

Introducción a los estudios genómicos en insuficiencia cardíaca

Sergio D. Sevilla*

Glosario

Alelo: formas alternativas de un gen.

Codón: secuencia de tres bases de DNA o RNA que especifica un aminoácido simple.

Genotipo: composición genética de un individuo que se refleja en su secuencia de ADN.

Haplotipo: grupo de alelos cercanos que se heredan en forma conjunta.

Heterozygota: que porta dos alelos diferentes de un locus cromosómico determinado.

Homozygota: que porta dos alelos idénticos en un locus cromosómico determinado.

Linkage disequilibrium: asociación, no azarosa de alelos cercanos.

Locus: ubicación cromosómica de un gen o un marcador.

Fenotipo: presentación clínica o expresión de un gen, genes, factores ambientales o ambos.

Polimorfismo: locus que está representado por un número diferente de alelos o haplotipos en la población.

En el año 2003, la revista *New England Journal of Medicine* publicó una editorial titulada "Bienvenidos a la Era Genómica". La fecha de comienzo de esta "era genómica" fue determinada el 14 de abril de aquel año, momento en el cual el esfuerzo internacional, conocido como *Human Genome Project* puso un cierre a la era *pre-genómica* con la conclusión del titánico objetivo de completar la secuencia completa del genoma humano¹. Desde entonces, la disciplina genómica (en su más amplio sentido) está avanzando sobre muchos ámbitos de la medicina, desde el conocimiento de nuevos patógenos, uso de patrones de expresión genética para guiar tratamientos, el conocimiento de nuevos factores genéticos de riesgo, la tipificación de riesgo de pacientes, el conocimiento de la respuesta a ciertas drogas hasta el diseño y desarrollo de drogas nuevas.

Aunque ya llevamos 4 años de esta declaración, es probable que muchos de los médicos dedicados a la actividad asistencial no tengan todavía una idea clara del impacto que está teniendo este cambio en el conocimiento biomédico y hasta en la práctica clínica cotidiana. Para muchos de nosotros (formados antes de los años '90), la *genética* como disciplina, parece sólo interesar a un grupo pequeño de cole-

gas que lidian con enfermedades raras.

Hemos escuchado muchas veces que "la *genética* es el futuro" o que "se viene la *genómica*"; pero hasta cierto punto, sospecho que aun no entendemos bien que significa esto.

Aunque, la *genómica* se presenta como un campo de conocimiento inabarcable, es interesante revisar algunos artículos que nos permitan acercarnos un poco a los estudios basados en esta disciplina que se hayan hecho en insuficiencia cardíaca.

En los próximos párrafos, se resumirán algunos conceptos básicos y definiciones.

Introducción a la genómica

Los genes son fragmentos de ADN que contienen información biológica y por lo tanto codifican ARN y/o una molécula polipeptídica. La genética tradicional se define como: el estudio de genes simples y su efecto. En cambio, el término *genómica*, acuñado hace alrededor de 20 años atrás, implica no sólo el estudio de genes simples; sino también, su función y la interacción de todos los genes en el genoma. Tiene la ambición de alcances mucho más amplios que la genética, descansa sobre el acceso experimental al genoma completo y se aplica a enfermedades comunes como el cáncer de mama, la enfermedad por HIV, la enfermedad de Alzheimer y la enfermedad coronaria, entre otras.

Estas enfermedades son causadas por la interacción de múltiples genes y factores ambientales y por lo tanto se consideran desordenes multifactoriales. Las variaciones genéticas en este tipo de enfermedades pueden tener un rol protector o patogénico en la expresión de la enfermedad.

Como sabemos, el componente genómico de las enfermedades comunes es bastante diferente a lo que ocurre en las conocidas como condiciones monogénicas o mendelianas. En estas últimas, la alteración de un solo gen se transmite en forma consonante con las leyes de la herencia descritas en el siglo XIX por Mendel. Actualmente, existe un compendio *online* de enfermedades mendelianas clasificadas (OMIM-Online *Mendelian Inheritance in Man*²). Este catálogo incluye las patologías heredadas por ADN mitocondrial y otros fenómenos más complejos.

La mayoría de las condiciones monogénicas son infrecuentes. Aun las más comunes como la hemocromatosis hereditaria, la fibrosis quística, el déficit de alfa-1 antitripsina o la neurofibromatosis afectan a un número total reducido de

*Médico Internista y Farmacólogo (UBA)
Exchange Research Scientist-Innate Immunity genomics.
National Institute of Environmental Health Science (NIH), USA.

Correspondencia: Dr. Sergio D. Sevilla
E-mail: sergiodsevilla@yahoo.com.ar

Trabajo recibido: 03/04/2007
Trabajo aprobado: 20/5/2007

pacientes. Sin embargo, la comprensión de la forma y mecanismo con los cuáles actúan estos trastornos genéticos, ha ayudado a comprender los procesos fisiopatológicos básicos que subyacen a enfermedades más comunes tales como la aterosclerosis y la acción de las estatinas³.

El Proyecto *Genoma Humano*^{4,5}

La característica crítica de la molécula de ADN es su secuencia de nucleótidos. Si la secuencia es conocida, entonces los genes que contiene pueden ser identificados y las actividades de aquellos genes pueden ser estudiadas en detalle.

Desde mediados de los años '70, los biólogos moleculares han logrado obtener secuencias de porciones de ADN cada vez más largas, completando en los '90 la primera secuencia de genoma completo de un organismo vivo. Sin duda, el proyecto más relevante en este aspecto es el dedicado al genoma humano.

El Proyecto *Genoma Humano* fue concebido en 1984 y comenzó a arrojar resultados en el año 1990. El objetivo primario de este esfuerzo era determinar la secuencia completa de nucleótidos del genoma nuclear humano. El proyecto ha sido solventado por gobiernos y otras instituciones de varios sitios del mundo y ha sido el plan colaborativo internacional más complejo jamás intentado en cualquier área de la ciencia. Por su parte, *Celera Genomics, Inc.* (Maryland, USA) lanzó su propio secuenciado de genoma en 1998. Ambos proyectos lograron completar un borrador de la secuencia en el año 2001 (se representaba sólo el 83-84% de la secuencia total) y los resultados fueron publicados en las revistas *Nature* y *Science* en Febrero de ese mismo año^{6,7}. Completar la secuencia del genoma humano requirió el mejoramiento permanente de las técnicas de secuenciamiento y aunque fue declarado como terminado en 2003, no fue hasta mayo de 2006, cuando se anunció la terminación de la secuencia final y completa de los 24 cromosomas humanos. La corrección de errores continuará en los años por venir. Para comprender la magnitud de este logro, hay que entender que secuenciar ADN significa determinar el orden exacto de 3 billones de bases, comúnmente abreviadas como A (adenina), T (timidina), C (citosina) y G (guanina) que componen el ADN de los 24 cromosomas humanos. Esto permitió estimar la existencia de 20.000 a 25.000 genes dentro del ADN.

Sabemos, actualmente, que menos del 2% del genoma humano codifica proteínas. El resto está compuesto de secuencias de distinto tipo. Las secuencias repetitivas no-codificantes (a veces subestimadas erróneamente) constituyen un archivo histórico de la evolución biológica humana, proveen una rica fuente de información sobre genética poblacional y médica y además son agentes regulatorios activos dentro del genoma.

Los genes se distribuyen en forma despareja dentro el genoma. Particularmente, algunos cromosomas, como el 17, 19 y 22, tienen una alta densidad de los mismos, en comparación con otros como el 4, 8, 13, 18 e Y. Así mismo, la densidad de genes varía dentro de un mismo cromosoma y

esto se encuentra asociado a las zonas ricas en C y G en comparación con A y T. También, sabemos que no todos los genes residen dentro de los cromosomas nucleares. Varias docenas de estos, relacionados al metabolismo energético, se encuentran dentro de las mitocondrias.

Actualmente, sólo una porción superior al 50% de los genes identificados tiene motivos (*motif*) conocidos o secuencias de ADN reconocidas que sugieran una posible función. Se cree además, que, gracias al mecanismo conocido como *splicing* alternativo, se puedan derivar más de 100.000 proteínas de los 20 a 25.000 genes supuestos.

Otros fenómenos como la metilación (inclusión de grupos químicos metilo) y la modificación de histonas (proteínas que se asocian al ADN) también pueden alterar el efecto de un gen. Estos se conocen como fenómenos epigenéticos y constituyen un complejo arreglo de señales moleculares que hace a genes específicos encontrarse en estado "on" u "off" y por lo tanto expresarse o no. En la actualidad, existen varias líneas de investigación intentando determinar si la interacción con factores del ambiente (exposición a ciertas drogas, etc.) pueden producir cambios en la expresión génica, que impliquen cambios fenotípicos en las generaciones inmediatas. Por ejemplo, que la exposición de las madres embarazadas a sustancias tóxicas o farmacológicas produzcan modificaciones epigenéticas en sus hijos que los hagan propensos a enfermedades comunes como el asma o la alergia^{8,9}.

Polimorfismos y mutaciones en el ADN

A pesar de que se refiere el genoma humano como una secuencia única de nucleótidos, en realidad existen tantas secuencias posibles como humanos (salvo en el caso de los gemelos idénticos). Las diferencias entre individuos se deben mayormente a mutaciones o alternativas de una sola base en la cadena (ej.: A por T). Estas mutaciones se conocen como SNPs (se pronuncia *snips*) y el nombre proviene del idioma inglés: *Single Nucleotide Polimorphism*¹⁰. Son, sin duda, las más comunes y en cuanto al efecto que producen, las hay de varios tipos. Las *missense mutations* (cambio de sentido) son las que producen la transcripción de un aminoácido alternativo, debido al cambio que se produce en la secuencia de 3 bases o codón que codifica tal aminoácido. Las *nonsense mutations* (sin sentido) son más deletéreas, ya que cambian el codón de transcripción por un *stop codon*; el cual causa la terminación de la transcripción de la proteína, en vez de un nuevo aminoácido. Otro tipo de mutación es la *frameshift mutation* (cambio de marco), la cual cambia el marco de lectura del gen y lleva generalmente a un *stop codon* prematuro.

En términos del efecto funcional, muchas mutaciones no tienen consecuencias en el fenotipo del individuo. Las conocidas como *silent mutations* (silenciosas) producen el reemplazo de una base por otra que resulta en un codón que codifica el mismo aminoácido. Así mismo, las mutaciones pueden no cambiar el fenotipo si el cambio de aminoácido en la proteína no genera cambios significativos en la función de la misma (*conservative mutations*). Sin embargo, las conocidas como *non-conservative mutations* reemplazan un aminoácido por uno muy

diferente y es más probable que afecten el fenotipo.

Las mutaciones pueden causar enfermedad por varios mecanismos. El más frecuente es la pérdida de función. En este caso, la mutación puede cambiar el fenotipo por disminución de la actividad funcional de la proteína que codifica dicho gen. Sin embargo, en otros casos, algunas mutaciones pueden causar enfermedad a través de la ganancia de función, cuando la proteína adquiere una función tóxica (ej.: enfermedad de Huntington¹¹).

Se podría suponer que las mutaciones que se encuentran en el 98,5% del genoma no-codificante, no deberían afectar el fenotipo. En efecto, muchas de estas mutaciones no tienen ningún impacto en el fenotipo; pero otras que se sitúan en sitios regulatorios, pueden ser importantes en la etiología de las enfermedades. Las mutaciones regulatorias actúan alterando la expresión de un gen, llevando a la inhibición de la expresión o a la expresión inesperada en un tejido donde debería ser silente, o finalmente a un cambio en el *timing* de expresión.

Algunas mutaciones, también, pueden reducir el riesgo de enfermedad, como la que ocurre en el gen del receptor de quimocinas CCR5 y que hace a los portadores homocigotas casi completamente resistentes a la infección por HIV y a los heterocigotos lentamente progresivos¹².

A pesar de que se estima que existen 10 millones de SNPs en el DNA, aun sólo se han identificado algo más de 1,4 millones. En promedio, uno cada 2,0 Kb (2000 bases) de secuencia. Alrededor de 60.000 SNPs se encuentran dentro de los genes y al menos algunos de éstos tienen impacto en la actividad de ese gen. Además, se estima que cada 2 Kb de secuencia existe un microsatélite (también llamado *short tandem repeat*) que está formado por una serie de nucleótidos repetitivos (CACACA-CACA) en la cual el número de repeticiones es variable para cada individuo.

Variación en el genoma humano

En promedio, dos personas no-relacionadas comparten 99,9% de su ADN. Sin embargo, dado que el ADN humano está compuesto por 3 billones de pares de bases, las secuencias de dos personas no-relacionadas varían en millones de bases¹⁰.

Se han hecho grandes esfuerzos en determinar y catalogar esas variantes (polimorfismos: SNPs, microsatélites, etc.) y correlacionar estas variaciones genotípicas con fenotipos determinados. En muchos casos, la correlación SNP-fenotipo ocurre como una consecuencia directa de esa mutación en la transcripción del DNA (ver arriba). Sin embargo, en la mayoría de los casos, dicho SNP es simplemente un marcador de biodiversidad que se encuentra cercano a un sitio que es el verdadero factor genético de la enfermedad.

El concepto de proximidad es sólo una medida grosera de cercanía física, en cambio connota que los eventos de recombinación ocurridos entre el sitio del SNP y el sitio preciso del factor genético han ocurrido sólo raramente durante la evolución.

La recombinación es el fenómeno de intercambio de un segmento de ADN entre cromosomas homólogos, durante la meiosis de las células eucariotas. De aquí surge el concepto de que dicho SNP y el sitio genético real de efecto se encuentra en *linkage disequilibrium*; es decir que se heredan juntos en un bloque. Conocer el polimorfismo nos permite ubicar aproximadamente el gen responsable de la variación fenotípica.

Como una extensión del esfuerzo de mapear y relacionar los SNPs con enfermedades, surge el concepto de haplotipo. Mientras, el SNP representa una variación puntual, el haplotipo representa una secuencia de nucleótidos considerablemente más larga (alrededor de 25.000 nucleótidos), que como vimos antes, tiende a heredarse en bloque con el SNP.

Actualmente, existe un proyecto colaborativo conocido como *HapMap*. Se trata de un catálogo de variantes genéticas comunes que ocurren en los seres humanos. Describe cuáles son las variantes, dónde ocurren en el ADN y cómo se distribuyen entre personas dentro de poblaciones y entre poblaciones en diferentes partes del mundo. Este proyecto internacional provee información para que otros científicos puedan hacer estudios de asociación entre enfermedades y variantes genéticas.

En la mayoría de los cromosomas, se encuentra sólo una pequeña cantidad de haplotipos. En una determinada población, el 55% de las personas puede tener una versión de haplotipo, el 30% puede tener otra, el 8% una tercera y el resto variantes menos comunes. El *HapMap Project* está tratando de identificar estos haplotipos comunes en 4 poblaciones de diferentes partes del mundo. Además, el proyecto está identificando *tag-SNPs* (SNPs etiqueta) que permitan identificar en forma única el haplotipo completo. Genotipificando un *tag-SNP*, los investigadores serán capaces de identificar la colección de haplotipos del DNA de una persona. Se estima que unos 300000 a 600000 SNPs podrían aportar información de los patrones de variación genética; lo cual es significativamente menos que los 10 millones de SNPs supuestos¹³.

Referencias bibliográficas

1. Guttmacher A et al. N Engl J Med 2003; 349:996-998, Sep 4, 2003.
2. Online Mendelian Inheritance in Man, OMIM (TM). Baltimore: McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University, 2000.
3. Goldstein JL. The metabolic & molecular bases of inherited disease. 8th ed. Vol. 2. New York: McGraw-Hill, 2001:2863-913.
4. Genomes Brown, T.A. New York and London: Garland Science; c2002. Online NCBI version.
5. Human Genome Project Information. http://www.ornl.gov/sci/techresources/Human_Genome/project/about.shtml
6. Lander ES et al. Nature 2001;409:860-921.
7. Venter JC et al. Science 2001;291:1304-1351.
8. Bousquet J et al. Allergy 2004 Feb;59(2):138-47.
9. Vercelli D et al. J Allergy Clin Immunol 2004 Mar;113(3):381-6; quiz 387.
10. Guttmacher A et al. N Engl J Med 2002; 347:1512-1520, Nov 7, 2002.
11. Goldberg YP et al. Nat Genet 1996;13:442-449.
12. Dean M et al. Science 1996;273:1856-1862.
13. International HapMap Project <http://www.hapmap.org/-Online>.