

Impacto del proyecto del genoma humano sobre la epidemiología y la salud pública

Antonio Velázquez*

Un genoma es la totalidad del material genético de un individuo o de una especie. La palabra se deriva de la fusión de los términos "genes" y "cromosomas". Comprende así a todos los genes y su ubicación en los diferentes cromosomas. Es comparable a un manual de instrucciones que guía a las células desde el óvulo fecundado por el espermatozoide, en su desarrollo hasta el final de la vida de un organismo.

Proyectos de investigación sobre genomas

A fines de la década del 80 y a lo largo de la década actual, un grupo cada vez más numeroso de países han establecido proyectos nacionales y locales para la investigación de diversos genomas, en particular el humano. Esto se ha acompañado de un espíritu de colaboración que ha dado origen a un proyecto internacional sobre el genoma humano (PGH), cuya meta es el conocimiento completo de la estructura del genoma humano, a más tardar para el año 2005.

Ya desde ahora, y creciendo en forma exponencial, han surgido un gran número de aplicaciones, particularmente en la medicina, además de enormes contribuciones al conocimiento de la biología humana. Para tener una idea de la magnitud de este proyecto, podemos comparar al genoma humano con nuestro planeta; si una célula fuese del tamaño de toda la Tierra, uno de los cromosomas correspondería a un continente, una banda cromosómica sería del tamaño de un país, un gen sería

similar a un municipio y una base nitrogenada equivaldría a uno de los habitantes humanos de la Tierra. En esta forma, el Proyecto del Genoma Humano equivaldría a identificar y localizar a cada uno de los habitantes de nuestro planeta.

Este proyecto tiene diferentes niveles de resolución, desde el conocimiento de que tenemos 23 pares de cromosomas, que se obtuvo en 1956, y que estos cromosomas tienen regiones específicas observables al microscopio de luz, bien definidas por los diferentes métodos de bandeo cromosómico, información que se obtuvo a partir de la década del setenta, hasta la determinación de la secuencia completa de los 3,000 millones de pares de bases nitrogenadas contenidas en el DNA humano. El PGH incluye la elaboración de diferentes mapas, tales como un mapa genético, un mapa físico y un mapa de genes.

Marcadores y mapas genéticos

En 1995 se logró completar el mapa genético de los 23 pares de cromosomas humanos, con un número suficiente de marcadores genéticos en cada uno de los cromosomas, como para permitir la localización de cualquier gen en el que uno esté interesado. Podemos imaginarnos a los mapas genéticos de los diferentes cromosomas, como mapas de las diferentes carreteras de México en las que, a trechos de un kilómetro de distancia, existen piedras a la vera del camino en las que está anotado el kilometraje desde el punto de partida.

*Unidad de Genética de la Nutrición, Instituto Nacional de Pediatría
Correspondencia y solicitudes de sobresos Dr. Antonio Velázquez. N.P. Insurgentes Sur N.º 3700 C, Insurgentes Cuicuilco 04530, México D.F.
Tel 606 3489

Esas piedras se conocen en castellano como mojeneras, (en inglés se llaman *milestones*); su equivalente en un mapa genético son los marcadores genéticos.

Un marcador genético puede ser un gen o simplemente una pequeña región del DNA del cual conocemos su ubicación exacta y que, además, tiene diferentes formas alternativas (alelos). Los primeros marcadores genéticos que se emplearon fueron los genes del grupo sanguíneo ABO. Un grupo de marcadores genéticos que han sido particularmente útiles por su asociación con enfermedades autoinmunes y por su gran variedad alélica, son los genes del complejo mayor de histocompatibilidad (HLA).

Con el descubrimiento de las endonucleasas de restricción y su propiedad de reconocer secuencias específicas de varios nucleótidos y de cortar el DNA en esos sitios, surgió una nueva clase de marcadores genéticos, definidos directamente a nivel del DNA. Se tienen ya localizados, a lo largo del todo el genoma humano, más de 5 mil de ellos. Se conocen como polimorfismos en la longitud de fragmentos de restricción (RFLP's es el acrónimo de su denominación en inglés), debido a que diferencias alélicas se manifiestan en fragmentos de DNA de diferente tamaño que, después de ser separados entre sí por electroforesis, pueden ser visualizados con facilidad (empleando sondas específicas o amplificándolos por medio de la reacción en cadena de la polimerasa: PCR). Muchos de estos sitios tienden a ser altamente polimórficos en las poblaciones naturales, incluyendo a las humanas; esto es, los diferentes alelos de un determinado marcador, cada uno representando un fragmento de DNA de diferente longitud, tienen frecuencias relativamente altas en la población (tienen que ser mayores del 1% para poder ser denominados *polimorfismos*).

Una variedad de estos RFLP's que ha resultado particularmente útil son los llamados microsatélites que consisten en segmentos de DNA en los que unos pocos nucleótidos—di, tri o tetranucleótidos—se encuentran repetidos muchas veces. Las diferentes formas alélicas de un determinado marcador microsatélite están dadas por el número de repeticiones de estos oligonucleótidos y varían grandemente entre diferentes individuos. Las enzimas de restricción cortan a ambos lados de los

microsatélites y, dependiendo del número de veces que los oligonucleótidos se encuentran repetidos, se observarán bandas de diferente tamaño.

Por medio del análisis de enlace (llamado también análisis de "ligamiento") se puede saber si dos diferentes marcadores genéticos están cercanos entre sí, así como la distancia aproximada a la que se encuentran. Mientras mayor sea la cercanía entre ellos en una región cromosómica dada, con más frecuencia se heredarán juntos (*co-segregación*). Esta cosegregación podrá ser observable únicamente si los individuos estudiados son heterocigotos para los marcadores en cuestión. Pero algunas veces, aun si están muy cerca uno del otro, se heredarán separadamente; en este caso el hijo que reciba uno no recibirá el otro, aunque en el cromosoma del progenitor hayan estado juntos. Esto se deberá a que ha ocurrido un suceso de recombinación genética, mediado por el mecanismo cromosómico de entrecruzamiento (*crossing over*). Naturalmente, mientras más alejados estén entre sí, más frecuentemente recombinarán, lo que hace posible estimar la distancia entre ellos en términos de esta frecuencia de recombinación. La unidad de medida de distancias genéticas es el "morgan" (en honor de Thomas H. Morgan, uno de los pioneros de la genética), y la que es más común que se utilice es el centimorgan (cM), que corresponde aproximadamente a un millón de pares de bases. Así se logró construir el mapa genético humano que consta de más de 5 mil marcadores, la mayor parte de los cuales es microsatélite. Este mapa tiene ya la densidad necesaria de marcadores para ubicar, con una precisión de aproximadamente 2 cM, cualquier gen en el que se esté interesado.

Estrategias para la búsqueda de genes relacionados a enfermedades

La biología humana y la medicina cuentan ahora con una herramienta extraordinariamente poderosa para la mejor comprensión de los mecanismos responsables de procesos biológicos y de enfermedades, por medio del descubrimiento, aislamiento y caracterización de sus genes. Esta estrategia se conoce con el nombre de donación *posicional* o, más coloquialmente, como "mapeo

genético". El contar con un mapa genético con un número suficiente de marcadores permite efectuar un escrutinio total del genoma (*whole genome scan*) y consiste en buscar, a lo largo de todos los 22 autosomas y del cromosoma X, aquellas regiones con marcadores que cosegreden con el o los genes de la enfermedad en cuestión. En la práctica, no es necesario usar los 5 mil marcadores; basta utilizar un panel de 300 a 400 de ellos para hacer un tamiz inicial, que identifique regiones promisorias, con alta probabilidad de contener a los genes responsables. En un segundo escrutinio, de mayor resolución, se usan nuevos conjuntos de marcadores localizados en cada una de esas regiones. En esta forma se está logrando ubicar a genes no conocidos con anterioridad, responsables de un número creciente de enfermedades, como la diabetes mellitus insulino dependiente y la malformación cavernosa del cerebro.

Sin embargo, la construcción de mapas genéticos y su utilización para el descubrimiento de nuevos genes es apenas el primer paso de esta nueva forma de acceder al conocimiento biológico, por medio de la Genómica. La siguiente etapa en el proceso de identificación de los genes requiere la utilización de mapas físicos. Un mapa físico consiste en tener a todo el genoma en fragmentos relativamente pequeños, ordenados entre sí, cuya ubicación precisa en su cromosoma correspondiente se conoce. Hay diferentes formas de construir mapas físicos.

Una es por clonación de cada uno de estos fragmentos, que se obtienen cortando el DNA total por medio de endonucleasas de restricción. Después se introducen, por ejemplo, en células bacterianas o de levaduras. Cada célula original, que después dará origen a una colonia, tendrá un diferente fragmento del genoma. Se tendrán tantas colonias bacterianas o de levaduras, como fragmentos del genoma humano se hayan tenido inicialmente, cada colonia conteniendo copia (clon) de un determinado fragmento. En esta forma, el conjunto total de colonias contendrá físicamente a todo el genoma humano; estos conjuntos también se denominan "bibliotecas genómicas".

Después es necesario ordenar los diferentes fragmentos entre sí, de tal manera que su orden corresponda al que originalmente tenían en cada uno de los cromosomas, esto es, en el genoma original. Esto se logra utilizando marcadores

genéticos para yuxtaponer a los fragmentos entre sí. Hay otra forma que puede tomar un mapa físico, sin necesidad de tener una de estas bibliotecas en forma concreta sino solamente virtual. En esta modalidad, los fragmentos que la integran son definidos por medio de secuencias conocidas de DNA, de una longitud corta, que se conocen como *STSs* (*sequence-tagged sites* o sitios marcados con secuencias).

Un mapa físico brinda el sustrato para obtener la secuencia de bases nitrogenadas del fragmento del DNA en donde se encuentra el gen buscado. Como estos fragmentos son muy grandes, de decenas o centenas de miles de bases nitrogenadas, todavía hay una etapa difícil y prolongada antes de la identificación inequívoca de la secuencia que corresponde al gen en cuestión.

Además del mapa genético, del mapa físico y del de secuencias, habrá también mapas de genes y mapas de genes expresados (cDNAs), los que variarán de un tejido a otro y entre diferentes etapas de la vida. Hay además el llamado mapa mórbido, que indica la posición en el genoma, de los genes responsables de las diferentes enfermedades.

El proyecto universitario del Genoma de la UNAM

México en general y la UNAM en particular habían carecido de una política científica relativa a la investigación genómica. Desde junio de 1994, un grupo de universitarios nos hemos dado a la tarea de explorar las posibilidades y las formas de participar en este importante proyecto científico internacional y hemos desarrollado una intensa labor de planificación, con el apoyo de las autoridades universitarias. En esta forma, se constituyó un consorcio de investigadores (Cuadro 1) pertenecientes a siete dependencias universitarias (Facultad de Medicina, cuatro institutos de la investigación científica, uno de humanidades y una Dirección General), que culminará en la creación del Proyecto Universitario del Genoma Humano (PUGH). Se han alcanzado los siguientes consensos. México puede y debe participar en el Proyecto Internacional del Genoma Humano y la UNAM es una institución idónea para incubar un proyecto

nacional, estableciendo vínculos con otras instituciones nacionales de salud y de educación superior. El Proyecto del Genoma Humano tiene dos grandes vertientes. La primera, que es sobre la que principalmente se ha difundido información, consiste, como se describió en la sección anterior, en descubrir el texto completo del genoma humano; esta parte del proyecto está siendo ejecutada por algunos de los países rícorde y en ella no podemos ni queremos participar. La otra vertiente, que es en realidad la más interesante e importante, consiste en "traducir y leer" este texto; tener la capacidad de encontrar en él la información que requiramos. Es en esta vertiente en la que los mexicanos podemos y debemos participar. Para ello es necesario estudiar a muchísimas familias informativas, las que se encuentran distribuidas en todo el mundo y no ubicadas sólo en los países ricos. En particular, México tiene extraordinarias oportunidades en sus poblaciones geográficamente aisladas y altamente consanguíneas, que son genéticamente muy homogéneas, además de que ello permite más fácilmente la identificación y la caracterización de genes y de sus mutaciones.

El objetivo general del PUGH es adquirir la capacidad ("saber hacer") conceptual y metodológica, para contribuir a la comprensión de una enfermedad o de un proceso biológico por medio de la identificación y caracterización de los genes participantes, en la forma que se describió con

anterioridad. Esto conducirá al análisis mutacional, estudio de su distribución en los diferentes grupos poblacionales de nuestro país, epidemiología molecular y desarrollo de procedimientos de diagnóstico y de terapia génica. Dadas las enormes implicaciones éticas, legales, sociales y políticas de la investigación del genoma humano y de las aplicaciones derivadas de su estudio, su estudio tendrá una alta prioridad dentro del PUGH.

De Inicio, el PUGH tendrá dos componentes: 1). Un sistema de información y comunicación entre los miembros del consorcio y con la comunidad en general, a través de Internet, que incluye una "página" en la World Wide Web y que conste de una base de datos con el inventario de todos los proyectos de genética y biología molecular que se llevan a cabo en la UNAM y que contiene información sobre los diferentes investigadores, su ubicación, los proyectos de investigación que se están realizando y las diferentes metodologías con las que se cuentan (<http://tztetza1.dcaa.unam.mx/genoma.html>). 2) Una infraestructura consistente en unidades centralizadas de apoyo metodológico de alta complejidad y costo ("laboratorios universitarios", equivalentes a las "core facilities" de países como Estados Unidos): un repositorio de líneas celulares y de DNA, con una base de datos genealógica; una unidad de recursos genómicos, así como una unidad de procesamiento y análisis matemático de datos.

Cuadro I. Consorcio del Proyecto Universitario del Genoma Humano

Larissa Adier	Instituto de Matemáticas Aplicadas y Sistemas
Rogelio Alonso	Facultad de Medicina
Luis Covarrubias	Instituto de Biotecnología
Gerardo Gamba	Instituto de Investigaciones Biomédicas
Alejandro García Carrancá	Instituto de Investigaciones Biomédicas
Diego González Halphen	Instituto de Fisiología Celular
Susana Kofman	Facultad de Medicina
Jaime Mas Oliva	Instituto de Fisiología Celular
Marcia Muñoz de Alba	Instituto de Investigaciones Jurídicas
Laura Riba	Instituto de Investigaciones Biomédicas
Xavier Soberón	Instituto de Biotecnología
Gerardo Vega	Dirección Gral. de Servicios de Cómputo Académico
Ma. Teresa Tusíé	Instituto de Investigaciones Biomédicas e Instituto Nacional de Pediatría
Antonio Velázquez	Instituto de Investigaciones Biomédicas e Instituto Nacional de Pediatría

Con objeto de no dispersar esfuerzos, se ha considerado iniciar este Proyecto Universitario con una investigación genómica multicolaborativa sobre una enfermedad que, por su importancia en el campo de la salud pública en México, por sus peculiaridades en la población mexicana y por la competitividad a nivel internacional que nos dé el contar con un gran número de familias multigeneracionales con individuos afectados, constituya un modelo adecuado para adquirir la experiencia de trabajar colaborativamente, desde múltiples dependencias e incluso instituciones, en la nueva disciplina de la "Genómica" y, al mismo tiempo, nos confiera una alta probabilidad de hacer una contribución científica de primer orden. La enfermedad seleccionada después de este intenso proceso de planeación, fue la diabetes mellitus no insulino-dependiente. En su determinación existe un componente genético muy importante, pero este es complejo y variable de distintas familias. Aún no se han descubierto muchos de los genes que con seguridad intervienen en su determinación, pero los datos que se tienen hacen suponer que los genes que con más frecuencia participan en la producción de diabetes en pacientes mexicanos, son diferentes a los que muestran una mayor ingerencia en la diabetes de otros diferentes grupos étnicos.

Referencias

- ASHG Committee. Statement of the ASHG on genetic testing for breast and ovarian cancer predisposition. *Am J Hum Genet.* 1994;55:i-iv.
- Beardsley T. Vital data. *Sci Amer.* 1996;March:7641.
- Dib C, Fauré S, Fizames C, et al. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature.* 1996;380:152-154.

- Clayton EW et al. Informed consent for genetic research on stored tissue samples. *J Am Med Assoc.* 1995;274:1786-1792.
- Collins F, Galas D. A New Five-Year Plan for the U.S. Human Genome Project. *Sci.* 1993;262:4346.
- Collins FS. BRCA1-Lots of mutations, lots of dilemmas. *New Eng J Med.* 1996;334:186-188.
- Cox David R et al. Assessing Mapping Progress in the Human Genome Project. *Sci.* 1994;265:2031-2032.
- Davies J L et al. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature.* 1994;371:130-136.
- Dubovsky J et al. Sets of short tandem repeat polymorphisms for efficient linkage screening of the human genome. *Hum Mol Genet.* 1994;4:449-452.
- Gunel M et al. A founder mutation as cause of cerebral cavernous malformation in Hispanic Americans. *New Eng J Med.* 1996;334:946-951.
- Housman D. Human DNA Polymorphism. *New Eng. J. Med.* 1995;332:318-320.
- Hudson K L et al. Genetic Discrimination and Health Insurance: An Urgent Need for Reform. *Sci.* 1995;270:391-393.
- Hudson T J et al. An STS-based map of the human genome. *Sci.* 1995;270:1945-1954.
- Knoppers B, Chadwick R. The human genome project: under an international ethical microscope. *Sci.* 1994;265:2035-2036.
- Levy-Lahad E et al. A Familial Alzheimer's Disease Locus on Chromosome 1. *Sci.* 1995;269:970-977.
- Moore A H. Genetics: the money rush is on Fortune. *Magazine* May 30. 1994;60-67.
- Motulsky A G. Predictive genetic diagnosis. *Am J Human Genet.* 1994;55:603-605.
- Murray J C et al. A Comprehensive Human Linkage Map with Centimorgan Density. *Sci.* 1994;265:2049-2054.
- Olson M V et al. A time to Sequence. *Sci.* 1995;270:394-396.
- Paterson A H et al. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 1988;335:721-726.
- Tugendreich S et al. Genes conserved in yeast and humans. *Hum Mol Genet.* 1994;3:1509-1517.
- Reiman E M et al. Preclinical evidence of Alzheimer's disease in persons homozygous for the E4 allele for apolipoprotein E. *New Eng J Med.* 1996;334:752-758.