



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN CIENCIAS BIOMÉDICAS
FACULTAD DE MEDICINA

**IDENTIFICACIÓN DE VARIANTES FUNCIONALES EN GENES CANDIDATO
ASOCIADOS CON EL RIESGO AL DESARROLLO DE DISLIPIDEMIA MIXTA EN
GRUPOS INDÍGENAS Y POBLACIÓN MESTIZA MEXICANA**

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:
PEDRO ALFONSO JURADO CAMACHO

DIRECTORA DE TESIS
DRA. LORENA SOFÍA OROZCO OROZCO
INSTITUTO NACIONAL DE MEDICINA GENÓMICA

COMITÉ TUTOR
DRA. SARA FRÍAS VÁZQUEZ
INSTITUTO NACIONAL DE PEDIATRÍA
DR. ÓSCAR ARMANDO PÉREZ MÉNDEZ
INSTITUTO NACIONAL DE CARDIOLOGÍA

CIUDAD DE MÉXICO, MÉXICO, NOVIEMBRE DE 2023



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.

Agradecimientos

A los miembros de mi comité tutor: Dra. Lorena Orozco, Dra. Sara Frías y Dr. Óscar Pérez.

A CONAHCyT por los apoyos económicos asignados con el número de becario 619190

Al Programa de Doctorado en Ciencias Biomédicas de la Universidad Nacional Autónoma de México por brindarme la oportunidad y el privilegio de iniciar y concluir los estudios de doctorado.

Al Instituto Nacional de Medicina Genómica, por proporcionar la infraestructura y los medios necesarios para el desarrollo de este trabajo

Al Dr. Emilio Córdova, por su valiosa asesoría y apoyo constante.

Contenido

1.	ÍNDICE DE TABLAS.....	I
2.	ÍNDICE DE FIGURAS.....	I
3.	LISTA DE ABREVIATURAS	II
4.	GLOSARIO	III
5.	RESUMEN	1
6.	ABSTRACT	2
7.	INTRODUCCIÓN	3
7.1.	ENFERMEDADES METABÓLICAS: DISLIPIDEMIAS	3
7.2.	EPIDEMIOLOGÍA	3
7.3.	ETIOLOGÍA DE LAS DISLIPIDEMIAS	5
7.4.	METABOLISMO DE LÍPIDOS Y LIPOPROTEÍNAS.....	6
7.5.	VÍAS METABÓLICAS DE LOS LÍPIDOS.....	8
7.5.1.	<i>Vía exógena</i>	<i>8</i>
7.5.2.	<i>Vía endógena.....</i>	<i>9</i>
7.5.3.	<i>Transporte reverso de colesterol.</i>	<i>10</i>
7.6.	GENÉTICA DE LAS DISLIPIDEMIAS	11
7.6.1.	<i>Dislipidemias monogénicas.....</i>	<i>11</i>
7.7.	DISLIPIDEMIAS COMO ENFERMEDADES COMPLEJAS	12
7.8.	ESTUDIOS EN POBLACIÓN MEXICANA	12
7.9.	POBLACIÓN MEXICANA	13
8.	JUSTIFICACIÓN.....	15
9.	HIPÓTESIS.....	15
10.	OBJETIVOS	15
10.1.	OBJETIVO GENERAL:	15
10.2.	OBJETIVOS ESPECÍFICOS:	15
11.	MATERIAL Y MÉTODOS	17
11.1.	ESTRATEGIA GENERAL.....	17
11.2.	POBLACIÓN DE ESTUDIO.....	18
11.2.1.	<i>Criterios de inclusión:</i>	<i>19</i>
11.2.2.	<i>Criterios de eliminación:.....</i>	<i>19</i>
11.3.	EXTRACCIÓN DE VARIANTES EN GENES CANDIDATO	19
11.3.1.	<i>Archivos de genotipos</i>	<i>19</i>
11.3.2.	<i>Archivos de fenotipos.....</i>	<i>20</i>
11.4.	ESTUDIO DE ASOCIACIÓN.....	20
11.4.1.	<i>Análisis por variante</i>	<i>20</i>
11.4.2.	<i>Estudio cuantitativo</i>	<i>21</i>
11.4.3.	<i>Análisis por haplotipos</i>	<i>21</i>
11.4.4.	<i>Análisis de variantes raras</i>	<i>21</i>
11.5.	REPLICACIÓN DE RESULTADOS	22
11.6.	ENSAYOS <i>IN SILICO</i>	23
11.7.	ANÁLISIS DE FRECUENCIAS ALÉLICAS	23
11.8.	ANÁLISIS ESTADÍSTICOS ADICIONALES.....	23

12. RESULTADOS.....	24
12.1. CARACTERÍSTICAS CLÍNICAS Y DEMOGRÁFICAS DE LA POBLACIÓN DE ESTUDIO	24
12.2. VARIANTES DE UN SOLO NUCLEÓTIDO EN LOS GENES CANDIDATO	25
12.3. CATÁLOGO DE LAS VARIANTES	26
12.4. ESTUDIO DE ASOCIACIÓN CASO-CONTROL.	27
12.5. ESTUDIO DE ASOCIACIÓN DE HAPLOTIPOS.	32
12.6. ANÁLISIS BASADOS EN GEN	34
12.7. ESTUDIO DE REPLICACIÓN EN AMERINDIOS.	36
13. DISCUSIÓN.....	38
14. CONCLUSIÓN	44
15. REFERENCIAS	45
16. PUBLICACIÓN.....	53
17. ANEXOS	62

1. Índice de Tablas

TABLA 1 CARACTERÍSTICAS CLÍNICAS Y DEMOGRÁFICAS DE LA POBLACIÓN DE ESTUDIO -----	24
TABLA 2 FRAGMENTO DEL CATÁLOGO DE VARIANTES -----	27
TABLA 3 TIPO DE VARIANTES INCLUIDAS EN EL ESTUDIO DE ASOCIACIÓN CASO CONTROL-----	27
TABLA 4 VARIANTES COMUNES ASOCIADAS CON DISLIPIDEMIAS EN LA COHORTE DE DESCUBRIMIENTO ----	28
TABLA 5 ANÁLISIS DE FRECUENCIAS DEL HAPLOTIPO EN HNF1A ENTRE LA POBLACIÓN DE DESCUBRIMIENTO Y OTRAS POBLACIONES -----	33
TABLA 6 RESULTADOS REPRESENTATIVOS DE ASOCIACIÓN POR HAPLOTIPOS -----	33
TABLA 7 RESULTADOS DE LA ASOCIACIÓN DICOTÓMICA DE VARIANTES RARAS EN LA COHORTE DE DESCUBRIMIENTO -----	34
TABLA 8 RESULTADOS DE LA ASOCIACIÓN CUANTITATIVA DE VARIANTES RARAS EN LA COHORTE DE DESCUBRIMIENTO -----	35
TABLA 9 RESULTADOS DE LAS SNV REPLICADAS -----	36
TABLA 10 ANÁLISIS DE FRECUENCIAS DE LAS VARIANTES REPLICADAS ENTRE LA POBLACIÓN DE DESCUBRIMIENTO Y OTRAS POBLACIONES -----	37
TABLA SUPLEMENTARIA 1 LISTA DE GENES CANDIDATO Y POSICIONES EN EL GRCh37	62
TABLA SUPLEMENTARIA 2 CATÁLOGO DE VARIANTES EN LOS GENES CANDIDATO	63
TABLA SUPLEMENTARIA 3 VARIANTES CON ASOCIACIÓN A RASGOS LIPÍDICOS.....	69
TABLA SUPLEMENTARIA 4 ASOCIACIÓN POR HAPLOTIPOS DE GENES CON VARIANTES	72

2. Índice de Figuras

FIGURA 1 PREVALENCIA DE LAS DISLIPIDEMIAS EN LAS ENSANUT. -----	4
FIGURA 2 DISLIPIDEMIAS EN LA ENSANUT POR GÉNERO. -----	5
FIGURA 3 METABOLISMO DE LIPOPROTEÍNAS. -----	10
FIGURA 4 TRANSPORTE REVERSO DE COLESTEROL -----	11
FIGURA 5 DISTRIBUCIÓN GEOGRÁFICA DE LA POBLACIÓN INDÍGENA. -----	14
FIGURA 6 ESTRATEGIA EXPERIMENTAL-----	17
FIGURA 7 DESCRIPCIÓN DE LAS VARIANTES POR FRECUENCIA. -----	25
FIGURA 8 ANOTACIÓN DE LAS VARIANTES -----	26
FIGURA 9 VARIANTES ASOCIADAS A HTG -----	29
FIGURA 10 VARIANTES ASOCIADAS A HTC -----	30
FIGURA 11 VARIANTES ASOCIADAS A HIPO HDL-C -----	30
FIGURA 12 VARIANTES ASOCIADAS A HIPER LDL-C -----	31
FIGURA 13 GENES CON VARIANTES ASOCIADAS A LOS RASGOS LIPÍDICOS-----	31
FIGURA 14 DESEQUILIBRIO DE LIGAMIENTO DE LAS SNVs EN HNF1A Y ASOCIACIÓN CON HIPO HDL-C -	32
FIGURA 15 DESEQUILIBRIO DE LIGAMIENTO PARA LAS SNV PRESENTES EN LOS BLOQUES DE HAPLOTIPOS	34
FIGURA 16 GENES ASOCIADOS A MÚLTIPLES COMPONENTES LIPÍDICOS-----	40
FIGURA 17 DISTRIBUCIÓN GEOGRÁFICA DE LAS VARIANTES REPLICADAS -----	41

3. Lista de abreviaturas

AHA/NHLBI: Asociación Americana del corazón/Instituto Nacional del corazón, pulmón y sangre, por sus siglas en inglés.

apo-: Apolipoproteína.

CAD: Enfermedad Cardiovascular, por sus siglas en inglés.

CPAS: Escaneo alélico cross poblacional, por sus siglas en inglés.

Da: Dáltones

DT2: Diabetes mellitus tipo 2.

Emets: Enfermedades metabólicas.

ENSANUT: Encuesta Nacional de Salud y Nutrición.

FFA: Ácido graso libre, por sus siglas en inglés

GWAS: Estudios de asociación amplios de genoma, por sus siglas en inglés

HDL-C: Lipoproteínas de alta densidad.

Hiper LDL-C: Niveles altos de LDL-C.

HTC: Hipercolesterolemia, por sus siglas en inglés

HTG: Hipertrigliceridemia, por sus siglas en inglés

IDL-C: Lipoproteínas de densidad intermedia, por sus siglas en inglés

IMC: Índice de masa corporal.

IMSS: Instituto Mexicano del Seguro Social

INEGI: Instituto Nacional de Estadística y Geografía.

INMEGEN: Instituto Nacional de Medicina Genómica.

ISSTE: Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado

LDL-C: Lipoproteínas de baja densidad.

Hipo HDL-C: Hipo α -lipoproteinemia.

MAIS: Por sus siglas en inglés de “Análisis Metabólicos en Población Indígena”; Cohorte de replicación

MAF: Frecuencia de alelo menor, por sus siglas en inglés

MEZ: Mestizos mexicanos; Cohorte de descubrimiento

NGS: Secuenciación de nueva generación, por sus siglas en inglés.

OR: Razón de momios, por sus siglas en inglés.

PCR: Reacción en cadena de la polimerasa, por sus siglas en inglés

SIGMA: Iniciativa Slim en Medicina Genómica para las Américas, por sus siglas en inglés

Smet: Síndrome Metabólico.

SNV: Variante de un solo nucleótido, por sus siglas en inglés

TC: Colesterol Total.

TG: Triacilglicéridos.

VEP: Predictor de efecto de variantes, por sus siglas en inglés

4. Glosario

Coefficiente de regresión beta (Beta): En los modelos aditivos de SNVs, se asocia con una variable predictora y representa el efecto de cada alelo menor extra (por ejemplo, un coeficiente de regresión positivo significa que la presencia del alelo menor aumenta el riesgo o la media del fenotipo cuantitativo).

Razón de momios: (OR, por sus siglas en inglés) es la medida de asociación entre una exposición y un resultado. Representa las probabilidades de que ocurra un resultado dado una exposición en particular, en comparación con las probabilidades de que ocurra el resultado en ausencia de dicha exposición.

De manera simplificada, podemos interpretar los resultados de OR de la siguiente forma:

OR=1 La exposición no afecta las probabilidades del resultado.

OR>1 La exposición se asocia con mayores probabilidades del resultado.

OR<1 La exposición se asocia con menores probabilidades del resultado.

Intervalo de confianza 95%: (95% CI, por sus siglas en inglés) es un parámetro empleado para estimar la precisión de un OR. Intervalos grandes indican bajos niveles de precisión del OR, mientras que intervalos pequeños indican mayor precisión. En la práctica, de manera frecuente se utilizan los CI del 95% como un indicador de presencia de significancia estadística si no se superpone al valor nulo (OR=1).

VEP: Predictor de efecto de variantes (Por las siglas en inglés) se trata de un programa bioinformático ampliamente utilizado para predecir los efectos funcionales y las consecuencias de las variantes genéticas, especialmente las variantes de nucleótido único (SNVs) y las variantes de inserción/delección (INDELS), en secuencias de ADN.

5. Resumen

En las últimas décadas, la disminución del riesgo de enfermedades infectocontagiosas, la rápida urbanización y el cambio del ritmo de vida a un esquema sedentario han conducido a que México se encuentre inmerso en un periodo de transición epidemiológico, presentando ahora a las enfermedades no transmisibles como el principal problema de salud pública (Bolio Galvis, 2014; Stevens et al., 2008)

En la actualidad, las enfermedades crónico-degenerativas han emergido como las principales causas de atención hospitalaria y de muerte, generando un impacto socioeconómico muy importante en la población mexicana. Dentro de estos padecimientos, las enfermedades metabólicas (EMets) se han destacado por su creciente prevalencia e incidencia. Este fenómeno conduce, a que por primera ocasión, nuestro país se encuentre en un proceso que compromete la esperanza de vida de futuras generaciones (Rull et al., 2005).

Dentro del conjunto de las EMets, las dislipidemias son uno de los padecimientos más frecuentes y constituyen un factor de riesgo mayor para el desarrollo de la Diabetes Mellitus tipo 2 (DT2) y de enfermedades cardiovasculares.

En este estudio, nuestro objetivo fue identificar la variación genética en los loci relacionados con los lípidos en la población mexicana y su asociación con la dislipidemia. Se utilizaron los datos de secuenciación del exoma de 968 individuos mestizos mexicanos previamente publicados, donde se llevó a cabo la búsqueda de variantes de un solo nucleótido en 177 genes candidatos relacionados con los lípidos y se replicó con una cohorte de 671 amerindios.

Posteriormente, con las variantes identificadas se llevó a cabo un estudio de casos y controles mediante una regresión logística y análisis cuantitativos de rasgos, utilizando el software PLINK.

Nuestros resultados revelaron 21 variantes en 6 genes que mostraron asociación significativa con al menos un rasgo lipídico. Las asociaciones más fuertes se observaron en el grupo de genes *APOA1/C3/A4/A5-ZPR1-BUD13* en el cromosoma 11. Dentro de estos se encontraron diversas variantes asociadas por primera vez con diferentes rasgos lipídicos, como la rs3825041 en *BUD13* asociada con hipertrigliceridemia y la rs11076176 en *CETP* asociada con un mayor riesgo de hipoalfalipoproteinemia.

La importancia de este trabajo reside en la identificación de variantes genéticas tanto nuevas como previamente descritas en otras poblaciones, asociadas al riesgo de desarrollar dislipidemia en la comunidad mexicana, una población genéticamente heterogénea y subrepresentada en los estudios genómicos. El incremento de este tipo de estudios en poblaciones multiétnicas permite profundizar en el conocimiento de la fisiopatología y la etiología de estas entidades, para incidir en su prevención, diagnóstico y tratamiento mediante una medicina de precisión.

6. Abstract

Because of the improvement of the economic status, the reduction in the risk of suffering infectious diseases, the rapid urbanization, and the change towards a sedentary lifestyle, Mexico is going through an epidemiological transition, now facing non-communicable diseases as the main public health challenge (Galvis, Hamdy, Pulido, & Haje, 2014; Stevens et al., 2008).

Nowadays, chronic-degenerative diseases have emerged as the leading causes of hospitalization and death, generating a significant socioeconomic impact on the Mexican population, particularly in terms of public expenditure.

Within these conditions, metabolic diseases (EMets) have stood out due to their increasing prevalence and rapid rise in incidence. This phenomenon marks the first time that, within chronic-degenerative entities in our country, public health systems face a process that jeopardizes the life expectancy of future generations.

Among EMets, dyslipidemias are one of the most frequent disorders, constituting a major risk factor for the development of Type 2 Diabetes Mellitus (DT2) and cardiovascular diseases (CAD).

In this study, our objective was to analyze genetic variation in loci related to lipid traits in the Mexican population and its association with dyslipidemia. We used previously published exome sequencing data from 968 Mexican mestizo individuals for the discovery of single nucleotide variants within 177 candidate lipid-associated genes. We subsequently, replicated our findings in a cohort of 671 Amerindians.

We conducted a case-control study using the identified variants and performed statistical analyses, including logistic regression and quantitative trait analysis, using PLINK software. Our results revealed 21 variants within 6 genes that exhibited significant associations with at least one lipid trait. The strongest associations were observed in the *APOA1/C3/A4/A5-ZPR1-BUD13* gene cluster on chromosome 11.

Within these, several novel variants were found associated with various lipid traits, such as rs3825041 in *BUD13*, associated with hypertriglyceridemia and rs11076176 in *CETP*, associated with an increased risk of hypoalphalipoproteinemia.

The importance of this work lies in the identification of genetic variants, both novel and previously described in other populations, associated with the risk of developing dyslipidemia in a population like Mexico, genetically heterogeneous and underrepresented in genomic studies. The increase in this type of studies in multiethnic populations allows us to deepen our knowledge of the pathophysiology and etiology of these entities, to influence their prevention, diagnosis, and treatment through precision medicine.

7. Introducción

7.1. Enfermedades metabólicas: Dislipidemias

El metabolismo es el proceso por el cual el cuerpo obtiene energía y crea moléculas esenciales a partir de los alimentos que consumimos. Este proceso se lleva a cabo mediante reacciones enzimáticas. Las enfermedades metabólicas (Emets) resultan de defectos genéticos en estas vías bioquímicas, lo que puede provocar una amplia variedad de efectos y síntomas.

Aunque las Emets están ligadas a errores en vías metabólicas, el organismo puede tolerar ciertos errores gracias a su complejo sistema de enzimas y regulación. Sin embargo, la presencia de alteraciones genéticas, capaces de alterar procesos enzimáticos, proteínas reguladoras o mecanismos de transporte puede causar interrupciones críticas en las vías metabólicas, provocando enfermedades graves.

Dentro de las Emets, las dislipidemias son el conjunto de enfermedades caracterizadas por alteraciones en el metabolismo de las lipoproteínas y su subsecuente manifestación clínica es observada en los niveles sanguíneos de lípidos, ya sea del colesterol total (TC), de las lipoproteínas de alta o de baja densidad (HDL-C y LDL-C, respectivamente) o bien de los triglicéridos (TG) (Dennis Kasper, Anthony Fauci, Stephen Hauser, Dan Longo, J. Larry Jameson, 2018)

Actualmente, está bien reconocido que estas entidades son determinantes de riesgo cardiovascular y que también contribuyen de manera importante en el síndrome metabólico (Smet), y otros desordenes metabólicos (Bai et al., 2019). Las dislipidemias más comunes son la hipo-a-lipoproteinemia (Hipo HDL-C, niveles séricos de HDL-C <50mg/dL en mujeres y <40mg/dL en hombre), hipertrigliceridemia (HTG, con valores de triglicéridos >150 mg/dL), hipercolesterolemia (HTC, con niveles de colesterol total >200 mg/dL) y la dislipidemia mixta compuesta por dos o más de estas alteraciones (Alberto Aguilar Salinas et al., 2004)

7.2. Epidemiología

El desarrollo de políticas de salud en México y la implementación de las Encuestas Nacionales de Salud y Nutrición (Ensanut) han permitido tener a disposición de directivos del sector salud, investigadores y académicos un conjunto de datos fidedignos que permiten conocer las condiciones de salud de la población y sus tendencias (Ensanut, 2012).

Como ya se mencionó, las dislipidemias comprenden factores de riesgo importantes para padecer enfermedades cardiovasculares. Desde 1988 se ha descrito a la HTC como un problema de salud pública en México, cuando se llevó a cabo uno de los primeros estudios con representatividad nacional sobre la salud del país. Posteriores determinaciones describieron un incremento en la prevalencia de las dislipidemias, fenómeno que ha sido atribuido también de acuerdo con los resultados nutrimentales, a un incremento en consumo de grasa como porcentaje de calorías totales (Barquera & Campos, 2009).

Algunos de los datos más representativos de las dislipidemias en México son los que comprenden los resultados obtenidos de la Ensanut 2006, donde se consiguió un conjunto de 4040 individuos entre 20 y 69 años con ayuno de 9 a 12 horas que fueron empleados para evaluar las dislipidemias en la población mexicana. En esta encuesta se encontró que la prevalencia de niveles bajos de HDL-C está presente en un 60.5%, niveles elevados de LDL-C en 46%, HTC en 43.6% y HTG en 31.5% (Aguilar-Salinas et al., 2010)(Figura 1a).

Basados en las mismas características descritas por Aguilar-Salinas et al., 2010, a partir de los resultados de la Ensanut 2018, de la base de datos con acceso público se puede obtener una muestra de 6151 individuos con los resultados bioquímicos completos. Con esta información, realizando una estadística descriptiva se puede observar que la prevalencia de estos desordenes metabólicos continúan siendo muy elevada, conservándose como la afección más frecuente los niveles bajos de HDL-C con un 58.2%, seguido por la HTG con 44.16%, HTC con 28.74% y finalmente niveles elevados de LDL-C con 20.83% (Figura 1b).

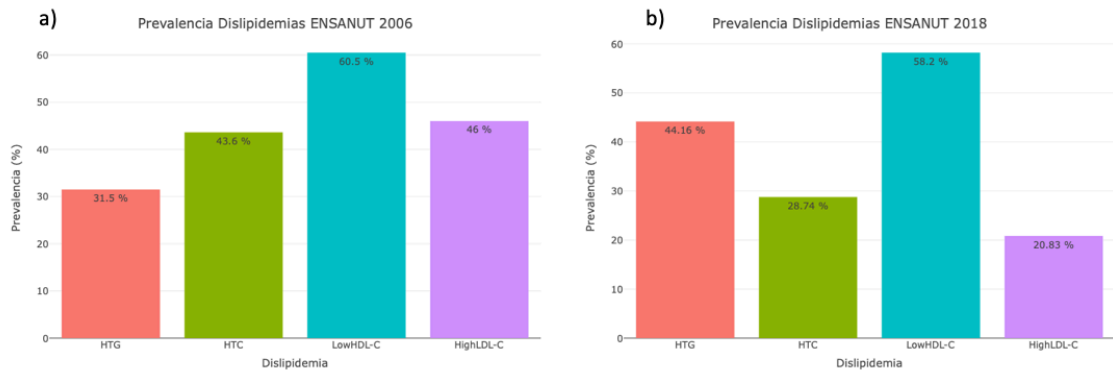


Figura 1 Prevalencia de las dislipidemias en las ENSANUT.

Los datos representados corresponden a ambos géneros y siguiendo los criterios del AHA/ATPIII para la asignación de los fenotipos. HTG: Hipertrigliceridemia; HTC: Hipercolesterolemia; Hipo HDL-C: Hipo α -lipoproteinemia; Hiper LDL-C: Niveles elevados de LDL-C.

Mas aún, comparando las prevalencias de las dislipidemias por género en ambas encuestas de salud, se observa el incremento de la HTG del 2006 respecto al 2018 (Figura 2) para ambos géneros, una disminución para el caso de HTC y de Hiper LDL-C en ambos géneros, sin embargo, lo que resulta alarmante es la proporción de mujeres que en la ENSANUT 2018 presentaron con niveles bajos de HDL-C, ya que se estima que siete de cada 10 participantes presentaban este tipo de dislipidemia Figura 2b, mientras que en los hombres se observó una reducción en la prevalencia de esta entidad patológica

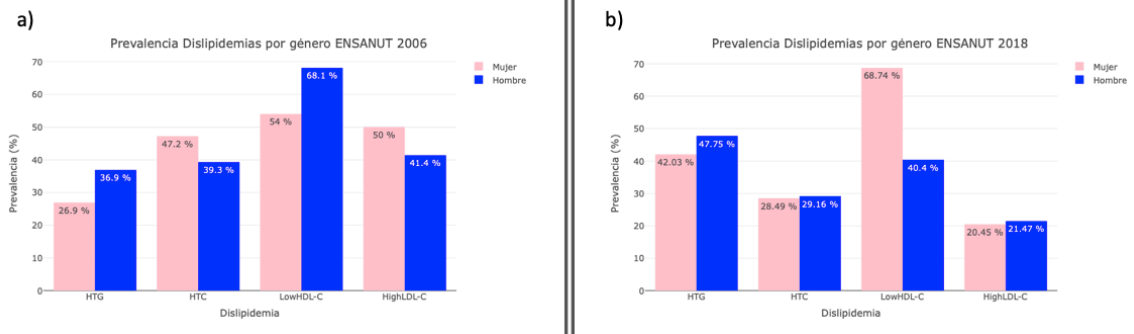


Figura 2 Dislipidemias en la ENSANUT por género.

a) Reporte de la prevalencia de las dislipidemias por género encontradas en la ENSANUT 2016, b) Prevalencia estimada de dislipidemias por género en la ENSANUT 2018, HTG: Hipertrigliceridemia; HTC: Hipercolesterolemia; Hipo HDL-C: Hipo α -lipoproteinemia; Hiper LDL-C: Niveles elevados de LDL-C.

De manera importante, la prevalencia observada de dislipidemias es mayor en México que la reportada en países de origen caucásicos, incluido Estados Unidos de América, donde también se ha reportado que el grupo hispano es el grupo étnico con la prevalencia más alta (Rodríguez et al., 2014).

7.3. Etiología de las dislipidemias

Los niveles circulantes de lípidos y lipoproteínas son determinados por una gran diversidad de factores, tanto ambientales como hereditarios. La mayoría de los individuos que cursan con dislipidemias presentan la combinación de una predisposición genética (usualmente poligénica) y una contribución del ambiente (dieta, estilo de vida, el estado de salud y el uso de algunos fármacos) para la manifestación de las dislipidemias. De esta forma, podemos agrupar a estas enfermedades metabólicas en dos grupos, dislipidemias primarias y secundarias.

Las dislipidemias primarias son debidas a defectos en un único gen o en un conjunto de ellos y generalmente se observan en una variedad de desórdenes metabólicos hereditarios. Por ejemplo, la Hipercolesterolemia Familiar es debida a una mutación en el gen que codifica el receptor de LDL-C. Las dislipidemias secundarias constituyen trastornos metabólicos cuya etiología se relaciona con condiciones médicas subyacentes o influencias externas. Estas condiciones pueden englobar patologías como la diabetes, el hipotiroidismo, trastornos renales, el consumo excesivo de alcohol o la administración de fármacos específicos. Los desequilibrios lipídicos subyacentes pueden, en su mayoría, atribuirse a factores de estilo de vida no saludables y condiciones médicas adquiridas, lo que engloba enfermedades subyacentes y la terapia con fármacos aplicada (Yanai & Yoshida, 2021).

En ausencia de causas secundarias de dislipidemias, las principales causas de estas patologías son debidas a defectos en el metabolismo de lipoproteínas y de manera menos frecuente a disrupción en el tráfico intracelular (Sulaiman, 2019).

Es decir, defectos de alta penetrabilidad en el metabolismo de lípidos están asociados con dislipidemias primarias, mientras que la presencia de variantes de baja penetrabilidad, en conjunción con la dieta y otros factores ambientales, está relacionada con dislipidemias secundarias.

7.4. Metabolismo de lípidos y lipoproteínas

Los lípidos son un conjunto diverso de biomoléculas que en común comparten su insolubilidad en agua. Sus funciones biológicas son tan variadas como lo es su estructura química. Las grasas y los aceites constituyen una de las principales fuentes de reserva energética en muchos organismos; los fosfolípidos y los esteroides son los principales componentes de elementos estructurales y de las membranas biológicas; mientras que una pequeña porción de estos compuestos actúa como moléculas de señalización, cofactores y pigmentos (Vinet & Zhedanov, 2011). El metabolismo de los lípidos involucra ciertas enzimas, subtipos de fracciones lipídicas y lipoproteínas. Las concentraciones de los distintos componentes lipídicos en plasma se encuentran mediadas, tanto por su ingesta a partir de la dieta como por su síntesis *de novo* (Figura 3), así como por su eliminación, consumo intracelular y secreción (Dennis Kasper, Anthony Fauci, Stephen Hauser, Dan Longo, J. Larry Jameson, 2018). Dentro del metabolismo del ser humano, el colesterol participa en la integridad de la membrana celular y como precursor de hormonas esteroideas, vitaminas y sales biliares; mientras que los TG son destinados como fuente de energía (tejido muscular) o de almacenamiento (tejido adiposo).

Con excepción de los FFA que pueden ser transportados en el plasma unidos a la albúmina, los lípidos circulantes en el organismo son transportados en complejos multiprotéicos y pseudo-micelares denominados lipoproteínas (Segrest et al., 1994).

Uno de los sistemas más empleados para clasificar a las lipoproteínas, parte de las densidades estimadas para estas partículas mediante ultra centrifugación y su posterior estudio en el laboratorio para determinar su composición química (Mahley et al., 1984), como se muestra a continuación:

- Quilomicrones (QM): Este grupo de lipoproteína tiene una densidad inferior a 1.006 g/mL. Están compuestos principalmente por triglicéridos (85%), mientras que las proteínas, los fosfolípidos, los ésteres de colesterol y el colesterol esterificado representan cantidades más bajas.
- Lipoproteínas de muy baja densidad (VLDL-C): Estas lipoproteínas tienen una densidad en el rango de 0.95 a 1.006 g/mL. Están compuestas por triglicéridos (50%), proteínas, fosfolípidos, ésteres de colesterol y colesterol esterificado en proporciones variables.
- Lipoproteína de densidad intermedia (IDL-C): Estas lipoproteínas tienen una densidad en el rango de 1.006 a 1.019 g/mL. Están compuestas por proteínas (20%), triglicéridos, fosfolípidos, ésteres de colesterol y colesterol esterificado en diferentes proporciones.

- Lipoproteínas de baja densidad (LDL-C): Estas lipoproteínas tienen una densidad en el rango de 1.019 a 1.063 g/mL. Está compuesta principalmente por proteínas (23%), ésteres de colesterol, fosfolípidos, triglicéridos y colesterol esterificado en diferentes proporciones.
- Lipoproteínas de alta densidad (HDL-C): Estas lipoproteínas tienen una densidad en el rango de 1.063 a 1.210 g/mL. Están compuesta principalmente por proteínas (55%), fosfolípidos, ésteres de colesterol, triglicéridos y colesterol esterificado en diferentes proporciones.

En esta clasificación, se muestra claramente las diferencias en la composición entre las lipoproteínas de diferentes densidades. Por ejemplo, se observa que las lipoproteínas de baja densidad (LDL-C) tienen un mayor porcentaje de proteínas y ésteres de colesterol en comparación con otras lipoproteínas, mientras que las lipoproteínas de alta densidad (HDL-C) tienen una proporción más alta de proteínas y fosfolípidos.

Las proteínas que forman parte estructural de las lipoproteínas reciben el nombre de apolipoproteínas (“apo-”). Estas proteínas son de naturaleza anfipática, es decir, poseen tanto un dominio hidrofóbico como uno hidrofílico, por lo cual pueden interactuar con los lípidos en las lipoproteínas y el medio acuoso que representa el plasma sanguíneo (Segrest et al., 1994). Además de su papel como vehículo lipídico, favorecido por la estructura de hélice alfa de estas proteínas, las apo- cumplen con mecanismos de regulación críticos en el metabolismo de los lípidos mediante la interacción con múltiples receptores y enzimas (Ramasamy, 2014).

Una descripción más detallada de las apo-, las lipoproteínas donde se encuentran, así como las funciones metabólicas particulares de cada una de ellas se presenta a continuación.

- ApoA1: Tiene un peso molecular de 28016 Da y se encuentra en las lipoproteínas HDL-C y QM. Funciona como componente estructural de HDL-C y también como activador de *LCAT*. Se sintetiza en el hígado y el intestino.
- ApoAII: Tiene un peso molecular de 17414 Da y se encuentra en las lipoproteínas HDL-C y QM. Actúa como proteína estructural de HDL-C y se sintetiza en el hígado.
- ApoAIV: Tiene un peso molecular de 46465 Da y se encuentra en las lipoproteínas HDL-C y QM. Participa en el ensamble y secreción de QM y se sintetiza en el intestino humano.
- ApoAV: Tiene un peso molecular de 39000 Da y se encuentra en las lipoproteínas HDL-C, VLDL-C y QM. Su función en la concentración de TG es compleja y variable. Actúa como activador intravascular de *LPL* y modula el metabolismo hepático de TG. Se sintetiza predominantemente en el hígado.
- ApoB48: Tiene un peso molecular de 264000 Da y se encuentra en las lipoproteínas QM. Es necesaria para el ensamble y secreción de QM en el intestino delgado. Su síntesis ocurre en el intestino.
- ApoB100: Tiene un peso molecular de 540000 Da y se encuentra en las lipoproteínas VLDL-C, IDL-C y LDL-C. Es necesaria para el ensamble y secreción de VLDL-C en el hígado. Actúa como proteína estructural de VLDL-C, IDL-C y LDL-C, y también como ligando del receptor de LDL-C. Se sintetiza en el hígado.

- ApoCI: Tiene un peso molecular de 6630 Da y se encuentra en las lipoproteínas QM, VLDL-C, IDL-C y HDL-C. Inhibe la unión de lipoproteínas con sus receptores y es un potente inhibidor de la *CETP*. Se sintetiza en el hígado.
- ApoCII: Tiene un peso molecular de 8900 Da y se encuentra en las lipoproteínas QM, VLDL-C, IDL-C y HDL-C. Actúa como activador de *LPL* y se sintetiza en el hígado.
- ApoCIII: Tiene un peso molecular de 8800 Da y se encuentra en las lipoproteínas QM, VLDL-C, IDL-C y HDL-C. Funciona como inhibidor de *LPL*, incrementa la secreción de VLDL-C y está involucrada en procesos relacionados con la aterogénesis e inflamación vascular. Su síntesis ocurre principalmente en el hígado y en menor medida en el intestino.
- ApoE: Tiene un peso molecular de 34145 Da y se encuentra en las lipoproteínas QM, VLDL-C, IDL-C y HDL-C. Actúa como receptor de LDL-C y ligando para remanentes de LDL-C y QM. También participa en el transporte reverso de colesterol. Se sintetiza predominantemente en el hígado.
- Apo(a): Tiene un peso molecular que varía de 250000 a 800000 Da y se encuentra en las lipoproteínas Lp(a). Su función aún es desconocida y se sintetiza en el hígado.
- ApoM: Tiene un peso molecular de 26000 Da y se encuentra principalmente en las lipoproteínas HDL-C, pero también se encuentra en CM, VLDL-C y LDL-C. Está involucrada en la síntesis de HDL-C y se sintetiza en el hígado y el riñón.

De acuerdo con el contenido de apo-, las lipoproteínas pueden clasificarse también en dos grupos: aquellas que contienen apoAI (HDL-C) y las que contienen apoB (QM, VLDL-C, IDL-C y LDL-C). A su vez, existen dos grupos de lipoproteínas que contienen apoB, las que presentan apoB48 y aquellas con apoB100. En el humano, apoB48 es secretada exclusivamente en el intestino y se encuentra presente en los QM; apoB100 por su parte es secretada exclusivamente en el hígado dentro de las VLDL-C. La apoB100 contiene 4536 residuos, en tanto que apoB48 presenta una identidad del 48% con el extremo N-terminal de apoB100 y es el producto de un proceso de edición del mRNA mediado por *APOBEC1* (Innerarity et al., 1987).

7.5. Vías metabólicas de los lípidos

El metabolismo normal de los lípidos está comprendido por tres rutas principales: la del componente de la dieta también conocida como vía exógena, el componente endógeno o vía endógena y el transporte reverso del colesterol (TRC).

7.5.1. Vía exógena

La degradación de la mayoría de los lípidos provenientes de la dieta ocurre en el epitelio intestinal, siendo convertidos en ácidos grasos libres (FFA) y monoacilgliceroles (MAG) por la acción de la lipasa pancreática (Figura 3). Los FFA son transportados a través de la membrana apical de los enterocitos mediante modelos de difusión pasiva o dependiente de proteínas, mientras que el colesterol es facilitado por las micelas de sales biliares y traspasado por el transportador Niemann-Pick C1-like 1 (*NPC1L1*) a través de la membrana de borde ciliado (Davis & Altmann, 2009). El exceso del colesterol de la dieta es regresado al lumen

intestinal por medio de los transportadores *ABCG5/ABCG8* (Vrins, 2010). Una vez dentro de los enterocitos, los FFA y los MAG son reensamblados en nuevos TGs por acción de la acil-CoA: colesterol aciltransferasa (*ACAT*) (Helgerud et al., 1982). Los ésteres de colesterol y TGs provenientes de la ingesta, son empaquetados junto con fosfolípidos y apolipoproteínas para formar quilomicrones (QM) por acción de la proteína microsomal de transferencia de TGs (*MTP*), todos ellos con una molécula de apoB48. Estos QM son secretados en el sistema linfático intestinal desde dónde entran a la circulación sanguínea a través del conducto torácico. Ya en circulación los QM nacientes son transportados al resto del cuerpo. En los capilares de los diferentes tejidos donde son transportados, la enzima Lipoproteína Lipasa (*LPL*), activada por la proteína apoCII presente en los QM, hidroliza los TGs a FFA y glicerol, los cuales son absorbidos por las células de los tejidos blanco, donde realizan funciones diversas (Ramasamy, 2014). En el músculo, las partículas lipídicas son oxidadas para obtener energía; mientras que en el tejido adiposo se re-esterifican para su almacenamiento como TG. Los remanentes de QM, casi sin la presencia de TG en ellos, contienen aún colesterol y las apo-. Estos agregados viajan al hígado, donde son endocitados mediante los receptores específicos de las apolipoproteínas, por medio de muchas vías que incluyen apoE, lipasa hepática y el *LDLR* (Mahley et al., 1984).

7.5.2. Vía endógena

Otra de las funciones claves de las lipoproteínas es el transporte de los lípidos hepáticos a los tejidos periféricos como una fuente de energía durante un periodo de ayuno (Figura 3). En estas condiciones, la lipólisis de los TGs en el tejido adiposo genera los FFA que son transportados al hígado, y también el hígado es capaz de sintetizar ácidos grasos a través de la lipogénesis *de novo*. Estos lípidos son esterificados en las VLDL-C en conjunto con apoB100, apoCI, apoCII, apoCIII, apoE, ésteres de colesterol, fosfolípidos y vitamina E; al igual que los QM, este proceso es dependiente de *MTP* (Dennis Kasper, Anthony Fauci, Stephen Hauser, Dan Longo, J. Larry Jameson, 2018). En la circulación sanguínea las VLDL-C sufren una hidrólisis similar a la de los QM por la acción de la *LPL* presente en los capilares de los tejidos periféricos y en el proceso son convertidas en IDL-C y finalmente en LDL-C. Las LDL-C son entonces llevadas a los diferentes tejidos extrahepáticos como músculo, glándulas suprarrenales y tejido adiposo, donde son captadas mediante el *LDLR* que reconoce apoB100. Las LDL-C que no son captadas por los tejidos periféricos regresan al hígado donde el colesterol contenido en ellas puede ser sustrato para la producción de nuevas membranas, convertido en sales biliares o ser re-esterificado y almacenado en el citosol en gotas lipídicas (Vinet & Zhedanov, 2011). El *LDLR* es reciclado y transportado de regreso a la superficie de la célula, mecanismo que es regulado por la proteína convertasa subtilisina/kexina tipo 9 (*PCSK9*). Durante este proceso, es sometido a degradación lisosomal, cuando *PCSK9* se une al receptor LDL-C, pero es reciclado de regreso a la superficie de los hepatocitos en la ausencia de *PCSK9*. La neutralización de *PCSK9* incrementa el total de la capacidad de unión de los hepatocitos, lo que conlleva la reducción en la concentración de LDL-C (Nassoury et al., 2007).

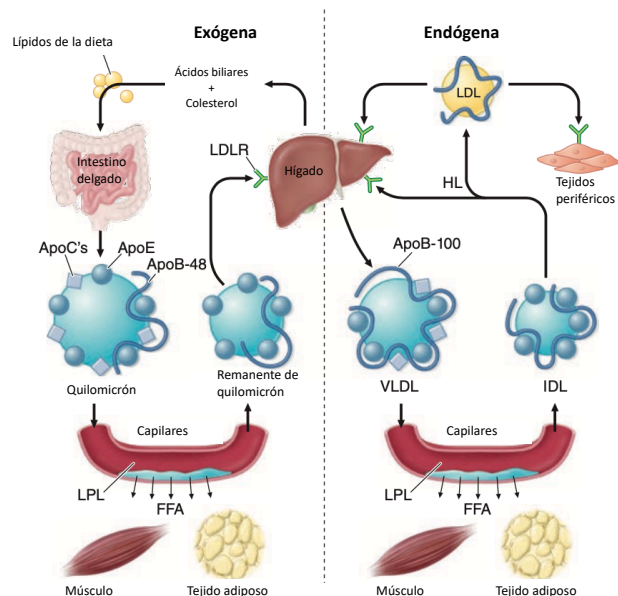


Figura 3 Metabolismo de lipoproteínas.

La vía exógena transporta los lípidos de la dieta a la periferia y al hígado. La vía endógena transporta lípidos hepáticos a la periferia. FFA: Ácido graso libre; HL: Lipasa hepática; IDL-C: Lipoproteínas de densidad intermedia; LDL-C: lipoproteínas de baja densidad; LDLR: receptor de LDL-C; LPL: Lipoproteín lipasa; VLDL-C: lipoproteína de muy baja densidad. Modificado de Harrison's Principles of Internal Medicine.

7.5.3. Transporte reverso de colesterol.

La síntesis de colesterol puede ocurrir en la mayoría de los tejidos, pero solo los enterocitos y hepatocitos tienen la capacidad de eliminar el exceso de colesterol de sus células, ya sea hacia el intestino o la vesícula biliar, respectivamente (Dennis Kasper, Anthony Fauci, Stephen Hauser, Dan Longo, J. Larry Jameson, 2018). Para transportar este exceso de colesterol desde las células de los tejidos periféricos de vuelta al hígado y prevenir la formación de placas de ateroma, las HDL-C desempeñan un papel crucial en lo que se conoce como el "transporte inverso del colesterol" (Figura 4). Este proceso se inicia con la síntesis de HDL-C recién formadas por hepatocitos y enterocitos, principalmente a través del flujo de apoAI y la lecitina-colesterol aciltransferasa (*LCAT*), que capturan colesterol y fosfolípidos para ensamblar partículas de HDL-C (Oram & Yokoyama, 1996). Dentro de la circulación, las HDL-C son enriquecidas con ésteres de colesterol por la acción de la proteína de transferencia de ésteres de colesterol (*CETP*) y la proteína transferidora de fosfolípidos (*PLTP*); *CETP* además regula la biosíntesis del colesterol en el hígado reduciendo la transcripción de la Hidroximetilglutaril-CoA reductasa (*HMGCR*), considerada como el paso limitante en la síntesis de nuevo colesterol endógeno (Jiang & Zhou, 2006). La participación de la proteína de unión de casete ATP A1 (*ABCA1*) resulta crucial para la transferencia del colesterol de las células de los tejidos periféricos (incluyendo los macrófagos en paredes arteriales) y la formación de las HDL-C maduras (Oram & Lawn, 2001). Posteriormente, la proteína de unión de casete ATP G1 (*ABCG1*) y el receptor "scavenger" clase BI (*SR-BI*) contribuyen enriqueciendo en colesterol a las HDL-C, en el caso de este último de manera bidireccional transfiriendo también colesterol desde las LDL-C (Stangl et al., 1999). En el hígado, la lipasa hepática hidroliza a los ésteres de colesterol de las HDL-C generando FFA

y colesterol. El colesterol entonces puede ser excretado vía biliar y las HDL-C resultantes pueden contribuir de nuevo con el transporte reverso del colesterol (Nghiem-Rao et al., 2014).

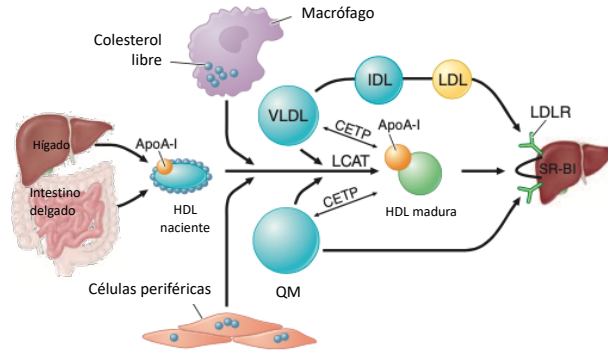


Figura 4 Transporte reverso de colesterol

CETP: Proteína transferidora de ésteres de colesterol, LCAT: Lecitin-colesterol aciltransferasa, SR-BI: Receptor "scavenger" clase B tipo I, LDLR: Receptor de LDL-C, HDL: Lipoproteínas de alta densidad, IDL: Lipoproteínas de densidad intermedia, LDL-C: Lipoproteínas de baja densidad, QM: Quilomicrones, VLDL-C: Lipoproteínas de muy baja densidad. Modificado de Harrison's Principles of Internal Medicine.

7.6. Genética de las dislipidemias

La heredabilidad de los niveles en plasma de los lípidos de acuerdo con estudios familiares se estima es de cerca del 32% (Herbeth et al., 2010) y puede incluso llegar a ser del 60% (C. Gao et al., 2018). Cabe mencionar que su prevalencia varía entre las poblaciones y los grupos étnicos. Por ejemplo, los trastornos lipídicos son altamente prevalentes en poblaciones con ascendencia amerindia (hasta un 85.9%), en comparación con la prevalencia entre individuos de otras ascendencias (por ejemplo, caucásica 31.2% y africana 41.1%) (Aguilar-Salinas et al., 2014; Noubiap et al., 2018).

Dentro de los factores genéticos asociados con el riesgo a desarrollar dislipidemias se encuentran la presencia de variantes en genes que codifican para enzimas claves en el metabolismo de los lípidos, su transporte intracelular y diversos factores de transcripción (Estrada et al., 2014). Dependiendo del impacto biológico que tengan estas variantes genéticas, su presencia se relaciona con dislipidemias monogénicas, también denominadas familiares, o con una mayor susceptibilidad a dislipidemias secundarias.

7.6.1. Dislipidemias monogénicas

Son muchas las dislipidemias de carácter monogénico que se han descrito (Sulaiman, 2019), por ejemplo, la HTG familiar se debe usualmente a variantes de pérdida de función en *LDLR*, el gen que codifica para el receptor de LDL-C, así como variantes en otros genes codificantes para proteínas que interactúan con este receptor entre los que se encuentran: *APOB*, *PCSK9* y *LDLRAP1*. La HTG monogénica es resultado de variantes en genes que regulan el metabolismo de las lipoproteínas ricas en triacilglicéridos como son: *LPL*, *APOC2*, *APOA5*, *LMF1* y *GPIHBP*. De manera similar, los niveles bajos de HDL-C pueden ser resultado de variantes que regulan el catabolismo o la producción de este componente: *APOA1*, la Lecitin:

colesterol aciltransferasa (*LCAT*) y el transportador de casete dependiente de ATP A1 (*ABCA1*) (Ramasamy, 2016)

7.7. Dislipidemias como enfermedades complejas

Recientemente, el desarrollo de estudios amplios del genoma (GWAS) y de secuenciación de nueva generación (NGS) ha señalado la importancia que presentan las variantes comunes en genes asociados con el metabolismo lipídico en el riesgo a desarrollar dislipidemias. Por lo general, las formas más comunes de dislipidemias como se indicó previamente son de origen multifactorial, donde los componentes genéticos, ambientales y la interacción entre ellos contribuyen a su patogenia (Khor & Goh, 2010).

Por ejemplo, un estudio del Consorcio Mundial de Genética de Lípidos (GCLG, por sus siglas en inglés), donde se identificaron 157 loci asociados con los niveles de lípidos a un valor de $P < 5 \times 10^{-8}$, incluyendo 62 loci que no habían sido previamente relacionados con los niveles de lípidos en seres humanos. En este trabajo, además, utilizando una genotipificación densa en individuos de ascendencia europea, asiática oriental, asiática meridional y africana, consiguieron reducir las señales de asociación en 12 loci (Willer et al., 2013).

Sin embargo, estos estudios se han centrado principalmente en poblaciones de origen europeo y asiático y no es claro como el conocimiento generado en estos estudios pueda ser trasladado a otras poblaciones como la nuestra.

7.8. Estudios en población mexicana

Existen pocos estudios en población mexicana sobre la relación existente entre las variantes génicas y la susceptibilidad genética al desarrollo de dislipidemias, sin embargo proporcionan datos que sugieren la existencia de variantes exclusivas de la población mexicana.

Uno de los más representativos fue la identificación de la variante R230C en el gen *ABCA1*, la cual ha mostrado asociación a niveles bajos de HDL-C, obesidad, DT2 y síndrome metabólico; se sabe que esta variante juega un papel muy importante en el transporte reverso de colesterol. Mas aún, el alelo menor de esta variante, que codifica para una cisteína parece ser exclusiva de poblaciones amerindias al no encontrarse en poblaciones europeas, asiáticas y africanas y presentar un rango de frecuencia de 0.18-0.28 en diversos grupos Amerindios como, Mayas, Purépechas, Yaquis y Teenek; mientras que en mestizos residentes en el centro de México presento una frecuencia 0.11 (Acuña-Alonzo et al., 2010).

El primer GWAS para dislipidemias en individuos mexicanos fue desarrollado en el 2011. En la fase inicial de este estudio se incluyeron 2240 muestras, genotipando 563,599 SNVs e imputando otros 769,042 SNVs. En la segunda fase se incluyeron 2,121 individuos y 1,235 SNVs. Las asociaciones más significativas se identificaron en los genes *APOA5*, *GCKR*, *LPL* y *NCPI* seguidos de *ANGPTL3*, *TIMD4*, *MLXIPL* y *CILP2*. La frecuencia de los alelos de susceptibilidad de *APOA5*, *ANGPTL3*, *NCPI* y *TIMD4* resultaron más altas en mexicanos que en las reportadas en otras poblaciones (Weissglas-Volkov et al., 2013). Posteriormente,

usando una estrategia denominada CPAS (por sus siglas en inglés de *Cross Population Allele Scanning*) empleando los mismos datos de genotipificación del GWAS de 2011, se mostró la asociación de los genes *RORA* y *SIK3* con niveles altos de triglicéridos en población mexicana. Estudios funcionales mostraron que después de un reto con una comida rica en grasa los portadores del alelo de riesgo en *SIK3* presentaron niveles más altos de triglicéridos en comparación a los individuos que no son portadores de este alelo, hallazgo que sugiere que la arquitectura genética amerindia conduce a una incidencia mayor de dislipidemias y obesidad en mexicanos (Ko et al., 2014).

Otro estudio en donde se usaron dos cohortes de mexicanos de la Ciudad de México comprendiendo un total 3097 individuos y usando una estrategia llamada *admixture mapping*, en 2017 identificó la asociación de una región genómica con importante ancestría nativa y que incluye los genes *BUD13/ZNF259/APOA5* con altas concentraciones de TG (Parra et al., 2017). Más recientemente en 2018, Gao y colaboradores llevaron a cabo un estudio empleando los datos de exoma de una cohorte de familias Mexicoamericana, donde los hallazgos más significativos recayeron en la asociación de regiones en *APOA5*, *TOMM40* y *GLRB/C4orf45* (C. Gao et al., 2018).

En suma, la evidencia derivada de estudios de asociación genética demuestra que existen variantes específicas de la población mestiza mexicana asociadas con la HTG y la Hipo HDL-C. Éste fenómeno fue señalado también en el proyecto SIGMA (siglas del inglés para *Slim Initiative in Genomic Medicine for the Americas*) para la DT2, donde se reportó un haplotipo de riesgo cuya frecuencia ronda alrededor del 50% en amerindios mexicanos, 12% en grupos asiáticos, <2% en europeos y ausente en africanos (Williams Amy et al., 2014a); o bien la identificó en el gen *HNF1A*, que codifica para el factor nuclear de hepatocitos 1 α y que es responsable de la diabetes tipo MODY3, a la variante p.E580K también asociada con riesgo a DT2, este es un SNV no sinónimo que muestra un riesgo relativo de 5.48 para el desarrollo esta enfermedad metabólica (Estrada et al., 2014).

Estas observaciones son acordes con el intenso proceso de selección que ha sufrido la población amerindia, por lo que se postula que dicho fenómeno puede ser resultado de un efecto de adaptación para la supervivencia. Esto es, que las variantes se conservaron en nuestra población porque confirieron alguna ventaja en el pasado (es decir, mantener el colesterol en el interior de las células durante tiempos de hambruna) y que este “genotipo ahorrador” ahora expuesto a un estilo de vida moderno, sin las mismas condiciones adversas y un aporte excesivo de nutrimentos resulta en el efecto opuesto al que dio origen a la fijación de la variante (NEEL, 1962).

Considerando que las variantes raras o exclusivas de algunas poblaciones son cruciales en la etiopatogenia de enfermedades consideradas como un problema de salud pública es indispensable la caracterización genómica específica de cada población.

7.9. Población mexicana

La mayoría de la población mexicana está formada por mestizos, que son un grupo resultado de la mezcla entre los habitantes originarios de América que vivían en México, los

8. Justificación

Las dislipidemias son un problema de salud pública creciente que aumentan el riesgo de desarrollar enfermedades cardiovasculares y cerebrovasculares. El diagnóstico y tratamiento tempranos son importantes para prevenir el desarrollo de estas enfermedades metabólicas y sus comorbilidades (H. Gao et al., 2021).

La existencia de marcadas diferencias genéticas entre poblaciones con distinto origen étnico ha demostrado que el número y la importancia relativa de algunos alelos asociados con enfermedades crónico-degenerativas, entre las que se incluyen las dislipidemias, puede variar entre los diversos grupos étnicos y que estas diferencias, aunadas con el ambiente, pueden ser las determinantes en las variaciones de la prevalencia y morbi-mortalidad de estos padecimientos alrededor del mundo (Martin et al., 2017).

La mayor parte de los estudios genómicos relacionados con desordenes lipídicos han sido realizados en poblaciones de origen europeo, y esto, aunado a que la gran mayoría de las plataformas de microarreglos utilizadas para el análisis masivo de estas enfermedades, han sido diseñadas con base en lo observado en poblaciones de origen caucásico, abre la posibilidad de que muchos de los factores genéticos responsables de estas enfermedades en la población mexicana aún no se hayan identificado, principalmente aquellos aportados por las poblaciones amerindias (Moreno-Estrada et al., 2014; Sirugo et al., 2019). Considerando que las variantes raras o exclusivas de algunas poblaciones son cruciales en la etiopatogenia de diversas enfermedades que constituyen un importante problema de salud pública, es indispensable la caracterización genómica específica de cada población. En términos de aprovechamiento de recursos, la estrategia ideal podría ser el empleo de datos generados previamente a través de estudios de secuenciación masiva, a partir de los cuales se pueda conocer la totalidad de las variantes presentes en genes candidato. Esto, sin duda, permitiría profundizar en el conocimiento del defecto básico de las dislipidemias y mejorar las estrategias para su manejo, prevención y tratamiento en la población mexicana.

9. Hipótesis

La presencia de variantes génicas propias de los pueblos indígenas originales en genes candidato es un factor de riesgo para el desarrollo de dislipidemias en la población mexicana moderna.

10. Objetivos

10.1. OBJETIVO GENERAL:

Identificar las variantes comunes y propias de la población mexicana asociadas con las dislipidemias.

10.2. OBJETIVOS ESPECÍFICOS:

- 1) Identificar en poblaciones amerindias, variantes en al menos 50 genes candidato asociados con las dislipidemias.
- 2) Determinar la asociación de las variantes nuevas con las dislipidemias mixtas en poblaciones amerindias y mestiza mexicana.
- 3) Evaluar a través de análisis *in silico* el posible efecto de las variantes asociadas con las dislipidemias sobre la expresión del mRNA o la función de la proteína.

11. Material y Métodos

11.1. Estrategia general

El estudio presentado es de carácter observacional en el que se comparan dos grupos: un grupo de personas con una enfermedad o condición (los casos) y un grupo de personas sin la enfermedad o condición (los controles). El estudio busca diferencias entre los dos grupos en términos de exposición a un factor particular, en nuestro caso una variante genética. En la Figura 6 se detalla la estrategia que se siguió.

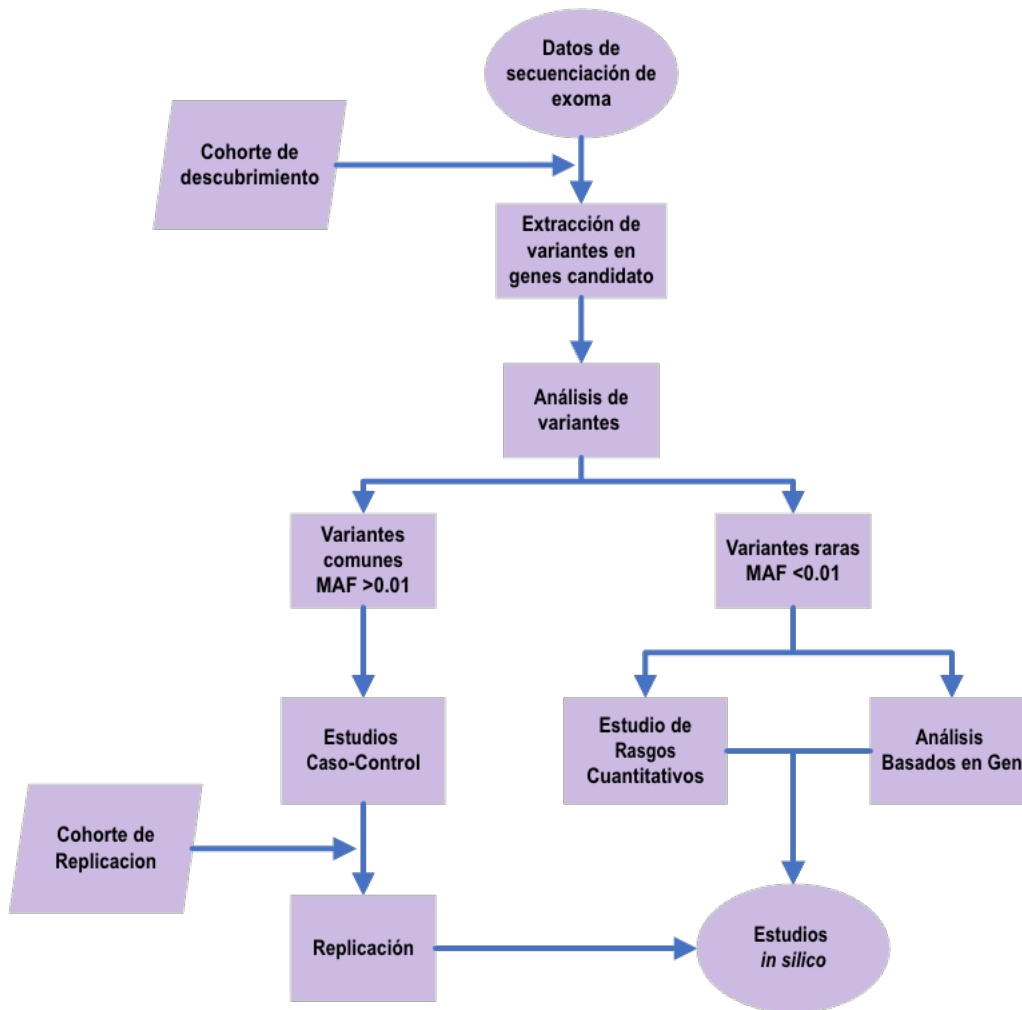


Figura 6 Estrategia experimental
MAF: Frecuencia del alelo menor

11.2. Población de estudio.

Esta investigación fue aprobada por los comités locales de Investigación, Ética y Bioseguridad del Instituto Nacional de Medicina Genómica y se llevó a cabo siguiendo los principios de la Declaración de Helsinki.

Para la realización de este estudio se utilizó una base de datos de exoma, que se obtuvieron de un total de 968 mestizos mexicanos captados en instituciones de tercer nivel (IMSS e ISSSTE) en el centro del país y que forma parte de la cohorte “Estudio de Diabetes en México” (DMS, por sus siglas en inglés). Esta muestra poblacional constituyó nuestra cohorte de descubrimiento y los datos fueron publicados por Estrada y cols. (Estrada et al., 2014)

Una segunda población conformada por 671 indígenas mexicanos captados directamente en sus comunidades de origen como parte de la Cohorte “Análisis Metabólicos en Población Indígena” (MAIS, por sus siglas en inglés), fue incluida para la replicación de los resultados.

La totalidad de las muestras utilizadas forman parte del biobanco del Laboratorio de Inmunogenómica y Enfermedades Metabólicas (LIyEM) del INMEGEN. De cada participante se cuenta con datos demográficos (edad, sexo, etnicidad y lugar de nacimiento), antropométricos (peso, talla, porcentaje de músculo y grasa) y bioquímicos (triglicéridos, colesterol total, HDL-C, glucosa y hemoglobina glicosilada), así como historia clínica. Los participantes firmaron una carta de consentimiento informado.

Para la captación de ambas poblaciones se cumplió con los siguientes criterios:

En el macro estudio, se reclutaron individuos procedentes de dos instituciones de nivel terciario (IMSS e ISSSTE) ubicadas en la Ciudad de México durante 2010-2011. El diagnóstico de la diabetes tipo 2 (T2D) se realizó siguiendo los criterios establecidos por la Asociación Americana de Diabetes (ADA). Se clasificaron como controles a 811 sujetos sanos no relacionados, mayores de 45 años, con niveles de glucosa en ayunas por debajo de 100 mg/dL. Por otro lado, se incluyeron como casos de T2D a 569 individuos no relacionados, mayores de 18 años, que presentaban diagnóstico previo de T2D o niveles de glucosa en ayunas superiores a 125 mg/dL. Se excluyeron a las personas con niveles de glucosa en ayunas en el rango de 100-125 mg/dL. Todos los participantes otorgaron su consentimiento informado, y el estudio se llevó a cabo con la aprobación de los Comités de Ética e Investigación de todas las instituciones involucradas.

11.2.1. Criterios de inclusión:

Individuos mayores de 30 años, no relacionados, sin antecedentes extranjeros reconocidos (padres y abuelos nacidos en México) y que cuenten con historia clínica completa con datos bioquímicos, clínicos y antropométricos.

- Cohorte MAIS:

Individuos nacidos en la comunidad indígena, que se auto reconozcan como indígenas y que sus padres y abuelos hablen la lengua indígena.

11.2.2. Criterios de eliminación:

Muestra biológica insuficiente o de calidad deficiente para todas las pruebas y falta de consentimiento informado. Que curse con algún padecimiento hormonal, enfermedades concomitantes, dislipidemias familiares diagnosticadas o bien, que el participante decida retirarse del estudio.

11.3. Extracción de variantes en genes candidato

Para la realización de este estudio se seleccionaron 177 genes previamente asociados a dislipidemias (Tabla suplementaria 1). De estos, 175 loci se seleccionaron de acuerdo con resultados previos de GWAS y metaanálisis, en particular los presentados por el GLGC (X. Lu et al., 2017; Willer et al., 2013); además de estos se incluyeron los genes *SLC16A11* y *HNF1A*, identificados en estudios previos como marcadores de susceptibilidad para el desarrollo de DT2 y cuya función potencial está involucrada en el transporte lipídico (Estrada et al., 2014; Williams Amy et al., 2014a). A partir de los datos de exoma como parte del proyecto SIGMA publicados por Estrada y cols. (Estrada et al., 2014) se extrajeron la totalidad de las variantes bialélicas presentes en los genes candidato en un total de 968 muestras con un componente amerindio evaluado por microarreglos, mayor a 60% y menor al 90%. Este grupo de muestras constituyó nuestra cohorte de descubrimiento.

Las ubicaciones físicas de los loci fueron mapeadas de acuerdo con el ensamblaje genómico GRCh37 del genoma de referencia, accesible en <http://grch37.ensembl.org/index.html>, En este proceso, se consideraron los transcritos canónicos o aquellos con una mayor cobertura, y se registraron con precisión las coordenadas de inicio y término de cada transcrito. El software GATK (Broad) fue utilizado para el llamado de variantes.

11.3.1. Archivos de genotipos

A partir de los archivos de llamado de variantes (VCF files, por sus siglas en inglés), de los cuales los controles de calidad, así como la profundidad de lectura, cobertura media y las tasas de datos faltantes se pueden consultar en Estrada et al., 2014, se realizó la extracción

de las regiones comprendidas por los genes candidato antes mencionados, usando línea de comando en Linux para generar un set de datos con las SNV bialélicas. Posteriormente se realizó la anotación de los SNV empleando la herramienta *Variant Effect Predictor* (Ensembl) disponible en <https://www.ensembl.org/info/docs/tools/vep/index.html>; versión 87) (McLaren et al., 2016).

11.3.2. Archivos de fenotipos

Como se mencionó anteriormente, el LlyEM cuenta con las bases de datos clínicas y antropométricas de los individuos voluntarios reclutados participantes. Los niveles de TG, CT y HDL-C se midieron a partir de muestras de sangre recolectadas después de un ayuno durante la noche, utilizando un Sistema de Analizador Synchron CX5 (Beckman Coulter, Fullerton, CA, Estados Unidos). Los valores de colesterol LDL (LDL-C) se calcularon mediante la fórmula de Friedewald, excluyendo las muestras con niveles de TG superiores a 400 mg/dl (Warnick et al., 1990). De la base se extrajeron los datos correspondientes a las cohortes de trabajo y se construyeron los fenotipos dicotómicos para dislipidemias siguiendo los criterios de la AHA-NHLBI (<http://www.nhlbi.nih.gov>), los cuales son:

- Hipertrigliceridemia: Se diagnostica cuando los niveles de triglicéridos (TG) en sangre son iguales o superiores a 150 mg/dl, o si el individuo está bajo tratamiento para este trastorno.
- Hipercolesterolemia total: Esta condición se establece si los niveles de colesterol total (TC) en la sangre son iguales o superiores a 200 mg/dl, o si el paciente se encuentra en tratamiento.
- Hipo HDL-C: Se considera que existe dislipidemia si los niveles de colesterol de lipoproteínas de alta densidad (HDL-C) son menores a 40 mg/dl en hombres o menores a 50 mg/dl en mujeres, o si el individuo está recibiendo tratamiento.
- Hiper LDL-C: Se diagnostica cuando los niveles de colesterol de lipoproteínas de baja densidad (LDL-C) son iguales o superiores a 130 mg/dl, o si el paciente está bajo tratamiento.

Para el abordaje de rasgos cuantitativos, los valores de los distintos tipos de lípidos fueron analizados como variables continuas, por lo que se realizó su transformación logarítmica para evitar sesgos por una distribución no uniforme.

11.4. Estudio de asociación

11.4.1. Análisis por variante

Utilizando la herramienta informática PLINK (v1.07) (Purcell et al., 2007), se realizó un estudio de casos y controles para cada uno de los tipos de dislipidemias (HTG, HTC, Hipo HDL-C e Hiper LDL-C) utilizando las SNVs identificadas a partir de los datos de exoma que presentaran un MAF > 0.01. Los datos de asociación se obtuvieron por medio de modelos de

regresión logística, bajo un modelo aditivo, ajustados por variables de confusión que incluyeron la edad, género, IMC, ancestría y el estatus DT2. Valores de $P < 0.05$ fueron considerados como significativos.

11.4.2. Estudio cuantitativo

En este enfoque, se incluyeron todas las variantes presentes en el conjunto de genes candidatos sin aplicar un filtro de frecuencia. Se llevó a cabo un análisis de asociación de tipo cuantitativo que implicó la evaluación de variables continuas (TG, TC, HDL-C y LDL-C) mediante regresión lineal, empleando un modelo aditivo. Además, se realizaron ajustes por variables de confusión, que abarcaron factores como la edad, el género, IMC, ancestría y el estatus DT2. Valores de $P < 0.05$ fueron considerados como significativos.

11.4.3. Análisis por haplotipos

La construcción de los haplotipos se realizó de seleccionando de manera manual un conjunto de marcadores para su subsecuente análisis, o bien utilizando el comando de PLINK (v1.07) (Purcell et al., 2007) para generar los bloques de haplotipos. Inicialmente, se evaluó el desequilibrio de ligamiento (LD), la estructura y frecuencia de los haplotipos por medio del software Haploview versión 4.2 (Barrett et al., 2005) utilizando los valores de R^2 . En este trabajo los valores de R^2 mayores a 0.8 fueron considerados como indicativos de variantes en alto LD. Los bloques de haplotipos se estimaron siguiendo el procedimiento predeterminado en Haploview, incluyendo en este análisis a los individuos sin fenotipos no faltantes. Por defecto, el desequilibrio de ligamiento (LD) se calcula solo para SNVs dentro de una ventana de 200 kb. El estudio de asociación fue realizado con el software PLINK por regresión logística en el caso de los fenotipos dicotómicos y por regresión lineal para los rasgos cuantitativos, ambos bajo un modelo aditivo ajustando por las variables de confusión antes mencionadas. Para este análisis también se consideró un valor de $P < 0.05$ como significativo.

11.4.4. Análisis de variantes raras

Para el análisis de las variantes raras se ocuparon las herramientas KGGseq (Li et al., 2012) con el algoritmo SKAT-o (*SNP-set/Sequence Kernel Association Test-optimal*) para ambos fenotipos (cuantitativo y binario). El SKAT busca detectar asociaciones genéticas al evaluar la contribución conjunta de grupos de variantes en lugar de considerarlas individualmente, lo que puede mejorar la capacidad para identificar efectos genéticos sutiles y complejos en enfermedades y rasgos fenotípicos. Para el estudio de variantes individuales se empleó el programa WISARD (Lee et al., 2018) utilizando el análisis de “*Score test*”, el utiliza en análisis de datos y estudios de asociación genética. Este tipo de prueba se utiliza para evaluar si un conjunto de datos es consistente con un modelo estadístico en particular (Fan et al., 2014). Valores de $P < 0.05$ fueron considerados como significativo.

11.5. Replicación de resultados

A partir de los resultados del estudio de asociación por variantes individuales realizado en la cohorte de descubrimiento fueron seleccionados 7 SNVs para su replicación en la población de validación. Los criterios para la selección de estas variantes fueron:

- 1) nivel de significancia P, es decir aquellos que por rasgo representen las asociaciones estadísticamente más representativas de los estudios de asociación
- 2) en la anotación funcional del SNV, se dio preferencia a aquellas variantes que tuvieran una mayor posibilidad de presentar impacto biológico, como lo puede ser por un cambio de aminoácido en la proteína del gen que codifica, y
- 3) que no hubiera una descripción extensa de la variante en la literatura.

La genotipificación de las variantes en la población de validación se realizó por el método de discriminación alélica por sondas de TaqMan (ThermoFisher) con el equipo QuantStudio 3 Real-Time PCR System (Applied Biosystems, Foster City, CA, EUA). La tasa de genotipificación que se observó fue mayor al 94% en todos los SNVs seleccionados. Para validar y confirmar los genotipos, se seleccionaron al azar el 10% de las muestras y se genotipificaron por duplicado.

La discriminación alélica por medio de ensayos TaqMan® es una aplicación de la PCR en tiempo real por la cual se pueden detectar los alelos presentes en una muestra de DNA para una SNV determinada. Los ensayos TaqMan® consisten en un sistema de sondas marcadas con fluoróforos en sus extremos 5', y una molécula que inhibe la emisión de fluorescencia (quencher) en los extremos 3'. En el caso de los ensayos TaqMan® para discriminación alélica se encuentran presentes dos sondas marcadas con diferentes fluoróforos, cada una de las cuales es complementaria a uno de los alelos alternativos. Las sondas entonces hibridan con la zona complementaria del producto específico de PCR (amplicón). La señal de fluorescencia se genera cuando la polimerasa interactúa con la sonda hibridada y la hidroliza mediante su actividad exonucleasa 5'-3'; lo que provoca la separación del quencher y el fluoróforo. Este proceso elimina la inhibición del fluoróforo, resultando en la emisión de una señal fluorescente, que es detectada por el sistema óptico de los termocicladores en tiempo real y de esta forma revelando los alelos identificados en la muestra de acuerdo con las especificaciones del fabricante.

En las condiciones para el ensayo de discriminación de las sondas empleadas se consideró la preparación de mezclas de reacción que incluyó los siguientes componentes en las cantidades especificadas: 2.626 µL unidades de agua libre de nucleasas (NFW), 2.30 µL de GMM (máster mix de genotipificación) y 0.074 µL de sonda 80X específica para cada variante seleccionada. Esta mezcla fue sometida a un protocolo estándar de PCR en punto final: brevemente, una etapa de 95°C por 10 min, 45 ciclos con una rampa de temperaturas con 95°C durante 15 segundos, 60°C durante 60 segundos, y un hold de 4°C para retirar la placa del termociclador.

Posteriormente se llevó a cabo la lectura de las placas post-PCR en el equipo QuantStudio 3 y los datos crudos una vez fueron analizados con el # Software Version Number: 3.4.5-Q1-18-build19 para asignar los alelos y esta información fue exportada en Excel para organizar los genotipos correspondientes a cada una de las muestras analizadas y cada una de las variantes replicadas.

11.6. Ensayos *in silico*.

Para revisar el posible efecto funcional de las variantes asociadas con las dislipidemias se ocupó el software VEP (McLaren et al., 2016) usando los módulos de SIFT (Ng & Henikoff, 2001) y PolyPhen-2 (Adzhubei et al., 2010) para las SNVs ubicadas en regiones codificantes. y las bases de datos miRBase (Griffiths-Jones et al., 2008), Patrocles, PolymiRTS, TargetScan y miRanda para las que se ubicaron en regiones 3'UTR.

11.7. Análisis de frecuencias alélicas

Para las variantes replicadas en el estudio, se determinaron las frecuencias alélicas y genotípicas comparando las frecuencias en la cohorte de descubrimiento y la de replicación. Además, se incluyeron comparaciones de frecuencias contra las poblaciones de referencia del proyecto 1000 genomas (Altshuler et al., 2010) empleando el programa de genética de poblaciones Genepop versión 4.7 (Rousset, 2008) usando la versión online disponible en <https://genepop.curtin.edu.au>

11.8. Análisis estadísticos adicionales

Los datos clínicos se presentan como valores de medias y desviación estándar. La significancia estadística se asumió a valores de $p < 0.05$. Tanto los valores como las gráficas presentadas fueron realizadas con códigos propios mediante el uso del ambiente R CRAN Project o Microsoft Excel.

12. Resultados

12.1. Características clínicas y demográficas de la población de estudio

Para la primera etapa se contó con los datos de la secuenciación del exoma (Estrada et al., 2014) de 968 individuos mestizos, pertenecientes a la cohorte DMS. Para realizar la replicación de los resultados de asociación en la población de descubrimiento y para la comparación de las frecuencias observadas entre las poblaciones mestiza y amerindia, se analizaron 671 muestras de DNA de individuos pertenecientes a 30 grupos étnicos del país, incluidos en la cohorte MAIS.

La muestra de Mestizos consistió en 968 individuos: 681 mujeres (70.3%) y 287 hombres (29.7%). La edad promedio de los participantes fue de 54.1 ± 9.8 años y el índice de masa corporal promedio fue de 28.4 ± 4.9 kg/m². Según los puntos de corte de la AHA/NHLBI, el rasgo lipídico con mayor prevalencia en esta muestra fue el Hipo HDL-C (70.8%), seguido por el HTC (65.1%), HTG (54.9%) y el Hiper LDL-C (37%). La muestra de AMERINDIOS estuvo conformada por 671 individuos: 491 mujeres (73.2%) y 180 hombres (26.8%). De manera similar a la cohorte de descubrimiento, los niveles bajos de HDL-C (72.9%) fueron el rasgo lipídico afectado más común, seguido por la HTG (59.2%), HTC (28.2%) y el Hiper LDL-C (14.6%)

La Tabla 1 muestra los parámetros bioquímicos de mestizos y amerindios, así como las proporciones de la presencia de dislipidemias en las muestras.

Dentro de los valores demográficos observados de las poblaciones de trabajo, pudimos ver que, aunque los niveles de glucosa y de LDL-C presentados en la cohorte de mestizos fueron ligeramente más elevados respecto a los presentes en la cohorte de amerindios, solo la proporción de HTC y de Hiper LDL-C fueron significativamente diferentes entre las dos cohortes estudiadas.

Tabla 1 Características clínicas y demográficas de la población de estudio

Parámetro	MESTIZOS n= 968	AMERINDIOS n= 671
Mujer/Hombre (%)	70.3/29.7	73.2/26.8
Edad (Años)	54.1 ± 9.8	51.1 ± 15.8
IMC (Kg/m ²)	28.4 ± 4.9	27.0 ± 4.9

Glucosa en ayuno (mg/dL)		125.2 ± 61.4	110.5 ± 59.3
HTG (%)		54.9	59.2
TG (mg/dL)		203.7 ± 129.3	203.11 ± 141
HTC (%)		65.1	28.2*
TC (mg/dL)		201.1 ± 43.8	178.8 ± 38.8
Hipo HDL-C (%)		70.8	72.9
HDL-C (mg/dL)	Mujer	39.7 ± 12.1	39.6 ± 11.9
	Hombre	37.4 ± 13.0	38.5 ± 14.1
Hiper LDL-C (%)		37	14.6*
LDL-C ¹ (mg/dL)		120.7 ± 35.3	102.5 ± 30.5

* = $P < 0.05$. Los valores de P fueron analizados con la prueba de Mann Whitney-Wilcoxon para datos paramétricos, z-test fue empleado para la comparación de proporciones. IMC= Índice de masa corporal, HDL-c = Colesterol HDL. LDL-C = Colesterol LDL. ¹ determinado con la fórmula de Friedewald. Los valores se presentan como promedio ± desviación estándar

12.2. Variantes de un solo nucleótido en los genes candidato

En el estudio de los 177 genes candidatos, se identificaron un total de 8,405 variantes (Figura 7). De las estas variantes, el 50.74% corresponden a SNVs denominadas *singletons*, que se presentan en un solo individuo y en una sola copia en el conjunto de variantes obtenidas a partir de la secuenciación del exoma de las muestras analizadas. Después de realizar las anotaciones empleando el programa VEP observamos que el 99.21% de las variantes ha sido reportadas previamente, mientras que el 0.79% restantes son SNVs potencialmente nuevas, es decir, sin reporte en la literatura.

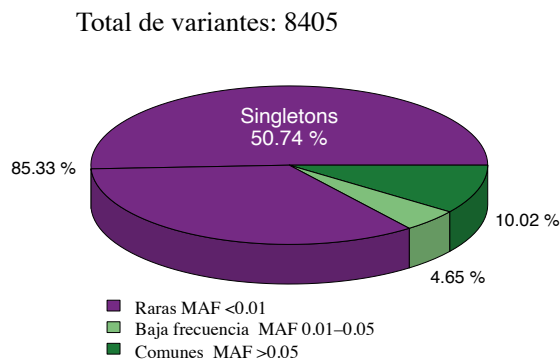


Figura 7 Descripción de las variantes por frecuencia.

Distribución acumulativa de variantes de acuerdo con la MAF). Las variantes raras incluyen Singletons, Doubletones y variantes con MAF <0.01. Las variantes de baja frecuencia son aquellas con MAF de 0.01 a 0.05. Las variantes comunes son aquellas con MAF >0.05.

Del total de variantes encontradas, el 58% fueron codificantes y el resto se encontraron en sitios no codificantes (Figura 8). Dentro de las variantes codificantes, las no sinónimas fueron las más frecuentes (32.25%), seguidas de las variantes sinónimas (23.22%), las no sinónimas/región *splicing* (0.84%), sinónimas/región *splicing* (0.69%), codones de paro ganados (0.49%), pérdidas de codones de inicio (0.06%), ganancia de codón de paro/región *splicing* (0.04%) y pérdida de codones de paro (0.01%) (Figura 8, barra izquierda). En las regiones no codificantes las variantes intrónicas fueron las más frecuentes (34.03%), seguidas de las variantes en región de *splicing*/intrón (3.67%), 3'UTR (1.3%), 5'UTR (1.26%), gen rio arriba (0.61%), intrón/transcrito no codificante (0.52%), exones transcritos no codificante/transcrito no codificantes (0.46%), gen rio abajo (0.21%), donador de *splicing* (0.2%), aceptor de *splicing* (0.11%), y finalmente región *splicing* /5'UTR (0.01%) (Figura 8, barra derecha).

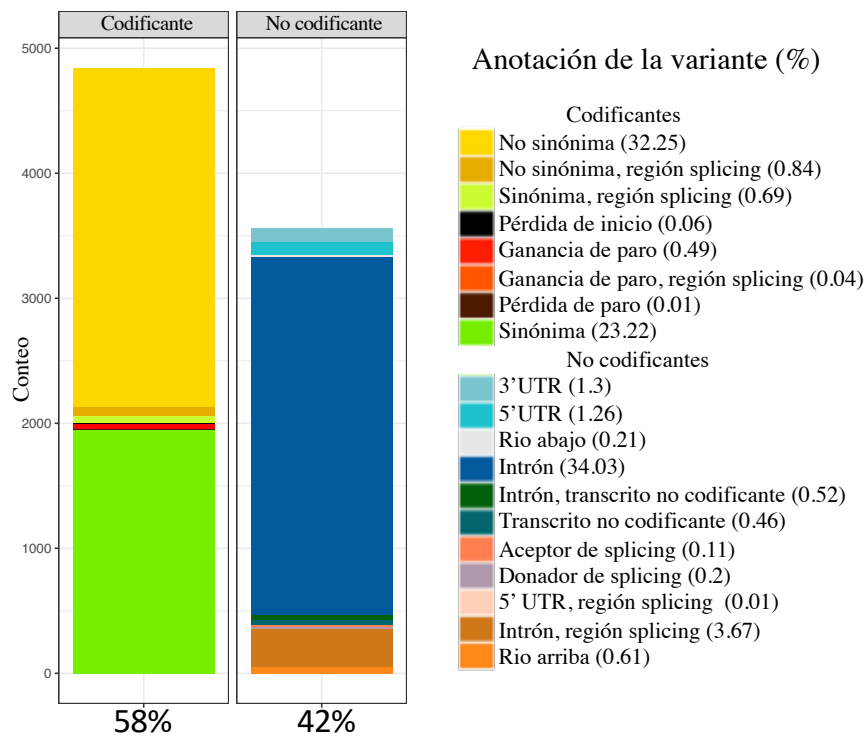


Figura 8 Anotación de las variantes

UTR: regiones no traducidas o regiones no codificantes; la proporción de cada variante se presenta como numero decimal entre paréntesis.

12.3. Catálogo de las variantes

Después de realizar la anotación de las variantes, se construyó un catálogo considerando el total de SNV presentes en cada región secuenciada de los genes candidato. Este catálogo consideró si cada variante ya había sido previamente reportada, así como su ubicación dentro de la secuencia génica.

A continuación, se presentan los resultados obtenidos en la Tabla 2, que muestra los primeros 5 genes con la mayor cantidad de variantes codificantes dentro del conjunto de datos:

Tabla 2 Fragmento del Catálogo de variantes

Gen	Total SNV	SNV reportadas	SNV sin reporte	Exón	3'UTR	5'UTR	Intrón	Intergénico	Longitud de transcrito*	SNV cod / transcrito	SNV / transcrito
<i>PLEC</i>	561	556	5	440	0	0	115	6	15249	27	35
<i>DNAH11</i>	344	341	3	202	1	0	141	0	14167	41	70
<i>APOB</i>	186	186	0	146	0	0	40	0	14121	76	97
<i>LRP1</i>	264	260	4	143	1	0	120	0	14897	56	104
<i>CUBN</i>	246	245	1	139	0	0	107	0	11949	49	86

SNV: Variantes de un solo nucleótido; SNV cod: SNV en región codificante; *: Las longitudes de los transcritos fueron obtenidos del browser ENSEMBL (http://grch37.ensembl.org/Homo_sapiens/Gene/)

En este catálogo, se detallan varios aspectos importantes relacionados con las variantes genéticas. Se registran el número total de SNV presentes en cada gen, la cantidad de SNV que ya ha sido reportada previamente y aquellos que no lo han sido. Además, se indica la posición de las variantes en relación con regiones clave de los genes, como exones, regiones 3'UTR, 5'UTR, intrones y regiones intergénicas.

12.4. Estudio de asociación Caso-Control.

Del total de las variantes detectadas en los 177 genes, 1233 presentaron un MAF mayor al 1% (Tabla 3) y fueron utilizadas para los estudios de asociación por medio de una regresión logística corrigiendo como se indica en la sección de Material y Métodos. El efecto de estas variantes predicho por el software VEP se presenta en la Tabla 3

Tabla 3 Tipo de variantes incluidas en el estudio de asociación caso control

Tipo de variante	Total	Tipo de variante	Total
Intrón	507	No sinónimas, región <i>splicing</i>	6
Sinónimas	330	Río abajo	3
No sinónimas	260	Aceptor de <i>splicing</i>	1
Transcrito de exón no codificante, transcrito no codificante	18	Pérdida de inicio	1
5'UTR	18	Ganancia de codón de paro	1
3'UTR	17	Donador de <i>splicing</i>	2
Región de <i>splicing</i> , intrón	51	Intrón, transcrito no codificante	7
Región <i>splicing</i> , sinónimas	11		

Los resultados del estudio de asociación para los diferentes rasgos lipídicos presentaron factores de inflación genómica entre 1-1.3 Los resultados de asociación para todas las variantes asociadas se pueden consultar en la Tabla suplementaria 3. Entre estas variantes, 23 (distribuidas en 15 genes) estaban asociadas con la HTG, 12 variantes (presentes en 10 genes) con la HTC, 28 variantes (presentes en 11 genes) Hipo HDL y 12 variantes (distribuidas en 8 genes) con Hiper LDL-C. A manera de resumen y considerando solamente aquellas variantes con un posible efecto funcional, en la Tabla 4 se presentan las variantes con mayores niveles de asociación a los rasgos lipídicos. Es importante mencionar que mantuvimos la variante en *BUD13* rs3825041 por presentar el OR de mayor magnitud y significancia.

Tabla 4 Variantes comunes asociadas con dislipidemias en la cohorte de descubrimiento

Chr	Gene	rsID	Consecuencia	MAF FA / FU	OR [95 % CI]	P	Beta (P)
Hipertrigliceridemia							
2	<i>GCKR</i>	rs1260326_T	No sinónima (L446P)	0.33 / 0.26	1.47 [1.17–1.83]	7.41E-04	18.17 ^b
11	<i>APOA4</i>	rs5104_C	No sinónima (S147N)	0.3 / 0.23	1.44 [1.14–1.81]	1.74E-03	21.66 ^a
	<i>APOA5</i>	rs651821_C	5'UTR	0.24 / 0.16	1.65 [1.28–2.13]	1.12E-04	21.26 ^b
	<i>APOC3</i>	rs5128_G	3'UTR	0.29 / 0.2	1.63 [1.28–2.06]	5.47E-05	20.82 ^b
	<i>BUD13</i>	rs3825041_T	Intrón	0.23 / 0.15	1.71 [1.32–2.22]	4.94E-05	21.25 ^a
		rs11820589_A	No sinónima (P148L)	0.19 / 0.14	1.45 [1.11–1.88]	5.48E-03	22.33 ^b
16	<i>CETP</i>	rs5882_G	No sinónima (V357I)	0.43 / 0.48	0.76 [0.61–0.92]	6.41E-03	-9.99
19	<i>ANGPTL4</i>	rs186754194_C	No sinónima (M41I)	0.03 / 0.05	0.54 [0.34–0.85]	8.19E-03	-24.66 ^b
Hipercolesterolemia							
2	<i>APOB</i>	rs1367117_A	No sinónima (T98I)	0.31 / 0.25	1.34 [1.09–1.65]	5.64E-03	6.56 ^b
7	<i>MLXIPL</i>	rs3812316_G	No sinónima (Q241H)	0.04 / 0.06	0.56 [0.36–0.86]	8.74E-03	-9.06 ^c
12	<i>PHCI</i>	rs1049925_G	No sinónima (T693A)	0.42 / 0.47	0.77 [0.63–0.93]	9.40E-03	-2.78
18	<i>LIPG</i>	rs34474737_G	5'UTR	0.09 / 0.12	0.64 [0.47–0.87]	4.28E-03	-6.27
19	<i>APOE</i>	rs429358_C	No sinónima (C130R)	0.12 / 0.09	1.51 [1.12–2.04]	6.65E-03	5.06
Hipo α-lipoproteinemia							
6	<i>HFE</i>	rs2071303_C	Splice región, intrón	0.4 / 0.5	0.66 [0.53–0.81]	1.12E-04	1.69 ^b
7	<i>CD36</i>	rs3173798_C	Splice región, intrón	0.21 / 0.14	1.45 [1.10–1.91]	7.15E-03	-2.1 ^a
	<i>TMEM176A</i>	rs10378_T	No sinónima (L187F)	0.06 / 0.1	0.6 [0.41–0.86]	6.28E-03	1.78
10	<i>GPAM</i>	rs2792751_T	No sinónima (I43V)	0.21 / 0.29	0.65 [0.51–0.82]	3.42E-04	1.36 ^c
11	<i>BUD13</i>	rs10488698_A	No sinónima (R120C)	0.13 / 0.2	0.51 [0.38–0.67]	1.59E-06	4.58 ^a
12	<i>HNF1A</i>	rs2464195_A	Splice región, intrón	0.35 / 0.44	0.66 [0.54–0.82]	1.56E-04	1.51 ^b
		rs2464196_A	No sinónima (S487N)	0.34 / 0.42	0.68 [0.55–0.84]	3.17E-04	1.39 ^b
		rs1169288_C	No sinónima (I27L)	0.33 / 0.4	0.69 [0.56–0.84]	3.88E-04	1.12 ^c
16	<i>CETP</i>	rs1532625_T	Splice región, intrón	0.37 / 0.46	0.65 [0.53–0.80]	6.78E-05	2.16 ^a
19	<i>ANGPTL8</i>	rs2278426_T	No sinónima (R59W)	0.37 / 0.28	1.41 [1.12–1.77]	3.37E-03	-1

Hiper LDL-C							
1	<i>ANXA9</i>	rs267733_G	No sinónima (D166G)	0.05/0.07	0.55 [0.36–0.85]	7.20E-03	-5.34
5	<i>HMGCR</i>	rs5909_A	3'UTR	0.2 / 0.15	1.47 [1.13–1.91]	3.40E-03	7.3 ^a
17	<i>PGSI</i>	rs2292642_C	Splice región, sinónimos	0.26 / 0.33	0.72 [0.57–0.89]	3.20E-03	-4.11 ^c
		rs2292643_G	Splice región, intrón	0.26 / 0.33	0.72 [0.57–0.89]	3.20E-03	-4.11 ^c

Chr: Cromosoma, MAF: Frecuencia de alelo menor; FA: frecuencia en individuos afectados; FU: Frecuencia en individuos no afectados. a: $P \leq 0.001$, b: $P \leq 0.01$, c: $P \leq 0.05$.

Los resultados del estudio de asociación en la cohorte de descubrimiento se plasman como gráficos de Manhattan Figura 9-13, donde se presenta la distribución de los valores de significancia de la asociación (valores de P) a lo largo de los cromosomas. En este tipo de gráfico, los cromosomas se muestran en el eje horizontal, mientras que en el eje vertical se representan los valores negativos del logaritmo decimal del valor de P .

Cada punto en el gráfico de Manhattan representa una SNV individual. La altura del punto en el eje vertical indica la significación estadística de la asociación de esa variante con un fenotipo específico. Las asociaciones más fuertes se representan como picos altos en el gráfico.

Para el caso de la HTG (Figura 9), destacan la señal de la región del Chr11, teniendo como pico más alto la asociación para el clúster de *Apo/BUD13* (Tabla 4).

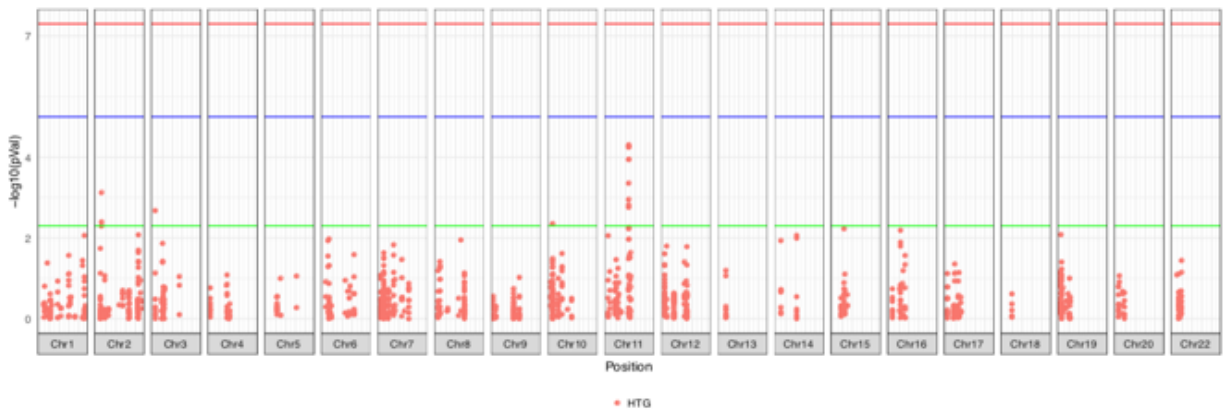


Figura 9 Variantes asociadas a HTG
Manhattan plot de los resultados del estudio de asociación para HTG. La línea verde representa $P=0.005$, línea azul $P=5E-05$, línea roja $P=5E-08$.

Respecto al caso de HTC (Figura 10), fueron pocas las variantes que superaron el umbral de $P=0.005$, siendo las variantes rs34474737 en el gen *LIPG* del cromosoma 18 y la variante rs1367117 en el gen *APOB* del cromosoma 2 los SNV potencialmente funcionales con mayor grado de asociación (Tabla 4). Por su parte, la asociación más significativa se presentó en el cromosoma 22 para la variante rs139047 del gen *PNPLA3* (Tabla suplementaria 2)

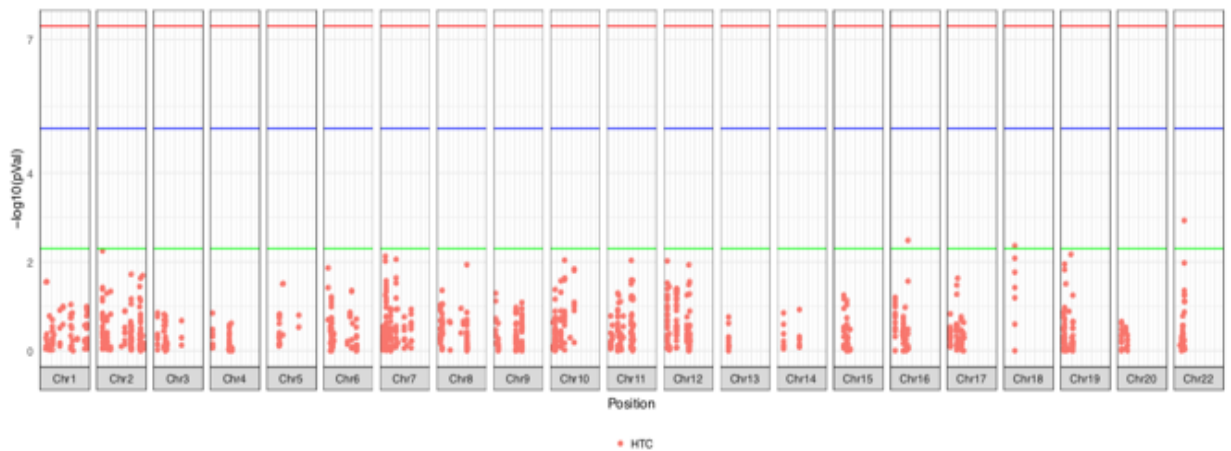


Figura 10 Variantes asociadas a HTC

Manhattan plot de los resultados del estudio de asociación para HTC. La línea verde representa $P=0.005$, línea azul $P= 5E-05$, línea roja $P= 5E-08$.

Para el fenotipo de Hipo HDL-C, los principales picos de señales presentados en el gráfico de Manhattan (Figura 11) se presentan en el cromosoma 11, y corresponden a dos variantes localizadas en el gen *BUDI3* una de las cuales su anotación es no sinónima (Tabla 4). Adicionalmente, en este rasgo se identificó una variante en el cromosoma 16 en el gen *CETP* (Tabla 4)

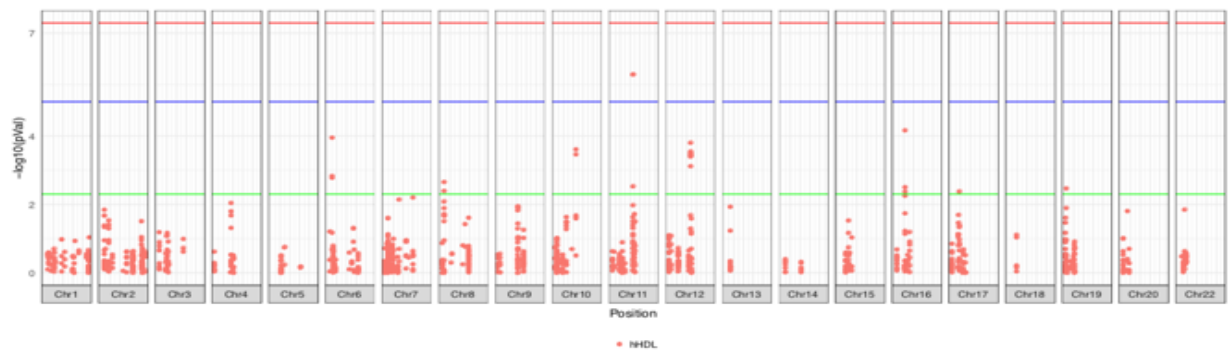


Figura 11 Variantes asociadas a Hipo HDL-C

Manhattan plot de los resultados del estudio de asociación para Hipo HDL-C. La línea verde representa $P=0.005$, línea azul $P= 5E-05$, línea roja $P= 5E-08$.

Finalmente, en la Figura 12 se presenta el Manhattan plot del rasgo Hiper LDL-C, donde se aprecia que las señales más altas son para las variantes en cromosoma 11 en el gen *FADS2* y en el cromosoma 16 en el gen *CMIP* (Tabla suplementaria 2), ambas variantes localizadas en una región intrónica. Para este rasgo, las variantes con un posible efecto funcional que presentaron un mayor nivel de asociación son las presentes en el cromosoma 17 en el gen *PGSI* (Tabla 4)

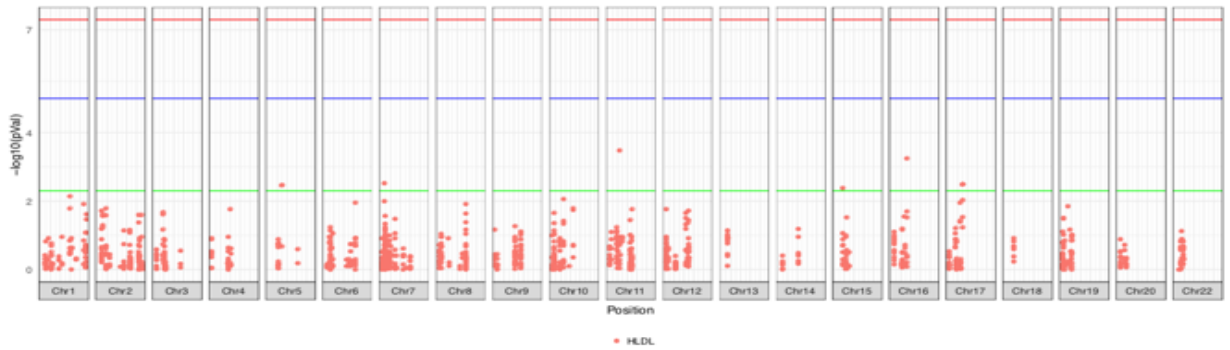


Figura 12 Variantes asociadas a Hiper LDL-C
 Manhattan plot de los resultados del estudio de asociación para Hiper LDL-C. La línea verde representa $P=0.005$, línea azul $P= 5E-05$, línea roja $P= 5E-08$.

De manera interesante, se pudo observar que del conjunto de genes candidato solamente seis genes presentan asociaciones hacia dos componentes en la cohorte de descubrimiento, como se puede observar en el diagrama de cuerdas de la Figura 13

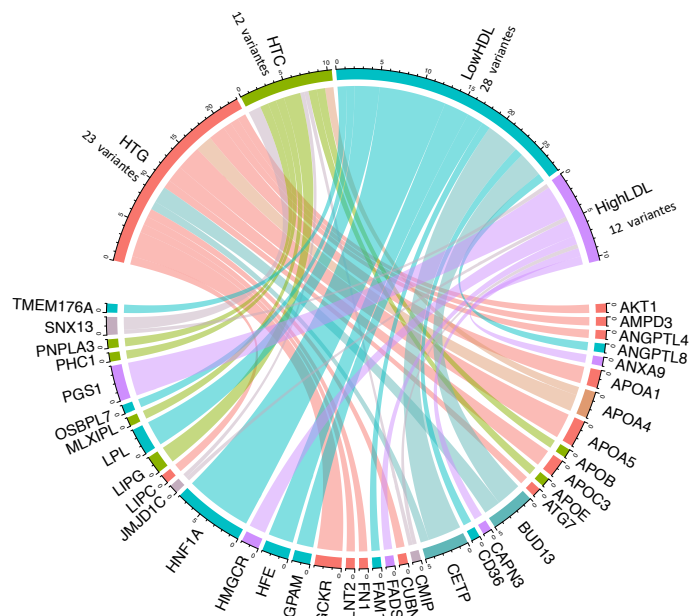


Figura 13 Genes con variantes asociadas a los rasgos lipídicos
 En la parte superior del diagrama de cuerdas se muestran los rasgos lipídicos y en la parte inferior los genes con los cuales se obtuvo una asociación genética. Las bifurcaciones en los segmentos que salen de los genes implican la asociación del gen a más de un componente, Hipo HDL-C: Hipo α -lipoproteinemia, HTC: Hipercolesterolemia, HTG: Hipertrigliceridemia, Hiper LDL-C

Los genes *BUD13* y *CETP* se encontraron asociados con los rasgos de HTG e Hipo HDL-C, en tanto que *SNX12*, *JMJD1C* y *CMIP* presentaron asociación con HTC e Hiper LDL-C. Por su parte, *APOA4* presentó asociación con HTC y HTG.

12.5. Estudio de asociación de haplotipos.

Se pudo observar que las variantes encontradas asociadas de manera individual a Hipo HDL-C en *HNF1A* presentan un LD alto para dos SNVs (rs2464195 y rs2464196) con una $R^2=89$ (Figura 14A), al realizar el estudio de asociación con las tres variantes, encontramos que el haplotipo CAA se asocia de manera significativa con protección ante Hipo HDL-C OR= como se observa en la Figura 14B. A partir de los datos públicos del proyecto 1000 genomas obtuvimos las frecuencias alélicas de estas tres variantes, y con ellas calculamos las frecuencias de los haplotipos (Figura 14C). Este haplotipo parece estar ausente en las poblaciones africanas, pero de una manera común se presenta en poblaciones con ancestría asiática, europea y en algunos grupos de América.

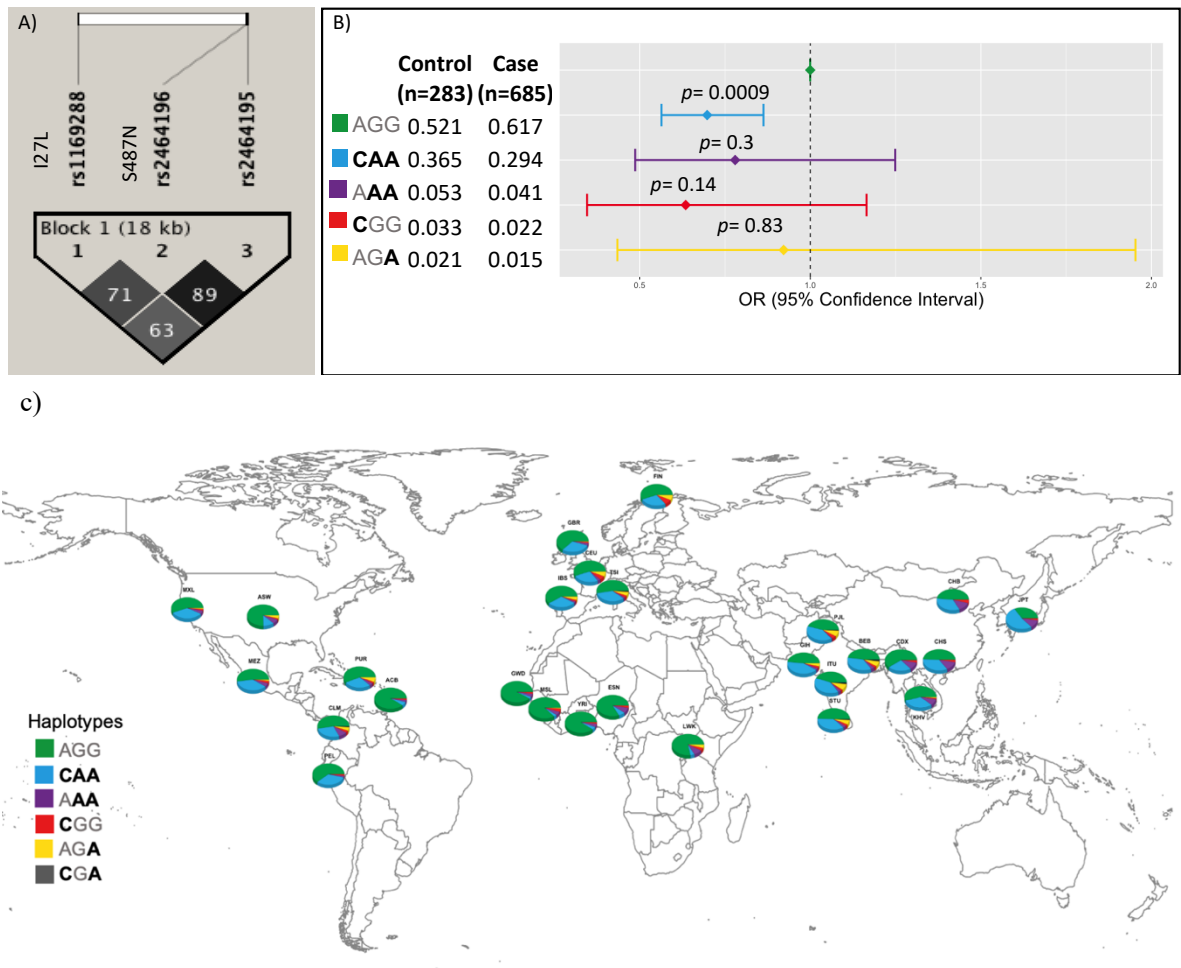


Figura 14 Desequilibrio de Ligamiento de las SNVs en *HNF1A* y asociación con Hipo HDL-C
 A) Variantes en el gen *HNF1A* incluidas en el estudio de haplotipo; B) estudio de asociación caso-control por haplotipo, los alelos alternativos se presentan en negritas; C) Distribución de las frecuencias de los haplotipos en las poblaciones del proyecto 1000 genomas AFR: Africana, AMR: Americano de ascendencia mixta, ASJ: Judía Ashkenazi, EAS: Asiática del Este, EUR: Europea, FIN: Finlandesa, NFE:

Europea no finlandesa, SAS: Asiática del Sur, ACB: Caribeña Africana, CDX: Chino Dai, GWD: Gambiana en Divisiones Occidentales en Gambia, IBS: Población Ibérica en España, JPT: Japonesa en Tokio, Japón, LWK: Luhya en Webuye, Kenia, MSL: Mende en Sierra Leona, MXL: Mexicano en Los Ángeles, EE: UU., PEL: Peruana en Lima, Perú, PJI: Punjabi en Lahore, Pakistán, STU: Tamil de Sri Lanka en el Reino Unido, TSI: Toscanos en Italia, YRI: Yoruba en Ibadán, Nigeria, BEB: Bengali en Bangladés, GBR: Británica en Inglaterra y Escocia, CLM: Colombiana en Medellín, Colombia, PUR: Puertorriqueña en Puerto Rico, CHB: Han Chino en Pekín, China

Al realizar el análisis de frecuencias para los haplotipos en este gen entre las poblaciones parentales y una muestra de mexicanos del proyecto 1000 genomas, se pudo observar que, en efecto, solo hay diferencias significativas comparando contra la población africana (Tabla 5)

Tabla 5 Análisis de frecuencias del haplotipo en HNF1A entre la población de descubrimiento y otras poblaciones

Haplotipo	MEZ (n=968)	MXL (n=67)	IBS (n=107)	YRI (n=108)
AGG	0.589	0.569	0.602	0.883*
CAA	0.315	0.342	0.293	0.045*
AAA	0.045	0.047	0.024	0.047
CGG	0.026	0.024	0.023	0.024
AGA	0.017	0.015	0.051*	0*

*=P<0.05 en la comparación contra la población de descubrimiento. Los valores de P fueron determinados a partir de la prueba de χ^2

Se llevó a cabo el estudio de asociación por haplotipos para los genes que presentaron más de una variante asociada al mismo rasgo lipídico, el resumen de la información se presenta en la Tabla 6 y la información en extenso está disponible en la Tabla suplementaria 4.

Tabla 6 Resultados representativos de asociación por haplotipos

Gen	Haplotipo	Frq	Rasgo	OR	P ajustada	Beta	P ajustada
APOA1	AA	0.237	HTG	1.65	3.32E-05	18.9	0.0054
APOA5	CG	0.611	HTG	0.548	3.44E-08	-31.6	1.22E-07
APOC3	TG	0.255	HTG	1.62	6.05E-05	20.8	0.00154
	CC	0.432	HTG	0.655	4.11E-05	-28.4	1.34E-04
BUD13	ATA	0.15	Hipo HDL-C	0.517	2.38E-06	4.58	8.20E-11
	CG	0.621	HTG	0.545	2.37E-08	-29.6	6.17E-07

Se presentan solo los resultados más significativos de acuerdo con el valor P del estudio de haplotipos. En negrito se señalan los resultados con mayor asociación. Frq: Frecuencia

En este caso, la asociación más fuerte fue un haplotipo de protección hacia HTG formado por las variantes de BUD13 rs3741300 y rs3825041; seguidos del formado en APOA5 por las variantes rs2075560 y rs12287066 Figura 15b; y en tercer lugar el haplotipo de BUD13 formado por las tres variantes rs17519093, rs918144 y rs10488698 (Figura 15d) asociado con Hipo HDL-C.

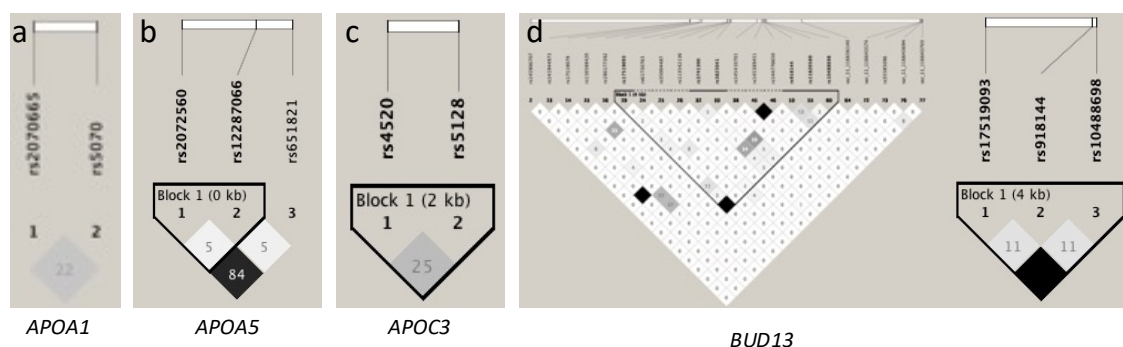


Figura 15 Desequilibrio de ligamiento para las SNV presentes en los bloques de haplotipos. La escala de grises representa un alto grado de LD, se presentan valores de R^2 . En el caso de BUD13 se presenta el diagrama completo y con las variantes agrupadas asociadas a Hipo HDL-C.

12.6. Análisis basados en gen

El análisis estadístico se realizó utilizando el método SKAT implementado usando KGGseq, que es un método robusto para detectar asociaciones genéticas en estudios con un gran número de variantes genéticas. Los resultados del análisis SKAT hacia los rasgos dicotómicos se muestran Tabla 7. Se presentan los valores P de SKAT-O dado que este es una versión corregida del valor P de SKAT que controla la inflación del tipo I debido a la correlación entre las variantes genéticas.

Tabla 7 Resultados de la asociación dicotómica de variantes raras en la cohorte de descubrimiento

CHR	Gene	#Var	SKAT_P	SKATO_P	Burden_P
HTG					
8	TRIB1	9	1.18E-02	3.02E-03	2.07E-03
16	CETP	34	3.27E-01	1.90E-02	9.92E-03
7	KLF14	8	4.61E-02	2.06E-02	1.29E-02
20	SNX5	23	6.50E-01	4.33E-02	2.45E-02
HTC					
15	FRMD5	18	1.57E-01	1.32E-03	6.72E-04
22	PNPLA3	31	5.74E-02	1.34E-02	7.65E-03
11	LRP4	81	6.63E-02	1.41E-02	7.87E-03
14	NYNRIN	48	2.58E-02	1.65E-02	1.27E-02
Hipo HDL-C					
11	BUD13	29	4.22E-03	4.13E-03	7.18E-03
11	PHLDB1	64	7.38E-01	5.49E-03	2.73E-03
9	APOB	150	9.39E-02	6.32E-03	3.17E-03
12	PDE3A	43	1.41E-02	1.62E-02	2.91E-02
Hiper LDL-C					
14	NYNRIN	48	1.21E-02	2.02E-03	1.40E-03
1	RP11-316M1.12 ^a	13	4.62E-01	8.52E-03	4.21E-03
12	A2ML1-AS1 ^a	7	2.09E-02	9.78E-03	6.54E-03
9	KLF4	22	1.06E-02	1.70E-02	4.99E-01

Se presentan solo los resultados más significativos de acuerdo con el valor P del estudio SKAT-O. #Var: Cantidad de variantes incluidas en el análisis; SKAT_P: Valor P del análisis SKAT, SKATO_P:

Valor *P* del análisis SKAT-O; Burden_ *P* Valor de *P* del análisis de agregación. *Nombre de los genes asignados por la base de referencia del programa KGGseq

Siguiendo este enfoque, se observó la asociación hacia HTG a nivel gen de *TRIB1* SKATO $P=3.02E-03$ al considerar un total de 9 variantes en el análisis, seguido del gen *CETP* SKATO $P=1.90E-02$ considerando 34 SNVs en el set evaluado. Para el caso de HTC, los genes *FRMD5* con 18 SNVs y *PNPLA3* con 31 SNV fueron los más asociados con este fenotipo (SKATO $P=1.32E-03$ y $1.34E-02$, respectivamente). Respecto al fenotipo de Hipo HDL-C dos genes en el cromosoma 11 representaron los resultados más significativos *BUD13* y *PHLDB1* (SKATO $P=4.13E-03$ y $5.49E-03$). Por su parte, el rasgo de Hiper LDL-C presentó al gen *NYNRIN* con 48 SNVs con un SKATO $P=2.02E-03$.

Respecto a lo encontrado en el estudio cuantitativo, se observaron valores de asociación de $P<1.0E-06$ que es el valor estándar comúnmente empleado en estudios de GWAS para los fenotipos de TG y HDL-C; en el caso de los niveles de TC y LDL-L no se encontró dicho nivel de asociación, presentando niveles de significancia las variantes identificadas de $P<2.04E-05$ y $P<1.55E-04$, respectivamente (Tabla 8)

Tabla 8 Resultados de la asociación cuantitativa de variantes raras en la cohorte de descubrimiento

CHR	SNV	GEN	Frq	Consecuencia	STAT	P
Niveles de TG						
8	<i>PLEC</i>	rs199983097	0.00155	No sinónima (R/W)	55.2098	1.08E-13
8	<i>PLEC</i>	rs367761849	0.00155	Sinónima (P)	55.2098	1.08E-13
8	<i>PLEC</i>	rs782426081	0.00155	Sinónima (E)	55.2098	1.08E-13
8	<i>PLEC</i>	rs201075382	0.001033	No sinónima (E/K)	52.2835	4.80E-13
6	<i>LPA</i>	rs751267533	0.001033	Intrón	49.0357	2.51E-12
Niveles de TC						
5	<i>ARL15</i>	rs547287139	0.0005165	No sinónima (A/V)	18.1515	2.04E-05
7	<i>DNAH11</i>	rs776256138	0.0005165	No sinónima (M/V), Splice región	18.1515	2.04E-05
7	<i>MET</i>	rs754467674	0.0005165	Sinónima (Y)	18.1515	2.04E-05
8	<i>PLEC</i>	rs782313043	0.0005165	No sinónima (H/R)	18.1515	2.04E-05
20	<i>SNX5</i>	rs1555786727	0.0005165	Intrón	18.1515	2.04E-05
Niveles de HDL-C						
11	<i>BUD13</i>	rs17519093	0.1498	Intrón	35.5183	2.53E-09
11	<i>BUD13</i>	rs10488698	0.1498	No sinónima (R/C)	35.5183	2.53E-09
6	<i>RSPO3</i>	rs773525574	0.001033	5'UTR	31.0259	2.55E-08
12	<i>BRAP</i>	rs3782886	0.001033	Sinónima (R)	31.0259	2.55E-08
8	<i>PLEC</i>	rs782468518	0.001033	Sinónima (P)	24.3496	8.03E-07
Niveles de LDL-C						
6	<i>LPA</i>	rs753180293	0.0005165	No sinónima (A/D), Splice región	14.3064	1.55E-04
7	<i>DNAH11</i>	rs111425487	0.0005165	Intrón	14.3064	1.55E-04
7	<i>MLXIPL</i>	rs782324167	0.0005165	No sinónima (R/C)	14.3064	1.55E-04
8	<i>PLEC</i>	rs782269448	0.0005165	Intrón	14.3064	1.55E-04
11	<i>FADS2</i>	rs776092981	0.0005165	Intrón	14.3064	1.55E-04

Se presentan solo los cinco resultados más significativos de acuerdo con el valor *P* del Score Test. CHR: Cromosoma, P: Valor *P* del análisis.

12.7. Estudio de replicación en Amerindios.

Después de realizar los estudios de asociación y ajustando como se indica en la sección de metodología, en la cohorte de replicación se obtuvieron los siguientes resultados:

La asociación con el fenotipo de hipo HDL-C de la variante no sinónima rs2792751 en el gen *GPAM* en el cromosoma 10 fue replicada en la cohorte de amerindios, presentando un OR=0.71; IC 95% [0.54-0.95]; P=0.021. Se observó la replicación también de la asociación de riesgo para el fenotipo de HTG de la variante rs3825041 en el gen *BUD13* localizado en el cromosoma 11 (OR=1.79; 95% CI [1.36-2.36]; P=2.7E-05). De manera similar, la variante rs1532625 del gen *CETP* localizado en el cromosoma 16 se encontró asociada a protección para el fenotipo de Hipo HDL-C (OR=0.61; 95% CI [0.47-0.81]; P=2.15E-04) en la cohorte de replicación (Tabla 9).

Por otro lado, las variantes rs1367117 en el gen *APOB*, rs11820589 y rs10488698 en el gen *BUD13*, y rs186754194 en el gen *ANGPTL4* no mantuvieron su asociación en la cohorte de replicación.

Tabla 9 Resultados de las SNV replicadas

CHR	Gen	rsID	Rasgo	Alelos	Anotación	Cohorte	MAF		OR	[95% CI]	P
							F	A/U			
10	<i>GPAM</i>	rs2792751	Hipo HDL-C	C/T	No sinónima	MEZ	0.21/0.29	0.65	[0.52 – 0.82]	3.42E-04	
						MAIS	0.24/0.30	0.71	[0.54 – 0.95]	0.021	
11	<i>APOB</i>	rs1367117	HTC	G/A	No sinónima	MEZ	0.31/0.27	1.34	[1.09 – 1.66]	5.64E-03	
						MAIS	0.34/0.29	1.22	[0.93 – 1.6]	0.14	
11	<i>BUD13</i>	rs11820589	HTG	G/A	No sinónima	MEZ	0.19/0.14	1.45	[1.10 – 1.89]	5.48E-03	
						MAIS	0.22/0.21	1.02	[0.77 – 1.34]	0.87	
		rs3825041	HTG	C/T	intrón	MEZ	0.23/0.15	1.71	[1.32 – 2.22]	4.94E-05	
						MAIS	0.28/0.18	1.79	[1.36 – 2.36]	2.70E-05	
rs10488698	Hipo HDL-C	G/A	No sinónima	MEZ	0.13/0.2	0.51	[0.39 – 0.68]	1.59E-06			
				MAIS	0.17/0.21	0.73	[0.53 – 1.006]	0.054			
16	<i>CETP</i>	rs1532625	Hipo HDL-C	C/T	región de <i>splicing</i>	MEZ	0.37/0.46	0.65	[0.53 – 0.81]	6.78E-05	
						MAIS	0.32/0.44	0.61	[0.47 – 0.79]	2.15E-04	
19	<i>ANGPTL4</i>	rs186754194	HTG	G/C	No sinónima	MEZ	0.03/0.05	0.54	[0.34 – 0.85]	8.19E-03	
						MAIS	0.06/0.08	0.72	[0.47 – 1.10]	0.13	

MEZ: Cohorte de descubrimiento, MAIS: Amerindios Mexicanos, MAF: Frecuencia de alelo menor; FA: frecuencia en individuos afectados; FU: Frecuencia en individuos no afectados. Hipo HDL-C: Hipo α -lipoproteinemia, HTC: Hipercolesterolemia, HTG: Hipertrigliceridemia

El análisis de las frecuencias mostró que las variantes rs2792751 de *GPAM*, rs11820589 en *BUD13*, rs1532625 en *CETP* y rs186754194 en *ANGPTL4*, presentan una diferencia significativa entre mestizos e indígenas (P<0.05, Tabla 10).

Al comparar las frecuencias de la cohorte de descubrimiento contra las poblaciones parentales de referencia del proyecto 1000 genomas observamos que no existe diferencias significativas contra la MXL, pero que si las hay comparando contra la población IBS para las tres variantes de *BUD13* y la de *ANGPTL4*. Así mismo, se encontró que sólo la variante

de *BUD13* rs11820589 no presenta diferencia significativa en su frecuencia al comparar con la población YRI, ya que el resto de las variantes incluidas en el análisis de frecuencias difieren contra esta población.

Tabla 10 Análisis de frecuencias de las variantes replicadas entre la población de descubrimiento y otras poblaciones

Gen	Variante	MEZ	MAIS	MXL	IBS	YRI
		n=968	n=671	n= 64	n= 107	n= 108
<i>GPAM</i>	rs2792751_T	0.24	0.29*	0.31	0.28	0.03*
<i>APOB</i>	rs1367117_A	0.28	0.28	0.26	0.29	0.06*
<i>BUD13</i>	rs11820589_A	0.17	0.22*	0.11	0.08*	0.20
	rs3825041_T	0.21	0.24	0.17	0.10*	0.00*
	rs10488698_A	0.15	0.18	0.12	0.06*	0.00*
<i>CETP</i>	rs1532625_T	0.40	0.35*	0.40	0.42	0.06*
<i>ANGPTL4</i>	rs186754194_C	0.04	0.071*	0.03	0.00*	0.00*

MEZ: Cohorte de descubrimiento, MAIS: Cohorte de replicación, MXL: mexicanos residentes de los Ángeles, IBS: Población ibérica en España, YRI= Yorubas de África

*= $P < 0.05$ en la comparación contra la población de descubrimiento. Los valores de P fueron determinados a partir de la prueba de χ^2

13. Discusión

Las dislipidemias son condiciones metabólicas que se caracterizan por la alteración de los niveles de lípidos en sangre. Estas patologías han emergido como uno de los desafíos en salud más apremiantes a nivel global. Su relevancia radica en su profunda conexión con enfermedades cardiovasculares, las cuales ocupan uno de los lugares como principales causas de morbilidad y mortalidad en todo el mundo (Rodríguez et al., 2014).

Las prevalencias de las dislipidemias observadas en las poblaciones de estudio de este trabajo, al igual que con lo observado con otras enfermedades metabólicas, fueron más altas a lo previamente reportado por la ENSANUT (Ensanut, 2012), y es consistente con publicaciones previas donde se reporta que los grupos de origen amerindio presentan una mayor prevalencia de enfermedades metabólicas si se comparan con otros de distinto origen étnico, aun si comparten el mismo espacio sociodemográfico (Mendoza-Caamal et al., 2020). Una comprensión profunda de la base genética y de las interacciones complejas que influyen en el desarrollo de estas entidades, son algunos de los motivantes para el estudio de este grupo de enfermedades metabólicas. En este contexto, la investigación exhaustiva y rigurosa de las asociaciones genéticas y los factores subyacentes a estas condiciones se convierte en un pilar fundamental para entender su etiología y su fisiopatología y brindar perspectivas esenciales para incidir en su prevención, diagnóstico y tratamiento mediante una medicina de precisión (Xue et al., 2012).

Se estima que dentro del genoma humano se presenta una SNV cada 300 pb (Ke et al., 2008). Uno de los objetivos de este trabajo comprendió el describir la variabilidad genética de los genes candidato asociados con la dislipidemia, de esta manera, la información contenida en el catálogo generado (Tabla suplementaria 2) puede servir para identificar los genes con un alto grado de variabilidad, y aquellos que presentan una mayor cantidad de variantes que no han sido descritas previamente, lo que puede ayudar en la toma de decisiones clínicas.

Han sido múltiples los esfuerzos a nivel mundial por tener una comprensión más completa de la genética de los trastornos lipídicos, no obstante, la mayoría de los estudios genómicos realizados presentan factores limitantes, como la subrepresentación de ciertas poblaciones, y el tipo de tecnología utilizada dentro de los estudios de asociación genética (Martin et al., 2017; Sirugo et al., 2019). Por ejemplo, algunos esfuerzos por entender la genética de las enfermedades metabólicas en poblaciones étnicamente heterogéneas, han estudiado muestras de individuos mestizos cuya ancestría amerindia oscila entre el 49-66 % (Aguilar-Salinas et al., 2014; Estrada et al., 2014; Weissglas-Volkov et al., 2013), y han empleado tecnologías como los microarreglos, que son pruebas de genotipificación en principio más económicas, pero con un alcance pobre en nuestra población, ya que contienen un número de variantes limitado y todas ellas son de origen europeo y asiático. Actualmente, gracias a los avances tecnológicos y a la capacidad de reunir grandes colecciones de pacientes, ahora somos capaces de identificar variantes asociadas a cualquier enfermedad y que son propias de cada una de las poblaciones. El incremento de este tipo de estudios en poblaciones multiétnicas permite profundizar en el conocimiento de la fisiopatología y la etiología de estas entidades, para incidir en su prevención, diagnóstico y tratamiento mediante una medicina de precisión.

En este trabajo, se observó que más del 50% de las variantes presentes en la base de datos del exoma del grupo de genes candidato analizado en este trabajo, fueron variantes genéticas que sólo se encuentran en un individuo o en un grupo muy pequeño de individuos y que se denominan Singletons. Este tipo de variantes potencialmente tiene mayor efecto funcional (variantes no sinónimas, en sitios de *splicing* o regiones 3'/5'UTR) (Ke et al., 2008).

En nuestro estudio, se logró replicar diversas asociaciones con diferentes rasgos lipídicos que habían sido previamente reportados en otros grupos étnicos. El ejemplo más claro se observó en la asociación de las variantes localizadas en el grupo de genes *APOA1/C3/A4/A5-ZPR1-BUD13*, como lo son: la SNV no sinónima rs2072560 en *APOA5*, la variante intrónica rs2070665 en *APOA1* y las SNVs rs5128 y rs651821, respectivamente localizadas en regiones reguladoras de *APOC3* y *APOA5*, todas ellas asociadas con HTG en nuestra población, al igual que en poblaciones europeas, asiáticas y africanas (Aung et al., 2014; Feng et al., 2016; Fu et al., 2015; Jasim et al., 2018; Ken-Dror et al., 2010; Song et al., 2015; Zhou et al., 2013).

APOC3 es reconocido como un elemento clave en la regulación de los niveles plasmáticos de TG, su función consiste en reducir la degradación de los restos de lipoproteínas ricas mediante la inhibición de la acción de la *LPL* (Chan et al., 2008). Con base en esto, una de las estrategias terapéuticas ha sido la terapia celular, bloqueando el mRNA de esta proteína en pacientes que cursan con HTG, mediante el uso de un oligonucleótido anti sentido (Gaudet et al., 2015).

Por otra parte, la asociación del SNV no sinónimo rs1367117 en *APOB* con HTC, previamente reportada en poblaciones de ascendencia europea (Y. Lu et al., 2010), también fue replicada en la población de descubrimiento. Otras asociaciones replicadas en nuestra población fueron entre los SNV no sinónimo rs10488698 en *BUD13* y rs9282541 en *ABCA1* con el fenotipo Hipo HDL-C, consistente con lo reportado en poblaciones latinoamericanas o asiáticas (Acuña-Alonzo et al., 2010; Zhang et al., 2017).

Del estudio de asociación con las variantes comunes, en la Figura 16 se presenta un resumen donde se observan los genes y las variantes con un posible efecto funcional que se relacionan con dos componentes lipídicos diferentes. Aquí resulta interesante señalar las variantes en *SNX13* rs2723501, *CMIP* rs4888167 y *JMJD1C* rs10740107 las cuales se asocian tanto a HTC como a hiper-LDL-C.

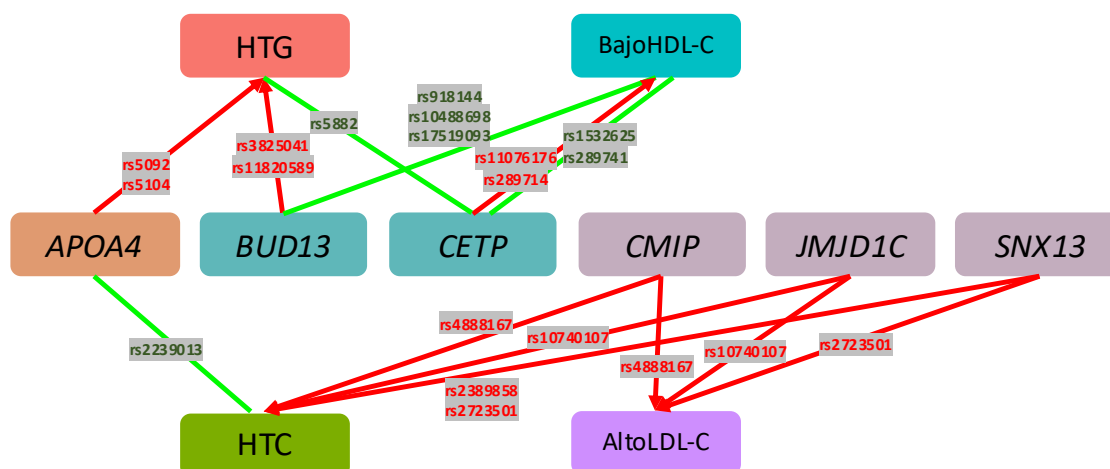


Figura 16 Genes asociados a múltiples componentes lipídicos
Las líneas rojas representan riesgo y las líneas verdes protección hacia los fenotipos.

Entre los resultados de la primera fase también observamos variantes funcionales asociadas con la protección a HTG, como la rs1260326 (P446L) en el gen *GCKR* (Tabla 4), el cuál promueve la captación y movilización de la glucosa, la síntesis de triglicéridos de novo y al mismo tiempo atenúa la oxidación de ácidos grasos (Ramamamy, 2014). La asociación de esta variante había sido descrita previamente en otros estudios en población mestiza (Weissglas-Volkov et al., 2013; Weissglas-Volkov & Pajukanta, 2010) y población amerindia mexicana (Cid-Soto et al., 2018). En este último estudio, los autores reportan que la frecuencia de esta variante en el grupo amerindio no presenta diferencias respecto al grupo mestizo, pero de manera interesante en los indígenas (cohorte MAIS) no se mantiene la asociación con la protección a HTG, demostrando que la asociación de las variantes depende del origen étnico de las poblaciones estudiadas. Así, dentro de las variantes asociadas a dislipidemias en nuestra población de descubrimiento (mestiza), se seleccionaron siete para llevar a cabo la replicación, observando que la variante rs3825041 en *BUDI3* mostró niveles de asociación con el rasgo de HTG más alto en la población indígena que en mestizos (Tabla 9). Otras dos variantes evaluadas en población indígena mantuvieron su asociación con la protección para Hipo HDL-C (rs2792751 en *GPAM* OR=0.71, IC95% [0.54–0.95] $P=0.021$ y rs1532625 en *CETP* OR= 0.61, IC95% [0.47–0.79] $P=2.15E-04$), mientras que la variante rs10488698 en *BUDI3* sólo mostró una tendencia (OR=0.73, IC95% [0.53–1.006] $P=0.054$). Los resultados obtenidos con el resto de las variantes sugieren que éstas se asocian con las dislipidemias de una forma dependiente al origen étnico.

Las frecuencias alélicas de las variantes cuya asociación fue replicada en la cohorte MAIS, presentaron diferencias entre mestizos y amerindios, sólo para *BUDI3* rs3825041 y *GPAM* rs2792751 (Tabla 10). La comparación de las frecuencias con las super poblaciones del proyecto 1000 genomas (Figura 17), muestra que la variante *BUDI3* rs3825041 se encuentra

enriquecida en grupos americanos y de manera importante en la población mexicana (MEZ y MAIS) (Figura 17).

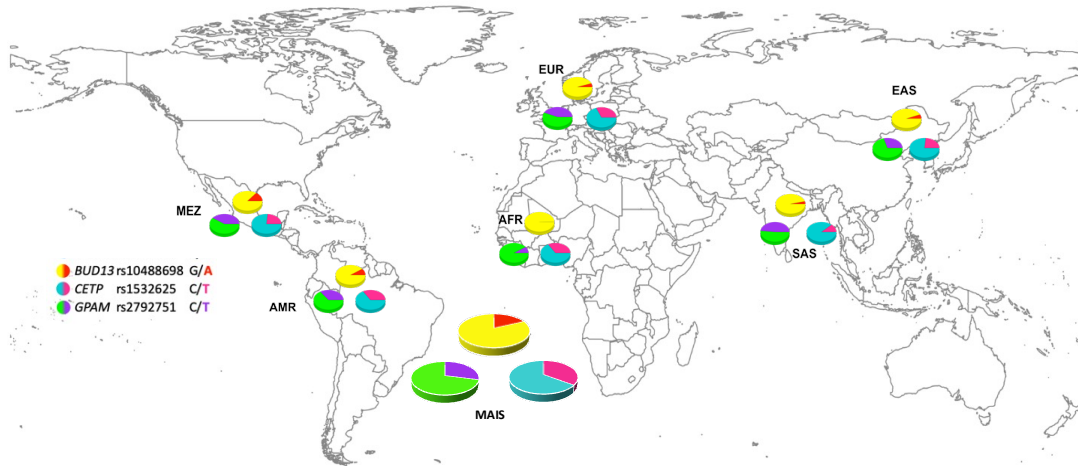


Figura 17 Distribución geográfica de las variantes replicadas

Se presentan las frecuencias alélicas de las variantes de las super poblaciones de referencia del proyecto 1000 genomas agrupadas por origen étnico. AFR: africano; AMR: americanos; EUR: europeo; MAIS: Cohorte de replicación; MEZ: cohorte de descubrimiento, SAS: asiático del sur; EAS: asiático oriental

El alelo derivado T de la variante rs1523625 en *CETP* cuya asociación fue replicada en la cohorte MAIS se encuentra en menor frecuencia en el grupo de asiáticos del sur, mientras que en las otras poblaciones presenta una frecuencia similar a la de nuestras cohortes de estudio, en línea con los hallazgos presentados un estudio de metaanálisis que involucró cohortes hispanas (C. Gao et al., 2018).

Por otra parte, múltiples variantes cercanas mostraron asociación con el mismo tipo de dislipidemia (Tabla 4), sugiriendo la presencia de haplotipos asociados. Existen evidencias que muestran el papel de haplotipos/variantes raras dentro de la susceptibilidad a las enfermedades, por lo que estos no pueden ser ignorados en los estudios genéticos de las enfermedades complejas (Liu et al., 2005). En este estudio, uno de los haplotipos con mayor interés se localizó en el gen *HNF1A*. El haplotipo conteniendo las variantes no sinónimas rs2464196 (S487N) y rs1169288 (I27L), en conjunto con la variante rs2464195 localizada dentro de una región de *splicing*, se encontró asociado con Hipo HDL-C. Cabe aclarar que este haplotipo no contiene la variante rs483353044, descrita previamente como propia de la población mexicana y asociada a DT2 (Estrada et al., 2014). Aún se desconoce el efecto funcional de este haplotipo y sus implicaciones en la fisiopatología de las dislipidemias. Otros estudios en población mexicana se han encontrado haplotipos altamente frecuentes en población mexicana, que se encuentran asociados con entidades metabólicas, como el haplotipo en *SLC16A11*, asociado a DT2 (Williams Amy et al., 2014b). Este gen es un transportador de mono carboxilatos acoplado a protones. Es importante mencionar que estudios funcionales mostraron que la presencia de este haplotipo disminuye la expresión y

la adecuada localización del transportador en la superficie celular a nivel hepático, disminuyendo la dinámica del metabolismo de lípidos y ácidos grasos libres (Rusu et al., 2017). De igual manera, se pudo observar concordancia con los resultados de la asociación de haplotipos para *APOA5*, para la HTG con otro trabajo del grupo de investigación donde se describió la asociación del Haplotipo **GTC** hacia HTG con un OR=1.57, IC95% [1.34–1.83] $P=1.34E-08$; Beta=27.4, IC95% [18.20–36.60] $P=6.28E-09$], muy similares a lo descrito en este trabajo (Tabla 6) (Cid-Soto, 2023)

Sin embargo, es importante destacar que, si bien el efecto de las variantes comunes sobre un fenotipo de interés puede detectarse evaluando los efectos marginales con relativa confianza, el panorama se complica cuando se trata de variantes raras. Las variantes raras, debido a su baja frecuencia en la población, a menudo sufren de una alta tasa de falsos negativos en los análisis y pueden pasar desapercibidas si no se utiliza una estrategia de muestreo adecuada (Goswami et al., 2021). Los algoritmos tradicionalmente utilizados para identificar asociaciones genéticas con variantes comunes pueden no responder de manera efectiva para variantes raras, debido a su baja frecuencia y al sesgo que puede introducirse en los resultados por la presencia de variantes raras en individuos con fenotipos extremos (Barnett et al., 2013). Cómo era de esperarse, usando los análisis de SKAT/SKAT-O y de colapso, se pudo observar asociación a nivel de gen por ejemplo para los genes *BUD13* $P_{SKAT}=4.22E-03$, $P_{SKAT-O}=4.13E-03$ y $P_{BURDEN}=7.18E-03$ y *CETP* $P_{SKAT}=3.27E-01$, $P_{SKAT-O}=1.90E-02$ y $P_{BURDEN}=9.92E-03$, que también fueron asociados por medio del estudio con las variantes comunes. Si el número de variantes genéticas causales en un locus genético es muy grande y cada variante contribuye con una pequeña cantidad a los rasgos, utilizar las pruebas de SKAT/SKAT-O proporciona resultados más confiables que modelos fijos, además las pruebas de análisis basadas en un gen presentan varias ventajas, como el incremento del poder estadístico (Chiu et al., 2017).

Interesantemente, el gen *PLEC* se asoció por el análisis cuantitativo a todas las dislipidemias, pero cada una de las asociaciones era con un SNV diferente; sin embargo, los resultados presentados en la Tabla 8 permiten observar que múltiples SNVs presentaron los mismos valores de asociación genética al hacer el estudio de las variantes individuales, considerando el conjunto total de variantes comunes y variantes raras. Esto puede ser debido a un efecto de sesgo por la presencia de dichas variantes en individuos con fenotipos extremos, ya que una MAF de 0.000516 implicó la presencia de un sólo alelo del SNV dentro de nuestra cohorte de descubrimiento. Este efecto no se observó para los casos de los niveles de HDL-C donde se pudo documentar también la asociación para las variantes del gen *BUD13*, aquí descritas como asociadas a Hipo HDL-C.

Interesantemente, en los genes candidatos seleccionados por presentar previas asociaciones con dislipidemia en otras poblaciones, se identificaron variantes nuevas que presentan las asociaciones más sólidas con los perfiles lipídicos en individuos de origen mexicano. Estas diferencias apoyan la noción de la diversidad genética en los loci relacionados con los lípidos

y que estos pueden ser específicos de las poblaciones. Es importante destacar que las nuevas asociaciones de variantes en genes previamente vinculados con dislipidemia, subrayan la relevancia de incluir todas las poblaciones en los estudios de asociación genómica, ya que las variantes asociadas pueden variar entre diferentes grupos poblacionales. Además, es notable que muchas de las variantes de riesgo previamente identificadas en relación con los perfiles lipídicos en diversas poblaciones étnicas, como las europeas, asiáticas y africanas, se han replicado en nuestro estudio. En conjunto, nuestros hallazgos indican que la base genética de las dislipidemias es compartida parcialmente entre grupos étnicos diversos.

14. Conclusión

Los resultados derivados de este trabajo contribuyen a la caracterización de las variantes asociadas con las dislipidemias. Es importante destacar que en este estudio se documenta la presencia de variantes nuevas en la población mexicana localizadas en genes previamente asociados con dislipidemias en otras poblaciones. Nuestra investigación mostró una fuerte asociación entre estas variantes propias de individuos mexicanos y rasgos lipídicos. Estas divergencias subrayan la alta heterogeneidad alélica presente en los loci relacionados con los lípidos en diversas poblaciones. Es digno de mención que nuestro trabajo replicó varias de las variantes de riesgo previamente asociadas con dislipidemias en distintos grupos étnicos, incluyendo poblaciones europeas, asiáticas y africanas.

En conjunto, estos hallazgos respaldan la idea de que existe una arquitectura genética compartida parcialmente entre diferentes grupos étnicos, resaltando así la necesidad de realizar análisis detallados en poblaciones específicas. Aunque se observa una convergencia en las asociaciones genéticas, se reconoce que estas pueden diferir significativamente según el origen étnico. Este enfoque de investigación no solo ofrece una perspectiva esencial para comprender mejor las bases genéticas de las enfermedades metabólicas, sino que también subraya la importancia de adaptar estrategias médicas y terapéuticas de manera más precisa y efectiva, considerando las variaciones específicas de cada población. El aumento de este tipo de estudios en poblaciones multiétnicas no solo respalda una comprensión más profunda de la fisiopatología y etiología de estas entidades, sino que también contribuye significativamente a su prevención, diagnóstico y tratamiento mediante una medicina de precisión. Este enfoque holístico, que abarca diversas poblaciones étnicas, es esencial para desarrollar estrategias de salud más efectivas y personalizadas.

15. Referencias

- Acuña-Alonzo, V., Flores-Dorantes, T., Kruit, J. K., Villarreal-Molina, T., Arellano-Campos, O., Hünemeier, T., Moreno-Estrada, A., Ortiz-López, M. G., Villamil-Ramírez, H., León-Mimila, P., Villalobos-Comparan, M., Jacobo-Albavera, L., Ramírez-Jiménez, S., Sikora, M., Zhang, L. H., Pape, T. D., de Ángeles Granados-Silvestre, M., Montufar-Robles, I., Tito-Alvarez, A. M., ... Canizales-Quinteros, S. (2010). A functional ABCA1 gene variant is associated with low HDL-cholesterol levels and shows evidence of positive selection in Native Americans. *Human Molecular Genetics*, *19*(14), 2877–2885. <https://doi.org/10.1093/hmg/ddq173>
- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nat Methods*, *7*(4), 248–249. <https://doi.org/10.1038/nmeth0410-248>
- Aguilar-Salinas, C. A., Gómez-Pérez, F. J., Rull, J., Villalpando, S., Barquera, S., & Rojas, R. (2010). Prevalencia de las dislipidemias en la Encuesta Nacional de Salud y Nutrición 2006. *Salud Publica de Mexico*, *52*(SUPPL.1), S44–S53. <https://doi.org/10.1590/S0036-36342010000700008>
- Aguilar-Salinas, C. A., Tusie-Luna, T., & Pajukanta, P. (2014). Genetic and environmental determinants of the susceptibility of Amerindian derived populations for having hypertriglyceridemia. *Metabolism: Clinical and Experimental*, *63*(7), 887–894. <https://doi.org/10.1016/j.metabol.2014.03.012>
- Alberto Aguilar Salinas, C., Javier Gómez Pérez, F., Lerman Garber, I., Vázquez Chávez, C., Pérez Méndez, Ó., Posadas Romero, C., Alberto Aguilar Salinas Instituto Nacional de Ciencias Médicas Nutrición, C., & Zubirán, S. (2004). Posición de la Sociedad Mexicana de Nutrición y Endocrinología Diagnóstico y tratamiento de las dislipidemias: posición de la Sociedad Mexicana de Nutrición y Endocrinología. *Revista de Endocrinología y Nutrición*, *12*(1), 7–41. <http://new.medigraphic.com/cgi-bin/resumen.cgi?IDARTICULO=406>
- Altshuler, D. L., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Collins, F. S., De La Vega, F. M., Donnelly, P., Egholm, M., Flicek, P., Gabriel, S. B., Gibbs, R. A., Knoppers, B. M., Lander, E. S., Lehrach, H., Mardis, E. R., McVean, G. A., Nickerson, D. A., ... Peterson, J. L. (2010). A map of human genome variation from population-scale sequencing. *Nature*, *467*(7319). <https://doi.org/10.1038/nature09534>
- Aung, L. H. H., Yin, R. X., Wu, J. Z., Wu, D. F., Wang, W., & Li, H. (2014). Association between the MLX interacting protein-like, BUD13 homolog and zinc finger protein 259 gene polymorphisms and serum lipid levels. *Scientific Reports*, *4*, 1–12. <https://doi.org/10.1038/srep05565>
- Bai, W., Kou, C., Zhang, L., You, Y., Yu, W., Hua, W., Li, Y., Yu, Y., Zhao, T., & Wu, Y. (2019). Functional polymorphisms of the APOA1/C3/A4/A5-ZPR1-BUD13 gene cluster are associated with dyslipidemia in a sex-specific pattern. *PeerJ*, *2019*(1), 1–15. <https://doi.org/10.7717/peerj.6175>

- Barnett, I. J., Lee, S., & Lin, X. (2013). Detecting Rare Variant Effects Using Extreme Phenotype Sampling in Sequencing Association Studies. *Genetic Epidemiology*, 37(2). <https://doi.org/10.1002/gepi.21699>
- Barquera, S., & Campos, I. (2009). Dislipidemias: epidemiología, evaluación, adherencia y tratamiento. En C. Oropeza (Ed.), *Instituto Nacional de Salud Pública* (1a ed., Vol. 1). Instituto Nacional de Salud Pública. https://www.insp.mx/images/stories/Produccion/pdf/110311_cp1.pdf
- Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21(2), 263–265. <https://doi.org/10.1093/bioinformatics/bth457>
- Bolio Galvis, A. (2014). Transcultural Diabetes Nutrition Algorithm: The Mexican Application. *Journal of Diabetes & Metabolism*, 05(09). <https://doi.org/10.4172/2155-6156.1000423>
- Chan, D. C., Chen, M. M., Ooi, E. M. M., & Watts, G. F. (2008). An ABC of apolipoprotein C-III: A clinically useful new cardiovascular risk factor? En *International Journal of Clinical Practice* (Vol. 62, Número 5). <https://doi.org/10.1111/j.1742-1241.2007.01678.x>
- Chiu, C. Y., Jung, J., Wang, Y., Weeks, D. E., Wilson, A. F., Bailey-Wilson, J. E., Amos, C. I., Mills, J. L., Boehnke, M., Xiong, M., & Fan, R. (2017). A comparison study of multivariate fixed models and Gene Association with Multiple Traits (GAMuT) for next-generation sequencing. *Genetic Epidemiology*, 41(1), 18–34. <https://doi.org/10.1002/gepi.22014>
- Cid-Soto, M. A. (2023). *VARIANTES ALÉLICAS DE LA POBLACIÓN INDÍGENA QUE CONTRIBUYEN AL DESARROLLO DEL SÍNDROME METABÓLICO EN LA POBLACIÓN MEXICANA*. UNAM.
- Cid-Soto, M. A., Martínez-Hernández, A., García-Ortíz, H., Córdova, E. J., Barajas-Olmos, F., Centeno-Cruz, F., Contreras-Cubas, C., Mendoza-Caamal, E. C., Ciceron-Arellano, I., Morales-Rivera, M. I., Jimenez-Ruiz, J. L., Salas-Martínez, G., Saldaña-Álvarez, Y., Revilla-Monsalve, C., Islas-Andrade, S., & Orozco, L. (2018). Gene variants in AKT1, GCKR and SOCS3 are differentially associated with metabolic traits in Mexican Amerindians and Mestizos. *Gene*, 679(4809), 160–171. <https://doi.org/10.1016/j.gene.2018.08.076>
- Consejo Nacional para la Población. (2016). *Hablantes De Lengua Indígena*. https://www.gob.mx/cms/uploads/attachment/file/121653/Infografia_INDI_FINAL_08082016.pdf
- Davis, H. R., & Altmann, S. W. (2009). Niemann-Pick C1 Like 1 (NPC1L1) an intestinal sterol transporter. En *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* (Vol. 1791, Número 7, pp. 679–683). <https://doi.org/10.1016/j.bbalip.2009.01.002>
- Dennis Kasper, Anthony Fauci, Stephen Hauser, Dan Longo, J. Larry Jameson, J. L. (2018). Harrison. Principios de Medicina Interna, 19e | AccessMedicina | McGraw-Hill Medical. En *Harrison*.
- Ensanut. (2012). *Ensanut 2012. Resultados Nacionales*. Instituto Nacional de Salud Pública.

<https://ensanut.insp.mx/encuestas/ensanut2012/doctos/informes/ENSANUT2012ResultadosNacionales.pdf>

- Estrada, K., Aukrust, I., Bjørkhaug, L., Burt, N. P., Mercader, J. M., García-Ortiz, H., Huerta-Chagoya, A., Moreno-Macías, H., Walford, G., Flannick, J., Williams, A. L., Gómez-Vázquez, M. J., Fernandez-Lopez, J. C., Martínez-Hernández, A., Jiménez-Morales, S., Centeno-Cruz, F., Mendoza-Caamal, E., Revilla-Monsalve, C., Islas-Andrade, S., ... MacArthur, D. G. (2014). Association of a low-frequency variant in HNF1A with type 2 diabetes in a latino population the SIGMA Type 2 Diabetes Consortium. *JAMA - Journal of the American Medical Association*, *311*(22), 2305–2314. <https://doi.org/10.1001/jama.2014.6511>
- Fan, R., Wang, Y., Mills, J. L., Carter, T. C., Lobach, I., Wilson, A. F., Bailey-Wilson, J. E., Weeks, D. E., & Xiong, M. (2014). Generalized functional linear models for gene-based case-control association studies. *Genetic Epidemiology*, *38*(7). <https://doi.org/10.1002/gepi.21840>
- Feng, D. W., Ma, R. L., Guo, H., He, J., Yan, Y. Z., Muratbek, Niu, Q., Li, S. G., Rui, D. S., Sun, F., Zhang, M., Zhang, J. Y., Ding, Y. S., Liu, J. M., Wang, K., & Guo, S. X. (2016). Association of APOA1 gene polymorphisms (rs670, rs5069, and rs2070665) with dyslipidemia in the Kazakhs of Xinjiang. *Genetics and Molecular Research*, *15*(2). <https://doi.org/10.4238/gmr.15028094>
- Fu, Q., Tang, X., Chen, J., Su, L., Zhang, M., Wang, L., Jing, J., & Zhou, L. (2015). Effects of polymorphisms in APOA4-APOA5- ZNF259-BUD13 gene cluster on plasma levels of triglycerides and risk of coronary heart disease in a Chinese Han population. *PLoS ONE*, *10*(9). <https://doi.org/10.1371/journal.pone.0138652>
- Gao, C., Tabb, K. L., Dimitrov, L. M., Taylor, K. D., Wang, N., Guo, X., Long, J., Rotter, J. I., Watanabe, R. M., Curran, J. E., Blangero, J., Langefeld, C. D., Bowden, D. W., & Palmer, N. D. (2018). Exome Sequencing Identifies Genetic Variants Associated with Circulating Lipid Levels in Mexican Americans: The Insulin Resistance Atherosclerosis Family Study (IRASFS). *Scientific Reports*, *8*(1). <https://doi.org/10.1038/s41598-018-23727-2>
- Gao, H., Wang, H., Shan, G., Liu, R., Chen, H., Sun, S., & Liu, Y. (2021). Prevalence of dyslipidemia and associated risk factors among adult residents of Shenmu City, China. *PLoS ONE*, *16*(5 May). <https://doi.org/10.1371/journal.pone.0250573>
- Gaudet, D., Alexander, V. J., Baker, B. F., Brisson, D., Tremblay, K., Singleton, W., Geary, R. S., Hughes, S. G., Viney, N. J., Graham, M. J., Crooke, R. M., Witztum, J. L., Brunzell, J. D., & Kastelein, J. J. P. (2015). Antisense Inhibition of Apolipoprotein C-III in Patients with Hypertriglyceridemia. *New England Journal of Medicine*, *373*(5). <https://doi.org/10.1056/nejmoa1400283>
- Goswami, C., Chattopadhyay, A., & Chuang, E. Y. (2021). Rare variants: data types and analysis strategies. *Annals of Translational Medicine*, *9*(12). <https://doi.org/10.21037/atm-21-1635>
- Griffiths-Jones, S., Saini, H. K., Van Dongen, S., & Enright, A. J. (2008). miRBase: Tools for microRNA genomics. *Nucleic Acids Research*, *36*(SUPPL. 1). <https://doi.org/10.1093/nar/gkm952>

- Helgerud, P., Petersen, L. B., & Norum, K. R. (1982). Acyl CoA:Retinol acyltransferase in rat small intestine: Its activity and some properties of the enzymic reaction. *Journal of Lipid Research*, 23(4), 609–618.
- Herbeth, B., Samara, A., Ndiaye, C., Marteau, J. B., Berrahmoune, H., Siest, G., & Visvikis-Siest, S. (2010). Metabolic syndrome-related composite factors over 5 years in the STANISLAS Family Study: Genetic heritability and common environmental influences. *Clinica Chimica Acta*, 411(11–12), 833–839. <https://doi.org/10.1016/j.cca.2010.02.070>
- Innerarity, T. L., Young, S. G., Poksay, K. S., Mahley, R. W., Smith, R. S., Milne, R. W., Marcel, Y. L., & Weisgraber, K. H. (1987). Structural relationship of human apolipoprotein B48 to apolipoprotein B100. *Journal of Clinical Investigation*, 80(6), 1794–1798. <https://doi.org/10.1172/JCI113273>
- Jasim, A. A., Al-Bustan, S. A., Al-Kandari, W., Al-Serri, A., & AlAskar, H. (2018). Sequence analysis of APOA5 among the Kuwaiti population identifies association of rs2072560, rs2266788, and rs662799 with TG and VLDL levels. *Frontiers in Genetics*, 9(APR). <https://doi.org/10.3389/fgene.2018.00112>
- Jiang, X. C., & Zhou, H. W. (2006). Plasma lipid transfer proteins. *Current Opinion in Lipidology*, 17(3), 302–308. <https://doi.org/10.1016/j.aca.2005.10.079>
- Ke, X., Taylor, M. S., & Cardon, L. R. (2008). Singleton SNPs in the human genome and implications for genome-wide association studies. *European Journal of Human Genetics*, 16(4), 506–515. <https://doi.org/10.1038/sj.ejhg.5201987>
- Ken-Dror, G., Goldbourt, U., & Dankner, R. (2010). Different effects of apolipoprotein A5 SNPs and haplotypes on triglyceride concentration in three ethnic origins. *Journal of Human Genetics*, 55(5). <https://doi.org/10.1038/jhg.2010.27>
- Khor, C. C., & Goh, D. L. M. (2010). Strategies for identifying the genetic basis of dyslipidemia: Genome-wide association studies vs. the resequencing of extremes. *Current Opinion in Lipidology*, 21(2), 123–127. <https://doi.org/10.1097/MOL.0b013e328336eae9>
- Ko, A., Cantor, R. M., Weissglas-Volkov, D., Nikkola, E., Reddy, P. M. V. L., Sinsheimer, J. S., Pasaniuc, B., Brown, R., Alvarez, M., Rodriguez, A., Rodriguez-Guillen, R., Bautista, I. C., Arellano-Campos, O., Muñoz-Hernández, L. L., Salomaa, V., Kaprio, J., Jula, A., Jauhiainen, M., Heliövaara, M., ... Pajukanta, P. (2014). Amerindian-specific regions under positive selection harbour new lipid variants in Latinos. *Nature Communications*, 5, 3983. <https://doi.org/10.1038/ncomms4983>
- Lee, S., Choi, S., Qiao, D., Cho, M., Silverman, E. K., Park, T., & Won, S. (2018). WISARD: Workbench for integrated superfast association studies for related datasets. *BMC Medical Genomics*, 11. <https://doi.org/10.1186/s12920-018-0345-y>
- Li, M. X., Gui, H. S., Kwan, J. S. H., Bao, S. Y., & Sham, P. C. (2012). A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. *Nucleic Acids Research*, 40(7). <https://doi.org/10.1093/nar/gkr1257>
- Liu, P. Y., Zhang, Y. Y., Lu, Y., Long, J. R., Shen, H., Zhao, L. J., Xu, F. H., Xiao, P., Xiong, D. H., Liu, Y. J., Recker, R. R., & Deng, H. W. (2005). A survey of haplotype variants at several disease candidate genes: The importance of rare variants for complex diseases. En *Journal of Medical Genetics* (Vol. 42, Número 3, pp. 221–227). <https://doi.org/10.1136/jmg.2004.024752>

- Lu, X., Peloso, G. M., Liu, D. J., Wu, Y., Zhang, H., Zhou, W., Li, J., Tang, C. S. M., Dorajoo, R., Li, H., Long, J., Guo, X., Xu, M., Spracklen, C. N., Chen, Y., Liu, X., Zhang, Y., Khor, C. C., Liu, J., ... Willer, C. J. (2017). Exome chip meta-analysis identifies novel loci and East Asian-specific coding variants that contribute to lipid levels and coronary artery disease. *Nature Genetics*, *49*(12), 1722–1730. <https://doi.org/10.1038/ng.3978>
- Lu, Y., Feskens, E. J. M., Boer, J. M. A., Imholz, S., Verschuren, W. M. M., Wijmenga, C., Vaarhorst, A., Slagboom, E., Müller, M., & Dollé, M. E. T. (2010). Exploring genetic determinants of plasma total cholesterol levels and their predictive value in a longitudinal study. *Atherosclerosis*, *213*(1). <https://doi.org/10.1016/j.atherosclerosis.2010.08.053>
- Mahley, R. W., Innerarity, T. L., Rall, S. C., & Weisgraber, K. H. (1984). Plasma lipoproteins: Apolipoprotein structure and function. En *Journal of Lipid Research* (Vol. 25, Número 12, pp. 1277–1294).
- Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., Daly, M. J., Bustamante, C. D., & Kenny, E. E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *American Journal of Human Genetics*, *100*(4), 635–649. <https://doi.org/10.1016/j.ajhg.2017.03.004>
- McLaren, W., Gil, L., Hunt, S., Riat, H. S., Ritchie, G., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *bioRxiv*, 042374. <https://doi.org/10.1101/042374>
- Mendoza-Caamal, E. C., Barajas-Olmos, F., García-Ortiz, H., Cicerón-Arellano, I., Martínez-Hernández, A., Córdova, E. J., Esparza-Aguilar, M., Contreras-Cubas, C., Centeno-Cruz, F., Cid-Soto, M., Morales-Marín, M. E., Reséndiz-Rodríguez, A., Jiménez-Ruiz, J. L., Salas-Martínez, M. G., Saldanña-Alvarez, Y., Mirzaeicheshmeh, E., Rojas-Martínez, M. R., & Orozco, L. (2020). Metabolic syndrome in indigenous communities in Mexico: A descriptive and cross-sectional study. *BMC Public Health*, *20*(1). <https://doi.org/10.1186/s12889-020-8378-5>
- Moreno-Estrada, A., Gignoux, C. R., Fernández-López, J. C., Zakharia, F., Sikora, M., Contreras, A. V., Acuña-Alonzo, V., Sandoval, K., Eng, C., Romero-Hidalgo, S., Ortiz-Tello, P., Robles, V., Kenny, E. E., Nuño-Arana, I., Barquera-Lozano, R., Macín-Pérez, G., Granados-Arriola, J., Huntsman, S., Galanter, J. M., ... Bustamante, C. D. (2014). The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science*, *344*(6189), 1280–1285. <https://doi.org/10.1126/science.1251688>
- Nassoury, N., Blasiolo, D. A., Tebon Oler, A., Benjannet, S., Hamelin, J., Poupon, V., McPherson, P. S., Attie, A. D., Prat, A., & Seidah, N. G. (2007). The cellular trafficking of the secretory proprotein convertase PCSK9 and its dependence on the LDLR. *Traffic*, *8*(6), 718–732. <https://doi.org/10.1111/j.1600-0854.2007.00562.x>
- NEEL, J. V. (1962). Diabetes mellitus: a “thrifty” genotype rendered detrimental by “progress”? *American journal of human genetics*, *14*, 353–362.
- Nghiem-Rao, T. H., Mavis, A., & Patel, S. B. (2014). Cholesterol and Lipoproteins. En *Reference Module in Biomedical Sciences* (Third Edit). Elsevier. <https://doi.org/10.1016/b978-0-12-801238-3.00045-3>

- Noubiap, J. J., Bigna, J. J., Nansseu, J. R., Nyaga, U. F., Balti, E. V., Echouffo-Tcheugui, J. B., & Kengne, A. P. (2018). Prevalence of dyslipidaemia among adults in Africa: a systematic review and meta-analysis. *The Lancet Global Health*, 6(9), e998–e1007. [https://doi.org/10.1016/S2214-109X\(18\)30275-4](https://doi.org/10.1016/S2214-109X(18)30275-4)
- Oram, J. F., & Lawn, R. M. (2001). ABCA1: The gatekeeper for eliminating excess tissue cholesterol. En *Journal of Lipid Research* (Vol. 42, Número 8, pp. 1173–1179).
- Oram, J. F., & Yokoyama, S. (1996). Apolipoprotein-mediated removal of cellular cholesterol and phospholipids. En *Journal of Lipid Research* (Vol. 37, Número 12, pp. 2473–2491).
- Parra, E. J., Mazurek, A., Gignoux, C. R., Sockell, A., Agostino, M., Morris, A. P., Petty, L. E., Hanis, C. L., Cox, N. J., Valladares-Salgado, A., Below, J. E., & Cruz, M. (2017). Admixture mapping in two Mexican samples identifies significant associations of locus ancestry with triglyceride levels in the BUD13/ ZNF259/APOA5 region and fine mapping points to rs964184 as the main driver of the association signal. *PLoS ONE*, 12(2). <https://doi.org/10.1371/journal.pone.0172880>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3). <https://doi.org/10.1086/519795>
- Ramasamy, I. (2014). Recent advances in physiological lipoprotein metabolism. En *Clinical Chemistry and Laboratory Medicine* (Vol. 52, Número 12, pp. 1695–1727). <https://doi.org/10.1515/cclm-2013-0358>
- Ramasamy, I. (2016). Update on the molecular biology of dyslipidemias. *Clinica Chimica Acta*, 454, 143–185. <https://doi.org/10.1016/j.cca.2015.10.033>
- Rodriguez, C. J., Daviglius, M. L., Swett, K., González, H. M., Gallo, L. C., Wassertheil-Smoller, S., Giachello, A. L., Teng, Y., Schneiderman, N., Talavera, G. A., & Kaplan, R. C. (2014). Dyslipidemia patterns among Hispanics/Latinos of diverse background in the United States. *American Journal of Medicine*, 127(12), 1186–1194.e1. <https://doi.org/10.1016/j.amjmed.2014.07.026>
- Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8(1), 103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Rull, J. A., Aguilar-Salinas, C. A., Rojas, R., Rios-Torres, J. M., Gómez-Pérez, F. J., & Olaiz, G. (2005). Epidemiology of type 2 diabetes in Mexico. En *Archives of Medical Research* (Vol. 36, Número 3, pp. 188–196). <https://doi.org/10.1016/j.arcmed.2005.01.006>
- Rusu, V., Hoch, E., Mercader, J. M., Gymrek, M., von Grotthuss, M., Fontanillas, P., Spooner, A., Altshuler, D. M., Florez, J. C., Jacobs, S. B. R., Clish, C. B., Tenen, D. E., Hartigan, C. R., DeRan, M., Guzman, G., Deik, A. A., Pierce, K. A., Dennis, C., Carr, S. A., ... Cortes, M. L. (2017). Type 2 Diabetes Variants Disrupt Function of SLC16A11 through Two Distinct Mechanisms. *Cell*, 170(1). <https://doi.org/10.1016/j.cell.2017.06.011>
- Segrest, J. P., Garber, D. W., Brouillette, C. G., Harvey, S. C., & Anantharamaiah, G. M. (1994). The amphipathic α helix: A multifunctional structural motif in plasma

- apolipoproteins. En *Advances in Protein Chemistry* (Vol. 45, pp. 303–369).
[https://doi.org/10.1016/s0065-3233\(08\)60643-9](https://doi.org/10.1016/s0065-3233(08)60643-9)
- Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). The Missing Diversity in Human Genetic Studies. *Cell*, *177*(1), 26–31. <https://doi.org/10.1016/j.cell.2019.02.048>
- Song, Y., Zhu, L., Richa, M., Li, P., Yang, Y., & Li, S. (2015). Associations of the APOC3 rs5128 polymorphism with plasma APOC3 and lipid levels: A meta-analysis. *Lipids in Health and Disease*, *14*(1). <https://doi.org/10.1186/s12944-015-0027-0>
- Stangl, H., Hyatt, M., & Hobbs, H. H. (1999). Transport of lipids from high and low density lipoproteins via scavenger receptor-BI. *Journal of Biological Chemistry*, *274*(46), 32692–32698. <https://doi.org/10.1074/jbc.274.46.32692>
- Stevens, G., Dias, R. H., Thomas, K. J. A., Rivera, J. A., Carvalho, N., Barquera, S., Hill, K., & Ezzati, M. (2008). Characterizing the epidemiological transition in Mexico: National and subnational burden of diseases, injuries, and risk factors. *PLoS Medicine*, *5*(6), 0900–0910. <https://doi.org/10.1371/journal.pmed.0050125>
- Sulaiman, R. A. (2019). Inherited metabolic disorders and dyslipidaemia. *Journal of Clinical Pathology*, 1–7. <https://doi.org/10.1136/jclinpath-2019-205910>
- Vinet, L., & Zhedanov, A. (2011). A “missing” family of classical orthogonal polynomials. En W. H. Freeman (Ed.), *Journal of Physics A: Mathematical and Theoretical* (6th ed., Vol. 44, Número 8). Macmillan Higher Education. <https://doi.org/10.1088/1751-8113/44/8/085201>
- Vrins, C. L. J. (2010). From blood to gut: Direct secretion of cholesterol via transintestinal cholesterol efflux. En *World Journal of Gastroenterology* (Vol. 16, Número 47, pp. 5953–5957). <https://doi.org/10.3748/wjg.v16.i47.5953>
- Warnick, G. R., Knopp, R. H., Fitzpatrick, V., & Branson, L. (1990). Estimating low-density lipoprotein cholesterol by the Friedewald equation is adequate for classifying patients on the basis of nationally recommended cutpoints. *Clinical Chemistry*, *36*(1), 15–19. <https://doi.org/10.1093/clinchem/36.1.15>
- Weissglas-Volkov, D., Aguilar-Salinas, C. A., Nikkola, E., Deere, K. A., Cruz-Bautista, I., Arellano-Campos, O., Muñoz-Hernandez, L. L., Gomez-Munguia, L., Ordoñez-Sánchez, M. L., Linga Reddy, P. M. V., Lusic, A. J., Matikainen, N., Taskinen, M. R., Riba, L., Cantor, R. M., Sinsheimer, J. S., Tusie-Luna, T., & Pajukanta, P. (2013). Genomic study in Mexicans identifies a new locus for triglycerides and refines European lipid loci. *Journal of Medical Genetics*, *50*(5), 298–308. <https://doi.org/10.1136/jmedgenet-2012-101461>
- Weissglas-Volkov, D., & Pajukanta, P. (2010). Genetic causes of high and low serum HDL-cholesterol. En *Journal of Lipid Research* (Vol. 51, Número 8, pp. 2032–2057). <https://doi.org/10.1194/jlr.R004739>
- Willer, C. J., Schmidt, Ellen M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., Andrea, G., Ingelsson, E., & Abecasis, G. R. (2013). Discovery and refinement of loci associated with lipid levels Global Lipids Genetics Consortium *. *Nature Gene*, *45*(11), 1274–1284. <https://doi.org/10.1038/ng.2797>
- Williams Amy, A. L., Jacobs Suzanne, S. B. R., Moreno-Macías, H., Huerta-Chagoya, A., Churchhouse, C., Márquez-Luna, C., Gómez-Vázquez, M. J., Burt Noël, N. P., Aguilar-Salinas, C. A., González-Villalpando, C., Florez, J. C., Orozco, L., Tusié-Luna, T., Altshuler, D., Ripke, S., Manning, A. K., García-Ortíz, H., Neale, B., Reich,

- D., ... Orozco Lorena, L. (2014a). Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature*, 506(7486), 97–101. <https://doi.org/10.1038/nature12828>
- Williams Amy, A. L., Jacobs Suzanne, S. B. R., Moreno-Macías, H., Huerta-Chagoya, A., Churchhouse, C., Márquez-Luna, C., Gómez-Vázquez, M. J., Burt Noél, N. P., Aguilar-Salinas, C. A., González-Villalpando, C., Florez, J. C., Orozco, L., Tusié-Luna, T., Altshuler, D., Ripke, S., Manning, A. K., García-Ortíz, H., Neale, B., Reich, D., ... Orozco Lorena, L. (2014b). Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature*, 506(7486), 97–101. <https://doi.org/10.1038/nature12828>
- Xue, Y., Chen, Y., Ayub, Q., Huang, N., Ball, E. V., Mort, M., Phillips, A. D., Shaw, K., Stenson, P. D., Cooper, D. N., & Tyler-Smith, C. (2012). Deleterious- and disease-allele prevalence in healthy individuals: Insights from current predictions, mutation databases, and population-scale resequencing. *American Journal of Human Genetics*, 91(6), 1022–1032. <https://doi.org/10.1016/j.ajhg.2012.10.015>
- Yanai, H., & Yoshida, H. (2021). Secondary dyslipidemia: its treatments and association with atherosclerosis. *Global Health & Medicine*, 3(1). <https://doi.org/10.35772/ghm.2020.01078>
- Zhang, L., You, Y., Wu, Y., Zhang, Y., Wang, M., Song, Y., Liu, X., & Kou, C. (2017). Association of BUD13 polymorphisms with metabolic syndrome in Chinese population: A case-control study. *Lipids in Health and Disease*, 16(1), 9–13. <https://doi.org/10.1186/s12944-017-0520-8>
- Zhou, L., He, M., Mo, Z., Wu, C., Yang, H., Yu, D., Yang, X., Zhang, X., Wang, Y., Sun, J., Gao, Y., Tan, A., He, Y., Zhang, H., Qin, X., Zhu, J., Li, H., Lin, X., Zhu, J., ... Wu, T. (2013). A genome wide association study identifies common variants associated with lipid levels in the Chinese population. *PLoS ONE*, 8(12). <https://doi.org/10.1371/journal.pone.0082420>

16. Publicación



OPEN ACCESS

Edited by:

Tommaso Pippucci,
Policlinico Sant'Orsola-Malpighi, Italy

Reviewed by:

Alfonsina Tirozzi,
University of Molise, Italy
Mark Z. Kos,
The University of Texas Rio Grande
Valley, United States

*Correspondence:

Emilio J. Córdova
ecordova@inmegen.gob.mx
Lorena Orozco
lorozco@inmegen.gob.mx

[†]These authors share last authorship

Specialty section:

This article was submitted to
Genetics of Common and Rare
Diseases,
a section of the journal
Frontiers in Genetics

Received: 02 November 2021

Accepted: 25 April 2022

Published: 20 May 2022

Citation:

Jurado-Camacho PA, Cid-Soto MA,
Barajas-Olmos F, García-Ortiz H,
Baca-Peynado P,
Martínez-Hernández A,
Centeno-Cruz F, Contreras-Cubas C,
González-Villalpando ME,
Saldaña-Álvarez Y, Salas-Martínez G,
Mendoza-Caamal EC,
González-Villalpando C, Córdova EJ
and Orozco L (2022) Exome
Sequencing Data Analysis and a Case-
Control Study in Mexican Population
Reveals Lipid Trait Associations of New
and Known Genetic Variants in
Dyslipidemia-Associated Loci.
Front. Genet. 13:807381.
doi: 10.3389/fgene.2022.807381

Exome Sequencing Data Analysis and a Case-Control Study in Mexican Population Reveals Lipid Trait Associations of New and Known Genetic Variants in Dyslipidemia-Associated Loci

Pedro A. Jurado-Camacho^{1,2}, Miguel A. Cid-Soto¹, Francisco Barajas-Olmos¹, Humberto García-Ortiz¹, Paulina Baca-Peynado^{1,2}, Angélica Martínez-Hernández¹, Federico Centeno-Cruz¹, Cecilia Contreras-Cubas¹, María Elena González-Villalpando³, Yolanda Saldaña-Álvarez¹, Guadalupe Salas-Martínez¹, Elvia C. Mendoza-Caamal⁴, Clicerio González-Villalpando³, Emilio J. Córdova^{5†*} and Lorena Orozco^{1†*}

¹Immunogenomics and Metabolic Diseases Laboratory, National Institute of Genomic Medicine, Mexico City, Mexico. ²Posgraduate in Biomedical Sciences, National Autonomous University of Mexico, Mexico City, Mexico. ³Centro de Estudios en Diabetes, Unidad de Investigación en Diabetes y Riesgo Cardiovascular, Centro de Investigación en Salud Poblacional, Instituto Nacional de Salud Pública, Mexico City, Mexico. ⁴Clinical Area, National Institute of Genomic Medicine, Mexico City, Mexico. ⁵Oncogenomics Consortium Laboratory, National Institute of Genomic Medicine, Mexico City, Mexico

Background: Plasma lipid levels are a major risk factor for cardiovascular diseases. Although international efforts have identified a group of loci associated with the risk of dyslipidemia, Latin American populations have been underrepresented in these studies.

Objective: To know the genetic variation occurring in lipid-related loci in the Mexican population and its association with dyslipidemia.

Methods: We searched for single-nucleotide variants in 177 lipid candidate genes using previously published exome sequencing data from 2838 Mexican individuals belonging to three different cohorts. With the extracted variants, we performed a case-control study. Logistic regression and quantitative trait analyses were implemented in PLINK software. We used an LD pruning using a 50-kb sliding window size, a 5-kb window step size and a r^2 threshold of 0.1.

Results: Among the 34251 biallelic variants identified in our sample population, 33% showed low frequency. For case-control study, we selected 2521 variants based on a minor allele frequency $\geq 1\%$ in all datasets. We found 19 variants in 9 genes significantly associated with at least one lipid trait, with the most significant associations found in the *APOA1/C3/A4/A5-ZPR1-BUD13* gene cluster on chromosome 11. Notably, all 11 variants associated with hypertriglyceridemia were within this cluster; whereas variants associated with hypercholesterolemia were located at chromosome 2 and 19, and for low high density lipoprotein cholesterol were in chromosomes 9, 11, and 19. No significant associated variants were found for low density lipoprotein. We found several novel variants associated with

different lipemic traits: rs3825041 in *BUD13* with hypertriglyceridemia, rs7252453 in *CILP2* with decreased risk to hypercholesterolemia and rs11076176 in *CETP* with increased risk to low high density lipoprotein cholesterol.

Conclusions: We identified novel variants in lipid-regulation candidate genes in the Mexican population, an underrepresented population in genomic studies, demonstrating the necessity of more genomic studies on multi-ethnic populations to gain a deeper understanding of the genetic structure of the lipemic traits.

Keywords: association study, genetic variants, dyslipidemia, exome analysis, mexican population

INTRODUCTION

Currently, metabolic diseases have become one of the most challenging health concerns worldwide; they account for nearly 45% of all deaths worldwide (Forouzanfar et al., 2016). Prevention and management of these diseases are complicated, due to their long latency periods, numerous risk factors, and the presence of co-morbidities (Sevick et al., 2007). Among the known metabolic diseases, dyslipidemias comprise a well-established risk factor for cardiovascular diseases, which are the leading cause of death worldwide (Libby, 2002).

Although factors like diet and lifestyle (Wietlisbach et al., 1997) are recognized as important determinants in the clinical development of dyslipidemias, these diseases have also a strong genetic component (Gao et al., 2018). For instance, lipid disorders are highly prevalent in populations with Amerindian ancestry (up to 85.9%), compared with the prevalence among individuals from other ancestries (e.g., Caucasian 31.2% and African 41.1%) (Aguilar-Salinas et al., 2014; Noubiap et al., 2018). In addition, a large number of genome-wide association studies (GWAS) performed in populations of different ancestries have described more than 175 loci associated with dysregulated levels of plasma lipids. However, the largest body of information generated to date has relied mostly on evidence from Caucasian and Asian cohorts, with very few studies analyzing Latin American populations (Teslovich et al., 2010; Asselbergs et al., 2012; Willer et al., 2013; Wu et al., 2013; Surakka et al., 2015).

Recent findings have demonstrated that genetic factors associated with metabolic traits, including dyslipidemias, exhibited significant heterogeneity in allele frequency and in variant effects across groups with different ancestries (Klarin et al., 2018; Martagón et al., 2018). A study analyzing populations from multiple ethnicities has found important differences in the levels of association, allele frequencies, and haplotype distributions of several lipid loci (Wu et al., 2013). Several population-specific signals at these loci have been reported in non-European populations. For example, the association between the regulatory variant, rs12740374, in the *CELSR2/PSRC1/SORT1* locus, and low-density lipoprotein cholesterol (LDL-C) is higher in African-American individuals than in individuals of European descent (Buyske et al., 2012). Furthermore, the R230C variant of the *ABCA1* gene, which is associated with low levels of high-density lipoprotein cholesterol (HDL-C), is private to individuals with Amerindian ancestry (Acuña-Alonzo et al., 2010). Thus, the

distribution of variants located in lipid-related genes might vary between populations of different ethnicities.

The modern Mexican population is mainly composed of Mestizo individuals, who are the result of the recent admixture of original Amerindians, Europeans (mainly Spaniards) and, to a lesser extent, sub-Saharan Africans (Moreno-Estrada et al., 2014). The recent and complex admixture in the Mexican Mestizo population might have produced a high level of genetic heterogeneity in variants in lipid-related loci. Therefore, this study aimed to determine the frequency distribution of variants located at genes related to lipid traits in Mexican individuals, followed by an association testing and the identification of potential variants for lipemic traits, using previously published exome sequencing data (Estrada et al., 2014; García-Ortiz et al., 2021).

MATERIAL AND METHODS

Study Populations

This study included 2838 Mexican individuals belonging to the Mestizo cohorts Diabetes in Mexico Study (DMS, $n = 968$) and Mexico City Diabetes Study (MCDS, $n = 796$), published previously as part of the Slim Initiative in Genomic Medicine for the Americas (SIGMA) Type 2 Diabetes Consortium (Estrada et al., 2014), as well as to the indigenous cohort Metabolic Analysis in an Indigenous Sample (MAIS, $n = 1074$) (García-Ortiz et al., 2021). The sample design was previously described in Estrada et al. and García-Ortiz et al. (Estrada et al., 2014; García-Ortiz et al., 2021). The MAIS sample belongs to 71 indigenous communities representing 60 ethnic groups from 10 linguistic families. All participants were unrelated volunteers and provided signed informed consent. This investigation was approved by the local ethics and research committees from the National Institute of Genomic Medicine and was conducted according to the principles of the Declaration of Helsinki. The genetic structure of DMS and MCDS population was previously described, with a mean (SD) proportion of Amerindian ancestry of $66 \pm 17\%$, whereas the proportion of Amerindian ancestry in the MAIS cohort was of $93.2 \pm 8.7\%$ (Estrada et al., 2014; García-Ortiz et al., 2021).

Lipid Measurements in Plasma

Levels of triglycerides (TG), total cholesterol (TC), and HDL-C were measured from blood samples collected after overnight

fasting, using a Synchron CX5 Analyzer System (Beckman Coulter Fullerton, CA, United States). LDL-C values were calculated with the Friedewald formula, excluding those samples with TG > 400 mg/dl (Warnick et al., 1990).

Each lipid disorder was diagnosed according to the American Heart Association and National Heart, Lung, and Blood Institute guidelines (AHA/NHLBI; <http://www.nhlbi.nih.gov>). An individual was diagnosed with a lipid disorder when serum levels showed any of the following: TG \geq 150 mg/dl (hypertriglyceridemia; HTG), TC \geq 200 mg/dl (hypercholesterolemia; HTC), LDL-C \geq 130 mg/dl (elevated LDL), or HDL-C \leq 50 mg/dl in females or \leq 40 mg/dl in males (low HDL-C). Individuals with desirable lipid values were assigned as controls. Data on lipid-lowering medications were available for over 80% of the participants and adjustment was done for TG and TC binary phenotypes.

Dataset Building and Single-Nucleotide Variant Annotations

First, we analyzed the Variant Call Format (VCF) files previously obtained from all participants. Quality controls for whole-exome sequencing (e.g., read depth, mean coverage, and missing rate) were described previously (Estrada et al., 2014). From these files, we extracted all the biallelic variants (mapped in the Genome Reference Consortium Human genome build 37) of 177 candidate genes for any lipid trait (**Supplementary Table S1**). These genes were selected from: 1) the Global Lipid Genetics Consortium ($n = 165$) (Willer et al., 2013), 2) meta-analysis studies ($n = 10$) (Teslovich et al., 2010; Weissglas-Volkov and Pajukanta, 2010), and 3) re-sequencing and clinical exome studies ($n = 2$) (Estrada et al., 2014; Williams-Amy et al., 2014). We annotated the variants with the ENSEMBL Variant Effect Predictor (<https://www.ensembl.org/info/docs/tools/vep/index.html>; version 87) (McLaren et al., 2016). Finally, we constructed PLINK files that comprised clinical-demographic and genotyping data for the association analyses, which employed VCF tools (v0.1.12b) (Danecek et al., 2011).

Statistical Analysis

A case-control study was conducted in each cohort to identify low-frequency SNVs [minor allele frequency (MAF) = 0.01–0.05] and common SNVs (MAF > 0.05) associated with lipid traits. To analyze associations between lipid components and alleles we estimated the odds ratio (OR) using logistic regression. We performed quantitative trait analyses with linear regression (Beta value). Both methods were performed using an additive model, adjusting by age, sex and the first 10 eigenvectors from the principal component analysis as covariates. All analyses were performed using PLINK v.1.9 software (Purcell et al., 2007).

Next, we performed a fixed effects model meta-analysis for each trait using a weighted inverse variance model in the software package METAL (Willer et al., 2010). Also, genomic control was applied to each study within METAL by adjusting for the genomic inflation factor prior to meta-analysis, to correct for possible residual population stratification.

Significant threshold was determined following the approach by Kanai et al. (Kanai et al., 2016). According to this approach an LD pruning was done with PLINK v.1.9 using a 50-kb sliding window size, a 5-kb window step size and a r^2 threshold of 0.1. According with this, a total of 780 independent SNVs were identified giving a significant threshold of 6.4×10^{-5} .

Significant threshold was determined following the approach by Kanai et al. (Kanai et al., 2016). According to this approach, high LD SNVs in 177 candidate lipids genes were filtered based on LD pruning performed in PLINK v.1.9 using a 50-kb sliding window size, a 5-kb window step size and a r^2 threshold of 0.1. According with this, a total of 780 independent SNVs were identified that were used to establishing the genome-wide significant threshold *via* a straight-forward Bonferroni correction $p < (6.4 \times 10^{-5})$.

RESULTS

Study Participants

Our study population comprised 2838 individuals belonging to three different cohorts, the Diabetes in Mexico Study (DMS), the Mexico City Diabetes Study (MCDS) and the indigenous cohort Metabolic Analysis in an Indigenous Sample (MAIS) (**Table 1**). The DMS sample was composed of 968 individuals: 681 females (70.3%) and 287 males (29.7%). The mean age of the participants was 54.1 ± 9.8 years and the mean body mass index BMI was 28.4 ± 4.9 kg/m². According to the AHA/NHLBI cutoff points, the lipid trait with the highest prevalence in this sample was low HDL-C (70.8%), followed by HTC (65.1%), HTG (54.9%), and high LDL-C (37%). The MCDS sample consisted of 796 individuals: 482 females (60.6%) and 314 males (39.4%) with a mean age of 62.9 ± 7.6 years and a mean BMI of 29.5 ± 4.9 kg/m². The most frequent lipid alteration in this population was low HDL-C (82.8%), followed by HTG (58.8%), HTC (44.1%), and high LDL-C (40.0%). The MAIS sample was composed of 1074 individuals: 679 females (63.2%) and 395 males (36.8%) with a mean age of 58.6 ± 12.1 years and a mean BMI of 27.5 ± 5.0 kg/m². Similar to the two previous cohorts, low HDL-C (75.2%) was the most common lipid trait, followed by HTG (66.8%), HTC (29.7%) and high LDL-C (14.9%) (**Table 1**).

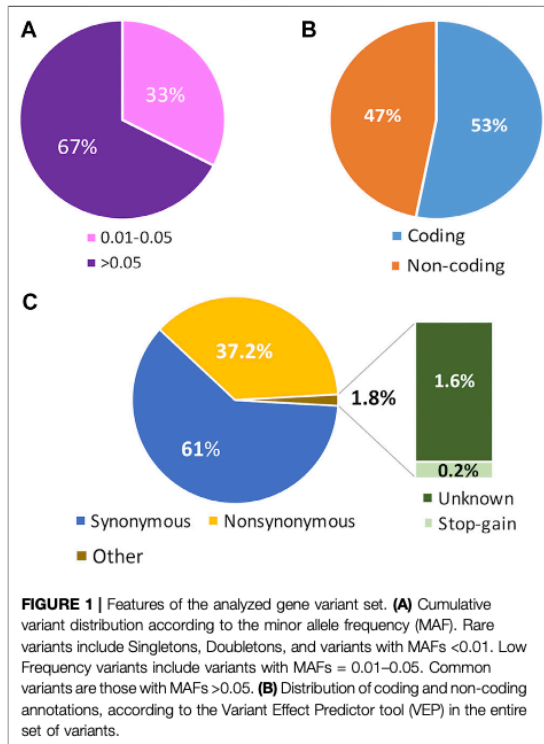
Analysis of the Variation in Candidate Dyslipidemia-Related Genes

After analyzing the exome data from the three analyzed cohorts, we found a total of 34251 biallelic variants with a MAF above 1% within the 177 genes previously related to dyslipidemia. Among these variants, 33% were detected at low frequency (MAF = 0.01–0.05), and 67% were common (MAF > 0.05) (**Figure 1A**). According to their position in the gene, 53% of the variants were in coding regions, and the remaining 47% were found in non-coding regions (**Figure 1B**). According to the predicted annotation obtained with the Variant Effect Predictor web tool (<https://grch37.ensembl.org/Tools/VEP>) 61% of the variants found in the coding regions were synonymous, 37.2% were non-synonymous, 0.2% were stop-gain and 1.6% were unknown (**Figure 1C**).

TABLE 1 | Clinical and demographic characteristics of the studied cohort.

Characteristic	DMS		MCDS		MAIS	
	n = 968		n = 796		n = 1074	
Women/Men (%)	70.3/29.7		60.6/39.4		63.2/36.8	
Age (Years)	54.1 ± 9.8		62.9 ± 7.6		58.6 ± 12.1	
BMI (kg/m ²)	28.4 ± 4.9		29.5 ± 4.9		27.5 ± 5.0	
FG (mg/dl)	125.2 ± 61.4		114.3 ± 54.2		114.2 ± 63.9	
HTG (%)	54.9		58.8		66.8	
Mean TG (mg/dl)	203.7 ± 129.3		176.3 ± 96.3		207.5 ± 119.5	
HTC (%)	65.1		44.1		29.7	
Mean TC (mg/dl)	201.1 ± 43.8		195.9 ± 36.5		182.8 ± 38.3	
Low HDL-C (%)	70.8		82.8		75.2	
Mean HDL-C (mg/dl)	Women	39.7 ± 12.1	Men	31.5 ± 8.5	40.6 ± 12.6	
		37.4 ± 13.0		28.6 ± 13.4	38.7 ± 12.6	
High LDL-C (%)	37		40.0		14.9	
Mean ¹ LDL-C (mg/dl)	120.7 ± 35.3		131.5 ± 33.6		103.7 ± 29.5	

Data are presented as the mean ± SD, or the percentage, as indicated. BMI, body mass index; FG, fasting glucose; HTG, hypertriglyceridemia; TG, triglycerides; HTC, hypercholesterolemia; TC, total cholesterol; HDL-C, high density lipoprotein; LDL-C, Low density lipoprotein. ¹LDL-C values were calculated with the Friedewald equation.



Low-Frequency and Common Variant Association Study

For the case-control study we used a meta-analysis approach to evaluate low-frequency and common variants pruned for LD (MAF >0.01; $n = 780$). We explored associations between individual variants and each different lipid phenotype with

binary logistic regression. Quantile-quantile (QQ) plots for each analyzed phenotype showed a genomic inflation factor for HTG, HTC, low HDL-C and high LDL-C of $\lambda = 1.0, 1.02, 1.03$, and 0.91 respectively (**Supplementary Figure S1**). A total of 19 variants in 9 genes achieved a $p < 6.4 \times 10^{-5}$ or higher significant association with at least one of the analyzed lipid components (**Figure 2**). Remarkably, the 11 variants that were associated with HTG were found on chromosome 11, located in *BUD13*, *APOA5*, *APOC3*, *APOA1*, and *APOA4* genes (**Supplementary Table S2**). In the case of HTC, associated variants were found on chromosome 2 and 19, in *APOB* and *CILP2* genes; whereas in the case of low HDL-C, two of the associated variants were found on chromosome 11, in *BUD13* gene, four on chromosome 16 in *CETP* and one on chromosome 9 in *ABCA1* (**Supplementary Table S2**). We did not find any variant associated with high LDL-C.

The 11 variants associated with HTG were all located in the *APOA1/C3/A4/A5-ZPR1-BUD13* gene cluster (**Figure 2; Supplementary Table S2**). The strongest signal for the variants associated with increased risk to HTG was detected for rs3825041 in *BUD13* (OR = 1.53, $p = 7.55 \times 10^{-9}$; $\beta = 22.24$ mg/dl) (**Figure 3; Supplementary Table S2**), followed by rs651821 (OR = 1.48, $p = 5.70 \times 10^{-8}$; $\beta = 21.40$ mg/dl) and rs2072560 in *APOA5* (OR = 1.48, $p = 6.18 \times 10^{-8}$; $\beta = 21.65$ mg/dl), rs2070665 in *APOA1* (OR = 1.41, $p = 3.79 \times 10^{-7}$; $\beta = 17.71$ mg/dl), rs5128 in *APOC3* (OR = 1.40, $p = 4.22 \times 10^{-7}$; $\beta = 18.31$ mg/dl), rs5104 (OR = 1.34, $p = 6.1 \times 10^{-6}$; $\beta = 19.61$ mg/dl) and rs5092 in *APOA4* (OR = 1.33, $p = 8.07 \times 10^{-6}$; $\beta = 18.97$ mg/dl) and rs11820589 in *BUD13* (OR = 1.35, $p = 5.71 \times 10^{-5}$; $\beta = 21.92$ mg/dl). In contrast, the variants rs4520 in *APOC3* (OR = 0.74, $p = 2.96 \times 10^{-7}$; $\beta = -19.03$ mg/dl) rs5070 in *APOA1* (OR = 0.78, $p = 4.27 \times 10^{-5}$; $\beta = -15.75$ mg/dl) and rs10488698 in *BUD13* (OR = 0.73, $p = 5.80 \times 10^{-3}$; $\beta = -20.98$ mg/dl) were associated with protection against HTG. All 11 variants were also significantly associated with TG levels in the quantitative trait analysis (**Supplementary Figure S2; Supplementary Table S2**).

Moreover, the variant rs1367117 in *APOB* (OR = 1.36, $p = 1.88 \times 10^{-6}$; $\beta = 4.89$) was significantly associated with HTC and with a



significant increase in serum TC levels in the quantitative trait analysis (Figure 3; Supplementary Table S2). On the other hand, the variant rs7252453 in *CILP2* (OR = 0.73, $p = 4.40 \times 10^{-5}$; $\beta =$

-5.17) was significantly associated with protection against HTC (Supplementary Figure S3; Supplementary Table S2). Regarding low HDL-C, the variants rs17519093 and

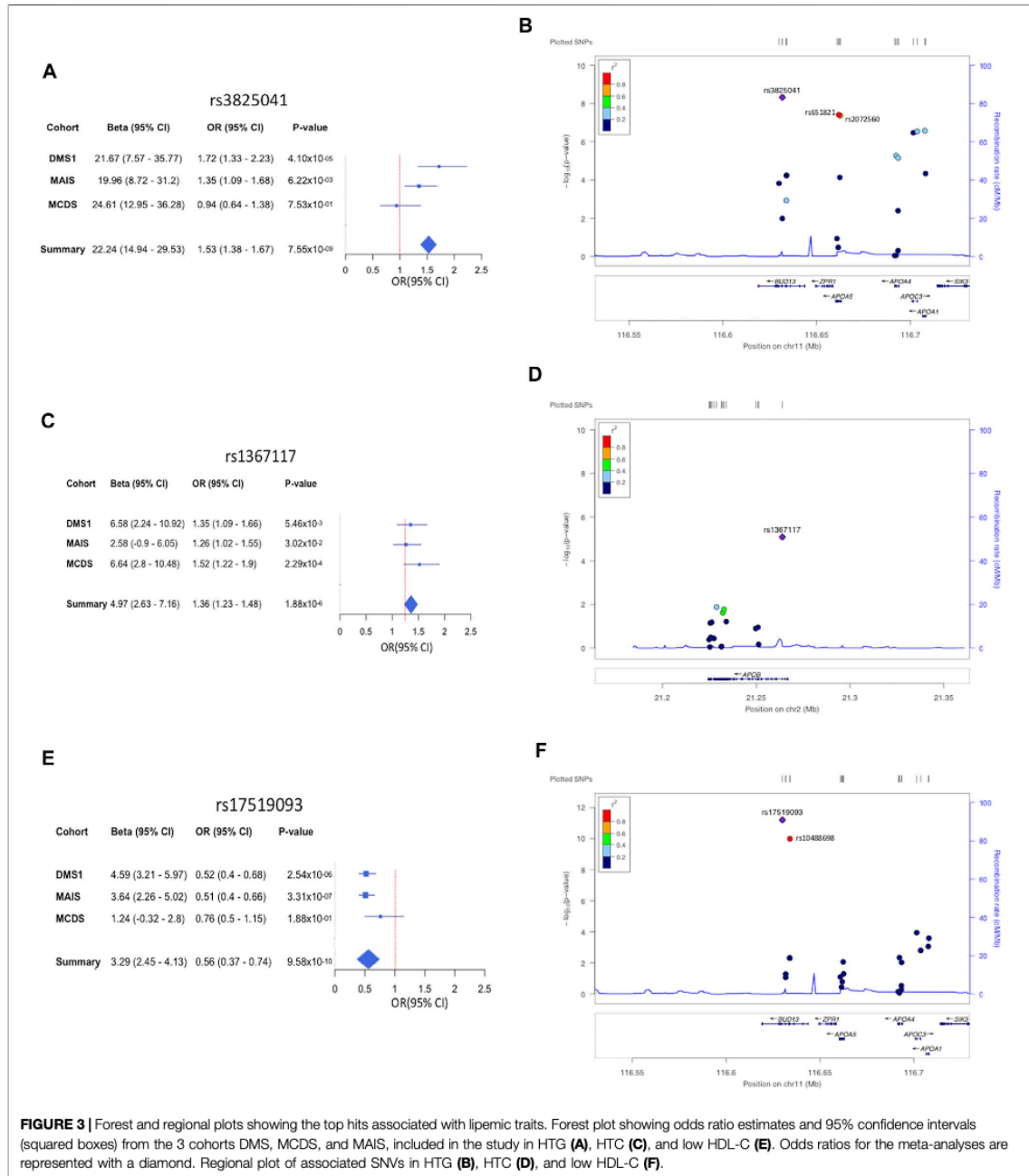


FIGURE 3 | Forest and regional plots showing the top hits associated with lipemic traits. Forest plot showing odds ratio estimates and 95% confidence intervals (squared boxes) from the 3 cohorts DMS, MCDS, and MAIS, included in the study in HTG (A), HTC (C), and low HDL-C (E). Odds ratios for the meta-analyses are represented with a diamond. Regional plot of associated SNVs in HTG (B), HTC (D), and low HDL-C (F).

rs10488698 in *BUD13* (OR = 0.56, $p = 9.58 \times 10^{-10}$, $\beta = 3.29$; OR = 0.58, $p = 7.34 \times 10^{-9}$; $\beta = 3.12$, respectively) and rs1532625 in *CETP* (OR = 0.72, $p = 6.56 \times 10^{-6}$; $\beta = 2.21$) were significantly associated with decreased risk in both the binary and the

quantitative trait analysis (Figure 3; Supplementary Figure S4; Supplementary Table S2). Finally, the variants rs708272, rs11076176, and rs289714 in *CETP* (OR = 1.32, $p = 1.75 \times 10^{-5}$, $\beta = -2.12$; OR = 1.32, $p = 3.64 \times 10^{-5}$, $\beta = -1.83$; OR = 1.32, $p = 4.03$

$\times 10^{-5}$, $\beta = -1.75$, respectively) as well as rs9282541 in *ABCA1* (OR = 1.67, $p = 3.34 \times 10^{-5}$; $\beta = -2.50$) showed a significant association with increased risk to low HDL-C in both the binary and the quantitative trait analysis (Supplementary Figure S4; Supplementary Table S2).

DISCUSSION

The contribution of genetic variation to human diseases is widely recognized (Xue et al., 2012). In particular, loss-of-function or modifier variants are usually related to changes in the biological activity of the corresponding gene. Several genomic regions are currently recognized as drivers of dyslipidemias in populations from several ethnicities, although in some populations, such as those of Amerindian or African origin, genomic studies have been scarce. Furthermore, many of these loci have displayed highly different associations with these entities across different ancestries (Lek et al., 2016; Martin et al., 2017). Therefore, lipid-associated genes should be analyzed in all populations.

Here, by analyzing SNVs within 177 candidate genes, we found several associations with different types of dyslipidemias in Mexican Mestizos, an admixed population with strong Amerindian (51%) and European (46%) components (Norris et al., 2018).

In our population we were able to replicate several associations with different lipemic traits previously reported in other ethnicities. For example, variants in the gene cluster *APOA1/C3/A4/A5-ZPR1-BUD13*, such as the regulatory SNVs rs5128 in *APOC3*, and rs651821 in *APOA5*, as well as the missense SNV rs2072560 also within *APOA5*, and the intronic variant rs2070665 in *APOA1*, were all associated with HTG in our population, as they are in European, Asian, and African populations (Feng et al., 2016; Fu et al., 2015; Jasim et al., 2018; Ken-Dror et al., 2010; Song et al., 2015; Zhou et al., 2013). Likewise, the association of the missense SNV rs1367117 in *APOB* with HTC has also been reported in populations of European ancestry (Lu et al., 2010), whereas the association of the missense SNVs rs10488698 in *BUD13* and rs9282541 in *ABCA1* with low HDL-C has also been observed in Asian and Latino American populations, respectively (Zhang et al., 2017; Acuña-Alonso et al., 2010). We also observed novel lipemic trait-associated variants in *BUD13*, such as rs3825041 and rs17519093, both localized within introns, associated with HTG and high levels of HDL-C, respectively. Using public data contained in the Common Metabolic Diseases Knowledge Portal (<https://t2d.hugeamp.org/>), we were able to confirm the associations of rs3825041 and rs17519093 with different lipemic traits. Thus, one of the genomic regions most consistently associated with lipid traits in human populations of diverse ethnic origins is the cluster *APOA1/C3/A4/A5-ZPR1-BUD13* (Teslovich et al., 2010; Willer et al., 2013; Parra et al., 2017; Bai et al., 2019). Importantly, all 11 variants associated with HTG in our study, were in this cluster. Among them, the novel variant rs3825041 in *BUD13* showed the strongest association. Notably, this variant showed high LD with rs651821 ($r^2 = 0.84$) and rs2072560 ($r^2 = 0.89$) within *APOA5*, which have both been previously reported as associated with HTG in several

populations of different ancestries, including in Mexicans (Ken-Dror et al., 2010; Kim et al., 2019). Taken together, these data provide more insights about the variants at the *APOA1/C3/A4/A5-ZPR1-BUD13* gene cluster as a relevant risk factor for dyslipidemias such as HTG and low HDL-C, and highlight the notion that these could be biomarkers for susceptibility to these traits.

Others novel association signals were observed with the synonymous SNV rs7252453 in *CILP2* and decreased risk to HTC and the intronic SNV rs11076176 in *CETP* and increased risk to low HDL-C serum levels. On *CETP* we also observed an association of the intronic variants rs708272 and rs289714 with high risk to low HDL-C, as well as the association of the intronic variant rs1532625 with protection to this dyslipidemia. These findings are in line with those reported previously in Mexican individuals (Vargas-Alarcon et al., 2018), in a meta-analysis involving six independent Hispanic cohorts (Gao et al., 2018) and in Chinese population (Guo et al., 2015).

In summary, despite our selecting candidate genes were previously associated with dyslipidemia in other populations, we were able to find additional variants showing the strongest associations with lipid traits in Mexican individuals. These differences support the notion that high allelic heterogeneity exists in lipid loci across populations. Remarkably, the new associations of variants in genes previously related to dyslipidemia, points out the importance of studying different ethnicities, since different associated variants within the same genes could be particular to one or another population ancestry. It is also worth to note that several of the risk variants previously associated with lipemic traits in different ethnic groups, including European, Asian and African populations were replicated in our study. Taken together, our results suggest that genetic architecture of dyslipidemias is partially share among different ethnic groups.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the European Variation Archive repository (EVA <https://www.ebi.ac.uk/eva/>), accession number PRJEB52611.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the local ethics and research committees from the National Institute of Genomic Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHORS CONTRIBUTIONS

LO and EC contributed to the conception and design of the work. PJ-C, FB-O, HG-O, and MC-S performed the bioinformatic

analyses for variant calling. PB-P, AM-H, MG-V, and CG-V contributed in acquisition of data. FC-C, CC-C, YS-A, GS-M, and EM-C participated in the association studies. PJ-C, LO, and EC wrote the manuscript.

FUNDING

This work was supported by the National Institute of Genomic Medicine (grant number INM 2014); and the CONACyT (grant numbers SEP-2014-243587 and PDCPN_2016_3251).

REFERENCES

- Acuña-Alonso, V., Flores-Dorantes, T., Kruit, J. K., Villarreal-Molina, T., Arellano-Campos, O., Hünemeier, T., et al. (2010). A Functional ABCA1 Gene Variant Is Associated with Low HDL-Cholesterol Levels and Shows Evidence of Positive Selection in Native Americans. *Hum. Mol. Genet.* 19 (14), 2877–2885. doi:10.1093/hmg/ddq173
- Aguilar-Salinas, C. A., Tusie-Luna, T., and Pajukanta, P. (2014). Genetic and Environmental Determinants of the Susceptibility of Amerindian Derived Populations for Having Hypertriglyceridemia. *Metabolism Clin. Exp.* 63, 887–894. doi:10.1016/j.metabol.2014.03.012
- Asselbergs, F. W., Guo, Y., Van Iperen, E. P. A., Sivapalaratnam, S., Tragante, V., Lanktree, M. B., et al. (2012). Large-scale Gene-Centric Meta-Analysis across 32 Studies Identifies Multiple Lipid Loci. *Am. J. Hum. Genet.* 91 (5), 823–838. doi:10.1016/j.ajhg.2012.08.032
- Bai, W., Kou, C., Zhang, L., You, Y., Yu, W., Hua, W., et al. (2019). Functional Polymorphisms of the APOA1/C3/A4/A5-ZPR1-BUD13 Gene Cluster Are Associated with Dyslipidemia in a Sex-specific Pattern. *PeerJ* 6, e6175. doi:10.7717/peerj.6175
- Buyse, S., Wu, Y., Carty, C. L., Cheng, I., Assimes, T. L., Dumitrescu, L., et al. (2012). Evaluation of the MetaboChip Genotyping Array in African Americans and Implications for Fine Mapping of Gwas-Identified Loci: The PAGE Study. *PLoS ONE* 7 (4), e35651. doi:10.1371/journal.pone.0035651
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The Variant Call Format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/bioinformatics/btr330
- Estrada, K., Aukrust, I., Bjørkhaug, L., Burt, N. P., Mercader, J. M., Garcia-Ortiz, H., et al. (2014). Association of a Low-Frequency Variant in HNF1A with Type 2 Diabetes in a Latino Population the SIGMA Type 2 Diabetes Consortium. *JAMA - J. Am. Med. Assoc.* 311, 2305–2314. doi:10.1001/jama.2014.6511
- Feng, D. W., Ma, R. L., Guo, H., He, J., Yan, Y. Z., Muratbek, et al. (2016). Association of APOA1 Gene Polymorphisms (Rs670, Rs5069, and Rs2070665) with Dyslipidemia in the Kazakhs of Xinjiang. *Genet. Mol. Res.* 15. doi:10.4238/gmr.15028094
- Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., et al. (2016). Global, Regional, and National Comparative Risk Assessment of 79 Behavioural, Environmental and Occupational, and Metabolic Risks or Clusters of Risks, 1990–2015: a Systematic Analysis for the Global Burden of Disease Study 2015. *Lancet* 388, 1659–1724. doi:10.1016/S0140-6736(16)31679-8
- Fu, Q., Tang, X., Chen, J., Su, L., Zhang, M., Wang, L., et al. (2015). Effects of Polymorphisms in APOA4-APOA5-ZNF259-BUD13 Gene Cluster on Plasma Levels of Triglycerides and Risk of Coronary Heart Disease in a Chinese Han Population. *PLoS ONE* 10. doi:10.1371/journal.pone.0138652
- Gao, C., Tabb, K. L., Dimitrov, L. M., Taylor, K. D., Wang, N., Guo, X., et al. (2018). Exome Sequencing Identifies Genetic Variants Associated with Circulating Lipid Levels in Mexican Americans: The Insulin Resistance Atherosclerosis Family Study (IRASFS). *Sci. Rep.* 8. doi:10.1038/s41598-018-23727-2
- García-Ortiz, H., Barajas-Olmos, F., Contreras-Cubas, C., Cid-Soto, M. Á., Córdova, E. J., Centeno-Cruz, F., et al. (2021). The Genomic Landscape of Mexican Indigenous Populations Brings Insights into the Peopling of the Americas. *Nat. Commun.* 12, 5942–5954. doi:10.1038/s41467-021-26188w
- Guo, S., Hu, Y., Ding, Y., Liu, J., and Zhang, M. (2015). Association between Eight Functional Polymorphisms and Haplotypes in the Cholesterol Ester Transfer Protein (CETP) Gene and Dyslipidemia in National Minority Adults in the Far West Region of China. *Int. J. Environ. Res. Public Health* 12, 15979–15992. doi:10.3390/ijerph121215036
- Jasim, A. A., Al-Bustan, S. A., Al-Kandari, W., Al-Serri, A., and Al-Askar, H. (2018). Sequence Analysis of APOA5 Among the Kuwaiti Population Identifies Association of Rs2072560, Rs2266788, and Rs662799 with TG and VLDL Levels. *Front. Genet.* 9. doi:10.3389/fgene.2018.00112
- Kanai, M., Tanaka, T., and Okada, Y. (2016). Empirical Estimation of Genome-wide Significance Thresholds Based on the 1000 Genomes Project Data Set. *J. Hum. Genet.* 61 (10), 861–866. PMID: 27305981; PMCID: PMC5090169. doi:10.1038/jhg.2016.72
- Ken-Dror, G., Goldbourt, U., and Dankner, R. (2010). Different Effects of Apolipoprotein A5 SNPs and Haplotypes on Triglyceride Concentration in Three Ethnic Origins. *J. Hum. Genet.* 55. doi:10.1038/jhg.2010.27
- Kim, H. K., Anwar, M. A., and Choi, S. (2019). Association of BUD13-Znf259-APOA5-APOA1-SIK3 Cluster Polymorphism in 11q23.3 and Structure of APOA5 with Increased Plasma Triglyceride Levels in a Korean Population. *Sci. Rep.* 9. doi:10.1038/s41598-019-44699-x
- Klarin, D., Damrauer, S. M., Cho, K., Sun, Y. V., Teslovich, T. M., Honerlaw, J., et al. (2018). Genetics of Blood Lipids Among ~300,000 Multi-Ethnic Participants of the Million Veteran Program. *Nat. Genet.* 50, 1514–1523. doi:10.1038/s41588-018-0222-9
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of Protein-Coding Genetic Variation in 60,706 Humans. *Nature* 536, 285–291. doi:10.1038/nature19057
- Libby, P. (2002). Inflammation in Atherosclerosis. *Nature* 420, 868–874. doi:10.1038/nature01323
- Lu, Y., Feskens, E. J., Boer, J. M., Imholz, S., Verschuren, W. M., Wijmenga, C., et al. (2010). Exploring Genetic Determinants of Plasma Total Cholesterol Levels and Their Predictive Value in a Longitudinal Study. *Atherosclerosis* 213, 200–205. doi:10.1016/j.atherosclerosis.2010.08.053
- Martagón, A. J., Bello-Chavolla, O. Y., Arellano-Campos, O., Almeda-Valdés, P., Walford, G. A., Cruz-Bautista, I., et al. (2018). Mexican Carriers of the HNF1A P. E508K Variant Do Not Experience an Enhanced Response to Sulfonylureas. *Diabetes Care* 41, 1726–1731. doi:10.2337/dc18-0384
- Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., et al. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* 107, 788–789. doi:10.1016/j.ajhg.2017.03.004
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., et al. (2016). The Ensembl Variant Effect Predictor. *Genome Biol.* 17 (1), 122. doi:10.1186/s13059-016-0974-4
- Moreno-Estrada, A., Gignoux, C. R., Fernández-López, J. C., Zakharia, F., Sikora, M., Contreras, A. v., et al. (2014). The Genetics of Mexico Recapitulates Native American Substructure and Affects Biomedical Traits. *Science* 344, 1280–1285. doi:10.1126/science.1251688
- Norris, E. T., Wang, L., Conley, A. B., Rishishwar, L., Mariño-Ramírez, L., Valderrama-Aguirre, A., et al. (2018). Genetic Ancestry, Admixture and

ACKNOWLEDGMENTS

PJ-C is a doctoral student of the Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and received a CONACyT fellowship (619190).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.807381/full#supplementary-material>

- Health Determinants in Latin America. *BMC Genomics* 19. doi:10.1186/s12864-018-5195-7
- Noubiap, J. J., Bigna, J. J., Nansseu, J. R., Nyaga, U. F., Balti, E. V., Echouffo-Tcheugui, J. B., et al. (2018). Prevalence of Dyslipidaemia Among Adults in Africa: a Systematic Review and Meta-Analysis. *Lancet Glob. Health* 6, e998–e1007. doi:10.1016/S2214-109X(18)30275-4
- Parra, E. J., Mazurek, A., Gignoux, C. R., Sockell, A., Agostino, M., Morris, A. P., et al. (2017). Admixture Mapping in Two Mexican Samples Identifies Significant Associations of Locus Ancestry with Triglyceride Levels in the BUD13/ZNF259/APOA5 Region and Fine Mapping Points to Rs964184 as the Main Driver of the Association Signal. *PLoS ONE* 12. doi:10.1371/journal.pone.0172880
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795
- Sevick, M. A., Trauth, J. M., Ling, B. S., Anderson, R. T., Piatt, G. A., Kilbourne, A. M., et al. (2007). Patients with Complex Chronic Diseases: Perspectives on Supporting Self-Management. *J. General Intern. Med.* 22, 438–444. doi:10.1007/s11606-007-0316-z
- Song, Y., Zhu, L., Richa, M., Li, P., Yang, Y., and Li, S. (2015). Associations of the APOC3 Rs5128 Polymorphism with Plasma APOC3 and Lipid Levels: A Meta-Analysis. *Lipids Health Dis.* 14. doi:10.1186/s12944-015-0027-0
- Surakka, I., Horikoshi, M., Mägi, R., Sarin, A. P., Mahajan, A., Lagou, V., et al. (2015). The Impact of Low-Frequency and Rare Variants on Lipid Levels. *Nat. Genet.* 47, 589–597. doi:10.1038/ng.3300
- Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., et al. (2010). Biological, Clinical and Population Relevance of 95 Loci for Blood Lipids. *Nature* 466, 707–713. doi:10.1038/nature09270
- Vargas-Alarcon, G., Perez-Mendez, O., Herrera-Maya, G., Garcia-Sanchez, C., Martinez-Rios, M. A., and Peña-Duque, M. A. (2018). CETP and LCAT Gene Polymorphisms Are Associated with High-Density Lipoprotein Subclasses and Acute Coronary Syndrome. *Lipids* 53, 157–166. doi:10.1002/lipd.12017
- Warnick, G. R., Knopp, R. H., Fitzpatrick, V., and Branson, L. (1990). Estimating Low-Density Lipoprotein Cholesterol by the Friedewald Equation Is Adequate for Classifying Patients on the Basis of Nationally Recommended Cutpoints. *Clin. Chem.* 36, 15–19. doi:10.1093/clinchem/36.1.15
- Weissglas-Volkov, D., and Pajukanta, P. (2010). Genetic Causes of High and Low Serum HDL-Cholesterol. *J. Lipid Res.* 51, 2032–2057. doi:10.1194/jlr.R004739
- Wietlisbach, V., Paccaud, F., Rickenbach, M., and Gutzwiller, F. (1997). Trends in Cardiovascular Risk Factors (1984–1993) in a Swiss Region: Results of Three Population Surveys. *Prev. Med.* 26, 523–533. doi:10.1006/pmed.1997.0167
- Willer, C. J., Li, Y., and Abecasis, G. R. (2010). METAL: Fast and Efficient Meta-Analysis of Genomewide Association Scans. *Bioinformatics* 26, 2190–2191. doi:10.1093/bioinformatics/btq340
- Willer, C. J., Schmidt, E. M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., et al. (2013). Discovery and Refinement of Loci Associated with Lipid Levels. *Nat. Genet.* 45, 1274–1285. doi:10.1038/ng.2797
- Williams Amy, A. L., Jacobs Suzanne, S. B. R., Moreno-Macías, H., Huerta-Chagoya, A., Churchhouse, C., Márquez-Luna, C., et al. (2014). Sequence Variants in SLC16A11 Are a Common Risk Factor for Type 2 Diabetes in Mexico. *Nature* 506, 97–101. doi:10.1038/nature12828
- Wu, Y., Waite, L. L., Jackson, A. U., Sheu, W. H. H., Buyske, S., Absher, D., et al. (2013). Trans-Ethnic Fine-Mapping of Lipid Loci Identifies Population-specific Signals and Allelic Heterogeneity that Increases the Trait Variance Explained. *PLoS Genet.* 9, e1003379. doi:10.1371/journal.pgen.1003379
- Xue, Y., Chen, Y., Ayub, Q., Huang, N., Ball, E. V., Mort, M., et al. (2012). Deleterious- and Disease-Allele Prevalence in Healthy Individuals: Insights from Current Predictions, Mutation Databases, and Population-Scale Resequencing. *Am. J. Hum. Genet.* 91, 1022–1032. doi:10.1016/j.ajhg.2012.10.015
- Zhang, L., You, Y., Wu, Y., Zhang, Y., Wang, M., Song, Y., et al. (2017). Association of BUD13 Polymorphisms with Metabolic Syndrome in Chinese Population: a Case-Control Study. *Lipids Health Dis.* 16, 127. doi:10.1186/s12944-017-0520-8
- Zhou, L., He, M., Mo, Z., Wu, C., Yang, H., Yu, D., et al. (2013). A Genome Wide Association Study Identifies Common Variants Associated with Lipid Levels in the Chinese Population. *PLoS ONE* 8. doi:10.1371/journal.pone.0082420

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Jurado-Camacho, Cid-Soto, Barajas-Olmos, García-Ortiz, Baca-Peynado, Martínez-Hernández, Centeno-Cruz, Contreras-Cubas, González-Villalpando, Saldaña-Álvarez, Salas-Martínez, Mendoza-Caamal, González-Villalpando, Córdova and Orozco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

17. Anexos

Tabla suplementaria 1 Lista de genes candidato y posiciones en el GRCh37

Gen ID	CHR	Inicio	Fin	Gen ID	CHR	Inicio	Fin	Gen ID	CHR	Inicio	Fin
A2ML1	12	8975216	9029383	FAM13A	4	89647104	89978336	NR0B2	1	27237979	27240466
ABCA1	9	107543286	107690436	FN1	2	216225162	216300793	NR1H2	19	50879713	50886285
ABCA8	17	66863427	66951533	FRK	6	116252311	116382054	NR1H3	11	47269850	47290584
ABCB11	2	169777290	169887834	FRMD5	15	44162958	44487492	NYNRIN	14	24868208	24888489
ABCG4	11	119019749	119033374	FTO	16	53737874	54155853	OR4C46	11	51515281	51516211
ABCG5	2	44039610	44065958	GALNT2	1	230193535	230417876	OSBPL7	17	45884737	45899160
ABCG8	2	44066102	44110127	GCK	7	44182811	44229038	PABPC4	1	40026484	40042521
ABO	9	136125787	136150617	GCKR	2	27719705	27746550	PCCB	3	135969166	136049013
ADH5	4	99992129	100009939	GPAM	10	113909621	113943537	PCSK9	1	55505894	55530526
AKR1C4	10	5238799	5260912	GPR146	7	1084181	1098905	PDE3A	12	20522178	20837041
AKT1	14	105235685	105262080	GSK3B	3	119540801	119813264	PDXDC1	16	15068591	15233267
ALOX5	10	45869613	45941567	HAS1	19	52216364	52227245	PEPD	19	33877854	34012799
AMPD3	11	10471867	10529126	HBS1L	6	135281516	135376036	PGS1	17	76374698	76420740
ANGPTL1	1	178818669	178840212	HDBGF	1	156711900	156722240	PHC1	12	9067104	9094060
ANGPTL3	1	63063157	63071976	HFE	6	26087508	26095469	PHLDB1	11	118477212	118528748
ANGPTL4	19	8429010	8439259	HLA-B	6	2615167	31324956	PIGV	1	27114453	27124894
ANGPTL8	19	11350294	11352619	HMGCR	5	74632992	74657941	PINX1	8	10622470	10697409
ANXA9	1	150954767	150968109	HNF1A	12	121416370	121440314	PLA2G6	22	38507501	38577837
APOA1	11	116706466	116708666	HNF4A	20	42984440	43061485	PLEC	8	144989314	145050913
APOA2	1	161192081	161193418	HPR	16	72097124	72111145	PLTP	20	44527258	44541003
APOA4	11	116691415	116694020	IKZF1	7	50343678	50472798	PMVK	1	154897207	154909484
APOA5	11	116660085	116663136	INSIG2	2	118845993	118868573	PNPLA3	22	44319618	44343448
APOB	2	21224300	21266945	INSR	19	7112265	7294425	PPARA	22	46546436	46639653
APOC3	11	116700622	116703788	IRF2BP2	1	234740014	234745271	PPP1R3B	8	8993763	9009152
APOE	19	45409005	45412652	IRS1	2	227596032	227663506	PXK	3	58318616	58411854
APOH	17	64208150	64225533	JMJD1C	10	64926980	65028983	RAB3GAP1	2	135809834	135928279
ARL15	5	53179771	53606403	KAT5	11	65479472	65487077	RAF1	3	12625099	12705700
ASAP3	1	23755055	23810750	KCNK17	6	39266776	39282237	RBM5	3	50126340	50156397
ATG7	3	11314009	11606126	KLF14	7	130417381	130418860	RSP03	6	127439815	127520626
BRAP	12	112079949	112123760	KLF4	9	110247132	110252050	SBNO1	12	123773655	123849756
BRCA2	13	32889616	32973809	KLHL8	4	88081255	88141770	SCARB1	12	125261402	125348410
BUD13	11	116618885	116643714	LACTB	15	63413998	63434260	SETD2	3	47057897	47205467
CAPN3	15	42651697	42704515	LCAT	16	67973786	67978015	SLC16A11	17	6944940	6947236
CDS36	7	80231503	80308593	LDLR	19	11200037	11244505	SLC16A13	17	6939393	6943440
CEP350	1	150937648	150947479	LDLRAP1	1	25870075	25895377	SLC39A8	4	103172197	103266639
CETP	16	56995834	57017757	LILRA3	19	54799849	54804265	SNX13	7	17830384	17980131
CILP2	19	19649073	19657468	LIPC	15	58724189	58862043	SNX5	20	17922239	17949634
CITED2	6	139692943	139695785	LIPG	18	47088400	47125554	SORT1	1	109852196	109940540
CMIP	16	81478774	81745367	LPA	6	160952514	161085291	SOX17	8	55370494	55373452
COMT6	3	32522803	32544344	LPL	8	19796763	19824770	SPTLC3	20	12989601	13149750
COBLL1	2	165514440	165698678	LRP1	12	57522281	57607125	SPTY2D1	11	18627949	18655889
CPS1	2	211342405	211543832	LRP4	11	46878267	46940101	ST3GAL4	11	126225539	126284536
CSNK1G3	5	122847792	122952738	LRPAP1	4	3505323	3534149	STAB1	3	52529353	52558511
CTF1	16	30907927	30914881	MAFB	20	39314487	39317876	STARD3	17	37793332	37820454
CUBN	10	16865964	17171816	MAMSTR	19	49215978	49222982	TIMD4	5	156346292	156390266
CYP11A1	15	75011882	75017869	MAP3K1	5	56111375	56191978	TMEM176A	7	150497856	150502206
CYP26A1	10	94833231	94837641	MARCH8	10	45950032	46090354	TOM1	22	35695267	35743987
CYP7A1	8	59402736	59412722	MC4R	18	58038294	58040008	TOP1	20	39657461	39753126
DAGLB	7	6448746	6487643	MET	7	116312249	116438440	TRIB1	8	126442599	126450647
DGAT2	11	75479777	75512581	MLXIPL	7	73007523	73038873	TRPS1	8	116420723	116681255
DLG4	17	7093209	7123369	MOGAT2	11	75428882	75443997	TTC39B	9	15170841	15307358
DNAH11	7	21582656	21941457	MARC1	1	220960254	220993001	UBASH3B	11	122526429	122685181
EHBP1	2	62900985	63273621	MPP3	17	41878166	41910562	UBE2L3	22	21903735	21978323
ERGIC3	20	34129777	34145405	MSL2	3	135867758	135915177	UGT1A1	2	234668918	234681945
EV15	1	92974252	93257961	MTMR3	22	30279157	30426857	VEGFA	6	43737945	43754223
FADS1	11	61567096	61584475	MVK	12	110011499	110035930	VIM	10	17270239	17279592
FADS2	11	61583674	61634826	MYLIP	6	16129316	16148478	VLDR	9	2621786	2660056
FADS3	11	61640997	61659017	NAT2	8	18248791	18258725	ZNF648	1	182023704	182030847
FAM117B	2	203499691	203634480	NPC1L1	7	44552134	44580914	ZNF664	12	124457761	124499986

Tabla suplementaria 2 Catálogo de variantes en los genes candidato

Gen	Total SNV	SNV reportadas	SNV sin reporte	Exón	3'UTR	5'UTR	Intrón	Intergénico	Longitud de transcrito*	SNV cod / transcrito	SNV/ transcrito
LILRA3	68	68	0	46	1	1	20	0	1604	35	24
PEPD	79	75	4	48	1	1	29	0	1910	40	24
HPR	48	48	0	24	0	0	7	17	1244	52	26
PLEC	561	556	5	440	0	0	115	6	15249	35	27
PINX1	54	52	2	26	4	2	22	0	1545	59	29
C19orf80	30	29	1	12	0	0	3	15	872	73	29
DAGLB	91	91	0	41	0	3	47	0	2897	71	32
APOH	34	34	0	17	2	2	13	0	1176	69	35
CETP	51	51	0	25	0	0	26	0	1791	72	35
HLA-B	42	42	0	20	0	2	20	0	1547	77	37
STAB1	207	204	3	115	2	1	89	0	7928	69	38
ST3GAL4	49	49	0	9	0	0	40	0	1967	219	40
DNAH11	344	341	3	202	1	0	141	0	14167	70	41
OR4C46	22	22	0	22	0	0	0	0	930	42	42
MPP3	51	51	0	23	2	0	26	0	2182	95	43
LIPC	48	48	0	28	1	0	19	0	2161	77	45
CAPN3	71	71	0	23	0	4	44	0	3228	140	45
APOA4	32	32	0	27	0	1	4	0	1470	54	46
GPR146	42	42	0	35	1	3	3	0	1966	56	47
PCCB	39	39	0	22	0	0	17	0	1841	84	47
PDXDC1	82	82	0	25	0	0	57	0	3884	155	47
MLXIPL	69	67	2	47	0	0	22	0	3278	70	48
MOGAT2	29	29	0	14	2	1	12	0	1404	100	48
ABCG8	55	55	0	34	1	0	20	0	2665	78	48
CUBN	246	245	1	139	0	0	107	0	11949	86	49
PXK	41	41	0	17	0	0	24	0	2037	120	50
AKR1C4	28	28	0	11	0	0	17	0	1414	129	51
TIMD4	27	27	0	18	2	0	7	0	1369	76	51
LDLR	71	71	0	41	2	0	28	0	3617	88	51
FADS3	34	34	0	20	2	0	12	0	1779	89	52
A2ML1	99	99	0	49	0	2	48	0	5274	108	53
PCSK9	67	67	0	42	0	1	24	0	3637	87	54

<i>ERGIC3</i>	25	24	1	7	1	4	13	0	1359	194	54
<i>ABCG5</i>	50	49	1	33	0	3	14	0	2740	83	55
<i>HAS1</i>	38	38	0	35	0	0	3	0	2087	60	55
<i>PGS1</i>	40	40	0	12	4	1	23	0	2201	183	55
<i>HNF1A</i>	62	62	0	34	1	1	26	0	3442	101	56
<i>ABCB11</i>	86	85	1	57	0	0	29	0	4775	84	56
<i>BUD13</i>	39	39	0	26	1	2	10	0	2196	84	56
<i>LRP1</i>	264	260	4	143	1	0	120	0	14897	104	56
<i>STARD3</i>	37	37	0	16	0	2	19	0	2132	133	58
<i>CD36</i>	40	40	0	20	0	0	20	0	2365	118	59
<i>OSBPL7</i>	60	60	0	24	0	2	34	0	3640	152	61
<i>ALOX5</i>	41	41	0	22	0	0	19	0	2504	114	61
<i>CERS2</i>	41	40	1	13	2	0	10	16	2505	193	61
<i>ABCA8</i>	92	91	1	58	1	0	33	0	5677	98	62
<i>PNPLA3</i>	45	44	1	25	2	2	16	0	2805	112	62
<i>NR1H2</i>	32	32	0	21	0	1	10	0	1998	95	62
<i>PLTP</i>	36	36	0	20	0	0	16	0	2255	113	63
<i>PLA2G6</i>	52	52	0	26	1	1	24	0	3271	126	63
<i>MVK</i>	31	31	0	16	0	0	9	6	1952	122	63
<i>LPA</i>	103	103	0	67	1	0	35	0	6490	97	63
<i>KCNK17</i>	27	27	0	18	1	0	8	0	1706	95	63
<i>CMIP</i>	68	65	3	16	1	0	51	0	4357	272	64
<i>CYP26A1</i>	35	35	0	20	0	3	12	0	2245	112	64
<i>SLC16A11</i>	28	25	3	23	0	3	2	0	1803	78	64
<i>TOM1</i>	37	37	0	20	1	1	15	0	2390	120	65
<i>FN1</i>	133	132	1	79	0	0	54	0	8905	113	67
<i>ABCA1</i>	150	149	1	83	1	2	64	0	10494	126	70
<i>ANXA9</i>	26	26	0	16	0	0	10	0	1826	114	70
<i>GCKR</i>	31	30	1	16	0	1	14	0	2186	137	71
<i>LCAT</i>	21	21	0	13	1	0	7	0	1507	116	72
<i>NPC1L1</i>	70	70	0	51	1	0	18	0	5048	99	72
<i>ANGPTL4</i>	26	26	0	18	2	0	6	0	1879	104	72
<i>NAT2</i>	18	18	0	17	1	0	0	0	1322	78	73
<i>CPS1</i>	79	78	1	35	0	1	43	0	5821	166	74
<i>LACTB</i>	27	27	0	18	1	0	8	0	2023	112	75

<i>ASAP3</i>	54	53	1	31	0	0	23	0	4050	131	75
<i>RAB3GAP1</i>	65	65	0	37	0	0	27	1	4914	133	76
<i>GCK</i>	36	35	1	13	1	2	20	0	2729	210	76
<i>APOB</i>	186	186	0	146	0	0	40	0	14121	97	76
<i>APOA1</i>	16	16	0	11	0	0	5	0	1227	112	77
<i>LRP4</i>	105	105	0	56	0	0	49	0	8076	144	77
<i>RBM5</i>	41	40	1	19	0	0	22	0	3188	168	78
<i>AMPD3</i>	51	51	0	27	0	1	23	0	3994	148	78
<i>KAT5</i>	29	27	2	7	1	0	20	1	2296	328	79
<i>PHLDB1</i>	72	72	0	47	0	0	25	0	5753	122	80
<i>SNX5</i>	28	27	1	10	1	0	17	0	2288	229	82
<i>CILP2</i>	51	49	2	40	1	2	8	0	4199	105	82
<i>CYP1A1</i>	32	32	0	28	2	0	2	0	2685	96	84
<i>TMEM176A</i>	18	18	0	11	0	0	7	0	1530	139	85
<i>PMVK</i>	15	15	0	9	0	1	5	0	1280	142	85
<i>DLG4</i>	45	45	0	13	0	0	32	0	3975	306	88
<i>INSR</i>	100	99	1	55	0	0	45	0	8954	163	90
<i>FAM13A</i>	65	65	0	32	2	0	31	0	5858	183	90
<i>NR1H3</i>	30	29	1	11	1	2	10	6	2731	248	91
<i>LPL</i>	42	42	0	24	2	1	15	0	3846	160	92
<i>APOC3</i>	6	6	0	2	1	0	3	0	550	275	92
<i>MYLIP</i>	18	18	0	14	0	0	4	0	1666	119	93
<i>BRCA2</i>	117	116	1	88	0	2	27	0	10984	125	94
<i>MOSC1</i>	23	23	0	16	0	0	7	0	2170	136	94
<i>APOA2</i>	5	5	0	3	0	0	2	0	474	158	95
<i>FADS2</i>	38	38	0	6	4	0	28	0	3630	605	96
<i>VIM</i>	22	21	1	9	0	4	9	0	2128	236	97
<i>GALNT2</i>	46	46	0	16	2	0	28	0	4454	278	97
<i>SLC16A13</i>	18	18	0	11	0	1	6	0	1748	159	97
<i>SCARB1</i>	34	34	0	16	0	0	18	0	3329	208	98
<i>DGAT2</i>	25	25	0	13	1	0	11	0	2453	189	98
<i>AKT1</i>	39	39	0	12	0	1	26	0	3916	326	100
<i>MAMSTR</i>	18	18	0	12	1	0	5	0	1837	153	102
<i>MC4R</i>	14	13	1	14	0	0	0	0	1438	103	103
<i>RAF1</i>	32	31	1	18	0	0	14	0	3300	183	103

<i>EHBP1</i>	50	50	0	25	0	0	25	0	5165	207	103
<i>UGT1A1</i>	23	23	0	16	0	0	7	0	2382	149	104
<i>KLF4</i>	26	25	1	20	0	0	6	0	2813	141	108
<i>CYP7A1</i>	27	27	0	18	0	0	9	0	2950	164	109
<i>LIPG</i>	37	37	0	19	2	1	15	0	4161	219	112
<i>PABPC4</i>	27	27	0	9	0	0	18	0	3048	339	113
<i>JMJD1C</i>	76	75	1	47	0	2	27	0	8666	184	114
<i>KLF14</i>	13	13	0	12	1	0	0	0	1488	124	114
<i>SPTLC3</i>	33	33	0	13	0	3	17	0	3830	295	116
<i>SOX17</i>	20	19	1	18	0	0	2	0	2342	130	117
<i>ABO</i>	54	53	1	0	0	0	54	0	6341	0	117
<i>HDGF</i>	20	20	0	7	1	0	12	0	2380	340	119
<i>APOA5</i>	16	16	0	10	2	1	3	0	1929	193	121
<i>APOE</i>	10	10	0	8	1	0	1	0	1208	151	121
<i>LDLRAP1</i>	24	24	0	9	1	0	14	0	2940	327	123
<i>IRF2BP2</i>	38	38	0	31	0	2	5	0	4663	150	123
<i>ABCG4</i>	31	31	0	16	0	0	15	0	3849	241	124
<i>SNX13</i>	50	50	0	21	0	0	29	0	6357	303	127
<i>MET</i>	52	52	0	41	0	0	11	0	6635	162	128
<i>ZNF648</i>	28	28	0	26	1	1	0	0	3649	140	130
<i>HNFB4A</i>	36	36	0	21	2	1	12	0	4694	224	130
<i>NYNRIN</i>	58	58	0	48	0	0	10	0	7857	164	135
<i>ADH5</i>	19	19	0	8	0	0	11	0	2612	327	137
<i>GPAM</i>	46	46	0	24	0	0	22	0	6369	265	138
<i>PDE3A</i>	52	50	2	40	1	0	11	0	7576	189	146
<i>ATG7</i>	34	33	1	17	2	1	14	0	4959	292	146
<i>UBASH3B</i>	46	46	0	27	2	1	16	0	6912	256	150
<i>ANGPTL1</i>	22	22	0	17	0	1	4	0	3348	197	152
<i>BRAP</i>	26	26	0	11	0	0	15	0	4026	366	155
<i>MAP3K1</i>	45	45	0	30	0	0	15	0	7011	234	156
<i>SETD2</i>	52	52	0	40	0	1	11	0	8142	204	157
<i>CITED2</i>	12	11	1	10	1	0	1	0	1905	191	159
<i>VLDLR</i>	57	57	0	24	4	0	29	0	9163	382	161
<i>NROB2</i>	7	7	0	6	0	0	1	0	1156	193	165
<i>SBNO1</i>	66	66	0	25	0	0	41	0	10981	439	166

<i>TOP1</i>	22	22	0	6	1	2	13	0	3738	623	170
<i>ANGPTL3</i>	16	16	0	11	0	1	4	0	2805	255	175
<i>SORT1</i>	38	38	0	23	0	0	15	0	7028	306	185
<i>LRPAP1</i>	42	42	0	25	5	0	12	0	7819	313	186
<i>MTMR3</i>	46	46	0	25	0	1	20	0	8906	356	194
<i>VEGFA</i>	18	18	0	3	0	0	14	1	3535	1178	196
<i>IKZF1</i>	28	28	0	16	1	0	11	0	5760	360	206
<i>HMGCR</i>	22	22	0	9	2	0	11	0	4585	509	208
<i>HBS1L</i>	34	34	0	14	1	1	18	0	7163	512	211
<i>TTC39B</i>	49	48	1	25	0	0	24	0	10483	419	214
<i>COBLL1</i>	42	42	0	21	0	0	21	0	9367	446	223
<i>EVI5</i>	33	33	0	21	2	0	10	0	7403	353	224
<i>IRS1</i>	43	43	0	43	0	0	0	0	9705	226	226
<i>PHC1</i>	19	18	1	8	0	1	10	0	4292	537	226
<i>FRMD5</i>	22	22	0	14	0	0	8	0	5011	358	228
<i>SPTY2D1</i>	23	23	0	13	1	0	9	0	5754	443	250
<i>HFE</i>	20	20	0	9	0	2	9	0	5286	587	264
<i>ARL15</i>	9	9	0	6	0	1	2	0	2486	414	276
<i>MSL2</i>	16	16	0	14	0	2	0	0	4692	335	293
<i>CMTM6</i>	13	13	0	6	1	1	5	0	3863	644	297
<i>UBE2L3</i>	10	9	1	1	0	0	9	0	3028	3028	303
<i>TRIB1</i>	12	12	0	9	0	0	3	0	3635	404	303
<i>SLC39A8</i>	13	13	0	7	0	0	6	0	4032	576	310
<i>INSIG2</i>	8	8	0	3	1	0	4	0	2590	863	324
<i>MARCH8</i>	17	17	0	9	0	2	6	0	5636	626	332
<i>CSNK1G3</i>	13	13	0	6	0	0	7	0	4462	744	343
<i>TRPS1</i>	27	27	0	22	1	1	3	0	9990	454	370
<i>PPARA</i>	27	27	0	17	1	1	8	0	10048	591	372
<i>FADS1</i>	12	12	0	1	0	0	11	0	4523	4523	377
<i>RSPO3</i>	7	7	0	4	0	1	2	0	2697	674	385
<i>FRK</i>	33	33	0	19	1	0	13	0	13230	696	401
<i>KLHL8</i>	14	14	0	6	0	0	8	0	5621	937	402
<i>CTF1</i>	4	4	0	1	0	0	3	0	1664	1664	416
<i>GSK3B</i>	18	18	0	8	0	2	8	0	7711	964	428
<i>PIGV</i>	6	6	0	6	0	0	0	0	2713	452	452

<i>FAM117B</i>	12	12	0	5	0	0	7	0	5763	1153	480
<i>PPP1R3B</i>	11	11	0	9	1	0	1	0	5548	616	504
<i>FTO</i>	23	23	0	14	0	0	9	0	11766	840	512
<i>MAFB</i>	6	6	0	6	0	0	0	0	3393	566	566
<i>ZNF664</i>	7	7	0	6	0	1	0	0	5108	851	730

SNV: Variantes de un solo nucleótido; SNV cod: SNV en región codificante; *: Las longitudes de los transcritos fueron obtenidos del browser ENSEMBL (http://grch37.ensembl.org/Homo_sapiens/Gene/)

Tabla suplementaria 3 Variantes con asociación a rasgos lipídicos

SNV	Gen	Consecuencia	A.A.	BP	CHR	A1	MAF	OR	L95	U95	OR_P	BETA	Beta_P
Hipertrigliceridemia													
rs3825041	<i>BUD13</i>	intrón	-	116631707	11	T	0.21	1.714	1.321	2.222	4.94E-05	21.25	2.86E-03
rs5128	<i>APOC3</i>	3 prime UTR	-	116703640	11	G	0.26	1.628	1.285	2.063	5.47E-05	20.82	1.52E-03
rs4520	<i>APOC3</i>	sinónima	G	116701535	11	C	0.43	0.6607	0.54	0.8083	5.63E-05	-28.27	1.41E-06
rs651821	<i>APOA5</i>	5 prime UTR	-	116662579	11	C	0.21	1.653	1.281	2.134	1.12E-04	21.26	2.48E-03
rs2070665	<i>APOA1</i>	intrón	-	116707684	11	A	0.25	1.603	1.261	2.037	1.14E-04	17.16	9.95E-03
rs2072560	<i>APOA5</i>	intrón	-	116661826	11	T	0.22	1.572	1.222	2.023	4.34E-04	18.12	1.02E-02
rs1260326	<i>GCKR</i>	no sinónima, Splice región	P/L	27730940	2	T	0.30	1.466	1.174	1.831	7.41E-04	18.17	4.17E-03
rs12287066	<i>APOA5</i>	sinónima	I	116662331	11	T	0.17	1.567	1.197	2.052	1.09E-03	29.68	6.54E-05
rs5092	<i>APOA4</i>	sinónima	T	116693464	11	C	0.28	1.44	1.149	1.803	1.52E-03	22.67	4.24E-04
rs5104	<i>APOA4</i>	no sinónima	N/S	116692334	11	C	0.27	1.441	1.147	1.812	1.74E-03	21.66	8.47E-04
rs2305686	<i>ATG7</i>	intrón	-	11350553	3	C	0.40	0.7249	0.5906	0.8897	2.08E-03	-3.454	5.69E-01
rs2293571	<i>GCKR</i>	intrón	-	27729480	2	G	0.38	1.352	1.101	1.661	3.98E-03	18.88	1.53E-03
rs780807	<i>CUBN</i>	intrón	-	16948177	10	G	0.46	0.7523	0.6184	0.9151	4.42E-03	-7.66	1.87E-01
rs813592	<i>GCKR</i>	intrón	-	27721971	2	T	0.37	1.344	1.094	1.652	4.95E-03	20.5	5.95E-04
rs11820589	<i>BUD13</i>	no sinónima	P/L	116633862	11	A	0.17	1.45	1.116	1.886	5.48E-03	22.33	2.28E-03
rs690	<i>LIPC</i>	sinónima	V	58834741	15	T	0.40	1.322	1.084	1.612	5.89E-03	3.558	5.40E-01
rs5070	<i>APOA1</i>	intrón	-	116708020	11	G	0.36	0.7513	0.6129	0.9209	5.90E-03	-22.94	1.39E-04
rs5882	<i>CETP</i>	no sinónima	I/V	57016092	16	G	0.45	0.7569	0.6196	0.9247	6.41E-03	-9.874	9.54E-02
rs186754194	<i>ANGPTL4</i>	no sinónima	M/I	8429328	19	C	0.04	0.5431	0.3454	0.8539	8.19E-03	-24.31	8.19E-02
rs6725958	<i>FNI</i>	intrón	-	216262621	2	A	0.28	0.7506	0.6065	0.929	8.37E-03	-9.13	1.52E-01
rs78030300	<i>GALNT2</i>	intrón	-	230410150	1	T	0.03	2.729	1.291	5.77	8.58E-03	51.81	3.72E-03
rs8192700	<i>AKT1</i>	intrón	-	105238670	14	T	0.05	1.921	1.181	3.125	8.58E-03	10.72	3.95E-01
rs2071016	<i>AMPD3</i>	intrón	-	10518521	11	T	0.36	1.32	1.073	1.625	8.76E-03	6.758	2.66E-01
Hipercolesterolemia													
rs139047	<i>PNPLA3</i>	intrón	-	44323074	22	A	0.18	1.497	1.173	1.909	1.16E-03	3.739	1.40E-01
rs4888167	<i>CMIP</i>	intrón	-	81736324	16	T	0.15	1.51	1.148	1.987	3.24E-03	7.133	1.33E-02
rs34474737	<i>LIPG</i>	5 prime UTR	-	47088655	18	G	0.11	0.6418	0.4734	0.87	4.28E-03	-6.276	5.45E-02
rs1367117	<i>APOB</i>	no sinónima	T/I	21263900	2	A	0.28	1.344	1.09	1.657	5.64E-03	6.561	3.11E-03
rs429358	<i>APOE</i>	no sinónima	C/R	45411941	19	C	0.11	1.513	1.122	2.041	6.65E-03	5.063	1.06E-01
rs2389858	<i>SNX13</i>	intrón	-	17908130	7	T	0.44	1.298	1.073	1.57	7.35E-03	2.902	1.53E-01
rs35978968	<i>LIPG</i>	intrón	-	47088826	18	T	0.11	0.6626	0.4883	0.899	8.20E-03	-5.503	9.32E-02
rs3812316	<i>MLXIPL</i>	no sinónima	Q/H	73020337	7	G	0.05	0.5603	0.3634	0.8639	8.74E-03	-9.065	4.92E-02
rs10740107	<i>JMJD1C</i>	intrón	-	64944317	10	A	0.46	1.281	1.063	1.542	9.11E-03	6.554	9.84E-04

rs2239013	<i>APOA4</i>	intrón	-	116693353	11	T	0.02	0.435	0.2324	0.8143	9.26E-03	-14.52	2.77E-02
rs1049925	<i>PHC1</i>	no sinónima	T/A	9086898	12	G	0.45	0.7714	0.6342	0.9383	9.40E-03	-2.782	1.85E-01
rs2723501	<i>SNX13</i>	sinónima	P	17854532	7	G	0.43	1.293	1.065	1.571	9.48E-03	3.629	8.00E-02
Hipo HDL-C													
rs10488698	<i>BUD13</i>	no sinónima	R/C	116633947	11	A	0.15	0.511	0.3884	0.6722	1.59E-06	4.583	8.20E-11
rs17519093	<i>BUD13</i>	intrón	-	116629905	11	A	0.15	0.511	0.3884	0.6722	1.59E-06	4.583	8.20E-11
rs1532625	<i>CETP</i>	Splice región, intrón	-	57005301	16	T	0.40	0.6589	0.5367	0.809	6.78E-05	2.165	2.33E-05
rs2071303	<i>HFE</i>	Splice región, intrón	-	26091336	6	C	0.43	0.6627	0.5378	0.8165	1.12E-04	1.693	1.11E-03
rs2464195	<i>HNFL1A</i>	Splice región, intrón	-	121435475	12	A	0.38	0.6671	0.5409	0.8229	1.56E-04	1.513	4.06E-03
rs2254537	<i>GPAM</i>	sinónima	P	113917085	10	T	0.24	0.6462	0.5117	0.8161	2.45E-04	1.314	2.86E-02
rs735396	<i>HNFL1A</i>	intrón	-	121438844	12	C	0.38	0.6854	0.5589	0.8404	2.83E-04	1.465	4.28E-03
rs2464196	<i>HNFL1A</i>	no sinónima	S/N	121435427	12	A	0.37	0.6842	0.5565	0.8412	3.17E-04	1.394	7.35E-03
rs1169301	<i>HNFL1A</i>	intrón	-	121431300	12	T	0.37	0.6857	0.5578	0.8428	3.38E-04	1.392	7.29E-03
rs2792751	<i>GPAM</i>	no sinónima	V/I	113940329	10	T	0.24	0.6547	0.5192	0.8255	3.42E-04	1.364	2.22E-02
rs2259820	<i>HNFL1A</i>	sinónima	L	121435342	12	T	0.37	0.6878	0.5597	0.8451	3.69E-04	1.418	6.21E-03
rs1169288	<i>HNFL1A</i>	no sinónima	I/L	121416650	12	C	0.35	0.6914	0.5639	0.8477	3.88E-04	1.123	2.93E-02
rs1169294	<i>HNFL1A</i>	intrón	-	121426594	12	A	0.37	0.6814	0.5512	0.8424	3.92E-04	1.238	2.01E-02
rs1169289	<i>HNFL1A</i>	sinónima	L	121416622	12	G	0.40	0.7106	0.5823	0.867	7.65E-04	1.135	2.36E-02
rs1572982	<i>HFE</i>	intrón	-	26094367	6	G	0.46	1.421	1.144	1.765	1.47E-03	-1.552	4.10E-03
rs1800708	<i>HFE</i>	intrón	-	26093303	6	C	0.29	0.696	0.5553	0.8723	1.66E-03	1.291	2.74E-02
rs248	<i>LPL</i>	sinónima	E	19810826	8	A	0.01	0.2971	0.1366	0.6464	2.21E-03	7.364	4.92E-04
rs918144	<i>BUD13</i>	sinónima	P	116633825	11	C	0.39	1.373	1.114	1.693	2.92E-03	-1.679	1.06E-03
rs11076176	<i>CETP</i>	intrón	-	57007446	16	G	0.37	1.38	1.115	1.708	3.12E-03	-1.637	1.86E-03
rs2278426	<i>C19orf80</i>	no sinónima	R/W	11350488	19	T	0.34	1.411	1.121	1.776	3.37E-03	-1.002	7.35E-02
rs301	<i>LPL</i>	intrón	-	19816934	8	C	0.19	0.6907	0.5367	0.8888	4.03E-03	1.716	8.51E-03
rs78312568	<i>OSBPL7</i>	intrón	-	45894777	17	T	0.14	0.6657	0.5041	0.879	4.12E-03	1.272	7.74E-02
rs289714	<i>CETP</i>	intrón	-	57007451	16	G	0.37	1.362	1.102	1.684	4.22E-03	-1.528	3.46E-03
rs289741	<i>CETP</i>	intrón	-	57017474	16	G	0.31	0.7317	0.5868	0.9123	5.52E-03	1.828	1.19E-03
rs10378	<i>TMEM176A</i>	no sinónima	L/F	150501455	7	T	0.07	0.6025	0.4189	0.8665	6.28E-03	1.787	6.72E-02
rs3173798	<i>CD36</i>	Splice región, intrón	-	80285850	7	C	0.19	1.454	1.107	1.91	7.15E-03	-2.101	9.74E-04
rs343	<i>LPL</i>	intrón	-	19810787	8	A	0.13	0.6789	0.5098	0.9041	8.06E-03	1.906	1.04E-02
rs61737091	<i>FAM13A</i>	sinónima	P	89671721	4	T	0.02	4.11	1.424	11.87	8.98E-03	-3.185	8.14E-02
Hiper LDL-C													
rs2072114	<i>FADS2</i>	intrón	-	61605215	11	G	0.41	0.6802	0.5511	0.8395	3.30E-04	-6.456	2.02E-04
rs4888167	<i>CMIP</i>	intrón	-	81736324	16	T	0.15	1.633	1.236	2.158	5.60E-04	6.043	1.18E-02

rs2723501	<i>SNXI3</i>	sinónima	P	17854532	7	G	0.43	1.372	1.114	1.69	3.00E-03	3.33	5.64E-02
rs2292642	<i>PGSI</i>	Splice región, sinónima	G	76395430	17	C	0.30	0.7169	0.5749	0.8942	3.20E-03	-4.118	2.36E-02
rs2292643	<i>PGSI</i>	Splice región, intrón	-	76395421	17	G	0.30	0.7169	0.5749	0.8942	3.20E-03	-4.118	2.36E-02
rs2292644	<i>PGSI</i>	intrón	-	76394449	17	C	0.30	0.7169	0.5749	0.8942	3.20E-03	-4.118	2.36E-02
rs11742194	<i>HMGCR</i>	intrón	-	74646878	5	T	0.17	1.473	1.137	1.91	3.40E-03	7.302	9.55E-04
rs5909	<i>HMGCR</i>	3 prime UTR	-	74656175	5	A	0.17	1.473	1.137	1.91	3.40E-03	7.302	9.55E-04
rs3743003	<i>CAPN3</i>	intrón	-	42691664	15	G	0.02	0.3059	0.1362	0.6867	4.10E-03	-10.36	6.90E-02
rs267733	<i>ANXA9</i>	no sinónima	D/G	150958836	1	G	0.06	0.5502	0.3558	0.8507	7.20E-03	-5.346	1.13E-01
rs10740107	<i>JMJDIC</i>	intrón	-	64944317	10	A	0.46	1.307	1.07	1.596	8.70E-03	4.894	3.70E-03
rs3817294	<i>PGSI</i>	intrón	-	76396963	17	A	0.31	0.7443	0.5959	0.9298	9.30E-03	-4.709	1.06E-02

Tabla suplementaria 4 Asociación por haplotipos de genes con variantes

Gen	Haplotipo	Frq	Rasgo	OR	P ajustada	Beta	P ajustada
<i>APOA1</i>	AA	0.237	HTG	1.65	3.32E-05	18.9	5.40E-03
	GG	0.355		0.73	1.57E-03	-22.7	2.00E-04
<i>APOA4</i>	CC	0.273	HTG	1.47	1.08E-03	22	7.09E-04
	TT	0.717		0.702	2.05E-03	-22.2	5.50E-04
<i>APOA5</i>	CT	0.172	HTG	1.55	1.45E-04	28.7	1.20E-04
	TG	0.216		1.57	4.74E-04	17.3	1.50E-02
	CG	0.611		0.548	3.44E-08	-31.6	1.22E-07
	CGT	0.764	Hipo HDL-C	0.764	1.31E-02	1.58	2.00E-03
<i>APOC3</i>	TG	0.255	HTG	1.62	6.05E-05	20.8	1.54E-03
	CC	0.432		0.655	4.11E-05	-28.4	1.34E-04
	TG	0.255	Hipo HDL-C	1.38	2.80E-02	0.67	ns
<i>BUD13</i>	ATA	0.15	Hipo HDL-C	0.517	2.38E-06	4.58	8.20E-11
	GCG	0.388		1.36	3.91E-03	-1.68	1.00E-03
	TG	0.205		1.33	3.42E-02	-1.61	1.11E-02
	CG	0.621		0.747	7.57E-03	1.9	3.04E-04
	ACTGA	0.15		0.517	2.38E-06	4.58	8.20E-11
	GTCGG	0.203		1.35	2.00E-02	-1.7	7.49E-03
	ATA	0.15	HTG	0.708	1.10E-02	-22.2	5.70E-03
	GCG	0.388		1.24	3.10E-02	1.78	ns
	CA	0.17		1.45	5.82E-03	22.6	2.19E-03
	TG	0.205		1.72	4.90E-05	21.3	2.87E-03
	CG	0.621		0.545	2.37E-08	-29.6	6.17E-07
	ACTGA	0.15		0.708	1.10E-02	-22.2	5.70E-03
	GCTAG	0.173		1.46	5.23E-03	22.7	2.04E-03
	GTCGG	0.203		1.75	3.03E-05	22.6	1.60E-03
	GCTGG	0.286		0.75	7.70E-03	-4.22	ns
<i>CETP</i>	GG	0.365	Hipo HDL-C	1.37	3.93E-03	-1.61	2.00E-03
	TA	0.628		0.735	4.24E-03	1.55	2.98E-03
	TTAG	0.231		0.65	5.44E-04	2.26	2.89E-04
	CGGA	0.321		1.43	2.39E-03	-1.83	1.24E-03
<i>GCKR</i>	TGT	0.299	HTG	1.46	7.93E-04	17.9	4.77E-03
	CAC	0.613		0.721	1.94E-03	-19.6	9.89E-04
<i>GPAM</i>	TT	0.235	Hipo HDL-C	0.653	3.25E-04	1.32	2.76E-02
	AC	0.76		1.54	2.51E-04	-1.36	2.30E-02
<i>HFE</i>	CCA	0.289	HTG	0.734	6.59E-03	-15.2	2.10E-02
	TTG	0.457	Hipo HDL-C	1.44	1.03E-03	-1.55	4.10E-03

	CCA	0.289		0.7	2.46E-03	1.29	2.74E-02
<i>HMGCR</i>	TA	0.173	HTC	1.32	2.70E-02	7.25	5.60E-03
	CG	0.827		0.758	2.70E-02	-7.25	5.60E-03
	TA	0.173	Hiper LDL-C	1.48	2.88E-03	7.49	7.62E-04
	CG	0.827		0.673	2.88E-03	-7.49	7.62E-04
<i>HNF1A</i>	GC	0.345	Hipo HDL-C	0.705	7.51E-04	1.06	3.97E-02
	CA	0.596		1.42	5.55E-04	-1.19	1.71E-02
	TTAAC	0.358		0.677	2.97E-04	1.53	4.13E-03
	CCGGT	0.613		1.45	3.55E-04	-1.4	6.23E-03
	GCTTAAC	0.315		0.697	9.03E-04	1.27	1.91E-02
	CACCGGT	0.554		1.46	1.99E-04	-1.36	6.63E-03
	GCATTAAC	0.308		0.703	1.47E-03	1.22	2.82E-02
	CAGCCGGT	0.549		1.46	2.14E-04	-1.38	6.14E-03
<i>LIPG</i>	GT	0.102	HTC	0.632	3.71E-03	-6.39	ns
	TC	0.889		1.53	5.36E-03	5.43	ns
<i>LPL</i>	CAC	0.0128	Hipo HDL-C	0.229	8.37E-04	8.12	3.84E-04
	AGC	0.123		0.682	1.12E-02	1.73	2.46E-02
	CGT	0.8		1.48	1.76E-03	-1.91	2.86E-03
	AGC	0.718		0.705	1.60E-02	-28.4	1.05E-03
<i>PGSI</i>	CGCA	0.297	Hiper LDL-C	0.72	4.09E-03	-3.95	3.30E-02
	TATG	0.681		1.37	5.62E-03	4.25	2.04E-02
<i>SNX13</i>	GT	0.414	HTC	1.28	1.14E-02	2.92	ns
	TA	0.548		0.771	6.86E-03	-3.6	ns
	GT	0.414	Hiper LDL-C	1.32	8.55E-03	2.63	ns
	TA	0.548		0.728	1.91E-03	-3.29	ns

Haplotipos con asociación a rasgos lipídicos. Los bloques de haplotipos fueron generados por el programa plink1.07. Frq: Frecuencia; ns: no significativo