

Identificación de genes de susceptibilidad para el cáncer colorrectal: aportación de los Genome-Wide Association Studies

SERGI CASTELLVÍ BEL Y ANTONI CASTELLS

Servicio de Gastroenterología, Hospital Clínic, Barcelona, España.



Ilustración: Roger Ballabrera

Puntos clave

- Los alelos que contribuyen a la predisposición genética en el cáncer colorrectal (CCR) familiar son más comunes que los de las formas hereditarias, pero moderadamente penetrantes, actúan en cooperación y contribuyen con su pequeño efecto individual a la acción conjunta de diversas variantes genómicas.
- La estrategia más apropiada para identificar múltiples componentes genéticos comunes y de penetrancia baja para CCR son los estudios de asociación genética que comparan su frecuencia entre casos y controles sin enfermedad.
- Algunos polimorfismos podrían modular la expresión diferencial de genes del metabolismo de carcinógenos, metilación, modificadores del microambiente colónico u oncogenes y genes supresores de tumor, que resultan en diferencias en la susceptibilidad al CCR.
- En la actualidad, la mejora reciente en el conocimiento de la variación genómica humana, los bloques haplotípicos HapMap y las tecnologías de genotipado permiten un tratamiento pangenómico en los estudios de asociación genética.
- Dos estudios pangenómicos de asociación genética para CCR realizados por grupos independientes han permitido identificar hasta el momento 6 variantes genómicas implicadas en riesgo para esta neoplasia.

El cáncer colorrectal (CCR) constituye un problema de salud pública de gran importancia en el ámbito mundial, representa en España el segundo cáncer más frecuente y la segunda causa de muerte relacionada con cáncer. En España se diagnostican cada año entre 30 y 40 casos nuevos de CCR por cada 10.000 habitantes¹. El riesgo de desarrollar esta neoplasia en la población general se sitúa en el 5%, pero aumenta de forma exponencial con la edad.

Cáncer colorrectal familiar

Las formas hereditarias clásicas bien conocidas con una fuerte agregación familiar, como la poliposis adenomatosa familiar y el síndrome de Lynch, corresponden a la minoría de los casos de CCR (5%). La agregación familiar, definida como la aparición de CCR en más miembros de una familia de lo que correspondería por azar, es bastante notable fuera de las formas hereditarias que predisponen al CCR, y se denomina CCR familiar. De hecho, esta forma correspondió al 30% de los casos de CCR en el estudio EPICOLON en población española². Por último, una mayoría de casos no muestra ningún tipo de agregación familiar y se engloba dentro del CCR esporádico.

La agregación familiar puede estar causada por una predisposición genética compartida o por factores ambientales comunes. En el CCR, los factores ambientales pueden ser relevantes, porque las familias comparten costumbres dietéticas y hábitos o estilos de vida. La obesidad, una vida sedentaria, el tabaco o una dieta rica en grasas y carne roja se encuentran entre los factores ambientales de riesgo alto que de forma más común se consideran en el CCR, y controlar su efecto real es también importante en cuanto a estrategias de prevención³.

Por otro lado, en estudios epidemiológicos se ha evaluado la contribución ambiental en diversas neoplasias utilizando estudios de gemelos para estimar si la concordancia para la presencia de cáncer es mayor entre gemelos monocigóticos, que comparten todo su genoma, que entre gemelos dicigóticos, que sólo comparten el 50% de sus genes. En un estudio en población escandinava se concluyó que los factores genéticos heredados parecen tener una contribución menor a la susceptibilidad en la mayoría de neoplasias, lo que indica que el ambiente sería la causa primordial. Sin embargo, este mismo estudio detectó un mayor efecto relativo de la heredabilidad en cáncer de próstata y CCR⁴. En un estudio más reciente en una población similar, se evaluó la extensión del riesgo de CCR debido a factores ambientales o heredados, y en el análisis se incluyó a todos los familiares de primer grado (progenitores, hermanos y descendientes) y consortes, con lo que se concluyó que los efectos genéticos son probablemente más importantes que los ambientales para el CCR⁵.

El cáncer colorrectal como enfermedad poligénica

Los síndromes hereditarios que predisponen al CCR son mayoritariamente monogénicos o causados por mutaciones en un único gen, siguiendo un patrón de herencia mendeliano. Los componentes genéticos involucrados en estas formas heredita-

rias poco frecuentes se identificaron con éxito en las 2 décadas anteriores y corresponden a alelos raros altamente penetrantes que predisponen al CCR⁶. En contraste, se postula que los alelos que contribuyen a la predisposición genética de las formas más frecuentes de CCR familiar serían más comunes que los de las formas hereditarias, pero moderadamente penetrantes, que actúan en cooperación y contribuyen con su pequeño efecto individual a la acción combinada de diversas variantes genómicas⁷. Así, este modelo poligénico de predisposición genética al CCR implicaría la coexistencia de cientos de alelos, cada uno ligado a un riesgo moderado-bajo. Estos alelos de riesgo para el CCR sumarían sus efectos individuales menores para ser causantes del desarrollo de esta neoplasia. Pese a que los estudios epidemiológicos resaltan su importancia, hasta el momento la identificación de estos elementos genéticos implicados en el CCR familiar ha sido poco exitosa.

Los análisis de ligamiento genético clásicos, a partir de familias que segregan para una enfermedad concreta, fueron la estrategia correcta para identificar los componentes genéticos del CCR menos comunes y de penetrancia alta, como *APC*, en los que inicialmente se señalaba una región cromosómica específica que contenía el gen causante y, más tarde, se caracterizaban e identificaban mutaciones en individuos afectados. Sin embargo, esta aproximación no es la adecuada para identificar a esos múltiples componentes genéticos para CCR comunes de penetrancia baja. Los estudios de asociación genética constituyen una estrategia mucho más apropiada para identificarlos. Estos estudios buscan detectar la asociación entre una o más variantes genómicas, en las que se compara su frecuencia entre diferentes grupos de individuos (casos frente a controles sin enfermedad) para determinar su implicación en la susceptibilidad a la enfermedad. La asociación alélica es positiva cuando la distribución de genotipos es diferente en casos y controles, lo cual aporta evidencia de que el *locus* estudiado, o un *locus* cercano, está relacionado con susceptibilidad a la enfermedad⁸⁻¹⁰.

Estudios preliminares de asociación genética para evaluar el riesgo de cáncer colorrectal

En los últimos 20 años, en varios estudios de asociación genética se ha evaluado el riesgo de CCR asociado con polimorfismos específicos. La mayoría de estos estudios de asociación de casos y controles se han centrado de forma simultánea en una o pocas variantes genómicas en genes candidatos con una significación en la biología del CCR (tabla 1)¹¹.

Las aminas heterocíclicas son carcinógenos formados durante la cocción de la carne. Estas moléculas se activan metabólicamente por la familia de las enzimas CYP (CYP1A1, CYP2D6, CYP2E1), mientras que se detoxifican por las N-acetiltransferasas (NAT1, NAT2) y las glutatión-S transferasas (GSTM1, GSTP1, GSTT1). Además, la metilación del ácido desoxirribonucleico (ADN) parece tener un papel central en el desarrollo del cáncer¹². Su aparición podría estar influida por la alteración de la disponibilidad de grupos donadores metilo, como los folatos. Las enzimas codificadas por

los genes *MTHFR* y *MTR* controlan los valores de folato. Por otro lado, las alteraciones del microambiente celular dentro de la cripta colónica podrían estar causadas por la síntesis de prostaglandinas (PLA2G2A) o por la excreción de ácidos biliares (APOE). Adicionalmente, hay activación de oncogenes (*HRAS1*) e inactivación de genes supresores de tumor (*APC*, *TP53*)¹³ a lo largo de la secuencia adenoma-carcinoma bien establecida para la progresión del CCR¹⁴. Por estas razones, algunos polimorfismos podrían tener un efecto en la expresión diferencial de genes del metabolismo de carcinógenos, genes implicados en metilación, modificadores del microambiente colónico, oncogenes y genes supresores de tumor, resultando en diferencias en la susceptibilidad al CCR.

La mayoría de los estudios de asociación genética, en los que se ha evaluado el riesgo de CCR asociado a polimorfismos en genes en las categorías anteriores, han sido bastante controvertidos y han concluido asociaciones positivas no replicadas en muchos casos en estudios independientes subsiguientes, o incluso se ha llegado a conclusiones opuestas en cohortes de CCR diferentes. La explicación más factible a esta falta de reproducibilidad, al considerar la misma variación genética ligada a riesgo de CCR, es probablemente un diseño del estudio inapropiado. Los errores comunes en estudios de asociación genética en el pasado han sido: un tamaño muestral pequeño, con la consiguiente falta de poder estadístico; una ausencia de replicación de los resultados observados en un nuevo grupo independiente de casos y controles; un grupo control pobremente apareado con el grupo de casos; un sesgo en la publicación de estudios positivos, y una estratificación en la población estudiada^{8,15}. Recientemente, un esfuerzo notable para usar las estrategias apropiadas en los estudios de asociación genética ha resuelto los problemas mencionados anteriormente, y se ha puesto más énfasis en un diseño metódico del estudio que en el valor de la significación estadística inicial¹⁶. En la tabla 1 se esquematizan resultados de estudios de asociación más recientes que ya han usado las recomendaciones metodológicas y sus resultados se pueden considerar más fiables¹⁷.

Tabla 1. Selección de genes y polimorfismos con estudios de asociación genética realizados para evaluar su hipotético riesgo de cáncer colorrectal

Gen	Polimorfismo
APC	E1317Q
APOE	E2/E3/E4
ARLTS1	C148R
CCDN1	A870G
CHEK2	1100delC
HRAS1	Repetición minisatélite
IRS1	G972R
IL6	-174G>C
IL8	-251T>A
MTHFR	C677T, A1298C
NAT1/NAT2	Alelos acetiladores rápidos
P53	R72P
TGFBR1	del(Ala) ₃
VDR	Bsml intrón 8, poliA corto/largo

Mejora en las herramientas moleculares utilizadas en los estudios de asociación genética

Durante la pasada década, tuvo lugar un avance enorme en las herramientas moleculares utilizadas en los estudios de asociación genética que, por ello, han permitido un adelanto notable en el campo de la susceptibilidad genética.

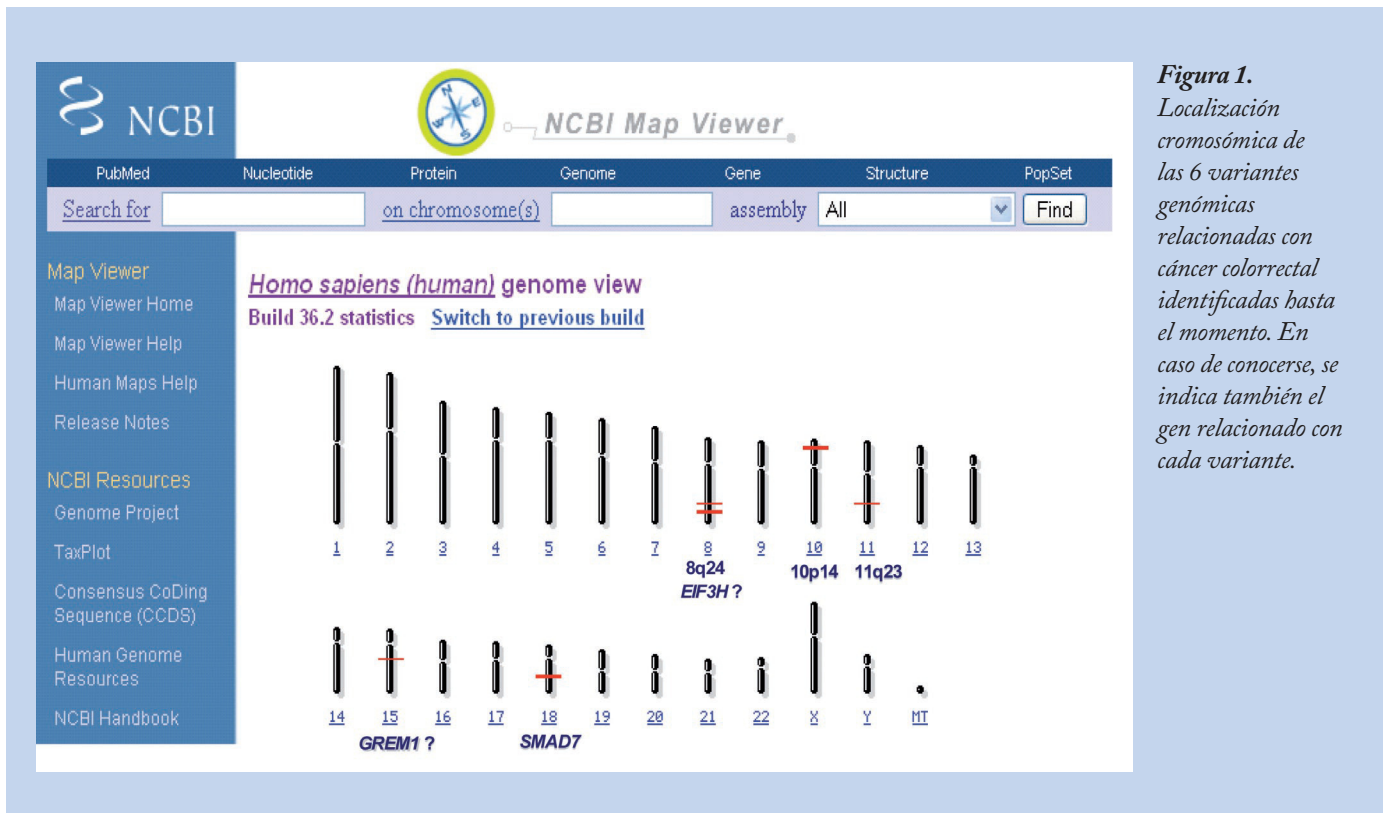
En primer lugar, los polimorfismos de un único nucleótido o SNP (del inglés, *single nucleotide polymorphism*), la variación genómica más común en el genoma humano, se han identificado de forma amplia, están catalogados y son accesibles públicamente en varias bases de datos. Una de ellas, la dbSNP (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp>) tiene en la actualidad más de 12.800.000 entradas humanas disponibles, según los datos aportados por la versión más reciente (*build 129*).

En segundo lugar, los SNP también se han ensamblado en bloques haplotípicos en el proyecto HapMap. Este proyecto es un catálogo de variantes genómicas comunes que existen en los seres humanos y describe el tipo de variaciones, donde ocurren en nuestro ADN, y cómo están distribuidas en los individuos de distintas poblaciones del mundo. Estos bloques haplotípicos cubren el genoma humano y, dentro de cada bloque, algunos SNP (tagSNP) son representativos de la variación genética en un gen o una región genómica. Así, usar tagSNP en los estudios de asociación genética se ha convertido en una estrategia que permite una reducción enorme en el número de SNP que se necesitarían estudiar para tener una cobertura pangenómica. En la fase 1 de este proyecto se genotiparon aproximadamente 1,3 millones de SNP, y muy recientemente, se han hecho públicos los resultados de la fase 2 con datos de 2,1 millones de SNP adicionales¹⁸.

Finalmente, en los últimos años ha existido también una gran mejora en las tecnologías de genotipado, lo cual hace que sean factibles las aproximaciones a gran escala o de carácter pangenómico en los estudios de asociación genética. En la actualidad, la tecnología de matrices o *arrays* y otras técnicas de rendimiento alto permiten un genotipado masivo en paralelo a un coste razonable para genotipar hasta 2 millones de SNP al mismo tiempo.

Resultados iniciales de los estudios de asociación pangenómicos para cáncer colorrectal

Recientemente se han abordado muchas enfermedades comunes mediante la metodología de los estudios de asociación pangenómicos (WGAS, del inglés *whole-genome association studies*) con un éxito significativo, como por ejemplo el infarto de miocardio agudo, la diabetes mellitus tipo 1 y la enfermedad inflamatoria intestinal¹⁹. Los WGAS han implicado típicamente el genotipado de un número considerable de SNP (típicamente 500.000) en un gran número de muestras para poder conseguir el poder estadístico necesario para ser capaces de detectar los efectos de suscep-



tibilidad genética menores. Es importante remarcar que la replicación en un grupo independiente de casos y controles es obligatoria para confirmar cualquier resultado positivo de un WGAS u otro estudio de asociación genética debido a la existencia de falsos positivos²⁰.

Aunque no tan numerosos como los de las enfermedades anteriormente mencionadas, 2 iniciativas WGAS centradas en la susceptibilidad al CCR realizadas por grupos independientes ya han mostrado resultados interesantes (fig. 1). El primer resultado WGAS identificó SNP asociados con riesgo de CCR en la región cromosómica 8q24 por parte de 3 grupos independientes²¹. Algunos de estos SNP ya habían sido implicados en susceptibilidad al cáncer de próstata, y asimismo demostraron la posibilidad de componentes genéticos compartidos entre neoplasias relacionadas.

Posteriormente, un segundo grupo de SNP en el gen *SMAD7* se asoció también con riesgo de CCR²². Este gen actúa como antagonista intracelular de la señalización TGF- β al unirse de forma estable al complejo receptor y bloquear la activación posterior de la cascada de señalización, y la alteración de su expresión se había relacionado anteriormente con CCR.

Adicionalmente, algunos SNP en la región cromosómica 15q13.3 también se han asociado con riesgo de CCR gracias a la estrategia WGAS²³. Previamente se sabía que esta región contenía un *locus* involucrado con un síndrome de poliposis mixta y CCR (MIM#601228).

Muy recientemente, se han publicado nuevos resultados de 2 WGAS independientes para CCR, que han mostrado 3 nuevos componentes genéticos comunes de baja penetrancia en las regiones cromosómicas 10p14 y 8q23.3²⁴ y en 11q23²⁵. Actualmente, se está realizando un metaanálisis de los datos

de estos últimos 2 estudios, que muy probablemente aumentará el número de variantes genómicas relacionadas con CCR identificadas.

Cabe resaltar que la gran mayoría de las variantes genómicas relacionadas con CCR identificadas hasta ahora tienen riesgos asociados moderados (1,1-2) y se encuentran ampliamente distribuidas en la población general con frecuencias por encima del 10% para el alelo menor. Por ello, un porcentaje alto de individuos de la población son portadores de genotipos de riesgo. Probablemente, teniendo en cuenta un modelo de herencia poligénico, cada alelo individual ligado a riesgo para el CCR con un pequeño efecto podría combinarse de forma aditiva o multiplicativa para producir riesgos mayores en conjunto. Los portadores de múltiples alelos de riesgo podrían tener ya un riesgo de CCR suficiente que justificaría intervenciones de cribado específicas y diferentes a las propuestas a la población general, u otras medidas preventivas. Asimismo, en la actualidad se está analizando si estas mismas variantes genómicas relacionadas con CCR podrían tener un papel modificador en las formas hereditarias de esta neoplasia.

Finalmente, también es necesario comentar que la mayoría de los componentes identificados por la aproximación WGAS para CCR u otras enfermedades complejas corresponden a tagSNP, debido a que las matrices de genotipado de alto rendimiento se diseñaron con este tipo de variación de forma preferente. Los tagSNP no suelen tener una importancia funcional por ellos mismos y, por esa razón, en la mayoría de casos no se conoce el mecanismo funcional real por el cual se explicaría el efecto en el riesgo de CCR, lo cual abre un nuevo frente en la investigación de la susceptibilidad genética en las enfermedades humanas.

Bibliografía



● Importante ●● Muy importante

■ Epidemiología

1. Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide. IARC CancerBase No. 5. version 2.0. Lyon: IARC Press; 2004. Disponible en: <http://www-dep.iarc.fr>
2. Piñol V, Castells A, Andreu M, Castellvi-Bel S, Alenda C, Llor X, et al. Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. *JAMA*. 2005;293:1986-94.
3. Johnson IT, Lund EK. Review article: nutrition, obesity and colorectal cancer. *Aliment Pharmacol Ther*. 2007;26:161-81.
4. ● Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo, et al. Environmental and heritable factors in the causation of cancer-analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000;343:78-85.
5. Hemminki K, Chen B. Familial risk for colorectal cancers are mainly due to heritable causes. *Cancer Epidemiol Biomarkers Prev*. 2004;13:1253-6.
6. Rustgi AK. The genetics of hereditary colon cancer. *Genes Dev*. 2007;21:2525-38.
7. Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. *Nat Genet*. 2003;33(Suppl):238-44.
8. Cardon LR, Bell JL. Association study designs for complex diseases. *Nat Rev Genet*. 2001;2:91-9.
9. ● Pharoah PD, Dunning AM, Ponder BA, Easton DF. Association studies for finding cancer-susceptibility genetic variants. *Nat Rev Cancer*. 2004;4:850-60.
10. Cordell HJ, Clayton DG. Genetic association studies. *Lancet*. 2005;366:1121-31.
11. Kemp Z, Thirlwell C, Sieber O, Silver A, Tomlinson I. An update on the genetics of colorectal cancer. *Hum Mol Genet*. 2004;13:R117-85.
12. Jones PA, Baylin SB. The epigenomics of cancer. *Cell*. 2007;128:683-92.
13. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990;61:759-67.
14. Morson BC. The poly-p-cancer sequence in the large bowel. *Proc R Soc Med*. 1974;67:451-7.
15. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. *Lancet*. 2003;361:865-72.
16. Hattersley AT, McCarthy MI. What makes a good genetic association study? *Lancet*. 2005;366:1315-23.
17. Castellvi-Bel S, Castells A, De Cid R, Muñoz J, Balaguer F, Gonzalo V, et al. Association of the ARLTS1 Cys148Arg variant with sporadic and familial colorectal cancer. *Carcinogenesis*. 2007;28:1687-91.
18. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007;449:851-61.
19. Kingsmore SF, Lindquist IE, Mudge J, Gessler DD, Beavis WD. Genome-wide association studies: progress and potential for drug discovery and development. *Nat Rev Drug Discov*. 2008;7:221-30.
20. NCI-NHGRI Working Group on Replication in Association Studies, Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, et al. Replicating genotype-phenotype associations. *Nature*. 2007;447:655-60.
21. ●● Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet*. 2007;39:984-8.
22. Broderick P, Carvajal-Carmona L, Pittman AM, Webb E, Howarth K, Rowan A, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat Genet*. 2007;39:1315-7.
23. Jaeger E, Webb E, Howarth K, Carvajal-Carmona L, Rowan A, Broderick P, et al. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat Genet*. 2008;40:26-8.
24. ●● Tomlinson IP, Webb E, Carvajal-Carmona L, Broderick P, Howarth K, Pittman AM, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet*. 2008;40:623-30.
25. Tenesa A, Farrington SM, Prendergast JG, Porteous ME, Walker M, Haq N, et al. Genome-wide association scan identifies a colorectal susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat Genet*. 2008;40:631-7.