

## **Cosas del genoma**

Carlos López-Fanjul  
1 junio, 2001

El 26 de junio de 2000, el presidente Clinton de los Estados Unidos y el primer ministro británico Blair anunciaron el remate de la secuenciación del genoma humano, reflejo directo de la nacionalidad de los principales centros de investigación embarcados en esta empresa. En realidad, se trataba de dos operaciones paralelas, una de ellas financiada con cargo a fondos públicos (Human Genome Project) y otra de carácter comercial privado (Celera Genomics), que utilizaron técnicas diferentes, durante períodos muy distintos, para obtener los resultados convergentes publicados el pasado mes de febrero en las revistas Nature y Science. El futuro económico de Celera reside en su capacidad de persuadir al posible cliente de la superioridad de sus servicios, pero no faltan quienes atribuyen esa pretendida mayor eficiencia a la ventaja que supone el libre acceso a los datos generados por los centros públicos durante la década previa. Dados los intereses en juego, es evidente que las espadas siguen desenvainadas.

La secuencia en cuestión, que es accesible vía Internet, no es otra cosa que una larguísima sucesión constituida por unos tres mil millones de signos, limitados a cuatro letras diferentes que representan las cuatro bases que componen la estructura química del ADN o material hereditario (A: adenina, C: citosina, G: guanina y T: timina). La mayor parte de esa lista que, en nuestra especie, abarca

aproximadamente tres cuartas partes del total, contiene excesivas repeticiones, no determina proteínas y proviene, en gran medida, de transferencias de ADN vírico acumuladas a lo largo de nuestro pasado evolutivo. Su función actual es desconocida pudiendo ser simplemente parasitaria, aunque cabe especular con otras posibilidades aún no suficientemente investigadas. La otra cuarta parte de la secuencia es transcrita en forma de ARN y posteriormente depurada y ajustada, por eliminación de ciertas partes (intrones) y el ensamblaje de las restantes (exones). Sólo esta última fracción, del orden del 1,5% de la totalidad del genoma (o el 6% de la porción transcrita), es leída de tres en tres letras, de manera que cada triplete especifica uno de los veinte aminoácidos distintos que forman las proteínas o componentes básicos del cuerpo de cualquier ser vivo.

Debe quedar claro que la secuenciación del genoma humano no es tanto un descubrimiento científico como una proeza tecnológica. De hecho, la descripción de genomas más cortos, como el del virus lambda, ya se había logrado en 1982. Lo que caracteriza a la tarea realizada es su gran escala, sólo posible por el desarrollo simultáneo de métodos automatizados innovadores, tanto para la obtención de muestras de ADN y la secuenciación de éstas, como para su posterior tratamiento informático que permite encajar las secuencias parciales en un todo. La revolución tecnológica que posibilita esta estrategia de tipo «bulldozer» precisa de un fuerte respaldo económico y, por esta razón, la investigación con potencialidad para producir beneficios comerciales está pasando rápidamente de las manos de los centros públicos y universidades a las de las grandes empresas multinacionales. Nos guste o no, la creciente dependencia de complejas y costosas herramientas anuncia la próxima extinción de una de las figuras más características de la creación intelectual: el investigador clásico que recolecta paciente y trabajosamente sus datos, para luego someterlos al minucioso análisis dictado por su propia intuición. Pero dejemos esto y veamos qué hemos sacado en limpio, de momento, de la secuenciación del genoma humano.

Su número de genes se cifra ahora entre treinta y cuarenta mil, frente a los seis mil de la levadura de cerveza (*Saccharomyces*), los trece mil de la mosca del vinagre (*Drosophila*), los dieciocho mil de un gusano (*Caenorhabditis*) o los veintiséis mil de una planta (*Arabidopsis*). El que sólo contemos con el doble o el triple de los genes de que disponen los animales comúnmente considerados inferiores, o unos pocos más que los de un vegetal de la misma familia que la col, parece haber preocupado a algunos, aunque no nos han dicho por qué ni, sobre todo, cuáles son las medidas que piensan tomar al respecto. Conviene recordar aquí, no sin cierto sonrojo científico, que la cifra mística de cien mil genes que se barajaba hasta hace unos meses, no era otra cosa que una aproximación grosera, obtenida por la simple división del número total de pares de bases del genoma (unos tres mil millones) por el que se atribuía al gen promedio (unos treinta mil), ignorando a sabiendas que la mayor parte del genoma no es funcional. Da la impresión de que los propios genéticos moleculares se sentían más cómodos pretendiendo ser portadores de más genes, signo flagrante de lo precario de nuestro conocimiento sobre la relación gen-individuo. Por otra parte, tampoco es una sorpresa que el tamaño total del genoma esté poco o nada relacionado con la complejidad del organismo pertinente. De hecho, el nuestro es unas doscientas veces mayor que el de la levadura y otras tantas menor que el de otro organismo unicelular (*Amoeba dubia*), debido a que la fracción no génica del ADN puede variar extraordinariamente de unas especies a otras.

La densidad génica del genoma humano, unos trece genes por millón de bases, es, con mucho, la

menor de todos los secuenciados hasta ahora, en los que esa proporción oscila entre cien y quinientos genes. No obstante, es muy posible que esta disparidad sea un mero subproducto de lo limitado de la comparación. Dicho de otro modo, la elección de las especies mencionadas (excluida la nuestra por razones obvias) se ha circunscrito a un conjunto caracterizado por un contenido de ADN comparativamente pequeño, con objeto de ahorrar esfuerzo experimental en la correspondiente secuenciación. Sólo un 9% de las cerca de mil trescientas familias de proteínas especificadas por nuestro genoma son exclusivas de los vertebrados, lo cual pone de manifiesto que la determinación de las funciones vitales básicas es muy semejante en todos los organismos y, también, que la evolución no ha permitido cambios mayores en este respecto. Por esta razón, la simple comparación del genoma del hombre con el del chimpancé, cuando este último se descifre, no va a decir gran cosa sobre las hasta ahora desconocidas causas últimas de la distinción entre ambas especies.

¿Cuáles podrían ser las consecuencias más o menos inmediatas de la secuenciación del genoma humano? Una de ellas resultaría de la descripción de las diferencias debidas al cambio de una sola base entre los genomas de distintos individuos, con el propósito de establecer si esas sustituciones producen una mayor resistencia o susceptibilidad a contraer una enfermedad determinada. En términos generales, dicha operación no es factible porque el número de cambios de este tipo, correspondientes a mutaciones ocurridas en el pasado evolutivo, se estima en torno al millón y medio. Esto implica que, al comparar las secuencias genómicas de dos individuos cualesquiera, se espera encontrar, aproximadamente, una base cambiada de cada dos mil. Por ello, el esfuerzo debe concentrarse en zonas genómicas muy concretas, donde la información previa sugiera que se pudieran lograr resultados de interés con mayor facilidad. Aun así, el asunto sigue siendo peliagudo. Por ejemplo, los individuos hemofílicos difieren de los sanos en unos doscientos cambios de base, todos ellos en el mismo gen, sin que por ahora se haya llegado a detectar cuáles son los responsables de la deficiente coagulación de la sangre de los afectados por la enfermedad. Dicho sea de paso, el 70% del ADN analizado en el proyecto «Genoma Humano» corresponde a un solo individuo.

Una cosa es establecer la secuencia de una variante génica patológica y otra, muy distinta, deducir a partir de este dato cuáles son las causas de la dolencia y su correspondiente terapia. En este sentido, una de las vías propuestas se orienta hacia el descubrimiento y posterior desarrollo de nuevos medicamentos, basados en la localización genómica de «dianas» sobre las que éstos puedan actuar. Por el momento, se conocen unas quinientas dianas y otros tantos específicos, pero la secuenciación podría permitir el hallazgo de otras distintas y sus correspondientes fármacos. Esto nos lleva al controvertido asunto de las patentes de moléculas que contienen información genética. Por una parte, la detección de un fragmento de ADN cuyo producto interviene en un determinado proceso no es fruto de la invención y, por ello, debería quedar excluido de lo patentable. Por otra, las multinacionales farmacéuticas no invertirían sin la protección que proporcionan las patentes. De hecho, en los Estados Unidos ya se han adjudicado miles de éstas, cuya utilidad futura es puramente especulativa, tanto de secuencias de ADN y mutaciones causadas por una sola base, como de la composición y estructura de los correspondientes productos proteicos. La aplicación indiscriminada de un concepto de propiedad de validez más que dudosa, obstaculizará el progreso sanitario y sólo beneficiará a las grandes compañías, únicas capaces de sostener largos e intrincados litigios.

La secuenciación del genoma humano es un paso importante para el avance de la investigación

genética, pero su publicación ha dado pie a numerosas conjeturas que transmiten la falsa impresión de que la esencia de nuestra naturaleza biológica está contenida (y nuestra sentencia final escrita) en la larga sucesión de bases de que están formados los genes. Este regusto determinista es lo que queda tras la lectura de la obra de Matt Ridley, *Genoma*. La autobiografía de una especie en 23 capítulos (Taurus, 2000), donde dicha secuencia se presenta como el conjunto de «las instrucciones casi completas de cómo construir y hacer funcionar un cuerpo humano» (pág. 11). Cada capítulo está dedicado a un cromosoma diferente y, en cada caso, se ha elegido un gen representativo (real o atribuido) para exponer con diferentes palabras un mismo argumento: la sublimación del poder de la herencia. Para mantener vivo el interés del lector (y, acaso, el volumen de ventas), es casi obligado que ese gen representativo no sea cualquiera. Quizás sea esta la razón por la que se califica de «aburrida» a la variante genética causante de una enfermedad no muy grave (alcaptonuria), que «no se puede vincular al CI [cociente intelectual] o la homosexualidad, no nos dice nada acerca del origen de la vida, no es un gen egoísta, no desobedece las leyes de Mendel, no puede matar ni mutilar» (pág. 68).

La secuela más visible de la herencia biológica es la tendencia de sus variantes a concentrarse en determinadas familias, porque los parientes comparten genes, en mayor medida cuanto más cercanos sean. Pero esa misma propensión también podría observarse si el atributo examinado estuviera enteramente libre de influjos genéticos, porque los parientes participan de muchas circunstancias ambientales, también en razón directa a la proximidad de su vínculo (como ocurre con la herencia testamentaria). Por esta razón, afirmaciones como «no hay lugar a dudas de que la homosexualidad tiene un alto grado de heredabilidad» (pág. 137), no resisten un examen técnico meticuloso. Los estudios de semejanza entre parientes, referidos a rasgos de comportamiento tan vidriosos como el citado, rara vez apoyan o refutan la hipótesis hereditaria o su alternativa; los genes pueden estar ahí o no, pero el procedimiento elegido para sacarlos a la luz suele carecer de poder discriminante. Ridley es perfectamente consciente de esta objeción y, por ello, acaba recurriendo a pruebas a primera vista más contundentes, es decir, a genes concretos. Pero, en el caso que nos ocupa, la demostración de la realidad de éstos es más que dudosa, debido a dificultades de tipo estadístico. Por ello, una vez examinada la candidatura de una variante de la región Xq28 del cromosoma X a la determinación parcial de la homosexualidad masculina, se deja caer que «en realidad, cabe la posibilidad de que la relación entre el Xq28 y la sexualidad sea una falsa apariencia» (pág. 139). Muchas de estas asociaciones estadísticas, aunque no todas, han acabado desvaneciéndose tras exámenes más rigurosos, pero lo que deja huella en el público es el sensacionalismo que acompaña a su primera publicación, no las retractaciones posteriores. Desde luego, cada vez conocemos más variantes genéticas, muchas de ellas perjudiciales. Esto hace posible la valoración del riesgo individual a contraer una determinada dolencia a un plazo dado o, globalmente, la esperanza de vida de una persona. Aplicado al individuo, el concepto de probabilidad conduce a predicciones muy imprecisas, pero no ocurre lo mismo cuando se atribuye a un grupo y, por tanto, las empresas y las compañías de seguros están comenzando a utilizar esa información, imponiendo un cariz determinista a lo que no pasa de mera verosimilitud.

*Genoma* transmite un mensaje en buena medida desproporcionado. Aunque sólo sea de pasada, Ridley reconoce que, en ocasiones, las cosas no son tan sencillas, que se trata de «un mundo de penumbras, de matices, de calificativos, de "depende"» (pág. 81). En muchos casos no será fácil que

las técnicas más perfeccionadas contribuyan a despejar esa incertidumbre, porque forma parte de la naturaleza intrínseca de la mayoría de los genes el que su expresión varíe (o, incluso, desaparezca) dependiendo del medio en que se expresen, y ello estará supeditado al particular conjunto de circunstancias experimentadas por el individuo a lo largo de su existencia.