones.

Algunos avances en la química analítica hacia finales del siglo pasado, han incrementado el conocimiento sobre la estructura de muchos metabolitos, esto ha servido para crear bases de datos que hacen posible la investigación del potencial de cada uno de ellos. Este progreso ha condicionado un salto gigantesco en la forma como se enfoca la nutrición moderna; sin embargo, todavía es muy temprano para emitir un concepto sobre su impacto en la salud pública.

La bioinformática se perfila como una de las herramientas más poderosas en la era post-genómica. Su aplicación en la genómica nutricional debe incluir tres áreas fundamentales: 1. Implementar análisis de perfiles genómicos multiparalelos que evalúen el efecto metabólico y nutricional de los metabolitos. 2. Evaluar el perfil genómico a lo largo del tiempo, con lo que se evalúa el efecto en el perfil genómico a metabolitos durante su consumo en dietas prolongadas. 3. Integrar los perfiles genómicos con la proteómica para entender más detalladamente el metabolismo humano.

### ***LOS ALIMENTOS Y EL FUTURO DE LA GENÓMICA NUTRICIONAL***

Si bien en un comienzo el cruce entre la dieta y el ADN se hará de manera amplia, la genómica permitirá afinar las prescripciones para una atención preventiva de la salud. Probablemente lleve muchos años, pero creo que la genómica nutricional permitirá organizar dietas que eviten o demoren la aparición de enfermedades graves y muy difundidas hoy, como el cáncer y algunas enfermedades degenerativas como el mal de Alzheimer. Se prevé además una dieta que tenga en cuenta a la vez la constitución genética y la actividad. Por ejemplo, la dieta de un atleta debe tener en cuenta también su constitución genética para lograr una eficacia máxima.

Ya se han dado algunos pasos para comercializar la investigación en genómica nutricional. Una empresa denominada Galileo Laboratories ya ha iniciado conversaciones con fabricantes de alimentos para llevar sus productos al mercado. Galileo se especializa en los problemas de óxido-reducción, perturbación del metabolismo energético celular que provoca enfermedades como la apoplejía, los ataques de corazón, las inflamaciones y la diabetes. La empresa dice que ha encontrado diversos compuestos que podrían prevenir o demorar la aparición de estas enfermedades.

La genómica nutricional desempeñará un papel importante para la seguridad porque permitirá comprobar en el laboratorio el efecto de nuevos componentes nutricionales. Pensemos por ejemplo en los ácidos grasos saturados en los alimentos: hace 10 ó 15 años las empresas productoras de alimentos decían que no eran un problema y no tenían efectos negativos sobre la salud y el bienestar. Ahora, sus riesgos son evidentes y la Asociación Cardiológica Norteamericana ha publicado un artículo en el cual se solicita que los productos alimenticios que los contienen lleven una etiqueta especial, porque pueden provocar el engrosamiento de las arterias.

En última instancia, la genómica nutricional es sólo una rama de la genómica, como la genómica de las proteínas (proteómica) y la genómica farmacológica (farmacogenómica), que han surgido después de establecido el mapa del genoma humano. Sin embargo, la interacción entre los alimentos y la expresión de los genes tendrá un efecto mucho más profundo en la sociedad puesto que todos tenemos que comer.

No obstante, se trata de un tema muy complejo que implica cuestiones de otro orden que habrá que resolver. El tema de los seguros es una de estas cuestiones. ¿Qué consecuencias tendrá para la póliza de seguro de vida el hecho de que el solicitante tenga una tendencia genética a las enfermedades cardiovasculares? ¿Qué consecuencias tendrá para el interesado el hecho de no observar una dieta que demore la aparición de la enfermedad?. Todos estos temas todavía imprecisos llegarán algún día a los tribunales. Lo único seguro es que, en el futuro, la dieta tendrá más vinculación con la constitución genética de cada uno y que la manzana diaria se verá reemplazada por algún aminoácido.

**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

## **LA TERAPIA GÉNICA**

*“Se diría que no hay aspecto de nuestras vidas que no esté dentro del territorio reclamado por el poder del DNA”*

**RICHARD LEWONTIN. “THE DREAM OF THE HUMAN GENOME”.  
(El sueño del Genoma Humano). 1992.**

Sin lugar a dudas una de las aplicaciones más directas de la era genómica ha sido la terapia génica. Ésta es la sustitución o reparación de genes defectuosos en células vivas humanas. A medida que se iba conociendo la constitución del genoma humano, se hizo posible la identificación de genes completos en los que algunas mutaciones producen enfermedades transmisibles genéticamente. Así las funciones que están alteradas por acción de un gene defectuoso son restauradas por la inserción, en la mayoría de los casos, de un vector quimérico que contenga una copia normal de su alelo mutante o gene terapéutico, (aquel que permite la restitución de la función afectada), la expresión del gene terapéutico se logra por diferentes metodologías.

La terapia génica ha demostrado ser utilizable tanto en modelos animales como en algunos experimentos controlados en humanos, en los que se ha entregado al genoma una copia no alterada de un determinado gene terapéutico. La introducción de una copia de un gene normal en un individuo afectado para ese gene se puede efectuar tanto en sus células somáticas como en sus células germinales. Estos dos enfoques determinan las dos posibilidades de terapia génica, la una es la denominada terapia génica somática mientras que la otra es la terapia génica a nivel germinal. Es importante destacar que actualmente sólo se han efectuado de manera controlada, experimentos de terapia génica somática, siendo un tema fundamental de discusión ética la terapia génica de la línea germinal.

## **SISTEMAS DE TRANSFERENCIA GENÉTICA**

La terapia génica involucra la utilización de una serie de modalidades para lograr la inserción del gene terapéutico dentro de las células blanco. En este sentido se han desarrollado diferentes estrategias de introducción de las cuales haremos una descripción general. En la tabla 10.1 se consignan las principales metodologías actualmente empleadas para la entrega de genes.

Dentro de lo expuesto en la tabla 10.1 es posible ver que los vectores virales representan un método biológico muy conveniente para la transferencia de genes. También se han ensayados otros sistemas de transferencia no biológica que incluyen varios métodos fisicoquímicos.

**Tabla 10.1.** Principales sistemas utilizados para la transferencia de genes.

<b>Sistema de transferencia</b>	<b>Tipos de instrumentos</b>
Viral	Adenovirus Retrovirus Virus adenoasociados (AAV) Herpes Virus
Químicos	Fosfato Cálculo Medios Catiónicos Polietilenglicol
Físicos	Electroporación Disparo de micropartículas de oro
Fusión	Liposomas cargados con DNA Policationes

## **ESTRATEGIAS DE TERAPIA GÉNICA SOMÁTICA**

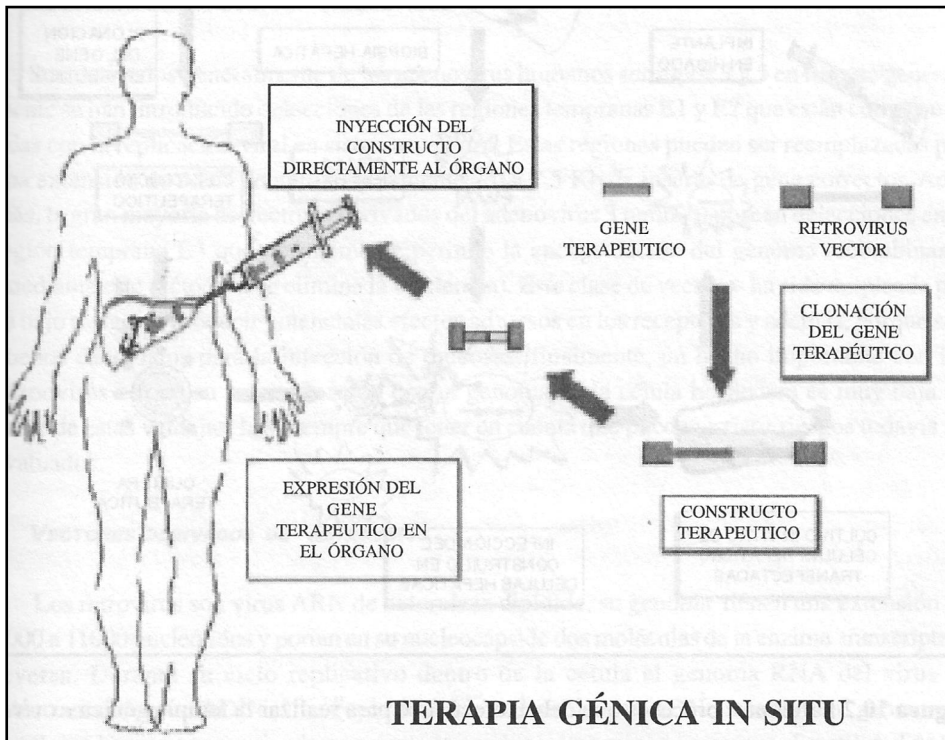
Tal vez el tipo de terapia génica que más futuro tiene en el tratamiento de enfermedades genéticas e infecciosas, tanto por sus implicaciones tecnológicas como éticas, es la terapia génica al nivel de las células somáticas. Con respecto a la terapia génica somática, es posible definir dos estrategias básicas viables: la genoterapia *ex vivo* y la *in situ* (figura 10.1).

La terapia *ex vivo* consiste en la introducción del gene corrector dentro de las células, en general linfocitos, de la persona enferma. Éstos son manipulados y puestos a proliferar en medios *in vitro* para constatar la integración estable y expresión del gen corrector. Posteriormente

los linfocitos corregidos son transfundidos de nuevo al receptor para lograr su infiltración en el órgano blanco y lograr la circulación de la proteína normal (figura 10.2). En contraste, la terapia *in situ* tiene por objeto introducir el gene corrector contenido dentro de un elemento construido (retrovirus quimérico) en un tejido específico para lograr la integración del gene corrector y la síntesis tejido específica de la proteína normal. Este último enfoque genera una serie de problemas tecnológicos y éticos que se discutirán mas adelante.

### **SISTEMAS DE TRANSFERENCIA MEDIADOS POR VIRUS**

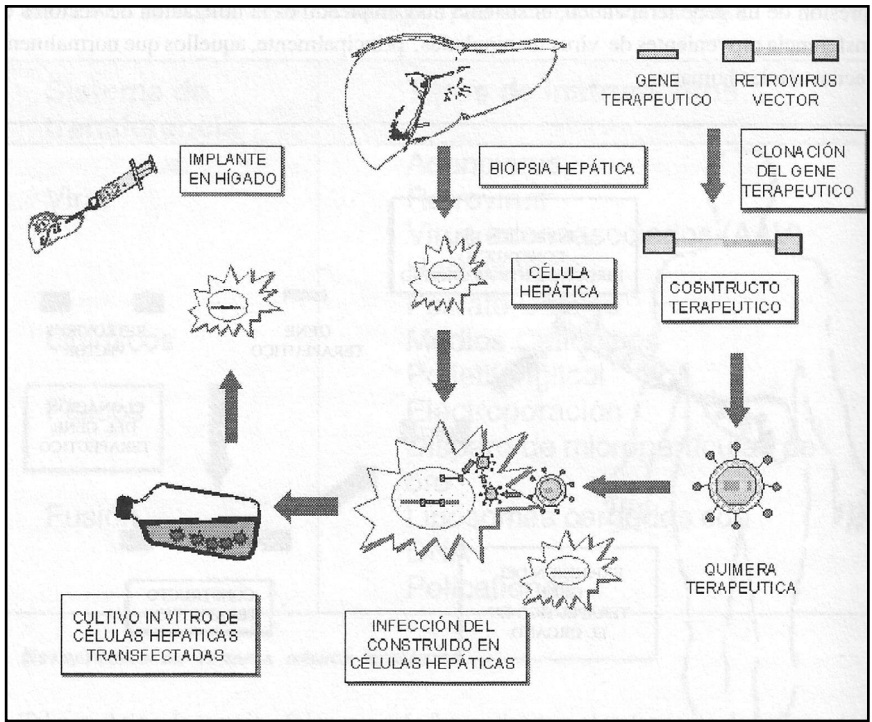
Tal como se ha descrito anteriormente, la terapia génica es una nueva estrategia capaz de alterar la función celular de un organismo con fines benéficos. Para lograr la inserción y expresión de un gene terapéutico, el sistema más empleado es la utilización de vectores de transferencia provenientes de virus manipulados; principalmente, aquellos que normalmente infectan células humanas.



**Figura 10.1** Terapia génica *in situ*



Existen cuatro grupos de vectores virales diseñados para la entrega efectiva de genes terapéuticos que actualmente se han empleado en protocolos de terapia génica: los derivados de los adenovirus 3 y 5 (Adv 3 o 5), los del virus herpes tipo 1 (HHSV-1), los derivados de retrovirus murinos y humanos (derivados del virus de la inmunodeficiencia humana tipo 1 VIH-1) y los virus adenoasociados (AAV). Recientemente se han construido vectores que tienen como base el virus del simio 40 (SV40) y el virus vaccinia, sin embargo todavía no han sido utilizados en protocolos en humanos. La tabla 10.2 muestra algunas de las características más importantes de los virus empleados como vectores para terapia génica en protocolos humanos.



**Figura 10.2** Diferentes procesos que deben efectuarse para realizar la terapia génica *ex vivo*.

**Tabla 10.2.** Virus como candidatos para la entrega de genes terapéuticos en células eucarióticas.

Virus	Ácido nucleico	Integración en genoma huésped	Capacidad de transporte de genes (Kb)	Efecto citopático de los vectores
Adenovirus	ADN	No se integra	4 a 8	Bajo
Herpesvirus	ADN	No se integra	20	Alto
AAV	ADN	Se integra	< 4	Bajo
Retrovirus	ARN	Se integra	4 a 8	Bajo
Virus de Vaccinia	ADN	No se integra	25 a 75	Alto

### **VECTORES DERIVADOS DE ADENOVIRUS**

Son derivados generalmente de los adenovirus humanos serotipos 3 y 5 en los que generalmente se han introducido deleciones de las regiones tempranas E1 y E2 que están comprometidas con la replicación viral en sistemas *in vitro*. Estas regiones pueden ser reemplazadas por una extensión de ADN de aproximadamente 7.0 a 7.5 Kb de inserto de gene corrector. Además, la gran mayoría de vectores derivados del adenovirus 3 también poseen deleciones en la región temprana E3 que normalmente permite la encapsidación del genoma recombinante (mediante este método se le elimina la virulencia). Este clase de vectores ha sido empleada por su bajo riesgo en producir potenciales efectos adversos en los receptores y además, porque son buenos candidatos para la infección de mucosas; finalmente, un hecho importante con los adenovirus es que su recombinación con el genoma de la célula hospedera es muy baja. A pesar de estas ventajas, hay siempre que tener en cuenta que pueden existir riesgos todavía no evaluados.

### **VECTORES DERIVADOS DE RETROVIRUS**

Los retrovirus son virus ARN de naturaleza diploide, su genoma tienen una extensión de 9000 a 11000 nucleótidos y portan en su nucleocápside dos moléculas de la enzima transcriptasa reversa. Durante su ciclo replicativo dentro de la célula el genoma RNA del virus es retrotranscrito a un ADNc. Una copia de doble cadena de este ADNc se integra al genoma de la célula blanco permaneciendo en un estado genético denominado provirus. La utilidad de los vectores retrovirales en terapia génica es como vector de integración al genoma de la célula blanco.

Una amplia gama de vectores de inserción empleados en terapia génica se derivan de retrovirus, murinos. En este sentido los vectores retrovirales presentan una serie de ventajas porque tienen capacidad de integrarse de manera estable dentro del genoma de la célula blanco. Además su LTR es un promotor fuerte que garantiza de manera independiente la expresión modulada del gene corrector clonado. En general la construcción de vectores retrovirales involucra la retirada de las regiones de su genoma que codifican por los genes virales gag, pol y env dejando solamente las porciones terminales LTR y la región que es muy importante para el empaquetamiento del genoma retroviral. Debido a la incapacidad de producir proteínas de las cápsides virales, estos vectores deben ser infectados en un sistema especial que incluye una célula empaquetadora la cual si produce todas las proteínas estructurales del virus.

Esta amplia gama de vectores de terapia génica presenta características de una elevada infección en células blanco. En condiciones controladas permite una integración regional de la secuencias del vector. Recientemente vectores derivados del virus de la inmunodeficiencia humana tipo 1 (VIH-1), han posibilitado la manipulación de linfocitos humanos para efectuar la terapia génica somática ex vivo.

### ***VECTORES DERIVADOS DEL VIRUS HERPES SIMPLEX HUMANO TIPO 1 (HSV-1)***

Una de las estrategias más promisorias para el desarrollo de terapia génica de enfermedades del sistema nervioso central son los vectores derivados del virus Herpes tipo 1. Este tipo de virus presenta, en su ciclo de infección natural, un neurotropismo por ciertas terminaciones nerviosas trigeminales. Como resultado de este proceso se produce una migración por las vías aferentes creando una condición de latencia y residencia del ADN viral en ciertos puntos del sistema nervioso central (SNC). Esta característica ha sido aprovechada de manera experimental para enviar genes correctores hacia células blanco en el SNC. El ADN de los herpes es muy grande llegando a tener hasta 150 Kb de tamaño. Su complejidad replicativa ha sido aprovechada para lograr “construidos” muy promisorios en la entrega de genes correctores en áreas centrales y periféricas del sistema nervioso. Un vector derivado de herpes puede contener un cassette de información correctora el cual se incluye dentro de la porción UL del genoma del herpes, la introducción de estos construidos permitió a De Luca a expresar proteínas de potencial utilización terapéutica.

### ***VECTORES DERIVADOS DE VIRUS ADENOSOCIADOS (AAV)***

Este grupo de vectores incluye una serie de construidos que contienen elementos de los virus adenoasociados. Estos virus se presentan en diferentes tipos de mamíferos y se incluyen dentro del género Dependovirus de la familia Parvoviridae. Son virus muy pequeños cuyo genoma es un ADN de cadena única de 9 Kb; éste se dobla para formar una estructura

terminal de doble cadena. Los virus adenoasociados tienen secuencias terminales cortas denominadas ITR que permiten establecer recombinación con otros ADN. Se ha demostrado que la integración de los adenoasociados humanos ocurre específicamente en el cromosoma 19 por lo que se constituye en una alternativa potencialmente más segura para la integración de una nueva información genética en las células humanas. Los vectores adenoasociados tienen como ventaja que se expresan en células que no se replican mitóticamente por lo que se les está utilizando para la introducción de genes en tejidos no proliferativos.

En la actualidad existen muchos protocolos de terapia génica los cuales incluyen diferentes metodologías de entrega de genes correctores. Dentro de las más utilizadas están los retrovirus, la lipofección, los adenovirus entre otros.

### ***ENTREGA NO VIRAL DE GENES***

Son varios los métodos que hemos descrito que se utilizan para la entrega de genes sin utilizar vectores como intermediarios. Además de la electroporación y el uso de liposomas, se ha intentado recientemente crear una “molécula vectora de genes” híbrida que pueda fijar ADN y lo suministre específicamente a una célula. Uno de los complejos mejor estudiados es el de la asialoglicoproteína-polilisina. La polilisina posee una carga neta positiva la cual actúa como región de unión del ADN de plásmidos a pH celular y también unirse a la asialoglicoproteína (ASGP). Las células hepáticas poseen un receptor para proteínas que han perdido ácido siálico entre las cuales está la ASGP. Este hecho permite la captación del complejo DNA-Polilisina-ASGP específicamente por el hepatocito. La inyección intramuscular de este complejo, puede permitir que el músculo se convierta en una fuente de expresión de la proteína correctora que pueda ser depositada en el sistema circulatorio y alcance los blancos celulares alterados. De esta manera se tendría una estrategia correctora para las deficiencias en factores de coagulación (hemofilias A y B).

### ***CONTROL DE LA EXPRESIÓN DE LOS GENES ENTREGADOS***

Uno de los puntos centrales de la terapia génica es lograr la expresión del gene corrector en el tejido alterado. Este objetivo debe incluir como requisito la existencia de promotores tejido específicos regulables que permitan la expresión del gene en un tejido en particular. Tal vez donde tiene gran importancia este tipo de promotores, es en la entrega de genes empleando cromosomas artificiales. Recientemente se ha reportado la construcción de un cromosoma artificial conteniendo elementos de expresión y estabilización mitótica de humanos. Todo cromosoma artificial es un elemento genético complejo compuesto, por lo mínimo, de los siguientes elementos. Un elemento de inicio de replicación (ori), un elemento de estabilización mitótica (regiones centroméricas) que permitan su estabilización como elementos genéticos durante la división celular y secuencias teloméricas. Un elemento importante de este

complejo es un promotor tejido-específico que asegure la expresión del gene corrector en un tejido específico. Normalmente estos cromosomas artificiales son un buen prospecto para la entrega de genes cuyo tamaño es grande (megagenes que tienen mas de 200 Kb de tamaño).

Los genes celulares se pueden agrupar en dos grupos principales; aquellos que se expresan en todas las células y corresponden a los denominados genes domésticos (“Housekeeping genes”) y los que se expresan en un tejido o un conjunto de ellos denominados genes histo-específicos. Algunos ejemplos de estos últimos son las globinas que se expresan sólo en el tejido eritoblástico. Otro ejemplo es la  $\alpha$ -fetoproteína, cuya expresión es abundante en hepatocitos. Otro factor importante de expresión de un gene en una célula es el estado de la cromatina. La porción relajada de la cromatina es la transcripcionalmente activa, por lo que se debe entregar el gene corrector en esta zona de la cromatina. La colocación de un gene en una región cromosómica transcripcionalmente activa es un arma de doble filo, puesto que es más probable que cause daño si se inserta en un sitio funcional importante. Algunos virus, entre los cuales se refieren los retrovirus, parecen tener una tendencia a la integración en porciones de la cromatina transcripcionalmente activas, sin embargo el mecanismo de tal preferencia todavía no ha sido explicado.

En general las células eucarióticas no toleran ADN que no sea parte de sus cromosomas. El destino de este tipo de ADN es la eliminación durante la división celular. Una vez más los virus han proporcionado la solución debida puesto que algunos de ellos tienen un estado episomal no integrado en la célula hospedera. EL virus de Epstein-Barr (EBV), un miembro de la familia Herpesviridae, consigue mantener fuera del cromosoma su propio ADN, en esta forma es capaz de persistir; también logra replicarse durante la división celular y ser transmitido a las células hijas (transmisión vertical). En el EBV se ha identificado un segmento corto de ADN que actúa como origen de replicación. Este segmento puede ser incorporado en un construido para terapia génica con el objetivo de mantener el ADN en un estado episomal estable.

Existen algunas células en las que la captación y mantenimiento de ADN no integrado son eficientes. Un ejemplo de ellas son las células musculares que aceptan fácilmente el ADN introducido por inyección, mantienen el estado episomal durante varios meses y permiten la expresión continua de la proteína correctora.

### ***ENFERMEDADES POTENCIALMENTE OBJETO DE TERAPIA GÉNICA SOMÁTICA***

Hasta los científicos más optimistas son absolutamente conscientes de que la terapia génica no será la panacea médica para las enfermedades genéticas que aquejan a nuestra especie. En la actualidad sólo una pequeña proporción de estas enfermedades tiene su defecto génico identificado. De éstas, unas 3000 tienen su etiología en el defecto de un solo gene por lo que se les denomina monogénicas y son transmisibles de manera mendeliana. Las enfermedades monogénicas son las más fuertes candidatas a poder ser manejadas mediante la genoterapia.

Otros enfoques hacen posible, aunque de manera muy remota, la curación de ciertos tipos de cáncer por terapia génica introduciendo genes “suicidas” que matarían a la célula maligna. Recientemente se han vislumbrado protocolos de terapia génica para la cura de la infección por el virus de la inmunodeficiencia humana (VIH-1) lo cual parece ser prometedor. El avance de la terapia génica en los años recientes ha sido enorme, en 1996 ya se habían efectuado alrededor de 100 ensayos in vivo en todo el mundo involucrando casi 450 pacientes afectados con algunas de las enfermedades monogénicas, con buenos resultados.

El empleo de la terapia génica como posibilidad de cura de enfermedades genéticas comenzó a ser enfocada en las dos últimas décadas del siglo XX. El primer intento de corrección terapéutica de enfermedades genéticas se hizo en un paciente joven que sufría del síndrome de inmunodeficiencia combinada severa (SCID), la causa etiológica de esta enfermedad es la deficiencia en la producción de la enzima Adenosin Desaminasa (ADA) y generalmente afecta el sistema inmunológico con serias consecuencias de infecciones recurrentes que le pueden causar la muerte en edades muy tempranas. En este caso el gene clonado correspondiente fue introducido en linfocitos que habían sido previamente obtenidos del paciente (terapia ex vivo). Después de su cultivo in vitro, aquellos linfocitos que expresaron la enzima, fueron transfundidos al paciente a intervalos regulares. Se observó que la producción de ADA conseguía aliviar la sintomatología del SCID; en la actualidad se tienen informes muy alentadores del aumento de la expectativa de vida de estos niños que cuentan en la actualidad con nueve años. Otro de los intentos de terapia génica en humanos se realizó inicialmente en una paciente afectada por la fibrosis quística, una enfermedad de curso muy complicado que puede causar la muerte de niños en edades muy tempranas. Esta enfermedad monogénica, que tiene una frecuencia de 1 en 2000 niños nacidos vivos caucásicos, es causada por mutaciones en el gene de la proteína reguladora de la conductancia de transmembrana de fibrosis quística (CFTR) que se localiza en células epiteliales y regula el transporte de agua y iones Cloruro (Cl<sup>-</sup>) a través de las membranas plasmáticas de células epiteliales de tejidos pancreáticos, renales, de intestino, de glándulas salivares y sudoríparas. Un defecto en este transporte ocasiona que la secreción de mucopolisacáridos sea mas espesa conllevando entre otros a problemas respiratorios afectando la fisiología del pulmón. También se han determinados cambios deletéreos en órganos como el páncreas, el hígado y el riñón. Mediante diferentes tipos de análisis genéticos se han identificado unas 50 mutaciones en el gene de FQ dentro de las cuales la más frecuente en la población (70%) es una delección de una tripleta de nucleótidos que ocasiona la ausencia de una fenilalanina en la posición 508 de la proteína (DF508). A inicios de la década de los 90, se inició la fase I para corregir, mediante terapia génica somática ex vivo, la corrección en un niño con FQ. Mediante una continua introducción de linfocitos T infiltrantes (TIL) se ha podido lograr la entrega y expresión en tejidos afectados. Este es tal vez uno de los avances en la medicina molecular empleando la terapia génica con el fin de corregir el defecto e introducir el alelo terapéutico (figura 10.2).

La aceptación de la terapia génica como una alternativa de cura para algunas enfermedades humanas, ha producido una explosión de intentos por corregir defectos genéticos somáticos. Al igual que toda investigación clínica, en la terapia génica se deben seguir cuatro fases sucesivas hasta su uso como alternativa terapéutica.

### ***¿POR QUÉ ES FACTIBLE ACTUALMENTE LA TERAPIA GÉNICA?***

Desde varios puntos de vista la terapia génica plantea dilemas éticos y dificultades e imprecisiones metodológicas que hacen que deba tomarse con reserva. Aunque éticamente la terapia génica somática no genera serios problemas, pues en la mayoría de los casos se trata de enfermedades fatales o incurables por métodos terapéuticos convencionales, si existen ciertos riesgos de tipo metodológico para su ejecución. El uso de vectores de inserción retrovirales plantea la posibilidad de producción de mutaciones por inserción que puedan tener el efecto de activar un gene nuevo, lo más arriesgado sería la activación de un protooncogene que cause un efecto de malignización celular, o por inactivación en la expresión de un gene activo, lo que puede traer consecuencias bastante serias e impredecibles.

La historia de la eugenesia sugiere que una vez que una característica humana ha sido etiquetada de defecto, es de esperar que se levanten voces dentro de la sociedad para pedir su eliminación en nombre de la higiene génica, esto es errado y peligrosamente ingenuo. La imperfección génica es una característica inevitable de los procesos hereditarios humanos. Además es fuente de variabilidad genética, de la cual depende la selección natural. A pesar de todo lo anterior, la medicina, no ha ignorado la parte germinal. El consejo genético es parte de las predicciones sobre la variabilidad genética en la línea germinal de nuestra especie. La tecnología de la manipulación con embriones humanos ya es una puerta abierta a la implementación de terapia génica en la línea germinal en humanos, los efectos del futuro de la terapia génica todavía están por evaluarse.

Uno de los mayores impedimentos éticos para efectuar terapia génica germinal es precisamente el futuro que seguirá la información en la descendencia. No sabemos de manera exacta como se integrará el gene corrector al genoma del óvulo. Además la microinyección de ADN en óvulos no es una estrategia segura por lo menos en las actuales condiciones investigativas.

## LA FARMACOGENÓMICA

*“Desde los días de Tintagel han surgido nuevos magos. Pero han cambiado las armaduras por batas de laboratorio. Y se están haciendo ya una pregunta a lo que antes no se le daba vueltas sino en el mundo del mito. Por ella son tantos los que se interesan en los problemas de la longevidad”*

**JOHN MEDINA “THE CLOCK OF AGES”  
(El reloj de la edad). 1996.**

Los médicos empiezan a recetar los medicamentos en función del perfil genómico del paciente. Una nueva disciplina de la era postgenómica, la farmacogenómica, pretende mejorar la eficacia de los fármacos y evitar sus efectos nocivos. Los así denominados súpermedicamentos serán diseñados para cada paciente y sin efectos secundarios.

### **GENES HUMANOS Y GENES MURINOS EN LA ERA DE LA GENÓMICA**

No parece haber muchas dudas del papel que van a jugar en el futuro la posibilidad de realizar diagnósticos presintomáticos (facilitados por el acceso rápido y barato a métodos de secuenciación del ADN) y las terapias génicas para muchas enfermedades. Sin embargo, el mayor impacto sobre la medicina podría venir de la colaboración entre la genética molecular y la farmacología, lo que se ha llamado farmacogenómica. ¿Qué es la farmacogenómica y cómo podría influir en la práctica médica? Muchos de los fármacos empleados en la actualidad se descubrieron, mediante un proceso de tamizado, en el que primero se aisló la sustancia y después se probaron sus efectos sobre animales. Uno de los objetivos de la investigación farmacogenómica será identificar qué proteínas de nuestro organismo están implicadas en el



desarrollo de una enfermedad. Los farmacólogos deberán entonces diseñar moléculas para inhibir la función de esas proteínas o para estimular su producción. Puesto que las proteínas están codificadas por genes, hablar aquí de una proteína es lo mismo que hacerlo del gen que la codifica. La secuencia de aminoácidos de una proteína determina su estructura tridimensional, y ésta nos da pistas muy fiables de su función. Así, en el mismo artículo que en 1996 publicaba la secuencia del segundo gen de la poliquistosis renal del adulto (situado en el cromosoma 4) se describían la secuencia de aminoácidos de la proteína (policistina 2, PKD) y su estructura tridimensional (proteína con seis dominios transmembranales) y se pronosticaba su función más probable, por analogía con otras proteínas conocidas podría ser un canal para el calcio. Nadie había visto esta proteína en la membrana, ni el flujo de calcio a través de ella. Ni tiene que decir que estos trabajos finalizan señalando las posibilidades terapéuticas del descubrimiento: en el caso de la poliquistosis renal podrían ensayarse agentes que regulen el transporte de calcio para tratar la enfermedad.

Una de las consecuencias más inmediatas de la secuenciación de un gen humano es la «fabricación» de un modelo animal, casi siempre el ratón, en el que la función del gen está modificada, en general silenciada (ratones «knock-out»). No olvidemos que la inmensa mayoría de los genes humanos tienen un homólogo murino y que las secuencias de nucleótidos y las funciones de las proteínas son más o menos parecidas entre las dos especies. Por utilizar un ejemplo reciente, tras la descripción de la secuencia del receptor de tipo 2 de la angiotensina se creó un ratón que carecía de copias funcionales de este gen, por lo que no expresaba este receptor. Así descubrió su efecto hipotensor, en contraposición al efecto hipertensor del receptor de tipo 1. La manipulación farmacológica de esta proteína se convierte así en una línea de investigación para el tratamiento de las patologías en las que juega un papel fundamental el sistema de la angiotensina, entre las que se incluyen enfermedades renales o cardiovasculares.

### **VARIACIÓN GENÉTICA Y NUEVOS BLANCOS TERAPÉUTICOS**

Por motivos obvios, la información que nos dan los modelos animales no podría ser completamente obtenida en el hombre. Sin embargo, hay una aproximación experimental para la búsqueda de nuevas vías en el tratamiento de las enfermedades que sí contempla el empleo de pacientes. La mayoría de las enfermedades tienen componentes ambientales y biológicos (es decir, genéticos). Éstos actúan sobre el origen de la enfermedad (riesgo de desarrollarla) o una vez iniciada, modulando su progresión. Así, la enfermedad coronaria es desencadenada por factores como el tabaco o la dieta. Sin embargo, también reconocemos que hay diferencias biológicas interindividuales, de forma que en el mismo ambiente algunas personas sufren un infarto más precozmente que otras. En otros casos, como el del Alzheimer, las personas nacerían con una predisposición a desarrollar la enfermedad y el ambiente sería un simple modulador del riesgo de que se manifieste más o menos pronto o progrese

más o menos lentamente (el caso del Alzheimer es representativo de una problemática que se nos «viene encima», ya que el principal factor de riesgo es ambiental pero inevitable: el envejecimiento). Por tanto, el problema sólo puede ser planteado en términos de cuáles son los factores biológico-genéticos que hacen que algunas personas no desarrollen la enfermedad o lo hagan a edad más tardía que otras. La secuencia a la que nos referimos al hablar del genoma humano es un patrón de referencia. En realidad tendríamos que hablar de tantas secuencias como seres humanos existen, han existido o existirán. No hay dos personas con la misma secuencia (salvo en el caso de los gemelos monocigóticos). Por término medio, uno de cada mil nucleótidos es variable, por lo que en el conjunto del genoma habría más de 4 millones de posiciones polimórficas en nuestro genoma. Por ejemplo, en el gen de la enzima convertidora de angiotensina (ECA) hay varias posiciones polimórficas. Para un determinado polimorfismo, las personas pueden ser heterocigotas u homocigotas. En el caso de la ECA, para el polimorfismo adenina (A) /timina (T) en la posición -240 (región reguladora de la expresión del gene), los tres genotipos son AA, AT y TT. Recordemos que esta característica genética la heredamos de nuestros padres. Aunque esta variación genética no se traduce en diferencias en la secuencia de aminoácidos de la proteína, si afecta a la expresión del gen y por tanto, la cantidad de la enzima que las células sintetizan. En otros casos, como el de la apolipoproteína E (APOE) y su bien documentado polimorfismo E2-E3-E4, sí se produce un cambio en la secuencia de aminoácidos que modifica la función de la proteína. Así, la variante E4 tiene mayor afinidad por el receptor de las lipoproteínas y en una dieta alta en grasas, sus portadores tienen niveles de colesterol en la sangre más altos que los homocigotas.

Los casos de la ECA y la APOE podrían generalizarse a todo el genoma, ya que la mayoría de los genes tiene al menos un polimorfismo que conlleva diferencias en la cantidad (expresión) o en la función de la proteína. Si queremos investigar el papel de esta variación en la progresión de una enfermedad en la que el gen (la proteína que codifica) esté implicado, procederemos a analizar enfermos. Si un genotipo es más frecuente en los enfermos que en los controles sanos, significa que esta variación genética (y por tanto este gen) influye sobre el origen de la enfermedad. Una conclusión inmediata es que cualquier fármaco (conocido o por descubrir) que actúe sobre la proteína codificada por el gen debería ser ensayado como posible tratamiento. Por ejemplo, puesto que el sistema de la angiotensina actúa a nivel cerebral y podría participar en determinadas enfermedades neurológicas, tiene sentido analizar el papel de la variación en el gen de la ECA en el desarrollo del Alzheimer. Al menos dos estudios recientes sobre tres poblaciones diferentes sugieren que las personas con un genotipo de expresión baja son más frecuentes entre los enfermos. Aunque podemos estar aún lejos de hallarle aplicación terapéutica a estos resultados, parece lógico iniciar líneas de investigación para definir si la «manipulación» farmacológica del sistema de la angiotensina podría contribuir al tratamiento de la enfermedad de Alzheimer.

## **LA INDIVIDUALIDAD GENÓMICA Y LA PROGRESIÓN DE LAS ENFERMEDADES**

Aunque la investigación biomédica persigue evitar la aparición de las enfermedades o conseguir la curación de los pacientes, estos objetivos no serán alcanzables a corto plazo. En la mayoría de las enfermedades existen grados de afectación o de progresión, que pueden ser debidos a factores ambientales o a diferencias genéticas entre los pacientes. Consideremos los casos de la poliquistosis renal o la nefropatía diabética. En las dos condiciones patológicas tenemos pacientes que progresan rápida o lentamente hacia la insuficiencia renal. Si estas diferencias en la progresión tienen una base genómica, descubrirla abriría nuevas expectativas en el tratamiento de las enfermedades. Si no podemos curar, tratemos al paciente para que la enfermedad progrese lo más lentamente posible y el paciente pueda tener un mejor estar de vida.

Para descubrir qué genes controlan la progresión (genes modificadores) analizamos la variación genética en los progresores rápidos y lentos. Así, varios estudios han concluido que el genotipo de producción alta de ECA es más frecuente entre los poliquísticos con una progresión más rápida hacia la insuficiencia renal (55 vs 65 años). El tratamiento de los pacientes poliquísticos con inhibidores de la ECA u otras moléculas que reduzcan la actividad del sistema de la angiotensina podría contribuir a ralentizar la progresión de su enfermedad. Esto sería extensible al caso de la nefropatía diabética y la glomerulonefritis, en las que los pacientes con genotipo de producción alta progresan más rápidamente hacia la insuficiencia renal. Sin embargo, no debemos olvidar que la probabilidad de tener éxito en el tratamiento de las enfermedades progresivas será tanto mayor cuanto más precozmente sea aplicado. Para descubrir qué variación genética contribuye a modular la progresión de cualquier enfermedad necesitaremos un gran número de pacientes correctamente clasificados. Esta tarea (la clasificación clínica) puede llegar a ser más difícil y laboriosa que el análisis genético, el cual se irá simplificando a medida que la tecnología progrese. Puesto que hemos llegado a un punto en el que fenotipificar es más difícil que genotipificar, es fundamental el papel de los investigadores clínicos en el desarrollo de la farmacogenómica y en general, en cualquier aplicación médica de la genética molecular.

## **VARIOS GENES, UNA ENFERMEDAD**

Quienes se dedican a investigar el papel de la variación genética en el origen o en la progresión de las enfermedades son con frecuencia tachados de reduccionistas. Les recordamos que, dejando de lado el papel del ambiente, la mayoría de las enfermedades son genéticamente complejas puesto que no dependen de un solo gen. Ésto es cierto y ni siquiera la mente más reduccionista dejará de reconocerlo. En muchas ocasiones la visión simplista de la genética no parte de los propios científicos, sino que obedece al gusto de los medios de comunicación

por la simplificación sensacionalista. Los ratones transgénicos que sobreexpresan el gen NMDA, que codifica el receptor del neurotransmisor N-metil-D-aspartato (una sustancia secretada por neuronas y que participa en la transmisión del impulso nervioso), tienen mejor memoria que sus hermanos no manipulados. Esta mejora en la capacidad de aprendizaje de los ratones fue presentada por algunas revistas divulgativas como el descubrimiento del gen de la inteligencia. El genotipo de baja producción de ECA es más frecuente entre la élite de los deportistas de resistencia (ciclistas y maratonianos), algo lógico si consideramos el papel del sistema de la angiotensina en el control de la fisiología cardiovascular. Aunque sea más probable que un campeón del mundo contrarreloj tenga un genotipo de baja producción de ECA, sólo los medios divulgativos pueden presentar estos resultados como el hallazgo del «gen del deportista». Sin embargo, no olvidemos que por muchos que sean los genes que actúen sobre el origen o la progresión de una enfermedad, ésta no deja de tener un componente genético. Muchos estudios están descubriendo que la variación en un gen puede no aparecer, por sí sola asociada a una enfermedad, pero cuando se analiza en combinación con la variación en otro gen podemos ver un efecto sinérgico. Así, el genotipo de alta producción de ECA confiere un riesgo bajo de desarrollar infarto de miocardio precoz (por debajo de los 50 años) entre los fumadores. La variación en el gen del receptor tipo 1 de la angiotensina tampoco aparece asociada al desarrollo precoz de esta enfermedad. Pero cuando se analizan los dos genes en conjunto aparece una combinación que incrementaría más de 5 veces el riesgo de desarrollar infarto de miocardio precoz (tengamos en cuenta que el sistema de la angiotensina lo componen varias proteínas que interactúan para producir una respuesta biológica). Una conclusión inmediata de estos estudios es que debería ensayarse la combinación de dos o más fármacos dirigidos a bloquear ambos componentes del sistema (como inhibidores de la ECA y antagonistas del receptor), para prevenir el desarrollo de la enfermedad coronaria. Otro ejemplo es la interacción entre la variación en los genes de la ECA y de la óxido nítrico sintetasa endotelial (ecNOS) (una enzima esencial del metabolismo de muchas células). Un polimorfismo en la región que regula la expresión de este gen (-786 T/C) ha sido asociado al desarrollo de espasmo coronario entre los japoneses. El alelo C determina menor expresión de la ecNOS y por tanto, menor producción de óxido nítrico (un agente vasodilatador). Los homocigotos CC parecen ser más susceptibles al desarrollo de espasmo coronario, y entre los fumadores este genotipo conferiría un riesgo moderado de desarrollar infarto de miocardio precoz. Sin embargo, tener este genotipo y el de alta producción de ECA incrementaría unas 4 veces el riesgo de sufrirlo. La combinación de fármacos que inhiban la actividad de la ECA y estimulen la actividad de la ecNOS podría tener así más valor terapéutico que el empleo de cualquiera de ellos por separado.

### **VARIACIÓN GENÓMICA Y RESPUESTA A LOS TRATAMIENTOS**

Hasta ahora nos hemos referido a una de las líneas de actuación de la genómica, consistente en la búsqueda de nuevos blancos terapéuticos a partir del descubrimiento del papel de un gen

en el origen o en la progresión de una enfermedad. Existe otra línea de trabajo consistente en investigar cómo la variación en un gen puede influir en la respuesta al tratamiento con un fármaco determinado. La farmacogenómica debe su nombre al hecho de que los pacientes responden de formas diferentes a los tratamientos y esta variación en la respuesta clínica a un fármaco tiene su origen tanto en factores ambientales como genéticos. Quizá, uno de los ejemplos más ilustrativos venga de la nefrología. Son varios los estudios que han demostrado que la respuesta a los inhibidores de la ECA es mucho peor entre aquellos pacientes con un genotipo de alta producción de la enzima. Así, después de ser tratados con Captopril, los pacientes con nefropatía diabética y un genotipo de producción baja de ECA presentan una menor proteinuria (alta cantidad de proteínas en la orina) que los pacientes con el genotipo de producción alta. Otro ejemplo lo proporciona un estudio reciente que ha descrito cómo el polimorfismo alanina/cisteína en la posición 1166 del gen del receptor tipo 1 de la angiotensina, condicionaría la respuesta al Losartán. Los estudios de la relación entre la variación genética y las respuestas a fármacos van a generalizarse en el futuro inmediato. Los primeros resultados obtenidos sugieren que el genotipado de los pacientes podría ser una práctica habitual para decidir qué tratamiento farmacológico aplicar o qué dosis administrar. La elaboración de un carnet o perfil genético de cada paciente para valorar cómo ha de ser tratado podría ser una práctica más cercana de lo que pensamos. Sin embargo, la complejidad de los ensayos clínicos requeridos y el volumen de la información necesaria para demostrar el papel de la variación genética en las respuestas a los tratamientos, dan de nuevo al clínico un papel central en el desarrollo de la farmacogenómica. En la tabla 11.1 presentamos un resumen de las características funcionales de algunas drogas sobre la acción de genes asociados con enfermedades en los humanos.

### **EL PROTAGONISMO DE LOS CLÍNICOS**

La investigación sobre el genoma humano nos permitirá disponer de unas herramientas de un poder nunca conocido. Quizá esto no sea más que una consecuencia del natural progreso hacia una interpretación de los fenómenos biológicos en términos físico-químicos, algo que nos debería permitir superar el empirismo que ha caracterizado a la medicina. Sin embargo, no olvidemos que aunque se puede determinar la secuencia del genoma del *Homo sapiens* a partir del ADN de un solo individuo de la especie, la aplicación de todo este conocimiento al tratamiento de las enfermedades va a requerir grandes esfuerzos en el campo de la investigación clínica.

**Tabla 11.1** Efecto de algunos fármacos sobre la acción de algunos genes implicados en enfermedades.

Enfermedad	Gen Implicado	Fármaco	Efecto genético
Infarto	N-acetiltransferasa (NAT2)	Procainamida	La procainamida puede propiciar graves infecciones en personas portadoras del gen defectuoso
Tuberculosis	N-acetiltransferasa (NAT2)	Isoniazida	El fármaco causa entumecimiento de las extremidades en pacientes con baja actividad de la NAT2
Esquizofrenia	Receptor de la serotonina (subtipo 5-HT2A)	Clozapina	Los individuos con dos copias de la forma C102 de este gen responden peor al tratamiento
Alzheimer	Apoproteína E (ApoE)	Tacrina	El fármaco se muestra menos eficaz en pacientes con la variante genética ApoE4
Depresión	CYP2D6	Desipramina	Una mutación en este gen ralentiza el metabolismo del medicamento y aumenta el riesgo de intoxicación
Asma	Receptor- $\beta$ adrenérgico	$\beta$ -agonistas	La variante del gen hace a los pacientes menos sensibles a la medicación
Colesterol alto	CEPT, LPL	Pravastatina	Mutaciones en estos genes restan eficacia al fármaco
Cáncer	Resolvasa	5-Fluoracilo	Una variante del gen hace que el medicamento cause graves trastornos gastrointestinales

**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

## EL PROYECTO PROTEOMA HUMANO

*“Aunque la química propia de la vida terrestre dependa en poca medida de la historia, todavía queda una enorme cantidad de complejidad efectiva dentro de la biología, muchas más que en la química o la física de la materia condensada”*

MURRAY GELL-MANN, “THE QUARK AND THE JAGUAR ADVENTURES  
IN THE SIMPLE AND COMPLEX”  
(El Quark y el Jaguar aventuras de lo simple y lo complejo). 1995.

En la terminología de la era postgenómica son muy abundantes los denominados “omas” que se refieren a términos conceptuales entre los que podríamos destacar genoma, transcriptoma, metaboloma y proteoma. En realidad, con ellos se trata de describir los diferentes niveles de análisis con los que podemos abordar el funcionamiento de nuestras células, órganos, tejidos u organismos. De este modo, el genoma es el conjunto de genes. Una parte de esos genes se transcriben en forma de ARN mensajero, cuyo conjunto sería el *transcriptoma*. Los ARN mensajeros codifican la síntesis de proteínas y el conjunto de todas las proteínas obtenidas directamente o tras diversas transformaciones constituye el *proteoma*, mientras que el término *metaboloma* se reservaría al conjunto de los diferentes metabolitos existentes, con diversas naturalezas químicas.

### **EL PROTEOMA**

El Proteoma, término acuñado en 1994 por Marc Wilkins de la empresa *Proteome Systems Limited* (PSL) con sede en Sydney (Australia) designa al «conjunto de las proteínas expresadas por un genoma». Se emplea también para referirse al arreglo de las proteínas de la célula y es definitivamente más complejo que el mismo genoma. Aunque un gen puede



codificar una proteína, las proteínas se modifican de muchas formas después que son construidas. De esta manera, cada producto del gen puede dar lugar a docenas de proteínas que han sido reorganizadas, fragmentadas o que se han modificado químicamente para producir una actividad apenas diferente. Se podría pensar que estas proteínas modificadas van a ser elementos claves para entender su función y eventualmente, su fisiología.

### **LA PROTEÓMICA**

Mientras que la genética ha estudiado los mecanismos de la herencia, que se expresan a través de los genes, la genómica, intenta conocer la naturaleza íntima de los genes en su conjunto y determinar su funcionamiento global. Los logros de la genómica han estado y estarán, en gran parte, íntimamente ligados a los del Proyecto Genoma Humano, lo que, sin duda repercutirá muy favorablemente en numerosas facetas relacionadas con nuestra vida, ambiente, bienestar, salud o enfermedades.

En un artículo titulado «Proteomics in Genomeland», publicado en el número del 16 de febrero de 2001, de la revista *Science*, Stanley Fields investigador del Instituto Howard Hughes escribe: *«En el país de las maravillas de secuencias completas, quedan muchas cosas que el estudio genómico no puede hacer y es por esto que el futuro pertenece al estudio proteómico, análisis total de los complementos de proteínas»*

Las proteínas son las verdaderas expresiones funcionales del genoma, de los genes y los conocimientos actuales hacen insostenible lo que fue denominado como uno de los dogmas de la biología, es decir, la idea de un gen-una proteína. El mundo de las proteínas se nos está apareciendo como un mundo cuantitativamente (y posiblemente, cualitativamente) mayor y más complejo que el de los genes y además, aún, nos es bastante desconocido. Ante una situación concreta, muchas veces no sabemos cuáles son las proteínas que se expresan, como se cuantifica esa expresión, ni las consecuencias que sobre la misma ejercen diversas modificaciones posibles, de sufrir por parte de una proteína original, que la lleva a convertirse en diferentes modalidades de esa proteína con funciones muy diversas. Por otra parte, cada vez es más evidente, que una misma forma de proteína, en un ambiente biológico determinado, puede tener una función que sea muy diferente a la que ejerce en otro ambiente distinto.

Por tanto, la situación es que estamos comenzando a descubrir como, desde un cierto número de genes se puede producir un número mucho mayor de proteínas y empezamos a comprender que los mecanismos de modificación postraduccionales (tras su biosíntesis) de las proteínas tales como su fosforilación, glicosilación, polimerización, entre otros, causan efectos en su función biológica, además, desconocemos casi totalmente aspectos tan importantes sobre la funcionalidad de las proteínas como son las influencias medioambientales o las relaciones multigénicas que subyacen en la mayoría de las enfermedades o en fenómenos como el del envejecimiento. Todo ello no podría aclararse por el escueto examen y conocimiento del genoma.

Para muchas personas quizá supuso una decepción saber que “sólo” poseemos unos 40.000 genes, un número no mucho mayor que el de otros organismos menos complejos en la escala evolutiva. Para tales personas podría servirles de consuelo, conocer que la relación entre número de genes y proteínas es sólo de 1:2 en bacterias, de 1:3 en levaduras, algo superior en organismos intermedios, pero en los humanos ese número es mucho más elevado, posiblemente en el rango de 1:6 a 8.

### ***PPH EL PROYECTO PROTEOMA HUMANO***

Con base en los hechos ya descritos, es evidente que la investigación postgenómica tiene que ir más allá del genoma. Decimos que se ha iniciado la era postgenómica y algunos científicos manejan el término de operómica, como el conjunto de abordajes de estudio a realizar durante todo el trayecto que va desde el ADN, pasando por el ARN, hasta las proteínas y el análisis molecular y celular de sus funciones.

Posiblemente quedará fijada la fecha de abril del 2001 como la del nacimiento del Proyecto Proteoma Humano (PPH), coincidente con la primera reunión internacional que se celebró para presentar a HUPO (“Human Proteome Organization), es decir, la organización mundial creada para coordinar y estimular todos los estudios proteómicos que se pretenden integrar dentro del PPH. Como presidente de HUPO se eligió a un prestigioso científico, el Dr. Sam Hanash, un pionero en la proteómica del cáncer, cuyas muy valiosas contribuciones científicas incluyen, desarrollos de complejas tecnologías analíticas, construcción de bancos de datos de las expresiones proteicas, integración de datos genómicos y proteómicos, hasta ciertas aplicaciones clínicas como las posibilidades proteómicas en la clasificación de las enfermedades, en el desarrollo de nuevos medicamentos o el descubrimiento de nuevos biomarcadores aplicables al diagnóstico precoz del cáncer.

### ***OBJETIVOS DEL PPH***

Partiendo de la realidad inicial, que las organizaciones ya integradas en HUPO cuentan con más de mil millones de dólares de financiación, el primer congreso sobre el PPH logró establecer dos objetivos más o menos generales a cumplirse en un periodo de cinco años (2001-2006): 1. Crear un catálogo general de proteínas humanas, en el que se incluyan todas las variantes posibles para cada proteína. 2. Estimular el conocimiento de las interacciones entre proteínas y proteínas o entre proteínas y ácidos nucleicos; sin olvidar las investigaciones para descubrir los mecanismos que gobiernan los niveles relativos de expresión y las formas de esa expresión de las proteínas de cada tejido u órgano en situaciones de salud, enfermedad o terapia.

## **LOS CHIPS DE PROTEÍNAS**

En los capítulos anteriores nos referimos a los microarreglos de ADN o los denominados Chips de ADN como instrumentos importantes para el análisis de la genómica funcional. Para la detección de péptidos y proteínas, además de las técnicas de electroforesis bidimensional (two-dimensional gel electrophoresis, o 2DE) y la cromatografía líquida con espectrometría de masas (liquid chromatography - mass spectrometry, o LC-MS), recientemente han aparecido al menos dos sistemas de biochips adaptados a la identificación de estas moléculas, cuyo desarrollo promete abaratar y agilizar los procesos de caracterización de estos biocompuestos. Ambos sistemas se fundamentan en la especificidad de la unión entre proteínas, algo que suele compararse con el ejemplo de una llave que hace juego con su cerradura. Para poder utilizar esta propiedad, se crean matrices de microsúperficies capaces de unirse de forma específica a diversos tipos de péptidos y proteínas.

Las diversas técnicas para lograr este propósito, así como los procedimientos utilizados en la detección de las mismas pueden diferir bastante, y por ello dar lugar a biochips también distintos.

En todos los casos la muestra a estudiar es preparada (a veces la preparación puede ser mínima en flúidos como suero, sangre, orina; pero si se trata de tejidos, estos se machacan y homogeneizan para crear una «sopa de proteínas»), a continuación se depositan unas pocas gotas sobre el chip, se deja transcurrir un tiempo variable para que ambos interactúen (tiempo durante el que las proteínas se unirán a las diversas regiones específicas de aquel por procesos de absorción según su afinidad con las superficies tratadas) y se pasa a la etapa de revelado o detección de peptídica.

Veamos con algo más de detalle uno de estos chips proteicos, el denominado Sistema ProteinChip que fue desarrollado por la compañía *Ciphergen Biosystems Ltd.*, de Palo Alto (California). Está formado de un sustrato de aluminio en el que existen matrices con multitud de pequeñas superficies de hasta 1 mm de diámetro que han recibido un recubrimiento químico (hidrófobo, hidrofílico, iónico, etc) o bioquímico (con receptores, anticuerpos, etc) para interactuar con determinadas proteínas, unirse a ellas y retenerlas.

## EL NEGOCIO DE LA GENÓMICA

*“El producto de las nuevas tecnologías es también información. Su inclusión en bienes y servicios, en decisiones, en procedimientos, es el resultado de la aplicación de su producción informacional, no de la producción en si misma”.*

LUIS JOYANES.

**“Cibersociedad los retos sociales ante un nuevo mundo digital”. 1997**

Es anecdótico que la emergencia de la primera empresa biotecnológica del mundo se hubiese dado entre cervezas que son el producto de la práctica biotecnológica más antigua del mundo, la fermentación. Es curioso que cuando se visita la *Genetech Inc.* se puede ver, a la entrada principal, una escultura de bronce de tamaño natural de dos personas sentadas en una mesa celebrando con bebidas en sus manos. Fue así como en 1976 Herb Boyer, profesor de la Universidad de California en San Francisco (UCSF) y Robert Swanson, un inversionista de capital de riesgo, acordaron crear la compañía Genetech (*Genetech Inc.*) para producir, mediante técnicas de ADN recombinante, proteínas útiles en su aplicación médica que pudiesen ser comercializadas. La insulina humana, la primera de estas proteínas elaborada por técnicas de ADN recombinante, apareció en el mercado en 1982 seis años después de la creación de la Genetech Inc. En la década de los noventa del siglo pasado, el negocio de la biotecnología parecía muy prometedor, prueba de ello es que a 1997 existían solo en Estados Unidos 1.200 compañías de biotecnología que empleaban una fuerza laboral de 140.000 personas, el negocio era tan prometedor que se creó una cornucopia de nuevos fármacos para desarrollar nuevas medicinas además de considerar importantes objetivos en la agricultura, produciendo plantas transgénicas. La empresa del ADN recombinante había llegado a ser un renglón importante en el esquema de producción capitalista.

## LA FIEBRE DEL ORO DE LA BIOTECNOLOGÍA

El encanto de la biotecnología ha atraído a muchos inversionistas, aunque a la gran mayoría no les ha ido bien, si tuvo un impacto grande en las bolsas de valores de todo el mundo. En el mercado de valores de alta tecnología el indicador de tecnología NASDAQ ha sufrido incrementos importantes a causa de la comercialización de productos recombinantes. Entre 1994 y 1999, en la bolsa de valores de Estados Unidos, el índice de valores de biotecnología (^ IXB) se elevó 24%, periodo en el cual un índice comparable de valores de las compañías farmacéuticas líderes (^ DRG) tuvo un incremento de 150% mientras que los promedios del Dow Jones industrial y del índice de tecnología NASDAQ fueron de 100%. Es importante advertir que la mayoría de las empresas farmacéuticas líderes desarrollan y comercializan productos a base de ingeniería genética. Desde su apogeo, pocas compañías biotecnológicas han fracasado, sin embargo la gran mayoría sobrevivió debido a la aplicación de un arsenal de "gimnasia financiera" que produjo fusiones con grandes empresas o inyección de capital de inversión a riesgo por parte de transnacionales de diferente tipo. Actualmente muchas de las compañías biotecnológicas iniciales son satélites de otras más grandes. Solo unas 20 de las 1200 compañías biotecnológicas de Estados Unidos han producido ingresos que en 1998 ascendían a 17.000 millones de dólares en comparación con los 25.000 millones de dólares de *Merck*, un reconocido fabricante de fármacos, o de los 20.000 millones de dólares de la *Dell Computer*, una empresa que se inició hace unos diez años. Si comparamos entre si estas cifras, el negocio de la biotecnología no ha sido tan exitoso como se pronosticaba inicialmente. Qué ha sucedido entonces, tal vez una palabra que define de manera más acertada este proceso, es fusión. Sí, muchas de las empresas de biotecnología ahora hacen parte de grandes consorcios, aun mas *Genetech Inc.* gracias a una alianza con *Roche Inc.*, un socio doméstico, adquirió una seguridad financiera en la producción de fármacos recombinantes.

¿Que ha impulsado a que la mayoría de las empresas farmacéuticas mundiales hayan entrado en la era de la biotecnología molecular?. Esta es una pregunta cuya respuesta tiene como base la salud humana. Mediante ingeniería genética se pueden clonar y producir en grandes cantidades proteínas humanas para su utilización terapéutica; prueba de ello son entre otras, la hormona de crecimiento y la vacuna contra el virus de la hepatitis B, que han sido producidas por estas metodologías. Sin embargo hay una razón muy poderosa para que estas industrias participen del negocio de la biotecnología, es su futuro frente a los resultados del PGH. El genoma humano, es tal vez, la mayor fuente de materia prima para las industrias farmacéuticas, representa la mina de oro de la cual dependen muchos planes futuros a mediano y largo plazo, el genoma se convirtió en pocos años en un objeto indispensable para el industrial de la genómica. Pero ¿que hay de interesante en el genoma que atrae a los inversionistas, que a pesar del alto riesgo de inversión que hoy en día representa la industria genómica, invierten en él? , la respuesta es la información genética codificada en su ADN. Tal como lo hemos expuesto a lo largo de este libro, el genoma humano puede ser manipulado y

moldeado para producir obras maestras de la farmacología en la cura de la infección por el virus de la inmunodeficiencia humana, del cáncer y de un gran número de enfermedades que actualmente no poseen tratamientos exitosos o el desarrollo de moléculas que destapan las arterias obstruidas. En otras palabras el ADN es el oro de la industria genómica.

## **LA NUEVA INDUSTRIA GENÓMICA**

Si la técnica del ADN recombinante generó un paso importante en la experimentación física con el genoma, el PGH abrió una nueva frontera, la de la información contenida en el genoma como parte del conocimiento en red. En una sociedad que se dirige a pasos agigantados hacia la globalización total, la información contenida en el genoma, representa un valor agregado importante para ser explotado comercialmente.

A principios de los años noventa del siglo pasado se fundaron dos compañías dedicadas a la genómica, la *Incyte Genome* y *Human Genome Science*. Una característica general de estas dos empresas era el almacenamiento y manejo de información de secuencias de ADN y de potenciales blancos farmacológicos, esto representó una revolucionaria estrategia en la industria de la genómica, no era preciso desarrollar una gran capacidad tecnológica en máquinas de secuenciación, el objetivo era la información y su manejo. Más de 20 compañías farmacéuticas compraron el acceso a las bases de datos de *Incyte*, entre ellas se incluían siete de las más grandes farmacéuticas a nivel mundial. En solo cuatro años *Incyte* generó sus primeras ganancias hacia el último trimestre de 1996 y se distinguió de otras empresas de investigación de biología que solo consumían dinero. El enfoque revolucionario de *Incyte* generó una nueva industria sin laboratorios, solo con computadores y una excelente red de distribución de la información; se constituyó en el ejemplo perfecto de la función de una compañía en la sociedad de la información; la globalización llegó para quedarse en la era postgenómica.

Las compañías genómicas que actualmente existen aplican tecnologías innovadoras para adquirir, almacenar, analizar o distribuir todo tipo de información molecular. Algunas de ellas, tal como describimos en la tabla 13.1, se dedican a analizar e identificar, mediante electroforesis capilar y espectrometría de masas, miles de secuencias de proteínas. Otras desarrollan herramientas bioinformáticas que identifican motivos funcionales en las secuencias de ADN o aun el modelamiento molecular de secuencias de aminoácidos. Otras más crean dispositivos muy sofisticados para detectar y rastrear la localización de miles de proteínas utilizando miles de anticuerpos diferentes. Las actuales compañías farmacéuticas y de biotecnología se han acercado a la genómica de manera poco usual, planteando programas locales y de socios múltiples para identificación de nuevos blancos farmacológicos que permitan la producción de nuevas drogas que se utilizarán en el manejo terapéutico más acertado y exitoso de muchas patologías.

**Tabla 13.1.** Principales compañías genómicas actuales

Compañía	Capital de Mercado (*)	Tipo de Actividad
Applied Biosystem (PEB)	20,4	Herramientas para secuenciación, genotipificación, etc.
Agilent Technologies	19,0	Herramientas para la expresión genética, genotipos, SNPs
Millenium Pharmaceutical	13,1	Desarrollo de información y nuevos fármacos
Human Genome Science	8,8	Desarrollo de información y nuevos fármacos
Celera Genomics	3,8	Información de secuencias, interacciones con proteínas.
Affymetrix (AFFX)	3,7	Microarreglos de ADN, Chips de ADN y proteínas
Incyte Genomics (INCY)	2,0	Información; secuencias y expresión genética
(*). En miles de millones de dólares. (Datos del último trimestre del año 2000).		

Muchos de los objetivos de cada una de estas empresas los hemos explicado a lo largo de los diferentes capítulos de este libro y son tema de debate actual no solo en lo financiero si no en lo ético y lo moral.

### **LAS VARIACIONES DE UN SOLO NUCLEÓTIDO EN EL GENOMA GENERAN UNA NOVEDOSA ESTRATEGIA**

En abril de 1999 un grupo de compañías farmacéuticas y una fundación de caridad de los Estados Unidos, anunciaron la creación de una empresa sin ánimo de lucro conocida como el *Consortio SNP* (Consortio de los Polimorfismo de un Solo Nucleótido). La meta de este consorcio era identificar y crear el mapa de 300.000 variaciones de un solo nucleótido conocidas como SNP (*“Single Nucleotide Polimorphisms”*) en el ADN humano en los siguientes dos años. La información de estos puntos variables en el genoma estaría disponible para su utilización por parte de los investigadores y no sería patentable. Las diez empresas farmacéuticas, entre las que no estaba *Merck*; aportaron un capital de 30 millones de dólares mientras que *Wellcome Trust* donó 14 millones de dólares al consorcio. El consorcio complementaría un proyecto del gobierno de los Estados Unidos conocido, como dbSNP, para crear una base de 160.000 SNP que había sido anunciado en Diciembre de 1997.

En julio de 1997 la empresa francesa *Genset* anunció que desarrollaría un mapa de 60.000 SNP del ADN humano. Este proyecto tenía fondos de una sociedad con los laboratorios *AB-*

*BOTT* que aportó 42,5 millones dólares, es de aclarar que esta compañía no hizo parte del *Consortio SNP*. Recientemente *Celera Genomics*, empresa dirigida por Craig Venter, ha anunciado su intención de desarrollar una mapa de 2,4 millones de SNP, este anuncio puso en alerta a las demás empresas que habían invertido en el negocio de los SNP, puesto que existía el antecedente de la carrera por la secuenciación completa del genoma humano, la cual *Celera* completo en un tiempo record poniendo aprietos al Instituto Nacional de Genoma, institución que se había encargados desde 1987 del PGH. La historia de las iniciativas de descubrimiento de SNP a gran escala se resume en la tabla 13.2

### **¿POR QUE ES ATRACTIVA LA INVERSIÓN EN LOS SNP?**

Siendo una inversión de tan alto riesgo, además de que muchas de las instituciones participantes proclaman la no patentabilidad de los SNP, entonces ¿Cual es el interés comercial por los SNP?, ¿Qué beneficio obtienen los inversionistas?, la respuesta está en la ciencia que subyace al SNP y en las implicaciones de su aplicación en la producción de una nueva generación de fármacos con un mayor éxito terapéutico; en este último punto si hay una prospección de cientos, quizás miles de millones de dólares en un futuro no muy lejano. La capacidad para detectar fácilmente las variaciones de los SNP y relacionarlas directa o indirectamente con los rasgos, enfermedades o susceptibilidades humanas representa un poder que debe manejarse con cuidado puesto que su utilización inadecuada tendría graves consecuencias para nuestra sociedad.

Un SNP es un sitio en el genoma donde por ejemplo 30% de los humanos tengan un T mientras que el 70% restante posean una C. Pero ¿Cuándo se considera que un sitio en el genoma se convierta en un SNP?, se ha establecido que para considerarlas como SNP es indispensable que mas del 1% de toda la población humana (aproximadamente 60 millones de personas) tenga una variación de nucleótidos en ese sitio. Los SNP son antiguos, se iniciaron tal vez como mutaciones y se diseminaron a través de la población en el curso de miles de millones de generaciones. Algunos grupos de SNP son tan antiguos que se han detectado tanto en simios como en humanos.

Existen dos grupos de SNP: 1. Aquellos que se asocian con un fenotipo definido y 2. Los que son aparentemente neutrales (no se asocian con ningún fenotipo). De acuerdo con la composición del genoma humano, que posee un elevado grado de repetición, la gran mayoría de los SNP no estarían asociados a un fenotipo especial. Estudios que tienen como base la secuenciación de porciones del genoma de muchos individuos, sugieren que existen aproximadamente unos 9 millones de SNP. Sin embargo en las secuencias codificadores de genes existirían aproximadamente unos 400.000 SNP de los cuales unos 200.000 pueden alterar la secuencia de aminoácidos de una proteína.



**Tabla 13.2** Descripción de las diferentes instituciones con iniciativas de descubrimiento y comercialización de los SNP a gran escala

Institución	Fecha	Objetivo	Reservas (*)'	Estrategia	Estado
Geneset/Abbott	Julio/97	60K SNP para mapeo	42.5	Mapeo de genes asociados a enfermedades	Nose conoce públicamente
Celera	Mayo/98	Catalogo de variantes humanas	Nose conoce	Por alineación de secuencias en escopeta	2.4 millones SNP únicos a partir de 9/2000
Incyte/Hexagen	Agosto /98	100 K SNP en cADN	30	La mayoría por alineación	> 70Ka partir de 8/2000
NHGRI/NCBI	Septiembre/98	60-100 K SNP	Nose conocele	Envíos de toda la comunidad científica	26.5 K envíos a partir de 3/2000
Curagen	Noviembre/98	60 K SNP	Nose conoce	Genómica integrada para el desarrollo de fármacos	> 120 K SNPc a partir de 4/2000
Consorcio SNP	Abril/99	300 K SNP distribuidos aleatoriamente en el genoma a2002	44 inversión inicial	Hacer que los datos SNP estén disponibles al público sin ninguna restricción	Mapeo de 296.990 SNP a partir de 8/2000
Japón (Varios ministerios)	Mayo/99	100 K -200 K SNP en regiones codificadoras de genes al 2001	Nose conoce	Crear bases de datos publicas a partir de SNP encontrados en 50 individuos japoneses	Nose conoce
HGS/Compugen	Marzo/2000	500 K SNP en cADN	Nose conoce	Vía ESTy alineación de ADN genómico	No se conoce

(\*). De conocerse se dan en millones de dólares

## EL GENOMA HUMANO FRENTE AL FUTURO DE LA HUMANIDAD

*“En términos negativos, la especie humana tiene que evitar la guerra aniquiladora, la tiranía generalizada y la omnipresencia continua de la pobreza extrema, así como la desastrosa degradación de la biosfera y la destrucción de la diversidad biológica y ecológica”.*

**MURRAY GELL-MANN, “THE QUARK AND THE JAGUAR ADVENTURES IN THE SIMPLE AND COMPLEX” (El Quark y el Jaguar aventuras de los simple y lo complejo). 1995.**

El 9 de diciembre de 1998, la Asamblea General de las Naciones Unidas, aprobó la declaración de 25 artículos titulada, *“DECLARACION UNIVERSAL SOBRE EL GENOMA HUMANO Y LOS DERECHOS HUMANOS”* promulgada por la UNESCO en 1997. Dicho documento establece las directrices acerca de las conductas éticas y morales sobre las cuales debe regirse la investigación y el conocimiento que se obtenga sobre el genoma humano. En sus primeros cuatro artículos, se establecen las bases para la dignidad humana con relación al su genoma. A continuación transcribimos textualmente algunos apartes de esta declaración así como los artículos más importantes:

*“ Recordando que en el Preámbulo de la Constitución de la UNESCO se invocan «los principios democráticos de la dignidad, la igualdad y el respeto mutuo de los hombres y de las razas», se indica «que la amplia difusión de la cultura y la educación de la humanidad para la justicia, la libertad y la paz son indispensables a la dignidad del hombre y constituyen un deber sagrado que todas las naciones han de cumplir con un espíritu de responsabilidad y de ayuda mutua», se proclama que «esa paz debe basarse en la solidaridad intelectual y*

moral de la humanidad» y se declara que la Organización se propone alcanzar «mediante la cooperación de las naciones del mundo en las esferas de la educación, de la ciencia y de la cultura, los objetivos de paz internacional y de bienestar general de la humanidad, para el logro de los cuales se han establecido las Naciones Unidas, como proclama su carta»,.....”Teniendo presentes y sin perjuicio de lo que dispongan, los instrumentos internacionales que pueden concernir a las aplicaciones de la genética en la esfera de la propiedad intelectual, en particular la Convención de Berna para la Protección de las Obras Literarias y Artísticas del 9 de septiembre de 1886 y la Convención Universal de la UNESCO sobre Derecho de Autor del 6 de septiembre de 1952, revisadas por última vez en París el 24 de julio de 1971, el Convenio de París para la Protección de la Propiedad Industrial del 20 de marzo de 1883, revisado por última vez en Estocolmo el 14 de julio de 1967, el Tratado de Budapest de la OMPI sobre el Reconocimiento Internacional del Depósito de Microorganismos a los fines del Procedimiento en materia de Patentes del 28 de abril de 1977, el Acuerdo sobre los Aspectos de los Derechos de Propiedad Intelectual relacionados con el Comercio (ADPIC) anexo al Acuerdo por el que se establece la Organización Mundial del Comercio que entró en vigor el 1 de enero de 1995” ..... “Teniendo presente también el Convenio de las Naciones Unidas sobre la Diversidad Biológica del 5 de junio de 1992 y destacando a este respecto que el reconocimiento de la diversidad genética de la humanidad no debe dar lugar a ninguna interpretación de tipo social o político que cuestione «la dignidad intrínseca y (...) los derechos iguales e inalienables de todos los miembros de la familia humana», de conformidad con el Preámbulo de la Declaración Universal de Derechos Humanos,” ..... “Reconociendo que las investigaciones sobre el genoma humano y sus aplicaciones abren inmensas perspectivas de mejoramiento de la salud de los individuos y de toda la humanidad, pero destacando que deben al mismo tiempo respetar plenamente la dignidad, la libertad y los derechos de la persona humana, así como la prohibición de toda forma de discriminación fundada en las características genéticas.

*Proclama los principios siguientes y aprueba la presente Declaración:*

## **A. LA DIGNIDAD HUMANA Y EL GENOMA HUMANO**

### **ARTÍCULO 1**

*El genoma humano es la base de la unidad fundamental de todos los miembros de la familia humana y del reconocimiento de su dignidad intrínseca y su diversidad. En sentido simbólico, el genoma humano es el patrimonio de la humanidad.*

### **ARTÍCULO 2**

*a) Cada individuo tiene derecho al respeto de su dignidad y derechos, cualesquiera que sean sus características genéticas.*

*b) Esta dignidad impone que no se reduzca a los individuos a sus características genéticas y que se respete el carácter único de cada uno y su diversidad.*

### **ARTÍCULO 3**

*El genoma humano, por naturaleza evolutivo, está sometido a mutaciones. Entraña posibilidades que se expresan de distinto modos en función del entorno natural y social de cada persona, que comprende su estado de salud individual, sus condiciones de vida, su alimentación y su educación.*

### **ARTÍCULO 4**

*El genoma humano en su estado natural no puede dar lugar a beneficios pecuniarios.”*

El desentrañar el genoma humano le aporta a nuestra sociedad un vasto conocimiento que, evidentemente, tiene una parte positiva. Descubrir el origen de las enfermedades, poder desarrollar curas con terapia génica son algunas de las posibilidades que se presentan. Sin embargo, debemos tener cuidado con el impacto y los problemas que podrían desatarse con su aplicación práctica. La falta de precauciones podría implicar verdaderas perversiones de este conocimiento. Es indispensable que reflexionemos acerca de cuán grave puede llegar a ser su mal uso que nos lleve a pensar en la discriminación, el reduccionismo y el determinismo genético. Elementos como estos pueden atentar contra los principios que motivan a la ciencia para alcanzar conocimientos como el de los genes humanos.

### **SONDEOS GENÉTICOS HACIA UNA VERDADERA MEDICINA PREDICTIVA**

El Proyecto Genoma Humano se basa en la suposición de que la información contenida en el gene nos permitirá diagnosticar muchas enfermedades genéticas en el útero o incluso antes, gracias a esto podremos tomar decisiones aun antes de la procreación. Uno de los principales problemas éticos que plantea el Proyecto Genoma Humano será que su desciframiento dará a los médicos poderosas herramientas para diagnosticar enfermedades genéticas, aun antes de que sus síntomas se produzcan. Esto podría parecer un hecho beneficioso; efectivamente, así sería., si existiese una cura o por lo menos un tratamiento contra esas enfermedades, lo cual en la mayoría de los casos no existe actualmente. Lo cierto es que, una vez descubierto el gen asociado a una enfermedad, el desarrollo de un método de diagnóstico prácticamente seguro al 100% es hoy en día trivial y casi inmediato. Sin embargo, el desarrollo de terapias contra dichas enfermedades puede llegar a ser muchísimo más largo. Los investigadores más optimistas predicen que, hasta dentro de al menos veinte años, no poseeremos métodos de lucha eficaces contra prácticamente ninguna de las enfermedades genéticas conocidas. Sin embargo, todas las predicciones que se han hecho en el campo de la biología molecular han resultado siempre ser excesivamente pesimistas. La ciencia avanza más rápido de lo que los propios expertos pueden llegar a imaginar, lo que me permite ser optimista en este aspecto.

Sin embargo, a pesar de los sorprendentes avances, por el momento, el diagnóstico precoz de las enfermedades genéticas en la mayoría de las veces sólo supone para el paciente afectado una ominosa carga que debe soportar. El conocer que, en el plazo de una década, posiblemente sufriremos una horrible enfermedad, sin poder hacer nada por evitarlo, lo único que nos trae, es amargarnos la existencia viviendo con una sensación de impotencia y ansiedad que puede llevarnos a no disfrutar de los años de vida feliz y sana que nos hubieran correspondido si hubiéramos dejado a la naturaleza seguir su curso.

### ***¿ESTÁ USTED EN LA LISTA DE LOS DEFECTUOSOS GENÉTICOS?***

Lo que resulta aún peor en algunas ocasiones, los gobiernos han promovido este tipo de conocimiento sesgado y mal intencionado entre la población. En el recuerdo de muchos negros americanos está el caso de la campaña para detectar portadores del gen de la anemia falciforme que fue promovida por algunos gobiernos estatales de los Estados Unidos durante los años setenta. Esta campaña aunque no tuvo nada que ver con el Proyecto Genoma Humano y se realizó incluso antes de que se desarrollara la ingeniería genética si nos alerta a uno de los potenciales riesgos que nos enfrenta el futuro del genoma humano.

Conviene que echemos un vistazo a la historia de este caso. Las pruebas para diagnosticar en esa época la anemia falciforme tenían como base análisis bioquímicos de la proteína hemoglobina. La idea había surgido a partir de una campaña para detectar la fenilcetonuria. Esta enfermedad posee ciertas características que la hacen especialmente adecuada para una campaña de predicción. Se puede detectar fácilmente mediante un sencillo análisis de sangre desde el primer día de vida del recién nacido (en la actualidad este análisis es obligatorio en la mayoría de los países desarrollados). Si un niño padece fenilcetonuria, el tratamiento es bien sencillo. Basta con seguir una estricta dieta, baja en el aminoácido fenilalanina, durante los primeros años de vida, aunque no totales, sus beneficios son enormes; por el contrario, si no se trata a tiempo, se produce retraso mental y muerte prematura. La fenilcetonuria afecta aproximadamente a uno de cada 12.000 niños, si bien en la actualidad, gracias a estas campañas de prevención obligatorias, es una de las enfermedades genéticas más benignas. La anemia falciforme es un caso bien distinto, pero los legisladores estadounidenses de los años setenta no supieron ver las diferencias, ni intuir las consecuencias de su falta de previsión. La anemia falciforme es mucho más frecuente que la fenilcetonuria; de hecho, es la enfermedad genética más frecuente entre la población negra, con un caso por cada 400 niños. Es una enfermedad recesiva bastante cruel; los enfermos que han tenido la mala suerte de heredar los dos alelos recesivos de sus padres no pueden realizar esfuerzos ya que corren un grave riesgo de sufrir una insuficiencia respiratoria aguda que les ocasione repentinamente la muerte. Los individuos heterocigóticos, por el contrario, son portadores absolutamente sanos y pueden seguir una vida perfectamente normal. Cuando se desarrolló la técnica de detección basada en los análisis de la hemoglobina, numerosos colectivos de personas de

raza negra sufrieron un ataque de falsas esperanzas, pensando quizás que los científicos habían encontrado la solución a la enfermedad. Sin embargo, la cruel diferencia con la fenilcetonuria era que no existía tratamiento alguno. Si alguien era diagnosticado de anemia falciforme no poseía la más mínima esperanza de curación. El diagnóstico no representaba beneficio alguno para el paciente. Además, en aquella época no existía ningún método para examinar al feto en el útero, para que los padres tuvieran la opción de interrumpir el embarazo si es que el feto estaba afectado. En cualquier caso, el aborto fue ilegal en Estados Unidos hasta 1973.

A pesar de todos estos inconvenientes que hacían desaconsejable la campaña de detección, el gobierno federal financió un programa a nivel nacional. En varios estados, se declaró como obligatorio realizar la prueba a los recién nacidos y a los escolares, todo ello sin un programa paralelo de orientación genética que pudiera ofrecer consejo a las familias afectadas.

Pero lo peor fue que el público comenzó a confundir a las personas portadoras con las enfermas, debido a la completa falta de una campaña informativa. Muchos padres llegaron a pensar que sus hijos, en realidad portadores sanos, padecían una terrible enfermedad debilitante y debían recibir cuidados especiales para que no agravase. Para complicarlo todo, Linus Pauling, que había descubierto el método de análisis de la hemoglobina, realizó unas desafortunadas declaraciones, sugiriendo que se «marcara» a los portadores para que no se casaran entre sí o al menos, que no tuvieran hijos.

En 1978, Loretta Kopelman denunciaba que la prueba utilizada en el estado de Nueva York detectaba tanto a los enfermos como a los portadores del gen y que, aunque la ley exigía solamente que quedaran registrados los casos de enfermedad, también se «fichaba» a los portadores del carácter. Dicha información pasaba a formar parte permanente del historial médico del niño. Las compañías de seguros comenzaron a negarse a formalizar el seguro si descubrían que su posible cliente padecía anemia falciforme o era portador del carácter.

También el mercado de trabajo discriminaba a los enfermos y portadores. A las personas de color que portaban el gen se les negaba el trabajo en compañías aéreas e incluso el ingreso en la Academia de las Fuerzas Aéreas, porque se creía, erróneamente, que su sangre reaccionaría mal a las bajas presiones que se experimentan al volar a gran altitud. La Academia no levantó las restricciones hasta 1981.

El programa de detección de la anemia falciforme en Estados Unidos acabó degenerando hasta desaparecer. En 1987, un equipo de expertos designados por los NIH desenterró el viejo fantasma, recomendando de nuevo que se practicara la prueba en recién nacidos. Esta vez, la motivación era diferente. La investigación clínica había demostrado que los niños menores de tres años con anemia falciforme tenían menos capacidad de defensa contra las infecciones bacterianas, existiendo un 15% de probabilidades de morir a causa de una infección durante los primeros años de su vida. También se demostró que ésto se podía evitar administrando penicilina a los niños. Los NIH recomendaron que se administrara penicilina a todos los niños con anemia falciforme desde los cuatro meses a los cinco años de edad. Esta vez el análisis tenía su justificación y reportaba un beneficio claro. La campaña se lleva practicando con éxito desde entonces.

## **EL GENOMA HUMANO UN CONFLICTO A NIVEL MUNDIAL**

Desde su inicio el Proyecto Genoma Humano, fue claro para los científicos, suficientemente escarmentados por las experiencias previas, que no se podían efectuar campañas a gran escala del tipo de la que se realizaron con la anemia falciforme, sin haber realizado previamente un minucioso estudio ético y social. Aproximadamente, el 5% de los fondos destinados a la financiación del PGH se utilizan para llevar a cabo estos estudios, que se engloban bajo las siglas ELSI (Ethical, Legal and Social Issues). Así el PGH constituye el primer programa científico a gran escala que dedica específicamente parte de sus fondos a este tipo de preocupaciones. Se espera que en los próximos años, el porcentaje de fondos dedicados a los estudios ELSI aumente considerablemente. Sólo en 1997, se tenía prevista una inversión de 11 millones de dólares en estos aspectos. En la actualidad, se está llevando a cabo un estudio exhaustivo sobre la posibilidad de llevar a cabo un programa de detección de portadores del gen de la fibrosis quística, la enfermedad genética más común entre la población blanca. No se quiere, de ningún modo, caer en los errores que se cometieron con el caso de la anemia falciforme. Al menos, existen dos ventajas, a priori, que no existían hace veinte años: la posibilidad de detectar la enfermedad en el feto y la aceptación social mayoritaria del aborto con fines terapéuticos.

En el otro extremo se encuentran los sondeos de genes implicados en enfermedades poligénicas aún no bien comprendidas, que pueden dar lugar a otros tipos de problemas, distintos de los anteriores, pero no menos importantes. En estos casos no se habla de que «el paciente padece la enfermedad» sino de que, por ejemplo, «el paciente posee un 60% de posibilidades de que en un futuro pueda desarrollar la enfermedad», ya que dicho porcentaje de las personas que poseen esa variante del gen, la padecen. La medicina entonces se torna predictiva y si es posible, preventiva. Entre las posibilidades más relevantes están algunas de las causas fundamentales de muerte en el hombre: el cáncer, las enfermedades cardiovasculares o la diabetes, junto con otras como el enfisema o enfermedades mentales como el Alzheimer o la esquizofrenia.

Cada una de estas enfermedades representa un problema científico complejo, porque la expresión de estos trastornos incluye frecuentemente un componente genético que interactúa con varios factores ambientales, como la dieta o el estrés.

Por ejemplo, unos veinte genes regulan el nivel de colesterol en la sangre. Determinadas combinaciones de variedades de estos genes sitúan a la persona en un grupo de riesgo mayor de padecer enfermedades tempranas de las arterias coronarias y ataques cardíacos. Si además, la persona lleva una dieta rica en grasas animales y una vida sedentaria, es muy posible que muera de infarto antes de los cincuenta años. El problema es que aún no conocemos qué combinaciones de genes son especialmente peligrosas. Quizás, después de todo, no sean una veintena, sino un centenar, los genes implicados. El PGH, posiblemente, nos dará la respuesta.

## **DOCTOR QUE ME ACONSEJA**

Los adelantos de la tecnología médica han hecho variar la situación de las personas con riesgo de transmitir enfermedades genéticas, en comparación con la que existía en los años setenta, al comienzo del programa contra la anemia falciforme. Desde mediados de los setenta, los médicos han podido informar a las madres de si el niño que esperaban estaba afectado de algún trastorno genético. El diagnóstico prenatal se basa en distintos métodos de extracción de muestras de células del feto; hoy en día, se puede incluso analizar la sangre del feto.

Tras la fecundación, el óvulo, denominado «embrión», se transforma en dos células, luego en cuatro y así sucesivamente. En cada etapa, los genes, de algún modo aún no muy bien entendido, reciben la señal de activarse o seguir dormidos, esperando su turno. El embrión crece en tamaño y complejidad. Ocho semanas después de la fertilización, ha adquirido la forma de un ser humano en miniatura, se denomina entonces «feto», que quiere decir «el descendiente». Es un tránsito peligroso pasar de huevo a feto. Antes de que se dispusiera de métodos hormonales altamente sensibles para detectar precozmente el embarazo, se estimaba que un 15% de los huevos fecundados sufrían aborto espontáneo. Hoy día, sabemos que más del 80% de los óvulos fecundados no sobreviven, sucumbiendo silenciosamente sin ser detectados.

Oculto en el vientre, el embrión o feto no puede ser examinado directamente. Pero mientras el feto crece, está rodeado de un medio protector de líquido amniótico. El feto pierde algunas de sus células, cuyos restos van a parar al fluido amniótico. El análisis del fluido amniótico o «amniocentesis», fue durante mucho tiempo, posible sólo a partir de la decimosexta semana de embarazo. Para recoger una muestra se recurría al procedimiento de insertar una larga aguja en el saco lleno de fluido que recubre al feto. Se utilizaba al principio para determinar el factor Rh y para realizar el cariotipo fetal, lo cual permitía conocer el sexo del futuro descendiente o cualquier anomalía cromosómica importante. Hoy en día, las técnicas de PCR permiten analizar datos genéticos a partir de una muestra de tejido recogido mediante un procedimiento menos intrusivo, denominado biopsia del vello coriónico, a partir de las ocho semanas de gestación. Es importante aclarar que también del vello coriónico se pueden obtener cariotipos fetales. En 1978, Y. W. Kan y Andrees Dozy, de la UCSF, comunicaron que habían desarrollado una técnica para detectar la presencia del gen mutante de la anemia falciforme mediante el análisis del ADN de las células del fluido amniótico. Fue el primer diagnóstico prenatal por análisis de ADN. Las parejas que podían estar en riesgo de tener hijos afectados de la terrible enfermedad ya podían disponer de un método para conocer la futura salud de su hijo en las primeras semanas de embarazo, para asegurarse que el feto era sano o en caso contrario, poder optar por la interrupción del embarazo. Se imponía, por vez primera, un «control de calidad» de los fetos, con opción a desechar los que no cumplan las condiciones necesarias para pasarlo.



El PGH hará posible el diagnóstico prenatal de, teóricamente, unas 6.000 enfermedades genéticas. En la práctica, ningún feto podrá ser completa y exhaustivamente analizado. Sin embargo, con seguridad, llegará el día en que algunas decenas de análisis genéticos prenatales se conviertan en técnicas de rutina en las consultas de ginecología. La gran mayoría de mujeres cuyo futuro hijo no sea capaz de pasar el control de calidad, decidirán abortarlo. Muchos grupos, contrarios al aborto, ven en este tipo de diagnóstico algo aborrecible y absolutamente antinatural.

Quizás sea así, pero aún pueden plantearse serias dudas sobre la posibilidad de que el poder genético sesgue nuestras ideas. Casi todos estamos de acuerdo en que un feto afectado del síndrome de Tay-Sachs, que tendría una esperanza de vida de algunos meses, puede ser abortado sin muchos problemas morales. Las controversias y dudas surgen a medida que vamos subiendo en la escala de benignidad de las enfermedades genéticas: ¿abortaríamos a un feto con anemia falciforme? ¿O a un feto que padecerá la enfermedad de Huntington a los cuarenta años? ¿O a un feto que posee el 60% de probabilidad de desarrollar Alzheimer a los 75? ¿Es «anormal» un niño que posee el 35% de posibilidades de desarrollar cáncer a los 60 o un niño que posiblemente tenga una inteligencia por debajo de la media? ¿Qué hay de un niño que padecerá daltonismo o miopía?

En 1871, Charles Darwin, en *«El linaje del hombre»*, afirmaba que *«los miembros débiles de la sociedad civilizada propagan su especie. Nadie que haya ayudado en la crianza de animales domésticos dudará de que esto debe ser altamente injurioso para la raza humana... casi nadie es tan ignorante como para permitir que sus peores animales se reproduzcan»*. El primo de Darwin, Francis Galton, apoyó fervientemente esta filosofía y le dió un nombre que para él y para muchos otros, durante algún tiempo, poseía connotaciones positivas: «eugenesia», que deriva del griego «eugenes» o «dotado por herencia de cualidades buenas». Se puede definir la eugenesia como «los métodos para mejorar la calidad de la raza humana, en especial mediante la reproducción selectiva». Galton había promocionado la «eugenesia positiva», la procreación entre personas con una dotación genética *supuestamente* buena. Sus sucesores, como el americano de Cold Spring Harbor, Charles Davenport, promocionaron la «eugenesia negativa», el impedimento del apareamiento de personas con características supuestamente no deseables. Para 1930, 24 estados de los Estados Unidos de América poseían leyes que permitían la esterilización de una amplia variedad de «indeseables»: epilépticos, «insanos» o criminales habituales. En abril de 1924, el presidente Calvin Coolidge transformó en ley el Acta de Inmigración, que establecía cuotas a otras nacionalidades. Antes había declarado: «los Estados Unidos deben ser mantenidos americanos... Las leyes biológicas muestran... que los nórdicos se deterioran cuando se mezclan con otras razas».

Cuando Adolf Hitler defendió la esterilización eugenésica en 1923, apoyado por cientos de miles de alemanes, sus enfoques eran acordes con los de muchos respetables genetistas de Estados Unidos. A esto le siguió, en 1933, una «ley para la prevención de

las enfermedades hereditarias en las generaciones futuras». En 1934, 56.000 órdenes de esterilización fueron emitidas en Alemania. En 1939, el Tercer Reich fue mucho más allá de la esterilización y la prohibición del matrimonio, hasta el asesinato sistemático de los enfermos mentales y todos los judíos.

¿No resulta razonable, considerando todos estos abusos, que la misma palabra «eugenesia» se haya convertido en algo peyorativo? James Watson admitió que «las sombras de los pasados abusos flota en el fondo de la investigación genética. Podemos impedir que tales atrocidades vuelvan a ocurrir si los científicos, los doctores y la sociedad en general rehúsan a ceder el control de los descubrimientos genéticos a aquellos que podrían usarlos mal».

En enero de 1989, el Parlamento Europeo modificó el texto de su proposición original sobre financiamiento del PGH; en el original, tal como fue propuesto por la Comisión Europea, se hablaba de «identificar a los individuos de alto riesgo» para «protegerlos de las enfermedades a las que son vulnerables genéticamente y cuando fuera apropiado, impedir la transmisión de las deficiencias genéticas a la generación siguiente».

Hoy en día, las pruebas científicas experimentales apuntan a un giro de 180° respecto a las ideas eugenésicas del siglo pasado, basadas en la existencia de una «raza superior» (por supuesto, siempre la raza a la que pertenece quien escribe las ideas eugenésicas). No existen individuos perfectos. Lo atestigua el hecho de que un 4% de los nacimientos de niños vivos poseen algún tipo de defecto genético. Dado que la gran mayoría son defectos recesivos, podemos concluir que todos nosotros acumulamos en nuestra dotación genética alelos defectuosos, que en un 96% de los nacimientos no llegan a manifestarse, pero que permanecen latentes y podemos transmitir a las generaciones futuras. La mejor forma de disminuir la probabilidad de incidencia de una enfermedad recesiva (aparte de los sondeos genéticos de portadores y el consejo genético), es la exogamia y el mestizaje. La experiencia demuestra que todos los grupos étnicos o sociales que practican la endogamia, no importa lo elevadas que fueran sus características originales, finalmente acaban degenerando y llegan a poseer tasas realmente altas de determinadas enfermedades. El estudio de la genética humana, permite la confirmación empírica del viejo refrán: «en la variedad está el placer» y la actual tendencia a la globalización de la cultura y de la sociedad permite contemplar nuevos y despejados horizontes para el futuro.

La UNESCO, con la Declaración Universal sobre el Genoma y Derechos Humanos de 1997, declaró al genoma, en sentido simbólico, *Patrimonio de la Humanidad*. El año pasado, tras el anuncio de la compañía de secuenciación Celera, Bill Clinton y Tony Blair apoyaron que los datos del genoma humano fueran públicos. Sin embargo, no fueron las empresas privadas, sino los Institutos Nacionales de la Salud, de Estados Unidos, los primeros en plantear la posibilidad de patentar la secuencia *en crudo* de los genes humanos, que sus científicos iban describiendo a principios de los años 90. Esta intención provocó, como ya lo dijimos, la dimisión del primer director del proyecto y premio Nóbel James Watson.

El ser humano tiene sólo el doble de genes que la *Drosophila*, un tercio más que el gusano nematodo *Chaenorabditis* y apenas 5.000 genes más que la planta *Arabidopsis*. El ADN humano es al menos en un 98% idéntico al de los chimpancés y otros primates, sólo un 2 a 5 % del genoma codifica proteínas y el 25% del genoma humano está casi desierto, existiendo largos espacios libres entre un gen y otro. Si entendemos que no es lícito discriminar a un ser humano en función de su raza, sexo o religión, tampoco lo es discriminarlo en función de sus genes, especialmente cuando sabemos que la información que de ellos se deriva puede dar una idea, una probabilidad, de que puede padecer una determinada enfermedad, pero no implica que necesariamente vaya a sufrirla. Por tanto, hacemos énfasis que cualquier práctica eugenésica o discriminatoria a nivel social, en función de los genes de una persona, no está en absoluto justificada. La dignidad del ser humano es inviolable por el hecho de ser persona y este derecho no puede menoscabarse por su perfil genético. Además, el perfil genético de cada individuo pertenece a dicho individuo, y es por ello *confidencial*. Enfatizar esta confidencialidad es importante, puesto que el acceso a esta información, por parte de compañías de seguros médicos, puede repercutir negativamente sobre el asegurado.

Hay algo sobre la especie humana que el conocimiento actual del genoma ha venido a ratificar. Algo que, lamentablemente, a lo largo de la historia de la Humanidad, ha producido un innumerable número de muertes y vejaciones, como la esclavitud, el nazismo y el racismo entre otros. En el inicio del tercer milenio de nuestra cultura occidental, todavía en muchos Países, lamentablemente, continuamos padeciéndolos. Sin embargo es claro actualmente que a la luz de los resultados del Proyecto Genoma Humano todos los seres de la especie humana nos parecemos en un 99.9 % en nuestro genoma. Las diferencias entre las personas son, en este sentido, superficiales. Existe la misma variación genética entre un africano y un caucásico que entre un indígena o un oriental, es decir, sólo un 0.1 %. Es con base en este 0.1% de diferencia del genoma, que la especie humana ha llegado a cometer las mas terribles atrocidades y ha permitido las peores violaciones a la igualdad social, étnica y cultural.

# GLOSARIO

## A

**Acido nucléico:** Nombre genérico que se aplica indistintamente al ADN o al ARN las dos moléculas informacionales de los seres vivos.

**Adenina:** Base púrica encontrada en ADN y ARN. En secuencias duplex de ADN la adenina se aparea con la timina. En secuencias duplex ADN-ARN la adenina (en ADN) se aparea con el uracilo (en ARN)

**ADN:** (Acido desoxirribonucleico). Polímero formado por la unión covalente de nucleótidos. Un nucleótido = base (Adenina, Timina, Guanina, Citosina) + azúcar (2' deoxiribosa) + fosfato. Las moléculas de ADN eucariontes y procariontes son de doble hebra antiparalelas que presentan en promedio una conformación de tipo B, definida por Watson y Crick al analizar los parámetros de hélice obtenidos por difracción de rayos X. Ambas hebras están unidas por enlaces de tipo no covalente: efecto hidrofóbico y puentes de Hidrógeno.

**ADN de relleno:** (“Junk DNA”). Ácido desoxirribonucleico altamente repetido muy abundante en el genoma de organismos eucarióticos superiores y que no codifica por ningún gen

**ADNc:** Es el ADN complementario que se obtiene *in vitro* mediante la enzima transcriptasa reversa que utiliza como molde un ARNm maduro. Por lo tanto no contienen intrones.

**ADN genómico:** Ácido desoxirribonucleico total del núcleo de un organismo

**ADN mitocondrial:** Ácido desoxirribonucleico contenido en las mitocondrias

**ADN recombinante:** Molécula híbrida de ADN formada por la unión covalente de secuencias de ADN de diferentes orígenes que se introduce en un organismo huésped.

**ADN Ligasa (Polinucleótido Ligasa):** Es una enzima que une covalentemente extremos 3' con extremos 5' de cadenas polinucleotídicas. Participa en replicación y reparación de ADN.

**ADN nativo:** ADN intacto, tal cual como se encuentra en el organismo.

**ADN Polimerasa:** Enzima que cataliza la síntesis de ADN. Requiere de ADN molde, de ARN iniciador y de los 4 dNTP (dATP, dGTP, dCTP, dTTP).

**Alelos:** Formas alternativas de un gen en un mismo locus. Por ejemplo 2 posibles alelos en el locus PAH (Fenilalanina hidroxilasa) son PAH el normal y PKU el alterado que produce la fenilcetonuria. El término de alelo fue acuñado por William Bateson; literalmente significa «forma alternativa».

**Aminoácido:** Monómero químico que al unirse covalentemente a otro mediante el enlace peptídico forma cadenas polipeptídicas de las proteínas. La secuencia de los aminoácidos en las cadenas polipeptídicas está determinada por el código genético.

**Anemia falciforme:** Alteración de la cadena beta de la hemoglobina, donde ha ocurrido una sustitución del aminoácido ácido glutámico por valina. En el gen de la beta globina ocurrió una mutación por transversión (cambio de AT por TA) y en el ARNm cambia A por U. La anemia falciforme se caracteriza por tener un efecto pleiotrópico, es decir, por afectar a una serie de características fenotípicas (Ej. alteraciones cardíacas, reumatismo, parálisis, dolor abdominal y otros).

**ARN: (ácido ribonucleico):** Molécula que transmite información genética. Además cumple con funciones estructurales y de acoplamiento en la maquinaria de traducción de la información.

**ARN mensajero:** (ARNm) Molécula de ARN, copia de un gene, que lleva la información desde el genoma hasta donde se realiza la traducción

**AUG:** Triplete de inicio de la traducción. En procariontes significa formilmetionina; en eucariontes es metionina. AUG interno en el mensaje significa metionina.

**Autosomas:** Cromosomas diferentes a los cromosomas sexuales. En seres humanos existen 22 pares de autosomas.

## B

**BAC:** (Cromosoma Artificial Bacteriano). Vector de clonación diseñado a partir de elementos de replicación y estabilización mitótica de bacterias. Se utiliza para insertar fragmentos de ADN de hasta 30 Kbp

**Biblioteca de genes:** Conjunto de clones recombinantes que contiene clonada la totalidad del genoma de un organismo.

**Biblioteca de ADNc:** Genoma clonado como ADNc en un conjunto de vectores. Así sólo se han clonado los genes que se expresan en ARNm.

**Biochip:** Uno de los primeros términos empleados para referirse a las micromatrices de material biológico. En la actualidad este término se emplea para referirse a los chips fabricados con la tecnología de posicionamiento electrónico de la empresa Nanogen. En ocasiones también se usa al hablar de los Genechips

**Bp:** Par de bases nitrogenadas complementarias de una molécula de ADN de doble cadena (1 Kbp = 1000 bp)

**Biocomputación:** Es la utilización del material biológico en sistemas computacionales, como por ejemplo las memorias basadas en proteínas o la computación con ADN.

**Bioinformática:** Disciplina científica dedicada a la investigación y desarrollo de herramientas útiles para llegar a entender el flujo de información desde los genes a las estructuras moleculares, a su función bioquímica, a su conducta biológica y, finalmente, a su influencia en las enfermedades y en la salud.

**Biología molecular:** Rama de la biología nacida a raíz de la identificación de la naturaleza química (molecular) del material genético. Hoy día, nos referimos a biología molecular cuando hablamos de estudios o técnicas centradas en los genes y sus productos inmediatos, las proteínas.

**Biosfera:** El conjunto total de los organismos que habitan el planeta tierra.

**Biotecnología:** Conjunto de técnicas biológicas desarrolladas y aplicadas a la investigación y desarrollo de productos.

## C

**Cariotipo:** Complemento cromosómico de una célula. Se realiza con cromosomas metafásicos, los que se ordenan en parejas de cromosomas homólogos de acuerdo bá-

sicamente a sus longitudes y posición del centrómero.

**Centrómero:** Región especializada del cromosoma eucarionte nuclear al cual las fibras del huso se unen durante la división celular. La ubicación del centrómero en el cromosoma determina si un cromosoma es telocéntrico (centrómero en extremo), acrocéntrico (centrómero cercano a extremo), submetacéntrico (centrómero cercano a posición media) ó metacéntrico (centrómero en posición media).

**Chips de ADN:** Matrices bidimensionales en las que se integran ADNc u oligonucleótidos con secuencias específicas, siendo utilizadas para analizar muestras de ADNc en cuanto a la presencia de variaciones de secuencia o patrones de expresión génica.

**Chips de proteínas:** Matrices bidimensionales en las que se integran polipéptidos utilizados para analizar las proteínas que son expresadas en un organismo completo o célula.

**Citosina (C):** Base pirimidínica que se encuentra en AND y ARN. En secuencias de doble hebra se une mediante tres enlaces por puente de hidrógeno con G.

**Clonación de ADN:** Proceso de creación de múltiples copias de una única molécula de ADN.

**Clonación de Genes:** Involucra la modificación del genoma de una célula(s) por incorporación de un gen de interés. Las

etapas comprenden la obtención del gen de interés, unión a vector de clonación, introducción a célula(s) huésped y selección de células huésped recombinantes.

**Clon:** Conjunto de individuos genéticamente idénticos derivados de una célula original mediante métodos asexuales ó parasexuales. Ej. Colonias de bacterias

**Código genético:** Es la correspondencia entre tripletes o codones en el ADN (ó ADN) y aminoácidos en las proteínas. Incluye a 64 tripletes, de los cuales 61 de ellos tienen sentido al codificar para alguno de los 20 aminoácidos conocidos. Tres tripletes no codifican para aminoácidos (UAA, UAG y UGA) y se conocen como tripletes de parada. El código genético se caracteriza por ser degenerado (un mismo aminoácido puede tener más de un código), no ambiguo (un triplete siempre tiene el mismo significado) y universal (con ciertas excepciones en genomas de mycoplasmas, ciliados y mitocondrias). El código tiene una polaridad de 5' a 3'. Así, por ejemplo el triplete GUU significa valina si se lee de izquierda a derecha; en cambio significa leucina si se lee de derecha a izquierda.

**Codón:** Secuencia de 3 nucleótidos (triplete) en la hebra codificadora del DNA ó en el ARNm que representa a un aminoácido específico en el código genético y se traduce en su aminoácido correspondiente en el proceso de traducción. También existen codones que no significan aminoácidos y funcionan como señales de término de la traducción.

**Contig:** Arreglo continuo de clones recombinantes de una porción de ADN cromosómico en la que no hay discontinuidades. La contig completa corresponde al total de las secuencias de nucleótidos de un cromosoma completo.

**Complejidad de Genoma:** Longitud total de diferentes secuencias de DNA contenidas en el genoma. Generalmente se expresa en miles de pares de bases (Kbp).

**Complementariedad de bases:** Afinidad química entre bases nitrogenadas y que ocurre mediante la formación de enlaces por puente de hidrógeno.

**Contenido GC:** Porcentaje total de pares de nucleótidos de Guanina y Citosina que tiene un genoma.

**Cromatina:** Complejo de ADN, proteínas histonas, proteínas no histonas y ARN presentes en el núcleo celular.

**Cromosoma:** Unidad genética constituida por una molécula de ADN. Su tamaño y número varían dependiendo de la especie de que se trate. Puede medir desde medio millón (en las bacterias más simples) hasta varios cientos de millones de pares de bases (en los organismos superiores). En los organismos eucariontes los cromosomas se condensan y hacen visibles en ciertos momentos del ciclo de reproducción celular. Los cromosomas eucariontes son lineales y poseen centrómero y generalmente se presentan en pares. El cromosoma procarionte es circular, no posee centrómero.

**Cromosomas artificiales:** Vectores de ADN que tienen elementos genéticos de los cromosomas naturales. Tienen un origen de replicación (Ori). Secuencias centroméricas (CEN) y Telómeros. Además poseen sitios únicos para la digestión con enzimas de restricción

**Cromosomas sexuales:** Cromosomas que son diferentes en los dos sexos y que están implicados en la determinación del sexo. Ej. cromosomas X e Y en humanos.

**Construido terapéutico:** Vector de entrega de genes en el que se ha clonado el alelo normal de un gene que está mutado y produce una proteína no funcional o anormal. Se utiliza en terapia génica para restaurar la función que está alterada en una enfermedad genética.

## D

**ddNTP: (didesoxirribonucleósido 5'-trifosfato)** Nucleótido modificado que en posición 3' de la desoxirribosa ha perdido el grupo OH. Se incorpora en la síntesis de ADN pero se inhibe el proceso de síntesis posterior o elongación.

**Denaturación de ácidos nucleicos:** Se refiere a la conversión de secuencias de DNA doble hebra o ARN doble hebra a secuencias de hebra simple. Generalmente ocurre por calentamiento.

**De novo:** Síntesis de producto a partir de la unión de todos los precursores. Ej: en transcripción la ARN polimerasa realiza síntesis

de novo. En replicación de ADN ninguna de las ADN polimerasas conocidas es capaz de síntesis de novo, se requiere de la síntesis de un «primer».

**Distancia genética:** Grado de semejanza entre dos moléculas de ADN o genes.

**dNTP: (Desoxirribonucleósido 5'-trifosfato):** Monómero del ADN formado por la unión covalente de una base (A,T,G,C) + azúcar (2' deoxiribosa) + 3 fosfatos.

**Desoxirribonucleótido :** Monómero del ADN formado por la unión covalente de una base (A,T,G,C) + azúcar (2' deoxiribosa) + fosfato

**Desoxirribosa :** Azúcar cíclico de 5 carbonos que se encuentra en los desoxirribonucleótidos del DNA.

## E

**Eco RI:** Enzima de restricción obtenida de *E. coli*. Reconoce la secuencia palíndrome GAATTC/CTTAAG y produce cortes de tipo escalonados (entre G y A).

**Enfermedad genética:** Alteración patológica que tiene su causa en la mutación de genes que codifican por proteínas

**Enlace fosfodiéster:** En ácidos nucleicos, es el enlace covalente que une a los grupos fosfatos de nucleótidos adyacentes, uniendo el C5' de una pentosa con el C3' de la otra pentosa. Los enlaces fosfodiesteres junto con



los azúcares forman el esqueleto azucar-fosfato de las moléculas de ácidos nucleicos.

**Enlace peptídico:** Enlace de tipo covalente que se forma entre el grupo carboxilo de un primer aminoácido con el grupo amino de un segundo aminoácido. Así, las cadenas polipeptídicas crecen con la polaridad de extremo amino libre hacia extremo carboxilo libre.

**EST:** (“Expressed Sequence Tag”). Secuencias de nucleótidos producidas a partir de ADNc que tienen un oligonucleótido etiquetador único. Fueron desarrolladas inicialmente por Craig Venter y se han convertido en un instrumento poderoso para analizar el transcriptoma de un organismo.

**Entropía:** El grado de desorden termodinámico de un sistema.

**Eugenesia:** Los métodos para mejorar la calidad de la raza humana, en especial mediante la reproducción selectiva. Actualmente es motivo de mucho debate por sus implicaciones negativas en la sociedad humana.

**Exón:** Secuencia de DNA que se transcribe y es parte de la molécula de proteína funcional.

**Enzima:** Proteína o ARN con actividad catalítica, es decir; que acelera una reacción química específica sin destruirse en el proceso. Las moléculas blanco de una enzima se denominan sustratos.

**Enzima de Restricción ó Endonucleasa de Restricción:** Enzima que reconoce una secuencia específica entre 4 a 6 pares de bases de ADN de doble cadena y corta un enlace fosfodiéster en cada una de las hebras. Generan segmentos discretos de ADN Según el tipo de enzima el corte puede generar moléculas de ADN con extremos cohesivos de tipo palíndrome («sticky ends») ó extremos romos («blunt ends»). Estas enzimas son esenciales para la ingeniería genética.

**Extremos cohesivos:** («sticky ends»). Terminales de cadena única protuberantes que se producen por la digestión de un ADN de doble cadena por ciertas endonucleasas de restricción.

## F

**Factores de transcripción:** Son proteínas específicas que requieren cada una de las RNA polimerasas eucariontes para el inicio de la transcripción. Cada polimerasa utiliza un conjunto específico de TFs.

**Familia de Genes:** Conjunto de genes cuyos exones están relacionados. Se forma por duplicación y variación de algún gen ancestral.

**Farmacogenómica:** Uso del conocimiento de la información genómica para descubrimiento y evaluación de nuevas dianas farmacológicas

**Fármaco:** Molécula que produce un efecto fisiológica que generalmente restaura la función normal de un sistema, órgano o tejido.

**Fragmentación exhaustiva del genoma («whole genome shotgunning»):** Técnica desarrollada por Craig Venter para clonar y secuenciar todo el genoma humano y el de la *Drosophila melanogaster*. Consiste en fragmentar un genoma, mediante enzimas de restricción, en fragmentos pequeños (aproximadamente 30.000 bp), agregarles secuencias etiquetadoras y clonarlos en un vector de ADN diseñado para tal fin puede ser un BAC o un YAC

**Gen:** Unidad física y funcional que ocupa una posición específica en el genoma. Para genomas eucariontes y procariontes corresponde a una secuencia de ADN de doble hebra, donde una de las hebras actúa como templado en el proceso de transcripción (síntesis de ARN) y la otra hebra se denomina codificadora. El producto transcripcional tendrá una secuencia complementaria a la hebra del ADN templado, y la misma secuencia de bases que la hebra codificadora, con la excepción que las timinas del ADN han sido reemplazadas por uracilo en el ARNA. Además el gen posee secuencias regulatorias que no se transcriben ó que se transcriben pero no están representadas en el producto final.

**Gen mutante:** Alteración de la unidad de información biológica que se hereda y que normalmente altera dicha función.

**Genes murinos:** Genes de los ratones (*Mus musculus*).

**Genes Ligados:** Se refiere a aquellos genes que están localizados en un mismo cromosoma. Ej. En genomas eucariontes to-

dos los genes en un mismo cromosoma están ligados. Los genes no ligados son aquellos localizados en diferentes cromosomas y segregan independientemente. Para genomas procarióticos que poseen un solo cromosoma todos sus genes están ligados.

**Gen Reportero:** Gen cuyo producto es fácilmente detectable. Ej. cloranfenicol transacetilasa.

**Genética:** Ciencia de la herencia y la variación; estudia los patrones de herencia de rasgos específicos.

**Genoma:** Contenido total de material genético de una célula u organismo. Incluye la material genético nuclear y citoplasmático. Por ejemplo, el tamaño del genoma haploide de una célula de cebada es  $5.3 \times 10^9$  pares de bases y el tamaño del genoma humano es  $3.1 \times 10^9$  pares de bases.

**Genómica:** Estudio de todo el ADN contenido en un organismo o una célula

**Genómica funcional:** Conjunto de metodologías para poder determinar la expresión *in toto* de los genes de un genoma.

**Genoterapia:** (Terapia génica). Corrección de defectos genéticos mediante la introducción del alelo normal que restablece la función biológica que está afectada.

**Guanina (G):** Base púrica presente en ADN y ARN. En secuencias duplex se aparea con citosina mediante tres enlaces por puente de H.

## H

**Hemofilia:** Característica ligada al cromosoma X en humanos que afecta la coagulación sanguínea.

**Heteroduplex (Híbrido):** Molécula de ácido nucleico de doble hebra, en la cual cada una de las hebras tiene diferente origen. Estas moléculas ocurren *in vivo* por ejemplo como intermediarios en procesos de recombinación génica.

**Hibridización *In situ*:** Técnica para la localización citológica de secuencias de DNA complementaria a otra secuencia determinada. La hibridización también se puede realizar en un filtro ó en solución.

**HUGO:** (“Human Genome Organization”): Organización del genoma humano

**Huella dactilar de ADN:** (“DNA fingerprinting”) Distribución especial de secuencias repetidas de ADN o VNTR de un genoma que no se repite en otros genomas

## I

**Ingeniería genética:** Conocida también como Tecnología del DNA Recombinante y como Clonación de genes (Ver clonación de genes).

***In silico:*** (de Silicio). Se refiere a la posibilidad de poder crear una serie de procesos biológicos con base en los microchips de silicio utilizados en los computadores actuales.

Se refiere a la potencial construcción de organismos híbridos de Carbono y Silicio

***In vivo:*** Significa que ocurre en el organismo vivo

**Intrón:** Clásicamente es un segmento de ADN dentro de un gen eucarionte, el que se transcribe pero es eliminado del transcrito primario mediante una modificación post transcripcional (splicing). Pero por efecto de un splicing alternativo un intrón determinado puede ser parte de la secuencia informacional del ARNm funcional. También los intrones se han descrito para algunos genes procariontes.

## K

**Kilobase (Kb):** Unidad de longitud de ácidos nucleicos correspondiente a 1000 nucleotidos. Se abrevia como kb para ácidos nucleicos de hebra simple y kbp (kilo par de bases) para ácidos nucleicos de doble hebra.

## M

**Mapa de ligamiento:** Representación ordenada de los genes en un cromosoma y se obtiene observando los porcentajes de recombinación. Para un genoma bien descrito el número de mapas = número de grupos de ligamiento = número de cromosomas.

**Mapa de restricción:** Mapa genético del ADN que muestra los sitios de reconocimiento de enzimas de restricción.

**Mapa Físico:** Representación ordenada de los genes en un cromosoma basado en unidades físicas (pares de bases de ADN) más que en recombinación.

**Mapa Genético:** Representación del orden y localización relativa de los genes

**Metaboloma:** Conjunto de los genes de un genoma que codifican por las enzimas del metabolismo.

**Micromatriz:** Se define como micromatriz a aquellas matrices de material biológico cuyo tamaño de los puntos es inferior a las 500  $\mu\text{m}$ . Son dispositivos miniaturizados en los que se integran decenas de miles de sondas de material biológico genético.

“**Microarrays**”: Ver Micromatriz.

**Microsatélite:** Ver SSR

**Mutación:** Cambio heredable en el material genético.

**Mutante:** organismo ó célula portadora de una mutación; alelo alterado con una función diferente al normal. Generalmente (no siempre) es recesivo.

**M13:** Bacteriófago que infecta a varias cepas de la bacteria *E. coli*

## N

**Nanobiología:** Desarrollo de dispositivos y procesos biológicos en escalas de tamaño submicroscópicas

**Nucleótido:** Unidad monomérica de ácidos nucleicos. En el ADN se encuentran los desoxiribonucleótidos, que están formados por la unión covalente de una base (Adenina, Guanina, Citosina, Timina) + azúcar (2' Desoxiribosa) + Fosfato. En el ARN se encuentran los ribonucleótidos, que están formados por la unión covalente de una base (Adenina, Guanina, Citosina, Uracilo) + azúcar (Ribosa) + Fosfato.

**Northern blot:** Técnica de análisis de ARN; las moléculas de ARN son separadas por tamaño (electroforesis) y transferidas a una membrana, para a continuación detectar secuencias específicas por medio de sondas marcadas de ADN complementario.

## O

**Oligonucleótido:** Secuencia lineal de nucleótidos (hasta 20).

**Oligonucleótido iniciador:** Se refiere a una secuencia corta (generalmente ARN) que se sintetiza a partir de la enzima primasa en la forma de primosoma, utilizando como molde cada una de las hebras del ADN que se autoduplicarán. El primer proporciona un extremo 3' OH libre para que las enzimas ADN polimerasas prosigan el proceso de síntesis de ADN. Finalmente los “primers” son eliminados y reemplazados por secuencias de ADN.

**Origen de replicación:** Secuencia específica de ADN necesaria para el inicio de la replicación tanto en procariotes como en eucariotes

## P

**Plásmido:** Molécula de ADN extracromosomal circular de las bacterias, que posee genes no esenciales y de acuerdo al tipo de plásmido puede replicarse de manera dependiente ó independiente de la replicación del cromosoma de la célula huésped.

**PCR:** (Reacción en Cadena de la Polimerasa). Es un método *in vitro* para la amplificación enzimática de secuencias específicas de ADN. Utiliza iniciadores y la ADN polimerasa Taq. La amplificación ocurre a través de ciclos de denaturación, anillamiento con primer y extensión con ADN pol. La secuencia de ADN se amplifica mas de un millón ( $10^6$ ) de veces.

**PGH:** (Proyecto Genoma Humano). Proyecto para obtener la secuencia completa de los 3 mil millones ( $3 \times 10^9$ ) de pares de bases del genoma humano y mapear todos sus genes, estimados en aproximadamente 30.000 a 40.000.

**PNA:** ("Peptide Nucleic Acid"). Análogo del ADN en el que el esqueleto no es azúcar y fosfato, sino entidades similares a péptidos o proteínas.

**Polilinker:** Secuencia de ADN obtenida *in vitro* que contiene múltiples sitios de reconocimiento para enzimas de restricción. Generalmente se incorporan a vectores de clonación.

**Polinucleótido:** Secuencia de 20 ó más nucleótidos unidos por enlaces fosfodiesteres 5' a 3'.

**Polipéptido:** Molécula formada por la unión covalente de aminoácidos. Este término se utiliza para indicar que la cadena polipeptídica está en su estructura primaria, que aún no se ha plegado para adquirir la configuración tridimensional funcional.

**Poli A:** Secuencia de 50-250 nucleótidos de adenina que se agregan como una modificación post transcripcional al extremo 3' de ARNm eucariontes mediante la enzima poli A polimerasa. La reacción se conoce como poliadenilación.

**Polimorfismo Cromosomal:** Secuencia de ADN que presenta dos ó más variantes en la población.

**Proteína:** Molécula compuesta por una (proteína monomérica) ó más cadenas polipeptídicas (proteína multimérica). Los aminoácidos dentro de la cadena polipeptídica están unidos por enlaces covalentes denominados enlaces peptídicos. Diferentes cadenas polipeptídicas están unidas principalmente por enlaces no covalentes.

**Proteoma:** Conjunto de las proteínas expresadas por un genoma

**Proteómica:** Técnicas desarrolladas para el estudio de las proteínas expresadas por un genoma

**pUC:** Plásmido utilizado como vector de clonación que fue desarrollado y patentado por la Universidad de California.

**Puente de Hidrógeno:** Atracción electrostática que estabiliza a las moléculas duplex de ácidos nucleicos. Ocurre entre un átomo de H enlazado a otro átomo fuertemente electronegativo (O ó N) con otro átomo que es electronegativo ó contiene un par de electrones sin compartir. Ejemplo: A se une con T mediante dos puentes de H. G se une con C mediante tres puentes de H.

**Purina:** Bases nitrogenadas. En ADN y ARN son adenina (A) y guanina (G). En secuencias de ácidos nucleicos de doble hebra una base púrica reconoce a una pirimidina, de manera que la púrica A reconoce a la pirimidina T (en DNA) y U (en ARN). La púrica G reconoce a la pirimidina C.

## R

**Ratón «knock-out»:** Ratón al que se le han silenciado algunos de sus propios genes. En algunos casos estos son reemplazados por otros de otra especie.

**Restricción calórica (RC):** Efecto de la ingestión de una dieta con bajo contenido en carbohidratos y lípidos

**Retrotransposón:** Elemento de ADN genómico que produce una copia de él y se inserta covalentemente al genoma de una célula.

**Retrovirus:** Clase de virus con envoltura que contiene un genoma de ARN de hebra simple (+) protegido por una envoltura icosaédrica. Todos los viriones contienen la enzima transcriptasa inversa, necesaria para

la síntesis de una copia de ADN a partir del genoma viral. Estos virus están ampliamente distribuidos entre los vertebrados y se relacionan con varias enfermedades.

**Ribosomas:** Organelas ribonucleo-proteicas formadas por dos subunidades y encargadas junto con ARNt y ARNm de la biosíntesis de proteínas

## S

**Screening:** Tamizado o escogencia de genes.

**Secuenciación de ADN:** Determinación del orden de nucleótidos (bases) en una muestra pura de moléculas de ADN.

**Secuencia Altamente Repetida:** Secuencia de nucleótidos en ADN repetida entre  $10^5$  a  $10^7$  veces en el genoma.

**SNP:** (Polimorfismo de nucleótido único). Mutaciones puntuales producidas por la sustitución de un nucleótido por otro en la misma posición de un determinado gene o secuencia del genoma y que tienen una representación poblacional

**Sonda:** Molécula utilizada para identificar a otra molécula. Por ejemplo, para identificar un gen se utiliza como sonda el ARNm correspondiente ó el ADNc. En ambos casos la sonda debe estar marcada (ej. radioisótopo). En este caso la identificación ocurre por complementariedad de bases.

**SSR:** (Secuencia Simple Repetida). También conocido como microsatélite. Tipo de marcador molecular que involucra amplificación por PCR de secuencias de ADN seleccionadas repetidas en tandem.

**Southern Blot :** Técnica desarrollada por E.M. Southern y utilizada para analizar fragmentos de ADN, los cuales son separados mediante electroforesis, transferidos a un filtro e incubados con sonda específica.

**STS:** (“Sequence Tagged Site”): Secuencias de porciones de genoma a las cuales se les ha adicionado un nucleótido etiquetado único.

## T

**Talasemia:** Tipo de anemia relacionada con la ausencia de la cadena alfa ó cadena beta de ls hemoglobina. Depende de un par de genes que se expresan sin dominancia. TT= fenotipo normal Tt= talasemia menor tt=talasemia mayor

**Termociclador:** Instrumento que permite ejecutar en forma automatizada la técnica de PCR. Mediante la aplicación de ciclos térmicos secuenciales ocurre la denaturación, hibridación y síntesis de ADN.

**Traducción:** Proceso celular de decodificación de la secuencia de codones del ARNm en la secuencia, descifrada vía código genético, de la correspondiente proteína.

**Transcriptasas reversas:** Enzimas, generalmente presentes en los retrovirus. Son

ADN polimerasas que utilizan como moldes moléculas de ARNm.

**Transcrito primario:** Es el producto de ARN que se obtiene inmediatamente después de la transcripción. En eucariontes todos los transcritos primarios requieren de modificación post transcripcional para la obtención de moléculas de RNA funcionales. En procariontes los ARNm son inmediatamente funcionales, no requieren de modificación post transcripcional, de manera que es posible que se acoplen los procesos de transcripción con traducción.

**Transgénico:** Se refiere a organismos modificados genéticamente mediante la introducción de nuevas secuencias de ADN en la línea germinal.

**Transposon (Tn):** Elemento genético móvil que contiene genes para su inserción en el cromosoma y para su movilización a otras localizaciones

**Transcriptoma:** El conjunto de todos los ARNm que se sintetizan a partir de un genoma.

**Trastornos poligénicos:** Patologías cuyo mecanismo de transmisión genética de complejo y en los que intervienen muchos tipos de genes, principalmente con efectos muy pequeños sobre la condición patológica

## V

**Vector de clonación:** Molécula de ADN utilizada para introducir ADN exógeno en

células huésped. Se originan a partir de plásmidos bacterianos, de virus o de ADN de un organismo superior en el que otro fragmento de ADN puede ser integrado.

**Vector de Expresión:** Plásmidos ó fagos que poseen regiones promotores para expresar las secuencias clonadas de ADN.

**Vector «Shuttle»:** Vector de clonación capaz de replicarse en dos ó más organismos huésped. Ej. En *E. coli* y *S. cerevisiae*.

**VNTR:** (secuencias minisatélite). Porciones de genoma que contienen números variables de unidades de repetición conformadas por ciertas secuencias de nucleótidos

**Virus:** Organismos no celulares que poseen sólo una clase de ácido nucleico, ADN ó ARN, nunca ambos y sólo pueden reproducirse dentro de una célula huésped (procarionte, eucarionte vegetal, eucarionte animal). Poseen material genético en la forma de ADN ó ARN. En ambos casos este puede ser de hebra simple ó hebra doble.

## Y

**YAC:** Cromosoma Artificial de Levadura. Vector de clonación en la forma de un cromosoma artificial de levadura. Construido utilizando elementos cromosomales que incluyen telómeros (de ciliados), centrómeros, origen de replicación y genes marcadores de levadura. Se utilizan para clonar largas secuencias de ADN eucariótico.



**PÁGINA EN BLANCO  
EN LA EDICIÓN IMPRESA**

## BIBLIOGRAFÍA DE CONSULTA

- AACH John , BULYK Martha, CHURCH George, COMANDER Jason, DERTI Adnan y SHENDURE Jay. “Computational comparison of two draft sequences of the human genome”. *Nature*. Londres. 409, 2001.pp 856 – 859.
- AGUIRRE Raul, ARMANET Leonor, CASTILLO Silvia, CIFUENTES Lucia, CRUZ-COKE Ricardo, ETCHEBERRY Alfredo y LLOP Elena. IDENTIFICACIÓN GENÉTICA Y PRUEBAS DE PATERNIDAD. Editorial Universitaria. Santiago de Chile. 1996. 89-103.
- ARCO Pedro. BASES MOLECULARES DE LA CARDIOLOGÍA. Editorial Médica Panamericana. Madrid. 1996. 17-319.
- AYALA Francisco. “El genoma humano y las grandes fronteras de la biología en el siglo XXI”. En: Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.
- BALTIMORE David. “Our genome unveiled “. *Nature*. Londres. 409. 2001.pp 814 – 816.
- BARRERA Luis Alejandro. Hitos en la bioquímica y genética relacionados con el genoma humano. En: Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.
- BARRERA Luis Alejandro. Usos y abusos en el patentamiento de genes humanos. En: Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.
- BEEBEE Trevor y BURKE Julian. GENE STRUCTURE AND TRANSCRIPTION. Second Edition. Oxford University Press. New York. 1992. 17. 78.

- BLANCHARD Raymond K., MOORE J. Bernadette, GREEN Calvert L., COUSINS Robert J. "Modulation of intestinal gene expression by dietary zinc status: Effectiveness of cDNA arrays for expression profiling of a single nutrient deficiency". *PNAS*. Vol. 98 (24). New York 2001. pp 13507-13513.
- BRONOWSKI Jacob. EL ASCENSO DEL HOMBRE. Fondo Educativo Interamericano S.A. Bogotá. 1983. 379-410.
- CAO Shelley X., DHAHBI Joseph M., MOTE Patricia L., SPINDLER Stephen R. "Genomic profiling of short- and long-term caloric restriction effects in the liver of aging mice". *PNAS*. Vol. 98 (19). New York 2001. pp 10630-10635.
- CANTOR Charles. The Challenge to technology and informatics. . En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.
- CLARK William R. THE NEW HEALERS. Oxford University Press. New York. 1997.176-231.
- COLLINS A., LONJOU C., MORTON N.E. "Genetic epidemiology of single-nucleotide polymorphism". *PNAS*. Vol. 96 (26). Washington 1999. pp 15173-15177
- COX Timothy y SINCLAIR John. BIOLOGÍA MOLECULAR EN LA MEDICINA. Editorial Médica Panamericana. Buenos Aires.1998. 25-346.
- CUMMINGS Craig A., RELMAN David A. "Using DNA Microarrays to Study Host-Microbe Interactions". *Genomics*. Vol. 6 (5). Washington 2000. pp513-525.
- DAMASIO Antonio. DECARTES' ERROR. G. P Putnam's Sons. New York. 1994. 52-252.
- DAVIES Paul. EL QUINTO MILAGRO. Crítica. Grijalbo Mondadori S.A. Barcelona. 2000. 78-115.
- DAWKINS Richard. THE BLIND WATCHMAKER. W. W. Norton and Company Inc. New York. 1987. 21-255.
- DAWKINS Richard. EL GEN EGOISTA. Salvat. Barcelona. 1985. 29-67.
- DEACON Terréense. THE SYMBOLIC SPECIES. W. W. Norton and Company Inc. New York. 1997. 21-122.

- DENNIS Carina y SURRIDGE Christopher. “*A. thaliana* genome”. *Nature*. Londres. **408.**, 2000. pp 791.
- DIEHI Frank, GRAHIMANN Sussane, BEIER Markus, HOHEISEL Jörg D. “Manufacturing DNA microarrays of high spot homogeneity and reduced background signal”. *Nucleic Acids Research*. Vol. 29. Oxford 2001. pp 7-38.
- DRYSDALE Connie M., MAcGRAW Dennis W., STACK Catharine B., STEPHENS J. Clairbone, JUDSON Richard S., NANDALABAN Krishnan, ARNOLD Kevin, RUANO Gualberto, LIGGETT Sthepehn B. “Complex promoter and region b<sub>2</sub>-adrenergic receptor haplotypes alter receptor expression and predict *in vivo* responsiveness”. *PNAS*. Vol. 97 (19). Washington 2000. pp 10483-10488.
- FOX Keller Evelyn. Nature, Nurture and the human genome project. En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.
- FRASER Claire M., EISEN Jonathan, FLEISCHMANN Robert D., KETCHUM Karen A., PETERSON Scott. “Comparative Genomics and Understanding of Microbial Biology”. *Genomics*. Vol. 6 (5). Washington 2000. pp505-512.
- FRASER Andrew, KAMATH Ravi , ZIPPERLEN Peder , MARTINEZ-CAMPOS Maruxa, SOHRMANN Marc y AHRINGER Julie “Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference”. *Nature*. Londres 2000. 408.pp 325 – 330.
- FUTREAL Andrew, KASPRZYK Arek, BIRNEY Ewan, MULLIKIN James., WOOSTER Richard y STRATTON Michael. “Cancer and genomics”. *Nature*. Londres. **409**, 2001. pp 850 – 852
- GARCÍA-VALLEJO Felipe. ¿Cómo descifrar el genoma? En: Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.
- GARCÍA-VALLEJO Felipe y BARRERA Luis Alejandro. La terapia Génica. En: Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.
- GARCÍA-VALLEJO Felipe. “¿Cómo descifrar el genoma?”. *Innovación y Ciencia*. IX. (3 y 4). Bogotá 2001. pp 42-51.

- GARCÍA-VALLEJO Felipe, BARRERA Luis Alejandro. “La terapia génica”. *Innovación y Ciencia*. IX. (3 y 4). Bogotá 2001. pp 88-97.
- GARCÍA-VALLEJO Felipe. *BIOLOGÍA MOLECULAR Y BIOTECNOLOGÍA EN MEDICINA*. Editorial Catorse. Cali Colombia. 2000. 3-254.
- GARCÍA-VALLEJO Felipe. *BIOLOGÍA MOLECULAR. PRINCIPIOS Y APLICACIONES*. Centro Editorial de Ciencias. Universidad del Valle. Cali. 1997. 9-232.
- GATES Bill. *CAMINO AL FUTURO*. Segunda Edición. McGraw-Hill. Bogotá 1997. 1- 68.
- GELL-MANN Murria. *EL QUARK Y EL JAGUAR. AVENTURAS DE LO SIMPLE Y LO COMPLEJO*. Tusquets Editores S.A. Barcelona. 1998. 141-346.
- GILBERT Walter. A vision of the grail. En: D. Kevels y L. Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.
- GHOSH Amitav. *EL CROMOSOMA CALCUTA*. Grupo Editorial Norma. Bogotá. 1997. 11-183.
- GONZALEZ DE CANCINO Emilssen. Genes y derecho. En: Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.
- GOULD Stephen Jay. La amoralidad de la naturaleza. En: M Gardner (ed). *El Escarabajo Sagrado (I)*. Salvat. Barcelona. 1986. 37-49.
- GRACE Eric. *LA BIOTECNOLOGÍA AL DESNUDO: PROMESAS Y REALIDADES*. Editorial Anagrama. Barcelona. 1997. 22-120.
- GREELY Henry. Health insurance, employment discrimination and genetic revolution. En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.
- HANS-PETER Klenk, REBECCA A. Clayton, JEAN-FRANCOIS Tomb, OWEN White, KAREN E. Nelson, KETCHUM Karen A. , DODSON Robert J. , GWINN Michelle , HICKEY Erin K. , PETERSON Jeremy D. , RICHARDSON Delwood L. , KERLAVAGE Anthony R. , GRAHAM David E. , KYRPIDES Nikos C. , FLEISCHMANN Robert D. , QUACKENBUSH John , LEE Norman H. , SUTTON Granger G. , GILL Steven , KIRKNESS Ewen F. , DOUGHERTY Brian A. , MCKENNEY Keith , AD-

AMS Mark D. , LOFTUS Brendan , PETERSON Scott , REICH Claudia I. , MCNEIL Leslie K. , BADGER Jonathan H. , GLODEK Anna , ZHOU Lixin , OVERBEEK Ross , GOCCAYNE Jeannine D. , WEIDMAN Janife F. , MCDONALD Lisa , UTTERBACK Teresa , COTTON Matthew D. , SPRIGGS Tracy , ARTIACH Patricia , KAINÉ Brian P. , SYKES Sean M. , SADOW Paul W. , D'ANDREA Kurt P. , BOWMAN Cheryl , FUJII Claire , GARLAND Stacey A. , MASON Tanya M. , OLSEN Gary J. , FRASER Claire M. , SMITH Hamilton O. , WOESE Carl R. y VENTER Craig. “The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*”. *Nature*. 390. Londres 1997. pp 364 – 370.

HARTL Daniel. “Fly meets shotgun: shotgun wins” *Nature. Genetics*. 24 (4). Londres 2001. pp 327 – 328.

HATTORI Masahira y TAYLOR Todd.”The human genome: Part three in the book of genes “. *Nature*. 414. Londres 2001. pp 854 – 855.

HERNANDEZ Bronchud Miguel. LA INDUSTRIA DEL DNA. PRINCIPIOS BÁSICOS Y APLICACIONES. Editorial MCR, S.A. 1992. 31-200.

HOOD Leroy. Biology and medicine in the twenty-first century. . En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.

HUXLEY Thomas Henry. Ciencia y Cultura. En: M Gardner (ed). *El Escarabajo Sagrado* (I). Salvat.Barcelona. 1986. 147-163.

HUXLEY Aldous. La ciencia en el mundo feliz. En: M Gardner (ed). *El Escarabajo Sagrado* (II). Salvat.Barcelona. 1986. 304- 315.

INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM . “A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms”. *Nature* 409, 2001.pp 928 – 933.

INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. “Initial sequencing and analysis of the human genome”. *Nature. Londres*. 409. 2001.pp 860 – 921.

ISLAS Maria Cristina. Biología molecular y sociedad. En: A Panduro (ed). *Biología Molecular en la Clínica*. McGraw-Hill Interamericana. México 2000. 11-21.

- JACOB François. EL RATÓN, LA MOSCA Y EL HOMBRE. Crítica. Grijalbo Mondadori S.A. Barcelona. 1998. 15-186.
- JIMENEZ-SANCHEZ Gerardo, CHILDS Barton y VALLE David. "Human disease genes". *Nature*. 409. Londres 2001. pp 853 – 855.
- JONATHAN Hodgkin. Genome sequencing: "A view of Mount *Drosophila*". *Nature* 404. Londres 2000. pp 442 – 443.
- JOYANES Luis. CIBERSOCIEDAD. McGraw-Hill. Madrid. 1997. 3-52.
- JUDSON Horace . A history of the science and technology behind gene mapping and sequencing. En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.
- KELLAM Paul. "Host-pathogen studies in the post-genomic era". *Genome Biology*. 1 (2). Washington 2000. pp 1009.1-1009.4.
- KEVELS Daniel J. Out of Eugenics: The historical politics of the human genome. En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.
- KEYEUX Genoveva. Regulación mundial sobre el conocimiento y la intervención del genoma humano. En: Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.
- KUPER Adam. EL PRIMATE ELEGIDO. Crítica. Grijalbo Mondadori S.A. Barcelona. 1996. 13-109.
- LEWONTIN Richard. EL SUEÑO DEL GENOMA HUMANO Y OTRAS ILUSIONES. Ediciones Paidós Ibérica S.A. Barcelona. 2001. 125-160.
- MARRA Marco, HILLIER Ladeana, KUCABA Tamara, ALLEN Mellisa, BARSTEAD Robert, BECK Catherine, BLISTAIN Angela, BONALDO María, BOWERS Yvette, BOWLES Louise, CARDENAS Marco, CHAMBERLAIN Ann, CHAPPELL Julie, CLIFTON Sandra, FAVELLO Anthony, GEISEL Steve, GIBBONS Marilyn, NJATA Harvey, HILL Francesca, JACKSON Yolanda, KOHN Sophie, LENNON Greg, MARDIS Elaine, MARTIN John, MILA Leeanne, MCCANN Rhonda, MORALES Richard, PAPE Deana, PERSON Barry, PRANGE Christa, RITTER Erika, SOARES Marcelo, SCHURK Rebecca, SHIN Tanya, STEPTOE Michele, SWALLER Timothy, THEISING Brenda, UNDERWOOD Karen, WYLIE Todd, YOUNT Tamara, WILSON

Richard y WATERSTON Robert. "An encyclopedia of mouse genes". *Nature. Genetics* 409. Londres 2001. pp 817.

NELKIN Dorothy. The social power of genetic information. En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.

OPPENHAIMER James Robert. . La física en el mundo contemporaneo. En: M Gardner (ed). *El Escarabajo Sagrado (I)*. Salvat.Barcelona. 1986. 147-163.

PANDURO Arturo, BASTIDAS RAMIREZ Blanca Stela y NUÑO GONZÁLEZ Patricia. DNA recombinante, sondas y técnicas elementales de biología molecular. En: A. Panduro (ed). *Biología Molecular en la Clínica*. McGraw-Hill Interamericana. México.2000. 11-21.

PANDURO Arturo, RIVAS Ana María y SANCHEZ Laura. REACCIÓN EN CADENA DE LA POLIMERASA (PCR). En: A Panduro (ed). *Biología Molecular en la Clínica*. McGraw-Hill Interamericana. México 2000. 11-21.

RASKÓ Istvan y DOWNES Stephen. GENES IN MEDICINE. MOLECULAR BIOLOGY AND HUMAN GENETIC DISORDERS. Chapman & Hall. Londres. 1995. 11-384.

RIDLEY Matt. GENOMA. Grupo Santillana de Ediciones S.A. Madrid 2000. 9-379.

RUIZ-GARCÍA Manuel y ÁLVAREZ Diana. Evolución comparativa de genomas animales y su relación con el genoma humano. En : Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.

RUSSEL Bertrand. La ciencia para salvarnos de la ciencia. M Gardner (ed). *El Escarabajo Sagrado (II)*. Salvat. Barcelona. 1986. 441- 449.

SANCHEZ Laura, RIVAS Ana María y PANDURO Arturo. Secuenciación de los ácidos nucleicos. ). En: A Panduro (ed). *Biología Molecular en la Clínica*. McGraw-Hill Interamericana. México.2000. 11-21.

SANCHEZ TORRES Fernando y ESCOBAR LLANO Alfonso. Implicaciones éticas del proyecto genoma humano. En: Asociación Colombiana para el Avance de las Ciencias (ed.), *El genoma Humano*. Panamericana Editorial. Bogotá. 2002.



- SHWARTZ COWAN Ruth. Genetic technology and reproductive choice: ethics for autonomy. En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.
- SILVER Lee M. VUELTA AL EDEN. MAS ALLÁ DE LA CLONACIÓN EN UN MUNDO FELÍZ. Grupo Santillana de Editores S.A. Madrid. 1998. 29-92.
- SOLARI Alberto Juan. GENÉTICA HUMANA. FUNDAMENTOS Y APLICACIONES EN MEDICINA. Editorial Médica Panamericana. Buenos Aires. 1996. 31-237.
- TUPLER Rossella, PERINI Giovanni y GREEN Michael. “Expressing the human genome”. *Nature*. 409. Londres 2001. pp 832 – 833.
- VAN GINKEL Frederik W., NGUYEN Huan H., MAGHEE Jerry R. “Vaccines for Mucosal Immunity to Combat Emerging Infectious Diseases”. *Synopses*. Vol. 6 (2). Washington 2000. pp 123-132.
- WALTERS Leroy y GAGE PALMER Julie. THE ETHICS OF HUMAN GENE THERAPY. Oxford University Press. New York. 1997. 3-143
- WATSON James Dewey. A personal view of the project. . En: D. Kevels y L Hood. (eds). *The Code of Codes*. Harvard University Press. Cambridge. 1992.
- WATSON James y CRICK Francis. “A structure for Deoxyribose Nucleic Acid”. *Nature*. Londres .171. 1953. pp737.
- WELLS Herbert George. La ciencia y la verdad suprema. M Gardner (ed). El Escarabajo Sagrado (II). Salvat. Barcelona. 1986. 365-370.
- WILSON Edward. LA DIVERSIDAD DE LA VIDA. Crítica, Grijalbo Comercial S.A. Barcelona. 1994. 83-99.
- ZWEIGER Harold. EL GENOMA. INFORMACIÓN, ANARQUÍA Y REVOLUCIÓN EN LAS CIENCIAS BIOMÉDICAS. McGraw-Hill. México. 2002. 29-304.

# MOTORES DE BÚSQUEDA DE PÁGINAS EN LA WEB

Se pueden emplear varios motores de búsqueda tales como

<http://www.altavista.com>

<http://www.infoseek.com>

<http://www.google.com>

Alimentándolos con palabras claves.

## PAGINAS DE CONSULTA GENERAL EN LA WEB DEL INTERNET

- [http:// www.hfbr.bnl.gov/dna.html](http://www.hfbr.bnl.gov/dna.html) (Contiene excelente material sobre la estructura y función del DNA)
- <http://ejb.ucv.cl/dna.asp> (Tiene un buen recurso sobre estructura del DNA)
- <http://www.tokyo-med.ac.jp/genet>. (Incluye material visual sobre la estructura y función de las moléculas de DNA y RNA)
- [http:// www.cbc.med.umn.edu](http://www.cbc.med.umn.edu)
- [http:// www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)
- [http:// stanford.edu/~waterman/filo.html](http://stanford.edu/~waterman/filo.html)
- [http:// www. gene.com:80/ae/wn/nm/imurphy.EMS.html](http://www.gene.com:80/ae/wn/nm/imurphy.EMS.html).
- <http://www.rcsb.org/pdb>.
- [http://www. embl. bioimage. org](http://www.embl.bioimage.org).
- <http://www.ncbi.nlm.nih.gov>
- <http://life.aun.edu.au>.
- <http://esg-www.mit.edu:8001/esgbio/lander>
- <http://www.accessexcellence.org/AB/IE>.
- [http://www.shamockbay.com /-jkimball/BiologyPages](http://www.shamockbay.com/~jkimball/BiologyPages)
- <http://www.mssm.edu/molbio>

### **FELIPE GARCÍA VALLEJO**

(PHD) Jefe y Profesor Titular del Departamento de Ciencias Fisiológicas y Director Científico del Laboratorio de Biología Molecular y Patogénesis de la Facultad de Salud de la Universidad del Valle. Ha sido investigador visitante del Centro de Enfermedades Crónicas Virales de la Facultad de Medicina de la Universidad de Kagoshima, Japón, Centro Internacional de Ingeniería Genética y Biotecnología (IGEB) de las Naciones Unidas en Nueva · Delhi y Tri este. De 1992 a 1995 fue investigador posdoctoral en el Departamento de Inmunología y Enfermedades Infecciosas de la Escuela de Salud Pública de la , Universidad de Harvard e investigador asociado del Harvard AIDS Institute. Durante este tiempo llevó a cabo investigaciones sobre la biología molecular y la epidemiología del Virus Linfotrópico Humano Tipo I (HTLV-1)

### **MARTHA C. DOMÍNGUEZ**

Es, desde el año de 1990, investigadora asociada al laboratorio de Biología Molecular y Patogénesis de la Facultad de Salud de la Universidad del Valle. Ha realizado varios cursos de entrenamiento científico de Posgrado: en el Instituto de Biociencias, Universidad de Sao Paulo, Brasil 1982; en el Instituto de Química, Universidad de Sao Paulo, Brasil; en el Instituto Oceanográfico, Universidad de Sao Paulo, Brasil; con la UNESCO/UNIDO/ONU, Universidad del Valle; Instituto Ludwin, Sao Paulo, Brasil. Entrenamiento en el Departamento de Biología del Cáncer, Escuela de Salud Pública de la Universidad de Harvard, Boston, USA.



## Programa Editorial




Ciudad Universitaria, Meléndez  
Cali, Colombia

Teléfonos: (+57) 2 321 2227

321 2100 ext. 7687

<http://programaeditorial.univalle.edu.co>  
[programa.editorial@correounivalle.edu.co](mailto:programa.editorial@correounivalle.edu.co)

**i S i g u e n o s !**

   [programaeditorialunivalle](#)