



# Del Proyecto Genoma a la medicina personalizada

Francisco Javier Novo Villaverde



**Biografía.** Es doctor en Medicina por la Universidad de Compostela, ciudad en la que también realizó la especialidad en Bioquímica Clínica. Durante dos años, trabajó en el Proyecto Genoma Humano en la Universidad de Cambridge, y posteriormente se trasladó a Londres para trabajar en un proyecto de terapia génica en una enfermedad genética. Actualmente es profesor Titular en el Departamento de Genética de la Universidad de Navarra, donde imparte docencia sobre Genética Humana y sobre Bioinformática. Su investigación se centra en la identificación de alteraciones genéticas que causan cáncer.

**Resumen.** En el año 2001 se abrió una nueva puerta en la lucha contra las enfermedades. En aquel año, el esfuerzo conjunto de miles de científicos de todo el mundo cristalizó en la publicación de la secuencia de los 3.000 millones de letras, los nucleótidos, de que consta nuestro manual de instrucciones, el genoma humano. Es éste un hallazgo de extraordinaria relevancia biomédica ya que, entre otras utilidades, ayudará a determinar nuestra distinta predisposición a sufrir determinadas enfermedades y la manera en que respondemos a los fármacos. En aquel momento comenzó también una nueva etapa encaminada a identificar todas las letras del genoma que pueden ser diferentes entre personas distintas. Las primeras investigaciones en esta área han mostrado que existe una cantidad de variación mucho mayor de la esperada: dos personas cualesquiera tienen alrededor de tres millones de letras diferentes en su genoma. Para que se pueda hacer realidad el enorme potencial que el estudio del genoma pone a nuestra disposición, la tecnología utilizada actualmente para leer su secuencia es insuficiente y debe ser sustituida por otra que reduzca tiempo y costes y permita leer la secuencia completa del genoma de miles de personas en los próximos años. Nadie puede anticipar cuánto tendremos que esperar hasta alcanzar ese conocimiento pero los avances en las tecnologías de secuenciación permiten pensar que sucederá antes de lo que muchos pensaban.

**Summary.** In 2001, a new door was opened in the fight against disease when the combined efforts of thousands of scientists all over the world crystallised in the publication of the sequence of the 3 billion letters - nucleotides - that form our instruction manual: the human genome. This is an immensely important biomedical discovery because, among other things, it will help determine the different propensity of individual human beings to certain diseases and the way we respond to medicinal drugs. That discovery also marked a new stage aimed at identifying all the letters in the genome that may be different in different people. Initial research in this field has shown that the differences are much greater than initially expected: any two people have around three million different letters in their genome. In order to realize the enormous potential afforded by human genome research, existing technology used to read its sequence is insufficient and must be replaced by other technology to reduce time and costs and decipher the complete sequence of the human genome in the coming years. Nobody knows exactly how long we will have to wait until we obtain that knowledge but advances in sequencing technologies suggest that this may be sooner than we think.

**Résumé.** En 2001, une nouvelle porte dans la lutte contre les maladies s'est ouverte. Cette année-là, l'effort conjoint de milliers de scientifiques du monde entier a cristallisé dans la publication de la séquence des 3 000 millions de lettres, les nucléotides, dont est composé notre manuel d'instructions, le génome humain. C'est une découverte d'une importance biomédicale extraordinaire étant donné qu'entre autres utilités, elle aidera à déterminer notre prédisposition différente à souffrir de certaines maladies et la manière dont nous répondons aux médicaments. À ce moment-là, a également commencé une nouvelle étape visant à identifier toutes les lettres du génome qui peuvent être différentes entre des personnes différentes. Les premières recherches dans ce secteur ont montré qu'il existe une quantité de variation beaucoup plus grande que celle attendue : deux personnes quelconques ont environ trois millions de lettres différentes dans leur génome. Pour que l'énorme potentiel que l'étude du génome met à notre disposition puisse se réaliser, la technologie utilisée actuellement pour lire sa séquence est insuffisante et doit être remplacée par une autre réduisant le temps et les coûts et permettant de lire la séquence complète du génome de milliers de personnes durant les prochaines années. Personne ne peut anticiper combien nous devrons attendre avant d'atteindre cette connaissance mais les progrès en technologies de séquençage permettent de penser que cela se produira plus vite qu'on ne le pense.



**P**ara celebrar su 18 cumpleaños, Andrea envió un raspado bucal al laboratorio y abonó los 600 euros que costaba el «regalo». Dos semanas después recibió un mensaje electrónico con un archivo que contenía la secuencia completa de su genoma, junto con una interpretación pormenorizada del mismo. Aunque había estudiado algunas nociones básicas de genética en el colegio, todo aquello se le hacía bastante incomprendible. Días después, cuando recibió la video-llamada en su dispositivo portátil a la hora convenida, pudo hablar cara a cara con el especialista en consejo genómico, el cual le explicó durante media hora cuáles eran sus principales características genéticas. En concreto, le recomendó un programa de dieta y ejercicio para retrasar el comienzo de un tipo de hipercolesterolemia al que mostraba cierta predisposición. Para la psoriasis que indudablemente desarrollaría, y de la que ya había empezado a notar algunos síntomas, le aconsejó la medicación más adecuada a su perfil genético. La peor noticia fue que tendría que comenzar un programa de revisiones periódicas para la detección precoz de cáncer de mama. «Ventajas de vivir en el 2019», pensó.

Así podría comenzar la típica novela cuya acción tiene lugar en un futuro más o menos lejano. Y, sin embargo, es cada vez más probable que el escenario descrito en esas líneas sea realidad en pocos años. Nuestra protagonista habría nacido en el año 2001, coincidiendo con la publicación de la secuencia del genoma humano. Cuatro años después, se comenzaron a catalogar las diferencias entre genomas de distintas personas y esto deparó algunas sorpresas. De hecho, a finales del año 2007, la tradicional votación de la revista *Science* concluyó que el avance científico más importante de ese año había sido el estudio de la variación genética en humanos por el impacto que indudablemente tendrá en nuestro conocimiento de los factores biológicos implicados en el desarrollo de enfermedades comunes. En el momento actual, estamos siendo testigos de la aparición de las primeras máquinas capaces de secuenciar un genoma completo en pocos meses y esto acelerará tremendamente el estudio de las variantes genéticas que determinan la predisposición de cada persona a padecer distintas enfermedades. Los desarrollos tecnológicos de los próximos años pueden hacer que el ideal de la «medicina personalizada» sea una realidad en el próximo decenio.

\* \* \*

Visto con la perspectiva de estos años, el Proyecto Genoma Humano fue un trabajo impresionante. Para reconstruir el orden preciso de los 3.000 millones de nucleótidos, «letras», de que consta nuestro genoma, aquellos pioneros tuvieron que dividir el genoma en trocitos, leer la secuencia de cada uno de los fragmentos y después volver a ensamblar todo poco a poco, como un gigantesco puzzle de la vida. De hecho, como en todo puzzle, el principal problema era saber a qué región del genoma correspondía cada trozo, por lo que un buen número de años se dedicó a construir mapas del genoma que permitiesen ordenar los fragmentos antes de leer sus secuencias. Al final, cientos de investigadores necesitaron unos trece años para obtener una secuencia de alta calidad de un genoma humano.

Una de las cuestiones que surgió ya en esos años era que esta secuencia «modelo» no se correspondía necesariamente con el genoma de ninguna persona concreta, puesto que entre los genomas de las personas «reales» deberían existir muchas pequeñas diferencias. Los primeros análisis confirmaron esta sospecha y animaron a la comunidad científica a

explorar con mayor detalle la «variación» que muestra el genoma cuando se comparan distintas personas. La idea era responder a una pregunta fundamental: ¿cuáles son las diferencias genéticas que nos hacen distintos unos de otros? Y distintos no sólo en el color del pelo, o en la estatura, sino sobre todo en cuestiones de gran relevancia biomédica: la predisposición a sufrir determinadas enfermedades, la distinta respuesta a fármacos, la susceptibilidad a sustancias tóxicas del medio ambiente. En definitiva, todas las diferencias que están en la base de nuestras peculiaridades biológicas.

Así comenzó una nueva etapa, encaminada a identificar y catalogar todas las letras del genoma que pueden ser diferentes entre personas distintas. Estas investigaciones mostraron que, cuando se leen los genomas de dos personas tomadas al azar, estos cambios aparecen en una de cada 1.000 letras. Dado el tamaño total del genoma, esto significa que esas dos personas tendrán alrededor de tres millones de letras diferentes. Si pensamos que el cambio de una sola letra puede en ocasiones provocar una enfermedad genética, tal cantidad de variación resulta extraordinaria. Por suerte, los estudios posteriores han mostrado que la mayor parte de estas variaciones en realidad no tiene efectos funcionales y que sólo un pequeño porcentaje de esos cambios afecta al funcionamiento de los genes. Además, el efecto de estas variaciones es generalmente pequeño: hacen que un gen funcione un poco mejor o un poco peor pero sin llegar a inactivarlo totalmente.

Estas investigaciones pusieron también de manifiesto la importancia de que los científicos de todo el mundo trabajen en colaboración, algo que ya había caracterizado al Proyecto Genoma Humano. Gracias a las sinergias resultantes de esta forma de trabajar, en poco tiempo se pudo concluir un primer catálogo de variaciones, que incluía un millón de cambios de secuencia en individuos de procedencia europea, asiática y africana. Después, otros estudios han ido ampliando este catálogo y la idea es que hacia finales de 2008 podamos disponer de una lista completa de los aproximadamente diez millones de variantes que, según las últimas estimaciones, aparecerán cuando se estudien individuos de todas las poblaciones del planeta. Este catálogo será de una utilidad biomédica extraordinaria ya que permitirá hacer dos cosas fundamentales: por un lado, esas variaciones se usarán como marcadores para detectar si cerca de ellas reside algún gen que juegue un papel importante en el desarrollo de las principales enfermedades humanas; además, será posible identificar cuáles son las variantes que realmente afectan a la función de esos genes y esto ayudará a conocer mejor los mecanismos por los que se originan esas enfermedades y a diseñar nuevas estrategias terapéuticas más eficaces y seguras.

Si todo esto ya supone un reto importante para los implicados en investigación genómica, los últimos dos años han traído nuevas sorpresas, una especie de «más difícil todavía» que nos recuerda la dificultad de prever lo que va a suceder en el mundo de la ciencia. Los nuevos protagonistas del drama de la variación genética son las llamadas «variantes estructurales», que precisamente fueron las merecedoras del galardón otorgado en el año 2007 por la revista *Science* al que antes se hacía mención. Como su nombre indica, este nuevo tipo de variación consiste en cambios estructurales, que afectan a regiones relativamente grandes del genoma, entre distintos individuos normales. Imaginemos, por ejemplo, un fragmento de varios miles de letras que aparece repetido dos, tres o cuatro veces (o está totalmente ausente) en los genomas de personas distintas. Por pura probabilidad, un fragmento de ese tamaño podría albergar uno o varios genes, lo que sig-

nifica que unas personas podrían llevar más copias de un gen que otras. Pero esto, que podría ser desastroso dependiendo del gen implicado, no es lo más sorprendente. Lo increíble es que existen unas doscientas regiones del genoma que muestran este tipo de variación entre personas aparentemente normales. Hace diez años, cualquier genetista habría dicho que tal cantidad de alteraciones es incompatible con la ausencia de enfermedad genética severa. Y, sin embargo, así es. Evidentemente, todos esos cambios pueden tener efectos considerables sobre las diferencias biológicas entre personas, su predisposición a desarrollar distintas patologías, la diferente manera de interactuar con el ambiente y de responder a estímulos o agresiones externas, etc. Pero los estudios que permitirán saber exactamente cuáles son los efectos de las variantes estructurales están todavía en sus fases iniciales. Esta parte de la historia no ha hecho más que empezar...

En cualquier caso, para que todo lo anterior pueda ser realidad algún día, parece claro que la tecnología «clásica» utilizada para leer la secuencia del genoma es totalmente insuficiente. Para llevar a cabo los estudios mencionados, será necesario leer la secuencia completa del genoma de miles de personas en todo el mundo. Buena muestra de ello es la nueva iniciativa internacional, hecha pública en enero de 2008, cuyo objetivo es leer la secuencia completa del genoma de mil personas en los próximos tres años. Evidentemente, no podemos tardar diez años en completar cada una de esas secuencias... Por tanto, éste es un punto crucial en el que los científicos tampoco se han quedado ociosos. De hecho, los avances en las tecnologías de secuenciación han sido espectaculares: si la secuencia del Proyecto Genoma Humano fue fruto de los esfuerzos de 2.800 personas durante trece años con un coste total cercano a los 2.700 millones de dólares, en abril de 2008 pudimos asistir a la publicación de la secuencia completa del genoma de un individuo llamado James Watson, conocido por ser uno de los descubridores de la estructura del ADN. La lectura del genoma de Watson se completó en cuatro meses, con la participación de 27 investigadores y un coste en torno a 1 millón de dólares. ¿Dónde está el secreto?

El secreto, como es habitual en estos casos, está parcialmente oculto por las patentes que rodean al competitivo mundo de la industria biotecnológica. Lo que está claro es que están emergiendo nuevas tecnologías, cuya estrategia consiste en realizar miles de reacciones simultáneas en pequeñísimos tubos de ensayo, del tamaño de una gota, para así acelerar y abaratar todo el proceso. Estas nuevas máquinas pueden llegar a hacer que la lectura de un genoma completo sea algo rutinario. El interés en este campo es tal, que la «X PRIZE Foundation» ha creado el Premio «Archon X PRIZE», una especie de concurso para ingenieros: el primero que sea capaz de leer la secuencia completa de cien genomas humanos en diez días con un coste de diez mil dólares por genoma, se llevará un premio de diez millones de dólares. Varias empresas participan en el concurso, entre ellas la misma que completó el genoma de Watson. El fundador de otra de las empresas participantes está seguro de ganar el premio, y afirma que la tecnología desarrollada por su empresa podrá leer un genoma humano completo «en cuestión de minutos, por unos pocos céntimos...»

\* \* \*

Y así volvemos al futuro. Todos los ingredientes están servidos para que la historia de Andrea sea realidad en 2019, si no antes. Está claro que queda todavía mucho camino por recorrer pero cada semana asistimos a la publicación de algún avance en este campo. A

no ser que se produzca algún nuevo descubrimiento revolucionario, algo que nunca se puede descartar del todo, parece que ya sabemos cuáles son las letras del genoma que codifican las diferencias biológicas entre personas. De entre todas ellas (millones, en total) habrá que identificar las que realmente tienen interés biomédico, las que permitirán predecir la predisposición a padecer enfermedades y la respuesta a fármacos o a los estímulos ambientales. Nadie puede anticipar cuánto tendremos que esperar para alcanzar ese conocimiento pero los avances actuales en las tecnologías de secuenciación permiten pensar que sucederá antes de lo que muchos pensaban.

Y sin embargo, todo esto puede fallar por un punto al que habitualmente se presta muy poca atención. Si algo puede retrasar la llegada de la medicina predictiva basada en el conocimiento del genoma, no será la falta de dinero o de investigadores o de tecnologías adecuadas. El gran peligro es que la sociedad, las personas que podrían beneficiarse de estos avances, no estén suficientemente preparadas para comprender el significado y las implicaciones de la ciencia genómica. Como tantas veces, puede fallar el aspecto peor previsto, el más olvidado: la educación. Estamos en un momento crucial en el que la sociedad debe tomar conciencia de la importancia de que todas estas nociones sean percibidas claramente y con corrección por todos los ciudadanos. Y esa conciencia deberá traducirse en acciones concretas para promover la cultura científica y la comprensión de los avances en nuestro conocimiento del genoma humano. La batalla por hacer realidad la medicina personalizada no se librará solamente en los centros de investigación genómica; tendrá lugar también en los colegios, en los libros de texto, en los medios de comunicación, en las asociaciones culturales, en las bitácoras científicas. Sólo si ganamos esta batalla la historia de Andrea podrá ser un día realidad. Sólo entonces podremos decir que los esfuerzos de miles de personas, durante años de investigación, han servido para avanzar en la lucha contra la enfermedad.

## **Bibliografía y fuentes de información**

- [1] Initial sequencing and analysis of the human genome. *Nature*. 2001 Feb 15; 409 (6822): 860-921.
- [2] A haplotype map of the human genome. *Nature*. 2005 Oct 27; 437 (7063): 1299-320.
- [3] Structural genomic variation and personalized medicine. *N Engl J Med*. 2008 Feb 14; 358 (7): 740-1.
- [4] The complete genome of an individual by massively parallel DNA sequencing. *Nature*. 2008 Apr 17; 452 (7189): 872-6.
- [5] Mapping and sequencing of structural variation from eight human genomes. *Nature*. 2008 May 1; 453 (7191): 56-64.

