

# Influencia de la combinación de SNPs sobre dislipidemias y parámetros lipídicos

Trabajo de Fin de Grado

Nerea Cerezuela Filgueras

Grado de Biotecnología



UNIVERSITAT  
ROVIRA i VIRGILI

Tutora: Dra. Katihuska Viviana Paredes Aguilar

Tutora de prácticas: Dra. Montserrat Guardiola Guionnet

Universitat Rovira i Virgili

Tarragona, junio 2019

# Índice

Agradecimientos.....	2
Índice de abreviaciones.....	3
Datos del centro .....	4
Resumen y palabras clave .....	5
<b>1. Introducción.....</b>	<b>6</b>
<b>1.1. Dislipidemias.....</b>	<b>7</b>
1.1.1. Dislipidemias primarias .....	8
1.1.2. Dislipidemias secundarias .....	9
1.1.3. Futuro de las investigaciones .....	10
<b>2. Hipótesis del trabajo y objetivos .....</b>	<b>12</b>
<b>3. Metodología .....</b>	<b>13</b>
<b>3.1. Población estudiada .....</b>	<b>13</b>
<b>3.2. Obtención de muestras y análisis .....</b>	<b>14</b>
3.2.1. Determinación de parámetros lipídicos y bioquímicos.....	14
3.2.2. Capa leucocitaria para obtención de ADN y realización del genotipado .....	15
<b>3.3. Selección de los SNPs para el genotipado .....</b>	<b>15</b>
<b>3.4. Determinación del grosor entre la íntima y la media (IMT).....</b>	<b>16</b>
<b>3.5. Distribución de grupos .....</b>	<b>16</b>
<b>3.6. Análisis estadístico .....</b>	<b>17</b>
<b>4. Resultados y discusión.....</b>	<b>18</b>
<b>4.1. Descripción de la población .....</b>	<b>18</b>
<b>4.2. Descripción de los genotipos de la población .....</b>	<b>20</b>
<b>4.3. Estudio de la relación entre los SNPs, las dislipidemias y los factores de riesgo cardiovascular .....</b>	<b>21</b>
4.3.1. Asociación de SNPs a dislipidemias.....	22
4.3.2. Asociación de SNPs a parámetros lipídicos.....	25
4.3.3. Estudio del efecto aditivo de varios SNPs sobre el cLDL.....	30
4.3.4. Distribución de las dislipidemias sobre los grupos de alelos .....	31
4.3.5. Asociación de SNPs a otros parámetros de las LDLs .....	33
<b>5. Conclusión.....</b>	<b>36</b>
<b>6. Bibliografía.....</b>	<b>38</b>
<b>7. Autoevaluación.....</b>	<b>44</b>
<b>8. Anexo .....</b>	<b>45</b>

## Agradecimientos

Me gustaría empezar agradeciendo a todas las personas que me han ayudado para que este trabajo sea posible.

En primer lugar, quiero agradecer a mi familia por haberme brindado la oportunidad de estar haciendo lo que me gusta y por haberme apoyado durante todo el proceso.

Gracias también a todo el equipo de la URLA por haberme acogido en el despacho con los brazos abiertos desde el primer momento. También quiero agradecerles que me hayan ofrecido su ayuda de manera desinteresada ante cualquier duda.

También quiero agradecer al Dr. Josep Ribalta por sus ideas de enfoque del estudio, así como sus comentarios cargados de experiencia.

Quiero destacar el esfuerzo de mi tutora, la Dra. Katihuska Paredes, por sus guías e indicaciones a lo largo de todo el trabajo. Estoy especialmente agradecida por las muchas tardes empleadas conmigo en las correcciones del trabajo.

Pero, sobre todo, gracias a mi tutora de prácticas la Dra. Montserrat Guardiola que me aceptó en el equipo y me ha enseñado tantísimo sobre el tema. Sin olvidar la gran ayuda durante todo el proyecto, los consejos y la calidad de todos sus comentarios.

## Índice de abreviaciones

<b>Nombre abreviado</b>	<b>Nombre completo</b>
<b>APOA1</b>	Apolipoproteína A-1
<b>APOB</b>	Apolipoproteína B
<b>cHDL</b>	Colesterol de las HDL
<b>cLDL</b>	Colesterol de las LDL
<b>Combinada</b>	Hiperlipidemia combinada
<b>DA</b>	Dislipidemia aterogénica
<b>DM2</b>	Diabetes tipo 2
<b>ECV</b>	Enfermedades cardiovasculares
<b>HDL</b>	Lipoproteína de alta densidad
<b>HiperCOL</b>	Hipercolesterolemia aislada
<b>HiperTG</b>	Hipertrigliceridemia aislada
<b>HW</b>	Hardy Weinberg
<b>IDL</b>	Lipoproteína de densidad intermedia
<b>IMC</b>	Índice de masa corporal
<b>IMT</b>	<i>Intima-media thickness</i>
<b>Q</b>	Quilomicrones
<b>LDL</b>	Lipoproteína de baja densidad
<b>MAF</b>	Frecuencia del alelo minoritario
<b>mRNA</b>	RNA mensajero
<b>PCR</b>	Proteína C reactiva
<b>cRLP</b>	Colesterol de las lipoproteínas remanentes
<b>SM</b>	Síndrome metabólico
<b>SNP</b>	Polimorfismo de un único nucleótido
<b>TG</b>	Triglicéridos
<b>TRL</b>	Lipoproteína rica en triglicéridos
<b>VLDL</b>	Lipoproteína de muy baja densidad

## Datos del centro

La realización del Trabajo de Fin de Grado ha sido posible gracias a un periodo de prácticas en la URLA (Unitat de Recerca en Lípids i Arteriosclerosi) de la URV, concretamente en la línea de investigación centrada en la genética y epigenética de las dislipidemias y riesgo cardiovascular. Este grupo de investigación se encuentra en la Facultad de Medicina y Ciencias de la Salud de Reus (C/ Sant Llorenç, 21 43201 Reus).

## Resumen y palabras clave

Las enfermedades cardiovasculares suponen uno de los mayores problemas en los países desarrollados. El aumento de los lípidos circulantes supone uno de los principales factores de riesgo para padecer este tipo de enfermedades. El control genético de las dislipidemias implica la participación de una o pocas mutaciones causales en alguno de los genes clave para el metabolismo lipídico (en las dislipidemias primarias), o bien, la participación de un conjunto de variaciones genéticas en distintos puntos del genoma implicados en la regulación de los lípidos circulantes junto con factores ambientales (en las dislipidemias secundarias).

Se ha seleccionado una población de 480 individuos que incluye tanto individuos sanos como pacientes con elevado riesgo cardiovascular, de los que se dispone de un perfil lipídico detallado. Se han seleccionado 45 SNPs localizados en 25 genes involucrados en el metabolismo lipídico, factores de inflamación y en la patogenia de la arteriosclerosis. Se han seleccionado con la finalidad de encontrar aquellos SNPs asociados a las principales dislipidemias y a los lípidos circulantes, y que puedan suponer buenos marcadores para predecir el riesgo a sufrir desordenes lipídicos y por consiguiente mayor riesgo de sufrir enfermedades relacionadas.

Se ha podido encontrar relación entre determinados SNPs y las dislipidemias más frecuentes: 4 SNPs relacionados con la hipercolesterolemia aislada; 2 SNPs relacionados con la hipertrigliceridemia aislada; 4 SNPs relacionados con la hiperlipidemia combinada; y 1 con la dislipidemia aterogénica.

Hemos encontrado que los SNPs rs2948080 (en el gen *KAT2B*), rs540796 y rs562556 (en *PCSK9*), rs1042034 (en *APOB*), rs3135506 y rs964184 (en *APOA5*) y rs4520 (en *APOC3*) se asocian significativamente a la concentración de cLDL y con el número de partículas LDL circulantes. También se ha podido observar como hay un efecto dosis-dependiente según el número de alelos de los SNPs relacionados con el cLDL que presentan los sujetos.

**Palabras clave:** Dislipidemia, riesgo cardiovascular, hipercolesterolemia, SNP, cLDL, *genetic risk score*.

## 1. Introducción

En las últimas décadas se ha incrementado el número de personas afectadas por diferentes desórdenes lipídicos, conocidos como dislipidemias. Éstos junto con la arteriosclerosis, la obesidad, la hipertensión arterial, la diabetes y un estilo de vida poco saludable, comportan un aumento en el riesgo de sufrir enfermedades cardiovasculares (ECV). (Zhong *et al.*, 2019)

Las ECV son una de las principales causas de mortalidad en todo el mundo. El Instituto Nacional de Estadística publicó que en España durante el año 2017 las enfermedades del sistema circulatorio fueron la primera causa de muerte (28.8% del total de defunciones), por delante de los tumores (26.7%). (Instituto Nacional de Estadística, 2017)

En relación a los principales factores de riesgo de ECV, el Ministerio de Sanidad, Consumo y bienestar social, sitúa como principal problema crónico de salud a la hipertensión, el cual sufre un 20% de la población. En tercera posición se sitúa el colesterol alto con un 18% y la diabetes con un 8%. La obesidad no está considerada como un problema de salud crónico, sin embargo, supone un factor muy preocupante que afecta al 17% de los adultos y a un 10% de las personas en edad infantil (entre 2 y 17 años). (Ministerio de Sanidad, Consumo y Bienestar social, Sanidad en datos: <https://www.mscbs.gob.es/estadEstudios/sanidadDatos/home.htm> (accedido 22 de Mayo, 2019))

La arteriosclerosis, principal responsable de las enfermedades cardiovasculares, es un proceso patológico de carácter inflamatorio y lipídico que afecta a los vasos sanguíneos de tamaño medio-grande. Es una enfermedad que progresa lentamente con el tiempo y no produce manifestaciones clínicas hasta llegar a estadios avanzados. Se caracteriza por la infiltración y acumulación de lípidos y componentes del sistema inmunitario, la proliferación de células musculares lisas y acumulación de componentes del tejido conectivo en la pared vascular. Esto da lugar a la formación de placas de ateroma, que

conllea una pérdida de elasticidad y un engrosamiento de la arteria hasta poder llegar a taponarlas. (Insull, 2009)

### 1.1. Dislipidemias

Los niveles de lípidos circulantes son uno de los principales factores de riesgo de ECV. Debido a que el desajuste en las concentraciones de lípidos se considera como un factor de riesgo para muchas enfermedades, sus causas y posibles tratamientos se han convertido en el foco de estudio. (Zhong *et al.*, 2019) Las dislipidemias son enfermedades relacionadas con un perfil lipídico alterado. Las lipoproteínas plasmáticas son partículas que tienen como función transportar las moléculas lipídicas hidrofóbicas a través de la sangre. Dependiendo de su composición o su origen, podemos clasificar las lipoproteínas en: quilomicrones (Q) (muy ricas en triglicéridos y se sintetizan en el intestino), VLDL (del inglés *very low-density lipoprotein*) (muy ricas en triglicéridos y se sintetizan en el hígado), IDL (del inglés *intermediate-density lipoprotein*) (transportan colesterol y triglicéridos), LDL (del inglés *low-density lipoprotein*) (transportan colesterol a los tejidos) y HDL (del inglés *high-density lipoprotein*) (transportan el exceso de colesterol desde los tejidos al hígado para ser eliminado).

Los parámetros básicos que se miden para el diagnóstico clínico de los desórdenes lipídicos son las concentraciones de colesterol LDL (cLDL), de colesterol HDL (cHDL) y triglicéridos (TG). No obstante, hay gran variedad de desórdenes lipídicos que se diferencian por las concentraciones de éstos y otros parámetros como la apolipoproteína B (APOB) o la apolipoproteína A-1 (APOA1) que son las principales proteínas de estas lipoproteínas y las responsables de la interacción de la lipoproteína con diferentes enzimas, receptores etc., y también se evalúan para un diagnóstico más completo. (Schaefer *et al.*, 2019)

Este tipo de enfermedades son muy complejas al tener una importante influencia ambiental y genética. La dieta, la edad, el sexo, el ejercicio, el tabaco, el alcohol o incluso ciertos medicamentos son algunos de los factores que pueden alterar las concentraciones de lípidos circulantes. Pero por otro lado, también existe un importante

componente genético como pueden ser mutaciones en genes clave o variantes genéticas que como resultado alteran los parámetros lipídicos. (Zhong *et al.*, 2019)

Las dislipidemias pueden clasificarse en dos tipos; las dislipidemias primarias que están causadas por una o pocas mutaciones causales y hay muy poca influencia ambiental, y las dislipidemias secundarias que tienen un origen más complejo debido a la interacción de varias variantes genéticas que actúan a la vez, junto con la influencia de factores ambientales.

#### 1.1.1. Dislipidemias primarias

Están causadas por mutaciones de uno o pocos genes, en estos casos el diagnóstico y el tratamiento son relativamente sencillos y ya han sido ampliamente estudiados.

En el caso de la hipercolesterolemia familiar (HF), el cribado diagnóstico suele implicar el estudio de mutaciones en varios genes: *LDLR* (responsable del 90% de los casos diagnosticados de HF); *APOB* (responsable del 2-5% de los casos); *PCSK9* (responsable en menos del 1%) y en menor frecuencia *LDLRAP1* y *LIPA*. (Sturm *et al.*, 2018) Mutaciones en estos genes hacen que se acumulen las lipoproteínas LDL en circulación aumentando los niveles de colesterol.

En el caso de la hipertrigliceridemia primaria, actualmente hay cinco genes identificados como principales responsables: *LPL*, *APOC2*, *GPIHBP1*, *APOA5*, y *LMF1*. Estos cinco genes codifican para proteínas clave en el proceso de lipólisis y por ello, mutaciones de pérdida de función en estos genes dificultan el metabolismo de los TG que se acaban acumulando en circulación. (Surendran *et al.*, 2012)

En estos casos la detección de las mutaciones se hace mediante un estudio familiar, secuenciando los genes susceptibles de haber sufrido algún cambio e identificando las mutaciones causales. Para casos en los que se sospecha de una mutación ya conocida y solo se desea realizar el diagnóstico se puede utilizar la técnica de Sanger o la *next-generation DNA sequencing*. (Berberich & Hegele, 2019a; Berberich & Hegele, 2019b) En el caso de la hipercolesterolemia familiar, existe un chip de diagnóstico comercial (Lipochip® Progenika Biopharma) que detecta las principales mutaciones descritas, facilitando el diagnóstico.

No obstante, las dislipidemias primarias son muy poco frecuentes en comparación con las dislipidemias secundarias. La forma homocigota de la hipercolesterolemia familiar tiene lugar en alrededor de 1/1.000.000 personas, y las hipertrigliceridemias primarias tienen una prevalencia de 1/100.000-1.000.000. Las dislipidemias secundarias son mucho más prevalentes y más desconocidas a nivel genético, y por este motivo este estudio se centra en ellas. (Catapano *et al.*, 2016; Lamiquiz-Moneo *et al.*, 2016)

### 1.1.2. Dislipidemias secundarias

Sin embargo, en la mayoría de los casos las dislipidemias tienen una compleja base genética que consiste en una acumulación de pequeños efectos causados por polimorfismos frecuentes, y de hecho, el efecto sumatorio de estos polimorfismos puede ser igual o superior al que tiene una mutación causal. (Futema *et al.*, 2015; Khera *et al.*, 2018) Además, no hay que olvidar la influencia ambiental, ya que un SNP que por ejemplo aumenta los niveles de TG, puede tener poco efecto en un individuo sano y si ese sujeto pasa a tener una peor alimentación o sufre alguna enfermedad metabólica, el efecto del SNP es mayor. (Ribalta *et al.*, 2002)

Los polimorfismos de un solo nucleótido (SNPs) son las formas más comunes de variaciones genéticas del genoma humano. Pueden tener implicaciones a nivel funcional como pueden ser el cambio de un aminoácido, un cambio en la estabilidad del mRNA, la creación o destrucción del lugar de unión de un microRNA o también puede producirse un cambio en la afinidad de unión a un factor de transcripción. No obstante, estos SNPs también pueden ser marcadores de alguna variante muy próxima pero desconocida. Por este motivo suelen ser las variaciones más utilizadas para los análisis genéticos de estas enfermedades. (Bush & Moore, 2012)

Hasta hace poco tiempo, el estudio de la genética detrás de estas dislipidemias más complejas se basaba en el análisis de los SNPs de genes candidatos conocidos individualmente intentando asociarlos a la enfermedad. No obstante, desde hace unos años se utilizan los *genome wide association studies* (GWAS) que analizan el efecto de los SNPs de todo el genoma para encontrar nuevas regiones genómicas o loci que

puedan estar asociados a ciertos rasgos o enfermedades. Los GWAS se pueden utilizar para relacionar varios SNPs al mismo tiempo con una misma enfermedad a través del estudio de su efecto combinado. (Berberich & Hegele, 2019a)

### 1.1.3. Futuro de las investigaciones

Actualmente, los mayores beneficios en la aplicación de estos estudios han sido en el área de la farmacología y la medicina personalizada. En farmacología el objetivo principal es identificar las variaciones genéticas asociadas al metabolismo de los medicamentos, su eficacia y sus efectos adversos. En algunos casos, diferentes variaciones pueden determinar una terapia particular y la dosis apropiada para cada paciente. (Bush & Moore, 2012)

Con respecto a la medicina personalizada estos estudios permitirían predecir y prevenir al paciente frente a ciertas enfermedades. En el caso de las dislipidemias, la medicina personalizada debe centrarse en las formas poligénicas muy frecuentes, pero poco estudiadas. Para ello sería necesario realizar un genotipado del paciente para observar que tipo de variaciones alélicas posee y relacionarlo con los conocimientos que se tienen sobre éstas. Para simplificar este proceso se pueden realizar algoritmos predictivos que incluyan todas las variantes genéticas implicadas en la patogenia de una determinada enfermedad o característica a nivel fenotípico. Estos algoritmos dan como resultado un valor “*genetic risk score*” que representa el efecto acumulativo de las variaciones estudiadas (normalmente SNPs) aumentando su poder predictivo a medida que se suman más polimorfismos.

A nivel clínico se están empezando a utilizar estos “*genetic risk scores*” que permiten mejorar el diagnóstico genético de las dislipidemias más severas y así, clasificar correctamente los pacientes, y con ello podemos evaluar mejor su posible riesgo cardiovascular asociado o sus complicaciones clínicas. (Rosenson *et al.*, 2017)

No obstante, estos algoritmos presentan ciertas limitaciones. Una de ellas es que se utilicen para el cálculo del *genetic risk score* una cantidad de variaciones alélicas muy inferior a la cantidad que realmente está causando el fenotipo. En consecuencia, será

muy difícil predecir realmente el riesgo de la persona de sufrir un determinado fenotipo. Además existe la limitación de utilizar para la elaboración de los algoritmos datos de asociaciones erróneas debido a que en el momento del estudio había otros parámetros que afectaban de forma indirecta o se trabajaba con un volumen de población demasiado bajo. (Holmes *et al.*, 2010) Por este motivo es importante disponer de la mayor cantidad de datos fiables que afectan al fenotipo estudiado antes de crear un algoritmo predictivo.

En resumen, el estudio de los SNPs candidatos de ser marcadores de dislipidemias es necesario para descubrir los mecanismos detrás de estas enfermedades y abrir las puertas a una medicina más personalizada y efectiva.

## 2. Hipótesis del trabajo y objetivos

### Hipótesis:

Cada vez hay más evidencias que indican que el estudio conjunto de varios SNPs, es muy informativo para la evaluación de la predisposición genética a enfermedades complejas. Queremos utilizar esta estrategia para conocer la variabilidad genética implicada en el proceso de arteriosclerosis. Por ello hipotetizamos que estudiar el efecto conjunto de varios SNPs en genes reguladores de los principales factores de riesgo arteriosclerótico (lípidos, inflamación, y patogenia de la arteriosclerosis) nos permitirá caracterizar mejor la predisposición genética en una cohorte con elevado riesgo cardiovascular.

### Objetivos:

- Identificar los SNPs que presentan una diferente frecuencia alélica según la clasificación de los sujetos en las principales dislipidemias (hipercolesterolemia aislada, hipertrigliceridemia aislada, hiperlipidemia combinada y dislipidemia aterogénica).
- Identificar los SNPs asociados a los principales factores de riesgo arteriosclerótico (cLDL, cHDL, cRLP, TG, IMT y PCR).
- Estudiar el efecto conjunto de los SNPs identificados, en relación a los factores de riesgo cardiovascular.

### 3. Metodología

#### 3.1. Población estudiada

Para este estudio se analizaron los datos de 480 individuos. La población está compuesta por sujetos sanos sin ECV ni enfermedades crónicas diagnosticadas (n=258). También se compone de pacientes con elevado riesgo cardiovascular (n=222) que sufren alguna alteración metabólica según los criterios del *Adult Treatment Panel III* (criterio para definir la presencia de síndrome metabólico (SM) y diabetes mellitus tipo 2 (DM2)).

Los participantes forman parte de otros estudios que han sido descritos en mayor detalle anteriormente. (Rock *et al.*, 2001; Guardiola *et al.*, 2015)

Para evitar que los medicamentos hipolipemiantes que tomaban los pacientes con alto riesgo cardiovascular interfiriesen con los resultados, se les pidió que dejaran de tomar la medicación durante las seis semanas antes de tomar las muestras para los análisis.

Todos los participantes en el estudio firmaron un consentimiento informado y los estudios fueron aprobados por el Comité Ético de los hospitales implicados en su reclutamiento.

De todos los participantes, disponemos de información básica de la población (edad, género, índice de masa corporal (IMC)...) así como de datos antropométricos y bioquímicos (glucosa, proteína C reactiva (PCR)). En relación al perfil lipídico disponemos de la concentración en plasma de colesterol total, cLDL, cHDL, TG, y del perfil lipídico determinado mediante resonancia magnética nuclear. Éste último permite obtener datos sobre el número y tamaño (grandes, medianas y pequeñas) de las principales lipoproteínas circulantes (VLDL, LDL y HDL).

De los pacientes con elevado riesgo cardiovascular disponemos de la medida del grosor entre la íntima y la media (IMT, del inglés intima-media thickness) que utilizamos como marcador de la arteriosclerosis subclínica.

## 3.2. Obtención de muestras y análisis

Tras una noche en ayunas de los participantes, se obtuvieron muestras de sangre venosa en tubos con EDTA. Se centrifugó inmediatamente a 1500g a 8°C durante 10 minutos. Se separó el plasma y se dividió en alícuotas que se guardaron a -80°C hasta la determinación de las variables analíticas. Para aislar el ADN se recogió la capa leucocitaria.

### 3.2.1. Determinación de parámetros lipídicos y bioquímicos

Para cuantificar el colesterol total y los triglicéridos se utilizó un método colorimétrico enzimático indirecto, mientras que para el cHDL se determinó con uno directo. El cLDL se calculó utilizando la fórmula de Friedewald. Las apolipoproteínas AI y B100 se midieron con inmunturbidometría utilizando antisueros específicos para apoA-I y apoB (Hoffman-La Roche). Los niveles de partículas de colesterol remanentes (RLPc) en plasma se midieron como se describe en el artículo de Nakajima *et al.*, (1993) utilizando el kit *RLP-Cholesterol Assay* (Jimro-II, Japan Immunoresearch Laboratories) En resumen, la técnica se basa en la adición de EDTA con plasma en un gel de suspensión con apoA-I y apoB-100 anti-humano de anticuerpos monoclonales de ratón unidos a Sepharose®. La fracción que no se ha unido será la que contenga las partículas remanentes de las que se toma una fracción para cuantificar el colesterol mediante un autoanalizador Cobas Mira (Roche).

Las concentraciones de subclases de partículas de lipoproteínas y el tamaño medio de las partículas de lipoproteínas se midieron utilizando espectrometría RMN (resonancia magnética nuclear) (LipoScience). Las concentraciones de las diferentes subclases y de los distintos tamaños se obtuvieron directamente midiendo las amplitudes de las señales del RMN. Los tamaños medios de las partículas de lipoproteína se obtuvieron de la suma del diámetro de cada subclase multiplicada por su porcentaje de masa relativo basado en la amplitud de su señal de RMN. También se midieron las concentraciones de las distintas subclases. (Mallol *et al.*, 2015)

### 3.2.2. Capa leucocitaria para obtención de ADN y realización del genotipado

Los pacientes fueron genéticamente caracterizados a través del ADN aislado de la capa leucocitaria. Los genotipos de los 45 SNPs escogidos se determinaron utilizando *Sequenom MassARRAY iPLEX platform* (Sequenom, San Diego, CA), basado en el método MALDI-TOF. Posteriormente el genotipado se llevó a cabo en Centro Nacional de Genotipado Español.

### 3.3. Selección de los SNPs para el genotipado

De todos los pacientes se disponía del genotipado de 45 polimorfismos (SNPs) presentes en 25 genes que pertenecen a genes clave en el metabolismo lipídico o por su implicación en la patogenia de la arteriosclerosis (Tabla Anexo 1).

Los genes se escogieron por su implicación en:

- El control de los principales parámetros lipídicos (cHDL, cLDL y TG): *APOA5*, *APOB*, *APOC3*, *APOE*, *CETP*, *GCKR*, *LIPC*, *LPL*, *MTTP* y *PCSK9*.
- El control de la síntesis y eliminación de lipoproteínas remanentes: *APOA5*, *APOE*, *CETP*, *FBLN2*, *FBNL5*, *LIPC*, *LPL*, *NDST1*, *SCD*, *SCARB1* y *VLDLR*.
- El control de la patogenia de la arteriosclerosis: *CASP3*, *FASLG*, *IL1A*, *IL1B*, *IL1RN*, *NLRP3*, *PCSK9*, *TLR4* y *VEGFA*.
- El control epigenético y que se asocia a una reducida enfermedad cardiovascular: *KAT2B*. (Pons *et al.*, 2011)

De todos estos genes se escogieron los SNPs que se encontraban ampliamente estudiados en estudios anteriores o que eran tagSNPs (International HapMap Consortium, 2005) (Tabla Anexo 1) y que pueden ser tanto marcadores como variantes funcionales.

En la Tabla Anexo 1 se puede observar una breve descripción de cada uno de los genes y sus funciones así como el número de personas de las que se obtuvieron resultados para cada SNP.

### 3.4. Determinación del grosor entre la íntima y la media (IMT)

Las fases iniciales de lesión arteriosclerótica se pueden diagnosticar midiendo el grosor entre la íntima y la media (IMT) que son las capas de la pared vascular más internas. La técnica para determinar el IMT es relativamente económica, no invasiva, fiable y reproducible para la detección precoz de la arteriosclerosis subclínica. Para ello se utilizó un sonógrafo MyLab 50 XVision (Esaote SA, España) equipado con una sonda de ultrasonido de matriz lineal de banda variable (5-12MHz) para identificar y digitalizar la pared de las arterias carótida común izquierda y derecha a 1cm de la parte proximal de la bifurcación; el bulbo carotideo en la bifurcación; y la arteria carótida interna a 1 cm de la parte distal de la bifurcación. Las medidas de IMT se representan como puntos predefinidos usando el procesador de imágenes Thicksoftware y se determina mediante medidas en tres imágenes de las arterias carotideas izquierda y derecha para obtener el IMT medio que es valor con el que trabajamos.

### 3.5. Distribución de grupos

Durante el estudio se ha clasificado la población en varios subgrupos en función de las dislipidemias más frecuentes y con más relevancia en cuanto al riesgo cardiovascular; hipercolesterolemia aislada (HiperCOL), hipertrigliceridemia aislada (HiperTG), hiperlipidemia combinada (Combinada) y la dislipidemia aterogénica (DA). La categoría HiperCOL está formada por los participantes que solo presentan concentraciones elevadas de colesterol en las LDL (cLDL). La categoría HiperTG está formada por aquellos participantes que solo presentan concentraciones elevadas de TG. En la Combinada encontramos los participantes que tienen a la vez concentraciones elevadas de cLDL y de TG. Y la DA contempla los participantes que además de niveles altos TG, también presentan niveles bajos de cHDL. (Taskinen, 2002)

Para definir cada uno de los grupos se han tenido en cuenta los valores que indican las guías clínicas tanto para el cLDL, como para TG y cHDL. Los valores se pueden ver en la Tabla 1. (Álvarez A *et al.*, 2012)

Tabla 1. Concentraciones de los parámetros de cLDL, cHDL, TG y cHDL para la definición de cada dislipidemia.

Tipo de desorden lipídico	Concentraciones de parámetros
<b>Hipercolesterolemia aislada (HiperCOL)</b>	cLDL > 3,3 mmol/L
	TG < 1,7 mmol/L
<b>Hipertrigliceridemia aislada (HiperTG)</b>	cLDL < 3,3 mmol/L
	TG > 1,7 mmol/L
<b>Hiperlipidemia combinada (Combinada)</b>	cLDL > 3,3 mmol/L
	TG > 1,7 mmol/L
<b>Dislipidemia aterogénica (DA)</b>	cLDL < 3,3 mmol/L
	TG > 1,7 mmol/L
	cHDL < 1,0 mmol/L en hombres cHDL < 1,2 mmol/L en mujeres

### 3.6. Análisis estadístico

Para el análisis se ha utilizado el programa estadístico “IBM SPSS Statistics” versión 25. Se ha empleado el cálculo de “frecuencias” para determinar la *minor allele frequency* (MAF) de la población.

Se han utilizado los estudios “ANOVA” para comparar medias entre dos grupos. Para determinar si había diferencias entre grupos de varias categorías se ha empleado el “test de Bonferroni”.

También se han utilizado las “tablas cruzadas” para encontrar diferencias significativas entre variables categóricas como entre las frecuencias de cada alelo en diferentes grupos de población, y también para el cálculo del equilibrio de Hardy Weinberg.

En los estudios de asociación de los SNPs con factores de riesgo cardiovascular, los resultados fueron ajustados por edad, sexo e IMC, ya que estos factores afectan de forma directa sobre las concentraciones de lípidos y proteínas en circulación. Para ello se empleó el “Modelo lineal general univariado” que permite definir variables dependientes (en este caso sería el parámetro a ajustar), factores fijos (el SNP del que se quiere estudiar el efecto) y covariables (en este estudio han sido la edad, el sexo y el IMC).

Se han considerado significativos valores de  $p \leq 0,05$ .

## 4. Resultados y discusión

### 4.1. Descripción de la población

Para este estudio hemos analizado datos de 480 individuos de los que 258 son sujetos sanos y 222 son pacientes con alto riesgo cardiovascular.

De forma conjunta, (Tabla 2) la población se encuentra en los  $53.60 \pm 14.81$  años de media. Está mayoritariamente compuesta por hombres (71.50%) debido a que, entre los sujetos sanos, hay mayoritariamente hombres. (Rock *et al.*, 2001)

Un grupo de 222 pacientes con riesgo cardiovascular formado por diabéticos tipo 2, pacientes con síndrome metabólico y obesos. En este grupo las mujeres representan un 45.50%, la edad media del grupo es de  $61.50 \pm 10.50$  años y el promedio del IMC es de  $31.47 \pm 4.91$  kg/m<sup>2</sup>. El segundo grupo está formado por 258 individuos sanos con un 13.95% de mujeres, la edad media del grupo es de  $46.81 \pm 14.61$  años y el IMC promedio es de  $24.86 \pm 2.62$  kg/m<sup>2</sup>. Teniendo en cuenta estas diferencias, los análisis estadísticos están ajustados por género, edad e IMC.

Lo primero que hicimos fue subdividir toda la población en varias categorías según las distintas dislipidemias estudiadas; HiperCOL (n=126), HiperTG (n=62), Combinada (n=74) y DA (n=22) según los criterios descritos en la Tabla 1. Para los datos del IMC se puede observar una media de  $27.91 \pm 5.07$  kg/m<sup>2</sup> en el total de la población. En cambio, para las categorías de HiperTG, Combinada y DA el valor es un 12.43%, 13.51% y un 15.91% superior respectivamente sobre el conjunto de la población. Esto se debe a que tanto para la HiperTG como la Combinada y la DA, el principal lípido que las define son los TG, y los TG están altamente correlacionados con el IMC. (Howard *et al.*, 2003) Contrariamente, en la categoría de HiperCOL donde la cantidad de TG es menor, el IMC es un 3.5% más bajo que para la población total.

Respecto al colesterol hemos estudiado varios parámetros de los cuales destacamos en la Tabla 2 el colesterol total, el cLDL y el cHDL y para el conjunto de toda la población se observan valores de  $5.20 \pm 1.17$ ,  $3.28 \pm 1.01$  y  $1.36 \pm 0.34$  mmol/L respectivamente. Para el colesterol total, en las categorías de HiperCOL y Combinada estos parámetros son superiores a los de la población total, en un 10.38% y 29.42% respectivamente. En

cambio, en la HiperTG y la DA son de un 2.31% y 21.54% menores a la población total respectivamente, asegurando así una buena definición de cada categoría.

Para cLDL las categorías HiperCOL y Combinada tienen un 22.87% y un 36.28% más que la población total, mientras que las categorías HiperTG y DA tienen un 17.07% y 20.73% menos que el conjunto total de la población.

Por último, para el cHDL las categorías HiperCOL y Combinada tienen una media un 2.94% y 1.47% superior al total de la población, mientras que HiperTG y DA tienen la media un 15.44% y un 27.94% inferior al conjunto de la población respectivamente.

Debido a que una parte de la población sufre DM2 hemos considerado importante mostrar los valores de glucosa en sangre. Debido a que los TG son el principal lípido alterado en pacientes DM2, (Taskinen, 2002) se ha podido observar la misma correlación para ambos parámetros (glucosa y TG). Mientras que el valor medio de glucosa para el conjunto de la población es de 142.2 mg/dL, la concentración de glucosa es superior al conjunto de la población para HiperTG (20.15%), Combinada (13.23%) y DA (20.40%) mientras que para HiperCOL es inferior (7.57%) al total de la población.

Finalmente hemos querido destacar en la Tabla 2 la cantidad de proteína C reactiva (PCR) debido a que es un importante marcador inflamatorio. Para el conjunto de la población el valor medio es de 1.70 mmol/L mientras que es superior para HiperTG (2.40%), Combinada (18.00%) y DA (20.80%) e inferior para HiperCOL (0.80% menos).

Tabla 2. Descripción de la población.

	<b>Total, n=480</b>	<b>HiperCOL, n=126</b>	<b>HiperTG, n=62</b>	<b>Combinada, n=74</b>	<b>DA, n=22</b>
<b>Edad (años)</b>	53.60 ± 14.81	55.84 ± 12.35	58.53 ± 11.60	57.46 ± 12.22	56.14 ± 11.29
<b>Género (%mujeres)</b>	28.5	28.57	25.81	39.19	36.36
<b>IMC (kg/m<sup>2</sup>)</b>	27.91 ± 5.07	26.96 ± 4.35	31.38 ± 4.17	31.68 ± 5.35	32.35 ± 4.80
<b>Colesterol total (mmol/L)</b>	5.20 ± 1.17	5.74 ± 0.81	5.08 ± 0.93	6.73 ± 1.10	4.60 ± 0.79
<b>cLDL (mmol/L)</b>	3.28 ± 1.01	4.03 ± 0.70	2.72 ± 0.39	4.47 ± 0.93	2.60 ± 0.36
<b>cHDL (mmol/L)</b>	1.36 ± 0.34	1.40 ± 0.31	1.15 ± 0.21	1.38 ± 0.30	0.98 ± 0.17
<b>TG (mmol/L)</b>	1.70 ± 1.83	1.07 ± 0.34	3.97 ± 3.38	3.15 ± 1.68	3.47 ± 2.96
<b>Glucosa (mg/dL)</b>	142.20 ± 64.46	131.43 ± 62.75	170.86 ± 79.34	161.02 ± 60.80	171.21 ± 110.94
<b>PCR (mg/L)</b>	2.50 ± 1.89	2.48 ± 1.81	2.56 ± 1.69	2.95 ± 1.97	3.02 ± 2.05

Los datos se muestran con la media ± la desviación estándar.

## 4.2. Descripción de los genotipos de la población

Disponemos de los genotipos de 45 SNPs distribuidos entre 25 genes, que han sido determinados en toda la población (Tabla 3).

Primero comprobamos si la frecuencia del alelo minoritario (MAF) de cada SNP en nuestra población era similar a la descrita en bases de datos públicas como el NCBI (<https://www.ncbi.nlm.nih.gov/snp/>) seleccionando la población de origen europeo (Tabla 3) y se observó que eran muy similares.

También se comprobó si las frecuencias genotípicas eran diferentes o no a las predichas mediante el equilibrio de Hardy Weinberg (HW). Como se puede ver en la Tabla 3, ninguno de los SNPs dio resultado significativo en el equilibrio HW.

Tabla 3. Comprobación del genotipado de los SNPs.

<b>SNP</b>	<b>Gen</b>	<b>Alelo mayoritario</b>	<b>Alelo minoritario</b>	<b>MAF población</b>	<b>MAF NCBI</b>	<b>Equilibrio HW*</b>
rs1042034	APOB	A	G	0.21	0.22	0.93
rs1260326	GCKR	C	T	0.47	0.40	0.22
rs2430347	FBLN5	G	A	0.25	0.26	0.95
rs2770146	TLR4	A	G	0.29	0.27	0.41
rs328	LPL	C	G	0.12	0.09	0.72
rs3732666	FBLN2	T	C	0.24	0.18	0.23
rs3771243	SCD1	C	T	0.43	0.37	0.08
rs505151	PCSK9	A	G	0.05	0.04	0.84
rs540796	PCSK9	G	A	0.17	0.16	0.53
rs6948	CASP3	C	A	0.45	0.44	0.34
rs1061376	FBLN2	C	T	0.32	0.30	0.46
rs10754558	NLRP3	C	G	0.43	0.41	0.99
rs1143627	IL1B	T	C	0.37	0.35	0.34
rs1143633	IL1B	G	A	0.34	0.33	0.46
rs11536896	TLR4	T	C	0.15	0.14	0.27
rs12678919	LPL	A	G	0.12	0.09	0.62
rs1304037	IL1A	A	G	0.29	0.29	0.55
rs1454626	VLDLR	A	C	0.29	0.28	0.09
rs17561	IL1A	G	T	0.29	0.30	0.62
rs1800588	LIPC	C	T	0.22	0.21	0.50
rs1800591	MTTP	G	T	0.26	0.24	0.72
rs1894399	IL1A	G	A	0.29	0.29	0.55
rs1927911	TLR4	C	T	0.25	0.27	0.13
rs1927914	TLR4	T	C	0.34	0.35	0.22
rs2284340	FBLN5	C	T	0.47	0.47	0.22

rs268	LPL	A	G	0.02	0.01	0.78
rs2705897	CASP3	G	T	0.21	0.26	0.31
rs2748217	NDST1	G	C	0.20	0.20	0.06
rs2948080	KAT2B	G	C	0.32	0.31	0.03
rs3135506	APOA5	G	C	0.07	0.06	0.70
rs35829419	NLRP3	C	A	0.03	0.05	0.45
rs3732164	SCD1	A	T	0.36	0.31	0.41
rs3797620	NDST1	C	T	0.35	0.42	0.45
rs439401	APOE	C	T	0.40	0.35	0.75
rs4520	APOC3	C	T	0.29	0.28	0.42
rs4986790	TLR4	A	G	0.07	0.06	0.51
rs5128	APOC3	C	G	0.08	0.09	0.96
rs562556	PCSK9	A	G	0.17	0.16	0.53
rs579543	IL1RN	C	T	0.30	0.27	0.27
rs5888	SCARB1	C	T	0.47	0.40	0.23
rs662799	APOA5	A	G	0.07	0.07	0.65
rs708272	CETP	C	T	0.40	0.42	0.30
rs763110	FASLG	C	T	0.42	0.32	0.90
rs833068	VEGFA	G	A	0.34	0.28	0.22
rs964184	APOA5	C	G	0.15	0.13	0.33

\*p-valor del equilibrio HW.

Para trabajar con un número de muestra más grande, agrupamos los genotipos en portadores y no portadores del alelo de menor abundancia.

Para este estudio no es necesario saber cuál es el tipo de cambio que comporta el SNP sobre la secuencia de ADN e incluso sobre las proteínas, ya que pretendemos utilizar los SNPs como marcadores de una zona del genoma. Por tanto, utilizamos los SNPs como marcadores de un locus del genoma que causa una determinada variación en los parámetros estudiados ya esté o no causado por el propio SNP.

#### 4.3. Estudio de la relación entre los SNPs, las dislipidemias y los factores de riesgo cardiovascular

Para identificar los SNPs que tuviesen relación con el riesgo cardiovascular seguimos dos estrategias distintas. La primera tenía como objetivo identificar si algún SNP se presentaba en mayor o menor frecuencia alélica según la categorización de las distintas dislipidemias. En la segunda, en cambio, buscábamos identificar los SNPs que se asocian significativamente a un parámetro de riesgo cardiovascular.

#### 4.3.1. Asociación de SNPs a dislipidemias

Para llevar a cabo la asociación comparamos la frecuencia del alelo minoritario de cada SNP en cada uno de los grupos de las diferentes dislipidemias y el grupo de sujetos sanos (Tabla 4). De esta forma se pone en evidencia a los SNPs que afectan a cada desorden lipídico.

En la Tabla 4 vemos que cuando se compara los sujetos clasificados con HiperCOL y los sujetos sanos, identificamos 4 SNPs cuyos alelos minoritarios se encuentran en menor frecuencia en la HiperCOL. Concretamente, rs540796 y rs562556 de *PCSK9*, rs1800588 de *LIPC* y rs1800591 de *MTTP*. Ambos SNPs de *PCSK9* están identificados como variantes genéticas protectoras en relación a la hipercolesterolemia familiar, según la base de datos ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>) que recoge toda la información disponible de variantes genómicas en relación a la salud en humanos. Esto confirma nuestros resultados. De los genes involucrados en la HiperCOL, *PCSK9* se confirma como gen clave regulando los niveles de cLDL. (Abifadel *et al.*, 2003; Benjannet *et al.*, 2004; Cohen *et al.*, 2005) De hecho, un posible tratamiento para los casos más severos de hipercolesterolemia familiar consiste en el tratamiento del paciente con inhibidores de *PCSK9*. (Sharifi *et al.*, 2019) Este tratamiento ha cambiado el paradigma de las terapias hipolipemiantes para prevenir el riesgo cardiovascular. Los resultados del estudio FOURIER mostraron el importante descenso de los niveles de cLDL y de eventos cardiovasculares conseguidos con los inhibidores de *PCSK9*. (Murphy *et al.*, 2019) Respecto al gen *LIPC* tiene una relación muy directa con la síntesis de LDLs a partir de IDLs y VLDLs. Por este motivo era de esperar obtener resultados significativos para la HiperCOL, y dado que se ha identificado que esta misma variante está asociada a una menor actividad de la lipasa hepática, confirmamos el sentido de nuestros resultados. (Silbernagel *et al.*, 2019) Finalmente, el gen *MTTP* ayuda a regular el metabolismo del colesterol y triglicéridos lo que se correlaciona con un trastorno que afecta a la concentración de ambos. Recientemente se han identificado mutaciones en este gen en pacientes con hipocolesterolemia, lo que va en el mismo sentido que nuestros resultados. (Blanco-Vaca *et al.*, 2019) Algo que sorprendió fue no obtener resultados significativos para SNPs de genes que codifican para apolipoproteínas, sobre todo para apoB100 que es la principal apolipoproteína en las LDLs. Aunque en otros estudios sí se

encontraron, es posible que para este caso no se hayan escogido los SNPs adecuados o que la muestra de la población no fuese la adecuada para este parámetro. (Ramasamy, 2016; Sharifi *et al.*, 2019)

Cuando comparamos la HiperTG y los sujetos sanos, se encuentran 2 SNPs más frecuentes en individuos sanos que en individuos con HiperTG, rs328 y rs12678919. Ambos pertenecen al gen LPL, un gen que hidroliza los TG circulantes, lo que se correlaciona con este tipo de dislipidemia. Y tal y como se ha comentado en la Introducción, el gen LPL es uno de principales genes causantes de las hipertrigliceridemias más severas. (Brown *et al.*, 2015)

Tabla 4. Resultados de la primera estrategia, asociación de SNPs a dislipidemias.

<b>SNP</b>	<b>Gen</b>	<b>Frecuencia del MAF en HiperCOL (%)</b>	<b>Frecuencia del MAF en sanos (%)</b>	<b>p valor</b>
rs540796	PCSK9	31.03	45.24	0.047
rs562556		30.77	45.24	0.044
rs1800588	LIPC	40.17	54.76	0.009
rs1800591	MTTP	38.89	52.38	0.045
<b>SNP</b>	<b>Gen</b>	<b>Frecuencia del MAF en HiperTG (%)</b>	<b>Frecuencia del MAF en sanos (%)</b>	<b>p valor</b>
rs328	LPL	11.29	33.33	0.011
rs12678919		9.68	31.58	0.009
<b>SNP</b>	<b>Gen</b>	<b>Frecuencia del MAF en Combinada (%)</b>	<b>Frecuencia del MAF en sanos (%)</b>	<b>p valor</b>
rs35829419	NLRP3	1.35	8.62	0.047
rs662799	APOA5	20.27	5.17	0.012
rs540796	PCSK9	16.22	39.66	0.001
rs562556		16.22	39.66	0.001
<b>SNP</b>	<b>Gen</b>	<b>Frecuencia del MAF en DA (%)</b>	<b>Frecuencia del MAF en sanos (%)</b>	<b>p valor</b>
rs708272	CETP	59.68	77.19	0.021

MAF: minor allele frequency

Cuando comparamos la frecuencia de los SNPs entre los sujetos con dislipidemia Combinada y los sujetos sanos, encontramos 4 SNPs: uno de ellos se encuentra con mayor frecuencia en individuos con Combinada que en individuos sanos (rs662799 del gen APOA5) y tres están asociados a individuos sanos (rs35829419 del gen NLRP3 y dos del gen PCSK9; rs540796 y rs562556). La identificación de variantes genéticas en estos 3 genes en relación a la Combinada es muy interesante. Tal y como hemos descrito anteriormente, la Combinada se caracteriza por niveles alterados de cLDL y de TG, y en

este sentido, como ya hemos comentado, PCSK9 es uno de los principales reguladores de los niveles de cLDL, y el gen de la APOA5 es uno de los principales genes reguladores de los TG. Variantes genéticas en el gen *APOA5* tienen importantes implicaciones en el metabolismo y la acumulación de los triglicéridos (TG). También se ha observado una influencia sobre la distribución de las lipoproteínas y en la producción de quilomicrones. Todo esto derivando en mayor riesgo de sufrir ECVs. (Guardiola & Ribalta, 2017) Además se encontró relación con un SNP de *NLRP3*, un importante marcador del estado inflamatorio. (Haneklaus & O'Neill, 2015) La hiperlipemia familiar combinada es una de las dislipidemias más prevalentes y recientemente se ha descrito que cursa con un estado inflamatorio que corrobora nuestros hallazgos. (Díaz-Ruiz *et al.*, 2019)

Finalmente, cuando se comparó a los individuos con DA con los sanos se observó que el SNP rs708272 del gen *CETP* se encuentra en mayor frecuencia en individuos sanos que en aquellos con DA. La CETP es una proteína importante para el metabolismo lipídico por su papel en el intercambio de colesterol y TG entre distintas lipoproteínas. La DA es otra forma de dislipidemia que, en este caso, incorpora el cHDL como parámetros clínicos para su diagnóstico. Es la forma de dislipidemia más frecuente en pacientes con diabetes, síndrome metabólico u obesidad.

Tras realizar estos primeros análisis nos llamó la atención que los SNPs rs540796 y rs562556 aparecen siempre en los mismos casos y con resultados significativos muy similares o idénticos. Por este motivo el primer paso fue verificar si los polimorfismos se encontraban cerca en la secuencia ya que ambos pertenecen al gen *PCSK9*. Se comprobó que se encontraban únicamente a 40pb de distancia en el brazo corto del cromosoma 1 (rs540796 a chr1:55058524 y rs562556 a chr1:55058564). El siguiente paso fue comprobar que efectivamente existía desequilibrio de unión causando esta herencia conjunta a través de un análisis realizado con la herramienta "*LDpair*" del National Institutes of Health (NIH) (<https://ldink.nci.nih.gov>). Tanto el valor  $r^2$  como el valor  $D'$  eran igual a 1, su valor máximo en ambos casos. Esto demuestra que existe un desequilibrio de unión claro entre ambos SNPs.

En resumen, de esta primera estrategia de análisis destacaríamos dos importantes observaciones.

Primero, hemos identificado un perfil de variantes genéticas en relación al metabolismo del colesterol, en una cohorte de sujetos que no presentan significativamente defectos importantes relacionados con el metabolismo del colesterol. Nuestra cohorte está formada mayoritariamente por pacientes con diabetes, síndrome metabólico y obesidad que, aunque tienen un elevado riesgo cardiovascular, no se explica por elevaciones importantes del cLDL.

Segundo, esta estrategia nos ha permitido identificar que actúan conjuntamente variantes genéticas tanto de genes clave en el metabolismo lipídico como de un gen clave en el proceso inflamatorio, en una cohorte de sujetos con alteraciones metabólicas que cursan con elevado riesgo cardiovascular.

Uno de los principales inconvenientes de esta estrategia es que como hemos comparado grupos, hemos reducido el tamaño muestral. En el caso de la HiperCOL se trabaja con 42 individuos sanos y 116 con la dislipidemia. En la HiperTG se trabaja con 57 individuos sanos y 62 con la dislipidemia. En la Combinada se trabaja con 58 individuos sanos y 74 con la dislipidemia. Finalmente se trabaja en la DA con 58 individuos sanos y 22 con la dislipidemia. Pero para esta estrategia creemos que el beneficio de comparar grupos extremos bien caracterizados para identificar los SNPs implicados es superior a estudiar toda la población conjuntamente.

#### 4.3.2. Asociación de SNPs a parámetros lipídicos

Para la segunda estrategia, hemos estudiado a toda la población conjuntamente. Esta estrategia tenía un objetivo más concreto puesto que en vez de relacionar los SNPs con todo el desorden lipídico, se busca la asociación con el parámetro que se ve afectado. Hemos estudiado la asociación de SNPs con la concentración de TG y cLDL porque son los parámetros clínicos de mayor importancia al diagnosticar a los pacientes con riesgo cardiovascular. Sin embargo, se han incluido los parámetros PCR, cHDL, el colesterol de las partículas de lipoproteínas remanentes (cRLP) y el IMT puesto que son otros factores de riesgo cardiovascular que ayudan a comprender mejor el desorden lipídico. La PCR es el parámetro clínico más utilizado para evaluar el estado inflamatorio de los sujetos. El cRLP es un parámetro que nos da una idea de la cantidad de lipoproteínas remanentes

que presenta un sujeto. Estas lipoproteínas remanentes se acumulan en el plasma de los pacientes diabéticos, con síndrome metabólico y obesos, y se ha asociado a un mayor riesgo cardiovascular. (Qin *et al.*, 2019) Y como ya hemos comentado anteriormente, el IMT es una medida de la arteriosclerosis subclínica. (Bots & Grobbee, 2002)

No se encontró una relación significativa para ninguno de los SNPs con respecto a las concentraciones de TG. Los resultados obtenidos para este parámetro no fueron los esperados en un principio, debido a que la elección de los SNPs incluyó algunos que podían estar involucrados en el metabolismo de los TG como es el caso del gen *APOC3* o *APOA5* del cual se escogieron dos SNPs. Tampoco se encontró ningún SNP significativamente relacionado con la concentración de cRPL. No obstante, debido a la estrecha relación entre los TG y el cRPL era de esperar que ambos vayan en la misma línea.

Identificamos 2 SNPs que estaban significativamente asociados a la concentración de PCR (Tabla 5). Uno de ellos es el rs10754558 del gen *NLRP3* en el que los portadores del alelo minoritario presentaban una concentración de PCR 19.38% inferior a la de los no portadores ( $p=0.026$ ). El otro SNP es el rs1143633 del gen *IL1B* que presentaba una concentración de PCR un 14.29% superior en los portadores del alelo minoritario frente a los no portadores ( $p=0.035$ ). Aunque ambos genes están implicados en la inflamación, no hemos encontrado ningún artículo que corrobore nuestra asociación del SNP rs10754558 con niveles de PCR, por lo tanto, somos los primeros en describir esta asociación. En el caso del rs1143633 en el gen *IL1B*, hemos encontrado un artículo que describe la asociación con PCR, pero en pacientes con enfermedad pulmonaria obstructiva. (Wang *et al.*, 2015) Estos resultados sugieren una relación de estos genes con la arteriosclerosis y supone un importante riesgo cardiovascular a elevadas concentraciones y de forma continuada. (Berberich & Hegele, 2019b)

En el caso del parámetro cHDL se encontraron 2 SNPs involucrados. El rs1800588 del gen *LIPC* presenta una concentración de cHDL en los portadores del alelo minoritario un 8.33% superior a la de los no portadores ( $p=0.004$ ). Por otro lado, el rs708272 del gen *CETP* presenta una concentración de cHDL un 11.81% superior en los portadores del alelo minoritario que en los no portadores ( $p<0.001$ ). Estos dos SNPs ya fueron identificados anteriormente por estar asociados a DA, caracterizada por alteraciones en

los niveles de TG y de cHDL, en pacientes infectados por el VIH. (Guardiola *et al.*, 2015b) Por lo tanto, como era de esperar, ambos genes están implicados en el metabolismo lipídico. De hecho, a nivel farmacológico, desde hace unos años se están intentando desarrollar fármacos inhibidores de CETP para conseguir aumentar los niveles de cHDL como método de prevención cardiovascular. (Armitage *et al.*, 2019)

Para el parámetro IMT se encontraron 4 SNPs asociados. Uno de ellos es el rs1042034 del gen *APOB*, el cual presentaba un valor del IMT un 7.35% superior en los portadores del alelo minoritario en comparación al de los no portadores ( $p=0.004$ ). Otro de los SNPs es el rs1454626 *VLDLR* que en portadores del alelo minoritario mostraba un valor de IMT un 4.35% superior a los sujetos no portadores ( $p=0.044$ ). El rs1800591 del gen *MTTP* mostró un valor un 7.35% superior en los individuos portadores del alelo minoritario que en los no portadores ( $p=0.006$ ). Finalmente, en el rs662799 del gen *APOA5* el valor del IMT en portadores del alelo minoritario era un 8.45% inferior a la de los no portadores ( $p=0.047$ ). No hemos encontrado ningún estudio que describa las asociaciones con IMT que hemos encontrado, por lo tanto, somos los primeros en describir estas asociaciones entre estos SNPs y la arteriosclerosis subclínica. Aunque en relación a algunos de los genes identificados, sí que se han descrito asociaciones con la arteriosclerosis subclínica y otros SNPs en estos genes. (Guardiola *et al.*, 2015c)

Por último, el parámetro cLDL mostró 7 SNPs con resultado significativo. El rs2948080 del gen *KAT2B* mostró una concentración de cLDL un 7.35% inferior en los portadores del alelo minoritario frente a los no portadores ( $p=0.009$ ). El rs540796 del gen *PCSK9*, al igual que el rs562556 también de *PCSK9*, presentan una concentración de cLDL un 8.63% inferior en los portadores del alelo minoritario que en los no portadores ( $p=0.012$  y  $p=0.011$  respectivamente). Otro SNP, el rs1042034 del gen *APOB* presenta una concentración de cLDL un 5.31% superior en los portadores del alelo minoritario que en los no portadores ( $p=0.048$ ). El rs3135506 del gen *APOA5* presenta una concentración un 10.84% inferior en los individuos portadores del alelo minoritario frente a los no portadores ( $p=0.002$ ). Al igual que el SNP anterior, el rs964184 también pertenece al gen *APOA5* y también presenta una concentración de cLDL inferior en los portadores del alelo minoritario que en los no portadores. En este caso la concentración es un 8.08% inferior ( $p=0.001$ ). Por último, el rs4520 del gen *APOC3* presentó en los portadores del

alelo minoritario una concentración de cLDL un 8.24% inferior a la de los individuos no portadores ( $p=0.001$ ). Todos los genes con excepción del *KAT2B* están involucrados en el metabolismo de lipoproteínas o bien en la síntesis o regulación de partículas de transporte de lípidos y colesterol. El gen *KAT2B*, en cambio, codifica para una proteína con una función regulatoria de la transcripción, y ya se ha asociado a riesgo cardiovascular asociado a mecanismos epigenéticos (Pons *et al.*, 2011) y al cLDL (Wei *et al.*, 2015).

Tabla 5. Resultados de la segunda estrategia, SNPs asociados a parámetros lipídicos.

<b>SNPs asociados a concentración de PCR (mg/L)</b>				
<b>SNP</b>	<b>Gen</b>	<b>Portadores MAF</b>	<b>No portadores del MAF</b>	<b>p valor</b>
rs10754558	<i>NLRP3</i>	2.33 ± 1.83	2.89 ± 1.99	0.026
rs1143633	<i>IL1B</i>	2.64 ± 2.05	2.31 ± 1.69	0.035
<b>SNPs asociados a IMT (mm)</b>				
<b>SNP</b>	<b>Gen</b>	<b>Portadores MAF</b>	<b>No portadores del MAF</b>	<b>p valor</b>
rs1042034	<i>APOB</i>	0.73 ± 0.17	0.68 ± 0.12	0.004
rs1454626	<i>VLDLR</i>	0.72 ± 0.16	0.69 ± 0.13	0.044
rs1800591	<i>MTTP</i>	0.73 ± 0.15	0.68 ± 0.13	0.006
rs662799	<i>APOA5</i>	0.65 ± 0.13	0.71 ± 0.14	0.047
<b>SNPs asociados a concentración de cHDL (mmol/L)</b>				
<b>SNP</b>	<b>Gen</b>	<b>Portadores MAF</b>	<b>No portadores del MAF</b>	<b>p valor</b>
rs1800588	<i>LIPC</i>	1.43 ± 0.36	1.32 ± 0.32	0.004
rs708272	<i>CETP</i>	1.42 ± 0.36	1.27 ± 0.28	0.000
<b>SNPs asociados a concentración de cLDL (mmol/L)</b>				
<b>SNP</b>	<b>Gen</b>	<b>Portadores MAF</b>	<b>No portadores del MAF</b>	<b>p valor</b>
rs2948080	<i>KAT2B</i>	3.15 ± 0.97	3.40 ± 1.07	0.009
rs540796	<i>PCSK9</i>	3.07 ± 0.92	3.36 ± 1.06	0.012
rs562556		3.07 ± 0.92	3.36 ± 1.06	0.011
rs1042034	<i>APOB</i>	3.37 ± 1.09	3.20 ± 0.98	0.048
rs3135506	<i>APOA5</i>	2.96 ± 0.84	3.32 ± 1.04	0.002
rs964184		3.07 ± 0.83	3.34 ± 1.08	0.006
rs4520		<i>APOC3</i>	3.12 ± 0.92	3.40 ± 1.09

Los datos se muestran con la media ± la desviación estándar. P valor ajustado por edad, sexo e IMC.

Llegados a este punto, nos preguntamos, si existían coincidencias entre los SNPs identificados mediante las dos estrategias.

De todos los 15 SNPs con resultados significativos encontrados mediante la segunda estrategia (Tabla 5), seis de ellos se encontraron también en la primera estrategia. Cuatro de los seis se encontraron relacionados con la HiperCOL, uno con la Combinada y el último con la DA. Los que se encontraron relacionados con la HiperCOL son los dos SNPs de *PCSK9* que han aparecido relacionados con la concentración de cLDL (rs540796

y rs562556), uno del parámetro cHDL (rs1800588) y otro del parámetro IMT (rs1800591). En los dos SNPs de *PCSK9* se puede observar como en ambos casos, en los portadores del alelo minoritario, tanto la frecuencia de individuos con HiperCOL como la concentración de cLDL se reduce. De manera que podemos confirmar la clara asociación entre estos SNPs en *PCSK9* con los niveles de cLDL circulante, mediante dos estrategias diferentes. El rs1800588 del gen *LIPC* que se encontraba asociado a la HiperCOL en la primera estrategia, se ha visto asociado a la concentración de cHDL en la segunda. En la primera estrategia se ha observado una asociación del alelo minoritario a una reducción de la HiperCOL y en la segunda a un aumento de la cHDL. Estos resultados son complementarios puesto que el cHDL se considera como la mejor forma de eliminar el exceso de colesterol en los tejidos debido a la elevada eficiencia de las partículas HDL para transportarlo. El cuarto SNP asociado a la HiperCOL, es el rs1800591 del gen *MTTP*, el cual se ha visto también asociado al valor del IMT. Mientras que en la primera estrategia se ha visto que el alelo minoritario puede ser beneficioso para reducir la frecuencia de la HiperCOL, en la segunda se ha visto una asociación del alelo minoritario a elevados valores del IMT lo cual no resulta beneficioso. El SNP rs662799 se ha encontrado relacionado con la Combinada y con el valor IMT. Al igual que en el caso anterior, una mayor frecuencia del alelo minoritario se ha visto asociada a por un lado una reducción de la Combinada pero también a un aumento del IMT. Futuros estudios en relación a este polimorfismo serían necesarios para confirmar estas observaciones. Finalmente, el rs708272 se ha visto relacionado con la DA y la concentración de cHDL. En este caso, una mayor frecuencia del SNP se ha visto asociado a una reducción en la frecuencia de la dislipidemia y a un aumento de la concentración de cHDL. Por tanto, tal y como se ha explicado para el SNP rs1800588 estos dos resultados son complementarios y muestran un efecto beneficioso del SNP rs708272 sobre el riesgo cardiovascular.

#### 4.3.3. Estudio del efecto aditivo de varios SNPs sobre el cLDL

Debido a la gran cantidad de SNPs que se encontraron relacionados con el cLDL plasmático se decidió estudiar el efecto conjunto de los diferentes SNPs. Queríamos simular la creación de un “*genetic risk score*” para el control de los niveles circulantes de cLDL. Para ello, el primer paso fue comprobar qué alelos causaban un efecto negativo (cLDL alto) y qué alelos daban un efecto positivo (cLDL bajo) (Tabla 5). Con estos datos se realizó un estudio dosis-efecto para comprobar si existe un efecto sumatorio a medida que los individuos presentan más de un SNPs (Gráfico 1). En nuestra cohorte, encontramos que los sujetos podían presentar de 0 hasta 7 alelos que aumentan el riesgo cardiovascular en estos SNPs relacionados con el cLDL. La cantidad de individuos para cada cantidad de alelos del 0 al 7 es de 7, 13, 37, 72, 88, 127, 97 y 19 respectivamente, sumando un total de 460 individuos estudiados. Como se puede observar en el Gráfico 1 existe la tendencia a aumentar la concentración de cLDL a medida que aumenta el número de alelos estudiados. No obstante, solo se obtienen diferencias significativas entre ser portadores de 1, 2, 3, 4 o 5 de los alelos frente a tener 7. Las concentraciones medias para las columnas del 0 al 7 se han incluido en el gráfico y son de 2.73, 2.79, 2.94, 3.08, 3.10, 3.33, 3.54 y 4.11 mmol/L respectivamente.

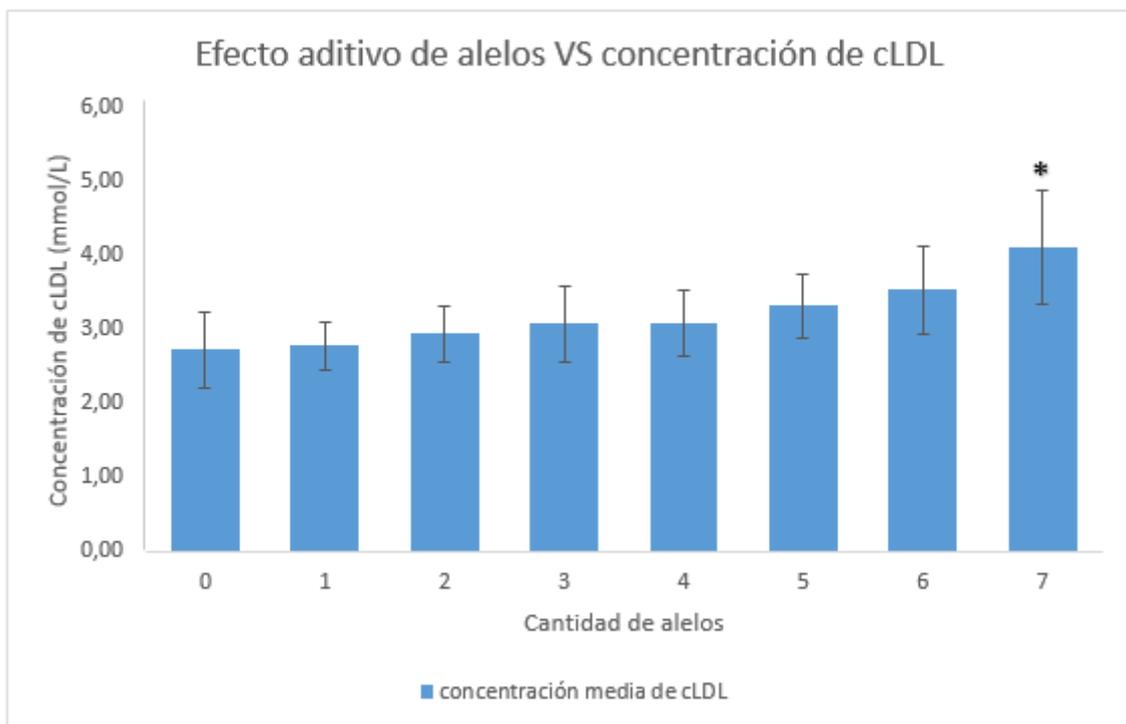


Gráfico 1. Dosis-efecto de la combinación de varios alelos que aumentan la concentración plasmática de cLDL. Los datos han sido ajustados por edad, sexo e IMC. \* Diferencia significativa con los grupos de 0, 1, 2, 3, 4 y 5 alelos.

#### 4.3.4. Distribución de las dislipidemias sobre los grupos de alelos

A continuación, se quiso conocer los SNPs más abundantes de las diferentes combinaciones. En primer lugar, cabe destacar que la mayoría de los casos estudiados (68.12%) de los 460 totales, contenían 4, 5 o 6 alelos considerados perjudiciales para la concentración de cLDL. La distribución de los 7 SNPs por las distintas categorías es uniforme y únicamente cabe destacar que el SNP rs3135506 del gen *APOA5* es el más frecuente en prácticamente todos los casos. Se encuentra en un 23.1%, 56.8%, 72.2%, 89.8%, 99.2% y un 100% de los casos de 0 a 7 alelos respectivamente. Este resultado nos sorprendió puesto que *APOA5* fue escogido por su papel regulando los niveles de TG. Además de su relevante papel regulando los niveles de TG, también se ha descrito que los polimorfismos en *APOA5* se asocian a un mayor número de partículas LDL circulantes. (Guardiola *et al.*, 2015c) Además, esta abundancia no es debida a la alta frecuencia de este SNP en el total de la población puesto que como se observa en la Tabla 3 solo se encuentra con una frecuencia del 0.06.

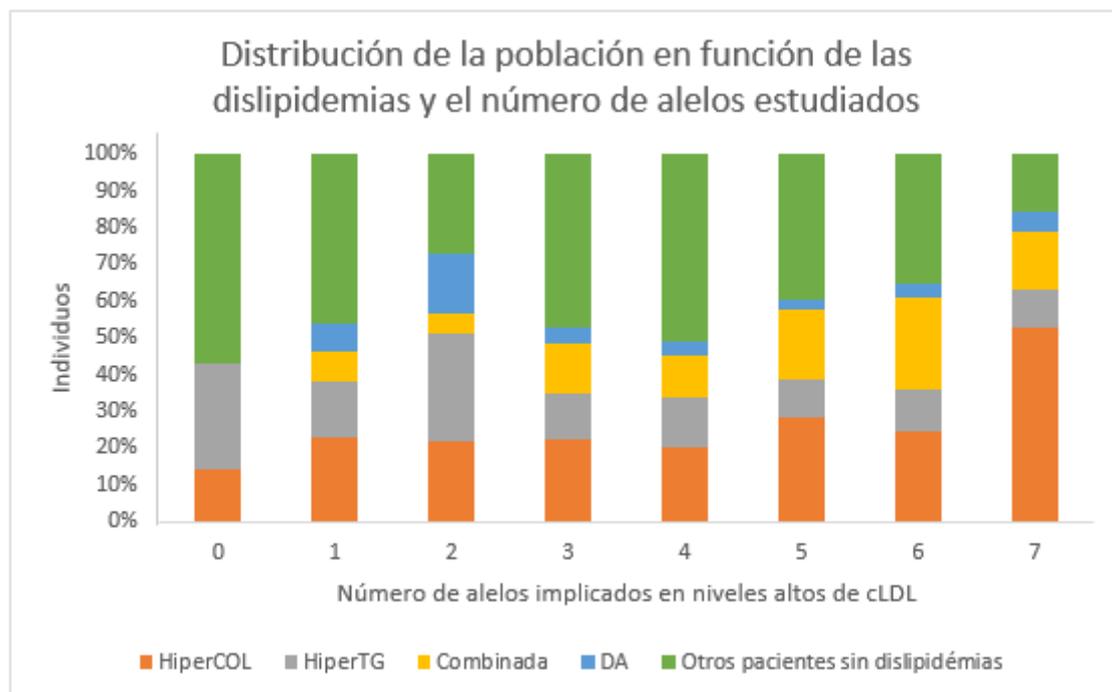


Gráfico 2. Distribución de la población en función del número de alelos estudiados y las distintas dislipidemias.

Otro aspecto que se observó en el Gráfico 2 fue la distribución de las distintas dislipidemias dentro de estas categorías prestando especial atención a la HiperCOL debido a su importante relación con el cLDL. Como se puede observar en el Gráfico 2, en cuanto a porcentajes, a grandes rasgos la HiperCOL se encuentra más representada en los casos portadores de mayor número de alelos estudiados (14%, 23%, 22%, 22%, 21%, 28%, 25% y 53% de las columnas 0 a la 7 respectivamente). Este mismo efecto se puede observar para la Combinada (0%, 8%, 5%, 14%, 12%, 19%, 25% y 16% de las columnas 0 a la 7 respectivamente); ambas dislipidemias caracterizadas por alteraciones en los niveles de cLDL. Contrariamente a HiperCOL y la Combinada, para la HiperTG, a grandes rasgos se observa que un mayor número de alelos implica menor porcentaje de individuos con este desorden metabólico (29%, 15%, 30%, 13%, 13%, 10%, 11% y 11% de las columnas 0 a la 7 respectivamente). Respecto a la DA debido a que se disponía de una cantidad de individuos muy baja no ha sido posible extraer conclusiones acerca de la influencia de estos SNPs sobre la dislipidemia.

Finalmente, cabe destacar que por lo general, un mayor número de alelos equivale a una mayor concentración de cLDL y un menor número de personas sin dislipidemias.

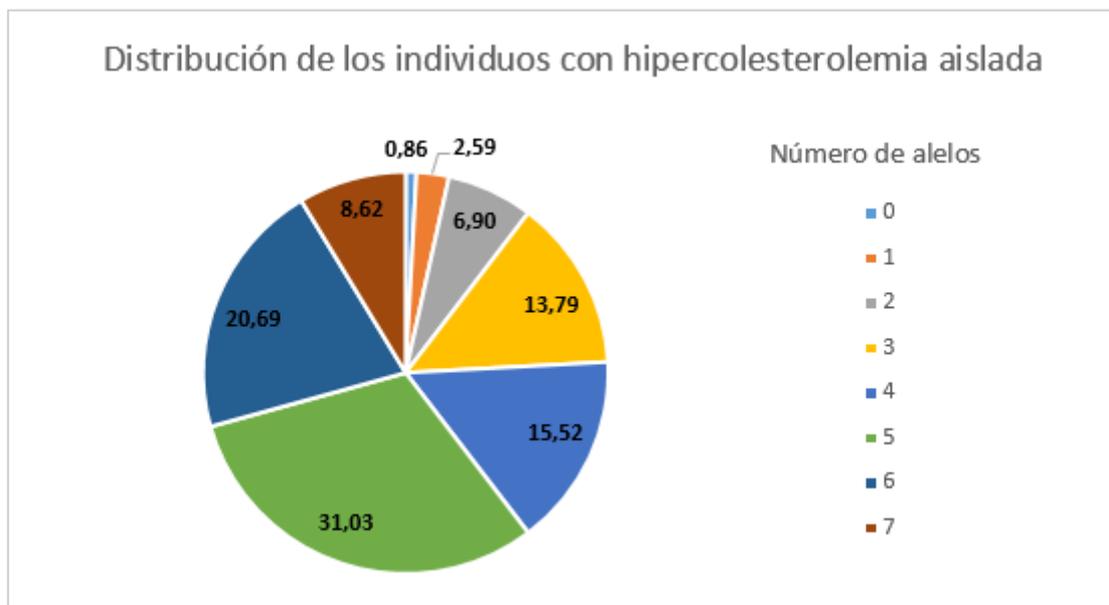


Gráfico 3. Distribución de los individuos con hipercolesterolemia aislada según la cantidad de alelos que aumentan la concentración de colesterol plasmático (en porcentaje).

Como se ha explicado anteriormente se prestó especial atención a los casos de HiperCOL, los cuales son un total de 116. Como se observa en el Gráfico 3, los casos con 5, 6 o 7 alelos suman un 60.34% de los casos totales de HiperCOL. En cambio, los casos con 0, 1 o 2 alelos únicamente suman el 10.35%. Esto nos lleva a pensar que en la población que se está estudiando hay un elevado porcentaje de pacientes con hipercolesterolemia poligénica. La hipercolesterolemia poligénica se trata de la forma más común de la hipercolesterolemia, afectando al 4% de la población española adulta. Se trata de una dislipidemia poco estudiada dada la complejidad de identificar los genes implicados pero con un gran interés clínico. (Saraju & Knowles, 2019)

#### 4.3.5. Asociación de SNPs a otros parámetros de las LDLs

Una vez identificados los SNPs asociados a cLDL, se quiso profundizar en su papel sobre el metabolismo lipídico. Por este motivo, se estudió si los SNPs también se asociaban a varios parámetros relacionados con las LDLs como el número de partículas y su tamaño. Concretamente se realizó el estudio de las partículas LDL grandes, medianas y pequeñas determinadas mediante resonancia magnética nuclear, para observar individualmente si la distribución aumentaba de forma similar o no a la cantidad de cLDL. Los resultados se pueden observar en los Gráficos 4, 5 y 6 ajustados para edad, género e IMC. En los tres casos el aumento de alelos estudiados se corresponde con un aumento de las concentraciones de partículas, siguiendo una tendencia muy similar a la que se veía en el Gráfico 1. No obstante, debido al pequeño impacto que presenta cada SNP por separado, en la mayoría de los casos no se encontraron diferencias significativas entre grupos.

Por tanto, se puede afirmar que el aumento en concentración de cLDL que se observó en el Gráfico 1 se debe a un aumento de todos los tamaños de partícula y no únicamente de las pequeñas que son las que causan efectos más negativos en cuanto al riesgo cardiovascular. (Shiffman *et al.*, 2017)

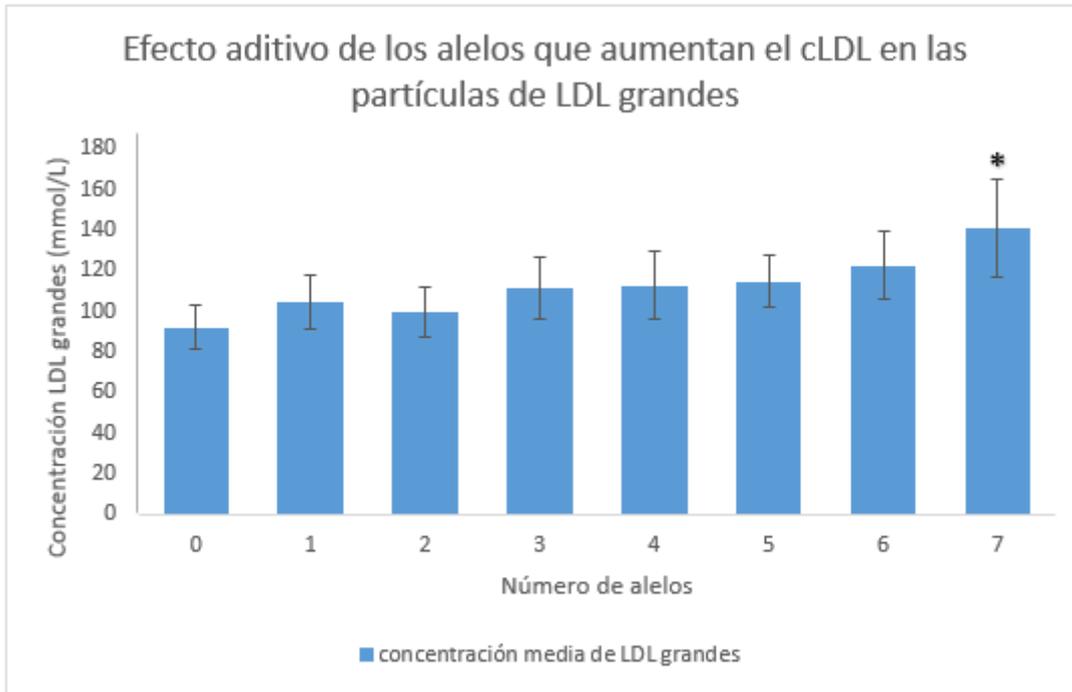


Gráfico 4. Concentración de partículas LDL grandes. Los datos han sido ajustados por edad, sexo e IMC. \* Diferencias significativas con los grupos de 0 y 2 alelos.

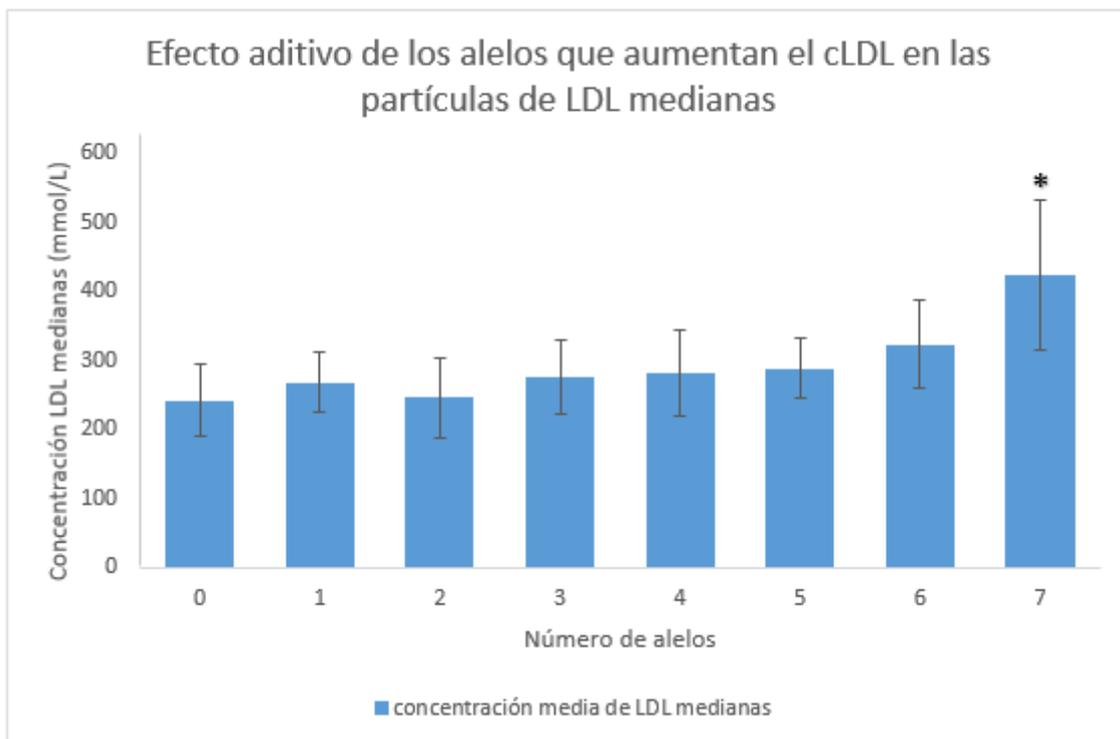


Gráfico 5. Concentración de partículas LDL medianas. Los datos han sido ajustados por edad, sexo e IMC. \*Diferencias significativas con los grupos de 2, 3, 4 y 5 alelos.

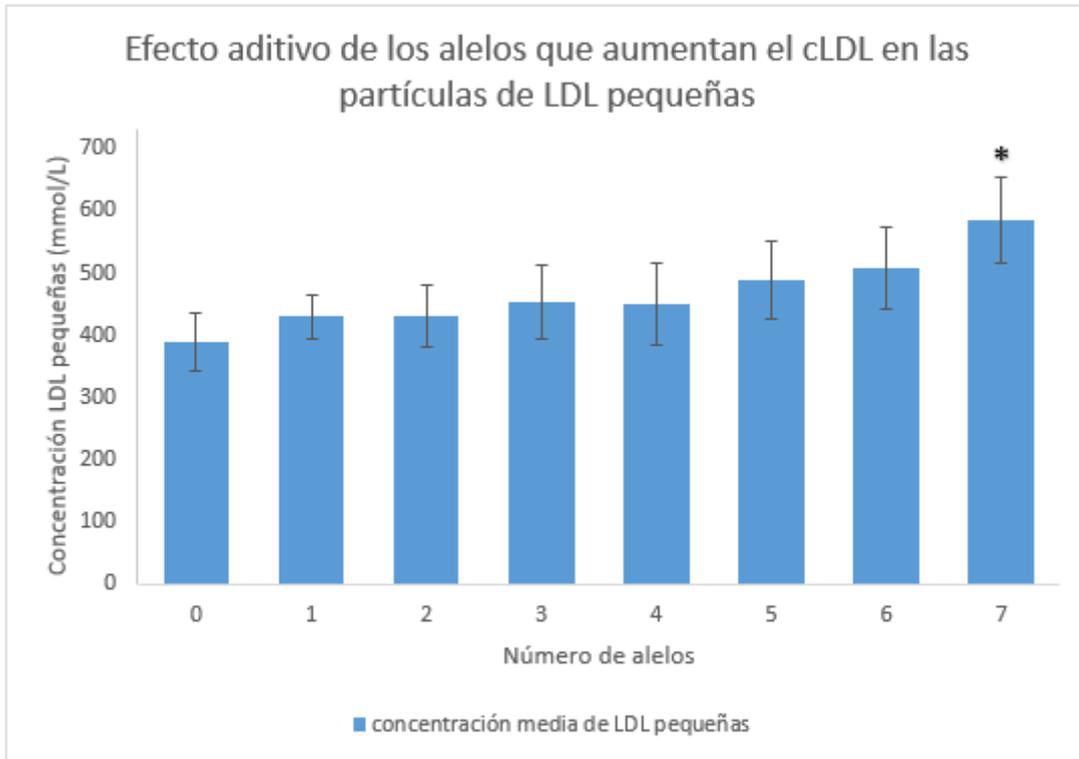


Gráfico 6. Concentración de partículas LDL pequeñas. Los datos han sido ajustados por edad, sexo e IMC. \* Diferencias significativas con los grupos de 2 y 4 alelos.

## 5. Conclusión

Hemos identificado en una población con un elevado porcentaje de personas con riesgo cardiovascular una combinación de 7 SNPs (rs2948080 de *KAT2B*, rs540796 y rs562556 de *PCSK9*, rs1042034 de *APOB*, rs3135506 y rs964184 de *APOA5* y rs4520 de *APOC3*) formada por variantes en genes implicados en el metabolismo lipídico y la inflamación, que se asocian a hipercolesterolemia, concentración de cLDL y número de partículas LDL.

- Más de la mitad de los sujetos estudiados presentaban 5 o más alelos que elevan el cLDL de la combinación de los 7 SNPs estudiados para este parámetro.
- Hemos identificado asociaciones entre polimorfismos y factores de riesgo cardiovascular no descritas anteriormente:
  - Hemos asociado por primera vez los SNPs rs10754558 de *NLRP3* y rs1143633 de *IL1B* a la PCR en pacientes con riesgo cardiovascular. Además, es la primera vez que se asocia la hiperlipidemia combinada a un estado inflamatorio (a través del gen *NLRP3*), lo cual va en la línea de recientes descubrimientos.
  - Hemos encontrado por primera vez relación entre los SNPs rs1042034 de *APOB*, rs1454626 de *VLDLR*, rs1800591 *MTTP* y rs662799 de *APOA5* y la concentración de IMT.
- Los SNPs rs540796 y rs562556 del gen *PCSK9* presentan un elevado desequilibrio de unión. Pudiéndolos utilizar indistintamente en futuros algoritmos para detectar el riesgo cardiovascular.
- Se demuestra que en las dislipidemias poligénicas es posible asociar una creciente concentración de cLDL en función de una mayor presencia de determinados SNP.
- Los datos obtenidos son muy útiles para la realización de algoritmos que permitan predecir el riesgo que tiene el paciente de sufrir un determinado desorden metabólico con una única prueba. Esta prueba permite actuar en consecuencia a través de la medicina personalizada. No obstante, es necesario que se determine primero el impacto que tiene cada SNP sobre el parámetro que se quiere estudiar ya que no todos afectan en la misma medida.

- Es importante seguir buscando más polimorfismos ya sean o no SNPs para poder ampliar la precisión de los algoritmos a la hora de determinar el “*risk score*” del paciente. Se pueden seguir buscando variaciones genéticas en los mismos genes que se han empleado en este estudio o bien en otros distintos. En cualquier caso, es necesario ampliar el número de individuos de la población estudiada ya que esto nos permitirá trabajar con polimorfismos menos frecuentes y acercarnos más a los valores de la población total.

## 6. Bibliografía

- Abhary S, Burdon KP, Gupta A, Lake S, Selva D, Petrovsky N, et al.** Common Sequence Variation in the VEGFA Gene Predicts Risk of Diabetic Retinopathy. *Investigative Ophthalmology & Visual Science* 2009; 50:5552.
- Abifadel M, Varret M, Rabès J-P, Allard D, Ouguerram K, Devillers M, et al.** Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nature Genetics* 2003; 34:154–156.
- Álvarez A, Blasco M, Ferreras JM, Lago F, Navarro E, et al.** Manejo de las dislipemias en Atención Primaria. 2012. ISBN: 9788415037279
- Armitage J, Holmes M V., Preiss D.** Cholesteryl Ester Transfer Protein Inhibition for Preventing Cardiovascular Events: JACC Review Topic of the Week. *Journal of the American College of Cardiology* 2019; 73:477–487.
- Benjannet S, Rhainds D, Essalmani R, Mayne J, Wickham L, Jin W, et al.** NARC-1/PCSK9 and its natural mutants: zymogen cleavage and effects on the low density lipoprotein (LDL) receptor and LDL cholesterol. *The Journal of Biological Chemistry* 2004; 279:48865–48875.
- Berberich AJ, Hegele RA.** The role of genetic testing in dyslipidaemia. *Pathology* 2019a; 51:184–192.
- Berberich AJ, Hegele RA.** The complex molecular genetics of familial hypercholesterolaemia. *Nature Reviews Cardiology* 2019b; 16:9–20.
- Blanco-Vaca F, Martin-Campos JM, Beteta-Vicente Á, Canyelles M, Martínez S, Roig R, et al.** Molecular analysis of APOB, SAR1B, ANGPTL3, and MTTP in patients with primary hypocholesterolemia in a clinical laboratory setting: Evidence supporting polygenicity in mutation-negative patients. *Atherosclerosis* 2019; 283:52–60.
- Bots ML, Grobbee DE.** Intima Media Thickness as a Surrogate Marker for Generalised Atherosclerosis. *Cardiovascular Drugs and Therapy* 2002; 16:341–351.
- Brown WV, Goldberg IJ, Young SG.** JCL Roundtable: Hypertriglyceridemia due to defects in lipoprotein lipase function. *Journal of Clinical Lipidology* 2015; 9:274-280.

- Bush WS, Moore JH.** Chapter 11: Genome-Wide Association Studies. *PLoS Computational Biology* 2012; 8:1–11.
- Catapano AL, Graham I, De Backer G, Wiklund O, Chapman MJ, Drexel H, et al.** 2016 ESC/EAS Guidelines for the Management of Dyslipidaemias. *European Heart Journal* 2016; 37:2999–3058.
- Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH.** Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nature Genetics* 2005; 37:161–165.
- Cuda C, Badawi A, Karmali M, El-Sohemy A.** Polymorphisms in Toll-like receptor 4 are associated with factors of the metabolic syndrome and modify the association between dietary saturated fat and fasting high-density lipoprotein cholesterol. *Metabolism* 2011; 60:1131–1135.
- Díaz-Ruiz M, Martínez-Triguero ML, López-Ruiz A, Fernández-de la Cruz F, Bañuls C, Hernández-Mijares A.** Metabolic disorders and inflammation are associated with familial combined hyperlipemia. *Clinica Chimica Acta* 2019; 490:194–199.
- Ding K, Kullo IJ.** Molecular population genetics of PCSK9: a signature of recent positive selection. *Pharmacogenetics and Genomics* 2008; 18:169–179.
- Echeverría P, Guardiola M, González M, Carles Vallvé J, Puig J, Bonjoch A, et al.** Association between lipid genetic and immunological status in chronically HIV-infected patients. *Journal of the International AIDS Society* 2014; 17:19555.
- Futema M, Shah S, Cooper JA, Li K, Whittall RA, Sharifi M, et al.** Refinement of variant selection for the LDL cholesterol genetic risk score in the diagnosis of the polygenic form of clinical familial hypercholesterolemia and replication in samples from 6 countries. *Clinical Chemistry* 2015; 61:231–238.
- Guardiola M, Cofán M, de Castro-Oros I, Cenarro A, Plana N, Talmud PJ, et al.** APOA5 variants predispose hyperlipidemic patients to atherogenic dyslipidemia and subclinical atherosclerosis. *Atherosclerosis* 2015c; 240:98–104.
- Guardiola M, Echeverría P, González M, Vallvé JC, Puig J, Clotet B, et al.** Polymorphisms in LPL, CETP, and HL Protect HIV-Infected Patients from Atherogenic Dyslipidemia in an Allele-Dose-Dependent Manner. *AIDS Research and Human Retroviruses* 2015b; 31:882–888.

- Guardiola M, Plana N, Ibarretxe D, Cabré A, González M, Ribalta J, et al.** Circulating PCSK9 levels are positively correlated with NMR-assessed atherogenic dyslipidaemia in patients with high cardiovascular risk. *Clinical Science* 2015a; 128:877–882.
- Guardiola M, Ribalta J.** Update on APOA5 Genetics: Toward a Better Understanding of Its Physiological Impact. *Current Atherosclerosis Reports* 2017; 19:30.
- Haneklaus M, O’Neill LAJ.** NLRP3 at the interface of metabolism and inflammation. *Immunological Reviews* 2015; 265:53–62.
- Holmes M V., Shah SH, Angelakopoulou A, Khan T, Swerdlow D, Kuchenbaecker K, et al.** A report on the Genetics of Complex Diseases meeting of the British Atherosclerosis Society, Cambridge, UK, 17–18 September 2009. *Atherosclerosis* 2010; 208:599–602.
- Howard B V, Ruotolo G, Robbins DC.** Obesity and dyslipidemia. *Endocrinology and Metabolism Clinics of North America* 2003; 32:855–867.
- Humphries SE, Whittall RA, Hubbart CS, Maplebeck S, Cooper JA, Soutar AK, et al.** Genetic causes of familial hypercholesterolaemia in patients in the UK: relation to plasma lipid levels and coronary heart disease risk. *Journal of Medical Genetics* 2006; 43:943–949.
- Instituto Nacional de Estadística.** Defunciones según la Causa de Muerte Principales causas de muerte por grupos de enfermedades 1. 2017.
- Insull W.** The Pathology of Atherosclerosis: Plaque Development and Plaque Responses to Medical Treatment. *The American Journal Medicine* 2009; 122:S3–14.
- International HapMap Consortium TIH.** A haplotype map of the human genome. *Nature* 2005; 437:1299–1320.
- Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, et al.** Common variants at 30 loci contribute to polygenic dyslipidemia. *Nature Genetics* 2009; 41:56–65.
- Khera A V., Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al.** Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics* 2018; 50:1219–1224.

- Klen J, Goričar K, Janež A, Dolžan V.** NLRP3 Inflammasome Polymorphism and Macrovascular Complications in Type 2 Diabetes Patients. *Journal of Diabetes Research* 2015; doi:10.1155/2015/616747.
- Lamiquiz-Moneo I, Blanco-Torrecilla C, Bea AM, Mateo-Gallego R, Pérez-Calahorra S, Baila-Rueda L, et al.** Frequency of rare mutations and common genetic variations in severe hypertriglyceridemia in the general population of Spain. *Lipids in Health and Disease* 2016; 15:82.
- Li T, Jing J, Sun L, Jiang B, Xin S, Yang J, et al.** TLR4 and MMP2 polymorphisms and their associations with cardiovascular risk factors in susceptibility to aortic aneurysmal diseases. *Bioscience Reports* 2019; doi:10.1042/BSR20181591.
- Lu Y, Dollé MET, Imholz S, van 't Slot R, Verschuren WMM, Wijmenga C, et al.** Multiple genetic variants along candidate pathways influence plasma high-density lipoprotein cholesterol concentrations. *Journal of Lipid Research* 2008; 49:2582–2589.
- Mallol R, Amigó N, Rodríguez MA, Heras M, Vinaixa M, Plana N, et al.** Liposcale: a novel advanced lipoprotein test based on 2D diffusion-ordered 1H NMR spectroscopy. *Journal of Lipid Research* 2015; 56:737–746.
- Murphy SA, Pedersen TR, Gaciong ZA, Ceska R, Ezhov M V., Connolly DL, et al.** Effect of the PCSK9 Inhibitor Evolocumab on Total Cardiovascular Events in Patients With Cardiovascular Disease. *JAMA Cardiology* 2019. doi:10.1001/jamacardio.2019.0886.
- Nakajima K, Saito T, Tamura A, Suzuki M, Nakano T, Adachi M, et al.** Cholesterol in remnant-like lipoproteins in human serum using monoclonal anti apo B-100 and anti apo A-I immunoaffinity mixed gels. *Clinica Chimica Acta* 1993; 223:53–71.
- Peng D, Jiang F, Zhang R, Tang S, Chen M, Yan J, et al.** Association of Toll-like Receptor 4 Gene polymorphisms with susceptibility to type 2 diabetes mellitus in the Chinese population. *Journal of Diabetes* 2015; 7:485–492.
- Peters BJM, Pett H, Klungel OH, Stricker BHC, Psaty BM, Glazer NL, et al.** Genetic variability within the cholesterol lowering pathway and the effectiveness of statins in reducing the risk of MI. *Atherosclerosis* 2011; 217:458–464.
- Pons D, Trompet S, Craen AJM de, Thijssen PE, Quax PHA, Vries MR de, et al.** Genetic variation in PCAF, a key mediator in epigenetics, is associated with

- reduced vascular morbidity and mortality: evidence for a new concept from three independent prospective studies. *Heart* 2011; 97:143–150.
- Pontillo A, Oshiro TM, Girardelli M, Kamada AJ, Crovella S, Duarte AJS.** Polymorphisms in inflammasome' genes and susceptibility to HIV-1 infection. *Journal of Acquired Immune Deficiency Syndromes* 2012; 59:121–125.
- Qin Z, Zhou K, Li Y, Wang J, Cheng W, Hu C, et al.** Remnant lipoproteins play an important role of in-stent restenosis in type 2 diabetes undergoing percutaneous coronary intervention: a single-centre observational cohort study. *Cardiovascular Diabetology* 2019; 18:11.
- Ramasamy I.** Update on the molecular biology of dyslipidemias. *Clinica Chimica Acta* 2016; 454:143–85.
- Ribalta J, Figuera L, Fernández-Ballart J, Vilella E, Castro Cabezas M, Masana L, et al.** Newly identified apolipoprotein AV gene predisposes to high plasma triglycerides in familial combined hyperlipidemia. *Clinical Chemistry* 2002; 48:1597–1600.
- Roberts CGP, Shen H, Mitchell BD, Damcott CM, Shuldiner AR, Rodriguez A.** Variants in scavenger receptor class B type I gene are associated with HDL cholesterol levels in younger women. *Human Heredity* 2007; 64:107–113.
- Rock E, Winklhofer-Roob BM, Ribalta J, Scotter M, Vasson MP, Brtko J, et al.** Vitamin A, vitamin E and carotenoid status and metabolism during ageing: functional and nutritional consequences (VITAGE PROJECT). *Nutrition, metabolism, and cardiovascular diseases: NMCD* 2001; 11:70–73.
- Rosenson RS, Najera SD, Hegele RA.** Heterozygous familial hypercholesterolemia presenting as chylomicronemia syndrome. *Journal of Clinical Lipidology* 2017; 11:294–296.
- Sarraju A, Knowles JW.** Genetic Testing and Risk Scores: Impact on Familial Hypercholesterolemia. *Frontiers in Cardiovascular Medicine* 2019; 6:5.
- Schaefer EJ, Geller AS, Endress G.** The biochemical and genetic diagnosis of lipid disorders. *Current Opinion in Lipidology* 2019; 30:56–62.
- Sharifi M, Futema M, Nair D, Humphries SE.** Polygenic Hypercholesterolemia and Cardiovascular Disease Risk. *Current Cardiology Reports* 2019; 21:43.

- Shiffman D, Louie JZ, Caulfield MP, Nilsson PM, Devlin JJ, Melander O.** LDL subfractions are associated with incident cardiovascular disease in the Malmö Prevention Project Study. *Atherosclerosis* 2017; 263:287–292.
- Silbernagel G, Scharnagl H, Kleber ME, Delgado G, Stojakovic T, Laaksonen R, et al.** LDL triglycerides, hepatic lipase activity, and coronary artery disease: An epidemiologic and Mendelian randomization study. *Atherosclerosis* 2019; 282:37–44.
- Sturm AC, Knowles JW, Gidding SS, Ahmad ZS, Ahmed CD, Ballantyne CM, et al.** Clinical Genetic Testing for Familial Hypercholesterolemia: JACC Scientific Expert Panel. *Journal of the American College of Cardiology* 2018; 72:662–680.
- Surendran RP, Visser ME, Heemelaar S, Wang J, Peter J, Defesche JC, et al.** Mutations in LPL , APOC2 , APOA5 , GPIHBP1 and LMF1 in patients with severe hypertriglyceridaemia. *Journal of Internal Medicine* 2012; 272:185–196.
- Taskinen M-R.** Diabetic dyslipidemia. *Atherosclerosis Supplements* 2002; 3:47–51.
- Teslovich TM, Musunuru K, Smith A V., Edmondson AC, Stylianou IM, Koseki M, et al.** Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010; 466:707–713.
- Vallvé J-C, Serra N, Zalba G, Fortuño A, Beloqui Ó, Ferre R, et al.** Two Variants in the Fibulin2 Gene Are Associated with Lower Systolic Blood Pressure and Decreased Risk of Hypertension. *PLoS ONE* 2012; 7:e43051.
- Wang Y, Shumansky K, Sin DD, Man SFP, Akhbar L, Connett JE, et al.** Associations of interleukin-1 gene cluster polymorphisms with C-reactive protein concentration and lung function decline in smoking-induced chronic obstructive pulmonary disease. *International Journal of Clinical and Experimental Pathology* 2015; 8:13125–13135.
- Wei FJ, Cai CY, Yu P, Lv J, Ling C, Shi WT, et al.** Quantitative candidate gene association studies of metabolic traits in Han Chinese type 2 diabetes patients. *Genetics and Molecular Research* 2015; 14:15471–15481.
- Zhong S, Li L, Shen X, Li Q, Xu W, Wang X, et al.** An update on lipid oxidation and inflammation in cardiovascular diseases. *Free Radical Biology and Medicine* 2019. doi:10.1016/J.FREERADBIOMED.2019.03.036.

## 7. Autoevaluación

Gracias a este trabajo he podido ver de primera mano cómo funciona un equipo de investigación desde dentro. He aprendido que hay múltiples formas de afrontar un mismo reto y que la combinación de varias estrategias puede llevar a obtener resultados más comprensibles.

Debido a la temática del proyecto, he enfocado mis conocimientos hacia un área que no había trabajado demasiado durante la carrera. Además, al tratarse de un estudio que investiga enfermedades de origen genético complejo, ha sido necesario trabajar con un volumen muy elevado de datos. Desde mi punto de vista, el mayor reto de todos ha sido aprender a manejar a través de programas estadísticos esta gran cantidad de datos y diseñar estrategias para buscar relaciones entre ellos.

Valoro muy positivamente la experiencia obtenida realizando este trabajo ya que me ha permitido poner en práctica muchos conocimientos que no había utilizado hasta el momento. Además, supone mi primera experiencia en el área de la investigación y me anima a seguir avanzando en esta dirección.

## 8. Anexo

Tabla Anexo 1. Breve descripción de cada gen y de la función de la proteína codificada. También se muestra la cantidad de individuos de los que se dispone de resultados para cada SNP

<b>Gen</b>	<b>Nombre completo del gen</b>	<b>Breve descripción del gen y función de la proteína</b>	<b>SNP</b>	<b>Referencias</b>	<b>n</b>
<b>APOA5</b>	<i>apolipoprotein A5</i>	Codifica para una apolipoproteína involucrada en la regulación de los niveles de TG en plasma. También es un componente de las HDL y en menor medida de las VLDLs y los quilomicrones.	rs3135506	(Guardiola <i>et al.</i> , 2015c)	461
			rs964184	(Kathiresan <i>et al.</i> , 2009)	461
			rs662799	(Guardiola <i>et al.</i> , 2015c)	480
<b>APOB</b>	<i>apolipoprotein B</i>	Codifica para la lipoproteína principal de los quilomicrones y de las LDL. En plasma presenta dos isoformas distintas: la apoB-48 y la apoB-100 (la primera producida por el intestino y la segunda por el hígado).	rs1042034	(Teslovich <i>et al.</i> , 2010)	460
<b>APOC3</b>	<i>apolipoprotein C3</i>	Codifica para un componente proteico de las TRLs incluyendo VLDLs, HDLs y quilomicrones. La proteína también participa en el metabolismo de los TRLs.	rs4520	(Echeverría <i>et al.</i> , 2014)	461
			rs5128	(Echeverría <i>et al.</i> , 2014)	461
<b>APOE</b>	<i>apolipoprotein E</i>	Codifica para una de las apoproteínas de los quilomicrones. Es esencial para el catabolismo normal de los constituyentes ricos en TG de las lipoproteínas.	rs439401	(Teslovich <i>et al.</i> , 2010)	461
<b>CASP3</b>	<i>caspase 3</i>	Codifica para una proteína con un papel central en la ejecución de la apoptosis.	rs6948	tagSNP	460
			rs2705897	tagSNP	461
<b>CETP</b>	<i>cholesteryl ester transfer protein</i>	Codifica para una proteína que se queda en plasma donde está involucrada en la transferencia de éster de colesterol de las HDL a otras lipoproteínas.	rs708272	(Guardiola <i>et al.</i> , 2015b)	461
<b>FBLN2</b>	<i>fibulin 2</i>	Codifica para una proteína de la matriz extracelular de la familia de las fibulinas. Esta proteína se une a varios ligandos extracelulares y al calcio.	rs3732666	(Vallvé <i>et al.</i> , 2012)	451
			rs1061376	(Vallvé <i>et al.</i> , 2012)	461
<b>FBLN5</b>	<i>fibulin 5</i>	Codifica para una proteína que promueve la adhesión de las células endoteliales a través de la interacción con integrinas. Se expresa sobre	rs2430347	tagSNP	460
			rs2284340	tagSNP	461

		todo en vasos sanguíneos en desarrollo, dañados o con lesiones ateroscleróticas.			
<b>FASLG</b>	<i>Fas ligand</i>	Es un gen de la superfamilia de factores de necrosis tumorales. Codifica para una proteína transmembranal que induce la apoptosis y es esencial para la regulación del sistema inmunitario.	rs763110	tagSNP	461
<b>GCKR</b>	<i>glucokinase regulator</i>	Codifica para una proteína reguladora que inhibe a la glucoquinasa en el hígado y en los islotes pancreáticos. Se considera una de las variantes del gen está involucrada en un tipo de diabetes y en la regulación de TG.	rs1260326	(Teslovich <i>et al.</i> , 2010)	460
<b>HL</b>	<i>hepatic lipase lipase C, hepatic type</i>	Codifica para una enzima producida en el hígado y que se libera al torrente sanguíneo donde ayuda en la conversión de VLDLs y IDLs a LDLs. La enzima también ayuda a transportar HDLs al hígado y a mantener las moléculas transportadoras de ácidos grasos en equilibrio regulando la formación de LDLs y HDLs.	rs1800588	(Guardiola <i>et al.</i> , 2015b)	461
<b>IL1A</b>	<i>interleukin 1 alpha</i>	Codifica para una citoquina involucrada en varias respuestas inmunitarias, procesos inflamatorios y hematopoyesis. Se produce especialmente cuando hay daño celular.	rs1304037	tagSNP	461
			rs17561	tagSNP	461
			rs1894399	tagSNP	461
<b>IL1B</b>	<i>interleukin 1 beta</i>	Codifica para una citoquina producida por macrófagos. Tiene un papel muy importante como mediadora de la respuesta inflamatoria, la proliferación celular, la diferenciación y la apoptosis.	rs1143627	tagSNP	461
			rs1143633	tagSNP	461
<b>IL1RN</b>	<i>interleukin 1 receptor antagonist</i>	Codifica para una proteína que inhibe la actividad de la IL1A y la IL1B, modulando muchas de las respuestas inmunitarias e inflamatorias reguladas por las interleucinas 1.	rs579543	tagSNP	461
<b>KAT2B</b>	<i>lysine acetyltransferase 2B</i>	Codifica para una proteína con un papel importante en la regulación transcripcional a través de la modificación de histonas. También se asocia a otras proteínas involucradas en el crecimiento y diferenciación celular.	rs2948080	(Pons <i>et al.</i> , 2011)	461
<b>LPL</b>	<i>lipoprotein lipase</i>	Codifica para una lipoproteína que se expresa en el corazón, el músculo y el tejido adiposo. La lipoproteína tiene como función hidrolizar TG y mediar en la absorción de lipoproteínas.	rs328	(Guardiola <i>et al.</i> , 2015b)	460
			rs12678919	(Teslovich <i>et al.</i> , 2010)	461
			rs268	(Lu <i>et al.</i> , 2008)	461

<b>MTTP</b>	<i>microsomal triglyceride transfer protein</i>	Codifica para una subunidad del <i>heterodimeric microsomal triglyceride transfer protein</i> , el cual tiene un papel central en el ensamblaje de las lipoproteínas.	rs1800591	(Echeverría <i>et al.</i> , 2014)	480
<b>NDST1</b>	<i>N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 1</i>	Codifica para una enzima transmembranal del aparato de Golgi que modifica el heparán sulfato, un polisacárido de reserva energética y estructural.	rs2748217	tagSNP	460
			rs3797620	tagSNP	460
<b>NLRP3</b>	<i>NLR family pyrin domain containing 3</i>	Codifica para una proteína que interacciona con un complejo que participa en la regulación de la respuesta inmunitaria y la apoptosis.	rs10754558	(Pontillo <i>et al.</i> , 2012)	461
			rs35829419	(Klen <i>et al.</i> , 2015)	461
<b>PCSK9</b>	<i>proprotein convertase subtilisin/kexin type 9</i>	Codifica para una proteína autocatalítica que se secreta de forma inactiva a la matriz extracelular. Se expresa en el hígado, el intestino y el riñón donde participan en el metabolismo del colesterol y los ácidos grasos en sangre.	rs505151	(Peters <i>et al.</i> , 2011)	459
			rs540796	(Humphries <i>et al.</i> , 2006)	460
			rs562556	(Ding & Kullo, 2008)	461
<b>SCARB1</b>	<i>scavenger receptor class B member 1</i>	Codifica para un receptor de membrana proteico plasmático para las HDLs. Esta proteína regula la transferencia de colesterol a las HDLs.	rs5888	(Roberts <i>et al.</i> , 2007)	461
<b>SCD1</b>	<i>stearoyl-CoA desaturase</i>	Codifica para una enzima involucrada en la biosíntesis de los ácidos grasos, principalmente del ácido oleico.	rs3771243	tagSNP	460
			rs3732164	tagSNP	461
<b>TLR4</b>	<i>toll like receptor 4</i>	Codifica para una proteína de la familia TLR, que tiene un papel fundamental en el reconocimiento y la activación de la inmunidad innata.	rs2770146	tagSNP	460
			rs1927911	(Peng <i>et al.</i> , 2015)	461
			rs11536896	(Peng <i>et al.</i> , 2015)	461
			rs1927914	(Li <i>et al.</i> , 2019)	461
<b>VEGFA</b>	<i>vascular endothelial growth factor A</i>	Codifica para un factor de crecimiento que induce a la proliferación y migración de las células endoteliales vasculares.	rs4986790	(Cuda <i>et al.</i> , 2011)	461
			rs833068	(Abhary <i>et al.</i> , 2009)	461
<b>VLDLR</b>	<i>very low-density lipoprotein receptor</i>	Codifica para una proteína involucrada en la endocitosis específica de ligandos. Tiene un papel importante en el metabolismo de los TG de las VLDLs.	rs1454626	(Echeverría <i>et al.</i> , 2014)	461

*n* representa el número de individuos. tagSNP: cuando se escoge una región del genoma en desequilibrio de unión.