



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

Facultad de Medicina

Identificación de marcadores genéticos asociados
al desarrollo y progresión de enfermedad
pulmonar obstructiva crónica mediante
genotipificación de regiones candidato.

TESIS

QUE PARA OBTENER EL GRADO ACADÉMICO DE

**MAESTRA EN CIENCIAS BIOLÓGICAS
(BIOLOGÍA EXPERIMENTAL)**

P R E S E N T A

GLORIA PÉREZ RUBIO

TUTOR PRINCIPAL DE TESIS:

DR. RAMCÉS FALFÁN VALENCIA

COMITÉ TUTOR: DR. OSVALDO MUTCHINICK BARINGOLTZ

DRA. ANA MARÍA SALAZAR MARTÍNEZ

TUTOR INVITADO: DRA. IRMA SILVA ZOLEZZI

MÉXICO, D.F.

MAYO DE 2011



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Dr. Isidro Ávila Martínez
Director General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día 28 de febrero de 2011, se aprobó el siguiente jurado para el examen de grado de **MAESTRA EN CIENCIAS BIOLÓGICAS (BIOLOGÍA EXPERIMENTAL)** de la alumna **PÉREZ RUBIO GLORIA** con número de cuenta 92285301 con la tesis titulada **"IDENTIFICACIÓN DE MARCADORES GENÉTICOS ASOCIADOS AL DESARROLLO Y PROGRESIÓN DE ENFERMEDAD PULMONAR OBSTRUCTIVA CRÓNICA MEDIANTE GENOTIPIFICACIÓN DE REGIONES CANDIDATO."**, realizada bajo la dirección del: **DR. RAMCES FALFÁN VALENCIA.**

Presidente: DR. MOISÉS SELMAN LAMA
Vocal: DRA. PATRICIA OSTROSKY SHEJET
Secretario: DR. OSVALDO MUTCHINICK BARINGOLTZ
Suplente: DRA. MARTHA E. PÉREZ RODRÍGUEZ
Suplente: DR. CITLALTEPETL SALINAS LARA

Sin otro particular, me es grato enviarle un cordial saludo.

Atentamente
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, D.F., a 11 de mayo de 2011

M. del Coro Arizmendi
DRA. MARIA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA

Agradecimientos

Primeramente al Posgrado en Ciencias Biológicas, UNAM por haberme permitido continuar con mi preparación académica.

Al CONACYT por haberme otorgado una beca como alumno de maestría, con número de registro de becario: 225531.

Al Fondo Sectorial de Investigación en Salud (FOSIS) por el apoyo otorgado para la realización de éste proyecto, clave: SALUD-2008-C01-87380

A los miembros del comité Tutor, Dr. Ramcés Falfán Valencia, Dr. Osvaldo Mutchinick Baringoltz, Dra. Ana María Salazar Martínez y Dra. Irma Silva Zolezzi por el apoyo recibido durante la elaboración de éste trabajo así como por el tiempo invertido y sus conocimientos compartidos.

Agradecimientos

A los miembros de mi jurado, Dr. Selman, Dra. Ostrosky, Dr. Mutchinick, Dra. Pérez y Dr. Salinas por el tiempo invertido para que éste trabajo sea mejor.

A la UNAM por ser responsable de mi progreso académico y personal.

Al INER por permitirme haber realizado éste trabajo.

Al QFB Ángel Camarena por su apoyo incondicional y aceptarme en su grupo de trabajo.

A la Dra. Alejandra Ramírez y todo el personal de la clínica de EPOC así como al Dr. Raúl Sansores y personas que forman parte de la clínica de ayuda para dejar de fumar por su invaluable apoyo para la realización de éste trabajo.

En el INMEGEN a la unidad de genotipificación y análisis de expresión Illumina y al laboratorio de genómica computacional, en especial a la M en C. Fabiola Morales y M en C. Juan Carlos Fernández, gracias por su apoyo.

Y a todas las personas que de forma indirecta o indirecta contribuyeron a la realización del presente trabajo.

Agradecimientos

A mi madre Sra. Gloria Rubio Gutiérrez por ser mi ejemplo de vida, por ese apoyo incondicional y por darme ese amor que afortunadamente no se te acaba. Te quiero.

A mi esposo, por ser parte de mi vida, por tu enorme confianza y por haber iniciado algo más que un proyecto de vida juntos. Te amo.

A mis hijos Danna e Ian que se han convertido en mi motor de vida, por haberme enseñado otra forma de amor y por soportar a la madre que les toco. Los amo.

A mi papá Sr. Victoriano Pérez Muñoz por tu cariño, por estar a mi lado. Te quiero.

A mis hermanos Víctor y Susana, porque juntos hemos vivido alegrías y tristezas, por sentirme querida y apoyada.

A mis amigos y compañeros de trabajo de los cuales sería muy difícil de nombrar a todos sin que alguno me hiciera falta, gracias por estar a mi lado.

Al resto de mi familia de nacimiento así como con aquella con la que me case, gracias por su cariño y por estar formando parte de mi vida,

“A los fracasos no hay que darles tiempo, es lo peor que puede hacerse porque enseguida lo impregnan y contaminan todo.

“Javier Marías.

Índice

Resumen	1
Abstract	2
1.0. Introducción	3
1.1. Epidemiología	4
1.2. Generalidades	5
1.3. Fenotipos	7
1.4. Etiopatogenia	11
1.4.1. Inflamación.....	12
1.4.2. Equilibrio alterado de proteasas-antiproteasas.....	15
1.4.3. Estrés oxidativo.....	17
1.5. Aspectos genéticos de la EPOC	18
1.5.1. Estudios asociación en la era genómica	22
1.5.2. Los estudios de asociación genética y EPOC.....	25
1.5.3. Los GWAS en el estudio de la EPOC.....	33
2.0. Hipótesis	35
3.0. Objetivo General	36
3.1. Objetivos Específicos.....	36
4.0. Material y Métodos	37
4.1. Población de estudio.....	37
4.1.1. Definición de caso y control:.....	37
4.1.2. Criterios de inclusión, no inclusión y exclusión empleados.....	37
4.2. Procedimientos de obtención de muestras biológicas, selección, diseño y genotipificación. 38	
4.2.1. Obtención de material genético y genotipificación.	38
4.2.2. Diseño del microarreglo.....	39
4.2.2.1. Selección de SNPs asociados a EPOC.....	39
4.2.2.2. Selección de SNPs etiqueta dentro de regiones genómicas de interés.....	40
4.2.3. Selección de SNPs marcadores informativos de ancestría (AIMs).....	42
4.3. Genotipificación por GoldenGate.....	44
5.0 Resultados	46
5.1. Participantes en el estudio.....	46
5.2. Obtención de genotipos.....	46
5.3. Generación de archivos.....	47
5.4. Análisis de resultados.....	48
5.4.1. Análisis de ancestría.....	48

5.4.2. Análisis de asociación.....	51
6.0 Discusión	57
7.0 Conclusiones.....	62
8.0 Bibliografía consultada	63
Anexo A.....	71
Anexo B.....	73

Índice de figuras

Figura 1. Diagrama de Venn propuesto por la ATS donde esquematizan lo difícil que puede ser definir la EPOC.	8
Figura 2. Etiopatogenia propuesta para el desarrollo de EPOC.	13
Figura 3. Esquema de la técnica de genotipificación masiva.	45
Figura 4. Genotipificación del software <i>GenomeStudio</i>	47
Figura 5. Análisis de ancestría por Eigensoft.	49
Figura 6. Empleo de ADMIXTURE para estimar estructura poblacional.	50
Figura 7. Bloques de haplotipos generados con los SNPs asociados a EPOC.	55

Índice de tablas

Tabla 1. Estudios de asociación positiva con EPOC usando la estrategia de gen funcional.....	28
Tabla 2. Estudios de asociación positiva con EPOC usando la estrategia de GWAS o mapeo fino en regiones candidato..	29
Tabla 3. Regiones evaluadas en Haploview para la búsqueda de SNPs etiqueta.	41
Tabla 4. Segmento de una tabla de trabajo para la búsqueda de SNPs etiqueta.	42
Tabla 5. Distribución en los cromosomas de los SNPs de ancestría empleados en este estudio ...	43
Tabla 6. Descripción de la población empleada para la genotipificación.....	46
Tabla 7. Comparación del valor F_{st} entre grupos de estudio vs caucásicos y zapotecos.....	49
Tabla 8. Proporciones ancestrales promedio empleando ADMIXTURE.....	51
Tabla 9. Resultados del análisis de asociación genética.....	53
Tabla 9. Continuación de resultados del análisis de asociación genética.	54
Tabla 10. Haplotipos asociados a EPOC encontrados en población mestiza mexicana.....	56

Resumen

Introducción: La Enfermedad Pulmonar Obstructiva Crónica (EPOC) es un padecimiento prevenible y tratable, se caracteriza por la limitación del flujo aéreo que no es completamente reversible. La OMS señala que para el 2020 la EPOC causará aproximadamente 10 millones de muertes al año, el factor de riesgo común para desarrollar EPOC es la exposición al humo del cigarro; no obstante, solo el 15-20% de los fumadores desarrollan la enfermedad; múltiples estudios han evaluado la participación individual de marcadores genéticos (como polimorfismos de un solo nucleótido, SNPs) en el desarrollo de la patología.

Objetivo: Identificar polimorfismos genéticos tipo SNP asociados al riesgo o protección de desarrollar enfermedad pulmonar obstructiva crónica por medio de mapeo fino de regiones genómicas candidato.

Material y métodos: Se seleccionaron dos grupos, uno 157 pacientes con EPOC secundaria a tabaquismo y 169 fumadores sanos como grupo control; se diseñó un microarreglo con 1536 SNPs, el cual incluyó 458 polimorfismos previamente reportados en la literatura con asociación a EPOC en otras poblaciones, 253 marcadores informativos de ancestría (AIMs) y 711 SNPs etiqueta en 24 regiones seleccionadas para mapeo fino. Para la genotipificación fue empleada la plataforma comercial de alta capacidad GoldenGate de Illumina.

Resultados: La ancestría poblacional se evaluó mediante el uso de AIMs, mostrando un índice de fijación <0.01 ($F_{st} = 0.000$), lo que indica que entre la población de casos y controles no existe estratificación poblacional. El análisis de asociación genética, mostró 36 SNPs con asociación genética en 17 regiones genómicas; de los cuales 18 se asocian a riesgo y 18 se asocian a protección, entre ellos se encuentran genes que codifican para mediadores de la inflamación como *TNF* (rs2844480), *TNFSF8* (rs11791315, rs979584) e *IL6R* (rs4072391); adicionalmente se encontró un SNP en *EPHX1* (rs2234922) y genes que intervienen en respuesta inmunológica como *SERPINE2* (rs7597833, rs16865390, rs7605945, rs7608941, rs6712954, rs6738983), *SERPINE1* (rs2227631), *ADAM33* (rs598418), *ADAM19* (rs6579952, rs17565170, rs17565365, rs17054697, rs6875485, rs11134819) e incluso genes de los cuales aún no se tiene claro el papel que desempeñan dentro de la patología como *RHOB*, *FAM13A*, *ARRDC3*, *CDKN1A*, *VPS13A* y *CBLN2*; todos los anteriores con valores de $p < 0.05$. La generación de haplotipos revela para población mestiza mexicana 4 bloques de los cuales 3 son de protección y uno de ellos es de susceptibilidad ($OR=1.61$), tal haplotipo se encuentra dentro del gen *RHOB* que codifica para una proteína involucrada en el proceso de apoptosis.

Conclusiones: Existe una gran cantidad de estudios que exploran asociaciones genéticas de polimorfismos aislados en la EPOC, principalmente evaluando los contextos "clásicos" de la enfermedad, entre estos el desequilibrio en los mecanismos de proteasas y antiproteasas, oxidantes y antioxidantes así como procesos inflamatorios anormales locales y sistémicos, adicionalmente existen estudios de genoma completo que revelan asociaciones en genes que previamente no habían sido relacionados a EPOC; realizados principalmente en poblaciones caucásicas y asiáticas; en el presente estudio se explora por primera vez en población mestiza mexicana la asociación de múltiples SNPs en genes candidato, así como en aquellos genes que han sido asociados por genoma completo. Se logran identificar nuevos marcadores asociados a la enfermedad, incluso se encuentran señales de asociación en genes que aún no está claro del papel que juegan en la patología; particularmente interesante resulta el hallazgo de factores genéticos involucrados en la apoptosis celular que pueden contribuir al establecimiento de la EPOC, mecanismo biológico poco explorado.

Abstract

Introduction: Chronic Obstructive Pulmonary Disease (COPD) is a preventable and treatable condition; airflow limitation not fully reversible is the principal pulmonary component. WHO denotes that by the year 2020, COPD will cause approximately 10 million deaths per year, exposure to cigarette smoke is the common risk factor for developing this disease; however, only 15-20% of smokers develop COPD, several studies have evaluated the individual participation of genetic markers (such as single nucleotide polymorphisms, SNPs) in the development of the pathology.

Aim: To identify single nucleotide polymorphisms associated either with risk or protection for developing chronic obstructive pulmonary disease by fine mapping of candidate genomic regions.

Material and methods: Two groups were included: 157 patients with COPD secondary to cigarette smoking and 169 smoking healthy people as control group, we designed a microarray which evaluates 1536 SNPs, including 458 polymorphisms previously reported in the literature to be associated with COPD in other populations, 253 ancestry informative markers (AIMs) for Mexican population, and 711 tag SNPs in 24 regions selected for fine mapping. genotyping was performed by the high-capacity commercial platform Illumina GoldenGate.

Results: The ancestry was assessed using AIMs, showing a fixation index <0.01 ($F_{st} = 0.000$), indicating that there is none stratification between the case and the control populations. Genetic association analysis showed 36 SNPs with genetic association through 17 genomic regions, 18 of them are associated to confer risk meanwhile the rest 18 are associated to confer protection, also, some of them correspond to genes that codify for inflammatory mediators such as TNF (rs2844480) TNFSF8 (rs11791315, rs979584) and IL6R (rs4072391), and additionally, we found a SNP in EPHX1 gene (rs2234922) and genes involved in the immune response as SERPINE2 (rs7597833, rs16865390, rs7605945, rs7608941, rs6712954, rs6738983), SERPINE1 (rs2227631) ADAM33 (rs598418), ADAM19 (rs6579952, rs17565170, rs17565365, rs17054697, rs6875485, rs11134819); and even some genes which still have no clear role in the pathology like RHOB, FAM13A, ARRDC3, CDKN1A, and CBLN2 VPS13A; all genes previously mentioned with p values <0.05 . The haplotypes generation for Mexican mestizo population revealed 4 blocks of which 3 are protective and one of them is associated to susceptibility (OR = 1.61), this haplotype is found within the RHOB gene that codes for a protein involved in the apoptosis process.

Conclusions: There are a number of studies exploring genetic associations of individual polymorphisms in COPD, mainly evaluating the "classic" contexts of the disease, such as the imbalance of proteases and antiproteases mechanisms, oxidants and antioxidants, abnormal inflammatory local and systemic phenomena; in addition, there are several genome-wide studies that reveal associations in genes that had not been previously linked to COPD, primarily conducted in Caucasian and Asian populations; so the present study explores for the first time in the Mexican mestizo population the association of multiple SNPs located in candidate genes and even in those genes that have shown association in complete genome studies. It is possible to identify new markers related to the disease, including association signs found in genes whose roles are not clear in the pathology yet; particularly interesting is the discovery of genetic factors involved in apoptosis that may contribute to the establishment of COPD, biological mechanism poorly explored.

1.0. Introducción

La Enfermedad Pulmonar Obstructiva Crónica (EPOC) es un padecimiento prevenible y tratable, su componente pulmonar es caracterizado por la limitación del flujo aéreo el cual no es completamente reversible, por lo general es progresiva y se asocia con una respuesta inflamatoria anormal en el pulmón ocasionado por la exposición a partículas o gases nocivos. Además se presentan efectos extra-pulmonares severos que pueden contribuir a la gravedad del paciente (1).

El factor de riesgo común en el mundo para padecer EPOC es la exposición al humo del cigarro; no obstante, el tabaco no es causa suficiente ya que solo el 15-20% de los fumadores desarrollan la enfermedad y tampoco es causa necesaria porque algunos enfermos nunca han sido fumadores. Los estudios epidemiológicos demuestran claramente que hay una agregación familiar por la enfermedad que no puede explicarse completamente por los factores ambientales. Hasta la fecha la única causa genética directamente asociada con el desarrollo de EPOC es la deficiencia de la enzima alfa-1-antitripsina (AAT), presente en un porcentaje muy pequeño de la población enferma, de manera que sólo se le atribuye una proporción reducida (1%) de los casos, particularmente en población europea (1).

En países en vías de desarrollo el uso de biomasa como combustible es común, el grupo de Pérez-Padilla en la Ciudad de México en 1996 realizó un estudio donde observaron que aquellas mujeres que se exponían mayor tiempo al humo de leña presentaban mayores síntomas respiratorios y reducción su función pulmonar en comparación con aquellas que la exposición a la biomasa era menor (2).

Se debe considerar evaluar a una persona para descartar el padecimiento de EPOC si el paciente presenta disnea, tos crónica, una historia de exposición a factores de riesgo,

especialmente al humo de tabaco y/o de leña; para ello es indispensable considerar el empleo de una espirometría (3).

La espirometría es una prueba simple que mide la cantidad de aire que una persona puede exhalar en función del tiempo. Con esta prueba se obtienen mediciones importantes tales como Capacidad Vital Forzada (FVC), Volumen Espiratorio Forzado en el primer segundo (FEV1) y la relación FEV1/FVC. Cada valor se debe ajustar según el paciente tomando en cuenta su género, edad, peso, estatura y origen ancestral (entendido como posibles contribuciones raciales parentales, en términos de antropología de poblaciones). El FEV1 indica el volumen espirado en el primer segundo desde que comienza la espiración; tal valor se encuentra disminuido en los procesos obstructivos clásicos como la EPOC. La FVC es el máximo aire que puede ser espirado de forma forzada tras una inspiración máxima. Y el cociente FEV1/FVC es el porcentaje de la capacidad vital forzada que se espira en el primer segundo. En condiciones normales éste porcentaje es del 75-80%; sin embargo en los pacientes con EPOC el valor obtenido post-broncodilatador es menor o igual al 70% (1).

1.1. Epidemiología

Como resultado del incremento en el consumo de tabaco, la EPOC es frecuente tanto en países en vías de desarrollo como desarrollados, sin embargo, los datos epidemiológicos sobre la enfermedad son escasos en la mayor parte del mundo (4). La OMS señala que para el 2020 la EPOC causará aproximadamente 10 millones de muertes al año (5).

Tovar y colaboradores publicaron en 2005 un análisis de datos sobre las muertes debidas a EPOC en México dentro del periodo de 1980-2002. Los resultados evidenciaron un incremento del 65% entre 1980 y 2002, en 1980 se reportaron 27 muertes por EPOC por cada 100,000 mujeres en comparación de 2002 en donde el número de muertes aumentó hasta 43 por cada 100,000. Además lograron ubicar las zonas geográficas con mayor

número de fallecimientos por EPOC siendo las zonas del norte y centro de la república las que presentaron mayor riesgo de tener dicho padecimiento y ello lo atribuyeron a que la región del norte tiene como vecino a los Estados Unidos, mientras que las zonas del centro tienen a su alrededor estados productores de tabaco. Los autores sugieren que la prevalencia y mortalidad de la enfermedad se encuentra subestimada en nuestro país pero aún así las cifras reflejan que el riesgo de padecer la enfermedad va en aumento (6). El reporte PLATINO (Proyecto Latinoamericano de Investigación en Obstrucción Pulmonar) de 2008 estima que en México la enfermedad se encuentra subdiagnosticada en un 6.9% y el mismo organismo reporta que la prevalencia de EPOC en la población mexicana mayor de 40 años es de aproximadamente 10% (5). Datos del INER ubican a la enfermedad en el cuarto lugar de morbi-mortalidad anual, cifras que se deben tomar en cuenta ya que es una enfermedad que representa altos costos para millones de individuos pues limitan al paciente en sus capacidades tanto funcionales como laborales (7). En el IMSS, en 2001 se calculó que la atención médica de la enfermedad al año por paciente cuesta en promedio \$73,303.00 pero puede aumentar dependiendo de la gravedad clínica del paciente (8).

1.2. Generalidades

En Estados Unidos, la Iniciativa Global para la Enfermedad Pulmonar Obstruictiva Crónica (GOLD, por sus siglas en inglés: *Global Initiative for Chronic Obstructive Lung Disease*) se encarga de concientizar sobre la importancia de la EPOC estableciendo estrategias de diagnóstico, tratamiento y prevención con la finalidad de implementar guías para su adecuado manejo.

La GOLD define a la EPOC como una enfermedad prevenible y tratable que se caracteriza por la limitación al flujo de aire, no totalmente reversible, progresiva y que se asocia con una respuesta inflamatoria anormal por parte del pulmón hacia las partículas o gases nocivos. Se clasifica en 4 estados en relación a los valores obtenidos en una

prueba de espirometría simple tras el uso de broncodilatador, con las siguientes características (1):

Estadio I

Limitación media del flujo de aire ($FEV1/FVC < 70\%$; $FEV1 \geq 80\%$ del predicho) y en ocasiones tos y producción de esputo. En tal estado el paciente no necesariamente sabe que su función pulmonar es anormal.

Estadio II

EPOC moderada, empeora la limitación del flujo de aire ($FEV1/FVC < 70\%$; $50\% \leq FEV1 < 80\%$ predicho), con limitación leve en el flujo de aire hasta que se convierte en problemática. En éste estado los pacientes por lo general buscan atención médica debido a síntomas respiratorios crónicos y en algunos casos la presencia de exacerbaciones.

Estadio III

EPOC severa, empeoramiento gradual a la limitación al flujo de aire ($FEV1/FVC < 70\%$; $30\% \leq FEV1 < 50\%$ predicho), mayor limitación en el flujo de aire, capacidad reducida para el ejercicio, exacerbaciones repetidas que deterioran la calidad de vida del paciente.

Estadio IV

EPOC muy severa ($FEV1/FVC < 70\%$; $FEV1 < 30\%$ predicha) Este estado se caracteriza por la falta de respiración crónica.

Los síntomas de la EPOC incluyen: tos, producción de esputo y disnea por esfuerzo

A la fecha se sabe que la EPOC es una enfermedad compleja con alto componente genético, además de que se relaciona a otros factores de riesgo exógeno tales como (1, 5): humo de tabaco, exposición ocupacional prolongada o intensa a vapores, irritantes y humos, exposición en lugares cerrados a la combustión de biomasa.

La EPOC es considerada una enfermedad inflamatoria multisistémica con órganos blanco extrapulmonares tales como corazón, músculo y esqueleto. Existe pérdida de masa muscular por una combinación asociada a hipoxemia crónica, inmovilidad y aumento de la tasa metabólica, lo que, en conjunto, contribuye a mayor disnea, menor resistencia y disminución en la calidad de vida de aquellas personas que la padecen (9).

El curso progresivo de la EPOC se complica con exacerbaciones cada vez más frecuentes. Durante éstas, el flujo espiratorio se reduce, pero el cambio fisiopatológico fundamental es un empeoramiento del intercambio gaseoso. El diagnóstico de una exacerbación se establece al existir empeoramiento clínico agudo, que necesita un cambio en el tratamiento regular (10).

Existen factores de riesgo que aumentan los ingresos hospitalarios de un paciente tales como una baja función pulmonar, pacientes con edad avanzada, hipertensión pulmonar, frecuencia de exacerbaciones anteriores (11), presencia de ansiedad y depresión todo ello contribuyendo a disminuir la calidad de vida del paciente así como de sus familiares (12).

1.3. Fenotipos

La EPOC es una enfermedad de fisiopatología heterogénea, lo cual afecta el curso clínico de la misma. En la figura 1 se presenta un diagrama clásico de Venn propuesto por la Sociedad Americana de Tórax (ATS, por las siglas en inglés de *American Thoracic Society*) donde muestra la superposición de la EPOC con otras patologías, por lo cual hace difícil la clasificación y el diagnóstico certero (13).

Debido a la dificultad para caracterizar a la EPOC, en numerosos trabajos se ha propuesto establecer fenotipos de la enfermedad. Un fenotipo para fines de la EPOC secundaria al consumo de tabaco se refiere a las manifestaciones que se pueden observar o medir; tales como síntomas, exacerbaciones, respuesta a la terapia, tasa de progresión o mortalidad de la enfermedad; lo anterior tiene como finalidad identificar los

fenotipos presentes en la EPOC para así tener una mejor comprensión de la enfermedad y contar con estrategias de tratamiento adecuadas según cada paciente (14).

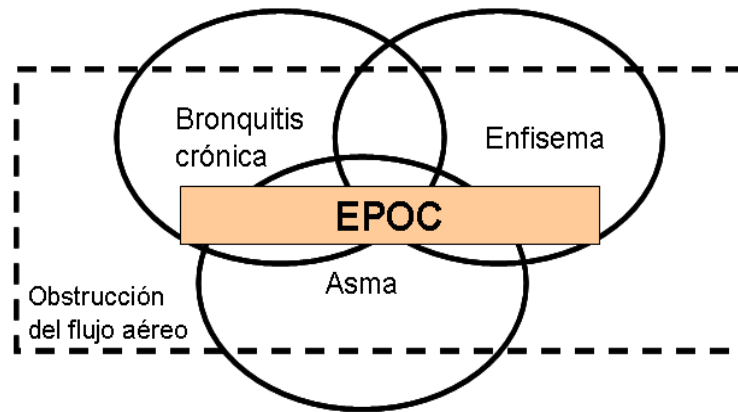


Figura 1 Diagrama de Venn propuesto por la ATS donde esquematizan lo difícil que puede ser definir la EPOC, tomada de (15).

Cabe mencionar que el fenotipo final que se presente en un sujeto depende de la genética, el ambiente y el entorno donde dicho individuo se desarrolle, puesto que se han identificado varios factores exógenos de riesgo ambientales que contribuyen a aumentar el riesgo de padecer EPOC (16).

Friedlander y colaboradores en 2007 proponen tres grandes grupos fenotípicos basados en rasgos clínicos, fisiológicos y radiológicos; en el grupo de fenotipos según rasgos clínicos existen a su vez subgrupos que indican estatus funcional como grado de disnea en base al FEV1, frecuencia de las exacerbaciones, índice de masa corporal y respuesta a corticosteroides inhalados; en tanto que para los fenotipos asignados según rasgos fisiológicos se encuentran los subgrupos según el grado en la limitación al flujo de aire, la declinación rápida de la función pulmonar e hipertensión pulmonar; por último el fenotipo asignado tomando en cuenta rasgos radiológicos basados en tomografía computarizada de alta resolución puede ser de tipo enfisematoso o bronquiolítico y a su vez dentro de

cada uno de ellos se encuentran clasificaciones adicionales según la zona pulmonar afectada (17).

Recientemente en un trabajo publicado en 2010 por Han y colaboradores señalan que la EPOC secundaria a consumo de tabaco puede ser considerada como un síndrome complejo con numerosos componentes pulmonares y extrapulmonares. Este grupo de trabajo propone que para establecer un fenotipo de la enfermedad es necesario tomar diversos puntos en base a manifestaciones clínicas, fisiológicas, caracterización radiológica, frecuencia de las exacerbaciones, evaluación de la inflamación sistémica en base a biomarcadores específicos como la proteína C reactiva y citocinas proinflamatorias así como comorbilidades tales como enfermedades cardiovasculares, osteoporosis, depresión entre otras, además proponen la importancia de validar el fenotipo asignado con la finalidad de definir subpoblaciones que puedan ser conducidas hacia una terapia específica y mejora en la calidad de vida(14).

Existe evidencia de que la expresión clínica de la enfermedad es diferente según el género. Laurin y colaboradores en 2007 señalan que las mujeres con EPOC tienen 1.5 veces más riesgo de presentar trastornos de ansiedad en comparación con los hombres con la misma patología. En cuanto a la depresión las mujeres tienen de 2 veces mayor probabilidad de formar dicho trastorno en comparación con los hombres; estos trastornos repercuten en la calidad de vida de las mujeres que padecen EPOC (18), En base al estudio PLATINO, el cual pone en evidencia que hay importantes diferencias en la expresión de EPOC según género, Lopez y colaboradores en 2010 publican que las mujeres presentan mayor disnea y limitación física en comparación con los varones estudiados independientemente del estado de la enfermedad, las mujeres con EPOC secundaria a tabaco presentan mayor obstrucción que aquellos hombres con similar exposición al cigarrillo (19); en cuanto a las comorbilidades, Barr y colaboradores en 2009 indican que predomina la depresión, osteoporosis y enfermedad cardiovascular en

mujeres en comparación de los hombres con EPOC (20). Mientras que Martínez y colaboradores en 2007 publican sus resultados que provienen de una población de mujeres y hombres caucásicos con EPOC secundaria a consumo de tabaco, ellos encuentran que las vías respiratorias pequeñas de las mujeres estudiadas presentan paredes más gruesas y el lumen de la vía respiratoria desproporcionadamente reducido en comparación con los hombres, sugieren que las partículas nocivas provenientes del cigarrillo se depositan de forma diferente dentro de las vías respiratorias dependiendo del género, aunado a lo anterior proponen que el carácter de la respuesta inflamatoria al humo del cigarrillo es probable que se vean afectados por las hormonas sexuales (21). Existe evidencia de que la progesterona induce cambios cíclicos en la producción de interleucina 8 (IL-8) y promueve una respuesta de citocinas que provienen de las células T colaboradoras tipo 2 (Th2) (22). Adicionalmente las variaciones en las concentraciones de estradiol en las mujeres, se han asociado con alteraciones en la densidad del receptor adrenérgico y en las concentraciones de moco, acetilcolina y prostaglandina (23).

Como se mencionó anteriormente, para caracterizar a un paciente en base a las manifestaciones fisiológicas se emplean valores que resultan de una espirometría simple, tales como FEV1, FVC y el cociente FEV1/FVC los cuales son usados para determinar la presencia de la enfermedad en estudio así como el grado de severidad y la respuesta del sujeto después de un broncodilatador. En cuanto a las manifestaciones radiológicas la propuesta es muy similar a la de Friedlander y colaboradores publicada en 2007. Para el fenotipo asignado en base a la frecuencia de las exacerbaciones se monitorea al paciente día a día a partir de una condición estable hasta la aparición de una exacerbación grave que requiera atención urgente. La presencia de inflamación sistémica puede ser un único fenotipo de la enfermedad, en tal fenotipo se determinan si los marcadores de la inflamación, como la proteína C reactiva, citocinas proinflamatorias y leucocitos se encuentran elevados en los pacientes con EPOC. Por último, pacientes con

comorbilidades incluyendo enfermedad cardiovascular, síndrome metabólico, osteoporosis, depresión, pérdida y disfunción del músculo esquelético alteran la historia natural de la EPOC así como su tratamiento (14).

Roy y colaboradores en 2009 emplea el análisis de componentes principales (PCA) para definir fenotipos de la EPOC en base a biomarcadores encontrados en el esputo de los pacientes; el PCA es el análisis más común usado que reduce un gran número de variables sin perder información y solo toma aquellas variables que representan una gran contribución al análisis. En el contexto de la EPOC los autores representan los componentes según los procesos fisiopatológicos responsables de la heterogeneidad de la enfermedad. El grupo de Roy propone 2 fenotipos en base a los biomarcadores encontrados en el esputo: en el primero existe presencia de neutrófilos que sugiere que hay inflamación de las vías aéreas asociada a inflamación sistémica, al medir los niveles de IL-8 observaron que tales niveles se asocian directamente con la gravedad de la obstrucción de flujo de aire y posterior disminución del FEV1; en el segundo fenotipo hay predominio de eosinofilia lo asocian con una mayor respuesta al tratamiento con corticosteroides (24).

Más reciente, se reportó otro PCA publicado por Burgel y colaboradores en 2010 con el objetivo de establecer fenotipos clínicos de la enfermedad, en este estudio se propone que los sujetos con EPOC se pueden clasificar dentro de 4 grupos; el primero incluye aquellos sujetos jóvenes con predominio severo o muy severo de enfermedad respiratoria; segundo: individuos mayores con limitación al flujo de aire leve así como síntomas y comorbilidades leves; el fenotipo tres incluye pacientes jóvenes con limitación al flujo de aire moderada o severa pero con síntomas y comorbilidades leves; en tanto que el cuarto fenotipo incluye sujetos mayores con limitación al flujo de aire severo a moderado, síntomas graves atribuidos en parte a la mayor presencia de comorbilidades (25).

1.4. ETIOPATOGENIA

La EPOC es resultado de la exposición crónica del pulmón a partículas y gases nocivos como el humo del cigarro o el humo procedente de la combustión de biomasa, lo que ocasiona que en los espacios aéreos se recluten células inflamatorias que posteriormente liberaran mediadores que en el microambiente ocasionan el aumento del estrés oxidativo con la consecuente degradación de la matriz extracelular y deterioro de la arquitectura pulmonar, debido a que los mecanismos de reparación se encuentran alterados. Las alteraciones anatomopatológicas características de la EPOC pueden encontrarse en las vías aéreas centrales y periféricas, el parénquima y la circulación pulmonar (26). Existe la hipótesis de que en la EPOC hay una respuesta inflamatoria alterada, un desequilibrio entre proteasas y antiproteasas así como de la concentración de oxidantes y antioxidantes en el microambiente pulmonar (Figura 2).

1.4.1. Inflamación

Los macrófagos alveolares son células que se encuentran en el tracto respiratorio inferior y participan en la iniciación, modulación y eventual resolución de la inflamación, tienen la capacidad de cambiar el fenotipo de sus moléculas de superficie en base a estímulos exógenos o bien procesos patológicos del parénquima pulmonar; ejemplos de receptores son aquellos que participan en la fagocitosis (CD44, CD36, CD51, CD61, CD14); moléculas del complejo principal de histocompatibilidad (HLA-DR); moléculas coestimuladores (CD80, CD86, CD40) o receptores tipo 3 del complemento (CD11b). Interesantemente, Pons en 2005 encontró que los macrófagos alveolares de pacientes con EPOC presentan en su superficie reducción en la expresión de HLA-DR, CD80 y CD44; tales observaciones pueden contribuir a la respuesta inflamatoria anormal de los pulmones debido a la exposición a agentes nocivos (27).

El mecanismo por el cual el humo de cigarro activa el sistema inmunológico aún se desconoce, pero se sabe que existe una glucoproteína aislada del tabaco que Francus empleó en 1988 para inmunizar ratones y observó que esta proteína tenía potente acción inmunoestimulante (28).

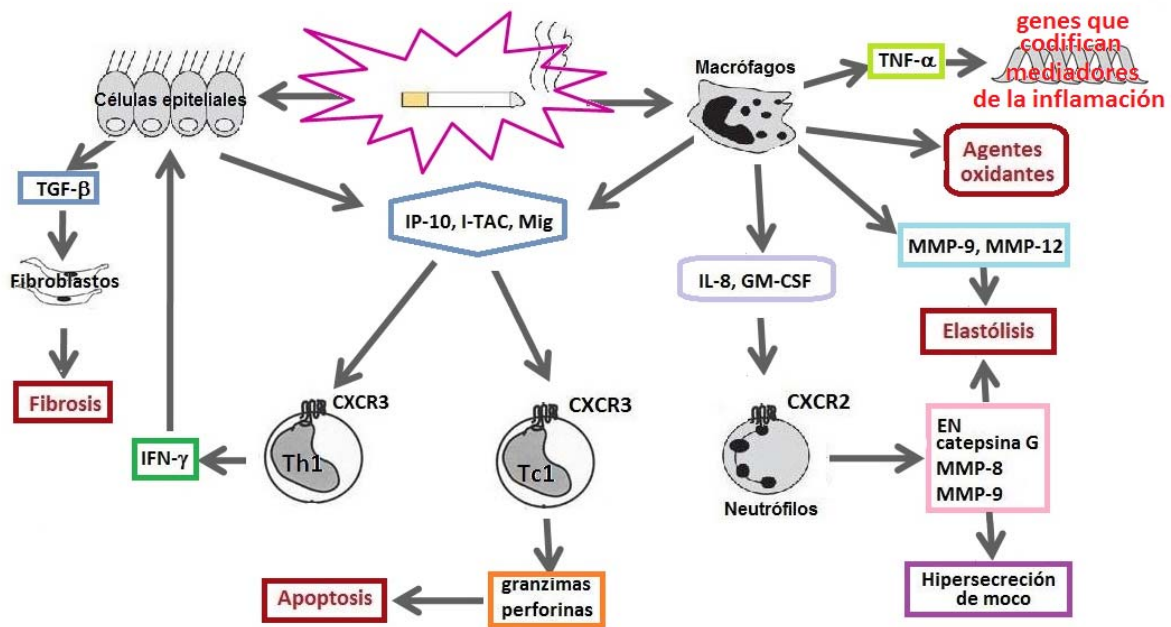


Figura 2 Etiopatogenia propuesta para el desarrollo de EPOC, modificado de Barnes y colaboradores, 2008 (29).

Por otro lado, los macrófagos al activarse con extracto de humo de tabaco liberan mediadores proinflamatorios entre ellos, factor de necrosis tumoral alfa ($TNF\alpha$), IL-8 y especies reactivas de oxígeno, quimiocinas como la proteína inducible de interferón γ (IP-10), quimiotáctico para células T inducibles por interferón (I-TAC) y monocina inducida por interferón γ (Mig), las cuales son quimiotácticas para células CD^+ Tc1, a través de la interacción con el receptor CXCR3 que se expresa en éstas células (30); también secretan enzimas elastolíticas como metaloproteinasas de matriz extracelular (MMP2, MMP9, MMP12) (31). La enzima elastolítica que predomina en pacientes con EPOC es la

MMP9 (32). Se ha observado que hay infiltración en pulmón de neutrófilos y células T citotóxicas CD8+ (Tc) y T colaboradoras CD4+ (Th) que actúan como mediadores inflamatorios, la cantidad de células activas se relaciona con el grado de factores proinflamatorios, mitogénicos y profibrosantes los cuales modulan el componente celular pulmonar (33). Entre las citocinas proinflamatorias que se han encontrado en mayor proporción en esputo y lavado bronquioalveolar (LBA) de personas con EPOC está el TNF α , interleucina 1 beta (IL-1 β) e interleucina 6 (IL-6), tales citocinas incrementan la inflamación a través de la activación del factor de transcripción NF-kB que posteriormente estimulará la expresión de otros genes inflamatorios (29). En pacientes con EPOC se ha observado que la liberación del factor estimulador de colonias de granulocitos y monocitos (GM-CSF) y de IL-8 por parte de los macrófagos alveolares contribuye a aumentar la supervivencia de neutrófilos y macrófagos en la vía aérea, primordialmente durante las exacerbaciones (29, 34). Los neutrófilos activados presentes en vías respiratorias y parénquima pulmonar secretan proteasas de serina, entre ellas elastasa de neutrófilos (EN), catepsina G y proteína 3 así como metaloproteinasa de matriz 8 y 9 (MMP8 y MMP9) las cuales podrían contribuir a la destrucción alveolar; es probable que la neutrofilia de vías respiratorias se vincule con la hipersecreción de moco (31).

En pacientes con EPOC las células que predominan son las T CD8+, y en pacientes con la enfermedad más severa se han encontrado folículos linfoides que contienen un centro con linfocitos B rodeados de células T y deposición de colágeno en la pared externa de la vía respiratoria. Es de esperar que en biopsias de pacientes con EPOC se encuentre un incremento de interferón gamma (IFN- γ) e interleucina 2 (IL-2) que proviene de las células Th1; el IFN- γ en pulmón de los pacientes con EPOC mantiene la respuesta de las células T y promueven el proceso de inflamación junto con neutrófilos y macrófagos (29). La presencia de células T citotóxicas liberan granzimas y perforinas que inducen apoptosis

(35). Las células Th1 que secretan IL-2 e IFN γ expresan distintos receptores, tales como CCR5 y CXCR3; en contraste las células Th2 producen interleucina 4 (IL-4) e interleucina 5 (IL-5) y expresan CCR4 y CCR3; el receptor fundamentalmente expresado en células T es el CCR3. Grumelli en 2004 publica sus hallazgos donde al caracterizar linfocitos T de pacientes con EPOC concluye que las células Th1 controlan la respuesta fisiológica y el cambio estructural a nivel pulmonar ya que son las que predominan en el microambiente (36).

Aunque el tabaquismo sea la principal causa de padecer EPOC, no parece que al abandonar el hábito se resuelva la respuesta inflamatoria en las vías respiratorias especialmente en etapas avanzadas de la enfermedad; esto sugiere que existen mecanismos que mantienen el proceso inflamatorio crónico cuando éste ya se ha establecido (31, 37).

1.4.2. Equilibrio alterado de proteasas-antiproteasas.

Se ha propuesto que varias proteasas descomponen los compuestos del tejido conectivo, especialmente la elastina dado que se produce una pérdida de elasticidad en el parénquima pulmonar (31). La exposición al humo de tabaco provoca que en el pulmón se incremente el número de neutrófilos y macrófagos, se sabe que en personas con deficiencia de (AAT) el daño es mayor ya que se trata de una proteína que inhibe la elastasa liberada por los neutrófilos y en el caso de células estimuladas por citocinas, se podría expresar sobre la superficie celular (36). La elastasa es una potente enzima elastolítica, en animales de experimentación, al ser inyectada les induce enfisema (38) y es capaz de inducir la expresión de IL8 en células epiteliales de la vía respiratoria (39). Los neutrófilos también almacenan otras proteasas de serina, catepsina G y proteinasa 3 en sus gránulos específicos con propiedades semejantes a la elastasa liberada por los neutrófilos (31).

Diversos estudios muestran que en pacientes con enfisema existe una mayor concentración de MMP9 (gelatinasa B), MMP1 (colagenasa), MMP8, MMP2 y MMP12 en comparación con un grupo control (40). Algunos mediadores proinflamatorios inducen la liberación de MMPs sin provocar aumento del inhibidor tisular de metaloproteinasa (TIMP), dando origen a un posible desequilibrio entre proteasas y antiproteasas, (41, 42) los TIMPs son inhibidores endógenos de MMPs (43). Los macrófagos alveolares de fumadores normales expresan más MMP9 que los sujetos normales y se presenta un aumento aún mayor en células de pacientes con EPOC (44). Finalay y su grupo en 1997 describieron por primera vez el incremento de MMP9 en pacientes con EPOC (45), desde entonces existen numerosos trabajos que refieren a la MMP9, esta proteína forma parte de una familia de enzimas colagenasas que regulan la homeostasis de la matriz pulmonar la cual se encuentra constituida mayoritariamente por colágena y elastina, la enzima puede ser producida por células alveolares tipo II, células endoteliales, leucocitos pulmonares, macrófagos, eosinófilos, y células NK (46).

Avilés y su grupo en un trabajo publicado en 2006 encontraron que la concentración de MMP9 en esputo de pacientes con EPOC es significativamente elevada en comparación con sujetos fumadores que no padecen EPOC también determinaron que la concentración del inhibidor tisular de MMP de tipo 1 fue menor en los sujetos enfermos (47). En un estudio realizado por Lomas y colaboradores en 2008 se evaluaron los niveles de proteína 16 secretada por células de Clara (CC-16); dichas células se encuentran en el recubrimiento epitelial de los bronquiolos, CC-16 proporciona protección contra agentes oxidantes. Los resultados muestran que en individuos fumadores, asmáticos, con bronquiolitis así como enfisematosos, los niveles de la proteína en suero se encuentran por debajo de los niveles normales en comparación con su grupo control por lo que proponen a la proteína como un biomarcador del daño del epitelio bronquial; interesantemente concluyen que, los niveles de la proteína en individuos con EPOC se

encuentran disminuidos y están relacionados con la gravedad de la enfermedad (48). Normalmente las proteasas se compensan con un exceso de antiproteasas endógenas, las principales son AAT en parénquima pulmonar e inhibidor de la leucoproteasa secretora (SLPI) derivado de epitelio de vías respiratorias, elafina y alfa 1 antitripsina (31).

1.4.3. Estrés oxidativo

El estrés oxidativo aparece cuando se producen más especies reactivas de oxígeno que mecanismos de defensa antioxidante, lo que provoca efectos perjudiciales como lesión al DNA, lípidos y proteínas, dichos eventos llevan a un aumento en la respuesta inflamatoria vía activación de factores de transcripción (31). La fuente endógena del estrés oxidativo son las células inflamatorias como neutrófilos y macrófagos que generan especies reactivas de oxígeno (ROS) y de nitrógeno (RNS); aunado a ello existen ROS y RNS exógenas que provienen del humo de cigarro que contribuyen a aumentar el ambiente oxidante. El sitio natural de producción de ROS dentro de las células son la mitocondria debido a la cadena respiratoria y los microsomas, la actividad fagocítica de macrófagos y neutrófilos también contribuye al aumento de ROS al ser estimuladas por las partículas que provienen del humo de cigarro y por mediadores de la inflamación, puesto que se activa el complejo nicotinamida adenina dinucleótido fosfato oxidasa (NADPH) generando así radicales libres (26). En pulmón se expresan los antioxidantes intracelulares en niveles relativamente bajos y no se inducen por el estrés oxidativo, más bien los antioxidantes que contribuyen a contrarrestar el exceso de oxidantes son extracelulares. En pulmón el sistema glutatión es el mecanismo antioxidante principal (49). La peroxidasa extracelular de glutatión (eGPx) es un antioxidante importante en pulmones y se podría secretar con células epiteliales y macrófagos, sobre todo como respuesta al humo de tabaco o el estrés oxidativo (31). El estrés oxidativo podría inducir apoptosis en células endoteliales y epiteliales, la apoptosis en neumocitos de tipo I podría contribuir al desarrollo del enfisema

y se podría inducir con linfocitos T citotóxicos o con la inhibición de receptores del factor de crecimiento vascular endotelial (29, 50).

En 2001 Montuschi y colaboradores emplearon el monóxido de carbono (CO) exhalado como un indicativo no invasivo del estrés oxidativo a nivel pulmonar y reportan que en los pacientes con EPOC el nivel del CO se encuentra incrementado 2.5 veces más en comparación con un grupo control de no fumadores. Bajo condiciones normales el pulmón y la sangre tienen mecanismos adecuados de protección contra los oxidantes, entre los más exitosos se encuentran las superóxido dismutasas (SODs), catalasa y enzima de la familia glutatión S transferasas (GSTs): en EPOC se sugiere un desequilibrio del nivel de oxidantes y antioxidantes que contribuyen al proceso inflamatorio y daño pulmonar (51).

1.5. Aspectos genéticos de la EPOC

Desde antes 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. 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 sólo nucleótido (SNPs por sus siglas en inglés), las cuales corresponden a sustituciones de un nucleótido por otro en la cadena de DNA (52). Se encuentra bien documentado que uno de los factores de riesgo genético para padecer EPOC es la deficiencia genética de AAT, sin embargo tal condición solo se presenta en el 1 a 2% de los pacientes con EPOC, la AAT es una glucoproteína responsable de la mayoría de la actividad antiproteasa, se encuentra en el espacio extracelular, en todos los

tejidos corporales, pero al parecer su relevancia fisiológica es en el pulmón, donde protege al alvéolo del daño proteolítico de la NE (53) que es capaz de digerir la elastina, las membranas basales y otros componentes de la matriz extracelular, también neutraliza las defensinas α del neutrófilo, el leucotrieno B4 (LTB4) y la IL8, potentes quimioatrayentes de neutrófilos, modula la adhesión de la NE al receptor fosfatidilserina de membrana del neutrófilo, proceso necesario para la apoptosis, sugiriendo un papel importante en la resolución de la inflamación. Además, posee 9 radicales de metionina, lo que le confiere una importante capacidad antioxidante (54). El gen que codifica a la proteína es altamente polimórfico pero solo algunos de ellos se asocian con bajas concentraciones de la proteína en suero (55). Cabe mencionar que históricamente las variantes se clasificaron de acuerdo con su velocidad de migración electroforética en un campo magnético con distintos gradientes de pH. Los investigadores pioneros en el tema denominaron M (del inglés, *medium*) a las de velocidad media, F (del inglés, *fast*) a las de velocidad rápida, y S (del inglés, *slow*) a las de migración lenta (54). En la actualidad, el genotipo "normal", presente en más del 90% de los sujetos sanos (94-96% en población caucásica), se denomina PiMM y se caracteriza por niveles séricos de la proteína entre 150-350 mg/100 mL; las variantes S (Glu264Val) y Z (Glu342Lys) constituyen el 95% de las mutaciones en pacientes con deficiencia severa de AAT; ambas son mutaciones sin sentido en el gen. Individuos con genotipos SS, SZ y ZZ expresan concentraciones séricas de la proteína de 85, 25 y 15%, respectivamente, con respecto a los niveles normales (genotipo MM) (56). El homocigoto para la variante S no ha sido asociado con enfermedad y son pocos los estudios enfocados a determinar el riesgo a desarrollar enfisema entre los genotipos SZ, (57) la mayoría de los estudios concluyen que el genotipo SZ es de menor importancia que ZZ, debido a que los pacientes con genotipo SZ desarrollan enfisema a una edad mayor que aquéllos portadores del ZZ (58) El deterioro progresivo de la función pulmonar en individuos homocigotos ZZ es bien conocido, y el

papel de los denominados genotipos de deficiencia intermedia, principalmente MZ y MS han sido poco estudiados. Dahl y colaboradores (59) en 2001 y, posteriormente en 2002 (60) investigaron el efecto de los genotipos de deficiencia intermedia en la población general danesa sobre la función pulmonar a través de medidas de espirometría (FEV1 y FVC) y genotipificación de los alelos de SERPINA1 en 9187 individuos; hallaron que los heterocigotos SZ tienen radios más bajos en la relación FEV1/FVC que aquellos individuos MM; mientras que los homocigotos ZZ tienen porcentajes más bajos de FEV1 y radios FEV1/FVC que el resto de los genotipos (MM, MS, SS y MZ), cuando estos resultados fueron estratificados por el estatus de fumador, la reducción en la función pulmonar en individuos SZ y ZZ vs MM fue estadísticamente significativa sólo entre fumadores y exfumadores. Los individuos MZ tienen ligera reducción de la función pulmonar comparada con aquellos pacientes MM. Esto sugiere que la heterocigocidad MZ actúa sólo en ciertos contextos, posiblemente aún no determinados. Esta observación puede explicar previos hallazgos contradictorios en diferentes estudios del genotipo MZ. En conclusión, comparado con el genotipo MM, el MZ se encuentra asociado con la reducción de la función pulmonar en individuos con EPOC, además, los genotipos SZ y ZZ se asocian con obstrucción de la vía aérea y reducción de la función pulmonar, particularmente en fumadores (59).

En una cohorte de origen suizo Senn y colaboradores, determinaron los niveles circulantes de AAT y establecieron determinantes de asociación con medidas de función pulmonar, entre sus resultados se encuentran asociaciones positivas con género femenino, ingesta de hormonas, presión sanguínea, edad y exposición al humo de cigarro activa y pasiva, mientras que la ingesta de alcohol, el índice de masa corporal correlacionan de manera inversa con los niveles de séricos de AAT. Además, encontraron una asociación inversa de AAT con FEV1 en el total de la población en estudio, misma que desaparece después de ajustar por valores de proteína C reactiva (PCR),

adicionalmente esta asociación se ve modificada por género, estatus de menopausia en mujeres y hábito de fumar. La fortaleza de este estudio radica en su diseño basado en la población general y la detallada caracterización de los participantes que permite investigar factores asociados con los niveles circulantes de AAT y medidas de función pulmonar. Basados en las determinaciones de PCR se pueden excluir sujetos con evidencia de inflamación al momento de la punción sanguínea y, de esta manera, controlar la inflamación sistémica menor como un potencial factor de confusión. Este estudio permite demostrar la complejidad de las interrelaciones entre la exposición al humo de tabaco, el género, los niveles circulantes de AAT y PCR (61).

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 enriquecimiento de las variantes en estos grupos lleva a la identificación de factores genéticos asociados al riesgo o a 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), 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 como es el caso de la EPOC (62).

Los estudios de asociación genética se enfrentan al reto de que las enfermedades complejas en relación a 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 (fenotipos), así como la existencia de interacciones de los genotipos con el medio ambiente (63). 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 novedosas para lograr incrementar el poder de este tipo de estudios (64).

1.5.1. Estudios asociación en la era genómica

Uno de los principales resultados del Proyecto del Genoma Humano (HGP, siglas en inglés de *Human Genome Project*) el cual tuvo como objetivo caracterizar la secuencia completa del DNA, fue la identificación de variantes genéticas que contribuyen a la individualidad genética humana. Adicionalmente, el proyecto HapMap (acrónimo 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 (65, 66). En lo que respecta a población mexicana, en 2009 el Instituto Nacional de Medicina Genómica

(INMEGEN) 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*), su éxito depende de una eficiente selección de SNPs (67). Los GWAS han generado conocimiento de importancia en la identificación de los riesgos genéticos a enfermedades complejas (65, 68); interesantemente, muchos de los genes recientemente identificados previamente no habrían sido “sospechosos” de tener algún papel importante en la biología de la enfermedad en estudio, incluso se han encontrado 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 de 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 (65). El diseño más frecuente de los GWAS es el de casos y controles, donde es crítico y fundamental seleccionar el número adecuado de participantes, que típicamente es de varios cientos o miles de individuos dentro de cada grupo, y el cual está en razón de la prevalencia de la enfermedad en la población en estudio, así como las estimaciones en relación a los posibles niveles de riesgo genético y 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 tener criterios estrictos para incluirlos dentro del grupo adecuado (63, 68).

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 variantes genéticas responsables de los efectos biológicos relacionados a la enfermedad, esto mediante la detección haplotipos (combinaciones lineales de SNPs dentro de un gen) de alto riesgo o bien de protección, estudios de asociación en genes y/o regiones candidato; así como estudios epidemiológicos para buscar identificar interacciones gen-gen, (69). En el análisis de los resultados de un GWAS es difícil 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 (LD, por sus siglas en inglés *linkage disequilibrium*). 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 inglés científico se denomina “*tagging*”, es decir etiquetado y a cada SNP con esta característica se le denomina SNP etiqueta. 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 (62). 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 SNPs etiqueta. Lo anterior se vuelve especialmente crítico en el diseño de estudios en regiones y genes candidato que buscan caracterizar con detalle las variantes genéticas asociadas a los rasgos y enfermedades, y aún más cuando se examinan poblaciones donde la estructura genética y las frecuencias de las variantes genéticas difieren de la población donde se identificaron originalmente (68). 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 resultan de una asociación espuria (69).

Para evaluar adecuadamente las contribuciones al riesgo de los factores genéticos, así como para identificar su interacción con factores ambientales, es deseable que 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 sobrevivencia, entre otros. Los estudios de cohorte presentan ventajas pero son costosos y su logística puede ser difícil de realizar (62).

A la fecha un gran número de GWAS han sido publicados y muchos más están en proceso (64), 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; este 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 (64).

1.5.2. Los estudios de asociación genética y EPOC.

Los estudios de asociación en casos y controles son los más empleados para estudiar enfermedades respiratorias como la EPOC (63, 68). Múltiples estudios han evaluado la participación individual de marcadores genéticos (SNP y otros marcadores polimórficos) en el desarrollo de EPOC y ya se han llevado a cabo algunos GWAS que buscan nuevos SNPs tratando de hallar hipótesis alternas y novedosas a los mecanismos genéticos subyacentes de la patología (16), algunos de los genes y SNPs más claramente asociados se muestran en las tablas 1 y 2, resulta importante destacar que las asociaciones de genes individuales estudiados a la fecha cubren vías biológicas que se sabe participan en el padecimiento de EPOC incluyendo, genes del metabolismo de oxidantes y xenobióticos, de la respuesta inmunológica incluyendo inflamación, genes que originan proteínas que participan en el sistema proteasas/antiproteasas, o en mecanismos de señalización involucrados en el establecimiento de daño pulmonar que lleva al desarrollo de la patología pero también hay regiones que previamente no habían sido relacionados a la patología.

Dentro de los genes estudiados y asociados a EPOC que participan en el metabolismo de xenobióticos se encuentran los miembros de la familia de las glutatión S-transferasas, las cuales son enzimas que conjugan compuestos electrofílicos a glutatión para su posterior metabolismo y excreción del organismo, la subfamilia omega (*GSTO*) tiene dos miembros funcionales, *GSTO1* y *GSTO2* ambos tienen sus genes ubicados en el cromosoma 10 humano y están separados por 7.5kb; el gen *GSTO1* tiene 6 exones y codifica una proteína de 27.5 kDa; el gen de *GSTO2* tiene 7 exones y codifica una proteína de 28 kDa, entre ellas tienen una homología del 64% (70); ambas *GSTO* se han estudiado en pacientes con EPOC, en 2009 Yanbaeva en un estudio de casos y controles en caucásicos donde los controles eran 195 exfumadores y los casos fueron 355 sujetos con EPOC estable, secundaria a tabaquismo con gravedad de moderada a severa, reportan

que el alelo G ubicado en el rs156697 de *GSTO2* se encuentra 1.4 veces más frecuente en pacientes con EPOC a comparación del alelo A que se encuentra en los controles, el SNP no fue posible asociarlo con los resultados de función pulmonar de cada grupo, sucede lo mismo cuando se generaran los haplotipos ya que los autores proponen un haplotipo compuesto por *GSTO1140Asp/GSTO2142Asp* asociado al riesgo de padecer EPOC (71).

Otro de gen asociado a EPOC es *GSTP1*, posee un cambio A/G en el exón 5 que da como resultado la sustitución de isoleucina por valina en la posición 105 de la proteína. Lakhdar en 2010 publica sus estudios realizados en población Tunecina de 182 controles y 234 casos con EPOC con FEV1 predicho <80% donde encuentra que aquellas personas con genotipo GG poseen un riesgo relativo 2.67 veces más de padecer EPOC que aquellos homocigotos AA (72). Dentro de los genes estudiados en la EPOC y que participan en el equilibrio de oxidantes y en controlar la inflamación se encuentra la superóxido dismutasa extracelular (*SOD3*), codifica para una enzima antioxidante tetramérica que se expresa en fluidos extracelulares de algunos tejidos y plasma, en pulmón es muy abundante (73).

Tabla 1. Estudios de asociación positiva con EPOC usando la estrategia de gen funcional.

Gen	SNP			Casos/ Controles	Fenotipo de EPOC	Valor de p	OR (IC 95%)	Referencia
	rs	Alelos	Asociado					
<i>GSTO2</i>	rs156697	A/G	G	355/195	FEV1 41.9% FEV1/FVC 42%	0.019	1.39 (1.00-1.93)	71
<i>GSTP1</i>	rs1696 [†]	G/A	A	234/182	FEV1 46.7% FEV1/FVC 69%	0.0011	1.60 (1.2-2.13)	72
<i>TNF</i>	rs1800629	A/G	G	298/125	FEV1 45.0% FEV1/FVC 48%	0.03	1.90 (1.10-3.40)	77
	rs361525	A/G	G					
<i>IL6</i>	rs1800796 [‡]	C/G	C	191/75	FVC 87.0% FEV1/FVC 51.0	0.014	0.30 (0.12-0.79)	78
	rs1818879	A/G	A	355/195	FVC 41.9% FEV1/FVC 42.0	0.02	1.49 (1.08-2.04)	79
<i>SFTPB</i>	rs3024791	A/G	A	389 EPOC	FEV1 28.0%	0.007	0.6 (0.4-0.9)	80
	rs2245121	A/G	A			0.01	1.3 (1.1-1.7)	
<i>SFTPD</i>	rs911887	A/G	G	389/470	FEV1 28.0%	0.006	1.4 (1.1-1.7)	81
	rs721917	C/T	C			0.006	1.4 (1.1-1.7)	

[†] En población tunecina. [‡] En población española. El resto de las asociaciones mostradas fueron realizadas en poblaciones caucásicas de diferentes países.

Tabla 2. Estudios de asociación positiva con EPOC usando la estrategia GWAS o mapeo fino de regiones candidato.

Gen o región genómica	SNP			Estrategia	Fenotipo de EPOC	Valor de p	OR (IC 95%) **β	Referencia
	rs	Alelos	Asociado					
<i>LOC123688</i>	rs8034191	C/T	C			1.48E ⁻¹⁰	1.39 (1.17-1.66)	
<i>CHRNA3</i>	rs1051730	C/T	T			5.74E ⁻¹⁰	1.39 (1.17-1.65)	
<i>ACVR1</i>	rs6720264	A/G	A	Validación con 127 familias.	FEV1 30.88%	1.22E ⁻⁵	1.03 (NR)	84
<i>HHIP</i>	rs1828591	A/G	A			1.47E ⁻⁷	0.70 (NR)	
	rs13118928	A/G	A			1.67E ⁻⁷	0.70 (NR)	
<i>HHIP</i>	rs13147758	A/G	G	Replicación con 853 individuos de familias.	FEV1/FVC <0.7	2E ⁻⁴	0.25**	85
<i>BICD1</i>	rs10844154	C/A	A	2542 sujetos divididos en 3 cohortes.	FEV1/FVC <0.7 con enfisema	5.5E ⁻⁷	1.46	90
	rs161981	C/T	T			1.20E ⁻⁶	1.44	
<i>STAT1</i>	rs13010343	A/G	G	Mapeo fino 953 casos 956 controles	FEV1 50.3% FEV1/FVC 0.51	0.0483	1.24 (1.01-1.55)	91
<i>NFKBIB/SIRT2</i>	rs2241704*	A/T	T			0.0214	0.81 (0.67-0.96)	

*El rs se encuentra en el límite entre ambos genes los cuales están muy cercanos entre sí.

EC-SOD es un gen localizado en el cromosoma 4 humano formado por tres exones, Juul y su grupo en 2006 publicaron sus estudios realizados sobre el polimorfismo (rs1799895) C/G que ocasiona que en la posición 231 de la proteína se codifique una glicina en lugar de una arginina ocasionando una disminución de afinidad de la enzima por su sustrato, empleó una cohorte de daneses con 7565 personas sanas y 1572 pacientes con EPOC clasificados según la GOLD en cuatro estados, donde el promedio de FEV1/FVC para los sujetos con EPOC era de 57% mientras que para los sujetos sin EPOC es de 80% y concluye que para aquellos fumadores heterocigotos funciona como un alelo de protección para desarrollar EPOC (74).

La epóxido hidrolasa microsomal 1 (*EPHX1*), es otro gen candidato, ya que en pulmón se expresa en cierto nivel, mismo que se ve afectado cuando la persona padece EPOC; es una enzima que participa en la detoxificación de sustancias inducidas por estrés oxidativo; a la fecha se han descrito dos variantes tipo SNP (rs1051740 y rs2234922) que se asocian con alteración en la actividad de la enzima. El rs1051740 corresponde a nivel de proteína a una tirosina en el aminoácido 113 pero si es sustituida por una histidina la actividad enzimática disminuye; por otro lado para el rs2234922, histidina en 139 que puede ser cambiada por arginina incrementa dicha actividad enzimática (75, 76).

Dentro de los genes que participan en el proceso de inflamación y que es de los más estudiados en EPOC se encuentra el *TNF*. Gingo y colaboradores en 2008 publican sus hallazgos realizados al estudiar caucásicos con EPOC (298 pacientes y 125 controles) cuyos valores promedio de FEV1/FVC eran de 45.0 para casos y 98.5 para controles y evaluar 6 de sus SNPs pudieron asociar dos de estos al riesgo de padecer EPOC (77).

Otro gen muy estudiado es el de *IL6*; Córdoba-Lanús en 2008 cuando publicó un estudio de casos y controles de España encontró polimorfismos con asociación significativa de susceptibilidad a padecer EPOC (78) mientras que en 2009 Yanbaeva llega a proponer haplotipos asociados a incrementar el riesgo de EPOC (79).

Con respecto a polimorfismos en genes que participan en la regulación de la proliferación y diferenciación celular, reparación tisular o bien producción de matriz extracelular, entre los más estudiados se encuentran: la proteína surfactante B (*SFTPB*) y D (*SFTPD*), la primera estudiada por Foreman en 2008, donde en sujetos con EPOC divididos en dos grupos uno de exacerbadores y el otro de no-exacerbadores encontró un SNP que altera los niveles del RNA mensajero, causa suficiente para que los niveles de proteína disminuyan, predisponiendo al individuo a un mayor número de exacerbaciones y por tal motivo disminuyendo la calidad de vida del paciente (80); nuevamente el mismo grupo de trabajo publica en 2010 sus hallazgos pero ahora con polimorfismos de *SFTPD* asociadas al riesgo de desarrollar EPOC (81).

En 2010 en Japón, Fujimoto publica su estudio realizado en 1335 autopsias de cadáveres encontrando que el rs975278 del gen *SERPINE2* está asociado a la presencia de enfisema cuando está presente el alelo A en comparación de los que poseen el alelo G ($p=0.001$) puesto que reportan un valor de OR de 2.61 (Intervalos de confianza inferior y superior de 1.45-4.76, respectivamente) (82).

1.5.3. Los GWAS en el estudio de la EPOC

En 2007, se publicó un estudio realizado en la cohorte de corazón de Framingham, en donde empleando estrategia de tipo GWAS identificaron dos SNPs asociados con las medidas de función pulmonar, localizados en dos genes que previamente no habían sido descritos, en tal estudio definieron 6 fenotipos de EPOC en base a las pruebas de función pulmonar y los participantes fueron agrupados dentro de uno de los fenotipos definidos según los resultados obtenidos por espirometría, en este estudio los autores sugieren que para aquellos SNPs con resultados significativos sobre todo para aquellos ubicados en genes que aún se desconoce su función biológica, se requiere replicación en otra muestra de individuos (83). Pillai y colaboradores en 2009 reportaron un GWAS en EPOC, en

donde usaron la estrategia de replicación en múltiples etapas. Inicialmente en este estudio se analizaron 538,030 SNPs en un diseño de casos y controles (n=1643), obteniéndose una lista de asociaciones de la que se eligieron los 100 SNPs más fuertemente asociados ($p < 0.00016$), los cuales se evaluaron en un estudio de familias adicional (n=1891), a lo que llamaron etapa de replicación I; de los resultados de la replicación I se seleccionaron los 7 SNPs con la mejor evidencia de asociación en el análisis conjunto de las dos etapas anteriores ($p < 1E^{-7}$), y 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 6 SNPs validados y con asociación importante a la patología. Es importante señalar que encontraron genes que no habían sido reportados previamente como asociados a EPOC, pero que en otras patologías como cáncer pulmonar tienen un papel importante, tal es el caso del gen *CHRNA3*. Otro gen con resultados importantes de asociación a EPOC es *HHIP* del cual se conoce poco, pero los autores sugieren que debido a sus resultados deberá ser estudiado a fondo en otras poblaciones (84). En 2009 Wilk y su grupo de trabajo publicaron un GWAS relacionado con medidas de función pulmonar en el estudio de corazón de Framingham donde la selección de participantes fue aún más estricta con respecto al trabajo de Pillai, fueron analizados 70,987 SNPs en 1220 individuos. Sus resultados dan evidencia de que el gen *HHIP* es importante en la función pulmonar y sugieren otros estudios que incluyan al gen de la glicoforina debido a la posición que ocupa con respecto a *HHIP*, ellos reportan para *HHIP* que el alelo G del rs13147758 reduce el riesgo de obstrucción al flujo aéreo en aquellos fumadores que no tienen dicho alelo (85). Se sabe que el gen *HHIP* influye en el desarrollo del pulmón de mamíferos en etapas embrionarias, hay reportes previos donde asocian a tal gen al riesgo de padecer cáncer pulmonar (86). Con respecto a la glicoforina, es un grupo de proteínas integrales con elevado contenido de ácido siálico presente en la membrana de los

eritrocitos y que contribuyen a su estabilidad (87), existen reportes previos que reportan alteraciones en las células rojas debido a la presencia de estrés oxidativo en pacientes con EPOC a nivel de morfología, rearrreglo del citoesqueleto, disminución en la expresión de glicoforina A entre otros; en éste tipo de pacientes la morfología del eritrocito es de vital importancia ya que son las células encargadas de transportar el oxígeno hacia todo el organismo (88, 89).

Otro GWAS publicado en 2010 por Kong y colaboradores fue realizado en pacientes caucásicos con enfisema (que es un fenotipo de la EPOC) de tres cohortes (2542 sujetos), reportaron al gen *BICD1* asociado con la susceptibilidad para padecer enfisema (90).

En 2010 Bakke y colaboradores al hacer una plataforma de genotipificación sobre diseño incluyó genes potencialmente asociados a EPOC empleando una población caucásica de 953 casos y 956 controles; en este estudio se reportaron nuevos SNPs asociados a padecer EPOC ubicados en genes como *SIRT2*, *STAT1* y *NFKB1B* que proponen deben tener relación para que se activen los mecanismos que lleven a padecer EPOC; en cuanto al gen *STAT1* aún falta establecer con claridad su participación, para el gen *NFKB1B* incrementa la evidencia de que el factor de transcripción NF-κB participa en procesos de señalización que contribuyen al establecimiento de la patología; mientras que para *SIRT2* se le relaciona con procesos degenerativos a través de la deacetilación de alfa-tubulina (91).

En 2010 se publicaron dos meta-análisis, uno de ellos publicado por Repapi (92) y el otro por Hancock (93) donde cada grupo por separado encontró *loci* asociados a disminución en valores de la función pulmonar ubicados en el cromosoma 4, como *HHIP* del cual ya se describió su función y *GSTCD*; en el cromosoma 5 *HTR4* codifica para una proteína G acoplada a membrana que participa en la regulación del AMP cíclico producido por serotonina, y *AGER* en el cromosoma 6 que codifica una proteína que se sabe participa

en la señalización que involucra la defensa del organismo, procesos de inflamación y remodelación de tejido; de todos estos por estudios de asociación de gen candidato no se tenía reporte previo, sin embargo por medio de estudios de genoma completo y meta-análisis ahora se sabe que se encuentran relacionados en la alteración de la función pulmonar, como consecuencia al obtener tales medidas se sabe que funcionan como importantes valores predictivos para diagnosticar la EPOC, así como su progresión no obstante es importante replicar éstas señales en otras poblaciones y de ser posible empleando diversas técnicas (92, 93).

2.0. Hipótesis

Existen polimorfismos tipo SNP en regiones genómicas “candidato” asociados al desarrollo de Enfermedad Pulmonar Obstructiva Crónica (EPOC) que pueden ser identificadas empleando estrategias masivas de mapeo fino, en población mestiza mexicana.

3.0. Objetivo General

Identificar polimorfismos genéticos tipo SNP asociados al riesgo o protección de desarrollar enfermedad pulmonar obstructiva crónica por medio de mapeo fino de regiones genómicas, utilizando la plataforma de alta capacidad GoldenGate de Illumina para regiones genómicas candidato.

3.1. Objetivos Específicos

3.1.1. Seleccionar SNPs con asociación previa a susceptibilidad a EPOC en estudios de asociación de genoma completo y gen candidato.

3.1.2. Seleccionar SNPs etiqueta en regiones genéticas con asociación previa a EPOC.

3.1.3. Evaluar la contribución de la variabilidad genética asociada al riesgo o protección de padecer EPOC en población mestiza mexicana.

4.0. Material y Métodos

4.1. Población de estudio

Todos los individuos fueron seleccionados de noviembre de 2008 a octubre de 2010 de la clínica de EPOC y de la clínica de ayuda para dejar de fumar del Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, para ello fueron empleados exámenes clínicos y radiológicos analizados por neumólogos especialistas,

4.1.1. Definición de caso y control:

Caso: Pacientes con Enfermedad Pulmonar Obstructiva Crónica (EPOC) confirmada por diagnóstico clínico y radiológico con antecedentes de tabaquismo activo.

Control: Individuos mayores de 50 años con hábito de tabaquismo activo por más de 10 años y que fumen o hayan fumado más de 10 cigarros por día y que posterior a las pruebas de función pulmonar (espirometría) se diagnosticaron como individuos sin EPOC.

4.1.2. Criterios de inclusión, no inclusión y exclusión empleados.

Criterios de inclusión:

Mestizos mexicanos por abolengo (padres y abuelos nacidos en México).

Mayores de 50 años de edad.

Fumadores activos que consuman más de 10 cigarrillos por día y al menos por 10 años de manera continua.

Exfumadores que hayan consumido más de 10 cigarrillos por día y por al menos 10 años de manera continua.

Seleccionados por exámenes clínicos y radiológicos por parte de neumólogos especialistas del INER.

Criterios de no inclusión:

Enfermedades pulmonares diferentes a EPOC que alteran la función pulmonar y sus medidas tales como asma, bronquiectasias, cáncer de la vía respiratoria

Sujetos con contraindicación para realizar la espirometría entre ellas, infarto al miocardio o con crisis cardiaca, cirugía reciente, embarazo avanzado o con complicaciones, neumotórax, tuberculosis activa sin tratamiento, infección contagiosa, hemoptisis, traqueotomía.

Que 4 horas previas a la espirometría haya usado broncodilatadores inhalados.

Criterios de exclusión:

Que no acepten la invitación como voluntarios para participar en el proyecto.

Falta de la carta de consentimiento informado con sus respectivos datos y firmas correspondientes.

Cuestionario de datos incompleto.

Muestra biológica insuficiente o de calidad analítica inapropiada.

Espirometría de calidad C o inferior.

4.2. Procedimientos de obtención de muestras biológicas, selección, diseño y genotipificación.

4.2.1. Obtención de material genético y genotipificación.

Extracción de DNA genómico

A partir de células de sangre periférica extraídas por punción venosa, se obtuvo DNA genómico utilizando el estuche comercial *BD tract DNA isolation kit* (Maxim Biotech, San Francisco CA, USA). El DNA fue cuantificado por micro espectrofotometría de absorción de luz ultravioleta a 260 nm de longitud de onda. La contaminación con proteínas se determinó a 280 nm y se consideró una muestra libre de contaminantes cuando la relación 260/280 estaba entre 1.7-2.0.

4.2.2. Diseño del microarreglo.

El microarreglo empleado pertenece a la plataforma de Illumina, fue elegida la modalidad de diseño a “la medida”, para ello fue necesario seleccionar y enviar la lista de SNPs de interés a la compañía para que aplicaran los criterios de control de calidad en base al éxito de genotipificación del SNP de interés por ésta técnica. Para nuestro estudio fue seleccionada la plataforma que permite explorar 1536 SNPs en una sola reacción por microarreglo (94).

4.2.2.1. Selección de SNPs asociados a EPOC

Fue consultada la base del NCBI usando las palabras clave: COPD, association, GWAS, smokers (95); del resultado obtenido seleccione los artículos publicados entre 2005 y 2010 con SNPs de asociación a algún fenotipo de la EPOC; las estrategias empleadas fueron la de gen candidato así como la de GWAS. Cada artículo fue analizado en base al fenotipo estudiado, la población empleada (origen étnico y número de sujetos empleados) y los resultados obtenidos, establecimos criterios muy estrictos en cuanto al valor de p para aquellas asociaciones positivas así como valor de OR o β según el análisis empleado en cada trabajo.

Genere una base de datos con los SNPs que resultaron con asociación positiva de al menos un valor de $p \leq 1E^{-05}$ a la patología, presento un fragmento de esta en el anexo A del presente documento en donde se muestran datos de utilidad tales como: gen o región genómica donde se ubica el SNP de interés, el cromosoma así como la posición en el mismo, el cambio de base, el alelo menor y el asociado, valor de p y OR obtenido en el estudio publicado, la frecuencia del alelo menor (MAF) en caucásicos del HapMap (96), así como en mestizos mexicanos del proyecto de diversidad genómica de los mexicanos (PDGM) del INMEGEN. De un total de 512 SNP de interés se estableció un segundo filtro para la selección de los SNPs a trabajar en nuestra población y que debían de cumplir

como una MAF $\geq 10\%$ en población mestiza mexicana, y que estuvieran en equilibrio de Hardy-Weinberg (HW) con un valor de $p \geq 1E^{-04}$.

Para la MAF en mestizos mexicanos, 10 SNPs no cumplieron con el criterio establecido, por tal motivo fueron excluidos del estudio; mientras que de 241 SNPs no se contaba con datos de frecuencia en mexicanos por lo que fueron conservados en el estudio. Posteriormente, esta lista de SNPs fue enviada a la compañía fabricante del microarreglo para su control de calidad, al realizar el análisis de control interno de calidad excluyeron 44 SNPs por no cumplir con los criterios mínimos establecidos; permaneciendo al final 458 SNPs en regiones génicas a explorar. Se realizó una selección de regiones con mayor cantidad de señales positivas reportadas para hacer el mapeo fino (97).

4.2.2.2. Selección de SNPs etiqueta dentro de regiones genómicas de interés.

Una vez seleccionadas las regiones con ayuda de la base de datos del HapMap Internacional (96) fue ubicada su posición en el cromosoma con el objetivo de cubrir señales de interés que posiblemente se encontrarían en regiones circundantes y ampliamos la zona de búsqueda de SNPs 2.5 kb antes y después de la región seleccionada, en la tabla 3 se muestran las regiones para buscar SNPs etiqueta, para ello fue empleado el software HaploView en población caucásica, asiática, africana, mexicana de Los Angeles del HapMap, y en población mestiza mexicana y zapoteca del PDGM (98).

Los criterios que fueron empleados para la búsqueda de SNP etiqueta y que se aplicaron en HaploView fueron: MAF $\geq 10\%$; desequilibrio de ligamiento: valor de $r^2 > 0.8$, la distancia mínima y máxima entra cada SNP etiqueta debía ser de 60 pb y 2000 pb respectivamente y que cumplieran con el equilibrio de HW con un valor de $p \geq 1E^{-04}$.

Tabla 3. Regiones evaluadas en HaploView para la búsqueda de SNPs etiqueta.

Gen	Cromosoma	Posición en cromosoma (pb)
<i>EPHX1</i>	1	196,537,000 a 196,544,000
<i>IL6R</i>	1	152,644,000 a 152,707,000
<i>SERPINE2</i>	2	216,693,000 a 216,710,000
<i>SFTPB</i>	2	85,792,000 a 85,794,000
<i>HHIP</i>	4	145,500,000 a 145,730,000
<i>GYPA</i>	4	145,480,000 a 145,500,000
<i>ADAM19</i>	5	152,020,000 a 152,025,000
<i>TNF</i>	6	31,675,000 a 31,677,000
<i>SERPINE1</i>	7	75,024,000 a 75,027,000
<i>IL6</i>	7	22,817,000 a 22,818,000
<i>TNFSF8</i>	9	116,747,000 a 116,761,000
<i>SFTPD</i>	10	81,687,000 a 81,699,000
<i>SFTPD</i> cercano	10	75,484,000 a 75,633,000
<i>MMP12</i>	11	102,733,000 a 102,746,000
<i>MMP12</i> cercano	11	98,690,000 a 98,654,000
<i>MMP15</i>	16	56,631,000 a 56,658,000
<i>TGFB1</i>	19	38,291,000 a 38,293,000
<i>CYP2A6</i>	19	46,040,000 a 46,050,000
<i>CYP2A6</i> cercano	19	38,000,000 a 38,050,000
<i>CYP2B6</i>	19	46,180,000 a 46,220,000
<i>CYP2B6</i> cercano	19	46,100,000 a 46,150,000
<i>ADAM33</i>	20	3,596,000 a 3,601,000
<i>ADARB1</i>	21	45,434,000 a 45,480,000
<i>ADARB1</i> cercano	21	46,493,000 a 46,647,000

Una vez obtenidos los SNPs etiqueta en cada población realice una base de datos para cada región monitoreada, con la finalidad de seleccionar los SNPs etiqueta que se incluirían en el microarreglo, en la tabla 4 se muestra un fragmento de una de las tablas

realizadas para una de las regiones genómicas de interés, la cual se empleó para la selección de los SNP etiqueta. Al finalizar obtuve 850 SNPs etiqueta en 24 regiones candidato incluidas en el diseño para mapeo fino; posterior al correspondiente control de calidad por parte del fabricante fue necesario seleccionar 711 de los 850 que pasaron dicho control para tener completo el microarreglo dando prioridad a la MAF en población mestiza mexicana.

Tabla 4 Segmento de una tabla de trabajo para la búsqueda de SNPs etiqueta.

SNP	Alelos	CEU-TSI		CH-JPT		YRI		MEX_LAT		MEX_PDGM	
rs512625	G:A		0.307		0.318		0.122		0.163		0.203
rs3746631	T:C		0.039		ND		0.178		0.067		0.093
rs2280089	G:A	0.149			0.124		0.096	0.067			0.006
rs2280091	A:G	0.149			0.121		0.126	0.067		0.061	
rs598418	A:G			0.366		0.388		0.317		0.327	0.465
rs3918395	C:A	0.146			0.115		0.104		0.067		0.066
rs511898	C:T			0.356		0.359		0.178		0.298	0.29
rs3918392	T:C		0.037			0.079		0.091		0.019	0.02
rs2853210	C:T			0.232		0.182		0.37		0.154	0.157
rs487377	C:T			0.217		0.359		0.483		0.337	0.32
rs554743	T:C			0.28		0.397		0.404		0.144	0.13

Se muestran los SNPs seleccionados, los alelos reportados; se indica la MAF en cada población. Los bloques de colores incluyen un SNP etiqueta (en negritas) y a los SNP etiquetados, aquellos que solo tienen una celda de color son SNP que no funcionan como etiqueta pero que ningún otro los puede etiquetar. ND: alelo no determinado.

4.2.3. Selección de SNPs marcadores informativos de ancestría (AIMs).

Con la finalidad de evaluar el origen poblacional de los individuos incluidos en el estudio (tanto casos como controles) y evitar que las asociaciones encontradas sean debidas a una estratificación poblacional, incluí AIMs que cumplieran con los siguientes criterios: ΔF (diferencia de la frecuencia del SNP entre dos poblaciones) >30%, r^2 entre AIMs <0.4 y la distancia entre AIMs >500 kb. Seleccione 260 AIMs sin embargo sólo 253 cumplieron con

los criterios de calidad de la compañía, tales SNPs se encuentran distribuidos desde el cromosoma 1 hasta el cromosoma 21, en la tabla 5 se indica de forma general la distribución de los SNPs de ancestría entre los cromosomas.

Tabla 5. Distribución en los cromosomas de los SNPs de ancestría empleados en este estudio.

Cromosoma	SNPs incluidos
1	34
2	32
3	23
4	16
5	16
6	23
7	19
8	10
9	7
10	12
11	10
12	9
13	15
14	5
15	6
16	1
17	2
18	6
19	1
20	1
21	5

Posteriormente, para estimar la ancestría de ambas poblaciones realice el cálculo del índice de fijación (F_{st}), fue estimada la diferenciación genética entre la población de casos y controles, el escenario ideal es obtener valores de F_{st} lo más cercano a cero, es decir, la diferencia entre ambas poblaciones es muy poca; para ello fue empleado el paquete computacional Eigensoft 3.0 que mediante el uso de “análisis de componentes principales” (PCA por sus siglas en inglés) estima la diferencia poblacional entre ambos grupos. El PCA es una técnica estadística de síntesis de la información, para ello se

construye una transformación lineal que escoge un nuevo sistema de coordenadas para el conjunto original de datos, las nuevas coordenadas son obtenidas en orden decreciente de importancia en base a la varianza de los datos. (99). La estimación ancestral por individuo fue realizada mediante el software ADMIXTURE (100), el algoritmo empleado se basa en la probabilidad de los genotipos observados (casos y controles) de tener una ascendencia poblacional diferencial, al comparar las frecuencias de los alelos empleados de cada población contra los datos de los mismos AIMS en poblaciones de referencia (caucásicos, japoneses-chinos, yorubas [todos los anteriores del proyecto del HapMap Internacional] mestizos mexicanos y zapotecas [del PDGM del INMEGEN]).

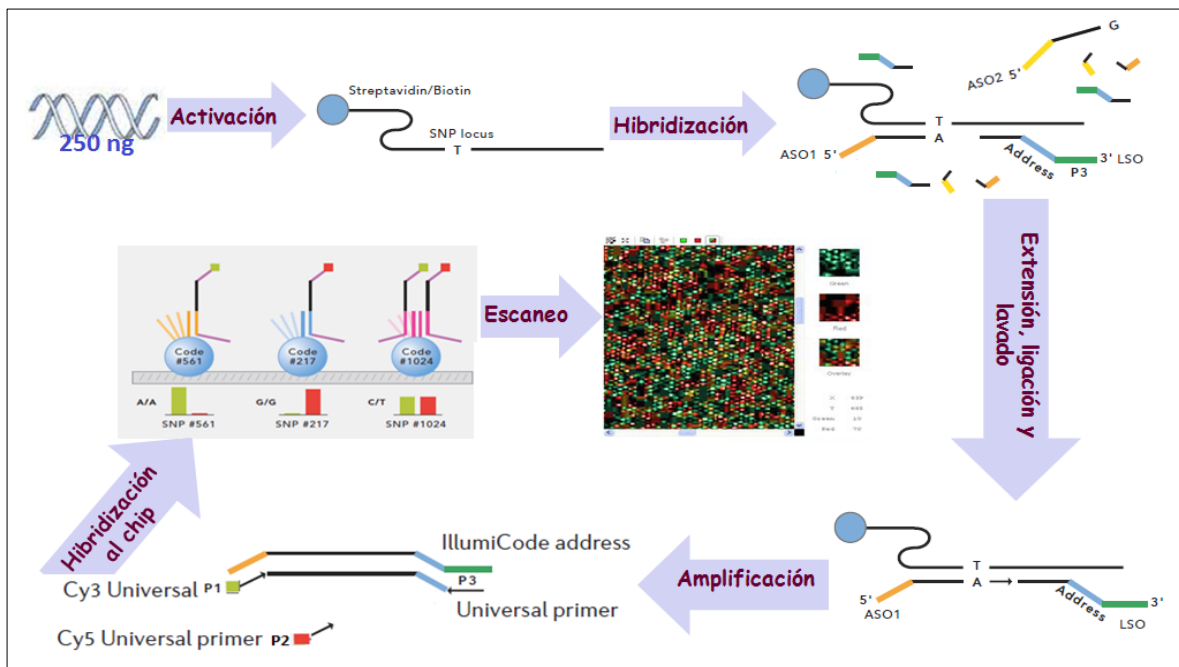
4.3. Genotipificación por GoldenGate

La genotipificación se llevo a cabo empleando la tecnología GoldenGate la cual se fundamenta en la activación del DNA con estreptavidina para que una vez activado pueda unirse a las partículas paramagnéticas y posteriormente sea alineado con los oligonucleótidos alelo específicos 1 (ASO1) y alelo específicos 2 (ASO2) que van a identificar al SNP y al oligonucleótido de direccionalidad, una vez alineado, es lavado para eliminar el exceso de reactivos y se realiza una ligación para que posteriormente se realice la amplificación del DNA por medio de una reacción de PCR; en la mezcla de ésta PCR se encuentran 3 oligonucleótidos universales, Cy3 que se incorpora a ASO1 y Cy5 que se incorpora a ASO2 y un anticuerpo biotinilado complementario al locus específico de direccionalidad, el cual posteriormente servirá para inmovilizar el producto de PCR sobre el microarreglo de genotipificación.

El escáner utilizado para estos microarreglos es el modelo denominado *Bead Array Reader*, el cual usa un laser para excitar la fluorescencia de la extensión de una sola base. Los fluoróforos presentan una imagen de alta resolución en cada sección de los microarreglos. Los datos de estas imágenes fueron analizados para determinar cada

genotipo de los SNPs probados usando el software *Beadstudio* con el modulo de genotipificación (Genotyping Module); el protocolo a detalle se muestra en el anexo B (101) y en la figura 3 se esquematizan los pasos a seguir para la genotipificación por GoldenGate.

Figura 3. Esquema de la técnica de genotipificación masiva.



Se muestra en resumen la técnica de genotipificación masiva empleando la plataforma comercial GoldenGate de Illumina.

5.0 Resultados

5.1. Participantes en el estudio

Se recolectó un total de 522 muestras, 247 procedentes de la clínica de EPOC y 205 de la clínica de ayuda para dejar de fumar, ambas ubicadas en el INER. Se excluyeron 196 muestras (90 y 106 respectivamente) por no cumplir con los criterios de inclusión antes mencionados. Finalmente se formaron 2 grupos, el de casos (n=157) y controles (n=169). La descripción de los datos demográficos, hábito tabáquico y medidas de función pulmonar de los participantes incluidos como grupos de estudio se muestran en la tabla 6.

Tabla 6. Descripción de la población empleada para la genotipificación

Variables	Casos (n=157)	Controles (n=169)
Edad (años)	66±8	54±7
Género		
Masculino (%)	122 (77.7)	81 (48.0)
Femenino (%)	35 (22.3)	88 (52.0)
Tabaquismo		
Años fumando	40±12	31±10
Cigarrillos/día	24±14	19±8
Paquetes/año	48±31	29±15
Función pulmonar		
FVC, ref (%)	79.3±22.1	97.3±15.7
FEV1, ref (%)	53.3±22.4	97.9±16.6
FEV1/FVC (%)	51.2±13.7	80.9±5.4

5.2. Obtención de genotipos

La obtención de genotipos se realizó usando el software *GenomeStudio*, diseñado y provisto por el fabricante, el cual muestra una gráfica (figura 4) para cada uno de los SNPs genotipificados con tres zonas definidas en base al *GenCall Score*, este último es un parámetro empleado por el software que toma en cuenta el ángulo, la dispersión y la intensidad de la señal emitida por el fluoruro, para así asignar el genotipo correspondiente en cada una de las tres zonas.

Figura 4. Genotipificación del software *GenomeStudio*.

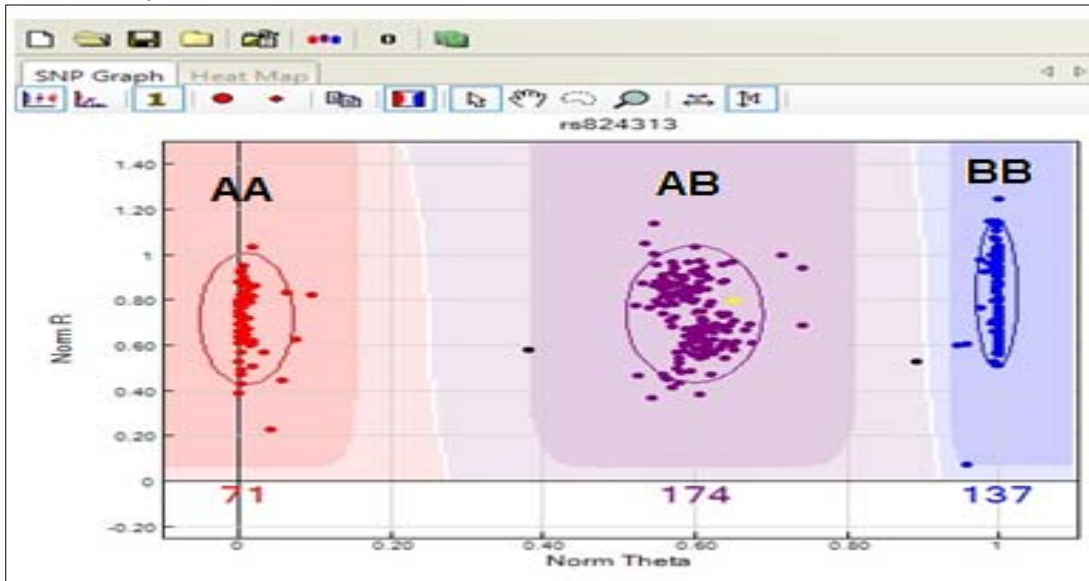


Gráfico con coordenadas polares que muestra por SNP el genotipo de los sujetos en estudio.

Otros criterios que fueron considerados para obtener los genotipos son, 1) la tasa de éxito de genotipificación (*CallRate*) la cual debe ser $\geq 95\%$ e indica que, aquellas muestras con un error de genotipificación total mayor al 5% serán eliminadas del análisis y 2) el *GeneTrain Score*, que es la medida de la calidad de agrupamiento para cada SNP. De ambos parámetros el software generó una tabla la cual se puede ordenar de manera ascendente o descendente en base a cualquiera de los dos parámetros descritos previamente y fueron eliminados aquellos que no cumplieran con los criterios de calidad; para nuestro diseño 24 SNPs no cumplieron con el *GeneTrain Score*, por tal motivo quedó un archivo con 1512 SNPs de los cuales 253 son de ancestría; de las 359 muestras genotipificadas 23 no cumplieron con el criterio de *CallRate* por lo tanto el archivo a analizar contiene 329 sujetos genotipificados.

5.3. Generación de archivos

Con el software GenomeStudio creamos los archivos tipo *ped* y tipo *map* (extensiones .ped y .map) que son aquellos que serán utilizados para el análisis de ancestría y de asociación.

El archivo *ped* se encuentra formado por 7 columnas que se describen a continuación:

1. Identificador por familia (0 si es un estudio de casos y controles).
2. Identificador por individuo.
3. Identificador del padre (0 si es un estudio de casos y controles).
4. Identificador de la madre (0 si es un estudio de casos y controles).
5. Identificador de género (1=masculino 2=femenino).
6. Estado de afección (1=no afectado, 2= afectado, 0= desconocido).
7. Genotipos del marcador, cada uno representado por dos columnas, una para cada alelo separado por un espacio.

El archivo *map* contiene información sobre los marcadores genéticos incluidos en el análisis:

1. Número de cromosoma.
2. Número de rs del SNP.
3. Distancia genética (Morgans).
4. Posición (pares de bases).

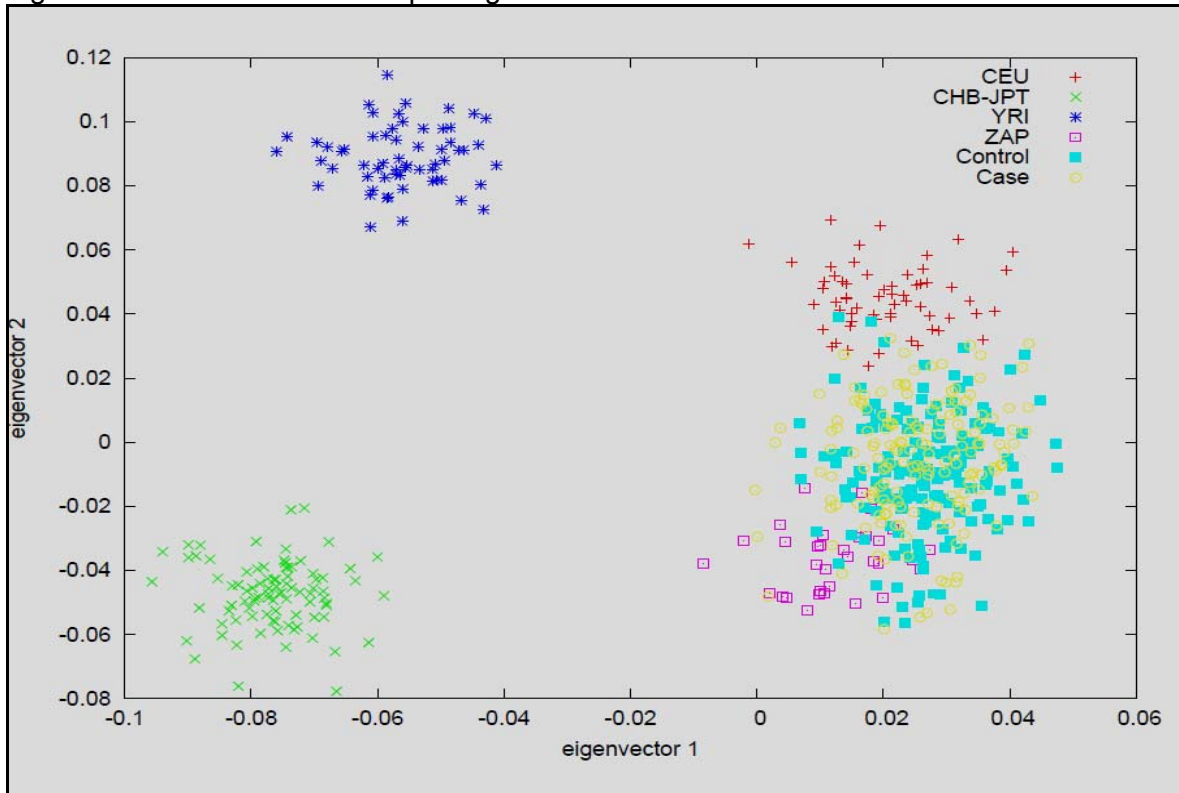
5.4. Análisis de resultados.

5.4.1. Análisis de ancestría.

Al separar los 253 SNPs de ancestría y analizarlos por Eigensoft se obtuvo la gráfica 5 de PCA en donde se observa que casos y controles se encuentran dentro de una misma zona de dispersión, distribuidos entre la población caucásica y zapoteca, las poblaciones de referencia empleadas fueron caucásicos, asiáticos y africanos del HapMap, el grupo de zapotecos fue tomado del PDGM. En la tabla 7 se muestra el valor de F_{st} para casos y

controles, adicionalmente fueron comparadas ambas poblaciones sólo con la población caucásica y zapoteca para poder calcular los valores de índice de fijación de las poblaciones más cercanas a la población de nuestro estudio.

Figura 5. Análisis de ancestría por Eigensoft.



Distribución por análisis de ancestría de los participantes agrupados en casos y controles (cuadros amarillos y cuadros aguamarina, respectivamente). Se observa que ambos grupos se distribuyen entre caucásicos (cruces rojas) y zapotecas (cuadros de línea rosa).

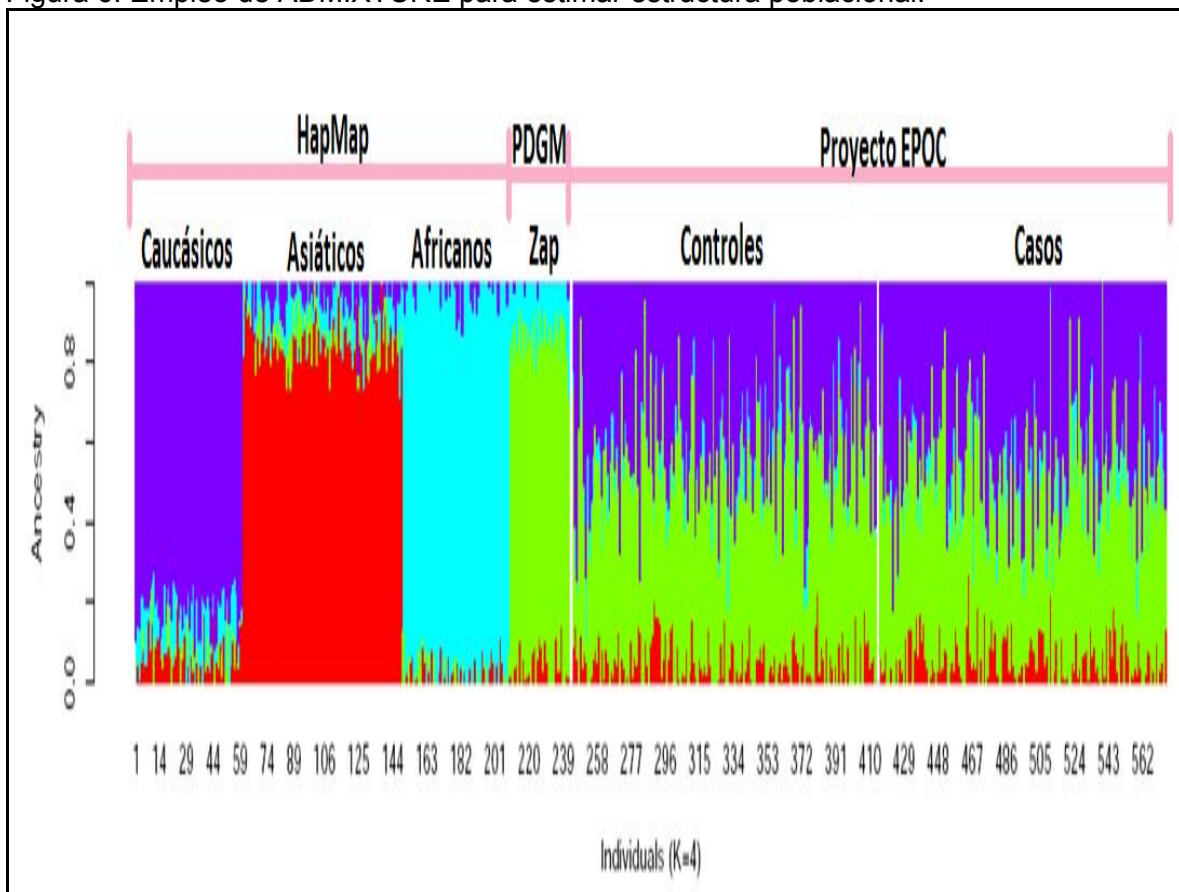
Tabla 7. Comparación del valor Fst entre grupos de estudio vs caucásicos y zapotecos.

Fst	ZAP	Controles	Casos
CEU	0.228	0.094	0.083
ZAP		0.109	0.115
Controles			0.000

Se pueden observar las diferencias en el valor Fst al extrapolar las filas con las correspondientes columnas, de esta manera encontramos que la distancia entre casos y controles es igual a 0.000.

Cuando empleamos ADMIXTURE para estimar la estructura poblacional de casos y controles fueron empleadas 3 poblaciones de referencia del HapMap y la población zapoteca del PDGM y generamos la gráfica de la figura 7, donde se señala para controles y casos la contribución de cada una de las poblaciones de referencia empleadas en el análisis (K=4) y en la tabla 8 se muestran las proporciones ancestrales promedio de las 4 poblaciones de referencia empleadas para casos y controles en donde se puede observar que los promedios entre las poblaciones de estudio es muy parecido para cada contribución ancestral.

Figura 6. Empleo de ADMIXTURE para estimar estructura poblacional.



Los códigos de colores empleados permiten evaluar fácilmente la contribución ancestral de las poblaciones en estudio, al comparar con las poblaciones descritas en el HapMap y Zapotecos del PDGM.

Tabla 8. Proporciones ancestrales promedio empleando ADMIXTURE.

Población	CEU	ASIA	YRI	ZAP
Casos	0.435±0.149	0.057±0.055	0.055±0.056	0.451±0.147
Controles	0.404±0.154	0.051±0.052	0.048±0.048	0.497±0.161

Se muestran las proporciones ancestrales promedio de los grupos en estudio con respecto a las 4 poblaciones de referencia.

5.4.2. Análisis de asociación.

Fue realizado con los archivos *.ped* y *.map* generados previamente, para ser usados en la herramienta bioinformática PLINK, para la comparación de frecuencias entre casos y controles fueron usados los modelos de χ^2 y posteriormente la prueba exacta de Fisher, ambas pruebas permitieron obtener valores de OR así como su intervalo de confianza (102).

El archivo *.ped* contiene 1259 SNPs, ya que en ésta parte no incluimos los SNPs de ancestría. Antes de hacer la prueba de asociación en PLINK también aplicamos criterios de calidad tales como: Exclusión de individuos con >10% de sus genotipos perdidos; exclusión de SNPs perdidos en más del 5% de los ensayos; eliminamos aquellos SNPs que en los controles no cumplieran con el Equilibrio de Hardy-Weinberg ($p > 0.00001$) y eliminamos SNPs con una MAF >0.1.

Aplicando todos los criterios antes mencionados, se obtuvo un archivo con 1188 SNPs y 326 sujetos (169 controles y 157 casos); con tales filtros de calidad fue posible aplicar la prueba de χ^2 y la exacta de Fisher y con esto obtuvimos el OR, los SNPs con asociaciones estadísticamente significativas se presentan en la tabla 9, en donde se muestran alelos de susceptibilidad y de protección. En ella podemos observar algunos SNPs previamente asociados en otras poblaciones principalmente caucásica y también

existen algunos de ellos en donde la región génica ya había sido asociada pero los SNPs asociados son otros, éste tipo de nuevos SNPs asociados en población mestiza mexicana fue identificado debido a que para el diseño de nuestro microarreglo fue empleada la estrategia de SNPs etiqueta en regiones de interés.

Al usar el software *HaploView* genere 4 bloques de haplotipos con un desequilibrio de ligamiento alto ($r^2 > 0.90$) asociados a EPOC que se muestran en la figura 8 y en la tabla 10 se observa claramente que de los 4 haplotipos generados el primero es de susceptibilidad a la enfermedad mientras que los tres últimos son de protección a la EPOC en población mestiza mexicana.

Tabla 9. Resultados del análisis de asociación genética.

Gen	SNP	Alelo Asociado	Casos	FA Controles	χ^2	p	OR	IC 95%	
IL6R	rs4072391	A	0.2229	0.1598	4.217	0.04573	1.509	1.017	2.238
EPHX1	rs2234922	G	0.1242	0.07692	4.054	0.04981	1.702	1.010	2.868
RHOB	rs12988451	A	0.3397	0.2337	8.956	0.00308	1.687	1.196	2.380
	rs9917322	A	0.3333	0.2278	8.991	0.00291	1.695	1.199	2.396
	rs4666308	G	0.3344	0.2367	7.647	0.00696	1.620	1.149	2.284
	rs4666402	G	0.3333	0.2411	6.751	0.01150	1.574	1.117	2.219
	rs10205221	A	0.3408	0.2367	8.62	0.00418	1.667	1.183	2.348
	rs7597833	G	0.4776	0.3817	6.097	0.01408	1.481	1.084	2.024
SERPINE2	rs16865390	A	0.2261	0.3195	7.132	0.00841	0.622	0.438	0.883
	rs7605945	G	0.3694	0.4556	4.985	0.02623	0.700	0.511	0.958
	rs7608941	C	0.3662	0.4586	5.72	0.01722	0.682	0.498	0.934
	rs6712954	A	0.2293	0.3491	11.32	0.00078	0.554	0.397	0.785
	rs6738983	A	0.4268	0.5149	5.058	0.02772	0.701	0.515	0.956
	rs1866153	A	0.1433	0.0858	5.351	0.02573	1.782	1.087	2.923
FAM13A	rs2869967	G	0.5064	0.4142	5.569	0.01859	1.451	1.065	1.977
GYPA*	rs6844670	A	0.379	0.3047	3.996	0.0476	1.392	1.006	1.927
ARRDC3	rs17641623	A	0.4199	0.5325	8.254	0.00466	0.635	0.466	0.866
	rs7703346	T	0.4103	0.5266	8.818	0.00359	0.625	0.458	0.853

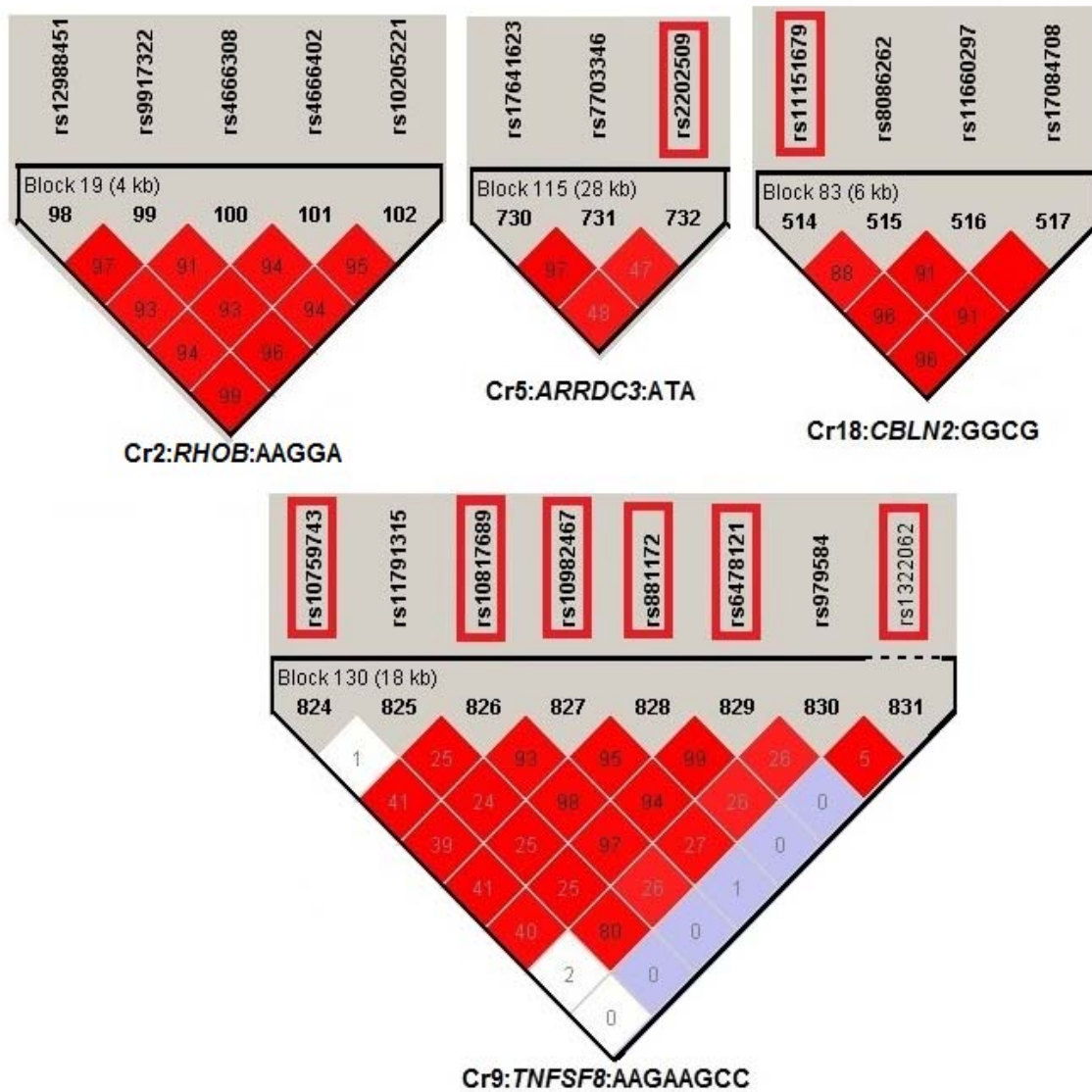
Se muestran 18 alelos localizados en 7 diferentes genes, de los cuales 11 presentan asociación al riesgo a desarrollar EPOC y 7 se asocian con la protección al desarrollo de la misma. FA: Frecuencia alélica del alelo asociado.

Tabla 9 continuación de resultados del análisis de asociación genética.

Gen	SNP	Alelo Asociado	FA		χ^2	p	OR	IC 95%	
			Casos	Controles					
ADAM19*	rs6579952	A	0.2484	0.1657	6.823	0.01151	1.664	1.133	2.444
	rs17565170	A	0.1847	0.1006	9.505	0.00231	2.026	1.285	3.192
	rs17565365	A	0.1688	0.247	6.004	0.01583	0.619	0.421	0.911
ADAM19	rs17054697	A	0.1401	0.08876	4.269	0.04762	1.673	1.023	2.736
	rs6875485	C	0.1369	0.08631	4.225	0.04537	1.680	1.020	2.765
	rs11134819	A	0.1019	0.05917	4.051	0.04414	1.804	1.009	3.226
TNF*	rs2844480	A	0.09236	0.1509	5.181	0.02362	0.572	0.353	0.929
CDKN1A	rs1801270	A	0.3279	0.4053	4.149	0.04993	0.715	0.518	0.988
SERPINE1	rs2227631	A	0.4045	0.3174	5.331	0.02203	1.461	1.058	2.016
VPS13A	rs4142603	G	0.4167	0.5208	7.046	0.00932	0.657	0.482	0.896
	rs12380218	G	0.4167	0.5	4.536	0.03388	0.714	0.524	0.974
TNFSF8*	rs11791315	A	0.08442	0.1399	4.919	0.03389	0.566	0.342	0.941
	rs979584	C	0.09677	0.1686	7.185	0.00792	0.528	0.329	0.847
CHRNA4	rs1316971	A	0.2962	0.3846	5.657	0.02063	0.673	0.486	0.933
	rs8086262	G	0.1433	0.2071	4.559	0.03947	0.640	0.425	0.966
CBLN2	rs11660297	C	0.1338	0.1923	4.068	0.04507	0.648	0.425	0.989
	rs17084708	G	0.1338	0.1923	4.068	0.04507	0.648	0.425	0.989
ADAM33	rs598418	G	0.4459	0.3698	3.901	0.04825	1.375	1.002	1.876

Se muestran 18 alelos localizados en 10 diferentes genes/regiones genómicas, de los cuales 7 presentan asociación al riesgo a desarrollar EPOC y 11 se asocian con la protección al desarrollo de la misma. FA: Frecuencia alélica del alelo asociado.*Localización cercana al gen.

Figura 7. Bloques de haplotipos generados con los SNPs asociados a EPOC.



El análisis de haplotipos empleando el software HaploView, determinó la existencia de 4 bloques de haplotipos en 4 diferentes cromosomas. En rectángulos de color rojo se encuentran los SNPs que no están asociados de forma significativa a EPOC pero muestran un desequilibrio de ligamiento superior al 90% ($r^2 > 0.90$).

Tabla 10. Haplotipos asociados a EPOC encontrados en población mestiza mexicana.

Haplotipo	Cr	Frecuencia en casos	p	OR	IC 95%	
AAGGA	2	0.322	0.0072	1.61	1.12	2.32
ATA	5	0.119	0.0096	0.57	0.36	0.89
AAGAAGCC	9	0.083	0.0245	0.57	0.40	0.96
GGCG	18	0.133	0.0402	0.64	0.41	0.99

Las letras en negritas del haplotipo son los alelos no asociados (de manera individual) a EPOC pero se encuentran en alto desequilibrio de ligamiento ($r^2 > 0.90$) con aquellos alelos asociados de forma significativa.

6.0 Discusión

Para este trabajo se han abordado diferentes estrategias de análisis con la finalidad de identificar SNPs asociados a EPOC en población mestiza mexicana, la primera de ellas, fue la evaluación de polimorfismos aislados asociados a la patología localizados en genes “candidato” que previamente se habían descrito en otras poblaciones; segunda, la evaluación de los SNPs que fueron identificados por medio de metodologías de genotipificación masiva (GWAS), y por último la selecciones de regiones genómicas para mapeo fino, dentro de las cuales se identificaron SNPs etiqueta que de ser posible proporcionaran nuevos polimorfismos (no descritos previamente) asociados a EPOC en nuestra población. Para ello se empleó el diseño de casos vs controles, recolectando un total de 326 muestras que cumplieron con los criterios de inclusión previamente establecidos, 157 casos procedentes de la clínica de EPOC y 169 controles de la clínica de ayuda para dejar de fumar, ambas del INER.

Una parte importante de nuestro estudio de asociación consistió en la caracterización de la contribución ancestral de ambos grupos en estudio, nuestro análisis muestra que no existen diferencias en cuanto a la ancestría poblacional entre el grupo de casos y controles empleado; dicha evaluación se llevó a cabo tomando en cuenta poblaciones que según el PDGM publicado en 2009 se pueden considerar parentales en el mestizaje de los mexicanos (caucásicos y zapotecos, estos últimos representativos de la contribución genética indígena del amerindio) (67), además de la comparación con poblaciones asiáticas (chinos-japoneses del HapMap) y africanas (yorubas del HapMap). Con tales pruebas de corrección poblacional podemos decir que nuestros hallazgos son debido a la asociación del SNP a la EPOC, eliminando así la posibilidad de resultados debidos a la estratificación poblacional.

El análisis de asociación originó en 17 regiones genéticas candidato 36 SNPs asociados a EPOC de los cuales 18 de ellos muestran asociación de riesgo a desarrollar la enfermedad mientras que los otros 18 se asocian a protección. Dentro de las señales de riesgo se encuentra el alelo A del rs4072391 ubicado en el gen *IL6R*, tal gen ha sido propuesto previamente por Wilk y colaboradores como un gen candidato en la enfermedad, se sabe que *IL6R* se expresa en pulmón y tiene una función importante en la respuesta inmunológica (85), sin embargo, este SNP en particular es reportado por primera vez en EPOC en población mestiza mexicana. Otra señal de riesgo reportada en el estudio es el alelo G del rs2234922 localizado en el gen *EPHX1*, previamente Hersh y colaboradores en población caucásica lo reportó como una señal de protección a padecer EPOC, no obstante, al tratar de replicar sus hallazgos en familias no obtuvo el mismo resultado (103). El alelo G del rs2869967 ubicado en el gen *FAM13A*, había sido asociado mediante GWAS a una disminución en la función pulmonar, posteriormente Hancock en un meta-análisis lo asocia a una obstrucción en el flujo de aire de sujetos europeos, en nuestro estudio resulta nuevamente ser una señal que indica susceptibilidad a padecer EPOC en población mestiza mexicana. Para el gen *GYPA* se tiene un panorama similar al gen anterior ya que, fue identificado mediante GWAS, Wilk en 2009 señala que en aquellos pacientes con EPOC la expresión de glicoforinas en la membrana de los glóbulos rojos se encuentra disminuida (85); nuestro estudio muestra un marcador distinto que en otras poblaciones no había sido reportado, el alelo A del rs6844670 (gen *GYPA*) es de riesgo a padecer EPOC en población mestiza mexicana.

Del gen *SERPINE2* se han encontrado diversos SNPs asociados a EPOC en poblaciones caucásicas y asiáticas, para nuestro estudio se incluyeron estas señales y además la región fue mapeada de tal forma que se seleccionaron SNPs etiqueta que cubrieran al máximo el gen y sus alrededores. En este sentido nuestros resultados muestran siete SNPs que no

habían sido reportados en otras poblaciones asociados a EPOC y que fue posible identificarlos ya que empleamos la estrategia de SNPs etiqueta, tomando como referencia la población mestiza mexicana y las poblaciones reportadas en el proyecto internacional del HapMap, los alelos G y A (rs7597833 y rs1866153 respectivamente) resultaron de riesgo mientras que los alelos A, G, C, A y A, de protección (rs16868390, rs7605945, rs7608941, rs6712954 y rs6738983, respectivamente) a la patología en población mestiza mexicana. Otro gen es *SERPINE1*, nuestro estudio muestra el alelo A del rs2227631 como un nuevo SNP dentro del gen con asociación a la susceptibilidad de padecer la enfermedad. Los genes que codifican proteínas de la familia de metaloproteasas y desintegrinas (ADAM), se han asociado a EPOC, múltiples estudios sugieren que *ADAM33* participa en la respuesta inmunológica pulmonar y en la remodelación del tejido, existen diversos polimorfismos dentro del gen asociados a la enfermedad o bien a una disminución en la función pulmonar realizados en población asiática y caucásica, nuestro estudio muestra que el alelo G del rs598418 es un marcador de riesgo a padecer EPOC en población mestiza mexicana ya que previamente no había sido reportado en otras poblaciones. Por otro lado, *ADAM19* no había sido estudiado previamente hasta que por medio de GWAS resultó ser un gen asociado de manera significativa a la enfermedad, nuestros datos muestran señales asociadas a EPOC dentro del gen y en una región cercana al mismo con asociaciones tanto de protección como de susceptibilidad, tales señales fueron identificadas por medio de la estrategia de búsqueda de SNPs etiqueta.

Debido a que la EPOC se caracteriza por su respuesta inflamatoria exagerada, otro gen candidato ampliamente estudiado es *TNF*, Keatings y colaboradores compararon biopsias de pacientes con EPOC y controles, encontraron niveles elevados de la proteína en el grupo de casos (104), nuestro estudio muestra un SNP cercano a *TNF* con asociación a protección en nuestra población que previamente no había sido reportado, éste último escenario se repite

para uno de los ligandos de *TNF*, el *TNFSF8* del cual también obtuvimos dos SNPs cercanos al gen del ligando que actúan como de protección; del rs11791315 no existen reportes de su frecuencia en población mestiza mexicana, sin embargo fue incluido ya que se tiene el dato de que es un SNP etiqueta en población caucásica y asiática. El gen *CDKN1A* codifica una proteína inhibidora de ciclinas dependientes de cinasa que regula la proteína p53 en el ciclo celular y específicamente, el rs1801270 fue asociado a EPOC por Lee y colaboradores en población asiática (105), nuestro estudio replica la asociación con tal SNP y contribuye a la propuesta realizada por algunos autores donde indican que en la EPOC también existe un componente que lleva al desequilibrio en el mecanismo de apoptosis. Los genes *VPS13A*, y *CHRNA4* fueron evidenciados mediante la estrategia de GWAS, del primero de ellos no queda clara su participación en la patología (106) mientras que el gen *CHRNA4* tiene varios SNPs que se han asociado al riesgo de adicción a la nicotina. En nuestra población de estudio ambos genes mostraron SNPs de protección a EPOC.

El análisis de haplotipos muestra 4 bloques asociados de forma significativa, siendo uno de ellos de riesgo a padecer EPOC (AAGGA, OR = 1.61 ubicado en el gen *RHOB*), mientras que los otros 3 (*TNFSF8*, *ARRDC3* y *CBLN2*) son de protección en población mestiza mexicana. Previamente ya se había discutido la participación del *TNFSF8* en la patología y nuestro estudio muestra un haplotipo asociado a la enfermedad. Por otro lado, el gen *ARRDC3* en nuestra población muestra un haplotipo asociado a la enfermedad, tal gen en EPOC no tiene reportes previos, hasta la fecha se ha asociado como un importante supresor de tumores (108), pero aún está por esclarecer su papel en la patología obstructiva crónica, incluso es posible que en la EPOC también exista un desequilibrio en el control de la proliferación celular. En cuanto al gen *CBLN2* se sabe que codifica para una proteína que se expresa en cerebro e incluso algunos polimorfismos del gen se han asociado a esclerosis múltiple, en EPOC está por esclarecer su papel pero, hay que recordar que el presente

estudio muestra señales asociadas a EPOC secundaria a tabaquismo y se sabe que el tabaquismo es considerada una enfermedad que tiene que ver con la adicción a nivel cerebral, entonces posiblemente el gen participe en el establecimiento de la adicción. En cuanto al gen *RHOB*, que muestra un haplotipo de susceptibilidad, en población caucásica empleando la estrategia de GWAS previamente ya se habían reportado algunos SNPs asociados a EPOC, nuestros resultados muestran estos mismos SNPs pero ahora en población mestiza mexicana. Cabe mencionar que el gen codifica para la proteína RhoB, la cual es miembro de la familia de las proteínas G, Srougi y colaboradores demostraron que la expresión del gen se estimula en respuesta a un ambiente de estrés celular, en tal contexto, la proteína es necesaria para inducir apoptosis en caso de que exista daño al DNA (109). Recientemente se han incrementado los trabajos que exploran la apoptosis como un mecanismo que contribuye a la patogénesis en la EPOC e incluso se llega a proponer un desequilibrio entre la apoptosis y mecanismos anti-apoptóticos y nuestro estudio aporta una evidencia más a tal hipótesis.

Finalmente, es importante señalar que la implementación de nuevas tecnologías, como la empleada en el presente trabajo permiten identificar marcadores genéticos que anteriormente hubiera sido difícil investigar como candidatos importantes en enfermedades complejas como la EPOC ampliando así nuestro campo de estudio en la enfermedad.

7.0 Conclusiones.

El presente trabajo muestra SNPs que previamente se han reportado asociados a EPOC, en su mayoría en población caucásica y que fue posible replicar la señal en población mestiza mexicana, entre ellos están el rs2234922 (*EPHX1*), rs12988451, rs9917322, rs4666308, rs4666402, rs10205221 (*RHOB*), rs2869967 (*FAM13A*), rs17641623, rs7703346 (*ARRDC3*), rs1801270 (*CDKN1A*), rs4142603 (*VPS13A*), rs1316971 (*CHRNA4*), rs8086262, rs11660297 y rs17084708 (*CBLN2*), tales hallazgos son una evidencia para validar la asociación de dichas señales en EPOC.

La estrategia empleada durante el estudio permitió encontrar nuevos SNPs asociados a la enfermedad que no han sido reportados en otras poblaciones, tales como el rs4072391 (*IL6R*), rs7597833, rs16865390, rs7605945, rs7608941, rs6712954, rs6738983, rs1866153 (*SERPINE2*), rs6844670 (*GYPA*), rs6579952, rs17565170, rs17565365 (región cercana a *ADAM19*), rs17054697, rs6875485, rs11134819 (*ADAM19*), rs2844480 (región cercana a *TNF*), rs2227631 (*SERPINE1*), rs12380218 (*VPS13A*), rs11791315, rs979584 (región cercana a *TNFSF8*) y rs598418 (*ADAM33*).

Nuestros hallazgos aportan conocimiento para la dilucidación del componente genético de la EPOC secundaria a tabaquismo en individuos mestizos mexicanos, que debido a su estructura poblacional, no pueden ser comparados con poblaciones de un origen ancestral distinto.

8.0 Bibliografía consultada

1. Global Initiative for Chronic Obstructive Lung Disease, Inc. 2010 <http://www.goldcopd.org>. Consultada 12-Enero-2011.
2. Pérez-Padilla R, Regalado J, Vedal S, Paré P, Chapela R, Sansores R, Selman M. Exposure to biomass smoke and chronic airway disease in Mexican women. A case-control study. *Am J Respir Crit Care Med*. 1996; 154(3 Pt 1):701-6.
3. Contreras Carreto y Martínez Quevedo. Utilidad de la espirometría en el paciente con EPOC. *Medica Sur México*. 2006; 13(4):156-160.
4. Organización Mundial de la Salud (OMS) <http://www.who.int/topics/tobacco/es/index.html>. Consultada 13 Enero 2011
5. Proyecto Latinoamericano de Investigación en Obstrucción Pulmonar. PLATINO <http://www.platino-alat.org/>, consultada 10 Enero 2011
6. Tovar Guzmán VJ, López Antuñano FJ, Rodríguez Salgado N. Recent trends in mortality due to chronic obstructive pulmonary disease (COPD) in Mexico, 1980-2002. *Arch Med Res*. 2005; 36(1):65-9.
7. INER. Clínica de EPOC, Consultada 12 Enero 2011 http://www.iner.salud.gob.mx/contenidos/Clinica_de_EPOC.html
8. Reynales-Shigematsu LM, Juárez-Márquez SA, Valdes-Salgado R. Costos en atención médica atribuibles al tabaquismo en el IMSS Morelos. *Salud Pública de México*. 2005; 47:451-457.
9. Vermeeren MA, Creutzberg EC, Schols AM, Postma DS, Pieters WR, Roldaan AC, Wouters EF; COSMIC Study Group. Prevalence of nutritional depletion in a large out-patient population of patients with COPD. *Respir Med*. 2006; 100(8):1349-55.
10. Rodríguez-Roisin R. COPD exacerbations.5: management. *Thorax*. 2006; 61(6):535-44.
11. Miravittles M, Guerrero T, Mayordomo C, Sánchez-Agudo L, Nicolau F, Segú JL. Factors associated with increased risk of exacerbation and hospital admission in a cohort of ambulatory COPD patients: a multiple logistic regression analysis. *Respiration* 2000; 67: 495–501.
12. Dahlén I, Janson C. Anxiety and depression are related to the outcome of emergency treatment in patients with obstructive pulmonary disease. *Chest*. 2002; 122(5):1633-7.
13. American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1995; 152(5 Pt 2):S77-121.
14. Han MK, Agusti A, Calverley PM, Celli BR, Criner G, Curtis JL, Fabbri LM, Goldin JG, Jones PW, Macnee W, Make BJ, Rabe KF, Rennard SI, Sciurba FC, Silverman EK, Vestbo J, Washko GR, Wouters EF, Martinez FJ. Chronic obstructive pulmonary disease phenotypes: the future of COPD. *Am J Respir Crit Care Med*. 2010; 182(5):598-604.
15. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was

adopted by the ATS Board of Directors, March 1995. *Am Respir Crit.Care Med.* 1995; 152:S77-S120.

16. Antón DE, Ruiz LD y Ancochea BJ, Herencia y ambiente en la EPOC. *Arch Bronconeumol.* 2007; 43 Supl 1:10-17.

17. Friedlander AL, Lynch D, Dyar LA, Bowler RP. Phenotypes of chronic obstructive pulmonary disease. *COPD.* 2007; 4(4):355-84.

18. Laurin C, Lavoie KL, Bacon SL, Dupuis G, Lacoste G, Cartier A, Labrecque M. Sex differences in the prevalence of psychiatric disorders and psychological distress in patients with COPD. *Chest.* 2007; 132(1):148-55.

19. Lopez Varela MV, Montes de Oca M, Halbert RJ, Muiño A, Perez-Padilla R, Tálamo C, Jardim JR, Valdivia G, Pertuzé J, Moreno D, Menezes AM; PLATINO Team. Sex-related differences in COPD in five Latin American cities: the PLATINO study. *Eur Respir J.* 2010; 36(5):1034-41.

20. Barr RG, Celli BR, Mannino DM, Petty T, Rennard SI, Sciruba FC, Stoller JK, Thomashow BM, Turino GM. Comorbidities, patient knowledge, and disease management in a national sample of patients with COPD. *Am J Med.* 2009; 122(4):348-55.

21. Martinez FJ, Curtis JL, Sciruba F, Mumford J, Giardino ND, Weinmann G, Kazerooni E, Murray S, Criner GJ, Sin DD, Hogg J, Ries AL, Han M, Fishman AP, Make B, Hoffman EA, Mohsenifar Z, Wise R; National Emphysema Treatment Trial Research Group. Sex differences in severe pulmonary emphysema. *Am J Respir Crit Care Med.* 2007; 176(3):243-52.

22. Piccinni MP, Giudizi MG, Biagiotti R, Beloni L, Giannarini L, Sampognaro S, Parronchi P, Manetti R, Annunziato F, Livi C. Progesterone favors the development of human T helper cells producing Th2-type cytokines and promotes both IL-4 production and membrane CD30 expression in established Th1 cell clones. *J Immunol* 1995; 155:128–133.

23. Becklake MR, Kauffmann F. Gender differences in airway behavior over the human life span. *Thorax* 1999; 54:1119–1138.

24. Roy K, Smith J, Kolsum U, Borrill Z, Vestbo J, Singh D. COPD phenotype description using principal components analysis. *Respir Res.* 2009; 29; 10:41.

25. Burgel PR, Paillasseur JL, Caillaud D, Tillie-Leblond I, Chanez P, Escamilla R, Court-Fortune I, Perez T, Carré P, Roche N; Initiatives BPCO Scientific Committee. Clinical COPD phenotypes: a novel approach using principal component and cluster analyses. *Eur Respir J.* 2010; 36(3):531-9.

26. Mak JC. Pathogenesis of COPD. Part II. Oxidative-antioxidative imbalance. *Int J Tuberc Lung Dis.* 2008; 12(4):368-74.

27. Pons AR, Noguera A, Blanquer D, Sauleda J, Pons J, Agustí AG. Phenotypic characterisation of alveolar macrophages and peripheral blood monocytes in COPD. *Eur Respir J.* 2005; 25(4):647-52.

28. Francus T, Klein RF, Staiano-Coico L, Becker CG, Siskind GW. Effects of tobacco glycoprotein (TGP) on the immune system. II. TGP stimulates the proliferation of human T

cells and the differentiation of human B cells into Ig secreting cells. *J Immunol*. 1988 Mar 15;140(6):1823-9. Erratum in: *J Immunol* 1988;140(12):4413.

29. Barnes PJ. The cytokine network in asthma and chronic obstructive pulmonary disease. *J Clin Invest*. 2008; 118(11):3546-56.

30. Sietta M, Mariani M, Panina-Bordignon P, Turato G, Buonsanti C, Baraldo S, Bellettato CM, Papi A, Corbetta L, Zuin R, Sinigaglia F, Fabbri LM. Increased expression of the chemokine receptor CXCR3 and its ligand CXCL10 in peripheral airways of smokers with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2002; 165(10):1404-9.

31. Barnes PJ, Shapiro SD, Pauwels RA. Chronic obstructive pulmonary disease: molecular and cellular mechanisms. *Eur Respir J*. 2003; 22(4):672-88.

32. Russell RE, Thorley A, Culpitt SV, Dodd S, Donnelly LE, Demattos C, Fitzgerald M, Barnes PJ. Alveolar macrophage-mediated elastolysis: roles of matrix metalloproteinases, cysteine, and serine proteases. *Am J Physiol Lung Cell Mol Physiol*. 2002; 283(4):L867-73.

33. Roth M. Pathogenesis of COPD. Part III. Inflammation in COPD. *Int J Tuberc Lung Dis*. 2008; 12(4):375-80.

34. Terashima T, Wiggs B, English D, Hogg JC, van Eeden SF. Phagocytosis of small carbon particles (PM10) by alveolar macrophages stimulates the release of polymorphonuclear leukocytes from bone marrow. *Am J Respir Crit Care Med*. 1997; 155(4):1441-7.

35. Barnes PJ, Cosio MG. Characterization of T lymphocytes in chronic obstructive pulmonary disease. *PLoS Med*. 2004; 1(1):e20.

36. Grumelli S, Corry DB, Song LZ, Song L, Green L, Huh J, Hacken J, Espada R, Bag R, Lewis DE, Kheradmand F. An immune basis for lung parenchymal destruction in chronic obstructive pulmonary disease and emphysema. *PLoS Med*. 2004; 1(1):e8.

37. Turato G, Di Stefano A, Maestrelli P, Mapp CE, Ruggieri MP, Roggeri A, Fabbri LM, Sietta M. Effect of smoking cessation on airway inflammation in chronic bronchitis. *Am J Respir Crit Care Med*. 1995; 152 (4 Pt 1):1262-7.

38. Janoff A, Sloan B, Weinbaum G, Damiano V, Sandhaus RA, Elias J, Kimbel P. Experimental emphysema induced with purified human neutrophil elastase: tissue localization of the instilled protease. *Am Rev Respir Dis*. 1977; 115(3):461-78.

39. Nakamura H, Yoshimura K, McElvaney NG, Crystal RG. Neutrophil elastase in respiratory epithelial lining fluid of individuals with cystic fibrosis induces interleukin-8 gene expression in a human bronchial epithelial cell line. *J Clin Invest*. 1992; 89(5):1478-84.

40. Finlay GA, O'Driscoll LR, Russell KJ, D'Arcy EM, Masterson JB, FitzGerald MX, O'Connor CM. Matrix metalloproteinase expression and production by alveolar macrophages in emphysema. *Am J Respir Crit Care Med*. 1997; 156(1):240-7.

41. Ohnishi K, Takagi M, Kurokawa Y, Satomi S, Konttinen YT. Matrix metalloproteinase-mediated extracellular matrix protein degradation in human pulmonary emphysema. *Lab Invest*. 1998 Sep; 78(9):1077-87.

42. Imai K, Dalal SS, Chen ES, Downey R, Schulman LL, Ginsburg M, D'Armiento J. Human collagenase (matrix metalloproteinase-1) expression in the lungs of patients with emphysema. *Am J Respir Crit Care Med*. 2001; 163(3 Pt 1):786-91.
43. Abboud RT, Vimalanathan S. Pathogenesis of COPD. Part I. The role of protease antiprotease imbalance in emphysema. *Int J Tuberc Lung Dis*. 2008; 12(4):361-7.
44. Russell RE, Culpitt SV, DeMatos C, Donnelly L, Smith M, Wiggins J, Barnes PJ. Release and activity of matrix metalloproteinase-9 and tissue inhibitor of metalloproteinase-1 by alveolar macrophages from patients with chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol*. 2002; 26(5):602-9.
45. Finlay GA, Russell KJ, McMahon KJ, D'arcy EM, Masterson JB, FitzGerald MX, O'Connor CM. Elevated levels of matrix metalloproteinases in bronchoalveolar lavage fluid of emphysematous patients. *Thorax*. 1997; 52(6):502-6.
46. Atkinson JJ, Senior RM. Matrix metalloproteinase-9 in lung remodeling. *Am J Respir Cell Mol Biol*. 2003; 28(1):12-24.
47. Avilés B, Belda J, Margarit G, Bellido-Casado J, Martínez-Brú C y Casan P, Marcadores de remodelado bronquial en el esputo inducido de fumadores sanos. *Arch. Bronconeumol*. 2006; 42(5):253-240.
48. Lomas DA, Silverman EK, Edwards LD, Miller BE, Coxson HO, Tal-Singer R; Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) investigators. Evaluation of serum CC-16 as a biomarker for COPD in the ECLIPSE cohort. *Thorax*. 2008; 63(12):1058-63.
49. Comhair SA, Erzurum SC. Antioxidant responses to oxidant-mediated lung diseases. *Am J Physiol Lung Cell Mol Physiol*. 2002; 283(2):L246-55.
50. Majo J, Ghezzi H, Cosio MG. Lymphocyte population and apoptosis in the lungs of smokers and their relation to emphysema. *Eur Respir J*. 2001; 17(5):946-53.
51. Montuschi P, Kharitonov SA, Barnes PJ. Exhaled carbon monoxide and nitric oxide in COPD. *Chest*. 2001; 120(2):496-501.
52. Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. *Lancet*. 2005; 366(9489):941-51. *Lancet* 2006;367:28.
53. Kalsheker N, Morley S, Morgan K. Gene regulation of the serine proteinase inhibitors alpha1-antitrypsin and alpha1-antichymotrypsin. *Biochem Soc Trans*. 2002; 30(2):93-8.
54. Vidal R, Blanco I, Casas F, Jardí R, Miravittles M; and Committee on the National Registry of Individuals with Alpha-1 Antitrypsin Deficiency. Guidelines for the diagnosis and management of alpha-1 antitrypsin deficiency. *Arch Bronconeumol* 2006; 42: 645-659.
55. Devereux G. ABC of chronic obstructive pulmonary disease. Definition, epidemiology, and risk factors. *BMJ*. 2006; 332(7550):1142-4.
56. Fregonese L, Stolk J, Frants RR, Veldhuisen B. Alpha-1 antitrypsin Null mutations and severity of emphysema. *Respir Med* 2008; 102:876-884.

57. Dahl M, Hersh CP, Ly NP, Berkey CS, Silverman EK, Nordestgaard BG. The protease inhibitor PI*S allele and COPD: a meta-analysis. *Eur Respir J* 2005; 26:67-76.
58. American Thoracic Society; European Respiratory Society. American Thoracic Society/European Respiratory Society statement: standards for the diagnosis and management of individuals with alpha-1 antitrypsin deficiency. *Am J Respir Crit Care Med* 2003; 168:818-900.
59. Dahl M, Nordestgaard BG, Lange P, Vestbo J, Tybjaerg-Hansen A. Molecular diagnosis of intermediate and severe alpha(1)-antitrypsin deficiency: MZ individuals with chronic obstructive pulmonary disease may have lower lung function than MM individuals. *Clin Chem* 2001; 47:56-62.
60. Dahl M, Tybjaerg-Hansen A, Lange P, Vestbo J, Nordestgaard BG. Change in lung function and morbidity from chronic obstructive pulmonary disease in alpha1-antitrypsin MZ heterozygotes: A longitudinal study of the general population. *Ann Intern Med* 2002; 136:270-279.
61. Senn O, Russi EW, Schindler C, *et ál.* Circulating alpha1-antitrypsin in the general population: determinants and association with lung function . *Respir Res* 2008; 9:35.
62. Cordell HJ, Clayton DG. Genetic association studies. *Lancet*. 2005; 366:1121-31.
63. Silverman EK, Palmer LJ. Case-control association studies for the genetics of complex respiratory diseases. *Am J Respir Cell Mol Biol* 2000; 22:645-8.
64. Pérez Rubio G, Silva-Zolezzi I, Ramírez-Venegas A, Sansores RH, Reséndiz-Hernández JM, Montaña M, Camarena A y Falfán-Valencia R. Aspectos actuales de los estudios de asociación de genoma completo (GWAS) en la enfermedad pulmonar obstructiva crónica. *Rev Inst Nal Enf Resp Mex*. 2009, 22(4):337-346.
65. Jimenez-Sanchez G, Silva-Zolezzi I, Hidalgo A, March S. Genomic medicine in Mexico: initial steps and the road ahead. *Genome Res* 2008; 18:1191-8.
66. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; 449:851–861.
67. 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; 106:8611-6.
68. Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res* 2009; 37:W600-5.
69. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA* 2008; 299:1335-44.
70. Board PG, Coggan M, Chelvanayagam G, Eastal S, Jermini LS, Schulte GK, Danley DE, Hoth LR, Griffon MC, Kamath AV, Rosner MH, Chrnyk BA, Perregaux DE, Gabel CA, Geoghegan KF, Pandit J. Identification, characterization, and crystal structure of the Omega class glutathione transferases. *J Biol Chem*. 2000; 275(32):24798-806.

71. Yanbaeva DG, Wouters EF, Dentener MA, Spruit MA, Reynaert NL. Association of glutathione-S-transferase omega haplotypes with susceptibility to chronic obstructive pulmonary disease. *Free Radic Res.* 2009;43(8):738-43.
72. Lakhdar R, Denden S, Knani J, Leban N, Daimi H, Hassine M, Lefranc G, Ben Chibani J, Haj Khelil A. Relationship between glutathione S-transferase P1 polymorphisms and chronic obstructive pulmonary disease in a Tunisian population. *Genet Mol Res.* 2010; 9(2):897-907.
73. Fattman CL, Schaefer LM, Oury TD. Extracellular superoxide dismutase in biology and medicine. *Free Radic Biol Med.* 2003; 35(3):236-56.
74. Juul K, Tybjaerg-Hansen A, Marklund S, Lange P, Nordestgaard BG. Genetically increased antioxidative protection and decreased chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 2006; 173(8):858-64.
75. Hersh CP, Demeo DL, Lazarus R, Celedon JC, Raby BA, Benditt JO, Criner G, Make B, Martinez FJ, Scanlon PD: Genetic association analysis of functional impairment in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2006; 173(9):977-984.
76. Hersh CP, DeMeo DL, Reilly JJ, Silverman EK: Xenobiotic metabolizing enzyme gene polymorphisms predict response to lung volume reduction surgery. *Respir Res* 2007; 8:59.
77. Gingo MR, Silveira LJ, Miller YE, Friedlander AL, Cosgrove GP, Chan ED, Maier LA, Bowler RP. Tumour necrosis factor gene polymorphisms are associated with COPD. *Eur Respir J.* 2008;31(5):1005-12.
78. Córdoba-Lanús E, de-Torres JP, López-Aguilar C, Rodríguez-Pérez MC, Maca-Meyer N, Montejo-de-Garcini A, Aguirre-Jaime A, Pérez-Méndez L, Casanova C. Association of IL-6 gene polymorphisms and COPD in a Spanish population. *Respir Med.* 2008;102(12):1805-11.
79. Yanbaeva DG, Dentener MA, Spruit MA, Houwing-Duistermaat JJ, Kotz D, Passos VL, Wouters EF. *IL6* and *CRP* haplotypes are associated with COPD risk and systemic inflammation: a case-control study. *BMC Med Genet.* 2009;10:23.
80. Foreman MG, DeMeo DL, Hersh CP, Carey VJ, Fan VS, Reilly JJ, Shapiro SD, Silverman EK. Polymorphic variation in surfactant protein B is associated with COPD exacerbations. *Eur Respir J.* 2008; 32(4):938-44.
81. Foreman MG, Kong X, Demeo DL, Pillai SG, Hersh CP, Bakke P, Gulsvik A, Lomas DA, Litonjua AA, Shapiro SD, Tal-Singer R, Silverman EK. Polymorphisms in Surfactant Protein D are Associated with COPD. *Am J Respir Cell Mol Biol.* 2010; 6.
82. Fujimoto K, Ikeda S, Arai T, Tanaka N, Kumasaka T, Ishii T, Kida K, Muramatsu M, Sawabe M. Polymorphism of *SERPINE2* gene is associated with pulmonary emphysema in consecutive autopsy cases. *BMC Med Genet.* 2010;11:159.
83. 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:S8.
84. 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.

85. 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:e100042
86. Lemjabbar-Alaoui H, Dasari V, Sidhu SS, Mengistab A, Finkbeiner W, Gallup M, Basbaum C. Wnt and Hedgehog are critical mediators of cigarette smoke-induced lung cancer. *PLoS One*. 2006;1:e93.
87. Marsh WL. Molecular defects associated with McLeod blood group phenotype. En. Salmon E, ed. *Blood groups and other red cells surface markers in health and disease* 1982:17-82.
88. Straface E, Matarrese P, Gambardella L, Forte S, Carlone S, Libianchi E, Schmid G, Malorni W. N-Acetylcysteine counteracts erythrocyte alterations occurring in chronic obstructive pulmonary disease. *Biochem Biophys Res Commun*. 2000;279(2):552-6.
89. Lucantoni G, Pietraforte D, Matarrese P, Gambardella L, Metere A, Paone G, Bianchi EL, Straface E. The red blood cell as a biosensor for monitoring oxidative imbalance in chronic obstructive pulmonary disease: an ex vivo and in vitro study. *Antioxid Redox Signal*. 2006; 8(7 8):1171-82.
90. Kong X, Cho MH, Anderson W, Coxson HO, Muller N, Washko G, Hoffman EA, Bakke P, Gulsvik A, Lomas DA, Silverman EK, Pillai SG; ECLIPSE Study NETT Investigators. Genome-wide Association Study Identifies BICD1 as a Susceptibility Gene for Emphysema. *Am J Respir Crit Care Med*. 2011; 183(1):43-9.
91. Bakke PS, Zhu G, Gulsvik A, Kong X, Agusti AG, Calverley PM, Donner CF, Levy RD, Make BJ, Paé PD, Rennard SI, Vestbo J, Wouters EF, Anderson W, Lomas DA, Silverman EK, Pillai SG. Candidate genes for COPD in two large data sets. *Eur Respir J*. 2011; 37(2):255-63.
92. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, Zhao JH, Ramasamy A, Zhai G, Vitart V, Huffman JE, Igl W, Albrecht E, Deloukas P, Henderson J, Granell R, McArdle WL, Rudnicka AR; Wellcome Trust Case Control Consortium, Barroso I, Loos RJ, Wareham NJ, Mustelin L, Rantanen T, Surakka I, Imboden M, Wichmann HE, Grkovic I, Jankovic S, Zgaga L, Hartikainen AL, Peltonen L, Gyllensten U, Johansson A, Zaboli G, Campbell H, Wild SH, Wilson JF, Gläser S, Homuth G, Völzke H, Mangino M, Soranzo N, Spector TD, Polasek O, Rudan I, Wright AF, Heliövaara M, Ripatti S, Pouta A, Naluai AT, Olin AC, Torén K, Cooper MN, James AL, Palmer LJ, Hingorani AD, Wannamethee SG, Whincup PH, Smith GD, Ebrahim S, McKeever TM, Pavord ID, MacLeod AK, Morris AD, Porteous DJ, Cooper C, Dennison E, Shaheen S, Karrasch S, Schnabel E, Schulz H, Grallert H, Bouatia-Naji N, Delplanque J, Froguel P, Blakey JD; NSHD Respiratory Study Team, Britton JR, Morris RW, Holloway JW, Lawlor DA, Hui J, Nyberg F, Jarvelin MR, Jackson C, Kähönen M, Kaprio J, Probst-Hensch NM, Koch B, Hayward C, Evans DM, Elliott P, Strachan DP, Hall IP, Tobin MD. Genome-wide association study identifies five *loci* associated with lung function. *Nat Genet*. 2010;42(1):36-44.
93. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marcianti KD, Franceschini N, van Durme YM, Chen TH, Barr RG, Schabath MB, Couper DJ, Brusselle GG, Psaty BM, van Duijn CM, Rotter JI, Uitterlinden AG, Hofman A, Punjabi NM, Rivadeneira F, Morrison AC, Enright PL, North KE, Heckbert SR, Lumley T, Stricker BH, O'Connor GT, London SJ. Meta-analyses of genome-wide association studies identify multiple *loci* associated with pulmonary function. *Nat Genet*. 2010; 42(1):45-52.

94. Illumina Technical note: DNA analysis <http://www.illumina.com.support.ilmn>, 2010, consultada 3 de Enero de 2011.
95. NCBI <http://www.ncbi.nlm.nih.gov/pubmed/>
96. HapMap Project <http://hapmap.ncbi.nlm.nih.gov/>
97. Paré P. A strategy for finding susceptibility genes and polymorphisms for COPD. *Exp Lung Res.* 2005 Sep;31 Suppl 1:76.
98. HaploView <http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview>
99. EIGENSOFT 3.0. <http://genepath.med.harvard.edu/~reich/Software.htm>.
100. ADMIXTURE <http://www.genetics.ucla.edu/software> for a free download.
101. Illumina, GoldenGateR Genotyping Assay, Automated. Experienced User Card (Universal BeadChip)
102. PLINK Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007). <http://pngu.mgh.harvard.edu/~purcell/plink/index.shtml>.
103. Hersh CP, Demeo DL, Lange C, Litonjua AA, Reilly JJ, Kwiatkowski D, Laird N, Sylvia JS, Sparrow D, Speizer FE, Weiss ST, Silverman EK. Attempted replication of reported chronic obstructive pulmonary disease candidate gene associations. *Am J Respir Cell Mol Biol.* 2005; 33(1):71-8.
104. Keatings VM, Collins PD, Scott DM, Barnes PJ. Differences in interleukin-8 and tumor necrosis factor- α in induced sputum from patients with chronic obstructive pulmonary disease or asthma. *Am J Respir Crit Care Med* 1996; 153:530–534.
105. Lee YL, Chen W, Tsai WK, Lee JC, Chiou HL, Shih CM, Wang YC. Polymorphisms of p53 and p21 genes in chronic obstructive pulmonary disease. *J Lab Clin Med.* 2006; 147(5):228-33.
106. Bierut LJ, Madden PA, Breslau N, Johnson EO, Hatsukami D, Pomerleau OF, Swan GE, Rutter J, Bertelsen S, Fox L, Fugman D, Goate AM, Hinrichs AL, Konvicka K, Martin NG, Montgomery GW, Saccone NL, Saccone SF, Wang JC, Chase GA, Rice JP, Ballinger DG. Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum Mol Genet.* 2007; 16(1):24-35.
107. Stevens VL, Bierut LJ, Talbot JT, Wang JC, Sun J, Hinrichs AL, Thun MJ, Goate A, Calle EE. Nicotinic receptor gene variants influence susceptibility to heavy smoking. *Cancer Epidemiol Biomarkers Prev.* 2008; 17(12):3517-25.
108. Draheim KM, Chen HB, Tao Q, Moore N, Roche M, Lyle S. ARRDC3 suppresses breast cancer progression by negatively regulating integrin beta4. *Oncogene.* 2010; 29(36):5032-47.
109. Srougi MC, Burrige K. The Nuclear Guanine Nucleotide Exchange Factors Ect2 and Net1 Regulate RhoB-Mediated Cell Death after DNA Damage. *PLoS One.* 2011; 6(2):e17108.

Anexo A

Fragmento de la tabla realizada donde fueron capturados datos importantes reportados en otros estudios sobre regiones génicas y SNPs asociados a EPOC, la información fue complementada con las frecuencias alélicas de los marcadores hallados consultando la base de datos del HapMap y del PDGM del INMEGEN.

Gen	Chr	SNP	Posición	Cambio	Alelo menor	Alelo asociado	p	OR	Frec alélica		REF
									CEU	MEX	
ADARB1	21	rs2225434	45458574 (+)	C/T	T	T	9.68 E-06	ND	0.367	ND	83
SYT10	12	rs1994169	33436783 (-)	A/C	C	C	9.55 E-06	ND	0.267	ND	
CCBL2	1	rs1409149	89201737 (+)	C/T	C	C	9.12 E-06	ND	0.492	ND	
intergénica	14	rs808225	57467669 (+)	A/G	G	G	7.38 E-06	ND	0.433	0.413	
NID2	14	rs730532	51588561 (-)	C/T	C	C	5.89 E-06	ND	0.217	ND	
KIAA1797	9	rs7851363	20748306 (+)	A/T	A	A	5.7 E-05	ND	0.192	ND	
RHBDD1	2	rs3820928	227481710 (-)	A/G	G	G	5.33 E-06	ND	0.392	0.460	
ETAA16	2	rs6740919	67473559 (+)	C/G	C	C	5.17 E-05	ND	0.280	0.382	
SNTG1	8	rs310558	51575144 (-)	A/G	G	G	5.14 E-06	ND	0.150	ND	
MTHFD1L	6	rs491552	151352748 (-)	C/G	C	C	3.86 E-05	ND	0.333	0.331	
Intergénica	3	rs4974153	55944898 (+)	G/T	T	T	9.55E-05	1.434	0.217	0.277	84
	1	rs3767943	41066006 (+)	A/G	A	A	9.18E-05	1.611	0.184	0.415	
	6	rs9350301	21048740 (+)	G/T	T	T	8.93E-05	0.681	0.209	0.319	
	6	rs4712564	21028102 (+)	A/G	A	A	8.79E-05	0.673	0.242	0.263	
	11	rs10769813	7598607 (+)	A/G	G	G	8.54E-05	0.709	0.483	0.388	
	14	rs200303	98472186 (+)	G/T	T	T	7.96E-05	1.418	0.367	0.384	
	18	rs7233241	49347873 (+)	C/T	T	T	7.76E-05	1.399	0.350	0.278	
	15	rs16943236	84034355 (+)	C/T	C	C	7.68E-05	0.346	0.102	0.222	
	22	rs2269640	24667911 (+)	C/T	C	C	7.54E-05	1.498	0.183	0.445	
2	rs935381	168532380 (+)	C/T	T	T	7.47E-05	1.731	0.250	ND		

<i>TNFSF8</i>	9	rs10817691	116760083 (+)	A/C	A	A	9.83E-05	0.068*	0.472	ND	85
<i>NR2C1</i>	12	rs7971079	93967620 (+)	A/G	G	G	9.83E-05	-0.076*	0.246	ND	
<i>ONECUT1</i>	15	rs2460	50860376 (+)	A/G	A	A	9.65E-05	-0.078*	0.258	ND	
<i>UTP20</i>	12	rs10507130	100278085 (+)	A/G	A	A	9.64E-05	0.078*	0.225	ND	
<i>UTP20</i>	12	rs11609516	100279051 (+)	C/G	G	G	9.54E-05	0.078*	0.225	0.339	
<i>GPR126</i>	6	rs7753012	142787576 (+)	G/T	G	G	9.50E-05	0.072*	0.297	0.531	
<i>SLC39A12</i>	10	rs2497782	18359047 (+)	A/G	G	G	9.44E-05	0.072	0.300	ND	
<i>UTP20</i>	12	rs7312071	98759143 (+)	A/G	G	G	9.41E-05	0.067*	0.383	ND	
<i>CBLN2</i>	18	rs638052	68379883 (+)	C/T	C	C	9.37E-05	-0.119*	0.108	ND	
<i>TMEM16B</i>	12	rs6489665	100223113 (+)	A/G	G	T	9.36E-05	0.21*	0.383	ND	
<i>DAAM2</i>	6	rs2395730	39502856 (+)	A/C	C	C	7.98E-08	0.044	0.396	0.407	92
<i>HTR4</i>	5	rs3995090	142992760 (+)	A/C	C	C	4.29E-09	0.038	0.479	0.345	
<i>TNS1</i>	2	rs2571445	210537107 (-)	C/T	C	G	1.11E-12	0.035	0.477	0.300	
<i>GPR126</i>	6	rs11155242	140257798 (+)	A/C	C	A	9.13E-09	-0.54	0.200	0.130	93
<i>ADAM19</i>	5	rs2277027	152020670 (+)	A/C	C	A	8.32E-10	0.49	0.308	0.450	
<i>NT5DC1</i>	6	rs1052443	114148966 (+)	A/C	A	A	7.50E-07	26.12	0.430	0.460	
<i>HHIP</i>	4	rs1032295	141165490 (+)	G/T	G	T	6.28E-09	-0.47	0.449	0.290	
<i>HTR4</i>	5	rs11168048	142989297 (+)	C/T	C	T	5.37 E-08	-0.43	0.381	ND	
<i>TSPYL4</i>	6	rs3749893	114151888 (+)	A/G	A	A	5.35E-07	-26.55	0.375	0.455	
<i>GPR126</i>	6	rs3817928	140316763 (+)	A/G	G	A	2.60E-10	-0.59	0.200	0.135	
<i>GPR126</i>	6	rs6937121	140273382 (+)	G/T	G	T	2.46E-09	-0.49	0.263	0.260	
<i>FAM13A</i>	4	rs6830970	85528992 (+)	A/G	G	A	1.92E-08	0.46	0.275	0.545	
<i>HTR4</i>	5	rs7735184	142991337 (+)	G/T	T	T	1.74E-07	0.4	0.367	0.345	

*El valor reportado en el artículo original es β . ND indica que su frecuencia alélica no está determinada en la población correspondiente.

Anexo B

GoldenGate Genotyping Assay, Automated.

Experienced User Card (Universal BeadChip)

Días antes de comenzar con la genotipificación debe de guardarse los mapas de distribución (beadmap) para cada microarreglo según el código de barras asignado por la empresa.

Preparación de las muestras

Descongelar a temperatura ambiente las muestras de DNA y homogeneizar en vortex.

Hacer el diseño para que en una placa de 96 pozos se coloquen las muestras a genotipificar.

Normalizar las muestras de DNA a 50 ng/ μ L con búfer TE en la placa de 96 pozos.

Colocar 5 μ L de cada una de las muestras de DNA en otra placa de 96 pozos que denominaremos placa SUD

Pre-Amplificación

Precalentar el termobloque a 95°C.

Encender el termosellador.

Descongelar a temperatura ambiente el reactivo MS1, agitar en vortex y depositar su contenido en un reservorio de plástico nuevo.

Agregar a cada pozo de la placa SUD con una pipeta multicanal, 5 μ L de reactivo MS1.

Colocar una cubierta de aluminio a la placa y sellar en el termosellador por 3 segundos.

Centrifugar un pulso a 250 xg la placa SUD.

Agitar en vortex a 2300 rpm por 20 segundos.

Centrifugar un pulso a 250 xg.

Colocar la placa SUD en el termobloque e incubar a 95°C por 30 min.

Una vez terminada la incubación colocar inmediatamente un bloque por 3 minutos para evitar que se deforme la placa.

Centrifugar a 3000 xg 2 minutos.

Precalentar el termobloque a 70°C.

Preparar el robot para ser usado en el modo "SUD/Precip SUD" e indicarle el número de placas a trabajar.

Mantener a temperatura ambiente la placa SUD y descongelar el tubo PS1

Colocar un reservorio de un cuarto o chico en la posición A del tablero del robot, de acuerdo con el mapa del monitor de la PC y dispensar PS1 como sigue:

- 2 mL para 1 placa.
- 2.5 mL para 2 placas.
- 3 mL para 3 placas

Colocar un reservorio mediano en la posición B indicada por el monitor de acuerdo con la posición del tablero del robot. Dispensar 2-propanol como se indica:

- 9 mL para una placa
- 17 mL para 2 placas
- 25 mL para 3 placas

Colocar la placa SUD en el robot de acuerdo a la posición indicada por el monitor.

Iniciar el robot presionando en run y tecleando la contraseña.

Cuando el robot finalice, sellar cada placa SUD con una cubierta adhesiva.

Agitar la placa a 1600 rpm por 20 segundos o hasta que los pozos se observen completamente azules.

Centrifugar la placa SUD a 3000 xg por 20 min. Al finalizar deben ser visibles los botones azules en la parte inferior de cada uno de los pozos.

Si no sucede este fenómeno se debe de centrifugar nuevamente la placa a 3000 xg por 10 minutos mas.

Remover la cubierta plástica y decantar el sobrenadante por inversión. Se debe de retirar el exceso de sobrenadante golpeando suavemente la placa invertida sobre un papel absorbente.

Colocar la placa boca arriba por 1 hora a temperatura ambiente, para permitir la evaporación del resto de sobrenadante y secar los botones precipitados de DNA.

Una vez terminado el tiempo de evaporación seleccionar desde la PC del robot, el modo "Select SUD/Resuspend SUD".

Teclear el número correcto de placas.

Colocar un reservorio de chico en la posición A que indica el robot y dispensar RS1 como sigue:

- 3.5 mL para 1 placa
- 6 mL para 2 placas
- 8.5 mL para 3 placas

Colocar la placa SUD en el tablero del robot de acuerdo al mapa del monitor para este procedimiento.

Iniciar el robot presionando en run y tecleando la contraseña.

Cuando el robot termine, sellar cada placa con una cubierta adhesiva translúcida.

Agitar en vortex la placa a 2300 rpm por 20 segundos o hasta que los pozos se observen uniformes en coloración azul.

Centrifugar la placa a 280 xg por 1 min.

Precalentar el termobloque a 70°C.

Descongelar los tubos de OPA y OB1 a temperatura ambiente y homogeneizar en vortex.

Tomar una nueva placa la cual identificaremos como ASE

Preparar el robot para su uso seleccionando el modo "SUD/Make SUD ASE" e indicándole el número de placas.

Centrifugar los tubos de OPA y OB1 a 280 xg un pulso.

Colocar los tubos de OPA y OB1 en la distribución que marca el monitor.

Centrifugar la placa SUD a 280 xg por 1 min.

Colocar las placas SUD y ASE en el tablero del robot de acuerdo a la posición indicada en el monitor de la PC.

Iniciar el robot presionando en run y tecleando la contraseña.

Cuando termine el robot, colocar una cubierta de aluminio a la placa ASE y sellar en el termosellador por 3 segundos.

Agitar la placa ASE en vortex a 1600 rpm por 1 min.

Colocar cada placa ASE en el termobloque previamente precalentado a 70°C y cerrar su tapa.

Inmediatamente cambiar la temperatura del termobloque a 30°C e incubar la placa ASE por 2 horas.

Una vez terminada la incubación precalentar el termobloque a 45°C y dejar que se equilibre.

Descongelar los reactivos MEL, AM1 Y UB1 a temperatura ambiente. Invertir 10 veces los tubos para mezclar su contenido.

Preparar el robot para su uso seleccionando "Biochem/Add MEL" e indicando el número de placas a trabajar.

Colocar el tubo MEL en la gradilla del robot de acuerdo a la distribución.

Colocar un reservorio pequeño en la posición A del robot y dispensar AM1 de la siguiente manera:

- 15 mL para 1 placa.
- 30 mL para 2 placas
- 45 mL para 3 placas.

Colocar un reservorio mediano en la posición B del robot y dispensar UB1 como sigue:

- 15 mL para 1 placa ASE
- 30 mL para 2 placas ASE

- 45 mL para 3 placas ASE

Centrifugar cada una de las placas ASE por 1 min a 280 xg.

Colocar la placa ASE en la barra magnética del robot de acuerdo al mapa de distribución de la pantalla.

Esperar a que todas las perlas sean capturadas por el magneto, el cual tarda cerca de 2 min. Entonces retirar la cubierta de la placa.

Iniciar el robot presionando en run y tecleando la contraseña.

Cuando de aviso el robot, retirar la placa ASE del tablero y sellarla con una cubierta adhesiva.

Agitar en vortex la placa ASE a 1600 rpm por 1 min.

Regresar la placa ASE a la posición indicada por el robot, retirar la cubierta plástica y presionar OK

Nuevamente cuando indique el robot, remover la placa ASE y sellarla con una cubierta adhesiva.

Agitar en vortex a 1600 rpm por 1 min, hasta que todas las perlas se encuentren completamente resuspendidas.

Regresar la placa a su posición correcta de acuerdo al tablero, retirar la cubierta adhesiva y presionar OK

Cuando finalice el robot, retirar la placa ASE

Sellar la placa con una cubierta adhesiva transparente.

Agitar la placa en vortex a 1725 rpm por 1 min.

En el robot presionar OK.

Colocar la placa ASE en el termobloque precalentado a 45°C, cerrar la tapa e incubar por exactamente 15 min.

Descongelar el tubo del reactivo MMP a temperatura ambiente y mezclar por inversión.

Preparar el robot para su uso.

Colocar el código de barras que indique PCR a cada placa nueva.

Desde la PC del robot, seleccionar "Biochem/Make PCR" e indicar el número de placas.

Agregar 64 μ L de DNA polimerasa a cada tubo de MMP.

Agregar 50 μ L de Uracil DNA Glucosilasa a cada tubo de MMP.

Tapar el tupo de MMP y mezclar por inversión 10 veces.

Colocar cada tubo MMP en la gradilla del robot de acuerdo a la pantalla correspondiente

Colocar una nueva placa que marcaremos como PCR en el robot de acuerdo a la distribución indicada en la pantalla de la PC.

Iniciar el robot presionando en run y tecleando la contraseña.

Una vez finalizado el proceso sellar la placa con una cubierta adhesiva translúcida.

Precalentar el termobloque a 95°C.

Descongelar el tubo de IP1 a temperatura ambiente y mezclar en vortex.

Seleccionar desde la PC del robot "Biochem/Inoc PCR" e indicar el número de placas con las que se está trabajando

Invertir 10 veces el tubo IP1 y centrifugar a 280 xg.

Colocar el tubo en la gradilla del robot de acuerdo a la distribución indicada.

Colocar un reservorio chico en la posición A del tablero del robot y dispensar el reactivo UB1 de la siguiente manera:

- Para 1 placa 10 mL
- Para 2 placas 15 mL
- Para 3 placas 20 mL

Colocar cada la placa ASE en el robot de acuerdo a la distribución indicada.

Colocar la placa de PCR en el robot de en su posición indicada en la PC.

Iniciar el robot presionando en run y tecleando la contraseña.

Cuando el monitor lo indique remover la placa ASE y sellarla con una adhesiva transparente.

Agitar en vortex a 1900 rpm por 1 min.

Incubar por 1 min exacto la placa en el termobloque a 95°C.

Regresar la placa a su posición original, remover la cubierta plástica y seleccionar Ok en la PC del robot para continuar con el proceso.

El robot esperará 1 min a que las perlas se fijen completamente y entonces incorporará el contenido desnatularizado de la placa ASE a placa de PCR.

Remover la placa ASE y de PCR del robot. Desechar la placa ASE y sellar la placa PCR con una cubierta especial para pasarla al termociclador.

Amplificación

Una vez colocada la placa de PCR en el termociclador usar el siguiente programa para la amplificación.

Programa	Temperatura	Tiempo	Ciclos
1	37°C	10 min	1
2	95°C	3 min	1
	95°C	35 segundos	
3	56°C	35 segundos	34
	72°C	2 min	
4	72°C	10 min	1
5	4°C	5 min	1

La placa se puede quedar durante toda la noche en el termociclador.

Post-Amplificación

Descongelar el reactivo MPB a temperatura ambiente.

Poner a temperatura ambiente y protegida de la luz cada placa de PCR por 15 min.

Seleccionar en el robot el modo "Post-PCR Task/Bind PCR" indicando el número de placas.

Agitar en vortex el tubo MPB por 30 segundos hasta que se resuspendan las perlas.

Colocar el tubo MPB en el robot de acuerdo a la distribución del robot.

Dispensar 5 mL de 2-butanol por cada placa PCR a procesar en un reservorio chico en la posición A.

Centrifugar la placa PCR a 3000 xg por 30 segundos.

Colocar la placa PCR en el robot de acuerdo a la distribución indicada la PC. Posteriormente retirar la cubierta plástica.

Iniciar el robot presionando en run y tecleando la contraseña.

Cuando el robot finalice, sellar con una cubierta adhesiva la placa de PCR.

Incubar la placa de PCR en un lugar seguro y protegida de la luz por 1 hora a temperatura ambiente.

Programar el horno de hibridación a 60°C

A una placa nueva de 96 pozos se le denominará INT mientras que a otra placa de 384 pozos se le denominará HYB.

Seleccionar desde la PC del robot, "Post-PCR Task/Make Hyb" e indicar el número de placas.

Centrifugar un pulso a 280 xg, el tubo del reactivo MH1.

Colocar el tubo MH1 de acuerdo a la distribución del robot.

Colocar un reservorio chico en la posición B del robot y agregar 10 mL de NaOH 0.1 N por placa de hibridación.

Colocar un reservorio chico en la posición A del robot con 10 mL del reactivo UB2 por placa de hibridación.

Colocar las placas PCR, INT e HYB de acuerdo a la distribución del robot. Remover la cubierta plástica de la placa de PCR.

Iniciar el robot presionando en run y tecleando la contraseña.

Remover la placa HYB y sellar con una cubierta adhesiva translúcida. Proteger la placa de la luz.

Centrifugar la placa HYB a 3000 xg por 4 min para remover las burbujas.

Proceder con el paso de hibridación al microarreglo.

Sacar los microarreglos del refrigerador 10 minutos antes de iniciar el proceso de hibridación, pero no sacar de su empaque, mantenerlo en la mesa de trabajo por no más de 24 horas en su empaque.

Preparación de las guías de alineación para puntas del robot:

Lavar y secar las guías alineadoras de puntas del robot previo a su uso.

Asegurarse de tener la guía alineadora correcta de puntas para el ensayo GoldenGate. El código de barras de la guía debe decir Guide-C.

Ensamblado de cámaras de hibridación:

Colocar todos los suplementos en la mesa de trabajo:

- Cámara de hibridación (1 por cada 4 microarreglos).
- Empaque de cámara de hibridación (1 por cada cámara).
- Insertos metálicos de microarreglos para la cámara de hibridación (1 por microarreglo).

Colocar el empaque a la cámara de hibridación.

Agregar 200µL de reactivo CHB en cada uno de los 8 pozos de cada cámara de hibridación por usar.

Cubrir la cámara de hibridación con su respectiva tapa.

Mantener cerrada la cámara de hibridación a temperatura ambiente hasta el momento de la incorporación de la muestra de DNA al microarreglo y de la incubación.

Remover los microarreglos de su empaque.

Para los microarreglos de 12 muestras colocar sobre cada alineador a utilizar 2 microarreglos con el código de barras apuntando hacia la parte inferior con las ranuras del alineador.

En la PC del robot seleccionar el modo “Task/Hyb Universal BC” señalando las placas con las que se está trabajando.

Elegir el microarreglo apropiado en la pantalla de diálogo.

Asegurarse que los códigos de barras sean claros. En el panel de “Basic Run Parameters”, elegir el número de placas de hibridación a procesar.

Colocar el alineador de los microarreglos en el robot de acuerdo al mapa de distribución de la PC. Colocando cada alineador con la parte superior izquierda en la posición marcada en el robot.

Si se están usando los microarreglos de 32 muestras, asegurarse que la flecha plateada este apuntando hacia la parte superior izquierda cuando sea colocado el alineador en el robot.

Iniciar el robot presionando en run y tecleando la contraseña.

Presionar OK para confirmar que han sido colocados correctamente los alineadores, microarreglos y la placa de muestras.

El robot transferirá las muestras de las placas a los microarreglos.

Al finalizar el proceso remover cuidadosamente cada guía alineadora de puntas del robot. Posteriormente remover el alineador de los microarreglos del robot y asegurarse que todas las líneas de los microarreglos se encuentren cubiertas de DNA, Recordar que algunas secciones del microarreglo pueden no estar cubiertas en su totalidad.

Nota: Al finalizar la hibridación deben ser lavados y secados inmediatamente los alineadores de las puntas.

Desechar la placa de hibridación una vez transferidas las muestras.

En una cámara de hibridación colocar los insertos en ésta y sobre colocar los microarreglos con el código de barras orientado hacia las ranuras de los insertos.

Asegurar la cámara de hibridación cerrándola adecuadamente.

Colocar la cámara bien sellada con los microarreglos en el horno de hibridación a 60°C.

Incubar por 30 min exactamente a 60°C y de inmediato cambiar a 45°C, dejar incubando mínimo 16 horas y máximo 18 horas.

Agregar 335 mL de etanol absoluto grado biología molecular a la botella de XC4. El volumen final será de 350 mL.

Cerrar la botella y agitar vigorosamente por 15 segundos.

Dejar la botella sobre la mesa en posición horizontal con la etiqueta mirando arriba toda la noche.

Antes de finalizar el tiempo de incubación mezclar nuevamente la botella de XC4 con etanol hasta asegurarse de disolver completamente el polímero.

Encender 2 h antes el escáner "*Bead Array Reader*".

Llenar 2 vasos de lavado con 300 mL de buffer PB1 y un vaso de lavado con XC4.

Sumergir las gradillas para microarreglos en uno de los 2 vasos de lavado de PB1, de tal manera que posteriormente puedan ser removidos fácilmente los microarreglos.

Una vez finalizado el tiempo de incubación para los microarreglos retirar del horno de hibridación las cámaras de hibridación que contienen los microarreglos.

Remover los microarreglos uno por uno y cámara por cámara.

Retirar la cubierta plástica con precaución de cada microarreglo en forma diagonal de manera consistente.

Colocarlos en la gradilla del vaso de lavado con PB1.

Cargar la gradilla con máximo 12 microarreglos. Colocar 6 en la primera parte y 6 en la segunda. Tratando de dejar un espacio entre cada laminilla.

Una vez que todos los microarreglos se encuentren dentro del vaso de lavado con PB1, mover de arriba a abajo 10 veces lentamente, para romper la tensión superficial.

Transferir la gradilla al segundo vaso de lavado con PB1 y dejar reposar por 5 minutos.

Transferir la gradilla al vaso con XC4 y mover de arriba a abajo lentamente 10 veces. Dejar reposar también por 5 min.

Preparar una gradilla para tubos de ensaye colocada sobre papel absorbente para evitar el escurrimiento del remanente de XC4 dentro de un desecador conectado al vacío.

Remover la gradilla con los microarreglos con un movimiento rápido y directo de tal manera que quede la gradilla con los microarreglos mirando hacia arriba.

Retirar cada microarreglo de la gradilla con las pinzas y colocar boca arriba sobre la gradilla del desecador.

Activar la cámara de vacío con al menos una presión de 508 mm de Mercurio (Hg) ó 0.68 barómetros.

Secar por aproximadamente 55 min.

Al finalizar el secado limpiar la parte posterior del microarreglo con un papel absorbente que no desprenda pelusa humedecida con etanol al 70% para remover el exceso de XC4.

Colocar los microarreglos sobre la base de lectura.

Abrir el programa en la computadora del escáner, el programa "BeadScan".

Seleccionar "Scan".

Señalar el tipo de microarreglo que se está usando.

Asegurarse que los mapas de distribución (beadmap) para cada microarreglo se encuentren en la computadora del escáner

Poner la liga exacta de la fuente de datos en el "Data repository" y en "Decode Data".

Cada archivo tiene que haber sido copiado en un folder específico; el archivo es "decode map" (*.dmap) para cada microarreglo.

Usando un lector de código de barras escanear cada laminilla.

Seleccionar el "Scan setting" para microarreglos de 32 muestras: "Universal X (GGGT)" y seleccionar "Scan"