

Estudios de asociación de genoma completo: no hay que juzgar un libro por la portada

Robert Roberts

Ruddy Canadian Cardiovascular Genetics Centre. Instituto del Corazón de la Universidad de Ottawa.
Ottawa. Ontario. Canadá.

La era de los estudios de asociación de genoma completo (*genome-wide association studies* [GWAS]) ya está aquí, y en el corto lapso de 5 años su prolífico rendimiento ha sido poco menos que extraordinario. Los GWAS serán probablemente uno de los hitos de la ciencia en el siglo XXI, tal como lo fue el Proyecto Genoma Humano en el siglo XX. El diseño de los GWAS proporciona un enfoque no sesgado para ubicar en los mapas cromosómicos las variantes genéticas de riesgo frecuentes^{1,2}. Una variante de riesgo frecuente se define arbitrariamente como la que tiene una frecuencia igual o superior al 5% en la población general³. Sin embargo, el límite inferior de la sensibilidad para la mayoría de los GWAS publicados hasta la fecha es de una frecuencia de tan sólo el 10%⁴. No se prevé que las variantes muy poco comunes (frecuencia $\leq 5\%$) se detecten mediante los GWAS, a menos que el tamaño muestral sea excepcionalmente grande. Un enfoque sistemático para detectar variantes muy poco comunes probablemente requerirá la secuenciación directa del ADN.

Una característica importante de los GWAS es la formación de grandes consorcios de investigación internacionales en torno a varias enfermedades. El más amplio de ellos, el Coronary Artery Disease Genome-wide Replication and Meta-Analysis (CARDIoGRAM)⁵, dedicado a la enfermedad coronaria (EC), ha determinado ya el fenotipo y el genotipo de una población de descubrimiento de 82.000 individuos y una población de replicación de resultados de más de 40.000. Estos consorcios reúnen los tamaños muestrales elevados necesarios no sólo para localizar en el mapa cromosómico las va-

riantes de riesgo, sino también para integrar y enriquecer las competencias y los recursos precisos para interpretar y analizar los datos. Como resultado de la actividad de estos grandes consorcios, se espera que la mayor parte de las variantes genéticas que predisponen a enfermedades frecuentes como la EC y el cáncer puedan ubicarse en el mapa cromosómico en un futuro próximo. A pesar del éxito obtenido por los GWAS en la detección de más de 400 *loci* asociados a enfermedades⁶, hay cierto escepticismo respecto a si los GWAS facilitarán un avance en el manejo de las enfermedades^{7,8}. Los mecanismos por los que se espera que las variantes de riesgo genético mejoren el manejo de estas enfermedades serían la detección genética sistemática con fines de diagnóstico y prevención, la farmacogenética para proporcionar el fármaco adecuado a la persona adecuada y en la dosis adecuada y, finalmente, su uso como dianas para el desarrollo de nuevos fármacos más apropiados. Debemos recordar también que puede haber tesoros ocultos que surjan de la mejora del conocimiento de la base molecular de la enfermedad, aparte de los ya esperados. Un ejemplo de ello podría ser el desarrollo de vacunas a menudo facilitado por el conocimiento de la secuencia del ADN.

La preocupación por las limitaciones de los GWAS continuará siendo legítima mientras no quede demostrado que su aplicación aporta efectos beneficiosos. Esta preocupación se intensifica con la observación de que, a pesar de que se han localizado en el mapa cromosómico muchas variantes de riesgo frecuentes para enfermedades como la EC, su riesgo relativo acumulativo no explica la mayor parte de la predisposición genética esperada^{7,9,10}. Según los datos actuales, el riesgo absoluto debido a esas variantes probablemente explique sólo un 20-30% del riesgo esperado. ¿A qué se deben los riesgos restantes? Ciertamente, la respuesta no es obvia en este momento, pero debemos recordar que han pasado menos de 5 años desde el inicio de la técnica de cartografía de *loci* frecuentes y que, por lo tanto, cabe prever que haya muchas más variantes por descubrir. Otra posibilidad es que haya

VÉASE ARTÍCULO EN PÁGS. 925-33

Correspondencia: Dr. R. Roberts.
President & Chief Executive Officer. Director.
Ruddy Canadian Cardiovascular Genetics Centre. University of Ottawa
Heart Institute.
40, Ruskin Street. Ottawa. ON K1Y 4W7. Canadá.
Correo electrónico: rroberts@ottawaheart.ca

Full English text available from: www.revespcardiol.org

variantes muy poco comunes¹¹, de las que esperaríamos un efecto muy superior, que aún están por identificar, puesto que los GWAS no disponen de la sensibilidad necesaria para detectar esas variantes raras. ¿Se comportan las variantes de riesgo de manera lineal o son interactivas entre sí y producen un efecto superior al de la simple suma del riesgo de cada variante individual? En este número de REVISTA ESPAÑOLA DE CARDIOLOGÍA, un estudio de Lluís-Ganella et al¹² intenta abordar esta cuestión. En dicho estudio, los investigadores llevaron a cabo un análisis *in silico* de una muestra de 7.368 casos de una base de datos del Wellcome Trust Case Control Consortium, formada por 1.988 casos de EC y 5.380 controles. A partir de los datos publicados en la literatura médica, se seleccionaron nueve variantes con una asociación probada con el aumento de riesgo de EC. Dado que cada individuo posee dos copias de cada variante, puede ser heterocigoto u homocigoto o no tener ninguno de los dos alelos que sea de riesgo. Así pues, el número total de posibles alelos de riesgo en un determinado individuo sería 18. Las nueve variantes seleccionadas manifiestan su riesgo de EC de manera independiente de los factores de riesgo clásicos, como colesterol, diabetes mellitus o hipertensión. Los investigadores presentan un análisis original e interesante que pone de relieve que el número de variantes de riesgo por individuo oscilaba entre 1 y 13, con una mediana de 7 para el número de variantes en un individuo. Los resultados de este análisis indican claramente que el riesgo es mayor cuanto mayor es el número de alelos de riesgo. El riesgo era aditivo, con una relación de tipo lineal. Dicha relación lineal descarta toda interacción entre genes e indica que el riesgo acumulativo es simplemente el total del riesgo que comporta cada una de las variantes individuales. Si estos resultados se confirman en otras poblaciones y para otros genes, se descartaría las interacciones entre genes, lo cual sería muy sorprendente. Se prevé que las interacciones entre genes produzcan, a su vez, un riesgo superior junto con una mayor respuesta ambiental. Sin embargo, es prematuro descartar la interacción entre genes sobre la base de este estudio solamente. La investigación no intentó determinar si la aplicación de la puntuación genética permitiría mejorar la estratificación del riesgo respecto a lo obtenido con los factores convencionales. Los investigadores recomiendan vivamente el estudio de una cohorte de población prospectiva como método más apropiado para intentar responder a esta cuestión.

Aunque la preocupación existente respecto a la aplicación y los beneficios aportados por la información sobre las variantes de riesgo genéticas es legítima y apropiada, más que impedir, debiera acelerar la búsqueda de más variantes frecuentes y

raras. La evidencia relativa a la existencia de una predisposición genética a la EC¹³ como la que se da para otras enfermedades procede fundamentalmente de estudios epidemiológicos de cohortes de población y de estudios familiares. Ahora disponemos de una prueba independiente y aditiva de esa predisposición genética. Hay centenares de *loci* cromosómicos para los que se ha demostrado una asociación con enfermedades mediante la aplicación de GWAS. En segundo lugar, tal como se había predicho, la predisposición genética a esas enfermedades es poligénica, de tal manera que cada alelo contribuye a producir el riesgo tan sólo mínima o moderadamente. Debemos recordar que no hay que juzgar un libro únicamente por la portada. Los *loci* que predisponen a la EC y otras enfermedades se han localizado en el mapa cromosómico a partir de la asociación entre los marcadores de polimorfismos de nucleótido único (SNP) situados en los *loci* y el fenotipo de la enfermedad. La mayor parte de los SNP son simplemente marcadores en un desequilibrio de ligamiento con el SNP causal real. Los *loci* actualmente localizados mediante la asociación de casos y controles (GWAS) tienen una longitud que oscila entre 50.000 y 100.000 pb. Sin embargo, sabemos que para la mayoría de los *loci* hay un único SNP que tiene un papel causal de la predisposición genética. En segundo lugar, la mayor parte de los *loci* no se encuentran en regiones de codificación de proteínas, sino en intrones y regiones promotoras^{8,9}. Además, no es infrecuente que las regiones incluyan moléculas de ARN no codificadoras. Un ejemplo de ello es la variante de riesgo 9p21^{14,15} que es un ARN no codificador largo, al que se denomina ANRIL¹⁶. Para evaluar y aplicar científica o clínicamente esta información, tendremos que identificar la secuencia causal y determinar su función. Parece razonable suponer que, aunque las pruebas genéticas de detección sistemática para la prevención vayan a ser útiles, posiblemente podría obtenerse un mayor beneficio con el empleo de las secuencias de las variantes como dianas para nuevos tratamientos.

BIBLIOGRAFÍA

1. Wang WY, Barratt B, Clayton DG, Todd JA. Genome-Wide Association Studies: Theoretical and practical concerns. *Nat Rev Genet.* 2005;6:109-18.
2. Hirshhorn JN, Daly MJ. Genome-wide Association Studies for common diseases and complex traits. *Nat Rev Genet.* 2005;6:95-108.
3. Hahn LW, Ritchie MD, Moore JH. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. *Bioinformatics.* 2003;19:376-82.
4. Dandona S, Stewart AFR, Roberts R. Genomics in coronary artery disease: Past, present and future. *Can J Cardiol.* 2010;26 Suppl A:A56-9.

5. CARDIoGRAM consortium. Design of the Coronary Artery Disease Genome-wide Replication and Meta Analysis (CARDIoGRAM) Consortium —a prospective meta-analysis of 13 genome-wide association studies. *Circ Cardiovas Genet*. 2009 [en prensa].
6. Donnelly P. Progress and challenges in genome-wide association studies in humans. *Nature*. 2008;458:728-31.
7. Goldstein DB. Common genetic variation and human traits. *New Engl J Med*. 2009;360:1699-701.
8. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindroff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461:747-53.
9. McClelland JH, King M-C. Genetic Heterogeneity in Human Disease. *Cell*. 2010;141:210-7.
10. Paynter NP, Chasman DI, Pare G, Buring J, Cook NR, Miletich JP, et al. Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA*. 2010;303:631-7.
11. Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science*. 2004;305:869-72.
12. Lluís-Ganella C, Lucas G, Subirana I, Sentí M, Jiménez-Conde J, Marrugat J, et al. Efecto aditivo de diferentes variantes genéticas en el riesgo de cardiopatía isquémica. *Rev Esp Cardiol*. 2010;63:925-33.
13. Roberts R. Personalized medicine: A reality within this decade. *J Cardiovas Transl Res*. 2008;1:11-6.
14. McKenna D, Summerour SR, Villareal FJ. Role of mechanical factors in modulating cardiac fibroblast function and extracellular matrix synthesis. *Cardiovasc Res*. 2000;46:257-63.
15. Slaughter LM. Your genes and privacy. *Science*. 2007;316:797.
16. Jarinova O, Stewart AFR, Roberts R, Wells G, Lau P, Naing T, et al. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler Thromb Vasc Biol*. 2009;29:1671-7.