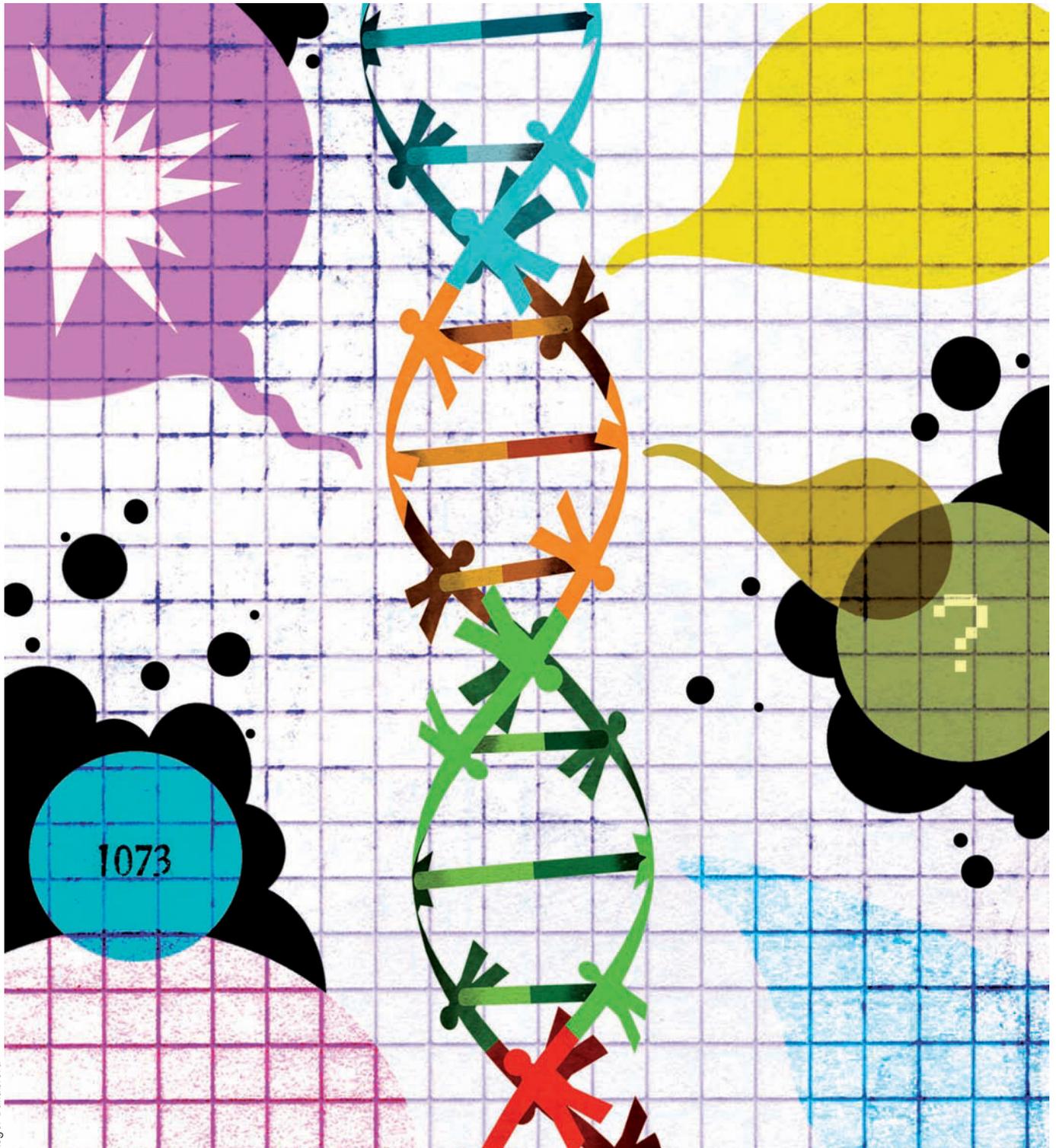


Hablemos de...

# El proyecto HapMap

ÁNGEL CARRACEDO

Fundación Pública Gallega de Medicina Genómica (SERGAS). Centro Nacional de Genotipado. Universidad de Santiago de Compostela. Facultad de Medicina. Santiago de Compostela. A Coruña. España.



Roger Ballabre

Los polimorfismos nucleotídicos simples (SNP, del inglés *single nucleotide polymorphisms*) son los marcadores ideales para, a través de estudios de asociación, encontrar los genes implicados en enfermedades comunes o genéticamente complejas.

Si bien a raíz del Proyecto Genoma Humano se descubrieron y validaron millones de SNP, este tipo de estudios era poco eficaz y económicamente casi inviable para los análisis genómicos globales, hasta que el descubrimiento de que el genoma humano está compuesto por bloques haplotípicos, donde apenas hubo recombinación separada por segmentos con altísima recombinación, abrió nuevas perspectivas. Si fuésemos capaces de identificar los bloques haplotípicos y encontrar los Tag SNP que definen los bloques, el análisis de SNP para realizar un análisis genómico global podrían pasar de 11 millones de SNP a alrededor de 400.000 SNP, lo que los haría mucho más asequibles y eficaces.

Este es el objetivo principal del proyecto HapMap, cuya primera fase acaba de concluirse y que consistió en el análisis de todos los SNP conocidos con una frecuencia esperada de al menos el 5% (esto es algo más de un millón de SNP) en cada una de 269 muestras seleccionadas para el proyecto de poblaciones del este de Asia, de origen europeo y africanas.

Los Tag SNP identificados y las herramientas informáticas desarrolladas en el proyecto HapMap, conjuntamente con la creación en algunos países de grandes centros de genotipado (CeGen en España), están permitiendo el sueño de abordar el análisis del componente genético de la enfermedad común en un gran número de enfermedades.

En 1980 Whyman y White<sup>1</sup> describieron el primer polimorfismo de un simple cambio de base en el ácido desoxirribonucleico (ADN). Este tipo de polimorfismos actualmente se denominan SNP y son, en general, polimorfismos bialélicos, lo cual significa que en una posición determinada del genoma podemos tener, por ejemplo, una A o una T, de modo que los individuos podemos ser AA, AT o TT. Desde entonces, se ha ido acumulando información creciente sobre este tipo de polimorfismos. Tras la finalización del Proyecto Genoma Humano, se comprobó que el número de SNP era muy abundante y actualmente se superan los 11 millones en el genoma humano recogidos en bases de datos públicas o privadas ([www.ncbi.nlm.nih.gov/projects/SNP/](http://www.ncbi.nlm.nih.gov/projects/SNP/)).

Los SNP actuarían como marcadores para localizar genes involucrados en enfermedades, gracias a la enorme ventaja de su densidad en el genoma, que es aproximadamente de un SNP cada 300 pares de bases. La idea sería realizar estudios de asociación con un grupo de pacientes y un grupo control de modo que se puedan encontrar SNP (y en consecuencia genes) asociados a una enfermedad. Por ejemplo, se trataría de recoger ADN de un grupo de personas que tuviera hipertensión arterial u otra enfermedad compleja, y un grupo de personas que no la tuviera. Si un SNP particular es más frecuente en la gente con hipertensión, podría utilizarse para localizar e identificar un gen implicado en la hipertensión. La localización de los genes involucrados en enfermedades complejas (esto es, las que tienen un componente multigénico y otro ambiental que interactúan entre sí) es tremendamente importante, ya que éstas representan el gran reto actual de la medicina y la farmacología. Estamos hablando de las enfermedades más comunes y de mayor mortalidad y morbilidad, como el cáncer, la hipertensión, la diabetes, las enfermedades psiquiátricas, las jaque-

cas, etc., en los que la farmacopea actual tiene una eficacia media limitada.

Si localizáramos los genes involucrados en estas enfermedades, podríamos subclasificarlas mejor y, en consecuencia, tratarlas mejor. Además, el descubrimiento de nuevos genes implicados en estas enfermedades permitiría obtener nuevas dianas a las que dirigir los fármacos. Por último, el hecho de descubrir variaciones en el ámbito del ADN que condicionen la respuesta a fármacos (farmacogenética) permitiría el desarrollo de la medicina personalizada.

Pero incluso con el enorme progreso tecnológico realizado en los últimos años, el análisis de un número tan enorme de SNP (imaginemos 10 millones) en un grupo de individuos (pongamos 2.000) supondrían 20.000 millones de genotipos para analizar el genoma completo, lo que sería muy difícil prácticamente, y casi inviable económicamente. Los análisis genómicos globales, que, respecto a analizar genes candidatos, tienen la ventaja de evitar sesgos hacia genes concretos, serían prácticamente inviables.

La evidencia de que el genoma humano está constituido por bloques haplotípicos<sup>2</sup>, con apenas recombinación dentro de cada bloque, abrió una nueva etapa de ilusión. Ahora sabemos que hay segmentos de cromosomas ancestrales que no se han roto por recombinación y que están separados por segmentos donde la recombinación es frecuente. Estos segmentos son los bloques haplotípicos.

Esto quiere decir que en un bloque donde hay, por ejemplo, 30 o 40 SNP, por definición todos polimórficos, en realidad sólo hay unas pocas variaciones haplotípicas. Así, en una población dada, el 55% de las personas podría tener una versión del haplotipo; el 30%, otra; el 8%, una tercera, y el resto, unos pocos haplotipos menos comunes.

Si conociésemos la estructura de estos bloques e identificásemos los Tag SNP que definen la variación en cada bloque, podríamos reducir muy significativamente el número de SNP necesarios para un estudio de asociación. Se calcula que hay aproximadamente de 300.000 a 600.000 Tag SNP, que es considerablemente menos que 11 millones de SNP, lo que posibilita hacer estudios de asociación con una eficacia mucho mayor y hace posible y asequible realizar incluso barridos genómicos globales.

Gracias al ADN mitocondrial sabemos que somos una especie relativamente joven que surgió en África hace algo menos de 150.000 años. Cuando los humanos migraron de África llevaron con ello parte de la variación que había en la población ancestral, aunque no toda. Como resultado de esto, los haplotipos vistos fuera de África tienden a ser subgrupos de los que se encuentran allí. Además, los bloques haplotípicos de las poblaciones no africanas tienden a ser más largos que los encontrados en poblaciones africanas, ya que su historia es mucho más larga y han tenido más tiempo de romperse por episodios de recombinación.

Dentro de un bloque haplotípico, las frecuencias de los haplotipos varían de región a región debido a variación al azar, fenómenos de selección y otros mecanismos genéticos. Al mismo tiempo, mutaciones en las secuencias de ADN han creado nuevos haplotipos que también tienen distribuciones poblacionales. Los más recientes, por ejemplo, no han tenido aún tiempo de extenderse fuera de la población y el área geográfica donde se han originado.

Y ésta fue la finalidad del proyecto internacional HapMap, que consiste no sólo en un catálogo de las variantes genéticas comunes simples que ocurren en el genoma humano, y en el análisis de cómo están distribuidas esas variantes en las distintas poblaciones humanas, sino también en el dibujo de los bloques haplotípicos en las principales poblaciones y en la identificación de los Tag SNP que definen la variación en cada bloque. La finalidad última del proyecto es encontrar, gracias a este esfuerzo, genes involucrados en enfermedades complejas y entender mejor la respuesta individual a los fármacos (farmacogenética).

Como ocurrió con el Proyecto Genoma Humano, el proyecto HapMap ([www.hapmap.org](http://www.hapmap.org)) es una colaboración de científicos y agencias de varios países. En este caso, cooperan Estados Unidos, Japón, el Reino Unido, Canadá, China y Nigeria. Oficialmente, el proyecto comenzó en octubre del 2002 y la primera fase terminó 3 años más tarde (Internacional HapMap Consortium, 2005)<sup>3</sup>.

El consorcio HapMap está analizando 269 muestras de 3 grandes grupos poblacionales: africanos (yorubas, ibadán, Nigeria), Este de Asia (japoneses de Tokio y chinos Han de Pekín) y europeos (residentes en Utah, con antepasados del Norte y Oeste de Europa). En el Instituto Coriell se inmortalizaron líneas celulares provenientes de estas muestras y se analizaron en centros de genotipado de los diferentes países del consorcio.

El objetivo de la fase I del proyecto fue genotipar al menos un SNP común (con una frecuencia esperada de 0,05 o más) cada 5 Kb del genoma en cada una de las 269 muestras (esto es más de un millón de SNP en cada una de las muestras). En la fase II del proyecto, que acaba de iniciarse, se quiere analizar 4,6 millones de SNP adicionales en cada una de las muestras HapMap.

Además de esto, en la fase I se eligieron 10 regiones de 500 Kb cada una (seleccionadas del proyecto ENCODE-Encyclopedia of DNA Elements) que se secuenciaron en 48 individuos y todos los SNP de esas regiones, bien descubiertos en el análisis, bien existentes en la base dbSNP (la base pública más importante de SNP), se genotiparon en el grupo completo de 269 muestras.

Gracias al HapMap, no sólo se pudo determinar con mayor precisión la frecuencia de los SNP analizados, sino, y esto es mucho más importante, se pudo conocer la estructura de los bloques haplotípicos. El desarrollo paralelo de excelentes software de análisis (Haploview) permite a los investigadores hacer diseños adecuados de sus estudios y elegir aquellos SNP (Tag SNP) que definen el bloque. También la información es esencial para poder elegir posibles SNP causales de una enfermedad con componente genético.

La primera fase del proyecto HapMap confirmó la generalidad de puntos calientes de recombinación junto con largos segmentos con un fuerte desequilibrio de ligamiento y diversidad haplotípica.

El proyecto HapMap está ya permitiendo acelerar la identificación de los factores genéticos implicados en enfermedades.

Su éxito o fracaso se medirá en términos de genes descubiertos y la mejora del conocimiento de las bases de las enfermedades y su tratamiento.

Pero para que esto sea un éxito, necesitamos un esfuerzo importante en otros ámbitos. En primer lugar, que los científicos sean capaces de organizar consorcios bien establecidos para estudios de asociación con diseños adecuados, en los que se recoja un número importante de muestras en pacientes y de controles con todas las garantías éticas y legales, y en los que se defina muy bien el rasgo fenotípico que queremos analizar. Se están encontrando muchísimas asociaciones entre genes y enfermedades, y muy pocas se están pudiendo replicar, pero estamos aprendiendo rápidamente. Hoy sabemos que la población importa, que podemos seleccionar poblaciones para obtener ventajas (con la obtención de ventajas del aislamiento, por ejemplo, o de la mezcla reciente), pero que también pueden causar problemas (como los derivados de posibles estratificaciones de la población). También sabemos que la enfermedad importa y que la definición de la enfermedad, sus síntomas o la población control es esencial y está lejos de ser fácil. Por último, sabemos que hay muchas causas genéticas que pueden alterar los resultados, por lo que la replicación del estudio en una población independiente es esencial.

En segundo lugar, es importante que los científicos puedan disponer de centros de genotipado, donde se puedan analizar de forma masiva SNP, y que dispongan de herramientas bioinformáticas que les permitan analizar una información tan compleja.

En España, afortunadamente, disponemos de un Centro Nacional de Genotipado ([www.cegen.org](http://www.cegen.org)), con una infraestructura y robótica muy buena, herramientas bioinformáticas excelentes y un personal muy preparado en genotipado y estudios de asociación, que utilizan la información proporcionada por HapMap para el diseño de los estudios. En este centro, se está realizando un importante número de proyectos por parte de consorcios de investigadores de España, desde grandes proyectos en los que se analiza todo el genoma, o proyectos medianos o pequeños que seleccionan genes candidatos.

Estos proyectos, al igual que otros que se están haciendo en otros centros de genotipado del mundo, están ya proporcionando importantes resultados, por lo que el éxito de las aplicaciones del proyecto HapMap parece garantizado, y parece claro que el conocimiento derivado del proyecto contribuirá, en los próximos años, a una comprensión y tratamiento mejores de las enfermedades.

## Bibliografía

1. Wyman AR, White R. A highly polymorphic locus in human DNA. *Proc. Natl Acad Sci USA*. 1980;77:6754-8.
2. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. *Science*. 2002;296:2225-9.
3. International HapMap Consortium. A haplotype map of the human genome. *Nature*. 2005;437:1299-320.

## Actualización **Metástasis hepáticas**

Metástasis hepáticas de carcinoma de origen desconocido:  
orientación diagnóstica

*Ferran Losa Gaspà*

Metástasis hepáticas de tumores neuroendocrinos:  
manejo medicoquirúrgico

*Alexandre Teulé Vega, Virginia Arrazubi Arrula, Gemma Soler González, Aleydis Pisa Gatell y Ramón Salazar Soler*

Metástasis hepáticas de carcinoma colorrectal:  
tratamiento quirúrgico

*Santiago López-Ben, Joan Figueras, Maria Teresa Albiol y Antoni Codina B*

Metástasis hepáticas de cáncer colorrectal:  
tratamientos regionales

*Antonio Viúdez y Javier Rodríguez*



## Implicaciones clínicas de la investigación básica

Función del CFTR. Más allá de la fibrosis quística

*Xavier Molero Richard*

## El lugar en terapéutica de...

Utilidad de las prótesis autoexpandibles digestivas

*Àngels Vilella Martorell y Carlos Dolz Abadía*

## Revisión técnica diagnóstica

Técnicas diagnósticas de hipertensión pulmonar en pacientes  
n lista de espera para trasplante hepático

*Francisco Acosta Villegas, Tomás Sansano Sánchez y Carlos García Palenciano*

## Ensayos clínicos y práctica clínica

Inyección lesional de esteroides en las estenosis esofágicas pépticas

*Livia de Rezende y David Nicolás Pérez*

## Prevención de...

Prevención del rechazo. ¿Es posible la tolerancia en el trasplante hepático?

*Pilar García, Marc Martínez-Llordella, Isabel Puig-Pey y Alberto Sánchez Fueyo.*

## Ventana a otras especialidades

Síndrome metabólico: de la resistencia a la insulina a la enfermedad cardiovascular

*Albert Lecube*

## Hablemos de...

La proteómica en investigación biomédica

*Fernando J. Corrales*