



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**Maestría y Doctorado en Ciencias Bioquímicas**

**CARACTERIZACIÓN DEL *LOCUS* DEL GEN *LRPPRC* Y SU PAPEL EN LA  
HIPERTRIGLICERIDEMIA EN HUMANOS**

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestra en Ciencias

PRESENTA:

QFB. Alejandra Yesamín Chávez Juárez

**TUTORA PRINCIPAL**

Dra. MARIA TERESA TUSIÉ LUNA  
IIB/INCMNSZ

**MIEMBROS DEL COMITÉ TUTOR**

Dra. LORENA AGUILAR ARNAL  
Instituto de Investigaciones Biomédicas

Dra. PATRICIA COELLO COUTIÑO  
Facultad de Química

Ciudad Universitaria, Ciudad de México. Noviembre 2022



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.

“CARACTERIZACIÓN DEL *LOCUS* DEL GEN *LRPPRC*  
Y SU PAPEL EN LA HIPERTRIGLICERIDEMIA EN  
HUMANOS”

*"It's okay not to know all the answers. It's better to admit our ignorance than to believe answers that might be wrong."*

*"Pretending to know everything, closes the door to finding out what's really there"*

*-Neil deGrasse Tyson.*

*(Cosmos)*

*"Nuestras ideas deben ser tan amplias como la Naturaleza, si aspiran a interpretarla"*

*-Sherlock Holmes.*

*(Sir Arthur Conan Doyle. Estudio en Escarlata.)*

*"La salud no lo es todo; pero sin ella, todo lo demás es nada"*

*-Arthur Schopenhauer*

## **Agradecimientos académicos**

Esta tesis se realizó bajo la dirección de la Dra. María Teresa Tusié Luna en la Unidad de Biología Molecular y Medicina Genómica del Instituto Nacional De Ciencias Médicas y Nutrición "Salvador Zubirán" (INCMNSZ), Unidad periférica del Instituto de Investigaciones Biomédicas, UNAM.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) que durante el desarrollo de esta tesis se contó con la beca de maestría otorgada con el número de registro 833753.

Se agradece al comité tutorial que asesoró el desarrollo de esta tesis que estuvo conformado por la Dra. María Teresa Tusié Luna (Instituto de Investigaciones Biomédicas, UNAM), la Dra. Lorena Aguilar Arnal (Instituto de Investigaciones Biomédicas, UNAM) y la Dra. Patricia Coello Coutiño (Facultad de Química, UNAM).

Se agradece la revisión, las observaciones y sugerencias de la presente tesis al jurado de examen conformado por:

- Presidente: Dra. Nimbe Torres y Torres, Instituto Nacional de Ciencias Médicas y Nutrición "Salvador Zubirán".
- Vocal: Dr. Víctor Julián Valdés Rodríguez, Instituto de Fisiología Celular, UNAM
- Vocal: Dra. Ileana Patricia Canto Cetina, Facultad de Medicina, UNAM
- Vocal: Dra. Myrian Velasco Torres, Instituto de Fisiología Celular, UNAM
- Secretario: Dra. Mayra Furlan Magaril, Instituto de Fisiología Celular, UNAM

A la M. en C. Yayoi Xóchitl Segura Kato, investigadora de la Unidad de Biología Molecular y Medicina Genómica del INCMNSZ/UNAM por su amplio, constante e invaluable apoyo tanto experimental como de diseño de experimentos.

A la Dra. Alicia Huerta Chagoya y Dra. Hortensia Moreno Macías, investigadoras de la Unidad de Biología Molecular y Medicina Genómica del INCMNSZ/UNAM por el apoyo teórico del proyecto.

Al M. en C. Ignacio Pacheco Bernal, doctorante en el Instituto de Investigaciones Biomédicas/UNAM por el apoyo en el manejo de datos bioinformáticos.

Al Biol. Saúl Cano Colín, Técnico Académico de la Unidad de Biología Molecular y Medicina Genómica en el apoyo técnico.

A la Dra. Ivette Cruz Bautista, investigadora de la Unidad de Investigación de Enfermedades Metabólicas del INCMNSZ y su equipo de trabajo por el apoyo en la captura de pacientes y manejo de muestras humanas.

Al programa de Maestría y Doctorado en Ciencias Bioquímicas de la Facultad de Química de la UNAM.

Al Instituto Nacional De Ciencias Médicas y Nutrición "Salvador Zubirán".

## **Agradecimientos personales**

A la memoria de mi padre, mi ejemplo de garra y pasión que se necesita en cualquier aspecto de la vida. Love you, mi papá. Always and forever.

A mi madre, por su apoyo incondicional, opiniones, consejos; que sin ellos y sin ella, yo no podría estar en donde estoy. Love you, mom.

A mis hermanas, Astrid y Ana, por su compañía, apoyo, complicidad y cariño; que sin los cuales no me sentiría completa. Sisternas ¡Cantemos!

A todo el equipo de la Unidad de Biología Molecular y Medicina Genómica:

A la Dra Tusié: muchas gracias por aceptarme en su equipo, por todo su apoyo y tutoría.

A Andrea: me faltan palabras para expresar todo lo que te agradezco, manis. Muchas gracias por tu apoyo, ayuda, *pep talk* y más que nada, tu amistad. También por todos los momentos amenos en el club nocturno, que no faltaron. Sin ti, esta tesis no se hubiera terminado.

A Yayoi: muchísimas gracias por todo su apoyo y amistad. Agradezco mucho su ayuda para mi crecimiento tanto académico como personal. Disfruté muchísimo compartir con ud. estos años en el lab. ありがとうございます

A Dany, Alice, Saúl, Dra. Hortensia, María Luisa, Ake, Tere Guerra, Jesús y Claudia.

*And last but not least*, a mi mejor amigo Antoño por todos sus excelentes consejos, sugerencias y apoyo.

A todo el equipo de la Unidad de Investigación de Enfermedades Metabólicas.

Al Dr. Rubén Gutiérrez y a la M. en C. Teresa Zariñán de la Unidad de Apoyo a la Investigación RAI del INCMNSZ.

# ÍNDICE

1. RESUMEN.....	11
2. ABSTRACT.....	12
3. INTRODUCCION.....	13
Hipertrigliceridemia.....	14
Componente genético de la HTG.....	16
2. Antecedentes directos.....	19
3. <i>LRPPRC</i> .....	20
Función de la proteína <i>LRPPRC</i> .....	22
<i>LRPPRC</i> en el metabolismo.....	24
<i>LRPPRC</i> en patologías humanas.....	27
Síndrome de Leigh.....	27
Enfermedad de Parkinson.....	27
Cáncer.....	28
Patologías varias.....	30
4. PREGUNTA DE INVESTIGACIÓN.....	32
5. HIPÓTESIS.....	32
6. OBJETIVO GENERAL.....	32
OBJETIVOS PARTICULARES.....	32
7. MATERIALES Y MÉTODOS.....	33
Participantes del estudio y muestras.....	33
Secuenciación Sanger.....	35
Diseño de oligonucleótidos.....	35
PCR de amplificación.....	35
PCR de secuenciación.....	36
Portales bioinformáticos para recolección de información.....	36
Cultivo celular.....	36
Extracción de ácidos nucleicos.....	37
Síntesis de cDNA.....	37
PCR cuantitativa.....	37
Análisis estadístico.....	37
8. RESULTADOS.....	38
Búsqueda de nuevas variantes en los exones del gen <i>LRPPRC</i> .....	38
Búsqueda de isoformas de <i>LRPPRC</i> .....	45
Estudios de niveles de expresión de genes en tejido hepático y línea celular humana HepG2.....	50
Nivel de expresión para <i>LRPPRC</i> .....	50
Validación del método para expresión de <i>LRPPRC</i> .....	50
PCR cuantitativa para la expresión de <i>LRPPRC</i> en muestras de tejido hepático de pacientes portadores y no portadores.....	52
Nivel de expresión para genes mitocondriales.....	54
Validación del método para expresión de genes mitocondriales.....	54
PCR cuantitativa para la expresión de genes mitocondriales en muestras de tejido hepático de pacientes portadores y no portadores.....	57
Búsqueda bioinformática de interacciones a larga distancia y co-localización con estados de la cromatina.....	59
Búsqueda de variantes en la región de interacción dentro del intrón 23 de <i>LRPPRC</i> .....	64
9. DISCUSIÓN.....	71



Variantes identificadas en la secuenciación de los exones del gen <i>LRPPRC</i> .....	71
Análisis de las posibles isoformas identificadas de <i>LRPPRC</i> .....	74
Estudios de niveles de expresión de genes en tejido hepático y línea celular humana HepG2.....	77
Nivel de expresión para <i>LRPPRC</i> .....	77
Nivel de expresión para genes mitocondriales.....	77
Búsqueda de interacciones a larga distancia y co-localización con estados de la cromatina.....	79
Análisis bioinformático de implicaciones a factores de transcripción por las variantes identificadas en la región del intrón 23 de <i>LRPPRC</i> y dos variantes del haplotipo de riesgo.....	81
10. CONCLUSIONES.....	88
11. ANEXO.....	89
Información suplementaria.....	89
Metodologías suplementarias.....	99
Consentimiento informado.....	104
Carta de aprobación del comité de ética.....	113
12. BIBLIOGRAFÍA.....	114

## INDICE DE FIGURAS

Figura 3.1. Signos clínicos del depósito de lípidos presente en la HTG severa.....	15
Figura 3.2. Complejidad de las bases genéticas de la HTG.....	19
Figura 3.3. Representación gráfica de la estructura del gen <i>LRPPRC</i> y su proteína.....	20
Figura 3.4. Las proteínas que contienen motivos PPR están involucrados en transcripción, traducción, procesamiento y estabilidad del RNA mitocondrial.....	23
Figura 3.5. Esquema del posible mecanismo de acción por el cual <i>LRPPRC</i> tiene influencia en la $\beta$ -oxidación de ácidos grasos y esteatosis hepática.....	26
Figura 3.6. Interacciones de <i>LRPPRC</i> con diferentes enfermedades y procesos biológicos.....	30
Figura 8.1. Búsqueda de isoformas de <i>LRPPRC</i> .....	46
Figura 8.2. Búsqueda de isoformas de <i>LRPPRC</i> con 3 condiciones de síntesis de cDNA.....	49
Figura 8.3. Curvas estándar a concentraciones decrecientes de cDNA de HepG2 para <i>LRPPRC</i> (gen blanco) y <i>GUS</i> (gen constitutivo).....	51
Figura 8.4. Expresión relativa de <i>LRPPRC</i> en muestras de tejido hepático de portadores y no portadores de las variantes de riesgo.....	53
Figura 8.5. Curvas estándar a concentraciones decrecientes de cDNA de HepG2 para 6 genes mitocondriales ( <i>ND1</i> , <i>ND6</i> , <i>COX1</i> , <i>CYTB</i> , <i>ATP6</i> y <i>ATP8</i> ) y 3 constitutivos ( <i>h18s</i> , <i>GAPDH</i> y $\beta$ -actina).....	56
Figura 8.6. Expresión relativa de genes mitocondriales en portadores de las variantes de riesgo tipo TAD presentes en el <i>locus</i> de <i>LRPPRC</i> .....	58
Figura 8.7. Estructuras tipo TAD presentes en el <i>locus</i> de <i>LRPPRC</i> .....	61
Figura 8.8. Mapeo del <i>locus</i> de <i>LRPPRC</i> .....	63
Figura 8.9. Esquema de la posición de elementos de respuesta a diferentes factores de transcripción dentro del <i>loci</i> de las 2 variantes VI1 y VI2.....	68
Figura 8.10. Esquema de la posición de elementos de respuesta a diferentes factores	

de transcripción dentro del loci de las variantes V9 y V15.....	70
---	----

INDICE DE TABLAS

Tabla 3.1. Variantes de genes asociados a concentración de niveles séricos de TG en una cohorte de población europea identificados mediante GWAS.....	17
Tabla 3.2. Variantes de genes asociados a HTG en una cohorte de población mexicana identificados mediante GWAS.....	18
Tabla 3.3. Implicaciones de la alta expresión de LRPPRC en diferentes tipos de cáncer.	29
Tabla 7.1. Información de los pacientes cuyas muestras de sangre fueron utilizadas para la secuenciación de exones de LRPPRC.....	34
Tabla 7.2. Información de los pacientes cuyas muestras de tejido hepático para estudios de expresión y para búsqueda de isoformas de LRPPRC.....	35
Tabla 8.1. Variantes identificadas en los exones y unión exón-intrón de LRPPRC en los 4 portadores homocigotos de las variantes de riesgo.....	39
Tabla 8.2. Recopilación de información para las variantes compartidas entre los 4 pacientes portadores.....	42
Tabla 8.3. Presencia de las 6 variantes seleccionadas en 19 muestras de no portadores sin HTG.....	43
Tabla 8.4. Pareo de los portadores y no portadores de las variantes de riesgo para muestras de tejido hepático.....	44
Tabla 8.5. Variantes identificadas en la región de interacción dentro del intrón 23 de LRPPRC en los 4 portadores homocigotos de las variantes de riesgo.....	65
Tabla 8.6. Búsqueda de las 2 variantes identificadas en la región 23 en muestras de no portadores sin HTG.....	65
Tabla 8.7. Recopilación de información para las 2 variantes compartidas de manera homocigota entre los 4 pacientes portadores.....	66
Tabla 9.1. Prevalencia de las variantes identificadas en las secciones 8. Búsqueda de nuevas variantes en los exones del gen LRPPRC y 8. Búsqueda de variantes en la región de interacción dentro del intrón 23 de LRPPRC en diferentes poblaciones.....	73
Tabla 9.2. Frecuencia de las variantes pertenecientes al haplotipo de riesgo en el estudio SIGMA.....	74

## ABREVIATURAS

Apo B: Apolipoproteína B

C: colesterol

CHD: Enfermedad de las Arterias Coronarias (*Coronary Heart Disease*)

ChIP: Inmunoprecipitación de la cromatina (*Chromatin immunoprecipitation*)

DNA: Ácido Desoxirribunocléico (*Deoxyribonucleic Acid*)

dsDNA: DNA de doble cadena (*double-stranded DNA*)

DT2: Diabetes Mellitus tipo 2, no insulinodependiente

ECV: Enfermedades Cardiovasculares

EHNA: Esteatohepatitis No Alcohólica

eQTL: *Loci* de expresión de rasgos cuantitativos (*Expression quantitative trait loci*)

EUR: población europea

FT: Factor de transcripción

GWA: Asociación a nivel genoma completo (*Genome-Wide Association*)

GWAS: Estudio de asociación a nivel de genoma completo (*Genome-Wide Association Study*)

H3K27ac: Acetilación de la lisina 27 de la histona 3

H3K4me1: Monometilación de la lisina 4 de la histona 3

HDL: Lipoproteína de Alta Densidad

HGNA: Hígado Graso No Alcohólico

HTG: Hipertrigliceridemia

HTGSF: Hipertrigliceridemia Severa Familiar

LD: Desequilibrio de ligamiento (*Linkage Disequilibrium*)

LDL: Lipoproteína de Baja Densidad

MAF: Frecuencia del alelo menos común (*Minor Allele Frequency*)

mRNA: RNA mensajero

mtDNA: DNA mitocondrial

MX: población mexicana

OMS: Organización Mundial de la Salud

PA: Pancreatitis aguda

PAD: Enfermedad Arterial Periférica (*Peripheral Arterial Disease*)

PPR: Repeticiones del motivo pentatricopéptido (*Pentatricopeptide repeat*)

Q: Quilomicrones

qRT-PCR: PCR de tiempo real

RNA: Ácido Ribunocléico (*Ribonucleic Acid*)

RNP: Ribonucleoproteína

rRNA: RNA ribosomal

SNP: Polimorfismo de un solo nucleótido (*Single Nucleotide Polymorphism*)

TADs: Dominios de Asociación Topológica (*Topologically associated domain*)

TG: Triglicéridos

TPR: Repeticiones del motivo tetratricopéptido (*Tetratricopeptide repeat*)

VLDL: Lipoproteína de Muy Baja Densidad

## 1. RESUMEN

La Hipertrigliceridemia Severa Familiar (HTGSF) es una dislipidemia primaria caracterizada por niveles de triglicéridos por arriba de 1000 mg/dL, niveles de Apo B normales y la presencia de por lo menos un familiar afectado con la misma condición. Nuestro grupo de investigación identificó mediante secuenciación de exomas de 68 individuos no relacionados con HTGSF, un haplotipo constituido por 3 variantes de secuencia en el gen *LRPPRC* asociado a un incremento de cerca de 6 veces el riesgo de desarrollar esta condición. En el presente trabajo damos seguimiento a este hallazgo inicial, analizando mediante secuenciación Sanger, la posible presencia de variantes de secuencia en la región codificante, que pudieran estar en desequilibrio de ligamiento con las variantes inicialmente identificadas. Excluimos la presencia de variantes adicionales en la región codificante del gen. Adicionalmente, la variante rs7593842 que forma parte del haplotipo de riesgo, se asoció a la presencia de isoformas específicas de RNAm de *LRPPRC* en el tejido hepático de portadores, mientras que no detectamos cambios en los niveles de RNAm asociados a la presencia de esta variante mediante RT-PCR cuantitativa. De forma interesante detectamos la disminución en la expresión de genes mitocondriales en el tejido hepático de portadores de la variante rs7593842. Finalmente, mediante análisis bioinformáticos identificamos 3 potenciales regiones regulatorias a distancia, una en el *locus* de *ABCG5/8*, otra en una región intergénica y la última dentro del *locus* de *LRPPRC*, en una región que abarca 2 de las 3 variantes del haplotipo de riesgo. Por último, mediante secuenciación Sanger se identificaron dos variantes (rs4263155 y rs4600713) dentro de los intrones de *LRPPRC* con potencial afectación en los motivos respuesta de factores de transcripción.

## 2. ABSTRACT

Familial Severe Hypertriglyceridemia (FSH) is a primary dyslipidemia characterized by triglycerides levels above 1000 mg/dL, accompanied by normal ApoB levels and the presence of a family member with the same condition. Through exome sequencing of 68 unrelated FSH probands, our research group identified the presence of an haplotype composed of three genetic variants at the *LRPPRC* locus, associated with a sixfold increase risk to develop FSH. In the present study we followed up on the initial finding analyzing through Sanger sequence of the 38 *LRPPRC* exons, the presence of additional variants in putative linkage disequilibrium with the initial haplotype. We excluded the presence of additional variants in the coding region of the *LRPPRC* gene. We then examined possible differences in the steady-state levels of mRNA between carriers and non-carriers of rs7593842 variant, one of the three variants of the risk haplotype. Although no significant differences were detected, we showed through quantitative RT-PCR that this variant is associated to a specific *LRPPRC* mRNA isoform in liver tissue from carriers. Interestingly, we also showed a decreased expression of mRNAs from several mitochondrial genes in carriers of the rs7593842 variant. Finally, through bioinformatic analyses we found three potential distance regulatory regions, one at the *ABCG5/8* locus, a second one located at an intergenic region and one more within the *LRPPRC* locus, the latest includes 2 of the 3 genetic variants part of the risk haplotype. Within the intergenic region two variants (rs4263155 and rs4600713) were found that putatively affect the binding motifs for transcription factors.

### 3. INTRODUCCIÓN

Las enfermedades cardiovasculares (ECV) son consideradas por la Organización Mundial de la Salud (OMS) como enfermedades no transmisibles y representan la principal causa de muerte en el mundo; en 2019, 17.8 millones de personas murieron debido a las ECV, esto representa el 32.2% de los decesos a nivel mundial<sup>1</sup> y se calcula que en México hay una probabilidad del 15.7% (en sujetos entre 30 y 70 años) de morir por una ECV<sup>2</sup>. Según el reporte del INEGI del 2018, las enfermedades isquémicas del corazón son la primera causa de muerte en la población mexicana<sup>3</sup>.

Las concentraciones anormales de lipoproteínas en la sangre son la característica principal de las dislipidemias, conjunto de enfermedades asintomáticas y principal factor modificable al riesgo cardiovascular; se caracterizan por valores anormales de colesterol (C) total (altos), colesterol de baja densidad (LDL) (altos), triglicéridos (TG) en la sangre (altos) o colesterol de alta densidad (HDL) (bajos). Se ha establecido que la elevación en la concentración de lípidos promueve el desarrollo de aterosclerosis, una condición multifactorial que daña el endotelio de manera crónica y está caracterizada por el depósito de lípidos en la íntima de las arterias de distribución<sup>4</sup>. Esta condición promueve las consecuencias clínicas de las ECV, tales como la enfermedad de las arterias coronarias (CHD, *Coronary Heart Disease*), infarto isquémico, enfermedad arterial periférica (PAD, *Peripheral Arterial Disease*), fallo cardíaco y muerte súbita cardiovascular<sup>4</sup>. La aterosclerosis es una de las principales causas para el desarrollo de las ECV y está involucrada en 3 de las principales causas de muerte en México<sup>5</sup>.

Además, a las dislipidemias se le han asociado a otras condiciones patológicas como a una incidencia mayor de Diabetes Mellitus tipo 2 (DT2), síndrome metabólico, resistencia a la insulina, intolerancia a la glucosa, insuficiencia arterial de miembros inferiores y ser la causa más frecuente de Pancreatitis Aguda (PA) en el caso de la Hipertrigliceridemia (HTG)<sup>6</sup>.

La prevalencia de dislipidemias varía entre grupos étnicos, además de que es influenciada por distintos factores ambientales, tales como sobrepeso, sedentarismo, ingesta elevada de productos hipercalóricos, consumo de tabaco y alcohol, así como falta de adherencia a recomendaciones médicas<sup>7</sup>. La predisposición de presentar una particular dislipidemia se hace evidente en ejemplos de poblaciones en las que se tiene la misma carga genética, pero los factores ambientales son los propicios para desarrollarla, como estrato socioeconómico y tipo de alimentación, por ejemplo, en grupos indígenas de población amerindia de Latinoamérica que permanecieron en su lugar de origen, el 20% presentan una dislipidemia, mientras los que migraron a una zona urbanizada, la prevalencia es cercana al 50%<sup>8</sup>. Otro componente importante es la etnicidad ya que se observan diferencias en las concentraciones de lípidos entre grupos étnicos, por ejemplo, las poblaciones caucásicas tienen una mayor concentración de C; la población asiática concentraciones bajas de HDL y la fracción de la población amerindia asentada en zonas urbanas tienen una mayor prevalencia de HTG, casi la mitad de los adultos la presenta<sup>8</sup>.

La hipoalfalipoproteinemia es la dislipidemia con el mayor prevalencia en adultos mexicanos<sup>9</sup>; es evidente una predisposición genética en la población mexicana a

desarrollar una dislipidemia, ya que casi la mitad de los adultos tiene HTG y/o concentraciones bajas de HDL. Esta carga genética está representada en la presencia de alelos de riesgo para desarrollar estas patologías, que en la población mestiza-mexicana tienen una alta frecuencia (Tabla 3.1)<sup>10</sup>.

Las dislipidemias son un problema de salud pública por ser el factor común entre un riesgo cardiovascular alto, incidencia en DT2, hipertensión arterial y sobrepeso, así como un papel determinante en 3 de las 5 causas más frecuentes de muerte en México<sup>8</sup>.

En años recientes, se ha acumulado evidencia que sugieren que la elevación de triglicéridos séricos está asociada de manera causal a la patofisiología de la aterosclerosis y a su vez a un aumento en el riesgo para padecer un evento cardiovascular<sup>11,12,13</sup>. Se han podido asociar mutaciones en un solo gen a la elevación severa de TG en genes esenciales para el metabolismo de lipoproteínas ricas en TG afectando en la actividad, ensamblaje y transporte de la enzima Lipoproteinlipasa (LPL)<sup>14</sup>; aunque en la mayoría de pacientes con HTG, su elevación de TG es causada por la combinación de variantes genéticas con efectos pequeños (Tabla 3.1, Tabla 3.2) y factores ambientales (volviendo a destacar niveles de glucosa alto, sobrepeso e inactividad física)<sup>6,8</sup>.

## Hipertrigliceridemia

El prefijo *híper-* significa: exceso o grado superior al normal; el sufijo *-emia*: en la sangre<sup>15</sup>; por lo que la HTG se entiende como: elevación de triglicéridos en la sangre. La legislación mexicana determina que niveles séricos en ayuno de TG por arriba de 150 mg/dL es diagnóstico de HTG<sup>5</sup>; se define HTG moderada con valores menores a 500 mg/dL y severa como valores mayores a 500 mg/dL<sup>16</sup>.

Cuando los niveles de TG rebasan los valores de referencia, se necesita determinar si la causa es primaria o secundaria. Las causas primarias suelen ser de origen genético, como pérdida de la función de LPL, deficiencia en la Apo C-II y defectos en LMF1 para maduración intracelular; generalmente observándose niveles de TG elevados como única característica<sup>14,17</sup>. Cuando los niveles elevados de TG no están aislados, puede sospecharse que otra enfermedad sea la causante de estas elevaciones; siendo causas secundarias de TG elevados la DT2, síndrome metabólico, consumo de alcohol, prescripción de estrógenos o glucocorticoides. El grado de HTG aumenta cuando se asocia con estas condiciones, aumentando además con obesidad<sup>8</sup>.

La HTG severa tiene distintas características: 1) niveles de TG en sangre variables de 500-2000 mg/dL<sup>16</sup>; 2) niveles de TG generalmente por arriba del percentil 90 por edad y género de la población; y 3) niveles normales o bajos de Apo B, colesterol LDL y colesterol HDL<sup>5</sup> todos asociados a signos clínicos observables, tales como episodios de dolor abdominal, xantelasmas (depósito amarillento de C en los párpados<sup>18</sup>), xantomas eruptivos (pápulas de 1-4 mm de diámetro distribuidas superficies extensoras de las extremidades, espalda, hombros, glúteos y manos<sup>19</sup>) xantomas tuberosos (nódulos de consistencia firme localizados en tendones extensores de las manos, las rodillas, los codos y el tendón de Aquiles), arco corneal (depósito de ésteres de C en el estroma corneal

formando posteriormente un arco periférico<sup>20</sup>), PA recurrente y hepatoesplenomegalia (incremento del tamaño del hígado y el bazo a sus dimensiones normales<sup>21</sup>). (Figura 3.1). Cuando los niveles de TG son extraordinariamente elevados, mayores a 900 mg/dL, es probable que coexistan con otra comorbilidad (DT2, síndrome metabólico) y factores ambientales que estén contribuyendo a la elevación de TG. La fisiopatología de la HTG tiene que ver con catabolismo alterado de las lipoproteínas quilomicrones (Q) y VLDL<sup>22</sup>. En la HTG severa ocurren varios procesos; hay un catabolismo deficiente de VLDL, baja actividad de LPL y baja depuración de Q y VLDL. Las VLDL alteradas (más grandes en tamaño) contienen una cantidad mayor de TG, causando que la hidrólisis de la LPL sea más lenta y a su vez una depuración retrasada, ya que tanto las VLDL como los Q compiten por el mismo mecanismo de depuración mediante la LPL. Por lo que, cuando la LPL está metabolizando las VLDL alteradas y los Q entran en el sistema circulatorio (generalmente después de una comida alta en lípidos) estos no son procesados y pueden permanecer en la circulación por días, en vez de horas, elevando la concentración sérica de TG; los Q pueden elevar los TG a niveles superiores 8000 mg/dL<sup>17</sup>.

Una variante de esta condición es la HTG familiar en el cual su diagnóstico se establece cuando el sujeto y uno o más de sus familiares de primer o segundo grado, siguen el mismo patrón de HTG severa<sup>8,23</sup>. Aunque se han realizado estudios de asociación genética para elevación de TG,<sup>24,25</sup> no se ha podido asociar variantes genéticas a la HTG familiar<sup>26</sup>. El diagnóstico diferencial es necesario ya que se necesita descartar otras dislipidemias con componente familiar como la hiperlipidemia familiar combinada, deficiencia de lipoproteinlipasa familiar, síndrome de quilomicronemia familiar o la disbetalipoproteinemia familiar. Más recientemente se ha probado que la HTG familiar se puede distinguir de otras dislipidemias por características bioquímicas presentes, tales como los niveles bajos de Apo B, concentraciones mayores de insulina, FGF-21 (*fibroblast growth factor 21*; el factor de crecimiento de fibroblasto 21; involucrado en la regulación del metabolismo de hepatocitos y adipocitos), ANGPTL3 (*angiopoietin-like proteins 3*; la proteína tipo angiopoietina 3; mayor inhibidor de la actividad de LPL) y apo A-II<sup>26</sup>.



Figura 3.1. **Signos clínicos del depósito de lípidos presente en la HTG severa.** A: Xantelasmas. B: Xantomas tuberosos. C: Arco corneal



El reporte nacional para la prevalencia de dislipidemias más detallado es la Encuesta Nacional de Enfermedades Crónicas (ENEC1993), en donde se reporta que la segunda anomalía más frecuente en la población mexicana es la HTG. La prevalencia de niveles de TG mayores a 150 mg/dL es de un 43% y de individuos con valores severos (mayores a 500 mg/dL) es del 2.9%<sup>27</sup>. Datos de la Encuesta Nacional de Salud del año 2000, señalan que la prevalencia de HTG aumentó del 43% a un 49.1% en un lapso de 7 años<sup>28</sup>. El siguiente estudio (ENSANUT 2006) estimó que la prevalencia de HTG disminuyó a un 31.5% y se convirtió en la tercera dislipidemia más frecuente en la población; también se observó en este estudio que la prevalencia de HTG es dos veces más grande en individuos con sobrepeso en comparación con personas delgadas<sup>29</sup>.

En Estados Unidos, los hispanos son el grupo étnico con mayor riesgo de padecer HTG<sup>30</sup> (40.4% para hombre y 34.9% para mujeres), en comparación con las personas caucásicas (39% para hombre y 31.6% para mujeres) y afroamericanas (19.5% para hombre y 13.6% para mujeres). Además, los hispanos tienen tres veces más probabilidades de padecer DT2 y presentar resistencia a la insulina<sup>31</sup>. La prevalencia de la HTG familiar no se ha podido calcular, ya que faltan los estudios de los niveles plasmáticos de TG en familiares en primer y segundo grado<sup>16</sup>.

### **Componente genético de la HTG**

Debido al componente familiar en diferentes condiciones asociadas a lípidos, se han realizado estudios para comprender el contexto genético que pudiera estar influyendo a la fisiopatología y se han asociado diferentes genes y variantes a una mayor probabilidad de desarrollar dislipidemias<sup>32</sup>. Mediante un estudio de asociación de genoma completo (GWAS) con una población de 100 000 individuos de distintas poblaciones europeas, se identificaron 95 *loci* asociados a niveles séricos de TG, C, LDL y HDL<sup>24</sup>. En lo que respecta a TG, se identificaron variantes en 32 genes (Tabla 3.1). Algunas de estas variantes también tienen asociado un eQTL (*Expression quantitative trait loci*; *loci* de expresión de rasgos cuantitativos) por lo menos con un gen dentro de una región de 500 kb y en por lo menos un tejido estudiado (hígado, grasa visceral y grasa subcutánea).

<i>Locus</i>	<i>Crom</i>	<i>SNP líder</i>	<i>Rasgo líder</i>	<i>Otro rasgo</i>	<i>Alelos/MAF</i>	<i>Frecuencia en mexicanos en EUA (MAF)*</i>	<i>Valor-p</i>	<i>eQTL</i>
<b>ANGPTL3</b>	1	rs2131925	TG	TC, LDL	T/G/0.32	0.336	9 x10 <sup>-43</sup>	Sí
<b>GALNT2</b>	1	rs4846914	HDL	TG	A/G/0.40	0.461	4x10 <sup>-21</sup>	
<b>APOB</b>	2	rs1042034	TG	HDL	T/C/0.22	0.273	1x10 <sup>-45</sup>	
<b>GCKR</b>	2	rs1260326	TG	TC	C/T/0.41	0.352	6x10 <sup>-133</sup>	Sí
<b>COBLL1</b>	2	rs10195252	TG		T/C/0.40	0.172	2x10 <sup>-10</sup>	Sí
<b>IRS1</b>	2	rs2972146	HDL	TG	T/G/0.37	0.164	3x10 <sup>-9</sup>	Sí
<b>MSL2L1</b>	3	rs645040	TG		T/G/0.22	0.227	3x10 <sup>-8</sup>	
<b>KLHL8</b>	4	rs442177	TG		T/G/0.41	0.312	9x10 <sup>-12</sup>	
<b>MAP3K1</b>	5	rs9686661	TG		C/T/0.20	0.289	1x10 <sup>-10</sup>	
<b>TIMD4</b>	5	rs6882076	TC	LDL,TG	C/T/0.35	0.250	7x10 <sup>-28</sup>	
<b>HLA</b>	6	rs2247056	TG		C/T/0.25	0.109	2x10 <sup>-15</sup>	
<b>TYW1B</b>	7	rs13238203	TG		C/T/0.04	0.008	1x10 <sup>-9</sup>	
<b>MLXIPL</b>	7	rs17145738	TG	HDL	C/T/0.12	0.055	6x10 <sup>-58</sup>	Sí
<b>PINX1</b>	8	rs11776767	TG		G/C/0.37	0.250	1x10 <sup>-8</sup>	
<b>NAT2</b>	8	rs1495741	TG	TC	A/G/0.22	0.328	5x10 <sup>-14</sup>	Sí
<b>LPL</b>	8	rs12678919	TG	HDL	A/G/0.12	0.055	2x10 <sup>-115</sup>	Sí
<b>TRIB1</b>	8	rs2954029	TG	TC, LDL, HDL	A/T/0.47	0.359	3x10 <sup>-55</sup>	Sí
<b>JMJD1C</b>	10	rs10761731	TG		A/T/0.43	0.281	3x10 <sup>-12</sup>	
<b>CYP26A1</b>	10	rs2068888	TG		G/A/0.46	0.461	2x10 <sup>-8</sup>	
<b>FADS1-2-3</b>	11	rs174546	TG	HDL, TC, LDL	C/T/0.34	0.312	5x10 <sup>-24</sup>	
<b>APOA1</b>	11	rs964184	TG	TC, HDL, LDL	C/G/0.13	0.281	7x10 <sup>-240</sup>	Sí
<b>LRP1</b>	12	rs11613352	TG	HDL	C/T/0.23	0.469	4x10 <sup>-10</sup>	
<b>ZNF664</b>	12	rs4765127	HDL	TG	G/T/0.34	0.312	3x10 <sup>-10</sup>	
<b>CAPN3</b>	15	rs2412710	TG		G/A/0.02	0.055	2x10 <sup>-8</sup>	
<b>FRMD5</b>	15	rs2929282	TG		A/T/0.05	0.125	2x10 <sup>-11</sup>	Sí
<b>LIPC</b>	15	rs1532085	HDL	TC, TG	G/A/0.39	0.320	3x10 <sup>-96</sup>	Sí
<b>CTF1</b>	16	rs11649653	TG		C/G/0.40	0.461	3x10 <sup>-8</sup>	Sí
<b>CETP</b>	16	rs3764261	HDL	TC, LDL, TG	C/A/0.32	0.336	7x10 <sup>-380</sup>	
<b>CILP2</b>	19	rs10401969	TC	TG, LDL	T/C/0.07	0.070	3x10 <sup>-38</sup>	Sí
<b>APOE</b>	19	rs439401	TG		C/T/0.36	0.500	1x10 <sup>-30</sup>	Sí
<b>PLTP</b>	20	rs6065906	HDL	TG	T/C/0.18	0.133	2x10 <sup>-22</sup>	
<b>PLA2G6</b>	22	rs5756931	TG		T/C/0.40	0.398	4x10 <sup>-8</sup>	

Tabla 3.1. **Variantes de genes asociados a concentración de niveles séricos de TG en una cohorte de población europea identificados mediante GWA.** Crom: Cromosoma, MAF: Frecuencia del alelo menos común. \*Proveniente del Proyecto *1000 Genomes* Fase 3. Modificado de Teslovich *et al*, 2010.

Cabe mencionar que este estudio se realizó exclusivamente en población europea y aunque se intentó extrapolar a otras poblaciones (africana y asiática) no se contaba con información genotípica suficiente. Posteriormente se realizó otro estudio de asociación a nivel genoma completo en la población mexicana (n=4361)<sup>25</sup>. Con este acercamiento se pudo determinar que aproximadamente el 50% de las variantes identificadas en población europea también está asociadas a la HTG en población mexicana, pero las variantes con mayor asociación en el estudio con mexicanos no se encontraron en población europea<sup>25</sup> (Tabla 3.3). Se hipotetizó que la diferencia en las frecuencias alélicas de las variantes entre grupos étnicos, la estructura de los bloques con desequilibrio de ligamiento entre los grupos étnicos y la existencia de variantes de riesgo exclusivas de ciertas poblaciones explica las diferencias en el componente genético de la HTG entre las poblaciones europeas y la población mestiza-mexicana<sup>25</sup>.

La población mestizo-mexicana es una población recientemente mezclada con componente Europeo y Nativo-Americano en proporciones similares, cercanas al 50%<sup>33</sup>. Debido a que el componente genético de poblaciones Nativo Americanas está

parcialmente explorado, el estudio de la población mestiza-mexicana representa una oportunidad para identificar genes nuevos y/o vías metabólicas relacionadas a la HTG no antes descubiertas.

Locus	SNP líder en mexicanos	MAF (MX/EUR)	Valor-p	¿Gen presente en el estudio europeo?
<b>APOA5</b>	rs964184	G (0.30/0.12)	5.5X10 <sup>-35</sup>	Sí, mismo SNP
<b>GCKR</b>	rs1260326	T (0.26/0.42)	2.2X10 <sup>-13</sup>	Sí, mismo SNP
<b>LPL</b>	rs12678919	G (0.05/0.12)	2.7X10 <sup>-10</sup>	Sí, mismo SNP
<b>MLXIPL</b>	rs2286276	T (0.12/0.26)	2.2X10 <sup>-06</sup>	Sí, diferente SNP líder SNP líder europeo MAF (MX/EUR) rs17145738 T (0.06/0.12)
<b>TIMD4</b>	rs2036402	C (0.45/0.25)	3.4X10 <sup>-06</sup>	Sí, diferente SNP líder SNP líder europeo MAF (MX/EUR) rs1363232 A (0.10/0.30)
<b>CILP2</b>	rs2228603	T (0.03/0.10)	3X10 <sup>-05</sup>	Sí, diferente SNP líder SNP líder europeo MAF (MX/EUR) rs10401969 C (0.05/0.10)
<b>ANGPTL3</b>	rs10889337	A (0.45/0.35)	3.3X10 <sup>-05</sup>	Sí, diferente SNP líder SNP líder europeo MAF (MX/EUR) rs2131925 G (0.45/0.33)

Tabla 3.2. **Variantes de genes asociados a HTG en una cohorte de población mexicana identificados mediante GWAS.** MX: población mexicana, EUR: población europea; MAF: Frecuencia del alelo menos común. Modificado de Aguilar *et al*, 2014.

Los genes identificados en los estudios GWAS de los dos estudios (población europea y población mexicana) codifican para proteínas que están encargadas del óptimo metabolismo de las lipoproteínas ricas en TG y algunas variantes de secuencia en estos genes son causantes de HTG monogénicas. Estos incluyen a *LPL*, *APOC2*, *APOA5*, *LMF1*, *GPIHBP1* y *GPD1*<sup>34</sup>; los cuales tienen un efecto directo o indirecto con la función de LPL y así reducir considerablemente el metabolismo de las lipoproteínas. Por otro lado, variantes monogénicas en hipotrigliceridemia (niveles bajos de TG séricos) se deben a mutaciones homocigotas en proteínas necesarias para el ensamblaje y secreción de lipoproteínas, como *APOB* y *ANGPTL3*, además de que *ANGPTL3* inhibe a LPL. En el caso del gen *GCKR* (*Glucokinase regulator*, Regulador de la Glucocinasa), se afecta el metabolismo de carbohidratos<sup>34</sup> ya que este gen codifica para la proteína inhibidora de la glucocinasa (enzima regulatoria de la absorción y almacenamiento de la glucosa proveniente de la dieta)<sup>35</sup>. En el hígado, este regulador forma un complejo inhibitorio con la glucocinasa y la variante funcional identificada atenúa la inhibición dependiente de la unión, lo cual promueve síntesis *de novo* de TG y atenuación de la oxidación de ácidos grasos<sup>36</sup>.

La mayoría de los casos de HTG, se debe a defectos poligénicos más que monogénicos; una mayor carga de variantes raras heterocigotas con efecto mayor, además de la acumulación de efectos pequeños causados por SNPs en más de 40 *loci* repartidos en el genoma<sup>14</sup>. La carga genética de niveles severos mayores de 800 mg/dL de TG en adultos es a una compleja mezcla de factores comunes y raros en el genoma de un individuo; estos pacientes son raros en la población porque una gran acumulación de variantes es

estadísticamente poco probable. Por lo regular, los niveles más severos de TG son en pacientes jóvenes con desórdenes monogénicos y con mínima contribución de factores no genéticos<sup>14</sup> (Figura 3.2). Una vez identificados estos *loci* y variantes genéticas es necesario dilucidar su mecanismo de acción y como alteran el metabolismo de TG.

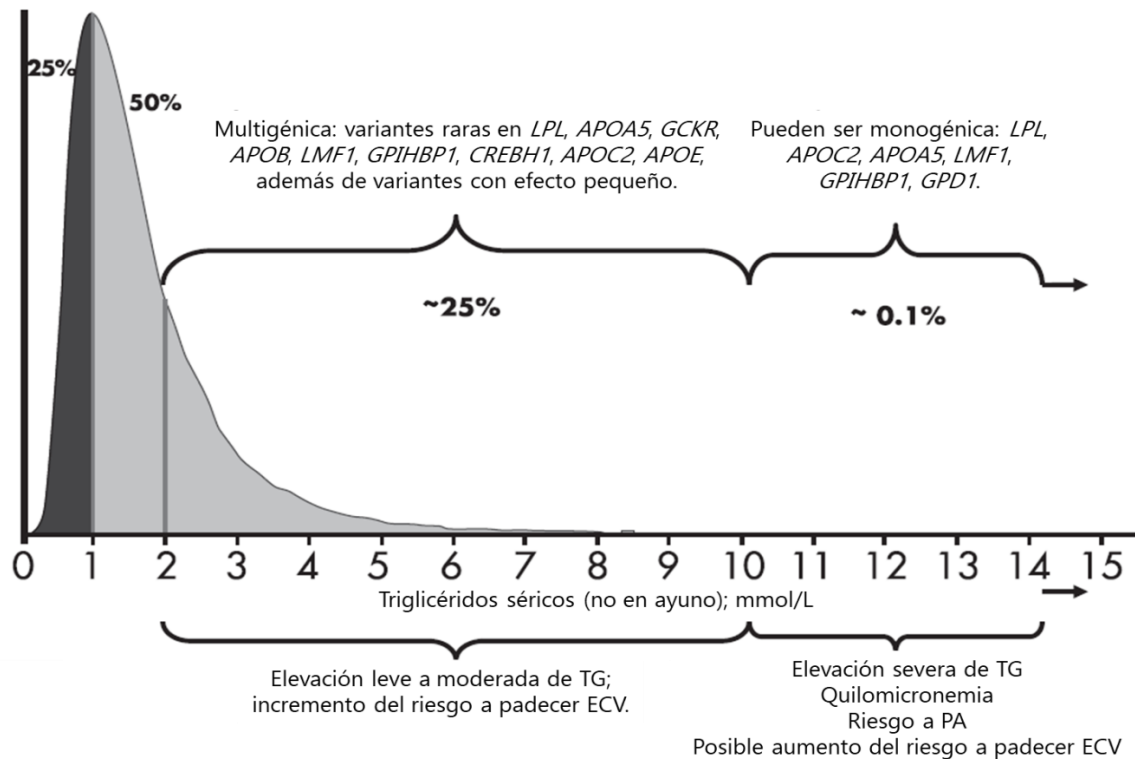


Figura 3.2. **Complejidad de las bases genéticas de la HTG.** Se muestra un aproximado del porcentaje de individuos presente en cada intervalo bioquímico. ECV: Enfermedades cardiovasculares, TG: Triglicéridos, PA: Pancreatitis aguda. Modificado de Lewis *et al*, 2015.

### Antecedentes directos

Recientemente, nuestro grupo de trabajo realizó un estudio de asociación GWAS posterior a la secuenciación masiva del exoma completo de pacientes mexicanos con HTG severa y de pacientes con niveles normales de triglicéridos circulantes para identificar genes y variantes asociadas a la elevación de TG. Mediante esta estrategia se identificó una región cromosómica dentro del gen *LRPPRC* asociado significativamente a un aumento importante en los niveles de triglicéridos séricos ( $p= 1 \times 10^{-6}$ ). En este estudio se identificó dentro del gen *LRPPRC* un haplotipo de tres variantes intrónicas (rs7593842, rs58811869 y rs72877186) y se encontró que ser portador de este haplotipo aumenta 6 veces el riesgo a desarrollar HTGSF (datos aún no publicados).

En otros experimentos funcionales, mediante un experimento de EMSA (*Electrophoresis Mobility Shift Assay*, Ensayo de movilidad electroforética) en la línea celular de hepatocarcinoma HepG2, se comprobó que la variante rs58811869 influye en la unión del

factor de transcripción CTCF, pudiendo tener afectaciones en expresión de genes o en organización de la cromatina, ya que CTCF tiene función como organizador de la cromatina<sup>37,38</sup>.

## LRPPRC

El gen *LRPPRC*, también conocido como *LRP130*, codifica para una proteína rica en leucinas que contiene repeticiones del motivo pentatricopéptido (*Leucine Rich Pentatricopeptide Repeat Containing, LRPPRC*) y se encuentra en la banda 21 del brazo corto del cromosoma 2 (2p21). Contiene 38 exones y 109.5 kb. Se codifica en la cadena antisentido y se conocen 71 distintas isoformas<sup>39</sup> (Tabla suplementaria 1.1). La proteína canónica codificada por *LRPPRC* tiene un tamaño de 130 kDa, contiene 1394 aminoácidos y es rica en leucina. Esta proteína está caracterizada por varios dominios PPR (*Pentatricopeptide repeat* o repeticiones del motivo pentatricopéptido) y por lo tanto pertenece a la familia de proteínas con dominios PPR (Figura 3.3).

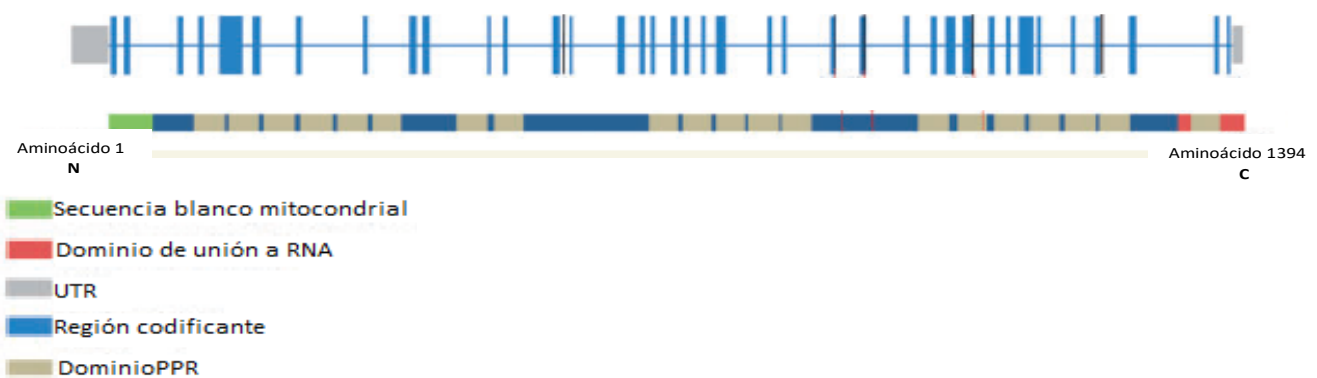


Figura 3.3. **Representación gráfica de la estructura del gen *LRPPRC* y su proteína.** Modificado de Oláhová *et al*, 2015.

Las proteínas con dominio PPR son parte de una familia de proteínas que se unen al RNA mitocondrial. Las proteínas pertenecientes a la familia PPR fueron descubiertas en estudios en plantas (*Arabidopsis thaliana*<sup>40</sup>) y se caracterizan por la presencia de arreglos en tándem de repeticiones derivados de 35 aminoácidos de 2 a 27 repeticiones por péptido<sup>40,41</sup>. El dominio PPR está vinculado a la repetición del motivo tetratricopéptido (*tetratricopeptide repeat, TPR*), aunque estas proteínas están más relacionadas a interacciones proteína-proteína. Se predijo que PPR forma dos  $\alpha$ -hélice antiparalelas que a su vez crearían una súper hélice, con arreglos en tándem y una cavidad central que podría involucrarse en la unión al mitRNA (dentro de la cavidad, la unión sería entre el fondo cargado positivamente al esqueleto de fosfatos cargado negativamente de la cadena de mitRNA) más que de proteína-proteína como los motivos TPR<sup>41</sup>. Aun así, estas interacciones son predictivas, ya que el mecanismo por el cual el dominio PPR interacciona con el RNA aún no ha sido dilucidado. En cuanto a *LRPPRC* se determinó que el sitio de

unión al mitRNA está en la fracción carboxilo terminal de la proteína y está formado por dos de los 22 dominios PPR de la proteína<sup>42</sup>. La mayoría de las proteínas PPR no tienen otro sitio activo, por lo que el procesamiento de RNA se lleve a cabo por otras proteínas que se asocian al complejo PPR-mitRNA<sup>43</sup> y el mitRNA se encuentra estabilizado estructuralmente.

En un inicio se pensó que esta familia de proteínas podía ser exclusiva de plantas, con más de 200 miembros y cuatro subclases (las cuales llegarían a contener de 400 a 600 proteínas), sólo 20 tipos en protistas y una minoría en eucariontes, sin embargo, se encontraron 7 proteínas PPR mitocondriales en humanos<sup>44</sup>. Estas son: **POLRMT** (*mitochondrial RNA polymerase*, RNA polimerasa mitocondrial); **LRPPRC**; **MRPP3** (*mitochondrial RNase P protein 3*, Proteína mitocondrial RNasa P 3); **MRPS27** (*mitochondrial ribosomal protein of the small subunit 27*, Proteína ribosomal mitocondrial de la subunidad pequeña 27); y tres proteínas que contienen dominios PPR (*PPR domain containing proteins*, **PTCD**)<sup>45</sup>. La proteína LRPPRC tiene sitios de fosforilación, que están ubicados en la región carboxilo terminal, en la secuencia de aminoácidos de 1026 al 1138, por lo que se podría inferir que su interacción con el mtRNA podría estar regulado por la fosforilación de serinas y treoninas<sup>46</sup>.

La proteína LRPPRC cumple su función de proveer estabilidad al mitRNA al formar un complejo proteína-ácido nucleico dentro de la mitocondria, pero su gen no está codificado en el mtDNA, sino en el núcleo, por lo que tiene que importarse hacia la mitocondria. Las proteínas que son importadas a la mitocondria tienen señales a lo largo de su secuencia de aminoácidos para ser identificadas, ya sean editables (presecuencias  $\alpha$ -hélices anfipáticas) o no editables (secuencias de aminoácidos y señales de anclaje)<sup>47</sup>. Estas proteínas tienen distintos métodos de importación dependiendo de la zona en la mitocondria a la cual van a realizar su función; cruzar la membrana externa y estar en el espacio intermembranal y/o cruzar la membrana interna. Estas maquinarias de importación están compuestas de varias subunidades, además de hacer uso de la bomba de electrones para su función<sup>47</sup>.

Para describir la importación de LRPPRC, en condiciones *in vitro* se marcó radioactivamente a la proteína y se incubó con mitocondrias aisladas de hígado de ratón, observándose la presencia de una LRPPRC más pequeña al tamaño predicho para esta proteína<sup>45</sup>; esto confirma la idea de que la secuencia objetivo para importe a la mitocondria se pierde a la entrada por la membrana mitocondrial.

En otra condición con un desacoplante de la fosforilación oxidativa como el CCCP (Carbonilcianuro-m-clorofenilhidrazona, inhibidor químico que produce que la membrana interna de la mitocondria sea permeable a protones), no hay corte de la secuencia objetivo para transporte a la mitocondria, por lo que se concluye que el procesamiento e importación de LRPPRC necesitan del potencial de membrana mitocondria<sup>45</sup>.

Por último, en esta evaluación con mitocondrias aisladas, se utilizó una condición de degradación mediante tripsina para determinar cuál de las dos variantes de LRPPRC (con o sin secuencia de importe) está presente dentro de la mitocondria; la cual fue la variante

sin secuencia de importe. No hubo presencia de LRPPRC sin corte, ya que no fue protegida contra la hidrólisis de la tripsina<sup>46</sup>.

Para determinar la localización subcelular de LRPPRC, Mili y colaboradores (2003) observaron mediante fraccionamiento celular que la proteína se encontraba en su mayoría en la mitocondria y en el núcleo, sin embargo, no se descartó una posible contaminación de mitocondria en la fracción nuclear<sup>48</sup>. Tsuchiya y colaboradores (2004) identificaron la presencia de LRPPRC en fracciones nucleares, membrana nuclear y de retículo endoplásmico. En estos trabajos se utilizaron técnicas de inmunoprecipitación por lo que los diferentes anticuerpos utilizados, así como posibles isoformas de LRPPRC pudieron haber influido para esta diferencia de co-localización subcelular<sup>49</sup>. También Sterky y colaboradores (2010) realizaron evaluaciones de la localización de LRPPRC y mediante transfección de plásmidos en células HeLa con expresión de LRPPRC acoplada a la proteína verde fluorescente y otro plásmido con una proteína rojo fluorescente dirigida a la matriz mitocondrial se mostró que existía una superposición perfecta de las señales fluorescente<sup>46</sup>; también mediante tinción inmunohistoquímica con anticuerpo contra LRPPRC y el colorante *Mitotracker* (tinción selectiva para la mitocondria) se demostró que la señal de LRPPRC sólo co-localiza en la mitocondria. Como último experimento, se realizó un fraccionamiento de mitocondrias de hígado de ratón y se identificó que LRPPRC se localiza específicamente en la matriz mitocondrial<sup>46</sup>.

## **Función de la proteína LRPPRC**

*LRPPRC* es un gen nuclear cuya proteína realiza su función dentro de la mitocondria; remodela estructuras secundarias en el mitRNA y mantiene laxa la estructura para facilitar la transcripción y poliadenilación mitocondrial<sup>50</sup>; por lo que a nivel mitocondrial ayuda al correcto funcionamiento de la respiración oxidativa, la oxidación de ácidos grasos y regulando el contenido de lípidos a nivel hepático<sup>51,52</sup>.

El DNA mitocondrial codifica para 37 genes requeridos para la producción de energía; 13 genes son para proteínas de la cadena respiratoria, 22 RNA de transferencia y 2 RNAs ribosomales<sup>53</sup>. La función del mitDNA en la célula depende tanto del número de copias del genoma mitocondrial así como de la integridad de cada molécula de mitDNA<sup>54</sup>.

Dada la participación de LRPPRC en estabilizar a los mitRNAs para su correcto procesamiento, se realizaron experimentos para evaluar si también la funcionalidad de la mitocondria estaría comprometida ante la disminución de la expresión de *LRPPRC*. Mediante *knockdown* graduales de *LRPPRC* con inactivación del 8% a 100% en células en cultivo MCH58 (fibroblastos inmortalizados de piel humana) y el análisis de enriquecimiento de conjuntos de genes (*GSEA*) se demostró que los transcritos de mRNA mitocondrial disminuyen de forma directamente proporcional a la disminución de *LRPPRC*. También se reportó una disminución de la respiración mitocondrial y un aumento en la actividad glucolítica, mientras que no hubo cambios en los niveles de rRNA mitocondriales<sup>55</sup>. Con uso de qPCR (PCR cuantitativa) dirigidos a los mRNA y rRNA mitocondriales se determinó que en las dos condiciones con un porcentaje de inactivación

mayor, había una reducción importante en cada uno de los trece mRNA mitocondriales y que no había afectación en ninguno de los dos rRNA ni en los transcritos policistrónicos. Por lo que se propuso que la proteína LRPPRC es necesaria para la expresión de todos los mRNA y su procesamiento<sup>55</sup>. (Figura 3.4).

Por medio de experimentos *knock-down* para *LRPPRC* en ratones se formaron heterocigotos *Lrpprc*<sup>+/-</sup> y su cruce no produjo homocigotos *Lrpprc*<sup>-/-</sup> viables, ya que en el estadio E8.5 la pérdida completa de *LRPPRC* produjo letalidad embrionaria<sup>56</sup>. Los transcritos extraídos del corazón de estos ratones demostraron una disminución marcada en la mayoría de los mRNA, desde las cuatro semanas de edad, produciendo una pérdida de coordinación en la traducción, la cual resultaba en colas de poliA aberrantes (más cortas) en los animales *LRPPRC*-deficientes. Los niveles de distintos transcritos, sin *LRPPRC*, fueron mucho más altos e incluso inestables, ya que la reserva de mRNAs extraribosomales desapareció y los mRNAs se unieron de manera aberrante a las subunidades de ribosomas libres. Aunque hubo un descenso marcado en niveles del mRNA por la pérdida de *LRPPRC* no hubo un aumento en la transcripción mitocondrial<sup>56</sup>.

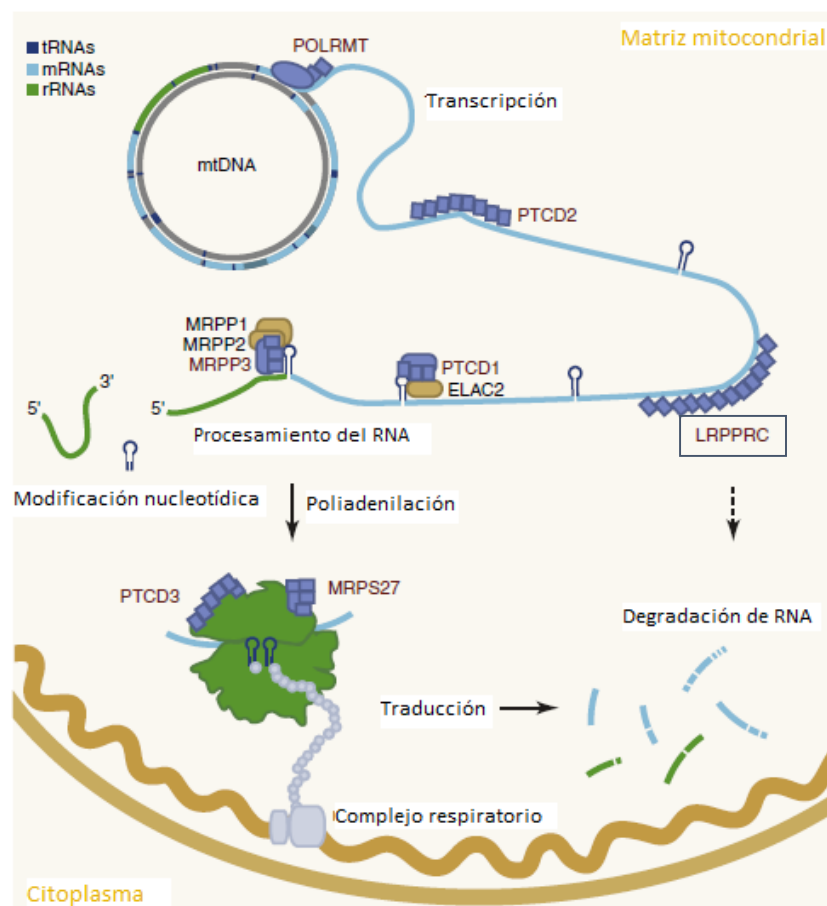


Figura 3.4. **Las proteínas que contienen motivos PPR están involucrados en transcripción, traducción, procesamiento y estabilidad del RNA mitocondrial.** En mamíferos, las siete proteínas identificadas que contienen motivos PPR realizan su función de regulación de expresión



de genes mitocondriales: POLRMT es responsable de la transcripción del genoma mitocondrial, PTC2 regula el procesamiento de del RNA del gen *Cyt b* ya que hay una estructura secundaria en su *locus*, LRPPRC está implicado en la estabilidad de los mRNA mitocondriales codificantes para proteínas de la cadena respiratoria, PTC1 se asocia a los tRNA leucina y actúa como un regulador negativo afectando la abundancia de proteínas codificadas en la mitocondria, MRPP3 es uno de los componentes de la RNasa P y es necesaria para el procesamiento del tRNA mitocondrial, PTC3 y MRPS27 se asocian con la subunidad pequeña del ribosoma mitocondrial teniendo impacto en la síntesis proteínas mitocondriales. Modificado de Rackham *et al*, 2011.

Otro estudio identificó a la proteína LRPPRC como una RNP (ribonucleoproteína) nuclear intermediaria<sup>57</sup>, el cual es un complejo de proteínas que se asocian con pre-mRNA y mRNA, regulando la estabilidad y el manejo de los mRNA maduros. Se demostró, mediante inmunoprecipitación, que LRPPRC se asocia a la proteína de 12 kDa SLIRP, asociada a homeostasis del RNA mitocondrial. Se realizaron las inmunoprecipitaciones con anticuerpos anti-LRPPRC y anti-SLIRP y el complejo entre las dos proteínas fue recuperado (así como mRNA), demostrando su interacción<sup>58</sup>. Se utilizó radiación UV para inducir enlaces cruzados entre RNA y proteínas en la línea celular HeLa, y se identificó a una proteína de 130 kDa, la cual fue sometida a un análisis de espectrometría de masas y se identificó como LRPPRC, además de su unión a la cola de poli-A del mRNA *in vivo*<sup>58</sup>.

### **LRPPRC en el metabolismo**

Debido a la función de brindar estabilidad al mitrRNA y consecuente expresión óptima de proteínas de la cadena respiratoria, la proteína de LRPPRC juega un papel importante en el metabolismo y la fosforilación oxidativa. Mediante el estudio de modelos *knock-down* y de sobre-expresión de LRPPRC en líneas celulares humanas y en ratones transgénicos hígado-específicos de LRPPRC, se evaluó la influencia de este gen en la fosforilación oxidativa y la oxidación de ácidos grasos<sup>51</sup>. En células con sobre-expresión de LRPPRC, hubo un incremento de la fosforilación oxidativa, que se ve reflejado en una mayor producción de ATP ya que LRPPRC no aumenta la cantidad de mitocondrias (biogénesis mitocondrial), pero sí en la estructuración de la matriz mitocondrial (una mayor densidad de crestas en la membrana interna)<sup>51</sup>. Para probar los efectos en la fosforilación oxidativa de LRPPRC *in vivo* se produjeron ratones transgénicos para LRPPRC y se evaluaron las vías metabólicas asociadas tanto a la fosforilación oxidativa, como a la oxidación de ácidos grasos y la cetogénesis. En ratones con deficiencia de LRPPRC se observó una disminución de  $\beta$ -hidroxibutirato (cuerpo cetónico más abundante producido en la cetogénesis). De manera inversa, en el ratón con sobre-expresión de LRPPRC, los niveles séricos de este cuerpo cetónico aumentaron; en ambos casos, los ácidos grasos libres séricos no mostraron cambios. La cetogénesis y la  $\beta$ -oxidación aumentaron en los hepatocitos aislados del ratón con un alto nivel de LRPPRC mientras que los niveles de mRNA nucleares de enzimas que controlan el equilibrio entre estos dos procesos no mostraron cambios; demostrando que los efectos de LRPPRC son independientes a la expresión de genes nucleares de la cetogénesis y la  $\beta$ -oxidación<sup>51</sup>. También se demostró el aumento de expresión de subunidades de la cadena respiratoria codificadas en el

mtDNA en células con una alta expresión de LRPPRC, por lo cual la proteína podría tener una función de co-activador de la transcripción mitocondrial<sup>51</sup>; proponiendo un efecto en el cual la presencia de la proteína estimula la respiración mitocondrial, mientras que su ausencia, la inhibe (Figura 3.5).

Para comprobar si el aumento en estas vías metabólicas de la cadena respiratoria (cetogénesis y  $\beta$ -oxidación) podrían estar influyendo los niveles séricos y hepáticos de TG y C se realizó un estudio con ratones transgénicos que tenían sobre-expresión de LRPPRC y ratones *knock-down* hígado-específicos<sup>59</sup>. Los resultados demostraron que con una dieta regular (*chow*) o con una dieta alta en grasas, el genotipo no influye en la ganancia de peso. Algunas diferencias en los ratones con modificación en la expresión de LRPPRC son: 1) el aumento del 17% en los niveles de ATP hepáticos en los ratones con sobre-expresión, 2) una disminución del 14% en niveles de ATP hepáticos en los ratones *knock-down* y 3) una disminución del 36.6% en el ritmo de respiración de la mitocondria en los ratones *knock-down*. En modelos murinos de sobre-expresión de LRPPRC existe un aumento de  $\beta$ -hidroxibutirato, así como un aumento en la expresión de las subunidades de la cadena respiratoria codificadas en la mitocondria. En contraste, el modelo murino con *knock-down* a LRPPRC se observó una disminución del 40% en la producción de CO<sub>2</sub> y una disminución del 19% en la captación de palmitato<sup>59</sup>, consecuencias esperadas, ya que hay una menor cantidad de subunidades de la cadena respiratoria y una menor tasa de fosforilación oxidativa.

En lo que respecta a lípidos séricos, en el mismo estudio, una dieta alta en grasas indujo una acumulación de TG y C en todos los genotipos estudiados. Sin embargo, en los ratones transgénicos se vio una acumulación menor en comparación con los ratones control, mientras que en los animales *knock-down* se presentaron niveles más altos (2.5 veces más a comparación con el ratón *wild type*)<sup>59</sup>. Otro efecto observado fue que LRPPRC provee protección contra lesiones de esteatosis en el hígado, ya que no hubo lesiones en tinciones de hígado de ratones transgénicos, pero sí en los ratones *knock-down*. No hubo cambios en niveles de ácidos grasos libres ni en genes que regulan el metabolismo de lípidos en ninguno de los dos tipos de ratones<sup>59</sup>, confirmando que la disminución de LRPPRC influye en los genes codificados en el mtDNA de la fosforilación oxidativa y no en los codificados en el DNA genómico<sup>51</sup>.

Akie y colaboradores, además de analizar la capacidad oxidativa de ácidos grasos en cultivos primarios de hepatocito que expresaban ectópicamente LRPPRC, evaluó si la fosforilación oxidativa exacerbada aumentaría la cantidad de NAD<sup>+</sup> (activador de las enzimas del Ciclo de Ácido Cítrico) y de esta forma promovía la oxidación completa de los ácidos grasos. Este estudio quería evaluar la hipótesis en la cual la Hígado Graso No Alcohólica (HGNA) es provocada por la resistencia a la insulina y la reducción de la proporción [NAD<sup>+</sup>]/[NADH] dependiente de la fosforilación oxidativa. Cuando se sobre-expresó LRPPRC en hígados de ratón, se observó todo lo contrario: con una dieta alta en grasa, hubo menos presencia de esteatosis en el hígado, menos inflamación, mayor sensibilidad a la insulina, disminución de especies reactivas del oxígeno (ROS), mayor liberación de CO<sub>2</sub> y una mayor cantidad de NAD<sup>+</sup> (últimas dos condiciones, indicativos de una mayor oxidación de palmitato); por lo que la activación de la fosforilación oxidativa

hepática influenciada por el aumento de la proteína de LRPPRC y aumento de NAD<sup>+</sup> promueve la oxidación completa de ácidos grasos y una protección contra HGNA<sup>52</sup>. (Figura 3.5).

Con los resultados anteriores, se puede inferir que el incremento en los niveles *LRPPRC*, resulta en el aumento en la expresión de genes codificados en la mitocondria relacionados con la fosforilación oxidativa. Esto lleva a un aumento en la captación y oxidación de ácidos grasos por los hepatocitos y a una reducción de TG y C tanto hepáticos como séricos. Por otro lado, si hay disminución de la expresión de *LRPPRC*, ocurre lo contrario; disminución en los niveles de mRNA de genes mitocondriales, disminución de actividad de fosforilación oxidativa, disminución de oxidación completa de ácidos grasos y la acumulación de lípidos en el hígado y su aumento en la circulación.

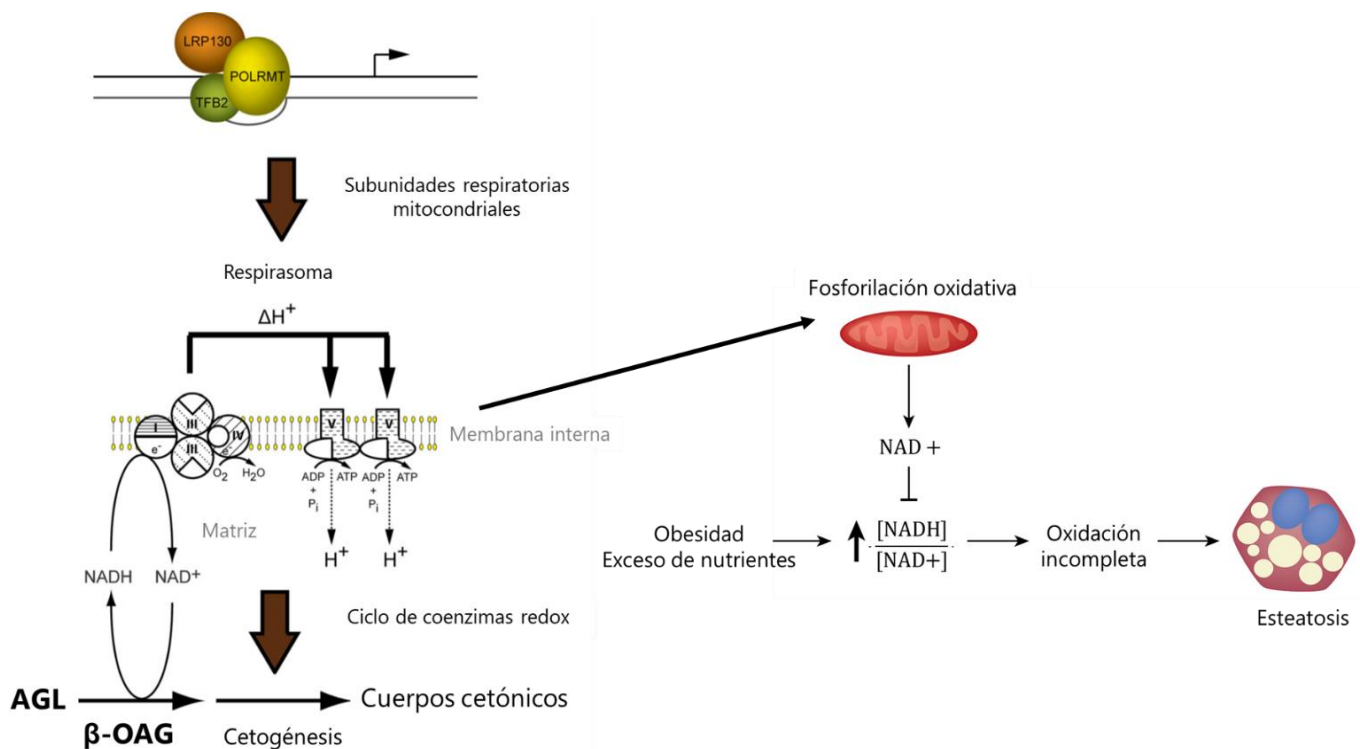


Figura 3.5. **Esquema del posible mecanismo de acción por el cual LRPPRC tiene influencia en la  $\beta$ -oxidación de ácidos grasos y esteatosis hepática.** LRPPRC (o LRP130) interactúa con POLRMT (RNA Polimerasa mitocondrial humana) para activar la transcripción mitocondrial; lo cual induce un aumento de la producción de péptidos codificada en el mtDNA y así promover la fosforilación oxidativa, oxidación hepática de ácidos grasos y cetogénesis; en el caso de exceso de nutrientes, hay un aumento de NADH reducido en la relación NADH/NAD<sup>+</sup> lo que inhibe la oxidación completa de ácidos grasos provocando esteatosis. Al tener una fosforilación oxidativa hepática aumentada, se aumenta el NAD<sup>+</sup> oxidado en la relación NAD<sup>+</sup>/NADH y así producir la oxidación completa de ácidos grasos. AGL: Ácidos grasos libres,  $\beta$ -OAG:  $\beta$ -oxidación de ácidos grasos. Modificado de Liu *et al*, 2011 y Akie *et al*, 2015.

## **Papel de *LRPPRC* en distintas patologías humanas**

### **Síndrome de Leigh**

Mutaciones presentes en *LRPPRC* están asociadas al síndrome de Leigh, particularmente a la variedad franco-canadiense<sup>60</sup>. El síndrome de Leigh es una condición neurodegenerativa muscular caracterizada por hipotonía, dimorfismo facial, y retraso en el desarrollo. En esta variedad de la enfermedad se detectaron niveles bajos de *COX* en el cerebro y el hígado pero no en tejidos renales ni musculares<sup>61</sup>. Los pacientes con esta condición, además de los síntomas neuronales, tienden a tener anormalidades en el metabolismo, reflejado en el aumento de lactato sérico y acidosis láctica, así como anormalidades en el sistema digestivo, como necrosis hepatocelular, aumento de lípidos intracelulares en los hepatocitos y esteatosis microvesicular hepática<sup>62,63</sup>. El síndrome de Leigh puede presentarse ante cualquier deficiencia en los complejos de la cadena de la fosforilación oxidativa, ya sea en genes nucleares o mitocondriales.

La integración de datos de DNA (mapeo en la región 2p16-21 con 15 genes candidatos), RNA (mRNA con patrones de expresión parecidos a genes mitocondriales) y proteína (proteómica de mitocondria aislada de la línea celular humana HepG2) se utilizaron para identificar el gen causal de esta variante del síndrome de Leigh<sup>60</sup>. Mediante la secuenciación de los exomas de *LRPPRC* en dos pacientes se encontró un cambio de base, en el exón 9, que resulta en un cambio en el aminoácido 354 (A354V) y al genotipar los pacientes restantes, 21 de 22 fueron homocigotos para la mutación, mientras el último fue heterocigoto en adición con una delección de 8 nucleótidos en el exón 35, resultando en un codón de terminación prematuro; la mutación no estaba presente en ninguno de los controles. Estas dos mutaciones proporcionan evidencia de la implicación de *LRPPRC* como la causante de la variante franco-canadiense del síndrome de Leigh<sup>60</sup>.

### **Enfermedad de Parkinson**

En otros estudios, se reportó asociación de *LRPPRC* con la enfermedad de Parkinson (EP). Esta es una condición neurodegenerativa caracterizada por temblor, rigidez y bradiquinesia (trastorno del movimiento provocado por un aumento de la rigidez muscular, causando lentitud motriz e incapacidad de realizar movimientos delicados). En 25% de los casos, existe más de un familiar afectado, por lo que hay una fuerte asociación genética.

De todos los mecanismos asociados a EP, la vía de PINK1/Parkin es de la más estudiada. En esta vía, existe fisión mitocondrial lo cual inicia la mitofagia por ubiquitinación de mitofusinas. En un estudio para estudiar más este mecanismo, se aislaron proteínas directamente asociadas con PINK1 y se identificaron 14 proteínas presentes en diferentes sitios celulares incluyendo la mitocondria, entre las que se encuentra *LRPPRC*. Como esta proteína está asociada a una patología neurodegenerativa (s. de Leigh), se secuenciaron los 38 exones y regiones intrónicas aledañas del gen de 46 pacientes con EP<sup>64</sup>. Se detectaron 4 variaciones sinónimas, 17 cambios intrónicos, 2 en la región 3'UTR y 1 en la

5'UTR, por lo que se concluyó que no hay una variante potencialmente patógena, pero no podrían descartarse variaciones en la dosis génica.

En un estudio posterior<sup>65</sup>, se estudiaron 7 de las 20 variantes no codificantes encontradas por Racovik y colaboradores. Con el ensayo de minigen se evidenciaron patrones aberrantes en el empalme de RNA (*splicing*) por la presencia de variantes en las secuencias regulatorias o zonas potenciadoras (*enhancer*) y silenciadoras (*silencer*) de este proceso. Se determinó que 4 de las variantes (rs35113761, rs3795859, rs76850904 y rs7593842) estudiadas influían en el patrón de empalme del RNA y que estos cambios eran tejido-específicos. Esto se demostró al tener diferentes patrones en el empalme del RNA y diferentes isoformas dependiendo si la variante estaba presente y de la línea celular humana (HeLa, HEK293T y SH-SY5Y) en la cual se expresó<sup>65</sup>. Se predijo bioinformáticamente que estos cambios y la formación de isoformas provocaría *skipping* de los exones aledaños a las variantes identificadas y posibles proteínas truncas; por lo que estas variantes podrían causar alteraciones en la estructura o la funcionalidad de la proteína.

## **Cáncer**

Varios estudios demuestran que cambios en la expresión de *LRPPRC*, sobretudo con sobre-expresión, están implicados en desarrollo de diferentes tipos de cáncer<sup>66</sup> (Tabla 3.4). Un ejemplo es en el cáncer de próstata, en el cual se ha observado que la elevación de niveles de *LRPPRC* está asociado positivamente con el grado del tumor, metástasis y niveles de antígeno prostático sérico, así como con un mal pronóstico.

En estudios con distintos modelos cancerígenos, se ha demostrado que la inhibición artificial de *LRPPRC* mediante siRNAs disminuyó el crecimiento de células gástricas cancerosas. Por su parte, en un modelo *knock-down* se disminuye la invasión, capacidad anti-apoptótica y la capacidad formadora de colonias *in vitro* en células de Linfoma de Hodgkin.

<b>Muestras de cáncer</b>	<b>Método</b>	<b>Expresión de LRPPRC</b>	<b>Ref.</b>
213 muestras de médula ósea proveniente de pacientes con leucemia mieloblástica crónica	PCR cuantitativa	El modelo con varios genes (incluyendo a <i>LRPPRC</i> ) está asociado a un aumento de riesgo a una recaída.	(67)
38 muestras de adenocarcinomas pulmonares	Inmunohistoquímica con acercamiento de Proteómica	Niveles de <i>LRPPRC</i> aumentados en tejidos cancerosos comparados a tejidos pareados sin signos de tumoración.	(68)
253 muestras de cáncer gástricos	Tinción inmunohistoquímica	Niveles de <i>LRPPRC</i> aumentados en tejidos cancerosos comparados a tejidos control pareados.	(69)
12 muestras de cáncer gástricos	Etiqueta de masa en tándem (TMT) y subsecuente espectrometría de masas	Niveles de <i>LRPPRC</i> aumentados en tejidos cancerosos comparados a tejidos sanos adyacentes.	(70)
112 muestras de cáncer de próstata	Tinción inmunohistoquímica	Niveles de <i>LRPPRC</i> aumentados en tejidos cancerosos comparados a tejidos HPB.	(71)
111 muestras de cáncer de próstata	Tinción inmunohistoquímica	Niveles de <i>LRPPRC</i> aumentados en tejidos cancerosos comparados a tejidos HPB.	(72)
198 muestras de cáncer de próstata	Tinción inmunohistoquímica	Niveles de <i>LRPPRC</i> aumentados en tejidos cancerosos comparados a tejidos HPB.	(73)
133 muestras de cáncer colorectal	Las etiquetas isobáricas para cuantificación relativa y absoluta y tinción inmunohistoquímica	Niveles de <i>LRPPRC</i> aumentados en tejidos cancerosos comparados a tejidos colorectales sanos.	(74)
20 muestras de carcinoma de células escamosas del epitelio de la lengua	Micro-ensayo de hibridación diferencial de metilación y secuenciación del genoma por bisulfito	Islas CpG asociadas a promotores de <i>LRPPRC</i> se encuentran hipermetiladas.	(75)
20 muestras de adenocarcinoma pulmonar, 9 de carcinoma de células escamosas del epitelio esofágico, 14 de cáncer gástrico, 7 de cáncer de colon, 11 de mamario, 13 de adenocarcinoma endometrial y 8 de linfoma difuso de células-B grandes	Tinción inmunohistoquímica	Tejidos cancerosos expresan <i>LRPPRC</i> de manera abundante, mientras que el tejido adyacente no-neoplásico tiene mínima expresión.	(76)

Tabla 3.3. **Implicaciones de la alta expresión de *LRPPRC* en diferentes tipos de cáncer.** HPB: Hiperplasia prostática benigna. Modificado de Cui et al, 2019.

## Patologías varias.

*LRPPRC* estaría implicado en otras condiciones que no están relacionadas directamente con condiciones neurológicas o cáncer<sup>66</sup>, por ejemplo, se ha documentado un riesgo incrementado de eventos cardiovasculares y tromboembolismo venoso. En un estudio de secuenciación de exoma completo de dos familias con Trombofilia, la variante rs372371276 (localizada en el intrón 13 de *LRPPRC*) se asoció al riesgo de padecer tromboembolia venosa<sup>77</sup>.

También se ha establecido su participación en infecciones virales. En dos estudios *knock-down* de *LRPPRC*, se demostró que hay una inhibición de la infección por virus de Hepatitis C y en segundo otro estudio se evidenció que se disminuye la replicación del virus ante una infección por HIV-1 en líneas celulares, pero no hay influencia en el ensamblaje de la cápside con el RNA<sup>78</sup>.

Adicionalmente, mediante un estudio de espectrometría de masas diferencial e inmunoprecipitación, se identificó la interacción de *LRPPRC* con el dominio de unión de la neurofibromina con la tubulina. La neurofibromina está codificada por el gen NF1 (Neurofibromina 1). Deleciones o mutaciones de pérdida de función en el gen NF1 es la causa de la neurofibromatosis tipo 1, el síndrome hereditario más común de predisposición para el desarrollo de tumores<sup>79</sup>.

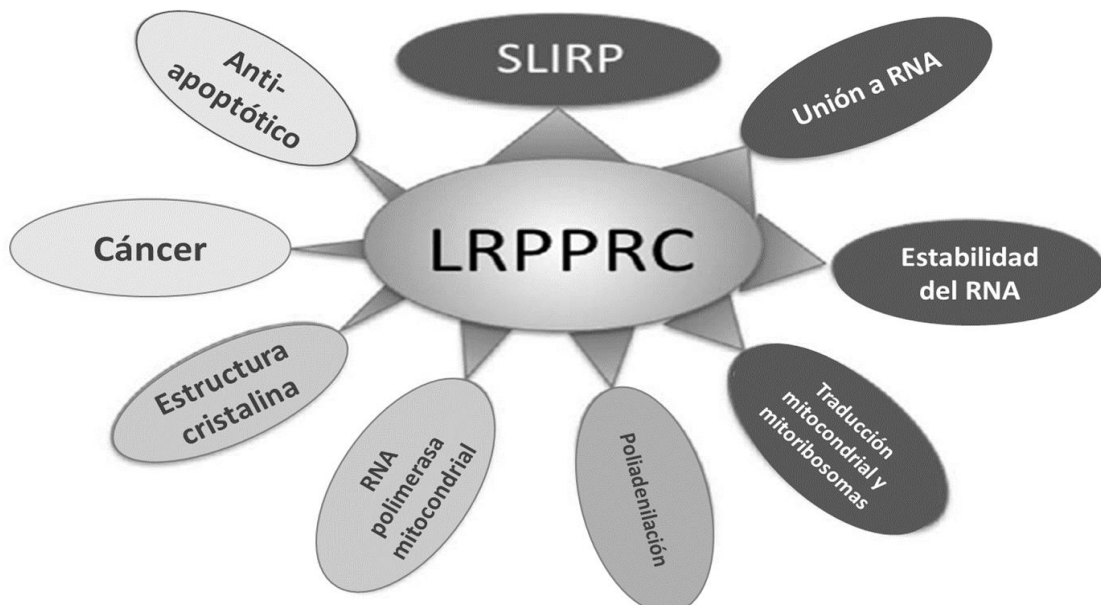


Figura 3.6. **Interacciones de *LRPPRC* con diferentes enfermedades y procesos biológicos.** Un mayor tamaño de la flecha y la intensidad del color del óvalo son indicadores de un mayor número de citas en la literatura que apoyan dichas relaciones con estas funciones. Modificado de Lightowlers *et al*, 2013.

Cabe señalar que la información acerca de las interacciones y funciones de *LRPPRC* no es completa (Figura 3.6), sin embargo, resulta evidente la importancia en el metabolismo y la estabilidad del RNA mitocondrial. También resulta interesante que la alteración en su expresión, ya sea al alta o a la baja, resulta en patología. Por lo anterior es importante estudiar el papel de *LRPPRC* en las diferentes patologías humanas, ya que eventualmente podría proponerse como un blanco terapéutico.



#### **4. PREGUNTA DE INVESTIGACIÓN**

¿Las variantes del haplotipo del gen *LRPPRC* asociadas a la hipertrigliceridemia severa familiar son variantes funcionales que pudieran contribuir a la elevación de triglicéridos séricos en humanos o bien existen otras variantes funcionales en desequilibrio de ligamiento con las inicialmente asociadas dentro del mismo *locus*?

#### **5. HIPÓTESIS**

Las variantes en el gen de *LRPPRC* asociadas a la hipertrigliceridemia familiar afectarán la función de *LRPPRC* a través de distintos mecanismos: a) modificando la estructura primaria de la proteína, b) alterando el empalme del mRNA c) alterando la expresión del gen, o d) alterando la regulación epigenética del *locus*. La alteración en *LRPPRC*, afectará de manera secundaria los niveles de mRNA mensajeros mitocondriales que codifican para las enzimas involucradas en la fosforilación oxidativa y la  $\beta$ -oxidación.

#### **6. OBJETIVO GENERAL**

Determinar si las variantes asociadas a la hipertrigliceridemia severa familiar alteran la expresión del gen *LRPPRC*; así como identificar posibles variantes adicionales en el *locus*. De manera conjunta, identificar alteraciones en la expresión de genes mitocondriales, relacionados a la disfunción de *LRPPRC* para contribuir a la elevación de triglicéridos séricos en humanos.

#### **OBJETIVOS PARTICULARES**

- Secuenciar los exones de *LRPPRC* para identificar posibles variantes en la región codificante que potencialmente afecten su función.
- Analizar la expresión de *LRPPRC* y genes mitocondriales a partir de biopsias hepáticas de pacientes portadores y no portadores de las variantes asociadas mediante cuantificación de mRNA.
- Búsqueda de posibles diferencias en los patrones de empalme del RNA a partir de biopsias hepáticas de pacientes portadores y no portadores de las variantes asociadas.
- Búsqueda de información bioinformática de posibles interacciones a largo alcance y de organización de la cromatina en el *locus* de *LRPPRC* a partir de datos públicos de Hi-C y marcas epigenéticas de la cromatina en tejido hepático.
- Búsqueda bioinformática de posibles alteraciones en la unión de factores de transcripción en el *locus* de las variantes que potencialmente afecten su unión.

## **7. MATERIALES Y MÉTODOS**

### **Participantes del estudio y muestras**

Para la identificación de posibles variantes en la región codificante de la proteína, analizamos muestras de sangre periférica

#### Población de estudio

Casos de HTG y controles normolipidémicos portadores y no portadores del haplotipo de riesgo no relacionados provenientes del estudio de asociación genómica mencionado en los antecedentes (Datos aun no publicados). (Tabla 7.1)

#### Criterios de inclusión

Hombres y mujeres mayores de 18 años no relacionados, casos de HTG (TG mayores a 1000mg/dL al menos una vez, con TG en la percentila 90 poblacional, con niveles de apo B debajo de la percentila 90 poblacional y con estudio familiar confirmatorio) y controles normolipidémicos (TG menores a 150mg/dL, HDL-C entre 40 mg/dL-60mg/dL y LDL-C menores a 190mg/dL sin uso de fármacos hipolipemiantes).

#### Criterios de exclusión

Relación familiar entre pacientes, otras causas primarias de HTG, embarazo, falla renal y/o hepática graves, consumo de alcohol mayor a 10 bebidas por semana, infección por VIH, cualquier enfermedad aguda 6 semanas previas a la toma de muestra e hipotiroidismo sin tratamiento.

Paciente	Sexo	Edad	Haplotipo de riesgo	HTG
1	M	54	Portador Homocigoto	Sí
2	M	35	Portador Homocigoto	No
3	F	37	Portador Homocigoto	No
4	M	40	Portador Homocigoto	No
5	H	82	No (WT)	No
6	M	45	No (WT)	No
7	H	56	No (WT)	No
8	M	48	No (WT)	No
9	M	45	No (WT)	No
10	H	52	No (WT)	No
11	H	65	No (WT)	No
12	M	64	No (WT)	No
13	M	49	No (WT)	No
14	M	55	No (WT)	No
15	M	50	No (WT)	No
16	H	45	No (WT)	No
17	H	71	No (WT)	No
18	M	57	No (WT)	No
19	H	54	No (WT)	No
20	H	47	No (WT)	No
21	M	48	No (WT)	No
22	H	47	No (WT)	No
23	H	64	No (WT)	No

Tabla 7.1. **Información de los pacientes cuyas muestras de sangre fueron utilizadas para la secuenciación de exones de LRPPRC.** WT: *Wild type* (no portador).

Para la búsqueda de posibles diferencias en los patrones del empalme del RNA y estudios de expresión se utilizaron biopsias hepáticas

#### Población de estudio

Pacientes portadores y no portadores de las variantes de riesgo sometidos a un *bypass* gástrico que justificaba la toma de biopsias de tejido hepático con obesidad mórbida pareados por edad ( $\pm 5$  años), sexo e IMC ( $\pm 5$  kg/m<sup>2</sup>) (Tabla 7.2).

<b>Muestra</b>	<b>Sexo</b>	<b>Edad</b>	<b>IMC</b>	<b>Haplotipo de riesgo</b>
<b>1</b>	M	38	50.5	Portador homocigoto
<b>2</b>	H	28	42.62	Portador homocigoto
<b>3</b>	H	31	40.16	Portador heterocigoto
<b>4</b>	H	43	44.44	No (WT)
<b>5</b>	M	41	51.25	No (WT)
<b>6</b>	M	38	51.00	No (WT)
<b>7</b>	H	33	42.61	No (WT)
<b>8</b>	H	47	42.78	No (WT)
<b>9</b>	M	53	52.12	No (WT)
<b>10</b>	H	58	41.68	No (WT)
<b>11</b>	H	37	41.70	No (WT)
<b>12</b>	M	46	49.55	No (WT)

Tabla 7.2. **Información de los pacientes cuyas muestras de tejido hepático para estudios de expresión y para búsqueda de isoformas de LRPPRC.** WT: *Wild type* (no portador).

### Criterios de inclusión

Hombres y mujeres mayores 18 años no relacionados portadores y no portadores de las variantes de riesgo provenientes de un protocolo de obesidad mórbida.

### Criterios de exclusión

Relación familiar entre pacientes, embarazo, falla renal y/o hepática graves, consumo de alcohol mayor a 10 bebidas por semana, infección por VIH, cualquier enfermedad aguda 6 semanas previas a la toma de muestra e hipotiroidismo sin tratamiento.

## **Secuenciación Sanger**

### **Diseño de oligonucleótidos**

El diseño de los oligonucleótidos para la reacción de PCR de amplificación y la reacción de PCR de secuenciación se realizó con el programa *OLIGO Primer Analysis Software Version 7 (Molecular Biology Insights, Inc)*. La secuencia de LRPPRC se obtuvo del portal *Ensembl* ([https://www.ensembl.org/Homo\\_sapiens/Transcript/Summary?db=core;g=ENSG00000138095;r=2:43886224-43996226;t=ENST00000260665](https://www.ensembl.org/Homo_sapiens/Transcript/Summary?db=core;g=ENSG00000138095;r=2:43886224-43996226;t=ENST00000260665)) y los oligonucleótidos se diseñaron para que hibridaran en las regiones intrónicas adyacentes a los exones de LRPPRC; para la región de interacción, regiones sobrelapadas. Las secuencias de los oligonucleótidos se encuentran en la Tabla suplementaria 2.1.

### **PCR de amplificación**

Las reacciones de PCR de amplificación se realizaron a partir de DNA genómico extraído de sangre periférica mediante el kit de extracción *QIAamp 96 DNA Blood Kit (Qiagen)* conforme al protocolo del fabricante. Para la amplificación de las regiones exónicas se

utilizó la enzima *HotStart Taq* Polimerasa (*Quiagen*) y para la región de interacción del intrón 23, la mezcla de polimerasas *LongRange PCR* (*Quiagen*).

### **PCR de secuenciación**

Después de la purificación del producto de PCR mediante el uso de las columnas *QIAquick Spin Columns* (*Qiagen*) y subsecuente electroforesis en gel de agarosa al 2%, se realizó la reacción de secuenciación mediante la técnica de Sanger con el terminador *BigDye* (*Applied Biosystems*). La reacción de secuencia fue purificada por medio de columnas con *Sephadex*, para después deshidratarla y enviar a procesamiento a la Unidad de Apoyo a la Investigación (RAI, INCMNSZ) mediante electroforesis capilar. Se recibía un archivo con el electroferograma de la secuencia. El análisis de la secuencia incluía revisión de la calidad del electroferograma, alineamiento con secuencia de referencia con la herramienta bioinformática de alineamiento *BLASTn* (*Basic Local Alignment Search Tool*, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) y búsqueda de señales heterocigotas.

### **Portales bioinformáticos para recolección de información**

Los portales de predicción de posibles afectaciones por presencia de variantes utilizados fueron: *Mutation Taster* (<http://www.mutationtaster.org/>), *Human Splicing Finder* (<http://www.umd.be/HSF3/HSF.shtml>) y *Spliceman Web Server* (<http://fairbrother.biomed.brown.edu/spliceman/index.cgi>).

Los portales para las interacciones a larga distancia y marcas en la cromatina fueron: *3DIV A 3D-genome Interaction Viewer and database* (<http://kobic.kr/3div/hic>) y *Cistrome Data Browser* (<http://cistrome.org/db/#/>)

Para los alineamientos de motivos de factores de transcripción se ocuparon: *HaploReg v4.1* (<https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>), *Jaspar 2020* (<http://jaspar.genereg.net/>), *LASAGNA-Search 2.0: Searching for transcription factor binding sites (TFBSs)* ([https://biogrid-lasagna.engr.uconn.edu/lasagna\\_search/index.php](https://biogrid-lasagna.engr.uconn.edu/lasagna_search/index.php)) y *Toolkit for Cistrome Data Browser* (<http://dbtoolkit.cistrome.org/>).

### **Cultivo celular**

Utilizamos la línea celular HepG2 como control positivo de la expresión de *LRPPRC*. Esta línea celular se creció en medio DMEM suplementado con Suero Fetal Bovino (SFB) al 10% y mezcla de antibióticos y antimicótico 100x al 0.5% (10'000 unidades/ml de penicilina, 10'000 µg/ml de estreptomina y 25 µg/ml de antimicótico *Fungizone*) en botellas de cultivo de 75 cm<sup>2</sup>. Los cultivos se mantuvieron en incubación a una temperatura de 37° C, atmósfera con CO<sub>2</sub> al 5% y humedad relativa del 100% hasta llegar a una confluencia aproximada de 70% (aproximadamente 7x10<sup>6</sup> células; 90 mil células/cm<sup>2</sup>).

## **Extracción de ácidos nucleicos**

A partir del cultivo de las células HepG2, se hizo una extracción simultánea de DNA genómico y RNA total mediante el *AllPrep DNA/RNA Mini Kit (Qiagen)* para extraer de manera simultánea DNA genómico y RNA total de las células. Se realizó el protocolo de manera establecida por el fabricante. Para las muestras de tejido hepático sólo se extrajo RNA total con el reactivo de lisis *QIAzol*.

## **Síntesis de cDNA**

Se utilizaron 3 métodos para la síntesis de cDNA a partir de 1 µg de RNA total extraído de tejido hepático y líneas celulares.

En el primero utilizamos el *Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Scientific)* que contiene oligonucleótidos aleatorios de secuencias desconocidas.

Para el segundo y tercer método utilizamos el *RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific)* en el que se realizó una síntesis de cDNA con un oligonucleótido con varias repeticiones del nucleótido T (dT)18 y otra síntesis con oligonucleótidos diseñados a unirse en una región determinada del mRNA de *LRPPRC*.

## **PCR cuantitativa**

Se utilizaron 2 métodos para la PCR cuantitativa a partir de cDNA proveniente de tejido hepático y de la línea celular HepG2.

Para el estudio de expresión de *LRPPRC* en cDNA de portadores y no portadores de las variantes de riesgo, se usó el ensayo *TaqMan* específico para *LRPPRC (TaqMan® Gene Expression Assays, Assay ID: Hs00939019\_m1)*, oligonucleótidos específicos para el gen de referencia *GUSB* y el reactivo *TaqMan Universal PCR Master Mix (Thermo Scientific)*. Las condiciones se realizaron por duplicado.

Para la evaluación de la expresión génica de genes mitocondriales, se utilizó el sistema *SYBR GREEN*, con oligonucleótidos específicos reportados en la literatura para cada gen. Los genes blanco fueron *ND1, ND6, CYTB, COX1, ATP6, ATP8* y los genes de referencia *h18s, β-actina* y *GAPDH*. Las condiciones se realizaron por triplicado.

Utilizamos el equipo *StepOne Plus (Applied Biosystems)*.

## **Análisis estadístico**

Para los análisis de expresión de genes mediante qPCR cuantitativa, se utilizó la prueba no paramétrica U de Mann Whitney para comparar diferencias de expresión entre portadores y no portadores (valor p menor a 0.05 es estadísticamente significativo).

## 8. RESULTADOS

### Búsqueda de nuevas variantes en los exones del gen *LRPPRC*

A partir de la identificación de las 3 variantes de riesgo (rs7593842, rs58811869 y rs72877186) en pacientes con HTGSF mediante un estudio de secuenciación de exoma completo, fue necesario averiguar si existen otras variantes que pudieran relacionarse funcionalmente con la HTG severa. Por lo tanto, el primer paso consistió en una búsqueda más específica y profunda en el *locus* de *LRPPRC* mediante la secuenciación de los exones del gen de *LRPPRC* para identificar posibles cambios de secuencia que pudieran causar un alto prematuro, un cambio en el marco de lectura o bien mutaciones puntuales tipo "missense" y así afectar la función de la proteína.

La muestra de estudio fue sangre periférica proveniente de pacientes portadores no relacionados de las variantes de riesgo que surgieron del estudio de descubrimiento y asociación de *LRPPRC* y mediante secuenciación por el método de Sanger, se obtuvieron las secuencias de regiones intrónicas adyacentes a los exones, la unión exón-intrón y la región exónica.

De estas secuenciaciones se identificaron 17 variantes (Tabla 8.1): 2 variantes exónicas y 15 variantes intrónicas. De las **17** variantes, **8** son homocigotas compartidas por los 4 portadores, **2** son compartidas por los 4 portadores (portadores 1, 2 y 3 son homocigotos y en el portador 4 es heterocigota), **2** son homocigotas compartidas por 3 portadores (portador 1, 2 y 3), **1** es heterocigota compartida por 2 portadores (portador 2 y 3), **1** es heterocigota compartida por 3 portadores (portador 2, 3 y 4) y **3** están presentes en 1 sólo portador (portadores 1 y 3).

De las 10 variantes compartidas (3 portadores homocigotos y 1 heterocigoto), una de las variantes es exónica (rs6741066). Esta se trata de una variante silenciosa/sinónima, ya que hay cambio en un nucleótido, pero no en el aminoácido (p.Q82Q). Las 9 variantes restantes son intrónicas, por lo que no habría una afectación directa a la secuencia codificante de la proteína; aunque pudieran influir en el empalme alternativo para alguna isoforma de *LRPPRC*.

De estos resultados se puede inferir que existen más variantes homocigotas compartidas en los portadores de las variantes originalmente asociadas pero que no afectan la estructura primaria de la proteína.

<b>Posición génica</b>	<b>Variación en el DNA</b>	<b>No. en NCBI</b>	<b>Presencia en portadores</b>	<b>Prevalencia en la población (AMR/MXL)</b>	<b>Base de datos</b>
<b>Exón 2</b>	c.246G>A (p.Q82Q)	rs6741066	4 portadores 3 homocigotos y 1 heterocigoto	0.445/0.461	1000 Genomes Fase 3
<b>Intrón 3</b>	IVS3-132C>G	rs6721144	4 portadores homocigotos	0.154/0.070	1000 Genomes Fase 3
<b>Intrón 6</b>	IVS6-70T>C	rs17031786	4 portadores homocigotos	0.156/0.070	1000 Genomes Fase 3
<b>Intrón 9</b>	IVS9+30A>G	*rs7593842	4 portadores homocigotos	0.154/0.070	1000 Genomes Fase 3
<b>Intrón 12</b>	IVS13-56T>G	rs10190161	3 portadores homocigotos	0.569/0.570	1000 Genomes Fase 3
<b>Intrón 15</b>	IVS15+11C>G	*rs58811869	4 portadores 3 homocigotos y 1 heterocigoto	0.144/0.078	1000 Genomes Fase 3
	IVS15+114	-	2 portadores heterocigotos	-	-
<b>Intrón 17</b>	IVS18-28T>G	*rs72877186	4 portadores homocigotos	0.156/0.078	1000 Genomes Fase 3
<b>Intrón 19</b>	IVS20-54C>A	rs79868165	1 portador heterocigoto	0.003/0.016	1000 Genomes Fase 3
<b>Intrón 20</b>	IVS20-40A>C	rs7594526	4 portadores homocigotos	0.411/0.477	1000 Genomes Fase 3
<b>Intrón 22</b>	IVS22+27T>G	rs28394191	4 portadores homocigotos	0.585/0.523	1000 Genomes Fase 3
	IVS22+41C>T	rs368657323	1 portador heterocigoto	5.812e <sup>-05</sup> /-	gnomAD exomes
<b>Intrón 28</b>	IVS28+21C>A	rs7568481	4 portadores homocigotos	0.592/0.555	1000 Genomes Fase 3
<b>Intrón 29</b>	IVS30-289C>T	rs4952691	3 portadores homocigotos	0.644/0.586	1000 Genomes Fase 3
<b>Exón 30</b>	c.3235T>A (p.S1079T)	rs769476622	1 portador heterocigoto	5.784e <sup>-05</sup> /-	gnomAD exomes
<b>Intrón 30</b>	IVS31-264C>T	-	3 portadores heterocigotos	-	-



<b>Intrón 37</b>	IVS37+37G>A	rs2955280	4 portadores homocigotos	0.640/0.578	1000 Genomes Fase 3
------------------	-------------	-----------	-----------------------------	-------------	---------------------------

Tabla 8.1. **Variantes identificadas en los exones y unión exón-intrón de LRPPRC en los 4 portadores homocigotos de las variantes de riesgo.** AMR: población americana; MXL: población con ancestría mexicana en Los Ángeles, California, EUA. /: Comparación entre las dos prevalencias; \*: Polimorfismos pertenecientes al haplotipo de riesgo para HTGSF.

Las variantes identificadas compartidas de manera homocigota por los 4 portadores podrían ser parte del haplotipo de riesgo, pero 9 de estas variantes identificadas son intrónicas, por lo que la posible implicación funcional no resulta obvia. Por esta razón se realizó una búsqueda bioinformática en portales de predicción de las posibles consecuencias de la presencia de estas variantes de secuencia.

Los portales de predicción utilizados fueron: *Mutation Taster*<sup>80</sup>, *Human Splicing Finder*<sup>81</sup> y *Spliceman Web Server*<sup>82</sup>. En estos sitios de predicción, se determina un valor umbral (*score*). Si la variante afecta las características de la proteína, transcritos o el proceso de empalme del RNA, el valor umbral de los programas es sobrepasado y se considera la variante como una variante con probable consecuencia funcional.

En la Tabla 8.2 se resume los resultados encontrados y tomándolos de manera conjunta, se estableció un criterio de inclusión para considerar la afectación como probable, el cual fue: 1) la cercanía de la variante con la unión exón-intrón, 2) que en el portal *Mutation Taster* se predijeran cambios importantes en la proteína o que hubiera omisión de exones, 3) varios cambios y cambios con *score* alto en el portal *Human Splicing Finder*, 4) así como un porcentaje alto en la predicción hecha en *Spliceman Web Server*. Tomando en cuenta estos criterios, se seleccionaron 6 variantes para experimentos funcionales posteriores: rs6721144 (intrón 3), rs7593842 (Intrón 9), \*rs58811869 (Intrón 15), rs7594526 (Intrón 20), rs28394191 (Intrón 22), rs7568481 (Intrón 28). A partir de ahora en este trabajo, las variantes se identificarán dependiendo del intrón donde se encuentren; V3, V9, V15, V20, V22 y V28, respectivamente. La variante de riesgo (rs72877186) no fue incluida ya que se predecían cambios sutiles por la presencia de esta variante.

Este análisis *in-silico* reveló que una posible afectación de las variantes intrónicas sería en el proceso de empalme del RNA en regiones importantes para el procesamiento de los intrones y posiblemente tener cambios a nivel de proteína.

<b>Variante</b>	<b>Distancia del sitio de unión exón/intrón (pb)</b>	<b>Mutation taster</b>	<b>Human Splicing Finder</b>	<b>Spliceman Web Server<sup>+</sup></b>
rs6721144	132	Ganancia de sitio donador de empalme del RNA.  Omisión de exones y características de la proteína afectada	Nuevo sitio de empalme del RNA e interrupción de otro  Roto el motivo del <i>branching</i>  Roto sitio de <i>enhancer</i> y nuevo sitio <i>silencer</i>	54%
rs17031786	70	Sitios aceptor y donador incrementados  Ganancia de sitio donador  Omisión de exones y características de la proteína afectada	Nuevo sitio <i>enhancer</i> en el alelo mutado	96%
*rs7593842	30	4 sitios aceptores sutilmente incrementados	Nuevo sitio de empalme del RNA	59%
rs10190161	56	Sin cambios en sitios de empalme del RNA	Roto el motivo del <i>branching</i>	71%
*rs58811869	11	Sin cambios en sitios de empalme del RNA	Nuevo sitio de empalme del RNA  Roto el motivo del <i>branching</i>  Nuevo sitio <i>enhancer</i> y roto un sitio <i>silencer</i>	61%
*rs72877186	28	Sin cambios en sitios de empalme del RNA	Nuevo sitio <i>silencer</i> y roto otro sitio <i>silencer</i>	86%
rs7594526	92	Sitio aceptor incrementado  Omisión de exones y características de la proteína afectada	Cambios en el sitio de <i>branching</i>  Nuevo sitio <i>enhancer</i> y roto otro sitio <i>enhancer</i>  Roto sitio <i>silencer</i>	75%
rs28394191	27	Sitio donador ligeramente incrementado	Nuevo sitio de empalme del RNA  Nuevo sitio <i>silencer</i>	74%
rs7568481	21	Un sitio donador incrementado y ganado otro sitio donador  Omisión de exones y características de la proteína afectada	Crea un nuevo sitio de empalme del RNA	62%
rs2955280	37	Sin cambios en sitio empalme del RNA	Nuevo sitio <i>silencer</i> , pero score bajo	80%

Tabla 8.2. **Recopilación de información para las variantes compartidas de entre los 4 pacientes portadores.** \*: Variantes pertenecientes al haplotipo de riesgo. SWS<sup>+</sup>: entre más alto el porcentaje, más probable es que la variante afecte el proceso de empalme del RNA.

Al seleccionar las 6 variantes que pudieran tener un impacto funcional, el siguiente paso fue determinar si pudieran pertenecer al haplotipo de riesgo; se esperaría entonces que las variantes no estuvieran presentes en no portadores, porque su ausencia confirmaría que las variantes son pertenecientes al haplotipo de riesgo.

Para este fin se seleccionaron participantes control del estudio de descubrimiento de las variantes de riesgo sanos (sin HTG) no relacionados y que fueran no portadores del haplotipo de riesgo, para la búsqueda de las variantes seleccionadas mediante secuenciación Sanger. Se utilizaron las mismas condiciones experimentales que las empleadas en los portadores homocigotos del haplotipo de riesgo.

De esta secuenciación se determinó que las variantes V22 y V28 se encuentran presentes en 9 muestras de los pacientes control de manera homocigota, en 7 de manera heterocigota y en 3 muestras no se encontraron. En relación a la variante 20, se identificaron 10 muestras que presentaban el polimorfismo de manera homocigota, 6 muestras de manera heterocigota y no se encontró en 3 muestras. La variante V3 no se encontró en las 4 muestras que se procesaron de las 19 muestras de estudio. (Tabla 8.3).

Las variantes que no se encontraron en los no portadores son la V9 y V15 (excepto una muestra, ya que el cambio para la variante V15 es heterocigoto). Teniendo en cuenta que se confirmó su ausencia en muestras de no portadores, se decidió continuar con más estudios funcionales con las variantes V9 y V15; 2 de las 3 variantes del haplotipo de riesgo.

Estos resultados muestran que la variante V9 es exclusiva de los portadores del haplotipo y que la frecuencia de la variante V15 es muy baja en no portadores.

<b>Variante</b>	<b>Presencia en no portadores a las variantes de riesgo Homocigotos/Heterocigotos/Ausencia</b>
<b>V3</b>	0/0/4 (4 muestras procesadas)
<b>*V9</b>	0/0/19
<b>*V15</b>	0/1/18
<b>V20</b>	10/6/3
<b>V22</b>	9/7/3
<b>V28</b>	9/7/3

Tabla 8.3. **Presencia de las 6 variantes seleccionadas en 19 muestras de no portadores sin HTG.** \*: Variantes pertenecientes al haplotipo de riesgo.

De acuerdo a lo reportado en la literatura, la presencia de la variante V9 tiene influencia en el patrón de empalme del RNA, originando el *skippin* del exón 9, que podría dar lugar a una proteína trunca<sup>45</sup> (Figura Suplementaria 1.2). En el caso de la variante V15, se demostró que su presencia tiene influencia en la unión de factores de transcripción<sup>37</sup> (Figura suplementaria 1.3), además de ser la variante más cercana a la unión exón-intrón, por lo que es posible también que afectara el empalme del RNA.

Por lo que como siguiente paso se buscó la posible presencia de transcritos alternativos derivados de las variantes intrónicas identificadas. Las muestras de tejido hepático provienen de biopsias de pacientes provenientes de un estudio de Obesidad Mórbida en las que se aisló RNA total (Tabla 8.4). Estas muestras de RNA también fueron utilizadas para experimentos de nivel de expresión posteriores (Sección 8. Estudios de niveles de expresión de genes en tejido hepático y línea celular humana HepG2).

Por la importancia del tejido hepático en el metabolismo de lípidos, se utilizaron muestras de RNA de la línea celular de hepatocarcinoma HepG2 como control positivo ya que en la literatura se reporta que expresa *LRPPRC* de manera endógena<sup>83</sup>.

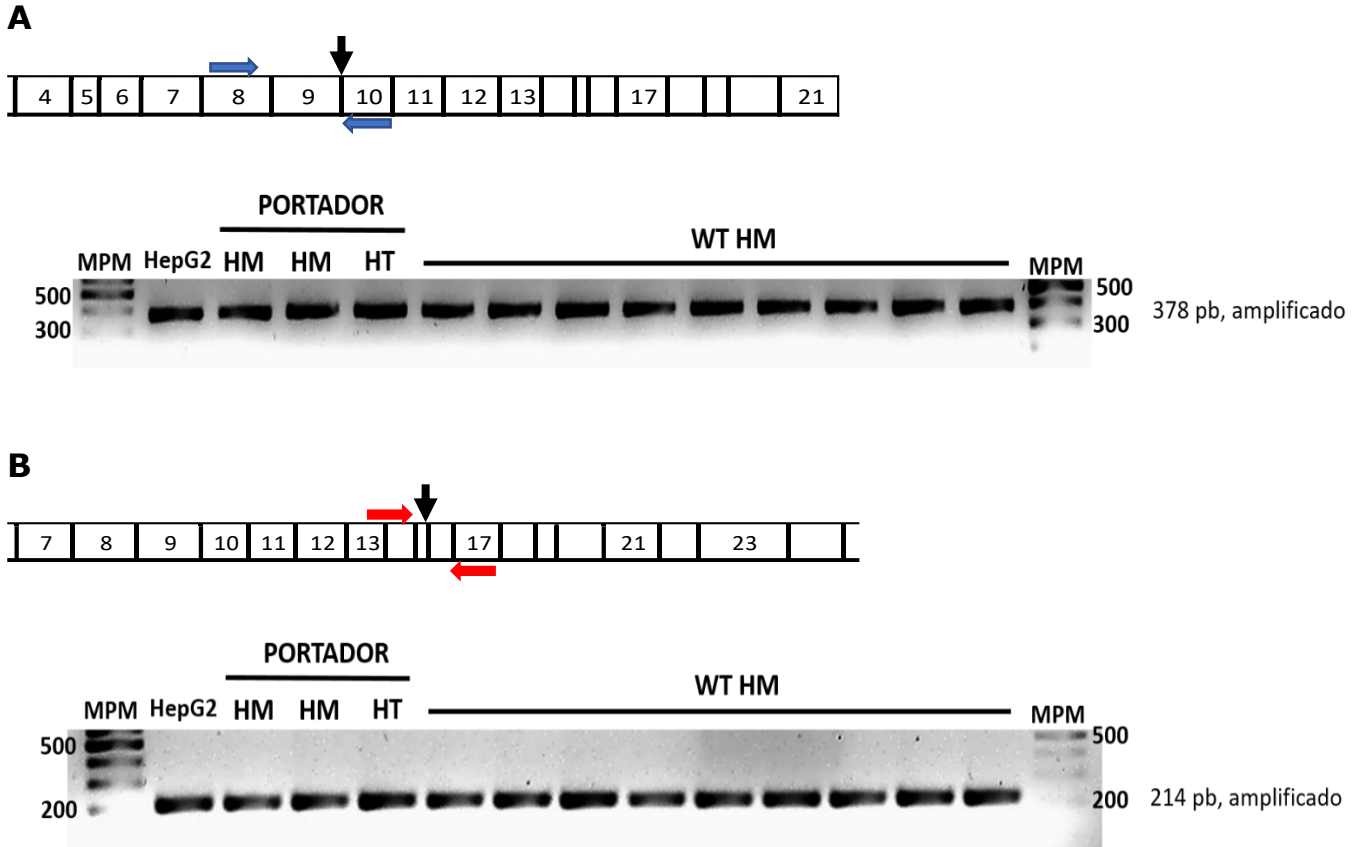
Variante	Portador	WT	Valor p
<b>Edad</b>	31.00 ± 5.13	43.00 ± 7.95	0.05179
<b>Sexo H/M (n)</b>	2/1	5/4	1
<b>IMC</b>	42.62 ± 5.38	44.43 ± 4.51	0.4818

Tabla 8.4. Pareo de los portadores y no portadores de las variantes de riesgo para muestras de tejido hepático.

### **Búsqueda de isoformas de *LRPPRC***

Procedimos al diseño de oligonucleótidos para la amplificación selectiva de estos posibles transcritos alternativos a partir de cDNA proveniente de RNA de tejido hepático de portadores y no portadores, así como de RNA de la línea celular HepG2. Pensando en las probables consecuencias de un defecto en el empalme del RNA, se diseñaron oligonucleótidos en los exones cercanos a las variantes a estudiar: exón anterior a la variante y un exón posterior (paneles superiores en Figura 8.1A y 8.1B).

Se observó que el peso de los amplicones era igual al esperado en las dos condiciones de búsqueda y en todas las muestras utilizadas, por lo que los resultados de estas amplificaciones no evidencian la presencia de isoformas derivadas de un defecto en el empalme del RNA de *LRPPRC* en los portadores y no portadores, tanto para la variante V9 como en la V15 con las condiciones utilizadas (Paneles inferiores en Figura 8.1A y 8.1B).



Considerando la posibilidad de los transcritos pequeños fueran inestables y escaparan su detección, decidimos ampliar la región de amplificación río arriba (Exón 6-Exón 8), río abajo (Exón 8-Exón 17) respecto a la variante V9 y ambas (Exón 6-Exón 17), conteniendo a las dos variantes (paneles superiores en Figura 8.2A, 8.2B y 8.2C). Además, utilizamos 3 métodos distintos para la síntesis de cDNA para encontrar la condición en la cual pudiera estar más enriquecidos transcritos alternativos: oligonucleótidos aleatorios, oligonucleótido (dT)18 (18 Timinas para hibridar con la cola de poli-A en mRNA maduros) y oligonucleótidos específicos para el mRNA de *LRPPRC*. Las muestras iniciales para evidenciar la presencia de un posible amplificado alternativo fueron cDNA de la línea celular HepG2, un portador heterocigoto (HT) y un individuo no portador (WT); ya que la cantidad de mRNA hepático era limitado.

En el panel inferior de la Figura 8.2A se muestra la amplificación de cDNA sintetizado con oligonucleótidos aleatorios y la presencia de 3 amplificados de diferente tamaño en las 3 muestras utilizadas. El amplificado de aproximadamente 320 pb (remarcado en la Figura 8.2A) se purificó y secuenció, para comprobar su identidad. Se identificó la secuencia como un fragmento del exón 8 (89 pb), el exón 9 completo (146 pb) y el inicio del exón 10 (26 pb). En los amplificados Exón 6-Exón 10 y Exón 6-Exón 17, no hay evidencia de una isoforma alternativa (Figura suplementaria 1.4)

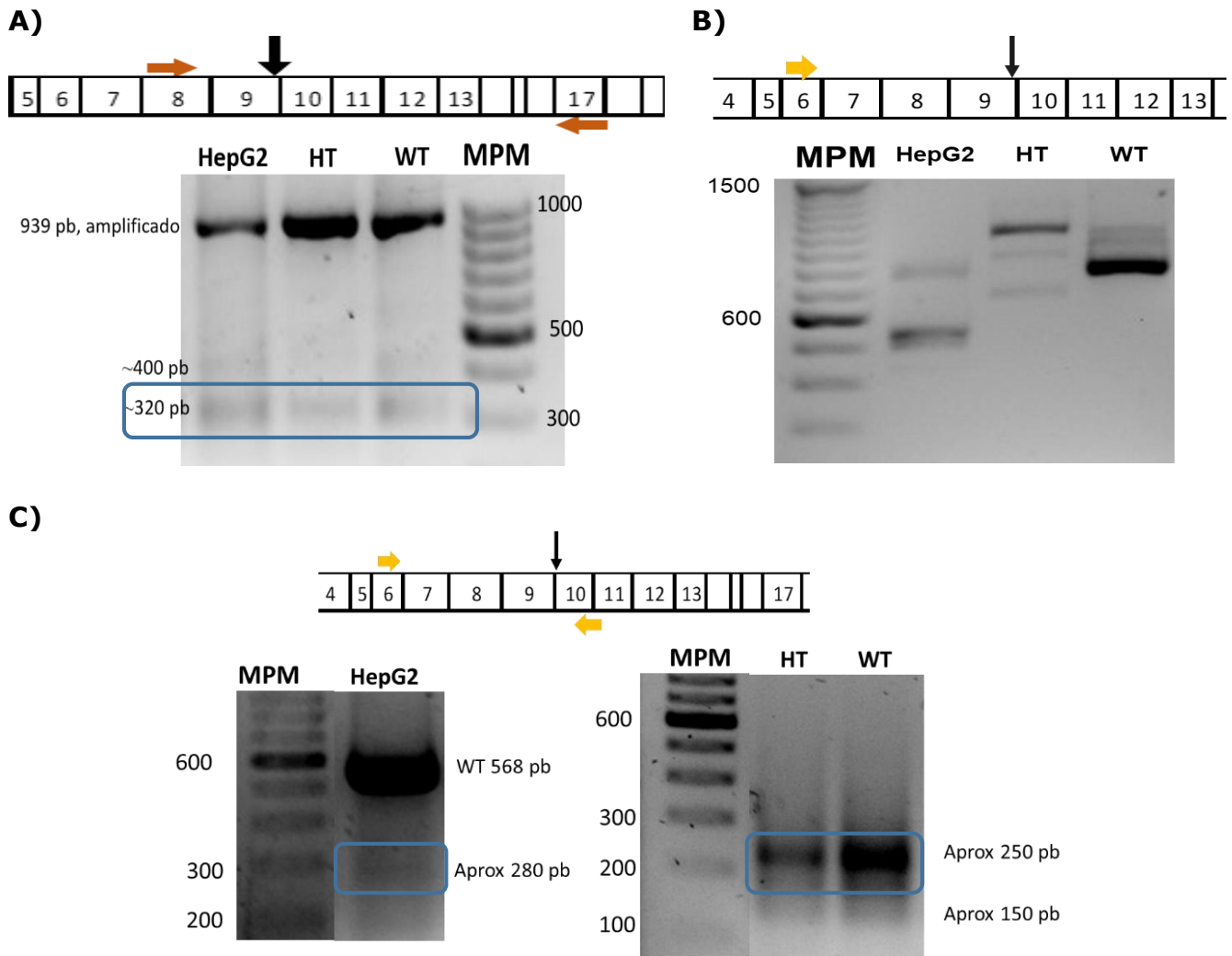
La siguiente condición de síntesis de cDNA fue con un oligonucleótido (dT)18 (Panel inferior en Figura 8.2B). La reacción de amplificación se realizó con el oligonucleótido sentido que se une al exón 6 (la región más río arriba) y el oligonucleótido (dT)18 (Panel superior en 8.2B) con las 3 muestras de cDNA anteriores (HepG2, HT y WT). Se obtuvieron tres amplificados en cada muestra, de diferentes tamaños y ninguno compartido por las 3 muestras. En HepG2 se obtuvieron 2 amplificados cercanos a 550 pb y uno de aproximadamente 800 pb; con la muestra HT se obtuvo 1 amplificado de aproximadamente 700 pb, 1 de aproximadamente 900 pb y un último de aproximadamente 1150 pb; con la muestra WT se obtuvieron dos amplificados cercanos a 1000 pb y uno de aproximadamente de 800 bp. Mediante secuenciación Sanger se identificó que ninguno de los 9 amplificados obtenidos corresponde a la secuencia de *LRPPRC*, por lo que se trata de amplificados inespecíficos.

El último método utilizado para sintetizar cDNA fue con oligonucleótidos específicos para el mRNA de *LRPPRC*; el oligonucleótido sentido se une al exón 6 y el antisentido al exón 17 (mismos sitios que el panel superior en Figura 8.2A). Amplificación con los mismos oligonucleótidos de síntesis obtuvo un amplificado de aproximadamente 800 pb, el cual fue complicado de purificar y secuenciar, por lo que se modificó la condición de amplificación y se utilizó el mismo oligonucleótido sentido que el de la síntesis y para el antisentido, se utilizó un oligonucleótido que se une al exón 10 (Panel superior en Figura 8.2C). En la muestra de HepG2 se identificaron dos amplificados, el esperado de 568 pb y uno de aproximadamente 280 pb. En las muestras de paciente (HT y WT) se identificó un amplificado de aproximadamente 250 pb y otro de aproximadamente de 150 pb (Panel inferior en Figura 8.2C). Se purificaron y secuenciaron los amplificados de 250 y 150 pb. Los dos amplificados se identifican como parte de *LRPPRC*, en el cual se lee el final del exón 6, los primeros 5 nucleótidos del exón 7, ausencia de los primeros 4 nucleótidos del



exón 10, una secuencia de 7 nucleótidos que son compartidos del exón 7 y 10 y continúa la secuencia del exón 10. Por lo anterior, en este amplificado se pierde un fragmento del exón 7, los exones 8 y 9 completos y los primeros 4 nucleótidos del exón 10 (Figura suplementaria 1.6).

Estos resultados sugieren la presencia isoformas de mRNA de *LRPPRC* no reportadas previamente, en 2 de las tres condiciones y en las tres muestras analizadas. Sin embargo, no identificamos ningún transcrito alternativo, exclusivo de la muestra del paciente portador heterocigoto. La variante V9 se ha probado que influye en patrones de empalme del RNA<sup>45</sup>, por lo que pudiera estar participando en estas diferentes isoformas identificadas.



**Figura 8.2. Búsqueda de isoformas de LRPPRC con 3 condiciones de síntesis de cDNA.** Amplificados de cDNA sintetizado con A) oligonucleótidos aleatorios, B) oligonucleótido poli (dT)18 y C) oligonucleótidos específicos para cDNA de LRPPRC. En los paneles superiores (A), (B) y (C) se muestra una representación esquemática de la zona donde se encuentra la variante de interés (flechas verticales negras) y la zona del cDNA de LRPPRC donde se estarían uniendo los oligonucleótidos (flechas horizontales; los colores son para representar los pares diferentes utilizados). Paneles inferiores (A), (B) y (C) se muestra el análisis de PCR de punto final de la retrotranscripción del mRNA de LRPPRC. Gel de agarosa al 3%. MPM: Marcador de peso molecular, HT: Heterocigoto, WT: *Wild Type* (no portador).

## **Estudios de niveles de expresión de genes mitocondriales en tejido hepático y la línea celular humana HepG2**

Decidimos también evaluar si existían diferencias en la abundancia de mRNA de *LRPPRC* y genes mitocondriales dependiendo de la presencia de los alelos de riesgo en tejido hepático. Se utilizó el mismo RNA proveniente de las muestras de biopsias hepáticas de pacientes no relacionados portadores y no portadores de las variantes de riesgo (Sección 8. Búsqueda de isoformas de *LRPPRC*). El estudio de expresión se dividió en 2 partes, la primera fue la validación del método y la segunda, el análisis de los ensayos de la PRC cuantitativa. Como control positivo de expresión de *LRPPRC* y genes mitocondriales se utilizó el cDNA de la línea celular humana HepG2 para la validación del método, incluyendo la compatibilidad de las sondas *TaqMan* y los oligonucleótidos específicos utilizados para los genes blanco y de referencia. Analizamos los datos obtenidos de las curvas de amplificación o Ct, que corresponde al ciclo en el cual la curva de detección cruza el umbral determinado para cada transcrito del estudio (Figura 8.3A, 8.3B).

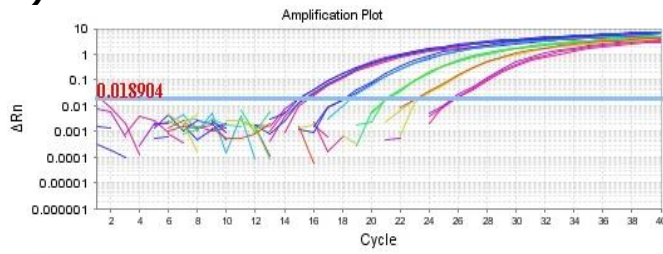
### **Nivel de expresión para *LRPPRC***

#### **Validación del método para expresión de *LRPPRC***

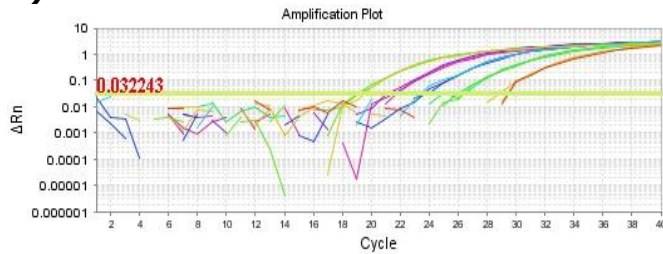
Para comprobar que la sonda de expresión para el gen constitutivo *GUSB* (gen codificante para una enzima expresada en la mayoría de los tejidos y está involucrada en la degradación de proteoglicanos dentro del lisosoma<sup>84</sup>) es compatible con la sonda de *LRPPRC*, se realizó una validación del método en el cual se calcula una regresión lineal de los  $\Delta Ct$  ( $Ct_{promedio_{LRPPRC}} - Ct_{promedio_{GUSB}}$ ) con el Log10 de la cantidad de cDNA de cada uno de los puntos de una curva estándar a partir de concentraciones decrecientes de cDNA. (Figura 8.3C). De la regresión se toma el valor de la pendiente y para que el ensayo se considere válido (e.g. que las sondas compatibles) se requiere un valor menor a 0.1. En este método, se asume que la cantidad de RNA utilizado en la conversión a cDNA se convierte en su totalidad a cDNA (eficiencia del 100%).

Con la primera regresión lineal (Figura 8.3D) el valor de la pendiente es de 0.2294, por lo que no se cumple la condición necesaria; en el reajuste de la Figura 8.3E se cumple el criterio establecido ( $m = -0.0729$ ). Por lo anterior, se pueden considerar las dos sondas compatibles y el método validado.

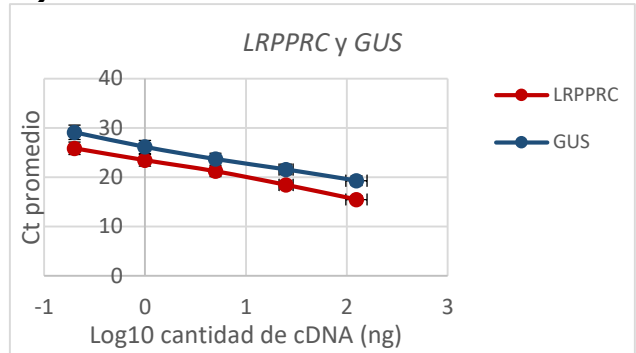
**A) LRPPRC**



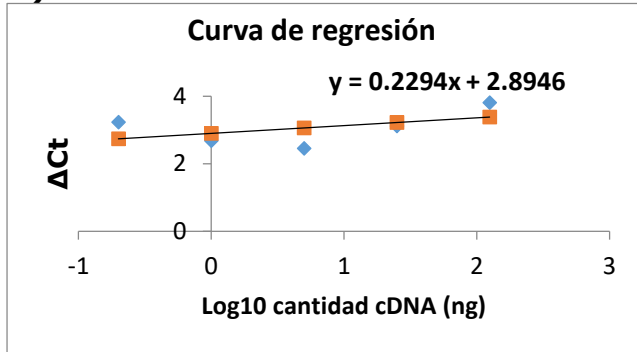
**B) GUSB**



**C)**



**D)**



**E)**

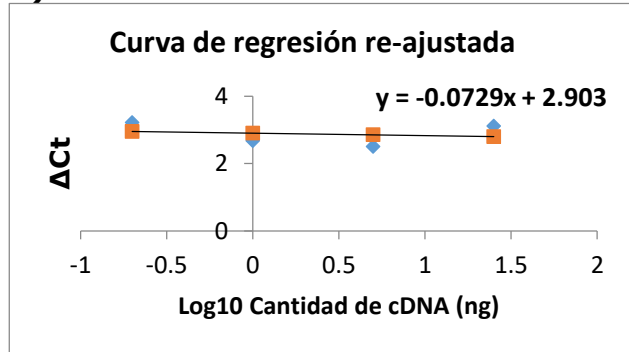


Figura 8.3. **Curvas estándar a concentraciones decrecientes de cDNA de HepG2 para LRPPRC (gen blanco) y GUSB (gen constitutivo).** A) Plot de amplificación de LRPPRC con 5 cantidades de cDNA B) Plot de amplificación de GUSB con 5 cantidades de cDNA C) comparación de las curvas estándar de LRPPRC y GUS D) Curva de regresión de la validación del método y E) curva de regresión re-ajustada de la validación del método. (Las diluciones fueron hechas 1:5 y cada condición se realizó por triplicado).

### **PCR cuantitativa para la expresión de *LRPPRC* en muestras de tejido hepático de portadores y no portadores**

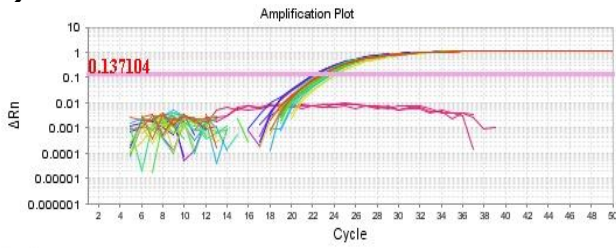
Una vez validado el método y la compatibilidad de las sondas para los genes *LRPPRC* y *GUSB*, se procedió a realizar la PCR cuantitativa en muestras de tejido hepático de portadores y no portadores del haplotipo de riesgo (Figura 8.4).

Se realizó la amplificación del gen de interés (*LRPPRC*, Figura 8.4A) y el gen constitutivo (*GUSB*, Figura 8.4B) con cada muestra de paciente por duplicado y a los valores de expresión relativa calculados mediante el método  $2^{-\Delta Ct}$  (Figura 8.4C) se les realizó una prueba de U de Mann Whitney para probar si había alguna diferencia significativa con los valores dependiendo del estado de portador o no portador. El valor p de esta prueba estadística fue de 1, por lo que los cambios de expresión genética de *LRPPRC* observados son azarosos (Figura 8.4E), representados en las barras de error en las medias de los valores de  $2^{-\Delta Ct}$  (Figura 8.4D).

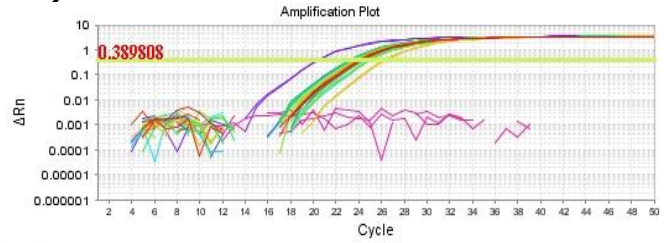
El valor de  $\Delta Ct$  de HepG2 no se consideró, ya que esta muestra representaba únicamente el control positivo de la expresión de *LRPPRC*.

Por lo tanto, por el tamaño de muestra, además de contar con dos muestras homocigotas, no fue posible evidenciar cambios en la expresión de *LRPPRC*, independientemente de ser portador o no del haplotipo de riesgo, ya que las diferencias en la expresión relativa no son estadísticamente significativas.

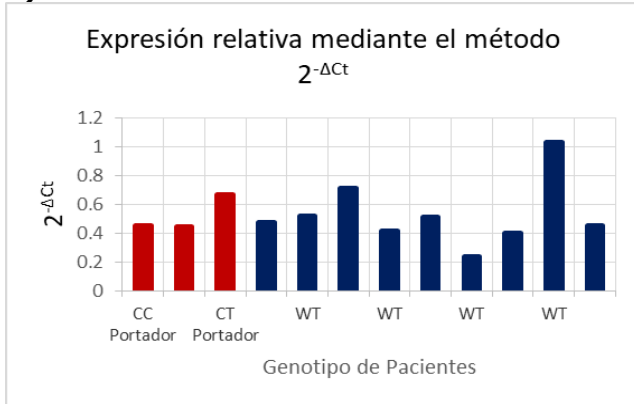
**A) LRPPRC**



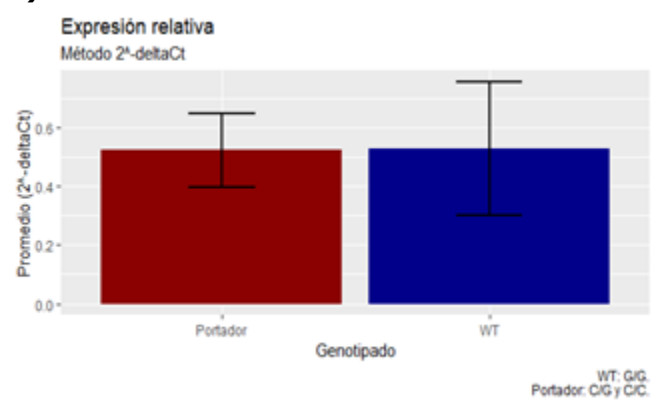
**B) GUSB**



**C)**



**D)**



**E)**

Wilcoxon rank sum test  
 data: deltaCt by Status  
 W = 14, **p-value = 1**  
 alternative hypothesis: true location shift is not equal to 0

Figura 8.4. **Expresión relativa de LRPPRC en muestras de tejido hepático de portadores y no portadores de las variantes de riesgo.** Plots de amplificación de A) LRPPRC y B) GUSB. C) Comparación de los valores de  $2^{-\Delta Ct}$  en muestras de portadores y no portadores, D) comparación de medias de los valores de  $2^{-\Delta Ct}$  en muestras de portadores y no Portadores y E) análisis estadístico U de Mann Whitney. WT: *Wild Type* (no portador).

## Nivel de expresión para genes mitocondriales

Al no encontrar diferencias significativas en la expresión de *LRPPRC* entre portadores y no portadores de las variantes de riesgo, se evaluó si pudieran existir cambios en la expresión de distintos genes mitocondriales, como se ha observado previamente en la literatura<sup>51,52,55</sup>. Para la qPCR se utilizó el reactivo *SYBR GREEN*, oligonucleótidos reportados en la literatura para la amplificación de los genes *ND1*, *ND6*, *COX1*, *CYTB*, *ATP6* y *ATP8*, así como cDNA proveniente de RNA total de tejido hepático de portadores y no portadores de las variantes de riesgo y de células HepG2.

## Validación del método para expresión de genes mitocondriales.

De igual forma fue necesario validar el método y compatibilidad de los oligonucleótidos utilizados con el sistema *SYBR GREEN*, ya que el sistema es diferente al previamente utilizado ensayo *TaqMan*.

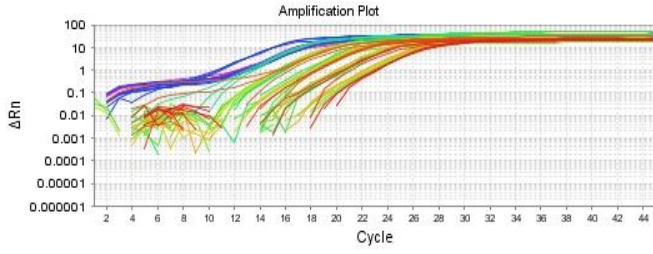
Para la validación se realizaron curvas estándar de concentraciones decrecientes de cDNA de HepG2 para los genes mitocondriales seleccionados y 3 diferentes genes constitutivos; *h18s* (gen para codificar el RNA ribosomal, componente estructural de la subunidad pequeña ribosomal 40S y componente básico de células eucariontes<sup>85</sup>), *GAPDH* (codificante para la enzima Gliceraldehído 3-fosfato deshidrogenasa, enzima clave para la glucólisis<sup>86</sup>) y *ACTB* (gen codificante para la  $\beta$ -actina, una de las dos proteínas del citoesqueleto que es expresada de manera ubicua<sup>87</sup>). (Figura 8.5).

De igual manera que en el ensayo *TaqMan*, la validación se realizó con las regresiones lineales de los valores de  $\Delta Ct$  ( $Ct_{promedio_{genblanco}} - Ct_{promedio_{genconstitutivo}}$ ) con los valores del Log10 de la cantidad de cDNA de las curvas estándar (Figura 8.5F, 8.5G y 8.5H). Se considera validado el método y oligonucleótidos compatibles si el valor de la pendiente en la regresión es menor a 0.1.

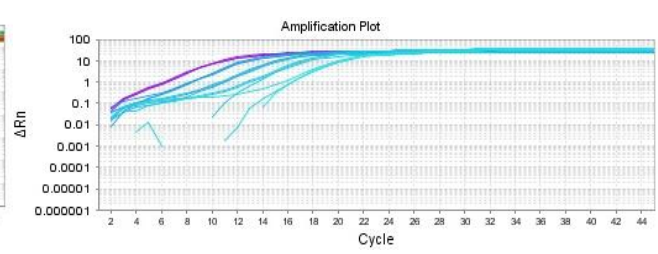
Los valores obtenidos de las pendientes en la regresión lineal (Figura 8.5I) para *h18s* de los 6 genes blanco son menores a 0.1; mientras que para *GAPDH*, 5 valores de las pendientes para 5 genes (*ND1*, *COX1*, *CYTB*, *ATP6* y *ATP8*) son menores a 0.1 y el obtenido para el gen *ND6* es mayor a 0.1. Para  $\beta$ -actina, todos los valores de pendiente obtenidos para los 6 genes son mayores a 0.1. Se realizó un reajuste de la regresión de  $\beta$ -actina, como con *LRPPRC* y *GUSB* (Figura 8.3E), al quitar el punto con la concentración más alta y aun así, todos los valores de las pendientes de los 6 genes blanco son mayores a 0.1.

Con esta validación se consideraron compatibles los distintos oligonucleótidos para los genes blancos mitocondriales (*ND1*, *ND6*, *COX1*, *CYTB*, *ATP6* y *ATP8*) y para el gen constitutivo *h18s*.

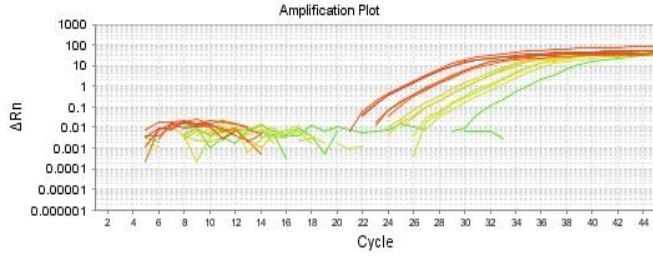
**A) ND1, ND6, COX1 y CYTB (4/6 genes blanco)**



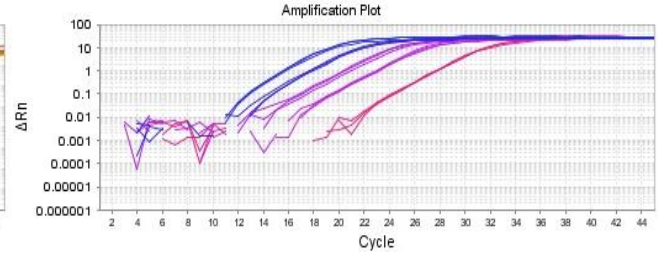
**B) h18s**



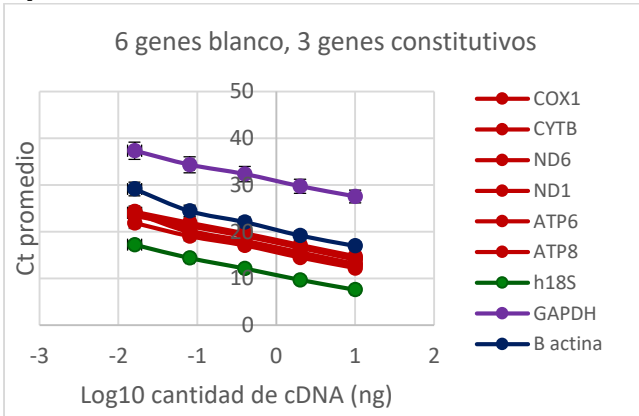
**C) GAPDH**



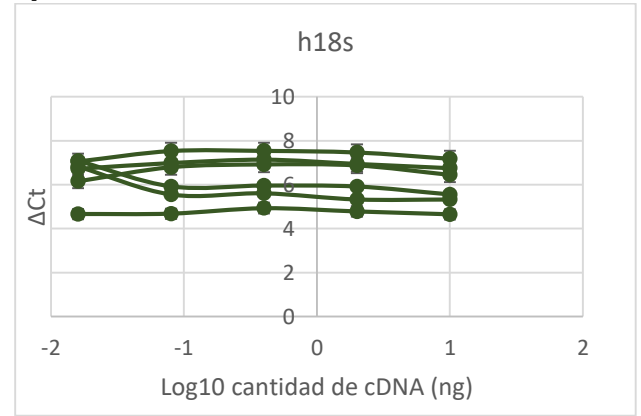
**D) β-ACTINA**



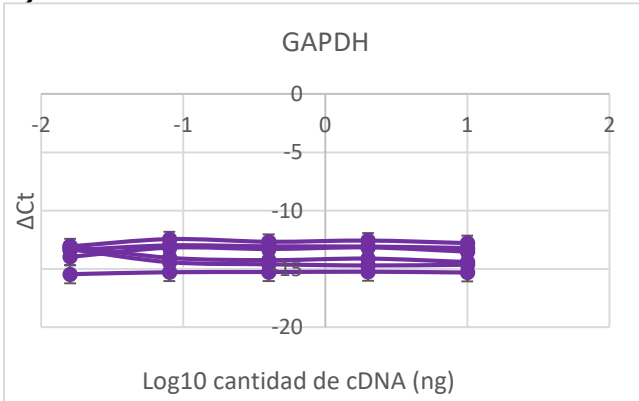
**E)**



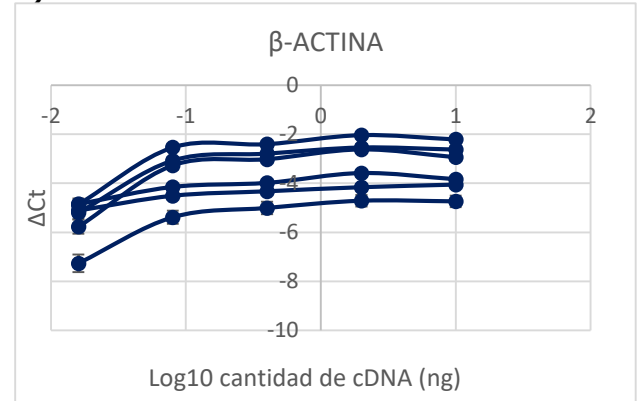
**F)**



**G)**



**H)**





I)

Gen mitocondrial	Ecuación de la recta ( $\Delta Ct$ vs LOG10 cantidad cDNA)			
	h18s	GAPDH	$\beta$ -ACTINA	$\beta$ -ACTINA reajustado con 4 puntos
ND1	$y = 0.0256x + 7.3639$	$y = 0.0625x - 12.674$	$y = 0.8335x - 2.4799$	$y = 0.1944x - 2.2875$
ND6	$y = 0.0952x + 6.6733$	$y = 0.1321x - 13.364$	$y = 0.9031x - 3.1704$	$y = 0.2009x - 2.959$
CYTB	$y = 0.0028x + 6.9143$	$y = 0.0396x - 13.123$	$y = 0.8107x - 2.9294$	$y = 0.2275x - 2.7539$
COX1	$y = 0.0109x + 4.7447$	$y = 0.0478x - 15.293$	$y = 0.8188x - 5.099$	$y = 0.3203x - 4.949$
ATP6	$y = -0.4376x + 5.9086$	$y = -0.4008x - 14.129$	$y = 0.3703x - 3.9351$	$y = 0.1918x - 3.8814$
ATP8	$y = -0.4577x + 5.5454$	$y = -0.4208x - 14.492$	$y = 0.3502x - 4.2984$	$y = 0.2178x - 4.2585$

Figura 8.5. **Curvas estándar a concentraciones decrecientes de cDNA de HepG2 para 6 genes mitocondriales (ND1, ND6, COX1, CYTB, ATP6 y ATP8) y 3 constitutivos (h18s, GAPDH y  $\beta$ -ACTINA).** Plots de amplificación de A) 4 de los 6 genes mitocondriales estudiados (ND1, ND6, COX1 y CYTB), B) h18s, C) GAPDH y D)  $\beta$ -ACTINA; (E) comparación de las curvas estándar de los genes blanco y los constitutivos; regresiones lineales de los valores de  $\Delta Ct$  para F) h18s, G) GAPDH y  $\beta$ -actina; (I) Ecuaciones de la recta de la regresión lineal de los 3 genes constitutivos para la validación del método y compatibilidad de los oligonucleótidos, Verde: método validado y oligonucleótidos compatibles, Rojo: no validado y oligonucleótidos no compatibles. (Las diluciones fueron hechas 1:5 y cada condición se realizó por triplicado).

## **PCR cuantitativa para la expresión de genes mitocondriales en muestras de tejido hepático de portadores y no portadores**

Tomando en cuenta las curvas estándar realizadas en la validación, se determinó que la cantidad de cDNA a utilizar con las muestras de tejido hepático para la evaluación de la expresión de genes mitocondriales sería de 0.4 ng. Se realizó el ensayo de qPCR para evidenciar posibles diferencias de expresión con el sistema *SYBR GREEN*, oligonucleótidos para los genes mitocondriales blanco y para el gen constitutivo *h18s*, así como cDNA de tejido hepático de portadores y no portadores. Cada condición se realizó por triplicado.

Para los valores de expresión relativa, se calcularon los valores de  $2^{-\Delta\Delta Ct}$  en donde el cálculo de  $\Delta\Delta Ct$  se obtiene mediante la diferencia del  $\Delta Ct_{\text{promedio}_{\text{genblanco}}}$  con el  $\Delta Ct_{\text{promedio}_{\text{genconstitutivo}}}$ . Para hacer las comparaciones de la expresión entre portadores y no portadores, se tomó la media de los valores de  $\Delta Ct$  en cada uno de los 6 genes en las muestras de no portadores para hacer el cálculo de  $\Delta\Delta Ct$  y el valor de  $2^{-\Delta\Delta Ct}$  en estas muestras se normalizó a 1 para comparar con los valores de  $2^{-\Delta\Delta Ct}$  calculados para portadores.

Con estas comparaciones de expresión relativa, se observa que en los sujetos portadores 1 y 3 (Figura 8.6A y 8.6C) no presentan diferencias significativas en la expresión relativa de genes mitocondriales, exceptuando el gen *COX1* cuya expresión está disminuida. En el caso del portador 3, se ve aumentada la expresión del gen *ND1*, en comparación de la normalización del no portador. Para el sujeto portador 2 (Figura 8.6B), se observan una expresión disminuida en aproximadamente un 50% de los genes mitocondriales analizados. En la Figura 8.6D se muestra la comparación con la normalización de no portadores y la media de los 3 valores de  $2^{-\Delta\Delta Ct}$  de los portadores con una disminución de la expresión de los 6 genes mitocondriales, a pesar de que la dispersión de los datos representada en las líneas de error de desviación estándar es alta. Llama la atención el valor de expresión relativa para el gen *COX1* es el que se encuentra más abatido, aunque el valor *p* (0.063) es limítrofe, aun así resulta estadísticamente no significativo.

La disminución en la expresión de genes mitocondriales observada en la muestra del portador 2 es consistente a lo observado en el modelo murino *knock-down* a *LRPPRC*.

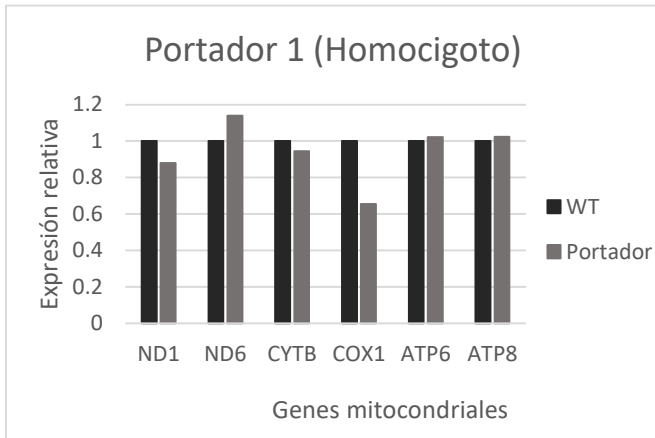
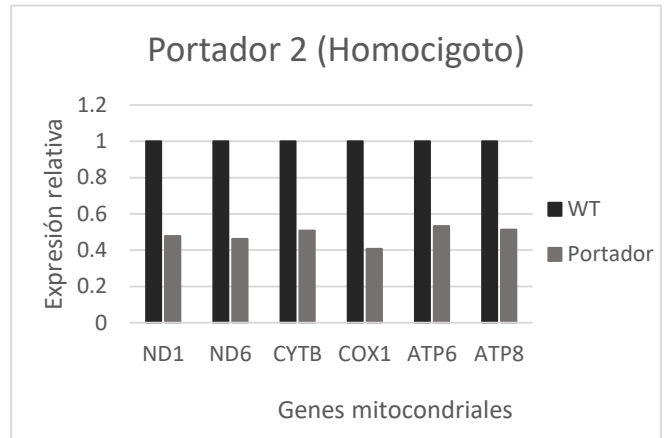
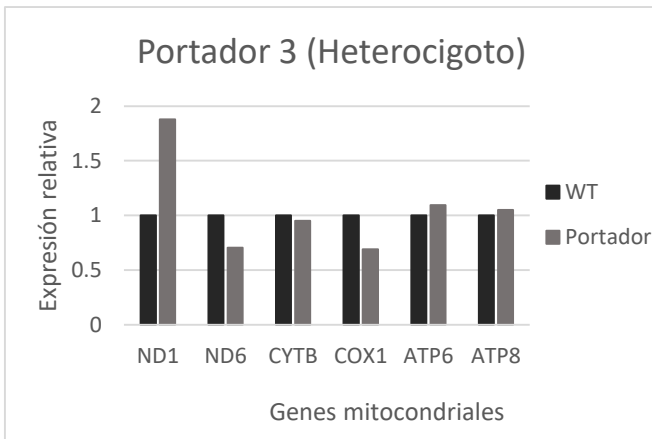
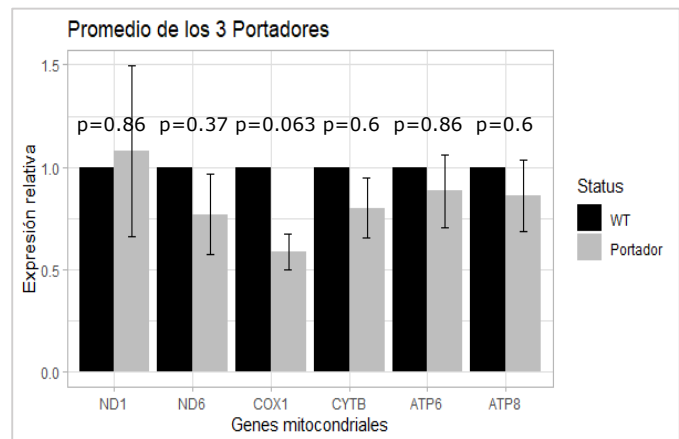
**A****B****C****D**

Figura 9.6. **Expresión relativa de genes mitocondriales en portadores de las variantes de riesgo.** Comparación de la expresión relativa normalizando la media de los valores de  $2^{-\Delta\Delta C_t}$  de los portadores a 1 y los valores de  $2^{-\Delta\Delta C_t}$  de los portadores; los homocigotos (A) y (B), el heterocigoto (C) y la media de los 3 portadores (D). WT: *Wild Type* (no portador).

## Búsqueda bioinformática de interacciones a larga distancia y co-localización con estados de la cromatina

Por último enfoque del proyecto, nos propusimos analizar posibles interacciones a larga distancia con otros genes del mismo *locus*.

Con este propósito se realizó una búsqueda en los datos públicos de Hi-C del *locus* de *LRPPRC* en tejido hepático sano, para evidenciar posibles Dominios de Asociación Topológica (TADs, *Topologically associated domain*) y estructuras presentes en la cromatina que pudieran estar cercanas espacialmente. Un ejemplo de estas regiones cercanas son los sub-TADs, los cuales darían un indicio de cercanía tridimensional y una posible interacción entre genes de la región.

Se inició con una ventana de 2 Mb (aproximadamente  $\pm 1$  Mb de distancia de *LRPPRC*) y una resolución en el *heatmap* de 30 kb para identificar el tamaño del TAD en el que estuviera contenido *LRPPRC*, el cual se identificó en la ventana chr2:43825732-45179748 (GRCh37/hg19). En esta región de aprox 1.4 Mb se reportan las siguientes estructuras genéticas: los genes codificantes ***PLEKHH2*** (*pleckstrin homology, MyTH4 and FERM domain containing H2*; dominio homólogo conteniendo a H2 de pleckstrina, MyTH4 y FERM), ***LOC728819***, ***DYNC2LI1*** (*dynein cytoplasmic 2 light intermediate chain 1*; cadena ligera intermediaria 1 de la dineína citoplasmática 2), ***ABCG5*** (*ATP binding cassette subfamily G member 5*; Miembro 5 de la subfamilia G de la familia de tipo casete de unión a ATP), ***ABCG8*** (*ATP binding cassette subfamily G member 8*, Miembro 8 de la subfamilia G de la familia de tipo casete de unión a ATP), ***LRPPRC***, ***PPM1B*** (*protein phosphatase, Mg<sup>2+</sup>/Mn<sup>2+</sup> dependent 1B*; fosfatasa 1B dependiente de Mg<sup>2+</sup>/Mn<sup>2+</sup>), ***SLC3A1*** (*solute carrier family 3 member 1*; miembro 1 de la familia de 3 de transportador de solutos), ***PREPL*** (*prolyl endopeptidase like*; parecida a prolil-endopeptidasa), ***CAMKMT*** (*calmodulin-lysine N-methyltransferase*; metiltransferasa de calmodulina-lisina), ***SIX3*** (*SIX homeobox 3*; factor homeótico SIX 3) y el miRNA ***MIR548AD***. Lo que respecta a ***7SK*** y ***Mir\_548***, se predice que son regiones con una estructura genética codificante. (Figura 8.7.B).

El gen de interés *LRPPRC* se encuentra en la región chr2:44113363-44223144 (GRCh37/hg19; aproximadamente 110 kb), cercano a la frontera río arriba de la estructura TAD. La señal de interacción (que varía en intensidad de rojo) está

### Arquitectura tridimensional del genoma

La organización espacial del genoma dentro del núcleo en eucariontes superiores no es azarosa y tiene un papel importante en funciones celulares como ciclo celular, diferenciación y senescencia.

El empaquetamiento de la cromatina provoca contacto de regiones distantes y este contacto físico entre genes y sus elementos regulatorios (marcas en la cromatina y proteínas) es crucial para mecanismos de regulación génica.

**TAD (Topologically Associated Domain)** Regiones en el genoma formadoras de asas con fronteras definidas con función organizacional y de regulación de expresión, debido a las interacciones del *loci* dentro de esta estructura y una mínima o nula interacción fuera de la misma.

**Hi-C (high-throughput chromosome conformation capture)** Método que utiliza ligación de regiones en la cromatina con contacto físico y una subsecuente secuenciación para identificar todas las interacciones que ocurren a través del genoma.

Pombo et al, 2015; Rocha et al, 2015; Rowley et al, 2018.

representado por el *heatmap* del Hi-C en la Figura 8.7.A. Esta representación muestra dos regiones con una mayor interacción dentro del TAD, evidenciando dos posibles estructuras sub-TAD de un tamaño aproximado de 600 kb. (Resaltado en la Figura 8.7.A).

Para las interacciones a larga distancia, se utilizó el inicialmente la misma ventana de 2MB y las predicciones de interacciones con la región promotora de *LRPPRC*. Los criterios para tomar en cuenta la interacción como probable fueron que la interacción se encontrara en una región dentro del TAD, un valor mayor a 2 en el valor normalizado de la interacción y que existiera algún indicio de una región reguladora; todo reportado por el portal *3DIV*. De esta forma, se identificaron 18 posibles interacciones con el promotor de *LRPPRC* (Figura 8.7.C): a) **1** con la región terminal del gen *DYNC2LI1*, b) **3** en tándem con los genes *ABCG5* y *ABCG8*, c) **4** en tándem dentro de *LRPPRC*, d) **3** en tándem con una **región intergénica** (sin genes reportados), e) **1** con el gen *PPM1B*, f) **3** con el gen *SLC3A1* y g) **3** con una región compartida por el gen *CAMKMT* y la secuencia para un miRNA *MIR548AD*. (Figura 8.C.D)

Con este acercamiento se evidencia que *LRPPRC* tiene posibles interacciones con 7 genes dentro del TAD que lo contiene, por lo tanto se postula que el *locus LRPPRC* representa una región de regulación a distancia.

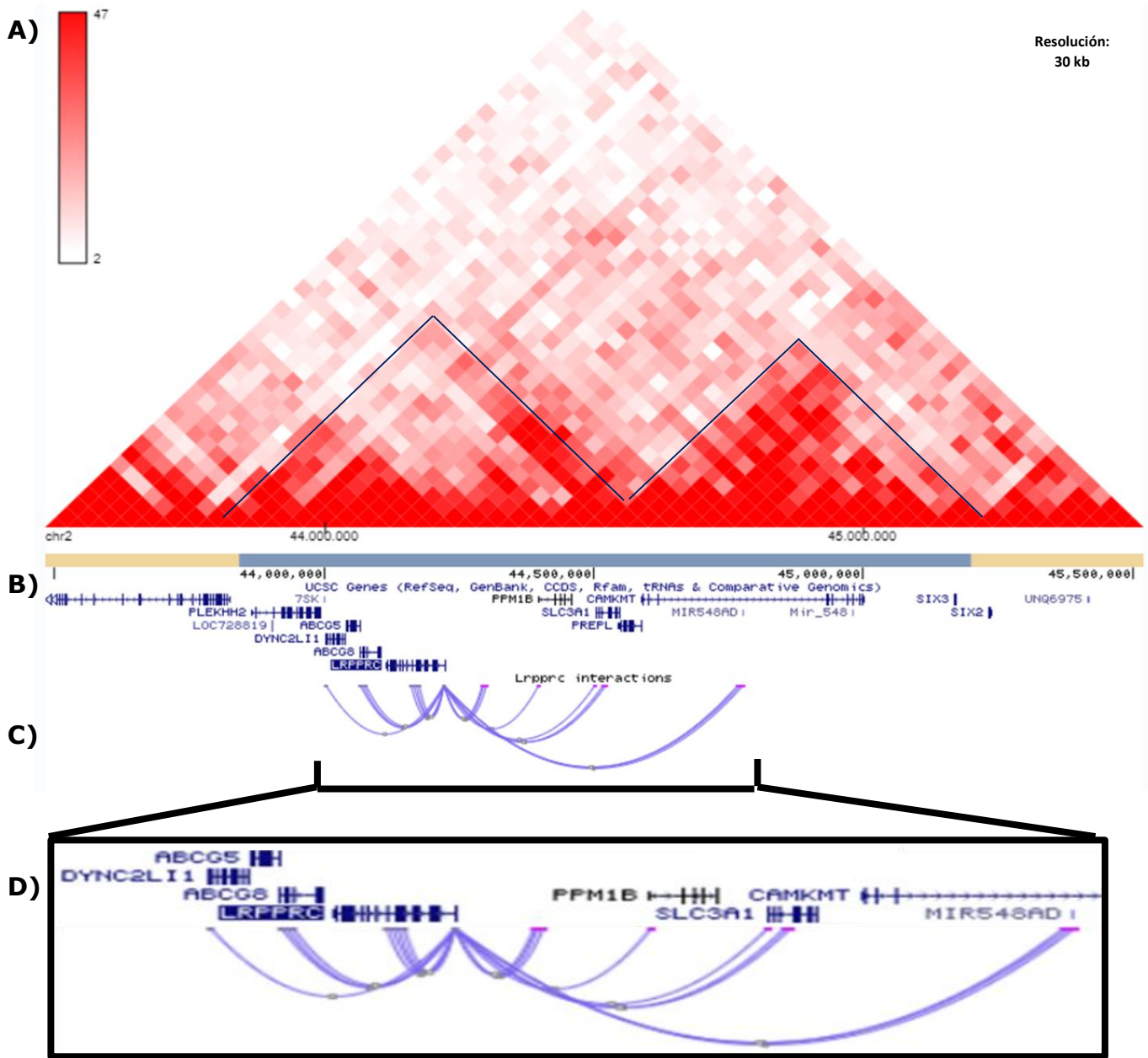


Figura 8.7. **Estructuras tipo TAD presentes en el locus de *LRPPRC*.** A) Matriz de datos públicos de Hi-C en tejido hepático sano, B) TADs y estructuras genéticas reportados en la región, C) Spider-plot de las predicciones de interacciones con la región promotora de *LRPPRC*, D) Acercamiento de la región de interacciones del promotor de *LRPPRC* con los genes aledaños. Resolución del *heatmap*: 30 kb.

Una vez identificada la posible interacción de la región promotora de *LRPPRC* con otros genes del mismo *locus*, se analizó si estas interacciones podrían tener un significado biológico por medio de un mapeo de datos bioinformáticos con co-localización de marcas en la cromatina. Para ello se utilizaron datos públicos de *ChIP-seq* para las marcas en las histonas de regiones *enhancer* (H3K27ac y H3K4), accesibilidad de la cromatina (DNAsa), unión de proteínas de la arquitectura nuclear que delimitan las regiones *enhancer* (CTCF y RAD21; subunidad de la cohesina) en tejido hepático sano y unión de proteínas que definen un *super-enhancer* (MED1) en la línea celular HepG2. Se delimitó la región de estudio a la establecida por la estructura TAD, de aproximadamente 1.4 Mb (chr2:43825732-45179748).

La señal de *ChIP-seq* está representada en picos que superan un score establecido para distinguirlas de una señal de fondo. Algunas son únicas y otras son un cúmulo de señal que se representa como distintos picos. Para la región en cuestión, la señal de la marca de la cromatina H3K27ac (Figura 8.8C) identifica 3 picos de señal únicos y 4 de múltiples picos, teniendo la más alta en una región intragénica entre los genes *LRPPRC* y *PPM1*. Por otra parte, en H3K4me1 (Figura 8.8D), hay más de 10 cúmulos de picos, donde la más alta se encuentra en el locus de *DYNC2LI1*. Asimismo, para la señal de la DNAsa (Figura 8.8E) se muestran 3 pequeños cúmulos de señal y un pico franco en el locus de *LRPPRC*. Adicionalmente, con las señales de CTCF (Figura 8.8F) se identifican 9 picos únicos, con el mayor en una región intergénica entre *PPM1B* y *SLC3A1* así como un patrón de picos muy parecido para RAD21 (Figura 8.8G). Por último, para la marca de la cromatina MED1 (Figura 8.8H), no se observan picos únicos ni con señal alta, sólo 8 pequeños cúmulos.

De esta forma se identificaron 11 interacciones que co-localizan con señales conjuntas de todas las marcas en la cromatina. Las cuales fueron: a) **3** interacciones en tándem en el locus que comprende a los genes ***DYNC2LI1, ABCG5, ABCG8***, b) **4** interacciones en tándem dentro del locus de ***LRPPRC*** y c) **3** interacciones en tándem en la **región intergénica** localizada en las coordenadas chr2:44290000-44305000 (GRCh37/hg19, 15 kb); entre los genes *LRPPRC* y *PPM1B*.

El cúmulo de estas señales muestra evidencia para que esta región sea considerada como de *super-enhancer*; es decir regiones transcripcionalmente activas por el reclutamiento de factores de transcripción que activan la transcripción de genes a distancia y son definidas por la unión de MED1 (*Mediator 1*; mediador 1)<sup>88</sup>.

Esta información sugiere que 11 de las 18 posibles interacciones identificadas podrían tener un potencial biológico de regulación a distancia.

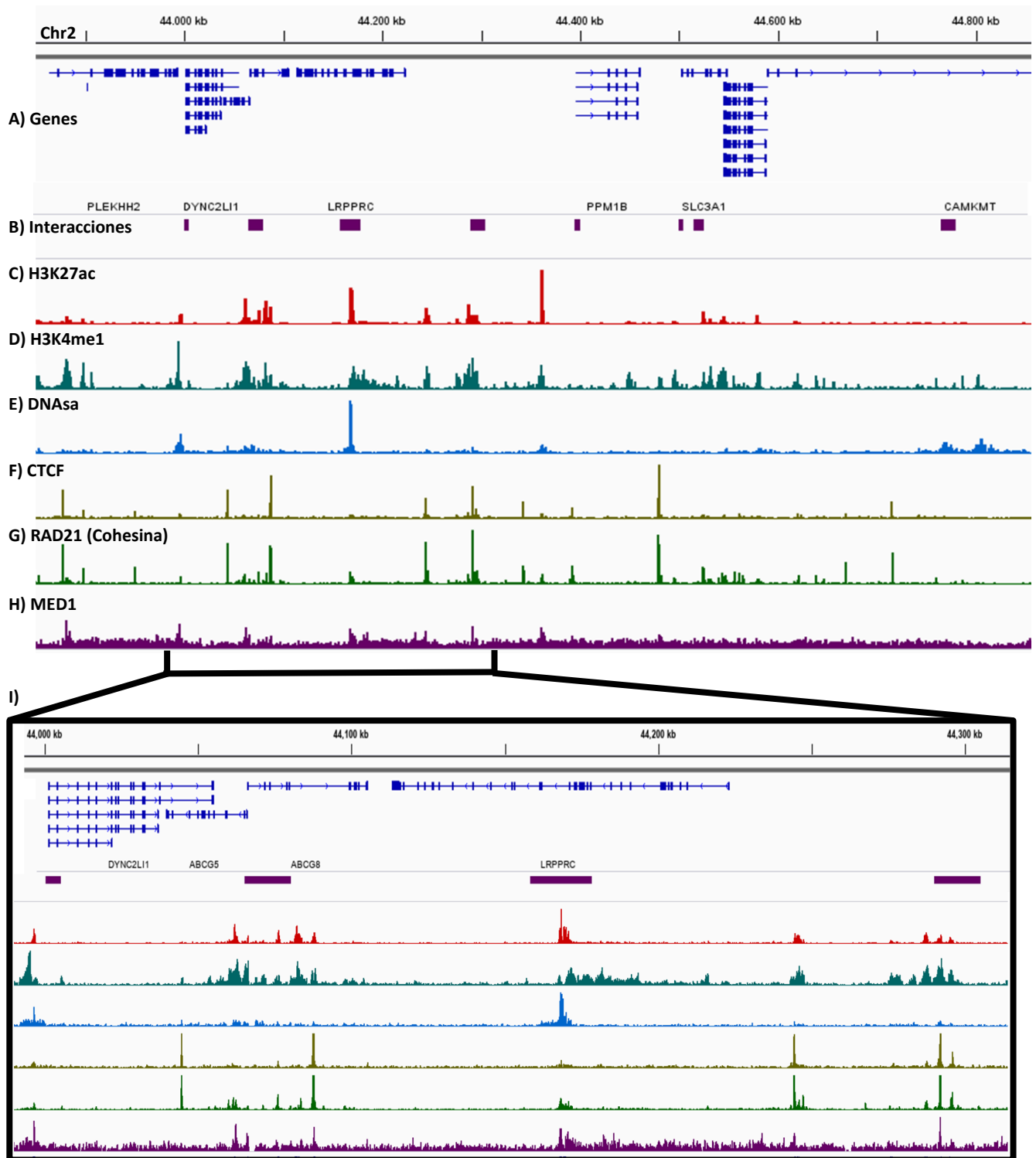


Figura 8.8. **Mapeo del locus de *LRPPRC***. A) genes de la región, B) predicción de interacciones, señales en tejido hepático sano de C) H3K27ac, D) H3K4me1, E) DNAsa, F) unión de CTCF y G) unión de RAD21 (subunidad de la cohesina). H) Señal en HepG2 de unión de MED1. I) acercamiento de la región de aproximadamente de 312 kb (chr2:43,999,679-44,311,991, GRCh37/hg19) mostrando la co-localización de predicción de interacciones y marcas en la cromatina.



## **Búsqueda de variantes en la región de interacción dentro del intrón 23 de *LRPPRC*.**

De estas interacciones, la más notable fue la identificada dentro del gen *LRPPRC*. Esta región de interacción contiene la señal más grande para accesibilidad en la cromatina, ya que la DNAsa corta el DNA en regiones laxas, abriendo la posibilidad de la entrada a factores de transcripción. La región, contiene cuatro interacciones en tándem que abarcan desde el intrón 14 hasta el intrón 25 de *LRPPRC* (20 kb; chr2:44158145-44178144, GRCh37/hg19), y en ésta, están incluidas 2 de las 3 variantes asociadas a riesgo de HTG, por lo que podría existir una variante en desequilibrio de ligamiento en la región con un potencial regulatorio.

Dentro de esta región, hay una zona que muestra una mayor evidencia de regulación y una posible unión de factores de transcripción dada la acumulación de señales en las marcas de la cromatina (H3K27ac y H3K4me1) y unión de proteínas (DNAsa, CTCF, RAD21 y MED1) analizadas (Figura 8.8.I), así como la co-localización con las 4 interacciones en tándem. Esta región de 5 kb (chr2:44167190-44171910, GRCh37/hg19) abarca parte del intrón 22, el exón 23 y parte del intrón 23. En la secuenciación de los exones de *LRPPRC* (Sección 8. Búsqueda de nuevas variantes en los exones del gen *LRPPRC*) ya se había analizado la secuencia de los exones 22 y 23 y en este análisis se identificó la variante V22 (rs28394191) en el intrón 22, cercana al exón 22, que es una variante identificada tanto en portadores (Tabla 8.1) como en no portadores (Tabla 8.3). Para explorar más esta zona de interés, se realizó la búsqueda de otras variantes que pudieran pertenecer al haplotipo de riesgo mediante la secuenciación en la región del intrón 22 (chr2:44,171,033-44,172,469, GRCh37/hg19) y una región del intrón 23 (chr2:44,166,932-44,170,824, GRCh37/hg19) en portadores de las variantes de riesgo (Sección 8. Búsqueda de nuevas variantes en los exones del gen *LRPPRC*).

De esta secuenciación, se identificaron 2 variantes en el intrón 23 de *LRPPRC*: rs4263155 y rs4600713 (Tabla 8.5), a las que nos referiremos como en Variante de Interacción 1 (VI1) y Variante de Interacción 2 (VI2), respectivamente. Además de su identificación, también se determinó que estas dos variantes son compartidas de manera homocigota por los 4 portadores del haplotipo de riesgo.

Posición génica	Variación en el DNA	No. En NCBI	Presencia en Portadores	Prevalencia en la población (AMR/MXL)	Base de datos
<b>Intrón 23</b>	IVS23+245	rs4263155 (VI1)	4 portadores homocigotos	0.157/0.086	1000 Genomes Fase 3
	IVS23+451	rs4600713 (VI2)	4 portadores homocigotos	0.159/0.086	1000 Genomes Fase 3

Tabla 8.5. **Variantes identificadas en la región de interacción dentro del intrón 23 de LRPPRC en los 4 portadores homocigotos de las variantes de riesgo.** AMR: población americana; MXL: población con ancestría mexicana residente en Los Ángeles, California, EUA.

Para identificar la posible presencia de las variantes en sujetos sanos no portadores del haplotipo, se realizó la búsqueda mediante secuenciación Sanger en muestras de no portadores (Tabla 8.3) y confirmamos que las variantes VI1 y VI2 son exclusivas de los sujetos portadores de las variantes de riesgo (Tabla 8.6).

Variante	Presencia en no portadores a las variantes de riesgo Homocigotos/Heterocigotos/Ausencia
<b>VI1</b>	0/0/12
<b>VI2</b>	0/0/12

Tabla 8.6. **Búsqueda de las 2 variantes identificadas en la región 23 en muestras de no portadores sin HTG.**

Para estas dos variantes, se realizó una búsqueda bioinformática de posibles alteraciones en el empalme del RNA utilizando los tres portales, *Mutation Taster*, *Human Splicing Finder* y *Spliceman Web Server*.

Para las variantes VI1 y VI2 hay predicciones sobre la alteración en proceso de *empalme del RNA*. El criterio que no se cumple con las predicciones anteriormente descritas (Tabla 8.2) fue la cercanía con la unión intrón-exón, sin embargo, los 3 portales predicen una posible implicación en el empalme del RNA. Para las predicciones anteriores, también se tomó en cuenta que las predicciones fueran compartidas, que los 3 portales tuvieran predicciones, por lo que esta recopilación de información se considera como probable (Tabla 8.7).

De esta forma mediante análisis *in silico*, se predice que las variantes VI1 y VI2 pudieran alterar el proceso de empalme del RNA.

<b>Variante</b>	<b>Distancia del sitio de unión exón/intrón (pb)</b>	<b><i>Mutation taster</i></b>	<b><i>Human Splicing Finder</i></b>	<b><i>Spliceman Web Server</i><sup>+</sup></b>
<b>VI1</b>	245	Sitio donador ligeramente incrementado  Omisión de exones y características de la proteína afectada	Sitio donador ligeramente incrementado para un potencial nuevo sitio de empalme del RNA	69%
<b>VI2</b>	451	Sitio aceptor incrementado  Omisión de exones y características de la proteína afectada	Sitio donador ligeramente incrementado para un potencial nuevo sitio de empalme del RNA  Nuevo sitio <i>enhancer</i> y roto otro sitio <i>enhancer</i>  Ligero incremento de sitio <i>silencer</i>	56%

Tabla 8.7. **Recopilación de información para las 2 variantes compartidas de manera homocigota entre los 4 pacientes portadores.** SWS<sup>+</sup>: entre más alto el porcentaje, más probable es que la variante afecte el proceso de empalme del RNA.

Adicionalmente, se realizó una búsqueda *in-silico* complementaria de posibles afectaciones en la unión de factores de transcripción mediante el uso de los portales *HaploReg v4.1*, *Jaspar 2020*, *LASAGNA-Search 2.0* y *Toolkit for Cistrome Data Browser*.

El análisis *in-silico* del *loci* de *LRPPRC*, reveló la presencia de 8 posibles elementos respuesta para diferentes factores de transcripción para la región de variante VI1: **CTCF** (*CCCTC-binding factor*; factor de unión a la secuencia CCCTC), **USF1** (*Upstream stimulating factor 1*; Factor estimulante 1), **E2F** (*E2 factor*; Factor E2), **GCNF** (*Germ Cell Nuclear Factor*; Factor nuclear de célula germinal), **NANOG** (*NANOG Homeobox*; factor homeótico NANOG), **NKX2** (*NK2 Homeobox 1*; factor homeótico NK2 1), **POU5F1** (*POU class 5 transcription factor 1*; factor de transcripción POU 1 clase 5) y **STAT** (*Signal transducer and activator of transcription*; Señal transductora y activadora de la transcripción) (Figura 8.9A).

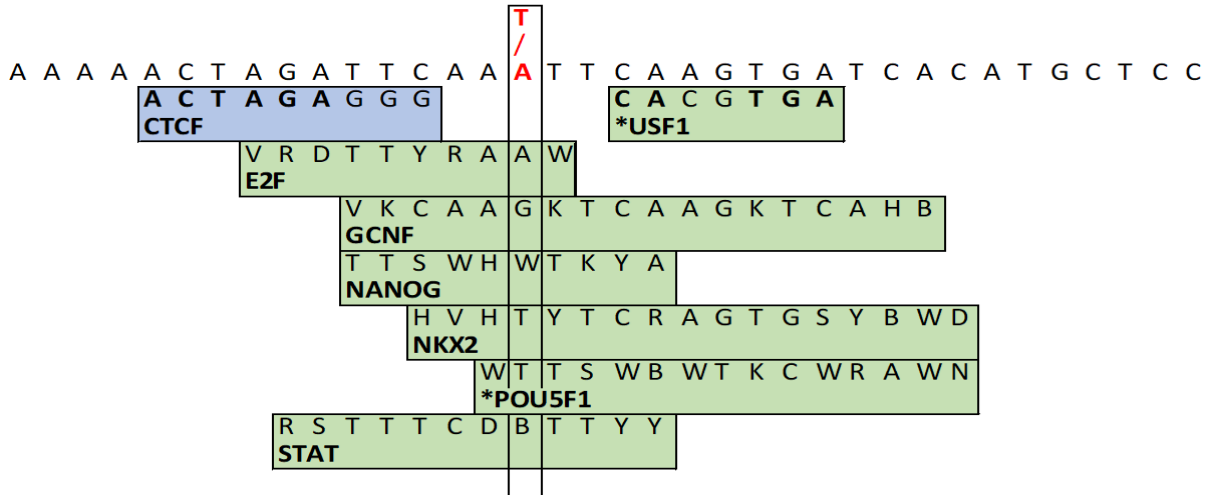
Mientras que para la región que comprende la VI2, se identificaron 6 posibles elementos respuesta para los factores de transcripción **NFE2L1** (*Nuclear Factor Erythroid-2 Related Factor 1*; factor 1 relacionado al factor nuclear eritroide-2), **MAFG** (*MAF bZIP transcription factor G*; factor de transcripción G con cremallera de leucina de la familia MAF), **NFE2** (*nuclear factor erythroid 2*; factor nuclear 2 eritroide), **FOXA** (*forkhead box A*; factor cabeza de horquilla A), **IRF** (*interferon regulatory factor*, Factor regulatorio de interferón), **C11orf30** (*chromosome 11 open reading frame 30*, marco abierto de lectura 30 en el cromosoma 11; codifica para el factor nuclear EMSY), **SS18** (*Synovial sarcoma translocation, chromosome 18*, traslocación del cromosoma 18 del sarcoma sinovial) y **STAT** (Figura 8.9B).

La región que comprende la variante VI1 afecta potencialmente 6 de los 8 elementos de respuesta para los factores de transcripción E2F, GNCF, NANOG, STAT, POU5F1 y NKX2 de manera directa.

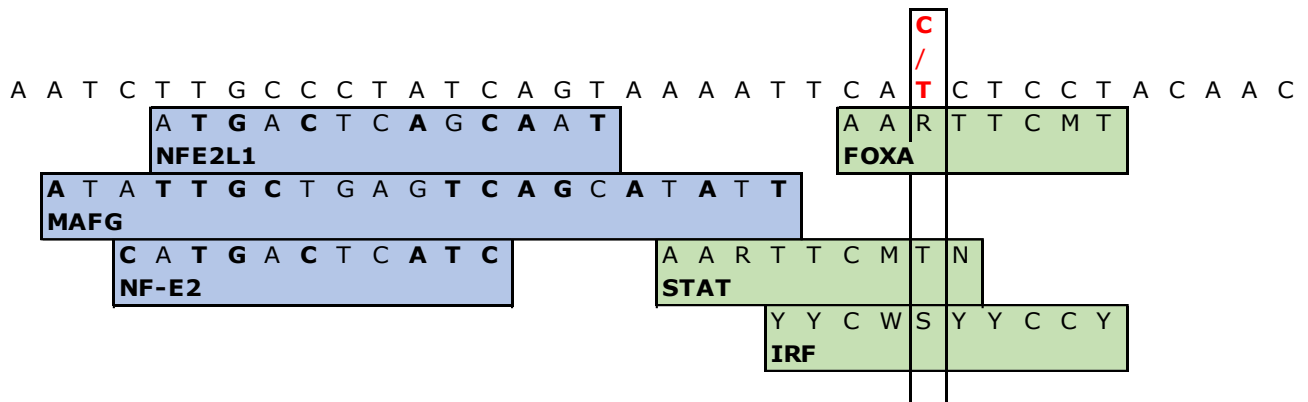
Para la región que comprende la variante VI2, habría afectación directa en los elementos respuesta para los factores de transcripción FOXA, IRF y STAT.

Estos análisis revelan una concentración de información de control transcripcional conglomerada alrededor de las variantes identificadas y abre la posibilidad de que el *locus* pudiera reclutar distintos factores de transcripción cuyos elementos modificaran la unión de distintos factores de transcripción dependiendo de la presencia de las variantes VI1 y VI2.

A)



B)



C11orf30  
SS18

Figura 8.9. **Esquema de la posición de elementos de respuesta a diferentes factores de transcripción dentro del *loci* de las 2 variantes VI1 y VI2.** La caja vertical resalta la posición del nucleótido que difiere entre ambos alelos. A) Locus de la variante VI1, B) locus de la variante VI2. Caja horizontal azul: predicción del portal *Toolkit for Cistrome Data Browser*. Caja horizontal verde: predicción del portal *Haploreg v4.1*. \*: Predicciones encontradas en los dos portales utilizados. Recuadros a la izquierda: factores que no cuentan con secuencia consenso para el motivo respuesta. Nomenclatura de las bases con ambigüedad en la secuencia de los elementos respuesta: R: base púrica; Y: base pirimídica; S: G o C; W: A o T; K: G o T; M: A o C; B: C, G o T; D: A, G o T; H: A, C o T; V: A, C o G.

A la vista de que existe una posible unión en factores de transcripción en variantes identificadas en la región de interacción, se complementó con una búsqueda bioinformática en los mismos sitios utilizados (*HaploReg v4.1*, *Jaspar 2020*, *LASAGNA-Search 2.0* y *Toolkit for Cistrome Data Browser*) para el *loci* de las variantes pertenecientes al haplotipo de riesgo V9 y V15 (Tabla 8.3 y sección 8. Búsqueda de isoformas de *LRPPRC*) y se realizó para identificar si existían datos de alteración en los motivos respuesta a factores de transcripción. No se realizó en las variantes V3, V20 y V22 ya que se demostró mediante secuenciación (Tabla 8.3) que también están presentes en no portadores.

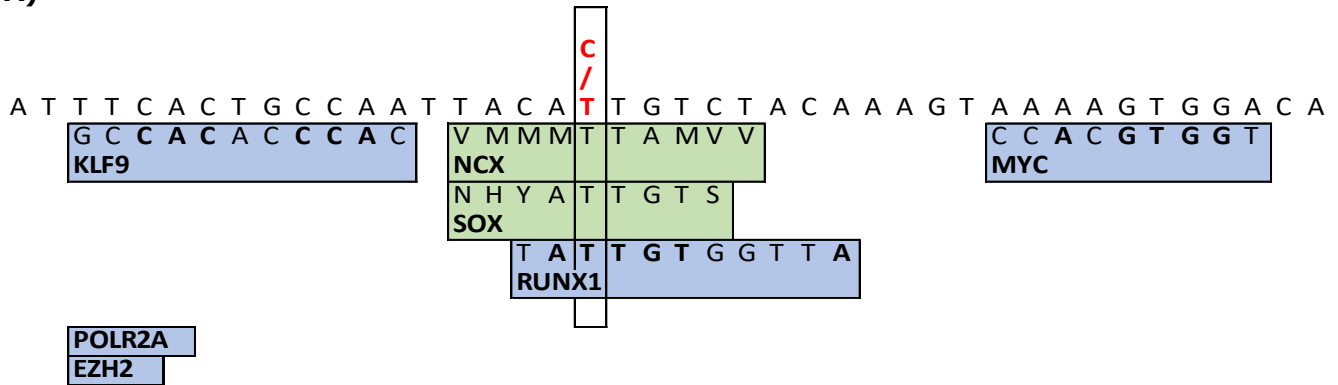
Este análisis reveló la predicción de 7 posibles elementos respuesta para diferentes factores de transcripción para el *locus* de la variante V9: **EZH2** (*Enhancer of zeste 2; enhancer 2* de la proteína *zeste*), **KLF9** (*Krüppel-like factor 9; factor 9* tipo Krüppel), **MYC** (*MYC Proto-Oncogene, BHLH Transcription Factor*), **POLR2A** (*RNA polymerase II subunit A, subunidad A* de la RNA polimerasa II), **RUNX1** (*RUNX Family Transcription Factor 1; factor de transcripción 1* de la familia RUNX), **NCX** (también conocido como *TLX2 T-cell leukemia homeobox 2; factor homeótico 2* de leucemia de células T) y **SOX** (*SRY-related HMG-box genes; factor con dominio de alta movilidad relacionado con la proteína región Y determinante del sexo*) (Figura 8.10A).

Mientras que para la V15 se identifican 20 posibles elementos respuesta: **FOXA2** (*Forkhead box A2; factor cabeza de horquilla A2*), **SMC1A** (*structural maintenance of chromosomes 1A; mantenimiento estructural de cromosomas 1A*), **CTCF**, **CTCFL** (*CCCTC-binding factor like; factor tipo CTCF*), **SMC3** (*structural maintenance of chromosomes 3; mantenimiento estructural de cromosomas 3*), **ZNF654** (*zinc finger protein 654; proteína dedos de zinc 654*), **RAD21** (*Double-strand-break repair protein, proteína de reparación del daño a cadena doble*), **STAG1** (*stromal antigen 1; antígeno estromal 1*), **ZNF766** (*zinc finger protein 766; proteína dedos de zinc 766*), **FAIRE** (*Formaldehyde-Assisted Isolation of Regulatory Elements; Elementos regulatorios aislados con ayuda de formaldehído*), **ARID3A** (*AT-rich interactive domain-containing protein 3A; dominio de interacción rico en AT de la proteína 3A*), **SMAD1** (*SMAD family member 1; miembro de la familia SMAD 1*), **SRF** (*serum response factor; factor de respuesta sérica*), **RCOR1** (*REST corepressor 1; correpresor REST 1*), **RFX5** (*Regulatory factor X-5; factor regulatorio X-5*), **ARID5A** (*AT-Rich Interaction Domain 5A; dominio de interacción rico en AT de la proteína 3A*), **GATA** (*GATA-binding factor; factor de unión a la secuencia GATA*), **NF-κB** (*nuclear factor kappa-light-chain-enhancer of activated B cells, factor nuclear potenciador de las cadenas ligeras kappa de las células B activadas*), **PAX5** (*Paired box 5; caja emparejada 5*), **TFIIA** (*General Transcription Factors IIA, factor de transcripción general IIA*) (Figura 8.10B).

La V9 afectaría de manera directa a 3 elementos (RUNX1, NCX y SOX), mientras que V15 por lo menos a 8 elementos (RFX5, ARID5A, GATA, NF-κB, PAX5, TFIIA, CTCF Y CTCFL).

Este análisis propone que estas variantes también podrían tener un efecto en el control transcripcional del *locus* a través de modificar la unión de factores de transcripción reclutados en la región.

A)



B)

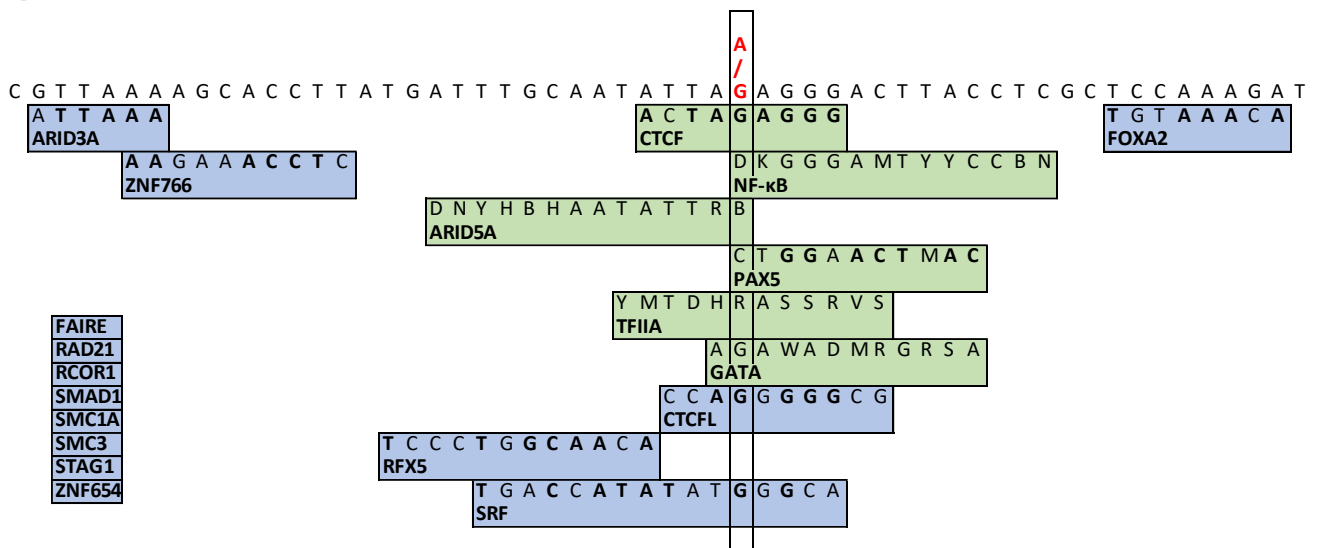


Figura 8.10. **Esquema de la posición de elementos de respuesta a diferentes factores de transcripción dentro del *loci* de las variantes V9 y V15.** La caja vertical resalta la posición del nucleótido que difiere entre ambos alelos. A) *Locus* de la variante V9, B) *locus* de la variante V15. Caja horizontal azul: predicción del portal *Toolkit for Cistrome Data Browser*. Caja horizontal verde: predicción del portal *Haploreg v4.1*. Recuadros a la izquierda: factores que no cuentan con secuencia consenso para el motivo respuesta. Nomenclatura de las bases con ambigüedad en la secuencia de los elementos respuesta: R: base púrica; Y: base pirimídica; S: G o C; W: A o T; K: G o T; M: A o C; B: C, G o T; D: A, G o T; H: A, C o T; V: A, C o G; N: cualquier base.

## 9. DISCUSIÓN

En este trabajo evaluamos las posibles implicaciones funcionales de distintas variantes genéticas identificadas en el gen *LRPPRC* asociadas a un riesgo incrementado para desarrollar hipertrigliceridemia severa familiar (HTGSF) en familias mexicanas (manuscrito en preparación). Tratándose de variantes intrónicas en desequilibrio de ligamiento, decidimos en primera instancia buscar mediante secuenciación Sanger, posibles mutaciones o variantes funcionales dentro de los exones y las uniones intrón-exón del gen *LRPPRC*, que pudieran ser parte del haplotipo de riesgo inicialmente identificado. Adicionalmente, analizamos posibles diferencias en la expresión del gen *LRPPRC* así como de distintos genes mitocondriales a partir de biopsias hepáticas de sujetos portadores y no portadores del haplotipo de riesgo. Analizamos también la presencia de posibles RNA mensajeros alternativos de *LRPPRC* entre portadores y no portadores. Por último, mediante estrategias bioinformáticas analizamos posibles interacciones a distancia y motivos de unión para distintos factores de transcripción presentes dentro del *locus* de *LRPPRC*.

Considerando que las variantes originalmente identificadas (haplotipo de riesgo conformado por las variantes rs7593842, rs58811869 y rs72877186) se encuentran dentro de regiones intrónicas de *LRPPRC*, se emplearon distintas estrategias tanto bioinformáticas como experimentales para identificar posibles variantes causales dentro del *locus*. A continuación se discuten los resultados obtenidos de las distintas estrategias así como las potenciales limitaciones de cada una de ellas:

### **Variantes identificadas mediante secuenciación de los exones del gen *LRPPRC*.**

En la Tabla 8.1 se muestran las variantes identificadas mediante secuenciación de los 38 exones del gen *LRPPRC* en los sujetos portadores de las variantes de riesgo. La variante encontrada en el exón 2 (rs6741066) en los portadores del haplotipo de riesgo es una variante silenciosa/sinónima. Por su parte, la variante del exón 30 solo se identificó en uno de los portadores del haplotipo de riesgo, portador 3 (rs769476622). Si bien esta variante resulta en el cambio de aminoácido (*missense*), mediante herramientas bioinformáticas (*PolyPhen*) se predice como un cambio benigno. Por lo tanto, no identificamos variantes tipo *missense* dentro de los exones; identificamos sin embargo variantes intrónicas adicionales. De estos resultados, concluimos que la estructura primaria de la proteína no se encuentra afectada y por lo tanto es posible que la alteración de la expresión tendría que ver con algún aspecto regulatorio, similar a lo observado en el modelo murino *knock-down* a *LRPPRC*<sup>55</sup>. De manera interesante Rakovic y colaboradores realizaron un análisis del mismo gen en pacientes con enfermedad de Parkinson. Este grupo secuenció las regiones exónicas y las uniones intrón-exon de *LRPPRC*. Identificaron 4 cambios sinónimos en la región codificante y 17 variantes intrónicas. De las 21 variantes identificadas en el estudio de Rakovic y colaboradores, en nuestro estudio identificamos 10 (1 exónica y 9 intrónicas), sin embargo, muchas eran compartidas entre portadores y no portadores del haplotipo de riesgo (Tabla 8.3).



Mediante el portal *Mutation Taster* analizamos si las variantes identificadas ya estaban reportadas en distintas bases de datos y si contaban con un identificador rs. En este portal también se reportaba si las distintas variantes podrían afectar el proceso de empalme del RNA y por lo tanto resultar en el *skipping* de secuencias exónicas o la formación de nuevos sitios donadores/aceptores. La Tabla 8.2 muestra la recopilación de información de 3 portales bioinformáticos utilizados y las diferentes predicciones sobre las implicaciones de la presencia de las distintas variantes. Esto en conjunto nos permitió seleccionar algunas variantes para ensayos funcionales.

La Tabla 8.3 muestra los resultados obtenidos de la secuenciación dirigida a buscar la presencia de las variantes seleccionadas en los sujetos no portadores. Es notorio de mención que la variante rs17031786, ubicada en el intrón 6, presenta el mayor porcentaje de probabilidad de afectación por el portal *Spliceman Web Server* (96%; Tabla 8.2), además de información en conjunto con los demás portales para que pudiera considerarse como una variante con probable afectación en el empalme del RNA, pero una vez que se identificó que la V9 no estaba presentes en los no portadores y la V15 tiene una baja frecuencia en no portadores, se decidió darle prioridad a estas dos variantes y continuar con su análisis funcional, dejando la secuenciación de rs17031786 en no portadores como perspectiva para un trabajo futuro, así como las muestras faltantes de la V3. Las otras variantes identificadas (V20, V22, V28) se descartaron al estar presentes en no portadores. En la sección 8. Búsqueda de nuevas variantes en los exones del gen *LRPPRC* se describen estos análisis.

En los estudios con variantes genéticas, un dato de vital importancia es la prevalencia de las mismas en diferentes tipos de poblaciones. Nuestra población mexicana es considerada mestizo-mexicana, ya que es resultado de una combinación reciente de población europea, nativa americana y en menor medida, africana. Se estima una proporción de la ancestría aproximadamente del 52% europea, 45% nativa americana (43% nativo norteamericano y 2% nativo sudamericano) y 4% africana<sup>33</sup>. Esta proporción es relevante al compararla con otras poblaciones porque existen diferencias incluso entre poblaciones latinas, como por ejemplo, en la población colombiana la proporción de ancestría son 71% europea, 19% nativo americana y 10% africana.

Estudios de asociación genética en poblaciones europeas o caucásicas no son completamente reproducibles en población mestiza-mexicana; variantes identificadas no están presentes en la población mestiza y/u otras variantes identificadas tienen una mayor asociación a la condición de estudio<sup>10,31</sup>. En el caso de las variantes de nuestro estudio, en la Tabla 9.1 se pueden observar esas diferencias en la prevalencia entre poblaciones. Se observa que unas variantes no están presentes en población asiática, pero sí en población europea y americana en proporciones similares (rs6721144, rs7593842, rs17031786, rs58811869, rs72877186, rs79868165, rs368657323, rs769476622 y rs4263155); reflejando el mestizaje y proporción de ancestría de la población mexicana.

No. en NCBI (alelo)	AFR	AMR (MXL)	EAS	EUR	SAS
rs6741066 (C)	0.320	0.445(0.461)	0.963	0.258	0.533
rs6721144 (C)	0.365	0.154(0.070)	0.000	0.171	0.025
rs17031786 (G)	0.368	0.156(0.070)	0.000	0.172	0.022
*rs7593842 (C)	0.367	0.154(0.070)	0.000	0.172	0.022
rs10190161 (C)	0.748	0.569(0.570)	0.045	0.748	0.487
*rs58811869 (C)	0.352	0.144(0.078)	0.000	0.153	0.022
*rs72877186 (C)	0.365	0.156(0.078)	0.000	0.172	0.026
rs79868165 (T)	0.000	0.003(0.016)	0.000	0.013	0.002
rs7594526 (T)	0.346	0.411(0.477)	0.043	0.586	0.460
rs28394191 (C)	0.552	0.585(0.523)	0.963	0.413	0.533
+rs368657323 (A)	0.000	5.812e-05(-)	0.000	8.846e-05	0.000
rs7568481 (T)	0.757	0.592(0.555)	0.960	0.405	0.502
rs4952691 (A)	0.900	0.644(0.586)	0.966	0.504	0.628
+rs769476622 (T)	0.000	5.784e-05(-)	0.000	2.644e-05	0.000
rs2955280 (T)	0.897	0.640(0.578)	0.956	0.500	0.642
rs4263155 (T)	0.373	0.157(0.086)	0.000	0.173	0.026

Tabla 9.1. **Prevalencia de las variantes identificadas en las secciones 8. Búsqueda de nuevas variantes en los exones del gen *LRPPRC* y 8. Búsqueda de variantes en la región de interacción dentro del intrón 23 de *LRPPRC* en diferentes poblaciones.** Valores obtenidos del estudio *1000 Genomes*, excepto + (Valores obtenidos del estudio *gnomAD genomes r3.0*). \*: Variantes pertenecientes al haplotipo de riesgo. AFR: población africana; AMR: población americana; MXL: población con ancestría mexicana en Los Ángeles, California, EUA; EAS: población del este de Asia, EUR: población europea, SAS: población del sur de Asia.

La presencia de mutaciones y las marcadas diferencias en las frecuencias alélicas se han asociado al mestizaje de cada población, ya que el proceso dinámico de la combinación de proporciones de componentes ancestrales, las combinaciones entre diferentes poblaciones y el paso del tiempo resulta en una organización específica a través del genoma; derivando en patrones de desequilibrio de ligamiento para cada población. Este patrón también es denominado como desequilibrio de ligamiento por mezcla (*ALD, Admixture Linkage Disequilibrium*), que consiste en grandes bloques de haplotipos y asociaciones no aleatorias a lo largo del cromosoma, pero con longitudes de los bloques diferenciadas entre población<sup>89</sup>.

En el trabajo de Weissglas-Volkov y colaboradores se llegó a la conclusión de que incluso si una variante tiene un efecto en todos los grupos ancestrales, esta podría tener una mayor prevalencia en una población en específico; sugiriendo que dependiendo de la frecuencia, la contribución de cada *locus* a la susceptibilidad a la condición de estudio pudiera variar entre poblaciones, haciendo diferente la asociación clínica. Al tener en cuenta estas diferencias en los estudios de asociación genética que no permiten extrapolar resultados a diferentes poblaciones, es necesario un estudio con integrantes de la población de interés y un tamaño de muestra (N) elevado para definir la frecuencia de mutaciones patógenas o para discernir variantes comunes exclusivas o más prevalentes en la población, revelar variantes que son nuevas, redefinir patrones en desequilibrio de ligamiento, prevalencia de enfermedades e interacciones gen-ambiente y sobretodo, una probable relevancia clínica.

## **Análisis de las posibles isoformas identificadas de *LRPPRC*.**

En el caso de la población mexicana y poblaciones latinas, uno de los estudios más importantes de asociación genética a nivel genoma completo para la identificación de variantes asociadas con enfermedades (como Diabetes) es el realizado por el Consorcio genético de Diabetes tipo 2 de la iniciativa Slim en Medicina Genómica para las Américas (*The Slim Initiative in Genomic Medicine for the Americas (SIGMA) Type 2 Diabetes Genetics Consortium*). En este estudio se identificó un haplotipo de variantes en el gen *SLC16A11* con una asociación a la Diabetes tipo 2 con significancia *genome-wide*<sup>90</sup>. Este haplotipo tiene una frecuencia de aproximadamente 50% en Nativos Americanos, 10% en Asiáticos del Este, pero muy rara en Europeos y Africanos. Además de esa importante asociación a una condición metabólica de gran importancia, es el estudio de asociación genética con la muestra de estudio más grande con mexicanos, haciéndolo una excelente base de datos de variantes comunes en mexicanos. Estos datos son fácilmente extrapolables para estudios genéticos en los que se tengan sujetos con ancestría mestizo-mexicana y poder dilucidar la frecuencia de variantes de estudio.

Tomando en cuenta la base de datos SIGMA, se consideró que la frecuencia del haplotipo de riesgo es relativamente alta en la población mestiza-mexicana (Tabla 9.2), por lo que fue posible identificar individuos portadores y no portadores de los cuales contábamos con biopsia hepática (individuos no pertenecientes al estudio de descubrimiento, pero sí a un estudio de Obesidad Mórbida en cual se justificó la toma de biopsia hepática).

<b>No. en NCBI</b>	<b>Intrón</b>	<b>Alelo</b>	<b>Prevalencia en SIGMA</b>
<b>rs72877186</b>	9	C	0.1162
<b>rs58811869</b>	15	C	0.1028
<b>rs7593842</b>	17	C	0.1194

Tabla 9.2. **Frecuencia de las variantes pertenecientes al haplotipo de riesgo en el estudio SIGMA.**

De esta forma, se seleccionaron muestras de portadores y no portadores de acuerdo al genotipo a *LRPPRC* considerando al mismo tiempo aparear por edad y sexo. Se obtuvieron 3 muestras de pacientes portadores (2 homocigotas y 1 heterocigota).

En el estudio de Gaweda-Walerych y colaboradores se demostró que la presencia de la variante rs7593842 (identificada como V9 en este documento) tiene implicación en el proceso de empalme del RNA al promover la formación de una isoforma de mRNA que presenta el *skipping* del exón 9 del transcrito de *LRPPRC*. Mediante un ensayo de mini-gen de la región de unión del intrón-exón, identificaron que esta isoforma de *LRPPRC* está presente en modelo de línea celular y cuando está presente la variante, hay una mayor abundancia de esta isoforma alternativa. También parecería que la presencia de esta isoforma es tejido específico, ya que al usar 2 diferentes líneas celulares humanas para sobre-expresar estas construcciones, sólo estuvo presente en las células HeLa (línea de cáncer cervicouterino utilizada con alta frecuencia en experimentos *in-vitro*) y Hek293T (línea de células embrionarias de riñón utilizada en experimentos de sobre-expresión génica). En el presente trabajo se utilizaron las mismas dos líneas celulares para intentar

identificar las isoformas reportadas por Gaweda-Walerych y colaboradores de manera WT y poder tomarlas como un control positivo, aunque no nos fue posible evidenciar la isoforma reportada en condiciones de amplificación y uso de oligonucleótidos diseñados (Figura suplementaria 1.2). Es digno de mención que estas dos líneas celulares no poseen el contexto metabólico del hígado, órgano del que provienen las muestras analizadas y nuestro interés en evidenciarla en ese tipo de tejido, pero se consideró pertinente utilizarlas para tratar de reproducir lo identificado por Gaweda-Walerych y colaboradores al intentar identificar posibles isoformas de *LRPPRC* con nuestras condiciones de amplificación del cDNA de *LRPPRC*.

En el presente trabajo probamos diferentes métodos de síntesis de cDNA y pares de oligonucleótidos para amplificar selectivamente las regiones exónicas aledañas, con la finalidad de identificar alguna isoforma que pudiera estar presente en el tejido hepático y enriquecerla con alguna de las condiciones utilizadas. Se utilizó además RNA de HepG2 como control.

En las Figuras 8.1A, 8.1B y 8.2A se muestran los amplificadores obtenidos del cDNA sintetizado mediante oligonucleótidos aleatorios. En la Figura 8.2A se muestra un amplificador alternativo de mayor tamaño que presenta secuencia del gen *LRPPRC*. Este amplificador de aproximadamente 320 pb, termina abruptamente después de los primeros 30 nucleótidos del exón 10. Si este transcrito alternativo llega a traducirse, daría como resultado a una proteína trunca conteniendo los primeros 399 aa de *LRPPRC* (*LRPPRC* completa tiene 1394 aa), con lo cual se perderían los motivos PPR que tienen unión al mRNA mitocondrial y de esta forma podría alterar su función de brindar estabilidad a los mtRNAs. Al conservar los primeros aminoácidos, se conserva la secuencia blanco para transporte a la mitocondria, por lo que esta proteína trunca podría llegar a la mitocondria y afectar la función óptima de la proteína completa. Es notorio mencionar una isoforma de *LRPPRC* reportada en una versión anterior del portal *Ensembl* que tiene una secuencia que termina en el exón 9 sin ningún codón de paro y codifica para una proteína de 359 aa. En una versión actualizada del portal *Ensembl*, la isoforma más pequeña (que comienza en el exón 1 como la proteína canónica) de *LRPPRC* termina hasta el exón 13, sí contiene un codón de paro y su proteína está formada por 531 aa; no se reporta mucha información de estas proteínas concerniente a su abundancia en tejidos en humano o funcionalidad. Al existir tantas isoformas reportadas para *LRPPRC* (Tabla suplementaria 1.1) que sí traducen proteínas, sería de interés evaluar la existencia de las proteínas propuestas tanto por Gaweda-Walerych y colaboradores como las presentadas en el presente trabajo, así como con el contexto hepático y posibles implicaciones por la presencia de la V9.

En la Figura 8.2B, se muestra el cDNA sintetizado utilizando un oligonucleótido con 18 timinas, para buscar hibridación con la cola de poli-A de los mRNA maduros. Además, se utilizó un oligonucleótido que hibrida en una región río arriba a los exones adyacentes a la variante V9, la isoforma previamente identificada (Figura 8.2.A). Sin embargo, con estas condiciones no era posible amplificar el transcrito canónico, ya que el tamaño esperado era de aproximadamente 4700 pb (secuencia codificante completa). Se obtuvieron múltiples amplificadores sin embargo, no identificó el amplificador esperado, por lo que la síntesis de cDNA con el oligonucleótido de 18 timinas no fue ilustrativo para nuestro

proyecto. Lo anterior podría deberse a inespecificidad en el alineamiento de los oligonucleótidos utilizados en la amplificación, ya que se utilizó una temperatura de alineamiento relativamente baja ( $T_a$ : 50 °C) e intermedia a las temperaturas *melting* de los oligonucleótidos ( $T_m$  oligonucleótido sentido: 55.6 °C y  $T_m$  oligonucleótido antisentido: 45.3 °C). Pudiera proponerse un nuevo diseño de oligonucleótidos así como una evaluación con incrementos de  $T_m$  para identificar una posible  $T_m$  óptima. Con esta condición no se pudo evidenciar ninguna de las isoformas de *LRPPRC* conocidas, a diferencia de las otras dos condiciones de síntesis de cDNA utilizadas.

La última condición utilizada para la síntesis de cDNA fue utilizando oligonucleótidos específicos para el transcrito canónico, además de amplificar una región mayor (del exón 6 al exón 17), bajo la premisa que un transcrito de mayor tamaño pudiera tener mayor estabilidad. Los amplificados obtenidos en las tres muestras utilizadas (HepG2, Heterocigoto y WT a las variantes de riesgo) se muestran en la Figura 8.2C y fueron identificados como *LRPPRC*. Estos transcritos muestran ausencia de secuencias exónicas (Figura suplementaria 1.6) y si llegaran a traducirse, se produciría una isoforma de la proteína; en el portal *Ensembl* se reportan isoformas con intrones retenidos (que no se llegan a traducir). La posible pérdida de secuencia codificante que identificamos no produce un cambio del marco de lectura ni un codón de paro prematuro, sin embargo, se pierde la secuencia del Asp249 a la Val387 de la proteína canónica. Este cambio identificado no afecta la región de transporte a la mitocondria (amino terminal) ni la región funcional (carboxilo terminal), pero sí podría afectar el óptimo plegamiento de la proteína y en consecuencia su funcionalidad. También en el portal *Ensembl* se reportan isoformas de proteínas con secuencias de aminoácidos faltantes con respecto a la proteína canónica de 1394 aa, pero ninguna con el cambio que nosotros proponemos, por lo que si la isoforma identificada llegara a traducirse, resultaría en una isoforma de la proteína *LRPPRC* aún no reportada.

De los amplificados alternativos identificados y secuenciados, ninguno es el reportado por Gaweda-Walerych y colaboradores, en el cual se reporta el *skipping* del exón 9. En el portal *Ensembl* no se reportan isoformas con *skipping* del exón 9, pero sí se reportan diferentes isoformas con alteraciones en los exones 19, 23 y 28 que son codificantes para proteínas, por lo que pudiera ser viable la traducción de estas dos isoformas propuestas; la propuesta en el presente trabajo y la propuesta por Gaweda-Walerych y colaboradores.

Estos cambios identificados en la secuencia de los transcritos amplificados fueron encontrados en las tres muestras utilizadas (portador heterocigoto, no portador y HepG2) y en los diferentes tamaños de amplificados (Figura 8.2.C). Aún con la diferencia de amplificados, se confirmó mediante secuenciación que los transcritos mostraban el mismo patrón en las secuencias (Figura suplementaria 1.6). No fue posible reproducir estas condiciones en el resto de las muestras de pacientes (2 homocigotos portadores y 8 no portadores, Figura suplementaria 1.5). Pudiera deberse a que la integridad del RNA no fuera la más óptima y sea necesario una nueva toma de muestra y nueva extracción de RNA; también existe la posibilidad que las dos muestras de pacientes analizadas donde fue identificada la isoforma sean atípicas y específicas para estas dos muestras; sería

necesario aumentar el número de muestras y sensibilidad de la técnica, como un *Northern-blot* o secuenciación de RNA (*RNAseq*).

Como perspectiva y considerando que el tejido hepático no es de fácil acceso, proponemos utilizar tejido adiposo de individuos portadores y no portadores del haplotipo de riesgo pertenecientes al estudio de familias con HTGSF, así como aumentar el número de muestras para poder evidenciar algún cambio en este tejido.

Aunque estos resultados son preliminares, proponemos que se promueva la síntesis de un mRNA alternativo en tejido hepático humano que de llegar a transcribirse produciría una proteína trunca; como propuesta se pudiera sobre-exresar el gen en líneas celulares para evaluar si se produce la proteína alternativa, así como evaluar si la presencia de la variante V9 tiene alguna afectación en este mRNA alternativo.

## **Estudios de niveles de expresión de genes en tejido hepático y línea celular humana HepG2**

### **Nivel de expresión para *LRPPRC*.**

Se ha demostrado mediante estudios de *Northern blot* que la expresión de *LRPPRC* es distinta entre tejidos. Los niveles de mRNA son más abundantes en algunos tejidos humanos entre los que se encuentran músculo esquelético, corazón y placenta<sup>91</sup>; aunque no siempre es lineal la abundancia de mRNA y proteína en todos los casos. Por ejemplo, se reportan niveles bajos de expresión de mRNA de *LRPPRC* en el hígado y el cerebro y estos son los tejidos mayormente afectados en el síndrome de Leigh con mutaciones en el gen *LRPPRC* y alteraciones en su proteína.

Es importante mencionar también que en pacientes con síndrome de Leigh, se demostró mediante inmunoensayos de *Western-blot* a partir de fibroblastos de piel y mitocondrias de tejido hepático, que los niveles de la proteína de *LRPPRC* están disminuidos. El cambio genético causante de la variante franco-canadiense de esta condición se encuentra en el exón 9 (C<sup>1119</sup>→T), a 123 nucleótidos de la variante V9 (rs7593842). Sin embargo, en el trabajo no evaluamos posibles cambios en los niveles de la proteína entre portadores y no portadores del haplotipo de riesgo, ni analizamos la presencia de una posible forma más corta o faltante de aminoácidos de la proteína mediante *Western-blot* en tejido hepático.

En la Figura 8.4 se muestra el análisis preliminar de expresión de *LRPPRC* realizado en muestras hepáticas de portadores y no portadores del haplotipo de riesgo. Sin embargo, es importante mencionar que no fue posible analizar un número suficiente de muestras de individuos portadores y por lo tanto no es posible concluir si existen o no diferencias en la expresión del gen entre estos grupos.

### **Nivel de expresión de distintos genes mitocondriales**

Una evaluación indirecta de la afectación en la función de *LRPPRC* es a través de demostrar cambios en los niveles de expresión de distintos genes mitocondriales, de

acuerdo a la función conocida del gen *LRPPRC*<sup>51-59</sup>. Se ha propuesto que *LRPPRC* participa en la estabilidad de los genes mitocondriales, los cuales codifican para proteínas de la cadena respiratoria, la cual está compuesta por 4 complejos: 3 bombas de protones y un enlace físico con el ciclo del ácido cítrico. En estos complejos es donde se realiza la oxidación completa de ácidos grasos mediante la fosforilación oxidativa.

En la Figura 8.6 se muestran los resultados obtenidos del ensayo de expresión genética de los genes mitocondriales de sujetos portadores y no portadores del haplotipo de riesgo. Para las evaluaciones estadísticas se tomó el cálculo de la media de los valores obtenidos de  $2^{-\Delta\Delta Ct}$  en cada gen de los no portadores y se normalizó a 1 para comparar con los valores  $2^{-\Delta\Delta Ct}$  obtenidos de los individuos portadores, aunque es de mencionarse que existe variabilidad entre los valores de expresión (Ct,  $\Delta Ct$ ) de los genes mitocondriales en los no portadores; hay poca variabilidad entre los valores de Ct de los triplicados de cada muestra para cada gen (Tabla suplementaria 1.2), pero sí en las medias de los valores  $\Delta Ct$  de los 9 no portadores (Tabla suplementaria 1.3).

En la Figura 8.6A y Figura 8.6C no observamos diferencias significativas, sin embargo cuando se comparan los resultados obtenidos de un portador homocigoto (Figura 8.6B) se observa la disminución en la expresión de los mRNA mitocondriales respecto a la normalización de los no portadores. Esta observación es similar a lo reportado para el modelo murino *knock-down* de *LRPPRC* donde la expresión de los genes mitocondriales involucrados en la fosforilación oxidativa se encuentra abatida<sup>59</sup>. Sin embargo, nuestros resultados corresponden a un solo portador homocigoto y es notoria la dispersión de los datos, representada en la desviación estándar en la Figura 8.6D. El único valor-p que resulta limítrofe a ser significativo es el encontrado para el gen *COX1* con las medias de los valores  $2^{-\Delta\Delta Ct}$  de portadores y no portadores y de manera interesante, mutaciones en este gen son las causantes del síndrome de Leigh (mutaciones en *LRPPRC* son las causantes de la variante franco-canadiense), por lo que afectaciones en la expresión de este gen o en cualquiera de los codificantes para la cadena respiratoria, resulta en patología. Pudiera evaluarse de manera funcional la eficiencia de la cadena respiratoria así como la oxidación de ácidos grasos en pacientes portadores en tejido hepático para determinar afectaciones debido a la presencia de las variantes de riesgo.

Es digno de mención el ambiente adverso que existiría en el tejido hepático de los pacientes con obesidad mórbida donadores de la biopsia de hígado. La obesidad es un estado crónico definido como exceso de grasa corporal perjudicial para la salud en la cual hay alteración en el metabolismo debido al exceso de nutrientes en el organismo. En este ambiente de exceso de nutrientes, la grasa depositada intracelularmente y secreciones producidas por la lipotoxicidad en tejido adiposo inducen a la resistencia a la insulina<sup>92</sup>. El exceso de depósito en el tejido adiposo visceral comúnmente es acompañado por acumulación de lípidos en tejidos que normalmente no acumulan grasa como el hígado, el corazón, riñones y páncreas y músculo esquelético<sup>93</sup>. La obesidad abdominal acentúa la condición de alta afluencia de ácidos grasos libres, citoquinas y hormonas hacia el hígado creando una distorsión en el metabolismo hepático aumentando la producción de VLDL y apoB<sup>91</sup>, además de crear lesiones de esteatosis y esteatohepatitis no alcohólica (forma más severa del hígado graso no alcohólico, caracterizada por necroinflamación hepática,

degeneración balonzante de los hepatocitos y fibrosis; pudiendo evolucionar a cirrosis hepática<sup>94</sup>). En la obesidad y exceso de nutrientes, la actividad de la fosforilación oxidativa está disminuida resultando en oxidación incompleta de ácidos grasos en los hepatocitos y en su acumulación que causa rompimiento de la membrana hepatocelular activando vías de inflamación<sup>52,95</sup>.

En un modelo murino se comprobó un aumento en la degradación de subunidades de la fosforilación oxidativa en el hígado que presentaba HGNA<sup>96</sup>, así como en el caso contrario, una mayor expresión de genes mitocondriales, induce una oxidación completa y una protección contra la HGNA<sup>52</sup>. Nakamuta y colaboradores observaron que en pacientes obesos con HGNA los procesos de oxidación de ácidos grasos y eliminación de especies reactivas del oxígeno eran insuficientes lo que resultó en un mayor daño hepático, esteatosis y resistencia a la insulina<sup>97</sup>.

La elección del tejido hepático es crucial para nuestro proyecto por su rol en el metabolismo de lípidos, pero al ser un tejido de difícil acceso, se realizaron estos experimentos preliminares con biopsias obtenidas de pacientes a los que se les realizaron una intervención quirúrgica que ameritaba la toma de muestra. Proponemos por lo tanto, aumentar el número de muestras y utilizar muestras de fácil acceso como sangre periférica de sujetos portadores y no portadores del haplotipo de riesgo, para poder confirmar lo observado en el portador homocigoto de la Figura 8.6D. Otro tejido de relativo fácil acceso es el tejido adiposo subcutáneo, de particular interés por su participación en la acumulación de lípidos; con una particular atención de que los miembros participantes portadores y no portadores de las variantes de riesgo no presenten antecedentes de obesidad, ya que adipocitos provenientes de individuos obesos han mostrado una formación inadecuada de proteínas de la cadena respiratoria y alteraciones en la función de la mitocondria<sup>98</sup>.

La HTG no es una enfermedad que sea exclusiva de pacientes obesos, por lo que es factible captar portadores sin estos antecedentes y tratar de reducir variables que puedan afectar la expresión de genes para que la variable a evaluar sea la expresión de *LRPPRC* y su afectación a genes mitocondriales y la evaluación de afectaciones con la presencia de las variantes de riesgo.

### **Búsqueda de interacciones a larga distancia y co-localización con estados de la cromatina**

Un último enfoque que se tomó para dilucidar una posible regulación génica para nuestro gen de interés, fue la búsqueda bioinformática de regulación epigenética. La regulación epigenética a larga distancia mediada por las estructuras TAD resulta esencial para establecer una regulación genética apropiada. Los TAD contribuyen a mantener dentro de los límites de su estructura interacciones regulatorias como reclutamiento de factores de transcripción, marcas en la cromatina y acercamiento espacial entre promotores y *enhancers*<sup>99</sup>; genes que se encuentran dentro del mismo TAD responden de manera similar a estímulos transcripcionales específicos<sup>99</sup>. La regulación dentro de cada TAD



también es específica, ya que no hay interacción con otros TADs vecinos, así como específica para patrones de TAD ya que son particulares para cada tejido.

En este trabajo presentamos una acumulación de información bioinformática que nos pudiera ilustrar como sería la organización para el *locus* de *LRPPRC* en un contexto hepático. Se observaron distintos genes y elementos de regulación presentes en el *locus* conteniendo a *LRPPRC*. Las funciones de estos genes son muy variadas, por ejemplo *PLEKHH2* codifica para una proteína con alta expresión en podocitos en el glomérulo del riñón y está implicada en estructuras de adhesión que unen a los podocitos con la membrana basal glomerular<sup>100</sup>, *DYNC2LI1* está implicado en ciliogénesis y es un componente del complejo del motor de microtubulo dineína 2<sup>101</sup>, *PPM1B* es un miembro de la familia de fosfatasa 2C y juega un rol en la regulación negativa de p53 y NF-κB posiblemente atenuando la expresión génica para la senescencia celular<sup>102</sup>, *SLC3A1* codifica para un glicoproteína de membrana tipo II que es uno de los componentes de transporte de aminoácidos en el túbulo renal y tracto intestinal<sup>103</sup>, *PREPL* una oligopeptidasa de serina involucrada en el llenado de acetilcolina en las vesículas sinápticas<sup>104</sup>, *CAMKMT* codifica a una proteína metiltransferasa clase I involucrada en la formación de la trimetil-lisina en la calmodulina<sup>105</sup> y *SIX3* es un factor de transcripción esencial para el desarrollo del sistema visual en vertebrados<sup>106</sup>.

De estos genes presentes en el *locus* de *LRPPRC*, llama la atención los genes *ABCG5* y *ABCG8*, los cuales están implicados en el metabolismo de lípidos al ser transportadores tipo ABC (*ATP binding cassette*). Estos genes codifican para dos diferentes proteínas, *ABCG5* codifica para Esterolina-1 y *ABCG8* para Esterolina-2, las cuales forman un heterodímero funcional. El heterodímero *ABCG5/8* promueve la eliminación de C y esteroides de plantas en el hígado y en el intestino<sup>107</sup>. En el acercamiento de la Figura 9.8.I se muestran 4 regiones de interacción con una co-localización de señales en todas las marcas de la cromatina seleccionadas y esto resultaría muy significativo ya que el acercamiento de forma tridimensional en la cromatina podría estar regulando la expresión de distintos genes. Una de las tres regiones identificadas como *super-enhancer* con interacción a la región promotora de *LRPPRC* es la que se encuentra en el *locus* de la región promotora de *ABCG8*. Esta interacción podría implicar una posible regulación conjunta entre estos genes (*LRPPRC-ABCG5-ABCG8*) y tener un impacto con el metabolismo de lípidos.

La siguiente región interesante es la encontrada dentro de *LRPPRC* en el intrón 23 (Figura 9.8.I) La predicción de las interacciones abarca una región de 20 kb (chr2:44158145-44178144, hg19), desde el intrón 14 hasta el intrón 23, la cual contendría a dos de las tres variantes inicialmente asociadas a HTG (rs58811869/V15 en el intrón 15 y rs72877186 en el intrón 17) y 2 de las variantes identificadas en la secuenciación de los exones de *LRPPRC* (rs7594526/V20 en el intrón 20 y rs28394191/V22 en el intrón 22; Tabla 8.1). La región de co-localización de las interacciones y el pico de señal en las marcas de *super-enhancer* es más pequeña y abarca aproximadamente 5 kb (4720 pb; chr2:44167190-44171910, hg19), por lo que la secuenciación realizada en esta región tuvo como objeto identificar alguna otra variante que pudiera ser funcional al interaccionar con la región promotora de *LRPPRC* y así afectar su expresión. También es

importante mencionar que en esta región de 5 kb se encuentra la señal más alta para unión de DNAsa, lo que se refleja en una mayor accesibilidad en la cromatina y en una posible unión de factores de transcripción. Las dos variantes identificadas (rs4263155/VI1 y rs4600713/VI2 en el intrón 23; Tabla 8.5) podrían estar implicadas en la regulación de genes de la región de interacción. En el portal *HaploReg v4.1* se reporta que la variante VI1 se encuentra en LD con la VI2 ( $r^2= 1$ ) y con las 3 variantes de riesgo ( $r^2= 0.99$  con rs72877186;  $r^2= 0.87$  con rs58811869;  $r^2= 0.95$  con rs7593842); además de que VI1 y VI2 fueron compartidas por los 4 portadores (Tabla 8.5) y no fueron identificadas en no portadores (Tabla 8.6).

Aunque es notorio la cantidad de interacciones del promotor de *LRPPRC*, el mapeo con las marcas en la cromatina es lo que hace que las interacciones tengan un potencial regulatorio, no sólo interacciones debido a la estructura de la cromatina. Este cúmulo de información muestra una región de alta expresión y un potencial regulatorio de genes a distancia de la región promotora de *LRPPRC* con la región intergénica río abajo de *LRPPRC*, la región *ABCG5/ABCG8* y la que se encuentra dentro del mismo gen.

Es necesario aún entender la posible implicación de los genes *ABCG5/8*, así como la región localizada en la región intergénica de interacción río abajo de *LRPPRC*. Se pudiera evaluar la expresión de *ABCG5/8* mediante RT-PCR en portadores y no portadores por posibles alteraciones y una posible expresión conjunta con *LRPPRC* considerando su interacción con su promotor. Para la región intergénica se pudieran evaluar posibles cambios epigenéticos como marcas en la cromatina y unión de factores de transcripción que pudieran estar en la región, así como posibles cambios por la presencia de las variantes de riesgo o el ambiente adverso en la condición de EHNA.

También sería importante comprobar experimentalmente estas interacciones así como evaluarlas en el contexto hepático sano y de EHNA de sujetos portadores y no portadores del haplotipo de riesgo, mediante un ensayo 3C (*chromosome conformation capture*, captura de la conformación cromosomal) en el cual se evalúa la frecuencia de contacto entre segmentos distantes de DNA de poblaciones celulares mediante fijación de regiones de contacto de DNA, digestión y ligación intramolecular, seguido de una amplificación de PCR punto final selectiva de la ligación y subsecuente gel de electroforesis<sup>108</sup>.

### **Análisis bioinformático de implicaciones a factores de transcripción por las variantes identificadas en la región del intrón 23 de *LRPPRC* y dos variantes del haplotipo de riesgo.**

Otra implicación funcional de variantes en el genoma es su afectación en la regulación epigenética, ya que en algunos casos alelo-específicos, la presencia de variantes cambia la secuencia de los elementos respuesta haciendo que no empaten con la secuencia consenso del factor de transcripción, provocando que la afinidad de unión del factor de transcripción aumente o disminuya por el *locus*, así como también eliminar la unión o generar una nueva; ocasionando que la región ya no sea regulada por el factor de transcripción determinado y se provoque una expresión génica alternativa, fenotipo y/o

susceptibilidad a exposición ambiental diferente a la original<sup>109</sup>. Estos factores de transcripción son proteínas o complejo de proteínas que reconocen pequeñas secuencias específicas de DNA (motivo de unión) a la cual acceden y se unen para actuar en el genoma. Dependiendo del factor de transcripción, su unión puede regular su gen blanco a la alza o a la baja. Pueden reclutar enzimas que catalizan la acetilación o desacetilación de histonas, a proteínas co-activadoras o co-represoras, funcionar como *insulators* o ayudar en la función de *enhancer*.

Al ser un elemento de expresión génica, los diferentes factores de transcripción pueden ser tejido específico o tener función en momentos específicos del desarrollo. De los factores identificados en nuestros análisis (Figuras 8.9 y 8.10) será importante determinar cuáles serían relevantes en nuestro estudio por su función y el tejido donde se exprese, así como los *loci* en donde cambien el motivo respuesta y pudieran tener afectación en la unión; basándose más en lo que ocurre *in-situ*, así como la influencia por el ambiente adverso de HTG y no tanto en las secuencias consenso, ya que estas están determinadas por la mayor probabilidad de unión a una base.

En nuestro estudio, iniciamos analizando mediante herramientas bioinformáticas si las variantes rs4263155 (VI1) y rs4600713 (VI2) identificadas en la región del intrón 23 reportadas en la Tabla 8.5 podrían representar variantes funcionales. Iniciamos por explorar posibles implicaciones en el proceso de empalme del RNA, similar al análisis que llevamos a cabo para las variantes intrónicas inicialmente identificadas en el gen *LRPPRC* (Tabla 8.2). En la Figura 8.8 reportamos la identificación de una región de alta accesibilidad en la cromatina por la marca de DNAsa, en esta región se encuentran variantes genéticas como parte de potenciales sitios de unión para distintos factores de transcripción.

Para el alineamiento de los motivos respuesta (Figura 8.9), se utilizaron datos de predicción basados en *ChIP-seq* para motivos de unión de factores de transcripción y datos de *ChIP-seq* para unión de proteínas en diferentes tejidos. Estos resultados aportan información sobre factores de transcripción que pudieran tener una influencia en la expresión de los genes y una posible regulación a la región que se unen o bien de genes a distancia. De los 16 motivos respuesta identificados para las 2 variantes VI1 y VI2, el único factor de transcripción que tiene datos experimentales de *ChIP-seq* directos provenientes del proyecto *ENCODE* es el factor USF1. Los datos existentes sobre los sitios de unión a USF1 se obtuvieron inicialmente de la línea celular H1-hESC, una línea de células troncales embrionarias humanas. Estas células son un modelo de ambiente celular pluripotencial, diferente al ambiente de una célula diferenciada como es el hepatocito, de interés particular en este trabajo. No obstante, los datos disponibles proveen evidencia experimental de unión de este factor de transcripción en la región que incluye la variante VI1. Sin embargo es importante demostrar experimentalmente esta interacción de USF1 y otros potenciales factores de transcripción a partir de tejido hepático de individuos portadores y no portadores del haplotipo de riesgo.

Entre otros factores que podrían ser de interés está CTCF, porque este factor tiene una implicación en la organización y estructura de la cromatina ya que promueve

interacciones a larga distancia al actuar como ancla entre las asas de la cromatina, además de delimitar las fronteras de los TADs. Adicionalmente CTCF tiene implicaciones en la regulación transcripcional al tener una función como *insulator* (complejos de DNA-proteína que actúan como inhibidores de la transcripción al prevenir interacciones con regiones adyacentes en el genoma, como un *enhancer* con un promotor<sup>110</sup>), además de que su unión es dependiente de la metilación en el DNA<sup>38</sup>. Además del alineamiento del elemento respuesta de CTCF con el *locus* de la variante VI1 (Figura 8.9A), también se confirma la presencia de este factor de transcripción en la co-localización de una señal (baja intensidad) de unión en el locus de la interacción intergénica dentro de *LRPPRC* (Figura 8.8I), por lo que si la variante VI1 afecta el elemento respuesta para la unión de CTCF podría existir una implicación en la arquitectura de la cromatina y posiblemente afectar la expresión de genes aledaños, ya sea *LRPPRC* o *ABCG5/8*. Se ha demostrado que si existe alguna delección en los límites de TADs o sub-TADs, crea nuevos dominios en la cromatina que podría causar nuevas interacciones entre genes y *enhancers* o activar zonas que fueron silenciadas, lo que podría llevar a expresión alterada de genes<sup>111</sup>.

Otro factor de transcripción que tiene conexión con metabolismo de lípidos es GCNF, también conocido como NR6A1 y que está codificado por el gen *NR6A1*. Un estudio demostró que NR6A1 regula el metabolismo de lípidos el ser un represor determinante en la lipogénesis y en la expresión del receptor de insulina en células HepG2<sup>112</sup>. En el modelo *knock-down* de NR6A1 se observa un incremento del contenido de lípidos en las células así como aumento en la expresión de los genes lipogénicos *DGAT2* (*diglyceride acyltransferase-2*; Acil-CoA:diacilglicerol aciltransferasa-2), *FASN* (*fatty acid synthase*; sintasa de ácido graso), *MCAD* (*medium-chain acylcoenzyme A dehydrogenase*; acil-coenzima A deshidrogenasa de cadena media), *ME1* (*malic enzyme 1*; enzima málica dependiente de NADP 1) y *MTTP* (*microsomal triglyceride transfer protein*; Proteína de transferencia de TG microsómica). En este modelo *knock-down* de NR6A1 también se observó aumento en la expresión del gen *PEPCK* (*phosphoenolpyruvate carboxykinase*; fosfoenolpiruvato-carboxiquinasa) que codifica para una enzima de la gluconeogénesis en el hígado, así como aumento de 18 veces del miRNA miR-205, el cual regula de manera indirecta la expresión del receptor de insulina<sup>112</sup>. Por lo que pudiera existir una regulación recíproca entre el factor nuclear NR6A1, el receptor de insulina en la membrana y miRN-205 para el control del metabolismo de lípidos en la línea celular HepG2. El enfoque del estudio era evaluar las implicaciones de anormalidades en el metabolismo de lípidos y su participación en cáncer, pero estas implicaciones podrían extrapolarse a nuestro modelo de estudio, ya que GCNF podría estar influyendo la función de miRNAs observados en la región (*MIR548AD* y *Mir\_548*; Figura 8.7) o bien, al tener interacción con la región promotora de *LRPPRC*, modificar su expresión. El cambio por la presencia de VI1 parecería que no tendría influencia, ya que la secuencia consenso del motivo respuesta no tenía unión en esta base. Aunque podría existir la posibilidad de unión en este *locus*, así como la posibilidad que la variante afecte la misma, ya que la variante se encuentra a la mitad del elemento respuesta de GCNF/NR6A1.

Por su parte, para el factor NFE2L1 existe información en la literatura donde parece jugar un papel distintos procesos metabólicos. La pérdida de expresión hepática de Nrf1 (otra manera para referirse a NFE2L1) se asocia con la acumulación de lípidos y cisteína, daño

hepatocelular y alteración en la composición de ácidos grasos<sup>113</sup>. El perfil del transcriptoma en el modelo murino *knock-down* al gen *Nrf1* muestra una disminución en la expresión de genes relacionados con el metabolismo de lípidos y reguladores transcripcionales. En otro trabajo, mediante estudios con un gen reportero y luciferasa, se identificaron a *Lipin1* y *Pgc-1β* como genes blanco de *Nrf1* sugiriendo que el papel en el metabolismo de lípidos es mediante el control de la expresión de reguladores transcripcionales para la expresión de genes codificantes de enzimas metabólicas<sup>114</sup>. En este mismo modelo se identificó que la cadena respiratoria mitocondrial se encuentra desacoplada como consecuencia de la deficiencia de *Nrf1*. Este factor es expresado de manera ubicua en distintos tejidos y líneas celulares, además se han descrito distintas isoformas y su expresión es inducida por señales de estrés y un amplio intervalo de estímulos<sup>115</sup>. La variante identificada VI2 podría afectar la unión de este factor de transcripción.

La literatura describe que uno de los factores presentes en el alineamiento (Figura 8.9), *USF1*, se ha relacionado con el riesgo al desarrollo de dislipidemias en diferentes grupos étnicos. Inicialmente se asoció a la región cromosómica 1q21-q23 en las poblaciones finlandesa, alemana, china y mexicana con Hiperlipidemia Familiar Combinada (HFC)<sup>116,117,118</sup>. Posteriormente la región de asociación se delimitó al gen *USF1*<sup>119</sup>. *USF1* codifica para una proteína de 43-kDa involucrada en metabolismo de glucosa y lípidos<sup>120</sup>. De los genes regulados por el heterodímero formado por *USF1* y *USF2* se encuentran *apoAV*, *apoC-III*, *ApoA-II*, *ApoE*, la lipasa sensible a hormonas, sintasa de ácidos grasos, acetilCoA carboxilasa, renina, angiotensinógeno, glucocinasa, receptor de glucagon, insulina, grelina, la desaturasa de ácidos grasos tipo 3 y la piruvato cinasa hepática<sup>121</sup>. También se asoció a *USF1* con la expresión del gen *FADS3*, el cual tiene asociación a nivel de genoma completo con niveles de TG<sup>122</sup> y también se asociaron variantes en este gen con niveles de TG en familias mexicanas con HFC<sup>123</sup>. La HFC también tiene afectaciones en los niveles de TG, además de tener un componente familiar pero se diferencia de la HTGSF por los niveles séricos elevados de ApoB y niveles anormales de C total sérico.

De los elementos respuesta identificados en el alineamiento, el factor *FOXA* es el único que se le ha relacionado con el tejido hepático. Este factor es esencial en el desarrollo del hígado durante el desarrollo embrionario ya que se ha propuesto que los factores *FOXA* compiten para abrir la cromatina compactada que contiene genes tejido hepático-específico y se demostró que los factores *Foxa1* y *Foxa2* son necesarios para la especificación hepática en el ratón<sup>124</sup>. Aunque también se demostró que tiene implicaciones en la vida adulta, ya que se identificó que las proteínas HNF3 (como también se les conoce a las proteínas *FOXA*) tienen un papel central en coordinar el control hormonal y metabólico de la expresión de genes regulados por insulina, glucagon o glucocorticoides; además de provocar la activación transcripcional de genes que responden a glucocorticoides<sup>125</sup>. Se han identificado mediante varias técnicas como EMSA, *ChIP*, co-transfecciones y modelos *knock-down* distintos genes que codifican para enzimas hepáticas y proteínas séricas que pudieran tener un papel en la respuesta hepática al ayuno, con sitios de unión para *Foxa* y *Foxo*<sup>126</sup>. Entre estos genes el que está relacionado con metabolismo de lípidos está LPL (Lipoproteína lipasa)<sup>127</sup>. De la misma manera que los factores anteriores, si la variante VI2 afecta la unión del factor, podría afectar su función.

Es importante señalar que la variante VI2 se encuentra en el 3er nucleótido del elemento respuesta para FOXA y aunque tiene libertad de unión de diferentes bases, el cambio no es una base púrica por lo que podría afectar la unión y de manera indirecta la expresión de los genes *LPL*, *LRPPRC* y/o *ABCG5/8*.

Al obtener esta información en la búsqueda bioinformática, se complementó para las dos variantes pertenecientes del haplotipo de riesgo que fueron exclusiva de portadores; V9 y V15. La variante V15 se encuentra dentro de la región que se identificó que interactúa con el promotor de *LRPPRC*, por lo que podría existir un reclutamiento de factores a su *locus* de la misma manera que VI1 y VI2. Algo muy notorio, es la gran acumulación de datos presente en el *locus* de la variante V15 (Figura 8.10.B). Algunos factores no cuentan con secuencia consenso, pero se reporta en el portal *Cistrome* que hay unión de múltiples proteínas en el *locus* (se definió la región de  $\pm 20$  pb a la variante).

De todos los factores identificados en el alineamiento esquematizado en la Figura 8.10, CTCF es el único que tiene datos provenientes de los portales *Haploreg* y *ENCODE* de unión de proteína ChIP-seq y unión en diferentes líneas celulares y tejidos. Para la variante V9, los factores identificados RUNX1, NCX (también conocido como TLX2<sup>128</sup>) y SOX están asociados al desarrollo embrionario y diferenciación de diferentes tejidos; en los 3 factores, la base WT de la V9 en el elemento respuesta es conservado y podría afectar la unión por la presencia de la variante. Podría evaluarse la presencia del factor SOX9, ya que este miembro de la familia SOX tiene influencia en células madre y progenitoras en hígado, intestino y páncreas adulto para producir hepatocitos, enterocitos y células acinares en situaciones de homeostasis o lesión. Podría existir alguna regulación en el *locus* en un contexto hepático; así como esa regulación alterada por la presencia de la variante.

El factor KLF9 tiene una implicación en el metabolismo, ya que es un factor inducible por glucocorticoides que incrementa la transcripción del receptor PGC1 $\alpha$ , resultando en gluconeogénesis<sup>129</sup>. KLFs regulan el uso y almacenamiento órgano-específico de nutrientes al influir en la homeostasis metabólica y hacer un ajuste fino de tránsito de nutrientes entre órganos en respuesta a cambios en el estado energético<sup>130</sup>. La pérdida de expresión hepática de KLF9 contribuye a la enfermedad de hígado graso no alcohólico (NAFLD) por la represión de genes de oxidación de ácidos grasos; se necesita un mayor estudio, pero pudiera parecerse al fenotipo identificado en el modelo murino deficiente de *LRPPRC*. La variante no se encuentra dentro del elemento respuesta a KLF9, pero se necesitaría evaluar experimentalmente si la presencia de V9 influye en la unión y expresión en esta región, así como evaluar si su función en esta región afecta el metabolismo de lípidos.

Así como en el alineamiento de VI2, en esta región también está presente un factor FOX (Figura 8.9.B), que ya se discutió que es necesario para el desarrollo embrionario hepático y que se expresen genes tejido hepático específicos. FOXA2 también es necesario para el desarrollo pancreático, así como para regular la expresión de múltiples factores de transcripción controlando el destino celular endócrino en el páncreas así como de la secreción de insulina<sup>131</sup>. Otro factor que también se identificó en la región de la VI1,

CTCF también está presente en este *locus*. En esta región se podría complementar la función de arquitectura en la cromatina por la presencia de los factores RAD21, SMAD1, SMC1A, SMC3 y STAG1. Estos últimos factores son subunidades del complejo de la cohesina, encargada de múltiples funciones dentro de la organización de la cromatina y expresión en el genoma, además de reparación en DNA dañado y división celular en la mitosis y meiosis<sup>132</sup>. Esta co-localización de elementos refuerza la idea de que esta región es imperativa para el mantenimiento de la estructura de la cromatina y pudiera coadyuvar en el acercamiento espacial que se necesita para la interacción que se identificó en la sección 8. Búsqueda bioinformática de interacciones a larga distancia y co-localización con estados de la cromatina (Figura 8.8.I). Si esta variante o la VI1 o la combinación de las dos (*Haploreg* reporta estas dos variantes en LD con una  $r^2 = 0.87$ ) cambia la afinidad por el sitio de unión para CTCF o alguna proteína estructural del complejo de la cohesina, podría existir una alteración en la estructura de la cromatina, así como una posible alteración en la interacción identificada y posiblemente una alteración en la expresión de *LRPPRC*. También se ha demostrado diferencia alelo-específica para la unión de CTCF por la presencia de esta variante V15 perteneciente al haplotipo de riesgo mediante un ensayo de retardo electroforético (EMSA)<sup>37</sup>. (Figura suplementaria 1.3)

Los demás factores identificados en los dos alineamientos (Figura 8.9 y Figura 8.10) tienen diversas funciones dentro de la célula. Hay factores involucrados en inmunología (STAT<sup>133</sup>, IRF<sup>134</sup>, PAX5<sup>135</sup>, NF- $\kappa$ B<sup>136</sup>, RFX5<sup>137</sup>), transcripción (TFIIA<sup>138</sup>, POLR2A<sup>139</sup>, FAIRE<sup>140</sup>), estado indiferenciado (NANOG<sup>141</sup>, NKX2<sup>142</sup>, POU5F1<sup>143</sup>), desarrollo y diferenciación de diferentes tejidos (MAFG<sup>144</sup>, NF-E2<sup>145</sup>, GATA<sup>146</sup>, factores ARID<sup>147</sup>, CTCFL<sup>148</sup>, RCOR1<sup>149</sup>), ciclo y división celular (MYC<sup>150</sup>, factores ARID, E2F<sup>151</sup>, EZH2<sup>152</sup>), así como desarrollo de cáncer (SS18<sup>153</sup>, C11orf30<sup>154</sup>, MYC, E2F, EZH2, NANOG); todas funciones importantes, pero tendría que evaluarse experimentalmente su unión en el contexto hepático adulto, así como alguna posible afectación debido a la presencia o ausencia del haplotipo de riesgo y finalmente, expresión de *LRPPRC* o genes aledaños.

De los dos alineamientos mostrados en las Figuras 8.9 y 8.10, los motivos respuesta a factores de transcripción relevantes para nuestro estudio muestran factores que son potencialmente excluyentes mutuamente. Por ejemplo, si se une GCNF, no podría haber unión de ningún otro factor de transcripción. Sin embargo, CTCF tiene posibilidad unirse con USF1. Para VI2 el factor FOXA, en principio no tendría competencia de unión con otro factor de transcripción. En el caso de V15 (Figura 8.10) hay una gran acumulación de factores que pudieran ser excluyentes entre sí, aunque las proteínas pertenecientes a la cohesina son necesarios para la formación del complejo, por lo que es esencial esta aglomeración de factores para que la cohesina cumpla su función de arquitectura en la cromatina junto con CTCF. Aun así, la distribución y superficie de contacto entre los factores de transcripción en el *locus* es importante para definir si la unión es continua o discontinua. En la unión continua, todos los sitios de unión están siendo ocupados por el factor de transcripción y en la discontinua, no todos los sitios son ocupados.

Los portales utilizados son de predicciones de unión, siendo USF1 y CTCF los únicos factores de transcripción para el cual existe información de *ChIP-seq* derivado del proyecto *ENCODE*. De estos datos concluimos que los sitios de unión se encuentran

dentro de los elementos respuesta que comprenden a VI1 y V15, ubicándose en el centro de la secuencia de potencial unión. Quedaría como perspectiva el identificar los factores de transcripción que se unen al *locus* en un contexto hepático, así como posibles diferencias en la unión como resultado de la presencia de las variantes.

Algo notorio es nuestra falta de un modelo celular óptimo, ya que por el momento se carecen de hepatocitos o su cultivo primario provenientes de individuos sanos o de pacientes portadores con HTGSF. En este estudio se utilizó material genético proveniente de sangre periférica, biopsias hepáticas y la línea celular HepG2, pero no fue el suficiente para intentar otras técnicas experimentales por lo que se podría proponer utilizar otros tejidos de fácil acceso como tejido adiposo subcutáneo. También se podría generar un modelo celular proveniente de fibroblastos de portadores y no portadores e indiferenciarlos a células pluripotentes inducidas (iPS) para generar tipos celulares hepáticos, musculares y pancreáticos; así sería posible evaluar todo el contexto genómico, transcripcional, proteínico y metabólico más cercano a lo que sucede *in vivo*.



## 10. CONCLUSIONES

El análisis del *locus* de *LRPPRC* tanto por estrategias experimentales como bioinformáticas sugiere que se trata de un *locus* regulatorio donde distintas variantes genéticas (parte del haplotipo de riesgo para la HTGSF) podrían afectar la expresión de este gen, así como afectar potenciales regiones de interacción a distancia con otros genes dentro del mismo *locus*.

Descartamos que existan mutaciones en la región codificante del gen y en las uniones intrón-exón en los sujetos portadores del haplotipo de riesgo que afecten la estructura primaria de la proteína de *LRPPRC*.

No identificamos alteraciones en los niveles de expresión del mRNA de *LRPPRC* a partir de muestras hepáticas de individuos portadores y no portadores del haplotipo de riesgo.

Mediante distintos análisis bioinformáticos identificamos que la región que contiene la variantes rs7593842 y rs58811869, originalmente asociadas a la HTGSF, podrían alterar el proceso de empalme del RNA aun cuando se encuentran distantes a las uniones intrón-exón. Además de potencialmente afectar la unión a diferentes factores de transcripción. Consistente con estas observaciones, identificamos una RNA mensajero más corto que carece de secuencias codificantes (exones 7, 8, 9 y 10) a partir de biopsias hepáticas de sujetos portadores del haplotipo de riesgo.

Adicionalmente, de acuerdo al papel conocido de la proteína *LRPPRC*, evidenciamos indicios de disminución de la expresión de genes mitocondriales en un sujeto portador del haplotipo de riesgo comparada con no portadores, a partir de muestras de biopsias hepáticas; semejante a lo reportado en el modelo murino *knock-down* de *LRPPRC*. Sin embargo, debido al reducido número de muestras analizadas no obtuvimos diferencias estadísticamente significativas.

Por último, respecto al potencial papel de este *locus* como una región de regulación transcripcional compleja, identificamos 3 regiones de interés, una de ellas afectando uniones a distancia con los genes *ABCG5* y *ABCG8*. También distintos factores de transcripción en tejido hepático podrían potencialmente afectar su unión por la presencia de dos variantes identificadas en el intrón 23, (rs4263155 y rs4600713) y dos de las variantes del haplotipo de riesgo (rs7593842 y rs58811869).

## 11. ANEXO

### Información suplementaria.

Nombre	Pares de bases	Biotipo	Proteína
LRPPRC-201	6603	Codifica proteína	1394aa
LRPPRC-217	5259	Codifica proteína	1369aa
LRPPRC-239	5169	Codifica proteína	1430aa
LRPPRC-241	5079	Codifica proteína	1404aa
LRPPRC-240	5071	Codifica proteína	1395aa
LRPPRC-222	5052	Codifica proteína	1393aa
LRPPRC-254	4973	Codifica proteína	1366aa
LRPPRC-249	4951	Codifica proteína	1363aa
LRPPRC-229	4757	Codifica proteína	1391aa
LRPPRC-248	4698	Codifica proteína	1285aa
LRPPRC-231	4513	Codifica proteína	1379aa
LRPPRC-220	4325	Codifica proteína	1400aa
LRPPRC-256	3514	Codifica proteína	913aa
LRPPRC-205	2789	Codifica proteína	837aa
LRPPRC-202	2663	Codifica proteína	841aa
LRPPRC-213	2541	Codifica proteína	795aa
LRPPRC-203	1875	Codifica proteína	531aa
LRPPRC-223	5264	Degradado por mutaciones terminadoras	1356aa
LRPPRC-259	5148	Degradado por mutaciones terminadoras	60aa
LRPPRC-268	5142	Degradado por mutaciones terminadoras	50aa
LRPPRC-257	5019	Degradado por mutaciones terminadoras	1277aa
LRPPRC-252	4981	Degradado por mutaciones terminadoras	50aa
LRPPRC-244	4543	Degradado por mutaciones terminadoras	50aa
LRPPRC-267	4440	Degradado por mutaciones terminadoras	430aa
LRPPRC-216	4365	Degradado por mutaciones terminadoras	50aa
LRPPRC-262	4125	Degradado por mutaciones terminadoras	54aa
LRPPRC-258	3965	Degradado por mutaciones terminadoras	739aa
LRPPRC-225	3765	Degradado por mutaciones terminadoras	776aa
LRPPRC-226	3328	Degradado por mutaciones terminadoras	266aa
LRPPRC-255	2182	Degradado por mutaciones terminadoras	422aa
LRPPRC-234	2028	Degradado por mutaciones terminadoras	281aa
LRPPRC-204	1282	Degradado por mutaciones terminadoras	167aa
LRPPRC-228	1188	Degradado por mutaciones terminadoras	33aa
LRPPRC-247	753	Degradado por mutaciones terminadoras	141aa
LRPPRC-263	524	Degradado por mutaciones terminadoras	115aa
LRPPRC-215	519	Degradado por mutaciones terminadoras	107aa
LRPPRC-242	369	Degradado por mutaciones terminadoras	55aa
LRPPRC-210	1920	Transcrito procesado	Sin proteína
LRPPRC-251	1802	Transcrito procesado	Sin proteína
LRPPRC-209	528	Transcrito procesado	Sin proteína
LRPPRC-266	10249	Intrón retenido	Sin proteína
LRPPRC-271	7255	Intrón retenido	Sin proteína
LRPPRC-253	6739	Intrón retenido	Sin proteína
LRPPRC-264	6182	Intrón retenido	Sin proteína
LRPPRC-246	5624	Intrón retenido	Sin proteína
LRPPRC-243	5290	Intrón retenido	Sin proteína
LRPPRC-250	5175	Intrón retenido	Sin proteína
LRPPRC-235	5084	Intrón retenido	Sin proteína
LRPPRC-224	4780	Intrón retenido	Sin proteína
LRPPRC-232	4510	Intrón retenido	Sin proteína
LRPPRC-261	4216	Intrón retenido	Sin proteína
LRPPRC-211	4189	Intrón retenido	Sin proteína
LRPPRC-219	3735	Intrón retenido	Sin proteína
LRPPRC-206	3549	Intrón retenido	Sin proteína
LRPPRC-245	3548	Intrón retenido	Sin proteína
LRPPRC-218	3437	Intrón retenido	Sin proteína





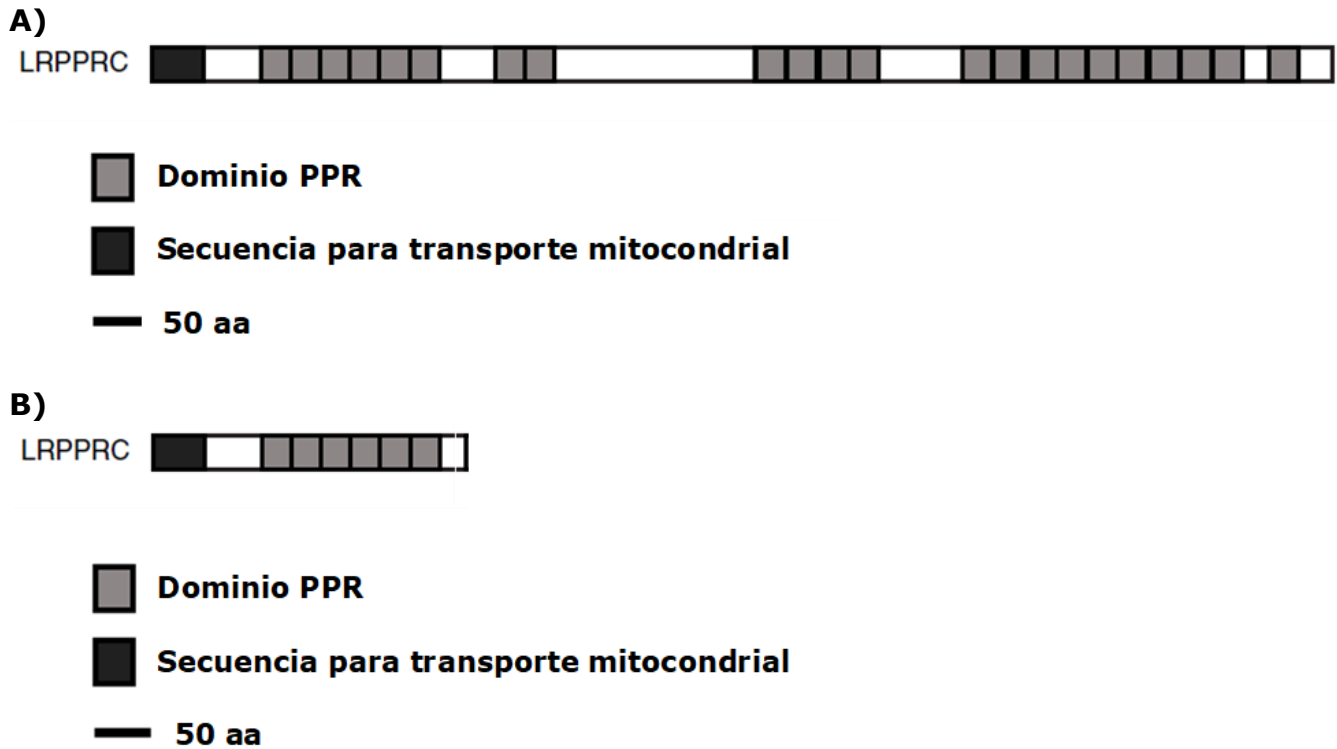


Figura suplementaria 1.2. **Representación esquemática de LRPPRC canónica y la propuesta de LRPPRC más corta.** A) LRPPRC completa (1394 aa), B) isoforma de LRPPRC trunca debido al *skipping* del exón 9, que causa un cambio en el marco de lectura y un paro prematuro (345 aa). Modificado de Gaweda-Walerych, 2016.

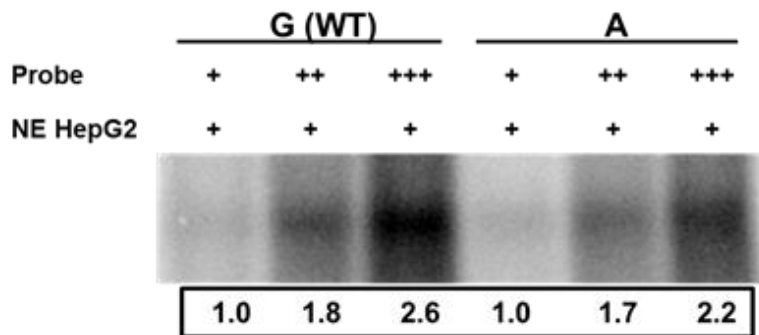


Figura suplementaria 1.3. **Ensayo EMSA donde se muestra una disminución de la unión alelo-específica de CTCF en el locus de rs58811869 (Intrón 15).** Probe: sonda marcada radioactivamente, NE: extracto nuclear de HepG2. Modificado de Chavez, 2018.

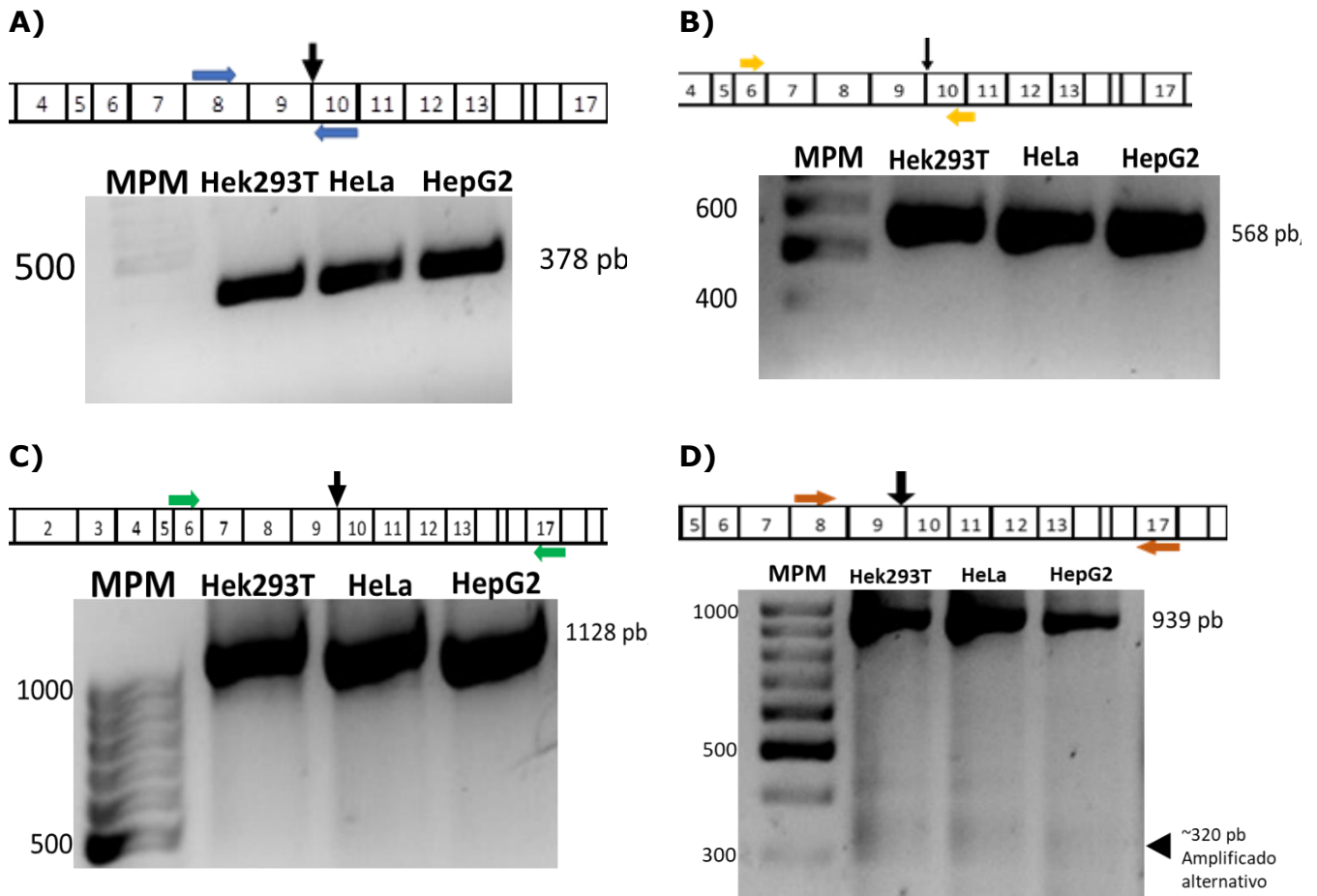


Figura suplementaria 1.4. **Búsqueda de isoformas de LRPPRC en muestras de líneas celulares.** En los paneles superiores (A), (B), (C) y (D) se muestra una representación esquemática de la zona donde se encuentra la variante de interés (flechas verticales) y la zona del cDNA de LRPPRC donde se estarían uniendo los oligonucleótidos (flechas horizontales; los colores son para representar los pares diferentes utilizados). Paneles inferiores (A), (B), (C) y (D) se muestra el análisis de PCR de punto final de la retrotranscripción del mRNA de LRPPRC. Gel de agarosa al 3%. MPM: Marcador de peso molecular.

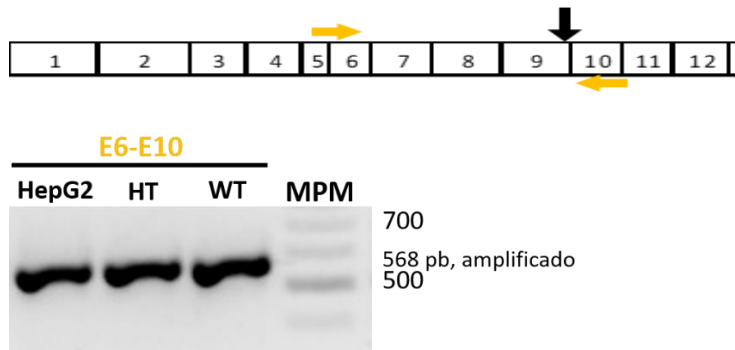
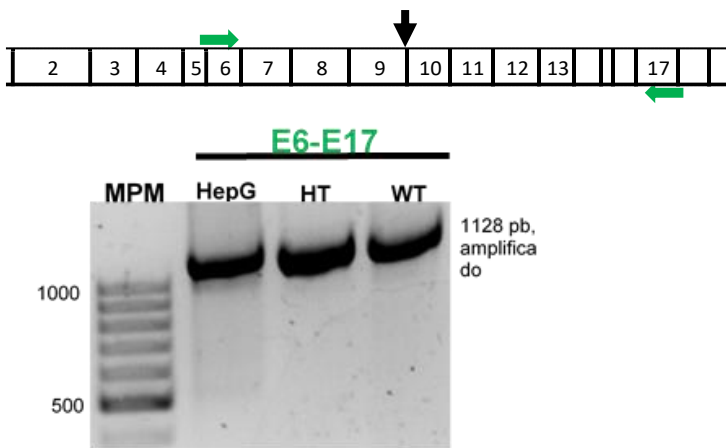
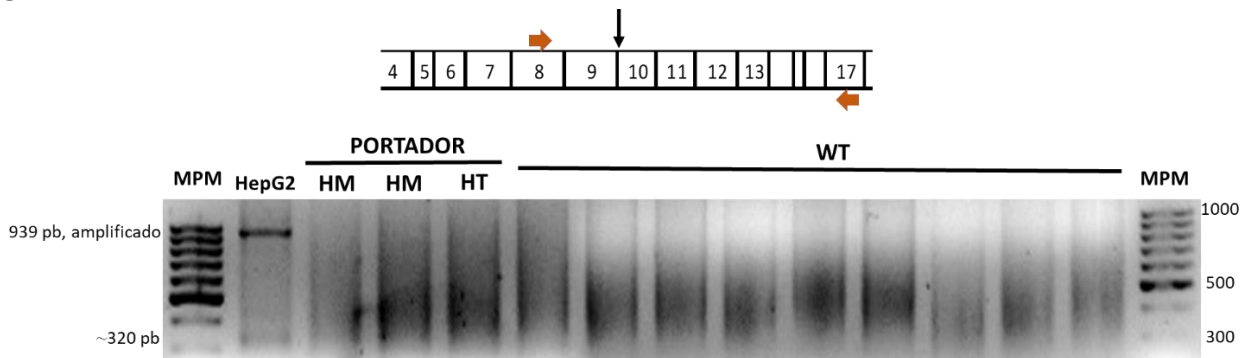
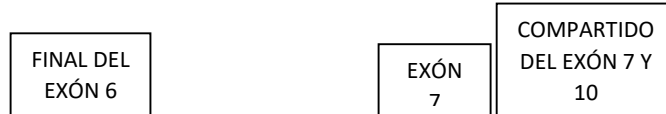
**A****B****C**

Figura suplementaria 1.5. **Búsqueda de isoformas de LRPPRC en muestras de pacientes.** En los paneles superiores (A), (B) y (C) se muestra una representación esquemática de la zona donde se encuentra la variante y la zona del cDNA de LRPPRC donde se estarían uniendo los oligonucleótidos. Paneles inferiores (A), (B) y (C) se muestra el análisis de PCR de punto final de la retrotranscripción del mRNA de LRPPRC. MPM: Marcador de peso molecular, HM: Homocigoto, HT: Heterocigoto, WT: Wild Type. Gel de agarosa al 3%.

A



GTATTCAGTGCCCTTGTGACAGGGCACGCCAGAGCTGGT **GATATGGAGAAGCTAACAGACTA**  
CTGTAAGAAGTTAAAGGAAGTCCAGATGCACTCCTTTCCTCTGCAG

EXÓN 10

B



Figura suplementaria 1.6. **Representación esquemática de la secuencia del amplificado alternativo identificado con ausencia de secuencias exónicas.** A) Construcción de la secuencia y B) esquema de la región en donde los oligonucleótidos hibridarían.

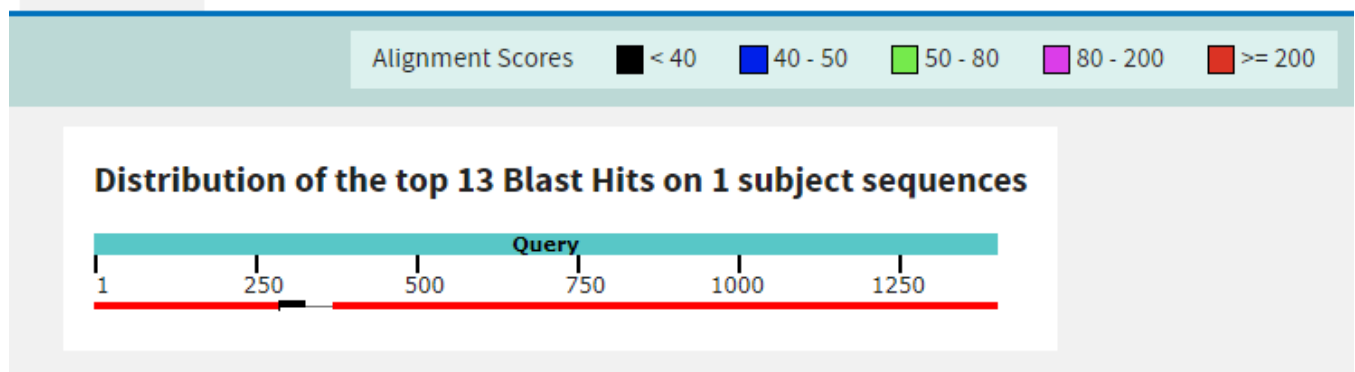
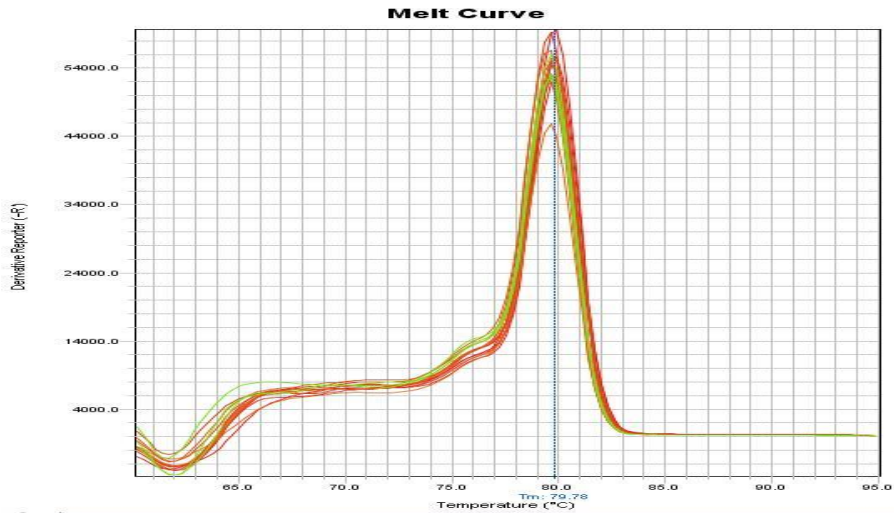


Figura suplementaria 1.7. **Imagen delimitada del resumen gráfico del alineamiento en la herramienta BLASTp.** Representación gráfica del alineamiento de las secuencias de aminoácidos de LRPPRC canónica y de la traducción hecha por en el portal *EXPASy* del posible transcrito de la Figura 9.2C.



A)



B)

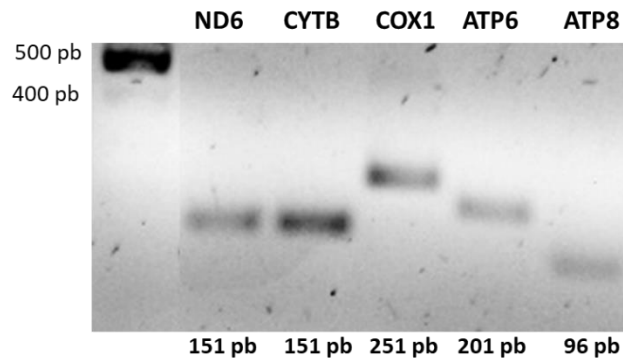


Figura suplementaria 1.8. **Comprobación de la especificidad en los ensayos de qPCR.** A) Imagen representativa de las *melting curve* de los ensayos de expresión relativa de *LRPPRC* y genes mitocondriales (gen ATP6 con las 3 repeticiones por las 5 concentraciones de cDNA utilizada), B) imagen representativa de algunos de los amplificadores de las reacciones de qPCR.

<b>Paciente</b>	<b>TG</b> (mg/dL)	<b>Colesterol</b> <b>total</b> (mg/dL)	<b>LDL</b> (mg/dL)	<b>HDL</b> (mg/dL)	<b>DMT2</b>
<b>1</b>	229.80	162.16	76.25	41.51	Sí
<b>2</b>	114	230	161	46	Sí
<b>3</b>	122	258	181	53	Sí
<b>4</b>	124	238	163	50	No
<b>5</b>	79.65	181.47	111.97	50.19	No
<b>6</b>	132.74	216.22	142.86	46.33	No
<b>7</b>	123.89	227.80	158.30	42.47	Sí
<b>8</b>	132.74	173.75	104.25	42.47	No
<b>9</b>	141.59	154.44	84.94	42.47	No
<b>10</b>	123.89	212.36	139.00	46.33	No
<b>11</b>	132.74	173.75	104.25	42.47	Sí
<b>12</b>	141.59	250.97	166.02	57.92	Sí
<b>13</b>	134.89	241.94	127.02	87.70	No
<b>14</b>	141.59	193.05	108.11	57.92	Sí
<b>15</b>	132.74	212.36	135.14	50.19	No
<b>16</b>	106.19	208.49	139.00	46.33	Sí
<b>17</b>	141.59	166.02	81.08	57.92	No
<b>18</b>	115.04	193.05	119.69	50.19	Sí
<b>19</b>	88.50	154.44	92.66	46.33	No
<b>20</b>	44.25	166.02	115.83	42.47	Sí
<b>21</b>	70.80	239.38	146.72	81.08	No
<b>22</b>	141.59	146.72	77.22	42.47	Sí
<b>23</b>	79.65	254.83	189.19	50.19	No

Tabla suplementaria 1.2. **Información de concentración de lípidos sanguíneos así como diagnóstico de diabetes de los pacientes de la secuenciación de los exones de LRPPRC** (Tabla 7.1, secciones 8. Búsqueda de nuevas variantes en los exones del gen LRPPRC y 8. Búsqueda de variantes en la región de interacción dentro del intrón 23 de LRPPRC). DMT2: Diabetes Mellitus tipo 2.

<b>Paciente</b>	<b>Peso</b> (kg)	<b>TG</b> (mg/dL)	<b>Colesterol</b> <b>total</b> (mg/dL)	<b>HDL</b> (mg/dL)	<b>LDL</b> (mg/dL)	<b>Apo B</b> (mg/dL)	<b>Estado</b> <b>hepático</b>	<b>DMT2</b>
<b>1</b>	116.6	236	184	27	110	131	EHNA	Sí
<b>2</b>	127.4	138	184	37	119.4	100.3		Sí
<b>3</b>	105.1	217	166	23	100	113	Frontera EHNA	Sí
<b>4</b>	130.7	150	188	38	120	120.4	Frontera EHNA	No
<b>5</b>	125.2	183	153	33	83	93.3	EHNA	Sí
<b>6</b>	135.5	163	200	40	127	143	Control	Sí
<b>7</b>	119.4	112	152	33	96.9	80.6	Frontera EHNA	Sí
<b>8</b>	122.9	115	178	39	116	104.5	Frontera EHNA	No
<b>9</b>	130.1	133	155	38	90	81	Control	No
<b>10</b>	116.25	133	93	30	36	60.7		Sí
<b>11</b>	117.4	109	141	37	82	81.8	Frontera EHNA	No
<b>12</b>	120.9	92	174	36	119.6	109.4		Sí

Tabla suplementaria 1.3. **Información de concentración de lípidos sanguíneos así como diagnóstico de diabetes y EHNA de los pacientes de la toma de biopsia hepática para los ensayos de búsqueda de isoformas de LRPPRC, expresión de LRPPRC y genes mitocondriales** (Tabla 7.2, secciones 8. Búsqueda de isoformas de LRPPRC y 8. Estudios de niveles de expresión de genes en tejido hepático y línea celular humana HepG2). DMT2: Diabetes Mellitus tipo 2. EHNA: Esteatohepatitis no alcohólica.

<b>M</b>	<b><math>\bar{x}</math> y <math>\sigma</math> ND1</b>	<b><math>\bar{x}</math> y <math>\sigma</math> ND6</b>	<b><math>\bar{x}</math> y <math>\sigma</math> CYTB</b>	<b><math>\bar{x}</math> y <math>\sigma</math> COX1</b>	<b><math>\bar{x}</math> y <math>\sigma</math> ATP6</b>	<b><math>\bar{x}</math> y <math>\sigma</math> ATP8</b>	<b><math>\bar{x}</math> y <math>\sigma</math> h18s</b>
<b>1</b>	19.80198797	20.76224899	19.43494733	17.83757782	18.4223849	18.54907799	11.98028183
<b>P</b>	$\pm 1.03729925$	$\pm 0.13418368$	$\pm 0.07983483$	$\pm 0.17822751$	$\pm 0.11210682$	$\pm 0.38667079$	$\pm 0.23097176$
<b>2</b>	20.8090712	19.9858863	19.3558369	17.8238818	18.4335887	18.5003255	11.89303239
<b>P</b>	$\pm 0.10709017$	$\pm 0.03161649$	$\pm 0.07160884$	$\pm 0.21355573$	$\pm 0.10726935$	$\pm 0.06308543$	$\pm 0.110157704$
<b>3</b>	21.3471127	20.9403712	19.9051882	18.1616141	19.0302684	19.153396	11.54958121
<b>P</b>	$\pm 0.03264536$	$\pm 0.02216021$	$\pm 0.03066778$	$\pm 0.02139357$	$\pm 0.05272405$	$\pm 0.16398009$	$\pm 0.08378904$
<b>4</b>	20.5237306	20.0269642	19.0424633	17.0785007	18.7381872	18.5894642	11.5385555
<b>WT</b>	$\pm 0.37546983$	$\pm 0.0403872$	$\pm 0.07392303$	$\pm 0.05608473$	$\pm 0.05331073$	$\pm 0.21396006$	$\pm 0.34788643$
<b>5</b>	20.285195	20.2058341	19.2390874	16.7094364	18.047849	18.1283951	11.4823469
<b>WT</b>	$\pm 0.0390903$	$\pm 0.08969729$	$\pm 0.04720165$	$\pm 0.09352769$	$\pm 0.15398331$	$\pm 0.04217553$	$\pm 0.19195948$
<b>6</b>	20.5915807	20.6530107	19.7044214	17.3819637	18.5901934	18.5724462	11.59625594
<b>WT</b>	$\pm 0.15537996$	$\pm 0.09371975$	$\pm 0.12321837$	$\pm 0.55844031$	$\pm 0.13154707$	$\pm 0.16522713$	$\pm 0.086212044$
<b>7</b>	20.7370122	20.8878969	19.9146709	17.2069244	18.6441924	18.6502838	12.59360981
<b>WT</b>	$\pm 0.06734488$	$\pm 0.05629171$	$\pm 0.20808397$	$\pm 0.14660196$	$\pm 0.1869186$	$\pm 0.15058531$	$\pm 0.125730187$
<b>8</b>	20.1128	18.7947985	18.5934188	17.1070251	17.4515851	17.6811523	11.46308136
<b>WT</b>	$\pm 0.10370427$	$\pm 0.10582561$	$\pm 0.12288648$	$\pm 0.01781789$	$\pm 0.06476831$	$\pm 0.0786446$	$\pm 0.106972546$
<b>9</b>	20.5997601	19.8841387	18.7953625	16.8824673	18.5507698	18.5651423	12.451088
<b>WT</b>	$\pm 0.24709339$	$\pm 0.03483949$	$\pm 0.1391345$	$\pm 0.23070745$	$\pm 0.17240258$	$\pm 0.11632977$	$\pm 0.15466238$
<b>10</b>	21.7563864	20.7667109	19.8998928	18.46897	19.2058239	19.2823334	12.32512728
<b>WT</b>	$\pm 0.08548733$	$\pm 0.18898377$	$\pm 0.10195767$	$\pm 0.07627473$	$\pm 0.11858092$	$\pm 0.09682783$	$\pm 0.136815032$
<b>11</b>	20.4173539	20.2282327	19.4288534	17.6046015	18.3098602	18.9078471	11.68710645
<b>WT</b>	$\pm 0.13978263$	$\pm 0.08360537$	$\pm 0.0970552$	$\pm 0.1737378$	$\pm 0.1635746$	$\pm 0.00654896$	$\pm 0.117692655$
<b>12</b>	20.7946256	20.3106772	19.0358423	16.6731952	18.8334141	18.6219514	12.0944796
<b>WT</b>	$\pm 0.10656928$	$\pm 0.0168503$	$\pm 0.14284569$	$\pm 0.58976837$	$\pm 0.12793101$	$\pm 0.12329235$	$\pm 0.1971311$

Tabla suplementaria 1.4. **Medias y desviaciones estándar de los valores de Ct del ensayo de expresión génica qPCR tiempo real de genes mitocondriales en las 12 muestras de tejido hepático.** M: Muestra; P: Portador; WT: *Wild type* (no portador).

<b>Status</b>	<b><math>\bar{x}</math> y <math>\sigma</math> ND1</b>	<b><math>\bar{x}</math> y <math>\sigma</math> ND6</b>	<b><math>\bar{x}</math> y <math>\sigma</math> CYTB</b>	<b><math>\bar{x}</math> y <math>\sigma</math> COX1</b>	<b><math>\bar{x}</math> y <math>\sigma</math> ATP6</b>	<b><math>\bar{x}</math> y <math>\sigma</math> ATP8</b>
<b>Portador</b>	8.845092138 $\pm 0.98982$	8.755203671 $\pm 0.64938183$	7.757692337 $\pm 0.51782531$	6.133392758 $\pm 0.41614277$	6.821115494 $\pm 0.57332306$	6.926634683 $\pm 0.58677095$
<b>WT</b>	8.731865989 $\pm 0.4065$	8.280734804 $\pm 0.5655299$	7.380262446 $\pm 0.52288149$	5.320159312 $\pm 0.63768154$	6.571136051 $\pm 0.43777183$	6.640818349 $\pm 0.43694315$

Tabla suplementaria 1.5. **Medias y desviaciones estándar de los valores de  $\Delta$ Ct del ensayo de expresión génica qPCR tiempo real de genes mitocondriales en las medias de muestras de portadores (N=3) y no portadores (N=9) de tejido hepático.** WT: *Wild type* (no portador).

## Metodologías suplementarias

### Secuenciación

Región (Exón)	Oligonucleótido Sentido 5'-3'	Oligonucleótido antisentido 5'-3'	Ta (°C)	Tamaño del amplificado (pb)
1	ACCTAAAACAATCCGTCGCTCC	CGCTTAACCCTGTCACCCGAA	63	466
2	CTTTCCATTAGTTCATACCTC	CTGTGACCAGCCAAAAGTCAA	54	393
3	TTTAGCCTTTGAATTGTAGC	TTATGAATCAAGTAGCCCTTC	52	551
4-5	CCTTTTAATATCCGAATAGGC	TATCTTTTCTCTTACACGA	49	709
6	AACAAAAGTAACTCAAGGC	ATTTTCTAATTTAAACCCC	49	287
7	AGGGTATGTTATTTAATGTC	TTTGGGCTTATTTAACTGAC	49	982
8	TTTAAAAGAAGTCGTAACCT	TTTGGGCTTATTTAACTGAC	52	414
9-11	TTGGACATTGGTTGCAGTG	AGATGTGCGTGCAAACCTGTC	60	828
12	TGGGCACATTGCTTAATTGT	CCATTTTGAGTCCTCAGTTCAA	59	500
13	TCAGGAAAACAGGATCTTTGG	CTTGCTTTTATTGCGTGAAC	59	262
14	TGCATTGTCTAATTTTGTATGAAAAC	AAACAAGCAAAGGTTATTTTCAGG	57	367
15	TGTCCTGAAAATTCATCCTGA	CAACATTAACCTCTGCCAG	59	300
16	TGGGGGTGGTTTGAATTAC	CCTCCAACCCTGTCTTTAG	60	409
17	CCACATAATCTCTATGCCTCA	AGGACAGGCATTATATAGCAA	54	299
18	TTGCTGCTTTTAAGTACAAC	CAGAATTAGAAAACACACGTA	50	548
19-20	TGACAGCCATTTAAAACGAA	TGGCATGTATGATAAACCT	51	1009
21	AACACATCTTTAAACCCCT	CCTTGCTGAACATAAGCCA	51	525
22	ACTTTGAATATAATTTCTGTGTA	GTGCAAATGTAATGACAAGC	53	473
23	AGTAATTAACCTGAAATCGGCAT	TATGTAAAGCCACAGGACCAC	56	530
24	GTACCCTTTAACACACGAA	AAAGTTTATCTGAATGTGCT	49	495
25	TATCTACTGCTCTTGAGGTCG	TTAATATAACAACCAGACTGCC	52	886
26	AAAATAATATTACGCCGAAG	CTTATCTGGCAAATGATCG	51	268
27	TTCAACTGCCCTTTGGTA	ACAGAATAAATGAAAGCGTGA	55	244
28-29	GGGCTTCTTGTTATTGCTTT	CACTTCTAATTGGTTGGGCTT	56	696
30	CTTTCCACACATTTTGCAT	TAAGCAAAATGTACTGCACT	51	717
31	GCACGCAGTTATTTTAGTCA	ACAAAAGGCTCTTGATAAGCA	54	522
32	TGACTTAAAGAAGAAATTCCTCG	CAATGCCATTCTAGTGAC	54	484
33-34	ATTCAGGGTTTCAGTTGCCTA	TAGTATTCACCACACGCCTA	58	645
35	CTTTGGTCTCTTTCTCAAGCG	TACTTCTGAGCAAGGCACGTA	58	696
36	TGCTGGTTTCTTTAGCTGT	AGTCACTTAAATAATAAAGCCAT	51	339
37	TGAATGTGGCTTCTTACTGA	GTAATCCATGCCTATAGCTG	54	536
38	CTCAGCCTTCATACCCAT	AACAGAGCAGGGAAACCCAA	57	636

Tabla suplementaria 2.1. **Oligonucleótidos y condiciones utilizadas para las reacciones de amplificación por PCR y secuenciación Sanger de las regiones exónicas de LRPPRC.**

Reactivo	Cantidad
<i>HotStarTaq Master Mix (Qiagen)</i>	5 µL
Oligonucleótido sentido (10 µM)	1 µL
Oligonucleótido antisentido (10 µM)	1 µL
DNAg	50 ng
Agua grado Biología Molecular	3 µL

Tabla suplementaria 2.2. **Cantidades de reactivos utilizadas para las reacciones de amplificación por PCR punto final.**

95 °C	94 °C	Ta	72 °C			72 °C	
10 min	30 seg	30 seg	30 seg	45 seg	1 min	10 min	20 min
			≤500 pb	600-800 pb	900-1000 pb	≤500 pb	≥600 pb
1 ciclo	40 ciclos					1 ciclo	

Tabla suplementaria 2.3. **Secuencia de ciclos utilizada en la reacción de amplificación por PCR punto final para las regiones exónicas de LRPPRC utilizando HotStarTaq Master Mix (Qiagen).**

Reactivo	Cantidad
BigDye Terminator v3.1 Cycle ( <i>Applied Biosystems</i> )	2 µL
Oligonucleótido Sentido o Antisentido (10 µM)	2 µL
BigDye Terminator v1.1, v3.1 5x Sequencing Buffer ( <i>Applied Biosystems</i> )	3 µL
Reacción de PCR purificado	3-6 µL
Agua grado Biología Molecular	11 µL

Tabla suplementaria 2.4. **Cantidades de reactivos utilizadas para las reacciones de secuenciación Sanger.**

93 °C	95 °C	52 °C	60 °C	60 °C
3 min	10 seg	15 seg	2 min	3 min
1 ciclo	35 ciclos			1 ciclo

Tabla suplementaria 2.5. **Secuencia de ciclos utilizada en la reacción de secuenciación Sanger**

Amplificado dentro de la región de interacción	Oligonucleótido Sentido 5'-3'	Oligonucleótido antisentido 5'-3'	Ta (°C)	Tamaño del amplificado (pb)
1	GTGATGCAAGTATCTT	GGATTCTATATAAAGTTGGCTA	51	2931
2	TTGGCATTAAAAGAATGAGT	GAAAGAGTCAATAGCCCTT	51	1899
3	TGTGCTTCTGATTATATCCTG	TAAAATTTACTCAAACATGCG	53	2243

Tabla suplementaria 2.6. **Oligonucleótidos y condiciones utilizadas para las reacciones de amplificación por PCR punto final de la región de interacción en el intrón 23 de LRPPRC.**

Reactivo	Cantidad (µL)
Solución Q, 5x	2.5
Oligonucleótido Sentido (10 µM)	1

Oligonucleótido Antisentido (10 µM)	1
Buffer LongRange PCR con Mg <sup>2+</sup> , 10x	1.25
Mezcla de dNTPs (10 mM cada uno)	0.6
Mezcla de enzima LongRange PCR	0.1
DNAg	50 ng
Agua grado Biología Molecular	2.5

Tabla suplementaria 2.7. **Cantidades de reactivos utilizados para las reacciones de PCR punto final con uso de la mezcla de enzimas *LongRange PCR* (Qiagen).**

93 °C	93 °C	Ta °C	68 °C	4 °C
3 min	15 seg	15 seg	4 min	Indefinido
1 ciclo		40 ciclos		

Tabla suplementaria 2.8. **Secuencia de ciclos utilizada en la reacción de amplificación por PCR punto final para las regiones exónicas de *LRPPRC* utilizando la mezcla de enzimas *LongRange PCR* (Qiagen).**

Amplificado dentro de la región de interacción	Oligonucleótidos para la secuenciación 5'-3'
1	Sentido. GTGATGCAAGTATCTT Sentido. ATGTTTTAATTTTATTATAGGGA
2	Sentido. TTGGCATTAAAAGAATGAGT Antisentido. GGATTTCTATATAAAGTTGGCTA Antisentido. GAAAGAGTCAATAGCCCTT
3	Sentido. TGTGCTTCTGATTATATCCTG Antisentido. TTCTTCCTTAATGTGACGG Antisentido. TAAAATTTACTCAAACATGCG

Tabla suplementaria 2.9. **Oligonucleótidos utilizados para las reacciones de secuenciación Sanger de la región de interacción en el intrón 23 de *LRPPRC*.**

### Amplificación para búsqueda de isoformas de *LRPPRC*

Región (Exón)	Oligonucleótido Sentido 5'-3'	Oligonucleótido antisentido 5'-3'	Ta (°C)	Tamaño del amplificado (pb)
V9	CGAGCTTCACCTTATGGAC	AAATCAGTTTTATTGGCGAGT	55	378
V15	GCAGTCTATAAGAAGTAGCCT	AAGTATTGTCTCAAATGCTCC	50	214
E6-E10	GTATTCAGTGCCTTGTGAC	AAATCAGTTTTATTGGCGAGT	55	568
E8-E17	CGAGCTTCACCTTATGGAC	CAGGAGATTACGAATGCCTC	53	969
V6	GTATTCAGTGCCTTGTGAC	TTTGCAGTAAATCACGGTCCA	55	223
E6-E17	GTATTCAGTGCCTTGTGAC	CAGGAGATTACGAATGCCTC	53	1128

Tabla suplementaria 2.10. **Oligonucleótidos y condiciones utilizadas para las reacciones de amplificación por PCR y secuenciación Sanger del mRNA de *LRPPRC* utilizando *HotStarTaq Master Mix* (Qiagen).**

## Estudios de expresión

Reactivo	Cantidad (µL)
TaqMan™ Universal PCR Master Mix, 2x	6.25
Sonda TaqMan	0.6
cDNA	2.5
Agua grado Biología Molecular	3.15

Tabla suplementaria 2.11. **Cantidades de reactivos utilizados para las reacciones de PCR cuantitativa con uso del ensayo TaqMan para LRPPRC.**

Reactivo	Cantidad (µL)
TaqMan™ Universal PCR Master Mix, 2x	6.25
Oligonucleótido Sentido (10 µM)	0.375
Oligonucleótido Antisentido (10 µM)	0.375
Sonda TaqMan	0.25
cDNA	2.5
Agua grado Biología Molecular	2.75

Tabla suplementaria 2.12. **Cantidades de reactivos utilizados para las reacciones de PCR cuantitativa con uso de oligonucleótidos y sonda TaqMan.**

50 °C	93 °C	Ta °C	68 °C	4 °C
2 min	15 seg	15 seg	4 min	Indefinido
1 ciclo		40 ciclos		

Tabla suplementaria 2.13. **Secuencia de ciclos utilizada en la reacción de PCR cuantitativa usando sondas TaqMan.**

Gen	Oligonucleótido Sentido 5'-3'	Oligonucleótido antisentido 5'-3'	Tamaño del amplificado (pb)
<b>ND1</b>	CCCTAAAACCCGCCACATCT	GAGCGATGGTGAGAGCTAAGGT	69
<b>ND6</b>	GGTTAGCGATGGAGGTAGGATTG	AAAGCCCCCGCACCAATA	151
<b>CYTB</b>	TCACCTCCCATTCCGATAAAATCAC	GGGTTGGCTAGGGTATAATTGTCTG	151
<b>COX1</b>	GACGTAGACACACGAGCATATTTCA	AGGACATAGTGAAGTGAGCTACAAC	251
<b>ATP6</b>	TAGCCATACACAACACTAAAGGACGA	GGGCATTTTTAATCTTAGAGCGAAA	201
<b>ATP8</b>	CAACTAAAAATATTAACACAACTACCACC	CGTTCATTTTGGTTCTCAGGG	96
<b>h18s</b>	CGGCTACCACATCCAAGGAA	CCTGTATTGTTATTTTTTCGTCACCT	86
<b>GAPDH</b>	CCCCGGTTTCTATAAATTGAGC	CTTCCCCATGGTGTCTGAG	364
<b>β-ACTINA</b>	ACCGCGAGAAGATGACCCAG	TGCTCGAAGTCCAGGGCGAC	331
<b>GUSB</b>	GAAAATATGTGGTTGGAGAGCTCATT	CCGAGTGAAGATCCCCTTTTFA	101

Tabla suplementaria 2.14. **Oligonucleótidos utilizados para las reacciones de amplificación por PCR cuantitativa del cDNA.**

Reactivo	Cantidad (µL)
SYBR GREEN	6.25
Oligonucleótido Sentido (10 µM)	0.375
Oligonucleótido Antisentido (10 µM)	0.375
cDNA (concentración variable)	2
Agua grado Biología Molecular	4

Tabla suplementaria 2.15. **Cantidades de reactivos utilizados para las reacciones de PCR cuantitativa con uso del sistema SYBR GREEN.**

50 °C	95 °C	60 °C	95 °C	60 °C	95 °C
2 min	15 seg	1 min	15 seg	1 min	15 seg
		Captación de señal		Captación de señal cada 0.3 °C	
1 ciclo		45 ciclos		Melting curve	

Tabla suplementaria 2.16. **Secuencia de ciclos utilizada en la reacción de PCR cuantitativa con el sistema SYBR GREEN.**



## Consentimiento informado



**CARTA DE CONSENTIMIENTO INFORMADO PARA PARTICIPAR EN EL PROYECTO:  
PARTICIPACIÓN DEL GEN LRPPRC EN LA HIPERTRIGLICERIDEMIA SEVERA:  
ESTUDIO DE LOS MECANISMOS FISIOPATOLÓGICOS SUBYACENTES  
MARZO 2018, VERSIÓN 3**

**Investigador principal:** Dra. Ivette Cruz Bautista  
**Dirección del investigador:** Vasco de Quiroga 15 col. Sección XVI Tlalpan CP 14080.  
**Teléfono de contacto del investigador (incluyendo uno para emergencias 24 horas):** móvil: 5539550154; oficina 54870900 ext. 6322  
**Investigadores participantes:** Dr. Carlos A. Aguilar Salinas, Dra. Teresa Tusié Luna, Dra. Hortensia Moreno, Dra. Alicia Chagoya, Dra. Claudia J Bautista C., Dra. Paloma Almeda Valdés, Dra. Roopa Mehta, Dr. Alexandro Martagón Rosado.  
**Nombre del patrocinador del estudio:** Unidad de Investigación de Enfermedades Metabólicas. Departamento de Biología Molecular.  
**Dirección del patrocinador:** MISMO  
**Versión del consentimiento informado y fecha de su preparación:** VERSION 3, marzo de 2018.

### **INTRODUCCIÓN:**

Este documento es una invitación a participar en un estudio de investigación del Instituto. Por favor, tome todo el tiempo que sea necesario para leer este documento; pregunte al investigador sobre cualquier duda que tenga.

Procedimiento para dar su consentimiento. Usted tiene el derecho a decidir si quiere participar o no como sujeto de investigación en este proyecto. El investigador le debe explicar ampliamente los beneficios y riesgos del proyecto sin ningún tipo de presión y **usted tendrá todo el tiempo que requiera para pensar, solo o con quien usted decida consultarlo, antes de decidir si acepta participar.** Cualquiera que sea su decisión no tendrá efecto alguno sobre su atención médica en el Instituto.

Con el fin de tomar una decisión verdaderamente informada sobre si acepta participar o no en este estudio, usted debe tener el conocimiento suficiente acerca de los posibles riesgos y beneficios a su salud al participar. Este documento le dará información detallada acerca del estudio de investigación, la cual podrá comentar quien usted quiera, por ejemplo un familiar, su médico tratante, el investigador principal de este estudio o con algún miembro del equipo de investigadores. Al final, una vez leída y entendida esta información, se le invitará a que forme parte del proyecto y si usted acepta, sin ninguna presión o intimidación, se le invitará a firmar este consentimiento informado.

Este consentimiento informado cumple con los lineamientos establecidos en el Reglamento de la Ley General de Salud en Materia de Investigación para la Salud, la Declaración de Helsinki, y a las Buenas Prácticas Clínicas emitidas por la Comisión Nacional de Bioética.

Al final de la explicación, usted debe entender los puntos siguientes:

- I. La justificación y los objetivos de la investigación.
- II. Los procedimientos que se utilizarán y su propósito, incluyendo la identificación de qué son procedimientos experimentales.
- III. Los riesgos o molestias previstos.
- IV. Los beneficios que se pueden observar.
- V. Los procedimientos alternativos que pudieran ser ventajosos para usted
- VI. Garantía para recibir respuestas a las preguntas y aclarar cualquier duda sobre los procedimientos, riesgos, beneficios y otros asuntos relacionados con la investigación y el tratamiento de la materia.
- VII. La libertad que tiene de retirar su consentimiento en cualquier momento y dejar de participar en el estudio, sin que por ello se afecte su atención y el tratamiento en el Instituto.
- VIII. La seguridad de que no se le va a identificar de forma particular y que se mantendrá la

- IX. confidencialidad de la información relativa a su privacidad.
- IX. El compromiso del investigador de proporcionarle la información actualizada que pueda ser obtenida durante el estudio, aunque esto pudiera afectar a su disposición para continuar con su participación.
- X. La disponibilidad del tratamiento médico y compensación a que legalmente tiene derecho, en el caso de que ocurran daños causados directamente por la investigación.

**Puede solicitar más tiempo o llevar a casa este formulario antes de tomar una decisión final en los días futuros.**

## **INVITACION A PARTICIPAR COMO SUJETO DE INVESTIGACION Y DESCRIPCIÓN DEL PROYECTO**

Estimado(a) Sr(a). \_\_\_\_\_

El Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ), a través del grupo de investigación, le invitan a participar como sujeto de investigación en este estudio que tiene como objetivo: conocer los mecanismos mediante los cuales el gen *LRPPRC* se asocia a hipertrigliceridemia severa (niveles de triglicéridos en sangre arriba de 1000 mg/dl), a través de diversos estudios. La duración total del estudio es de 6 visitas en el lapso de un mes (se agendarán las fechas de las visitas según su conveniencia y disponibilidad).

Su participación en el estudio tendrá una duración variable por visita, dependerá del tipo de estudio para el cual usted acuda. El estudio más largo será el de la evaluación de la producción de lipoproteínas (grasas) por el hígado con duración de 12 horas seguido del estudio de lipemia postprandial (nivel de triglicéridos después de comer una comida) con duración de 8 horas.

El número aproximado de participantes será: 25. Usted fue invitado al estudio debido a que tiene las siguientes características:

1. Se le han detectado niveles de triglicéridos en sangre de más de 1000 mg/dl al menos en una ocasión.
2. Tiene triglicéridos debajo de 150 mg/dl
3. Puede ser portador o no portador del haplotipo de riesgo del gen *LRPPRC*

### **PROCEDIMIENTOS DEL ESTUDIO**

En este estudio no será aplicado ningún tratamiento experimental por lo que usted no será asignado al azar a algún grupo de estudio.

Los procedimientos del estudio incluyen la realización de:

- 1) Aplicación de un cuestionario estandarizado con duración de 15 min. En él se le solicitarán datos como edad, sexo, antecedentes familiares de enfermedad cardiovascular, pancreatitis, diabetes, elevación de colesterol y/o triglicéridos, enfermedades que padezca, medicamentos que esté tomando y dosis de los mismos
- 2) Medición de peso, talla, cintura, cadera y presión arterial
- 3) Se le pedirá seguir su dieta habitual antes de iniciar los estudios.
- 4) Se le pedirá que mantenga su actividad física habitual.
- 5) Se le pedirá que llene un registro de su dieta y cuestionario que mide actividad física cuatro semanas después.
- 6) Visita 1. Se le citará con un ayuno de 12 h, deberá acudir al sanitario antes de iniciar el estudio, Se tomarán muestras de sangre: química sanguínea (para evaluar sus niveles de azúcar en sangre y función renal), perfil de lípidos (para evaluar la cantidad de grasas presentes en la sangre como el colesterol y triglicéridos), hemoglobina glucosilada (para saber en caso de que usted padezca diabetes como esta su control de la enfermedad), nivel de insulina (que valora el funcionamiento de su páncreas). Se guardará suero de la toma inicial de sangre para medir niveles de otras proteínas en sangre que participan en el metabolismo de las grasas en el cuerpo (colesterol, triglicéridos). En esta misma visita se realizara una prueba de tolerancia a la

su páncreas). Se guardara suero de la toma inicial de sangre para medir niveles de otras proteínas en sangre que participan en el metabolismo de las grasas en el cuerpo (colesterol, triglicéridos). En esta misma visita se realizara una prueba de tolerancia a la glucosa por vía intravenosa para saber el grado de funcionamiento de su páncreas (órgano que produce la insulina). Se le colocaran 2 catéteres (uno en cada antebrazo). En un brazo tomaremos muestras de sangre y en el otro se le administrará glucosa, se le tomaran muestras en el minuto -10, -5, 0, 2, 3, 4, 5, 6, 8, 12, 14, 16, 19, 20 (en el minuto 20 se aplicara una dosis baja de insulina rápida calculada de acuerdo a su peso) 22, 24, 25, 27, 30, 40, 50, 60, 70, 90, 100, 120, 140, 160,180 para medir sus niveles de glucosa (azúcar) e insulina en algunos. La cantidad de sangre total será de 219 ml (22 cucharadas soperas).

- D) En esta misma visita, se realizara un estudio llamado absorciometría dual de rayos X (DXA). Con este estudio conoceremos la cantidad de grasa de su cuerpo en total y la que está en el abdomen (visceral). Se le pedirá llegar con ropa ligera y cómoda (por ejemplo pants), que no se marquen los resortes, no deberá haberse realizado ningún estudio con medio de contraste 7 días antes, no estar embarazada, no portar aretes, o piercings, anillos, pulseras o collares (nada de metal), tener ayuno de 4hrs, deberá pasar al baño antes del estudio. El estudio durara aproximadamente 40 minutos. Se retirará los zapatos, se acostará de espalda o boca arriba en una cama acojinada en una posición cómoda por unos minutos, mientras que el brazo del aparato le va tomando fotografías a todo lo largo de su cuerpo. **NO** se deberá mover durante el estudio. La radiación que recibirá su cuerpo es mínima, es mucho menor que una radiografía, es igual a un día de sol estando en el exterior.
- E) En esta misma visita se realizara un ultrasonido doppler (que mide el flujo de sangre en sus arterias y si hay algún contenido de grasa depositado) en el cuello
- F) En esta misma visita se le realizara la medición de grasa en su hígado , este estudio se identificará la cantidad de grasa y fibrosis en el hígado.La fibrosis hepática es la acumulación de tejido cicatrizal, resultado de la inflamación del hígado. Se produce en la mayoría de las enfermedades crónicas del hígado. Para la detección de la grasa hepática y fibrosis se le realizará un estudio que evalúa la elasticidad o rigidez del hígado llamado elastografía transitoria (Fibroscan). El Fibroscan funciona de manera similar a un aparato de ultrasonido. Esta prueba se realiza en la consulta externa. Usted deberá acudir con ayuno de cuarto horas. Las mediciones se realizan en el lado derecho del hígado; se apoyará la sonda sobre la piel entre las costillas y se realizarán 10 mediciones en el mismo punto sin que esto produzca molestias.
- 4) En la visita 2 acudirá a realizarse un estudio llamado curva de lipemia postprandial (después de alimentos). Mediremos su nivel de triglicéridos (grasas en sangre) después de darle una comida proporcionada por nosotros que deberá ingerir en 20 minutos. La comida consistirá en una hamburguesa con queso, con mayonesa, papas fritas y leche entera. La duración de este estudio será de 8 horas, permanecerá en la Unidad de Investigación de Enfermedades Metabólicas (UIEM). Se le tomarán muestras en los tiempos 0, 3, 4, 6 y 8 horas posteriores a la ingesta de la comida mixta. La cantidad de sangre tomada será de 30 ml, es decir, el equivalente a 3 cucharadas soperas)
- 5) En la visita 3 se realizarán una biopsia de grasa subcutánea alrededor del ombligo y/o de músculo (del muslo derecho o izquierdo). Deberá presentarse con ayuno de 4 horas en la UIEM. El procedimiento tendrá duración de 15 minutos. Un cirujano endocrinólogo será el encargado de realizar las biopsias. Se recostará en la mesa quirúrgica, se le aplicará anestesia local alrededor de su ombligo y se realizará una pequeña incisión de 1-2 cm de largo para extraer aproximadamente 2 gramos de tejido graso. La herida será cerrada usando pocos puntos de sutura. El mismo procedimiento se realizará en la parte anterior del muslo para obtener un fragmento de musculo. La cantidad de tejido que se obtendrá no representa riesgos para su salud, permanecerá la herida del muslo vendada por 48 h. Las muestras serán usadas para extraer RNA (material genético). Las biopsias se congelarán para analizarlas posteriormente, serán identificadas con un código con el fin de mantener su identidad confidencial. Se le

solicitará que acuda para revisar su herida y retirar los puntos de sutura 8 días después. Se le proporcionarán los números celulares de los investigadores para cualquier problema relacionado con este u otro estudio y se le proporcionará un analgésico para tomarlo en caso necesario.

- 6) En la visita 4 ( subgrupo seleccionado) . Se le podrá solicitar ( no a todos los pacientes) que acuda para toma de muestra adicional con la finalidad de realizar un estudio para evaluar la función respiratoria de sus células encargadas de producir energía ( llamadas mitocondrias) ya que el gen que se esta estudiando (LRPPRC) puede afectar su correcto funcionamiento. Esto no tiene una repercusión directa en su salud ya que usted cuenta con mecanismos compensatorios, los estudios se realizan exclusivamente para fines de investigación.
- 7) Visita 5 ( subgrupo seleccionados) . En un subgrupo de pacientes, usted puede ser elegido y se le podría citar una noche antes para darle una cena estandarizada y al día siguiente se iniciara el estudio de infusión de trazadores (sustancias como aminoácidos que son componentes de proteínas y grasas marcadas) para conocer como su hígado produce y elimina las grasas. Se tomarán 8 muestras de sangre en el transcurso de 12 horas. La cantidad de sangre que se obtendrá es de aproximadamente 150 ml (15 cucharadas soperas). Al final del estudio se le proporcionará un refrigerio.
- 8) Entrega de resultados.

Los estudios se completarán 5 sesiones (de acuerdo a la conveniencia y disponibilidad del participante). Las evaluaciones se programarán en un período de 30 días. La dosis total de sangre obtenida en todos los procedimientos es de 273 ml (equivalente a la mitad de lo que se obtiene cuando usted dona sangre, 27 cucharadas soperas).

Las intervenciones incluidas en el estudio que son parte de su tratamiento habitual (estándar) requerido para su condición son: la medición de sus niveles de glucosa, perfil de lípidos.

Las responsabilidades de los participantes incluyen: Reportar cambios de domicilio y se si agrego algún medicamento no consignado en su historia clínica inicial, avisar si se hospitaliza por cualquier causa y si será intervenido quirúrgicamente por cualquier causa

Se podrá pagar su pasaje ida y vuelta (solo de usted como paciente, no de acompañantes) en caso de que usted viva en (Edo de México, Cuernavaca, Toluca), y decida venir en transporte público.

### **RIESGOS E INCONVENIENTES**

El Reglamento de la Ley General de Salud en Materia de Investigación para la Salud, señala que la obtención de muestras biológicas representa un riesgo mínimo dentro de la investigación. Los riesgos de la toma de muestra sanguínea son: posibilidad de sangrado ligero o moretón en el sitio de la punción, mareo o sensación de desmayo y raramente puede producirse punción arterial. El personal que extraerá la muestra sanguínea está entrenado para ello, lo que minimizará los riesgos de complicaciones. Los datos acerca de su identidad y su información médica no serán revelados en ningún momento como lo estipula la ley, por tanto, en la recolección de datos clínicos usted no enfrenta riesgos mayores a los relativos a la protección de la confidencialidad la cual será protegida mediante la codificación de las muestras y de su información.

1. Muestras sanguíneas: Los riesgos de la toma de muestra sanguínea son: posibilidad de sangrado ligero o moretón en el sitio de la punción, mareo o sensación de desmayo y raramente puede producirse punción arterial. El personal que extraerá la muestra sanguínea está entrenado para ello, lo que minimizará los riesgos de complicaciones.

2. Las biopsias de grasa subcutánea y de musculo: pueden causar dolor, sangrado e infección. El dolor se minimizará indicándole la toma de una dosis baja de analgésico. Los de infección y sangrado son minimizados utilizando técnica quirúrgica estéril por personal calificado. Puede quedar una pequeña cicatriz en los sitios de la biopsia de 1-2cm aproximadamente.

3. Los estudios radiológicos (DXA y fibroscan) no tienen riesgos inherentes.

### **BENEFICIOS POTENCIALES**

Este estudio no está diseñado para beneficiarle directamente. Sin embargo, la búsqueda de mecanismos fisiopatológicos sobre la participación de este gen en la hipertrigliceridemia severa podría permitir desarrollar nuevos blancos terapéuticos y con esto hacer un tratamiento más personalizado de esta enfermedad. Además, gracias a su participación altruista, su comunidad se puede beneficiar significativamente al encontrar nuevas formas de atender esta complicación médica. A todos los participantes se les instruirá sobre la adopción de un estilo de vida saludable. A los pacientes con hipertrigliceridemia se ajustará el tratamiento farmacológico con el fin de alcanzar el mejor control posible y se dará orientación nutricional por una nutrióloga capacitada.

### **CONSIDERACIONES ECONÓMICAS**

No se cobrará ninguna tarifa por participar en el estudio ni se le hará pago alguno.

### **COMPENSACION**

Si llegara a presentarse alguna complicación como resultado directo de su participación en este estudio, por parte del protocolo le proporcionaremos el tratamiento inmediato y lo referiremos, si lo amerita, al especialista médico que requiera.

### **ALTERNATIVAS A SU PARTICIPACIÓN:**

Su participación es voluntaria. Por lo que usted puede elegir no participar en el estudio. En caso de decidir no participar, usted como derechohabiente del Instituto, seguirá recibiendo el tratamiento o manejo habitual (estándar) para su enfermedad, el cual consiste en tratamiento farmacológico hipolipemiente (fibratos), tratamiento nutricional.

### **POSIBLES PRODUCTOS COMERCIALES DERIVABLES DEL ESTUDIO:**

Los resultados o materiales obtenidos en el estudio serán propiedad del INCMNSZ. Si un producto comercial es desarrollado como resultado del estudio, tal insumo será propiedad del Instituto o quienes ellos designen. En tal caso, usted no recibirá un beneficio financiero por el mismo.

### **ACCIONES A SEGUIR DESPUÉS DEL TÉRMINO DEL ESTUDIO:**

Usted puede solicitar los resultados de sus exámenes clínicos a la Dra. Ivette Cruz Bautista (54870900 ext 6321, 6322 ). Cabe mencionar que la investigación es un proceso largo y complejo, por tanto, el obtener los resultados finales del proyecto puede tomar varios meses.

### **PARTICIPACIÓN Y RETIRO DEL ESTUDIO:**

Su participación es VOLUNTARIA. Si usted decide no participar, no se afectará su relación con el Instituto Nacional de Ciencias Médicas y Nutrición, Salvador Zubirán (INCMNSZ) o su derecho para recibir atención médica o cualquier servicio al que tenga derecho en caso de que usted tenga expediente en el Instituto. Si decide participar, tiene la libertad para retirar su consentimiento e interrumpir su participación en cualquier momento sin perjudicar su atención en el Instituto Nacional de Ciencias Médicas y Nutrición, Salvador Zubirán (INCMNSZ). Se le informará a tiempo si nueva información es obtenida que pueda afectar su decisión para continuar en el estudio.

El investigador o el patrocinador del estudio puede excluirlo del estudio si presenta otras alteraciones como son: nivel alto de apolipoproteína B, tiene pancreatitis aguda.

El estudio puede ser terminado en forma prematura si se presentan dificultades económicas para mantener operando el estudio.

Los procedimientos que serán necesarios si usted termina su participación en el estudio son: ninguno.

### **CONFIDENCIALIDAD Y MANEJO DE SU INFORMACIÓN**

Su nombre no será usado en ninguno de los reporte públicos del estudio. Las muestras biológicas obtenidas no contendrán ninguna información personal y serán codificadas con un número de serie

Los resultados de estudios genéticos podrían ser causa de discriminación para las personas que tengan alguna anomalía que los predisponga a una enfermedad. Tomaremos las precauciones y acciones necesarias para evitar que su información sea conocida por terceros que pudieran tomar acciones discriminatorias contra usted.

Los resultados de los estudios genéticos no serán incluidos en su expediente del Instituto, a menos que tengan implicaciones para su tratamiento.

#### **IDENTIFICACIÓN DE LOS INVESTIGADORES:**

En caso de que usted sufra un daño relacionado al estudio, por favor póngase en contacto con la Dra. Ivette Cruz Bautista, Investigadora C, adscrita al Departamento de Endocrinología y Metabolismo en el INCMNSZ al teléfono: 54870900 ext. 6322. Si usted tiene preguntas acerca de sus derechos como participante en el estudio, puede hablar con el Presidente del Comité de Ética en Investigación del INCMNSZ (Dr. Arturo Galindo. Teléfono: 54870900 ext. 6101).

#### **DECLARACIÓN DEL CONSENTIMIENTO INFORMADO**

He leído con cuidado este consentimiento informado, he hecho todas las preguntas que he tenido y todas han sido respondidas satisfactoriamente. Para poder participar en el estudio, estoy de acuerdo con todos los siguientes puntos:

Estoy de acuerdo en participar en el estudio descrito anteriormente. Los objetivos generales, particulares del reclutamiento y los posibles daños e inconvenientes me han sido explicados a mi entera satisfacción.

Estoy de acuerdo en donar de forma voluntaria mis muestras biológicas (sangre total, suero, plasma, tejido adiposo periumbilical) para ser utilizadas en éste estudio. Así mismo, mi información médica y biológica podrá ser utilizada con los mismos fines.

Estoy de acuerdo, en caso de ser necesario, que se me contacte en el futuro si el proyecto requiere coleccionar información adicional o si encuentran información relevante para mi salud.

Mi firma también indica que he recibido un duplicado de este consentimiento informado.

Por favor responda las siguientes preguntas:

	<b>SÍ</b> <b>(marque</b> <b>por</b> <b>favor)</b>	<b>NO</b> <b>(marque</b> <b>por</b> <b>favor)</b>
a. ¿Ha leído y entendido el formato de consentimiento informado, en su lengua materna?	<input type="checkbox"/>	<input type="checkbox"/>
b. ¿Ha tenido la oportunidad de hacer preguntas y de discutir este estudio?	<input type="checkbox"/>	<input type="checkbox"/>
c. ¿Ha recibido usted respuestas satisfactorias a todas sus preguntas?	<input type="checkbox"/>	<input type="checkbox"/>
d. ¿Ha recibido suficiente información acerca del estudio y ha tenido el tiempo suficiente para tomar la decisión?	<input type="checkbox"/>	<input type="checkbox"/>
e. ¿Entiende usted que su participación es voluntaria y que es libre de suspender su participación en este estudio en cualquier momento sin tener que justificar su decisión y sin que esto afecte su atención médica o sin la pérdida de los beneficios a los que de otra forma tenga derecho?	<input type="checkbox"/>	<input type="checkbox"/>
g. ¿Entiende los posibles riesgos, algunos de los cuales son aún desconocidos, de participar en este estudio?	<input type="checkbox"/>	<input type="checkbox"/>

	SÍ (marque por favor)	NO (marque por favor)
h. ¿Entiende que puede no recibir algún beneficio directo de participar en este estudio?	<input type="checkbox"/>	<input type="checkbox"/>
i. ¿Ha discutido usted otras opciones de tratamiento con el médico participante en el estudio y entiende usted que otras opciones de tratamiento están a su disposición?	<input type="checkbox"/>	<input type="checkbox"/>
j. ¿Entiende que no está renunciando a ninguno de sus derechos legales a los que es acreedor de otra forma como sujeto en un estudio de investigación?	<input type="checkbox"/>	<input type="checkbox"/>
k. ¿Entiende que el médico participante en el estudio puede retirarlo del mismo sin su consentimiento, ya sea debido a que usted no siguió los requerimientos del estudio o si el médico participante en el estudio considera que médicamente su retiro es en su mejor interés?	<input type="checkbox"/>	<input type="checkbox"/>
l. ¿Entiende que el estudio puede ser suspendido por el patrocinador del estudio en cualquier momento?	<input type="checkbox"/>	<input type="checkbox"/>
m. ¿Entiende que usted recibirá un original firmado y fechado de esta Forma de Consentimiento para sus registros personales?	<input type="checkbox"/>	<input type="checkbox"/>

**Declaración del paciente:** Yo, \_\_\_\_\_ declaro que es mi decisión participar como sujeto de investigación clínica en el estudio. Mi participación es voluntaria. Se me ha informado que puedo negarme a participar o terminar mi participación en cualquier momento del estudio sin que sufra penalidad alguna o pérdida de beneficios. Si suspendo mi participación, recibiré el tratamiento médico habitual al que tengo derecho en el Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ) y no sufriré perjuicio en mi atención médica ni en futuros estudios de investigación. Yo puedo solicitar información adicional acerca de los riesgos o beneficios potenciales derivados de mi participación en este estudio. También puedo obtener los resultados de mis exámenes clínicos si así los solicito.

Si tengo preguntas sobre el estudio, puede ponerse en contacto la Dra. Ivette Cruz Bautista, responsable del proyecto en el Departamento de Endocrinología del INCMNSZ teléfono: 54870900 ext. 6322. Si usted tiene preguntas sobre sus derechos como participante en el estudio, problemas, preocupaciones o preguntas, obtener información, y ofrecer información que puede hablar con el Presidente del Comité de Ética de Investigación de INCMNSZ Dr. Arturo Galindo. Teléfono: 54870900 ext. 6101).

Debo informar a los investigadores de cualquier cambio en mi estado de salud (por ejemplo, uso de nuevos medicamentos, cambios en el consumo de tabaco) o en la ciudad donde resido, tan pronto como sea posible.

He leído y entendido toda la información que me han dado sobre mi participación en el estudio. He tenido la oportunidad para discutirlo y hacer preguntas. Todas las preguntas han sido respondidas a mi satisfacción. He entendido que recibiré una copia firmada de este consentimiento informado.

Tengo claro que en caso de tener preguntas sobre mis derechos como sujeto de investigación clínica en este estudio, problemas, preocupaciones o dudas, y deseo obtener información adicional, o bien, hacer comentarios sobre el desarrollo del estudio, tengo la libertad de hablar con el presidente del Comité de Ética en Investigación del INCMNSZ Dr. Arturo Galindo Fraga, teléfono: 54870900. Ext. 6101).

\_\_\_\_\_  
Nombre del / de la Participante

\_\_\_\_\_  
Firma del / de la Participante

\_\_\_\_\_  
Fecha

\_\_\_\_\_  
Coloque la huella digital del participante sobre esta línea si no sabe escribir



\_\_\_\_\_  
Nombre del representante legal (si aplica)

\_\_\_\_\_  
Firma del representante legal

\_\_\_\_\_  
Fecha

\_\_\_\_\_  
Nombre del Investigador  
que explicó el documento

\_\_\_\_\_  
Firma del Investigador

\_\_\_\_\_  
Fecha

\_\_\_\_\_  
Nombre del Testigo 1

\_\_\_\_\_  
Firma del Testigo 1

\_\_\_\_\_  
Fecha

\_\_\_\_\_  
Relación con el participante:

Dirección: \_\_\_\_\_  
\_\_\_\_\_

\_\_\_\_\_  
Nombre del Testigo 2

\_\_\_\_\_  
Firma del Testigo 2

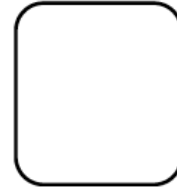


\_\_\_\_\_  
Nombre del / de la Participante

\_\_\_\_\_  
Firma del / de la Participante

\_\_\_\_\_  
Fecha

\_\_\_\_\_  
Coloque la huella digital del participante sobre esta línea si no sabe escribir



\_\_\_\_\_  
Nombre del representante legal (si aplica)

\_\_\_\_\_  
Firma del representante legal

\_\_\_\_\_  
Fecha

\_\_\_\_\_  
Nombre del Investigador  
que explicó el documento

\_\_\_\_\_  
Firma del Investigador

\_\_\_\_\_  
Fecha

\_\_\_\_\_  
Nombre del Testigo 1

\_\_\_\_\_  
Firma del Testigo 1

\_\_\_\_\_  
Fecha

\_\_\_\_\_  
Relación con el participante:

Dirección: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_  
Nombre del Testigo 2

\_\_\_\_\_  
Firma del Testigo 2

## Carta de aprobación anual comité de ética ultimo año 2021



**SALUD**  
SECRETARÍA DE SALUD



INSTITUTO NACIONAL DE  
CIENCIAS MÉDICAS  
Y NUTRICIÓN  
SALVADOR ZUBIRÁN

Ciudad de México, a 27 de noviembre de 2020  
Oficio No. Mcontrol-1536/2020  
REG. CONBIOÉTICA-09-CEI-011-20160627

**DRA. IVETTE CRUZ BAUTISTA**  
**INVESTIGADORA PRINCIPAL**  
**UNIDAD DE INVESTIGACIÓN EN ENFERMEDADES METABÓLICAS**  
**INSTITUTO NACIONAL DE CIENCIAS MÉDICAS Y NUTRICIÓN SALVADOR ZUBIRÁN**  
**AV. VASCO DE QUIROGA No. 15**  
**COL. BELISARIO DOMÍNGUEZ SECCIÓN XVI**  
**DEL. TLALPAN, CIUDAD DE MÉXICO, C.P. 14080**  
**PRESENTE**

En respuesta a su atenta del pasado 03 de noviembre, en relación al Protocolo de Investigación Clínica, titulado:

**"PARTICIPACIÓN DEL GEN LRPPRC EN LA HIPERTRIGLICERIDEMIA SEVERA: ESTUDIO DE LOS MECANISMOS FISIOPATOLÓGICOS SUBYACENTES"**

**REF. 2050**

---

## 12. BIBLIOGRAFÍA

- <sup>1</sup> Organización Mundial de la Salud (OMS), Global Health Estimates 2019: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. Geneva, World Health Organization; 2020.
- <sup>2</sup> OMS, World health statistics 2016: monitoring health for the SDGs, sustainable development goals.
- <sup>3</sup> Instituto Nacional de Estadística y Geografía (INEGI), Principales causas de mortalidad por residencia habitual, grupos de edad y sexo del fallecido. Año de registro: 2018. Tabuladores básicos.
- <sup>4</sup> Rygiel, Katarzyna. Hypertriglyceridemia - Common Causes, Prevention and treatment Strategies. *Current Cardiology Reviews*, 2018, 14, 67-76
- <sup>5</sup> Secretaría de Salud. NORMA Oficial Mexicana NOM-037-SSA2-2012, Para la prevención, tratamiento y control de las dislipidemias.
- <sup>6</sup> Aguilar Salinas, CA. et al. Diagnóstico y tratamiento de las dislipidemias: posición de la Sociedad Mexicana de Nutrición y Endocrinología. *Revista de Endocrinología y Nutrición* Vol. 12, No. 1, Enero-Marzo 2004, pp 7-41.
- <sup>7</sup> Catapano, A. et al. ESC Scientific Document Group, 2016 ESC/EAS Guidelines for the Management of Dyslipidaemias, *European Heart Journal*, Volume 37, Issue 39, 14 October 2016, Pages 2999-3058, <https://doi.org/10.1093/eurheartj/ehw272>
- <sup>8</sup> Aguilar Salinas, Carlos Alberto, *Dislipidemias. De lo clínico a lo molecular*. Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INSMNSZ), (México: Intersistemas, 2008).
- <sup>9</sup> Aguilar-Salinas, et al., "Prevalence of dyslipidemias in the 2006" *Encuesta Nacional de Salud y Nutrición: 2009 Salud Pública Méx. 2010; 52(sup1): 44-53*.
- <sup>10</sup> Weissglas-Volkov, D. et al. Investigation of variants identified in Caucasian genome-wide association studies for plasma HDL cholesterol and triglycerides levels in Mexican dyslipidemic study samples. *Circ Cardiovasc Genet*. 2010 February 1; 3(1): 31-38. doi:10.1161/CIRCGENETICS.109.908004.
- <sup>11</sup> Roubille, F. et al. Is hypertriglyceridemia atherogenic? *Presse Med.* (2018), <https://doi.org/10.1016/j.lpm.2018.08.009>
- <sup>12</sup> Peng et al. Hypertriglyceridemia and atherosclerosis. *Lipids in Health and Disease* (2017) 16:233 DOI 10.1186/s12944-017-0625-0
- <sup>13</sup> Do R, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet*. 2013;45:1345-52
- <sup>14</sup> Johansen, T; Hegele, R. Allelic and phenotypic spectrum of plasma triglycerides. *Biochimica et Biophysica Acta* 1821 833-842 (2012) doi:10.1016/j.bbali.2011.10.007
- <sup>15</sup> Real Academia Española (RAE), *Diccionario de la lengua española*, 23ª ed. <http://www.rae.es/rae.html>
- <sup>16</sup> Aguilar, C, Tusié T, Pajukanta, P. Genetic and environmental determinants of the susceptibility of Amerindian derived populations for having hypertriglyceridemia. *Metabolism* (2014), <http://dx.doi.org/10.1016/j.metabol.2014.03.012>
- <sup>17</sup> Durrington, P, "Dyslipidaemia", *The Lancet*, Vol. 362, agosto 2003.
- <sup>18</sup> Diccionario médico. Clínica Universidad de Navarra. Disponible en: <http://www.cun.es/diccionario-medico>
- <sup>19</sup> Andrea Sacchi, Liliana Olivares, Viviana Leiro. Xantomatosis y dislipoproteinemias. *Dermatol. Argent*, 2013, 19(3): 174-182
- <sup>20</sup> Corral. P. Arco corneal: ¿signo del paso del tiempo o de riesgo cardiovascular? Artículo de revisión. *Rev Fed Arg Cardiol*. 2015; 44(2): 78-81
- <sup>21</sup> Harrison. Principios de Medicina Interna. Vol. 2, 18ª ed. (México: MacGraw Hill, 2012).
- <sup>22</sup> Brunzell JD. "Familial Lipoprotein Lipase Deficiency", University of Washington <https://www.ncbi.nlm.nih.gov/books/NBK1308/>
- <sup>23</sup> Aguilar, CA. et al. "La hipertrigliceridemia familiar no se asocia a mayor prevalencia de complicaciones macrovasculares en la diabetes mellitus tipo 2", *Gaceta Médica de México* vol.141 no.3 may./jun. 2005
- <sup>24</sup> Teslovich, T., Musunuru, K., Smith, A. et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 466, 707-713 (2010). <https://doi.org/10.1038/nature09270>
- <sup>25</sup> Weissglas-Volkov, D. et al. Genomic study in Mexicans identifies a new locus for triglycerides and refines European lipid loci. *Med Genet* 2013;50:298-308. doi:10.1136/jmedgenet-2012-101461
- <sup>26</sup> Cruz-Bautista et al. Familial hypertriglyceridemia: an entity with distinguishable features from other causes of hypertriglyceridemia. *Lipids in Health and Disease* (2021) 20:14. <https://doi.org/10.1186/s12944-021-01436-6>
- <sup>27</sup> Encuesta Nacional de Enfermedades Crónicas 1993 ENEC93, Instituto Nacional de Salud Pública

- 
- <sup>28</sup> Encuesta Nacional de Salud 2000 (ENSA 2000), Instituto Nacional de Salud Pública
- <sup>29</sup> Encuesta Nacional de Salud y Nutrición 2006 (ENSANUT2006), Instituto Nacional de Salud Pública
- <sup>30</sup> Ford ES, et al., "Hypertriglyceridemia and its pharmacologic treatment among US adults", *Archives of internal medicine* 169, 2009:572–578.
- <sup>31</sup> Duggirala Ravindranath., et al., "A Major Susceptibility Locus Influencing Plasma Triglyceride Concentrations Is Located on Chromosome 15q in Mexican Americans", *American Journal of Human Genetics*. 66, 2000:1237–1245.
- <sup>32</sup> Aguilar S, C. et al. GENÉTICA DE LAS DISLIPIDEMIAS EN MÉXICO. (2016) MENSAJE BIOQUÍMICO, VOL. XL
- <sup>33</sup> Price, A.L. et al. A Genomewide Admixture Map for Latino Populations. *Am. J. Hum. Genet.* 2007;80:1024–1036. DOI: 10.1086/518313
- <sup>34</sup> Lewis, G. et al. Hypertriglyceridemia in the Genomic Era: A New Paradigm. *Endocrine Reviews*, , 36(1):131–147 (2015) doi: 10.1210/er.2014-1062
- <sup>35</sup> Raimondo, A. et al. Glucokinase regulatory protein: complexity at the crossroads of triglyceride and glucose metabolism. *Curr Opin Lipidol.* 2015 April ; 26(2): 88–95. doi:10.1097/MOL.0000000000000155
- <sup>36</sup> Beer, NL. Et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver, *Hum. Mol. Genet.* 18 (2009) 4081–4088
- <sup>37</sup> Chávez, A (2018). Evaluación de la interacción entre el SNP intrónico rs58811869 del gen LRPPRC con el factor de transcripción CTCF. (Tesis de pregrado) Facultad de Química, UNAM. México
- <sup>38</sup> Ohlsson, R. et al., "CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Review", *TRENDS in Genetics*, 2001, 17, 9
- <sup>39</sup> Ensembl. Human (GRCh38.p13). Gene: LRPPRC ENSG00000138095  
[https://www.ensembl.org/Homo\\_sapiens/Gene/Summary?db=core;g=ENSG00000138095;r=2:43886224-43996226](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000138095;r=2:43886224-43996226)
- <sup>40</sup> Small, Ian D., N. Peeters, "The PPR motif - a TPR-related motif prevalent in plant organellar proteins, *Trends Biochem*", *Sci* 25 (2000): 46-47.
- <sup>41</sup> Saha, D., et al., "Pentatricopeptide repeat proteins and their emerging roles in plants. Review". *Plant Physiology and Biochemistry* 45 (2007): 521-534.
- <sup>42</sup> Mili, Stravoula et al., "LRP130, a Pentatricopeptide Motif Protein with a Noncanonical RNA-Binding Domain, Is Bound In Vivo to Mitochondrial and Nuclear RNAs", *Molecular and cellular biology* 14, 23 (julio 2003): 4972-4982.
- <sup>43</sup> Delannoy, Etienne et al., "Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles", *Biochemical Society Transactions* 35, part 6 (2007): 1643-1647. doi: 10.1042/BST0351643.
- <sup>44</sup> Lightowers, Robert et al., "PPR (pentatricopeptide repeat) proteins in mammals: important aids to mitochondrial gene expression", *Biochemical Journal* 416 (2008):e5-e6. doi: 10.1042/BJ20081942.
- <sup>45</sup> Rackham, Oliver et al. "The role of mammalian PPR domain proteins in the regulation of mitochondrial gene expression", *Biochimica et Biophysica Acta* 1819 (2011): 1008–1016.
- <sup>46</sup> Sterky, Fredrik et al. "LRPPRC is a mitochondrial matrix protein that is conserved in metazoans", *Biochemical and Biophysical Research Communications* 398 (2010): 759–764.
- <sup>47</sup> Avedaño-Monsalve, MC. et al. Viaje al centro de la mitocondria: importación de proteínas, sus alteraciones y enfermedades relacionadas. *TIP Revista Especializada en Ciencias Químico-Biológicas*, 24: 1-25, 2021. <https://doi.org/10.22201/fesz.23958723e.2021.370>
- <sup>48</sup> Mili, Stravoula et al., "LRP130, a Pentatricopeptide Motif Protein with a Noncanonical RNA-Binding Domain, Is Bound In Vivo to Mitochondrial and Nuclear RNAs", *Molecular and cellular biology* 14, 23 (julio 2003): 4972-4982.
- <sup>49</sup> Tsuchiya, N. et al. LRP130, a single-stranded DNA/RNA-binding protein, localizes at the outer nuclear and endoplasmic reticulum membrane, and interacts with mRNA in vivo. *Biochemical and Biophysical Research Communications* 317 (2004) 736–743
- <sup>50</sup> Sira, S. et al. LRPPRC-mediated folding of the mitochondrial transcriptome. *NATURE COMMUNICATIONS*. (2017) 8: 1532. DOI: 10.1038/s41467-017-01221-z
- <sup>51</sup> Liu, L. et al., "LRP130 Protein Remodels Mitochondria and Stimulates Fatty Acid Oxidation", *The Journal of Biological Chemistry* 286, 48 (diciembre 2011): 41253–41264.
- <sup>52</sup> Akie TE, et al. OXPHOS-Mediated Induction of NAD+ Promotes Complete Oxidation of Fatty Acids and Interdicts Non-Alcoholic Fatty Liver Disease. *PLoS ONE*, (2015, 10(5): e0125617. doi:10.1371/journal.pone.0125617
- <sup>53</sup> Young, M. Copeland, W. Human mitochondrial DNA replication machinery and disease. *Current Opinion in Genetics & Development* 2016, 38:52–62 <http://dx.doi.org/10.1016/j.gde.2016.03.005>

- <sup>54</sup> Miller, F. et al. Precise determination of mitochondrial DNA copy number in human skeletal and cardiac muscle by a PCR-based assay: lack of change of copy number with age. *Nucleic Acids Research*, 2003, Vol. 31, No. 11 e61 DOI: 10.1093/nar/gng060
- <sup>55</sup> Gohil, V. M. et al., "Mitochondrial and Nuclear Genomic Responses to Loss of LRPPRC Expression", *The Journal of Biological Chemistry* 285, 18 (abril 2010): 13742–13747.
- <sup>56</sup> Ruzzenente, B. et al., "LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs". *The EMBO Journal* 31(2012): 443–456. doi:10.1038/emboj.2011.392.
- <sup>57</sup> Mili, Stavroula et al., "Distinct RNP Complexes of Shuttling hnRNP Proteins with Pre-mRNA and mRNA: Candidate Intermediates in Formation and Export of mRNA", *Molecular and Cellular Biology* 21, 21 (noviembre 2001): 7307–7319.
- <sup>58</sup> Sasarman, Florine et al., "LRPPRC and SLIRP Interact in a Ribonucleoprotein Complex That Regulates Posttranscriptional Gene Expression in Mitochondria", *Molecular Biology of the Cell* 21: 1315–1323.
- <sup>59</sup> Lei, S. et al., "Increased Hepatic Fatty Acids Uptake and Oxidation by LRPPRC-Driven Oxidative Phosphorylation Reduces Blood Lipid Levels", *Front. Physiol* 7(2016):270. doi: 10.3389/fphys.2016.00270.
- <sup>60</sup> Mootha, V.K., P. Lepage, K. Miller et al., "Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics". *Proceedings of the National Academy of Sciences of the United States of America* 100, 2 (2003): 605–610. doi:10.1073/pnas.242716699.
- <sup>61</sup> Morin, C. et al., "Clinical, metabolic, and genetic aspects of cytochrome C oxidase deficiency in Saguenay-Lac-Saint-Jean", *American Journal of Human Genetics* 53, 2 (1993):488–496.
- <sup>62</sup> Leigh syndrome, French Canadian type. Genetic Testing Registry. NCBI.
- <sup>63</sup> Jin, T. et al., "Clinical, Pathological, and Neuroimaging Analyses of Two Cases of Leigh Syndrome in a Chinese Family", *Journal of Child Neurology* 29, 11. doi: 10.1177/0883073813512524.
- <sup>64</sup> Rakovic, A. et al. PINK1-Interacting Proteins: Proteomic Analysis of Overexpressed PINK1. *Parkinson's Disease Volume 2011*, Article ID 153979, 9 pages. doi:10.4061/2011/153979
- <sup>65</sup> Gaweda-Walerych, K. et al. Parkinson's disease-related gene variants influence pre-mRNA splicing processes. *Neurobiology of Aging* 47 (2016) 127e138
- <sup>66</sup> Cui J. et al, LRPPRC: A Multifunctional Protein Involved in Energy Metabolism and Human Disease. *Front. Physiol.* 10:595. (2019) doi: 10.3389/fphys.2019.00595
- <sup>67</sup> Oehler, V. G. et al. Differential gene expression associated with chronic myeloid leukemia (Cml) progression predicts relapse and survival prior to allogeneic transplantation in chronic phase Cml patients. *Blood* 116, (2010).
- <sup>68</sup> Fahrman, J. F. et al. Proteomic profiling of lung adenocarcinoma indicates heightened DNA repair, antioxidant mechanisms and identifies LASP1 as a potential negative predictor of survival. *Clin. Proteomics* 13, 1–12, (2016). doi: 10.1186/s12014-016-9132-y
- <sup>69</sup> Li, X. et al. Reversal of multidrug resistance in gastric cancer cells by LRPPRC downregulation. *J. Dig. Dis.* 15:84, (2014).
- <sup>70</sup> Gao, W et al. Mitochondrial proteomics approach reveals voltage-dependent anion channel 1 (VDAC1) as a potential biomarker of gastric cancer. *Cell. Physiol. Biochem.* 37, 2339–2354. (2015) doi: 10.1159/000438588
- <sup>71</sup> Jiang, X et al. Elevated levels of mitochondrion-associated autophagy inhibitor LRPPRC are associated with poor prognosis in patients with prostate cancer. *Cancer* 120, 1228–1236. (2014) doi: 10.1002/cncr.28551
- <sup>72</sup> Jiang, X. et al. Autophagy defects suggested by low levels of autophagy activator Map1s and high levels of autophagy inhibitor LRPPRC predict poor prognosis of prostate cancer patients. *Mol. Carcinog.* 54, 1194–1204. (2015) doi: 10.1002/mc.22193
- <sup>73</sup> Zhang, H. Y. et al. Elevated levels of autophagy-related marker ULK1 and mitochondrion-associated autophagy inhibitor LRPPRC are associated with biochemical progression and overall survival after androgen deprivation therapy in patients with metastatic prostate cancer. *J. Clin. Pathol.* 70, 383–389. (2017) doi: 10.1136/jclinpath-2016-203926
- <sup>74</sup> Nishio, T. et al. Immunohistochemical expression analysis of leucine-rich PPR-motif-containing protein (LRPPRC), a candidate colorectal cancer biomarker identified by shotgun proteomics using iTRAQ. *Clin. Chim. Acta* 471, 276–282. (2017) doi: 10.1016/j.cca.2017.06.011
- <sup>75</sup> Bhat, S. et al. Gene promoter-associated CpG Island hypermethylation in squamous cell carcinoma of the tongue. *Virchows Arch.* 470, 445–454. (2017) doi: 10.1007/s00428-017-2094-2
- <sup>76</sup> Tian, T. et al. Role of leucine-rich pentatricopeptide repeat motif-containing protein (LRPPRC) for anti-apoptosis and tumorigenesis in cancers. *Eur. J. Cancer* 48, 2462–2473. (2012) doi:10.1016/j.ejca.2012.01.018
- <sup>77</sup> Cunha, M. L. R. et al. Whole exome sequencing in thrombophilic pedigrees to identify genetic risk factors for venous thromboembolism. *PLoS ONE* 2017, 12:e0187699. doi: 10.1371/journal.pone.0187699

- <sup>78</sup> Schweitzer, C. J. et al. Knockdown of the cellular protein LRPPRC attenuates HIV-1 infection. *PLoS ONE* 7, 2012, doi: 10.1371/journal.pone.0040537
- <sup>79</sup> Arun, V. et al. A novel neurofibromin (Nf1) interaction with the leucine-rich pentatricopeptide repeat motif-containing protein links neurofibromatosis type 1 and the French Canadian variant of Leigh's syndrome in a common molecular complex. *J. Neurosci. Res.* 2013, 91, 494–505. doi: 10.1002/jnr.23189
- <sup>80</sup> Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the Deep-sequencing age. *Nat Methods.* 2014 Apr;11(4):361-2
- <sup>81</sup> FO Desmet, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acid Research*, 2009
- <sup>82</sup> K.H. Lim, W.G. Fairbrother. Spliceman - A computational web server that predicts sequence variations in pre-mRNA splicing. *Bioinformatics*, 2012; DOI: 10.1093/bioinformatics/bts074
- <sup>83</sup> Hou, J. et al. Molecular cloning and expression of the gene for a major Leucine-rich protein from human hepatoblastoma cells (HepG2), *In Vitro Cell. Dev. Biol.* 1994, 30A:III-II4, <https://doi.org/10.1007/BF02631402>
- <sup>84</sup> National Center for Biotechnology Information (NCBI), "GUSB glucuronidase beta [Homo sapiens (human)]" Gene ID: 2990. <https://www.ncbi.nlm.nih.gov/gene/2990>
- <sup>85</sup> HUGO Gene Nomenclature Committee. HGNC "Gene group: 18S ribosomal RNAs (RNA18S). <https://www.genenames.org/data/genegroup/#!/group/1383>
- <sup>86</sup> National Center for Biotechnology Information (NCBI), "GAPDH glyceraldehyde-3-phosphate dehydrogenase [Homo sapiens (human)]" Gene ID: 2597, <https://www.ncbi.nlm.nih.gov/gene/2597>
- <sup>87</sup> National Center for Biotechnology Information (NCBI), "ACTB actin beta [Homo sapiens (human)]" Gene ID: 60. <https://www.ncbi.nlm.nih.gov/gene/60>
- <sup>88</sup> Pott, S y Lieb, JD. What are super-enhancers?. *Nature Genetics*, 2015, Vol 47, No 1. doi:10.1038/ng.3167
- <sup>89</sup> Vergara, C. et al. Mapeo Genético mediante desequilibrio de ligamiento por mezcla y su aplicación en enfermedades complejas. *Rev Cubana Invest Bioméd*, jun 2008, Ciudad de la Habana, v. 27, n. 2. Disponible en <[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0864-03002008000200005&lng=es&nrm=iso](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-03002008000200005&lng=es&nrm=iso)>.
- <sup>90</sup> The SIGMA (Slim Initiative in Genomic Medicine for the Americas) Type 2 Diabetes Consortium. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico, *Nature*. 2014 Feb 6; 506(7486): 97–101. PMID: 24390345. doi: 10.1038/nature12828
- <sup>91</sup> Xu, F. et al. The role of the LRPPRC (leucine-rich pentatricopeptide repeat cassette) gene in cytochrome oxidase assembly: mutation causes lowered levels of COX (cytochrome c oxidase) I and COX III mRNA. *Biochem. J.* (2004) 382, 331–336
- <sup>92</sup> Haslam, D, James P. Obesity. *Seminars. Lancet* 2005; 366: 1197–209
- <sup>93</sup> Piché, M et al. Obesity Phenotypes, Diabetes, and Cardiovascular Diseases. *OBESITY, METABOLIC SYNDROME, AND CVD COMPENDIUM. Circulation Research.* 2020;126:1477–1500. DOI: 10.1161/CIRCRESAHA.120.316101
- <sup>94</sup> Pérez-Aguilar, F. et al. Esteatohepatitis no alcohólica: consideraciones fisiopatológicas, clínicas y terapéuticas. *Rev. esp. enferm. dig.* vol.96 no.9 Madrid sep. 2004
- <sup>95</sup> Dornas, W y Schuppan, D. Mitochondrial oxidative injury: a key player in nonalcoholic fatty liver disease. Mini-Review. *AJF. Liver and Biliary Tract Physiology/Pathophysiology.* Volume 319, Issue 3, September 2020 <https://doi.org/10.1152/ajpgi.00121.2020>. PubMed32597705
- <sup>96</sup> Lee, K. et al. epatic Mitochondrial Defects in a Nonalcoholic Fatty Liver Disease Mouse Model Are Associated with Increased Degradation of Oxidative Phosphorylation Subunits. 2018, *Molecular & Cellular Proteomics* 17, 2371–238. <https://doi.org/10.1074/mcp.RA118.000961>
- <sup>97</sup> Nakamuta, M. et al. The significance of differences in fatty acid metabolism between obese and non-obese patients with non-alcoholic fatty liver disease. 2008. *INTERNATIONAL JOURNAL OF MOLECULAR MEDICINE* 22: 663-667. DOI: 10.3892/ijmm\_00000070
- <sup>98</sup> Fischer, B. et al. Inverse relationship between body mass index and mitochondrial oxidative phosphorylation capacity in human subcutaneous adipocytes. *AJP. Endocrinology and Metabolism.* Volume 309. Issue 4 August 2015 <https://doi.org/10.1152/ajpendo.00524.2014>
- <sup>99</sup> Rocha, P. et al. Breaking TADs: insights into hierarchical genome organization. *Epigenomics* (2015) 7(4), 523–526
- <sup>100</sup> Perisic, L. et al. Plekhh2, a novel podocyte protein downregulated in human focal segmental glomerulosclerosis, is involved in matrix adhesion and actin dynamics, *Kidney International* (2012) 82, 1071–1083; doi:10.1038/ki.2012.252
- <sup>101</sup> DYNC2LI1 dynein cytoplasmic 2 light intermediate chain 1 [Homo sapiens (human)] NCBI Gene. Gene ID: 51626, updated on 5-Apr-2020
- <sup>102</sup> Park JH. et al. PPM1B depletion induces premature senescence in human IMR-90 fibroblasts. *Mech Ageing Dev.* 2014 Jun;138:45-52. doi: 10.1016/j.mad.2014.03.003

- <sup>103</sup> SLC3A1 solute carrier family 3 member 1 [Homo sapiens (human)] NCBI Gene. Gene ID: 6519, updated on 13-Mar-2020
- <sup>104</sup> Silva, S. et al. The second point mutation in PREPL: a case report and literature review. *Journal of Human Genetics*, 2018, 63(5), 677–681. doi:10.1038/s10038-018-0426-y
- <sup>105</sup> CAMKMT calmodulin-lysine N-methyltransferase [Homo sapiens (human)] NCBI Gene. Gene ID: 79823, updated on 5-Apr-2020
- <sup>106</sup> Gallardo, E. Genes humanos de la familia six/sine oculis, genética molecular e implicaciones en malformaciones congénitas. Tesis. <https://dialnet.unirioja.es/servlet/tesis?codigo=197858>
- <sup>107</sup> Lee, J. et al. Crystal structure of the human sterol transporter ABCG5/ABCG8. *Nature*, 2016, 533(7604), 561–564. doi:10.1038/nature17666
- <sup>108</sup> Han, J. et al. 3C and 3C-based techniques: the powerful tools for spatial genome organization deciphering. *Molecular Cytogenetics* (2018) 11:21 <https://doi.org/10.1186/s13039-018-0368-2>
- <sup>109</sup> Buroker, N. Regulatory SNPs and transcription factor binding sites in ADRBK1, AKT3, ATF3, DIO2, TBXA2R and VEGFA. 2014, *Transcription*, e964559-2. <http://dx.doi.org/10.4161/21541264.2014.964559>
- <sup>110</sup> Gaszner, M. y Felsenfeld, G. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* 7, 703–713 (2006). <https://doi.org/10.1038/nrg1925>
- <sup>111</sup> Kaiser, V. B. y C. A. Semple, "When TADs go bad: chromatin structure and nuclear organization in human disease", *F1000Research* 2017, 6: *F1000 Faculty Rev-314*. doi:10.12688/f1000research.10792.1.
- <sup>112</sup> Wang, Y. et al. NR6A1 regulates lipid metabolism through mammalian target of rapamycin complex 1 in HepG2 cells. *Cell Communication and Signaling*, 2019, 17:77 <https://doi.org/10.1186/s12964-019-0389-4>
- <sup>113</sup> Xu, Z, et al.,. Liver-specific inactivation of the Nrf1 gene in adult mouse leads to nonalcoholic steatohepatitis and hepatic neoplasia. 2005, *Proc. Natl. Acad. Sci. U. S. A.* 102, 4120–4125.
- <sup>114</sup> Hirotsu, Y. et al. NF-E2-Related Factor 1 (Nrf1) Serves as a Novel Regulator of Hepatic Lipid Metabolism through Regulation of the Lipin1 and PGC-1 $\beta$  Genes. *Molecular and Cellular Biology*, 2012, p 2760–2770, doi:10.1128/MCB.06706-11
- <sup>115</sup> Ming Kim, N. et al. Nuclear Factor Erythroid-2 Like 1 (NFE2L1): Structure, function and regulation *Gene* 584 (2016) 17–25 <http://dx.doi.org/10.1016/j.gene.2016.03.002>
- <sup>116</sup> Pajukanta P. et al. Linkage of familial combined hyperlipidemia to chromosome 1q21–q23. *Nat Genet.* 1998; 18: 369–373
- <sup>117</sup> Pei W. et al. Support for linkage of familial combined hyperlipidemia to chromosome 1q21–q23 in Chinese and German families. *Clin Genet.* 2000; 57: 29–34.
- <sup>118</sup> Huertas-Vazquez A. et al. Contribution of Chromosome 1q21–q23 to Familial Combined Hyperlipidemia in Mexican Families. *Ann Hum Genet.* 2004; 68: 419–427
- <sup>119</sup> Pajukanta, P. et al. Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). *Nat Genet.* 2004; 36: 371–376
- <sup>120</sup> Guo, J. et al. Upstream stimulating factor 1 suppresses autophagy and hepatic lipid droplet catabolism by activating mTOR. *FEBS Letters* 592 (2018) 2725–2738 doi:10.1002/1873-3468.13203
- <sup>121</sup> Aguilar, C. et al. Cincuenta años de estudio de las hiperlipidemias primarias: El caso de la Hiperlipidemia familiar combinada. *Invest Clin* 2010, 51(2): 145 - 158,
- <sup>122</sup> Kathiresan, S. et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet.* 2009 January; 41(1): 56–65. doi:10.1038/ng.291
- <sup>123</sup> Plaisier, C. et al. A Systems Genetics Approach Implicates USF1, FADS3, and Other Causal Candidate Genes for Familial Combined Hyperlipidemia, 2009, *PLoS Genet* 5(9): e1000642. doi:10.1371/journal.pgen.1000642
- <sup>124</sup> Lee, C. et al. The initiation of liver development is dependent on Foxa transcription factors. *Nature LETTERS*, 16 June 2005, Vol 435 doi:10.1038/nature03649
- <sup>125</sup> Kaestner, K. et al. The Hepatocyte Nuclear Factor 3 (HNF3 or FOXA) Family in Metabolism. 2000, *TEM Vol.* 11, No. 7
- <sup>126</sup> Le Lay J, Kaestner KH. The Fox Genes in the Liver: From Organogenesis to Functional Integration. *Physiol Rev* 90: 1–22, 2010; doi:10.1152/physrev.00018.2009
- <sup>127</sup> Wang, H y Eckel, R. Lipoprotein lipase: from gene to obesity. *Am J Physiol Endocrinol Metab* 297: E271–E288, <https://doi.org/10.1152/ajpendo.90920.2008>
- <sup>128</sup> Borghini, S. et al. The TLX2 homeobox gene is a transcriptional target of PHOX2B in neural-crest-derived cells. 2006, *Biochem. J.* 395, 355–361 doi:10.1042/BJ20051386
- <sup>129</sup> Cui, A. et al. Dexamethasone-induced Krüppel-like factor 9 expression promotes hepatic gluconeogenesis and hyperglycemia. 2019, *Clin Invest.*;129(6):2266–2278. <https://doi.org/10.1172/JCI66062>
- <sup>130</sup> Sweet, D. et al. Taking KLF9 to "Cort" for crimes against metabolism. 2009, *J Clin Invest*;129(6):2178–2180. <https://doi.org/10.1172/JCI128481>

- <sup>131</sup> Elsayed et al. Aberrant development of pancreatic beta cells derived from human iPSCs with FOXA2 deficiency. 2021, *Cell Death and Disease*, 12:103. <https://doi.org/10.1038/s41419-021-03390-8>
- <sup>132</sup> Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. 2009, *Annu Rev Genet*, 43:525–58.
- <sup>133</sup> Mitchell, T.J y John, S. Signal transducer and activator of transcription (STAT) signalling and T-cell lymphomas. 2005, *Immunology*, 114, 301–312
- <sup>134</sup> Yanai, H. et al. The IRF family of transcription factors Inception, impact and implications in oncogenesis. Nov 2012. *OncoImmunology* 1:8, 1376–1386
- <sup>135</sup> Cobaleda, C. et al. Pax5: the guardian of B cell identity and function. REVIEW, *NATURE IMMUNOLOGY*, VOLUMEN 8, No 5, MAYO 2007, 463
- <sup>136</sup> Gilmore, TD. Introduction to NF-κB: players, pathways, perspectives. *Oncogene*. 2006, 25, 6680–6684
- <sup>137</sup> Chen, D.B. et al. RFX5 promotes the progression of hepatocellular carcinoma through transcriptional activation of KDM4A. *Scientific Reports*. 2020, 10:14538. <https://doi.org/10.1038/s41598-020-71403-1>
- <sup>138</sup> Høiby, T. et al. A facelift for the general transcription factor TFIIA, *Biochimica et Biophysica Acta* 1769 (2007) 429–436
- <sup>139</sup> Hansen, A.W. et al. Germline mutation in POLR2A: a heterogeneous, multi-systemic developmental disorder characterized by transcriptional dysregulation. January 2021, *Human Genetics and Genomics Advances* 2, 100014. <https://doi.org/10.1016/j.xhgg.2020.100014>
- <sup>140</sup> Giresi, P.G. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. 2007, *Genome Research*, 1088-9051/07. <http://www.genome.org/cgi/doi/10.1101/gr.5533506>
- <sup>141</sup> Gawlik-Rzemieniewska, N y Bednarek I. The role of NANOG transcriptional factor in the development of malignant phenotype of cancer cells. 2016. *CANCER BIOLOGY & THERAPY*, VOL. 17, NO. 1, 1–10 <http://dx.doi.org/10.1080/15384047.2015.1121348>
- <sup>142</sup> Wang, C.C. et al. Conserved linkage of NK-2 homeobox gene pairs Nkx2-2/2-4 and Nkx2-1/2-9 in mammals. 2000, *Mammalian Genome* 11, 466–468
- <sup>143</sup> Phillips, K. et al. The Virtuoso of Versatility: POU Proteins that Flex to Fit. *J. Mol. Biol.* 2000, 302, 1023–1039. doi:10.1006/jmbi.2000.4107
- <sup>144</sup> Blank, V. et al. Molecular Characterization and Localization of the Human MAFG Gene. 1997, *GENOMICS* 44, 147–149
- <sup>145</sup> Luk, A.D.W. et al. NF-E2 mutation as a novel cause for inherited thrombocytopenia. 2020, *Br J Haematol*. Apr;189(2):e41-e44. doi: 10.1111/bjh.16438. Epub 2020 Jan 17. PMID: 31951293; PMCID: PMC7187305.
- <sup>146</sup> Ko, L. et al. DNA-Binding Specificities of the GATA Transcription Factor Family. *MOLECULAR AND CELLULAR BIOLOGY*, Julio 1993, Vol 13, No 7, p. 4011-4022
- <sup>147</sup> Patsialou, A. et al. DNA-binding properties of ARID family proteins. *Nucleic Acids Research*, 2005, Vol. 33, No. 1, 66-80 doi:10.1093/nar/gki145
- <sup>148</sup> Loukinov, D. et al., "BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma", *PNAS* 99, 10 (may 2002). doi/10.1073/pnas.092123699.
- <sup>149</sup> Monaghana. C.E. et al. REST corepressors RCOR1 and RCOR2 and the repressor INSM1 regulate the proliferation-differentiation balance in the developing brain.
- <sup>150</sup> Bretones, G. et al. Myc and cell cycle control. 2015 May, *Biochim Biophys Acta.*;1849(5):506-16. doi: 10.1016/j.bbagr.2014.03.013.
- <sup>151</sup> Johnson DG, Schneider-Broussard R. Role of E2F in cell cycle control and cancer. *Front Biosci*. 1998 Apr 27;3:d447-8. DOI: 10.2741/a291
- <sup>152</sup> Duan, R. et al. EZH2: a novel target for cancer treatment. 2020, *Journal of Hematology & Oncology*, 13:104 <https://doi.org/10.1186/s13045-020-00937-8>
- <sup>153</sup> Carmody, E.E. et al. Loss of SS18-SSX1 Inhibits Viability and Induces Apoptosis in Synovial Sarcoma. 2014, *Clin Orthop Relat Res*, 472:874–882 DOI 10.1007/s11999-013-3065-9
- <sup>154</sup> Esparza G, J. et al. A common variant on chromosome 11q13 is associated with atopic dermatitis. MAY 2009 *NATURE GENETICS*, VOLUME 41, NUMBER 5