

Desarrollo histórico de la genética humana

De los “factores” de Mendel al desciframiento del genoma humano

La Genética es la ciencia que estudia los fenómenos de la *herencia* y la *variación*. Estos fenómenos son complejos, y su análisis experimental sólo fue fructífero a partir del momento en que se contó con un marco conceptual adecuado, que fue provisto por el monje austriaco Juan Gregorio Mendel (1822-1884), aunque sus concepciones permanecieron sin uso hasta su redescubrimiento en el año 1900.

La Genética Humana tardó también mucho tiempo en establecerse sobre bases sólidas; tanto es así que recién en 1956 se comprobó fehacientemente el número de cromosomas de la especie humana, que es 46.¹

Las grandes dificultades que presentaba la especie humana para realizar análisis genéticos (imposibilidad de realizar experimentos de cruzamiento u otros tipos de experimentación, relativamente escaso número de progenie, número de cromosomas relativamente alto, etc.) han sido totalmente superadas en la segunda mitad del siglo XX, para pasar a convertirse, en la actualidad, en una de las especies mejor estudiadas. El adelanto de la Genética Humana ha tomado un enorme impulso con la concreción del Proyecto Genoma Humano que desde el año 1990 se desarrolló en Estados Unidos con la cooperación de

muchos institutos de varios países, y cuya culminación, en el año 2003, marca un hito en esta disciplina. Este desarrollo explosivo de la Genética Humana ha tenido repercusiones en múltiples campos, desde el Derecho y la ciencia política hasta la Psicología, pero principalmente ha afectado (y seguirá haciéndolo) el campo de la Medicina. En todas esas áreas existen perspectivas significativas de beneficio, pero también hay posibilidades de confusiones y de perjuicios.

En la coyuntura actual es conveniente realizar una somera recopilación del desarrollo histórico de algunos conceptos de la Genética, antes de pasar a enumerar los principios básicos de la Genética Humana. Este resumen es necesariamente esquemático y resalta sólo algunos de los conceptos básicos, para al mismo tiempo ejemplificarlos con casos concretos e introducir términos genéticos usuales.

El desarrollo histórico de algunos conceptos de la Genética

El trabajo presentado por Mendel en la Sociedad de Ciencias Naturales en Brno (actual República Checa) en 1865 y publicado el año siguiente contiene los postulados teóricos de la Genética, deducidos por Mendel a partir de sus experiencias de hibridación con plantas (fig. 1-1).

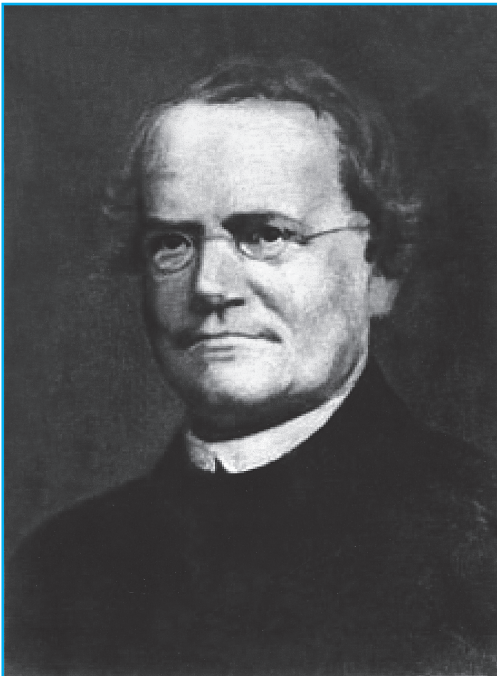


Fig. 1-1. Juan Gregorio Mendel, monje agustino y eminente biólogo, cerca de 1865. En ese año presentó su célebre trabajo sobre los híbridos, en la Sociedad de Ciencias Naturales de Brno (actualmente, República Checa).

En esa época aún no se conocía la meiosis, la fertilización era escasamente comprendida y la mitosis no había sido analizada aún en detalle, de modo que los postulados de Mendel no tenían una base celular y habían sido deducidos abstractamente de los resultados de sus experiencias. Por otra parte, Mendel usó una metodología estadística para establecer reglas cuantitativas para sus resultados de hibridación, efectuados en varios miles de plantas. Para explicar sus resultados, Mendel imaginó “factores” abstractos (décadas después estos factores serían llamados genes) que podían existir en *estados alternativos* (por ejemplo, un “factor” o gen para el color verde de la semilla, y un estado alternativo de ese factor, para el color amarillo). Actualmente, los estados alternativos o diferentes de un gen se denominan *alelos*, término introducido por el genetista W. Johannsen en 1909. En realidad, hoy sabemos que los alelos son todas las variantes que puede pre-

sentar un gen por mutación, pero en una simplificación podemos considerar que básicamente hay dos alelos para un factor: el normal o silvestre (generalmente simbolizado con un signo positivo) y el anormal o mutado. Mendel asumió que el origen de las variaciones radicaba en la existencia de “alelos” (normal y mutado, por ejemplo el color usual y otro color) y que los progenitores contribuían al descendiente con un alelo cada uno. Hoy sabemos que efectivamente nuestros organismos tienen en cada célula sus cromosomas por pares, es decir que nuestros 46 cromosomas son 23 pares, y que por consiguiente tenemos también nuestros genes por pares, condición que se llama diploidía (término introducido recién en 1905 por el citólogo alemán E. Strasburger). Por otra parte, hoy se sabe que efectivamente cada progenitor contribuye con *uno solo* de cada par de factores o genes, porque las células sexuales (o gametos) sólo poseen un juego de cromosomas en vez de un par de juegos (los gametos humanos tienen 23 cromosomas en vez de los 46 cromosomas de las demás células; son por consiguiente “haploides”, de *haplos* = mitad, en griego, término también introducido por Strasburger en 1905). Hasta aquí, los postulados puramente hipotéticos de Mendel estaban prediciendo los mecanismos aún no descubiertos de la fertilización y de la meiosis, y, al presumir la existencia de los “factores” (genes) que podían adoptar estados alternativos, predecía los estados de los genes, o normales o mutados (fig. 1-2).

Después de formular esos postulados, Mendel analizó la descendencia cuando los progenitores poseían “alelos” diferentes, por ejemplo color usual (+) y color mutante (m). Una de sus conclusiones capitales es que los alelos no se *fusionaban* en el descendiente y que, aunque un alelo m (“recesivo”) no tuviera un efecto evidente en ese descendiente, ese “factor” (gen) permanecía intacto en ese individuo y en la siguiente generación (llamada “filial 2” o F2) *podía aparecer de nuevo*. Además, cuando ese factor que no se expresó en la “filial 1” (F1 o primera descendencia) aparecía en un individuo de la F2, ese individuo era “puro”, *porque sus dos alelos eran iguales* (el individuo se llama *homocigótico* cuando tiene esa condición). Esa separación de los alelos, el silvestre + y el mutante m que

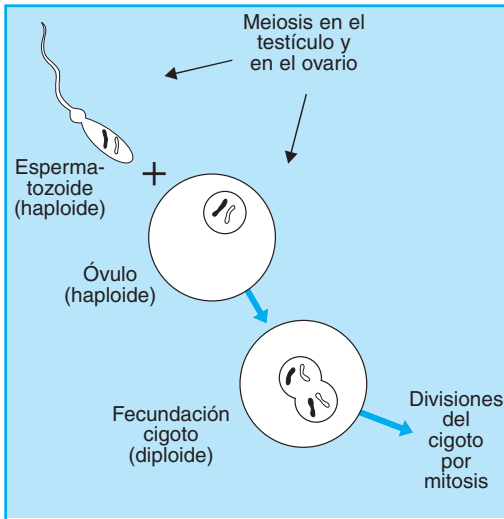


Fig. 1-2. Postulados predichos por Mendel que se comprobaron más tarde por observaciones citológicas. La meiosis (desconocida en tiempo de Mendel) reduce a la mitad el número de cromosomas en los gametos, espermatozoide y óvulo, que son haploides. Al producirse la fecundación, en el cigoto se reconstituye el número de pares de cromosomas, 23 pares en la especie humana (aquí se representan sólo dos pares de cromosomas).

están juntos en la F1 y se separan en la F2, es la *segregación* de los factores o genes, que tiene su base material en la separación de los dos miembros de cada par de cromosomas que ocurre en la meiosis (no descubierta aún en esa época).

Estas consecuencias, y otras que también dedujo Mendel, marcaron los caminos de la Genética en el siglo siguiente. En este libro de Genética Humana, daremos un ejemplo solamente aproximado a las experiencias de Mendel con plantas, para cuyo fin usaremos el árbol genealógico de la familia real inglesa del siglo XIX, en la cual se dieron una serie de casos de hemofilia, enfermedad grave con defecto de la coagulación de la sangre. Este ejemplo se aproxima a esas experiencias, por la (eugénicamente mala) costumbre de los miembros de las familias reales de casarse entre ellos, con un número no muy escaso de hijos, y que reciben cuidados tales que ciertos enfermos pueden sobrevivir hasta la edad reproductiva; seguramente Mendel hubiera apreciado estas observaciones (fig. 1-3).

La reina Victoria tuvo nueve hijos, cuatro varones, de los cuales uno solo fue enfermo de

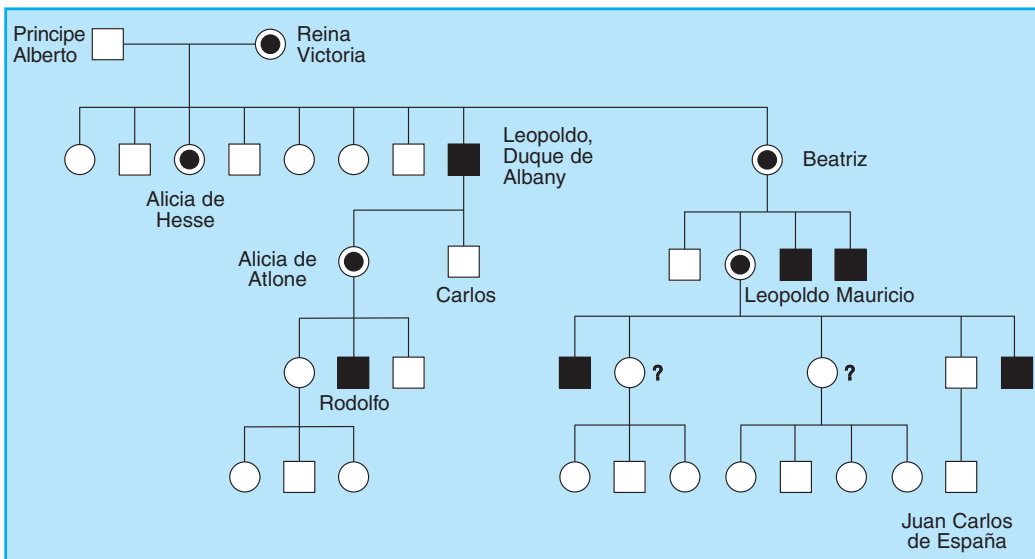


Fig. 1-3. Genealogía (parcial) de la descendencia de la reina Victoria de Inglaterra (1819-1901). En esta descendencia es evidente la transmisión hereditaria de la hemofilia. Se señala en particular el hijo menor de la reina, Leopoldo (hemofílico) y sus descendientes, y su hermana Beatriz (portadora) con los suyos. (Símbolo para varones: cuadrado; para mujeres, círculo; cuadrado llenado, enfermo; círculo parcialmente llenado, portador; signo ? = estado de portadora o no portadora incierto).



Fig. 1-4. Leopoldo, duque de Albany, fue el único hijo varón de la reina Victoria que fue hemofílico y transmitió el gen mutado a su hija Alicia, quien fue portadora del gen mutado; Rodolfo, hijo de Alicia también fue hemofílico, lo que demuestra la transmisión del gen y su segregación mendeliana.

hemofilia, y cinco hijas, de las cuales dos fueron “portadoras” del factor de la hemofilia pero sin mostrar evidencias de la enfermedad.

De acuerdo con los postulados de Mendel, el factor para la hemofilia debe ser “recesivo” frente al alelo normal (además hoy sabemos que este gen es “ligado al sexo” porque está en el cromosoma sexual X; véase el capítulo respectivo). El único hijo varón enfermo, Leopoldo de Albany (fig. 1-4), a pesar de los cuidados recibidos, murió a los 30 años, pero previamente se casó con la princesa Helena, con la cual tuvo dos descendientes: Alicia de Athlone (1883-1981) y Carlos Eduardo de Coburgo (1884-1954), *que no evidenciaron ningún signo de hemofilia*. Sin embargo, Alicia de Athlone tuvo tres descendientes, de los cuales uno, Rodolfo, *resultó enfermo de hemofilia*. Es decir, el gen mutado causan-

te de hemofilia, *reapareció en la generación F2*, es decir se produjo la “segregación” de este gen mutado (recesivo) y su alelo normal (dominante), tal como lo predice Mendel. Por otra parte, el factor para la hemofilia se transmite ligado al sexo, es decir que son los varones quienes expresan la enfermedad porque, como tienen un solo cromosoma sexual X, cuando este cromosoma lleva el alelo mutado lo expresan; en cambio las mujeres, que poseen dos cromosomas X, son “portadoras” porque el segundo cromosoma X con el alelo normal “oculta” los efectos del gen mutado (el individuo que lleva dos alelos diferentes se llama *heterocigótico* para ese gen). Generalmente los varones afectados con hemofilia no llegan a tener descendencia, como se aprecia en la descendencia de Beatriz (véase fig. 1-3).

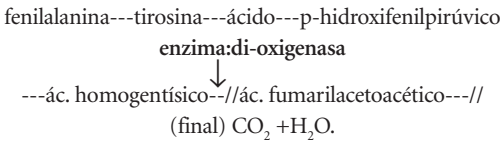
Varios miles de enfermedades humanas se deben a la mutación de un único gen; cuando no hay otros factores que perturben su expresión, esas enfermedades debidas a la mutación de un solo gen se comportan como los rasgos estudiados por Mendel, y son llamadas enfermedades *mendelianas* o monogénicas.

Archibald Garrod: el nacimiento de la genética bioquímica en Medicina

Poco después del redescubrimiento del trabajo de Mendel en el año 1900, el médico y profesor universitario Archibald Garrod realizó estudios sobre el modo de herencia de una peculiaridad que, en ciertas familias, originaba un cambio notable en el color de la orina, volviéndola negruzca, condición que se llama “alcaptonuria” (véase cap. 14) (fig. 1-5).

Garrod era un distinguido médico clínico pero además era versado en los recientes avances de la Bioquímica y también en los principios que se empezaban a perfilar de la Genética. En 1902 publicó sus resultados sobre alcaptonúricos, demostrando que en la orina de estos individuos había gran cantidad de ácido homogentísico, un producto del metabolismo de los aminoácidos (abreviado: aa) tirosina y fenilalanina, que normalmente no se encuentra en la orina. También

mostró que en las familias a las que pertenecían esos individuos, la alcaptonuria se heredaba como un “factor” mendeliano recesivo. Garrod supuso que la ausencia de ácido homogentísico en la orina normal se debe a que es degradado por una enzima, y que la falta de esa enzima es lo que provoca la alcaptonuria. Esta conclusión es exacta, como lo muestra el camino metabólico actualmente conocido que se expresa a continuación:



En ausencia de la enzima, se acumula ácido homogentísico, que se transforma espontáneamente en un polímero negrozco que oscurece la orina. Garrod propuso que la falta de la enzima (proteína) se debía a la mutación del gen normal para esa enzima. Garrod estudió otros rasgos con similares características, como la cistinuria y el albinismo, y llegó a la conclusión clarividente de que *cada una de esas enzimas se correspondía con un gen*. No sólo analizó las enfermedades hereditarias de enzimas; consideró que cada individuo debería tener un metabolismo diferente, por la gran cantidad de factores o genes involucrados y la gran variedad de cambios en cada factor o gen (alelos). Es también propia de Garrod la expresión “defectos congénitos del metabolismo” con la cual se engloba hoy a las enzimopatías genéticas.

Thomas Hunt Morgan: el ordenamiento lineal de los genes

El biólogo norteamericano Thomas H. Morgan (1866-1945) (fig. 1-6) se desvió de sus estudios iniciales de embriología para estudiar el ciclo vital y la herencia en la mosca de la fruta *Drosophila melanogaster*. Aprovechando las ventajas experimentales de esta mosca, demostró la existencia de la *recombinación* entre los genes, que ocurre en la profase de la meiosis, antes de la formación de gametos. Morgan y cols. demostraron que *los genes estaban colocados en un orden*

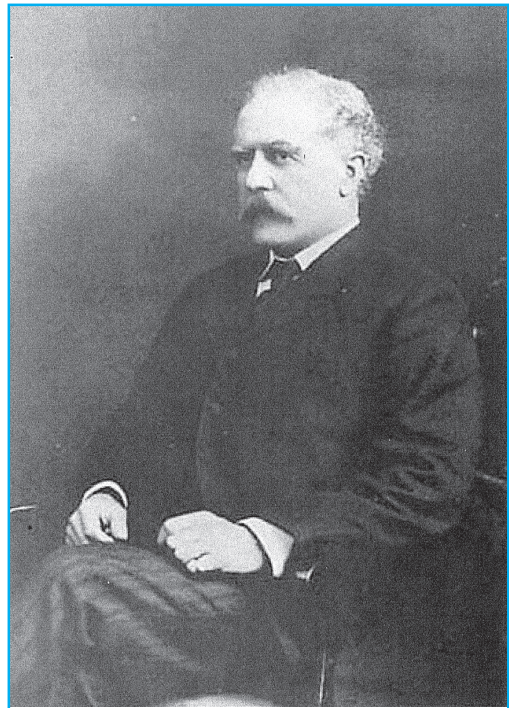


Fig. 1-5. Archibald Garrod (1858-1935). Médico inglés que propuso la relación entre genes y enzimas, y que fundamentó el estudio de los “errores congénitos del metabolismo”, como llamó Garrod a las mutaciones de genes para enzimas.

lineal, en cada cromosoma, aunque la naturaleza de esa línea no era conocida (hoy sabemos que esa línea se corresponde con la molécula bihelicoidal de ácido desoxirribonucleico (abreviado: ADN). La realización de los primeros *mapas de genes*, colocados en orden lineal y con una estimación de las distancias que los separan, tuvo gran repercusión y estableció la metodología para realizar estos *mapas de ligamiento* (ligamiento es la tendencia a permanecer juntos que tienen los genes que están localizados en el mismo cromosoma; véase el capítulo respectivo). Los trabajos de Morgan y cols. y de otros genetistas coetáneos demostraron además, en forma fehaciente, que los genes estaban localizados en los cromosomas, confirmando así la *teoría cromosómica de la herencia*, que da nacimiento a la Citogenética.



Fig. 1-6. Thomas H. Morgan, biólogo norteamericano que realizó los primeros mapas de ligamiento de genes y muchos otros adelantos, con el modelo de la mosca de la fruta.



Fig. 1-7. Rosalind Franklin (1920-1958). Doctora en Física, realizó el estudio cristalográfico que permitió demostrar la estructura molecular del ADN, que J. D. Watson y F. H. Crick usaron (sin su autorización) para proponer el modelo de doble hélice del ADN. Franklin falleció de cáncer de útero en 1958.

La conexión con el ácido desoxirribonucleico (ADN) y el modelo de la doble hélice

Hacia mediados del siglo XX la Genética estaba ingresando en una orientación hacia la Física y la Química. El gran interrogante acerca de la *base material de los genes* empujaba en esa dirección, y lo mismo hacían los avances más notables que ocurrían en la Genética de microorganismos y en especial de los virus. La conjunción del interés por la Genética despertado entre físicos y químicos más los resultados de la Genética microbiana llevaron en 1944 a la primera demostración de que la información hereditaria (es decir los genes) se encontraba en los ácidos nucleicos. Esta demostración fue realizada por el médico Oswald T. Avery junto a C. M.

MacLeod y Maclyn McCarty, en Estados Unidos, trabajando con bacterias (neumococos), comprobando que cierta información hereditaria de los neumococos residía en su ADN. La estructura molecular del ADN pasó entonces a ser un centro del interés de distinguidos físicos y químicos puros, tales como Linus Pauling, J. T. Randall y otros. El método más poderoso para investigar esa estructura molecular era el análisis por difracción de rayos X de cristales puros de esa sustancia en estado nativo. En el laboratorio dirigido por el físico J. T. Randall (King's College, en Londres) se realizaron durante 1952 y 1953 los más detallados y precisos estudios de difracción de rayos X de muestras de ADN, hechos por la joven física británica Rosalind Franklin (1920-1958) (fig. 1-7), a la sazón de 32 años de edad, quien trabajaba como investigadora asociada, en

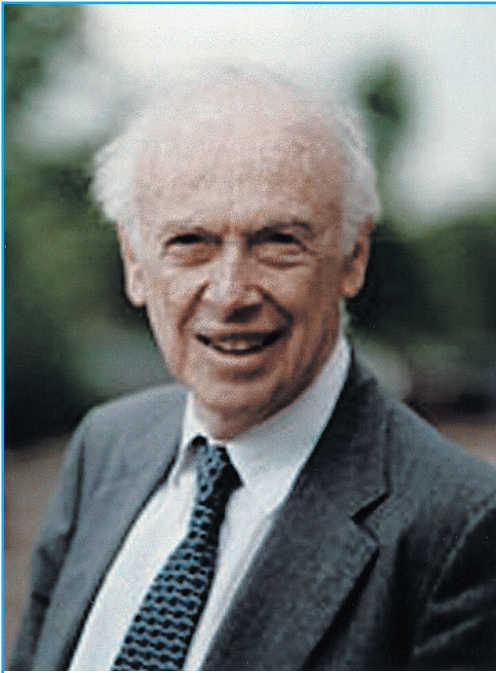


Fig. 1-8. James D. Watson, zoólogo norteamericano, que en colaboración con el físico inglés Francis Crick diseñó el modelo molecular de doble hélice del ADN en 1953, durante su estadía en Cambridge (Inglaterra).

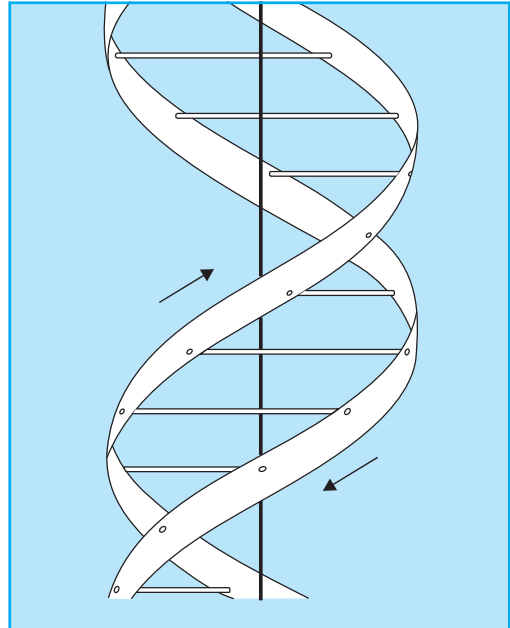


Fig. 1-9. Modelo original de la doble hélice del ADN, de Watson y Crick (Nature, 25 de abril de 1953). Cada hélice, representada aquí por una cinta, es una cadena de polinucleótidos. Las barras transversales representan pares de bases. La línea central representa el eje ideal sobre el cual las dos hélices se enrollan. Las flechas indican el sentido inverso de las dos hélices.

el mismo entorno en que lo hacía el físico neocelandés Maurice H. F. Wilkins, también interesado en el ADN.

Rosalind Franklin descubrió las formas A y B del ADN, demostró que las bases nitrogenadas estaban dirigidas hacia adentro y dio los detalles más precisos de las distancias moleculares en el ADN, pero no elaboró un esquema funcional de la molécula. Sus datos, incluidos sus cálculos y diagramas, fueron llevados por Wilkins a un par de investigadores, el biólogo norteamericano James D. Watson (nacido en 1928) (fig. 1-8) y el físico británico Francis H. C. Crick (1916-2004), quienes estaban en un laboratorio similar, en Cambridge, pero que estaban impedidos de realizar tareas experimentales y dedicaban su tiempo a proponer modelos moleculares teóricos.

Sobre la base de los datos de Franklin y las proporciones de las bases encontradas por el aus-

tríaco-norteamericano Erwin Chargaff (1905-2002), Watson y Crick elaboraron un modelo molecular del ADN, que consiste en dos hélices que corren con sentidos opuestos, con las bases de cada hélice dirigidas hacia el centro y guardando una relación espacial de estricta complementariedad (que justifica las relaciones halladas por Chargaff) y que sugiere de inmediato una forma simple de replicación de la molécula, por separación de las hélices y síntesis de las hélices complementarias (fig. 1-9).²

Confirmación y consecuencias del modelo de la doble hélice

El modelo de la doble hélice representó una forma intuitivamente fácil de representar las propiedades que debía tener el sustrato

material de la herencia: *autorreplicable* (por separación de las dos hélices), *lineal* (como lo es la doble hélice a lo largo de su eje), *localizado en los cromosomas* (como ya había sido demostrado por citólogos y bioquímicos) y capaz de contener la *información hereditaria*, lo cual era sugerido por una *secuencia de bases* con enorme cantidad de variaciones a lo largo de cada una de las hélices. Posteriormente a 1953 estas propiedades fueron definitivamente confirmadas por muchos experimentos. Particularmente la última de las citadas propiedades era central para la Genética y llevó al descubrimiento del llamado *código genético*, así como al mecanismo por el cual dicha información hereditaria depositada en el ADN era utilizada por las células, es decir el *flujo de la información genética*.

Las dimensiones de la molécula de ADN

Siendo la base material de la herencia y un componente de los cromosomas, la molécula de

ADN planteó algunos interrogantes, que fueron explicados durante el lapso de 1955-1975. En primer lugar, cuántas moléculas de ADN hay en un cromosoma; en segundo lugar, cuáles eran sus dimensiones, en especial su longitud; y en tercer lugar, cómo se encontraba dispuesto en los cromosomas. Cada cromosoma consta de dos *cromátides hermanas* iguales, y los datos de replicación del ADN mostraron que cada cromátide tiene *una única molécula* de ADN (de doble cadena). A su vez, la molécula de ADN desafió la noción química de que las moléculas son entidades submicroscópicas en todas sus dimensiones: la molécula única de cada cromosoma humano puede llegar (si estuviera totalmente desenrollada) a cerca de siete centímetros de longitud: se trata de moléculas de una longitud gigantesca a pesar de su anchura submicroscópica de sólo 23 nm (fig. 1-10). La molécula de ADN es flexible y puede curvarse, lo cual le permite adoptar sucesivos órdenes de curvatura, que a su vez permiten empaquetar estas moléculas larguísimas en los pequeños cromosomas.

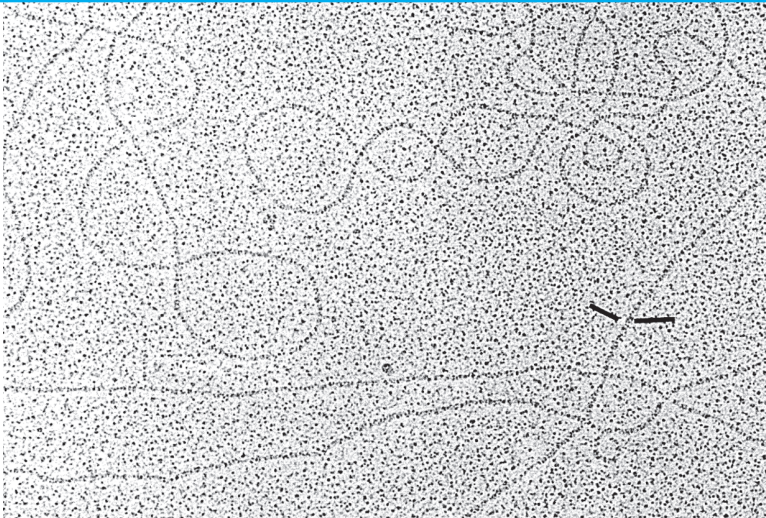


Fig. 1-10. Con microscopía electrónica se pudo demostrar que la molécula de ADN es muy larga, de una anchura constante (de 23 nm), señalada entre las dos líneas, y además que es flexible y esto le permite curvarse. Aumento: 52.000. (Solari AJ, Proc. Nat. Acad. Sci. USA, 1965; 53:503-511).

La codificación de la herencia

Cuando se demostró que la base material de la herencia era el ADN, surgió el problema de cómo estaba representada, en el ADN, esa determinación de la herencia. Dado que todas las células de los organismos vivos funcionan básicamente por sus proteínas, que forman no sólo los bloques estructurales con los que se organiza la célula, sino que también a través de proteínas enzimáticas las células movilizan las reacciones químicas del metabolismo, la asignación de la herencia a una sustancia poco reactiva como el ADN planteaba algunas dudas conceptuales que ya habían sido resueltas, y otras dudas nuevas. Desde hacía décadas se había demostrado que el “puente” entre una generación y la siguiente depende sólo de los gametos, y dentro de los gametos, de sus cromosomas. De modo que la base material de la herencia implica una miniaturización de esa base material, cualquiera sea su naturaleza. Por otra parte, se había demostrado que las unidades de la herencia, los genes, estaban colocados en orden lineal, que podía ser fraccionado por recombinación, y que además eran susceptibles de variar (por *mutación*). Siendo los componentes del ADN solamente de tres tipos, fosfato, desoxirribosa y bases nitrogenadas (éstas son cuatro), la única forma imaginable de acumular la información hereditaria en el ADN sería por el *orden secuencial de las bases nitrogenadas, adenina (A), guanina (G), citosina (C) y timina (T)*. El físico ruso-norteamericano George Gamow (1904-1968) propuso que las instrucciones hereditarias estaban *codificadas* en el ADN de modo que las variaciones de las secuencias de los cuatro tipos de bases (con repeticiones) determinaban el orden, a su vez, de los aminoácidos en cada proteína. Aunque el esquema de Gamow no era correcto en detalle, la idea estimuló a muchos otros científicos, entre ellos Francis Crick y el británico Sydney Brenner, hasta que la idea de un “código” de tres bases seguidas, es decir un “tripleto” se afirmó como el representante de cada uno de los 20 aminoácidos. Finalmente, en 1966 se determinó experimentalmente el “código genético”, por parte de un conjunto de estudios en los cuales intervinieron el bioquímico nortea-

mericano Marshall Nirenberg (1927-), el bioquímico hindú-americano Har G. Khorana (1922-) y el bioquímico español-norteamericano Severo Ochoa (1905-1993). El paso de la información desde el ADN hasta las proteínas era intermediado por otra sustancia, el ácido ribonucleico (ARN) del tipo que se llamó “mensajero” (ARNm).

Consecuencias del desciframiento del código genético

A partir de la confirmación experimental del código genético, gran parte de la Genética inició un viraje, desde los estudios estructurales hacia los estudios de descodificación, que se extiende hasta la actualidad. El objetivo, al parecer inalcanzable en 1960, de determinar la identificación y el orden de todas las bases nitrogenadas a lo largo del ADN (es decir la *secuenciación* del ADN), se fue volviendo factible gracias a adelantos técnicos. Los resultados por lograr se planificaron en gran escala bajo la forma del llamado *Proyecto Genoma Humano*, cuya finalización se realiza en el año 2003.

Ya finalizada la secuenciación del ADN humano, se abre una etapa diferente, de combinación de estudios estructurales (de *proteínas*, actualmente) y sobre todo de estudios de interacciones entre ácidos nucleicos y proteínas, y de proteínas entre sí, con el objetivo de dilucidar los *procesos de regulación génica* y de *función de las proteínas* dentro del organismo, que tiene enormes repercusiones en Medicina.

El Proyecto Genoma Humano de secuenciación y análisis del ADN humano

Durante la década de 1980-1990 un cierto número de instituciones académicas, organismos de política científica e investigadores individuales de Estados Unidos llegaron a la conclusión de que el desciframiento completo del ADN humano, con su conjunto de genes (o *genoma*) era un proyecto tecnológicamente factible, si se asignaban los considerables recursos económicos y académicos necesarios. La decisión finalmente

fue tomada por las autoridades respectivas de Estados Unidos, asignando fondos crecientes a dos instituciones gubernamentales, el Instituto Nacional de la Salud (NIH) y el Departamento de Energía (DOE) de Estados Unidos, así como a numerosas universidades de ese país, con el objetivo declarado de *descifrar la secuencia total de bases del ADN humano*. Este proyecto, que comenzó oficialmente en 1990, tenía metas precisas para cada quinquenio, que fueron ampliamente superadas en la realidad. En el año 2001, este proyecto llevaba consumidos alrededor de 3.000 millones de dólares desde el año 1990, sin por ello afectar la asignación de recursos a otros proyectos científicos, y había llegado a lograr la secuencia de casi todo el genoma (o ADN) humano. Paralelamente, una empresa privada llamada Celera y presidida por el investigador Craig Venter había logrado en el año 2000 una secuenciación casi total del ADN humano, por lo cual el entonces presidente de los Estados Unidos, Clinton, con la presencia conjunta de Venter y del director del proyecto gubernamental, Francis Collins, presentó en público la descodificación del ADN humano “en borrador”, el 26 de junio del año 2000. También se anunció que las secuencias “definitivas” serían publicadas en 2003 (la diferencia entre “secuencias en borrador” y secuencias definitivas o finales es el grado de error que pueden contener, que en las últimas debe ser menor a una base errónea entre 10.000).

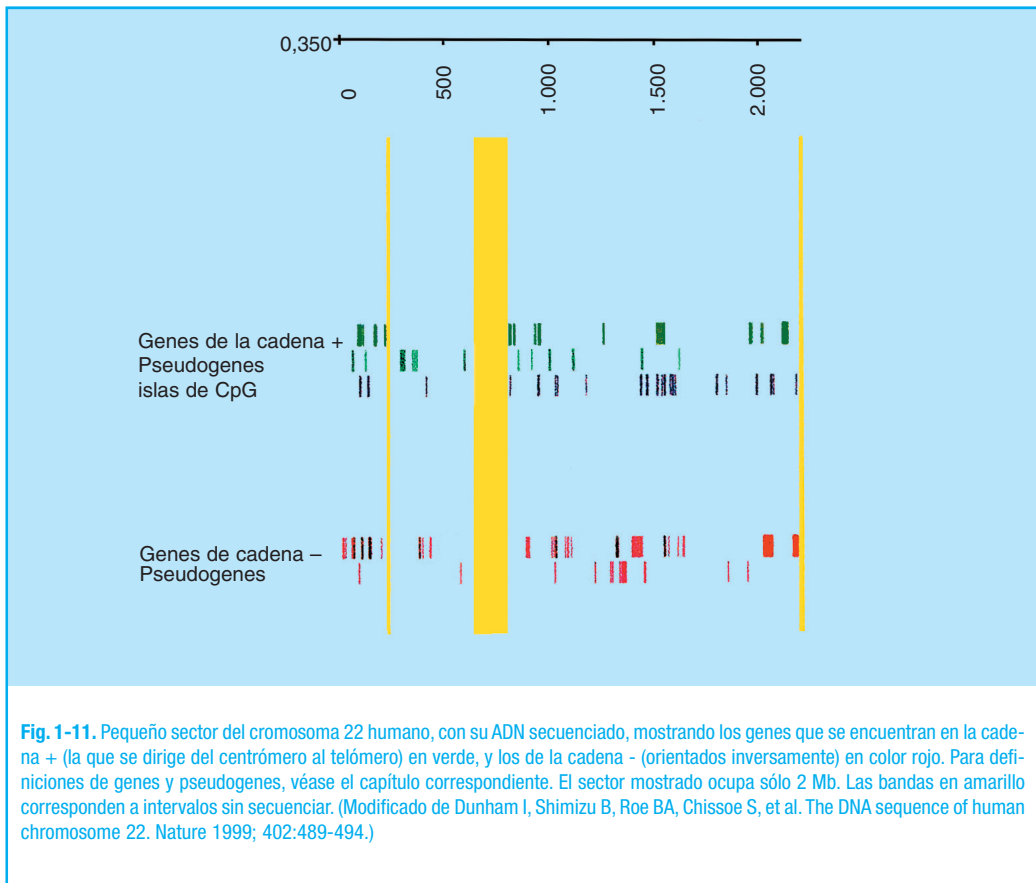
Consecuencias e implicancias de la secuenciación del ADN humano (Proyecto Genoma Humano)

La consecuencia inmediata del completamiento de la secuenciación del genoma humano es que se dispone de la secuencia de los 3.000 millones de bases nitrogenadas que constituyen en ADN humano. Esas tres mil Mb (abreviación de millón de bases) están distribuidas en los 23 pares de cromosomas humanos, proporcionalmente al tamaño de cada uno de éstos; de esa manera, el cromosoma 21, que es el más pequeño, tiene un ADN con aproximadamente 41 Mb, y el 22 tiene 45 de las cuales 33,4 Mb fueron secuenciadas³ (fig. 1-11).

La enorme cantidad de información acumulada en la preparación de la secuenciación y en sus resultados ha impulsado el desarrollo de métodos informáticos especializados para el manejo de esta información, y esto ha concluido en el establecimiento de una disciplina nueva, la *bioinformática*. A través de procedimientos de bioinformática se puede *acceder a bases de datos* y disponer inmediatamente de la secuencia de bases de cualquier sector del ADN de todos los cromosomas humanos. El Centro Nacional de Información Biotecnológica de Estados Unidos (NCBI; <http://www.ncbi.nlm.nih.gov/>) dispone de la base de datos más completa para estos fines.

Los resultados de la secuenciación en borrador fueron publicados simultáneamente por parte del proyecto oficial y el de compañía Celera en el año 2001.^{3,4} Uno de los objetivos de la secuenciación, que es obtener el número total de genes humanos, sólo pudo obtenerse en forma aproximada: alrededor de 25.000 genes. Esto se debe a que para demostrar fehacientemente que una secuencia es un gen es preciso saber si *funciona* como un gen, y demostrar el funcionamiento es más difícil que la secuenciación. Muchos genes importantes ya habían sido aislados individualmente, en laboratorios universitarios, antes de la conclusión del Proyecto Genoma Humano. Este proyecto confirmó casi totalmente las secuencias de esos genes conocidos, pero además se encontraron gran cantidad de genes de función totalmente desconocida, y muchos otros genes de los cuales se presupone con alguna probabilidad cuál es su función por ciertas características de la proteína que codifican.

Uno de los resultados más importantes de la secuenciación es que se dispone ahora de la codificación de varios miles de proteínas desconocidas y la secuencia codificada de muchas proteínas incompletamente conocidas. Se ha estimado que existirían alrededor de un millón de proteínas humanas, dado que cada gen codifica varias proteínas usando diferentes combinaciones de sus regiones codificantes, ya sea por empalme alternativo, por el uso de diferentes promotores, por modificaciones del transcripto o por modificaciones postraduccionales.⁵



Dado que el estudio de la estructura y función de cada proteína es sumamente complejo, se estima que el estudio de las propiedades fisiológicas y farmacológicas del millón de proteínas humanas (que colectivamente forman el llamado *proteoma humano*) constituye la tarea inmediata de las próximas décadas.⁶

Otro aspecto de la secuenciación del ADN es la obtención de *pruebas diagnósticas* de un número de enfermedades hereditarias, incluso para “portadores” que no presentan ninguna sintomatología. Estas pruebas diagnósticas, que ya empezaron a estar disponibles a medida que ciertos genes importantes en Medicina fueron aislados, han ido gradualmente incrementando en número, a medida que se conocieron los cambios de bases, pérdidas de bases u otros cambios

en la secuencia del ADN correspondiente a un gen que impiden su función normal (es decir “mutaciones”) y que están asociados a una determinada enfermedad. En el [cuadro 1-1](#) se enumeran algunas de estas pruebas genéticas.

Identificación de personas

La secuenciación total del ADN humano fue realizada con especímenes derivados de unas pocas personas; por ejemplo, el mapa génico de Celera (dirigida por Craig Venter) fue realizado con el ADN de sólo cinco personas.³ Sin embargo, uno de los propósitos del Proyecto Genoma Humano es también determinar las *variaciones* en la secuencia del ADN que se encuentran en distintos individuos, de diferentes poblaciones y

Cuadro 1-1. Algunas de las pruebas diagnósticas de ADN usadas actualmente para demostrar la presencia de mutaciones asociadas con la presencia real de una enfermedad o la portación de un gen mutado que eventualmente puede derivar en la enfermedad en el sujeto o en su descendencia.

Deficiencia de antitripsina-alfa 1 (enfermedad hepática y enfisema)

Esclerosis lateral amiotrófica (parálisis progresiva y eventualmente letal)

Enfermedad de Alzheimer

Ataxia-telangiectasia (enfermedad encefálica progresiva con pérdida del control muscular y desarrollo de tumores malignos)

Enfermedad de Gaucher (agrandamiento del hígado y del bazo, con anomalías de la médula ósea)

Cáncer de mama y de ovario, tipo hereditario

Cáncer de colon hereditario no poliposo (tumores de desarrollo temprano del colon y otros órganos)

Enfermedad de Charcot-Marie-Tooth (pérdida de sensibilidad en las regiones distales de los miembros)

Hiperplasia suprarrenal congénita (déficit hormonal que produce el desarrollo de genitales ambiguos y pseudohermafroditismo)

Enfermedad fibroquística o fibrosis quística (enfermedad de los pulmones y del páncreas, con acumulación de mucus espeso e infecciones crónicas)

Distrofias musculares de Duchenne y Becker (debilidad y atrofia muscular, grave o moderada)

Anemia de Fanconi (tipo C) (anemia y leucemia con deformidades óseas)

Factor V-Leyden (de coagulación)

Síndrome de X frágil (importante factor de retraso mental genético)

Hemofilia A; Hemofilia B (déficit de coagulación sanguínea)

Enfermedad de Huntington (enfermedad neurológica, de inicio a mediana edad y letal)

Distrofia miotónica (debilidad muscular progresiva)

Neurofibromatosis, tipo I (tumores múltiples benignos)

Fenilcetonuria (retraso mental por déficit enzimático; corregible por dieta)

Poliquistosis renal del adulto (enfermedad renal y hepática)

Síndromes de Prader-Willi y Angelman (déficit intelectual y motor)

Anemia falciforme (anemia que cursa con dolor e infecciones)

Ataxia espinocerebelosa, tipo 1 (movimientos involuntarios, alteraciones del lenguaje y de los reflejos)

Atrofia muscular espinal (atrofia muscular grave y progresiva)

Talasemias (anemias variadas por reducción de la síntesis de una hemoglobina)

Enfermedad de Tay-Sachs (neuropatía)

razas. Si bien el estudio de las variaciones no está completado, se demostró que hay numerosas variantes de la secuencia de bases, en especial en regiones no génicas, que son características de cada individuo y que se heredan como “caracteres” mendelianos. Esto hace que sea posible, con muestras de ADN de una persona y sus parien-

tes, establecer la filiación, la ascendencia y la identidad de las personas (para este fin también es usado el pequeño ADN de las mitocondrias, que se transmite exclusivamente por vía materna). Estos adelantos ya han sido introducidos en la práctica jurídica, aunque su validez es siempre determinada por los tribunales.

A su vez, la posibilidad de identificar a las personas por su ADN y la posibilidad adicional de determinar si es portador de genes mutados que eventualmente pueden determinar una enfermedad en el sujeto o en sus descendientes han originado una serie de problemas de índole bioética, legal y social. El Proyecto Genoma Humano de Estados Unidos (HUGO) contiene

un sector especial dedicado a analizar este tipo de problemas (llamado el *subprograma ELSI = ethical, legal, social issues*), convocando a reuniones de discusión multidisciplinaria, en las cuales intervienen juristas, teólogos, sociólogos, psicólogos y especialistas en bioética (esta última es una disciplina en rápido crecimiento) (panel 1-1).

Panel 1-1. Proteoma, transcriptoma y varioma

Después del completamiento del Proyecto Genoma Humano se realizaron las secuenciaciones de alrededor de dos centenares de otros organismos, que incluyen los genomas de mamíferos cercanos a la especie humana (el ratón, el chimpancé, el perro, y otros (véase <http://www.ncbi.nlm.nih.gov/sites/entrez>, GENOME PROJECTS). Ello acrecentó la posibilidad de comparar nuestro genoma con el de otros mamíferos y organismos inferiores, y detectar segmentos génicos conservados (es decir, poco variables en diferentes especies), lo cual es útil para distinguir secuencias esenciales para procesos celulares básicos.

Adicionalmente, se creó una disciplina en rápido crecimiento, la proteómica, cuyo objetivo es el estudio del conjunto completo de proteínas de cada organismo, es decir de los proteomas. Se estima que el conjunto de las proteínas humanas, es decir el proteoma, contiene aproximadamente un millón de proteínas.⁷ Esto significa que el proteoma es unas 40 veces mayor que el número de genes codificantes de proteína. Este hecho, a su vez, señala la importancia de las variantes producidas por empalme alternativo (5 a 7 son comunes), las modificaciones post-traduccionales y otros mecanismos de variación. El estudio del proteoma se lleva a cabo en centros especializados y con tecnologías fisicoquímicas, en especial la espectrometría de masa. La proteómica humana tiene también como objetivos caracterizar el conjunto de proteínas de cada tipo celular (p. ej., el hepatocito) de tal manera que existen numerosos subproyectos dirigidos hacia tipos celulares, órganos específicos o proteínas presentes en líquidos biológicos normales (plasma, líquido cefalorraquídeo, etc.). A su vez, otros subproyectos se dirigen a identificar las diferencias de composición proteica de células normales y células patológicas, pero estos estudios sólo están en sus inicios.⁸

El transcriptoma es el total de ARN transcritos en un organismo determinado.

El estudio del transcriptoma humano ha deparado varias sorpresas: la mayor parte del genoma humano es transcrita por la ARN polimerasa II, la misma que fabrica los transcritos de los genes codificantes de proteínas.⁹ A esta extensión de la transcripción por fuera de los genes de proteínas se la ha calificado de "pervasiva" (en el sentido de excesiva) y da lugar a la producción de numerosos ARN transcritos no codificantes de proteína. Estos transcritos no merecieron la atención de los investigadores hasta la primera década del siglo XXI, cuando se empezó a tomar conciencia de su importancia en la regulación génica. El *XIST* es un típico "gen de ARN", cuyo papel regulador de la inactivación del cromosoma X está muy estudiado (véase cap. 11). Actualmente se sabe que en el transcriptoma humano solamente el 2,5% del total de transcritos corresponde a genes de proteínas, mientras que el resto corresponde a transcritos de ARN no codificante de proteínas, ya sea largos (100 kilobases o más), cuyo papel regulatorio es intensamente estudiado, y ARN "cortos", que también son funcionales, puesto que intervienen en los fenómenos de interferencia de ARN (véase cap. 7).

El "varioma" es el conjunto de mutaciones que afectan los genes del organismo humano. Existe un Proyecto del Varioma Humano,¹⁰ que tiene por objetivo construir una base de datos de todas las variantes génicas humanas y establecer sus relaciones con los fenotipos correspondientes.



RESUMEN



Un bosquejo histórico del desarrollo de los conceptos de la Genética muestra que Juan Gregorio Mendel, en su trabajo de 1865, predijo con exactitud un número de fenómenos biológicos de gran importancia y el mecanismo de transmisión de los factores hereditarios. Su concepto de “factores” hereditarios, puramente teórico y deductivo, se corresponde con los genes conocidos como tales en la actualidad. La “segregación” de los genes se corresponde con la separación de los cromosomas homólogos de la entonces desconocida meiosis, y la correspondiente separación de los “alelos” o variantes génicas de esos cromosomas. Esta segregación de los alelos es visible en las “genealogías”, por ejemplo, de los miembros de la familia real británica portadores del gen para la hemofilia. El médico inglés Garrod dedujo acertadamente que los factores o genes mendelianos eran responsables de enfermedades hereditarias del metabolismo y predijo también acertadamente que cada “enzima” era determinada por un gen, realizando estudios en varios defectos congénitos del metabolismo, como la alcaptonuria. El ADN fue por primera vez demostrado como el material hereditario en la bacteria neumococo, por Avery y McCarty en 1944. En 1953 Rosalind Franklin consiguió las pruebas experimentales de la estructura del ADN y, en el mismo año, basándose en esos datos y en las proporciones de bases demostradas por Chargaff, J. D. Watson y F. H. C. Crick propusieron el modelo de estructura del ADN consistente en dos hélices de sentido inverso con las bases dirigidas hacia adentro, que establece una complementariedad entre una hélice y la opuesta. Esta estructura, que sugiere el mecanismo de la autorreplicación del ADN y su función de reservorio de información a través de la secuencia de bases, fue verificada por numerosos experimentos. La información hereditaria está codificada en tripletes de bases; este código fue determinado finalmente en 1966 por Nirenberg, Khorana, Ochoa y otros. La descodificación de todo el ADN, es decir su secuencia de bases, fue encarada por el Proyecto Genoma Humano que comenzó en 1990 y termina con éxito en 2003. Esta secuenciación tiene numerosas consecuencias científicas y sociales, y abre el paso a una nueva etapa de estudio del conjunto de las proteínas humanas (proteoma humano). La proteómica ya forma una disciplina propia, y junto al estudio del conjunto de transcritos o transcriptoma, y el de las variantes génicas o varioma, son objeto de investigación en la actualidad.

REFERENCIAS

1. Tjio HJ, Levan A. The chromosome numbers of man. *Hereditas* 1956; 42:1-6.
2. Watson JD, Crick FHC. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature* 1953;171:737-738.
3. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001; 409:860-921.
4. Venter C, et al. The sequence of the human genome. *Science* 2001; 291:1304-1351.
5. Pennisi E. The human genome. *Science* 2001; 291:1177-1180.
6. Baker D, Sali A. Protein structure prediction and structural genomics. *Science* 2001; 294:93-96.
7. Katz-Jaffe MG, McReynolds S, Gardner DK, Schoolcraft WB. The role of proteomics in defining the human embryonic secretome. *Mol Hum Reprod* 2009;15:271-277.
8. Rai AJ, Merlini G. A focus on recent advances in proteomics-one step closer to entrance into the clinical arena. *Clin Cem Lab Med* 2009; 47:625-626.
9. Amaral PP, Dinger ME, Mercer TR, Mattic JS. The eukaryotic genome as an RNA machine. *Science* 2008;319:1787-1789.
10. Cotton RGH, Auerbach AD, Axton M, et al. The human variome project. *Science* 2008;322:861-862.

BIBLIOGRAFÍA ADICIONAL

Micklos DA, Freyer GA. *DNA Science. A first course in recombinant DNA technology.* Cold Spring Harbor Laboratory Press, 1990.