



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS QUÍMICAS Y FARMACÉUTICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR

Análisis bioinformático comparativo del gen que codifica para el receptor de lipoproteínas de baja densidad oxidadas LOX1 en rata, ratón y humano

MEMORIA PARA OPTAR AL TÍTULO DE BIOQUÍMICA

VALENTINA ALEJANDRA CAMUS PARADA

Directora y Patrocinante

Dra. Lorena García Nannig

Laboratorio de Inflamación y Estrés Oxidativo

Cardiovascular

Departamento de Bioquímica y Biología Molecular

Facultad de Ciencias Químicas y Farmacéuticas

Universidad de Chile

Co-director

Dr. Sergio Lobos Camus

Laboratorio de Biología Molecular

Departamento de Bioquímica y Biología Molecular

Facultad de Ciencias Químicas y Farmacéuticas

Universidad de Chile

Santiago de Chile, 2015

**A mis hermanos,
Quienes me inspiran a seguir el camino de las ciencias.**

*“El corazón del entendido adquiere sabiduría;
Y el oído de los sabios busca la ciencia”*

Proverbios 18:15 RVR 1960

*“Adquiere sabiduría, adquiere inteligencia; No te olvides ni te apartes de las razones de mi
boca; No la dejes, y ella te guardará; Amala, y te conservará.*

*Sabiduría ante todo; adquiere sabiduría; y sobre todas tus posesiones adquiere
inteligencia.*

Engrandécela, y ella te engrandecerá; Ella te honrará, cuando tú la hayas abrazado.

Adorno de gracias dará a tu cabeza; Corona de hermosura te entregará”

Proverbios 4:5-9 RVR 1960

FINANCIAMIENTO

Proyecto FONDECYT 1140713 (Dra. Lorena García Nannig, investigadora responsable).

Proyecto FONDAP 15130011 (Dra. Lorena García Nannig, investigadora asociada).

Esta memoria de título fue realizada en el Laboratorio de Inflamación y Estrés Oxidativo Cardiovascular, Departamento de Bioquímica y Biología Molecular de la Facultad de Ciencias Químicas y Farmacéuticas de la Universidad de Chile.

Esta memoria dio origen a la siguiente presentación a congreso (comunicación libre modalidad oral):

Camus, Valentina., Lobos, Sergio., Garcia, Lorena. "Comparative bioinformatic study of the transcriptional regulation of the receptor for oxidized low-density lipoprotein LOX1 in rat, mouse and human and its potential role in heart diseases". XXXVIII Reunión Anual de la Sociedad de Bioquímica y Biología Molecular, 22 al 25 de Septiembre del 2015, Puerto Varas, Chile.

ÍNDICE GENERAL

ÍNDICE GENERAL	I
ÍNDICE DE FIGURAS.....	IV
ÍNDICE DE TABLAS.....	VI
ABREVIATURAS.....	VIII
GLOSARIO DE TÉRMINOS.....	XI
RESUMEN	XIV
SUMMARY	XV
1. INTRODUCCIÓN.....	1
1.1. ENFERMEDAD CARDIOVASCULAR	1
1.2. LDL.....	1
1.3. EL RECEPTOR DE OXLDL, LOX1 (“LECTIN-LIKE OXIDIZED LDL RECEPTOR-1”)	3
1.4. LOX1 EN FIBROBLASTOS CARDIACOS.....	5
1.5. GEN QUE CODIFICA A LOX1	6
1.6. ESTUDIO <i>IN SILICO</i>	8
2. OBJETIVOS.....	10
2.1. OBJETIVO GENERAL	10
2.2. OBJETIVOS ESPECÍFICOS.....	10
3. METODOLOGÍA.....	11
3.1. RECOPIACIÓN DE DATOS.....	11
3.2. ANÁLISIS DE SECUENCIAS	11
3.2.1. CLC SEQUENCE VIEWER 7.5	11
3.3. PREDICCIÓN DE SITIOS DE UNIÓN A FACTORES DE TRANSCRIPCIÓN (TFBSs).....	12
3.3.1. LASAGNA-SEARCH 2.0	13
3.4. CORRELACIÓN ENTRE TFBSs Y SNP.....	13
3.5. PREDICCIÓN DE SITIOS DE POLIADENILACIÓN ALTERNATIVA (PAS)	14
3.5.1. DNAFSMINER.....	14
3.6. CONSTRUCCIÓN DE DENDOGRAMAS FILOGENÉTICOS.....	14
3.6.1. MEGA 6	15

3.7.	MODELAMIENTO DE PROTEÍNAS.....	15
3.7.1.	SERVIDOR EN LÍNEA BLAST.....	15
3.7.2.	SERVIDOR EN LÍNEA CLUSTAL W2	16
3.7.3.	SERVIDOR EN LÍNEA SWISSMODEL.....	16
4.	DIAGRAMA DE FLUJO DE LA INVESTIGACIÓN BIOINFORMÁTICA DE OLR1 Y LOX1 EN HUMANO, RATÓN Y RATA.....	17
5.	RESULTADOS Y DISCUSIÓN	18
5.1.	GEN OLR1 EN HUMANO, RATÓN Y RATA.....	18
5.1.1.	SINÓNIMOS DE LOX1 PARA GEN Y PROTEÍNA.....	18
5.1.2.	SECUENCIAS DEL GEN OLR1 EN HUMANO, RATA Y RATÓN	20
5.1.3.	TAMAÑO DE SECUENCIAS Y COMPOSICIÓN (%GC) DEL GEN OLR1 EN HUMANO, RATA Y RATÓN.	23
5.1.4.	CONTEXTO GENÓMICO	24
5.1.4.1.	<i>HOMO SAPIENS</i>	24
5.1.4.2.	<i>MUS MUSCULUS</i>	28
5.1.4.3.	<i>RATTUS NORVEGICUS</i>	30
5.1.5.	SOLAPAMIENTO DE GENES	31
5.1.6.	ALINEAMIENTO DE GENES	34
5.1.7.	DISCUSIÓN SECCIÓN SECUENCIAS GEN OLR1 EN HUMANO, RATÓN Y RATA	38
5.2.	TRANSCRITOS DE OLR1 EN HUMANO, RATÓN Y RATA.....	49
5.2.1.	SITIOS DE UNIÓN A FACTORES DE TRANSCRIPCIÓN EN EL PROMOTOR DE OLR1	49
5.2.2.	SITIOS DE UNIÓN A FACTORES DE TRANSCRIPCIÓN ASOCIADOS A POLIMORFISMOS DE UN SOLO NUCLEÓTIDO (SNP) EN HUMANO ⁵¹	
5.2.3.	SECUENCIAS DE ARN MENSAJERO OLR1	53
5.2.4.	SITIOS DE POLIADENILACIÓN (PAS) EN HUMANO.....	55
5.2.5.	DISCUSIÓN SECCIÓN REGULACIÓN DE LA TRANSCRIPCIÓN GEN OLR1	60
5.3.	PROTEÍNA LOX1 EN HUMANO, RATÓN Y RATA	67
5.3.1.	SECUENCIAS DE PROTEÍNAS LOX1	67
5.3.2.	PROPIEDADES FÍSICOQUÍMICAS DE LAS ISOFORMAS DE LOX1	69
5.3.3.	ANÁLISIS FILOGENÉTICOS	73
5.3.4.	ALINEAMIENTO DE SECUENCIAS PROTEICAS LOX1	76
5.3.5.	MODELAMIENTO DE PROTEÍNAS.....	85
5.3.6.	DISCUSIÓN SECCIÓN PROTEÍNA LOX1	92

5.	LIMITACIONES	111
6.	CONCLUSIONES	112
7.	PROYECCIONES.....	113
8.	REFERENCIAS.....	114
9.	ANEXOS.....	130

ÍNDICE DE FIGURAS

<i>Figura 1. Esquema de la oxidación de LDL por oxidación lipídica secundaria</i>	2
<i>Figura 2. Esquema de LOX1</i>	3
<i>Figura 3. Esquema de LOX1 y sus ligandos</i>	4
<i>Figura 4. Esquema de fibrosis cardíaca por isquemia</i>	5
<i>Figura 5. Contexto genómico de OLR1, secuencia AJ131757 de humano</i>	25
<i>Figura 6. Contexto genómico de OLR1, secuencia NG_016743 de humano</i>	25
<i>Figura 7. Contexto genómico de OLR1, secuencia AC_000144 de humano</i>	25
<i>Figura 8. Contexto genómico OLR1, secuencia CM000263 de humano</i>	26
<i>Figura 9. Contexto genómico OLR1, secuencia NC_000012 de humano</i>	26
<i>Figura 10. Contexto genómico OLR1, secuencia NC_018923 de humano</i>	27
<i>Figura 11. Contexto genómico OLR1, secuencia AC_000028 de ratón</i>	28
<i>Figura 12. Contexto genómico OLR1, secuencia CM000214 de ratón</i>	28
<i>Figura 13. Contexto genómico OLR1, secuencia NC_000072 de ratón</i>	29
<i>Figura 14. Contexto genómico OLR1, secuencia AC_000072 de rata</i>	30
<i>Figura 15. Contexto genómico OLR1, secuencia NC_005103 de rata</i>	30
<i>Figura 16. Heterogeneidad de anotaciones de gen OLR1 en humano</i>	36
<i>Figura 17. Heterogeneidad de anotaciones de gen OLR1 en ratón y rata</i>	37
<i>Figura 18. Heterogeneidad de anotaciones de gen OLR1 en humano, ratón y rata</i>	37
<i>Figura 19. Esquema representativo del vecindario genómico de gen OLR1 en humano, ratón y rata</i>	41
<i>Figura 20. Heterogeneidad de anotaciones de gen OLR1 en rata</i>	42
<i>Figura 21. Mecanismo de origen de solapamiento OLR1 y TMEM52B en humano</i>	46
<i>Figura 22. Mecanismo de origen de solapamiento OLR1 y CLEC9A en ratón</i>	46
<i>Figura 23. Mapa de sitios poly(A) de gen OLR1 exhibido en APASdb</i>	57
<i>Figura 24. Sitios heterogéneos de corte según sitio de poli(A) gen OLR1</i>	58
<i>Figura 25. Lecturas de APAs de gen OLR1 según tejido</i>	59
<i>Figura 26. Esquema de la regulación de la transcripción de gen OLR1 a receptor LOX1 vía factores transcripcionales propuestos</i>	64
<i>Figura 27. Dendograma generado a partir del agrupamiento de secuencias de LOX1 de humano, ratón y rata usando algoritmo Neighbor joining</i>	74
<i>Figura 28. Dendograma generado a partir del agrupamiento de secuencias de LOX1 de humano, ratón y rata usando algoritmo UPGMA</i>	75
<i>Figura 29. Alineamiento de secuencias proteicas de LOX1 en humano</i>	81
<i>Figura 30. Alineamiento de secuencias proteicas de LOX1 en ratón</i>	82

<i>Figura 31. Alineamiento de secuencias proteicas de LOX1 en rata</i>	83
<i>Figura 32. Alineamiento de secuencias Isoforma 1 de LOX1 en humano, ratón y rata</i>	84
<i>Figura 33. Modelos proteicos de isoforma 1 LOX1 para humano, ratón y rata</i>	88
<i>Figura 34. Modelos proteicos de isoformas LOX1 para humano</i>	89
<i>Figura 35. Modelos proteicos de isoformas LOX1 para ratón</i>	90
<i>Figura 36. Esquema de residuos de LOX1 humano</i>	103
<i>Figura 37. Redes de regulación transcripcional para OLR1 comunes en humano, ratón y rata</i>	131
<i>Figura 38. Redes de regulación transcripcional para OLR1 en humano</i>	133
<i>Figura 39. Redes de regulación transcripcional para OLR1 en ratón</i>	134
<i>Figura 40. Redes de regulación transcripcional para OLR1 en rata</i>	135
<i>Figura 41. Longitud de isoformas proteicas de LOX1 en humano</i>	146
<i>Figura 42. Longitud de isoformas proteicas de LOX1 en ratón</i>	146
<i>Figura 43. Longitud de isoformas proteicas de LOX1 en rata</i>	147
<i>Figura 44. Contenido de Aminoácidos de Isoformas LOX1 en humano</i>	147
<i>Figura 45. Contenido de Aminoácidos de Isoformas LOX1 en ratón</i>	148
<i>Figura 46. Contenido de Aminoácidos de Isoformas LOX1 en rata</i>	148

ÍNDICE DE TABLAS

<i>Tabla 1. Sinónimos del gen OLR1 para Homo sapiens, Mus musculus y Rattus Norvegicus.</i>	<i>19</i>
<i>Tabla 2. Sinónimos de la proteína LOX1 para Homo sapiens, Mus musculus y Rattus Norvegicus.</i>	<i>19</i>
<i>Tabla 3. Información de secuencias de gen OLR1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>21</i>
<i>Tabla 4. Información de secuencias de gen OLR1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>22</i>
<i>Tabla 5. Tamaño y composición en gen OLR1.</i>	<i>23</i>
<i>Tabla 6. Solapamiento de OLR1 en humano y ratón.</i>	<i>32</i>
<i>Tabla 7. Tamaño de genes e isoformas de OLR1 y genes que lo solapan.</i>	<i>33</i>
<i>Tabla 8. Tabla resumen de la proporción de sitios informativos a partir del alineamiento de secuencias gen OLR1 en humano, ratón y rata.</i>	<i>35</i>
<i>Tabla 9. Correlación entre TFBSs y SNP en promotor de OLR1 humano.</i>	<i>52</i>
<i>Tabla 10. Sitios de señalización de poliadenilación hipotéticos en humano obtenidos con DNAFSMiner.</i>	<i>56</i>
<i>Tabla 11. Propiedades Físicoquímicas isoformas proteicas de LOX1 en humano, ratón y rata.</i>	<i>72</i>
<i>Tabla 12. Detalle de valores obtenidos en el modelamiento de isoformas de LOX1 en humano, ratón y rata.</i>	<i>91</i>
<i>Tabla 13. Información de secuencias de gen OLR1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>130</i>
<i>Tabla 14. Información de secuencias de gen OLR1 para humano, ratón y rata, relativas a base de datos NCBI y quienes las originaron.</i>	<i>131</i>
<i>Tabla 15. Factores de Transcripción a humano, ratón y rata.</i>	<i>136</i>
<i>Tabla 16. Información de secuencias de ARNm de OLR1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>137</i>
<i>Tabla 17. Información de secuencias de ARNm de OLR1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>138</i>
<i>Tabla 18. Información de secuencias de ARNm de OLR1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>139</i>
<i>Tabla 19. Exones secuencias ARNm OLR1. Se muestran solo las secuencias que reportan dicha información.</i>	<i>140</i>
<i>Tabla 20. Región regulatoria, 5'UTR, 3'UTR y Sitio poliA de ARNm OLR1.</i>	<i>140</i>
<i>Tabla 21. Información de secuencias proteicas LOX1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>141</i>
<i>Tabla 22. Información de secuencias proteicas LOX1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>142</i>

<i>Tabla 23. Información de secuencias proteicas LOX1 para humano, ratón y rata, relativas a base de datos NCBI.</i>	<i>144</i>
<i>Tabla 24. Variaciones de aminoácidos en LOX1 de humano, según secuencia proteica.</i>	<i>149</i>
<i>Tabla 25. Variaciones de aminoácidos en LOX1 de ratón, según secuencia proteica.</i>	<i>151</i>
<i>Tabla 26. Variaciones de aminoácidos en LOX1 de rata, según secuencia proteica.</i>	<i>151</i>

ABREVIATURAS

α -SMA	Actina de músculo liso tipo alfa
ADN	Ácido desoxirribonucleico
ADNc	Ácido desoxirribonucleico codificante
AP-1, AP-2	Proteína activadora 1, Proteína activadora-2
ARN	Ácido ribonucleico
ARNds	Ácido ribonucleico de doble hebra
ARNi	Ácido ribonucleico interferente
ARNlnc	Ácido ribonucleico largo no codificante
ARNm	Ácido ribonucleico mensajero
ARNmi	Micro ácido ribonucleico
Ang-II	Angiotensina II
ACS	<i>"Acute coronary syndrome"</i> , síndrome coronario agudo
APA	<i>"Alternative polyadenylation"</i> , poliadenilación alternativa
ApoB-100	Apolipoproteína B 100
Bach1	<i>"Basic Leucine Zipper Transcription Factor 1"</i>
BLAST	<i>"Basic Local Alignment Search Tool"</i>
BRCA1	<i>"Breast cancer type 1 susceptibility protein"</i>
CLEC	Receptores de lectina tipo C
CDS	<i>"Coding DNA Sequence"</i> , secuencia de ADN codificante
c-JUN	Subunidad del dímero de AP-1
CPSF	Factor específico de corte de poliadenilación
CUTs	Transcritos crípticos inestables
CVD	<i>"Cardiovascular Diseases"</i> , enfermedades cardiovasculares

EPD	<i>"Eukaryotic Promoter Database"</i>
FASTA	Formato de fichero informático basado en texto utilizado para representar secuencias de ácidos nucleicos, péptidos, y en el que los pares de bases o los aminoácidos se representan usando códigos de una única letra
GMQE	<i>"Global Model Quality Estimation"</i> , estimación global de calidad del modelo. Indica estabilidad de la estructura proteica hipotética, por tanto calidad del modelo
ICAM-1	Molécula de adhesión intercelular-1
JNK	Quinasa amino terminal Jun
JPEG	<i>"Joint Photographic Experts Group"</i> , formato de archivos de imágenes
LASAGNA	<i>"Length-Aware Site Alignment Guided by Nucleotide Association"</i> (Alineamiento de sitio consecuente con la longitud, guiado por la asociación de nucleótidos)
Lnc-ARN	ARN largo no codificante
LOX1	<i>"Lectin-type oxidized LDL Receptor 1"</i> , receptor de lipoproteínas oxidadas-1 en humano y/o de forma genérica
Lox1	<i>"Lectin-type oxidized LDL Receptor 1"</i> , receptor de lipoproteínas oxidadas 1 en ratón y rata
LPS	Lipopolisacárido
MAPK	<i>"Mitogen activated protein kinase"</i>
MEGA6	<i>"Molecular Evolutionary Genetics Analysis"</i> versión 6
MRTF-A	Cofactor transcripcional de NKX1-3
NAT	<i>"Natural Antisense Transcripts"</i> , transcritos antisentido naturales
NCBI	<i>"National Center for Biotechnology Information"</i>
NF-κB	<i>"Nuclear factor kappa-light-chain-enhancer of activated B cells"</i> , factor nuclear potenciador de las cadenas ligeras kappa de las células B activadas
NKX1-3	Factor transcripcional NFX1-3
NOX1	NADPH oxidasa-1

OLR1	<i>"Oxidized Low-density Lipoprotein Receptor 1 Gene"</i> , gen del Receptor de Lipoproteínas Oxidadas
oxLDL	Lipoproteínas de baja densidad oxidadas
ORF	<i>"Open Reading frame"</i> , marco de lectura abierto de un gen
PAS	<i>"Polyadenylation Signal"</i> , sitio señal de poliadenilación
PDB	<i>"Protein Data Bank"</i> , Base de datos de proteínas
QMEAN	Función de puntuación compuesta capaz de derivar estimaciones de errores globales (para toda la estructura) y locales (por residuo) sobre la base de un modelo único de una proteína. Indica estabilidad de la estructura proteica hipotética, por tanto calidad del modelo
ROS	<i>"Reactive oxygen species"</i> , especies reactivas del oxígeno
RSCB PDB	<i>"Research Collaboratory for Structural Bioinformatics Protein Data Base"</i>
SMAD2/3	Proteína intracelular que transduce señales extracelulares ligadas a TGF- β 1
SNP	<i>"Single nucleotide polymorphism"</i> , polimorfismo de un solo nucleótido
STS	<i>"Sequence-Tagged Sites"</i> , sitios de etiquetado de secuencias. Marcadores genéticos usados para mapeo de genoma y genotipificación
TMEM52B	Proteína de membrana 52B en humano
Tmem52b	Proteína de membrana 52b en ratón y rata
TNF- α	<i>"Tumor necrosis factor alpha"</i> , factor de necrosis tumoral alfa
TFs	<i>"Transcription factors"</i> , factores de transcripción
TFBSs	<i>"Transcription factors binding sites"</i> , sitios de unión a factores de la transcripción
TGF- β 1	Factor de crecimiento transformante beta 1
TSS	<i>"Transcription Start Site"</i> , sitio de inicio de la transcripción
Uniprot	<i>"Universal Protein Resource"</i> , base de datos
VCAM-1	Molécula de adhesión celular vascular 1

GLOSARIO DE TÉRMINOS

Células Natural Killer (NK): También conocidas como células K o células asesinas, son un tipo de linfocitos y componente del sistema inmune innato.

Contig: Corresponde a un conjunto de fragmentos de secuencias de ADN solapadas que en conjunto representan una región consenso de ADN. También puede definirse como los clones que se solapan formando un mapa físico del genoma, usado para guiar la secuenciación y ensamblado.

Enhancer: Región corta de ADN (50-1500 pb) a la que se pueden unir proteínas (activadores) para activar la transcripción de un gen.

Genomic scaffold: Corresponde a “contigs” solapados separados por intervalos de largo conocido.

Hardware: Conjunto de los componentes que integran la parte material de una computadora.

In silico: Expresión utilizada para referirse a “realizado en un computador o vía simulación computacional”. La frase fue acuñada en 1989 como alusión a las frases en latín “*in vivo*”, “*in vitro*” e “*in situ*” comúnmente usadas en biología.

In situ: Palabra de origen latino; en el lugar original, en la posición apropiada.

Neighbor Joining: Algoritmo utilizado para construir relaciones filogenéticas. El algoritmo comienza con un árbol completamente sin resolver, cuya topología corresponde a la de una red en estrella, luego itera hasta que el árbol esté completamente resuelto y todas las longitudes de rama sean conocidas.

Primer extension: Técnica usada en biología molecular que permite identificar el extremo 5' de un ARNm, y por tanto el sitio de inicio de la transcripción de un gen. El protocolo comienza con un partidor (“*primer*”) de largo aproximado de 20 residuos, marcado en su extremo 5', complementario a la región ubicada ~50-150 nucleótidos anterior al extremo 5' de la secuencia de ARNm con la que hibrida. Se cataliza la elongación del partidor hacia el extremo 5' del ARNm añadiendo transcriptasa inversa (RT) y reactivos adecuados. El ADN complementario (ADNc) marcado resultante se analiza en un gel de electroforesis, seguido de autoradiografía. El tamaño de la banda detectada provee una medida de la distancia que existe entre el extremo 5' del oligonucleótido sintético hasta el inicio del transcrito de ARNm. En teoría, el extremo 3' del ADNc coincidirá con el extremo 5' del ARNm, por lo tanto, el tamaño del ADNc marcado debería representar la distancia que existe en el extremo 5' del primer hasta el 5' del ARNm.

Promotor bashing: Técnica usada en biología molecular para identificar las regiones de una hebra de ADN, generalmente en un promotor, que afectan la transcripción de genes río abajo. En el ensayo “*promotor bashing*” se realiza una mutación sitio dirigida o deleciones en regiones determinadas del promotor y luego se mide la transcripción del gen. Si la mutación varía los niveles de transcripción, entonces dicha región del promotor corresponden a un elemento regulatorio o un sitio de unión a factores de transcripción.

Región regulatoria *cis*: Segmento de ADN que regula la transcripción. Estos segmentos pueden ubicarse de forma discontinua colindante a la región 5' del sitio de transcripción, Sin embargo, algunos segmentos individuales pueden residir en intrones, 5' y 3' UTRs, o decenas de kilobases a ambos lados del gen que regulan.

Regulación en *trans*: Regulación mediada por agentes localizados alejados del gen de interés. En términos prácticos, se refiere a cualquier parte del genoma excepto las regiones cercanas al gen regulado.

S1 mapping analysis: Preparaciones de ARN que contienen un ARNm de interés se hibridan a una sonda de ADN monocatenario complementario. Se utilizan nucleasas S1 para degradar las regiones no hibridadas de la sonda al finalizar el periodo de reacción. Los supervivientes híbridos ADN-ARN son separados por electroforesis en gel y se visualizan por autorradiografía o hibridación Southern. El método se puede usar para cuantificar ARN, mapear posiciones de intrones y para identificar las ubicaciones de los extremos 5' y 3' de ARNm en moldes de ADN clonados.

Scavenger receptor: Receptores de superficie celular que unen típicamente múltiples ligandos y promueven la remoción de blancos alterados propios y no propios. Generalmente funcionan vía mecanismos que incluyen endocitosis, fagocitosis, adhesión, y señalización que desencadena la respuesta de eliminación y degradación de sustancias dañinas.

Software: Conjunto de programas, instrucciones y reglas informáticas para ejecutar ciertas tareas en una computadora.

Técnica shotgun: El ADN a ser secuenciado es digerido por endonucleasas no específicas en forma controlada, fraccionado por tamaño y clonado generalmente en vectores fago como hebra simple. Luego los clones son tomados de forma aleatoria y usados como molde para la secuenciación. Programas computacionales leen los extremos superpuestos de diferentes lecturas para ensamblarlos en una secuencia continua.

UPGMA: “*Unweighted Pair Group Method with Arithmetic Mean*”. Algoritmo utilizado para construir relaciones filogenéticas. En cada paso, las dos agrupaciones más cercanas se combinan en un grupo de mayor nivel. La distancia entre dos grupos A y B se considera

como el promedio de todas las distancias medias entre los elementos de cada agrupación.

.

RESUMEN

Las enfermedades cardiovasculares (CVD) representan la principal causa de muerte a nivel mundial. El principal factor de riesgo de esta patología es el colesterol constituido por lipoproteínas de baja densidad (LDL, "*Low Density Lipoprotein*"). En condiciones patológicas el LDL puede oxidarse originando LDL oxidadas (oxLDL). LOX1 es el receptor de oxLDL perteneciente a la familia Lectina tipo C. En fibroblasto cardiaco, la activación de LOX1 aumenta el crecimiento celular y promueve la síntesis de colágeno bajo condiciones patológicas.

El objetivo general de esta memoria fue analizar y comparar el gen OLR1 de humano, rata y ratón a nivel de genes, transcripción y proteínas, utilizando herramientas bioinformáticas.

Se identificaron vecindades genéticas de OLR1 conservadas en las tres especies en estudio, destacándose el solapamiento del gen OLR1 con el gen TMEM52B en humano, y el solapamiento del gen OLR1 con el gen CLEC9A en ratón. Éstos influyen en la regulación de transcritos de OLR1, debido a la gran variedad de ARNm encontrados. Asimismo, esta diversidad se explicaría por fenómenos de poliadenilación alternativa (APA) presentes en el gen OLR1 de acuerdo a los sitios de señalización de APA encontrados en humano. La transcripción de OLR1 también es regulada por su región promotora, en la que se identificaron secuencias de unión a los siguientes factores de transcripción (TFs) comunes a las tres especies: AP-1, BRCA1, NKX1-3 y BACH1. Los factores de transcripción potencialmente afectados por polimorfismos de un solo nucleótido (SNP) en humano son Hand1, Pax-4, p53 y BRCA1.

Desde la perspectiva proteica, en humano y ratón existen 3 isoformas para el receptor LOX1, mientras que en rata sólo existe una isoforma. Las isoformas 2 y 3 de LOX1 de humano no unen oxLDL y no forman dímeros.

Para validar algunos resultados y predicciones de este estudio se proyecta obtener evidencia experimental en el laboratorio.

SUMMARY

Comparative bioinformatic analysis of gene encoding the receptor for oxidized low density lipoprotein LOX1 in rat, mouse and human.

Cardiovascular diseases (CVD) are the leading cause of death globally. The main risk factor for these diseases is LDL (Low Density Lipoprotein) cholesterol. In pathological conditions, these lipoproteins can undergo oxidation resulting in oxidized LDL (oxLDL). LOX1 is the oxLDL receptor belonging to the C-type lectin family. Increases in LOX- 1 activation promote cell growth and collagen synthesis under cardiac pathological conditions in cardiac fibroblasts.

In this work a genetic, transcriptomic and protein structure of the OLR1 gene was performed, which codes for the LOX1 protein. A bioinformatic search and analysis of sequences associated with the OLR1 gene and LOX1 receptor in humans, mice and rats was performed.

The genomic neighborhoods of the OLR1 gene is conserved among the three species under study. In human the TMEM52B gene overlaps the OLR1 gene, while the CLEC9A gene overlaps the OLR1 gene in mice. The neighboring genes influence OLR1 transcription, giving rise to a variety of mRNAs. Also, mRNA diversity is explained by alternative polyadenylation (APA) present in the OLR1 gene, according to APA signaling sites found in humans. OLR1 transcription is also regulated by its promoter region. Common transcription factor binding sequences were identified among the three species: AP-1, BRCA1, NKX1-3 and BACH1 transcription factors. Transcription factors (TFs) potentially affected by a single nucleotide polymorphisms (SNP) in humans are Hand1, Pax-4, p53 and BRCA1.

Three isoforms exist for the LOX1 receptor in human and mouse, while in rat only one isoform is found. Isoforms 2 and 3 of human and mouse LOX1 do not bind oxLDL and do not form dimers.

In order to validate some results and predictions of this study is necessary to obtain experimental evidence in the laboratory.

1. INTRODUCCIÓN

1.1. Enfermedad cardiovascular

Las enfermedades cardiovasculares (CVD) representan la principal causa de muerte a nivel mundial y el 30% de las muertes en el continente Americano (OPS, Organización Panamericana de la Salud). Específicamente, la enfermedad isquémica de corazón es la primera causa de muerte en América (11,69 % de las defunciones en el 2010). Se caracteriza por una obstrucción del riego arterial al músculo cardiaco causando hipoxemia, déficit de sustratos necesarios para la producción de ATP y un depósito anormal de productos de desecho del metabolismo celular. En Chile, el 45,6% de las defunciones (7.933 muertes anuales) son causadas por este tipo de patología.

Entre los factores de riesgo asociados a patologías cardiacas se encuentra el consumo de cigarrillo, diabetes mellitus, hipertensión y dislipidemia. Este último factor se relaciona con altos índices de colesterol total. El colesterol constituido por lipoproteínas de baja densidad (LDL o "*Low Density Lipoprotein*") está estrechamente relacionado a la enfermedad isquémica de corazón (Verhoye & Langlois, 2009). En condiciones patológicas, las LDL se pueden modificar por oxidación de sus componentes (Metha J, 2005) dando origen a LDL oxidadas (oxLDL).

1.2. LDL

Las partículas de LDL están formadas por un núcleo hidrofóbico de ácidos grasos poliinsaturados, moléculas de colesterol, y un revestimiento constituido por apolipoproteína B-100 (ApoB-100), fosfolípidos y colesterol no esterificado (Verhoye & Langlois, 2009). Se pueden modificar mediante oxidación inmediata y preferencial de residuos aminoacídicos de ApoB-100, en ausencia de antioxidantes lipídicos solubles, y/o por peroxidación lipídica (Soler-Cantero A, 2012). También, es posible que la pérdida de antioxidantes propios de su estructura, como el alfa-tocoferol y ubiquinol-10, produzcan la primera oxidación lipídica de gran parte de los ácidos grasos poliinsaturados, generando hidroperóxidos como productos. Éstos pueden fragmentarse, originando una oxidación

lipídica secundaria cuyos productos pueden reaccionar con residuos de lisina de ApoB-100, dando lugar a oxLDL (Soler-Cantero A, 2012) (Figura 1).

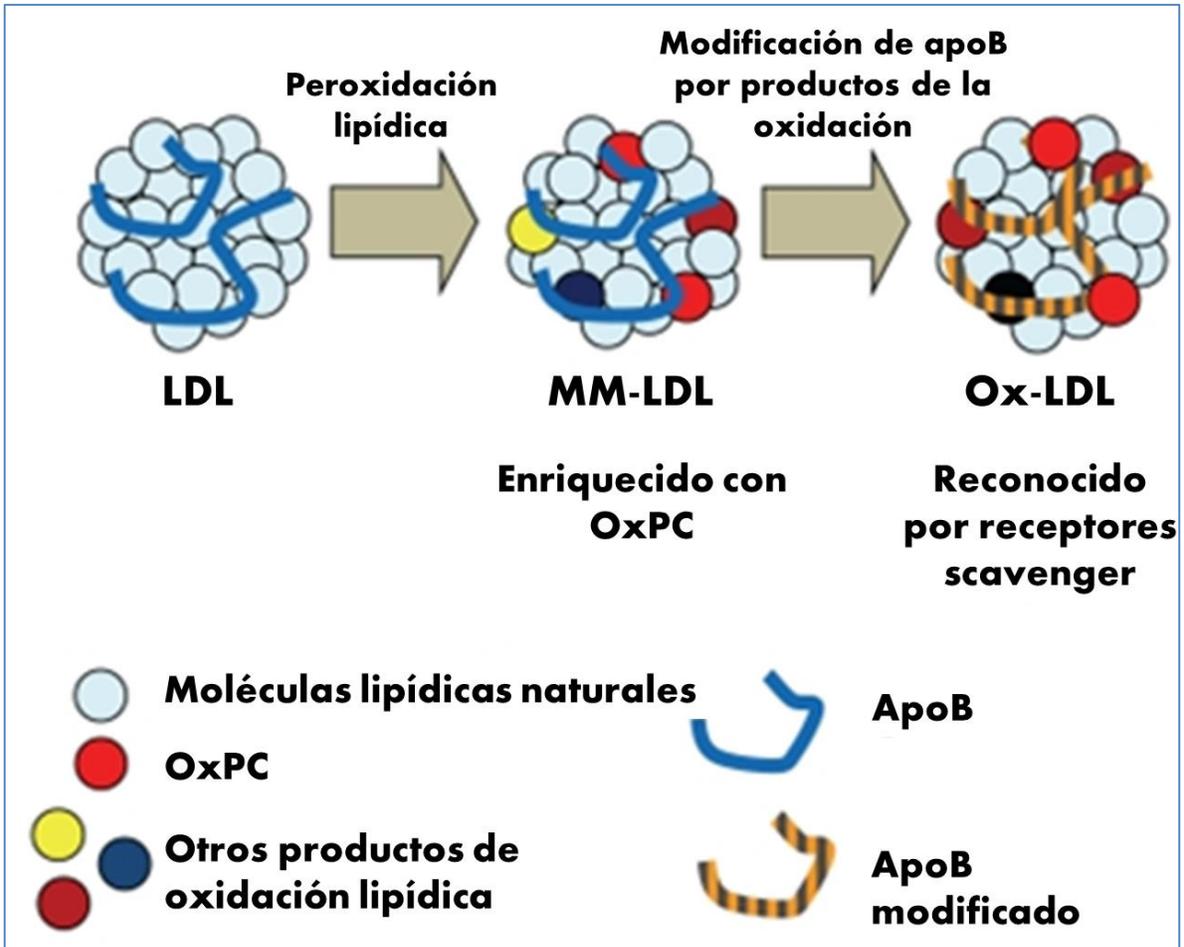


Figura 1. Esquema de la oxidación de LDL por oxidación lipídica secundaria. LDL: lipoproteínas de baja densidad; MM-LDL: lipoproteínas de baja densidad mínimamente modificadas; oxLDL: lipoproteínas de baja densidad oxidadas; OxPC: productos de oxidación lipídica; apoB: apolipoproteína B. (Figura adaptada y traducida de).

1.3. El receptor de oxLDL, LOX1 (“Lectin-like Oxidized LDL Receptor-1”)

LOX1 corresponde a una glicoproteína de transmembrana, de secuencia aminoacídica conservada entre mamíferos, cuyo peso molecular es 50 kDa (Yoshimoto R, 2011). Tiene un

dominio N-terminal citoplasmático, un único dominio transmembrana y uno extracelular que comprende una región tipo cuello, seguido del dominio C-terminal tipo lectina (Yoshimoto R, 2011) (**Figura 2**). Este último, une ligandos y también participaría en el inicio de la internalización y fagocitosis del ligando (Yoshimoto R, 2011). En 1997, Sawara y sus colegas clonaron por primera vez este receptor a partir de células de endotelio aórtico de bovino (Metha J, 2005).

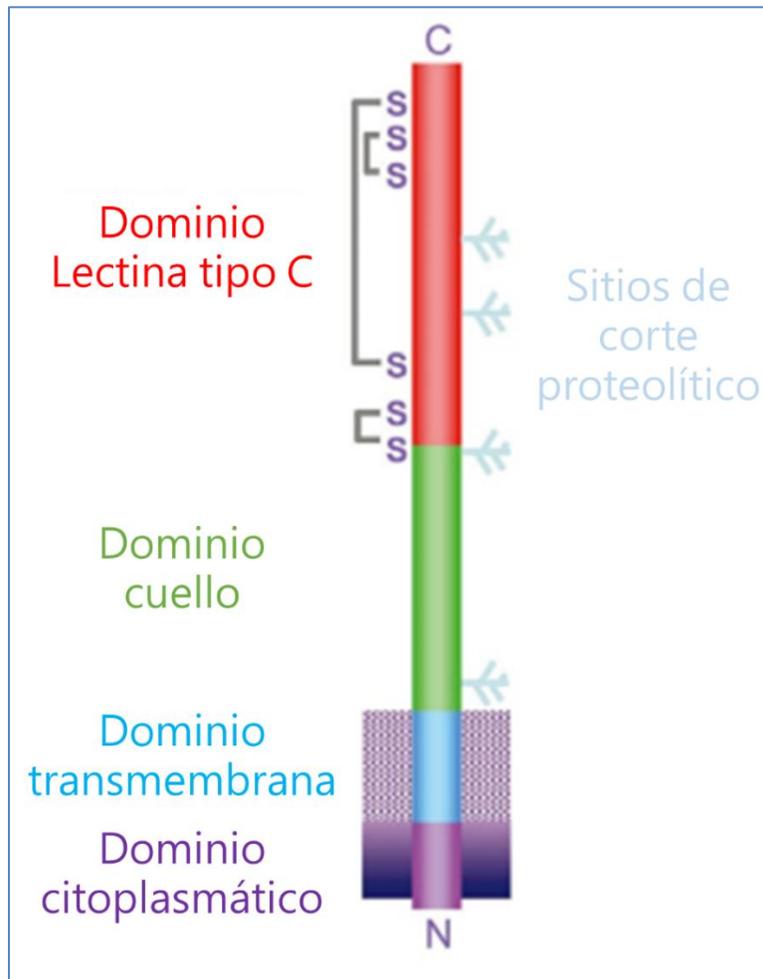


Figura 2. Esquema de LOX1. Se muestra el dominio citoplasmático en morado, el dominio transmembrana en celeste, el dominio cuello en verde y el dominio Lectina tipo C en rojo. Se indican los sulfuros correspondientes a residuos de triptófano que forman puentes disulfuros para estabilizar la proteína. En flechas celestes se indican los sitios de corte de la proteína que generan LOX1 soluble. (Figura adaptada y traducida de .

En placas de aterosclerosis coronaria se han detectado niveles séricos elevados de una versión soluble de LOX1 en el sistema circulatorio de pacientes con síndrome coronario

agudo (ACS). El aumento de actividad de serina proteasas y metaloproteinasas de matriz extracelular escinden LOX1 y promueven la liberación de su forma soluble al sistema circulatorio .

Se ha reportado que LOX1 se expresa en células endoteliales, macrófagos, plaquetas, músculo liso, ovario y fibroblastos cardiacos (Yoshimoto R, 2011), (Metha J, 2005). A pesar de ser considerado como un receptor tipo “scavenger”, el análisis de secuencia y estructura sugiere que se parece más a receptores característicos de células “Natural Killer” (NK) (Metha J, 2005). LOX1 puede unir otros tipos de ligandos, además de oxLDL, tales como lipopolisacárido (LPS), plaquetas, células muertas, químicos polianiónicos y fosfolípidos aniónicos (Figura 3).

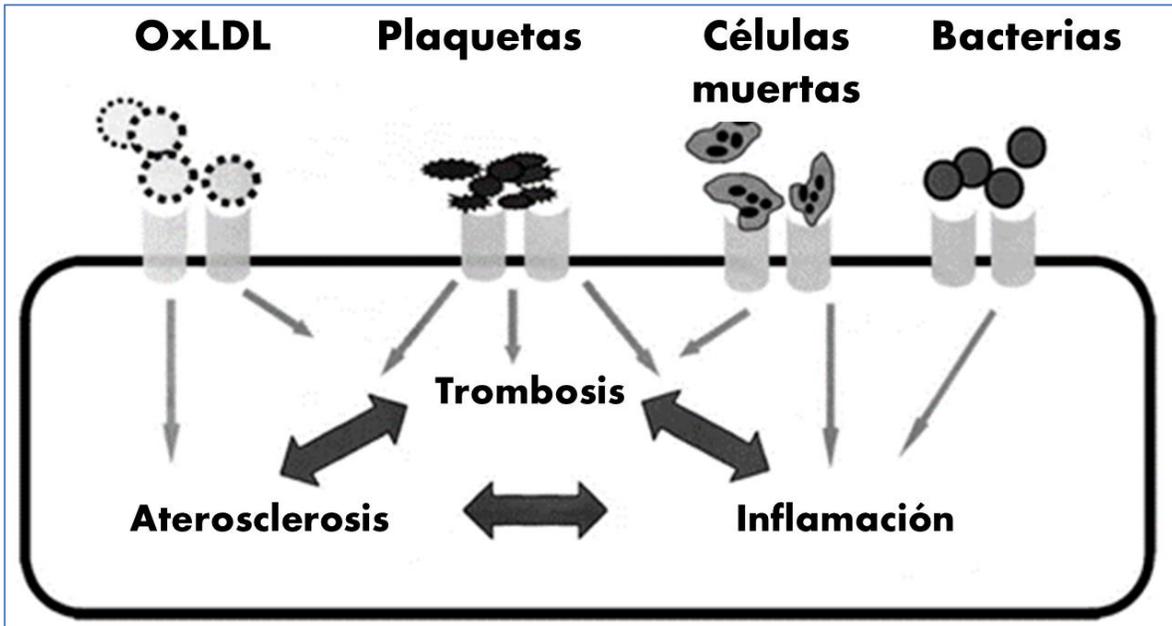


Figura 3. Esquema de LOX1 y sus ligandos. LOX1 no es un receptor exclusivo de oxLDL, pero es el único receptor conocido que une este tipo de moléculas. Además de unir oxLDL, LOX1 une plaquetas, LPS de bacterias y reconoce células muertas, entre otros ligandos. (Figura adaptada y traducida de).

1.4. LOX1 en fibroblastos cardiacos

El tejido cardiaco está formado por varios tipos celulares, entre los que se encuentran a los miocitos cardiacos y fibroblastos. Los fibroblastos cardiacos representan

aproximadamente el 25% del volumen del tejido de miocardio y aproximadamente el 60% de todas las células en el corazón (Wang X, 2013).

El desarrollo de tejido fibrótico involucra un desequilibrio entre la síntesis y degradación de colágeno por fibroblastos, activación del sistema inmune, inflamación y disfunción vascular. Esto es comúnmente observado en patologías cardíacas (Del Principe D, 2013) (**Figura 4**).

La activación de LOX1 en el fibroblasto cardíaco gatilla un aumento en el crecimiento celular y promueve la síntesis de colágeno (Wang X, 2013). Este es un proceso típico de remodelación del tejido, que ocurre en alteraciones relacionadas con la senescencia en el corazón y enfermedades cardíacas como isquemia o alzas de presión, pudiendo proceder a síndrome de falla cardíaca (Wang X, 2013). La fibrosis de miocardio es un proceso complejo que involucra a cada componente celular del tejido del miocardio. Los miofibroblastos cardíacos tienen un rol central en este proceso al incrementar la producción de colágeno y otros componentes de la matriz extracelular, bajo la influencia de señales como apoptosis de miocitos, citoquinas proinflamatorias y especies reactivas del oxígeno (ROS) (**Figura 4**).

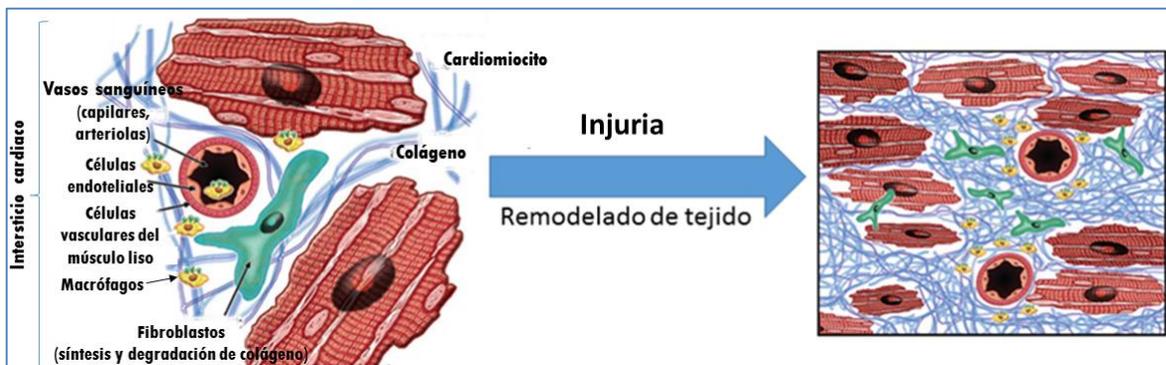


Figura 4. Esquema de fibrosis cardíaca por isquemia. A la izquierda se representa el tejido cardíaco sano, con poblaciones celulares de cardiomiocitos, miofibroblastos, macrófagos, etc; la imagen a la derecha representa el tejido cardíaco luego de sufrir injuria al corazón y su consecuente remodelado caracterizado por un aumento en las fibras de colágeno, infiltración de macrófagos y cambio en el fenotipo del miofibroblasto. Esquema adaptado y traducida de .

LOX1 es parcialmente responsable de la mantención del fenotipo en fibroblasto cardíaco en relación a la organización del citoesqueleto y proliferación celular, de acuerdo a lo que el equipo de Wang y cols. demostró en el año 2003. Específicamente, la delección del gen que codifica a LOX1 resulta en una alteración severa de la distribución de F-actina,

falla en la mitosis, poliploidía y aumento del tamaño de este tipo celular. Al transfectar los fibroblastos cardiacos knockout para este receptor con el gen que codifica para LOX1 se restaura la morfología, organización del citoesqueleto y actividad mitótica. Además, observaron que los fibroblastos senescentes eran mucho más grandes, tenían un citoesqueleto anormal y secretaban más colágeno que los fibroblastos en corazones de ratones jóvenes (Wang X, 2013).

La expresión de proteína LOX1 puede aumentar por estímulos pro-ateromatosos como TNF- α , angiotensina II (Ang-II), oxLDL y estrés mecánico por fluidos (Zhu H, 2005); además, de estímulos presentes en patología diabética como altos niveles de glucosa y ROS (Yan M, 2011). *In vitro*, la expresión basal del gen codificante para LOX1 es baja, pero puede ser fuertemente inducida por estímulos pro-inflamatorios y pro-oxidativos en células endoteliales, de músculo liso y macrófagos (Yoshimoto R, 2011).

1.5. Gen que codifica a LOX1

En *Homo sapiens* (humano) existe una sola copia del gen que codifica para el receptor LOX1 ubicado en el cromosoma 12p12.3-p13.2, y es conocido como gen OLR1 . OLR1 se encuentra inserto en el cluster de genes de la subfamilia de receptores de lectina tipo C (CLEC), involucrados en el sistema inmune innato agrupado en un complejo de genes que codifican receptores asociados a células “Natural Killer” . La estructura del gen OLR1 consiste en 6 exones interrumpidos por 5 intrones, abarcando cerca de 7000 pares de bases (pb). Los exones 1 al 5 de OLR1 tienen un largo aproximado de 102 pb a 246 pb, mientras que el exón 6 es más largo con 1722 bp . La secuencia del exón 1 de OLR1 incluye la región 5'-UTR (región no traducida 5') y parte del dominio citoplasmático de la proteína LOX1, el exón 2 codifica el resto del dominio citoplasmático y el dominio transmembrana de LOX1, el exón 3 de OLR1 codifica el dominio tipo cuello de la proteína, y los exones 4 a 6 codifican el dominio tipo lectina y 3'-UTR (región no traducida 3') de LOX1 .

Tanto el equipo de Aoyama y cols. (1999) como el grupo de Yamanaka y cols. (1998) se dedicaron a estudiar la región regulatoria o promotora proximal del gen que codifica a LOX1. Ambos coinciden en que se localizan cajas TATA y CAAT a 29pb y 99pb,

respectivamente, río arriba del sitio de inicio de la transcripción; además, identificaron una variedad de potenciales sitios de unión a factores de transcripción utilizando análisis computacionales, atribuyendo el carácter de “*enhancer*” a la secuencia . Metha y cols. (2005) identificaron un sitio de unión al factor de transcripción octamérico (Oct-1, se une a secuencias "ATTTGCAT") utilizando “*promoter bashing*”, comprendido en la región promotora entre los nucleótidos (nt) -1494 y -1599, como necesario para la activación del promotor de LOX1 inducida por oxLDL (Metha J, 2005). De la misma forma, identificaron un sitio de unión activo para NF-kB necesario para la activación en *trans* del promotor LOX1 inducido por Ang-II, ubicado las posiciones -2247 y -2131 (Metha J, 2005).

Se han reportado 446 polimorfismos de un solo nucleótido (SNP) de LOX1 en la población humana, de los cuales aproximadamente 19 involucran cambios conservativos en aminoácidos y 16 cambios no conservativos . El equipo de Mango y cols. (2003) realizó un estudio en la población italiana y asoció un SNP alojado en la región 3'UTR de OLR1 a enfermedad cardiaca: se encontró un 91,3% de prevalencia de SNP (C>T, a 188pb de la señal de término de la traducción) en pacientes con infarto agudo al miocardio en comparación una prevalencia de 73,8% en pacientes sanos . Biocca y cols. (2009) también reportaron un SNP relacionado con infarto agudo al miocardio en OLR1; en este caso, el SNP (501G>C) se aloja en una región codificante produciendo un cambio no-sinónimo en la posición 167 de la proteína LOX1 (p.K167N), alterando la afinidad del receptor por el ligando oxLDL (Biocca S, 2009).

El gen que codifica LOX1 en ratón es conocido como Olr1 . Está ubicado en el cromosoma 6, tiene un tamaño de 22109 pb, y codifica para una proteína de 366 aminoácidos .

Al igual que en ratón, el gen que codifica a LOX1 en rata es conocido como Olr1 . Está ubicado en el cromosoma 4, tiene 22085 pb de tamaño y codifica para una proteína de 368 aminoácidos .

En 1998, el equipo liderado por Miki Nagase (1998) aisló y caracterizó el gen que codifica para LOX1 en rata. De acuerdo a su estudio, Olr1 consiste en 8 exones ordenados en correspondencia a los dominios funcionales de la proteína: exones 3, 4 y 5 presentan un motivo de 46 aminoácidos triplicado correspondiente al dominio extracelular (no se ha

encontrado este motivo repetido en bovino o humano) . Además, utilizando técnicas experimentales como “*primer extension*” y “*S1 mapping analysis*”, localizaron el sitio de inicio de la transcripción 62 pb río arriba del sitio de inicio de la traducción, 25 pb río debajo de un motivo TATA-like box y 93 bp río debajo de un sitio CAAT-like box . Al analizar la región promotora se encontró la presencia de potenciales elementos *cis* que pueden actuar como reguladores del gen, entre ellos sitios de unión a NF-kB, AP-1 y AP-2. .

1.6. Estudio *in silico*

Hoy en día existe una gran cantidad de datos genómicos disponibles en línea, lo que permite que las ciencias biomédicas se acerquen paulatinamente, pero cada vez más, a una perspectiva computacional. El análisis de datos existentes en línea permite realizar nuevos descubrimientos. Por primera vez en la historia, los investigadores reúnen rutinariamente más información de la que pueden analizar de manera significativa. Como resultado, la ciencia es ahora rica en datos, pero pobre en descubrimientos. Para las personas que trabajan en convertir los datos en información útil, el éxito a menudo significa desarrollar y compartir nuevas herramientas, metodologías y prácticas que permiten responder a las preguntas de investigación a una escala que no era posible anteriormente.

Esta memoria abordará el estudio de los diferentes datos genómicos disponibles, utilizando diversos algoritmos y herramientas bioinformáticas con el objetivo de aportar un análisis genético, transcriptómico y proteico de los ortólogos del gen que codifica para el receptor de oxLDL en *Homo sapiens* (humano), *Mus musculus* (ratón) y *Rattus norvegicus* (rata), que permita escoger modelos de estudio animales lo más cercano a posible a lo que ocurre en patologías cardíacas en humano. Para esto se seleccionaron las secuencias de genes, transcritos y proteínas en bases de datos referentes a gen OLR1 y proteína LOX1 de humano, rata y ratón. Con esta información se analizaron los contextos genómicos de gen OLR1, se buscaron sitios de unión a factores de transcripción en el promotor de gen en las tres especies y se correlacionaron con SNP presentes en gen OLR1 humano. Se realizaron alineamientos de las secuencias proteicas de LOX1 de las tres especies y se construyeron

modelos de las isoformas para identificar similitudes y diferencias entre las estructuras de una misma especie y entre cada especie.

Estudios como el presentado en esta memoria se han realizado en otro tipo de patologías. “The project's cancer group” está buscando sistemáticamente a través de los genes secuenciados con la finalidad de compararlos con aquellos presentes en células cancerosas. Las diferencias genéticas aparentemente son pistas hacia los mecanismos biológicos del desarrollo de cáncer y podrían ayudar a encontrar nuevas herramientas de diagnóstico y tratamiento. Mike Stratton (2000), líder del proyecto cáncer, dice: “En 20 años más el tratamiento del cáncer se transformará” (Yamey, 2000). Las expectativas de esta memoria de título son similares, pero con un objetivo centrado en el estudio, desciframiento y tratamiento de patologías cardíacas en un futuro cercano. Generalmente, los estudios genómicos asociadas a patologías se centran en mutaciones que afectan regiones codificantes de una proteína, dejando de lado la importancia de mutaciones alojadas en regiones no codificantes . Específicamente, a través de este estudio esperamos entender de mejor forma el rol de regiones regulatorias en OLR1, y su influencia en patología cardíaca.

2. OBJETIVOS

2.1. Objetivo general

Realizar análisis bioinformáticos comparativo del gen que codifica para el receptor de lipoproteínas de baja densidad oxidadas LOX1 de humano, rata y ratón, a nivel de genes, transcripción y proteínas, mediante la utilización de herramientas bioinformáticas.

2.2. Objetivos específicos

1. Seleccionar las secuencias de genes, transcritos y proteínas en bases de datos referentes a LOX1 de humano, ratón y rata.
2. Analizar el contexto genómico del gen de LOX1 (OLR1) de humano, ratón y rata para observar similitudes y diferencias.
3. Estudiar la regulación de gen OLR1 mediante la búsqueda de sitios de unión a factores de transcripción en los promotores de OLR1 de humano, ratón y rata, y su correlación con SNP en humano.
4. Realizar un análisis comparativo de las secuencias y estructuras proteicas de LOX1 de humano, ratón y rata.
5. Establecer potenciales sitios de poliadenilación alternativa en el gen OLR1 humano, usando DNAFSMiner en línea.

3. METODOLOGÍA

3.1. Recopilación de datos

La búsqueda de información se realizó en bases de datos concernientes al contexto del tema en estudio (NCBI, GeneCards, Uniprot, PDB, etc), utilizando las abreviaturas “OLR1” y “LOX1” como palabras claves en los motores de búsqueda. La información fue revisada, catalogada, filtrada y acotada de acuerdo al enfoque de este trabajo, relativo a la búsqueda de secuencias genéticas, transcritos y de proteínas considerando como especies objetivo de este estudio a *Homo sapiens*, *Mus musculus* y *Rattus norvegicus*. En casos particulares, se contactó directamente vía email a la base de datos (NCBI, DDBJ) para obtener la primera fecha de disponibilidad pública de secuencias sin especificaciones.

3.2. Análisis de secuencias

Las secuencias ortólogas de LOX1 en *Homo sapiens*, *Mus musculus* y *Rattus norvegicus* se sometieron a un análisis comparativo con la finalidad de caracterizar las similitudes y diferencias en cuanto a contextos genómicos, tamaño de secuencias, número y tamaño de exones e intrones, variaciones genómicas, identificación de regiones promotoras, contenido de nucleótidos y aminoácidos según corresponda. También, se caracterizaron el punto isoeléctrico, índice alifático y masas moleculares de isoformas. Se ejecutó un informe estadístico comparativo de las secuencias y se confeccionaron alineamientos de secuencias empleando CLC Sequence Viewer 7.5 . Los análisis filogenéticos y de evolución molecular se hicieron usando MEGA version 6 . Los resultados se resumieron, contrastaron y agruparon en tablas y figuras.

3.2.1. CLC Sequence Viewer 7.5

CLC Sequence Viewer genera una interfaz de software que permite realizar un gran número de análisis bioinformáticos, combinado con la gestión de datos y excelentes opciones de visualización y salida de gráficos. El paquete de software es una edición comunitaria, disponible para Windows, Mac OS X y Linux. La descarga es gratuita. .

Se utilizó el buscador de secuencias de CLC Sequence Viewer 7.5 usando la palabra clave “OLR1” y la nomenclatura binominal de la especie objetivo. Se identificaron las secuencias que incluyen el gen OLR1, y se descargaron para su análisis. Se excluyeron secuencias tipo CDS (Coding DNA Sequence), andamio (*genomic scaffold*, porciones de secuencia del genoma reconstruida a partir de secuencias de clones generados por técnica *shotgun*), ensamblados, microARN, secuencias patentadas; y se conservaron secuencias descritas como ADN, ARNm y aminoácidos que coinciden con los objetivos de búsqueda.

Se seleccionó la opción para visualizar la secuencia del gen OLR1, su vecindad, transcritos y exones con el software CLC Sequence Viewer 7.5 (©2014 CLC bio, a QIAGEN Company). Los resultados se organizaron en tablas resumen, que incluyen información directa desde GenBank y CLC Sequence Viewer 7.5 .

Para realizar los alineamientos, se descargó la secuencia del gen OLR1, o de la proteína LOX1, en formato FASTA desde la base de datos de NCBI. Se cargó en CLC Sequence Viewer 7.5 y se procedió con el alineamiento en modo lento y muy preciso. Los archivos generados se guardaron en formato FASTA y las gráficas en formato JPG. El análisis de alineamiento se realizó en MEGA 6 .

3.3. Predicción de sitios de unión a factores de transcripción (TFBSs)

Se identificaron potenciales secuencias cortas (6-12pb) que podrían actuar como TFBSs en la región considerada como promotora de LOX1, (1000pb previas al TSS [Sitio de inicio de la transcripción]), utilizando LASAGNA-Search 2.0 como herramienta bioinformática. Esto se realizó para las tres especies en estudio, los resultados se presentaron en diagramas de redes de regulación de factores transcripcionales (Networks de regulación) y se agruparon los resultados comunes a las especies en estudio en una tabla resumen.

3.3.1. LASAGNA-Search 2.0

LASAGNA-Search 2.0 es una herramienta integrada para la búsqueda y visualización de sitios de unión a factores de transcripción (TFBS), basada en el algoritmo LASAGNA .

Para efectos de esta investigación, se trabajó tanto con modelos derivados de matrices (Matrices TRANSFAC y JASPAR CORE) como con modelos de alineamiento LASAGNA (TFBSs TRANSFAC y ORegAnno). Se usaron las secuencias de ARNm NM_002543.3 (Isoforma 1 LOX1) para humano, NM_138648 (Isoforma 1 LOX1) para ratón y NM_133306 (Isoforma 1) para rata.

3.4. Correlación entre TFBSs y SNP

Una vez identificadas las potenciales secuencias que cumplen con los requisitos para ser consideradas TFBSs en humano, se compararon con las regiones en las cuales se alojan SNPs en la misma especie. De esta forma, se determinó si existen SNP alojados en los TFBSs descritos que pudiendo afectar la transcripción del gen OLR1. Para esto, se utilizó el visualizador de genes de la base de datos NCBI y se estableció como punto de nuevo origen el inicio de transcripción de NM_2543.3. También, se marcó la región promotora de OLR1 con marcadores de posición, ubicando el marcador 0 al “*Transcription Start Site*” (TSS) y como marcador 1000 a la posición ubicada 1000 pares de base río arriba del TSS. Se efectuó un “*zoom*” a nivel de nucleótidos en la región promotora de OLR1, se limpió el área de visualización dejando a la vista sólo la secuencia de genes y dbSNP 142 (*Homo sapiens* Annotation Realease 106) All data¹, se recorrió el área del promotor de OLR1 en búsqueda de los polimorfismos, se correlacionó su posición con los TFBS previamente identificados y los resultados de agruparon en tablas.

3.5. Predicción de sitios de poliadenilación alternativa (PAS)

Utilizando DNAFSMiner como herramienta bioinformática, se identificaron potenciales secuencias de señalización de poliadenilación en secuencias de humano con el fin de examinar poliadenilación alternativa en el gen OLR1 de humano.

¹ “dbSNP 142 (*Homo sapiens* Annotation Release 106) All data” corresponde a la base de datos de polimorfismos de un solo nucleótido específica para humano, sin exclusión de datos, cuya liberación de anotaciones corresponde a la número 106.

Además, se consultó la base de datos concerniente a poliadenilación alternativa APASdb en búsqueda de información sobre OLR1. Los resultados de esta búsqueda se esquematizaron en una figura.

3.5.1. DNAFSMiner

Esta es una herramienta bioinformática en línea con utilidad para reconocer sitios funcionales en secuencias de ácidos nucleicos. Actualmente, este software proporciona dos herramientas: TIS Miner y Poly(A) Signal Miner. La primera se puede utilizar para predecir sitios de iniciación de traducción en el ADN de vertebrados / secuencias de ARNm / ADNc, mientras que Poly(A) Signal Miner permite identificar PAS en secuencias de ADN humano.

En esta memoria de título se utilizó Poly(A) Signal Miner para identificar secuencias AATAA y ATAAA en el gen OLR1 humano como PAS. Las secuencias del gen OLR1 fueron ingresadas en formato FASTA en el servidor en línea y se seleccionó la secuencia a identificar en las opciones del servidor. Los resultados obtenidos fueron agrupados en una tabla y filtrados de acuerdo a su puntaje ("*score*", considerándose como PAS más probables a aquellos con puntaje igual o superior a 0,6 de acuerdo al manual del programa).

3.6. Construcción de dendogramas filogenéticos

A partir de los alineamientos de proteínas generados en CLC Sequence Viewer , se construyeron árboles de distancia filogenética usando Mega 6 seleccionando los algoritmos UPGMA y Neighbor Joining. En ambos casos se seleccionó un "bootstrapping" de 10.000 y los resultados se expresaron en cladogramas.

3.6.1. MEGA 6

MEGA es una herramienta integrada para realizar alineamientos de secuencias, inferir árboles filogenéticos, estimar tiempos de divergencia, minería de datos en línea, estimar tasas de evolución molecular, inferir secuencias ancestrales, y probar hipótesis evolutivas.

Se cargó en MEGA 6 los alineamientos de proteína, previamente hechos en CLC Sequence Viewer, en formato FASTA. Se seleccionó la variable "protein sequence" en "*Input*

Data” y se marcaron las secuencias variables. Los resultados de las variaciones en las secuencias se representaron en tablas y figuras.

3.7. Modelamiento de proteínas

Las secuencias blanco (“*target*”) a ser modeladas se descargaron en formato FASTA desde la base de datos de NCBI y se alinearon de forma múltiple en el servidor BLAST. A partir de este alineamiento, se modeló la proteína utilizando los servidores en línea SwissModel. Para esto, se actualizó y adaptó el protocolo descrito por Bordoli y cols. en el año 2009. Los mejores modelos fueron seleccionados de acuerdo a los índices QMEAN y GMQE. Se utilizó software VMD 1.9.2. para gestionar los modelos obtenidos.

3.7.1. Servidor en línea BLAST

Siguiendo el protocolo descrito por Bordoli y cols (2009), la secuencia objetivo se ingresó en formato FASTA en la página web del servidor en línea BLAST. Se seleccionó la base de datos PDB en la configuración de búsqueda, se escogió el algoritmo BLASTp para el alineamiento múltiple y se ejecutó el servidor. Una vez que los resultados del alineamiento múltiple fueron expuestos, se seleccionaron aquellas secuencias que contaran con estructuras descritas por cristalografía de Rayos-X, resolución igual o menor a 2Å, un mínimo de identidad de secuencias de 70% y un E “*value*” menor a 1.

3.7.2. Servidor en línea CLUSTAL W2

ClustalW2 es un programa de alineamiento múltiple de secuencias de proteínas o ADN para propósitos generales. Una vez seleccionada la secuencia molde (“*template*”) pertinente según BLASTp, se alineó con la secuencia objetivo utilizando CLUSTAL W2 online de acuerdo a las indicaciones del protocolo de modelamiento de proteínas descrito por Bordoli y cols. . Los alineamientos se descargaron en formato FASTA.

3.7.3. Servidor en línea SwissModel

SwissModel es un servidor de modelamiento de estructuras proteicas por homología, accesible a través del servidor web ExPASy, o desde el programa DeepView (Swiss Pdb-

Viewer). El propósito de este servidor es hacer el modelamiento de proteínas accesible a todos los bioquímicos y biólogos moleculares .

El alineamiento generado a través de CLUSTAL W2 se cargó en el servidor SwissModel, se asignó un nombre de proyecto y se procedió a construir el modelo de acuerdo al protocolo del software.

A continuación se presenta un esquema que sintetiza el trabajo realizado en esta memoria para optar al título profesional de bioquímica (**Diagrama 1**)

4. Diagrama de flujo de la investigación bioinformática de OLR1 y LOX1 en humano, ratón y rata

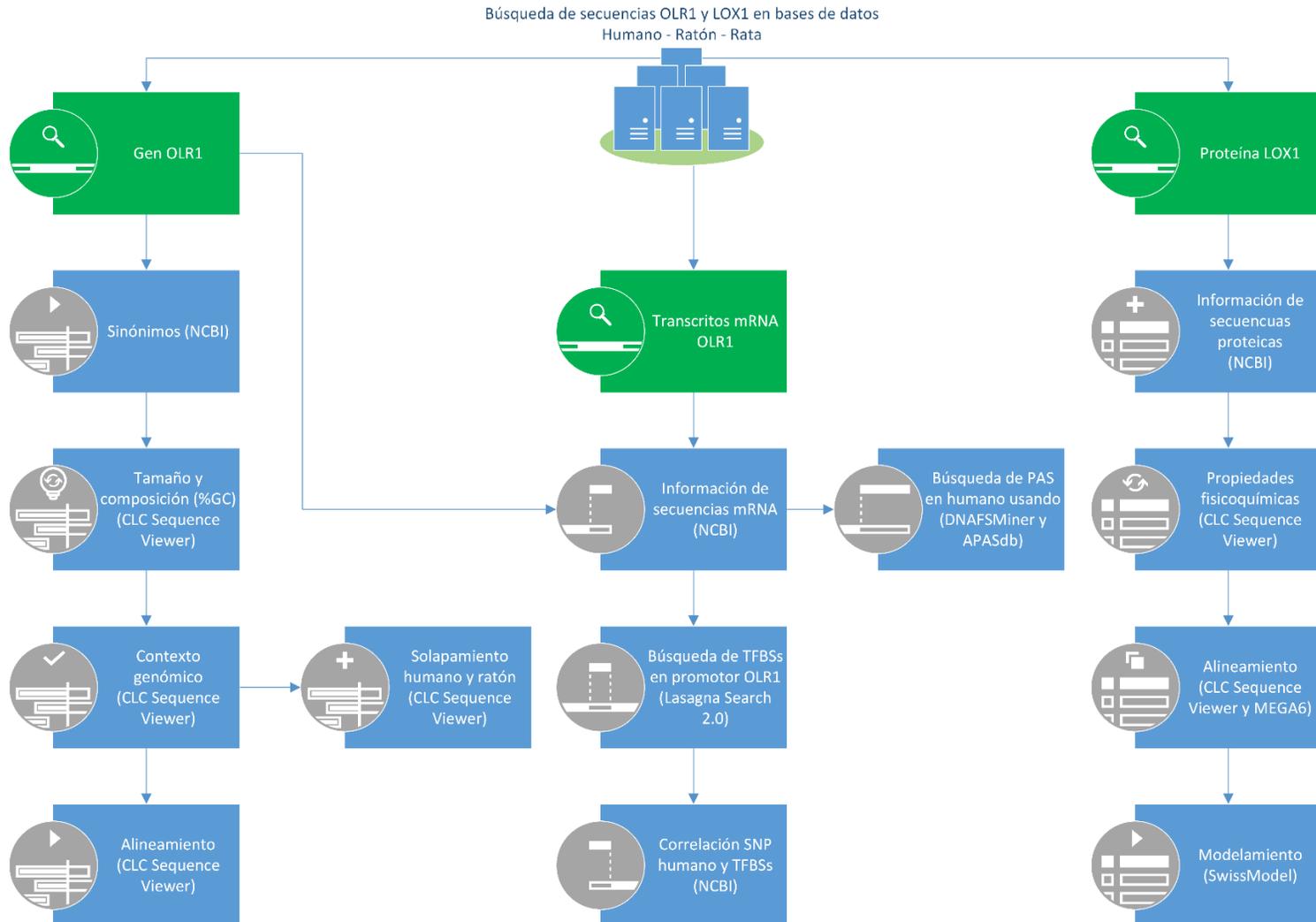


Diagrama 1. Esquema de trabajo de tesis. Se observa división de la investigación en etapas según tema, las operaciones realizadas y el programa y/o base de datos ocupado.

5. RESULTADOS Y DISCUSIÓN

5.1. Gen OLR1 en humano, ratón y rata

5.1.1. Sinónimos de LOX1 para gen y proteína

El primer objetivo de este trabajo fue seleccionar las secuencias de genes, transcritos y proteínas en bases de datos referentes a LOX1 de humano, ratón y rata. Para ello se buscaron las secuencias del gen OLR1 humano en la base de datos de NCBI, . Se encontró una página sinopsis que agrupa la descripción resumida de las características del gen OLR1 y su vecindad, ARNm, proteínas, SNPs, entre otra información.

El gen humano OLR1 codifica al receptor LOX1, perteneciente a la superfamilia lectina tipo C (Yoshimoto R, 2011) y está ubicado en el brazo corto del cromosoma 12, específicamente en la posición p13.2 (Gene ID: 4973, 2015) como copia única en el genoma haploide (OLR1 Gene, GeneCards, 2015). Presentó ortólogos en organismos mamíferos como chimpancé (*Pan troglodytes*), vaca (*Bos taurus*), perro (*Canis familiaris*), ornitorrinco (*Ornithorhynchus anatinus*), ratón (*Mus musculus*), rata (*Rattus norvegicus*) y también reptiles como lagartija (*Anolis carolinensis*) (MSN Gene, 2015). En ratón, el gen se ubica en el cromosoma 6 (Gene ID: 108078, 2015), y en rata en el cromosoma 4 (Gene ID: 140914, 2015); en ambos casos está presente como copia única en el genoma haploide.

En la **Tabla 1** se agruparon los sinónimos que recibe el gen OLR1 según especie, y se indica la página web de referencia correspondiente. Como se observa, no existe un sinónimo común a las tres especies, sólo “LOX-1” es sinónimo común entre ratón y rata. En la **Tabla 2** se agruparon los sinónimos que recibe la proteína LOX1 según especie, indicando la página web de referencia correspondiente. Se registraron tres sinónimos comunes para LOX1 a las tres especies (“*lectin-type oxidized LDL receptor 1*”, “*oxidized low-density lipoprotein receptor 1*” y “*ox-LDL receptor 1*”), dos sinónimos comunes a rata y ratón (“*lectin-like oxidized LDL receptor 1*” y “*lectin-like oxLDL receptor 1*”) y uno común para humano y rata (“*oxidized low density lipoprotein (lectin-like) receptor 1*”).

Tabla 1. Sinónimos del gen OLR1 para *Homo sapiens*, *Mus musculus* y *Rattus Norvegicus*.

Nombre Gen	Organismo	Referencia - Web
OLR1	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
CLEC8A	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
LOX1	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
LOXIN	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
SCARE1	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
SLOX1	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
LOX1	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
LOX-1	<i>Mm/Rn</i>	http://www.ncbi.nlm.nih.gov/gene/108078 http://www.ncbi.nlm.nih.gov/gene/140914
SR-EI	<i>Mm</i>	http://www.ncbi.nlm.nih.gov/gene/108078
Scare1	<i>Mm</i>	http://www.ncbi.nlm.nih.gov/gene/108078
Oldlr1	<i>Rn</i>	http://www.ncbi.nlm.nih.gov/gene/140914
Oldr1	<i>Rn</i>	http://www.ncbi.nlm.nih.gov/gene/140914

Hs: *Homo sapiens*; *Mm*: *Mus musculus*; *Rn*: *Rattus norvegicus*

Tabla 2. Sinónimos de la proteína LOX1 para *Homo sapiens*, *Mus musculus* y *Rattus Norvegicus*.

Nombre Proteína	Organismo	Referencia - Web
C-type lectin domain family 8 member A	<i>Hs</i>	
hLOX-1	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
oxidized low-density lipoprotein receptor 1, soluble form	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
scavenger receptor class E, member 1	<i>Hs</i>	http://www.ncbi.nlm.nih.gov/gene/4973
lectin-type oxidized LDL receptor 1	<i>Hs/Mm/Rn</i>	http://www.ncbi.nlm.nih.gov/gene/4973 http://www.ncbi.nlm.nih.gov/gene/108078 http://www.ncbi.nlm.nih.gov/gene/140914
oxidized low-density lipoprotein receptor 1	<i>Hs/Mm/Rn</i>	http://www.ncbi.nlm.nih.gov/gene/4973 http://www.ncbi.nlm.nih.gov/gene/108078 http://www.ncbi.nlm.nih.gov/gene/140914
ox-LDL receptor 1	<i>Hs/Mm/Rn</i>	http://www.ncbi.nlm.nih.gov/gene/4973
oxidised low density lipoprotein (lectin-like) receptor 1	<i>Hs/Rn</i>	http://www.ncbi.nlm.nih.gov/gene/4973 http://www.ncbi.nlm.nih.gov/gene/140914
lectin-like oxidized LDL receptor 1	<i>Mm/Rn</i>	http://www.ncbi.nlm.nih.gov/gene/108078 http://www.ncbi.nlm.nih.gov/gene/140914
lectin-like oxLDL receptor 1	<i>Mm/Rn</i>	http://www.ncbi.nlm.nih.gov/gene/108078 http://www.ncbi.nlm.nih.gov/gene/140914
lectin-like oxidized low-density lipoprotein receptor	<i>Rn</i>	http://www.ncbi.nlm.nih.gov/gene/140914
Lectin-like oxidized low-density lipoprotein receptor-1	<i>Rn</i>	http://www.ncbi.nlm.nih.gov/gene/140914

Hs: *Homo sapiens*; *Mm*: *Mus musculus*; *Rn*: *Rattus norvegicus*

5.1.2. Secuencias del gen OLR1 en humano, rata y ratón

La búsqueda de secuencias de OLR1 a través de CLC Sequence Viewer 7.5 arrojó gran cantidad de anotaciones de tipo genómicas, ARNm y proteicas para OLR1 y LOX1. La información sobre secuencias de tipo genéticas de OLR1 se agrupó en las tablas a continuación.

Se puede observar en las **Tabla 3** y **Tabla 4** (**Tabla 13** y **Tabla 14** en **ANEXOS**), que existe un mayor número de secuencias génicas de OLR1 descritas para humano, que para ratón y para rata; también se indican varios códigos para una misma secuencia, dependiendo de su origen. En todas las tablas se registra información incompleta respecto de las anotaciones. En la **Tabla 14** (**ANEXOS**) se indica que la anotación más antigua de gen OLR1 data del año 1999 (AJ131757.1) y la más reciente del 2014 (NC_005103.4).

La caracterización de un gen incluye la información respecto de su tamaño, inicio y final de transcripción. La **Tabla 3** da cuenta del tamaño, posición de inicio, posición final, secuencias génicas asociadas y muestra biológica que originó la secuencia de gen OLR1. Existe una diferencia significativa en tamaño entre anotación NG_016743.1 y el resto de las secuencia de gen OLR1 en humano. También se indican diferencias en los puntos de inicio y final de las secuencias del gen, así como en el origen de las muestras (cuando se reporta).

Los genes eucarióticos incluyen una región promotora que incluye secuencias regulatorias de la transcripción, secuencias de intrones, secuencias de exones y secuencias de poliadenilación que indican el sitio de corte y unión de cola poliA al transcrito. En la **Tabla 4** se muestran las secuencias NG_016743.1 y AJ131757.1 como únicas anotaciones que reportaron número de exones en humano. Sin embargo, solo los exones 2, 3 y 5 son similares entre ambas secuencias de gen OLR1. El resto de los exones difiere en 20 o más pares de bases de longitud. En rata y ratón no se reporta información respecto de los exones.

Tabla 3. Información de secuencias de gen OLR1 para humano, ratón y rata, relativas a base de datos NCBI.

Organismo	GenBank (NCBI Reference Sequence)	Tamaño	Especificaciones muestras	Inicio	Final	Secuencia asociada
<i>Hs</i>	NG_016743.1	20892bp	Células endotelio vascular	5001	19892	No Reportado
	AJ131757.1	15387bp	No especificado	0	15387	No Reportado
	CM000263.1	13841bp	No especificado	15471561	15485401	REEMPLAZA A CH003483 (05/02/2004)
	AC_000144.1	13889bp	El ensamblaje HuRef representa una versión haploide, compuesta de la secuencia del genoma diploide de un solo individuo. Nombre del Donante: J. Craig Venter Fecha de nacimiento: 14 de octubre 1946 Sexo: Masculino Caucásico Descendencia: Europea - Inglaterra.	10054056	10067944	No Reportado
	NC_018923.2	13882bp	Línea celular CHM1htert; mola hidatiforme (tumor uterino).	10280199	10204080	No Reportado
	NC_000012.1 2	13892bp	No especificado	10158300	10172191	NC_000012 (29/08/2002)
<i>Mm</i>	NC_000072.6	22109bp	Cepa C57BL/6J	129485247	12950735 5	NC_000072 (02/09/2004)
	AC_000028.1	23746bp	Cinco cepas diferentes de ratón: A/J, DBA/2J, 129X1/SvJ, 129S1/SvImJ and C57BL/6J.	131207291	13123103 6	No Reportado
	CM000214.2	23752bp	Contiene 27 millones de lecturas de Celera, a partir de cuatro cepas de <i>Mus musculus</i> (129X1/SvJ, 129S1/SvImJ, DBA/2J y A/J), 13 millones de lecturas de cepas C57BL/6 a partir de NCBI Trace Archive, 0,4 millones de secuencias terminadas de BAC desde TIGR, y secuencias BACs terminadas y no terminadas a partir de NCBI. (Nature 2002. 420:520-562).	128207285	12823103 6	CM000214.1 (08/08/2005)
<i>Rn</i>	NC_005103.4	22085bp	Hembras de cepas BN/SsNHsdMCW, más un macho cepa SHR (también conocida como SHR-Akr)	163239853	16326193 7	NC_005103 (16/10/2003)
	AC_000072.1	22076bp	Brown Norway y Sprague-Dawley.	151610779	15163285 4	No Reportado

Contenido de cada columna de izquierda a derecha: especie; nombre de la secuencia en GenBank; tamaño de la secuencia del gen; especificaciones de la muestra, tales como de que tejido proviene, casta del organismo, etc; ubicación de la primera base nitrogenada, de acuerdo a lo reportado en NCBI para cada secuencia; ubicación de la última base nitrogenada, de acuerdo a lo reportado en NCBI para cada secuencia; asociaciones de otras secuencias relativas a la secuencia en cuestión, según se reporta en NCBI. *Hs*: Homo sapiens; *Mm*: Mus musculus; *Rn*: Rattus norvegicus; pb: pares de bases.

Tabla 4. Información de secuencias de gen OLR1 para humano, ratón y rata, relativas a base de datos NCBI.

Organismo	GenBank (NCBI Reference Sequence)	N° Exones	Exon 1	Exon 2	Exon3	Exon 4	Exon 5	Exon 6
<i>Hs</i>	NG_016743.1	6	5001-5190 (189 bp)	8017-8118 (101 bp)	10235-10480 (245 pb)	16266-16407 (141 bp)	16730-16845 (115 pb)	17171-19892 (2721 pb)
	AJ131757.1	6	3192-3267 (75 pb)	6092-6194 (102 bp)	8311-8557 (246 pb)	14361-14480 (119 pb)	14805-14920 (115 pb)	15246-15387 (141 pb)

En cada columna de izquierda a derecha se indica: Especie; nombre de la secuencia en GenBank; n° de exones reportados en NCBI para cada secuencia, en caso de que corresponda; columnas 4-9 indican posición de primera y última base nitrogenada de cada exón, y tamaño de exones, según reporta NCBI. *Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Rn*: *Rattus norvegicus*; pb: *pares de bases*.

5.1.3. Tamaño de secuencias y composición (%GC) del gen OLR1 en humano, rata y ratón.

En las tres especies en estudio, la búsqueda del gen OLR1 arrojó anotaciones suscritas al mismo gen y especie por distintos grupos de investigadores. Las diferencias entre estas anotaciones, en cuanto a tamaño y composición (guanina-citosina), se agruparon en la **Tabla 5**. Como se observa, no existen grandes diferencias entre anotaciones de gen OLR1 en cuanto a tamaños reportados (exceptuando la anotación AJ131757.1 en humano) en una misma especie, pero sí existen diferencias al comparar entre especies. Se distingue que la composición es conservada en humano, ratón y rata.

Tabla 5. Tamaño y composición en gen OLR1.

GenBank	Tamaño (pb)	Composición GC
<i>Homo sapiens</i>		
NG_016743.1	13.892	0.387
AJ131757.1	15.387	0.390
CM000263.1	13.841	0.388
AC_000144.1	13.889	0.387
NC_018923.2	13.882	0.387
NC_000012.12	13.892	0.387
<i>Mus musculus</i>		
NC_000072.6	22.109	0.386
AC_000028.1	23.746	0.386
CM000214.2	23.752	0.386
<i>Rattus norvegicus</i>		
NC_005103.4	22.085	0.390
AC_000072.1	22.076	0.390

Se muestra a que especies pertenecen las anotaciones en celdas amarillas. Cada columna, de izquierda a derecha, indica: el nombre de acceso en GenBank de cada anotación del gen OLR1; la tamaño en pares de bases (pb) de la secuencia; la frecuencia de guanina y citosina en la secuencia de gen OLR1. pb: pares de bases; GC: guanina y citosina.

5.1.4. Contexto genómico

El contexto genómico en que está inmerso un gen específico permite identificar sus genes vecinos, y por tanto comparar la conservación de la vecindad genómica entre diferentes especies. Esta información puede ser interesante para determinar posible regulación coordinada en la vecindad y el conjunto de circunstancias que condicionan la transcripción y traducción del gen.

En las **Figura 5** a la **Figura 15** se esquematizaron los contextos genómicos del gen OLR1 en humano, rata y ratón. Se muestran gráficamente los genes vecinos, los exones y ARNm de OLR1 según la información que reporta visualmente cada secuencia encontrada, seleccionada y descargada desde la base de datos de NCBI , utilizando CLC Sequence Viewer. Se establece sintenia entre las tres especies en estudio, por tanto existe regiones genómicas ortólogas, pudiendo reflejar relaciones funcionales importantes entre los genes de esta región.

5.1.4.1. *Homo sapiens*

Se puede observar que existen dos secuencias génicas que representan a OLR1 sin vecindad genómica y entregan información sobre sus exones (**Figura 5** y **Figura 6**), y solo la secuencia AJ131757 entrega información sobre el promotor de OLR1 (**Figura 5**). Se distingue la vecindad genómica de OLR1 y solapamiento antisentido de OLR1 humano por su región promotora y 5' con el gen vecino TMEM52B (**Figura 7**, **Figura 9** y **Figura 10**) y hCG_2039306 (**Figura 8**).

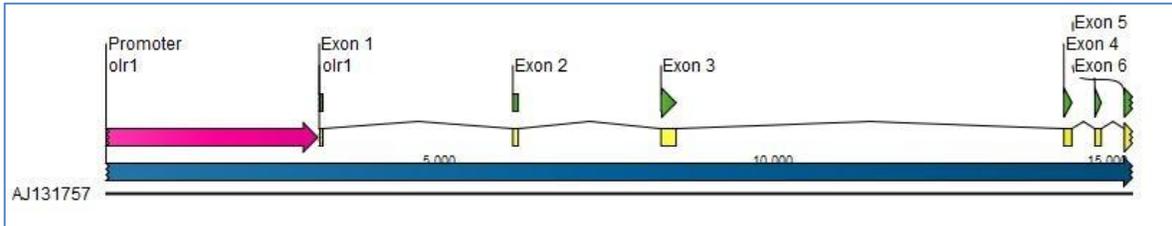


Figura 5. Contexto genómico de OLR1, secuencia AJ131757 de humano. Se muestra el gen OLR1 (azul), promotor de gen OLR1 (fucsia), ARNm (amarillo) y exones (verde). Esta secuencia no reporta genes vecinos a OLR1. Secuencia descargada desde NCBI .

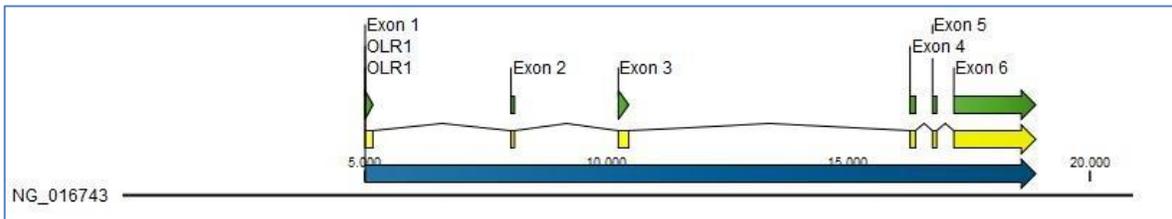


Figura 6. Contexto genómico de OLR1, secuencia NG_016743 de humano. Se muestra gen OLR1 (azul), ARNm (amarillo) y exones (verde). Esta secuencia no reporta genes vecinos a OLR1. Secuencia descargada desde NCBI .

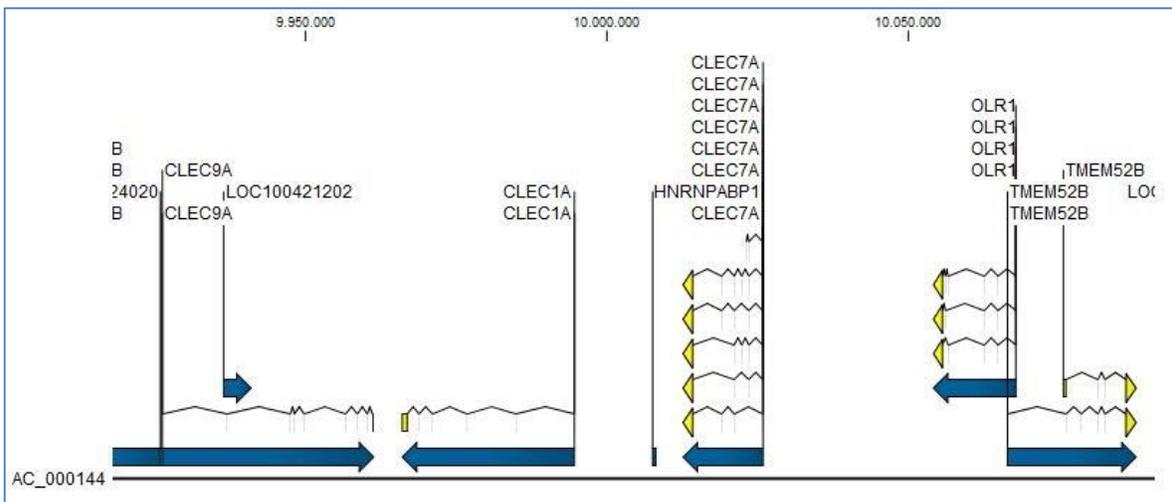


Figura 7. Contexto genómico de OLR1, secuencia AC_000144 de humano. Se muestra genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de OLR1, de izquierda a derecha: CLEC9A, CLEC1A, CLEC7A, OLR1 Y TMEM52B. Se visualiza solapamiento de genes OLR1 y TMEM52B, por la parte 5' UTR de ambos genes. Secuencia descargada desde NCBI .

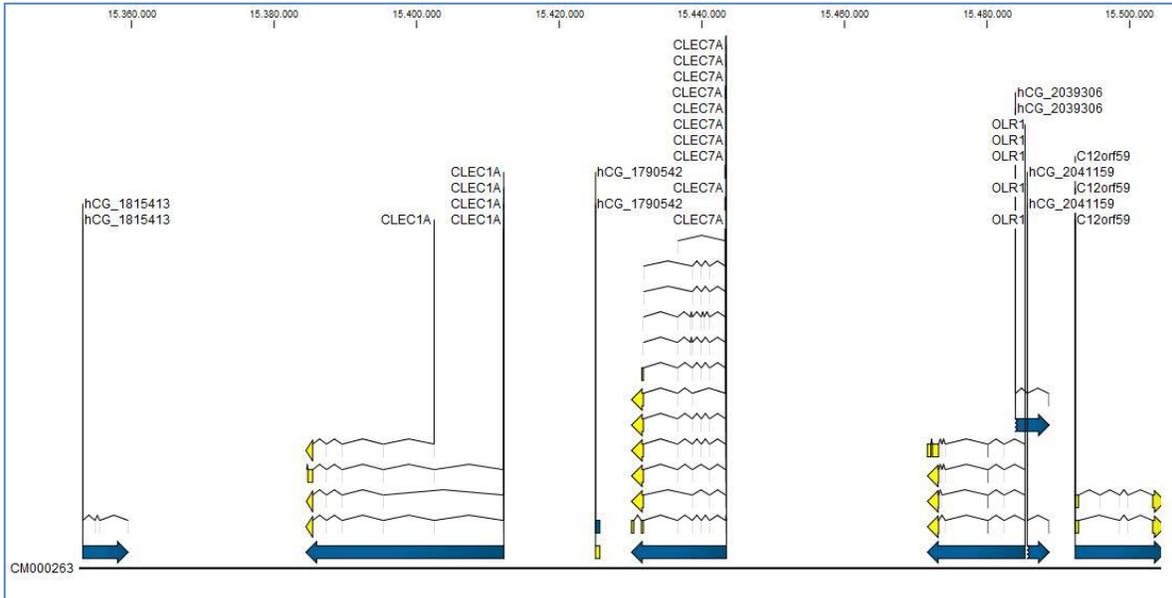


Figura 8. Contexto genómico OLR1, secuencia CM000263 de humano. Se muestran genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de OLR1, de izquierda a derecha: hCG_1815413, CLEC1A, hCG_1790542, CLEC7A, OLR1, hCG_2039306 y C12orf59. Se visualiza solapamiento de genes OLR1 y hCG_2039306, por la parte 5' UTR de ambos genes. Esta secuencia no reporta exones. Secuencia descargada desde NCBI .

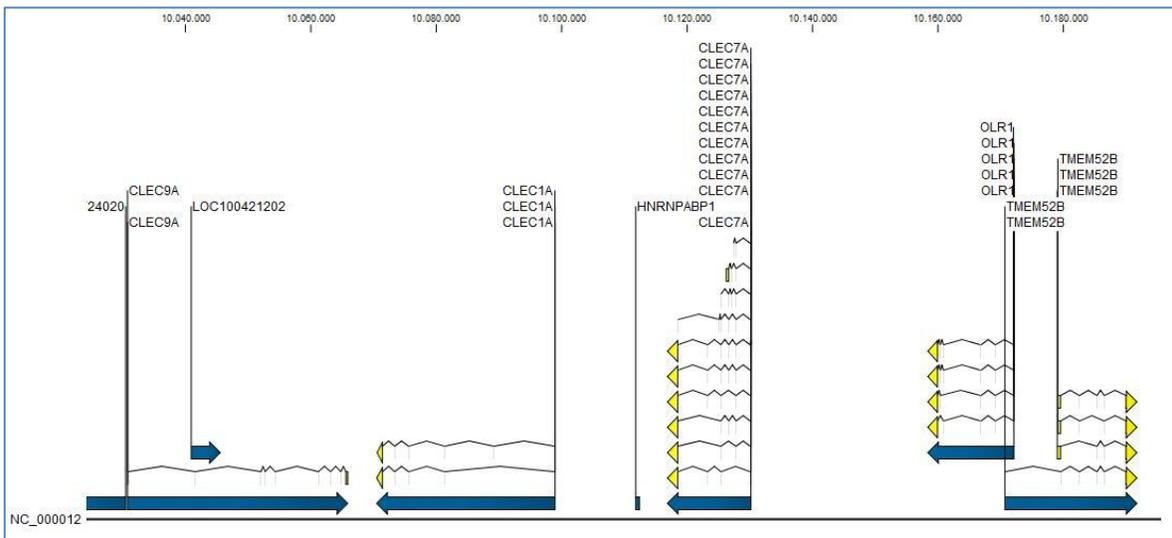


Figura 9. Contexto genómico OLR1, secuencia NC_000012 de humano. Se muestran genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de OLR1, de izquierda a derecha: CLEC9A, LOC100421202, CLEC1A, HNRNPABP1, CLEC7A, OLR1, TMEM52B. Se visualiza solapamiento de genes OLR1 y TMEM52B, por la parte 5' UTR de ambos genes. Esta secuencia no reporta exones. Secuencia descargada desde NCBI .

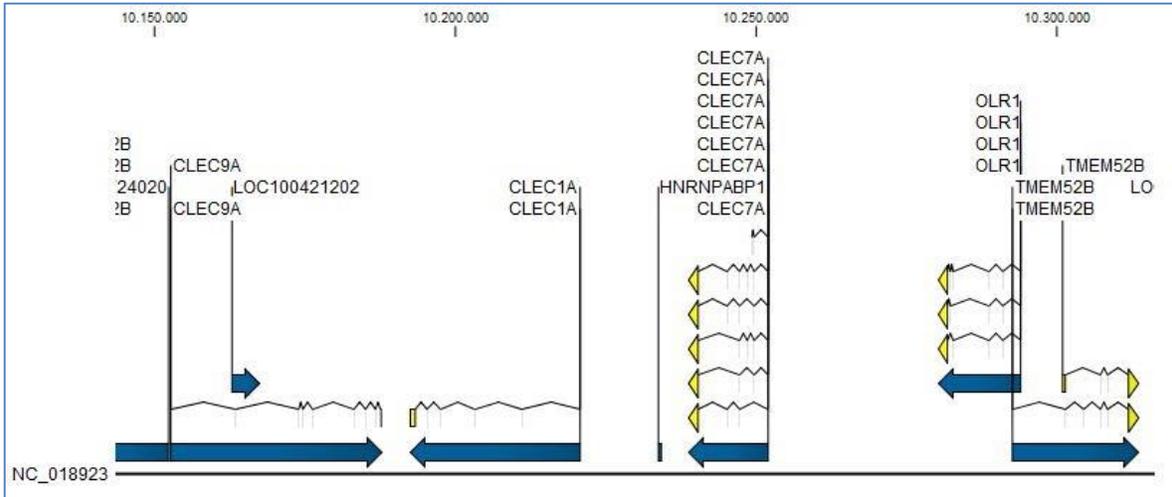


Figura 10. Contexto genómico OLR1, secuencia NC_018923 de humano. Se muestran genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de OLR1, de izquierda a derecha: CLEC9A, LOC100421202, CLEC1A, HNRNPABP1, CLEC7A, OLR1, TMEM52B. Se visualiza solapamiento de genes OLR1 y TMEM52B, por la parte 5' UTR de ambos genes. Esta secuencia no reporta exones. Secuencia descargada desde NCBI .

5.1.4.2. *Mus musculus*

En las **Figura 11**, **Figura 12** y **Figura 13**, correspondientes a ratón se muestran los transcritos originados por las secuencias de gen OLR1. Ninguna secuencia descargada presenta información respecto al promotor y/o exones de gen OLR1, así como tampoco existe una secuencia exclusiva para el gen OLR1. Además, en todas las figuras se distingue la vecindad genómica de OLR1, con vecinos comunes en todas las secuencias. Sólo en la **Figura 13**, correspondiente a la secuencia NC_000072, se distingue solapamiento antisentido de la región 3' de OLR1 y la región 5' de CLEC12B con el gen vecino CLEC9A, y genes CLEC1A y CLEC7A anidados en CLEC9A.

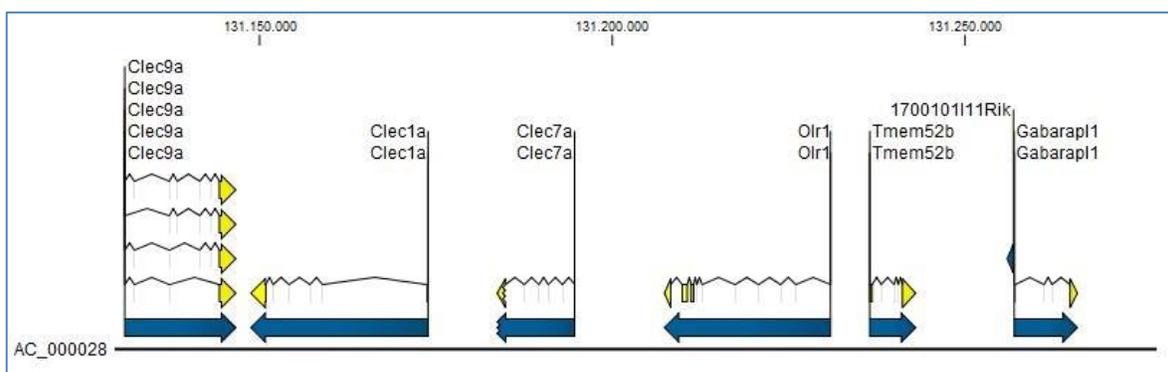


Figura 11. Contexto genómico OLR1, secuencia AC_000028 de ratón. Se muestran genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de Olr1, de izquierda a derecha: Clec9a, Clec1a, Clec7a, Olr1, Tmem52b, 170010111Rik y Gabarapl1. Esta secuencia no reporta exones. Secuencia descargada desde NCBI .

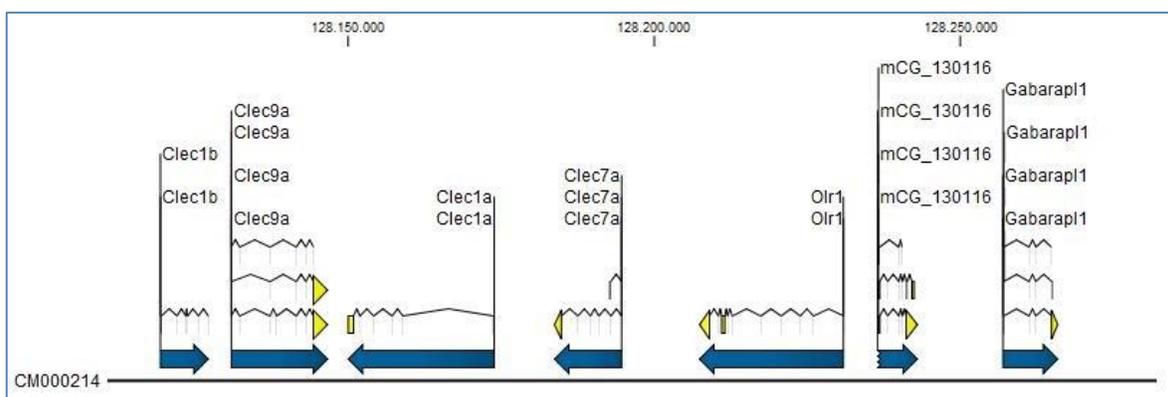


Figura 12. Contexto genómico OLR1, secuencia CM000214 de ratón. Se muestran genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de OLR1, de izquierda a derecha: Clec1b, Clec9a, Clec1a, Clec7A, Olr1, mCG_130116 y Gabarapl1. Esta secuencia no reporta exones. Secuencia descargada desde NCBI .

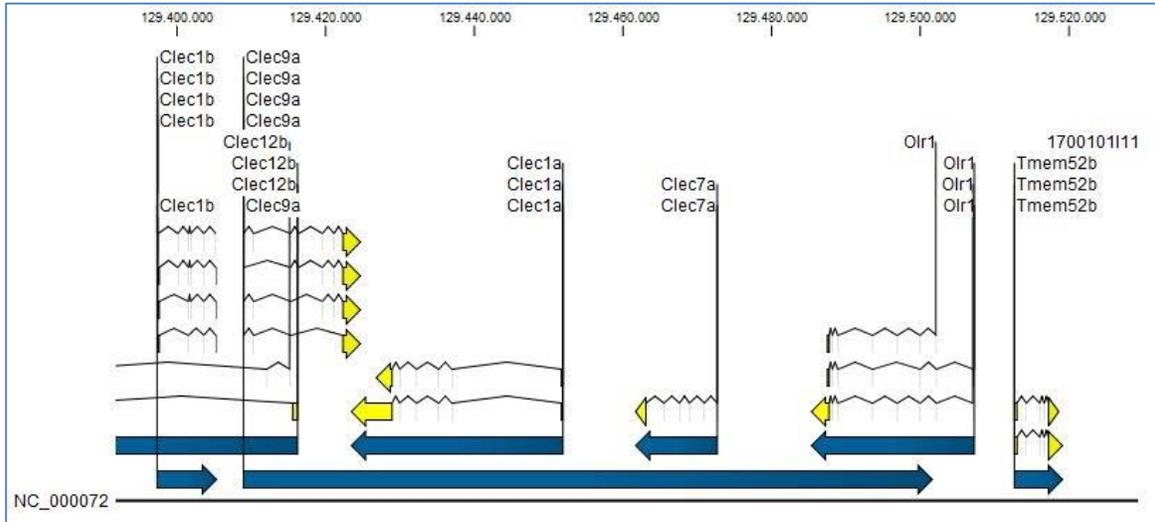


Figura 13. Contexto genómico OLR1, secuencia NC_000072 de ratón. Se muestran genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de OLR1, de izquierda a derecha: Clec1b, Clec9a, Clec12b, Clec1a, Clec7a, Olr1 y Tmem52b. Se observa solapamiento de genes Clec1b con Clec12b; y de Clec12b, Clec1a, Clec7a y Olr1 con Clec9a. Esta secuencia no reporta exones. Secuencia descargada desde NCBI .

5.1.4.3. *Rattus norvegicus*

En las **Figura 14** y **Figura 15** se muestran los transcritos originados por las secuencias de gen OLR1 correspondientes a rata. Ninguna secuencia descargada presenta información respecto a solapamientos, promotor y/o exones de gen OLR1, así como tampoco existe una secuencia exclusiva para gen OLR1. Además, se distinguen las vecindades genómicas de OLR1 con vecinos comunes en ambas secuencias, que parecieran ser idénticas a simple vista.

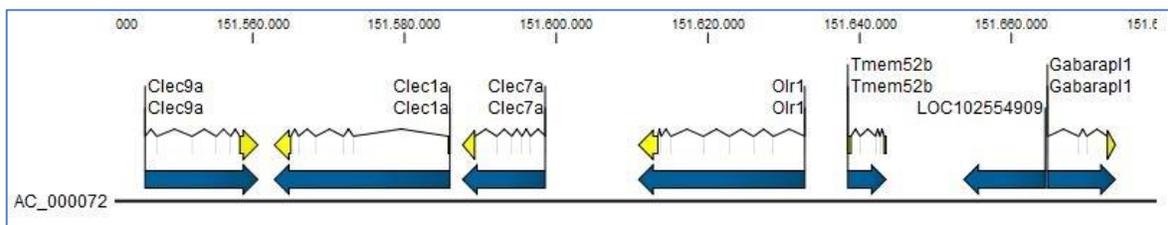


Figura 14. Contexto genómico OLR1, secuencia AC_000072 de rata. Se muestran genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de OLR1, de izquierda a derecha: Clec9a, Clec1a, Clec7a, Olr1, Tmem52b y Gabarapl1. Esta secuencia no reporta exones. Secuencia descargada desde NCBI .

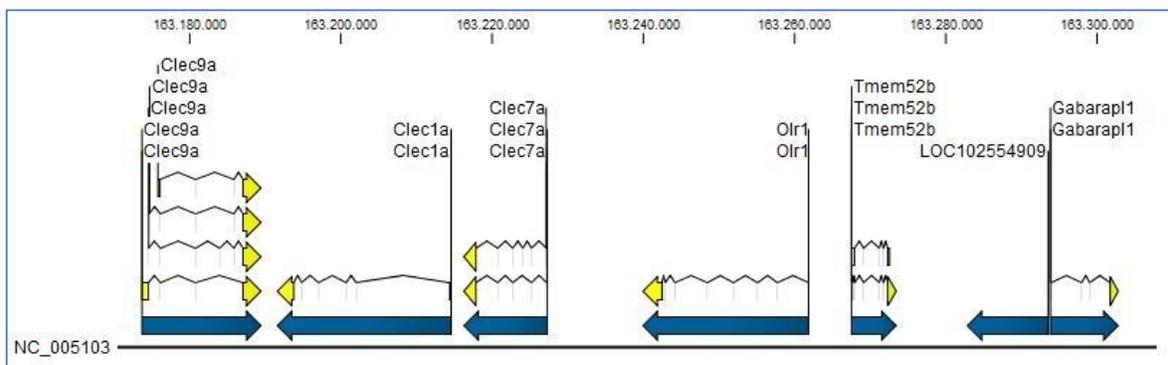


Figura 15. Contexto genómico OLR1, secuencia NC_005103 de rata. Se muestran genes (azul) y ARNm (amarillo). Esta secuencia muestra los genes vecinos de OLR1, de izquierda a derecha: Clec9a, Clec1a, Clec7a, Olr1, Tmem52b y Gabarapl1. Esta secuencia no reporta exones. Secuencia descargada desde NCBI .

5.1.5. Solapamiento de genes

Se analizaron los segmentos comunes a genes solapados de acuerdo a cada secuencia en que se encuentra presente este fenómeno. Los resultados resumidos en la **Tabla 6** indican que en humano la región solapada entre el gen OLR1 y el gen TMEM52B es de 1.593pb en hebras opuestas, mientras que entre el gen OLR1 y el gen hCG_2039306 es de 1.593pb de tamaño en hebras opuestas. En ratón los genes OLR1 y CLEC9A comparten 16.455pb solapadas.

De acuerdo a la base de datos Uniprot y NCBI , TMEM52B en humano traduce 2 isoformas de la proteína de transmembrana 52, su sinónimo proteico es C12orf59 y no se le conoce función. La única prueba de su existencia es el estudio proteómico desarrollado en exosomas urinarios por Gonzales en el año 2009 . Además, existen 2 secuencias de carácter predictivo en NCBI.

En ratón el gen es conocido como *tmem52b*, traduce la proteína transmembrana 52B sin función conocida y su existencia se atribuye al estudio del transcriptoma realizado por Diez-Roux en el año 2011 . En rata se conoce a la forma proteica como *Tmem52B* y/o RCG30157, no tiene función descrita y no existen estudios asociados a *Tmem52B* o su existencia.

En humano, ratón y rata hay evidencias de que *Clec9A* codifica una proteína cuya función corresponde a un receptor de endocitosis en un subgrupo de células mieloides y que participa en la toma y procesamiento de material proveniente de células muertas, incluyendo el reconocimiento de formas filamentosas de actina asociadas a proteínas del citoesqueleto con dominios de unión a actina .

Tabla 6. Solapamiento de OLR1 en humano y ratón.

Organismo	Secuencia	Genes solapados	Largo solapamiento	Inicio	Final
<i>Hs</i>	AC_000144	TMEM52B/OLR1	1593 pb	10066352	10067944
	CM000263	hCG_2039306/OLR1	1509 pb	15483393	15485401
	NC_000012	TMEM52B/OLR1	1593 pb	10170599	10172191
	NC_018923	TMEM52B/OLR1	1593 pb	10292488	10294080
<i>Mm</i>	NC_000072	Clec9A/OLR1	16455 pb	129485247	129501701

Se indica el organismo; nombre de la secuencia; nombre de los genes solapados separados por una diagonal; largo de la segmento de genes solapados en pares de bases; la base de inicio y fin del solapamiento en gen OLR1. *Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Rn*: *Rattus norvegicus*; pb: pares de bases.

Siguiendo el mismo protocolo de trabajo utilizado en esta memoria, se analizaron los tamaños de genes e isoformas proteicas de los genes solapados con OLR1. Los resultados expresados en la **Tabla 7** indican que el número de isoformas para una proteína tiene relación con el solapamiento encontrado a nivel genético: en humano LOX1 presenta tres isoformas proteicas, TMEM52B dos y CLEC9A una; en ratón LOX1 presenta tres isoformas proteicas, Tmem52b una y CLEC9A cinco; en rata existe 1 isoforma por cada tipo proteico, coincidente con la ausencia de solapamientos.

Al analizar el número de aminoácidos de las isoforma 1 de proteína LOX1 entre especies se distinguen tamaños relativamente similares, independiente del tamaño del gen OLR1. Las diferencias en tamaño del gen se relacionan con la pérdida de intrones en OLR1 humano y ganancia de intrones en CLEC9A y TMEM52B humano y CLEC9A de ratón (**Tabla 7, Figura 10, Figura 13, Figura 14 y Figura 19**).

Tabla 7. Tamaño de genes e isoformas de OLR1 y genes que lo solapan.

Organismo	Gen								
	OLR1		Clec9A		TMEM52B				
	Tamaño gen (pb)	Número de aminoácidos proteína	Tamaño gen (pb)	Número de aminoácidos proteína	Tamaño gen (pb)	Número de aminoácidos proteína			
<i>Hs</i>	13892	Isoforma		35290	Isoforma		21206	Isoforma	
		1	273		-	241		1	183
		2	181					2	163
		3	189						
<i>Mm</i>	22109	1	363	92840	1	238	6679	P	186
		2	155		2	237			
		3	189		3	173			
					4	264			
					5	148			
<i>Rn</i>	22085	-	364	15936	-	241	6119	P	192

Información extraída desde base de datos Uniprot y NCBI. Se denota con una letra P a las isoformas Predictivas. *Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Rn*: *Rattus norvegicus*; pb: pares de bases; aa: aminoácidos; P: predictivo.

5.1.6. Alineamiento de genes

Con la finalidad de determinar si existen diferencias en secuencias del gen OLR1 reportadas para humano, ratón y rata se realizaron alineamientos de las secuencias de genes para cada especie.

De acuerdo a los resultados abreviados registrados en la **Tabla 8**, las secuencias del gen OLR1 con menor variación en ácidos nucleicos corresponden a rata y ratón (**Figura 17**). En humano se distingue una menor proporción de sitios conservados y mayor proporción de sitios variables e informativos según parsimonia, indicando gran variabilidad de secuencias del gen OLR1 en la especie humana, incluso al descartar la secuencia con mayor tamaño y antigüedad descrita para OLR1 (AJ131757.1) (**Tabla 8** y **Figura 16**). Al realizar un alineamiento de las secuencias de gen OLR1 en humano, ratón y rata se observa escasa cobertura, proporción de sitios conservados, proporción de sitios variables y gran proporción de sitios únicos, indicando extensas regiones de delección en el alineamiento de la secuencia de gen OLR1 de las 3 especies (**Figura 18**). Para comparar las secuencias de genes entre las tres especies se escogió la secuencia más representativa de OLR1 para cada especie. Para esto se consideró la información entregada en las **Tabla 3**, **Tabla 4** y **Tabla 14 (ANEXOS)** de este trabajo.

En la sección **ANEXOS** de esta memoria de título, en su versión digital, se muestran en detalle los alineamientos realizados. (**Archivo vinculado 1**, **Archivo vinculado 2**, **Archivo vinculado 3**, **Archivo vinculado 4** y **Archivo vinculado 5**)

Tabla 8. Tabla resumen de la proporción de sitios informativos a partir del alineamiento de secuencias gen OLR1 en humano, ratón y rata.

Alineamiento	Proporción de sitios conservados	Proporción de sitios variables	Proporción de sitios informativos de parsimonia	Proporción de sitios únicos	Proporción de cobertura
Secuencias de humano	0.495	0.375	0.351	0.024	0.646
Secuencias de humano sin AJ131757.1	0.585	0.414	0.413	0.001	0.818
Secuencias ratón	0.983	0.005	0.000	0.005	0.913
Secuencias rata	0.999	0.000	0.000	0.000	0.999
Secuencias tres especies	0.471	0.371	0.000	0.304	0.555

De izquierda a derecha se detalla en cada columna el tipo de alineamiento realizado; proporción de nucleótidos conservados en cada alineamiento relativo al total de nucleótidos; proporción de nucleótidos variables por alineamiento relativo al total de nucleótidos; proporción de sitios informativos de parsimonia, que contienen al menos dos tipos de nucleótidos distintos repitiéndose en al menos dos secuencias, relativo al total de nucleótidos; proporción de sitios únicos (singleton), que contienen al menos dos tipos de nucleótidos con uno de ellos repitiéndose varias veces, relativo al total de nucleótidos; proporción de sitios con cobertura de nucleótidos sin ambigüedades, es decir sin transversiones, relativo al total de nucleótidos (Tabla construida a partir de alineamientos en **Archivo vinculado 1**, **Archivo vinculado 2**, **Archivo vinculado 3**, **Archivo vinculado 4** y **Archivo vinculado 5**).

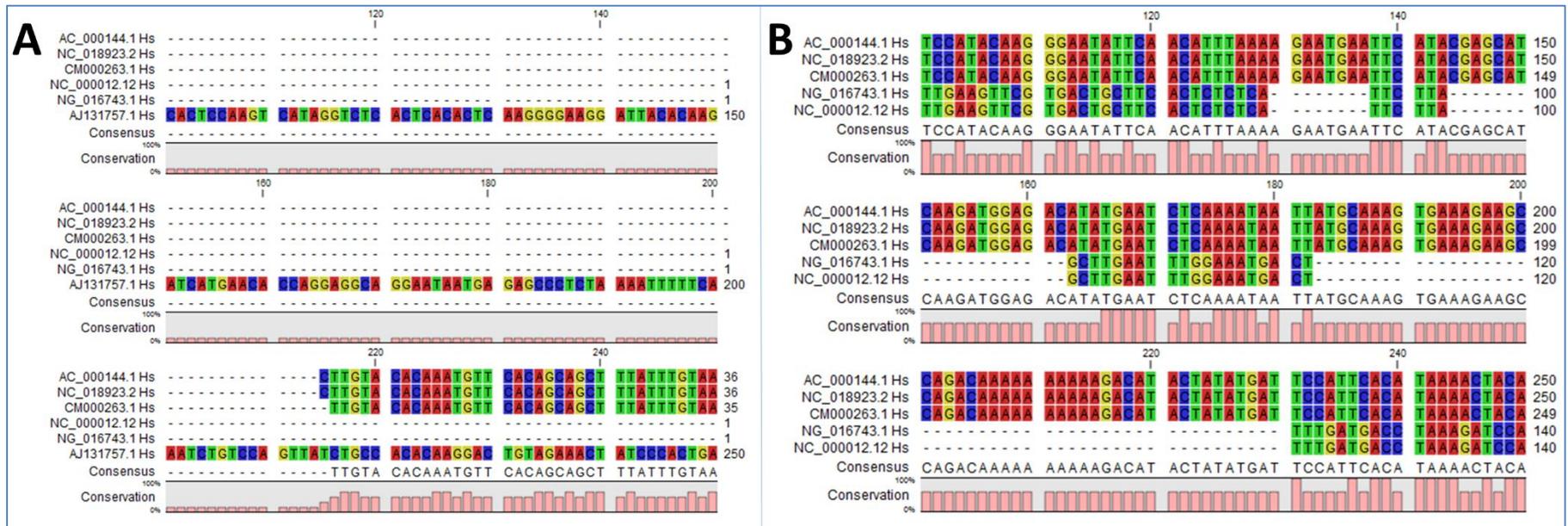


Figura 16. Heterogeneidad de anotaciones de gen OLR1 en humano. (A) Segmento de alineamiento de secuencias de gen OLR1 de humano comprendida entre 100 pb y 250 pb. **(B)** Segmento de alineamiento de secuencias de gen OLR1 humano excluyendo anotación AJ131757, comprendida entre 100 pb y 250 pb.

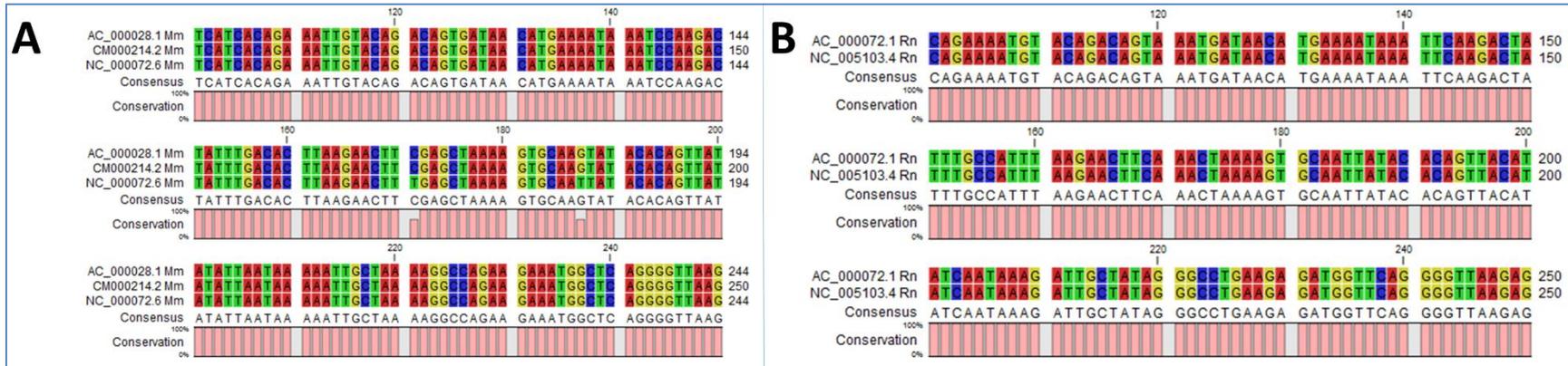


Figura 17. Heterogeneidad de anotaciones de gen OLR1 en ratón y rata. (A) Segmento de alineamiento de secuencias de gen OLR1 en ratón comprendida entre 100 pb y 250 pb. **(B)** Segmento de alineamiento de secuencias de gen OLR1 de rata comprendida entre 100 pb y 250 pb.



Figura 18. Heterogeneidad de anotaciones de gen OLR1 en humano, ratón y rata. Segmento de alineamiento secuencias gen OLR1 de humano, ratón y rata comprendida entre 100 pb y 250 pb.

5.1.7. Discusión sección secuencias gen OLR1 en humano, ratón y rata

Información de secuencias gen OLR1.

Desde una perspectiva general es importante recalcar la cantidad de información que se encuentra en bases de datos. La utilización de sinónimos para un gen y/o proteína entorpece la búsqueda de información, dificulta la masificación y enlace de resultados generados en el mundo (**Tabla 1** y **Tabla 2**). La cantidad de datos que deben ser reportados, catalogados, revisados y posteriormente publicados como parte de una base de datos involucra un trabajo enorme que debe llevarse a cabo con paciencia, conocimiento y gran organización. Llegar a consensos para determinar una secuencia como referencia entre todos los trabajos presentados (con origen en distintos laboratorios, metodologías, líneas de investigación, entre otros) se convierte en una labor gigantesca, poco definida, y con criterios propios de cada base de datos.

En forma particular, la cantidad de información relativa a gen OLR1 (o su parte proteica LOX1) asociada a enfermedades cardiovasculares es limitada y poco minuciosa. En la **Tabla 14** en **ANEXOS** se registra falta de uniformidad en la información de cada secuencia reportada. No se indica en todas las secuencias una definición clara y dirigida a OLR1 o LOX1; de un total de 11 secuencias, solo 3 reportan presentación directa (direct submission); las secuencia humana AJ131757.1 y de ratón CM000214.2 no refieren nombre de ensamblado (Assembly Name, información que notifica sobre el grupo que secuenció), a pesar de ser citadas en investigaciones científicas como indica la columna N° Referencias. Asimismo, no existe explicación alguna en que consiste la actualización de secuencias: ¿Muestras nuevas? ¿Nueva técnica? ¿Revisión?. En la **Tabla 3** llama la atención los distintos puntos de inicio y término de una secuencia para cada nomenclatura, designación arbitraria por cada grupo investigador; así como también la existencia de secuencias asociadas que generalmente fueron reemplazadas por la secuencia detallada ¿Cuál es el criterio para un reemplazo de secuencias?; y la pobre descripción de las muestras que originaron las secuencias, que en el caso de ser especificadas ¿son realmente representativas de la especie en cuestión? Es

difícil estipular los requisitos de publicación de secuencias en bases de datos, pues no es la idea condicionar la forma de secuenciar.

En la **Tabla 7** se muestran los tamaños de genes OLR1 y las isoformas proteicas en las 3 especies en estudio. Se señala que OLR1 en humano es 10.000 pb aproximadamente menor en tamaño que en rata y ratón, pero al comparar la versión proteica de las isoformas 1 no demuestra haber perdido una gran cantidad de aminoácidos (100 aminoácidos aprox. menor que en rata y ratón). Por lo tanto, la disminución en tamaños de OLR1 en humano se explicaría por una pérdida mayoritaria de intrones.

Contexto genómico del gen OLR1.

De un total de 11 secuencias entre las 3 especies en estudio, solo una presenta información sobre el promotor de OLR1 (AJ131757 en humano, **Figura 5**) y solo dos sobre exones descritos para el gen (AJ131757 en **Figura 5** y NG_016743 en **Figura 6** en humano). Ambas secuencias mencionadas anteriormente no reportan información sobre los vecinos genómicos de OLR1. Por el contrario, las 9 secuencias restantes informan sobre los transcritos de ARNm generados a partir del gen, y su vecindad genómica, como se muestra en las **Figura 7** a la **Figura 15**.

En general tanto para humano, ratón y rata se registran contextos genómicos de OLR1 conservados con vecindad genética común. En todas las figuras se observan los distintos tipos de transcritos originados a partir de los genes representados. De acuerdo a las **Figura 7** a la **Figura 15** el ordenamiento de los genes en la vecindad sería Clec9A – Clec1A – Clec7A – OLR1 – TMEM52B, lo que se representa en la **Figura 19**. En casos particulares como los de las **Figura 8** y **Figura 12** el lugar de TMEM52B es ocupado por los genes hCG_2039306 y mCG_130116, respectivamente. Ambas secuencias genómicas están notificadas por el mismo investigador, por lo que es de esperarse que no sean reconocidas como TEMEM52B. A partir de los tamaños reportados en la **Tabla 6**, y la posición en la vecindad, se infiere que se trata del mismo gen con distinto nombre.

Existe sintenia conservada entre humano, ratón y rata en cuanto al vecindario genómico de OLR1, lo que podría indicar un sistema de co-expresión de genes con similar función

(Boldogkoi, 2012). Al-Shahrour y cols. construyeron una cartografía funcional detallada de 8 genomas eucarióticos, entre ellos humano, ratón y rata, revelando que gran parte del ADN está arreglado en vecindades de genes funcionalmente relacionados (Boldogkoi, 2012).

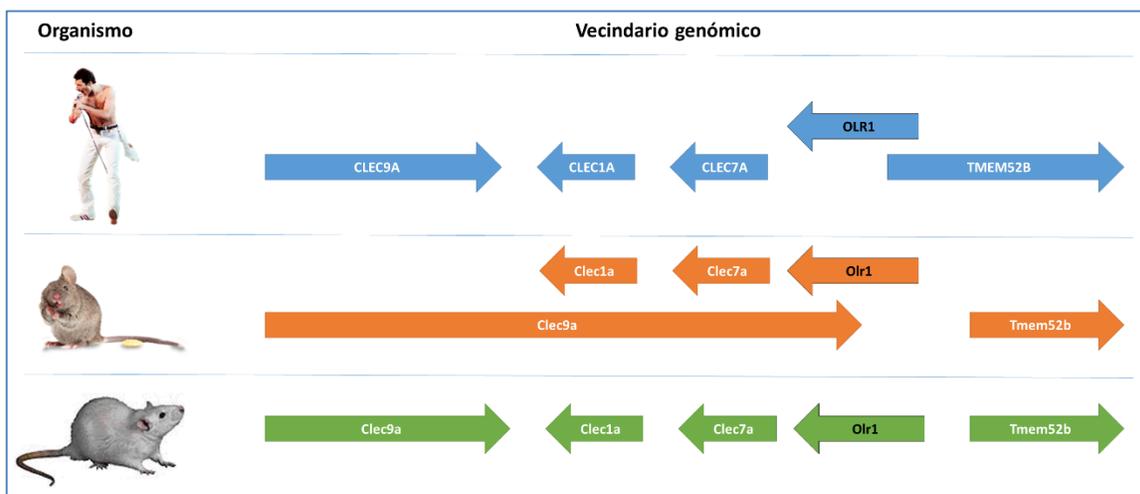


Figura 19. Esquema representativo del vecindario genómico de gen OLR1 en humano, ratón y rata. En orden descendente se indica el esquema del vecindario genómico de humano (azul), ratón (naranja) y rata (verde) respetando la nomenclatura de genes asociada a cada especie. El gen OLR1/Olr1 se enfatiza en letras negras.

Alineamiento de gen OLR1

Los alineamientos de gen OLR1 en humano indicaron secuencias poco similares entre sí, con grandes regiones de delección en las primeras 3000pb y las últimas 2000pb, y continuidad en la región central del gen (**Figura 16 y Archivo vinculado 1 y 2**). La secuencia AJ131757 es la que se distingue menos conservada, también es la primera secuencia en ser anotada en NCBI (**Tabla 14**), por lo que la tecnología y forma de obtención de la anotación pueden ser responsables de estas diferencias. Al hacer un alineamiento descartando AJ131757 se encuentra mayor continuidad entre las secuencias, observándose los mismos patrones de secuencias entre anotaciones AC_000144.1, NC:018923.2 y CM000263.1; y entre NG_016743 y NC_000012.12 (**Figura 16 y Archivo vinculado 2**). Las diferencias en pequeñas regiones del gen entre estos dos grupos pueden explicarse por la diversidad en el origen de las muestras, pudiendo ser polimorfismos característicos de cada muestra.

El alineamiento de OLR1 en ratón (**Figura 17 y Archivo vinculado 3**) indica gran similitud entre las secuencias génicas, exceptuando las regiones cercanas a los 2000pb y 3600pb en

que se ve una deleción en NC_000072.6. Probablemente esta deleción sea una característica de la cepas C57BL/6J utilizada como muestra (**Tabla 3**). Las muestras utilizadas para las anotaciones AC_000028.1 y CM000214.2 son más representativas de la especie *Mus musculus* que NC_000072.6 al incluir distintos tipos de cepas de ratón (**Tabla 5**), por lo que es de esperarse que sean más similares entre sí, como indica el alineamiento (**Archivo vinculado 3**).

En cuanto a rata, el alineamiento es completamente parejo con excepción de la región comprendida en la posición 10845 y 10864 en que en AC_000072.1 no indica preferencia por un nucleótido determinado (expresado como N) y una deleción en la misma anotación entre posiciones 10.865 y 10.875 (**Figura 20** y **Archivo vinculado 4**). Esto puede deberse a errores en la anotación, metodologías distintas, muestras variables y/o polimorfismos de la muestra.

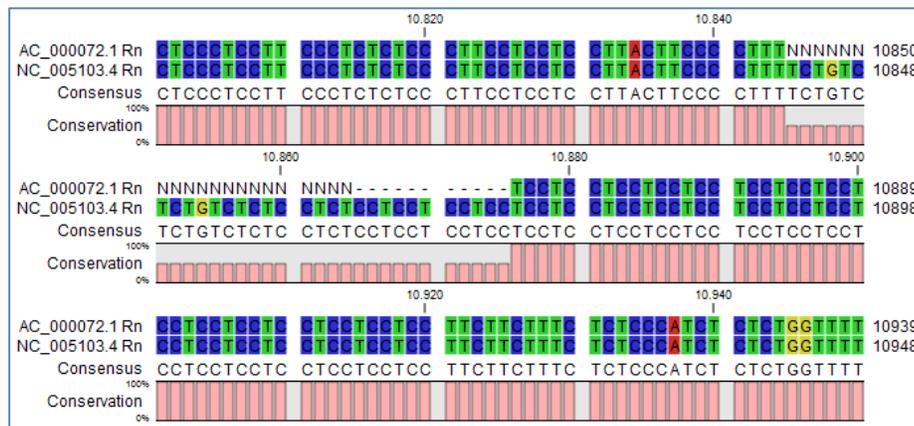


Figura 20. Heterogeneidad de anotaciones de gen OLR1 en rata. Segmento del alineamiento de secuencias del gen OLR1 de rata comprendida entre 10.800 pb y 10.950 pb.

Solapamiento de gen OLR1

De acuerdo con Sanna y cols. (Sanna C., 2008), existen dos tipos principales de solapamiento: 1) Solapamiento en la misma hebra en cual ambos genes involucrados son transcritos a partir de la misma hebra; y 2) Solapamiento de hebras diferentes en el que ambos genes involucrados son transcritos a partir de hebras opuestas. En el caso de OLR1 en humano y ratón se presentan solapamientos de hebras diferentes (**Figura 7, Figura 10** y **Figura 13**). También se conoce este fenómeno como “Transcritos antisentido naturales” (NAT) (Dahary D., 2005). Puede afectar la regulación de expresión génica a nivel de la

transcripción, procesamiento de ARNm, splicing y traducción (Sanna C., 2008), ocurrir en promotores, enhancers, intrones, exones, secuencias intergénicas y regiones del genoma sin traducción (Boldogkoi, 2012) y puede jugar un rol clave en enfermedades humanas (Dahary D., 2005). Este tipo de solapamiento es más común que el solapamiento de misma hebra. Según el estudio de Sanna y cols. aproximadamente el 90% de los genes solapados se presentan en hebras opuestas. En eucariontes es más común presentar solapamientos que en procariontes. Esto podría darse por un efecto de acoplamiento de las distintas estructuras genéticas complejas presentes en organismos eucariontes con genomas más grandes, jugando un rol en la regulación de la expresión de genes (Sanna C., 2008).

De acuerdo a las **Figura 7** a la **Figura 10** OLR1 se encontró solapado con TMEM52B en humano por la región 5' de ambos genes, compartiendo regiones codificantes. Este tipo de solapamiento se puede clasificar como “cola a cola” (Makalowska I., 2005), o divergente (Sanna C., 2008).

En el caso de ratón también se distinguió solapamiento en gen CLEC9A con genes Clec12b, Cleca1A, Clec7A y OLR1 (**Figura 13**). Los genes Clec12b, Cleca1A, Clec7A se encuentran anidados en CLEC9A, mientras que OLR1 se encuentra solapado por la región 3' de ambos genes. Este último tipo de solapamiento se puede clasificar como “cabeza a cabeza” (Makalowska I., 2005) o como convergente (Sanna C., 2008). Sin embargo, de las 3 visiones disponibles del contexto genómico de OLR1 en ratón (**Figura 11, 12 y 13**), solo se muestra el solapamiento en una figura. Por lo tanto, puede tratarse de un error de anotación y/o método. Aunque de las 3 secuencias expuestas es la más reciente según la **Tabla 14**.

De acuerdo con Makalowska y cols. (Makalowska I., 2005) los solapamientos divergentes se presentan en el 30,81% y 36,92% de los casos en humano y ratón, respectivamente. Mientras que los solapamientos convergentes con un 15% y 5,53%, respectivamente. Sanna y cols. tienen una opinión diferente, asignando cerca de un 21% a solapamientos divergentes y 50% a genes solapados convergentes (Sanna C., 2008). Por lo tanto, debido a lo contradictorio de la información, no podemos asignar a los solapamientos de OLR1 en humano y rata un carácter estadístico.

Dahary y cols. (Dahary D., 2005) estudió si efectivamente el solapamiento de genes es un fenómeno real o una “fuga” transcripcional. Sus resultados arrojan que en general el antisentido observado en secuencias que se expresan representa un solapamiento efectivo, más que una falla en el sistema de transcripción, y concluye que en humano cerca del 10% de todos los genes presentan solapamientos, pudiendo imponer una restricción importante en la evolución del orden de los genes y el aumento de la distancia entre genes (Dahary D., 2005). Por su parte, Liang y Landweber (2006) predicen que cerca del 7% de los genes humano contienen exones con splicing alternativo solapado, basándose en la similitud de proteínas conocidas a nivel de secuencias, estructuras secundarias y terciarias, así como también la existencia de motivos conocidos en genes funcionales (Sabath N., 2010).

Sanna y cols. (2008) estudiaron la ganancia o pérdida de solapamientos de distinta hebra en mamíferos encontrando que en humano hubo ganancia de 154 solapamientos, de los cuales 50 ocurren en la línea ancestral del ser humano y macacos; 121 solapamientos ganados en ratón, de los cuales 20 coinciden con la línea ancestrales de ratón y rata; 13 pérdidas de solapamientos en ratón, de los cuales 10 ocurren en la línea ancestral de ratón y rata; y 4 pérdidas de solapamientos en humano, de los cuales 3 ocurren en las líneas ancestrales de humano y macaco (Sanna C., 2008). Desde esta perspectiva, y de acuerdo con las

Figura 5-Figura 10 y Figura 13-Figura 15, el solapamiento presentado en OLR1 corresponde a un fenómeno reciente. Sin embargo, estudios con otros tipos de mamíferos (incluyendo macaco y líneas ancestrales de humano, rata y ratón) son necesarios, así como también mayor cantidad de secuencias genómicas para asegurar tal aseveración y localizar los solapamientos experimentados por OLR1. Aun así, podríamos fallar en tal aseveración si consideramos que los genes ortólogos de OLR1 pueden estar mal anotados en bases de datos.

Se postula que uno de los mecanismos involucrados en el cambio de genes vecinos a solapados puede ser el rearrreglo genómico en el caso de genes que no sean vecinos en determinada especie y aparezcan solapados en otra (Sanna C., 2008). Este no es el caso de

OLR1 dado lo conservado de la vecindad genómica a lo largo de las 3 especies, por lo tanto el mecanismo más probable es un corrimiento de las regiones UTR del gen (Sanna C., 2008). De acuerdo a este mecanismo, en el caso del ser humano el solapamiento con TMEM52B podría explicarse por un corrimiento de la región 5'UTR de TMEM52B, lo cual concuerda con el mayor tamaño de gen TMEM52B en humano con respecto a rata y ratón, según se registra en la **Tabla 7 (Figura 21)**. De la misma forma el solapamiento de OLR1 y CLEC9A en ratón (y los genes anidados en CLEC9A) puede explicarse por un corrimiento de la región 3'UTR de CLEC9A, coincidentemente con el gran aumento en el tamaño del gen CLEC9A respecto a humano y ratón (**Tabla 7 y Figura 22**). De acuerdo a Sanna y cols. la velocidad de transición debido a cambio en la región 3'UTR es de 2,6 pares de genes por millón de años aprox., y debido a cambio en la región 5'UTR es de 1,7 pares de genes por millón de años aprox. Por tanto, la ganancia o pérdida de UTR tienen una velocidad de cambio un poco menor que la velocidad de cambio por mutaciones puntuales (Sanna C., 2008). Otro mecanismo propuesto a través del cual se originarían los solapamientos de genes tiene relación con el proceso de "overprinting", en el cual se originan nuevos genes a partir de secuencias de nucleótidos preexistentes. Según Makalowska y cols., este proceso tiene lugar luego de la divergencia de mamíferos a partir de los pájaros y los genes solapados representarían genes jóvenes y filogenéticamente restringidos que codifican proteínas con diversas funciones, y por tanto especializados en el estilo de vida del organismo en que se encuentran (Makalowska I., 2005). Se deben realizar estudios más detallados secuencia a secuencia para discriminar entre mecanismos, pero el más probable parece ser un corrimiento de las regiones UTR.

OLR1 y TMEM52B: Corrimiento región 5'UTR, nuevo START POINT en TMEM52B Hs - Pérdida intrones en OLR1



Figura 21. Mecanismo de origen de solapamiento OLR1 y TMEM52B en humano. Se indica la especie con una figura representativa en la parte superior, y en la parte inferior se muestra un esquema de la distribución de los genes OLR1 y TMEM52B por especie.

OLR1 y CLEC9A: Corrimiento región 3'UTR, nuevo STOP POINT CLEC9A en Mm y Hs - Pérdida intrones en OLR1

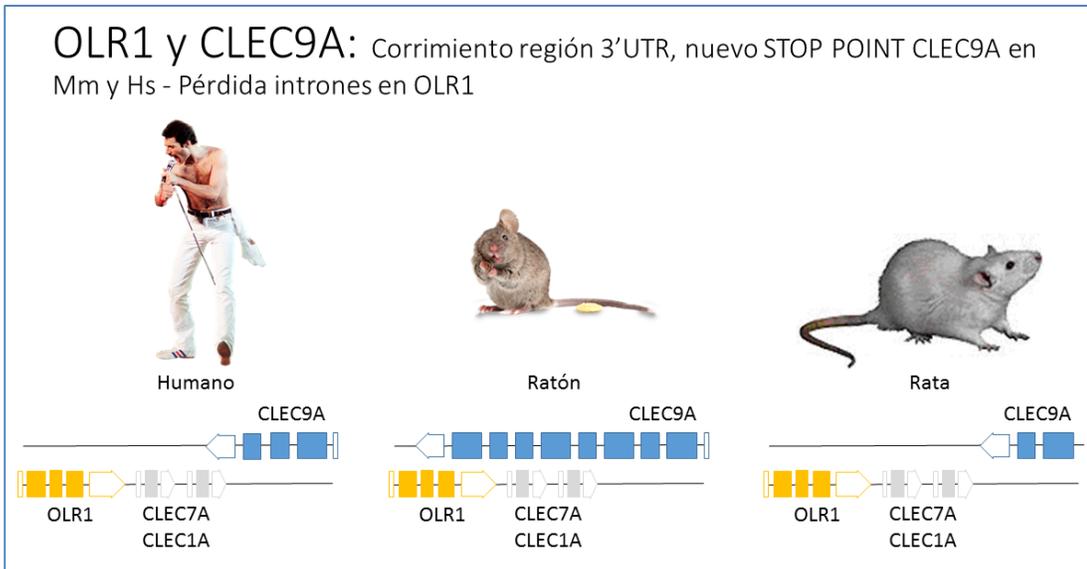


Figura 22. Mecanismo de origen de solapamiento OLR1 y CLEC9A en ratón. Se indica la especie con una figura representativa en la parte superior, y en la parte inferior se muestra un esquema de la distribución de los genes OLR1, CLEC9A, CLEC7A y CLEC1A por especie.

En las especies en que se presentan solapamientos, también se presentan isoformas distintas para las proteínas solapadas (**Tabla 7**). Esto puede ser explicado por interferencias generadas por el solapamiento al momento de transcribir, el procesamiento de ARNm, splicing y traducción (Makalowska I., 2005). Se ha presentado que mientras mayor sea el

largo de secuencias compartidas en el solapamiento en genes solapados de forma *cis* (en hebras opuestas), mayor será el grado de influencia en la interferencia transcripcional de ambos genes (Sanna C., 2008). Desde este punto de vista, y tomando la información contenida en la **Tabla 6**, la interferencia entre OLR1 y CLEC9A (y sus genes anidados) en ratón al momento de la transcripción debe ser mucho mayor que la causada entre OLR1 y TMEM52B en humano.

Makalowska y cols. (2005) asocian la interferencia en la transcripción a secuencias de ARNm complementarias con roles funcionales en la regulación de la misma, pudiendo ejercer control a distintos niveles. En algunos casos el transcrito complementario no codifica un producto proteico pero regula la expresión mediante hibridación con el ARNm del gen regulado. Este último punto, puede ser la explicación para los transcritos de TMEM52B, cuyo producto proteico ha sido encontrado solo por el trabajo de Diez-Roux y cols. (Diez-Roux G, 2011). Además, podría explicar la variedad de transcritos de ARNm detallados en las **Tabla 16-Tabla 18**, y la multiplicidad en los tamaños de los transcritos encontrados en humano y ratón (**Tabla 14**), que varían desde 215pb (AF079165) a 17.734pb (DQ314885) en el primer caso; y desde 1.208pb (XM_006505304) hasta 3.581pb (NM_138648) en el segundo.

La forma en que estos transcritos regulan la expresión de su contraparte solapada, según Makalowska y cols, podría ser vía interferencia transcripcional, a través de la cual mientras uno de los genes solapados se estaría transcribiendo el otro quedaría privado de la transcripción disminuyendo el número de transcritos posibles; enmascaramiento de ARN al formar dupletes entre el transcrito sentido y antisentido, enmascarando las señales regulatorias para ambos transcritos; y mecanismos de ARN de doble hebra (dsARN) y ARN interferente, activando vías dependientes de dsARN incluyendo edición de ARN o silenciamiento por ARN interferente (Makalowska I., 2005). Boldogkoi explica la hipótesis del ruido transcripcional en que las células se beneficiarían de alguna forma al permitir transcripción inespecífica no codificante a niveles bajos, como parte de un circuito complejo de regulación a un estrato oculto, controlando la expresión genética a múltiples niveles. Además, incluye como mecanismos posibles de interferencia a microARN (miARN)

generando modificaciones de la cromatina y/o ocluyendo los sitios de inicio de la transcripción; ARN no codificante de más de 100pb de tamaño (lnc-ARN) sin marco de lectura abierto (ORF), que suprimen la expresión de ARNm de su contraparte solapada; transcritos crípticos inestables (CUTs) originados a partir del promotor bidireccional compartido por los genes solapados (Boldogkoi, 2012) como en el caso de OLR1 y TMEM52B en humano.

Todos los mecanismos de interferencia en la regulación de transcritos de ARNm de OLR1 ayudan a la generación de las isoformas proteicas de LOX1, y pueden tener un rol asociado en el desarrollo de patologías cardiacas.

El fenómeno de solapamiento de genes y su correspondiente implicancia en la regulación de la expresión del gen OLR1 permite cuestionar los modelos de investigación utilizados al estudiar a LOX1. Como se informa en esta memoria en humano el gen OLR1 se encuentra solapado de una forma distinta al solapamiento presente en ratón, y el fenómeno se encuentra ausente en rata. ¿Podemos extrapolar los sucesos relativos a OLR1 y LOX1 ocurridos en humano a un modelo murino con distinta/incierta forma de regulación transcripcional y estructura genética?

5.2. Transcritos de OLR1 en humano, ratón y rata

5.2.1. Sitios de unión a factores de transcripción en el promotor de OLR1

Para la búsqueda de TFBSs en el promotor de OLR1 se utilizó el servidor en línea Lasagna Search . Se utilizaron dos estrategias de búsqueda distintas: 1.- Modelos de matrices, usando matrices disponibles en las base de datos TRANSFAC y JASPER CORE; y 2.- Modelos de alineamientos tipo LASAGNA, usando base de datos TRANSFAC y ORegAnno.

Para el estudio en humano, se usó un modelo de matrices y se identificaron 24 potenciales TFBSs de un total de 146 TFBSs correspondientes a la base de datos JASPAR CORE y 37 TFBSs potenciales identificados de un total de 295 TFBSs a la base de datos TRANSFAC. Al utilizar modelos de alineamiento tipo LASAGNA se identificaron 2 TFBSs potenciales de un total de 10 TFBSs correspondientes a la base de datos ORegAnno y 6 TFBSs potenciales de un total de 63 TFBSs correspondientes a la base de datos TRANSFAC (**Figura 38 en ANEXOS**).

Usando modelo de matrices en ratón, se identificaron 31 potenciales TFBSs de un total de 146 TFBSs correspondientes a la base de datos JASPAR CORE y 21 TFBSs potenciales identificados de un total de 295 TFBSs correspondientes a la base de datos TRANSFAC. Usando modelos de alineamiento tipo LASAGNA se identificó 1 TFBSs potencial de un total de 3 correspondientes a la base de datos ORegAnno y 8 TFBSs potenciales de un total de 53 TFBSs correspondientes a la base de datos TRANSFAC (**Figura 39 en ANEXOS**). En rata se identificaron 21 potenciales TFBSs de un total de 146 TFBSs correspondientes a la base de datos JASPAR CORE y 43 TFBSs potenciales identificados de un total de 295 TFBSs correspondientes a la base de datos TRANSFAC usando modelo de matrices. Al utilizar modelos de alineamiento tipo LASAGNA se identificó 1 TFBSs potencial de un total de 24 TFBSs correspondientes a la base de datos TRANSFAC, mientras que la búsqueda utilizando la base de datos ORegAnno no arrojó resultados (**Figura 40 en ANEXOS**).

La **Tabla 15 en ANEXOS** informa 7 TFs comunes a las tres especies en estudio con sus respectivos TFBSs (AP-1, Bach1, BRCA1, Gata-1, NKX3-1, Prrx2 y SRY [**Figura 37 en**

ANEXOS]]), con su correspondiente información sobre la secuencia de unión de los factores de transcripción, posición en que se encuentra en el promotor de cada especie, base de datos de TFBSs, descripción del TFs asociado y enfermedad relacionada a dicho TFs. Entre los TFs reportados existe una fuerte asociación a patologías oncológicas. También se reconoce gran densidad de TFBSs en la región comprendida entre 930pb y 1000pb del promotor de gen OLR1, considerando como posición 0 al TSS.

5.2.2. Sitios de unión a factores de transcripción asociados a polimorfismos de un solo nucleótido (SNP) en humano

Una vez que se identificó los TFBSs de OLR1 humano, se procede a cruzar con la información de SNP presente en la base de datos NCBI . Para esto, se buscó la posición de cada SNP a lo largo del promotor hipotético de largo 1000pb de OLR1, luego se correlacionó la posición del SNP con la posición del TFBSs en el promotor de OLR1. A partir de los resultados expuestos en la **Tabla 9** podemos deducir que los TFBSs con mayor colocalización con SNP en promotor de OLR1 humano son Hand1 (rs35628316 y rs2742112), Pax-4 (rs16910148 y rs35628316) y p53 (rs555390119 y rs2742112) de un total de 12 SNP cotejados con TFBSs. Todos los TFBSs y SNPs mencionados anteriormente se localizan en la región comprendida entre -338pb y -585pb desde el TSS. El TFBSs común a las 3 especies BRCA1 también presenta colocalización con el SNP rs41401144 en posición -117 del promotor de OLR1.

Tabla 9. Correlación entre TFBSs y SNP en promotor de OLR1 humano

SNP	Posición en el promotor	Tipo	TF	TFBS	Posición en el promotor	Hebra	Modelo
rs550761119	-830	G/T	POU2F2 (Oct-2.1)(T00646)	TGCTCATGAAAAATAG	831	+	TRANSFAC TFBS
rs561726794	-781	C/T	STAT1(MA0137.1)	AGAAACTGAAATTC	782	-	JASPAR
rs182387346	-679	C/G	Cart-1(T03978)	GTCTGCTAATGAAATTAGC	685	+	TRANSFAC TFBS
rs190510437	-649	A/G	NF-Y(M00209)	TCTGATTCGTAAAA	654	+	TRANSFAC
rs16910948	-585	A/G	Pax-4(M00378)	GTGATTCACCC	589	+	TRANSFAC
rs37321185	-433	A/G	FOXM1 (HNF3)(ORegAnno_9606_32)	CCAA	433	+	OREGANNO
rs35628316	-399, -400	-/G	Hand1::Tcfe2a(MA0092.1)	TGTCTGGAAA	394	-	JASPAR
			Pax-4(M00377)	CTAATTTCC	390	+	TRANSFAC
rs555390119	-349	C/T	p53(M00272)	AGACATGTC	340	+	TRANSFAC
rs2742112	-338	C/T	Evi-1(M00081)	TGACATGTCTGGAAA	335	-	TRANSFAC
			Hand1::Tcfe2a(MA0092.1)	TGTCTGGAAA	335	-	JASPAR
			Hand1:E47(M00222)	TGACATGTCTGGAAAT	334	-	TRANSFAC
			TP53(MA0106.1)	AATGACATGTCTGGAAATT	333	-	JASPAR
			TP53(MA0106.1)	AATGACATGTCTGGAAATT A	332	-	JASPAR
			Nr2e3(MA0164.1)	TAAGCTT	263	-	JASPAR
rs140754411	-185	C/T	HNF-4alpha1(M00411)	AGTACAAAGCTTGT	181	+	TRANSFAC
rs41401144	-117	G/T	BRCA1(MA0133.1)	TCAACTG	116	+	JASPAR

El contenido de cada columna de izquierda a derecha es: polimorfismo indicado por su código de anotación; posición del polimorfismo en el promotor, considerando la posición 0 al TSS; Tipo de cambio de nucleótido indicado por la letra símbolo de nucleótido original seguido por el nucleótido cambiado, separados por un / diagonal; factor transcripcional (TF) asociado; secuencia de unión a la cual se une el factor transcripcional (TFBS); posición del TFBS en el promotor indicado como valor absoluto; hebra en la que está presente el TFBS; modelo de búsqueda utilizado en Lasagna Search. A: Adenina; C: Citosina; G: Guanina; T: Timina.

5.2.3. Secuencias de ARN mensajero OLR1

Se encontró información pertinente a las secuencias de ARNm del gen OLR1 para cada especie. En las **Tabla 16** a la **Tabla 20** en **ANEXOS** se muestran los datos asociada a este tipo de secuencias para cada especie.

En las **Tabla 16** a la **Tabla 18** se señala que existe mayor cantidad de transcritos de OLR1 asociados a humano (11), luego ratón (5) y finalmente rata (2), y que existen varios códigos para una misma secuencia dependiendo de su origen.

En la **Tabla 16** se muestran distintas formas para nombrar transcritos de OLR1 de acuerdo al sistema de nomenclatura utilizado y la categoría a la que pertenecen en GenBank, clasificando cada secuencia de transcrito al orden de la especie a la que pertenecen con las siglas PRI (primates) o ROD (rodentia). También se observa que los transcritos NM_002543.3, NM_138648.2 y NM_133306.2 cuentan con versiones nuevas de la anotación, y que los transcritos AF079164.1/AH007081.1, AF079165.1, AF079166.1, AF079167.1 y NM_002543.3 presentan loci diferentes del código de GenBank, sin especificar su significado.

En la **Tabla 17** se reporta información relativa a la definición, fechas y quien originó la secuencia de los transcritos de OLR1. Se establece que todas las secuencias asocian directamente a OLR1 y al receptor LOX1 en su definición. Solo 8 secuencias reportan quien las anotó (Direct submission), y del total de anotaciones de transcritos de OLR1 3 secuencias tienen carácter predictivo basado en las secuencias especificadas en las columna “derivada a partir de” (XM_006719081.1, XM_006505303.1 y XM_006505304.1 en ratón). Se indica que la secuencia de transcrito para OLR1 más antigua data del año 1998 (AF079164.1/AH007081.1, AF079165.1, AF079166.1 y AF079167.1) y la más reciente del 2014 (XM_006505304.1).

La **Tabla 18** da cuenta del tamaño, variante del transcrito, ID de proteína correspondiente a LOX1, especificaciones de la muestra y comentarios sobre la anotación de transcrito de OLR1. Los resultados muestran que 9 anotaciones no reportan asociación con alguna variante proteica LOX1 en particular, 1 anotación no reporta correlación alguna

con una versión proteica (DQ314885.1) y solo 7 reportan especificaciones de la muestra que originó la secuenciación del transcrito de OLR1. Se observa una diferencia entre los tamaños de transcritos de OLR1 en humano: el transcrito de menor tamaño alcanza las 215pb (AF079165.1) mientras que el de mayor largo alcanza 17.734pb (DQ314885.1). En ratón sucede lo mismo, pero de forma menos drástica: el transcrito de OLR1 con menor tamaño tiene 1208pb y el de mayor tamaño 3581pb. En rata solo hay información relativa a dos transcritos para OLR1 de 3.751pb (NM_133306.2) y 2.985pb (BC097290.1), respectivamente.

Existe gran y variado número de transcritos para OLR1. 10 de un total de 18 anotaciones de secuencias de ARNm encontrados en bases de datos ofrecen información sobre los exones (**Tabla 19**), y solo en humano se reporta información incompleta sobre la región regulatoria, 5'UTR, 3'UTR y sitio de poliadenilación de OLR1 (**Tabla 20**).

5.2.4. Sitios de poliadenilación (PAS) en humano

Reportes recientes indican que en genes humanos ocurre un fenómeno conocido como poliadenilación alternativa (APA) en que existen sitios variados de poliadenilación (PAS) con señalización y sitios de corte de la secuencia de ARNm . El uso de PAS representan una regulación adicional durante la expresión de genes, resultando en formación de transcritos con diferentes extremos 3'. Éstos pueden diferir en su región codificante o en el largo de 3'UTR, afectando la regulación de la expresión génica en la localización de transcritos, su estabilidad, eficiencia en la traducción y naturaleza de la proteína codificada. Los perfiles de APA son específicos para cada tejido y parecen ser fuertemente regulados durante el desarrollo y diferenciación celular. Para dilucidar si OLR1 contiene secuencias PAS que unen a CPSF ("*cleavage polyadenylation specificity factor*"), las secuencias del gen en humano se sometieron a escrutinio utilizando el servidor en línea DNAFSMiner. Los resultados se muestran a continuación.

De acuerdo a los resultados expuestos en la **Tabla 10**, las secuencias AC_00144.1, CM_000263.1 y NC_018923.2 comparten los mismos PAS de secuencias AATAAA (3 sitios) y ATATAA (2 sitios) con puntajes similares e igual cantidad totales de PAS encontrados en la secuencia del gen OLR1. En el caso de PAS de secuencia AATAA se distinguió 1 sitio de señalización proximal (cercana al 5'UTR) y dos sitios ubicados de forma distal (lejano de 5'UTR) en el gen OLR1, mientras que ambos PAS de secuencia ATATAA se ubican de forma proximal.

Las secuencias de gen OLR1 NC_000012.12 y NG_016743.1 también presentaron similar ubicación de sitios de poliadenilación en su estructura: 2 de un total de 16 PAS con secuencia AATAAA ubicados en la misma posición de forma proximal y distal, e igual número de PAS de secuencia ATATAA ubicados de forma proximal en el gen OLR1.

En el caso de secuencia AJ1317571.1, se reconoció un PAS distal de secuencia AATAAA y otro proximal de secuencia ATATAA.

La búsqueda en bases de datos APASdb arrojó 3 APAs y 4 PAS ubicados en la región 3'UTR de OLR1 en humano sin asociaciones a tejido coronario, 4 transcritos de distinto tamaño y combinación de exones, sitios de corte heterogéneos por cada APA, y preferencia por la señal de poliadenilación AATAAA ubica de forma distal en posición 10310904 en el cromosoma 12 humano al realizar lecturas de los transcritos obtenidos de 22 tejidos humano (**Figura 23**, **Figura 24** y **Figura 25**). Los tejidos con mayor cantidad de lecturas de transcritos de OLR1 son placenta (148 lecturas), pulmón (68 lecturas) y riñón (19 lecturas) .

Tabla 10. Sitios de señalización de poliadenilación hipotéticos en humano obtenidos con DNAFSMiner.

Secuencia gen OLR1	No.of AATAAA(s) from the 5'end	Score	Position (bp)	Total AATAAA(s) in the query sequence	No.of ATTAAA(s) from the 5'end	Score	Position (bp)	Total ATTAAA(s) in the query sequence
AC_00144.1	9	0.872	4682	23	9	0.692	10029	9
	23	0.754	13013		1	0.605	3502	
	21	0.691	10073					
AJ131757.1	11	0.957	7774	18	1	0.659	3881	9
CM000263.1	9	0.889	4688	23	9	0.692	10034	9
	23	0.754	13018		1	0.604	3500	
	21	0.691	10078					
NC_000012.12	7	0.945	4699	16	2	0.797	2609	10
	15	0.666	13472		1	0.659	806	
NC_018923.2	9	0.872	4676	23	9	0.692	10022	9
	23	0.754	13006		1	0.605	3496	
	21	0.691	10066		3	0.359	6005	
NG_016743.1	7	0.945	4699	16	2	0.797	2609	10
	15	0.666	13472		1	0.659	806	

Contenido de cada columna de izquierda a derecha: código y/o nombre de la secuencia gen OLR1 procurado; número de secuencia AATAAA reconocida a partir del extremo 5'; score, o valoración del sitio de señalización de poliadenilación entregado por DNAFSMiner (Liu H., 2005); posición del sitio de señalización de poliadenilación en la secuencia procurada; número de sitios totales de secuencia AATAAA encontrados en la secuencia procurada; número de secuencia ATTAAA reconocida a partir del extremo 5'; score, o valoración del sitio de señalización de poliadenilación entregado por DNAFSMiner (Liu H., 2005); posición del sitio de señalización de poliadenilación en la secuencia procurada; número de sitios totales de secuencia ATTAAA encontrados en la secuencia procurada.

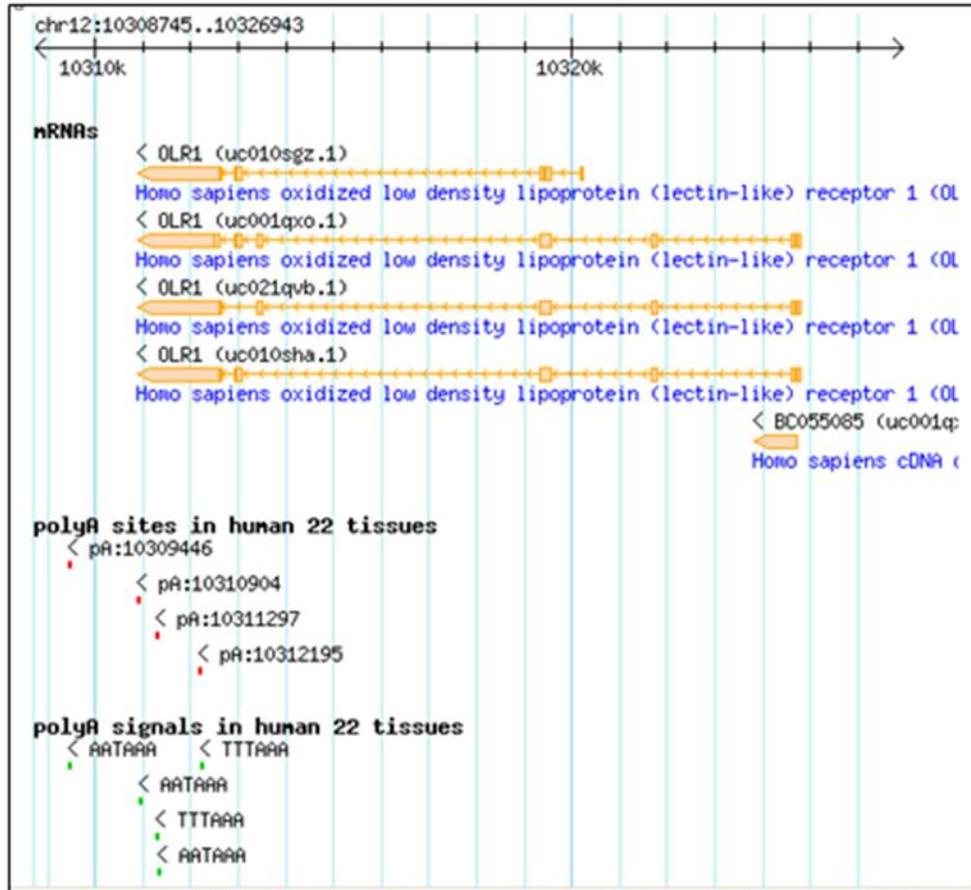


Figura 23. Mapa de sitios poly(A) de gen OLR1 exhibido en APASdb. En orden descendente: se muestra la región del cromosoma 12 humano en que se ubica gen OLR1; variedad de transcritos de ARNm correspondientes a OLR1 con diferentes largos y exones; ubicación de APAs identificados en 22 tejidos humanos; y ubicación de señales de poli(A) (PAS) identificados en 22 tejidos humanos.

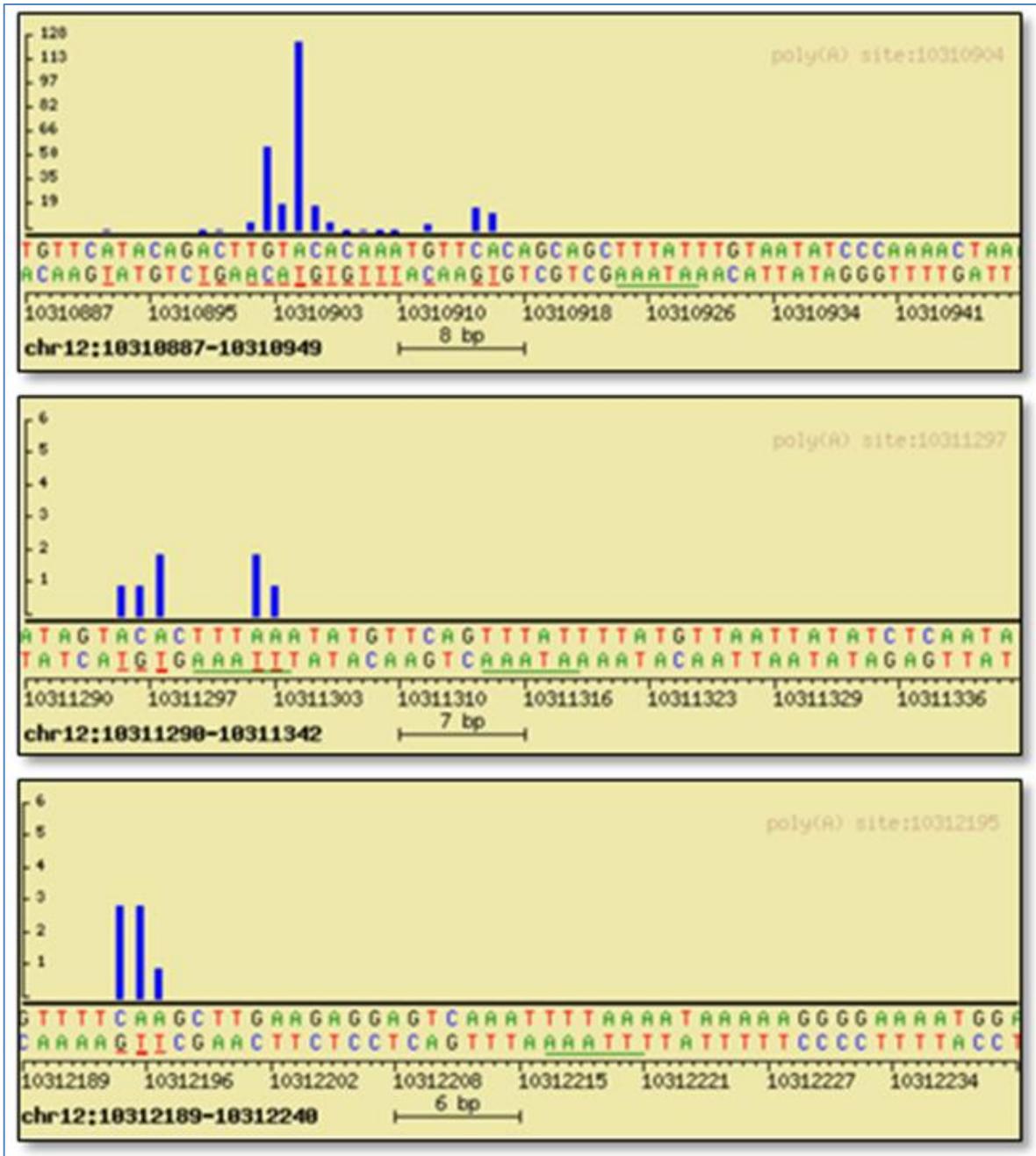


Figura 24. Sitios heterogéneos de corte según sitio de poli(A) gen OLR1. Desde recuadro superior a inferior se encuentran los sitios poly(A) (APAs) en posiciones 10310904, 10311297 y 10312195 del cromosoma 12 humano. En el eje X se muestran las secuencias de gen OLR1, se indican las bases nitrogenadas sitios de corte subrayados en rojo, indicando el sitio corte más frecuente subrayado en negrita, y las secuencias señal de corte subrayadas en verde; en el eje Y se indica el número de lecturas normalizadas. Información obtenida desde APASdb.

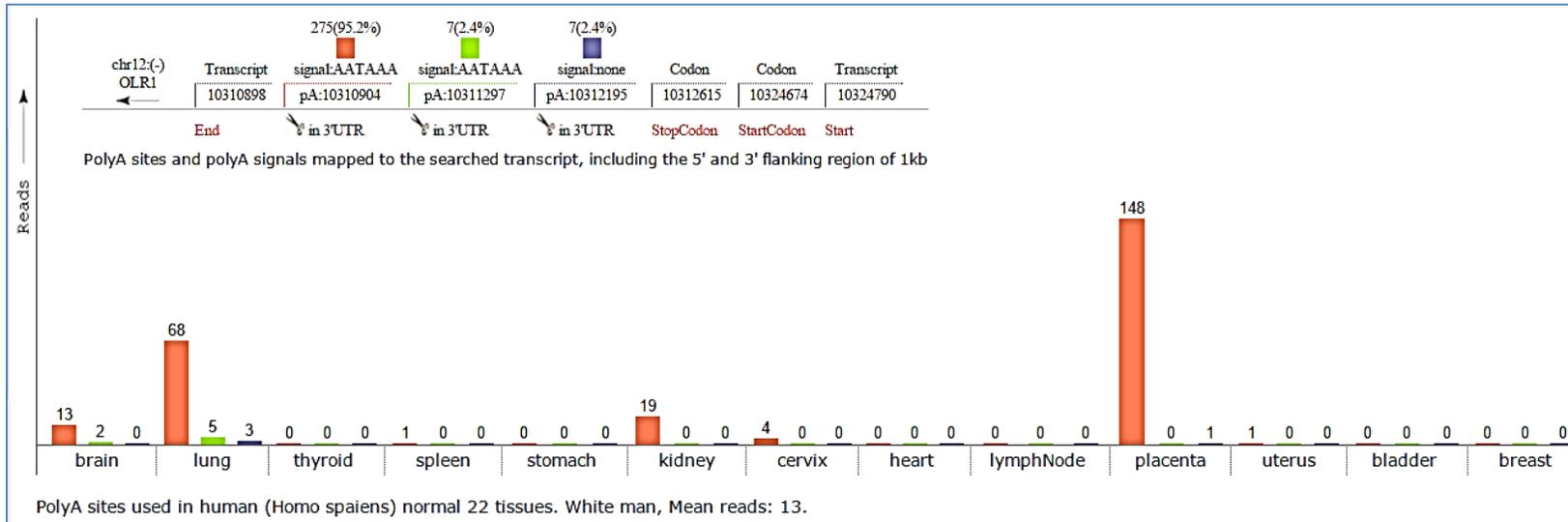


Figura 25. Lecturas de APAs de gen OLR1 según tejido. En la parte superior de la figura se indican las posiciones de inicio, codón de inicio, codón de término, APAs 10312195, 10311297 y 10310904, y el fin del gen OLR1 en el cromosoma 12 humano; además se indican las secuencias de cada APA y el número y porcentaje de lecturas encontrados a lo largo de los 22 tejidos en humano. En la parte inferior del panel (C) se detallan las lecturas de cada APA por tejido investigado, de forma parcial. Información obtenida desde APASdb.

5.2.5. Discusión sección regulación de la transcripción gen OLR1

Factores de Transcripción asociados al promotor de OLR1

La forma más estudiada de regular la expresión de genes es a través de factores transcripcionales (TFs) que se unen al promotor. En este estudio, se consideró como promotor de gen OLR1 al segmento de secuencias de ADN ubicada 1000pb río arriba del punto de inicio de la transcripción del gen en estudio. Se realizaron estudios de secuencias de unión a factores de transcripción utilizando el servidor en línea Lasagna Search, y los algoritmos detallados en la sección 2.3. De acuerdo a la **Tabla 15 en ANEXOS**, todos los TFs que podrían regular la expresión de OLR1 están asociados a cáncer, con excepción de SRY. No es de extrañar este hecho, pues OLR1 y LOX1 han sido asociados a cáncer (Hirsch A., 2010) (González-Chavarría I., 2014). Sin embargo, al extrapolar estos hallazgos a tejido coronario con bajos niveles de proliferación y frecuencia tumoral, debemos preguntarnos si efectivamente los sitios de unión a TFs se encuentran activos o no, o si los TFs encontrados en esta memoria han sido estudiados en el contexto del sistema cardiovascular.

Mediante la búsqueda en la literatura asociada a sistema cardiovascular de los TFs detallados en la **Tabla 15** comunes para humano, ratón y rata, se encontró el estudio realizado por Aneja y cols. en el año 2004. Este describe a AP-1 como un factor de la transcripción dimérico compuesto por las subunidades Jun y Fos. También describen un aumento del AP-1 y NF- κ B en área infartadas del corazón de rata por inducción de isquemia del miocardio y reperusión de la zona lesionada, asociando un efecto sinérgico de ambos factores de transcripción frente al daño (Aneja R., 2004). Es probable que la activación de la transcripción de OLR1 mediada por AP-1 en fibroblasto cardiaco sea activada mediante la vía no canónica gatillada por TGF- β 1, por activación de JNK (MAPK) que fosforila c-JUN, activando AP-1 (Aneja R., 2004). A su vez, la vía no canónica de TGF- β 1 puede activarse por especies reactivas de oxígeno (ROS) que se producen típicamente en patologías cardiacas, como también por ROS generado a partir de la activación de la NADPH oxidasa gatillado por la unión de oxLDL y su receptor LOX1, potenciando en un loop positivo la expresión de LOX1

y estimulando la fibrosis por vía canónica y no canónica (Taye A, 2013). Dichos antecedentes tienen correlación con los datos obtenidos en nuestro laboratorio respecto de ROS, TGF- β 1 y LOX1.

Es muy probable que la activación de la transcripción vía AP-1 sea efectiva, debido a que es el TF con mayor número de sitios de unión en el promotor de OLR1. Se proyecta realizar estudios en nuestro laboratorio para confirmar esta hipotética vía de señalización.

Por su parte, Bach1 ha sido identificado como un represor de la transcripción de genes, descrito principalmente como represor de la transcripción de la enzima Heme Oxigenasa 1, encargada de degradar Heme Prooxidante a monóxido de carbono y biliverdina/bilirrubina, ambos con efectos antiinflamatorios y antioxidantes en el ventrículo izquierdo del corazón (Mito S., 2008). En condiciones de estrés oxidativo Bach1 pierde la capacidad de unir ADN, dejando libre sitios de unión que activan la transcripción de genes, generando un efecto benéfico en el corazón isquémico. Mito y cols. Demostraron que en ratones KnockOut de Bach1 tanto la fibrosis intersticial como la hipertrofia de ventrículo derecho inducida por constricción de aorta disminuyen (Mito S., 2008). Si efectivamente Bach1 se une al promotor de OLR1 se explicaría porqué la expresión basal de OLR1 en tejido cardiaco es menor que en condiciones de estrés oxidativo (Hu C, 2008), pues Bach1 actuaría como represor de la transcripción de OLR1 y al generarse ROS en condiciones patológicas cardiacas se liberaría del promotor, permitiendo la transcripción y traducción a LOX1.

BRCA1 es un regulador del ciclo celular y de vías de reparación de ADN dañado (Li H., 2009). De acuerdo a lo estudios realizados por Li y cols (Li H., 2009) la pérdida o disminución de BRCA1 inhibe la vía de señalización de TGF- β 1 vía Smad3 durante estrés oxidativo en cáncer. Como se mencionó anteriormente, la fibrosis característica de eventos patológicos cardiacos es gatillada por TGF- β 1 a través de sus vías canónica y no canónica. En la vía canónica tenemos participación de Smad 2 y Smad 3, que potencian la producción de α -SMA, proteínas de adhesión y colágeno (Taye A, 2013). Ambas vías pueden activarse por efecto de ROS, de acuerdo a los resultados obtenidos en nuestro laboratorio. Por lo tanto, extrapolando al contexto de patologías cardiacas y LOX1, considerando el antecedente publicado por Li y cols., es probable que BRCA1 participe en las vías de señalización

mediadas por TGF- β 1 en respuesta a estrés oxidativo condicionado por patologías cardíacas, afectando la transcripción de OLR1. Dicha extrapolación es posible porque se ha reportado con anterioridad vías comunes de transcripción y señalización de transcripción en cáncer y metabolismo de lípidos (Hirsch A., 2010), actuando LOX1 en ambas patologías.

La búsqueda en literatura de GATA1 asociado a patologías cardíacas no dio resultados. Este TF es uno de los 6 miembros perteneciente a la familia de TFs GATA, está asociado a la regulación de líneas celulares sanguíneas en formación, y su mutación provoca desórdenes sanguíneos caracterizados por citopenias (Ciovacco W., 2008). Los miembros GATA4, GATA5 y GATA6 están presentes en el corazón en desarrollo (embrión y posnatal) (Charron F., 1999). Es probable que la activación de OLR1 se dé mediante este TF en la diferenciación de fibroblasto a miofibroblasto, gatillada por TGF- β 1, a través de un efecto directo o indirecto mediado por cofactores inducibles, que jueguen un rol en la regulación tejido específico en respuesta a estímulos extracelulares (Charron F., 1999). Sin embargo, no se encontró en literatura asociaciones entre GATA1 y TGF- β 1 o ROS. Experimentación *in situ* es necesaria para saber si GATA1 se une al promotor de OLR1, y si activa la expresión de éste en patologías cardíacas.

Evidencia reciente sugiere que las redes de regulación de genes están alteradas bajo condiciones de estrés en el corazón adulto (Mikhailov A., 2012). Los coactivadores transcripcionales también participan de los circuitos reguladores de genes como objetivos primarios de señales fisiológicas y patológicas (Mikhailov A., 2012). Una de las familias de estos cofactores transcripcionales son las miocardinas (MYOCD), y específicamente los miembros MRTF-A y MRTF-B activan la expresión de genes involucrados en crecimiento cardíaco y adaptación al estrés a través de distintos mecanismos (Mikhailov A., 2012). Éstas últimas se encuentran aumentadas en remodelamiento de corazón. Específicamente, al forzar la expresión de MRTF-A se activan genes profibróticos y la diferenciación de fibroblasto a miofibroblasto. Las MYOCD presentan un sitio de reconocimiento a NKX1-3 (Mikhailov A., 2012), por lo tanto el remodelamiento que ocurre posterior a isquemia de corazón podría estar sujeto a la acción de NKX1-3 y su cofactor MRTF-A. Asimismo, los TFBSs en promotores de genes que se encuentran sobre activados en etapas tardías de

aterosclerosis comparten motivos ricos en Myf, MZF1, FOXL1 y NKX3-1, elucidando un posible rol esencial en la aterosclerosis de carótida considerada como patología cardíaca (Lin M., 2014).

Chen y VanBuren (2014) identificaron a Prrx2 como uno de los factores transcripcionales participantes en las fases tempranas y tardías del desarrollo del corazón, junto con otros TFs. Su equipo propone que los distintos TFs trabajan juntos a lo largo de las distintas fases de desarrollo del tejido coronario, de forma fase específica para regular su expresión y mediar las transiciones entre fases durante el normal desarrollo cardíaco (Chen H., 2014). Este antecedente nos indica que Prrx2 se expresa en el corazón durante su desarrollo. No se encontró en literatura asociación de Prrx2 a fenómenos de fibrosis o vías de señalización sujetas a TGF- β 1. Por lo que no podemos hacer especulaciones sobre su transcripción, o si regula la transcripción de OLR1 en la diferenciación de fibroblasto a miofibroblasto durante fibrosis en patologías cardíacas.

El factor transcripcional SRY es el único que no está asociado a cáncer. Éste se ubica en el cromosoma Y, por lo que es específico para individuos de sexo masculino, y determina el desarrollo de los testículos y la subsecuente secreción de hormonas sexuales. Además, influencia la expresión de otros genes en cromosomas autosomales (Miller V. K. J., 2011). Las hormonas sexuales influyen directamente en el desarrollo celular en periodos críticos, que también pueden ser gatillados por factores ambientales (Miller V. K. J., 2011). Variantes genéticas de SRY predisponen a hipertensión en rata, y en humano se ha descrito que los padres hipertensos, y no las madres, determinan la hipertensión en su progenie (Kostrzewa G., 2013). Además, SRY afecta la expresión de tirosina hidroxilasa, enzima que convierte tirosina en dihidrofenilalanina (DOPA) requerida para la síntesis del neurotransmisor adrenérgico norepinefrina que regula la función cardíaca (Miller V. B. P., 2011). De acuerdo a esto, SRY tiene un rol en patologías cardíacas pero se desconoce el mecanismo molecular que podría estar involucrado en patologías cardíacas agudas como infarto agudo al miocardio, isquemia, etc. Por lo tanto, se desconoce la forma en que podría activar la diferenciación del fibroblasto en patologías cardíacas.

Resumiendo, se propone que los potenciales TFs que activarían la transcripción de OLR1 son: AP-1, mediante la vía no canónica de TGF- β 1 estimulada por ROS; BRCA1, mediante la vía canónica de TGF- β 1 activado por Smad3; NKX1-3, uniéndose al co factor MRTF-2, al activar la vía de TGF- β 1. Por otra parte, BACH1 reprimiría la expresión de OLR1 al unirse al promotor, pero liberándose en condiciones de ROS. Es necesaria investigación *in situ* para evaluar la unión efectiva de estos TFs a sus TFBSs, y si regulan la expresión de OLR1 (Figura 26).

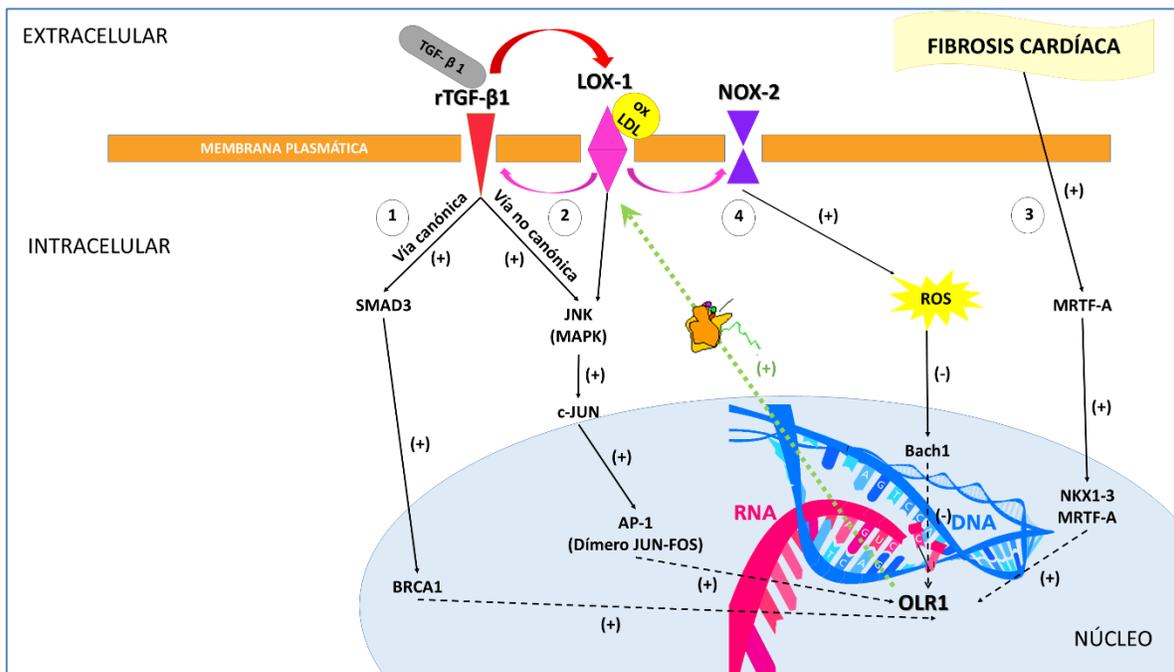


Figura 26. Esquema de la regulación de la transcripción de gen OLR1 a receptor LOX1 vía factores transcripcionales propuestos. 1) Activación de la transcripción de gen OLR1 a través de factor transcripcional BRCA1 activado por SMAD3 en la vía canónica de señalización de TGF- β 1. 2) Activación de la transcripción de gen OLR1 a través de factor transcripcional AP-1 activado por fosforilación por quinasa JNK de la subunidad c-JUN de AP-1, en la vía no canónica de señalización de TGF- β 1. 3) Activación de la transcripción de gen OLR1 a través de factor transcripcional NKX1-3 activado por cofactor MRTF-A en condición de fibrosis cardiaca. 4) Activación de la transcripción de gen OLR1 mediante liberación del factor transcripcional inhibitorio Bach1 en presencia de ROS.

TFBSs y SNP en humano

De acuerdo a la **Tabla 9** los TFBSs afectados por SNP con mayor frecuencia son Hand1, p53 y Pax-4.

Hand1 se encuentra disminuido en pacientes con cardiomiopatías como también en roedores con hipertrofia cardiaca, sugiriendo una correlación entre los niveles de Hand1 y el remodelamiento coronario (Lu S., 2011). Lo anterior sugiere que en pacientes que presenten SNPs rs35628316 y rs2742112 la unión de Hand1 al promotor de OLR1 se vería afectada, repercutiendo en su transcripción.

En el caso de p53 se ha visto que en fibroblastos hipóxicos, o extraídos desde corazón infartado, hay un aumento de la expresión de este TF, disminución del nivel de senescencia y citoquinas en el fibroblasto, y aumento de la expresión de colágeno (Zhu F., 2013). Por lo tanto, p53 probablemente sea un regulador de la transcripción de OLR1, y los SNP rs555390119 y rs2742112 afecten la unión de p53 al promotor de OLR1. Asimismo, el SNP rs41401144 que afecta al TFBS de BRCA1 (asociado a cáncer) debe afectar en algún nivel la expresión de OLR1.

Todos los SNPs mencionados podrían estar afectando el proceso de fibrosis característico de remodelamiento coronario, influyendo en la traducción de LOX1 (Hu C, 2008).

En el caso de Pax-4 no se encontró literatura asociada a patologías cardiacas o a su expresión en corazón, está estrechamente relacionado con patologías pancreáticas, aunque la familia de TFs Pax está asociados a organogénesis, proliferación, diferenciación celular y cáncer (Mansouri A., 1996).

Para confirmar si los SNP mencionados afectan realmente la transcripción de OLR1, y por tanto la fibrosis coronaria, se debe investigar estadísticamente en pacientes con dichos SNP y en el laboratorio comprobar el aumento o disminución de la afinidad de los distintos TFs al promotor de OLR1.

PAS en OLR1

Otro fenómeno que podría ser responsable de la generación de isoformas proteicas de LOX1 es APA. Los resultados expuestos en esta memoria de título indican que gen OLR1 presenta APA generando transcritos variables de acuerdo a la **Figura 23** y la gran variedad de secuencias de ARNm expuestos en las **Tabla 16** a la **Tabla 20** en **ANEXOS** podrían explicarse de la misma forma. Un fenómeno similar fue descrito al analizar el transcriptoma de 15 líneas celulares de humano, en que aparecieron 25 transcritos diferentes producidos por 1 solo gen, y 12 de esos transcritos producidos en una línea celular en particular (de Klerk E., 2015).

Lamentablemente aún faltan estudios, antecedentes en bases de datos y algoritmos ajustados a ratón y rata para comparar si existe PAS en regiones semejantes de gen OLR1 entre especies. En general, se sabe que APA está a lo largo de todos los eucariontes examinados hasta la fecha y está presente en ~70-75% de los genes en humano, considerándose como un mecanismo regulatorio extendido capaz de controlar la expresión de genes y expandir la diversidad proteica que incluso afecta a ARNm no codificantes de gran tamaño (lncARN) (Elkon R., 2013). Una posibilidad es que la secreción de la forma soluble de proteína LOX1 se genere por APA, pues se ha encontrado este mecanismo actuando en proteínas de membrana plasmática con isoformas secretadas en forma soluble (como inmunoglobulina M, IgM) (Elkon R., 2013).

Estudios han revelado una correlación inversa entre el largo de 3'UTR y los niveles de expresión de proteínas: algunos tejidos humanos están enriquecidos en transcritos altamente abundante con 3'UTR cortos (cerebro, testículos, pulmón, mamas), mientras que otros contienen muchos transcritos de baja abundancia con 3'UTR largos (corazón y músculo esquelético) (de Klerk E., 2015). En condiciones de hipertrofia cardiaca se encuentran potenciado el corte en sitios poly(A) proximales de forma global en humano (Elkon R., 2013). Si existe preferencia de transcritos largos de OLR1, con predilección en corte y poliadenilación de forma distal, o viceversa, en tejido coronario en condiciones saludables y/o patológicas es un misterio a resolver.

5.3. Proteína LOX1 en humano, ratón y rata

5.3.1. Secuencias de proteínas LOX1

La búsqueda de secuencias de OLR1 a través de CLC Sequence Viewer 7.5 arrojó gran cantidad de secuencias de tipo genómicas, ARNm y proteicas. La información sobre secuencias proteicas se agrupó en las tablas a continuación.

Los resultados expresados en la **Tabla 21** indican un total de 24 secuencias proteicas descritas para LOX1, de las cuales 14 pertenecen a humano, 7 a ratón y 3 a rata, categorizadas en la división de GenBank como PRI para humano y ROD en el caso de ratón y rata. En total son 9 las secuencias proteicas predictivas de LOX1 derivadas de análisis computacional (indicadas con un asterisco). Existe 1 secuencia de LOX1 por especie que difiere en el locus especificado y su código de GenBank (P78380.1 humano, Q9EQ09.2 en ratón y O70156 en rata).

En la definición de cada anotación de las secuencias proteicas de LOX1 se hace referencia al receptor de lipoproteínas oxidadas y/o LOX1, pero no a gen OLR1 (**Tabla 22**). De las 24 anotaciones indicadas en la **Tabla 22**, 12 manifiestan la solicitud de anotación directa, aportando información sobre quién secuenció la proteína. Se registran secuencias derivadas a partir de otras: 5 en total, 1 en humano (XP_006719144.1), 3 en ratón (NP_619589.2, NP_001288025.1 y NP_001288023.1) y 1 en rata (Q9EQ09.2). De acuerdo a las fechas de actualización de la anotación en NCBI, todas las secuencias proteicas de LOX1 publicadas parecen ser revisadas y/o actualizadas de forma regular. La secuencia de LOX1 publicada con mayor antigüedad data del año 1998 en humano (AAC97927.1), 2001 en ratón (NP_619589.2) y 1998 en rata (O70156). Mientras que las más actuales corresponden al año de publicación 2010 para humano (NP_001166104.1), 2014 para ratón (NP_001288025.1 y NP_001288023.1) y 2002 para rata (NP_579840.2). Del total de anotaciones indicadas, P78380.1 es la que ha sido citada con mayor frecuencia en literatura especializada al ser incluida en un total de 28 referencias de publicaciones científicas a la fecha.

A partir de la **Tabla 23** se puede deducir que las anotaciones predictivas de LOX1 EAW96159.1 y EAW96156.1 corresponden a la misma secuencia, pues presentan la misma información y corresponden a la misma isoforma obtenida mediante análisis computacional. Las secuencias correspondientes a la isoforma 1 en humano son NP_002534, BAC81565, CAB38175, AAH22295, P78380 y secuencias predictivas AAC97927, ABC40744, EAW96156, EAW96159 ya que todas tienen 273aa de tamaño de secuencia proteica. NP_001166103 y la secuencia predictiva EAW96157 también tienen igual tamaño, por lo tanto se deduce que la isoforma 2 y la isoforma CRA_b corresponden a la misma secuencia de LOX1. Lo mismo se observa para las secuencias NP_001166104 y EAW96158, derivando que la isoforma 3 y la isoforma CRA_c de LOX1 corresponden a la misma secuencia obtenida por métodos empíricos y computacionales. La isoforma X1 (XP_006719144.1) presenta tamaño de secuencia aminoacídica única (**Figura 41** en **ANEXOS**).

A partir de los resultados expuestos en la **Tabla 23** se deduce que en ratón la isoforma 1 de LOX1 puede asignarse a secuencias NP_619589, Q9EQ09 y secuencia predictiva EDK99928 de acuerdo al tamaño idéntico de sus secuencias aminoacídicas (**Figura 42** en **ANEXOS**). Las isoformas 2 (NP_001288025), 3 (NP_001288023), X1 (XP_006505366) y X2 (XP_006505367) no presentan tamaños iguales con otras secuencias, por lo tanto se deduce que existe una única secuencia por cada una de estas isoformas de LOX1. En rata no existen isoformas ni secuencias predictivas asociadas a LOX1: los tamaños de las 3 secuencias descritas son iguales, por lo tanto se deduce que corresponden a la isoforma única de LOX1 en rata (**Figura 43** en **ANEXOS**).

De un total de 24 anotaciones presentadas en la **Tabla 23** solo 13 reportan información sobre a qué isoforma de LOX1 corresponden, específicamente 8 anotaciones de humano, 5 en ratón y ninguna reporta dicha información en rata. Solo 5 anotaciones proteicas de humano se encuentran asociadas a algunas secuencias de ADN en particular, mientras que 15 se encuentran asociadas a secuencias de ARNm (8 de humano, 5 en ratón y 2 en rata). En cuanto a la forma de secuenciación de las anotaciones proteicas de LOX1 solo 4 secuencias de humano indican el tejido desde el cual se tomó la muestra para la

secuenciación. Comentarios sobre la anotación son detallados en 10 secuencias de humano, 6 en ratón y 2 en rata.

5.3.2. Propiedades fisicoquímicas de las isoformas de LOX1

En la **Tabla 11** se especifica la masa molecular, punto isoeléctrico e índice alifático de cada isoforma proteica de LOX1 por especie, construida a partir de la información entregada en el informe estadístico comparativo de secuencias proteicas de LOX1 obtenido de CLC Sequence Viewer 7.5. Se establece que las secuencias NP_2534, P78380, AAH22295, CAB3817, BAC81565 y secuencias predictivas EAW96156, EAW96159, ABC40744 y AAC97927 tienen los mismos valores de masa, punto isoeléctrico e índice alifático, deduciéndose que corresponden a la isoforma 1 humana de LOX1. Lo mismo sucede con secuencias NP_001166103 y EAW96157, derivando que se trata de la isoforma 2 humana de LOX1. Siguiendo el mismo razonamiento, también se infiere que las secuencias NP_001166104 y EAW96158 corresponden a isoforma 3 humana de LOX1.

En ratón los resultados expresados en la **Tabla 11** indican que las anotaciones NP_619589 y Q9EQ09 corresponden a isoforma 1 de LOX1, debido a que comparten iguales valores en sus propiedades fisicoquímicas. En rata no se observan grandes diferencias respecto de las propiedades fisicoquímicas entre las distintas anotaciones, por lo que se deduce que todas corresponden a la misma isoforma codificada por OI1.

A pesar de tener iguales tamaños entre distintas anotaciones es posible que los aminoácidos que componen la secuencia aminoacídica de las distintas versiones de LOX1 difieran entre anotaciones, otorgando propiedades fisicoquímicas y estructuras proteicas distintas. En las **Figura 44** a la **Figura 46** en **ANEXOS** se presentan las frecuencias de aminoácidos de las distintas isoformas para humano, ratón y rata, observándose diferencias menores en la composición aminoacídica de cada isoforma de LOX1 por especie.

Tabla 11. Propiedades Físicoquímicas isoformas proteicas de LOX1 en humano, ratón y rata.

GenBank	Masa (kDa)	Punto Isoeléctrico (pH)	Índice alifático	Isoforma
<i>Homo sapiens</i>				
NP_002534	30.959	7.54	81.538	1
P78380	30.959	7.54	81.538	No especificada
AAH22295	30.959	7.54	81.538	No especificada
CAB38175	30.959	7.54	81.538	No especificada
BAC81565	30.945	6.86	81.538	No especificada
NP_001166103	20.227	5.63	91.657	2
NP_001166104	21.424	5.85	87.778	3
XP_006719144	24.372	5.42	88.157	X1
EAW96156	30.959	7.54	81.538	CRA_a
EAW96159	30.959	7.54	81.538	CRA_a
EAW96157	20.227	5.63	91.657	CRA_b
EAW96158	21.424	5.85	87.778	CRA_c
ABC40744	30.959	7.54	81.538	No especificada
AAC97927	30.959	7.54	81.538	No especificada
<i>Mus musculus</i>				
NP_619589	41.643	8.22	82.287	1
NP_001288025	17.707	8.88	66.129	2
NP_001288023	21.573	8.79	82.063	3
XP_006505366	36.303	8.74	82.208	X1
XP_006505367	32.162	7.62	77.786	X2
EDK99928	42.1	7.6	82.725	No especificada
Q9EQ09	41.643	8.22	82.287	No especificada
<i>Rattus norvegicus</i>				
NP_579840	42.378	6.12	82.473	No especificada
AAH97290	41.89	6.28	82.308	No especificada
O70156	41.89	6.28	82.308	No especificada

Se muestra a que especies pertenecen las isoformas de LOX1 en celdas amarillas; celdas blancas corresponden a isoformas encontradas en muestras biológicas, mientras que las celdas en celeste corresponden a secuencias proteicas obtenidas mediante métodos computacionales, según lo reportado en la Tabla 15. Cada columna, de izquierda a derecha, indica: el nombre de acceso en GenBank de cada isoforma de LOX1; la masa molecular de cada isoforma; el punto isoeléctrico de acuerdo a la escala de pH desde máximo de acidez (0) hasta el máximo de basicidad; porcentaje índice alifático isoformas; variante de la isoforma, según lo especificado en la Tabla 20. *kDa*: KiloDalton, unidad de medida de masa molecular.

5.3.3. Análisis filogenéticos

Los mecanismos de variación molecular que se consideran en filogenia incluyen mutaciones, duplicación de genes, reorganización de genomas, y recombinación. La filogenética molecular se encarga del estudio de filogenias y procesos evolutivos por medio del análisis de ADN o secuencias aminoácidas. Para inferir relaciones filogenéticas existen diversas metodologías que utilizan diferentes algoritmos computacionales, entre ellos algoritmos de agrupamiento y de optimización. UPGMA y Neighbor joining corresponden al primer grupo de algoritmos, emplean métodos de distancia que operan con distancias ultraméricas (igual largo de ramas) y distancias aditivas (variación en el largo de ramas) respectivamente.

Con la finalidad de establecer que secuencias proteicas son más cercanas entre sí, se construyeron dendogramas utilizando dos algoritmos: Neighbor joining y UPGMA. Se utilizó secuencia proteica de LOX1 de lagartija como grupo control externo. Los resultados se muestran a continuación en las **Figura 27** y **Figura 28**.

De acuerdo al dendograma generado usando Neighbor joining en la **Figura 27** se muestran todas las isoformas correspondientes a humano en un mismo clado con un 76% de probabilidad. Las anotaciones BAC81565 y NP_001166103 (isoforma 2 de LOX1) se agrupan como monofiléticos, y agrupadas como grupo hermano de NP_001166104 (isoforma 3 de LOX1). Los roedores surgen de un nodo común con probabilidad del 67%, y se agrupan en clados por especie correspondientes a ratón con un 99% de probabilidad y rata con un 72%. Las isoformas 2 y 3 de ratón son monofiléticas entre sí. La lagartija, grupo control externo, se ubica en una rama lejana de las otras taxas.

En la **Figura 28** se muestra el dendograma generado con algoritmo UPGMA. Se observa que la isoforma 2 de LOX1 humana se presenta como grupo basal en relación al resto de las secuencias. Luego, con un 98% de probabilidad se reconoce la división en clados correspondiente a humano y otro a roedores. El clado humano incluye todas las anotaciones correspondientes a isoforma 1 y 2, agrupando a isoforma 1 y 2 de LOX1 como monofiléticas entre sí. En roedores las secuencias proteicas de LOX1 de ratón se agrupan en un mismo clado, y lo mismo se reconoce en rata. En ratón las isoformas 2 y 3 de LOX1 se

agrupan monofiléticas entre sí, y como grupo hermano de las anotaciones correspondientes a isoforma 1. La lagartija, grupo control externo, se ubica en una rama aparte de las otras taxas.

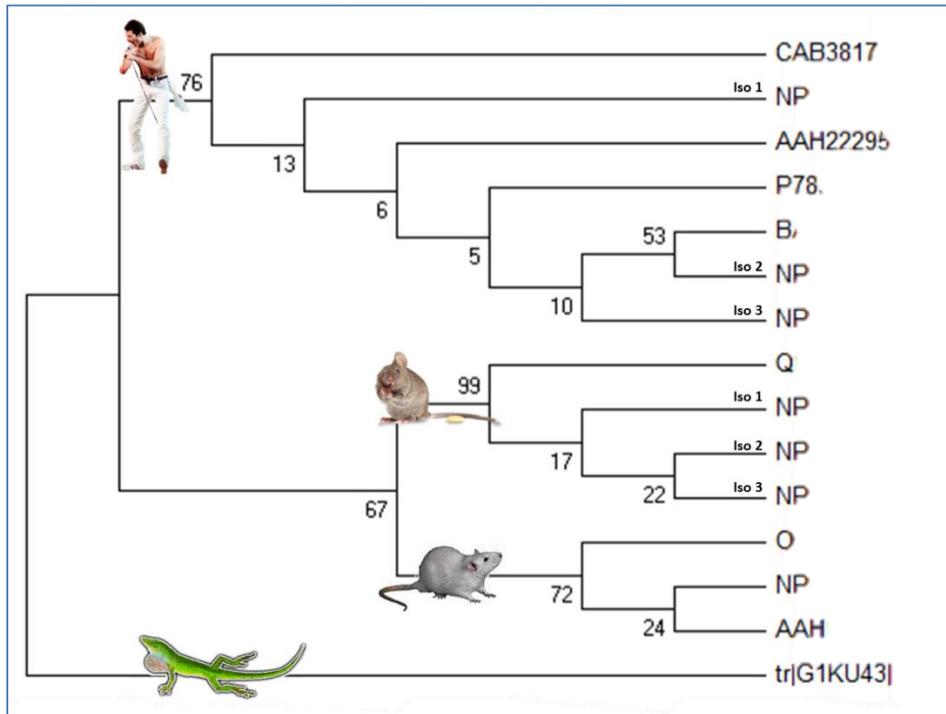


Figura 27. Dendograma generado a partir del agrupamiento de secuencias de LOX1 de humano, ratón y rata usando algoritmo *Neighbor joining*. Se indica con una imagen el organismo al que pertenece cada rama; la proteína correspondiente al final de cada rama; la probabilidad de ocurrencia en cada nodo; y en letras negras, junto a el nombre proteico, la isoforma en caso de que corresponda. Se utiliza LOX1 de lagartija como grupo control externo.

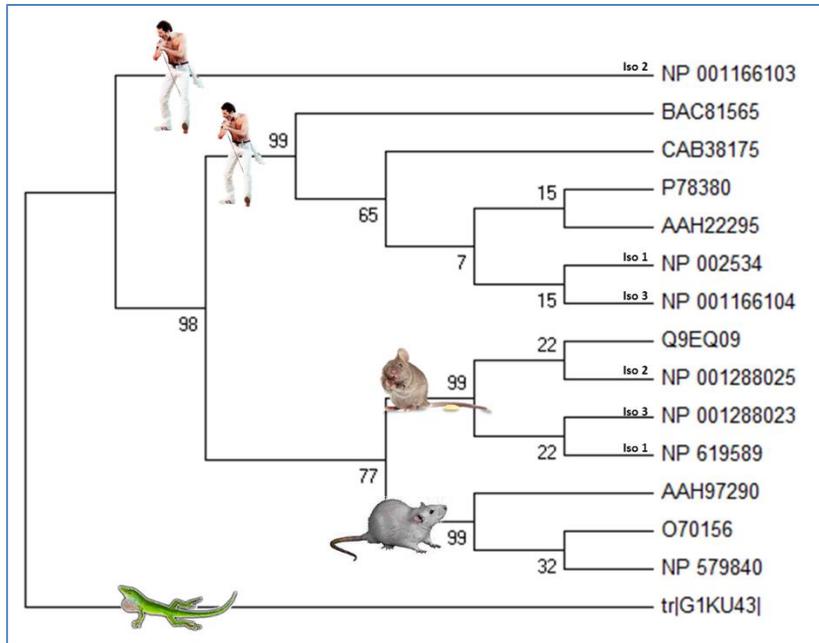


Figura 28. Dendograma generado a partir del agrupamiento de secuencias de LOX1 de humano, ratón y rata usando algoritmo *UPGMA*. Se indica con una imagen el organismo al que pertenece cada rama; la proteína correspondiente al final de cada rama; la probabilidad de ocurrencia en cada nodo; y en letras negras, junto a el nombre proteico, la isoforma en caso de que corresponda. Se utiliza LOX1 de lagartija como grupo control externo.

5.3.4. Alineamiento de secuencias proteicas LOX1

La **Figura 29** muestra el alineamiento de secuencias proteicas de LOX1 de humano y la **Tabla 24** en **ANEXOS** detalla los cambios determinados. Las secuencias de aminoácidos correspondientes a las isoformas 1 son prácticamente iguales (P78380, NP_002534, AAH22295, CAB38175, BAC81585), con excepción de BAC81585 en que se ve una variación de asparragina (N) por lisina (K) con respecto al resto de las isoformas 1 de LOX1. En isoforma 2 de LOX1 (NP_001166103) las diferencias más importantes encontradas son una deleción entre posiciones 143 a las 149 de la región cuello de LOX1, deleción mayor en posiciones 175-203, 210-216 y 225-273 en el dominio lectina tipo C, un cambio por leucina en la posición W150, y pérdida de las argininas en posición 208, 229, 231 y 248. En isoforma 3 de LOX1 (NP_001166104) se observan como diferencias más destacables una deleción mayoritaria de la región lectina tipo C en posiciones 190-273, con la consecuente pérdida de las argininas en posición 208, 229, 231 y 248, pero conservación de triptófano en posición 150.

Los resultados expuestos en las **Figura 30** y **Tabla 25 (ANEXOS)** detallan el alineamiento realizado para secuencias proteicas de LOX1 excluyendo secuencias hipotéticas para ratón. Se observa que ambas anotaciones correspondientes a isoforma 1 de LOX1 presentan igual secuencia aminoacídica (Q9EQ09, NP_619589). En isoforma 2 (NP_001288025) se reconoce una deleción entre posiciones 26-233 abarcando dominio citoplasmático, transmembrana, cuello, repeticiones típicas del receptor en roedores y región coiled coil, pero se mantiene intacto el dominio Lectina tipo C. En isoforma 3 se señala variación de leucina por cisteína en posición 159, y deleción entre posiciones 60-233 abarcando dominio cuello, repeticiones y coiled coil, se mantiene intacto dominios citoplasmático, transmembrana y lectina tipo C.

En rata se establecen secuencias proteicas similares a lo largo de todos los dominios de LOX1, y no se observan isoformas codificadas por gen *Olr1*. La anotación NP_579840 presenta 4 aminoácidos extra (N asparragina, L leucina, E ác. glutámico y M metionina) en posición 2-5 con respecto a las otras dos secuencias (**Figura 31** y **Tabla 26** en **ANEXOS**).

En la **Figura 32** se muestra el alineamiento de las isoformas 1 de LOX1 de humano, ratón y rata [secuencias que se asumen como isoforma1: O70156 de rata, Q9EQ09 de ratón y P78380 de humano de acuerdo a UniProt]. El análisis con MEGA 6 arroja 227 aminoácidos conservados a lo largo de las 3 especies alineadas, de un total de 366 analizados. Las regiones conservadas se observan a lo largo de todos los dominios descritos, incluso en algunos casos se señalan bloques conservados como en la región cuello en posiciones 62 a la 67 y los registrados en el dominio Lectina tipo C en posiciones 235-240, 245-250. 345-350 y 352-359. En humano se distingue una delección de las tres repeticiones características de roedores ubicadas en posiciones 96-233 de proteínas LOX1.

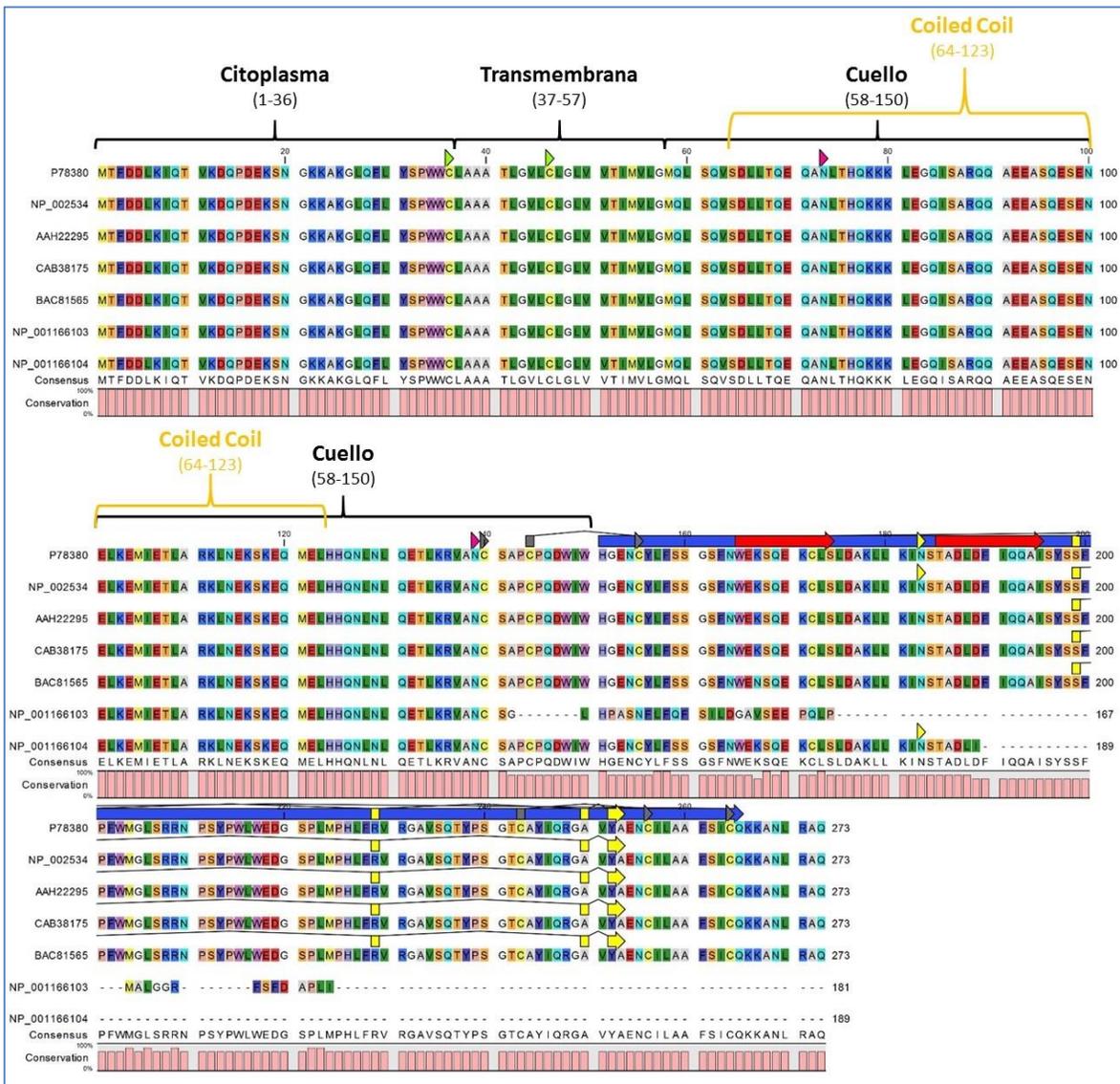


Figura 29. Alineamiento de secuencias proteicas de LOX1 en humano. Se indica cada residuo aminoacídico con su letra símbolo y un color identificativo; regiones citoplasmática, transmembrana, cuello y motivo coiled coil se indican con llaves; puentes disulfuros en flechas grises sobre el alineamiento; dominio C Tipo Lectina con flecha azul; sitios de glicosilación en flechas fucsia; regiones tipo helicoidal en flechas burdeo; y flechas en verde cata indican sitios de unión a lípidos. Secuencias que se asumen como isoforma1: P78380, NP_002534, AAH22295, CAB38175, BAC81565; secuencias correspondiente a isoforma2: NP_001166103; secuencia correspondiente a Isoforma3: NP_001166104.

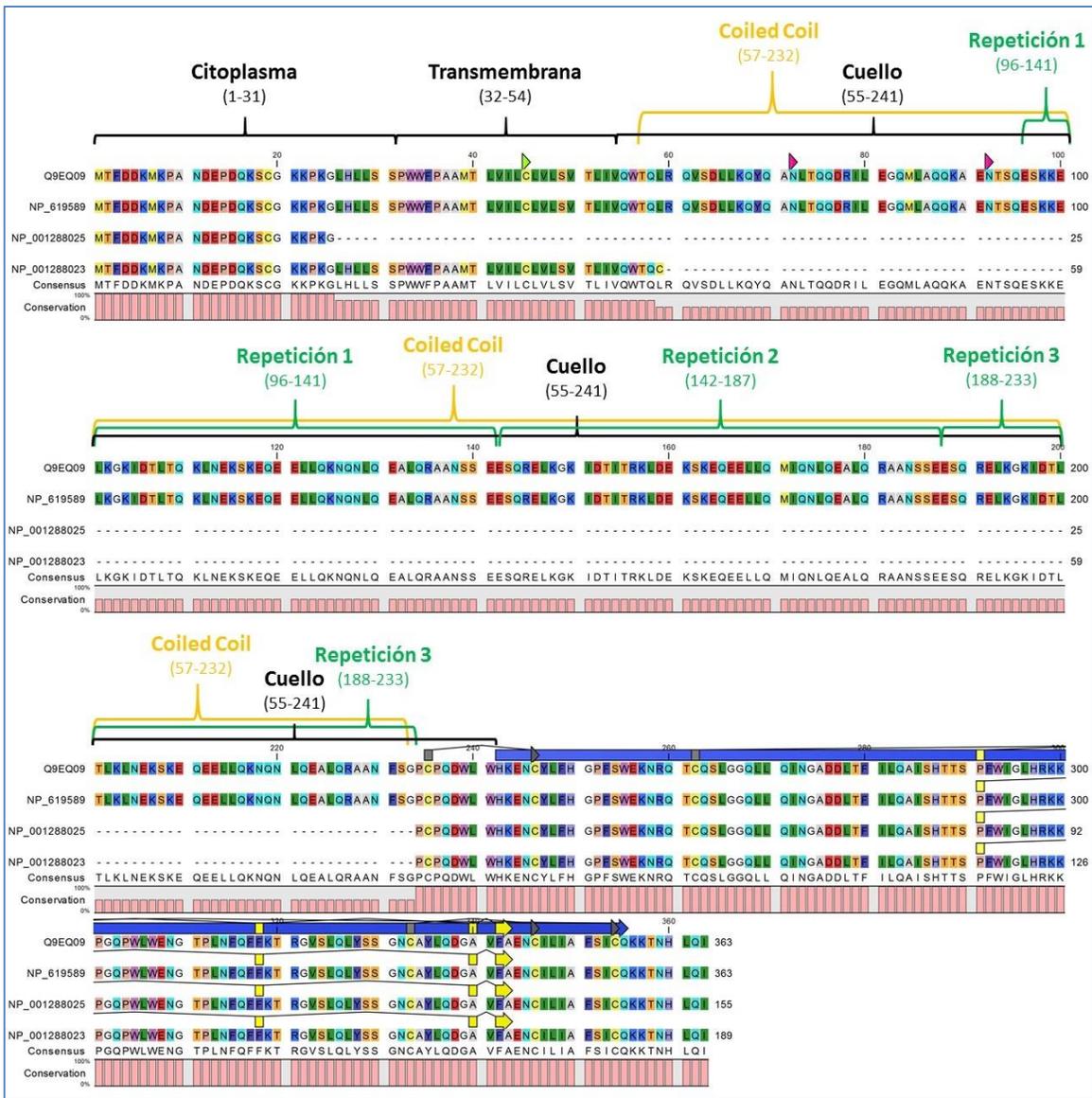


Figura 30. Alineamiento de secuencias proteicas de LOX1 en ratón. Se indica cada residuo aminoacídico con su letra símbolo y un color identificativo; regiones citoplasmática, transmembrana, cuello, motivo coiled coil y repeticiones típicas en ratón se indican con llaves; puentes disulfuros en flechas grises sobre el alineamiento; dominio C Tipo Lectina con flecha azul; sitios de glicosilación en flechas fucsia; regiones tipo helicoidal en flechas burdeo; y flechas en verde cata indican sitios de unión a lípidos. Secuencias que se asumen como isoforma1: Q9EQ09, NP_619589; secuencias correspondiente a isoforma2: NP_001288025; secuencia correspondiente a Isoforma3: NP_001288023.

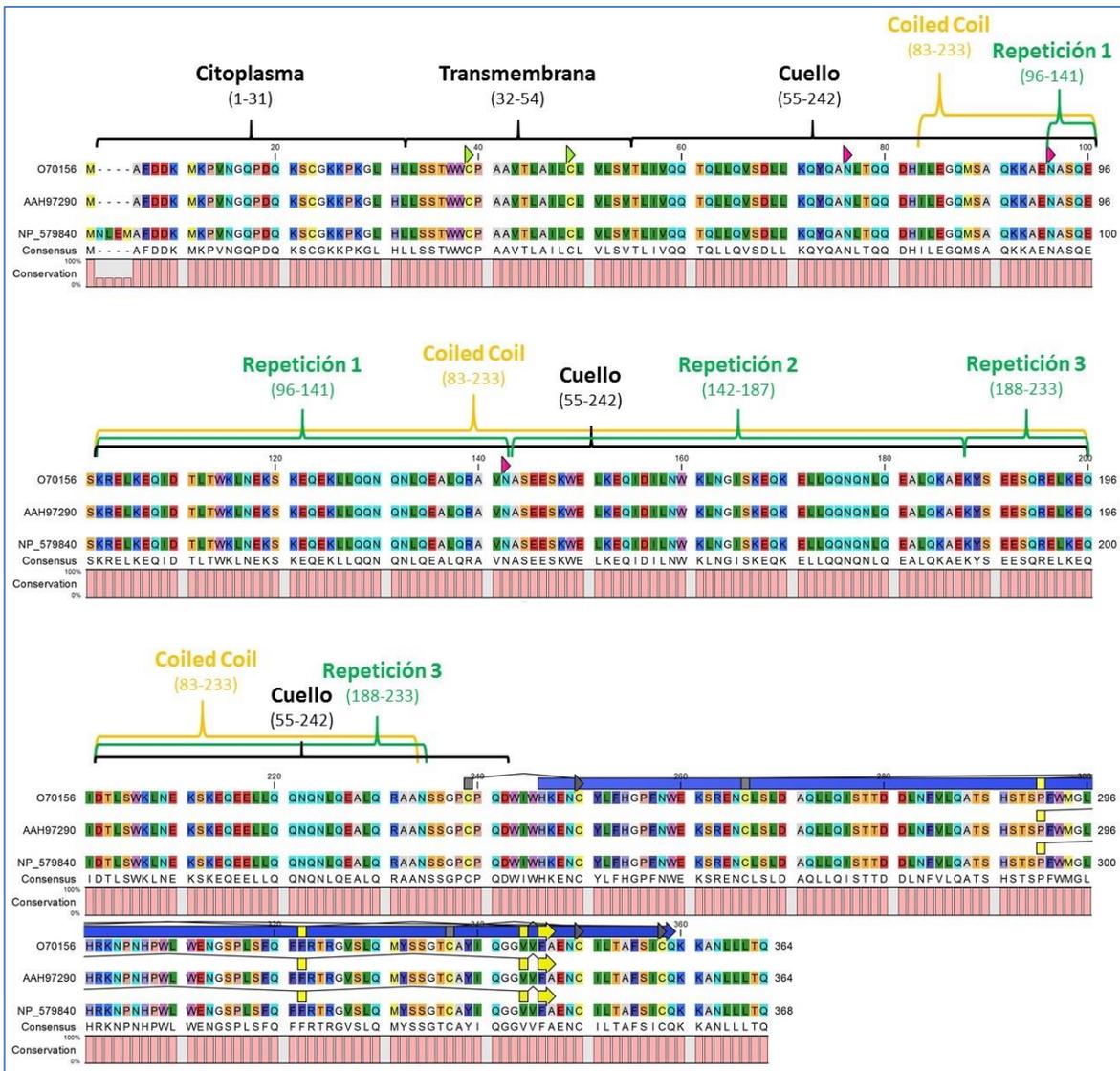


Figura 31. Alineamiento de secuencias proteicas de LOX1 en rata. Alineamiento realizado en CLC Sequence Viewer. Se indica cada residuo aminoacido con su letra símbolo y un color identificativo; regiones citoplasmática, transmembrana, cuello, motivo coiled coil y repeticiones típicas en rata se indican con llaves; puentes disulfuros en flechas grises sobre el alineamiento; dominio C Tipo Lectina con flecha azul; sitios de glicosilación en flechas fucsia; regiones tipo helicoidal en flechas burdeo; y flechas en verde cata indican sitios de unión a lípidos.

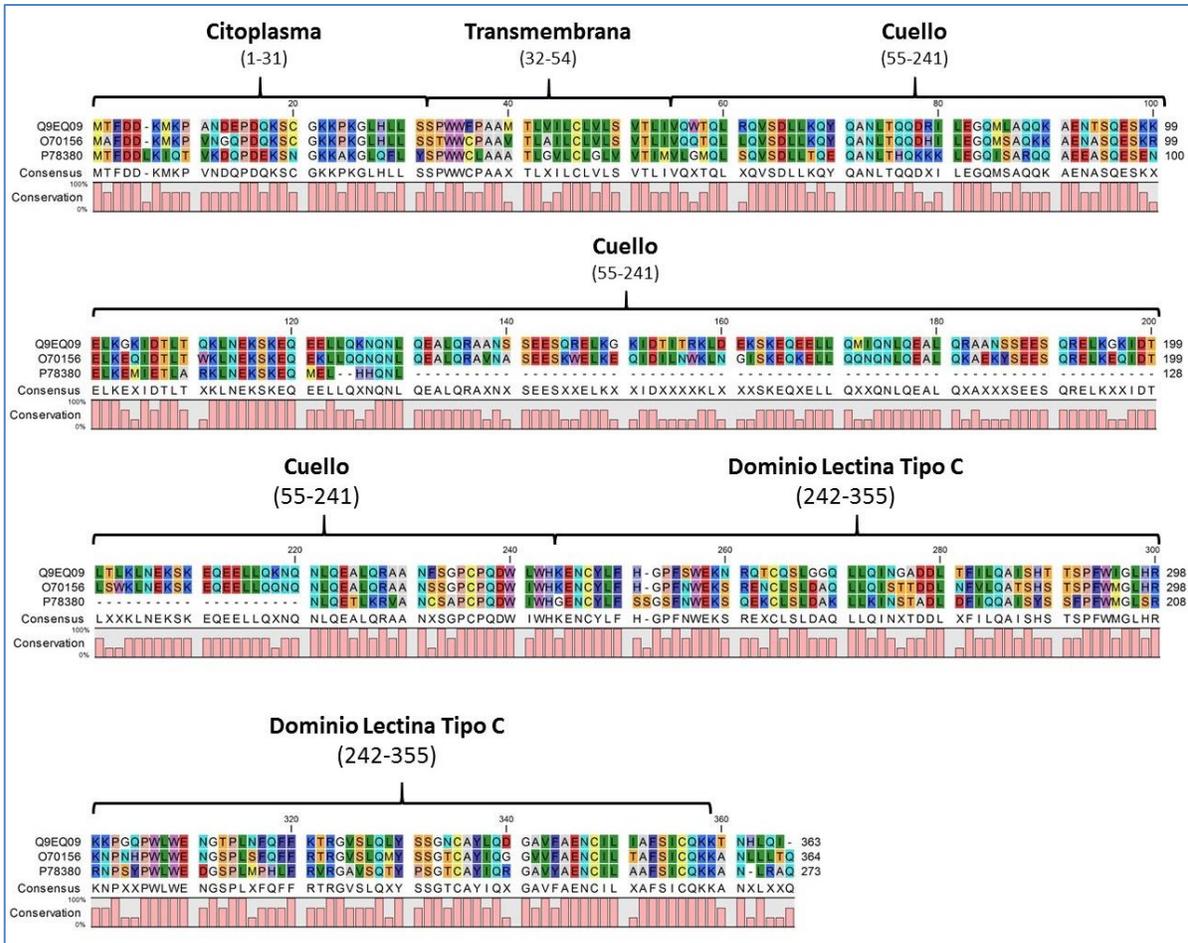


Figura 32. Alineamiento de secuencias Isoforma 1 de LOX1 en humano, ratón y rata. Alineamiento realizando en CLC Sequence Viewer. Se indica cada residuo aminoacido con su letra símbolo y un color identificativo; regiones citoplasmática, transmembrana, cuello y dominio Lectina Tipo C de rata se indican con llaves.

5.3.5. Modelamiento de proteínas

Con el fin de visualizar las diferencias a nivel de estructura secundaria y terciaria de LOX1 en las distintas especies, y entre las isoformas para una misma especie, se realizó modelamiento de isoformas de LOX1 usando SwissModel. Las diferencias en estructura primaria podrían alterar la estructura del receptor LOX1, y con eso su funcionalidad. La estructura utilizada como molde corresponde a la región extracelular de la isoforma 1 de LOX1 humana, descrita por cristalografía de rayos X , a excepción de casos particulares en que se detalla la estructura utilizada como molde.

El detalle de cada estructura hipotética para isoforma 1 de las tres especies en estudio se muestra en la **Figura 33**. Los porcentajes de identidad de secuencias, similitud y cobertura de las isoformas 1 de LOX1 de ratón, rata y humano fueron bastante similares por lo que construir un modelo en base a la estructura extracelular de isoforma 1 de LOX1 humano es una aproximación adecuada. Los índices de evaluación (GMQE y QMEAN) cercanos a 1 de la estructura compuesta para cada isoforma indican que los modelos construidos son una aproximación cercana a las estructuras probables para isoforma 1 de LOX1 en rata y ratón, pero no son el mejor modelo (**Tabla 12**).

En el caso del modelamiento de isoforma 2 para humano el alineamiento múltiples generado en BLAST arrojó mayor similitud de secuencias con una proteína llamada moesina, proveniente de la especie *Spodoptera frugiperda*. Los resultados del modelamiento usando a moesina como molde para la isoforma 2 de LOX1, y usando el molde regular, se muestran en la **Figura 34** y **Tabla 12**. De acuerdo a los porcentajes de identidad, similitud, cobertura e índices GMQE y QMEAN el modelo construido usando la secuencia de moesin como molde es de mejor calidad que el construido usando la secuencia de isoforma 1 de LOX1 humana, por lo tanto la estructura real de la isoforma 2 de LOX1 humana debe tener mayor similitud a moesina que a isoforma 1 humana. Sin embargo, ambos modelos son de baja calidad según a índices QMEAN y GMQE. A partir de los resultados expuestos para isoforma 3 de LOX1 en cuanto a los porcentajes de identidad, similitud, cobertura, índices GMQE y QMEAN se deduce que el modelo construido es de baja calidad, lo que indica que isoforma

3 debe tener un tipo de estructura distinto de isoforma 1 de LOX1 humana, o que no compartan gran similitud en la región extracelular (coincidente con los observado en el alineamiento de la **Figura 29**)

En la **Figura 35** se muestran las estructuras modeladas para isoformas 1, 2 y 3 de ratón. Los resultados indican que los modelos construidos para isoforma 2 y 3 tienen alto porcentaje de identidad de secuencias, similitud, cobertura e índice GMQE indicando buena calidad del modelo. Sin embargo, el índice QMEAN con valor negativo indica inestabilidad en su estado energético total y por tanto baja calidad del modelo (**Tabla 12**).

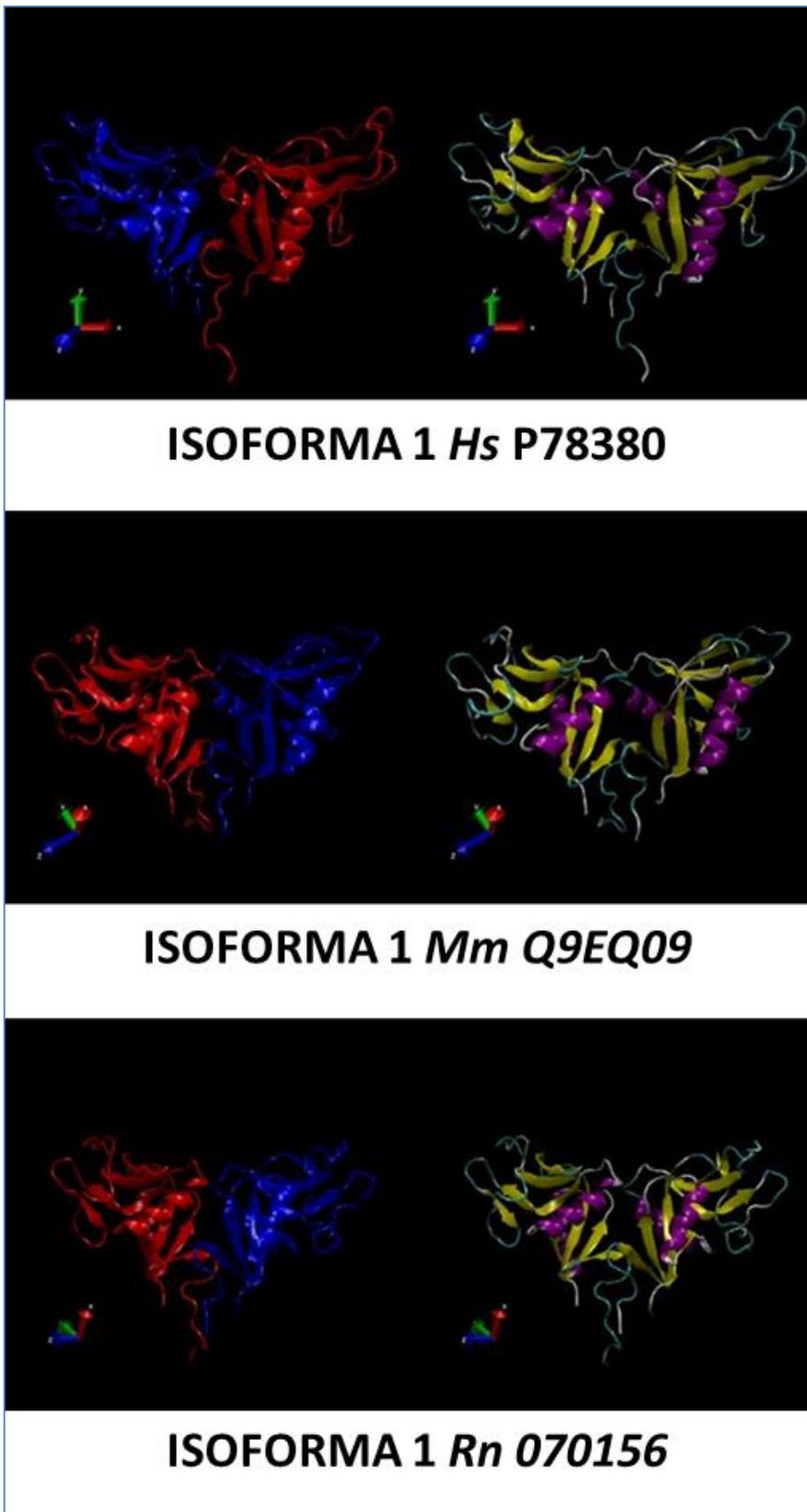


Figura 33. Modelos proteicos de isoforma 1 LOX1 para humano, ratón y rata. Al costado izquierdo se muestran los modelos construido indicando en rojo y azul las cadenas que forman el dímero, mientras que en el costado derecho se muestra el modelo construido enfatizando la estructura secundaria (hélices α en morado, sábanas β en amarillo).

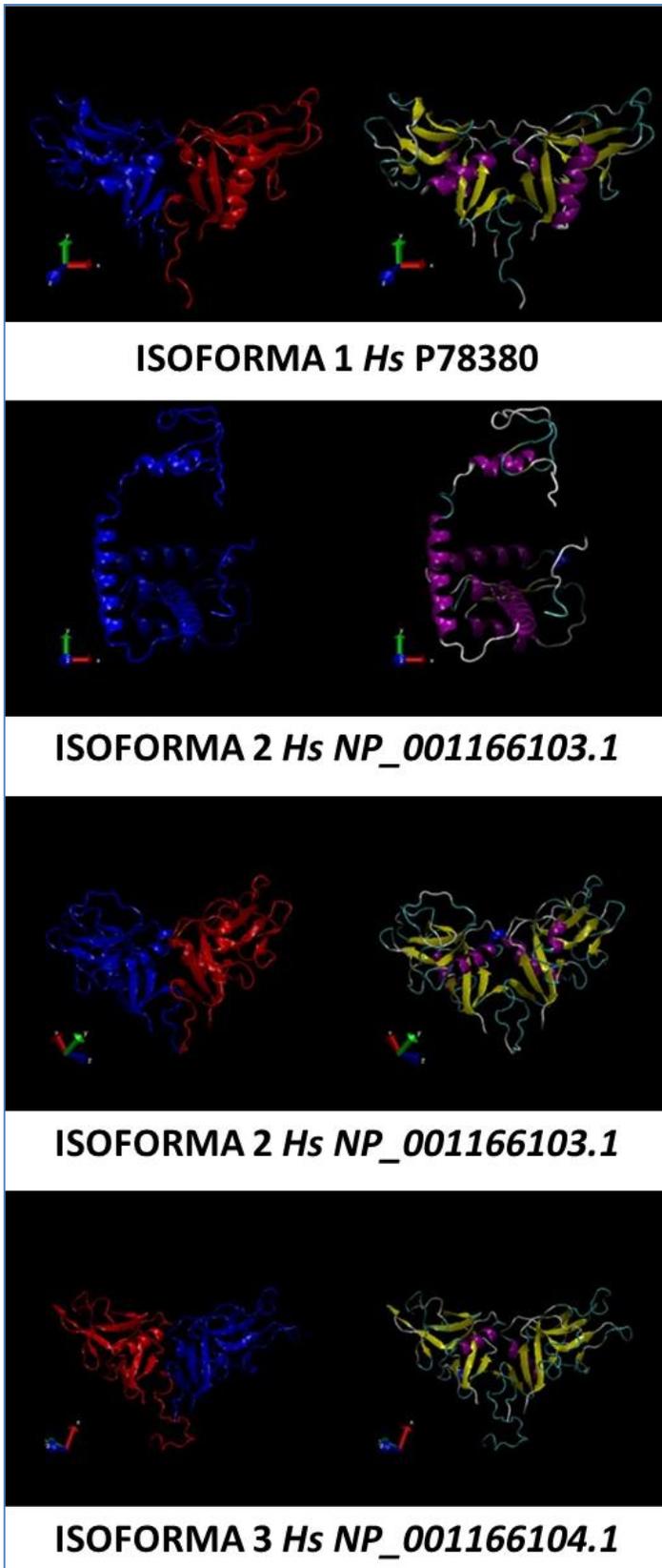


Figura 34. Modelos proteicos de isoformas LOX1 para humano. Al costado izquierdo se muestran los modelos construido indicando en rojo y azul las cadenas que forman el dímero, mientras que en el costado derecho se muestra el modelo construido enfatizando la estructura secundaria (hélices α en morado, sábanas β en amarillo). En el caso de isoforma 2, se muestra primero el modelo construido usando la secuencia molde de *moesin*, seguido del modelo construido usando como molde la secuencia de la isoforma 1 de LOX1 humano.

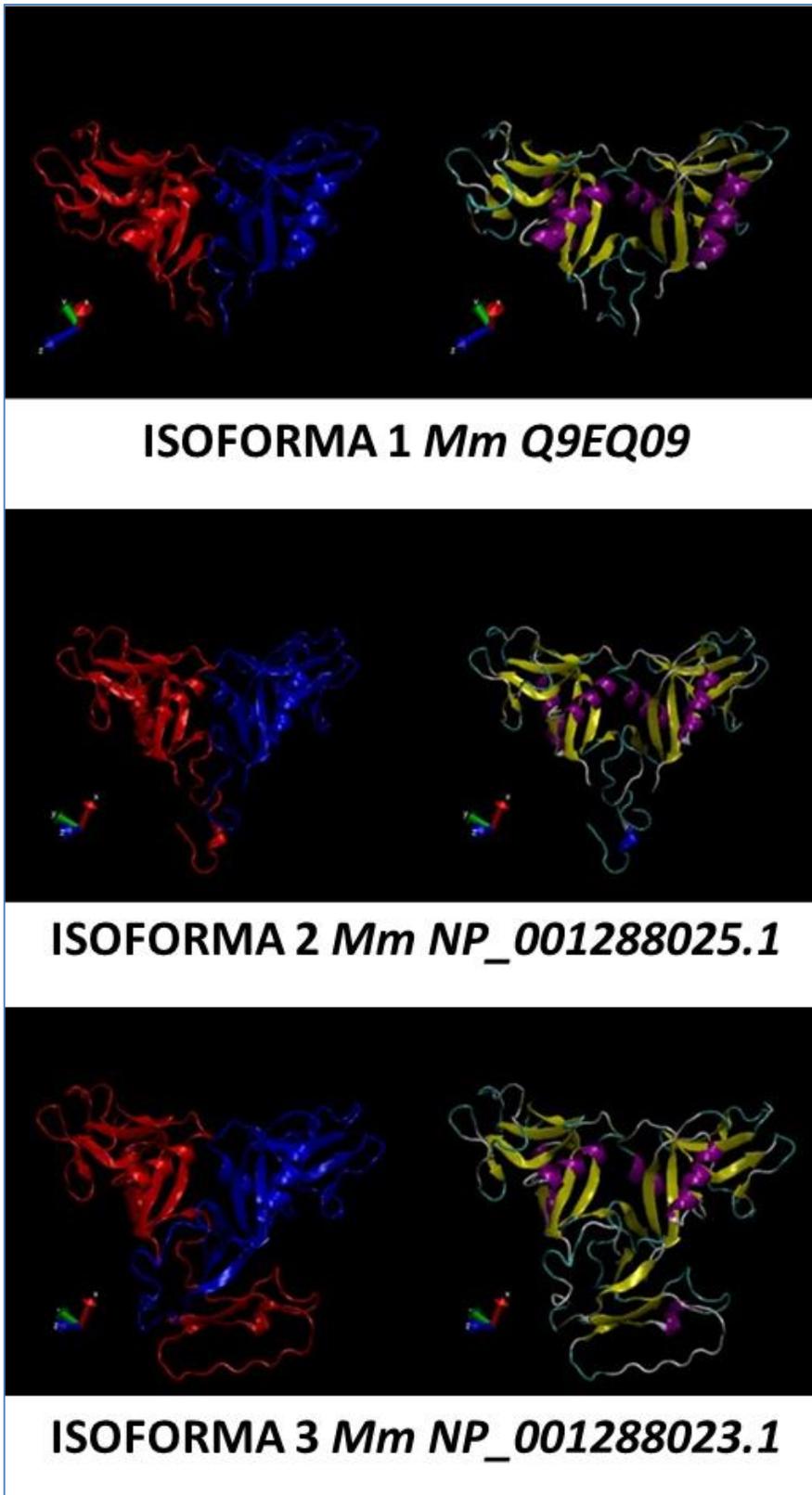


Figura 35. Modelos proteicos de isoformas LOX1 para ratón. Al costado izquierdo se muestran los modelos construido indicando en rojo y azul las cadenas que forman el dímero, mientras que en el costado derecho se muestra el modelo construido enfatizando la estructura secundaria (hélices α en morado, sábanas β en amarillo).

Tabla 12. Detalle de valores obtenidos en el modelamiento de isoformas de LOX1 en humano, ratón y rata.

Model	Template	Seq. Identity	Oligo state	Resolution	Seq. Similarity	Coverage	GMQE	QMEAN4
<i>Hs</i> Iso 1	1ypq.1.A	100.00	Homo-dimer	1.40Å	0.63	0.49	0,71	-0,36
<i>Hs</i> Iso 2	2i1j.1.A (moesin)	23.76	Monomer	2.10Å	0.32	1.00	0,48	-5,43
<i>Hs</i> Iso 2	1ypq.1.A	12.59	Homo-dimer	1.40Å	0.26	0.75	0,38	-10,62
<i>Hs</i> Iso 3	1ypq.1.A	18.52	Homo-dimer	1.40Å	0.28	0.71	0,41	-10,42
<i>Mm</i> Iso 1	1ypq.1.A	100.00	Homo-dimer	1.40Å	0.63	0.49	0,71	-0,36
<i>Mm</i> Iso 2	2ypq.1.A	60.45	Homo dimer	1.40Å	0.50	0.86	0,79	-4,19
<i>Mm</i> Iso 2	1ypq.1.A	60.45	Homo-dimer	1.40Å	0.50	0.71	0,73	-6,74
<i>Rn</i> Iso 1	1ypq.1.A	70.90	Homo-dimer	1.40Å	0.53	0.37	0,50	-1,55

Se muestran las especificaciones del modelo construido: isoforma modelada y especie correspondiente en cursiva; molde (template) usado; identidad de secuencias; estado del oligo (homo dímero o monómero); resolución de la estructura molde; similitud de secuencias molde y objetivo; grado de cobertura de la secuencia objetivo sobre la secuencia molde; índices GMQE y QMEAN4.

5.3.6. Discusión sección proteína LOX1

Los datos agrupados en las **Tabla 21**, **Tabla 22**, **Tabla 23** en **ANEXOS** hacen referencia a las distintas secuencias de LOX1 encontradas en bases de datos. Se detallan secuencias encontradas en muestras de tejidos y obtenidas mediante uso de algoritmos computacionales. Éste último punto es de relevancia, ya que de acuerdo a lo expresado en dichas tablas, 6 de 14 secuencias de humano y 3 de 7 secuencias en ratón son hipotéticas. Sin embargo, al examinar más de cerca dichas secuencias se ven similitudes entre las hipotéticas y reales. Por ejemplo en humano las secuencias CRA son hipotéticas pero tienen tamaños de secuencias (**Figura 41** en **ANEXOS**), masa, punto isoeléctrico, e índice alifático (**Tabla 11**) iguales a las isoformas empíricas (CRA_a comparte los mismos datos con Isoforma 1; CRA_b con Isoforma 2; y CRA_c con Isoforma 3). De la misma forma, secuencias hipotéticas ABC40744 y AAC97927 con isoforma 1 NP_002534 y secuencias P78380, AAH2295, CAB38175, BAC81565. En ratón, sucede lo mismo entre isoforma 1 NP_619589 y secuencia Q9EQ09. Es de suponer que se traten de la misma secuencia con información faltante al momento de la anotación, y que en el caso de las secuencias hipotéticas el programa arroje estas secuencias como probables, lo que refuerza la confianza en este tipo de metodología sujeta a comprobación.

En el caso de las isoformas predictivas denotadas con una X (isoforma X1 en humano, X1 y X2 en ratón) en las **Tabla 21**, **Tabla 22** y **Tabla 23** no existe relación con alguna de las isoformas encontradas en muestras biológicas. Por lo tanto, efectivamente se trataría de isoformas hipotéticas.

La variación en las propiedades fisicoquímicas de las distintas isoformas expuestas en la **Tabla 11** encuentra explicación en las **Figura 44**, **Figura 45** y **Figura 46** de este trabajo. Variaciones en la cantidad de leucina, cisteína, triptófano y fenilalanina serían las responsables de los cambios en el punto isoeléctrico e índice alifático de isoformas de humano; mientras que en ratón los aminoácidos responsables serían ácido glutámico, fenilalanina, leucina, glutamina y valina. Las diferencias de masa entre isoformas se explican principalmente por las variaciones en los tamaños de secuencias de aminoácidos (**Tabla 11**)

y composición, fenómeno ilustrado en los alineamientos de las **Figura 29, Figura 30 y Figura 31**.

En rata no existen isoformas proteicas por lo que se asume que todas corresponden a la misma isoforma. La variación de 4 aminoácidos en tamaño (**Tabla 11**), presumiblemente sea un error de anotación o polimorfismo típico de la muestra secuenciada.

La estructura de LOX1 ha sido descrita por Park y cols. mediante cristalografía de Rayos X identificando a la estructura extracelular de LOX1 como monómero, y asignando a la estructura de dímero la capacidad de unir oxLDL. De acuerdo a su experimentación, las estructuras necesarias para mantener la formación del dímero se resumen en: 1) Puente disulfuro formado entre los monómeros gracias a la cisteína en posición 140 (C140); 2) triptófano en posición 150; 3) puentes salinos entre asp147 del monómero A y his151 del monómero B, y viceversa; 4) puentes de hidrógeno; 5) red de puentes de hidrógeno entre 9 aminoácidos y residuos de agua proteína-proteína. Los estudios realizados por Nakano y cols. en mutantes reafirman los supuestos anteriores, pues el mutante W150A es incapaz de formar el dímero de forma estable a pesar de la formación del puente disulfuro en C140, perdiendo su funcionalidad. El trabajo realizado por Ohki y cols. afirma que la sección cuello característica de LOX1 es necesaria para mantener el dímero funcional (**Figura 36**). Por lo tanto, de acuerdo los alineamientos mostrados en las **Figura 29 y Figura 30**, la isoforma 2 en humano e isoformas 2 y 3 en ratón son incapaces de formar dímero. En el caso humano, en isoforma 2 el residuo W150 se reemplazado por un residuo de leucina según la **Tabla 24** en **ANEXOS**. En ratón en la isoforma 2 existe una delección de la región citoplasmática, transmembrana y gran parte del cuello, y en la isoforma 3 existe delección de la región cuello.

La funcionalidad de LOX1 no solo se ve truncada al no poder formar el dímero. También es necesario tener los residuos aminoacidicos necesarios para atraer los residuos de oxLDL al receptor y lograr la unión. Estos residuos forman una “espina básica” formada por argininas en la superficie de LOX1 (**Figura 36**). En la isoforma 2 humana la “espina básica” se ve delecionada, obstaculizando la atracción entre residuos positivos de oxLDL y los residuos negativos de la superficie de oxLDL. Específicamente, isoforma 2 no contiene los residuos R229, R231 y R248 que forman la “espina básica”, y el residuo R208 se encuentra

reemplazado por glicina (**Tabla 24** en **ANEXOS**). Isoforma 3 contiene los residuos necesarios para formar el dímero, pero sufre una deleción en el segmento lectina tipo C resultando en una isoforma no funcional al perder los residuos de la “espina básica”.

Se descarta la posibilidad de que las isoformas 2 y 3 de humano correspondan a LOX1 soluble, de acuerdo al alineamiento mostrado en la **Figura 29** ambas isoformas han perdido el sitio de corte en la región “*coiled coil*” del cuello que genera la forma soluble de LOX1 (**Figura 36**).

En el caso de ratón se requieren estudios más profundos respecto a la identificación de la ubicación de los residuos de la “espina básica” en isoforma 2 y 3.

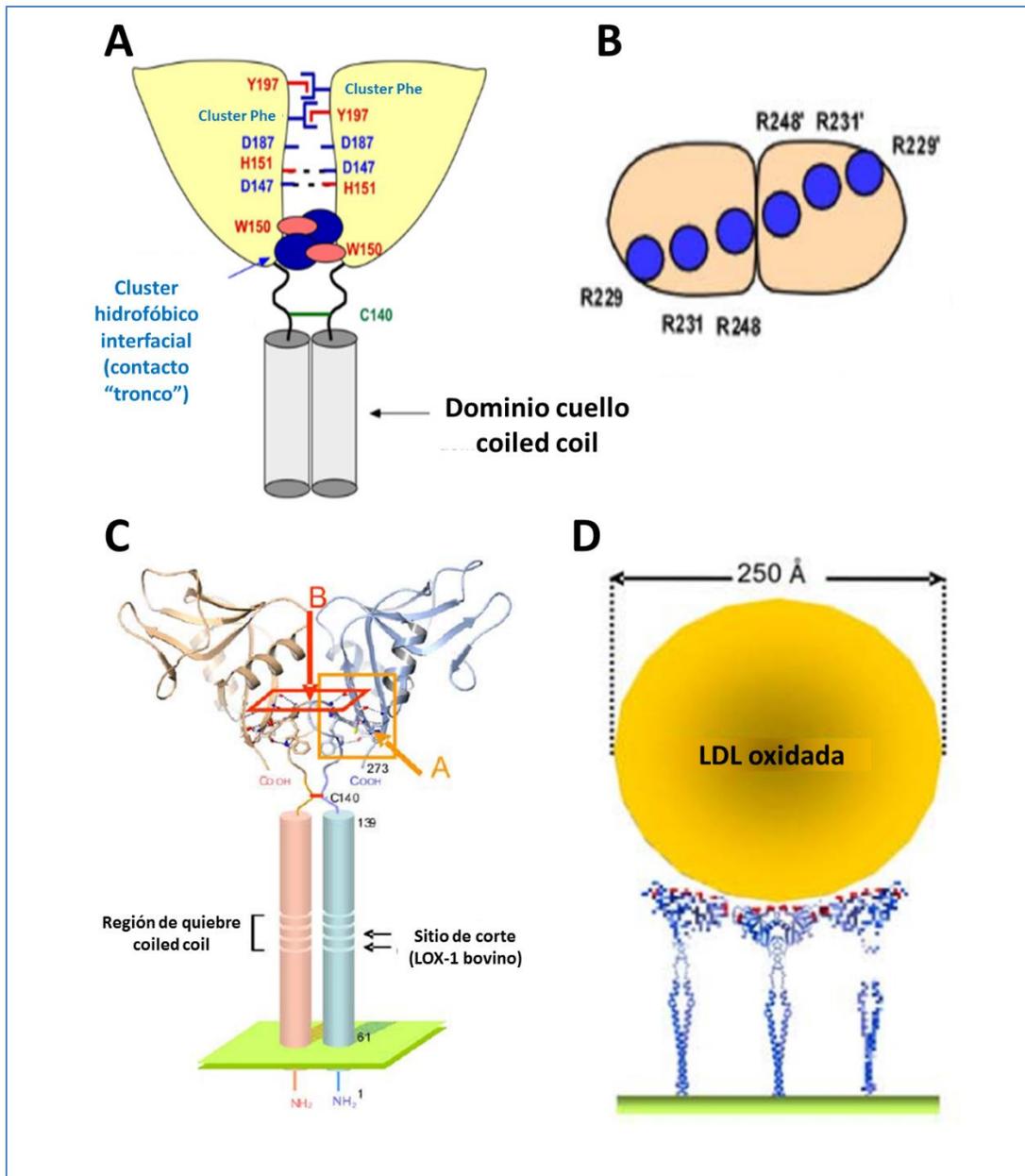


Figura 36. Esquema de residuos de LOX1 humano. **A)** Interacciones inter dominio observadas en estructura de dímero LOX1. Las interacciones inter dominio hidrofóbicos mediados por el anillo W150, que forma una estructura tipo tronco, estabiliza el dímero canónico asociado a interacciones de H151 y Y197. La unión disulfuro en C140 está ubicado en una región relativamente flexible. Por lo tanto, el puente disulfuro por sí solo no estabiliza el dímero, pero restringe el dominio a un espacio limitado que facilita la formación del dímero por asociación. El dominio cuello y coiled-coil también están involucrados en la proximidad espacial del dominio lectina tipo C. El dímero canónicamente estable mantiene la “espinas básicas” en la superficie. Imagen adaptada de . **B)** Esquema de la superficie de LOX1, y los residuos de arginina formando la “espinas básicas”. Imagen adaptada de . **C)** Esquema de la parte extracelular de LOX1, indicando los sitios de corte en la zona coiled coil del cuello, pudiendo originar LOX1 soluble. Imagen adaptada de . **D)** Representación de la estructura completa de LOX1 en la superficie de la célula, basado en la estructura cristalizada del dominio lectina tipo C y el modelo estructural de la región cuello. Imagen adaptada de .

En el alineamiento de la **Figura 32** se muestran las isoformas 1 correspondiente a cada especie en estudio. Se distingue que el residuo W150 se encuentra conservado en las 3 especies, así como también los residuos R231, R229 y R208 correspondientes a humano. No sucede lo mismo con el residuo R248. Por lo tanto, es de esperarse que la capacidad de formar dímeros esté intacta en la isoforma 1 de las 3 especies, así como también la capacidad de unir oxLDL a través de la “espina básica”. Sin embargo, la forma en que las diferencias aminoacídicas, tanto en tipo como cantidad de aminoácidos, afectan la estructura de LOX1 dependerá de la función específica y las características del ligando de LOX1 en cada especie. Específicamente, se observa en humano una región cuello mucho más corta que en ratón y rata, pudiendo afectar la dimerización del receptor.

Para examinar los cambios en las estructuras extracelulares de las isoformas de LOX1 por especie se realizaron modelamiento en SwissModel. La **Figura 33** muestra los modelos generados para las Isoformas 1 de cada especie. No se establecen grandes diferencias estructurales entre los modelos de isoforma 1 en humano, ratón y rata. Pero dada la carencia de las repeticiones típicas de ratón y rata en humano, el modelo en roedores está más compactado.

La **Figura 34** se muestra el modelamiento realizado para las isoformas en humano. No se ven gran diferencia entre las isoformas, pese a lo expuesto en el alineamiento de la **Figura 29**. Lo mismo sucede en la **Figura 35**, que muestra las isoformas de LOX1 en ratón. Esto se explica por el uso del mismo molde para el modelamiento, correspondiente a isoforma 1 humana, la baja similitud e identidad con la secuencia molde de algunas isoformas, y la baja calidad de los modelos construidos, evaluado por índice QMEAN y GMQE (**Tabla 12**).

En el caso del modelamiento de la isoforma 2 en humano al realizar el BLAST de la secuencia proteica arroja mayor similitud con la proteína moesina de *Spodoptera frugiperda* perteneciente a la familia de proteínas ezrin/radixin/moesin (ERM), que proporcionan un enlace conformacionalmente regulado desde las fibras de actina del citoesqueleto a la membrana plasmática. Este tipo de proteínas es capaz de unir directamente proteínas integrales de membrana como ICAM1-3, por lo que el rol funcional

de la isoforma 2 de LOX1 podría tener relación con la unión de proteínas integrales de membrana como ICAM, VCAM, colágeno y otras participantes del proceso de fibrosis en patologías cardíacas. Además, esta familia podría ser regulada por la fosforilación de una treonina en el dominio C-terminal, al igual que LOX1 de acuerdo a los antecedentes obtenidos en nuestro laboratorio usando análisis bioinformático de secuencias.

El modelo construido usando como molde el alineamiento con moesina de *Spodoptera frugiperda* es de mejor calidad que el modelo usando LOX1 como molde, de acuerdo a los índices QMEAN, GMQE, identidad, similitud y cobertura (**Tabla 12**). Por lo tanto, la estructura y función de la isoforma 2 en humano no debe distar mucho del modelo construido usando moesina como molde.

De acuerdo al alineamiento de la **Figura 29**, es probable que la isoforma 3 humana tenga forma y función similar a la isoforma 2, debido a la delección del segmento lectina tipo C capaz de formar el dímero y unir oxLDL. Lamentablemente esto no se ve reflejado en el modelamiento proteico, aunque al realizar el BLAST moesina de *Spodoptera frugiperda* aparece dentro de los primeros resultados, mostrando similitud en secuencias.

Las **Figura 27** y **Figura 28** muestran los dendogramas construidos usando las secuencias proteicas de LOX1 de las tres especies en estudio. En general, tanto el dendograma construido usando Neighbor Joining como UPGMA son bastante similares, con excepción de la isoformas 2 en humano. Dependiendo el algoritmo que se use aparece la isoforma 2 con origen anterior al resto de las isoformas. Coincidente con lo encontrado al realizar BLASTp para el modelamiento de proteínas, isoforma 2 parece ser la isoforma más alejada y/o distinta al resto de las versiones proteicas de LOX1. Las diferencias en secuencia y/o anotaciones mostradas en los alineamientos, BLASTp y árboles filogenéticos pueden reflejar su función y estructura más primitiva.

Las diferencias entre algoritmos, en la agrupación de clados en secuencias de una misma especie, están asociado al desigual tipo de método de agrupación característico de cada algoritmo.

En ambos algoritmos se muestra a rata y ratón con un origen común, apuntando a la agrupación en la casta de roedores; a humano en una rama distinta con origen común a la rama que diverge a roedores, indicando origen mamífero; y en una rama completamente distinta y alejada del resto de los grupos en estudio se encuentra LOX1 de lagartija, con un origen común al comienzo del árbol.

El rol de isoforma 1, 2 y 3 de LOX1 humano en fibroblasto cardiaco en condiciones saludables y patológicas sigue sin dilucidar. No obstante, se podría conjeturar un rol asociado a las vías de señalización intracelulares a las isoformas 2 y 3 de LOX1, específicamente asociado a la remodelación del citoesqueleto observado en los fibroblastos cardiacos en condiciones patológicas cardiacas bajo el efecto de TGF- β 1, dada la conservación de la región intracelular de acuerdo a los alineamientos expuestos y a la semejanza con la proteína moesin. Así como también se pueden asociar a un rol de anclaje al citoesqueleto de proteínas de adhesión integrales de membrana.

En resumen, la cantidad de información relativa a gen OLR1 (o su parte proteica LOX1) asociada a enfermedades cardiovasculares es limitada. Esto alienta al trabajo en esta línea de investigación. Justifica trabajos de indagación como este que ayudan a crear nuevas investigaciones innovadoras y creativas a partir de bases teóricas utilizando un computador como laboratorio, y algoritmos como reactivos. Sin embargo, y de acuerdo a lo expuesto en esta memoria de título, ¿los modelos utilizados en investigación *in situ* son apropiados para extrapolar los resultados a humano?

La mayor interferencia entre el modelo de miofibroblasto murino y el miofibroblasto característico de patologías cardíacas humana probablemente se dé a nivel de regulación de la transcripción del gen OLR1 y en la interacción intracelular del receptor LOX1, que es donde se presentan las diferencias más significativas entre especies. Además, la presencia de isoformas en humano, distintas a las presentes en ratón, y la ausencia de éstas en rata impide la simulación más cercana de la condición patológica y/o saludable en el miofibroblastos.

Es probable que tanto la interferencia de transcritos producto de solapamiento, manifestaciones de APA e isoformas jueguen un rol directo o indirecto en las vías de señalización del fibroblasto que influya en la diferenciación a miofibroblasto, síntesis de proteínas de adhesión y colágeno típico de patologías cardíacas de humano. El efecto coercitivo de todos estos fenómenos humano no puede simularse en modelo murino, pues es intrínseco a *Homo sapiens*. Sin embargo, a falta de un modelo más cercano *in vivo* e *in vitro* es un buen acercamiento dado lo conservado de las vecindades genéticas y similitudes en la región extracelular de isoformas 1 en humano, ratón y rata.

5. LIMITACIONES

La principal restricción de este trabajo es el carácter hipotético de sus resultados. No es posible asegurar que cierto fenómeno sea certeramente de la forma en que se expresan en los resultados de este trabajo. Comprobación con trabajo de laboratorio debe realizarse para validación en un modelo tangible. Sin embargo, ayuda a delimitar y guiar el trabajo de laboratorio.

Otro tipo de limitación refiere a la cantidad de secuencias descritas en bases de datos para OLR1 y/o LOX1. Análisis bioinformáticos requieren de gran cantidad de secuencias para lograr muestras representativas de la especie y objeto en estudio. Debido a lo poco explotado del tema de investigación de esta memoria, los resultados se ven restringidos por el bajo número de secuencias registradas.

Asimismo, la caracterización de estructura cristalina de LOX1 solo en humano dificulta el modelamiento de proteínas en otras especies con mayor cantidad de aminoácidos. Además, la existencia de estructura cristalina solo en la región extracelular de LOX1 impide la correcta asignación funcional y de vías de señalización mediadas por su parte intracelular. Qué sucede en la región transmembrana e intracelular de LOX1 es un misterio por el momento.

6. CONCLUSIONES

1. Existen varias anotaciones para gen OLR1, transcritos y proteína LOX1, entregando información distinta. El largo, secuencia, información contenida y versión de las anotaciones dependen del equipo investigador, técnica utilizada y tipo de muestra empleada al momento de secuenciar.
2. El contexto genómico de OLR1 es conservado a lo largo de las tres especies en estudio, de acuerdo a los genes vecinos de OLR1 y contenido de GC. Pero presenta diferencias en largos de secuencias y solapamientos especie específico.
3. La regulación de la transcripción de LOX1 es tejido y célula específico, involucra fenómenos de interferencia de transcritos según el contexto genómico en que está inmerso el gen OLR1, TFBSs asociados al promotor del gen OLR1 y sitios de poliadenilación alternativa. Estos fenómenos actuando de forma coercitiva potencialmente son responsables de la generación de isoformas de LOX1 en humano, rata y ratón y su rol en patologías cardíacas.
4. La isoformas 1 de LOX1 es conservada en humano, ratón y rata. Las isoformas 2 y 3 han perdido la capacidad de unir oxLDL por tanto tienen una función distinta a la isoforma 1 de LOX1, desconocida hasta el momento en patologías cardíacas.
5. Gen OLR1 sufre de poliadenilación alternativa (APA), generando transcritos de variado tamaño y tejido específico en humano

7. PROYECCIONES

Las proyecciones de este trabajo *in silico* prioritariamente tienen relación con la comprobación empírica de sus resultados. Dependiendo del enfoque que se desee dar a la investigación será el resultado a tratar. De esta forma, en el ámbito de la biología molecular se puede estudiar la comprobación de unión de TFs como AP-1, BRCA1, NKX1-3 y BACH1 al promotor de LOX1 utilizando footprinting; regulación de la transcripción de LOX1 mediante los TFs antes mencionados, utilizando Western blot y RTqPCR para cuantificar los niveles de transcripción de LOX1; coexpresión de LOX1 y TMEM52B en humano, y/o de CLEC9A en ratón; niveles de expresión de isoformas de LOX1; interferencia de ARNm; sitios de poliadenilación alternativa activos y perfil de transcritos de LOX1 en condiciones saludables y patológicas, etc.

Desde una perspectiva proteica se pueden realizar estudios respecto a las estructuras, inserción en membrana, localización en la célula, unión a proteínas de adhesión, unión a blancos farmacológicos, cinética, unión a oxLDL, etc de las distintas isoformas de LOX1 en las 3 especies.

En cuanto a vías de señalización es viable la investigación de los factores de transcripción hipotéticos y los blancos río arriba en la cascada de señalización, construyendo la red desde el núcleo hacia el citoplasma, uniéndose con los antecedentes ya acreditados.

Investigaciones como esta permiten tener una base científica para la comparación (o limitación) del estudio en distintos modelos animales. Probablemente en un futuro próximo será imprescindible realizar investigaciones *in silico* para elegir el modelo de estudio animal más adecuado, ya sea con fines éticos para con los animales y/o para resguardar la validez de los resultados obtenidos en la investigación científica en un laboratorio.

8. REFERENCIAS

- ©2014 CLC bio, a QIAGEN Company. *CLC Sequence Viewer 7.5*. Aarhus, s.f.
- American Heart Association. «Heart Disease and Stroke Statistics—2012 Update.» *Circulation*, 2012: e2–e220.
- Anderson, Stephen. «Shotgun DNA sequencing using cloned DNase I-generated fragments.» *Nucl. Acid Res.* 9, nº 13 (1981): 3015-3027.
- Aneja R., Hake P., Burroughs T., Denenberg A., Wong H., Zingarelli B. «Epigallocatechin, a green tea polyphenol, attenuates myocardial ischemia reperfusion injury in rats.» *Molecular Medicine* 10, nº 1-6 (2004): 55-62.
- APASdb. *Searching APA Sites, Alternative Polyadenylation Sites Database (APASdb)*. 2015. http://mosas.sysu.edu.cn/utr/search_APASdb.php?method=quick&id=49543&seqkeywords=uc010sha.1&db=APASdb/Homo.sapiens_hg19_human-all22-tissues (último acceso: Julio de 2015).
- Bastidas, R. y Zavala, Y. *Principios de Entomología Agrícola*. Sol de Barro, 1995.
- Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachil, Lipman DJ, Ostell J, Sayers EW. «GenBank.» *Nucleic Acids Res* 41 (2013): D36-42.
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., Schwede, T. «SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information.» *Nucleic Acids Research* 42, nº W1 (2014): W252-W258.
- Biocca S, Falconi M, Filesi I, Baldini F, Vecchione L, Mango R, Romeo F, Federici G, Desideri A, Novelli G. «Functional analysis and molecular dynamics simulation of LOX-1 K167N polymorphism reveal alteration of receptor activity.» *PLoS ONE* 4, nº 2 (2009): e4648.
- Boldogkoi, Z. «Transcriptional interference networks coordinate the expression of functionally related genes clusters in the same genomic loci.» *Frontiers in genetics* 3, nº 122 (2012): 1-17.

Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J., Schwede, T. «Protein structure homology modeling using SWISS-MODEL workspace.» *Nature Protocols* 4, nº 1 (2009): 1-14.

C.M. Millward, Mary Hayes. *A Biography of the English Language*. Cengage Learning, 2011.

Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., & Madden T.L. «BLAST+: architecture and applications.» *BMC Bioinformatics* 10 (2008): 421.

Carey, MF., Peterson, C.L., Smale, S. T. «The primer extension assay.» *Cold Spring Harb Protoc.* 2 (2013): 164-173.

Cavin Périer, R., Junier, T., Bucher, P. «The Eukaryotic Promoter Database EPD.» *Nucleic Acids Res* 26 (1998): 353-357.

Ciovacco W., Raskind W., Kacena M. «Human phenotypes associated with GATA-1 mutations.» *Gene* 427, nº 1-2 (2008): 1-6.

Cold Spring Harbor Laboratory Press. «Mapping RNA with nuclease S1.» *Nature Methods* 2, 2005: 397-398.

Chandler, Vicki. «Turning data into discovery.» 2015.

Charron F., Nemer M. «GATA transcription factors and cardiac development.» *Cell & Development Biology* 10 (1999): 85-91.

Chen H., VanBuren V. «A provisional gene regulatory atlas for mouse heart development.» *PLoS ONE* 9, nº 1 (2014): e83364.

Chen M., Masaki T., Sawamura T. «LOX-1, the receptor for oxidized low-density lipoprotein identified from endothelial cell: implications in endothelial dysfunction and atherosclerosis.» *Pharmacology & Therapeutics*, 2002: 89-100.

Chih Lee, and Chun-Hsi Huang. «LASAGNA-Search: an integrated web tool for transcription factor binding site search and visualization.» 54, nº 3 (Marzo 2013): 141-153.

Dahary D., Elroy-Stein O., Sorek R. «Naturally Occurring antisense: Transcriptional leakage or real overlap?» *Genome Research*, nº 15 (Febrero 2005): 364-368.

dbSNP . *NCBI National Center for Biotechnology Information*. 25 de Mayo de 2015. <http://www.ncbi.nlm.nih.gov/snp>.

- de Klerk E., Hoen P. «Alternative mRNA transcription, processing, and translation: insights from RNA sequencing.» *Trends in genetics* 31, nº 3 (2015).
- Dear, P.H. «Genome Mapping.» En *Encyclopedia of Life Sciences*. 2005.
- Del Principe D, Lista P, Malorni W, Giammarioli A. «Fibroblast autophagy in fibrotic disorders.» *J Pathol.* 229 (2013): 208-220.
- Departamento de estadísticas e información de salud, MINSAL. *Principales causas de muerte tasas según sexo Chile 1997-2012*. <http://www.deis.cl/?p=2543>, 2015.
- Diez-Roux G, Banfi S, Sultan M, Geffers L, Anand S, Rozado D, Magen A. «A high-resolution anatomical atlas of the transcriptome in the mouse embryo.» *PLoS Biol.* 9, nº 1 (2011): E1000582.
- Elkon R., Ugalde A., Agami R. «Alternative cleavage and polyadenylation: extent, regulation and function.» *Nature Reviews Genetics* 14 (2013): 496-506.
- Fullwood, M.J., Wei, C., Liu, E.T. «Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses.» *Genome Research* 19, nº 4 (2009): 521-532.
- GenBank: AAC97927.1. *lectin-type oxidized LDL receptor [Homo sapiens]*. 1998. <http://www.ncbi.nlm.nih.gov/protein/AAC97927.1>.
- GenBank: AAH22295.1. *Oxidized low density lipoprotein (lectin-like) receptor 1 [Homo sapiens]*. 2006.
- GenBank: AAH97290.1. *Oxidized low density lipoprotein (lectin-like) receptor 1 [Rattus norvegicus]*. 2006. <http://www.ncbi.nlm.nih.gov/protein/AAH97290.1>.
- GenBank: AB102861.1. *Homo sapiens OLR1 mRNA for oxidised low density lipoprotein (lectin-like) receptor 1, complete cds.* 7 de Agosto de 2003. <http://www.ncbi.nlm.nih.gov/nucleotide/AB102861.1>.
- GenBank: ABC40744.1. *oxidised low density lipoprotein (lectin-like) receptor 1 [Homo sapiens]*. 2005.
- GenBank: AF079164.1. *Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, exon 1.* 23 de Diciembre de 1998. <http://www.ncbi.nlm.nih.gov/nucleotide/AF079164.1>.

GenBank: AF079165.1. *Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, exon 2.*
23 de Diciembre de 1998. <http://www.ncbi.nlm.nih.gov/nuccore/AF079165.1>.

GenBank: AF079166.1. *Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, exon 3.*
23 de Diciembre de 1998. <http://www.ncbi.nlm.nih.gov/nuccore/AF079166.1>.

GenBank: AF079167.1. *Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, exons 4, 5, and 6, and complete cds.* 23 de Diciembre de 1998.
<http://www.ncbi.nlm.nih.gov/nuccore/AF079167.1>.

GenBank: AH007081.1. *Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, complete cds.* 23 de Diciembre de 1998.
<http://www.ncbi.nlm.nih.gov/nuccore/AH007081.1>.

GenBank: AJ131757.1. *Homo sapiens olr1 gene.* 14 de Noviembre de 2006.
<http://www.ncbi.nlm.nih.gov/nuccore/AJ131757>.

GenBank: BAC81565.1. *oxidised low density lipoprotein (lectin-like) receptor 1 [Homo sapiens].* 2003. <http://www.ncbi.nlm.nih.gov/protein/BAC81565.1>.

GenBank: BC022295.1. *Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1, mRNA (cDNA clone MGC:22491 IMAGE:4722086), complete cds.* 22 de Diciembre de 2006. <http://www.ncbi.nlm.nih.gov/nuccore/BC022295.1>.

GenBank: BC097290.1. *Rattus norvegicus oxidized low density lipoprotein (lectin-like) receptor 1, mRNA (cDNA clone MGC:114271 IMAGE:7458637), complete cds.* 17 de Julio de 2006. <http://www.ncbi.nlm.nih.gov/nuccore/BC097290.1>.

GenBank: CAB38175.1. *LOX1 [Homo sapiens].* 1999.
<http://www.ncbi.nlm.nih.gov/protein/CAB38175.1>.

GenBank: CM000214.2. *Mus musculus chromosome 6, whole genome shotgun sequence.* 23 de Marzo de 2015.
<http://www.ncbi.nlm.nih.gov/nuccore/CM000214.2?report=genbank&from=128207285&to=128231036>.

GenBank: CM000263.1. *Homo sapiens chromosome 12, whole genome shotgun sequence.* 15 de Marzo de 2015.

<http://www.ncbi.nlm.nih.gov/nuccore/CM000263.1?report=genbank&from=15471561&to=15485401>.

GenBank: DQ314885.1. *Homo sapiens oxidised low density lipoprotein (lectin-like) receptor 1 (OLR1) gene, complete cds.* 23 de Diciembre de 2005. <http://www.ncbi.nlm.nih.gov/nuccore/DQ314885.1>.

GenBank: EAW96156.1. *oxidised low density lipoprotein (lectin-like) receptor 1, isoform CRA_a [Homo sapiens].* 2005. <http://www.ncbi.nlm.nih.gov/protein/EAW96156.1>.

GenBank: EAW96157.1. *oxidised low density lipoprotein (lectin-like) receptor 1, isoform CRA_b [Homo sapiens].* 2005. <http://www.ncbi.nlm.nih.gov/protein/EAW96157.1>.

GenBank: EAW96158.1. *oxidised low density lipoprotein (lectin-like) receptor 1, isoform CRA_c [Homo sapiens].* 2005. <http://www.ncbi.nlm.nih.gov/protein/EAW96158.1>.

GenBank: EAW96159.1. *oxidised low density lipoprotein (lectin-like) receptor 1, isoform CRA_a [Homo sapiens].* 2005. <http://www.ncbi.nlm.nih.gov/protein/EAW96159.1>.

GenBank: EDK99928.1. *oxidized low density lipoprotein (lectin-like) receptor 1 [Mus musculus].* 2007. <http://www.ncbi.nlm.nih.gov/protein/EDK99928.1>.

GenBank: P78380.1. *RecName: Full=Oxidized low-density lipoprotein receptor 1; Short=Ox-LDL receptor 1; AltName: Full=C-type lectin domain family 8 member A; AltName: Full=Lectin-like oxidized LDL receptor 1; Short=LOX-1; Short=Lectin-like oxLDL receptor 1; Short=hLOX-1; Alt.* 2014. <http://www.ncbi.nlm.nih.gov/protein/p78380>.

Gene ID: 108078. *Olr1 oxidized low density lipoprotein (lectin-like) receptor 1 [Mus musculus (house mouse)].* 19 de Marzo de 2015. <http://www.ncbi.nlm.nih.gov/gene/108078>.

Gene ID: 140914. *Olr1 oxidized low density lipoprotein (lectin-like) receptor 1 [Rattus norvegicus (Norway rat)].* 5 de Abril de 2015. <http://www.ncbi.nlm.nih.gov/gene/140914>.

Gene ID: 4973. «OLR1 oxidized low density lipoprotein (lectin-like) receptor 1 [Homo sapiens (human)].» *National Center for Biotechnology Information, NCBI.* Abril de 2015. <http://www.ncbi.nlm.nih.gov/gene/4973>.

- Gonzales PA, Pisitkun T, Hoffert JD, Tchapyjnikov D, Star RA, Kleta R, Wang NS and Knepper MA. «Large-scale proteomics and phosphoproteomics of urinary exosomes.» *J. Am. Soc. Nephrol.* 20, nº 2 (2009): 363-379.
- González-Chavarría I., Cerro R., Parra N., Sandoval F., Zuñiga F., Omazábal V., Lamperti L., Jiménez S., Fernandez E., Gutiérrez N., Rodriguez F., Onate S., Sánchez O., Vera J., Toledo J. «Lectin-like oxidized LDL receptor-1 is an enhancer of tumor angiogenesis in human prostate cancer cells.» *PLOS ONE* 9, nº 8 (2014): e106219.
- Gregory, S. «Contig Assembly.» En *Encyclopedia of Life Sciences*. 2005.
- H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. «Protein Data Bank PDB.» *Nucleic Acids Research* 28 (2000): 235-242.
- Hirsch A., Iliopoulos D., Joshi A., Zhang Y., Jaeger S., Bulyk M., Tschlis P., Liu S., Strhl K. «A transcriptional signature and common gene networks link cancer with lipid metabolism and diverse human diseases.» *Cancer Cell* 4, nº 17 (Abril 2010): 348-361.
- Hu C, Dandapat A, Sun L, Khan J, Liu Y, Hermonat P, Mehta J. «Regulation of TGF β -1-mediated collagen formation by LOX-1: Studies base don forced overexpression of TGF β -1 in wild-type and LOX-1 knock-out mouse cardiac fibroblasts.» *J. Biol. Chem.* 283 (2008): 10226-10231.
- Itabe H., Obama T., Kato r. «The dynamics of oxidized LDL during atherogenesis.» *J Lipids*, 2011: 418313.
- Josefson, Deborah. «US company claims to have 90% of human genome on its database.» *BMJ* 320 (Enero 2000): 206.
- Kelley, L, Sternberg, M. «Protein structure prediction on the Web: a case study using the Phyre server.» *Nature Protocols* 4, nº 3 (2009): 363-371.
- Kostrzewa G., Broda G., Konarzewska M., Krajewki P., Ploski R. «Genetic polymorphism of human Y chromosome and risk factors for cardiovascular diseases: a study in WOBASZ cohort.» *PLoS ONE* 8, nº 7 (2013): e68155.
- Larkin, Blackshields, Brown, Chenna, McGettigan, McWilliam, Valentin, Wallace, Wilm, Lopez, Thompson, Gibson y Higgins. «Clustal W and Clustal X version 2.0.» *Bioinformatics* 23, nº 21 (2007): 2947-2948.

- Li H., Seng S., Avraham S., Avraham HK. «BRCA1 interacts with Smad3 and regulates Smad3-mediated TGF- β signaling during oxidative stress responses.» *PLoS ONE* 4, nº 9 (2009): e7091.
- Li Q., Nance M., Kulikauskas R., Nyberg K., Fehon R., Karplus A., Bretscher A., Tesmer J. «Self-masking in an intact ERM-nerlin protein: an active role for the central Alpha-helical domain.» *J Mol Biol.* 365, nº 5 (2007): 1446-1459.
- Lin M., Zhao L., Zhao W., Weng J. «Dissecting the mechanism of carotid atherosclerosis from the perspective of regulation.» *International Journal of Molecular Medicine* 34 (2014): 1458-1466.
- Liu H., Han H., Li J., Wong L. «DNAFSMiner: a web-based software toolbox to recognize two types of functional sites in DNA sequences.» *Bioinformatics* 21, nº 5 (2005): 671-673.
- Loots, G. «Genomic identification of regulatory elements by evolutionary sequence comparison and functional analysis.» *Advances in Genetics* 61 (2008): 269-292.
- Lu S., Nie J., Luan Q., Feng Q., Chang Z., Shan C., Hess D., Hemmings B., Yang Z. «Phosphorylation of the Twist1-Family Basic Helix-Loop-Helix transcription factors is involved in pathological cardiac remodeling.» *PLoS ONE* 6, nº 4 (2011): e19251.
- Lynch, M. «The origins of eukaryotic gene structure.» *Mol Biol Evol* 23, nº 2 (2006): 450-468.
- Madhusoodanan, Jyoti. «Expanding ENCODE.» *The Scientist Magazine*®, Agosto 2014.
- Makalowska I., Lin C., Makalowski W. «Overlapping genes in vertebrate genomes.» *Computational Biology and Chemistry*, nº 29 (2005): 1-12.
- Mango R, Clementi F, Borgiani P, Forleo G, Federici M, Contino G, Giardina E, Garza L, Fahdi I E. «Association of single nucleotide polymorphisms in the oxidised LDL receptor 1 (OLR1) gene in patients with acute myocardial infarction.» *J Med Genet* 40 (2003): 933-936.
- Mansouri A., Hallonet M., Gruss P. «Pax genes and their roles in cell differentiation and development.» *Cell Differentiation* 8 (1996): 851-857.

- Matsukura, S., Stellato, C., Plitt, J.R., Bickel, C. Miura, K. Georas, S. N., Casolaro, V., Schleimer, R. P.. «Activation of Eotaxin Gene Transcription by NF- κ B and STAT6 in Human Airway Epithelial Cells.» *J Immunol*, 1999: 6876-6883.
- Metha J, Chen J, Hermonat P, Romeo F, Novelli G. «Lectin-like, oxidized low-density lipoprotein receptor-1 (LOX-1): A critical player in the development of atherosclerosis and related disorders.» *Cardiovascular Research* 69 (2005): 36-45.
- Mewton N., Ying L., Liu Y., Croisille P., Bluemke D., Lima J. «Assessment of myocardial fibrosis with cardiac magnetic resonance.» *J Am Coll Cardiol.*, 2011: 891-903.
- Michener, C.D., Sokal, R.R. «A quantitative approach to a problem of classification.» *Evolution*, 1957: 11:490–499.
- Mikhailov A., Torrado M. «Review Article. In search of novel targets for heart disease: myocardin and myocardin-related transcriptional cofactors.» *Biochemistry Research International*, 2012: 1-11.
- Miller V., Best P. «Implications for reproductive medicine: Sex differences in cardiovascular disease.» *Sex Reprod Menopause* 9, n° 3 (2011): 21-28.
- Miller V., Kaplan J., Schork N., Ouyang P., Berga S., Wenger N., Shaw L., Webb C., Mallampalli M., Steiner M., Taylor D., Merz N., Reckelhoff J. «Strategies and methods to study sex differences in cardiovascular structure and function: a guide for basic scientists.» *Biology of Sex Differences* 2 (2011): 14-32.
- Misaka T., Suzuki S., Sakamoto N., Yamaki T., Sugimoto K., Kunii H., Nakazato K., Saitoh S., Sawamura T., Ishibashi T., Takeishi Y. «Significance of soluble lectin-like oxidized LDL receptor-1 levels in systemic and coronary circulation in acute coronary syndrome.» *BioMed Research International*, 2014.
- Misaka T., Suzuki S., Sakamoto N., Yamaki T., Sugimoto K., Kunii H., Nakazato K., Saitoh S., Sawamura T., Ishibashi T., Takeishi Y. «Significance of soluble Lectin-Like Oxidized LDL Receptor-1 Levels in Systemic and coronary circulation in acute coronary syndrome.» *BioMed Research International*, 2014.
- Mito S., Ozono R., Oshima T., Yano Y., Watari Y., Yamamoto Y., Brydun A., Igarashi K., Yoshizumi M. «Myocardial protection against pressure overload in mice lacking

- Bach1, a transcriptional repressor of Heme Oxygenase-1.» *Hypertension* 51 (2008): 1570-1577.
- MSN Gene. *GeneCards*. 25 de Mayo de 2015. <http://www.genecards.org/cgi-bin/carddisp.pl?gene=MSN&search=11a863c4b77e9b12aeaaa8f9c9ee531b>.
- Murthy K., Manley J. «The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'end formation.» *Genes & Dev.*, 1995: 2672-2683.
- Nagase M, Abe J, Takahashi K, Ando J, Hirose S, Fujita T. «Genomic Organization and Regulation of Expression of the Lectin-like Oxidized Low-density Lipoprotein Receptor (LOX-1) Gene.» *J. Biol. Chem.* 273 (1998): 33702-33707.
- Nakano S., Sugihara M., Yamada R., Katayanagi K., Tate S. «Structural implication for the impaired binding of W150A mutant LOX-1 to oxidized low density lipoprotein, oxLDL.» *Biochimica et Biophysica Acta* 1824 (2012): 739-749.
- NCBI National Center for Biotechnology Information. *OLR1 oxidized low density lipoprotein (lectin-like) receptor 1 [Homo sapiens (human)]*. 5 de Abril de 2015. <http://www.ncbi.nlm.nih.gov/gene/4973>.
- NCBI. *OLR1 oxidized low density lipoprotein (lectin-like) receptor 1 [Homo sapiens (human)]*. 22 de Agosto de 2014. <http://www.ncbi.nlm.nih.gov/gene/4973>.
- . *Olr1 oxidized low density lipoprotein (lectin-like) receptor 1 [Rattus norvegicus (Norway rat)]*, *Gene ID: 140914*. 22 de Agosto de 2014. <http://www.ncbi.nlm.nih.gov/gene?cmd=Search&doptcmdl=EntrezGene&term=140914>.
- NCBI Reference Sequence: AC_000028.1. *Mus musculus strain mixed chromosome 6, alternate assembly Mm_Celera, whole genome shotgun sequence*. 9 de Febrero de 2015. http://www.ncbi.nlm.nih.gov/nuccore/AC_000028.1?report=genbank&from=131207291&to=131231036.
- NCBI Reference Sequence: AC_000072.1. *Rattus norvegicus strain BN; Sprague-Dawley chromosome 4, alternate assembly Rn_Celera, whole genome shotgun sequence*. 7 de Agosto de 2014.

http://www.ncbi.nlm.nih.gov/nuccore/AC_000072.1?report=genbank&from=151610779&to=151632854.

NCBI Reference Sequence: AC_000144.1. *Homo sapiens chromosome 12, alternate assembly HuRef, whole genome shotgun sequence*. 3 de Febrero de 2014. http://www.ncbi.nlm.nih.gov/nuccore/NC_000012.12?report=genbank&from=10158300&to=10176261.

NCBI Reference Sequence: NC_000012.12. *Homo sapiens chromosome 12, GRCh38.p2 Primary Assembly*. 12 de Marzo de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NC_000012.12?report=genbank&from=10158300&to=10176261.

NCBI Reference Sequence: NC_000072.6. *Mus musculus strain C57BL/6J chromosome 6, GRCm38.p3 C57BL/6J*. 9 de Febrero de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NC_000072.6?report=genbank&from=129485247&to=129507165.

NCBI Reference Sequence: NC_005103.4. *Rattus norvegicus strain mixed chromosome 4, Rnor_6.0*. 7 de Agosto de 2014. http://www.ncbi.nlm.nih.gov/nuccore/NC_005103.4?report=genbank&from=163239853&to=163261937.

NCBI Reference Sequence: NC_018923.2. *Homo sapiens chromosome 12, alternate assembly CHM1_1.1, whole genome shotgun sequence*. 12 de Marzo de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NC_018923.2?report=genbank&from=10280199&to=10294080.

NCBI Reference Sequence: NG_016743.1. *Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), RefSeqGene on chromosome 12*. 14 de Mayo de 2014. http://www.ncbi.nlm.nih.gov/nuccore/NG_016743.

NCBI Reference Sequence: NM_001172632.1. *Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant 2, mRNA*. 15 de Marzo de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NM_001172632.1.

NCBI Reference Sequence: NM_001172633.1. *Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant 3, mRNA*. 15 de Marzo de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NM_001172633.1.

NCBI Reference Sequence: NM_001301094.1. *Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant 3, mRNA*. 15 de Febrero de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NM_001301094.1.

NCBI Reference Sequence: NM_001301096.1. *Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant 2, mRNA*. 15 de Febrero de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NM_001301096.1.

NCBI Reference Sequence: NM_002543.3. *Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant 1, mRNA*. 15 de Marzo de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NM_002543.

—. *Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant 1, mRNA*. 15 de Marzo de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NM_002543.3.

NCBI Reference Sequence: NM_133306.2. *Rattus norvegicus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), mRNA*. 4 de Septiembre de 2014. http://www.ncbi.nlm.nih.gov/nuccore/NM_133306.2.

NCBI Reference Sequence: NM_138648.2. *Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant 1, mRNA*. 15 de Febrero de 2015. http://www.ncbi.nlm.nih.gov/nuccore/NM_138648.2.

NCBI Reference Sequence: NP_001166103.1. *oxidized low-density lipoprotein receptor 1 isoform 2 [Homo sapiens]*. 2014. http://www.ncbi.nlm.nih.gov/protein/NP_001166103.1.

NCBI Reference Sequence: NP_001166104.1. *oxidized low-density lipoprotein receptor 1 isoform 3 [Homo sapiens]*. 2010. http://www.ncbi.nlm.nih.gov/protein/NP_001166104.1.

NCBI Reference Sequence: NP_001288023.1. *oxidized low-density lipoprotein receptor 1 isoform 3 [Mus musculus]*. 2014.
http://www.ncbi.nlm.nih.gov/protein/NP_001288023.1.

NCBI Reference Sequence: NP_001288025.1. *oxidized low-density lipoprotein receptor 1 isoform 2 [Mus musculus]*. 2014.
http://www.ncbi.nlm.nih.gov/protein/NP_001288025.1.

NCBI Reference Sequence: NP_002534.1. *oxidized low-density lipoprotein receptor 1 isoform 1 [Homo sapiens]*. 1999.
http://www.ncbi.nlm.nih.gov/protein/NP_002534.1.

NCBI Reference Sequence: NP_579840.2. *oxidized low-density lipoprotein receptor 1 [Rattus norvegicus]*. 2002. http://www.ncbi.nlm.nih.gov/protein/NP_579840.2.

NCBI Reference Sequence: NP_619589.2. *oxidized low-density lipoprotein receptor 1 isoform 1 [Mus musculus]*. 2001.
http://www.ncbi.nlm.nih.gov/protein/NP_619589.2.

NCBI Reference Sequence: XM_006505303.1. *PREDICTED: Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant X1, mRNA*. 27 de Diciembre de 2013.
http://www.ncbi.nlm.nih.gov/nucleotide/XM_006505303.1?report=genbank.

NCBI Reference Sequence: XM_006505304.1. *PREDICTED: Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant X2, mRNA*. 27 de Diciembre de 2013.
http://www.ncbi.nlm.nih.gov/nucleotide/XM_006505304.1?report=genbank.

NCBI Reference Sequence: XM_006719081.1. *PREDICTED: Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant X1, mRNA*. 3 de Febrero de 2014.
http://www.ncbi.nlm.nih.gov/nucleotide/XM_006719081.1?report=genbank.

NCBI Reference Sequence: XP_006505366.1. *PREDICTED: oxidized low-density lipoprotein receptor 1 isoform X1 [Mus musculus]*. 2013.
http://www.ncbi.nlm.nih.gov/protein/XP_006505366.1?report=genpept.

- NCBI Reference Sequence: XP_006505367.1. *PREDICTED: oxidized low-density lipoprotein receptor 1 isoform X2 [Mus musculus]*. 2013.
http://www.ncbi.nlm.nih.gov/protein/XP_006505367.1?report=genpept.
- NCBI Reference Sequence: XP_006719144.1. *PREDICTED: oxidized low-density lipoprotein receptor 1 isoform X1 [Homo sapiens]*. 2000.
http://www.ncbi.nlm.nih.gov/protein/XP_006719144.1?report=genpept.
- NCBI, Gene ID: 108078. *Olr1 oxidized low density lipoprotein (lectin-like) receptor 1 [Mus musculus (house mouse)]*. Gene ID: 108078. 22 de Agosto de 2014.
<http://www.ncbi.nlm.nih.gov/gene?cmd=Search&doptcmdl=EntrezGene&term=108078>.
- Ogura S., Kakino A., Sato Y., Fujita Y., Iwamoto S., Otsui K., Yoshimoto R., Sawamura T. «LOX-1 The multifunctional receptor underlying cardiovascular dysfunction.» *Curc J*, 2009: 1993-1999.
- Ohki I., Ishigaki T., Oyama T., Matsunaga S., Xie Q., Ohnishi-Kameyama M., Murata T., Tsuchiya D., Machida S., Morikawa K., Tate S. «Crystal structure of human lectin-like, oxidized low-density lipoprotein receptor 1 ligand binding domain and its ligand recognition mode to oxLDL.» *Structure* 13 (2005): 905-917.
- OLR1 Gene, GeneCards. *OLR1 Gene*. 2015. <http://www.genecards.org/cgi-bin/carddisp.pl?gene=OLR1&search=a2c4de52dc9d26c1839c0a6453efbca8>.
- OPS, ORGANIZACIÓN PANAMERICANA DE LA SALUD. *Cardiovascular Diseases (CVD) mortality by subgroups of causes*. 27 de Abril de 2014.
http://www.paho.org/hq/index.php?option=com_content&view=article&id=6682&Itemid=2391&lang=es.
- . *Leading causes of death*. 27 de Abril de 2014.
http://www.paho.org/hq/index.php?option=com_content&view=article&id=3501&Itemid=2391&lang=es.
- Oxford Dictionaries. *In silico - definicion*. 1980.
<http://www.oxforddictionaries.com/es/definicion/ingles/in-silico> (último acceso: Julio de 2015).

- . *In situ - definicion*. 2015. <http://www.oxforddictionaries.com/es/definicion/ingles/in-situ?q=in+situ> (último acceso: Julio de 2015).
- Park H., Adsit F., Boyington J. «The 1.4A Crystal structure of the human oxidized low density lipoprotein receptor Lox-1.» *The Journal of Biological Chemistry* 280, nº 14 (2005): 13593-13599.
- Pennacchio, L. A., Bickmore, W., Dean, A., Nobrega, M. A., Bejerano, G. «Enhancers: Five essential questions.» *Nature Reviews Genetics* 14, nº 4 (2013): 288-295.
- PrabhuDas M., Bowdish D., Drickamer K., Febbraio M., Herz J., Kobzik L., Krieger M., Loike J., Means T., Moestrup S., Post S., Sawamura T., Silverstein S., Wang X., Khoury J. «Standardizing Scavenger Receptor Nomenclature.» *J Immunol* 195, nº 5 (2014): 1997-2006.
- RAE, Real Academia Española. *Hardware*. 2014. <http://lema.rae.es/drae/?val=hardware> (último acceso: Julio de 2015).
- . *Software*. 2014. <http://lema.rae.es/drae/?val=software> (último acceso: Julio de 2015).
- Sabath N., Graur D. «Detection of functional Overlapping Genes. Simulation and Case Studies.» *J Mol Evol* 71, nº 4 (Octubre 2010): 308-316.
- Saitou N., Nei M. «The neighbor-joining method: a new method for reconstructing phylogenetic trees.» *Mol Biol Evol*, 1987: 406-425.
- Sanna C., Li W., Zhang L. «Overlapping genes in the human mouse genomes.» *BMC Genomics* 9, nº 169 (Abril 2008).
- Science Daily. *Natural killer cell*. 2015. http://www.sciencedaily.com/terms/natural_killer_cell.htm (último acceso: 2015).
- Smoczynski, Chandra Shekhar Pareek & Rafal. «Sequencing technologies and genome sequencing.» *J Appl Genetics* 52 (2011): 413-435.
- Soler-Cantero A, Jové M, Cacabelos D, Boada J, Naudí A, et al. «Plant-derived phenolics inhibit the accrual of structurally characterised protein and lipid oxidativa modifications.» *PLoS One* 7, nº 8 (2012): e43308.

- Tamura K, Stecher G, Peterson D, Filipinski A, and Kumar S. «MEGA6: Molecular Evolutionary Genetics Analysis version 6.0.» *Molecular Biology and Evolution* 30 (2013): 2725-2729.
- Taye A, El-Sheikh A. «Lectin-like oxidized low-density lipoprotein receptor 1 pathways.» *Eur J Clin Invest* 43, nº 7 (2013): 740-745.
- The GeneCards Human Gene Database. *Crown Human Genome Center, Department of Molecular Genetics, the Weizmann Institute of Science* . 22 de Agosto de 2014. <http://www.genecards.org/cgi-bin/carddisp.pl?gene=OLR1&snp=446&search=olr1%20olr-1#snp>.
- The UniProt Consortium. «UniProt: a hub for protein information.» *Nucleic Acids Res* 43 (2015): D204-D212 .
- UniProtKB/Swiss-Prot: O70156.1. *RecName: Full=Oxidized low-density lipoprotein receptor 1; Short=Ox-LDL receptor 1; AltName: Full=Lectin-like oxidized LDL receptor 1; Short=LOX-1; Short=Lectin-like oxLDL receptor 1; AltName: Full=Lectin-type oxidized LDL receptor 1; Contains: RecName: F. 1998.* <http://www.ncbi.nlm.nih.gov/protein/O70156>.
- UniProtKB/Swiss-Prot: Q9EQ09.2. *RecName: Full=Oxidized low-density lipoprotein receptor 1; Short=Ox-LDL receptor 1; AltName: Full=Lectin-like oxidized LDL receptor 1; Short=LOX-1; Short=Lectin-like oxLDL receptor 1; AltName: Full=Lectin-type oxidized LDL receptor 1; Contains: RecName: F. 2005.* <http://www.ncbi.nlm.nih.gov/protein/Q9EQ09.2>.
- Verhoye, E., y M. Langlois. «Review: Circulating oxidized low-density lipoprotein: a biomarker of atherosclerosis and cardiovascular risk?» *Clin. Chem. Lab. Med* 47, nº 2 (2009): 128-137.
- Wang X, Khaidakov M, Ding Z, Dai Y, Mercanti F, Mehta J. «. LOX-1 in the maintenance of cytoskeleton and proliferation in senescent cardiac fibroblast.» *Journal of Molecular and Cellular Cardiology* 60 (2013): 184-190.
- Wray, G. «The evolutionary significance of cis-regulatory mutations.» *Nature Reviews Genetics* 8 (2007): 206 -216.

- Yamey, Gavin. «Scientists unveil first draft of human genome.» *BMJ* 321 (2000): 7.
- Yan M, Metha J, Zhang W, Hu C. «LOX-1, oxidative stress and inflammation: A novel mechanism for diabetic cardiovascular complications.» *Cardiovasc Drugs Ther.* 25, nº 5 (2011): 451-459.
- Yoshimoto R, Fujita Y, Kakino A, Iwamoto S, Takaya T, Sawamura T. «The Discovery of LOX-1, its ligands and clinical significance.» *Cardiovasc Drugs Ther* 25 (2011): 379-391.
- You L., Wu J., Feng Y., Fu Y., Guo Y., Long L., Zhang H., Luan Y., Tian P., Chen L., Huang G., Huang S., Li Y., Li J., Chen C., Zhang Y., Chen S., Xu A. «APASdb: a database describing alternative poly(A) sites and selection of heterogeneous cleavage sites downstream of poly(A) signals.» *Nucleic Acids Research*, 2015: D59-D67.
- Zhu F., Li Y., Zhang J., Piao C., Liu T., Li H., Du J. «Senescent cardiac fibroblast is critical for cardiac fibrosis after myocardial infarction.» *PLoS ONE* 8, nº 9 (2013): e74535.
- Zhu H, Xia M, hou M, Tang Z, Li Y, Ma J, Ling W. «Ox-LDL plays a dual effect in modulating expression of inflammatory molecules trough LOX-1 pathways in human umbilical vein endothelial cells.» *Front. Biosci.* 1, nº 10 (2005): 2585-2594.

9. ANEXOS

Tabla 13. Información de secuencias de gen OLR1 para humano, ratón y rata, relativas a base de datos NCBI.

Organismo	GenBank (NCBI Reference Sequence)	Definición	Accesion	Version	GI	GenBank Division
Hs	NG_016743.1	<i>Homo sapiens</i> oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), RefSeq Gene on chromosome 12.	NG_016743	NG_016743.1	291084732	CON
	AJ15151.1	<i>Homo sapiens</i> Olr1 gene	AJ15151	AJ15151.1	4468343	PR1
	CM00263.1	<i>Homo sapiens</i> chromosome 12, whole genome shotgun sequence.	U15471561 REGION: 15471561..15485401	CM00263.1	74273670	CON
	AC_000144.1	<i>Homo sapiens</i> chromosome 12, alternate assembly HUKet, whole genome shotgun sequence.	AC_U0144 REGION: 10054056..10067944	AC_000144.1	157704453	CON
	NC_018923.2	<i>Homo sapiens</i> chromosome 12, alternate assembly CHM11_1.1, whole genome shotgun sequence.	NC_018923 REGION: 10280199..10294080 GPC_000001171	NC_018923.2	528476594	CON
	NC_000012.12	<i>Homo sapiens</i> chromosome 12, GRCh38 Primary Assembly.	NC_000012 REGION: 10158300..10172191 GPC_000001304	NC_000012.12	568815586	CON
Mm	NC_000072.6	<i>Mus musculus</i> strain C57BL/6J chromosome 6, GRCh38p2 C57BL/6J.	NC_000072 REGION: 129485247..129507355 GPC_00000079	NC_000072.6	372099104	CON
	AC_000028.1	<i>Mus musculus</i> strain mixed chromosome 6, alternate assembly Mn_Celera, whole genome shotgun sequence.	AC_U0028 REGION: 131207291..131231036	AC_000028.1	83280979	CON
	CM00214.2	<i>Mus musculus</i> chromosome 6, whole genome shotgun sequence.	U128207285 REGION: 128207285..128231036	CM00214.2	74229906	CON
Rn	NC_005103.4	<i>Rattus norvegicus</i> strain mixed chromosome 4, Rnor_60.	NC_005103 REGION: 163239853..163261937 GPC_000001525	NC_005103.4	666184249	CON
	AC_000072.1	<i>Rattus norvegicus</i> strain BN, Sprague-Lawley chromosome 4, alternate assembly Rn_Celera, whole genome shotgun sequence.	AC_000072	AC_000072.1	109652486	CON

Contenido de cada columna de izquierda a derecha: especie; nombre de la secuencia en GenBank; definición presente en la base de datos NCBI; identificador único de la secuencia, o accession; versión de la secuencia, indicada como "secuencia.n" donde n es la versión de la secuencia; código GI (Gene Info Identifier, nomenclatura usada antiguamente por GenBank); y la división de GenBank a la que pertenece la secuencia. Hs: Homo sapiens; Mm: Mus musculus; Rn: Rattus norvegicus; CON: secuencias tipo "Contig" (correspondientes a ensamblajes de secuencias o mosómicos); PR1: secuencias correspondientes a primates.

Tabla 14. Información de secuencias de gen OLRI para humano, ratón y rata, relativas a base de datos NCBI y quienes las originaron.

Oganismo	GenBank (NCBI Reference Sequence)	Definición	Solicitud Anotación Directa (Direct submission)	Fecha primera versión anotación NCBI	Fecha última actualización NCBI	Assembly Name	Nº Referencias
Hs	NG_016743.1	Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLRI), RefSeq Gene on chromosome 12.	No reportado	16-03-2010	14-05-2014	RefSeqGene	5
	AI131757.1	Homo sapiens olri gene	REFERENCE 2 (bases 1 to 12196), AUTHORS Millar DS., TITLE Direct Submission, JOURNAL Submitted (16-FEB-1999) Millar DS., Institute of Medical Genetics, University of Wales College of Medicine, Heath Park, Cardiff, CF44 3N, UNITED KINGDOM	02-03-1999 (ENA)	14-11-2006	No reportado	2
	CM00263.1	Homo sapiens chromosome 12, whole genome shotgun sequence.	REFERENCE 2 (bases 1 to 133516800), AUTHORS Istrail y cols., TITLE Direct Submission, JOURNAL Submitted (25-NOV-2003) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA REFERENCE 3 (bases 1 to 133516800), AUTHORS Mural R.J. y cols., TITLE Direct Submission, JOURNAL Submitted (05-JUL-2005) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA	08-09-2005	29-06-2009	Celera Genomics WGS human genome sequencing project	3
	AC_000144.1	Homo sapiens chromosome 12, alternate assembly HUKet, whole genome shotgun sequence.	No reportado	27-12-2007	03-02-2014	HRef	1
	NC_018923.2	Homo sapiens chromosome 12, alternate assembly CHMI_1.1, whole genome shotgun sequence.	No reportado	19-10-2012	03-02-2014	CHMI_1.1	0
	NC_000012.12	Homo sapiens chromosome 12, GRCh38 Primary Assembly.	No reportado	03-02-2014	03-02-2014	GRCh38 Primary Assembly	2
	NC_000072.6	MUS MUSCULUS strain C57BL/6J chromosome 6, GRCh38p2C57BL6J.	No reportado	23-02-2012	27-DIC2013	GRCh38p2C57BL6J	2
Mm	AC_000028.1	MUS MUSCULUS strain mixed chromosome 6, alternate assembly Mm_Celera, whole genome shotgun sequence.	No reportado	05-12-2005	27-DIC2014	Mm_Celera	1
	CM002142	Mus musculus chromosome 6, whole genome shotgun sequence.	REFERENCE 2 (bases 1 to 23752), AUTHORS Mural R.J., Adams MD., Myers EW., Smith HO. and Venter J.C., TITLE Direct Submission, JOURNAL Submitted (05-JUL-2005) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA REFERENCE 3 (bases 1 to 23752), AUTHORS Mural R.J., Adams MD., Myers EW., Smith HO. and Venter J.C., TITLE Direct Submission, JOURNAL Submitted (02-SEP-2005) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA	07-09-2005	17-04-2007	No reportado	2
Rn	NC_010105.4	Rattus norvegicus strain mixed chromosome 4, Rnor_60.	No reportado	07-08-2014	07-08-2014	Rnor_60	1
	AC_000072.1	Rattus norvegicus strain BN; Sprague-Jawley chromosome 4, alternate assembly Rn_Celera, whole genome shotgun sequence.	No reportado	02-06-2006	07-08-2014	Rn_Celera	0

Contenido de cada columna de izquierda a derecha: Especie; nombre de la secuencia en GenBank; definición presente en la base de datos NCBI; solicitud de publicación directa, se indica el número de referencia, título, fecha y autores de quienes solicitaron la publicación de la secuencia, cuando fue reportado; fecha de la primera versión de la anotación de la secuencia en NCBI; fecha de la última actualización de la anotación de la secuencia del gen en NCBI; assembly name, correspondiente a la organización que secuenció la anotación publicada en NCBI, según corresponda; y la cantidad de referencias que han citado dicha secuencia según NCBI. Hs: Homo sapiens; Mm: Mus musculus; Rn: Rattus norvegicus

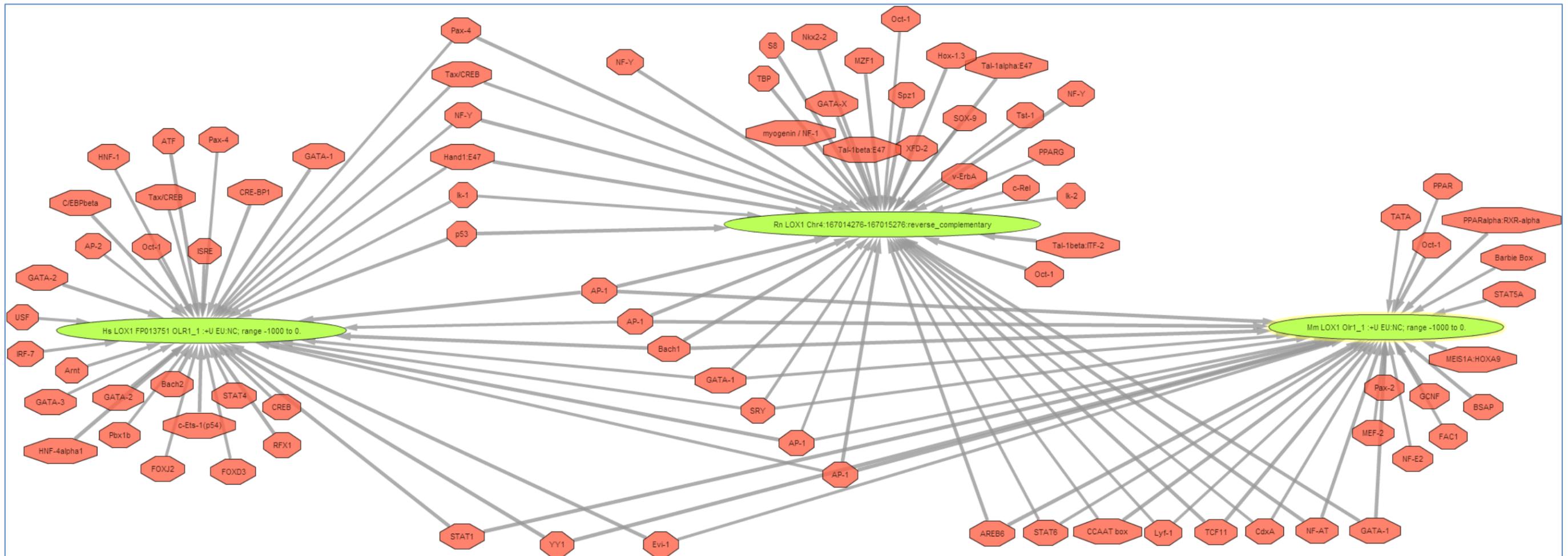


Figura 37. Redes de regulación transcripcional para OLR1 comunes en humano, ratón y rata. En verde se muestra el gen OLR1 y la especie, mientras que en rojo se muestran los factores de transcripción hipotéticos que regularían la expresión de OLR1. TFs que solo regulan una especie se muestran con una sola flecha en dirección a OLR1 en la especie correspondiente; TFs que regulan dos especies se muestran con dos flechas en dirección a OLR1 de las especies que corresponden; mientras que TF que regulan a las 3 especies muestran 3 flechas apuntando a cada diagrama de OLR1 en verde.

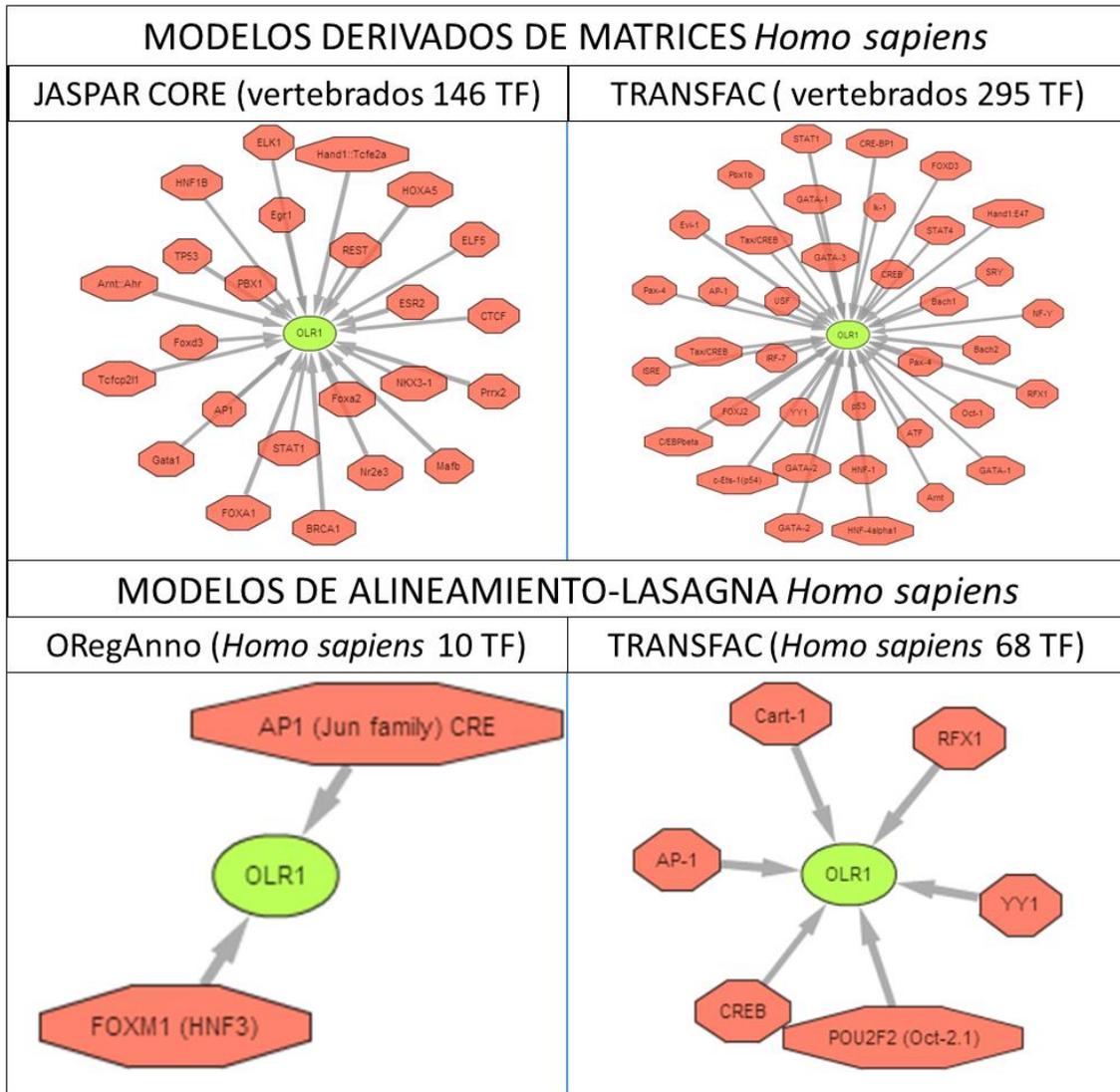


Figura 38. Redes de regulación transcripcional para OLR1 en humano. En la porción superior de cada imagen se especifica la estrategia de búsqueda, modelo, base de datos y cantidad de Factores de Transcripción (TF) utilizados. En verde se muestra el gen OLR1, mientras que en rojo se muestran los factores de transcripción hipotéticos que regularían la expresión de OLR1.

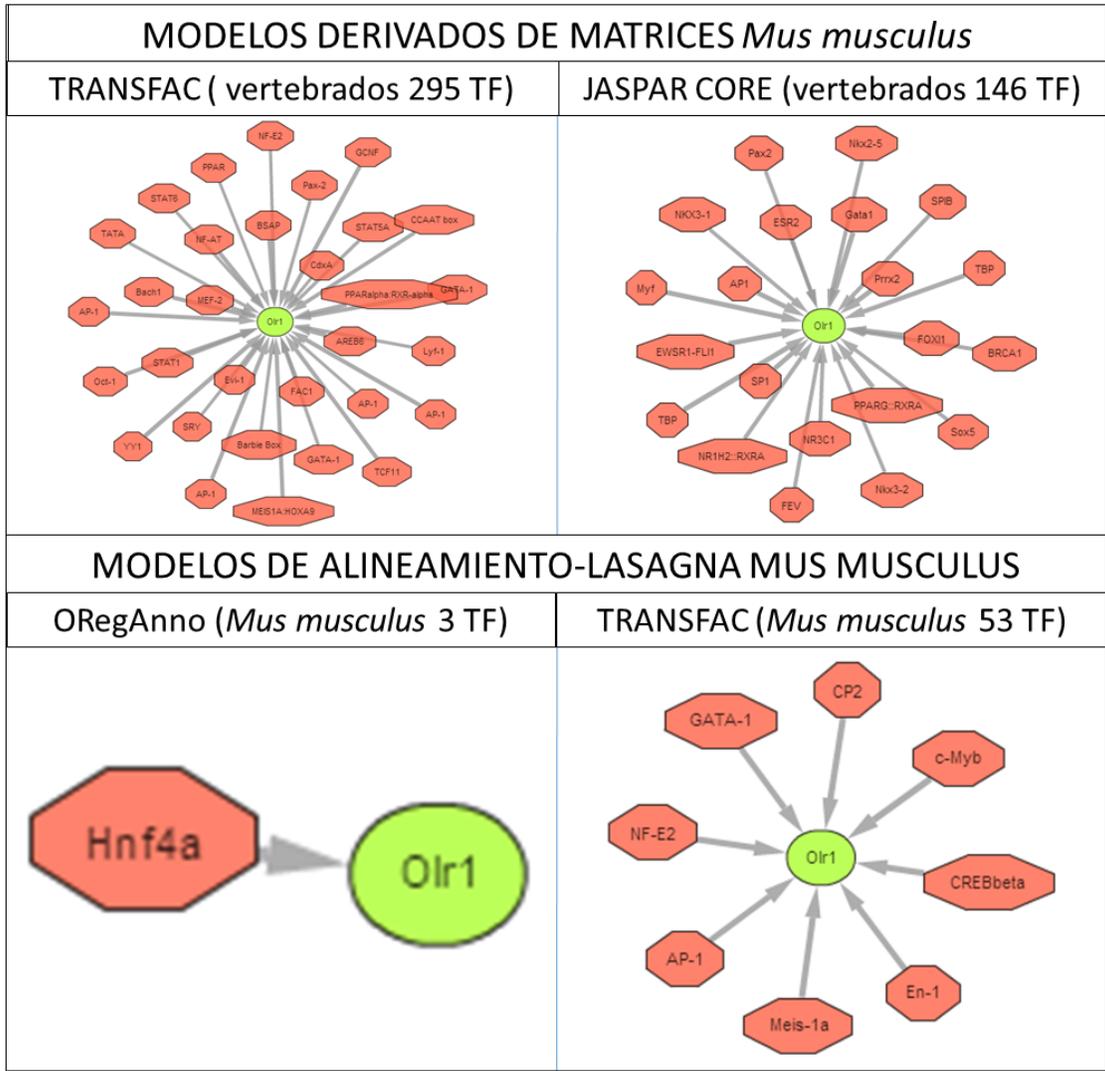


Figura 39. Redes de regulación transcripcional para OLR1 en ratón. En la porción superior de cada imagen, se especifica la estrategia de búsqueda, modelo, base de datos y cantidad de Factores de Transcripción (TF) utilizados. En verde se muestra el gen OLR1, mientras que en rojo se muestran los factores de transcripción hipotéticos que regularían la expresión de OLR1.

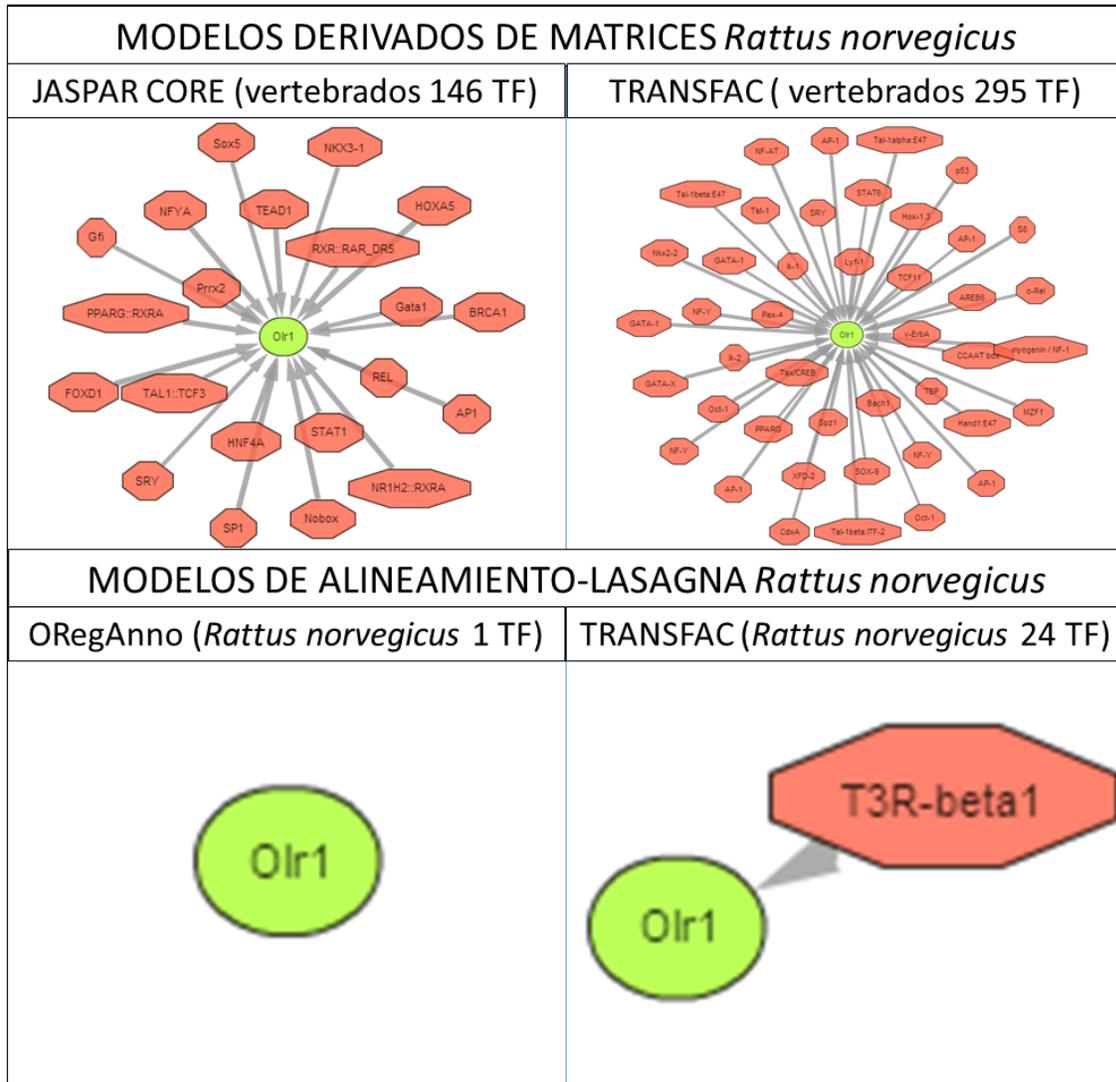


Figura 40. Redes de regulación transcripcional para OLR1 en rata. En la porción superior de cada imagen se especifica la estrategia de búsqueda, modelo, base de datos y cantidad de Factores de Transcripción (TF) utilizados. En verde se muestra el gen OLR1, mientras que en rojo se muestran los factores de transcripción hipotéticos que regularían la expresión de OLR1.

Tabla 15. Factores de Transcripción a humano, ratón y rata.

Nombre	Secuencia	Posición (0-based)	Hebra	p-Value	Modelo	Especie	Descripción	Desorden/Enfermedad					
AP1 (MA0099.2)	TGACTCA	11	-	0	JASPAR CORE	<i>Hs</i>	Posee actividad como co-activador de la transcripción y como factor de transcripción con unión específica a ciertas secuencias de ADN	Tumor; necrosis; leucemia de células T; inflamación; cáncer; leucemia de célula mieloides; transformación neoplásica; cáncer de mama; leucemia promielocítica; cáncer de colon.					
		935	-	0		<i>Mm</i>							
		940	-	0		<i>Rn</i>							
		970	-	0		<i>Hs</i>							
AP-1 (M00174)	GATGACTCAAC	933	-	0.000075	TRANSFAC	<i>Hs</i>							
	AATGACTCAAT	939	-	1		<i>Mm</i>							
AP-1 (M00188)	AATGACTCAAT	969	-	1		<i>Rn</i>				<i>Hs</i>			
	GATGACTCAAC	933	-	0.000025		<i>Mm</i>							
		938	-	0.0001		<i>Rn</i>							
AP-1 (M00199)	ATGACTCA	935	-	0.000025		<i>Hs</i>				<i>Mm</i>			
		10	+	0.000025									
		940	-	0.000025									
AP-1 (M00517)	GGTTGAGTCATCC	970	-	0.000025		<i>Rn</i>				<i>Hs</i>			
		932	+	0.000075									
		938	+	1							<i>Mm</i>		
Bach1 (M00495)	GACATGAGTCAATCA	968	+	1		TRANSFAC				<i>Rn</i>	Regulador negativo de la transcripción; unión a proteínas; unión a ADN; potenciador distal de ARN polimerasa II; actividad como factor de transcripción de secuencias específicas de ADN	Monosomía; eritroleucemia	
	CGGTTGAGTCATCCT	7	+	0			<i>Hs</i>						
		936	+	0			<i>Mm</i>						
BRCA1 (MA0133.1)	TCAACTG	966	+	0		JASPAR CORE	<i>Rn</i>	Unión a ADN dañado; Unión a ARN; unión a ADN; Actividad coactivadora de la transcripción; Contribuye a la actividad ubiquitinadora de proteínas; unión a regiones reguladoras de las transcripción en ADN.		Cáncer de ovario; cáncer de mama; mutaciones de líneas germinales; cáncer de mama bilateral; cáncer; carcinoma de mama; anemia de Danconis; tumores; carcinoma de ovario;			
		57	+	0							<i>Hs</i>		
		936	-	0.0001	<i>Mm</i>								
Gata1 (MA0035.2)	CCTGATAACAT	966	-	0.0001	JASPAR CORE	<i>Hs</i>	Unión a ADN; potenciador distal de ARN polimerasa II; actividad como factor de transcripción de secuencias específicas de ADN	Leucemia megacarioblastica aguda; anemia diseritropoyetica; eritroleucemia; síndrome de Down; desordenes mieloproliferativos; trombocitopenia 1; leucogenesis; leucemia aguda					
	TCAGATAAGGG	308	-	0.0001					<i>Mm</i>				
		249	-	0.0001					<i>Rn</i>				
GATA-1 (M00347)	CCTGATAACA	771	-	5.00E-05	TRANSFAC	<i>Hs</i>	Factor transcripcional que puede actuar como represor de la transcripción. Actúa como controlador de la supresor de tumores en carcinogénesis de próstata	Cáncer de próstata; neoplasia intraepitelial prostática; carcinoma y adenocarcinoma de próstata; tumores; hiperplasia benigna de próstata; tumor primario; metástasis					
		309	-	0.000075					<i>Mm</i>				
		250	-	0.000075					<i>Rn</i>				
NKX3-1 (MA0124.1)	GTACTTG	984	+	0.000075	JASPAR CORE	<i>Hs</i>	Factor de transcripción miembro de la familia perteneciente a la caja del Grupo de Alta movilidad (HMG)-box de proteínas que unen ADN. Esta proteína es un factor determinante de testículos (TDF), los cuales inician la determinación sexual masculina. También es capaz de unir Calmodulina	Síndrome de Swyer; hermafroditismo; gonadoblastoma; amenorrea primaria; displasia campomeliaca; digénesis gonadal mixta; síndrome de Turner; virilización; síndrome de regresión testicular					
	ATACTTA	120	-	7.50E-05					<i>Mm</i>				
		195	-	0.0001					<i>Rn</i>				
		440	-	0.0001					<i>Hs</i>				
Prrx2 (MA0075.1)	CATTAA	446	-	0.0001	JASPAR CORE	<i>Mm</i>	Proteína asociada a ADN, miembro de la familia de proteínas homeobox. Presenta actividad de unión a ADN						
		323	+	0					<i>Hs</i>				
		39	+	0									
		632	-	0									
		264	+	0									
		SRY (M00148)	AATGTGC	373					+	0	TRANSFAC	<i>Rn</i>	
				584					+	0			
				332					-	0			
				301					+	0			
				366					-	0			
74	+			0									
685	-	0.0001	<i>Hs</i>										
641	+	0.00005	<i>Mm</i>										
720	+	0.0001	<i>Rn</i>										
748	+	0.0001	<i>Hs</i>										

Datos obtenidos usando Lasagna Search, modelos de matrices, bases de datos TRANSFAC y JASPAR. Se indica, en columnas de izquierda a derecha, nombre del factor transcripcional y el código de Logo; Secuencia de ADN a la que se une el TF; Posición del promotor en que se presenta el TFBS; Hebra de ADN en la que está presente el TFBS; p-Value arrojado por Lasagna Search para dicho TFBS en el promotor; Modelo de Matriz utilizado (TRANSFAC o JASPAR CORE); Especie en la que está presente el TFBS; Descripción del TF; Desorden o enfermedades asociadas a dicho TF. *Hs*: Homo sapiens; *Mm*: Mus musculus; *Rn*: Rattus norvegicus.

Tabla 16. Información de secuencias de ARNm de OLR1 para humano, ratón y rata, relativas a base de datos NCBI.

Organismo	GenBank (NCBI Reference Sequence)	Locus	Accession	Version	GI	GenBank Division
<i>Hs</i>	BC022295.1	BC022295	BC022295	BC022295.1	18490152	PRI
	AB102861.1	AB10261	AB102861	AB102861.1	33468735	PRI
	AF079164.1/AH007081.1	HSOLR1	AF079164	AF079164.1	4049999	PRI
	AF079165.1	HSOLR2	AF079165	AF079165.1	4050000	PRI
	AF079166.1	HSOLR3	AF079166	AF079166.1	4050001	PRI
	AF079167.1	HSOLR4	AF079167	AF079167.1	4050002	PRI
	DQ314885.1	DQ31485	DQ314885	DQ314885.1	83699676	PRI
	NM_002543.3	NP_002534	NM_002543	NM_002543.3	119392084	PRI
	NM_001172632.1	NM_001172632	NM_001172632	NM_001172632.1	290654341	PRI
	NM_001172633.1	NM_001172633	NM_001172633	NM_001172633.1	290654343	PRI
XM_006719081.1*	XM_006719081	XM_006719081	XM_006719081.1	578822878	PRI	
<i>Mm</i>	NM_138648.2	NM_138648	NM_138648 XM_984831	NM_138648.2	134053880	ROD
	NM_001301094.1	NM_001301094	NM_001301094	NM_001301094.1	666875832	ROD
	NM_001301096.1	NM_001301096	NM_001301096	NM_001301096.1	666875862	ROD
	XM_006505303.1*	XM_006505303	XM_006505303	XM_006505303.1	568940171	ROD
	XM_006505304.1*	XM_006505304	XM_006505304	XM_006505304.1	568940173	ROD
<i>Rn</i>	NM_133306.2	NM_133307	NM_133308	NM_133306.2	404247469	ROD
	BC097290.1	BC097291	BC097292	BC097290.1	67678320	ROD

Contenido de cada columna de izquierda a derecha: especie; nombre de la secuencia en GenBank; identificador único con nomenclatura locus de la secuencia; identificador único de la secuencias, o accession; versión de la secuencia, indicada como "secuencia.n" donde n es la versión de la secuencia; código GI (Gene Info Identifier), corresponde a la nomenclatura usada antiguamente por GenBank; y la división de GenBank a la que pertenece la secuencia. *Hs*: Homo sapiens; *Mm*: Mus musculus; *Rn*: Rattus norvegicus; PRI; Primate; ROD: roedor. *Secuencias predictivas derivadas de análisis computacional.

Tabla 17. Información de secuencias de AFN de OLR1 para humano, ratón y rata, relativas a base de datos NCBI.

Organismo	GenBank (NCBI Reference Sequence)	Definición	Solicitud Anotación Directa (Direct submission)	Fecha primera versión anotación NCBI	Fecha última actualización NCBI	Deriva a partir de	N° Referencias	
<i>Hs</i>	BC022295.1	Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1, mRNA (cDNA clone MGC:22491 IMAGE:4722086), complete cds.	REFERENCE 2 (bases 1 to 2480), CONSRM NIH MGC Project, TITLE Direct Submission, JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA, REMARK NIH-MGC Project URL: http://mgc.ncbi.nih.gov	04-02-2002	22-12-2006	No Reportado	2	
	AB102861.1	Homo sapiens OLR1 mRNA for oxidised low density lipoprotein (lectin-like) receptor 1, complete cds	REFERENCE 2 (bases 1 to 822), AUTHORS Matsuoka,R., Furutani,M. and Furutani,Y., TITLE Direct Submission, JOURNAL Submitted (06-FEB-2003) Rumiko Matsuoka, Tokyo Women's Medical University, The Heart Institute of Japan; 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan (E- mail:rumiko@imcir.twmu.ac.jp, Tel:81-3-3353-8111(ex.24067), Fax:81-3-3352-3088)		07-08-2003	No Reportado	2	
	AF079164.1/AH007081.1	Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, exon 1	REFERENCE 2 (bases 1 to 2057), AUTHORS Yamanaka,S., TITLE Direct Submission, JOURNAL Submitted (20-JUL-1998) Pharmacology, Osaka City University Medical School, 1-4-54 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan	23-12-1998	23-12-1998	No Reportado	2	
	AF079165.1	Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, exon 2	REFERENCE 2 (bases 1 to 215), AUTHORS Yamanaka,S., TITLE Direct Submission, JOURNAL Submitted (20-JUL-1998) Pharmacology, Osaka City University Medical School, 1-4-54 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan	23-12-1998	23-12-1998	No Reportado	2	
	AF079166.1	Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, exon 3	REFERENCE 2 (bases 1 to 215), AUTHORS Yamanaka,S., TITLE Direct Submission, JOURNAL Submitted (20-JUL-1998) Pharmacology, Osaka City University Medical School, 1-4-54 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan	23-12-1998	23-12-1998	No Reportado	2	
	AF079167.1	Homo sapiens lectin-type oxidized LDL receptor (OLR1) gene, exons 4, 5, and 6, and complete cds	REFERENCE 2 (bases 1 to 215), AUTHORS Yamanaka,S., TITLE Direct Submission, JOURNAL Submitted (20-JUL-1998) Pharmacology, Osaka City University Medical School, 1-4-54 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan	23-12-1998	23-12-1998	No Reportado	2	
	DQ314885.1	Homo sapiens oxidised low density lipoprotein (lectin-like) receptor 1 (OLR1) gene, complete cds	REFERENCE 1 (bases 1 to 17734), AUTHORS Livingston,R.J., Rieder,M.J., McFarland,I.W., da Ponte,S.H., Aklestad,J., Swanson,J., Johnson,E. and Nickerson,D.A., TITLE Direct Submission, JOURNAL Submitted (05-DEC-2005) Genome Sciences, University of Washington, 1705 NE Pacific, Seattle, WA 98195, USA	23-12-2005	23-12-2005	No Reportado	1	
	NM_002543.3	Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant 1, mRNA	No Reportado		24-03-1999	05-05-2014	No Reportado	10
	NM_001172632.1	Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant 2, mRNA.	No Reportado			16-03-2014	No Reportado	10
	NM_001172633.1	Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant 3, mRNA	No Reportado		05-03-2010	16-03-2014	No Reportado	10
XM_006719081.1*	PREDICTED: Homo sapiens oxidized low density lipoprotein (lectin-like) receptor 1 (OLR1), transcript variant X1, mRNA	No Reportado		29-02-2000	03-02-2014	No Reportado	0	
<i>Mm</i>	NM_138648.2	Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant 1, mRNA.	No Reportado		02-01-2001	22-07-2014	No Reportado	10
	NM_001301094.1	Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant 3, mRNA	No Reportado		22-07-2014	22-07-2014	No Reportado	10
	NM_001301096.1	Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant 2, mRNA	No Reportado		22-07-2014	22-07-2014	No Reportado	10
	XM_006505303.1*	PREDICTED: Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant X1, mRNA.	No Reportado		27-12-2013	27-12-2013	NT_039353.8	0
	XM_006505304.1*	PREDICTED: Mus musculus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), transcript variant X2, mRNA.	No Reportado		27-12-2014	27-12-2013	NT_039353.8	0
<i>Rn</i>	NM_133306.2	Rattus norvegicus oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), mRNA	No Reportado		27-02-2002	04-09-2014	AB005900.1	10
	BC097290.1	Rattus norvegicus oxidized low density lipoprotein (lectin-like) receptor 1, mRNA (cDNA clone MGC:114271 IMAGE:7458637), complete cds	REFERENCE 2 (bases 1 to 2985), CONSRM NIH MGC Project, TITLE Direct Submission, JOURNAL Submitted (01-JUN-2005) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA, REMARK NIH-MGC Project URL: http://mgc.ncbi.nih.gov		17-07-2006	No Reportado	2	

Contenido de cada columna de izquierda a derecha: Especie; nombre de la secuencia en GenBank; definición presente en la base de datos NCBI; solicitud de publicación directa, se indica el número de referencia, título, fecha y autores de quienes solicitaron la publicación de la secuencia, cuando fue reportado; fecha de la primera versión de la anotación de la secuencia en NCBI; fecha de la última actualización de la anotación de la secuencia del gen en NCBI; desde que anotación deriva la presente secuencia, en caso de que se reporte; y la cantidad de referencias que han citado dicha secuencia según NCBI. *Hs*: Homo sapiens; *Mm*: Mus musculus; *Rn*: Rattus norvegicus. *Secuencias predictivas derivadas de análisis computacional.

Tabla 18. Información de secuencias de AFN de CLR1 para humano, ratón y rata, relativas a base de datos NCBI.

Orgnismo	GenBank (NCBI Reference Sequence)	Tamaño	Variante del transcrito	ID Proteína correspondiente	Especificaciones muestras	Comentario
Hs	BC022295.1	2480 pb	No Reportado	AAH22295.1	Con MGC:22491 IMAGE:4722086 Tejido: placenta Librería don: NIH_MGC_79	Incluye diferencias entre la secuencia del clon y el genoma <i>Hs</i> de referencia (constructo 36), y el genoma de chimpanzee (constructo 2)
	AB102861.1	822 pb	No Reportado	BAU81565.1	No Reportado	Informe sobre SNP que causa una variación aminoácida. Secuencia referencia: AF09161/
	AF09164.1/ AF007081.1	2057 pb	No Reportado	AAC97927.1	No Reportado	No Reportado
	AF09165.1	215 pb	No especificado	AAC97927.1	No Reportado	No Reportado
	AF09166.1	328 pb	No especificado	AAC97927.1	No Reportado	No Reportado
	AF09167.1	2681	No especificado	AAC97927.1	No Reportado	No Reportado
	U614885.1	17734 pb	No Reportado	No Reportado	No Reportado	Indica regiones repetidas y variaciones de la secuencia
	NM_002543.3	233 pb (242 pb según Ensembl)	1	NP_002534.1	No Reportado	Indica Sequence-Tagged Sites (STS), CDS y otros
	NM_001172632.1	2393 pb (950 pb según Ensembl)	2	NP_001166103.1	No Reportado	Indica Sequence-Tagged Sites (STS), CDS y otros Transcript Variant: This variant (2) lacks an exon in the coding region, which results in a frameshift and an early stop codon, compared to variant 1. The encoded isoform (2) is shorter and has a distinct C-terminus, compared to isoform 1.
	NM_001172633.1	2417 pb (1008 pb según Ensembl)	3	NP_001166104.1	No Reportado	Indica Sequence-Tagged Sites (STS), CDS y otros Transcript Variant: This variant (3) lacks an exon in the coding region, which results in a frameshift and an early stop codon, compared to variant 1. The encoded isoform (3) is shorter and has a distinct C-terminus, compared to isoform 1.
XM106719081.1*	2431 pb	X1	XP_006719144.1	No Reportado	Indica CDS y STS. mRNA derivado de análisis computacional usando Gnomon	
Mm	NM_138648.2	3581 pb	1	NP_619589.2	cepaC57BL/6J	Indica CDS, STS y otros This variant (1) encodes the longest isoform (1)
	NM_001301094.1	3059 pb	3	NP_001288023.1	cepaC57BL/6J	Indica CDS, STS y otros Transcript Variant: This variant (3) lacks exons in the coding region, compared to variant 1. The encoded isoform (3) is shorter, compared to isoform 1.
	NM_001301096.1	2957 bp	2	NP_001288025.1	cepaC57BL/6J	Indica CDS, STS y otros This variant (2) lacks exons in the coding region, compared to variant 1. The encoded isoform (2) is shorter, compared to isoform 1.
	XM106705083.1*	1457 pb	X1	XP_006505366.1	cepaC57BL/6J	Indica CDS y STS. mRNA derivado de análisis computacional usando Gnomon
	XM106705084.1*	1208 pb	X2	XP_006505367.1	cepaC57BL/6J	Indica CDS y STS. mRNA derivado de análisis computacional usando Gnomon
Rn	NM_133306.2	3751 pb	No especificado	NP_579840.2	No Reportado	Indica CDS, STS y otros
	BC097290.1	2985 pb	No especificado	AAH97290.1	clonación MGC:1142/1 IMAGE:7458637	Indica CDS

Contenido de cada columna de izquierda a derecha: Especie; nombre de la secuencia en GenBank; tamaño de la secuencia en pares de bases; isoforma variante que origina el transcrito, en caso de que sea reportado; identidad de la secuencia proteica que origina el transcrito; especificaciones de la muestra, como número de don, tejido, librería, etc; comentario sobre la secuencia. *Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Rn*: *Rattus norvegicus*; pb: pares de bases. *Secuencias predictivas derivadas de análisis computacional.

Tabla 19. Exones secuencias ARNm OURL. Se muestran solo las secuencias que reportan dicha información.

Oganismo	GenBank (NCBI Reference Sequence)	Nº Exones	Exon1	Exon2	Exon3	Exon4	Exon5	Exon6	Exon7	Exon8
<i>Hs</i>	AF079164.1/AH007081.1	1	1889..2025	No Reportado						
	AF079165.1	1	No Reportado	50..151	No Reportado					
	AF079166.1	1	No Reportado	No Reportado	52..297	No Reportado				
	AF079167.1	3	No Reportado	No Reportado	No Reportado	51..190	515..630	955..2681	No Reportado	No Reportado
	NM_002543.3	6	1..190	191..292	293..538	539..678	679..794	795..2516	No Reportado	No Reportado
	NM_001172632.1	5	1..190	191..292	293..538	539..654	655..2376	No Reportado	No Reportado	No Reportado
	NM_001172633.1	6	1..190	223..285	191..292	293..538	539..678	679..2400	No Reportado	No Reportado
<i>Mm</i>	NM_138648.2	8	1..131	132..233	234..479	480..617	618..755	756..892	893..1008	1009..3569
	NM_001301094.1	5	1..131	132..233	234..370	371..486	487..3047	No Reportado	No Reportado	No Reportado
	NM_001301096.1		1..131	132..268	269..384	385..2945	No Reportado	No Reportado	No Reportado	No Reportado

Contenido de cada columna de izquierda a derecha: Especie; nombre de la secuencia en GenBank; n° de exones reportados en NCBI para cada secuencia, en caso de que corresponda; columnas 4-11 indican posición de primera y última base nitrogenada de cada exón, y tamaños de exones, según reporta NCBI. Las secuencias en rata no presentan información sobre exones. *Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Rn*: *Rattus norvegicus*

Tabla 20. Región regulatoria, 5'UTR, 3'UTR y Sitio poliA de ARNm OURL.

Oganismo	GenBank (NCBI Reference Sequence)	Región regulatoria	5'UTR	3'UTR	Sitio poliA
<i>Hs</i>	AF079164.1/AH007081.1	1..1888	1889..1949	No Reportado	No Reportado
	AF079167.1	No Reportado	No Reportado	1097..2681	No Reportado
	DQ314885.1	No Reportado	No Reportado	No Reportado	No Reportado
	NM_002543.3	2486..2491	No Reportado	No Reportado	2516
	NM_001172632.1	2346..2351	No Reportado	No Reportado	2376
	NM_001172633.1	2370..2375	No Reportado	No Reportado	2400

Se muestran solo las secuencias que reportan dicha información. Contenido de cada columna de izquierda a derecha: Especie; nombre de la secuencia en GenBank; región regulatoria, indicando el sitio de ubicación de la primera y última base nitrogenada de la región; región 5'UTR, indicando el sitio de ubicación de la primera y última base nitrogenada de la región; región 3'UTR, indicando el sitio de ubicación de la primera y última base nitrogenada de la región; sitio de poliadenización, indicando la ubicación de la base nitrogenada donde ocurre. Las secuencias en rata y ratón no presentan información asociada. *Hs*: *Homo sapiens*.

Tabla 21. Información de secuencias proteicas LOX1 para humano, ratón y rata, relativas a base de datos NCBI.

Organismo	GenBank (UniProtKB/ Swiss-Prot)	Locus	Accession	Version	GI	GenBank Division
<i>Hs</i>	P78380.1 (UniProtKB/Swiss-Prot)	ORI_HUMAN	P78380	P78380.1	73621335	PRI
	AAH22295.1 (GenBank)	AAH22295	AAH22295	AAH22295.1	18490153	PRI
	EA W96159.1 (GenBank)*	EA W96159	EA W96159	EA W96159.1	119616565	PRI
	EA W96156.1 (GenBank)*	EA W96156	EA W96156	EA W96156.1	119616562	PRI
	EA W96157.1 (GenBank)*	EA W96157	EA W96157	EA W96157.1	119616563	PRI
	EA W96158.1 (GenBank)	EA W96158	EA W96158	EA W96158.1	119616564	PRI
	ABC40744.1 (GenBank)*	ABC40744	ABC40744	ABC40744.1	83699677	PRI
	AAC97927.1 (GenBank)*	AAC97927	AAC97927	AAC97927.1	4050004	PRI
	CAB38175.1 (GenBank)	CAB38175	CAB38175	CAB38175.1	4468344	PRI
	BAC81565.1 (GenBank)	BAC81565	BAC81565	BAC81565.1	33468736	PRI
	NP_002534.1 (NCBI Reference Sequence)	NP_002534	NP_002534	NP_002534.1	4505501	PRI
	NP_001166104.1 (NCBI Reference Sequence)	NP_001166104	NP_001166104	NP_001166104.1	290654344	PRI
	NP_001166103.1 (NCBI Reference Sequence)	NP_001166103	NP_001166103	NP_001166103.1	290654342	PRI
XP_006719144.1 (NCBI Reference Sequence)*	XP_006719144	XP_006719144	XP_006719144.1	578822879	PRI	
<i>Mm</i>	NP_619589.2 (NCBI Reference Sequence)	NP_619589	NP_619589 XP_989925	NP_619589.2	134053881	ROD
	NP_001288025.1 (NCBI Reference Sequence)	NP_001288025	NP_001288025	NP_001288025.1	666875863	ROD
	NP_001288023.1 (NCBI Reference Sequence)	NP_001288023	NP_001288023	NP_001288023.1	666875833	ROD
	XP_006505366.1 (NCBI Reference Sequence)*	XP_006505366	XP_006505366	XP_006505366.1	568940172	ROD
	XP_006505367.1 (NCBI Reference Sequence)*	XP_006505367	XP_006505367	XP_006505367.1	568940174	ROD
	EDK99928.1 (GenBank)*	EDK99928	EDK99928	EDK99928.1	148667512	ROD
	Q9EQ09.2 (UniProtKB/Swiss-Prot)	OLR1_MOUSE	Q9EQ09	Q9EQ09.2	341941209	ROD
<i>Rn</i>	NP_579840.2 (NCBI Reference Sequence)	NP_579840	NP_579840	NP_579840.2	404247470	ROD
	AAH97290.1 (GenBank)	AAH97290	AAH97290	AAH97290.1	67678321	ROD
	O70156 (GenBank)	OLR1_RAT	O70156	O70156.1	73621339	ROD

Contenido de cada columna de izquierda a derecha: Especie; nombre de la secuencia en GenBank, UniProt KB o Swiss-Prot; identificador único con nomenclatura Locus de la secuencia; identificador único de la secuencias, o accession; versión de la secuencia, indicada como “secuencia.n” donde n es la versión de la secuencia; código GI (Gene Info Identifier), corresponde a la nomenclatura usada antiguamente por GenBank; y la división de GenBank a la que pertenece la secuencia. *Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Rn*: *Rattus norvegicus*; PRI: primates; ROD: roedor. *Secuencias predictivas derivadas de análisis computacional.

Tabla 22. Información de secuencias proteicas LOX1 para humano, ratón y rata, relativas a base de datos NCBI.

Organismo	GenBank (UniProtKB/Swiss-Prot)	Definición	Solicitud Anotación Directa (Direct submission)	Fecha primera versión anotación NCBI	Fecha actualización anotación NCBI	Derivada de	Nº Referencias
Hs	P78380.1 (UniProtKB/Swiss-Prot)	RecName: Full=Oxidized low-density lipoprotein receptor 1; Short=Ox-LDL receptor 1; AltName: Full=C-type lectin domain family 8 member A; AltName: Full=Lectin-like oxidized LDL receptor 1; Short=LOX-1; Short=Lectin-like oxLDL receptor 1; Short=hLOX-1; AltName: Full=Lectin-type oxidized LDL receptor 1; Contains: RecName: Full=Oxidized low-density lipoprotein receptor 1, soluble form.	REFERENCE 6 (residues 1 to 273), AUTHORS Millar,D.S., TITLE Direct Submission, JOURNAL Submitted (??-FEB-1999), REMARK NUCLEOTIDE SEQUENCE [GENOMIC DNA]. REFERENCE 7 (residues 1 to 273), AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D., TITLE Direct Submission, JOURNAL Submitted (??-APR-2003), REMARK NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3).; TISSUE=Placenta REFERENCE 9 (residues 1 to 273), CONSRM NHLBI resequencing and genotyping service (RS&G), TITLE Direct Submission, JOURNAL Submitted (??-DEC-2005), REMARK NUCLEOTIDE SEQUENCE [GENOMIC DNA]. REFERENCE 11 (residues 1 to 273), AUTHORS Mural,R.J. y cols., TITLE Direct Submission, JOURNAL Submitted (??-JUL-2005), REMARK NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].	No Reportado	26-11-2014	No Reportado	28
	AAH22295.1 (GenBank)	Oxidized low density lipoprotein (lectin-like) receptor 1 [Homo sapiens]	REFERENCE 2 (residues 1 to 273), CONSRM NIH MGC Project, TITLE Direct Submission, JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA, REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov	No Reportado	22-12-2006	No Reportado	2
	EAW96159.1 (GenBank)*	oxidised low density lipoprotein (lectin-like) receptor 1, isoform CRA_a [Homo sapiens]	REFERENCE 2 (residues 1 to 273), AUTHORS Mural,R.J y cols., TITLE Direct Submission, JOURNAL Submitted (05-JUL-2005) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA	29-07-2005	14-05-2014	No Reportado	2
	EAW96156.1 (GenBank)*	oxidised low density lipoprotein (lectin-like) receptor 1, isoform CRA_a [Homo sapiens]	REFERENCE 2 (residues 1 to 273), AUTHORS Mural,R.J. y cols., TITLE Direct Submission, JOURNAL Submitted (05-JUL-2005) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA	29-07-2005	14-05-2014	No Reportado	2
	EAW96157.1 (GenBank)*	oxidised low density lipoprotein (lectin-like) receptor 1, isoform CRA_b [Homo sapiens]	REFERENCE 2 (residues 1 to 181), AUTHORS Mural,R.J. y cols., TITLE Direct Submission, JOURNAL Submitted (05-JUL-2005) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA	29-07-2005	14-05-2014	No Reportado	2
	EAW96158.1 (GenBank)*	oxidised low density lipoprotein (lectin-like) receptor 1, isoform CRA_c [Homo sapiens]	REFERENCE 2 (residues 1 to 189), AUTHORS Mural,R.J. y cols., TITLE Direct Submission JOURNAL Submitted (05-JUL-2005) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA	29-07-2005	14-05-2014	No Reportado	2
	ABC40744.1 (GenBank)*	oxidised low density lipoprotein (lectin-like) receptor 1 [Homo sapiens]	REFERENCE 1 (residues 1 to 273), AUTHORS Livingston,R.J., Rieder,M.J., McFarland,I.W., da Ponte,S.H., Aklestad,J., Swanson,J., Johnson,E. and Nickerson,D.A., TITLE Direct Submission, JOURNAL Submitted (05-DEC-2005) Genome Sciences, University of Washington, 1705 NE Pacific, Seattle, WA 98195, USA	23-12-2005	23-12-2005	No Reportado	1
	AAC97927.1 (GenBank)*	lectin-type oxidized LDL receptor [Homo sapiens]	REFERENCE 2 (residues 1 to 273), AUTHORS Yamanaka,S., TITLE Direct Submission, JOURNAL Submitted (20-JUL-1998) Pharmacology, Osaka City University Medical School, 1-4-54 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan	23-12-1998	23-12-1998	No Reportado	2
	CAB38175.1 (GenBank)	LOX1 [Homo sapiens]	No reportado	2 marzo 1999 (ENA)	14-11-2006	No Reportado	1
	BAC81565.1 (GenBank)	oxidised low density lipoprotein (lectin-like) receptor 1 [Homo sapiens]	REFERENCE 2 (residues 1 to 273), AUTHORS Matsuoka,R., Furutani,M. and Furutani,Y., TITLE Direct Submission, JOURNAL Submitted (06-FEB-2003) Rumiko Matsuoka, Tokyo Women's Medical University, The Heart Institute of Japan; 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan (E-mail:rumiko@imcir.twmu.ac.jp, Tel:81-3-3353-8111(ex.24067), Fax:81-3-3352-3088)	No Reportado	07-08-2003	No Reportado	2
	NP_002534.1 (NCBI Reference Sequence)	oxidized low-density lipoprotein receptor 1 isoform 1 [Homo sapiens]	No reportado	24-03-1999	05-05-2014	No Reportado	10
	NP_001166104.1 (NCBI Reference Sequence)	oxidized low-density lipoprotein receptor 1 isoform 3 [Homo sapiens]	No reportado	05-03-2010	16-03-2014	No Reportado	10
	NP_001166103.1 (NCBI Reference Sequence)	oxidized low-density lipoprotein receptor 1 isoform 2 [Homo sapiens]	No reportado	No Reportado	16-03-2014	No Reportado	10
XP_006719144.1 (NCBI Reference Sequence)*	PREDICTED: oxidized low-density lipoprotein receptor 1 isoform X1 [Homo sapiens]	No reportado	29 DEBRERO 2000	03-02-2014	NT_009714	0	
Mm	NP_619589.2 (NCBI Reference Sequence)	oxidized low-density lipoprotein receptor 1 isoform 1 [Mus musculus]	No reportado	02-01-2001	22-07-2014	AF303744	10
	NP_001288025.1 (NCBI Reference Sequence)	oxidized low-density lipoprotein receptor 1 isoform 2 [Mus musculus]	No reportado	22-07-2014	22-07-2014	AC161602.4	10

Continuación Tabla 24.

						BX513847.1	
	NP_001288023.1 (NCBI Reference Sequence)	oxidized low-density lipoprotein receptor 1 isoform 3 [Mus musculus]	No reportado	22-07-2014	22-07-2014	AC161602.4 BX513847.1	10
	XP_006505366.1 (NCBI Reference Sequence)*	PREDICTED: oxidized low-density lipoprotein receptor 1 isoform X1 [Mus musculus]	No reportado	27-12-2013	27-12-2013	No Reportado	0
	XP_006505367.1 (NCBI Reference Sequence)*	PREDICTED: oxidized low-density lipoprotein receptor 1 isoform X2 [Mus musculus]	No reportado	27-12-2013	27-12-2013	No Reportado	0
	EDK99928.1 (GenBank)*	oxidized low density lipoprotein (lectin-like) receptor 1 [Mus musculus]	REFERENCE 2 (residues 1 to 367), AUTHORS Mural,R.J., Adams,M.D., Myers,E.W., Smith,H.O. and Venter,J.C., TITLE Direct Submission, JOURNAL Submitted (05-JUL-2005) Celera Genomics, 45 W. Gude Dr., Rockville, MD 20850, USA	No Reportado	07-06-2007	No Reportado	2
	Q9EQ09.2 (UniProtKB/Swiss-Prot)	RecName: Full=Oxidized low-density lipoprotein receptor 1; Short=Ox-LDL receptor 1; AltName: Full=Lectin-like oxidized LDL receptor 1; Short=LOX-1; Short=Lectin-like oxLDL receptor 1; AltName: Full=Lectin-type oxidized LDL receptor 1; Contains: RecName: Full=Oxidized low-density lipoprotein receptor 1, soluble form.	REFERENCE 2 (residues 1 to 363), AUTHORS Park,S.-H., Ahn,H.-J. and Cho,J.-J., TITLE Direct Submission, JOURNAL Submitted (??-SEP-2000), REMARK NUCLEOTIDE SEQUENCE [MRNA]., REFERENCE 3 (residues 1 to 363), AUTHORS Caminci P. y cols., CONSRTM FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group), TITLE The transcriptional landscape of the mammalian genome JOURNAL Science 309 (5740), 1559-1563 (2005), PUBMED 16141072, REMARK NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].; STRAIN=NOD Errtum:[Science. 2006 Mar 24;311(5768):1713]	2005	29-10-2014	No Reportado	3
Rn	NP_579840.2 (NCBI Reference Sequence)	oxidized low-density lipoprotein receptor 1 [Rattus norvegicus].	No reportado	27-02-2002	04-09-2014	AC115156.5 BC097290.1 AABR07073062.1	10
	AAH97290.1 (GenBank)	Oxidized low density lipoprotein (lectin-like) receptor 1 [Rattus norvegicus]	REFERENCE 2 (residues 1 to 364), CONSRTM NIH MGC Project, TITLE Direct Submission JOURNAL Submitted (01-JUN-2005) National Institutes of Health, Mammalian Gene Collection, (MGC), Bethesda, MD 20892-2590, USA, REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov	No Reportado	17-07-2006	No Reportado	2
	O70156 (GenBank)	RecName: Full=Oxidized low-density lipoprotein receptor 1; Short=Ox-LDL receptor 1; AltName: Full=Lectin-like oxidized LDL receptor 1; Short=LOX-1; Short=Lectin-like oxLDL receptor 1; AltName: Full=Lectin-type oxidized LDL receptor 1; Contains: RecName: Full=Oxidized low-density lipoprotein receptor 1, soluble form.	No reportado	1998	26-11-2014	No Reportado	4

Contenido de cada columna de izquierda a derecha: Especie; nombre de la secuencia en GenBank, UniProt/Bo Swiss-Prot; definición presente en la base de datos NCBI; solicitud de publicación directa, se indica el número de referencia, título, fecha y autores de quienes solicitaron la publicación de la secuencia, cuando fue reportado; fecha de la primera versión de la anotación de la secuencia en NCBI; fecha de la última actualización de la anotación de la secuencia del gen en NCBI; desde que anotación deriva la presente secuencia, en caso de que se reporte; y la cantidad de referencias que han citado dicha secuencia según NCBI. *Hs*: Homo sapiens; *Mm*: Mus musculus; *Rn*: Rattus norvegicus. *Secuencias predictivas derivadas de análisis computacional.

Tabla 23. Información de secuencias proteicas LOXL para humano, ratón y rata, relativas a base de datos NCBI.

Organismo	GenBank (UniProtKB/ Swiss-Prot)	Longitud	Variante	ID DNA correspondiente	ID mRNA correspondiente	Forma de obtención	Tejido	Comentario
<i>Hs</i>	P78380.1 (UniProtKB/Swiss-Prot)	273 aa	No especificado	No Reportado	No especificado	No Reportado	No Reportado	Sin Comentarios
	AAH22295.1 (GenBank)	273 aa	No especificado	No Reportado	BC022295	No Reportado	Placenta	Sin Comentarios
	EAW96159.1 (GenBank)*	273 aa	CRA_a	CH471096 WGS Scaffold record	No especificado	Method: conceptual translation.	No Reportado	Method: conceptual translation.
	EAW96156.1 (GenBank)*	273 aa	CRA_a	CH471096 WGS Scaffold record	No especificado	Method: conceptual translation.	No Reportado	Method: conceptual translation.
	EAW96157.1 (GenBank)*	181 aa	CRA_b	CH471096 WGS Scaffold record	No especificado	Method: conceptual translation.	No Reportado	Method: conceptual translation.
	EAW96158.1 (GenBank)*	189 aa	CRA_c	CH471096 WGS Scaffold record	No especificado	Method: conceptual translation.	No Reportado	Method: conceptual translation.
	ABC40744.1 (GenBank)*	273 aa	No especificado	No especificado	DQ314885	Method: conceptual translation supplied by author	No Reportado	Method: conceptual translation supplied by author
	AAC97927.1 (GenBank)*	273 aa	No especificado	AF079164 AF079165 AF079166 AF079167	AF079165	Method: conceptual translation	No Reportado	Method: conceptual translation
	CAB38175.1 (GenBank)	273 aa	No especificado	No Reportado	No especificado	No Reportado	Leucocito	Sin Comentarios
	BAC81565.1 (GenBank)	273 aa	No especificado	No Reportado	AB102861	No Reportado	No Reportado	Sin Comentarios
	NP_002534.1 (NCBI Reference Sequence)	273 aa	Isoforma 1	No Reportado	NP_002534	From transcript and genomic sequence data to make the sequence consistent with the reference genome assembly; Molecular cloning, using expression cloning strategy	vascular endothelium and vascular-rich organs	The displayed sequence is further processed into a mature form. [Uniprot]
	NP_001166104.1 (NCBI Reference Sequence)	189 aa	Isoforma 3	No Reportado	NM_001172633	From transcript and genomic sequence data to make the sequence consistent with the reference genome assembly; Molecular cloning, using expression cloning strategy	vascular endothelium and vascular-rich organs	This variant (3) lacks an exon in the coding region, which results in a frameshift and an early stop codon, compared to variant 1. The encoded isoform (3) is shorter and has a distinct C-terminus, compared to isoform 1.
	NP_001166103.1 (NCBI Reference Sequence)	181 aa	Isoforma 2	No Reportado	NM_001172632	From transcript and genomic sequence data to make the sequence consistent with the reference genome assembly; Molecular cloning, using expression cloning strategy	No Reportado	This variant (2) lacks an exon in the coding region, which results in a frameshift and an early stop codon, compared to variant 1. The encoded isoform (2) is shorter and has a distinct C-terminus, compared to isoform 1
	XP_006719144.1 (NCBI Reference Sequence)*	217 aa	Isoforma X1	No Reportado	XM_006719081	MODEL REFSEQ: This record is predicted by automated computational analysis. This record is derived from a genomic sequence (NT_009714.18) annotated using gene prediction method: Gnomon, supported by EST evidence.	Hipotético	MODEL REFSEQ: This record is predicted by automated computational analysis. This record is derived from a genomic sequence (NT_009714.18) annotated using gene prediction method: Gnomon, supported by EST evidence.
<i>Mm</i>	NP_619589.2 (NCBI Reference Sequence)	363 aa	isoforma 1	No Reportado	NM_138648		No Reportado	Transcript Variant: This variant (1) encodes the longest isoform (1). [NCBI] The displayed sequence is further processed into a mature form. [Uniprot]
	NP_001288025.1 (NCBI Reference Sequence)	155 aa	isoforma 2	No Reportado	NM_001301096	From transcript and genomic sequence data to make the sequence consistent with the reference genome assembly; Molecular cloning, using expression cloning strategy	No Reportado	Transcript Variant: This variant (2) lacks exons in the coding region, compared to variant 1. The encoded isoform (2) is shorter, compared to isoform 1. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.

Continuación Tabla 25.

	NP_001288023.1 (NCBI Reference Sequence)	189 aa	isoforma 3	No Reportado	NM_001301094	From transcript and genomic sequence data to make the sequence consistent with the reference genome assembly; Molecular cloning, using expression cloning strategy	No Reportado	<p>Transcript Variant: This variant (3) lacks exons in the coding region, compared to variant 1. The encoded isoform (3) is shorter, compared to isoform 1.</p> <p>Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments.</p> <p>Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.</p>
	XP_006505366.1 (NCBI Reference Sequence)*	317 aa	Isoforma X1	No Reportado	XM_006505304	No Reportado	No Reportado	<p>MODEL REFSEQ: This record is predicted by automated computational analysis. This record is derived from a genomic sequence (NT_039353.8) annotated using gene prediction method: Gnomon, supported by EST evidence.</p>
	XP_006505367.1 (NCBI Reference Sequence)*	280 aa	isoforma X2	No Reportado	XM_006505305	No Reportado	No Reportado	<p>MODEL REFSEQ: This record is predicted by automated computational analysis. This record is derived from a genomic sequence (NT_039353.8) annotated using gene prediction method: Gnomon, supported by EST evidence.</p>
	EDK99928.1 (GenBank)*	367 aa	No Reportado	No Reportado	No Reportado	No Reportado	No Reportado	Method: conceptual translation
	Q9EQ09.2 (UniProtKB/Swiss-Prot)	363 aa	No Reportado	No Reportado	No Reportado	No Reportado	No Reportado	Sin Comentarios
Rn	NP_579840.2 (NCBI Reference Sequence)	368 aa	No Reportado	No Reportado	NM_133306	From transcript and genomic sequence data to make the sequence consistent with the reference genome assembly;	No Reportado	<p>VALIDATED REFSEQ: This record has undergone validation or preliminary review. The reference sequence was derived from AC115156.5, BC097290.1 and AABR07073062.1. On Sep 11, 2012 this sequence version replaced gi:18959248.</p> <p>Summary: receptor for oxidized low-density lipoprotein that may be involved in pathogenesis of hypertension as well as atherosclerosis [RGD, Feb 2006].</p> <p>Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments.</p> <p>Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.</p> <p>##Evidence-Data-START## Transcript exon combination :: AB005900.1, BC097290.1 [ECO:0000332] ##Evidence-Data-END##</p>
	AAH97290.1 (GenBank)	364 aa	No Reportado	No Reportado	BC097291	No Reportado	No Reportado	Sin Comentarios
	O70156 (GenBank)	364 aa	No Reportado	No Reportado	No Reportado	No Reportado	No Reportado	Cepa 10116

Contenido de cada columna de izquierda a derecha: Especie; nombre de la secuencia en GenBank, UniProtKB o Swiss-Prot; longitud de la secuencia en aminoácidos; variante o isoforma, en caso de que sea reportado; identidad de la secuencia de DNA que origina a la proteína; identidad de la secuencia de mRNA que origina a la proteína; forma de obtención de la proteína; especificaciones del tejido desde el cual se tomó la muestra; comentario sobre la secuencia. *Hs*: Homo sapiens; *Mm*: Mus musculus; *Rn*: Rattus norvegicus; *ccr*: aminoácidos. *Secuencias predictivas derivadas de análisis computacional.

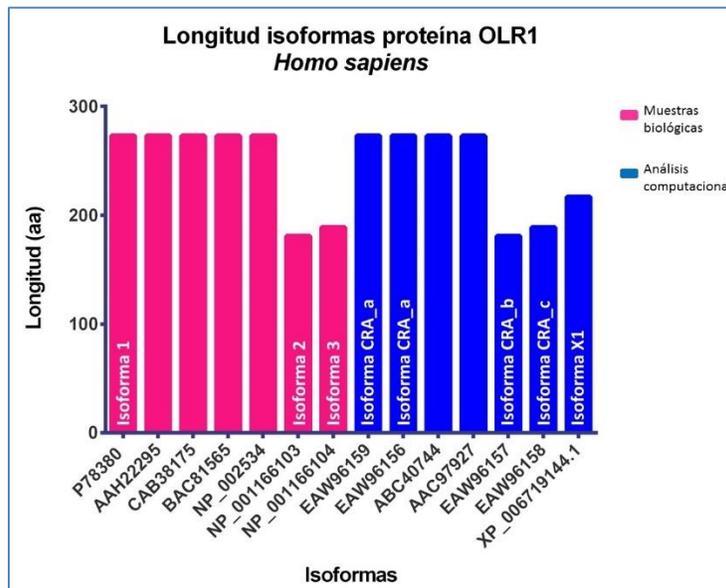


Figura 41. Longitud de isoformas proteicas de LOX1 en humano. En el eje X se identifica cada Isoforma con el nombre de la secuencia presentado en base de datos, mientras que en el eje Y se muestra el largo de la secuencia proteica en aminoácidos (aa). Las barras fucsia corresponden a isoformas encontradas en muestras biológicas, mientras que las barras en azul corresponden a secuencias proteicas obtenidas mediante métodos computacionales, según lo reportado en la **Tabla 21**. También se indica la variante de la isoforma en letras blancas en cada barra, según lo especificado en la **Tabla 21**.

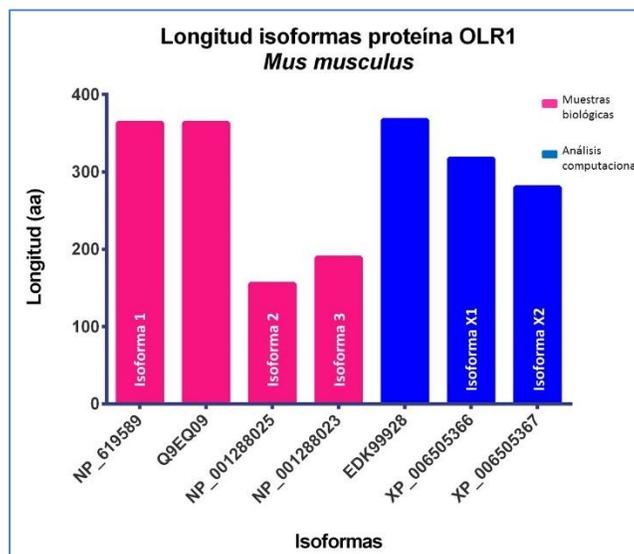


Figura 42. Longitud de isoformas proteicas de LOX1 en ratón. En el eje X se identifica cada Isoforma con el nombre de la secuencia presentado en base de datos, mientras que en el eje Y se muestra el largo de la secuencia proteica en aminoácidos (aa). Las barras fucsia corresponden a isoformas encontradas en muestras biológicas, mientras que las barras en azul corresponden a secuencias proteicas obtenidas mediante métodos computacionales, según lo reportado en la **Tabla 21**. También se indica la variante de la isoforma en letras blancas en cada barra, según lo especificado en la **Tabla 21**.

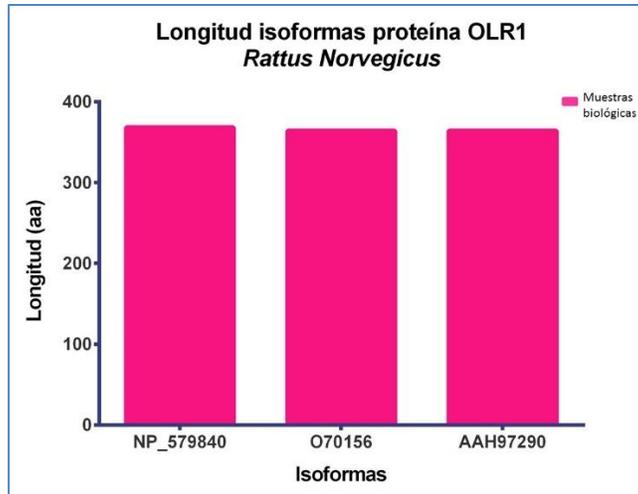


Figura 43. Longitud de isoformas proteicas de LOX1 en rata. En el eje X se identifica cada Isoforma con el nombre de la secuencia presentado en base de datos, mientras que en el eje Y se muestra el largo de la secuencia proteica en aminoácidos (aa). Las barras fucsia corresponden secuencias encontradas en muestras biológicas. No existen secuencias proteicas obtenidas mediante métodos computacionales o isoformas asociadas a esta especie, según lo reportado en la **Tabla 21**.

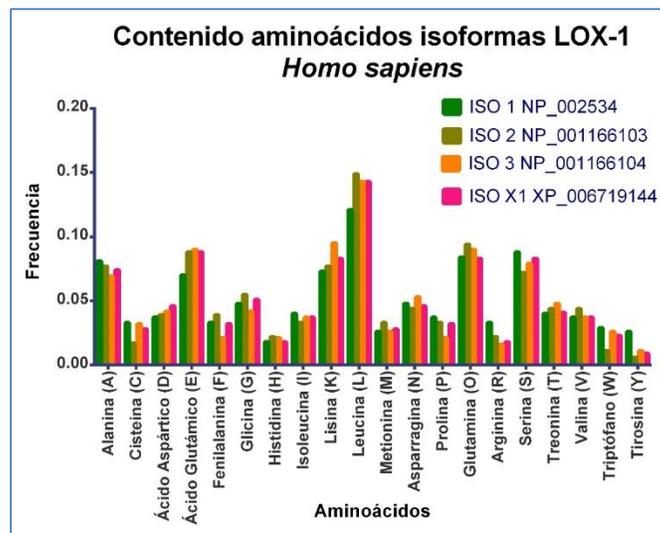


Figura 44. Contenido de Aminoácidos de Isoformas LOX1 en humano. En el eje X se identifica cada aminoácido con su nombre y letra representativa, mientras que en el eje Y se muestra la frecuencia en porcentaje en que se encuentra en la secuencia. Para facilitar la comparación, solo se muestran la isoformas más relevantes para este trabajo, según lo especificado en la **Tabla 21**.

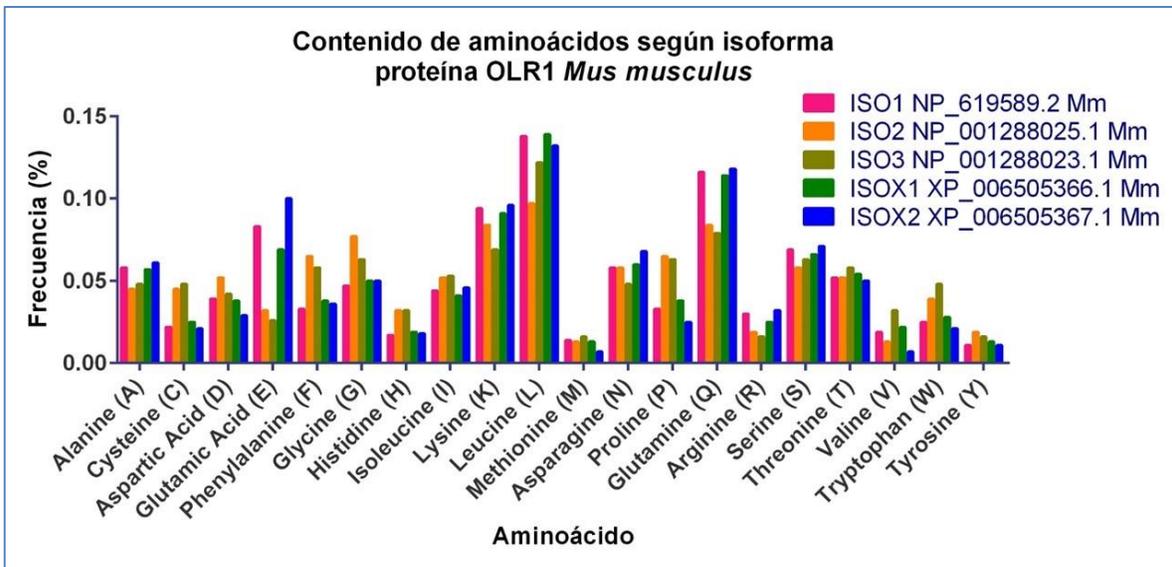


Figura 45. Contenido de Aminoácidos de Isoformas LOX1 en ratón. En el eje X se identifica cada aminoácido con su nombre y letra representativa, mientras que en el eje Y se muestra la frecuencia en porcentaje en que se encuentra en la secuencia. Para facilitar la comparación, solo se muestran la isoformas más relevantes para este trabajo, según lo especificado en la **Tabla 21**.

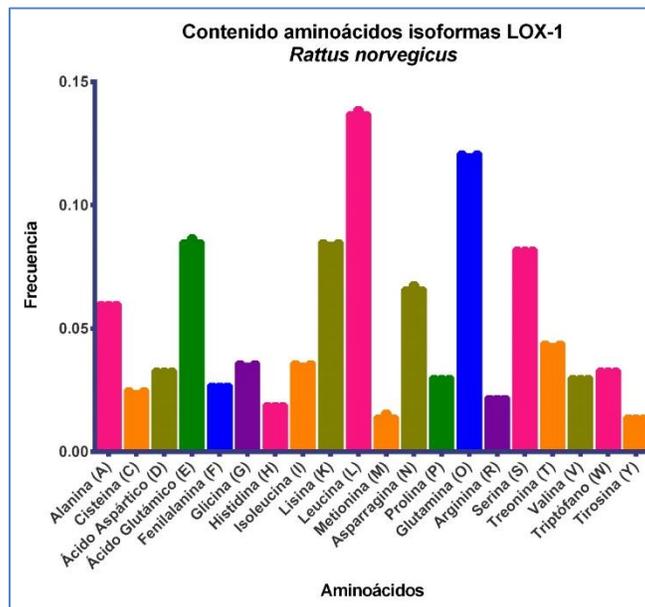


Figura 46. Contenido de Aminoácidos de Isoformas LOX1 en rata. En el eje X se identifica cada Aminoácido con su nombre y letra representativa, mientras que en el eje Y se muestra la frecuencia en porcentaje en que se encuentra en la secuencia. Para facilitar la comparación, solo se muestran la isoformas más relevantes para este trabajo, según lo especificado en la **Tabla 21**.

Tabla 24. Variaciones de aminoácidos en LOX1 de humano, según secuencia proteica.

Secuencia	Posición variación	Cambio Aminoácido (1° letra corresponde a P78380)	Dominio/Región	Tipo aminoácido original	Tipo aminoácido variante
NP_001166103	142	A/G	Cuello/ Extracelular	No polar	Polar no cargado
	143-149	Delección	Cuello/ Extracelular	-	-
	150	W/L	Cuello/ Extracelular	No polar	No polar
	152	G/P	Lectina Tipo C/ Extracelular	Polar no cargado	No polar
	153	E/A	Lectina Tipo C/ Extracelular	Ácido	No polar
	154	N/S	Lectina Tipo C/ Extracelular	Polar no cargado	Polar no cargado
	155	C/N	Lectina Tipo C/ Extracelular	Polar no cargado	Polar no cargado
	156	Y/F	Lectina Tipo C/ Extracelular	Polar no cargado	No polar
	159	S/Q	Lectina Tipo C/ Extracelular	Polar no cargado	Polar no cargado
	160	S/F	Lectina Tipo C/ Extracelular	Polar no cargado	No polar
	161	G/S	Lectina Tipo C/ Extracelular	Polar no cargado	Polar no cargado
	162	S/I	Lectina Tipo C/ Extracelular	Polar no cargado	No polar
	163	F/L	Lectina Tipo C/ Extracelular	No polar	No polar
	164	N/D	Lectina Tipo C/ Extracelular	Polar no cargado	Ácido
	165	W/G	Lectina Tipo C/ Extracelular	No polar	Polar no cargado
	166	E/A	Lectina Tipo C/ Extracelular	Ácido	No polar
	167	K/V	Lectina Tipo C/ Extracelular	Básico	No polar
	169	Q/E	Lectina Tipo C/ Extracelular	Polar no cargado	Ácido
	171	K/P	Lectina Tipo C/ Extracelular	Básico	No polar
	172	C/Q	Lectina Tipo C/ Extracelular	Polar no cargado	Polar no cargado
	174	S/P	Lectina Tipo C/ Extracelular	Polar no cargado	No polar
	175-203	Delección	Lectina Tipo C/ Extracelular	-	-
	205	G/A	Lectina Tipo C/ Extracelular	Polar no cargado	No polar
207	S/G	Lectina Tipo C/ Extracelular	Polar no cargado	Polar no cargado	

	208	R/G	Lectina Tipo C/ Extracelular	Básico	Polar no cargado
	210-216	Delección	Lectina Tipo C/ Extracelular	-	-
	217	W/F	Lectina Tipo C/ Extracelular	No polar	No polar
	218	E/S	Lectina Tipo C/ Extracelular	Ácido	Polar no cargado
	219	D/F	Lectina Tipo C/ Extracelular	Ácido	No polar
	220	G/D	Lectina Tipo C/ Extracelular	Polar no cargado	Ácido
	221	S/A	Lectina Tipo C/ Extracelular	Polar no cargado	No polar
	224	M/I	Lectina Tipo C/ Extracelular	No polar	No polar
	225-273	Delección	Lectina Tipo C/ Extracelular	-	-
BAC81565	167	K/N	Lectina Tipo C/ Extracelular	Básico	Polar no cargado
NP_001166104	189	D/I	Lectina Tipo C/ Extracelular	Ácido	No polar
	190-273	Delección	Lectina Tipo C/ Extracelular	-	-

Las distintas secuencias alineadas se compararon con la secuencia de referencia P78380, se anotaron las variaciones y se categorizaron de acuerdo a la región en que se encuentran, el cambio de aminoácido, y el tipo de aminoácido distinto. A: Alanina; C: Cisteína; D: Ácido Aspártico; E: Ácido Glutámico; F: Fenilalanina; G: Glicina; I: Isoleucina; K: Lisina; L: Leucina; M: Metionina; N: Asparragina; P: Prolina; Q: Glutamina; R: Arginina; S: Serina; V: Valina; W: Triptófano; Y: Tirosina.

Tabla 25. Variaciones de aminoácidos en LOX1 de ratón, según secuencia proteica.

Secuencia	Posición variación	Cambio Aminoácido (1° letra corresponde a Q9EQ09)	Dominio/Región	Tipo aminoácido original	Tipo aminoácido variante
NP_001288025 (Isoforma 2)	26- 233	Delección	Cuello/ Repeticiones/ Citoplasmático- Transmembrana- Extracelular	-	-
NP_001288023 (Isoforma 3)	159	L/C	Coiled Coil/ Cuello/ Repetición 2/ Extracelular	No polar	Polar no cargado
	60-233	Delección	Sin Coiled coil/ Sin Cuello/ Sin Repeticiones/ Transmembrana- Extracelular	-	-

Las distintas secuencias alineadas se compararon con la secuencia de referencia Q9EQ09, se anotaron las variaciones y se categorizaron de acuerdo a la región en que se encuentran, el cambio de aminoácido, y el tipo de aminoácido distinto. C: Cisteína; L: Leucina.

Tabla 26. Variaciones de aminoácidos en LOX1 de rata, según secuencia proteica.

Secuencia	Posición variación	Cambio Aminoácido (1° letra corresponde a Q9EQ09)	Dominio/Región	Tipo aminoácido original	Tipo aminoácido variante
NP_579840	2-5	Inserción N, L , E, M	Citoplasmático	-	Polar no cargado, No polar, Ácido, No polar

Las distintas secuencias alineadas se compararon con la secuencia de referencia O70156, se anotaron las variaciones y se categorizaron de acuerdo a la región en que se encuentran, el cambio de aminoácido, y el tipo de aminoácido distinto. E: Ácido Glutámico; L: Leucina; M: Metionina; N: Asparragina.



**Alineamiento gen
OLR1 Humano.tif**

Archivo vinculado 1. Alineamiento de secuencias de gen OLR1 de humano. Las secuencias fueron descargadas desde base de datos NCBI. Se indica el nombre de la secuencia, y la especie a la que pertenece (*Hs* para humano, *Homo sapiens*), a la izquierda del alineamiento; a la derecha del alineamiento se indica la posición de la última base nitrogenada de la secuencia en esa línea del alineamiento; bajo cada línea del alineamiento se indica el grado de conservación de las secuencias en gráfica de barras que van desde el 0% (ninguna base nitrogenada común en las secuencias analizadas) hasta 100% de conservación (bases nitrogenadas iguales en todas las secuencias analizadas). Versión digital: doble click sobre el objeto para acceder al archivo de 32 hojas.



**Alineamiento gen
OLR1 Humano exclu**

Archivo vinculado 2. Alineamiento de secuencias de gen OLR1 de humano, excluyendo la secuencia AJ131757. Las secuencias fueron descargadas desde base de datos NCBI. Se indica el nombre de la secuencia, y la especie a la que pertenece (*Hs* para humano *Homo sapiens*), a la izquierda del alineamiento; a la derecha del alineamiento se indica la posición de la última base nitrogenada de la secuencia en esa línea del alineamiento; bajo cada línea del alineamiento se indica el grado de conservación de las secuencias en gráfica de barras que van desde el 0% (ninguna base nitrogenada común en las secuencias analizadas) hasta 100% de conservación (bases nitrogenadas iguales en todas las secuencias analizadas). Versión digital: doble click sobre el objeto para acceder al archivo de 24 hojas.



**Alineamiento gen
OLR1 Ratón.tif**

Archivo vinculado 3. Alineamiento de secuencias de gen OLR1 de ratón. Las secuencias fueron descargadas desde base de datos NCBI. Se indica el nombre de la secuencia, y la especie a la que pertenece (*Mm* para ratón, *Mus musculus*), a la izquierda del alineamiento; a la derecha del alineamiento se indica la posición de la última base nitrogenada de la secuencia en esa línea del alineamiento; bajo cada línea del alineamiento se indica el grado de conservación de las secuencias en gráfica de barras que van desde el 0% (ninguna base nitrogenada común en las secuencias analizadas) hasta 100% de conservación (bases nitrogenadas iguales en todas las secuencias analizadas). Versión digital: doble click sobre el objeto para acceder al archivo de 30 hojas.



**Alineamiento gen
OLR1 Rata.tif**

Archivo vinculado 4. Alineamiento de secuencias de gen OLR1 de rata. Las secuencias fueron descargadas desde base de datos NCBI. Se indica el nombre de la secuencia, y la especie a la que pertenece (*Rn* para rata, *Rattus norvegicus*), a la izquierda del alineamiento; a la derecha del alineamiento se indica la posición de la última base nitrogenada de la secuencia en esa línea del alineamiento; bajo cada línea del alineamiento se indica el grado de conservación de las secuencias en gráfica de barras que van desde el 0% (ninguna base

nitrogenada común en las secuencias analizadas) hasta 100% de conservación (bases nitrogenadas iguales en todas las secuencias analizadas). Versión digital: doble click sobre el objeto para acceder al archivo de 24 hojas.



**Alineamiento gen
OLR1 Humano, Rató**

Archivo vinculado 5. Alineamiento secuencias gen OLR1 de humano, ratón y rata. Se indica el nombre de la secuencia, y la especie a la que pertenece (*Hs* para humano, *Mm* para ratón y *Rn* para rata), a la izquierda del alineamiento; a la derecha del alineamiento se indica la posición de la última base nitrogenada de la secuencia en esa línea del alineamiento; bajo cada línea del alineamiento se indica el grado de conservación de las secuencias en gráfica de barras que van desde el 0% (ninguna base nitrogenada común en las secuencias analizadas) hasta 100% de conservación (bases nitrogenadas iguales en todas las secuencias analizadas).