



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Escola Tècnica Superior
d'Enginyeria Agronòmica i del Medi Natural

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Agronómica
y del Medio Natural

Análisis transcriptómico del hígado de dos líneas de conejo
divergentes seleccionadas por el contenido de grasa
intramuscular

Trabajo Fin de Grado

Grado en Biotecnología

AUTOR/A: Peñuelas Cumplido, Ángel

Tutor/a: Hernández Pérez, María del Pilar

Director/a Experimental: Valdés Hernández, Jesús

CURSO ACADÉMICO: 2023/2024

RESUMEN

El contenido de grasa intramuscular (GIM) es un indicador importante de la calidad de la carne y afecta las propiedades nutricionales, sensoriales y tecnológicas de la carne. En la Universitat Politècnica de València se realizó un experimento de selección divergente de GIM en el músculo Longissimus thoracis et lumborum en conejos durante 10 generaciones. El experimento fue exitoso, obteniendo una respuesta directa a la selección en la 10ª generación de 3,8 desviaciones típicas del carácter. El GIM es un rasgo complejo de naturaleza poligénica, lo que significa que está controlado por múltiples genes, cada uno con un pequeño efecto. En este estudio, el objetivo fue identificar genes expresados diferencialmente (DEGs) en el hígado de dos líneas de conejo seleccionadas de manera divergente para GIM (líneas alta y baja), y estudiar aquellos que modulan la deposición de grasa. Se utilizó un total de 48 conejos (equilibrados entre sexo y línea) pertenecientes a la novena generación de selección para GIM. Se utilizaron muestras de hígado de estos animales para determinar los niveles de expresión génica mediante la metodología 3' RNA-Seq, la cual implica la amplificación del extremo 3' del ARN mensajero. La secuenciación de librerías se realizó en la plataforma Illumina NovaSeq X (2 × 150 bp paired-end con un promedio de 5 millones de reads/muestra). Tras aplicar un pipeline bioinformático desarrollado *in-house* se obtuvo la matriz de counts. El análisis de expresión diferencial se realizó mediante un modelo binomial negativo, seguido de un análisis funcional de los DEGs detectados. Nuestros resultados indicaron un total de 420 DEGs entre las líneas divergentes (cutoff de fold change = 1.5 y BH $P_{adj} \leq 0.05$), de los cuales 168 mostraron un incremento de expresión y 252 una disminución de la expresión en la línea alta versus la baja. Entre los DEGs, detectamos genes candidatos asociados en general con el metabolismo lipídico, lipólisis, lipogénesis, adipogénesis, beta oxidación de ácidos grasos, biosíntesis de ácidos grasos, entre otros. Detectamos una sobrerrepresentación de 38 términos funcionales que incluyeron 28 pathways, 8 procesos biológicos y 2 funciones moleculares, entre los detectamos la vía de señalización del receptor activado por proliferador de peroxisomas (PPAR), elongación de ácidos grasos, degradación de ácidos grasos, biosíntesis de ácidos grasos insaturados, metabolismo del ácido araquidónico y del retinol. El estudio de la expresión de genes hepáticos reveló información relevante sobre genes candidatos para el depósito de GIM en conejos, pero también sobre bioprocesos y vías metabólicas asociadas con el metabolismo de lípidos y carbohidratos.

Palabras clave: Metabolismo lipídico hepático, Deposición de grasa, 3' Tag-Seq y *Oryctolagus cuniculus*

ABSTRACT

The intramuscular fat (IMF) content is an important indicator of meat quality, affecting the sensory properties of meat. IMF is a complex trait with polygenic nature, meaning that it is controlled by multiple genes, each one with a small effect. In this study, the aim was to identify differentially expressed (DE) genes in the liver of two rabbit lines divergently selected for IMF (high and low lines), and to study those that modulate fat deposition. We used a total of 48 rabbits (balanced between sex and line) belonging to the ninth generation of selection for IMF. Liver samples from these animals were used to determine the gene expression levels by 3' Tag-Seq. This method involves the amplification of the 3' end of the messenger RNA using a combination of polyT primers and random primers. Sequencing of the library was performed on the Illumina NovaSeq X platform with 150 bp paired-end chemistry. A differential expression analysis using a negative binomial model was performed, followed by a functional analysis of the DE genes detected. Our results indicated a total of 551 DE genes between the divergent lines, of which 235 were up-regulated and 316 down-regulated in the high line. Among the DE genes, we detected candidate genes associated with lipolysis, lipogenesis, adipogenesis, fatty acid beta oxidation, fatty acid biosynthesis, lipid transport, or bile acid secretion (e.g., ALOXE3, ANKRD1, ARSI, BACH2, CPT1B, CROT, CUBN, CYP4A6, ELOVL6, FABP4, FITM1, GCKR, GGT1, HADHB, NR4A1, NR4A2, PLIN2, PPP1R3B, SLC44A4, SLC51A, SLC51B), among others. Likewise, we detected an overrepresentation of 38 functional annotation terms including 28 pathways, 8 biological process and 2 molecular function, among which the peroxisome proliferator-activated receptor signaling (PPAR), fatty acid elongation, fatty acid degradation, biosynthesis of unsaturated fatty acids, arachidonic and retinol metabolism were detected. The liver gene expression study revealed valuable information about candidate genes for IMF deposition in rabbits, but also about bioprocesses and pathways associated with lipid and carbohydrate metabolisms.

Key words: hepatic lipid metabolism; fat deposition; 3' Tag-Seq; *Oryctolagus cuniculus*

Este trabajo se relaciona con el ODS 2. Hambre cero cumpliendo la primera meta, para 2030, poner fin al hambre y asegurar el acceso de todas las personas, en particular las personas en situación de pobreza y vulnerables, incluidos los lactantes, a una alimentación sana, nutritiva y suficiente durante todo el año. También relacionado con el ODS 3. Salud y bienestar en concreto para 2030, reducir en un tercio la mortalidad prematura por enfermedades no transmisibles mediante la prevención y el tratamiento y promover la salud mental y el bienestar y por último con el ODS 12. Producción y consumo responsable, en particular, de aquí a 2030, reducir a la mitad el desperdicio de alimentos per capita mundial en la venta al por menor y a nivel de los consumidores y reducir las pérdidas de alimentos en las cadenas de producción y suministro, incluidas las pérdidas posteriores a la cosecha

Objetivos de Desarrollo Sostenible	Alto	Medio	Bajo	No procede
ODS 1. Fin de la pobreza.				X
ODS 2. Hambre cero.	X			
ODS 3. Salud y bienestar.	X			
ODS 4. Educación de calidad.				X
ODS 5. Igualdad de género.				X
ODS 6. Agua limpia y saneamiento.				X
ODS 7. Energía asequible y no contaminante.				X
ODS 8. Trabajo decente y crecimiento económico.				X
ODS 9. Industria, innovación e infraestructuras.				X
ODS 10. Reducción de las desigualdades.				X
ODS 11. Ciudades y comunidades sostenibles.				X
ODS 12. Producción y consumo responsables.		X		
ODS 13. Acción por el clima.				X
ODS 14. Vida submarina.				X
ODS 15. Vida de ecosistemas terrestres.				X
ODS 16. Paz, justicia e instituciones sólidas.				X
ODS 17. Alianzas para lograr objetivos.				X

AGRADECIMIENTOS

En primer lugar, agradecer a mi tutora, Pilar Hernández Pérez, por haberme dado la oportunidad de realizar este TFG y su ayuda a lo largo de todo este proceso.

También agradecer enormemente al tutor experimental, Jesús Valdés Hernández por todo su apoyo y dedicación enseñándome como funciona el mundo de analizar datos de secuenciación, así como su conocimiento y experiencia han sido imprescindibles para la realización de este trabajo.

Por otro lado, muchas gracias a mis compañeros de carrera, sobre todo a mis amigos por ese apoyo y ayuda en momentos en los que más lo necesitaba

Agradecer a Francisca Herrera Cortés su labor a la hora de revisar y mejorar este documento.

Y por último, pero no menos importante agradecer a mis padres todo su apoyo incondicional y su fe en mí a lo largo de estos años. Habéis sido mi inspiración y mi fuerza en los momentos más difíciles. Gracias por inculcarme los valores fundamentales por los que he conseguido alcanzar todas mis metas. Este Trabajo Fin de Grado es tanto mío como vuestro porque sin vosotros esto no habría sido posible.

ÍNDICE

1	INTRODUCCIÓN	1
1.1	Situación actual de la producción de carne de conejo	1
1.2	Importancia nutricional y composición de la carne de conejo	2
1.3	Importancia de la grasa Intramuscular (GIM).....	3
1.4	El papel del hígado en el metabolismo	5
1.5	Regulación de la expresión génica.....	7
1.6	Secuenciación del ARN (RNA-Seq) y sus aplicaciones	8
1.7	Análisis y visualización de genes diferencialmente expresados	11
1.8	Hipótesis y objetivos.....	12
2	MATERIALES Y MÉTODOS	13
2.1	Animales y diseño experimental.....	13
2.2	Análisis bioinformáticos y estadísticos	13
3	RESULTADOS Y DISCUSIÓN	14
3.1	Descripción del control de calidad de la secuenciación	14
3.2	Análisis de expresión génica diferencial.....	15
3.3	Análisis funcionales de genes expresados diferencialmente	21
4	CONCLUSIONES	23
5	BIBLIOGRAFÍA	24
6	ANEXOS	31

ÍNDICE DE FIGURAS

Figura 1. Proporción de producción de Carne de conejos y liebres, fresca o refrigerada por región. De: FAOSTAT, 2024	1
Figura 2. Distribución de la producción de carne de conejo por comunidades autónomas en el año 2021 (toneladas). De: “EL SECTOR CUNÍCOLA EN CIFRAS: Principales Indicadores Económicos”, Subdirección General de Producciones Ganaderas y Cinegéticas, Dirección General.....	2
Figura 3. Principales fuentes de ácidos grasos hepáticos. Adaptado de Alves-Bezerra & Cohen, 2018.	5
Figura 4. Mapa general para el análisis computacional de un RNA-Seq. Los tres bloques principales son preanálisis, análisis central y análisis avanzado. A) incluye diseño experimental, estrategia de secuenciación y control de calidad. B) Los análisis principales	8
Figura 5. Preparación de librerías para realizar un RNA-Seq. Adaptado de: Moll et al., 2014	10
Figura 6. Cada molécula está identificada con un marcador molecular propio, permitiendo diferenciar aquellos duplicados sintéticos de fragmentos con coordenadas idénticas. De: faqs.lexogen.com (https://faqs.lexogen.com/faq/what-are-unique-molecular-identifiers-umis).....	10
Figura 7. Estadísticos obtenidos en cuanto a la calidad de los alineamientos generados, midiendo las lecturas que no se encuentran dentro de regiones transcritas, o lecturas no alineadas.	14
Figura 8. Porcentaje de lecturas únicas y lecturas duplicadas dentro del total de secuencias disponibles para la muestra.....	15
Figura 9. Genes diferencialmente expresados provenientes del 3`RNA-Seq de hígado entre líneas altas y bajas para IGM. Las líneas discontinuas indican los valores del umbral.	16
Figura 10. Representación del network de los DEGs después de comparar las expresiones de genes en el RNA-Seq de las diferentes líneas de conejos.	21
Figura 11. Agrupamiento de los genes DE en los diferentes clusters y porcentaje de genes en cada uno.	22

ÍNDICE DE TABLAS

Tabla 1. Top 20 de genes diferencialmente expresados (por valor P) en hígado entre líneas divergentes según contenido de grasa intramuscular.	17
--	----

SIGLA Y ABREVIATURAS

AG: Ácidos grasos

ARNm: ARN mensajero

ARNr: ARN ribosómico

DEG: Genes diferencialmente expresados

DHA: Ácido graso docosahexaenoico

EPA: Ácido graso eicosapentaenoico

FC: Fold change

FDR: False discovery rate

GIM: Grasa intramuscular

GLM: Modelos lineales generalizados

GWAS: Estudio de asociación del genoma completo

LCEH: 2-enoil-CoA hidratasa

LCHAD: 3-hidroxiacil-CoA deshidrogenasa

LCKT: 3-cetoacil-CoA tiolasa

LDLR: Receptor de lipoproteínas de baja densidad

MUFA: Ácidos grasos monoinsaturados

NGS: Next generation sequence

Padj: Valor P ajustado

Pb: Pares de base

PUFA: Ácido grasos polinsaturados

qPCR: PCR cuantitativa en tiempo real

RNA-Seq: Secuenciación del ARN

RXR: Receptor retinoide X

SFA: Ácidos grasos saturados

TG: Triglicéridos

UMI: Unique molecular identifier

VLDL: Lipoproteínas de muy baja densidad

1 INTRODUCCIÓN

1.1 Situación actual de la producción de carne de conejo

La producción de carne de conejo en Europa en el año 2022 fue de 219791 toneladas anuales, lo que representa un 20,30% de la producción total de esta carne en el mundo (FAOSTAT, 2022). Además, la producción de carne de conejo representa una importante fuente de proteína en Europa con aproximadamente 90.786.000 de cabezas sacrificadas anuales. Dentro de la Unión Europea (UE) destacan como países líderes en la producción de carne de conejo, España, Francia e Italia. Mientras que fuera de la UE principalmente destacan Rusia y Ucrania. Sin embargo, a nivel mundial la máxima producción de carne de conejo se produce sobre todo en Asia, donde llega a representar el 69,8% del total mundial (figura 1).

El consumo de carne de conejo se ha visto reducido en los últimos años, de hecho, se ha reducido un 19,25% el consumo desde 2021 al 2022 (MAPA, 2024). También se ha reducido la cantidad de cabezas sacrificadas un 5,60% lo cual se podría deber al aumento de precio de los piensos de engorde, mismo que en los últimos 5 años han subido un 4,50% (MAPA, 2024).

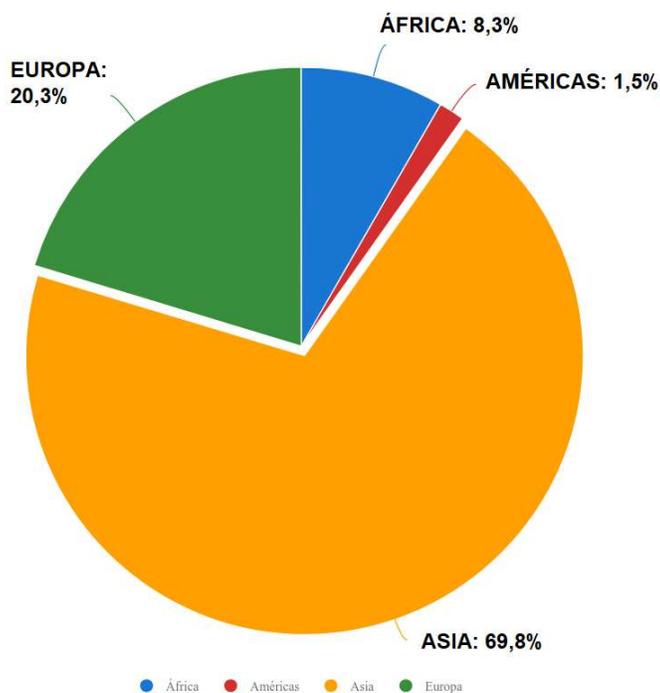


Figura 1. Proporción de producción de carne de conejos y liebres por región. Fuente: FAOSTAT, 2024

Centrándonos en la producción de carne de conejo a nivel nacional, se observa un fuerte decrecimiento de la misma. Aun así, las principales comunidades donde se centra la producción son Castilla y León, Cataluña, Castilla La Mancha y Comunidad Valenciana (figura 2). Sin embargo, los datos existentes indican que el pasado año bajó la producción en Castilla La Mancha y Comunidad Valenciana, pero aumentó en Aragón, Murcia, Andalucía y Cataluña (MAPA, 2024).

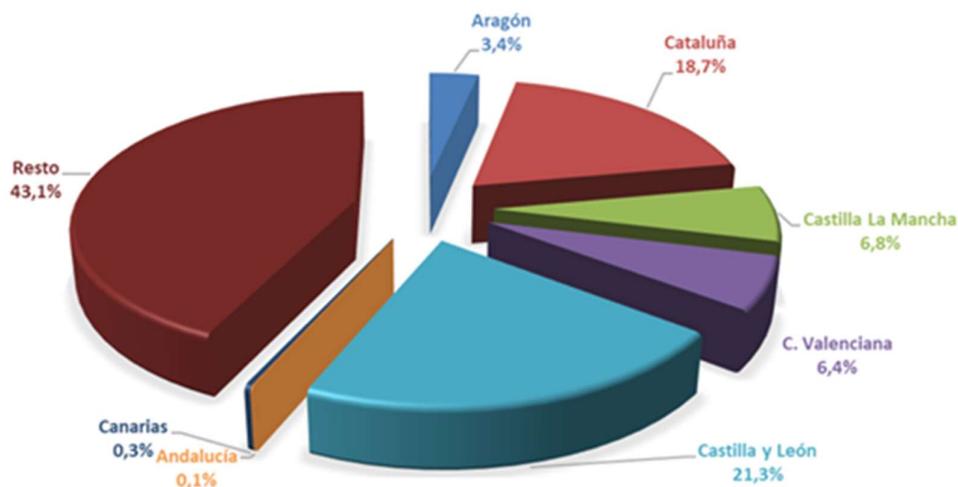


Figura 2. Distribución de la producción de carne de conejo por comunidades autónomas en el año 2021 (toneladas). Fuente: “EL SECTOR CUNÍCOLA EN CIFRAS: Principales Indicadores Económicos”, Subdirección General de Producciones Ganaderas y Cinegéticas, Dirección General.

La carne de conejo representa el quinto tipo de carne más consumido tras el porcino, aves, vacuno y ovino-caprino, aunque guarda un importante componente cultural. España es uno de los principales productores mundiales de carne de conejo, liderando junto con Italia y Francia la producción en la UE (MAPA, 2024).

En los últimos años la evolución productiva de este sector ha venido condicionada por una profesionalización de la actividad, con un descenso de la producción local ligada al autoconsumo y a circuitos cortos de distribución, lo que ha resultado en una concentración del eslabón productor y una reducción considerable del número de explotaciones. Esta situación ha venido acompañada de cambios en las tendencias de sacrificios y producción, en descenso desde 2015 (MAPA, 2024).

1.2 Importancia nutricional y composición de la carne de conejo

La carne de conejo presenta unos valores nutricionales excelentes, entre estos destaca la cantidad de proteínas y altos niveles de aminoácidos (Zotte, 2004), asimismo posee vitaminas y minerales. Por otro lado, la carne de conejo no contiene ácido úrico y tiene bajos niveles de purinas (Hernández et al., 2008). Cabe destacar también la composición de ácidos grasos de la carne de conejo, ya que se caracteriza por un alto porcentaje de grasos poliinsaturados (PUFAs), además presenta menos colesterol que la carne de otros animales como cerdo, vacuno o pollo (Hernández et al., 2008).

La grasa de la carne de conejo contiene principalmente ácidos grasos saturados (SFAs) y PUFA, ambos con porcentajes cercanos al 35%. Sin embargo, los ácidos grasos monoinsaturados (MUFAs) están menos representados (28,50%). En general, los ácidos grasos más abundantes son el oleico (C18:1), palmítico (C16:0) y linoleico (C18:2) con un porcentaje mayor al 20% (Hernández et al., 2008).

Entre los PUFA, el linoleico (C18:2) y el linolénico (C18:3) son conocidos como ácidos grasos esenciales porque los organismos animales no pueden sintetizarlos. El ácido linoleico es el precursor de la familia omega 6 de PUFA, mientras que el ácido linolénico cumple la misma función para la familia omega3, especialmente para los ácidos grasos eicosapentaenoico (EPA) y docosahexaenoico (DHA). De nota, el porcentaje de ácido graso linoleico es aproximadamente diez veces mayor en la carne de conejo que en la de vacuno y cordero y aproximadamente el doble que la cantidad reportada para la carne de cerdo (Enser et al., 1996).

1.3 Importancia de la grasa intramuscular (GIM)

El contenido de grasa intramuscular desempeña un papel esencial en la calidad de la carne ya que influye en su calidad sensorial, tecnológica y nutricional (Wood et al., 2008). La jugosidad y la terneza de la carne son las principales características sensoriales afectadas por la grasa intramuscular. Además, el contenido de grasa intramuscular y su composición en ácidos grasos determinan el valor nutricional de la carne (Wood et al., 2003). Así pues, estrategias para aumentar el contenido en grasa intramuscular pueden considerarse una forma de incrementar la calidad de la carne.

La grasa intramuscular presenta una heredabilidad de moderada a alta (Sellier, 1998) y una alta variabilidad por lo que es posible la mejora genética de este carácter mediante selección (Martínez-Álvaro et al., 2016a). Sin embargo, los programas genéticos requieren la medición de los caracteres de interés en un gran número de animales. Además, el contenido de grasa intramuscular generalmente se determina en animales sacrificados y, por lo tanto, la selección se basa en datos de hermanos completos o medios hermanos. Estas desventajas hacen que la mejora genética del carácter sea menos óptima (Zomeño et al., 2013).

Se han realizado diversos experimentos de selección por GIM, en cerdos (Schwab et al., 2009), ganado vacuno (Sapp et al., 2002) y pollos (Zhao et al., 2007). Después de siete generaciones de selección en cerdos (Schwab et al., 2009), cinco generaciones en pollos (Zhao et al., 2007) y sólo una generación en ganado vacuno (Sapp et al., 2002), todos estos experimentos mostraron importantes respuestas directas a la selección.

Por otra parte, en la Universitat Politècnica de València se realizó un experimento de selección divergente por GIM en el músculo *Longissimus thoracis et lumborum* en conejos durante 10 generaciones. El experimento fue exitoso, obteniendo una respuesta directa a la selección en la 10ª generación de 3,8 desviaciones típicas del carácter. La composición de ácidos grasos tanto del músculo como del hígado se modificó después de la selección (Zubiri-Gaitán et al., 2022). Además, se observó una respuesta correlacionada en el metaboloma plasmático (Zubiri-Gaitán et al., 2023).

Se ha observado que tras la selección de animales para experimentación en base a su GIM sí que existe una relación entre este valor, la jugosidad y lo sabroso de la carne (Martínez-Álvaro et al., 2018). Tales experimentos han sido realizados en cerdo (Schwab et al., 2009), en ganado (Sapp et al., 2002) y en pollo (Zhao et al., 2007).

Las líneas divergentes de conejos fueron criadas contemporáneamente bajo las mismas condiciones ambientales y alimentadas con la misma dieta, por lo que las diferencias encontradas entre ellas y en su metabolismo pueden atribuirse directamente a su composición genética. El conejo es un buen modelo para estudios genéticos en comparación con otras especies de ganado, debido a su corto intervalo generacional y al bajo costo de su canal. Estas líneas divergentes constituyen un material muy valioso para estudiar el efecto genético de los caracteres seleccionados y correlacionados, como la composición de ácidos grasos y la deposición de grasa en la canal.

En este experimento de selección además de obtenerse una respuesta directa a la selección, como hemos comentado anteriormente, se estudiaron respuestas correlacionadas en distintos caracteres. Se encontraron respuestas correlacionadas positivas en el contenido de GIM de otros músculos con diferentes patrones oxidativos y en otros depósitos de grasa (Martínez-Álvaro et al., 2018). A pesar de la diferencia en el contenido de GIM, no se encontraron diferencias en los atributos sensoriales del músculo *Longissimus* a las 9 semanas, lo que puede atribuirse al hecho de que dichos atributos se asocian principalmente a niveles más altos de GIM (Hocquette et al., 2010; Listrat et al., 2016)). Sin embargo, cuando se midió a las 13 semanas de edad si se encontraron diferencias en diversos atributos sensoriales (Martínez-Álvaro et al., 2016b).

También se estudió el metabolismo lipídico de músculos con diferentes patrones oxidativos, tejidos grasos e hígado y encontrándose diferencias metabólicas entre las líneas seleccionadas por alta y baja GIM (Martínez-Álvaro et al., 2017; Martínez-Álvaro et al., 2018). Además, la composición de ácidos grasos tanto del músculo como del hígado se modificó como consecuencias de la selección (Zubiri-Gaitán et al., 2022).

En estas líneas divergentes también se realizaron tres análisis genómicos: dos estudios de asociación del genoma completo (GWAS) (Laghouaouta et al., 2020; Sosa-Madrid et al., 2020a) y un estudio de huellas de selección (Sosa-Madrid et al., 2020b). En artículos de nuestro grupo de investigación (Laghouaouta et al., 2020; Sosa-Madrid et al., 2020a) ha sido reportado que la GIM no depende únicamente de un gen, sino que tiene una naturaleza poligénica y, por tanto, está regulada por la expresión simultánea de múltiples genes en diferentes tejidos lipogénicos (p.ej. tejido adiposo, hígado y músculo), entre ellos el hígado es de gran interés para llevar a cabo este trabajo. Por último, se realizó un análisis metagenómico del contenido del ciego que confirmó la relevancia de la composición del microbioma en la deposición de GIM, identificando cambios tanto en la composición del microbioma como en su funcionalidad (Martínez-Alvarado et al., 2021).

1.4 El papel del hígado en el metabolismo

El hígado posee un papel clave en el metabolismo, actúa como un centro de control para conectar varios tejidos entre ellos el adiposo y el músculo esquelético. Tras la digestión todos los componentes básicos provenientes de la dieta (azúcares, aminoácidos, ácidos grasos, etc.) son transportados al hígado donde tras una serie de reacciones químicas ellos son transformados en otros sustratos. Los sustratos posteriormente tras ciertas actividades serán liberados o almacenados según la necesidad fisiológicas del organismo. Las tres fuentes principales de ácidos grasos (AGs) hepáticos son los lípidos de la dieta, los AGs derivados del tejido adiposo y los AGs sintetizados vía *de novo* (Alves-Bezerra & Cohen, 2018). Los lípidos derivados de la dieta se hidrolizan en la luz intestinal. Tras la absorción intestinal, los AGs se reesterifican para formar moléculas de triglicéridos (TGs), que se empaquetan en quilomicrones y se transportan principalmente al músculo y al tejido adiposo. Sin embargo, los TGs presentes de los quilomicrones son transportados al hígado y procesados intracelularmente, lo que provoca la liberación de AGs en los hepatocitos (Figura 3).

Por otra parte, los carbohidratos y en particular la glucosa, se utilizan en la lipogénesis *de novo* hepática para la producción de AGs. Para ser metabolizados, los AGs se activan para formar moléculas de acil-CoA, las cuales pueden oxidarse o incorporarse a lípidos complejos. Los TGs sintetizados localmente pueden almacenarse en gotas de lípidos intracelulares o empaquetados en lipoproteínas de muy baja densidad (VLDL) y secretados al plasma.

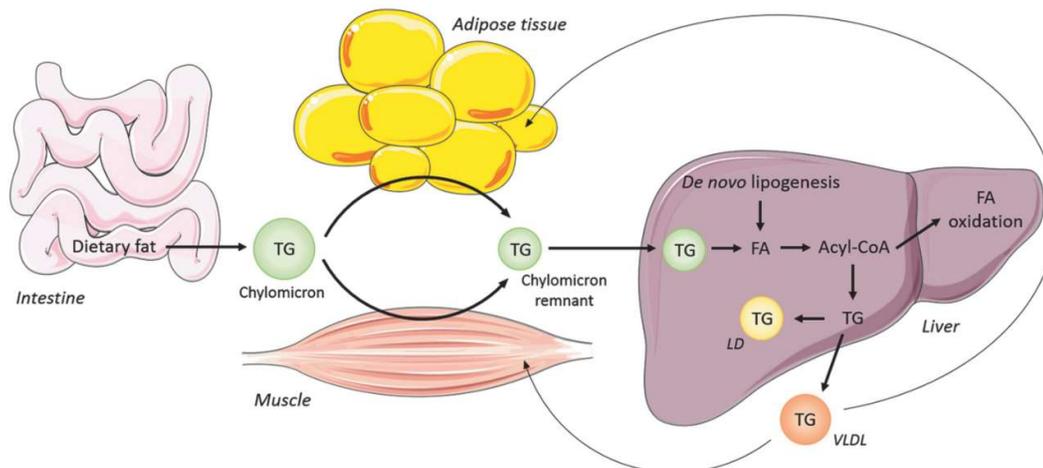


Figura 3. Principales fuentes de ácidos grasos hepáticos. Adaptado de Alves-Bezerra & Cohen, 2018.

Como proceso destacado esta la beta oxidación de ácidos grasos que se produce en las mitocondrias hepáticas (Rui, 2014; Gondret et al., 1997). Además, otro proceso fundamental es la síntesis *de novo* (o lipogénesis) de ácidos grasos, dado que el hígado es el principal órgano encargado de transformar los carbohidratos en ácidos grasos. Este proceso consta de una serie de pasos los cuales son primeramente una glicólisis mediante la cual la glucosa se convierte en piruvato en el citoplasma, tras ello ha de entrar a la mitocondria donde se transforma en acetil-CoA gracias a la enzima piruvato deshidrogenasa. A continuación, esta se combina con oxalacetato para formar citrato el cual sale de la mitocondria, una vez en el citoplasma se vuelve a descomponer en acetil-CoA y oxalacetato por acción de la enzima ATP-citrato liasa. Después se carboxila en acetil-CoA en malonil-CoA por la acetil-CoA carboxilasa y finalmente el enzima ácido graso sintasa une una molécula de acetil-CoA y varias de malonil-CoA para elongar la cadena de carbonos y llegar a formar un ácido palmítico (Song et al., 2018).

La lipogénesis es un proceso que está controlado por varios elementos reguladores tales como factores de transcripción, los co-reguladores y las hormonas. Comenzando por los factores de transcripción resaltamos la familia SREBP, formada por SREBP-1a, -1c y -2 que son reguladores importantes del metabolismo lipídico (Rui, 2014). Tanto SREBP-1a como SREBP-1c están codificados por un mismo gen y tienen diferentes extremos amino terminales. SREBP-2 está codificado por un gen independiente, y interesantemente tanto SREBP-1c como SREBP-2 se expresan abundantemente en el hígado. SREBP-1c activa los genes que controlan la síntesis de ácidos grasos y TGs, y SREBP-2 activa los genes que controlan la biosíntesis de colesterol. Mientras que SREBP-1b promueve la síntesis de ácidos grasos y colesterol. También podemos destacar el rol del factor de transcripción ChREBP, el cual se une y activa el promotor de "L-PK" en los hepatocitos. L-PK es una enzima glicolítica clave. ChREBP también estimula la expresión de genes lipogénicos, incluida la enzima málica, y tres genes como *ACL*, *ACC*, *FAS*, *SCD1* y *ELOVL5* (Lizuka et al., 2004). Asimismo, otro factor de transcripción a destacar sería PPAR γ , el cual promueve la expresión de genes relacionados con el almacenamiento de grasa y la sensibilidad a la insulina (Rui, 2004). Por lo que respecta a los reguladores destacamos los LXRs, de estos se conocen dos isoformas (α y β) en roedores, y cada isoforma forma heterodímeros con el receptor de retinoide X (RXR) para activar sus genes dianas, esta activación se produce mediante metabolitos del colesterol que se llaman oxisteroles. Estos reguladores, también estimulan la biosíntesis *de novo* de ácidos grasos (Calkin & Tontonoz, 2012). De hecho, LXR se une directamente al promotor de SREBP1 y aumenta la expresión de SREBP-1c. También estimula la expresión de ChREBP (Rui, 2004). Otro coregulador interesante es PGC-1 α , el cual actúa juntamente con PPAR γ para modular el metabolismo energético (Zhang et al., 2004). Finalmente, en cuanto a las hormonas resaltamos a la insulina que activa la acetil-CoA carboxilasa, y como antagonistas a este proceso están el glucagón y la adrenalina que la inhiben (Song et al., 2018).

Además, el metabolismo hepático está muy controlado ya sea por los múltiples nutrientes, las hormonas y las señales neuronales. En relación con lo último se describe como el sistema nervioso regula de forma directa el metabolismo del hígado por dos vías, en la primera vía actúa el sistema nervioso simpático que favorece la movilización de los sustratos a tejidos extrahepáticos, y la segunda vía es el sistema nervioso parasimpático cuya función es la antagonista de la primera (Rui, 2014).

1.5 Regulación de la expresión génica

La regulación de la expresión génica es un proceso complejo que permite a las células controlar qué genes se activan o reprimen, y en qué medida. La expresión de un gen eucariota implica una serie de pasos previos a la síntesis de una proteína. Estos pasos incluyen la transcripción del gen en el producto primario de ARN, el procesamiento del transcrito inicial para eliminar las secuencias de intrones y crear el extremo 3' maduro, el transporte del transcrito de ARNm procesado al citoplasma y, a continuación, la traducción del ARN mensajero en proteína. Con pocas excepciones, todos los genes que codifican proteínas siguen esta vía. En una célula típica de mamífero, el ARN mensajero (ARNm) que se traduce en proteínas constituye alrededor del 4% de la masa total de ARN y, aparte del 80% del ARN ribosómico (ARNr), otros ARN operativos constituyen el resto (Wu et al., 2014). No obstante, las velocidades y eficiencias del paso transcripcional, el procesamiento del ARN, el transporte nuclear y la traducción de proteínas pueden diferir considerablemente de un gen a otro.

Asimismo, regulación de la expresión génica implica una serie de eventos moleculares que actúan a nivel transcripcional, postranscripcional, traslacional y traduccional y postraduccional. Entre los elementos claves de este proceso se encuentran los factores de transcripción, modificadores epigenéticos, enzimas de procesamiento del ARN, ribosomas y proteínas quinasas/fosfatasa. Sobre todo, el resultado de la regulación de la expresión génica es la producción de proteínas con funciones específicas necesarias para el funcionamiento celular y el desarrollo. A nivel transcripcional el control de la expresión génica está modulado por proteínas que pueden clasificarse en dos categorías: proteínas de unión al ADN específicas de la secuencia, entre las que se incluyen los factores de transcripción, y otros factores reguladores claves como las proteínas de unión TATA o TATA-binding. Sin embargo, la regulación postranscripcional ha cobrado importancia en muchos procesos biológicos por su capacidad para obtener una respuesta rápida a señales celulares y ambientales.

Además, la expresión de un gen puede controlarse en múltiples niveles, incluyendo el empalme del ARNm, la estabilidad del ARNm, la traducción y los eventos postraduccionales como la estabilidad y la modificación de proteínas (Day & Tuite, 1998) o incluso diferentes niveles (p.eje. genoma, transcriptoma, proteoma, efectos ambientales, metaboloma, etc.). Por tanto, es importante valorar la función del ARNm para saber si esta es determinante en la expresión génica o no, esto se analiza mediante la correlación entre los valores de ARNm y los niveles de proteína. Es razonable asumir que la cantidad de ARNm tiene una correlación con la expresión de la proteína que codifica, sin embargo, esta relación no es perfecta, ya que existen otros procesos relacionados con la transcripción (splicing, modificaciones postraduccionales, etc.) poseen un papel muy relevante en este proceso (Buccitelli & Selbach, 2020).

1.6 Secuenciación del ARN (RNA-Seq) y sus aplicaciones

Las técnicas de secuenciación masiva o de next-generation sequence (NGS) han revolucionado los estudios genómicos. Los estudios de expresión génica implican el análisis de la actividad de los genes en las células, tejidos u organismos, utilizándose para investigar una amplia gama de procesos biológicos. Existen diferentes técnicas para estudiar la expresión génica, como por ejemplo el análisis de microarrays, secuenciación del transcriptoma (también conocida como RNA-Seq) y PCR cuantitativa en tiempo real (qPCR). Estos métodos han sido usados para medir la abundancia de diferentes moléculas de ARN en una muestra e identificar qué genes se expresan y a qué nivel. En la última década la aparición de técnicas de secuenciación del RNA (del inglés RNA-Seq) nos ha permitido analizar cómo cambian los niveles de expresión y se ha convertido en una herramienta muy poderosa, pero que a su vez todavía tiene un coste elevado muestra. No obstante, el RNA-Seq puede utilizarse en combinación con otros métodos de genómica funcional para enriquecer el análisis de la expresión génica (Figura 4).

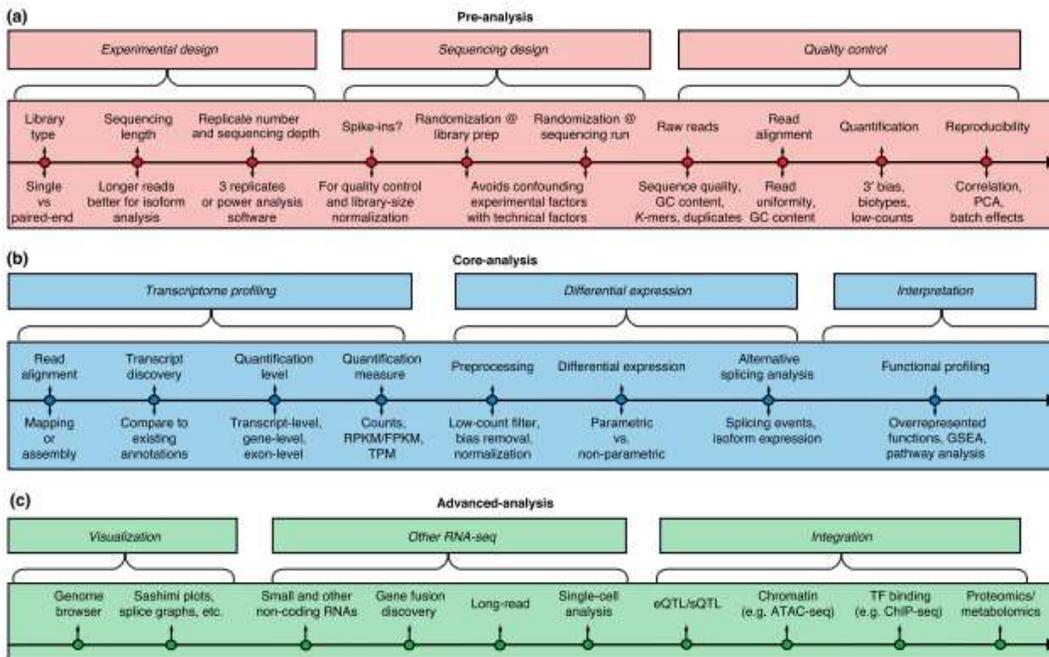


Figura 4. Mapa general para el análisis computacional de un RNA-Seq. Los tres bloques principales son preanálisis, análisis central y análisis avanzado. A) incluye diseño experimental, estrategia de secuenciación y control de calidad. B) Los análisis principales incluyen perfiles de transcriptoma, expresión genética diferencial y perfiles funcionales. C) El análisis avanzado incluye visualización, otras tecnologías de secuenciación de ARN e integración de datos. De: (Conesa et al., 2016)

Por otro lado, tenemos las aplicaciones del RNA-Seq, entre ellas encontramos la posibilidad de descubrir nuevos transcritos, aunque este proceso es realmente un reto complejo debido a las lecturas cortas también es un campo interesante, ya que permite conocer o indagar sobre transcritos desconocidos que pueden tener una aplicación importante en el desarrollo de un organismo. Además, existe la posibilidad de hacer una reconstrucción *de novo* de los transcritos, esto es útil

cuando el genoma de referencia no está disponible o está incompleto. Otra aplicación y probablemente la más común es realizar una cuantificación de los transcritos ya que mediante esta se puede llegar a conclusiones acerca del estado del organismo de interés, así como acerca de si algún gen se encuentra con una expresión incrementada o disminuida y entonces asociarlo con el fenotipo que presenta (Conesa et al., 2016).

A partir de la cuantificación mencionada previamente se puede realizar un análisis diferencial de la expresión de genes según diferentes muestras de interés, así pues, observando la magnitud de tales diferencias e intentando buscar posibles explicaciones para las mismas. Otras de las utilidades del RNA-Seq es que permite encontrar alternativas de splicing e isoformas. Análogamente también se descubren fusiones de genes que pueden tener un alto impacto en las funciones biológicas y, a su vez si explicase la coexpresión de alguna proteína. Por último, encontramos la posibilidad de encontrar RNAs de pequeño tamaño (small RNAs), en este grupo entran muchos tipos de RNA los cuales ejercen funciones muy diferentes y permiten la modulación de proteínas y el silenciamiento, entre otras funciones.

Acorde al interés de este estudio, existen metodologías alternativas para cuantificar los niveles de expresión génica a un menor coste. En concreto, la secuenciación de regiones no traducidas del extremo 3' UTRs (3' RNA-Seq) es considerada como un proxy para la expresión del gen completo.

Actualmente la técnica 3' RNA-Seq se está usando sobre todo para estudios de identificación de genes diferencialmente expresados (DEGs). De hecho, esta técnica permite obtener información exacta de la expresión de un gen con menos lecturas (McClure et al., 2023). Asimismo, esta metodología es especialmente útil para cuantificar los niveles de expresión génica en muestras con fragmentación y degradación del ARN, en las que los métodos tradicionales de RNA-Seq suelen tener dificultades para generar datos fiables (Tandonnet & Torres, 2017).

El proceso de laboratorio mediante el cual se amplifica la región 3' de los fragmentos de mRNA es mostrado en la siguiente figura. A modo de resumen la preparación de librería implica la amplificación de los fragmentos del RNA mensajero mediante el uso de random primers y un primer poliT. Tras una primera PCR para la amplificación de los fragmentos de mRNA, prosigue una segunda PCR que contiene el adaptador necesario para el posterior paso de secuenciación.

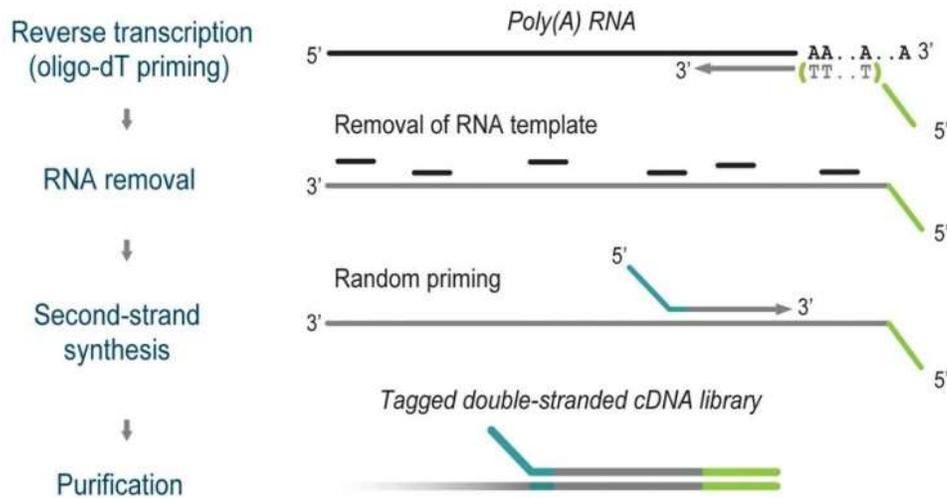


Figura 5. Representación de los pasos principales para amplificar la región 3' UTR de los fragmentos de mRNA. Fuente: Adaptado de: Moll et al. (2014).

Debido a que el protocolo enriquece una región concreta de los transcritos, el número de duplicados que pueden aparecer es bastante alto, especialmente en muestras de baja calidad, donde la cantidad de fragmentos disponibles puede ser menor. Para solventar esta posible situación, el random primer que amplifica desde el extremo 5' contiene ligado un marcador de molécula única (UMI – Unique Molecular Identifier), compuesto por 6 pb y separado del primer por un separador de secuencia conocida. La presencia de estos identificadores moleculares permite en el análisis bioinformático posterior diferenciar entre duplicados de una molécula o fragmentos distintos, evitando sesgar la información durante el paso de cuantificación de fragmentos por gen (figura 6).

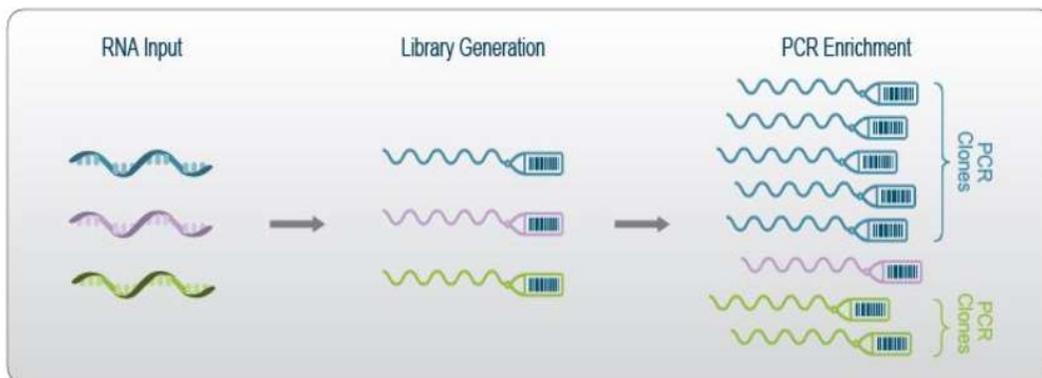


Figura 6. Cada molécula está identificada con un marcador molecular propio, permitiendo diferenciar aquellos duplicados sintéticos de fragmentos con coordenadas idénticas. Fuente: faqs.lexogen.com (<https://faqs.lexogen.com/faq/what-are-unique-molecular-identifiers-umis>).

1.7 Análisis y visualización de genes diferencialmente expresados

Para analizar los datos obtenidos por RNA-Seq no existe una guía única, de hecho, hay varias variaciones entre los diferentes pipelines que se usan para el tratamiento y visualización de los datos (Conesa et al., 2016).

Para el análisis de los genes diferencialmente expresados (DEGs) se ha de seguir una serie de pasos en los cuales el primero es realizar un control de calidad, eliminación de outliers y realizar un filtrado. Existen dos aproximaciones la identificación de los DEGs, la primera es por comparación a pares entre 2 grupos y la segunda es por la varianza entre grupos (Conesa et al., 2016).

DESeq2 ofrece una solución integral y versátil para el análisis a nivel genético de datos de RNA-Seq. Además, DESeq2 ofrece un rendimiento consistente en una amplia gama de tipos de datos, pudiéndose aplicar tanto para estudios pequeños con pocas réplicas como para estudios observacionales grandes.

La heurística de DESeq2 para la detección de valores atípicos ayuda a reconocer genes para los cuales las suposiciones del modelo no son adecuadas y, por lo tanto, evita errores de tipo I causados por estos. La incorporación de estas estrategias en el marco de los modelos lineales generalizados (GLM) permite el tratamiento de diseños tanto simples como complejos (Love et al., 2014). Por ejemplo, la herramienta DESeq2 permite detectar y corregir las estimaciones de dispersión comparándolas con la media de todas las muestras, mediante DESeq2 se obtienen varios parámetros estadísticos con el P-value, el logaritmo del Fold Change (LFC), el Valor P ajustado (P_{adj}) que tiene en cuenta el control de la tasa de falsos positivos, todos los cuales son de gran interés para identificar si un gen está diferencialmente expresado. De hecho, el paquete DESeq2 está bien integrado en la infraestructura de Bioconductor y viene con documentación extensa, incluida una viñeta que demuestra un análisis completo paso a paso y analiza casos de uso avanzados.

Un punto crítico es el estimador de contracción para el fold change en el análisis de expresión diferencial, el cual ofrece una solución sólida y estadísticamente bien fundada. Esto es valioso para muchas tareas de análisis posteriores, incluida la clasificación de genes para estudios de seguimiento y la asociación de cambios con otras variables de interés. Además, la transformación rlog, que implementa la reducción del fold change por muestra, facilita la visualización de diferencias, por ejemplo, en heatmaps, y permite la aplicación de una amplia gama de técnicas que requieren datos de entrada homocedásticos, incluidos las técnicas de aprendizaje automático u ordenación, como el análisis de componentes principales y la agrupación (Love et al., 2014).

Sin embargo, existen diferentes softwares online que permiten realizar el análisis de datos de RNA-seq. Independientemente de la herramienta usada hay que tener en cuenta que se ha de realizar un filtrado previo de los conteos bajos, este proceso se puede realizar a través funciones implementadas en el R base, para así tener a punto los datos antes de introducirlos en el programa elegido. Para dichos análisis por están disponibles varias herramientas como pueden ser edgeR, DESeq y cuffdiff (Koch et al., 2018).

Por último, se puede realizar un análisis de enriquecimiento funcional de los términos de ontología génica (GO) y pathways. Para ello se usan herramientas como son Enrichr, DAVID GOrilla, STRING, ClueGO y Cytoscape. Los términos GO aportan información acerca de procesos biológicos, funciones moleculares y componentes celulares, en estos tres grupos clasifica los genes. Estas herramientas de análisis se basan en un test hipergeométrico que tiene en cuenta el total de genes, el número de genes objetivo, los que están asociados a un término GO y la intersección de los genes objetivo y los que están en la lista de GO. Este agrupamiento al final se acompaña del FDR que permite distinguir los agrupamientos significativos de aquellos que no lo son (Koch et al., 2018).

1.8 Hipótesis y objetivos

Teniendo presente todo el background anterior en esta investigación la hipótesis planteada fue que las diferencias de deposición de grasa intramuscular en dos líneas de conejos divergentes (alta y baja) pueden estar asociadas con la expresión génica diferencial en el hígado. Para testear esta hipótesis, se determinaron los niveles de expresión en muestras de hígado mediante la metodología conocida como 3' RNA-Seq. Subsecuentemente, el objetivo principal de esta investigación fue identificar genes diferencialmente expresados entre las dos líneas divergentes para grasa intramuscular, y sobre todo aquellos genes que se asocien la variación del contenido de grasa intramuscular.

2 MATERIALES Y MÉTODOS

2.1 Animales y diseño experimental

En este experimento se utilizaron un total de 48 conejos seleccionados divergentemente por GIM, 24 de la línea alta y 24 de la línea baja, dentro de cada una con 12 machos y 12 hembras, respectivamente). Todos los animales fueron provenientes de la novena generación de la selección, sacrificados a las 9 semanas de edad tras 4 horas de ayuno. En efecto, el sacrificio de los animales se llevó a cabo en el propio matadero de la Universitat Politècnica de València y mediante desangrado tras un previo aturdimiento eléctrico.

Inmediatamente después del sacrificio se tomaron las muestras de hígado las cuales fueron alicuotadas en tubos criogénicos de 2 mL añadiendo 500 µl de RNA later (QIAGEN). Entonces las muestras se almacenaron a -80 °C para posteriores análisis.

Las muestras fueron procesadas por la empresa biotecnológica Seqplexing. El RNA total fue aislado del hígado usando el Kit "MagMAX mirVana Total RNA Isolation" (Thermo Fisher Scientific, Waltham, MA, USA) siguiendo las recomendaciones del fabricante. La calidad del ARN se comprobó con el sistema bioanalizador agilent 2100 (Hercules, CA, EE.UU.), y las muestras con un número de calidad del ARN (RQN) superior a 7 se seleccionaron para el experimento de secuenciación. Las librerías fueron preparadas usando la metodología de amplificación de 3´mRNA con una combinación de primers poliT y primers aleatorios. La calidad de las librerías fue evaluada usando QIAxcel Advanced System (QIAGEN), la secuenciación de la librería fue al 10x de profundidad usando la plataforma de Illumina NovaSeq X y de tipo paired-end con 150 pb (5 millones de lecturas por muestra en promedio). Los datos crudos obtenidos en formato FASTQ se usaron como input en los subsiguientes análisis bioinformáticos.

2.2 Análisis bioinformáticos y estadísticos

Las secuencias crudas sin procesar (archivos raw FASTQ) se procesaron utilizando un sistema bioinformático interno desarrollado por la empresa Seqplexing con parámetros predeterminados. Brevemente, la metodología consistió en: (1) control de calidad con el software FastQC v0.12.1 (Andrews, 2010) y trimmeado de FASTQ con la herramienta Cutadapt v1.2.0 (Martin, 2011), (2) procesamiento UMI (Unique Molecular Identifier) utilizando la herramienta umi v0.2.3 (Smith et al., 2017), (3) alineación contra el genoma de referencia (*Oryctolagus Cuniculus*) "UM_NZW_1.0 (RefSeq GCF_009806435.1)" utilizando el software STAR v2.7.11a (Dobin et al., 2013), y (4) cuantificación de la expresión mediante el programa HTSeq-count v2.0.5 (Anders et al., 2015).

La matriz de counts obtenida se utilizó para el análisis de expresión diferencial utilizando el paquete DESeq2 v1.42.0 R (Love et al., 2014). Se aplicó un prefiltrado de genes siguiendo criterios estándar (Chung et al., 2021), aquellos que no tenían al menos 10 counts por gen fueron eliminados, así como considerando sólo genes que codifican proteínas; finalmente se retuvieron 15430 genes. La normalización de los counts se realizó utilizando el método de mediana de las ratios implementado en el paquete DESeq2. En interés se centró en determinar las diferencias de expresión génica entre líneas divergentes para GMI aplicando un modelo binomial negativo, incluyendo como efectos fijos el sexo (2

niveles) y la línea (2 niveles). Todos los genes con un cambio absoluto de fold change (FC) de al menos 1,5 ($|\log_2FC| \geq 0,58$) y un FDR $P_{adj} \leq 0,05$ se seleccionaron como DEGs. Además, la lista de DEGs se sometió a análisis funcionales usando el plug-in ClueGO v2.5.10 en el software Cytoscape v3.10.1 (Bindea et al., 2009; Shannon et al., 2003) para identificar términos de ontología genética (GO) y rutas metabólicas sobrerrepresentadas. En efecto, la significancia estadística se evaluó mediante el test hipergeométrico utilizando el método de Benjamini y Hochberg para corrección por múltiples tests.

3 RESULTADOS Y DISCUSIÓN

3.1 Descripción del control de calidad de la secuenciación

En el experimento de secuenciación se generaron datos 3' RNA-Seq del hígado de 48 conejos divergentes para GIM (24 de la línea alta y 24 de la línea baja). Tras del control de calidad observamos una media de 30,24 millones de lecturas trimmeadas según total de muestras y una media del 92,82 % (rango entre el 91,37 % y 93,88 %) de lecturas asignadas de forma única.

La evaluación del número de duplicados tras el procesamiento de los UMIs se muestra en la figura 8. La media de fragmentos únicos se encuentra alrededor del 45%, lo que entra dentro de los parámetros obtenidos regularmente. El uso de UMIs garantiza que los reads duplicados que se encuentran en el análisis pertenecen a moléculas distintas, asegurando una mayor confianza en los resultados finales.

Tras la realización de los alineamientos con el software STAR, se obtienen distintas métricas en función de la especie analizada (figura 7). En nuestro estudio obtuvimos porcentajes de asignación única a genes desde aproximadamente 70%. Este valor depende enormemente de la anotación existente, especialmente en relación con las regiones 3' UTR.

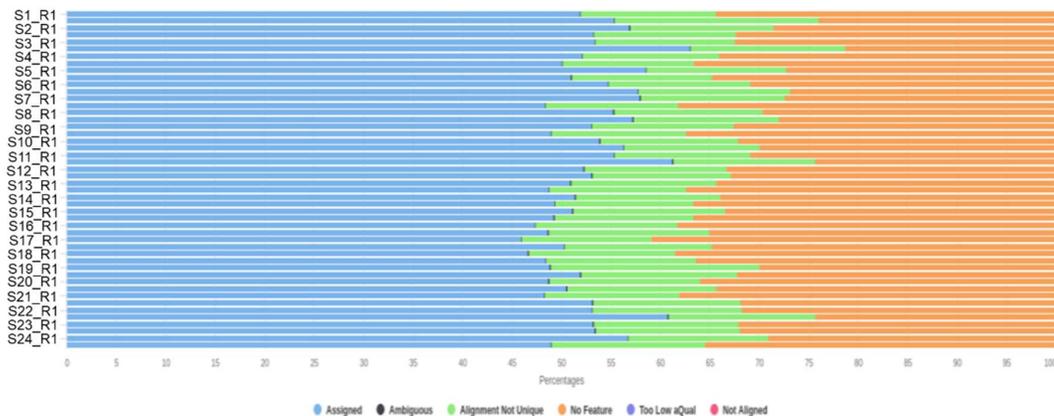


Figura 7. Estadísticos obtenidos en cuanto a la calidad de los alineamientos generados, midiendo las lecturas que no se encuentran dentro de regiones transcritas, o lecturas no alineadas.

Estudios anteriores han demostrado que la metodología de 3' RNA-Seq proporciona una alternativa al análisis de transcritos completos teniendo niveles similares de reproducibilidad que el método estándar de RNA-Seq. Además, el método de 3' RNA-Seq es útil en la detección de transcritos cortos (Ma et al., 2019), así como superior al RNA-Seq estándar en casos de datos dispersos (sparse data), ya que 3' RNA-Seq se basa en el enriquecimiento de los extremos 3' del transcrito y su uso se considera como un proxy de la expresión de todo el gen (McClure et al., 2023).

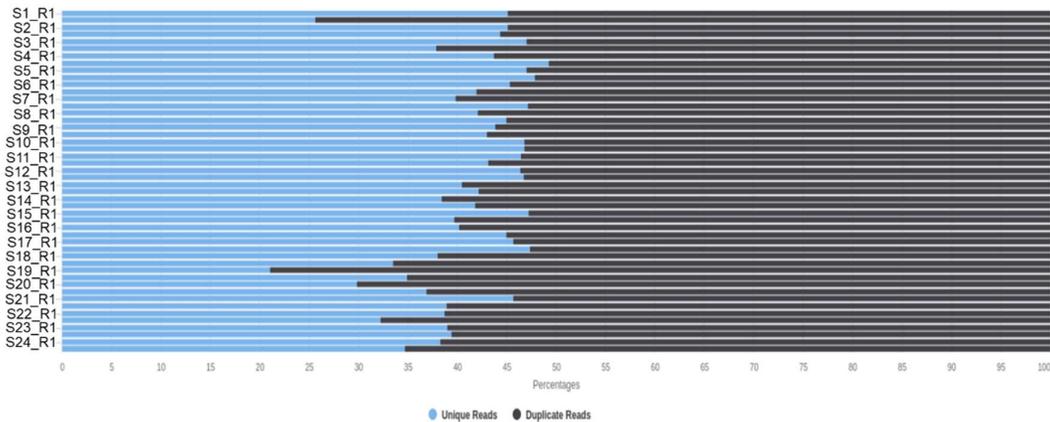


Figura 8. Porcentaje de lecturas únicas y lecturas duplicadas dentro del total de secuencias disponibles para la muestra.

3.2 Análisis de expresión génica diferencial

Un total de 15430 transcritos fueron sometidos al análisis de expresión diferencial usando el paquete DESeq2 (Anexo 1). Este trabajo se enfoca en estudiar las diferencias de expresión génica entre líneas divergentes (alta y baja) según el contenido de GIM. Se identificaron un total de 420 DEGs (FC de 1,5 y $P_{adj} < 0,05$) entre las dos líneas divergentes (Figura 9). Se identificaron un total de 420 DEGs (FC de 1,5 y $P_{adj} < 0,05$) entre las dos líneas divergentes (Figura 9). El gráfico volcano muestra la comparación por pares (alta versus baja), 168 DEGs mostraron una expresión incrementada (168/420, puntos verdes) mientras que 252 DEGs (252/420, puntos rojos) exhibieron una expresión disminuida, y el resto fueron los genes ($n = 15010$) que no superaron el umbral (puntos negros).

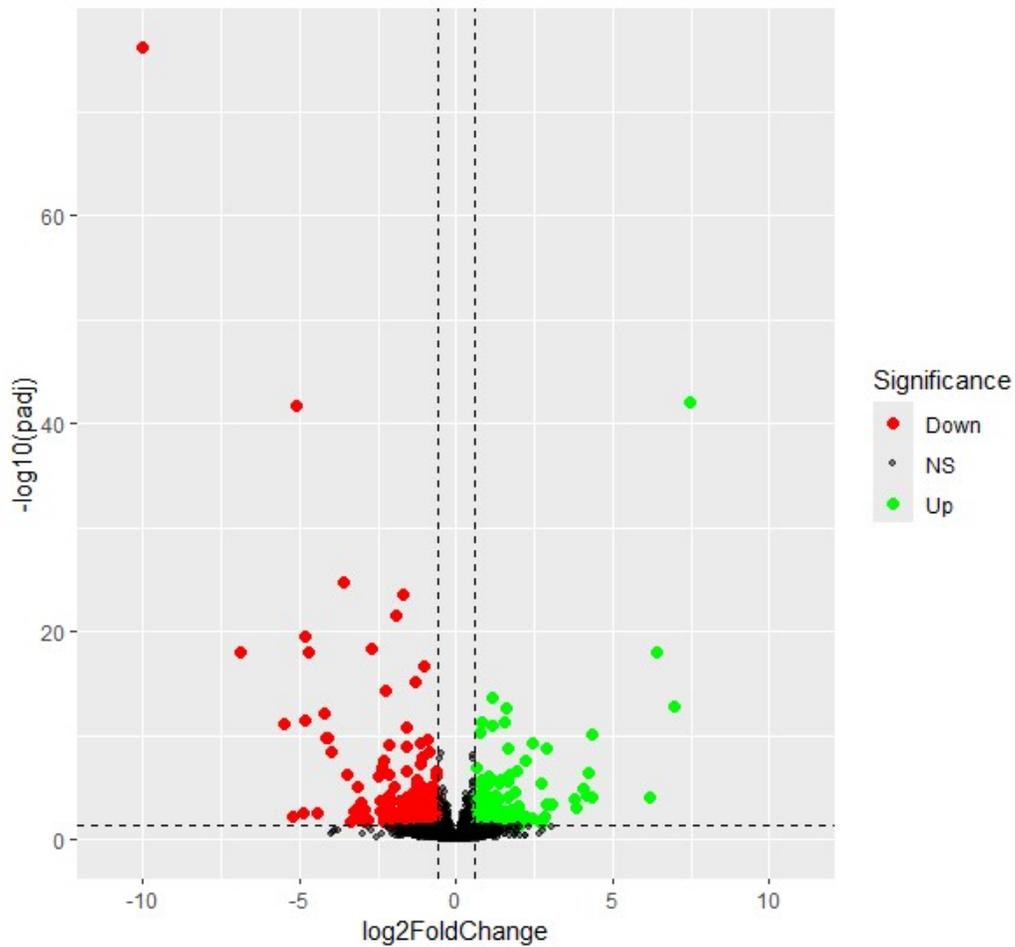


Figura 9. Genes diferencialmente expresados provenientes del 3'RNA-Seq de hígado entre líneas altas y bajas para IGM. Las líneas discontinuas indican los valores del umbral.

En general, observamos más genes con expresión disminuida versus genes con expresión aumentada. El top 20 de DEGs con menor valor P ajustado (P_{adj}) por el método vía Benjamini–Hochberg es listado en la Tabla 1. Los resultados completos para el total de 15430 transcritos pueden ser consultados en el anexo 1.

Tabla 1. Top 20 de genes diferencialmente expresados (por valor P) en hígado entre líneas divergentes según contenido de grasa intramuscular.

Symbol	Log ₂ FC	Padj	Descripción
LOC100343656	-10,00	8,16E-77	butyrophilin subfamily 1 member A1
LOC127489014	-5,11	1,87E-42	kelch-like protein 11
LOC100343733	-3,60	1,85E-25	pancreatic alpha-amylase
LOC100351163	-1,72	2,81E-24	HLA class II histocompatibility antigen, DQ beta 1 chain
GTF2H4	-1,92	2,51E-22	general transcription factor IIH subunit 4
LOC127486729	-4,84	2,61E-20	zinc finger protein 717
MRPL35	-2,72	4,30E-19	mitochondrial ribosomal protein L35
SLC44A4	-4,72	9,24E-19	solute carrier family 44 member 4
MRPL15	-6,90	9,44E-19	mitochondrial ribosomal protein L15
TYRO3	-1,03	2,01E-17	TYRO3 protein tyrosine kinase
LOC100342501	7,49	7,87E-43	polyubiquitin-B
LOC100347496	6,41	9,24E-19	trifunctional enzyme subunit beta, mitochondrial
RPS14	1,12	2,41E-14	ribosomal protein S14
LOC127486578	6,94	1,81E-13	intraflagellar transport protein 46 homolog
TPMT	1,61	2,34E-13	thiopurine S-methyltransferase
C19H17orf58	1,53	4,77E-12	chromosome 19 C17orf58 homolog
UFC1	0,82	5,14E-12	ubiquitin-fold modifier conjugating enzyme 1
PRRG4	1,17	1,25E-11	proline rich and Gla domain 4
TMUB2	0,78	5,07E-11	transmembrane and ubiquitin like domain containing 2
LOC127490114	4,33	8,98E-11	formin-binding protein 4-like

Examinando el top 20 de DEGs según las comparaciones entre la línea alta versus la baja, detectamos que 9 de ellos comienzan con el prefijo "LOC" existiendo escasa información en la literatura sobre ellos. No obstante, basándonos en los cambios de expresión y en la descripción de los genes, discutiremos algunos de ellos como sigue. El gen butyrophilin subfamily 1 member A1 (*LOC100343656* o *BTN1A1*) mostró notables diferencias de expresión ($\log_2FC = -10,00$) entre las dos líneas y fue el gen más significativo estadísticamente ($P_{adj} = 6,16E-77$). *BTN1A1* permite la unión de receptores de señalización y participa en la regulación de la secreción de gotas de leche. Las moléculas butirofilinas han sido reconocidas como una nueva familia de inmunorreguladores (Abeler-Dörner et al., 2012). Además, ha sido reportado que el knockout de *BTN1A1* altera la formación de gotas de lípidos y la composición de fosfolípidos en células epiteliales mamarias bovinas (Han et al., 2020). El gen trifuncional enzyme subunit beta, mitocondrial (*LOC100347496*) mostró una expresión incrementada ($\log_2FC = 6.41$ y $P_{adj} = 9, 24E-19$) en la línea alta versus línea baja. Este gen facilita la actividad aciltransferasa y es implicado en la beta-oxidación de ácidos grasos (<https://www.ncbi.nlm.nih.gov/gene/?term=LOC100347496+gene>). La proteína trifuncional mitocondrial es heterotrimérica y consiste en cuatro subunidades α y cuatro subunidades β y cataliza tres de las cuatro reacciones de acortamiento de cadenas en la β -oxidación mitocondrial de los ácidos grasos de cadena larga (Rector et al., 2008).

Considerando esto la identificación de genes candidatos del metabolismo lipídico o carbohidratos, deposición de grasa o transportadores, en el estudio encontramos, por ejemplo, que el gen perilipin2 (*PLIN2*) es interesante para explicar la variación en GIM. Consistente con nuestros resultados la expresión *PLIN2* resultó disminuida en la línea alta ($\log_2FC = -1.35$ y $P_{adj} = 4,75E-03$, anexo 1). En estudios previos ha sido notado que el nivel de proteína de *PLIN2* desempeña un rol clave en el control dinámico de las gotas lipídicas y su relación con las mitocondrias (Xu et al., 2019). Interesantemente, *PLIN2* es conocido como uno de miembro de los guardianes (gatekeepers) de la lipólisis intracelular. Además, en un GWAS previo sobre la composición de ácidos grasos usando dos poblaciones divergentes de conejos (incluyendo los 48 animales de nuestro estudio), *PLIN2* fue reportado como uno de los genes candidatos funcionales (Laghouaouta et al., 2020). Otro gen detectado con un incremento de expresión en la línea alta ($\log_2FC = 0.82$ y $P_{adj} = 1,35E-02$, anexo 1) fue *ELOVL6* fatty acid elongase 6 (*ELOVL6*). *ELOVL6* participa en la elongación de los ácidos grasos de cadena larga, pero también juega un papel muy importante en el metabolismo lipídico y la regulación de las funciones celulares (Matsuzaka and Shimano, 2009). Respecto a moléculas transportadoras, detectamos que varios miembros de la superfamilia de transportadores de solutos (SLC) fueron DEGs (por ejemplo, *SLC26A6*, *SLC16A11*, *SLC44A4*, entre otros). La superfamilia SLC comprende genes que codifican para transportadores unidos a membranas, los cuales desempeñan funciones esenciales en el transporte de múltiples sustratos (por ejemplo, aminoácidos, glucosa, ácidos grasos y lípidos, acetil coenzima A y vitaminas) a través de las membranas celulares (He et al, 2009). Entre los miembros de SLC, resaltamos los cambios de expresión del gen solute carrier family 44 member 4 (*SLC44A4*) el cual fue nominado dentro del top 20 de DEGs ($\log_2FC = -4.72$ y $P_{adj} = 9.24E-19$, tabla 1). En la literatura *SLC44A4* es también conocido con varios alias (*C6orf29*, *CTL4*, *DFNA72*, *FLJ14491*, *NG22*, *TPPT*). Hasta donde nosotros sabemos la función de *SLC44A4* en el hígado no ha sido

previamente descrita. Sin embargo, la proteína codificada por este gen puede ser una proteína transportadora (transmembrana dependiente de sodio) implicada en la captación de colina-acetilcolina. Además, *SLC44A4* ha sido caracterizado como un sistema de transporte predominante para la captación de pirofosfato de tiamina (TPP, forma fosforilada de vitamina B1) generada por la microbiota en los colonocitos tanto en humanos como en ratones (Sabui et al., 2024). También, el gen adipocyte fatty acid binding protein (*FABP4*) es otro candidato interesante en el transporte y metabolismo de los ácidos grasos. Los niveles de expresión de *FABP4* fueron inferiores en la línea alta versus la baja ($\log_2FC = -1.52$ y $P_{adj} = 2,35E-02$, anexo 1). *FABP4* (también conocido como *A-FABP* o *aP2*) codifica para una proteína de unión a ácidos grasos que se encuentra en los adipocitos. En tanto, *FABP4* se une a los ácidos grasos de cadena larga y a otras moléculas lipofílicas, facilitando su transporte dentro de las células. Asimismo, este gen puede actuar como proteína transportadora de los ácidos grasos generados por la lipólisis a partir de gotas de lípidos para su utilización extracelular y/o intracelular en otros tejidos (Hao et al., 2015; Furuhashi et al., 2014).

Por otra parte, nosotros encontramos que otros genes asociados con la beta oxidación ácidos grasos (*HADHB* y *CPT1B*, anexo 1) mostraron una disminución de la expresión ($\log_2FC = -1.25$ y $P_{adj} = 1,82E-06$; $\log_2FC = -1.36$ y $P_{adj} = 2,47E-05$, respectivamente) en la línea alta en comparación con la baja. El gene hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit beta (*HADHB*) codifica la subunidad beta de la proteína trifuncional mitocondrial. La proteína codificada también puede unirse al ARN y disminuir la estabilidad de algunos ARN mensajeros. En particular la subunidad β cataliza la actividad 3-cetoacil-CoA tiolasa (LCKT) participando en el paso final de la beta oxidación mitocondrial de los ácidos grasos de cadena larga (Dagher et al., 2021). La proteína codificada el gen carnitine palmitoyltransferase 1B (*CPT1B*) pertenece a la familia carnitina/colina acetiltransferasas, dicha enzima controla la velocidad de la vía de señalización de la beta oxidación de los ácidos grasos de cadena larga en las mitocondrias. Asimismo, esta enzima es necesaria para el transporte neto de acil-CoAs de ácidos grasos de cadena larga desde citoplasma a la mitocondria. De hecho, *CPT1B* es un miembro de la familia de proteínas "CPT1" incluyendo 3 isoformas codificadas por diferentes genes (*CPT1A*, *CPT1B* y *CPT1C*) y reconocidos como importantes mediadores de la oxidación de ácidos grasos (Schreurs et al., 2010). La expresión de *CPT1B* es crucial en la regulación de la degradación/breakdown de las grasas, suministro de energía, así como en la homeostasis energética (Schlaepfer & Joshi, 2020). Sin embargo, respecto a la alta sensibilidad de *CPT1B* a la inhibición por malonil-CoA, dicha inducción hepática no parece ser fisiológicamente relevante (Schreurs et al., 2010).

Otro DEG asociado con en el metabolismo del acetyl-CoA y de los ácidos grasos fue el gen peroxisomal succinyl-coenzyme A thioesterase (*LOC100344005*, también conocido como *ACOT4*). Consistente con nuestros hallazgos, *ACOT4* fue identificado con una disminución de la expresión en la línea alta ($\log_2FC = -1.18$ y $P_{adj} = 6,46E-04$, anexo 1). *ACOT4* facilita la actividad acil-CoA hidrolasa y la actividad succinil-CoA hidrolasa. También interviene en procesos metabólicos del ácido carboxílico y ácidos monocarboxílicos saturados, y en el proceso metabólico del succinil-CoA. La identificación específica de succinil-CoA tioesterasa en peroxisomas sugiere en gran medida que la β -oxidación

peroxisomal de los ácidos dicarboxílicos conduce a la formación de succinato, al menos bajo ciertas condiciones; pero también que *ACOT4* junto con *ACOT8* sean responsables de la terminación de la β -oxidación de los ácidos dicarboxílicos de cadena media con la liberación concomitante de los ácidos libres correspondientes (Westin et al., 2005). Aunque se considera que las mitocondrias son el principal sitio de la β -oxidación de los ácidos grasos para la utilización de energía, la importancia de los peroxisomas en el metabolismo lipídico está bien establecida. Los peroxisomas contienen aproximadamente 60 enzimas involucradas en vías metabólicas lipídicas. Asimismo, se han identificado varias enzimas acil-CoA tioesterasas peroxisomales que catalizan la hidrólisis de acil-CoA (ya sea de cadena corta, media, larga y muy larga), ácidos-CoAs biliares y metil-CoA ramificados (Hunt et al., 2012).

Cabe destacar también al gene glycerol-3-phosphate acyltransferase 3 (*GPAT3*) por su rol en la adipogénesis. *GPAT3* mostró una disminución de sus niveles de expresión en la línea alta ($\log_2FC = -1.02$ y $P_{adj} = 6,76E-03$, anexo 1). Este gen codifica para un miembro de la familia de proteínas aciltransferasas, la proteína codificada es una enzima que cataliza la conversión de glicerol-3-fosfato en ácido lisofosfatídico en la síntesis de triacilglicerol. En efecto, el triacilglicerol también puede servir como neutralizador para ácidos grasos potencialmente tóxicos y acil-CoAs dentro de las células (Takeuchi & Reue, 2009). Otro gen candidato interesante fue el myosin regulatory light chain interacting protein (*MYLIP*), el cual también fue detectado con una disminución de la expresión ($\log_2FC = -0.84$ y $P_{adj} = 3,52E-09$, anexo 1) en la línea alta versus la baja. *MYLIP* es también conocido como IDOL [inducible degrader of the low density lipoprotein receptor (LDLR)]. IDOL es un modulador postraduccional de los niveles de LDLR, pero también una diana directa para la regulación por receptores X del hígado (LXRs), así como su expresión es responsable regular el nivel del esteroles celular independientemente de las proteínas de unión al elemento de respuesta al esteroles (sterol-response element-binding proteins) (Hong et al., 2010; Weissglas-Volkov et al., 2011). Además, *MYLIP* ha sido involucrado en la regulación del metabolismo del colesterol y transporte de lípidos (Xu et al., 2023; Lindholm et al., 2009).

Por último, el gen cytochrome P450 4A6 (*CYP4A6*) también exhibió cambios de expresión diferencial comparando la línea alta versus la baja ($\log_2FC = -1.12$ y $P_{adj} = 3,72E-03$, anexo 1). *CYP4A6* es inducido por ácido clofíbrico en el hígado y el riñón del conejo, donde cataliza la omega hidroxilación de los ácidos araquidónico, láurico y palmítico (Muerhoff et al., 1992). La familia de enzimas del citocromo P450 4 (*CYP4*) son responsables de la activación o desactivación de numerosos sustratos endógenos. Dicha familia de enzimas son monooxigenasas responsables de la ω -oxidación de ácidos grasos endógenos y eicosanoides, así como desempeñan un rol crucial en la regulación de numerosas vías de señalización de eicosanoides (Zhou et al., 2023).

3.3 Análisis funcionales de genes expresados diferencialmente

Los resultados del análisis de sobrerrepresentación con el plug-in ClueGO revelaron el enriquecimiento de un total de 38 términos funcionales ($P_{adj} < 0,05$), de los cuales 28 fueron pathways, 8 pertenecieron a procesos biológicos y 2 a funciones moleculares (Figura 10). Para más detalles sobre los términos de ontología génica y pathways y los genes encontrados en cada uno de ellos consulte el anexo 2.

En el siguiente gráfico podemos apreciar como interaccionan diferentes términos funcionales, así como su agrupación en clusters (Figura X). En network se observan tres clusters mayores agrupando varios términos funcionales, así como otros nodos interconectados (por ejemplo, “biosynthesis of unsaturated fatty acids–fatty acid elongation”; “fatty acid degradation–PPAR signaling pathway”). Complementariamente, se muestra otra representación (gráfico de barras) de los resultados para denotar el número de genes encontrados en cada uno de los términos funcionales (Figura 11).

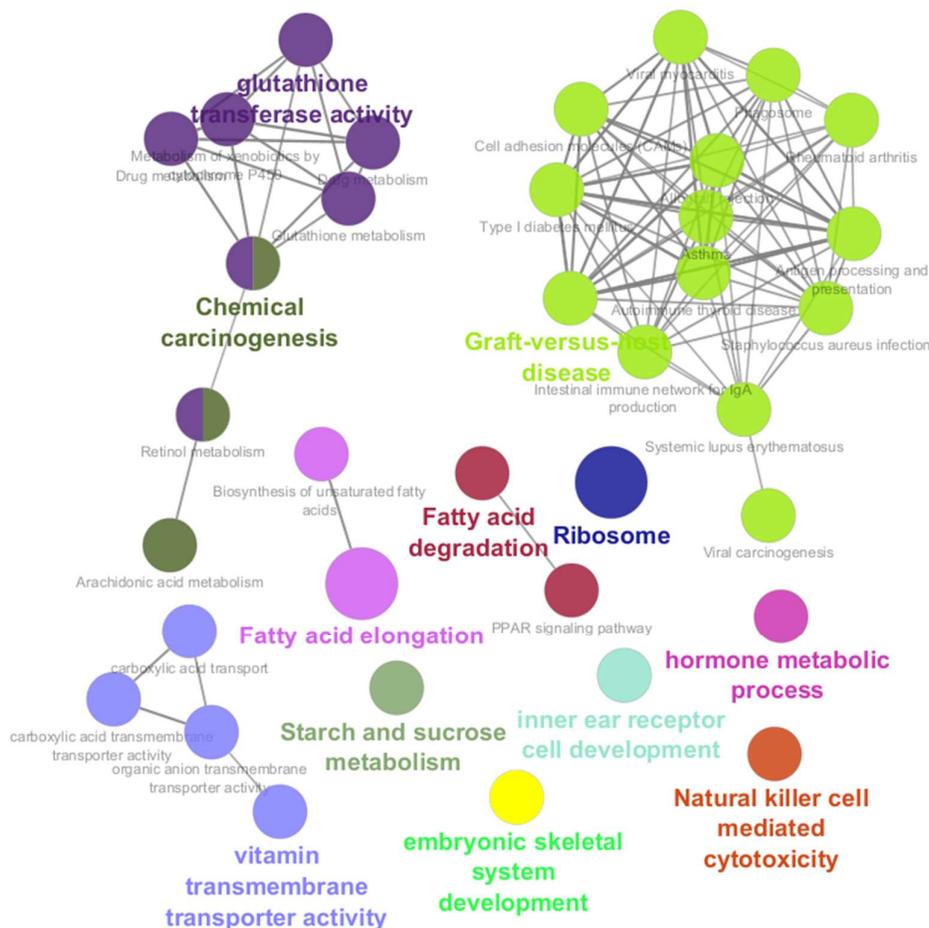


Figura 10. Network de las ontologías génicas y pathways enriquecidas usando el listado de DEGs tras comparar perfiles de expresión génica por 3' RNA-Seq de las líneas divergentes de conejos.

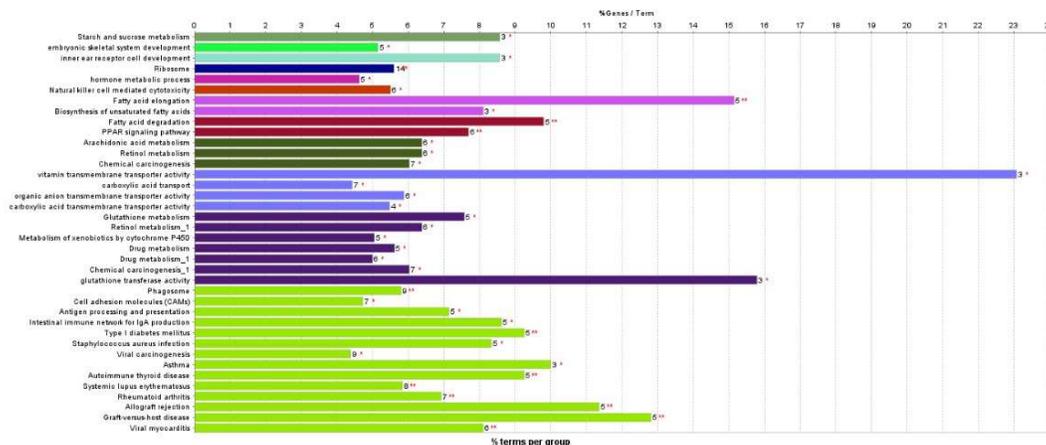


Figura 11. Términos funcionales agrupados en función de las similitudes entre ellos, conteniendo también el número de genes asociados encontrados.

Acorde al interés de nuestro estudio ejemplificaremos algunos de los términos de ontología génica y KEGG pathways (Figura 11). Respecto a términos KEGG, en la elongación de ácidos grasos (KEGG:00062) encontramos 5 genes (*ELOVL6*, *HADHB*, *LOC100344005*, *LOC100344509*, *LOC100347496*), mientras que en la biosíntesis de ácidos grasos insaturados (KEGG:01040) detectamos un total de 3 genes (*ELOVL6*, *LOC100344005*, *LOC100344509*). En el caso de la degradación de ácidos grasos (KEGG:00071) localizamos 5 genes (*CPT1B*, *CYP4A6*, *HADHB*, *LOC100341018*, *LOC100347496*), mientras que en la vía de señalización PPAR (KEGG:03320) detectamos 6 genes (*CPT1B*, *CYP4A6*, *FABP4*, *LOC100341018*, *LOC100342501*, *PLIN2*). Por otro lado, los resultados de los términos de ontología génica indicaron 3 genes asociados (*SLC19A3*, *SLC2A1*, *SLC44A4*) con la actividad del transportador transmembrana de vitaminas (GO:0090482), en tanto el transportador de ácido carboxílico (GO:0046942) contenía 7 genes relacionados (*CROT*, *FABP4*, *PLIN2*, *SLC16A11*, *SLC26A6*, *SLC51A*, *SLC51B*). También encontramos el transportador transmembrana de aniones orgánicos (GO:0008514) representados 6 genes asociados (*SLC16A11*, *SLC26A6*, *SLC2A1*, *SLC44A4*, *SLC51A*, *SLC51B*), así como 4 genes asociados (*SLC16A11*, *SLC26A6*, *SLC51A*, *SLC51B*) con termino de transportador de transmembrana de ácido carboxílico (GO:0046943). Entro todos los términos funcionales es interesante resaltar la vía de señalización PPAR, así pues, los PPARs son receptores nucleares que funcionan como factores de transcripción activados por ligandos, los cuales regulan el metabolismo de los ácidos grasos y modulan la expresión génica de los genes dianas (Wagner and Wagner, 2020).

4 CONCLUSIONES

En esta investigación identificamos un total de 420 genes diferencialmente expresados (DEGs) entre líneas seleccionadas divergentemente por grasa intramuscular (GIM). Entre estos DEGs detectamos genes candidatos interesantes para la deposición del contenido de GIM en conejos (por ejemplo, *ELOVL6*, *FABP4*, *PLIN2*, *CPT1B*, *CYP4A6*, *HADHB* y *SLC44A4*).

Asimismo, detectamos la sobrerrepresentación de 38 términos funcionales incluyendo 28 KEGG pathways, 8 procesos biológicos y 2 funciones moleculares. Entre ellos identificamos la vía de señalización del receptor activado por el proliferador de peroxisomas (PPAR), la elongación de ácidos grasos, la degradación de ácidos grasos y la biosíntesis de ácidos grasos insaturados.

5 BIBLIOGRAFÍA

Abeler-Dörner, L., Swamy, M., Williams, G., Hayday, A. C., & Bas, A. (2012). Butyrophilins: an emerging family of immune regulators. *Trends in Immunology*, 33(1), 34-41. <https://doi.org/10.1016/j.it.2011.09.007>

Alves-Bezerra, M., & Cohen, D. E. (2017). Triglyceride metabolism in the liver. En *Comprehensive Physiology* (Vol. 8, Número 1, pp. 1-22). Wiley. <https://doi.org/10.1002/cphy.c170012>

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* (Oxford, England), 31(2), 166-169. <https://doi.org/10.1093/bioinformatics/btu638>

Andrew, S. (s. f.). FastQC: a quality control tool for high throughput sequence data [Internet]. Scienceopen.com. Recuperado 25 de junio de 2024, de <https://www.scienceopen.com/document?vid=009384c0-b1d9-4494-87c4-364b6a729ff6>

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* (Oxford, England), 25(8), 1091-1093. <https://doi.org/10.1093/bioinformatics/btp101>

Buccitelli, C., & Selbach, M. (2020). mRNAs, proteins and the emerging principles of gene expression control. *Nature Reviews. Genetics*, 21(10), 630-644. <https://doi.org/10.1038/s41576-020-0258-4>

Calkin, A. C., & Tontonoz, P. (2012). Transcriptional integration of metabolism by the nuclear sterol-activated receptors LXR and FXR. *Nature Reviews. Molecular Cell Biology*, 13(4), 213-224. <https://doi.org/10.1038/nrm3312>

Chung, M., Bruno, V. M., Rasko, D. A., Cuomo, C. A., Muñoz, J. F., Livny, J., Shetty, A. C., Mahurkar, A., & Dunning Hotopp, J. C. (2021). Best practices on the differential expression analysis of multi-species RNA-seq. *Genome Biology*, 22(1). <https://doi.org/10.1186/s13059-021-02337-8>

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., & Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17(1). <https://doi.org/10.1186/s13059-016-0881-8>

Dagher, R., Massie, R., & Gentil, B. J. (2021). MTP deficiency caused by HADHB mutations: Pathophysiology and clinical manifestations. *Molecular Genetics and Metabolism*, 133(1), 1-7. <https://doi.org/10.1016/j.ymgme.2021.03.010>

Day, D. A., & Tuite, M. F. (1998). Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview. *The journal of endocrinology*, 157(3), 361-371. <https://doi.org/10.1677/joe.0.1570361>

- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* (Oxford, England), 29(1), 15-21. <https://doi.org/10.1093/bioinformatics/bts635>
- Enser, M., Hallett, K., Hewitt, B., Fursey, G. A. J., & Wood, J. D. (1996). Fatty acid content and composition of english beef, lamb and pork at retail. *Meat Science*, 42(4), 443-456. [https://doi.org/10.1016/0309-1740\(95\)00037-2](https://doi.org/10.1016/0309-1740(95)00037-2)
- Furuhashi, M., Saitoh, S., Shimamoto, K., & Miura, T. (2014). Fatty acid-binding protein 4 (FABP4): Pathophysiological insights and potent clinical biomarker of metabolic and cardiovascular diseases. *Clinical Medicine Insights. Cardiology*, 8s3(Suppl 3), CMC.S17067. <https://doi.org/10.4137/cmc.s17067>
- Gondret, F., Mouro, J., & Bonneau, M. (1997). Developmental changes in lipogenic enzymes in muscle compared to liver and extramuscular adipose tissues in the rabbit (*Oryctolagus cuniculus*). *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, 117(2), 259-265. [https://doi.org/10.1016/s0305-0491\(97\)00049-7](https://doi.org/10.1016/s0305-0491(97)00049-7)
- Han, L., Zhang, M., Xing, Z., Coleman, D. N., Liang, Y., Loo, J. J., & Yang, G. (2020). Knockout of butyrophilin subfamily 1 member A1 (BTN1A1) alters lipid droplet formation and phospholipid composition in bovine mammary epithelial cells. *Journal of Animal Science and Biotechnology*, 11(1). <https://doi.org/10.1186/s40104-020-00479-6>
- Hao, Y., Ma, X., Luo, Y., Hu, X., Pan, X., Xiao, Y., Bao, Y., & Jia, W. (2015). Associations of serum adipocyte fatty acid binding protein with body composition and fat distribution in nondiabetic Chinese women. *The Journal of Clinical Endocrinology and Metabolism*, 100(5), 2055-2062. <https://doi.org/10.1210/jc.2014-4373>
- He, L., Vasiliou, K., & Nebert, D. W. (2009). Analysis and update of the human solute carrier (SLC) gene superfamily. *Human Genomics*, 3(2), 195. <https://doi.org/10.1186/1479-7364-3-2-195>
- Hernández, P., Xicato, G., Trocino, A., & Lukefahr, S. (2008). Enhancement of nutritional quality and safety in rabbit meat. <https://www.semanticscholar.org/paper/b012f4fd444f48d44cf9062f28976f80f267cac6>
- Hocquette, J. F., Gondret, F., Baéza, E., Médale, F., Jurie, C., & Pethick, D. W. (2010). Intramuscular fat content in meat-producing animals: development, genetic and nutritional control, and identification of putative markers. *Animal: An International Journal of Animal Bioscience*, 4(2), 303-319. <https://doi.org/10.1017/s1751731109991091>
- Hong, C., Duit, S., Jalonen, P., Out, R., Scheer, L., Sorrentino, V., Boyadjian, R., Rodenburg, K. W., Foley, E., Korhonen, L., Lindholm, D., Nimpf, J., van Berkel, T. J. C., Tontonoz, P., & Zelcer, N. (2010). The E3 ubiquitin ligase IDOL induces the degradation of the low density lipoprotein receptor family members VLDLR and ApoER2. *The Journal of Biological Chemistry*, 285(26), 19720-19726. <https://doi.org/10.1074/jbc.m110.123729>

- Hunt, M. C., Siponen, M. I., & Alexson, S. E. H. (2012). The emerging role of acyl-CoA thioesterases and acyltransferases in regulating peroxisomal lipid metabolism. *Biochimica et Biophysica Acta. Molecular Basis of Disease*, 1822(9), 1397-1410. <https://doi.org/10.1016/j.bbadis.2012.03.009>
- Iizuka, K., Bruick, R. K., Liang, G., Horton, J. D., & Uyeda, K. (2004). Deficiency of carbohydrate response element-binding protein (ChREBP) reduces lipogenesis as well as glycolysis. *Proceedings of the National Academy of Sciences of the United States of America*, 101(19), 7281-7286. <https://doi.org/10.1073/pnas.0401516101>
- Koch, C. M., Chiu, S. F., Akbarpour, M., Bharat, A., Ridge, K. M., Bartom, E. T., & Winter, D. R. (2018). A beginner's guide to analysis of RNA sequencing data. *American Journal of Respiratory Cell and Molecular Biology*, 59(2), 145-157. <https://doi.org/10.1165/rcmb.2017-0430tr>
- Kumar, S. A., Kim, H.-J., Jayasena, D. D., & Jo, C. (2023). On-farm and processing factors affecting rabbit carcass and meat quality attributes. *Food Science of Animal Resources*, 43(2), 197-219. <https://doi.org/10.5851/kosfa.2023.e5>
- Laghouaouta, H., Sosa-Madrid, B. S., Zubiri-Gaitán, A., Hernández, P., & Blasco, A. (2020). Novel genomic regions associated with intramuscular fatty acid composition in rabbits. *Animals: An Open Access Journal from MDPI*, 10(11), 2090. <https://doi.org/10.3390/ani10112090>
- Li, W. V., & Li, J. J. (2018). Modeling and analysis of RNA-seq data: a review from a statistical perspective. *Quantitative Biology (Beijing, China)*, 6(3), 195-209. <https://doi.org/10.1007/s40484-018-0144-7>
- Lindholm, D., Bornhauser, B. C., & Korhonen, L. (2009). Mylip makes an Idol turn into regulation of LDL receptor. *Cellular and Molecular Life Sciences: CMLS*, 66(21), 3399-3402. <https://doi.org/10.1007/s00018-009-0127-y>
- Listrat, A., Leuret, B., Louveau, I., Astruc, T., Bonnet, M., Lefaucheur, L., Picard, B., & Bugeon, J. (2016). How muscle structure and composition influence meat and flesh quality. *TheScientificWorldJournal*, 2016, 1-14. <https://doi.org/10.1155/2016/3182746>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12). <https://doi.org/10.1186/s13059-014-0550-8>
- Ma, F., Fuqua, B. K., Hasin, Y., Yukhtman, C., Vulpe, C. D., Lusic, A. J., & Pellegrini, M. (2019). A comparison between whole transcript and 3' RNA sequencing methods using Kapa and Lexogen library preparation methods. *BMC Genomics*, 20(1). <https://doi.org/10.1186/s12864-018-5393-3>
- Manzhen, Z., Junda, L., Jinyi, X., Lufeng, Z., & Shengtao, X. (2023). Exploring human CYP4 enzymes: Physiological roles, function in diseases and focus on inhibitors. *Drug Discovery Today*, 28(5), 103560. <https://doi.org/10.1016/j.drudis.2023.103560>

- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, 17(1), 10-12. <https://journal.embnet.org/index.php/embnetjournal/article/view/200/479>
- Martínez-Álvaro, M., Agha, S., Blasco, A., & Hernández, P. (2017). Muscle lipid metabolism in two rabbit lines divergently selected for intramuscular fat¹. *Journal of Animal Science*, 95(6), 2576-2584. <https://doi.org/10.2527/jas.2017.1371>
- Martínez-Álvaro, M., Hernández, P., & Blasco, A. (2016a). Divergent selection on intramuscular fat in rabbits: Responses to selection and genetic parameters¹. *Journal of Animal Science*, 94(12), 4993-5003. <https://doi.org/10.2527/jas.2016-0590>
- Martínez-Álvaro, M., Paucar, Y., Satué, K., Blasco, A., & Hernández, P. (2018). Liver metabolism traits in two rabbit lines divergently selected for intramuscular fat. *Animal: An International Journal of Animal Bioscience*, 12(6), 1217-1223. <https://doi.org/10.1017/s1751731117002695>
- Martínez-Álvaro, Marina, Zubiri-Gaitán, A., Hernández, P., Greenacre, M., Ferrer, A., & Blasco, A. (2021). Comprehensive functional core microbiome comparison in genetically obese and lean hosts under the same environment. *Communications Biology*, 4(1). <https://doi.org/10.1038/s42003-021-02784-w>
- Martínez-Álvaro, M., Penalba, V., Blasco, A., & Hernández, P. (2016b). Effect of divergent selection for intramuscular fat on sensory traits and instrumental texture in rabbit meat¹. *Journal of Animal Science*, 94(12), 5137-5143. <https://doi.org/10.2527/jas.2016-0850>
- Matsuzaka, T., & Shimano, H. (2009). Elovl6: a new player in fatty acid metabolism and insulin sensitivity. *Journal of Molecular Medicine*, 87(4), 379-384. <https://doi.org/10.1007/s00109-009-0449-0>
- McClure, R. S., Rericha, Y., Waters, K. M., & Tanguay, R. L. (2023). 3' RNA-seq is superior to standard RNA-seq in cases of sparse data but inferior at identifying toxicity pathways in a model organism. *Frontiers in bioinformatics*, 3. <https://doi.org/10.3389/fbinf.2023.1234218>
- Moll, P., Ante, M., Seitz, A., & Reda, T. (2014). QuantSeq 3' mRNA sequencing for RNA quantification. *Nature Methods*, 11(12), i-iii. <https://doi.org/10.1038/nmeth.f.376>
- Muerhoff, A. S., Griffin, K. J., & Johnson, E. F. (1992). The peroxisome proliferator-activated receptor mediates the induction of CYP4A6, a cytochrome P450 fatty acid omega-hydroxylase, by clofibric acid. *The journal of biological chemistry*, 267(27). <https://pubmed.ncbi.nlm.nih.gov/1326542/>
- Rector, R. S., Payne, R. M., & Ibdah, J. A. (2008). Mitochondrial trifunctional protein defects: Clinical implications and therapeutic approaches. *Advanced Drug Delivery Reviews*, 60(13-14), 1488-1496. <https://doi.org/10.1016/j.addr.2008.04.014>
- Rui, L. (2014). Energy metabolism in the liver. En *Comprehensive Physiology* (Vol. 4, Número 1, pp. 177-197). Wiley. <https://doi.org/10.1002/cphy.c130024>

Sabui, S., Anthonymuthu, S., Ramamoorthy, K., Skupsky, J., Jennings, T. S. K., Rahmatpanah, F., Fleckenstein, J. M., & Said, H. M. (2024). Effect of knocking out mouse *Slc44a4* on colonic uptake of the microbiota-generated thiamine pyrophosphate and colon physiology. *American Journal of Physiology. Gastrointestinal and Liver Physiology*, 327(1), G36-G46. <https://doi.org/10.1152/ajpgi.00065.2024>

Sabuncuoglu, N., Coban, O., Lacin, E., Ceylan, Z. G., Ozdemir, D., & Ozkan, A. (2011). Effect of pre-slaughter environment on some physiological parameters and meat quality in New Zealand rabbits (*Oryctolagus cuniculus*). *Tropical Animal Health and Production*, 43(2), 515-519. <https://doi.org/10.1007/s11250-010-9725-9>

Sapp, R. L., Bertrand, J. K., Pringle, T. D., & Wilson, D. E. (2002). Effects of selection for ultrasound intramuscular fat percentage in Angus bulls on carcass traits of progeny^{1,2}. *Journal of Animal Science*, 80(8), 2017-2022. <https://doi.org/10.1093/ansci/80.8.2017>

Schlaepfer, I. R., & Joshi, M. (2020). CPT1A-mediated fat oxidation, mechanisms, and therapeutic potential. *Endocrinology*, 161(2), bqz046. <https://doi.org/10.1210/endocr/bqz046>

Schreurs, M., Kuipers, F., & Van Der Leij, F. R. (2010). Regulatory enzymes of mitochondrial β -oxidation as targets for treatment of the metabolic syndrome. *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity*, 11(5), 380-388. <https://doi.org/10.1111/j.1467-789x.2009.00642.x>

Schwab, C. R., Baas, T. J., Stalder, K. J., & Nettleton, D. (2009). Results from six generations of selection for intramuscular fat in Duroc swine using real-time ultrasound. I. Direct and correlated phenotypic responses to selection¹. *Journal of Animal Science*, 87(9), 2774-2780. <https://doi.org/10.2527/jas.2008-1335>

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), 2498-2504. <https://doi.org/10.1101/gr.1239303>

Smith, T., Heger, A., & Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Research*, 27(3), 491-499. <https://doi.org/10.1101/gr.209601.116>

Song, Z., Xiaoli, A. M., & Yang, F. (2018). Regulation and metabolic significance of DE Novo lipogenesis in adipose tissues. *Nutrients*, 10(10), 1383. <https://doi.org/10.3390/nu10101383>

Sosa-Madrid, B. S., Varona, L., Blasco, A., Hernández, P., Casto-Rebollo, C., & Ibáñez-Escriche, N. (2020a). The effect of divergent selection for intramuscular fat on the domestic rabbit genome. *Animal: An International Journal of Animal Bioscience*, 14(11), 2225-2235. <https://doi.org/10.1017/s1751731120001263>

Sosa-Madrid, Bolívar S., Hernández, P., Blasco, A., Haley, C. S., Fontanesi, L., Santacreu, M. A., Pena, R. N., Navarro, P., & Ibáñez-Escriche, N. (2020b). Genomic regions influencing intramuscular fat in divergently selected rabbit lines. *Animal Genetics*, 51(1), 58-69. <https://doi.org/10.1111/age.12873>

- Takeuchi, K., & Reue, K. (2009). Biochemistry, physiology, and genetics of GPAT, AGPAT, and lipin enzymes in triglyceride synthesis. *American Journal of Physiology. Endocrinology and Metabolism*, 296(6), E1195-E1209. <https://doi.org/10.1152/ajpendo.90958.2008>
- Tandonnet, S., & Torres, T. T. (2017). Traditional versus 3' RNA-seq in a non-model species. *Genomics Data*, 11, 9-16. <https://doi.org/10.1016/j.gdata.2016.11.002>
- Wagner, N., & Wagner, K.-D. (2020). The role of PPARs in disease. *Cells (Basel, Switzerland)*, 9(11), 2367. <https://doi.org/10.3390/cells9112367>
- Weissglas-Volkov, D., Calkin, A. C., Tusie-Luna, T., Sinsheimer, J. S., Zelcer, N., Riba, L., Tino, A. M. V., Ordoñez-Sánchez, M. L., Cruz-Bautista, I., Aguilar-Salinas, C. A., Tontonoz, P., & Pajukanta, P. (2011). The N342S MYLIP polymorphism is associated with high total cholesterol and increased LDL receptor degradation in humans. *The Journal of Clinical Investigation*, 121(8), 3062-3071. <https://doi.org/10.1172/jci45504>
- Westin, M. A. K., Hunt, M. C., & Alexson, S. E. H. (2005). The identification of a succinyl-CoA thioesterase suggests a novel pathway for succinate production in peroxisomes. *The Journal of Biological Chemistry*, 280(46), 38125-38132. <https://doi.org/10.1074/jbc.m508479200>
- Wood, J. D., Enser, M., Fisher, A. V., Nute, G. R., Sheard, P. R., Richardson, R. I., Hughes, S. I., & Whittington, F. M. (2008). Fat deposition, fatty acid composition and meat quality: A review. *Meat Science*, 78(4), 343-358. <https://doi.org/10.1016/j.meatsci.2007.07.019>
- Wu, J., Xiao, J., Zhang, Z., Wang, X., Hu, S., & Yu, J. (2014). Ribogenomics: The science and knowledge of RNA. *Genomics, Proteomics & Bioinformatics*, 12(2), 57-63. <https://doi.org/10.1016/j.gpb.2014.04.002>
- Xu, J., Mao, F., Lu, Y., Liu, T., Li, X., & Li, Y. (2023). Hepatic transcriptomics reveals reduced lipogenesis in high-salt diet mice. *Genes*, 14(5), 966. <https://doi.org/10.3390/genes14050966>
- Xu, S., Zou, F., Diao, Z., Zhang, S., Deng, Y., Zhu, X., Cui, L., Yu, J., Zhang, Z., Bamigbade, A. T., Zhang, H., Wei, X., Zhang, X., Liang, B., & Liu, P. (2019). Perilipin 2 and lipid droplets provide reciprocal stabilization. *Biophysics Reports*, 5(3), 145-160. <https://doi.org/10.1007/s41048-019-0091-5>
- Zhang, Y., Castellani, L. W., Sinal, C. J., Gonzalez, F. J., & Edwards, P. A. (2004). Peroxisome proliferator-activated receptor- γ coactivator 1 α (PGC-1 α) regulates triglyceride metabolism by activation of the nuclear receptor FXR. *Genes & Development*, 18(2), 157-169. <https://doi.org/10.1101/gad.1138104>
- Zhao, G. P., Chen, J. L., Zheng, M. Q., Wen, J., & Zhang, Y. (2007). Correlated responses to selection for increased intramuscular fat in a Chinese quality chicken line. *Poultry Science*, 86(11), 2309-2314. <https://doi.org/10.1093/ps/86.11.2309>

Zomeño, C., Hernández, P., & Blasco, A. (2013). Divergent selection for intramuscular fat content in rabbits. I. Direct response to selection¹. *Journal of Animal Science*, 91(9), 4526-4531. <https://doi.org/10.2527/jas.2013-6361>

Zotte, A. D. (2004). Dietary advantages: rabbit must tame consumers. https://www.researchgate.net/publication/285199973_Dietary_advantages_rabbit_must_tame_consumers

Zubiri-Gaitán, A., Blasco, A., Ccalta, R., Satué, K., & Hernández, P. (2022). Intramuscular fat selection in rabbits modifies the fatty acid composition of muscle and liver tissues. *Animals: An Open Access Journal from MDPI*, 12(7), 893. <https://doi.org/10.3390/ani12070893>

Zubiri-Gaitán, A., Blasco, A., & Hernández, P. (2023). Plasma metabolomic profiling in two rabbit lines divergently selected for intramuscular fat content. *Communications Biology*, 6(1), 1-10. <https://doi.org/10.1038/s42003-023-05266-3>

6 ANEXOS

Anexo 1. Resultados de DESeq2 con todos los genes.

Se añade un link de consulta ya que se trata de una tabla de Excel extensa:

<https://docs.google.com/spreadsheets/d/1dls-4h4xmlDvga-AvP82FoDFskIEYUG/edit?usp=sharing&oid=101903546731178448600&rtpof=true&sd=true>

Anexo 2. Resultados del análisis de ontología génica y pathways.

Se añade un link de consulta ya que se trata de una tabla de Excel extensa:

https://docs.google.com/spreadsheets/d/1yjbZLNgh33rp0_LM8ffTMTEzYYpk4zPL/edit?usp=sharing&oid=101903546731178448600&rtpof=true&sd=true

Anexo 3. Informe Turnitin



Recibo digital

Este recibo confirma que su trabajo ha sido recibido por Turnitin. A continuación podrá ver la información del recibo con respecto a su entrega.

La primera página de tus entregas se muestra abajo.

Autor de la entrega: Ángel Peñuelas Cumplido
Título del ejercicio: Curso 24_25 GENERAR INFORME TURNITIN TFGTFM ETSEAMN
Título de la entrega: 49352524:TFG_DEF.docx
Nombre del archivo: 23_Assignments_68ece32f-feba-4f26-be09-cedc2c1abcf3_TF...
Tamaño del archivo: 1.58M
Total páginas: 39
Total de palabras: 11,754
Total de caracteres: 68,256
Fecha de entrega: 07-jul.-2024 01:25p. m. (UTC+0200)
Identificador de la entrega: 2413344099

RESUMEN

El contenido de grasas intramusculares (GMI) es un indicador importante de la calidad de la carne y afecta las propiedades nutricionales, sensoriales y tecnológicas de la carne. En la Universidad Politécnica de Valencia se realizó un experimento de selección divergente de GMI en el modelo *Longissimus thoracis et lumborum* en conejos durante 10 generaciones. El experimento fue exitoso, obteniendo una respuesta directa a la selección en la 10ª generación de 3,8 desviaciones típicas del control. El GMI es un rasgo complejo de naturaleza poligénica, lo que significa que está controlado por múltiples genes, cada uno con un pequeño efecto. En este artículo, se reportan los resultados de genes expresados diferencialmente (DEGs) en el hígado de dos líneas de conejos seleccionados de manera divergente para GMI (líneas alta y baja), y estudiar aquellos que modulan la deposición de grasa. Se utilizaron un total de 46 conjuntos de datos de RNA-seq entre estas y líneas parentales, a lo largo de generaciones de selección para GMI. Se utilizaron muestras de hígado de estos animales para determinar los niveles de expresión génica mediante 3' RNA-seq. La secuenciación de la biblioteca se realizó en la plataforma Illumina NovaSeq X con química de lectura única de 150 pb. Se realizó un análisis de expresión diferencial mediante un modelo lineal mixto, seguido de un análisis funcional de los genes DEGs obtenidos. Nuestros resultados indican un total de 402 genes DEGs entre las líneas divergentes, de los cuales 156 estaban regulados positivamente y 242 regulados negativamente en la línea alta. Entre los genes DEGs, destacamos genes candidatos asociados en general con el metabolismo lipídico (fosfolípidos, triglicéridos, ácidos grasos, beta-oxidación de ácidos grasos, docosahexaenoícos, entre otros). Destacamos una sobreexpresión de 38 genes de acción funcional que incluyen 28 vías, 4 procesos biológicos y 2 funciones moleculares, entre los que se encuentran la beta-oxidación de ácidos grasos, degradación de ácidos grasos, desacetilación de ácidos grasos metilados, síntesis de lípidos, metabolismo energético y otros. El estudio de la expresión de genes hepáticos reveló información valiosa sobre genes candidatos para el depósito de GMI en conejos, pero también sobre lipoproteínas y sus interacciones con el metabolismo de lípidos y carbohidratos.

Palabras clave: GMI, RNA-seq, Oryctolagus cuniculus y FLM2