

Aspectos actuales de los estudios de asociación del genoma completo (GWAS) en la enfermedad pulmonar obstructiva crónica

GLORIA PÉREZ-RUBIO^{*,†}
IRMA SILVA-ZOLEZZI[§]
ALEJANDRA RAMÍREZ-VELEGAS^{*}
RAÚL H. SANSORES^{*}
JUAN M. RESÉNDIZ-HERNÁNDEZ^{*,†}
MARTHA MONTAÑO^{*}
ÁNGEL CAMARENA^{*}
RAMCÉS FALFÁN-VALENCIA^{*} ✉

^{*} Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas.

[†] Posgrado en Ciencias Biológicas, UNAM.

[§] Instituto Nacional de Medicina Genómica (INMEGEN).

Trabajo recibido: 11-XI-2009; aceptado: 13-IV-2010

Conflicto de interés: ninguno

337

RESUMEN

Se considera que la exposición al humo de cigarro es el principal factor de riesgo para padecer enfermedad pulmonar obstructiva crónica (EPOC); aun así, sólo entre 10-20% de los fumadores desarrollan la enfermedad. El grado de tabaquismo (paquetes-año, tiempo de fumador) se asocia

Palabras clave: sólo al 15% de la variación de la función pulmonar, indicando que deben existir otros factores involucrados en la susceptibilidad a desarrollar EPOC. Esta patología se considera una entidad única entre las enfermedades genéticas complejas en la cual, el factor de riesgo más importante es conocido y los niveles de exposición al mismo pueden ser documentados con cierta precisión. La elevada mortalidad y morbilidad asociada a EPOC, así como su naturaleza crónica y progresiva, ha promovido el uso de estudios de genética molecular para intentar identificar factores de susceptibilidad adicionales. Los estudios de ligamiento genético son muy poderosos para el estudio de enfermedades monogénicas, pero su poder es limitado en enfermedades genéticas complejas en las que muchos genes participan y se presentan patrones de herencia poco claros. Actual-

Key words: Genetic susceptibility, COPD, GWAS.

La elevada mortalidad y morbilidad asociada a EPOC, así como su naturaleza crónica y progresiva, ha promovido el uso de estudios de genética molecular para intentar identificar factores de susceptibilidad adicionales. Los estudios de ligamiento genético son muy poderosos para el estudio de enfermedades monogénicas, pero su poder es limitado en enfermedades genéticas complejas en las que muchos genes participan y se presentan patrones de herencia poco claros. Actual-

ABSTRACT

Cigarette smoking is considered as the major risk factor for chronic obstructive pulmonary disease (COPD); however, only 10-20% of smokers develop the disease. The extent of cigarette smoking (pack-year, duration of smoking habit) only accounts for 15% of the variation in lung function, indicating that factors other than tobacco habit must be involved in susceptibility to COPD. This disease is unique among complex genetic illnesses in that the environmental risk factor is known and the level of exposure can be documented with some precision. The high rates of mortality and morbidity associated with COPD, as well as its chronic and progressive nature, has prompted the use of molecular genetic studies in an attempt to identify additional susceptibility factors. Genetic linkage studies are very powerful for the study of monogenic disorders, but have limited power in complex genetic disorders where many genes are likely to be acting and there are no clear inheritance patterns. Currently, with the completion of the human genome project, the human genetic sequence is known and publicly available for consultation. These advances in the knowledge, along with technological progresses that allow high throughput

mente, con la conclusión del proyecto del genoma humano se conoce la secuencia del genoma de nuestra especie, además de que ésta es pública y está disponible para su consulta. Estos avances en el conocimiento, en conjunto con los progresos tecnológicos que permiten la genotipificación de alto rendimiento de miles de polimorfismos en cientos o miles de individuos, han permitido la generación de herramientas excepcionales para el descubrimiento de genes/polimorfismos de gran relevancia para la investigación biomédica. En la presente revisión, describimos el escenario más reciente en cuanto a estudios de asociación genética en EPOC se refiere.

ENFERMEDAD PULMONAR OBSTRUCTIVA CRÓNICA

La enfermedad pulmonar obstructiva crónica (EPOC) es un padecimiento pulmonar frecuente en el mundo, con repercusiones sobre la tasa de mortalidad y elevados costos de atención. El término EPOC agrupa un conjunto de procesos que tienen en común la disnea como principal síntoma, se acompaña de obstrucción crónica al paso de aire por los pulmones, no es completamente reversible y comúnmente es progresiva. Se asocia con una respuesta inflamatoria anormal de los pulmones por la exposición a partículas nocivas y gases, es más frecuente en personas mayores de 45 años. Los pacientes pueden desarrollar bronquitis crónica, enfisema o una combinación de ambas.^{1,2}

A pesar de la complejidad para medir la prevalencia en EPOC, se puede afirmar que en muchos países desarrollados está aumentando. Es mayor en hombres fumadores en comparación con exfumadores y en personas mayores de 40 años; además, tiende a incrementar con la edad. De acuerdo con el estudio PLATINO (Proyecto Latinoamericano en Investigación en Obstrucción Pulmonar) en poblaciones de América Latina (Brasil, Uruguay, Chile, México y Venezuela) con personas mayores de 60 años se podría tener una prevalencia que oscila en un rango de 18.4% (la más baja en México) a 32.1% (la más alta en Uruguay).³ El incremento en la mortalidad por EPOC a nivel mundial ha sido sustancial en los últimos 30 años; en 1990 ocupaba el sexto lugar

genotyping of thousands of polymorphisms on hundreds or thousands of individuals, have provided a remarkable set of tools to discover genes/polymorphisms relevant for biomedical research. In this review we describe the most recent scenario related to genetic association studies in COPD.

y se considera que para el 2020 ocupará el tercer lugar. En términos de morbilidad, para ese mismo año, se estima que ocupará entre el tercero o cuarto lugar a nivel mundial.⁴

En nuestro país en el Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas (INER), la EPOC ocupa el segundo lugar en visitas a urgencias y el cuarto lugar como causa de egreso hospitalario.⁵

Entre los factores condicionantes de riesgo a padecer EPOC se incluyen los factores genéticos y ambientales, originando una enfermedad compleja, multifactorial y probablemente multigénica.^{2,6} El factor de riesgo ambiental más frecuente y claramente asociado es el tabaquismo y en países en vías de desarrollo como México, Nepal, Nueva Guinea y Colombia, la exposición al humo de leña también se ha documentado como factor de riesgo a EPOC. La exposición en el ambiente laboral a polvos, gases y humos constituye otros factores de riesgo extrínsecos adicionales.²

Hay estudios donde la relación dosis-respuesta entre el humo de cigarro y la función pulmonar está bien establecida, existe una considerable variabilidad en el grado de obstrucción del flujo aéreo que ocurre en respuesta al humo de cigarro.^{7,8} El bajo porcentaje de variación en la función pulmonar explicada por humo de cigarro (alrededor de 15%), y la existencia de casos de inicio temprano y función pulmonar reducida severamente, sugieren variabilidad en la susceptibilidad genética individual a los efectos del humo de cigarro.⁸⁻¹⁰ Adicionalmente, hay que tomar en

cuenta que sólo entre el 10 y 20% de los casos que desarrollan EPOC son o fueron fumadores,^{6,11} lo cual indica que existen otros factores incluidos los genéticos que confieren susceptibilidad al desarrollo de la enfermedad. Diversos estudios han encontrado que después de ajustar por edad, género, hábito de fumar cigarro e índice de masa corporal, existe una fuerte correlación en el nivel de función pulmonar entre individuos biológicamente relacionados.^{7,9} Esto se apoya en estudios basados en familias, que han demostrado un aumento del riesgo al daño en la función pulmonar en familiares relacionados en primer grado con pacientes fumadores o exfumadores.⁸ Estudios de la función pulmonar en gemelos proveen evidencia adicional del papel de los factores genéticos en la determinación de la función pulmonar en población general.¹²

LOS ESTUDIOS DE ASOCIACIÓN COMO ABORDAJE EN EL ESTUDIO GENÉTICO DE LAS ENFERMEDADES COMPLEJAS

Desde antes de que existiera la posibilidad de analizar el material genético, es decir, el ácido desoxirribonucleico (DNA por sus siglas en inglés), los científicos han intentado relacionar la variabilidad genética (genotipo) a los rasgos biológicos o a las enfermedades (fenotipo). El DNA es el material biológico contenido en las células que posee las instrucciones necesarias para que los organismos vivan y cumplan todas sus funciones biológicas. Es una molécula conformada por compuestos químicos llamados nucleótidos a los que conocemos por sus iniciales: A (adenina), C (citosina), G (guanina) y T (timina). La secuencia del genoma humano corresponde al orden y posición de tales nucleótidos en la molécula de DNA. La variabilidad genética corresponde al conjunto de cambios a nivel de secuencia en el DNA que participan de manera importante en las diferencias biológicas observadas entre los seres vivos. Las variantes genéticas más estudiadas en los seres humanos son los polimorfismos de un solo nucleótido o SNP (del inglés *single nucleotide polymorphism*, en la jerga científica se pronuncia *snip* y al plural se le denomina SNPs), las cuales corresponden a sustituciones de un nucleótido por otro en la cadena del DNA.¹³ Los SNPs for-

man hasta el 90% de todas las variaciones genómicas humanas, y aparecen cada 100 a 300 bases en promedio a lo largo del genoma humano. En las bases de datos públicos internacionales se encuentran más de 6 millones de SNPs y es probable que existan más de 10 millones de SNPs con frecuencias alélicas superiores al 1%.

La idea de poder conocer los patrones genéticos que se encuentran detrás del desarrollo de las enfermedades y que son compartidos por individuos, familias y/o poblaciones es muy atractiva; ya que abre la posibilidad de identificar variaciones genéticas, genes y procesos biológicos que son información fundamental para el avance en el conocimiento de la fisiopatología de las enfermedades y la potencial aplicación de este conocimiento en el desarrollo de sistemas diagnóstico, terapias o políticas de salud.

Dentro de la epidemiología genética, existen dos abordajes para la búsqueda de los genes y variantes genéticas que participan en los rasgos y enfermedades comunes: a) los estudios de ligamiento, los cuales se basan en el estudio de la herencia de marcadores genéticos y fenotipos dentro de familias con individuos afectados; y b) los estudios de asociación genética basados en la identificación de diferencias significativas en la frecuencia de variantes genéticas entre grupos de estudio, generalmente un grupo de casos (individuos enfermos) y uno de controles (individuos sanos), el conocimiento de las variantes en estos grupos lleva a la identificación de factores genéticos asociados a riesgo o protección. El primer abordaje ha sido el más exitoso en la búsqueda e identificación de genes y variantes causantes de enfermedades monogénicas (fibrosis quística, hemofilia y la mayoría de los errores innatos del metabolismo, entre otras). Todas ellas, enfermedades poco frecuentes en las que los factores genéticos dan lugar a alteraciones biológicas severas que causan enfermedad. En contraste, los estudios de asociación genética han sido la estrategia de elección en la búsqueda de los genes y variantes relacionadas a la susceptibilidad al desarrollo de enfermedades complejas o multifactoriales. En este último grupo de enfermedades se encuentran las más comunes y que hoy en día afectan a un porcentaje amplio de la población, como son la diabetes, el cáncer, las enfermeda-

des cardiovasculares y un importante número de enfermedades respiratorias.¹⁴

Los estudios de asociación genética se enfrentan al hecho de que las enfermedades complejas en relación con sus bases genéticas presentan un escenario más problemático que las enfermedades monogénicas. Puesto que incluyen la participación simultánea de varios genes y variantes genéticas (heterogeneidad genética y alélica, respectivamente), la existencia de patologías con un diagnóstico similar pero con una construcción biológica distinta (fenocopias), el hecho de que una variante genética en particular no es suficiente ni necesaria para que se presente la enfermedad (penetrancia incompleta), y a la existencia de interacciones de los genotipos con el medio ambiente.¹⁵ Todo lo anterior ha complicado históricamente la caracterización de los factores genéticos relacionados a las enfermedades complejas, y ha impulsado la generación de abordajes y herramientas novedosos para lograr incrementar el poder de este tipo de estudios.

Los estudios de asociación en la era genómica

Uno de los principales resultados del Proyecto del Genoma Humano (del inglés *human genome project* [HGP]), el cual tuvo como objetivo caracterizar la secuencia completa del DNA humano, fue la identificación de variantes genéticas que contribuyen a la individualidad genética humana. Adicionalmente, el proyecto HapMap (acrónimo en inglés del proyecto internacional de "mapa de haplotipos") enfocado a la caracterización de diversidad genética en diferentes grupos poblacionales humanos, en conjunto con el desarrollo de tecnologías avanzadas de genotipificación masiva han hecho posible el planteamiento de abordajes y herramientas que han permitido un importante avance en la identificación de las bases genómicas de las enfermedades complejas.^{16,17} En lo que respecta a la población mexicana, en mayo de 2009 el Instituto Nacional de Medicina Genómica publicó los primeros resultados del Proyecto de Diversidad Genómica de Poblaciones Mexicanas, el cual se suma a los proyectos internacionales en la generación de información y herramientas para el desarrollo de la investigación

genómica en medicina. Estos proyectos han dado fundamento al desarrollo de estrategias para acelerar la identificación de genes asociados a enfermedades complejas; tales como los estudios de asociación de genoma completo GWAS (del inglés *genome-wide association studies*) y el mapeo por mestizaje AM (del inglés *admixture mapping*). El éxito de los GWAS depende de una eficiente selección de SNPs, en tanto que el AM de la disponibilidad de marcadores con diferentes frecuencias entre las poblaciones parentales, conocidos como marcadores informativos de origen ancestral AIMs (del inglés *ancestral informative markers*).¹⁸ Los GWAS han generado conocimiento de importancia en la identificación de los riesgos genéticos a enfermedades complejas.^{16,19} En muchos de los genes identificados recientemente no se tenía la "sospecha" de que tuvieran algún papel importante en la biología de la enfermedad en estudio e incluso se han descrito asociaciones con regiones genómicas con función aún desconocida. Estos resultados se han logrado en buena medida debido a que los GWAS interrogan desde cientos de miles hasta millones de SNPs a lo largo de todo el genoma, con estrategias que proveen una oportunidad única para la exploración del genoma libre de hipótesis; es decir, independientemente de la naturaleza o función de las regiones en donde se encuentran los SNPs analizados.¹⁶ El diseño más frecuente de los GWAS es el de casos y controles, donde es crítico seleccionar el número adecuado de participantes, que típicamente es de varios cientos o miles de individuos dentro de cada grupo. Dicha selección está en razón de la prevalencia de la enfermedad en la población en estudio, de las estimaciones en relación con los posibles niveles de riesgo genético (razones de momios), así como de su frecuencia en la población (frecuencia alélica). Adicionalmente, en estos diseños es importante cuidar que los casos y controles de estudio provengan de poblaciones con una estructura genética similar, y posean criterios estrictos para incluirlos dentro del grupo adecuado.^{15,19}

Entre las aplicaciones más importantes de los GWAS se encuentran la generación de hipótesis para la realización de estudios que permitan determinar con mayor detalle cuáles son las varian-

tes genéticas responsables de los efectos biológicos relacionados a la enfermedad. Lo cual se logra detectando haplotipos (combinaciones particulares de alelos en una región definida de un cromosoma) de alto riesgo, estudios de asociación en genes y/o regiones candidato;²⁰ así como estudios epidemiológicos para buscar identificar interacciones gen-gen.²⁰ En el análisis de los resultados de un GWAS es difícil o prácticamente imposible conocer cuál es el SNP o la variación genética realmente responsable del efecto biológico, ya que las asociaciones genéticas identificadas con esta estrategia son generalmente indirectas debido a un fenómeno que conocemos como desequilibrio de ligamiento (del inglés *linkage disequilibrium* [LD]). El LD entre variantes genéticas sucede como resultado de la no segregación independiente de las mismas en los cromosomas de una población, lo que ocasiona que dos o más variantes en posiciones físicas independientes se comporten como si fueran la misma, proceso que en el inglés científico se denomina *tagging*, es decir, etiquetado y a cada SNP con esta característica se le denomina *taggingSNP* (aunque diferentes autores prefieren el término *tagSNP*). Este fenómeno permite minimizar el número de variantes caracterizadas en el laboratorio para obtener el máximo de cobertura en las regiones genómicas analizadas, reduciendo así tiempos y costos.¹⁴ Las correlaciones entre variantes genéticas son particulares de una población, por lo que el conocimiento de la estructura genética de la misma es de fundamental importancia para una óptima selección de *taggingSNP*. Lo anterior se vuelve especialmente crítico en el diseño de estudios en regiones y genes candidato que buscan caracterizar con detalle las variantes genéticas asociada a los rasgos y enfermedades; más aún, cuando se examinan poblaciones donde la estructura genética y las frecuencias de las variantes genéticas difieren de la población donde se identificaron originalmente.¹⁹ En estos abordajes existe una gran variedad de diseños, que en su mayoría emplean estudios de múltiples etapas para disminuir los falsos positivos mediante análisis de replicación y validación de las señales de asociación. Diferencias adicionales entre los grupos de comparación (edad, talla, género, origen ancestral, etc.) deben ser consideradas en el análisis

para evitar el riesgo de identificar asociaciones genéticas que no corresponden a la enfermedad, sino que resultan de un factor confusor.²⁰

Para evaluar adecuadamente las contribuciones al riesgo de los factores genéticos, así como para identificar su interacción con factores ambientales es deseable que adicionalmente a los estudios de casos y controles se realicen estudios de cohorte, los cuales debido a su naturaleza longitudinal (los participantes son seguidos a lo largo de periodos de tiempo) permiten evaluar la capacidad predictiva de los marcadores genéticos, medir incidencias y hacer análisis de sobrevida, entre otros. Los estudios de cohorte presentan ventajas, pero son costosos y su logística puede ser difícil de realizar.¹⁴

A la fecha, un gran número de GWAS han sido publicados y muchos más están en proceso.²¹ Múltiples asociaciones entre miles de SNPs y distintas enfermedades han sido identificadas y confirmadas utilizando estos abordajes. La adecuada replicación y exploración fina y detallada de las regiones genéticas en poblaciones distintas a aquella en la que fueron identificadas mediante GWAS, es de mucha importancia en la identificación de las variantes genéticas directamente relacionadas a los efectos biológicos asociados a enfermedad. Lo anterior se logra mediante abordajes de estudios de asociación de región o genes candidato. Para el diseño de estos estudios en enfermedades complejas existe una gran cantidad de literatura que puede ser utilizada para identificar polimorfismos puntuales, genes o incluso grandes regiones cromosómicas de interés en la patología. Ese tipo de literatura incluye GWAS realizados en los últimos cinco años, estudios genéticos de ligamiento (cohortes, casos y controles, etc.), así como evidencias de investigación biomédica básica referente a la patogénesis de la enfermedad. De tal manera que los investigadores pueden tener un amplio pero finito grupo de genes que resultan de interés.

Los estudios de asociación genética y EPOC

Los estudios de asociación en casos y controles son los más aplicados para estudiar enfermedades respiratorias como la EPOC.^{15,19} Múltiples estudios han evaluado la participación individual de

marcadores genéticos (SNP y otros marcadores polimórficos) en el desarrollo de EPOC. De esta manera, se han llevado a cabo algunos GWAS que buscan nuevos SNPs tratando de hallar hipótesis alternativas o novedosas a los mecanismos genéticos subyacentes de la patología.⁹

Una gran variedad de genes han sido propuestos como posibles participantes en la susceptibilidad al desarrollo de EPOC. Algunos de los genes más claramente asociados, tanto en estudios previos a los GWAS como en estos mismos se enlistan en la Tabla I. En el caso de la deficiencia de alfa 1-antitripsina, relacionada a variantes genéticas en el gen de SERPINA1, ésta se considera un claro factor genético de riesgo a desarrollar EPOC, pero sólo explica aproximadamente el 1% de los casos, sobre todo aquéllos de inicio temprano en la vida.^{6,11} Tan clara es la asociación de esta deficiencia que ha sido utilizada como criterio de exclusión en el reclutamiento de pacientes para GWAS en EPOC.³² Lo anterior con el fin de optimizar la búsqueda de factores genéticos adicionales a SERPINA1 que confieren susceptibilidad a esta enfermedad. Otros genes han mostrado resultados contradictorios en estudios previos a los GWAS, como la α 1-antiquimotripsina (SERPINA3),³³⁻³⁵ factor de necrosis tumoral,³⁶⁻³⁸ Decorina (DCN),²⁴ la proteína de unión a vitamina D,³⁹ y el ligando 1 del motivo C-C quimiocina (CCL1).⁴⁰ Lo anterior, hace suponer que algunas o todas estas asociaciones pudiesen

ser falsos positivos. Resulta importante destacar que las asociaciones de genes individuales cubren vías biológicas con funciones dispares, incluyendo la respuesta inmune, inflamación, metabolismo de xenobióticos, antioxidantes y sistema de proteasas-antiproteasas. Hersh *et ál*, en 2005 realizaron un estudio de replicación basado en familias y de forma paralela en casos y controles con la estrategia de gen candidato. Utilizando asociaciones individuales relacionadas al desarrollo de EPOC, evaluaron 29 polimorfismos (24 tipo SNP, 1 inserción/delección, 3 repeticiones cortas en tándem y 1 delección nula) en 12 genes candidato reportados previamente en la literatura. Encontraron asociaciones significativas relacionadas con distintos fenotipos cualitativos y/o cuantitativos con diferentes SNPs, entre ellos: SFTPB (Thr131Ile), HMOX1 (GT31), EPHX1 (His139Arg). Estos hallazgos permitieron la replicación exitosa de algunos SNPs previamente asociados y la descripción de otros no reportados con anterioridad.⁴¹

LOS GWAS EN EL ESTUDIO DE LA EPOC

En 2007 se publicó un estudio realizado en la cohorte de corazón de Framingham, donde por medio de GWAS identificaron dos SNPs asociados con las medidas de función pulmonar, localizados en dos genes que previamente no habían sido descritos. En dicho estudio los participantes fueron agrupados por fenotipo según los resulta-

Tabla I. Principales genes encontrados por medio de estrategia de gen candidato con asociación positiva a EPOC.

Gen			Población estudiada	(n empleada) casos/controles	Referencia
Nombre	Símbolo	Cr			
Inhibidor de serpin peptidasa E, miembro 2	SERPINE2	2	Caucásica	304/441	22
Epóxido hidrolasa microsomal 1	EPHX1	1	Noruega	244/248	23
Factor de crecimiento transformante beta-1	TGFB1	19	Caucásica	188/1156	24
Dominio 33 de metaloproteínasa ADAM	ADAM33	20	Caucásica	287/311	25
Proteína surfactante B	SFTPB	2	Mexicana	97/99	6
Interleucina-6	IL6	7	Española	191/296	26
Interleucina-13	IL13	5	Japonesa	96/61	27
Glutación S-transferasa mu 1	GSTM1	1	Eslovaca	217/160	28
Glutación S-transferasa pi 1	GSTP1	11	Japonesa	50/53	29
Metalopeptidasa de matriz 9	MMP9	20	China	100/98	30
Receptor beta-2 adrenérgico	ADRB2	5	Alemana	190/172	31

La asociación con EPOC se da tomando como base valores de *p* significativos. Cr: Cromosoma.

dos de espirometría, los autores sugirieron que para aquellos SNPs con resultados significativos sobre todo para los ubicados en genes que aún se desconoce su función biológica, se requería replicación en otra muestra de individuos.⁹ Pillai *et al*, en 2009 reportaron un GWAS en EPOC, donde usaron la estrategia de replicación en múltiples etapas. Inicialmente en este estudio se analizaron 538,030 SNPs en un grupo de casos y controles ($n = 1,643$), obteniéndose una lista de asociaciones de la que se eligieron los 100 SNPs más claramente asociados ($p < 0.00016$) que fueron caracterizados en una cohorte de familias ($n = 1,891$) adicional, a lo que llamaron etapa de replicación I. De los resultados de la replicación I se seleccionaron los siete SNPs con la mejor evidencia de asociación en el análisis conjunto de las dos etapas ($p < 10^{-7}$), que fueron analizados en una población adicional de casos y controles ($n = 861$ sujetos) llamando a esta etapa replicación II. Finalmente, se hizo un estudio para la validación de resultados en una cohorte con individuos con EPOC ($n = 949$) dando como resultado seis SNPs validados y con asociación importante a la patología (Tabla II). Es importante señalar que encontraron genes que no habían sido reportados con asociados a EPOC, pero que en otras patologías como cáncer pulmonar tienen un papel importante, tal es el caso del gen CHRNA3/5. Otro gen con resultados importantes de asociación a EPOC es el HHIP del cual se conoce poco, pero los autores sugieren que debido a

sus resultados deberá ser estudiado a fondo.³² En 2009, Wilk *et al* publicaron un GWAS relacionado con medidas de función pulmonar en la cohorte del estudio de corazón de Framingham donde la selección de participantes fue más estricta con respecto al trabajo de Pillai; fueron analizados 70,987 SNPs en 1,220 individuos. Sus resultados proponen al gen HHIP como uno de los posibles reguladores en EPOC y sugieren al gen de la glicoforina como importante en la patología debido a la posición que ocupa con respecto a HHIP y hallazgos previos de tal proteína en pacientes con EPOC. En cuanto al gen CHRNA3/5 no obtuvieron resultados significativos, los autores señalan que la selección de la población es importante.⁴²

Repapi y un extenso grupo de colaboradores recientemente identificaron cinco loci genéticos asociados a diferentes medidas de la función pulmonar.⁴³ Basándose en el hecho ya conocido de que dichas medidas son rasgos hereditarios que predicen la morbilidad y mortalidad⁴⁴⁻⁴⁷ relacionada al desarrollo de EPOC. Emplearon un diseño tipo GWAS asociado a las medidas de función pulmonar (FEV_1 y FEV_1/FVC) en el consorcio Spiro-Meta ($n = 20,288$ individuos con ancestría europea). Adicionalmente realizaron un metaanálisis con las señales de asociación más fuertes procedentes de genotipificación directa ($n \leq 32,184$ individuos adicionales); una asociación sumaria de datos *in silico* del consorcio CHARGE ($n = 21,209$) y finalmente incluyeron datos del estudio

Tabla II. Principales genes encontrados por medio de estrategia GWAS con asociación positiva a EPOC.

Gen		Cr	Población estudiada	(n empleada) casos/controles	Referencia
Nombre	Símbolo				
Proteína de interacción Hedgehog	HHIP	4	Caucásica	20,288 casos*	43
Receptor nicotínico colinérgico alfa 3	CHRNA3	15	Caucásica	823/810	32
Receptor nicotínico colinérgico alfa 5	CHRNA5	15			
Receptor 126 acoplado a proteína G	GPR126	6	Caucásica	389/472	49
Dominio 19 de metalopeptidasa	ADAM19	5			
Dominio PDZ 3	PDZRN3	3	Caucásica	225/610	42
Metalopeptidasa de matriz 15	MMP15	16			
Glicoforina A	GYPA	4			

La asociación con EPOC se da tomando como base valores de p significativos. Cr: Cromosoma.

* Se trata de un estudio de cohorte, no se emplearon controles.

Health 2000 ($n \leq 883$). Con los grupos de estudio anteriormente mencionados, los autores lograron confirmar la asociación previamente reportada del locus 4q31. Además de las asociaciones con la FEV_1 o el radio FEV_1/FVC y variantes comunes en cinco loci: 2q35 en TNS1 ($p = 1.11 \times 10^{-12}$), 4q24 en GSTCD (2.18×10^{-23}), 5q33 en HTR4 ($p = 4.29 \times 10^{-9}$), 6p21 en AGER ($p = 3.07 \times 10^{-15}$) y 15q23 en THSD4 ($p = 7.24 \times 10^{-15}$). El análisis de RNAm mostró expresión de TNS1, GSTCD, AGER, HTR4 y THSD4 en tejido pulmonar humano. Finalmente, sugieren que estas asociaciones ofrecen señales de los mecanismos que regulan la función pulmonar e indican blancos potenciales para intervenciones terapéuticas en el tratamiento de la patología respiratoria.⁴³ En general, observaron 17 regiones asociadas de manera independiente ($p < 1 \times 10^{-5}$) con FEV_1 y 23 asociadas con el valor del radio FEV_1/FVC , incluyendo tres regiones (4q24 en GSTCD, 4q31 cercano a HHIP y 15q23 en THSD4), que alcanzan valores de $p < 5 \times 10^{-8}$ en la etapa 1 del GWAS inicial, mismo que corresponde a un valor marginal de $p < 0.05$ después de ajustar por 1 millón de pruebas independientes. El rs12504628 se encontró asociado con ambos rasgos FEV_1/FVC ($p = 6.48 \times 10^{-13}$) y FEV_1 ($p = 1.50 \times 10^{-10}$) se localiza en una región intergénica cercana a HHIP, cubriendo ~300 kb de la región 4q31 que previamente había sido asociada con la función pulmonar y EPOC por Wilk y su grupo.⁴² La familia de genes Hedgehog, de la cual HHIP es miembro, codifica para moléculas involucradas en la regulación de morfogénesis pulmonar, sugiriendo otros mecanismos subyacentes para tales asociaciones.⁴⁸

En ese mismo número de *Nature Genetics*, Hancock *et al* publicaron un trabajo en donde analizan las medidas de la función pulmonar empleando un diseño tipo metaanálisis y GWAS. Este estudio incluyó 20,890 participantes procedentes del consorcio CHARGE, de ancestría europea (el mismo utilizado en una etapa del estudio de Repapi *et al*).⁴³ Estos autores lograron definir ocho loci asociados con el radio FEV_1/FVC (HHIP, GPR126, ADAM19, AGER-PPT2, FAM13A, PTCH1, PID1 y HTR4) y un locus asociado con el valor de FEV_1 (INSTS12-GSTCD-NPNT).⁴⁹ Varios SNPs cercanos a la región codificante para la proteína de interacción Hedgehog

(HHIP) fueron asociados con el radio FEV_1/FVC en los consorcios CHARGE y SpiroMeta, confirmando hallazgos previos realizados en el estudio de corazón de Framingham.⁹ Como se mencionó anteriormente, la vía de señalización Hedgehog es crucial en varios procesos del desarrollo embrionario, incluyendo la morfogénesis pulmonar.

CONCLUSIONES

Los estudios de asociación genética son herramientas útiles en el entendimiento de la patogénesis de las enfermedades complejas, particularmente si existe un fondo genético subyacente. Los avances recientes tecnológicos que incluyen metodologías de genotipificación masiva y análisis bioinformáticos confiables contribuyen en dicha labor. Sin embargo, es necesario establecer diseños apropiados que tomen en cuenta criterios de inclusión adecuados tanto para los casos como para los controles, homogeneidad y tamaño poblacional apropiados en los grupos de estudio con base en la prevalencia de la enfermedad estudiada. Finalmente, realizar replicaciones posteriores en otras poblaciones con el objetivo de validar los hallazgos encontrados.

Agradecimientos:

Los autores agradecemos la colaboración del servicio de la Biblioteca "Horacio Rubio Palacios" del Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, particularmente a las licenciadas Beatriz Ayala Robles, Adriana Reyes Tsubaki e Isabel Carrillo Montiel por su invaluable apoyo en la localización de los artículos fundamentales para el desarrollo del presente artículo.

REFERENCIAS

1. Gordillo CR, Gómez ML, Hipólito CR, Lamuña EM, Pérez CR. *Enfermedad pulmonar obstructiva crónica*. Rev Asoc Mex Med Crit Ter Int 2002;16:201-210.
2. Sansores MR, Ramírez-Venegas A. *Enfermedad pulmonar obstructiva crónica y la celebración de su primer Día Mundial* (editorial). Rev Inst Nal Enf Resp Mex 2002;15:199-200.
3. Consenso Mexicano de EPOC. *Panorama epidemiológico e impacto económico actual*. Neumol Cir Torax 2007;66 Supl 2:13-16.

4. Pauwels RA, Rabe KF. *Burden and clinical features of chronic obstructive pulmonary disease (COPD)*. Lancet 2004;364:613-620.
5. Portal INER. *Clínica de EPOC*. Fecha de acceso: 23 de noviembre, 2009. Accesible en: http://portal.iner.gob.mx/inerweb/GEN_cont_esp.jsp?contentid=1987&version=1&channelid=3
6. Guo X, Lin HM, Lin Z, et al. *Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population*. Eur Respir J 2001;18:482-490.
7. Burrows B, Knudson RJ, Cline MG, Lebowitz MD. *Quantitative relationships between cigarette smoking and ventilatory function*. Am Rev Respir Dis 1977;115:195-205.
8. Silverman EK, Chapman HA, Drazen JM, et al. *Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis*. Am J Respir Crit Care Med 1998;157(6 Pt 1):1770-1778.
9. Wilk JB, Walter RE, Laramie JM, Gottlieb DJ, O'Connor GT. *Framingham Heart Study genome-wide association: results for pulmonary function measures*. BMC Med Genet 2007;8 Suppl 1:8.
10. Silverman EK, Speizer FE. *Risk factors for the development of chronic obstructive pulmonary disease*. Med Clin North Am 1996;80:501-522.
11. Sandford AJ, Weir TD, Spinelli JJ, Paré PD. *Z and S mutations of the alpha1-antitrypsin gene and the risk of chronic obstructive pulmonary disease*. Am J Respir Cell Mol Biol 1999;20:287-291.
12. Redline S, Tishler PV, Rosner B, et al. *Genotypic and phenotypic similarities in pulmonary function among family members of adult monozygotic and dizygotic twins*. Am J Epidemiol 1989;129:827-836.
13. Burton PR, Tobin MD, Hopper JL. *Key concepts in genetic epidemiology*. Lancet 2005;366:941-951.
14. Cordell HJ, Clayton DG. *Genetic association studies*. Lancet 2005;366:1121-1131.
15. Silverman EK, Palmer LJ. *Case-control association studies for the genetics of complex respiratory diseases*. Am J Respir Cell Mol Biol 2000;22:645-648.
16. Jimenez-Sanchez G, Silva-Zolezzi I, Hidalgo A, March S. *Genomic medicine in Mexico: initial steps and the road ahead*. Genome Res 2008;18:1191-1198.
17. International HapMap Consortium; Frazer KA, Ballinger DG, Cox DR, et al. *A second generation human haplotype map of over 3.1 million SNPs*. Nature 2007;449:851-861.
18. Silva-Zolezzi I, Hidalgo-Miranda A, Estrada-Gil J, et al. *Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico*. Proc Natl Acad Sci USA 2009;106:8611-8616.
19. Xu Z, Taylor JA. *SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies*. Nucleic Acids Res 2009;37(Web Server issue):600-605.
20. Pearson TA, Manolio TA. *How to interpret a genome-wide association study*. JAMA 2008;299:1335-1344.
21. Kruglyak L. *The road to genome-wide association studies*. Nat Rev Genet 2008;9:314-318.
22. Demeo DL, Mariani TJ, Lange C, et al. *The SERPINE2 gene is associated with chronic obstructive pulmonary disease*. Am J Hum Genet 2006;78:253-264.
23. Brögger J, Steen VM, Eiken HG, Gulsvik A, Bakke P. *Genetic association between COPD and polymorphisms in TNF, ADRB2 and EPHX1*. Eur Respir J 2006;27:682-688.
24. van Diemen CC, Postma DS, Vonk JM, Bruinenberg M, Nolte IM, Boezen HM. *Decorin and TGF-beta1 polymorphisms and development of COPD in a general population*. Respir Res 2006;7:89.
25. Sadeghnejad A, Ohar JA, Zheng SL, et al. *Adam33 polymorphisms are associated with COPD and lung function in long-term tobacco smokers*. Respir Res 2009;10:21.
26. Córdoba-Lanús E, de-Torres JP, López-Aguilar C, et al. *Association of IL-6 gene polymorphisms and COPD in a Spanish population*. Respir Med 2008;102:1805-1811.
27. Homma S, Sakamoto T, Hegab AE, et al. *Association of phosphodiesterase 4D gene polymorphisms with chronic obstructive pulmonary disease: relationship to interleukin 13 gene polymorphism*. Int J Mol Med 2006;18:933-939.
28. Zidzik J, Slabá E, Joppa P, et al. *Glutathione S-transferase and microsomal epoxide hydrolase gene polymorphisms and risk of chronic obstructive pulmonary disease in Slovak population*. Croat Med J 2008;49:182-191.
29. Ishii T, Matsuse T, Teramoto S, et al. *Glutathione S-transferase P1 (GSTP1) polymorphism in patients with chronic obstructive pulmonary disease*. Thorax 1999;54:693-696.
30. Zhou M, Huang SG, Wan HY, Li B, Deng WW, Li M. *Genetic polymorphism in matrix metalloproteinase-9 and the susceptibility to chronic obstructive pulmonary disease in Han population of south China*. Chin Med J (Engl) 2004;117:1481-1484.
31. Vacca G, Schwabe K, Dück R, et al. *Polymorphisms of the beta2 adrenoreceptor gene in chronic obstructive pulmonary disease*. Ther Adv Respir Dis 2009;3:3-10.
32. Pillai SG, Ge D, Zhu G, et al. *A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci*. PLoS Genet 2009;5:e1000421.
33. Poller W, Faber JP, Scholz S, et al. *Mis-sense mutation of alpha 1-antichymotrypsin gene associated with chronic lung disease*. Lancet 1992;339:1538.
34. Poller W, Faber JP, Weidinger S, et al. *A leucine-to-proline substitution causes a defective alpha 1-antichymotrypsin allele associated with familial obstructive lung disease*. Genomics 1993;17:740-743.
35. Ishii T, Matsuse T, Teramoto S, et al. *Association between alpha-1-antichymotrypsin polymorphism and susceptibility to chronic obstructive pulmonary disease*. Eur J Clin Invest 2000;30:543-548.
36. Huang SL, Su CH, Chang SC. *Tumor necrosis factor-alpha gene polymorphism in chronic bronchitis*. Am J Respir Crit Care Med 1997;156:1436-1439.
37. Sakao S, Tatsumi K, Igari H, Shino Y, Shirasawa H, Kuriyama T. *Association of tumor necrosis factor alpha*

- gene promoter polymorphism with the presence of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001;163:420-422.
38. Vernooy JH, Küçükaycan M, Jacobs JA, et al. Local and systemic inflammation in patients with chronic obstructive pulmonary disease: soluble tumor necrosis factor receptors are increased in sputum. *Am J Respir Crit Care Med* 2002;166:1218-1224.
 39. Ito I, Nagai S, Hoshino Y, et al. Risk and severity of COPD is associated with the group-specific component of serum globulin 1F allele. *Chest* 2004;125:63-70.
 40. Takabatake N, Shibata Y, Abe S, et al. A single nucleotide polymorphism in the CCL1 gene predicts acute exacerbations in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2006;174:875-885.
 41. Hersh CP, Demeo DL, Lange C, et al. Attempted replication of reported chronic obstructive pulmonary disease candidate gene associations. *Am J Respir Cell Mol Biol* 2005;33:71-78.
 42. Wilk JB, Chen TH, Gottlieb DJ, et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. *PLoS Genet* 2009;5:e1000429.
 43. Repapi E, Sayers I, Wain LV, et al. Genome-wide association study identifies five loci associated with lung function. *Nat Genet* 2010;42:36-44.
 44. Myint PK, Luben RN, Surtees PG, et al. Respiratory function and self-reported functional health: EPIC-Norfolk population study. *Eur Respir J* 2005;26:494-502.
 45. Schünemann HJ, Dorn J, Grant BJ, Winkelstein W Jr, Trevisan M. Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study. *Chest* 2000;118:656-664.
 46. Strachan DP. Ventilatory function, height, and mortality among lifelong non-smokers. *J Epidemiol Community Health* 1992;46:66-70.
 47. Young RP, Hopkins R, Eaton TE. Forced expiratory volume in one second: not just a lung function test but a marker of premature death from all causes. *Eur Respir J* 2007;30:616-622.
 48. Miller LA, Wert SE, Clark JC, Xu Y, Perl AK, Whitsett JA. Role of Sonic hedgehog in patterning of tracheal-bronchial cartilage and the peripheral lung. *Dev Dyn* 2004;231:57-71.
 49. Hancock DB, Eijgelsheim M, Wilk JB, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet* 2010;42:45-52.

✉ **Correspondencia:**

Dr. Ramcés Falfán-Valencia,
Laboratorio de HLA, Departamento
de Inmunogenética. Instituto
Nacional de Enfermedades
Respiratorias Ismael Cosío Villegas.
Calzada de Tlalpan 4502, colonia
Sección XVI. México, D.F., 14080
Teléfono (52 55) 5487 1700,
extensión 5152. Correo electrónico:
dcb_rfalfanv@hotmail.com