



MÁSTER INTERUNIVERSITARIO EN MEJORA GENÉTICA
ANIMAL Y BIOTECNOLOGÍA DE LA REPRODUCCIÓN

**Análisis metabolómico no dirigido en dos líneas de
conejos seleccionadas de forma divergente para
grasa intramuscular**

Tesis de Máster
Valencia, Septiembre 2018
Agostina Zubiri Gaitán

Directora:
Pilar Hernández

RESUMEN

La metabolómica es una ciencia emergente que está siendo cada vez más utilizada en el campo de la genética animal ya que ha abierto nuevas posibilidades para el estudio de los caracteres de interés. Los métodos de análisis multivariantes han ayudado en el desarrollo de esta ciencia, ya que, a partir de la reducción de la dimensionalidad, facilita el análisis y la interpretación de la gran cantidad de variables generadas.

Se realizó un análisis metabolómico no dirigido en muestras de plasma de dos líneas de conejos seleccionadas divergentemente por el carácter grasa intramuscular (GIM) mediante espectrometría de masas acoplada a cromatografía líquida (LC-MS) y espectrometría de masas acoplada a electroforesis capilar (CE-MS), con el objetivo de estudiar los mecanismos implicados en la diferenciación entre líneas de este carácter. Para el análisis de los datos se aplicaron distintas técnicas de análisis multivariantes: análisis de componentes principales (PCA), proyección sobre estructuras latentes–análisis discriminante (PLS-DA) y proyección ortogonal sobre estructuras latentes– análisis discriminante (OPLS-DA). La identificación de los metabolitos y su implicancia en las rutas metabólicas se realizó cruzando información de varias bases de datos (KEGG, HMDB, Metlin y LipidMaps).

Los métodos de análisis multivariantes han permitido reducir la complejidad de las grandes matrices de datos, con gran cantidad de información correlacionada, obtenidas a partir de los análisis metabolómicos. De esta manera, se encontraron diversos metabolitos diferencialmente expresados entre las líneas. Mediante LC-MS se encontraron en mayor cantidad en la línea de baja grasa intramuscular, varios metabolitos implicados en la ruta del metabolismo de los glicerofosfolípidos: fosfocolinas (PC), fosfoetanolaminas (PE) y lisofosfocolinas (LysoPC). Mediante CE-MS se encontró mayor contenido de ácido aspártico, nicotinamida, taurina y L-carnitina en la línea de alta GIM y menor de Dimetil-L-arginina.

Se pudieron identificar algunas rutas metabólicas potencialmente implicadas en la diferenciación del carácter grasa intramuscular, siendo las más destacadas la ruta del metabolismo de los glicerofosfolípidos, la ruta de la pentosa 5-fosfato, la biosíntesis de ácidos biliares y la β oxidación. La integración multiómica entre los resultados genómicos y metabolómicos podrán clarificar los mecanismos implicados en la deposición de la grasa intramuscular.

ABSTRACT

Metabolomics is an emerging field of study in animal breeding and genetics since it provides new possibilities for the study of targeted traits. Multivariate statistical analysis promoted the development of metabolomics by using a dimensionality reduction method to easily analyze and interpret the large amount of data generated.

Untargeted metabolomics analysis was performed in two rabbit lines divergently selected by intramuscular fat (IMF) using liquid chromatography-mass spectrometry (LC-MS) and capillary electrophoresis-mass spectrometry (CE-MS) techniques. The objective was to study the mechanisms related to the differentiation of the character between selected lines. Data analysis was performed using several multivariate statistical analyses: principal component analysis (PCA), projection to latent structures-discriminant analysis (PLS-DA) and orthogonal projection to latent structures-discriminant analysis (OPLS-DA). The identification of the metabolites and the study of its implications in the metabolic pathways were performed using several databases (KEGG, HMDB, Metlin and LipidMaps).

Multivariate analysis allowed to reduce the complexity of the large amount, highly correlated data, obtained from the metabolomic analysis. Several metabolites were found to be differentially expressed between lines. The metabolites found through LC-MS were higher in low IMF line and they were implicated in the glycerophospholipids metabolism: phosphocholine (PC), phosphoethanolamine (PE) and lysophosphocholine (LysoPC). CE-MS found larger amounts of aspartic acid, nicotinamide, taurine and L-carnitine and lower amount of Dimethyl-L-arginine in the high IMF line.

Some metabolic pathways were identified as potentially implicated in the differentiation of intramuscular fat phenotype. The most important were glycerophospholipids metabolism, pentose phosphate pathway, primary bile acids biosynthesis and β oxidation. Multiomics studies integrating genomic and metabolomic results could clarify the mechanisms implicated in intramuscular fat deposition.

ÍNDICE

RESUMEN	1
ABSTRACT	2
1. INTRODUCCIÓN	5
1.1 Calidad de carne y grasa intramuscular (GIM)	5
1.2 Metabolómica	6
1.3 Técnicas analíticas	10
1.4 Métodos de análisis multivariantes	13
1.4.1 Análisis de componentes principales	15
1.4.2 Proyección sobre estructuras latentes (PLS)	18
1.4.3 Proyección a estructuras latentes ortogonales (OPLS)	20
2. OBJETIVOS	21
3. MATERIALES Y MÉTODOS	22
3.1 Animales y recolección de muestras	22
3.2 Análisis de las muestras	23
3.2.1 Técnica analítica	23
3.2.2 Preparación de muestras para control de calidad (QC).	23
3.2.3 Cromatografía líquida – Espectrometría de masas (LC-MS).	23
3.2.4 Electroforesis capilar – Espectrometría de masas (CE-MS).	24
3.3 Tratamiento de datos	25
3.4 Análisis estadístico y selección de variables	25
3.5 Identificación de metabolitos	26
4. RESULTADOS Y DISCUSIÓN	27
4.1 Resultados de la novena generación de selección por grasa intramuscular	27
4.2 Análisis exploratorios	27
4.3 Análisis discriminantes	33
4.3.1 Proyección sobre estructuras latentes–análisis discriminante (PLS-DA)	33

4.3.2	Proyección ortogonal sobre estructuras latentes- análisis discriminante (OPLS-DA)	37
4.4	Selección e identificación de variables	43
4.5	Metabolitos diferencialmente expresados entre las líneas de alta y baja GIM	47
5.	CONCLUSIONES	51
6.	REFERENCIAS.....	52
	ANEXOS.....	63

1. INTRODUCCIÓN

1.1 Calidad de carne y grasa intramuscular (GIM)

La grasa intramuscular (GIM) es uno de los principales parámetros que determina la calidad de la carne, ya que afecta la jugosidad, la ternura y el sabor de la misma (Wood et al., 2008). Además, la GIM afecta a la calidad nutricional de la carne. Altos niveles de consumo de grasa y ácidos grasos saturados (SFA) han sido relacionados con obesidad y enfermedades coronarias (Wood et al., 2003; Zong et al., 2016). Un incremento de GIM modifica el perfil de ácidos grasos de la carne al causar un mayor aumento de los SFA y monoinsaturados (MUFA) que de los ácidos grasos poliinsaturados (PUFA), generando una disminución en la relación PUFA/SFA (De Smet et al., 2004).

El contenido de GIM es fácilmente modificable por selección, debido a su alta heredabilidad y variabilidad. En los estudios encontrados, las estimas de las heredabilidades oscilan entre 0.26 y 0.86 (Ciobanu et al., 2011) y entre 0.34 y 0.77 (Mateescu, 2015) en cerdos y en vacuno, respectivamente y en conejos se estimó una heredabilidad de 0.50 (Martínez-Álvaro et al., 2016). En la literatura se encuentran tres experimentos de selección por contenido de GIM: Schwab et al., 2009 en cerdos, Sapp et al., 2002 en bovino y Zhao et al., 2007 en pollos. Todos estos experimentos han mostrado buena respuesta a la selección.

En la Universitat Politècnica de València (UPV) se ha desarrollado un experimento de selección divergente por GIM en conejos (Zomeño et al., 2013 y Martínez-Álvaro et al., 2016). Dicho experimento presentó una buena respuesta a la selección, simétrica para ambas líneas. En la octava generación de selección la respuesta directa fue de 2.4 desviaciones típicas (DT) del carácter, lo que representa un progreso genético del 5% de la media en cada generación (Martínez-Álvaro et al., 2016).

Los experimentos de selección divergente presentan algunas ventajas a la hora de estudiar la composición genética de los caracteres. Por un lado, la diferenciación genética y fenotípica entre las líneas se logra más rápidamente que utilizando una población control sin seleccionar, por lo que, a igual número de generaciones entre ambos tipos de experimentos, la divergencia entre las líneas será mayor. Esto repercute directamente en el número de individuos que se necesitarán para obtener una buena precisión en las estimas. Por otro lado, los cambios ambientales afectan, por término medio, a las dos líneas por igual y esto, sumado al hecho que ambas están siendo seleccionadas por el mismo carácter,

supone que las líneas difieren únicamente en su composición genética; esto permite que, tanto la respuesta directa como las correlacionadas, puedan estimarse mediante la comparación del valor medio de los caracteres en las líneas alta y baja luego de la selección. La simetría en la selección se puede ver comparando las diferencias fenotípicas calculadas, como se mencionó anteriormente, con la respuesta estimada mediante modelos que separan los efectos genéticos de los ambientales (Blasco et al., 2017). Los experimentos de selección divergente tienen la ventaja de poder generar dos poblaciones que difieren únicamente en los efectos de los genes de los caracteres seleccionados y correlacionados, lo que los hace particularmente útil para muchos estudios, entre ellos, la metabolómica.

En el experimento de selección divergente por GIM desarrollado en la UPV, se ha obtenido la respuesta directa a la selección por GIM y las respuestas correlacionadas en diversos caracteres de calidad de la canal y de la carne. Además, en los animales de la novena generación se ha realizado un estudio genómico de las líneas con un chip de 200K, cuyos resultados están siendo analizados actualmente. El análisis del metaboloma junto la información genómica y fisiológica podrá aportar información sustancial sobre las rutas metabólicas que controlan la deposición de grasa en el músculo.

1.2 Metabolómica

La metabolómica es una ciencia emergente cuyo objetivo es la caracterización del conjunto de moléculas de bajo peso molecular (< 1500 Da) en una muestra biológica (Zhang et al., 2012). El conjunto de metabolitos expresados en una muestra biológica se denomina metaboloma. Aunque el análisis de metabolitos específicos ha existido durante mucho tiempo, limitaciones en las técnicas analíticas disponibles no han permitido el avance de la metabolómica hasta hace un par de décadas. El desarrollo de tecnologías capaces de extraer y analizar estos datos moleculares ha revolucionado la aplicación de la metabolómica en el análisis de organismos y procesos biológicos relevantes (Zhang et al., 2012).

Los metabolitos surgen como producto final de las reacciones o rutas metabólicas que ocurren en el organismo. Al ser el último paso en el flujo de información que va desde el genotipo hasta el fenotipo (o función) de un individuo (figura 1), los productos metabólicos pueden considerarse como la última respuesta a los cambios ambientales y a la composición genética de los individuos (Fiehn, 2002) y pueden por lo tanto dar información de su

actividad bioquímica. Este hecho convierte al metaboloma en un objeto de estudio de gran interés, ya que puede ser más fácilmente correlacionado con el fenotipo (Patti et al., 2013).

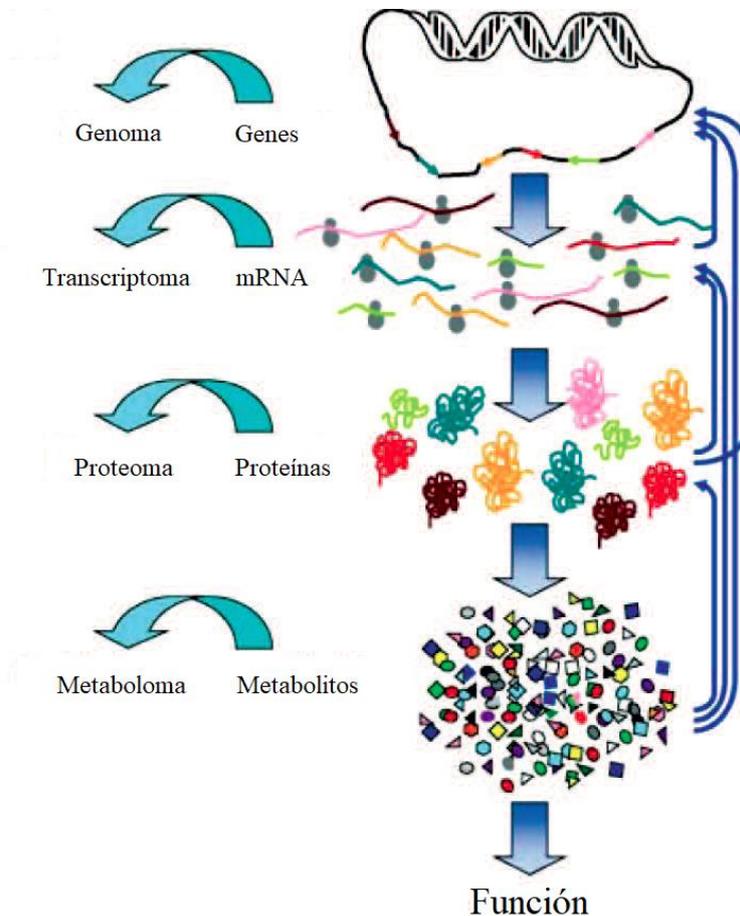


Figura 1. Esquema del flujo de la información desde el genoma hasta la función. Las flechas de la derecha intentan demostrar que el flujo no es unidireccional, sino que podría representarse como una red interconectada (adaptado de Goodacre, 2005).

Dentro de la metabolómica, podemos diferenciar entre la dirigida y la no dirigida. En la metabolómica dirigida (*targeted metabolomics*) el objetivo es detectar y cuantificar de manera precisa uno o un pequeño conjunto de metabolitos conocidos. En la metabolómica no dirigida (*untargeted metabolomics*), en cambio, el análisis intenta captar las señales de todos los metabolitos presentes en el metaboloma. También permite la cuantificación, aunque de manera menos precisa. Esta cuantificación será útil en la comparación entre individuos o grupos. Un resumen del flujo de trabajo en ambos casos puede verse en la figura 2.

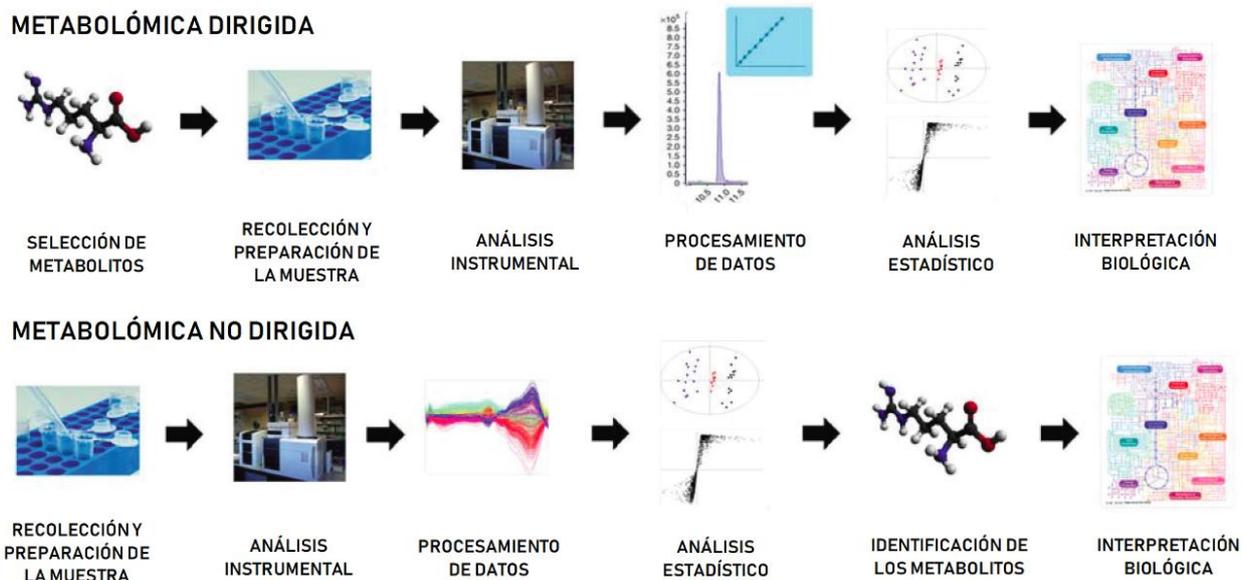


Figura 2. Esquema simplificado de los pasos a seguir en un análisis metabolómico dirigido y en uno no dirigido (adaptado de Canuto et al., 2018).

La metabolómica no dirigida incluye un paso adicional que es la identificación de los metabolitos. En este paso, todas las señales generadas por el análisis son ingresadas en bases de datos donde se buscará reconocer a qué metabolito pertenece. En la tabla 1 se resumen las bases de datos comúnmente utilizadas para la identificación. Es importante mencionar, además, una base de datos recientemente desarrollada para especies ganaderas: “Livestock Metabolome Database” (LMDB) (Goldansaz et al., 2017).

Tabla 1. Bases de datos de metabolitos utilizadas para la identificación.

Base de datos	Espectros	Sitio web	Referencia
HMDB	MS/NMR	http://www.hmdb.ca	Wishart et al., 2013
LMSD	MS	http://www.lipidmaps.org	Sud et al., 2007
METLIN	MS	http://metlin.scripps.edu	Tautenhahn et al., 2012
TOCCATA COLMAR	NMR	http://spin.ccic.ohio-state.edu	Robinette et al., 2008
MassBank	MS	http://www.massbank.jp	Bingol et al., 2015
Golm metabolome	GC-MS	http://gmd.mpimp-golm.mpg.de	Horai et al., 2010
BMRB	NMR	http://www.bmrwisc.edu	Hummel et al., 2007
Madison	NMR	http://mmcd.nmrfam.wisc.edu	Ulrich et al., 2008
NMRShiftDB	NMR	http://nmrshiftdb.nmr.uni-koeln.de	Cui et al., 2008
RIKEN	MS/NMR	http://prime.psc.riken.jp	Steinbeck et al., 2012
Birmingham Metabolite Library	NMR	http://www.bml-nmr.org	Akiyama et al., 2008
LMDB	MS/NMR	http://www.lmdb.ca	Sakurai et al., 2013
			Ludwig et al., 2012
			Goldansaz et al., 2017

Adaptado de Alonso et al. (2015)

En el análisis no dirigido muchos de los metabolitos detectados no han sido registrados todavía. La correlación de estos metabolitos con el fenotipo en estudio puede ser relevante y esto supondría el desarrollo de nuevos trabajos para su identificación y para el estudio de las vías metabólicas donde estén implicados. Aún sin conocer el nombre de los metabolitos relevantes que se detecten en el análisis, estos pueden ser utilizados como biomarcadores.

En ciencia animal, la metabolómica está siendo cada vez más utilizada, tanto en investigación como en producción, a pesar de que su aplicación se encuentra retrasada con respecto a otras áreas, como la medicina o la mejora genética vegetal (Goldansaz et al., 2017). En los últimos años ha habido un gran número de publicaciones de estudios metabolómicos aplicados a la producción animal (Goldansaz et al., 2017). Como ejemplos se pueden mencionar estudios relacionados con la eficiencia alimentaria (Karisa et al., 2014), sensibilidad a enfermedades (LeBlanc et al., 2005; Hailemariam et al., 2014), deposición lipídica (Weikard et al., 2010), fertilidad (Chapinal et al., 2012), calidad de la leche (Melzer et al., 2013), etcétera.

En los programas de mejora muchas veces la obtención de los fenotipos es un factor limitante, debido a su elevado coste económico o a la dificultad en la toma de datos o a que es un carácter que no se expresa hasta avanzada edad, o que solo puede medirse tras el sacrificio, etcétera. En este sentido, la metabolómica se puede convertir en una herramienta

clave para refinar y mejorar la descripción de los caracteres y, por consiguiente, mejorar la predicción de los valores genéticos en los programas de selección (Fontanesi, 2016).

1.3 Técnicas analíticas

Como ya se mencionó, el objetivo de los análisis metabolómicos no dirigidos es identificar la mayor cantidad de metabolitos posibles. Uno de los grandes desafíos de estos análisis se debe a la gran diversidad química que presentan los metabolitos que componen el metaboloma. Por este motivo, esta disciplina necesita el uso de varias técnicas (Krastanov, 2010). En los últimos años, las dos técnicas más utilizadas para la obtención de datos metabolómicos fueron la Resonancia Magnética Nuclear (*nuclear magnetic resonance*, NMR) y la espectrometría de masas (*mass spectrometry*, MS) (Fuhrer and Zamboni, 2015), aunque el mayor crecimiento de la metabolómica de los últimos tiempos se ha debido al desarrollo de la MS (Putri et al., 2013a). Para este trabajo, nos centraremos en esta última técnica.

La espectrometría de masas es una técnica analítica basada en la detección de iones que se separan en función de su relación masa/carga (m/z), al aplicar un campo electromagnético (Artati et al., 2012). Los elementos que conforman un espectrómetro de masas son: un inyector de muestra, una fuente de ionización, un analizador de masas, un detector y un sistema de análisis de datos (figura 3). El análisis comienza con la inyección de la muestra en el equipo. Esta muestra pasa por la fuente de ionización que generará los iones que serán luego separados en el analizador, en función de su relación m/z . La fuente de ionización les da carga a las moléculas y las fragmenta, siguiendo un patrón de ionización característico para cada metabolito, que se utiliza posteriormente en la identificación (Dunn et al., 2013). Por último, el detector es el mecanismo que capta las moléculas cargadas y devuelve los resultados en un espectrograma de masas. En estos espectros el eje X representa la relación m/z y el eje Y la abundancia relativa de los metabolitos detectados (Martín Gómez and Ballesteros González, 2010). La identificación de los compuestos se logra mediante la asociación de las masas obtenidas con masas conocidas e identificadas y mediante el patrón de ionización característico (Dunn et al., 2013).

Las variantes en la espectrometría de masas se deben al uso de distintas fuentes de ionización (p.ej. Ionización electrónica (EI), ionización por electrospray (ESI), etcétera) y

de distintos analizadores de masas (p. ej. Cuadripolo simple (Q), triple cuadripolo (QQQ), tiempo de vuelo (TOF), etcétera.). Las fuentes de ionización se diferencian en el rango de peso molecular y de polaridad de los analitos que detectan y los analizadores difieren en el rango de masas detectado, en la velocidad de medición de cada espectro de masas, en la precisión y en la resolución de masas (es decir, la capacidad de separar moléculas con masa similar).

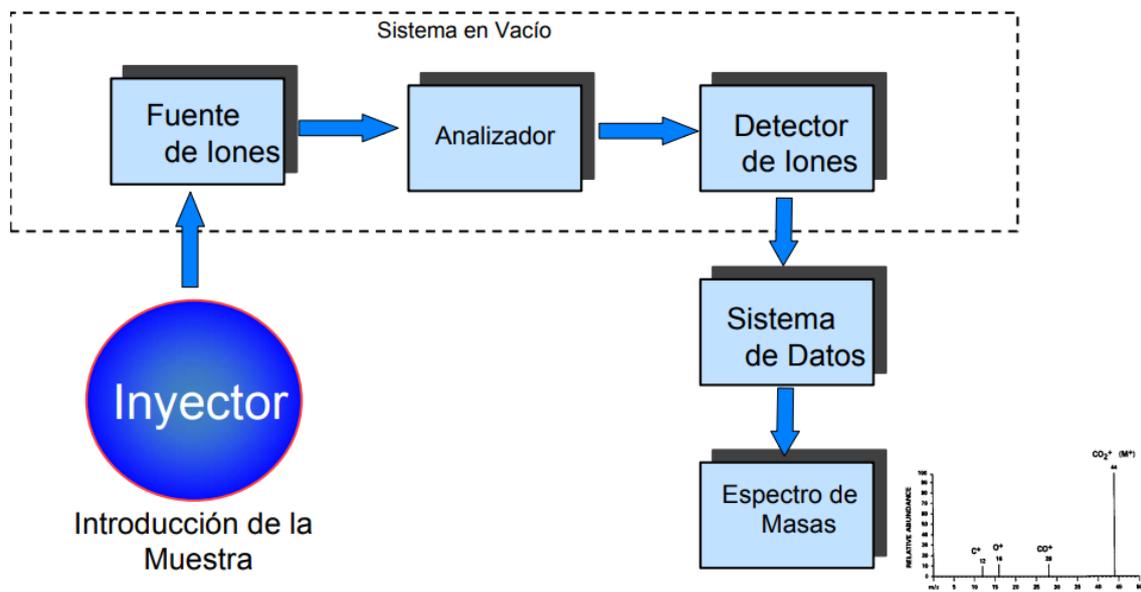


Figura 3. Esquema de análisis por espectrometría de masas. (<https://www.agilent.com/>)

Dada la complejidad de las muestras, la resolución de esta técnica mejora si se acopla con procedimientos previos de separación de los metabolitos. Las técnicas de separación utilizadas normalmente son la cromatografía de gases (*gas chromatography*, GC), la cromatografía líquida (*liquid chromatography*, LC) y la electroforesis capilar (*capilar electrophoresis*, CE) (Naz et al., 2014a). Tras de la separación y el análisis por el espectrómetro de masas, cada metabolito tendrá una relación m/z y un tiempo de retención característico, que se corresponde al tiempo transcurrido desde que se inyecta la muestra en el sistema de separación hasta que es detectado en el espectrómetro. Esta información adicional es de gran utilidad para ayudar a la identificación de los metabolitos (Gil de la Fuente et al., 2018).

La cromatografía gaseosa es capaz de separar compuestos de baja polaridad y volátiles, es decir, compuestos con bajo punto de ebullición que estarán en fase gaseosa en el rango de temperatura de 50-350°C. Es importante que sean térmicamente estables entre estas temperaturas (Mastrangelo et al., 2015). También es posible detectar moléculas que se volatilicen mediante cambios químicos, proceso conocido como derivatización. Para la separación, las muestras son primero vaporizadas en un horno y luego inyectadas en una columna cromatográfica por la fase móvil, que en la cromatografía gaseosa es un gas. El flujo gaseoso que sale de la columna es introducido en el espectrómetro de masas, directamente en la fuente de ionización. La fuente de ionización más utilizada es la ionización electrónica, debido a su robustez y repetibilidad (Putri et al., 2013b). Las ventajas más importantes de la GC-MS son la gran reproducibilidad de los tiempos de retención y de los perfiles de fragmentación, lo que permite obtener un espectro de masa único que facilita la identificación de los compuestos (Putri et al., 2013b). La desventaja de esta técnica está en la necesidad de incluir etapas de derivatización, que pueden generar una reducción de la concentración de los analitos y un aumento de artefactos o de ruido (Gromski et al., 2015).

La cromatografía líquida es una técnica de gran versatilidad. Permite detectar una gran cantidad de metabolitos, de amplio rango de peso molecular. La fase móvil utilizada en esta técnica es un líquido y la fase estacionaria es un sólido poroso. La HPLC (*high performance liquid chromatography*) fue un gran avance en este tipo de tecnología, que permitió aumentar la eficacia de la técnica. Tiene la particularidad que las partículas que recubren el interior de la columna cromatográfica tienen un tamaño muy pequeño, de 3-10 µm aproximadamente. Este tamaño de partículas requiere altas presiones del sistema para una separación óptima. Se diferencian dos tipos de cromatografía líquida según las fases utilizadas: en fase normal y en fase reversa. La LC en fase normal fue la primera en desarrollarse y utiliza una fase estacionaria polar y una fase móvil no polar; en la LC en fase reversa, en cambio, la fase estacionaria es no polar y la fase móvil es polar. El flujo que sale de la columna es introducido en el espectrómetro de masas, directamente a la fuente de ionización, que generalmente es por electrospray (Soriano Lopes et al., 2017).

La electroforesis capilar permite separar metabolitos altamente polares y con carga a través de la aplicación de altos voltajes (García et al., 2017). Los metabolitos se separan según su movilidad electroforética en una solución electrolítica y esta depende de la carga y el tamaño de la molécula. El flujo que sale del capilar se introduce directamente en la fuente de

ionización del espectrómetro de masas. Comúnmente para esta técnica se usa la ionización por electrospray (Hirayama et al., 2014). La baja repetibilidad es una de las mayores desventajas de este método y se debe a la variación en los tiempos de migración por su sensibilidad a los cambios de temperaturas (Putri et al., 2013b).

1.4 Métodos de análisis multivariantes

Los análisis metabolómicos no dirigidos generan cientos de señales por cada muestra analizada. En las matrices de datos obtenidas ($\mathbf{X}_{n \times k}$), cada fila corresponde a una muestra y cada columna contiene la abundancia relativa de un compuesto detectado. Son matrices de gran dimensión, con más variables que observaciones y altamente colineales. Hay varios métodos para trabajar con datos de estas características, pero aún presentan algunas limitaciones, como veremos más adelante. Los métodos de análisis multivariantes permiten analizar grandes matrices de datos a partir de la reducción de la dimensionalidad (Varmuza and Filzmoser, 2009). Cada variable se representa en un eje en un espacio de k dimensiones y cada muestra (en adelante observación) es un punto en dicho espacio. El objetivo, entonces, es transformar la matriz $\mathbf{X}_{n \times k}$ en una nueva matriz con dimensiones $n \times a$, donde $a \ll k$, que conserve la mayor cantidad de información posible de los datos originales, facilite su interpretación y la extracción de conclusiones. Las a nuevas dimensiones se denominan variables latentes y forman un nuevo hiperplano en el espacio donde las n observaciones originales son proyectadas. A estas proyecciones se les denomina *scores* (t). Las variables latentes se calculan de forma tal que no estén correlacionadas entre sí (son ortogonales), por ende, la información recogida por cada variable será independiente de las demás. Los métodos de análisis multivariantes se suelen denominar “data driven”, ya que la información se extrae directamente de los datos observados, sin hacer asunciones previas sobre los efectos que están influyendo; al contrario, en los métodos estadísticos clásicos, o métodos “model driven”, el modelo planteado es decisivo para la inferencia (Varmuza and Filzmoser, 2009).

Previo al análisis, es necesario realizar un pretratamiento de los datos importante para asegurar la calidad y facilitar la interpretación de los resultados (Karaman, 2017), mediante escalado y/o transformación. En primer lugar, el escalado se utiliza debido a que los rangos de medidas de las variables originales pueden ser muy dispares y se debe evitar que variables

muy abundantes enmascaren el efecto de las variables con abundancia intermedia o baja. En metabolómica, el escalado de Pareto es el más comúnmente utilizado ya que suele dar los mejores resultados. Este escalado reduce la importancia de las variables más abundantes y les da mayor importancia a las variables de abundancia intermedia y baja (Karaman, 2017). En segundo lugar, las transformaciones se utilizan, para convertir relaciones multiplicativas en aditivas y para hacer las distribuciones de las variables más simétricas (Karaman, 2017). En la tabla 2 se muestran algunos métodos de escalado y de transformación, con sus ventajas y desventajas.

Tabla 2. Métodos de escalado y transformación. O significa unidad original y (-) sin unidad.

Method	Formula	Unit	Goal	Advantages	Disadvantages
Autoscaling	$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i}$	(-)	Compare metabolites based on correlations	All metabolites become equally important	Inflation of the measurement errors
Range scaling	$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{(x_{i_{\max}} - x_{i_{\min}})}$	(-)	Compare metabolites relative to the biological response range	All metabolites become equally important. Scaling is related to biology	Inflation of the measurement errors and sensitive to outliers
Pareto scaling	$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{\sqrt{s_i}}$	O	Reduce the relative importance of large values, but keep data structure partially intact	Stays closer to the original measurement than autoscaling	Sensitive to large fold changes
Vast scaling	$\tilde{x}_{ij} = \frac{(x_{ij} - \bar{x}_i)}{s_i} \cdot \frac{\bar{x}_i}{s_i}$	(-)	Focus on the metabolites that show small fluctuations	Aims for robustness, can use prior group knowledge	Not suited for large induced variation without group structure
Level scaling	$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{\bar{x}_i}$	(-)	Focus on relative response	Suited for identification of e.g. biomarkers	Inflation of the measurement errors
Log transformation	$\tilde{x}_{ij} = 10 \log(x_{ij})$ $\hat{x}_{ij} = \tilde{x}_{ij} - \bar{\tilde{x}}_i$	Log O	Correct for heteroscedasticity, pseudo scaling. Make multiplicative models additive	Reduce heteroscedasticity, multiplicative effects become additive	Difficulties with values with large relative standard deviation and zeros
Power transformation	$\tilde{x}_{ij} = \sqrt{(x_{ij})}$ $\hat{x}_{ij} = \tilde{x}_{ij} - \bar{\tilde{x}}_i$	√O	Correct for heteroscedasticity, pseudo scaling	Reduce heteroscedasticity, no problems with small values	Choice for square root is arbitrary.

Adaptado de van den Berg et al. (2006).

Hay distintos métodos de análisis multivariante y cada uno tiene objetivos diferentes (Worley and Powers, 2012). Para este trabajo nos centraremos en el análisis de componentes principales (PCA), proyección sobre estructuras latentes (PLS), proyección ortogonal sobre estructuras latentes (OPLS) y los métodos discriminantes proyección sobre estructuras latentes–análisis discriminante (PLS-DA) y proyección ortogonal sobre estructuras latentes-

análisis discriminante (OPLS-DA). Para evaluar cómo el modelo calculado por el método se adapta a los datos, hay dos parámetros importantes: la capacidad de ajuste (R^2) y la capacidad de predicción (Q^2), que se explicarán más adelante.

1.4.1 Análisis de componentes principales

El PCA es el método de reducción de dimensionalidad más utilizado, tanto en metabolómica como en otras áreas (Worley and Powers, 2012). Es un método potente y claro para observar relaciones entre las variables y es utilizado también como un análisis exploratorio, ya que permite ver patrones, tendencias y detectar outliers. Un método estadístico tradicional para observar relaciones entre variables es la matriz de correlaciones, pero debido a la gran cantidad de datos de los análisis metabolómicos resulta difícil de interpretar.

En el PCA las variables latentes (también denominadas componentes principales) son rectas en el espacio \mathbf{X} que maximizan la varianza de las observaciones proyectadas (*scores*, t). La proyección se realiza conservando al máximo la distancia entre los puntos. De esta manera, la primera variable latente ($\mathbf{p1}$) es la recta en el espacio \mathbf{X} que mejor aproxima los datos recogiendo la máxima variación (Varmuza and Filzmoser, 2009). La segunda variable latente ($\mathbf{p2}$) será ortogonal a la primera y su dirección será la que capte la segunda mayor variación de los datos, independiente de la anterior (figura 4). Los elementos de cada vector \mathbf{p} son los pesos o *loadings* de cada variable en esa variable latente. Cada nueva componente calculada, explicará un porcentaje determinado de la varianza de los datos de \mathbf{X} , que se estima a partir del cálculo del coeficiente de determinación R^2 .

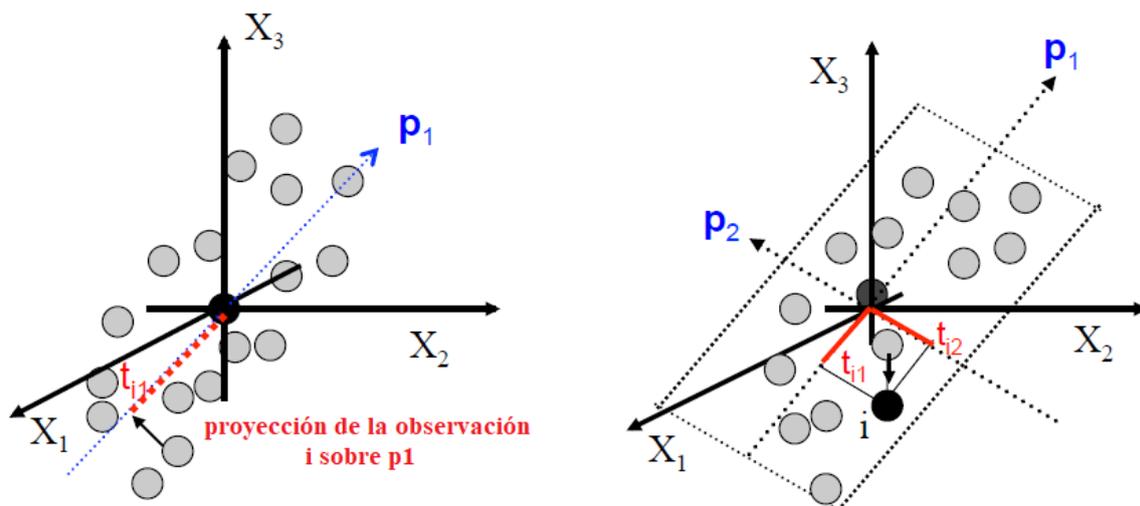


Figura 4. Trazado de la primera ($\mathbf{p1}$, izquierda) y segunda ($\mathbf{p2}$, derecha) componente principal y proyecciones de las observaciones sobre las componentes.

El número de componentes óptimos puede estimarse a partir de métodos de validación cruzada, que calculan el poder predictivo del análisis. Estos métodos se basan en la eliminación de grupos de datos, el cálculo de un nuevo modelo a partir de los datos restantes y la predicción de los valores de los datos eliminados. Se realizan muchas iteraciones del mismo proceso y, a partir de los datos predichos y observados, se calcula el *predictive residual sum of squares* (PRESS). A partir de los resultados de todas las iteraciones se calcula la nueva R^2 correspondiente, que en este caso se le llama Q^2 y representa la capacidad de predicción del modelo (Varmuza and Filzmoser, 2009). Este parámetro es un buen indicador para decidir hasta qué componente incluir en el modelo. Un aumento del R^2 acompañado de una disminución o estabilización del valor de Q^2 significa que hay un sobreajuste.

Las a nuevas variables latentes son combinaciones lineales de las k variables originales. De esta manera, la matriz de datos \mathbf{X} se descompone en *loadings* (\mathbf{p}) y *scores* (\mathbf{t}) como se muestra en la ecuación 1. Se puede comprobar que esta descomposición es la clásica descomposición en vectores y valores propios de la matriz de varianzas-covarianzas de \mathbf{X} (Abdi and Williams, 2010). Los vectores de *loadings* son los vectores propios asociados a los a mayores valores propios de la matriz de varianzas-covarianzas y, a su vez, estos valores propios son las varianzas explicadas por cada componente. El vector de *loadings* $\mathbf{p1}$ ($k \times 1$, siendo k : número de variables) resume las variables y define la dirección que toma la primera variable latente y el vector de *scores* $\mathbf{t1}$ ($n \times 1$, siendo n : número de muestras)

corresponde a las proyecciones de cada una de las observaciones sobre la primera variable latente (figura 4). La matriz **E** contiene la variación no explicada por el modelo (residuos).

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}'_1 + \mathbf{t}_2 \mathbf{p}'_2 + \dots + \mathbf{t}_a \mathbf{p}'_a + \mathbf{E} = \mathbf{TP}' + \mathbf{E} \quad (1)$$

Una de las ventajas de este tipo de análisis es la posibilidad de descomponer las grandes matrices de datos y visualizarlas en gráficos que resumen el comportamiento de las observaciones y de las variables. A partir del gráfico de dispersión de los *scores*, donde en cada eje está representada una variable latente, se pueden visualizar los outliers. La herramienta utilizada para su detección es la T^2 de Hotelling y es la generalización multivariante de la distribución t de Student. La T^2 calcula la distancia desde la proyección de una observación hasta el centro del hiperplano formado por las nuevas componentes. La elipse de T^2 de Hotelling al 95% provee una región por fuera de la cual se considera que las observaciones forman parte del 5% de los casos extremos (figura 5). El gráfico de *loadings*, por su parte, también es un gráfico bidimensional en donde cada eje representa un componente y permite visualizar el peso que tienen las variables en cada uno (figura 6).

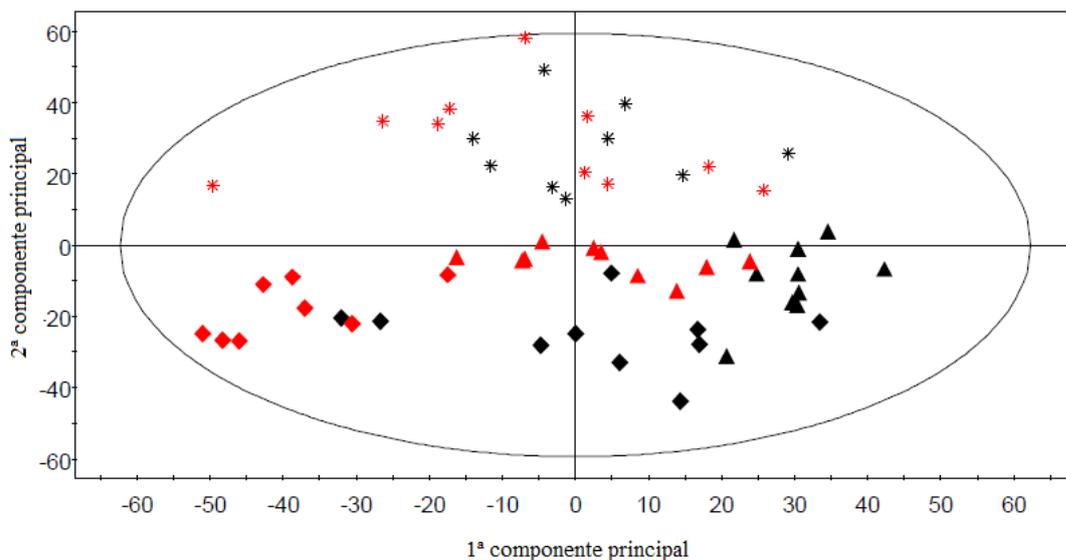


Figura 5. Ejemplo de un gráfico de *scores*, Asteriscos, rombos y triángulos rojos y negros indican grupos diferentes (Curso “Multivariate Data Analysis for Omics”, Umetrics).

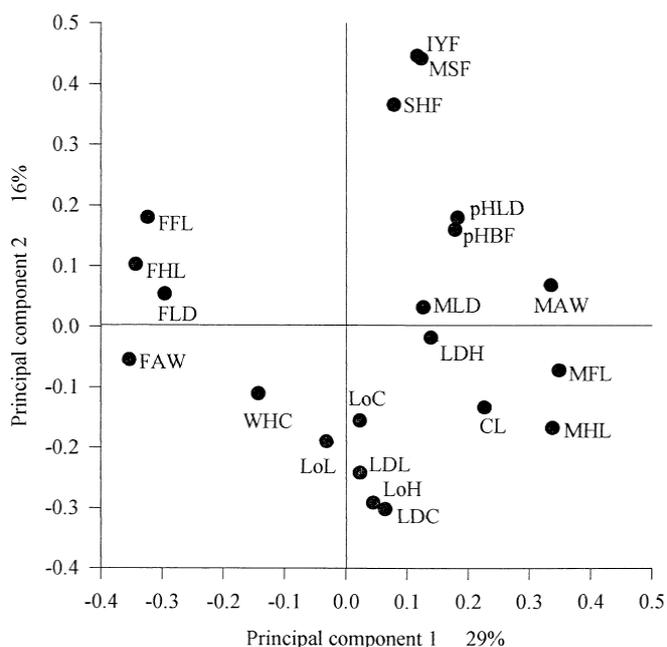


Figura 6. Ejemplo de gráfico de *loadings*, con diferentes variables de calidad de carne (Hernández et al., 1998).

El PCA capta la mayor variación de los datos, sin poder discriminar a qué se debe dicha variación. Sin embargo, frecuentemente lo que interesa es investigar la relación que existe entre los datos contenidos en \mathbf{X} con una variable y . Para esto se desarrollaron los métodos de análisis supervisado, que reciben este nombre debido a que se les da una información (y) que guiará el trazado de las componentes principales. El análisis proyección sobre estructuras latentes, y su variante la proyección ortogonal sobre estructuras latentes, son ejemplos de este tipo de métodos.

1.4.2 Proyección sobre estructuras latentes (PLS)

El objetivo del PLS es encontrar las relaciones entre un conjunto de datos \mathbf{X} y una variable y . Las ecuaciones de regresión múltiple utilizadas normalmente presentan un problema cuando hay más variables que observaciones y cuando presentan colinealidad. Varios métodos se han desarrollado para tratar con estos inconvenientes, como la *ridge regression*, métodos bayesianos (p.ej. Bayes A, Bayes B), etcétera, pero estos presentan limitaciones. Por ejemplo, cuando las variables están muy correlacionadas los resultados son más difíciles de interpretar, ya que los coeficientes no son proporcionales al efecto de las variables, mientras que en los análisis multivariantes las componentes son independientes entre sí;

además, la colinealidad genera grandes errores estándar por las dificultades de invertir la matriz. La solución multivariante a este problema, el PLS, fue originalmente desarrollado por Herman Wold (Wold, 1966). El procedimiento consiste en encontrar las direcciones en el espacio que maximicen la covarianza entre las variables latentes del conjunto de datos \mathbf{X} y la variable y (figura 7). Como en el PCA, estas direcciones (componentes) formarán un nuevo hiperplano donde se proyectarán las variables. Para la descomposición de la matriz \mathbf{X} para el PLS, el primer componente se calcula como una variable latente que tenga la máxima covarianza entre los *scores* y la variable y . A continuación, la información de esa componente se resta de la matriz de datos \mathbf{X} y, a partir de la matriz residual \mathbf{X} , se calculará el siguiente componente, de la misma manera que se calculó el primero. El procedimiento continúa hasta que no se obtenga ninguna mejora en el modelado de la variable y a partir de la matriz \mathbf{X} (Varmuza and Filzmoser, 2009). Este análisis también puede aplicarse cuando la variable y es un conjunto de datos multivariantes (\mathbf{Y}).

Los parámetros de ajuste utilizados por el PLS son similares a los del PCA, con la diferencia que el R^2 se denomina R^2_Y y está determinado por la covarianza entre las variables latentes del conjunto de datos en \mathbf{X} y la variable y . El Q^2 se calcula de la misma manera que para el PCA. Ambos parámetros se utilizan para estimar el número de componentes óptimo del modelo.

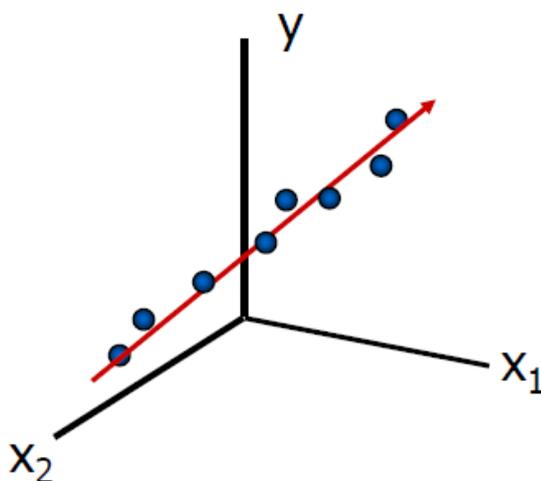


Fig. 7. Primer componente del PLS. La dirección que toma es aquella que maximiza la covarianza entre la matriz \mathbf{X} y la variable y .

1.4.2.1 Proyección sobre estructuras latentes–análisis discriminante (PLS-DA).

El PLS-DA es un caso particular del PLS, donde se busca discriminar el conjunto de datos \mathbf{X} , en dos o más clases. Tendremos, en este caso, un vector \mathbf{y} compuesto de 1 y 0, que clasificará las observaciones por grupos. Los análisis discriminantes con más de 2 clases, aunque son posibles, son más complejos de interpretar. En este caso, la dirección de las componentes serán las que logren la mayor separación de las medias, teniendo en cuenta la variabilidad (figura 8).

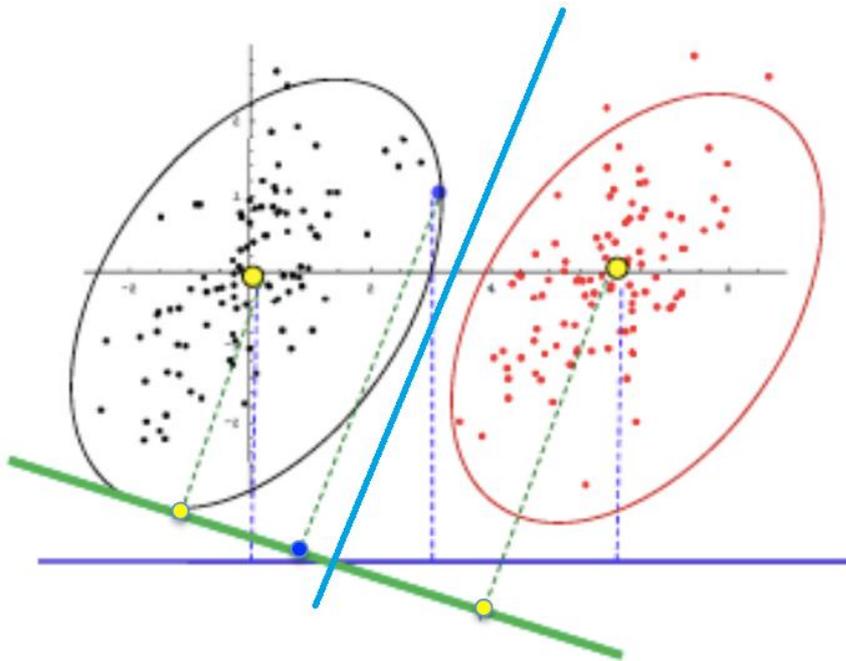


Figura 8. Discriminación en dos grupos (puntos rojos y negros). La recta verde representa la dirección que mayor separación logra. Los puntos amarillos en el centro de las elipses representan la media (Curso “Análisis, monitorización y diagnóstico de procesos multivariantes”. Ferrer Riquelme, Alberto J.).

1.4.3 Proyección a estructuras latentes ortogonales (OPLS)

El OPLS es una extensión del PLS, que resulta útil para mejorar la interpretación. Utiliza la información de la variable \mathbf{y} para descomponer la matriz \mathbf{X} en dos bloques, uno que contiene la información correlacionada (variación predictiva) y otro de información no correlacionada a \mathbf{y} (ortogonal) (Bylesjo et al., 2006). Lo desarrollaron Trygg y Wold en el 2002 (Trygg and Wold, 2002) y sirve para centrar la información correlacionada en una componente predictiva, mientras que la no correlacionada se recogerá en las componentes ortogonales. El OPLS y el PLS tienen la misma capacidad predictiva, pero la interpretación

se facilita, particularmente cuando hay grandes cantidades de datos correlacionados con la y .

1.4.3.1 Proyección ortogonal sobre estructuras latentes–análisis discriminante (PLS-DA).

Al igual que en el PLS-DA, el OPLS-DA discrimina el conjunto de datos X en dos o más clases. En la figura 9 se muestra un ejemplo simple de discriminación en 2 clases, donde se ve cómo el OPLS-DA (derecha) permite diferenciar los dos grupos con la información recogida por la componente predictiva.

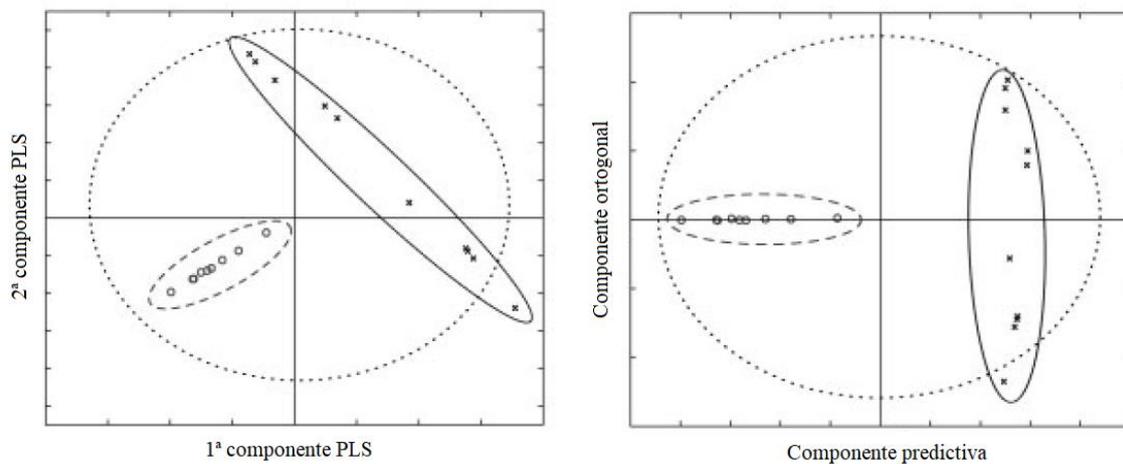


Figura 9. Conjunto de datos discriminados mediante PLS-DA (izquierda) y OPLS-DA (derecha) (Bylesjo et al., 2006)

2. OBJETIVOS

Los objetivos de este trabajo fueron:

- Implementar métodos multivariantes para el análisis de datos metabolómicos obtenidos de un estudio realizado en dos líneas divergentes de conejos seleccionadas por grasa intramuscular.
- Buscar metabolitos diferencialmente expresados entre las líneas de alta y baja grasa intramuscular que permitan identificar las rutas metabólicas involucradas en la diferenciación de este carácter.

3. MATERIALES Y MÉTODOS

3.1 Animales y recolección de muestras

El material animal utilizado para este estudio proviene de un experimento de selección divergente por grasa intramuscular (GIM) realizado durante 9 generaciones, en conejos. Se crearon dos líneas, una de alto contenido de GIM y otra de bajo contenido de GIM. En la novena generación las líneas se conformaron de 10 machos y 60 hembras. Los animales fueron criados en una granja de la Universitat Politècnica de València con ambiente controlado, régimen de iluminación de 16:8 hs, ventilación controlada y sistema de refrigeración en verano. En el momento del nacimiento las camadas se homogeneizaron a 9 gazapos mediante adopciones. Desde el destete hasta el sacrificio, a las 4 y a las 9 semanas de edad, respectivamente, se mantuvieron en jaulas comunes con 9 conejos por jaula. Su alimentación fue *ad libitum*, con pienso comercial compuesto de 15.1% de proteína cruda, 14.5% de fibra cruda y 2.47% de grasa. El contenido de ácidos grasos del pienso fue (expresados como % del total) 0.49% de C14:0, 19.4% de C16:0, 0.68% de C16:1, 2.77% de C18:0, 20.5% de C18:1n-9, 48.1% de C18:2n-6, 6.80% de C18:3n-3 y 1.26% de C > 20.

Los animales fueron sacrificados mediante exanguinación previo aturdimiento eléctrico. El contenido de GIM fue medido en el músculo *Longissimus thoracis y lumborum* (LTL) en dos hermanos completos, un macho y una hembra, del primer parto de cada hembra. El músculo LTL fue disecado, molido y liofilizado previo al análisis por espectroscopía de infrarrojo cercano (modelo 5000, NIRSystem INC., Silver Spring, MD, USA), utilizando las ecuaciones de calibración calculadas por Zomeño et al. (2011). El valor de GIM fue expresado en g/100 g de músculo fresco. Más detalles del experimento de selección divergente pueden encontrarse en Zomeño et al. (2013) y Martínez-Álvaro et al. (2016).

Para este experimento se escogieron 20 individuos de cada línea. Para la elección, se tuvieron en cuenta las distribuciones de las poblaciones de alta y baja grasa intramuscular. Se dividió cada distribución en cuartiles y se tomaron individuos de cada uno de estos cuartiles al azar, proporcionalmente al porcentaje representado. Además, se tomó como criterio de elección seleccionar dos muestras por cada familia de macho, intentando conservar la proporción de sexos. El 55% de los 40 individuos fueron machos y el 45% restante hembras. Se tomó 1 mL de sangre en un tubo de heparina en el momento del sacrificio, durante la exanguinación. Se separó el plasma mediante centrifugación a 3000

rpm durante 10 min a temperatura ambiente y fue almacenado a -80°C hasta el envío al Centro de Metabolómica y Bioanálisis (CEMBIO), Facultad de Farmacia, Universidad CEU San Pablo, Madrid, para su posterior análisis.

3.2 Análisis de las muestras

3.2.1 Técnica analítica

Se realizó un análisis metabolómico no dirigido. La técnica utilizada fue la espectrometría de masas (MS) acoplada a dos técnicas previas de separación: cromatografía líquida (LC) y electroforesis capilar (CE).

3.2.2 Preparación de muestras para control de calidad (QC).

Las muestras para los QC se prepararon mezclando alícuotas de las 40 muestras de plasma. Para cada técnica se preparó un conjunto independiente de muestras de QC. Las muestras para QC se inyectaron de forma continua al principio de cada análisis hasta equilibrar el sistema y luego a intervalos regulares entre las muestras para proveer evidencia del buen funcionamiento y estabilidad de la técnica. También se agregó una muestra blanco en cada análisis.

3.2.3 Cromatografía líquida – Espectrometría de masas (LC-MS).

Para la precipitación de proteínas y extracción de metabolitos de las muestras, de los individuos y de los controles de calidad, se mezclaron 100 μl de plasma con 300 μl de una mezcla de metanol/etanol 1:1 (v/v) a baja temperatura. Las muestras fueron almacenadas durante 5 minutos en hielo y posteriormente centrifugadas durante 20 minutos a 15000 g y 4°C . El sobrenadante se transfirió directamente a los viales para su análisis.

Las muestras, de los individuos, de los controles de calidad y el blanco, fueron analizadas en un sistema HPLC 1290 Infinity acoplado a un espectrómetro de masas QTOF 6520, con una fuente de ionización por electrospray (ESI) con tecnología Jet Stream (Agilent Technologies, Waldbronn, Germany), con la cual se trabajó en polaridad negativa. Para la separación se utilizó una columna de fase reversa (Supelco Discovery HS C18 150 mm \times 2.1 mm, 3 μm , Sigma-Aldrich) y se mantuvo a 40°C durante todo el procedimiento. El flujo del sistema fue constante a 0.6 mL/min, con una fase móvil compuesta por dos solventes: agua con 0.1% de ácido fórmico (solvente A) y acetonitrilo con 0.1% de ácido fórmico (solvente B). El gradiente comenzó con 2% de solvente B (0-5 min), seguido de un gradiente lineal de 2 a 50% de B (6-15 min), que se mantuvo a 50% por 2 min (15-17 min), luego

aumentó a 95% de B (17-18 min) y se retuvo así por 2 min (18-20 min). Finalmente, se retornó a las condiciones iniciales (2% de B) en 1 min (20-21 min) y estas se mantuvieron con la finalidad de re-equilibrar la columna, hasta completar los 25 min. El detector realizó un barrido de 100 a 1100 m/z, a una velocidad de 1 espectro por segundo. El voltaje del capilar fue de 4000 V. La presión del nebulizador fue de 35 psi y el flujo de gas de 11 L/min. El voltaje del fragmentador fue de 75V y el voltaje del octopolo de 750 V (Ciborowski et al., 2010). La precisión de las medidas m/z tomadas por el equipo fue controlada mediante un sistema de calibración automático que inyectaba de forma continua dos soluciones de referencia que contenían masas de 112.9856 m/z (ácido trifluoroacético) y 922.0098 m/z (HP-0921). El orden de inyección fue aleatorio, incluyendo cada cinco muestras de individuos una de QC.

3.2.4 Electroforesis capilar – Espectrometría de masas (CE-MS).

Para la preparación de las muestras de los individuos, de los controles de calidad y del blanco, se mezclaron 100 µl de plasma con 100 µl de una solución 0.2 M de ácido fórmico que contenía acetonitrilo al 5%, y 0.4 mM de metionina sulfona como estándar interno (IS) (Naz et al., 2013). La mezcla se agitó vigorosamente, se transfirió a una unidad de ultrafiltración Centrifree (Milipore Ireland, Eire) con un filtro de 30kDa y se centrifugó a 2000 g, durante 70 minutos, a 4°C. El filtrado obtenido se transfirió directamente a los viales para su análisis.

El análisis se realizó en un sistema de electroforesis capilar Agilent 7100 acoplado a un espectrómetro de masas TOF 6224 (Agilent Technologies, Waldbronn, Germany), con una fuente de ionización por electrospray. La separación se llevó a cabo usando un capilar de sílice fundida de 100 cm de longitud y 50 µm de diámetro (Agilent Technologies, Waldbronn, Germany). El capilar se acondicionó previo al análisis durante 30 min con una solución NaOH 1M, seguido de 30 min con agua MilliQ® y solución electrolítica (BGE, por sus siglas en inglés) (ácido fórmico 0.8 M en metanol 10%), a 20°C, en polaridad normal. Antes de cada inyección el capilar fue lavado durante 5 min (a 950 mBar de presión) con BGE. Las muestras se inyectaron hidrodinámicamente mediante la aplicación de 50 mBar de presión durante 50 s. Luego de cada inyección, se inyectó BGE a 100 mBar durante 10 s para mejorar la repetibilidad. Las condiciones para la separación fueron 25 mBar de presión y 30 kV de voltaje durante 40 min. La espectrometría de masas se realizó en polaridad positiva, el detector realizó un barrido de 80 a 1000 m/z y las condiciones fueron las siguientes: flujo y temperatura del gas de secado de 10L/min y 200°C, respectivamente,

presión del nebulizador 10 psi, voltaje del capilar 3.5kV, fragmentador 125V y skimmer 65V. El líquido auxiliar utilizado para la detección estaba compuesto de metanol/agua 1:1 (v/v), 1.0 mM de ácido fórmico y 2 estándares de referencia que permitieron una corrección constante de la m/z detectada: 121.0509 m/z (purina) y 922.0098 m/z (HP-0921) (Naz et al., 2013). El orden de inyección fue aleatorio, incluyendo cada seis muestras de individuos una de QC.

3.3 Tratamiento de datos

Los datos obtenidos fueron procesados usando el software MassHunter Profinder (B.08.00, Agilent Technologies, Waldbronn, Germany) para obtener una matriz de datos estructurada y en formato adecuado. Los datos crudos fueron analizados mediante el algoritmo RFE (*recursive feature extraction*), que reduce la cantidad y la complejidad de los datos eliminando información no específica y extrayendo las señales importantes. Este algoritmo abarca para cada señal la abundancia de la molécula, la masa y el tiempo de retención en cada una de las muestras. El alineamiento y filtrado de los datos fue realizado en el software Mass Profiler Professional 14.8 (Agilent Technologies, Waldbronn, Germany). La calidad de los datos en LC-MS y CE-MS se aseguró al mantener solamente aquellas señales que, en las muestras de QC, tenían un coeficiente de variación (CV) por debajo del 30% y que estuvieran presentes en el 75% de la totalidad de las muestras en estudio. Además, se eliminaron las señales que presentaban en el blanco una abundancia relativa a la abundancia promedio de las muestras mayor al 20%, según la ecuación 2. Los valores faltantes se estimaron usando el algoritmo *k*NN (*k*-nearest neighbors) (Troyanskaya et al., 2001).

$$\frac{\text{abundancia señal}_k \text{ en blanco}}{\text{abundancia promedio señal}_k \text{ de todas las muestras}} \times 100 > 20\% \quad (2)$$

3.4 Análisis estadístico y selección de variables

Los datos fueron analizados usando análisis univariante (AUV; software MATLAB R2015a (Mathworks, Inc., Natick, USA)) y multivariantes (AMV; software SIMCA P+ 15.0.1, Umetrics (Umea, Sweden)).

Para el AUV se utilizó el test-t en el análisis LC-MS y la prueba no paramétrica de Kruskal-Wallis en el caso de CE-MS, para obtener las señales significativas. En ambos casos se aplicó la corrección para multi-test de Benjamini-Hochberg (FDR, *false discovery rate*) (Benjamini and Hochberg, 1995). El nivel de significación establecido fue del 95% ($p < 0.05$).

Para el AMV los datos fueron previamente transformados y escalados. Se probaron distintas combinaciones de transformación y escalado de datos y se utilizó la transformación logarítmica y el escalado por Pareto. Como análisis exploratorio se realizó un análisis de componentes principales (PCA) para evaluar la calidad de los datos, detectar outliers y buscar patrones o tendencias en las muestras. Como análisis discriminantes se realizaron la proyección sobre estructuras latentes- análisis discriminante (PLS-DA) y la proyección ortogonal sobre estructuras latentes- análisis discriminante (OPLS-DA), clasificando los individuos por la línea de grasa intramuscular a la que pertenecen (alta o baja GIM).

Para seleccionar las variables que contribuían a la discriminación se utilizó el parámetro VIP (*variable importance in projection*) (Galindo-Prieto et al., 2014) y se evaluó la relación señal/ruido a partir del $p(\text{corr})$ del gráfico-S (Wiklund et al., 2008). Se calculó también el porcentaje de cambio (FC por sus siglas en inglés, ecuación 3) de cada variable entre las líneas de alta y baja grasa intramuscular, como forma de evaluar la relevancia biológica de las variables.

$$\frac{\text{abundancia promedio de la señal en línea alta GIM}}{\text{abundancia promedio de la señal en línea baja GIM}} \times 100 - 100 \quad (3)$$

Se resta 100 para que la igualdad de abundancia entre las líneas tenga valor 0. De esta manera, los valores positivos indican mayor expresión en la línea de alta GIM y los valores negativos indican mayor expresión en la línea de baja GIM. Se tomó como criterio de selección de variables un FC% de $\pm 15\%$.

3.5 Identificación de metabolitos

La identificación tentativa de los metabolitos para LC-MS y CE-MS se realizó sobre las masas experimentales de las variables que fueron seleccionadas. Se buscaron estas masas en bases de datos públicas disponibles online: KEGG, Metlin, LipidMaps y HMDB, usando la herramienta web CEU Mass Mediator (Gil de la Fuente et al., 2018).

4. RESULTADOS Y DISCUSIÓN

En primer lugar, se ofrecen los resultados de la selección divergente en la novena generación de selección de las líneas de alta y baja GIM, de las que provienen los datos. A continuación, se exponen los resultados del análisis de componentes principales (PCA), utilizado como análisis exploratorio y de los análisis discriminantes proyección sobre estructuras latentes– análisis discriminante (PLS-DA) y proyección ortogonal sobre estructuras latentes- análisis discriminante (OPLS-DA). Finalmente, se indican las variables más importantes en la discriminación de las dos líneas y se presentan conjeturas sobre su importancia biológica.

4.1 Resultados de la novena generación de selección por grasa intramuscular

Para la estimación de la respuesta a la selección se utilizaron 729 animales. La respuesta directa a la selección en la novena generación fue de 0.44 g/100 g de músculo, lo que representa 3.1 desviaciones típicas del carácter. La media del carácter fue de 1.09 y la desviación típica fue de 0.12. El progreso genético fue del 41% de la media tras 9 generaciones de selección.

4.2 Análisis exploratorios

Se realizó un PCA para comprobar la presencia de outliers e identificar patrones en los datos. El objetivo del PCA es resumir la información de un gran número de variables en unas pocas variables latentes, intentando evitar que haya un sobreajuste a medida que se agregan nuevas componentes. En la espectrometría de masas acoplada a cromatografía líquida (LC-MS) se obtuvieron 2687 variables y tras el control de calidad de los datos, se obtuvo una matriz con 456 variables. En la figura 10 se muestran los parámetros R^2 y Q^2 de las 4 primeras componentes principales. Tras evaluar el incremento de la R^2 y la Q^2 a medida que se añaden las componentes principales, se seleccionaron las tres primeras componentes, que explican una variabilidad del 60% (R^2) y tienen una capacidad de predicción del 50% (Q^2). Se evaluaron los parámetros de ajuste de la tercera y cuarta componente y se concluyó, por un lado, que el valor Q^2 se había estabilizado y, por otro lado, que la variabilidad explicada por la cuarta componente era muy pequeña y no aportaba nuevos elementos que ayudaran a la interpretación de los resultados (Abdi and Williams, 2010).

En las figuras 11, 12 y 13 se muestran los gráficos de dispersión de los *scores*. Para la representación y la detección de outliers se construye una elipse con 95% de nivel de confianza a partir del estadístico T^2 de Hotelling. Las proyecciones de los controles de calidad se encuentran agrupadas y centradas en la elipse; esto se debe a que, dado que son un pool de todas las muestras, recogen toda la variabilidad presente en ellas (Naz et al., 2014b). Los *scores* que caen fuera de la elipse forman parte del 5% de los casos extremos. Los outliers se evaluaron individualmente y no se evidenció ningún error o problema en el análisis de esas muestras. Además, se realizaron los análisis con y sin outliers y no se obtuvieron diferencias importantes, por lo tanto, dado que se dispone sólo de 20 muestras por grupo, se conservaron los outliers. Los *scores* en la gráfica están nombrados según el número experimental de cada individuo, cuyos detalles se encuentran en la tabla 1 en Anexos.

Cuando se observa la 1ª con la 2ª componente (figura 11) y la 1ª con la 3ª (figura 12), se advierte un patrón de discriminación entre las muestras correspondientes a las líneas de alta y de baja GIM, que permite prever que el análisis posterior podrá discriminar entre los grupos. Al representar la 2ª con la 3ª componente (figura 13) no se observa ninguna diferenciación entre estos grupos. Se observa que la 1ª componente es la que recoge la mayor información responsable de la separación de los grupos.

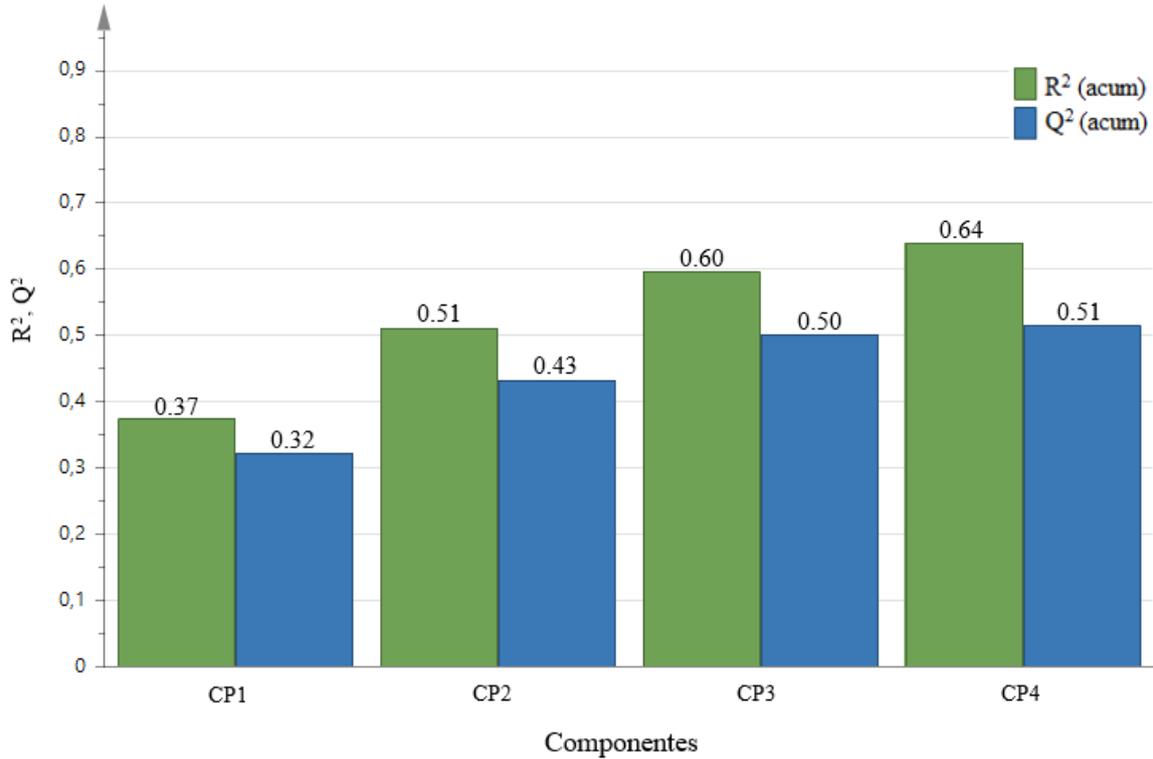


Figura 10. Variabilidad explicada (R^2) y capacidad de predicción (Q^2) de las primeras 4 componentes principales del PCA realizado con los datos obtenidos de la LC-MS. R^2 (acum): R^2 acumulada, Q^2 (acum): Q^2 acumulada.

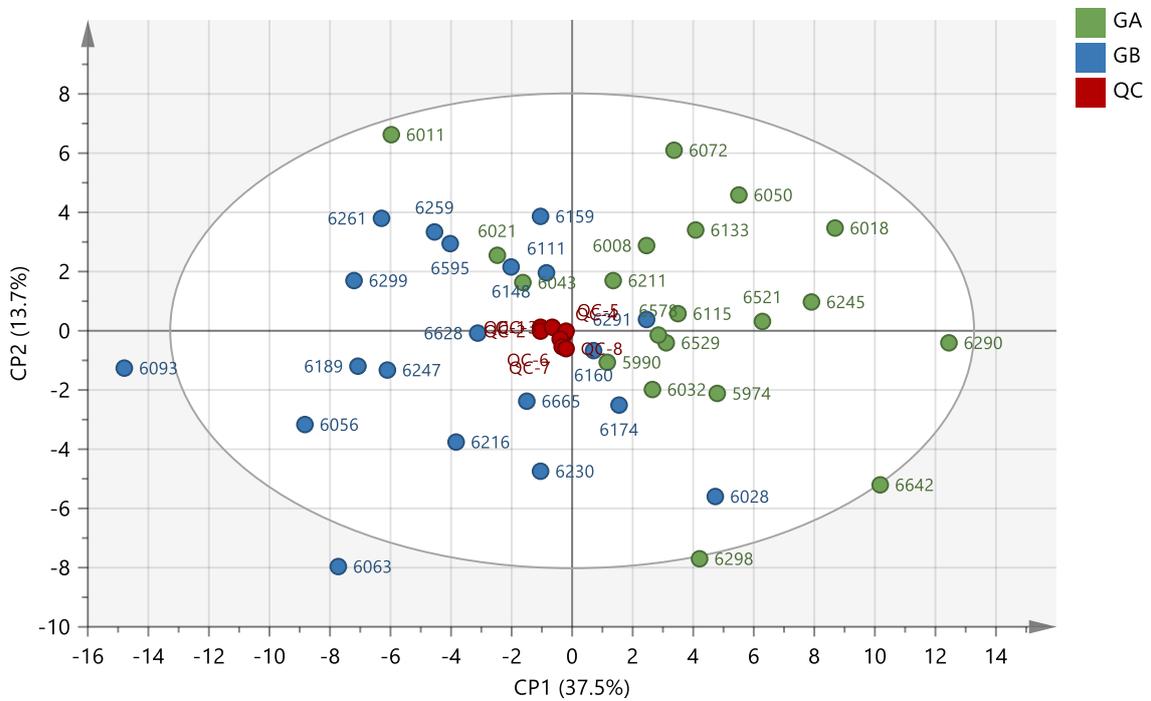


Figura 11. Gráfico de dispersión de los *scores* de la 1ª vs 2ª componente principal para los resultados de LC-MS. Elipse T^2 de Hotelling (95%). Cada punto corresponde a una muestra. CP: componente principal; entre paréntesis, la variabilidad explicada por cada componente; QC: control de calidad; GA, GB: línea de alta y baja grasa intramuscular respectivamente.

En la espectrometría de masas acoplada a electroforesis capilar (CE-MS) se obtuvieron 447 variables y después de aplicar los controles de calidad a los datos, se obtuvo una matriz con 208 variables. En la figura 14 se exponen los parámetros de ajuste (R^2 y Q^2) de las 3 primeras componentes principales. De la misma manera que se explicó anteriormente, se evaluó el incremento de la R^2 y la Q^2 y se seleccionaron las dos primeras componentes, que explican una variabilidad del 52% (R^2) y tienen una capacidad de predicción del 44% (Q^2). Se observó que el valor Q^2 se había estabilizado y que la variabilidad explicada por la tercera componente no aportaba nuevos elementos que ayudaran a la interpretación de los resultados (Abdi and Williams, 2010).

En la figura 15 se presenta el gráfico de dispersión de los *scores* de la 1ª con la 2ª componente. De nuevo, para la representación y la detección de outliers se construye una elipse con 95% de nivel de confianza a partir del estadístico T^2 de Hotelling. Las proyecciones de los controles de calidad se encuentran agrupadas y centradas en la elipse. Nuevamente, se evaluaron los outliers y, al igual que ocurrió en el análisis anterior, no se evidenció ningún error o problemas en el análisis de esas muestras. Tampoco se detectaron diferencias importantes al realizar pruebas con y sin outliers. Por los mismos motivos utilizados anteriormente, se decidió que los outliers podían conservarse. A diferencia de los resultados obtenidos con los datos de la LC-MS, con esta técnica no se aprecia diferenciación entre los grupos de alta y baja GIM. Este tipo de análisis es capaz de mostrar separación entre grupos solamente cuando la variabilidad dentro y entre grupos lo permite (Worley et al., 2013), ya que su objetivo no es la discriminación, sino encontrar unas pocas direcciones de proyección de las variables que permitan explicar la máxima variabilidad, sin tener en cuenta a qué se debe dicha variabilidad.

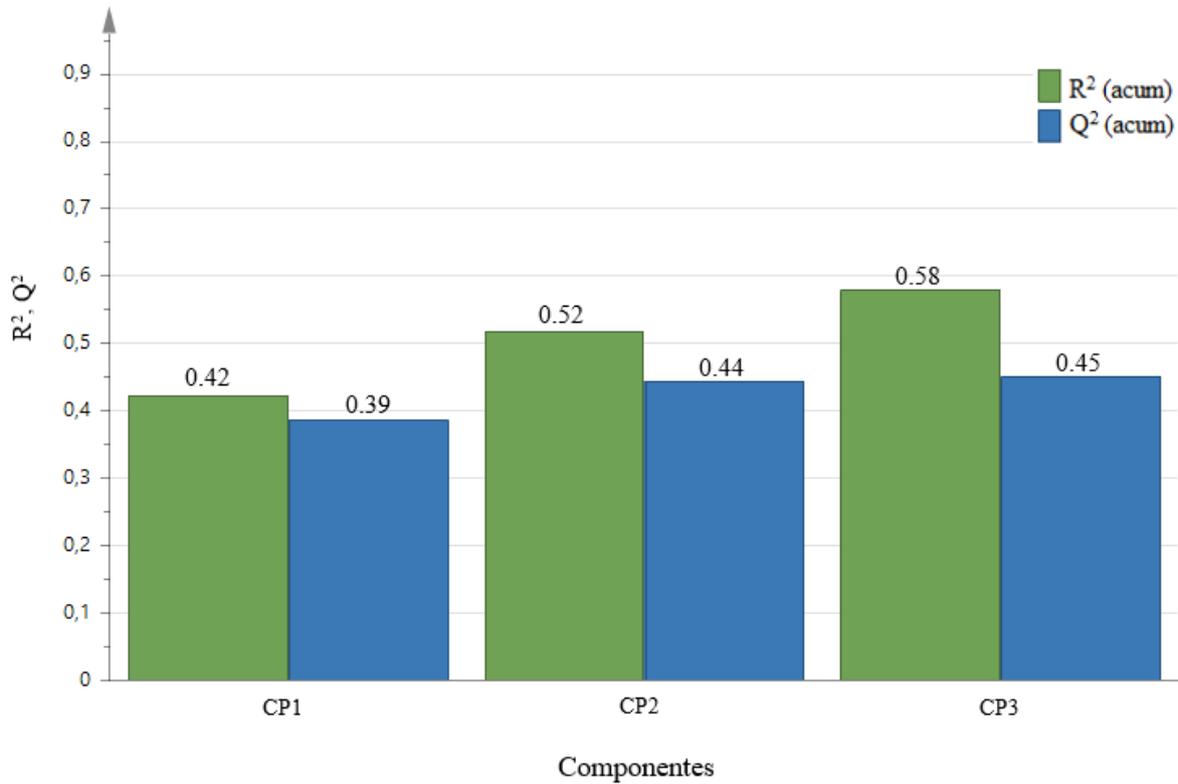


Figura 14. Variabilidad explicada (R^2) y capacidad de predicción (Q^2) de las primeras 3 componentes principales del PCA realizado con los datos obtenidos de la CE-MS. R^2 (acum): R^2 acumulada, Q^2 (acum): Q^2 acumulada.

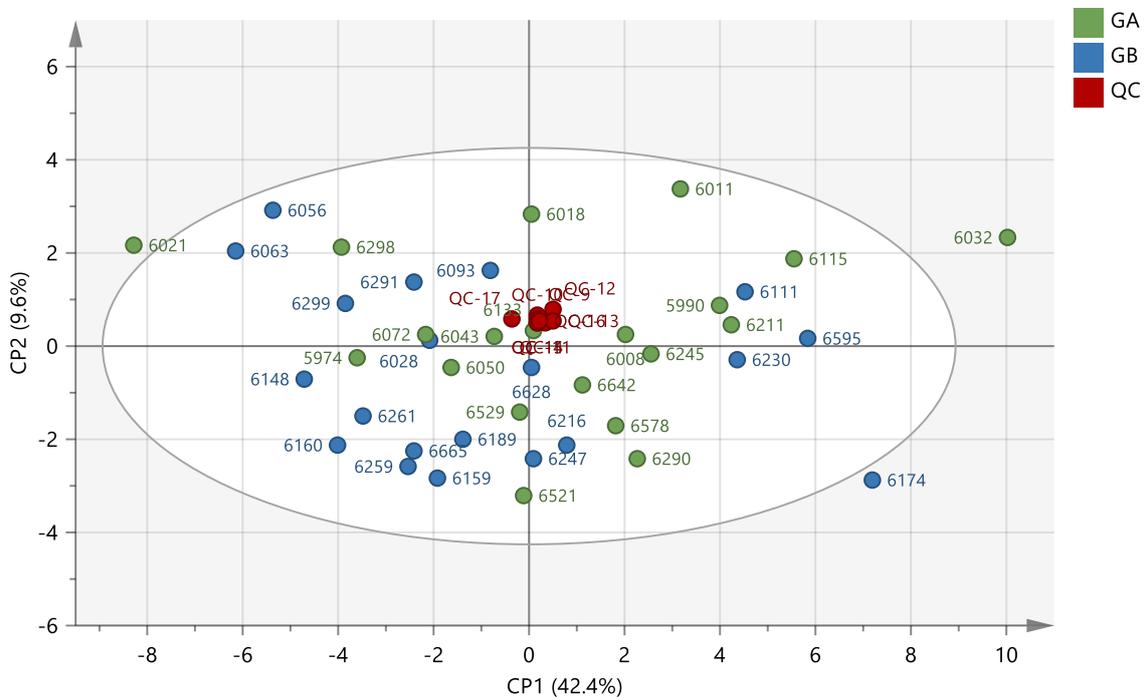


Figura 15. Gráfico de dispersión de los *scores* de la 1ª con la 2ª componente principal para los resultados de CE-MS. Elipse T^2 de Hotelling (95%). Cada punto corresponde a una muestra. CP: componente principal; entre paréntesis, la variabilidad explicada por cada componente; QC: control de calidad; GA, GB: línea de alta y baja grasa intramuscular respectivamente.

4.3 Análisis discriminantes

Para estos análisis, las muestras correspondientes a los controles de calidad no se incluyen en la matriz de datos, dado que no aportan información para el cálculo de la covarianza entre las variables de la matriz \mathbf{X} y el vector \mathbf{y} .

4.3.1 Proyección sobre estructuras latentes–análisis discriminante (PLS-DA)

Para los datos obtenidos de la LC-MS se utilizaron las tres primeras componentes del PLS (R^2_Y : 0.912, Q^2 : 0.786). En las figuras 16, 17 y 18 se muestran los gráficos de dispersión de los *scores*. Al igual que en el PCA, para la representación se construyó una elipse con 95% de nivel de confianza a partir del estadístico T^2 de Hotelling. En la figura 16, en plano formado por la 1ª y 2ª componente, se observa una clara diferenciación entre los grupos de alta y baja GIM, pero no se puede recoger la información en una única componente, ambos ejes están interviniendo. El plano formado por la 1ª y 3ª componente mostrado en la figura 17 tiene menor capacidad de separación de los grupos. Por último, en la figura 18, en el plano que conforman la 2ª y 3ª componente no es posible visualizar una distinción de líneas. Como ya se mencionó, las nuevas componentes que se agregan recogen información independiente a la de las componentes anteriores, por lo que es de esperar que cada vez abarquen menos variabilidad relacionada con \mathbf{y} , introduciendo más ruido en las gráficas.

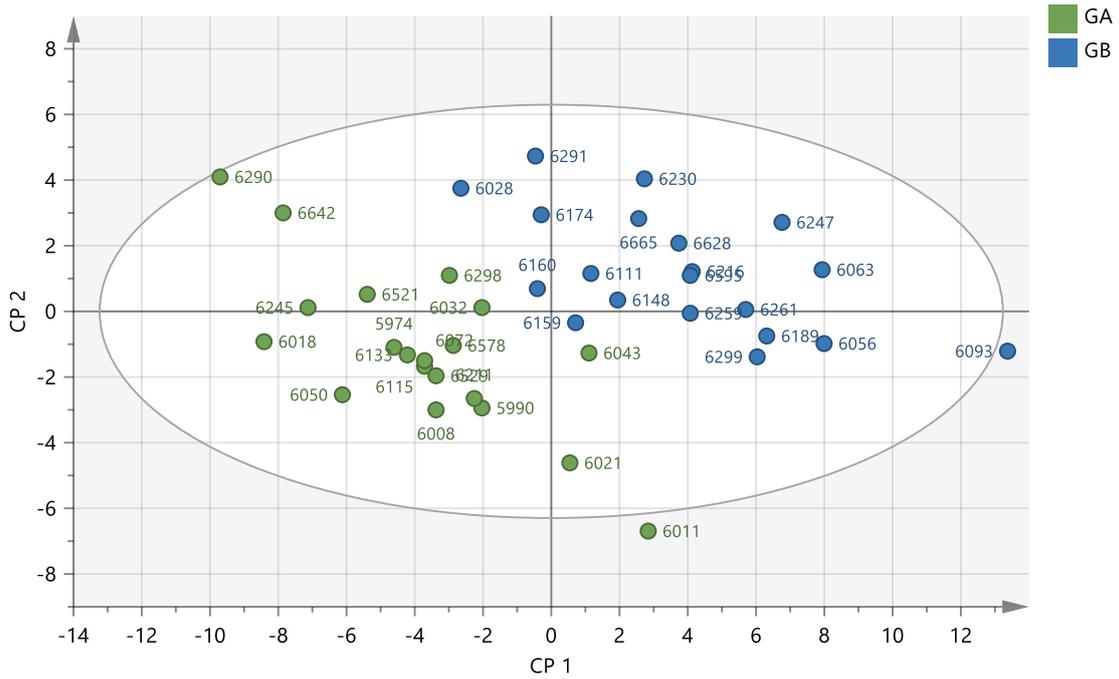


Figura 16. Gráfico de dispersión de los *scores* de la 1ª con la 2ª componente del PLS-DA para los resultados de LC-MS. Elipse T² de Hotelling (95%). Cada punto corresponde a un individuo. CP: componente principal; GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

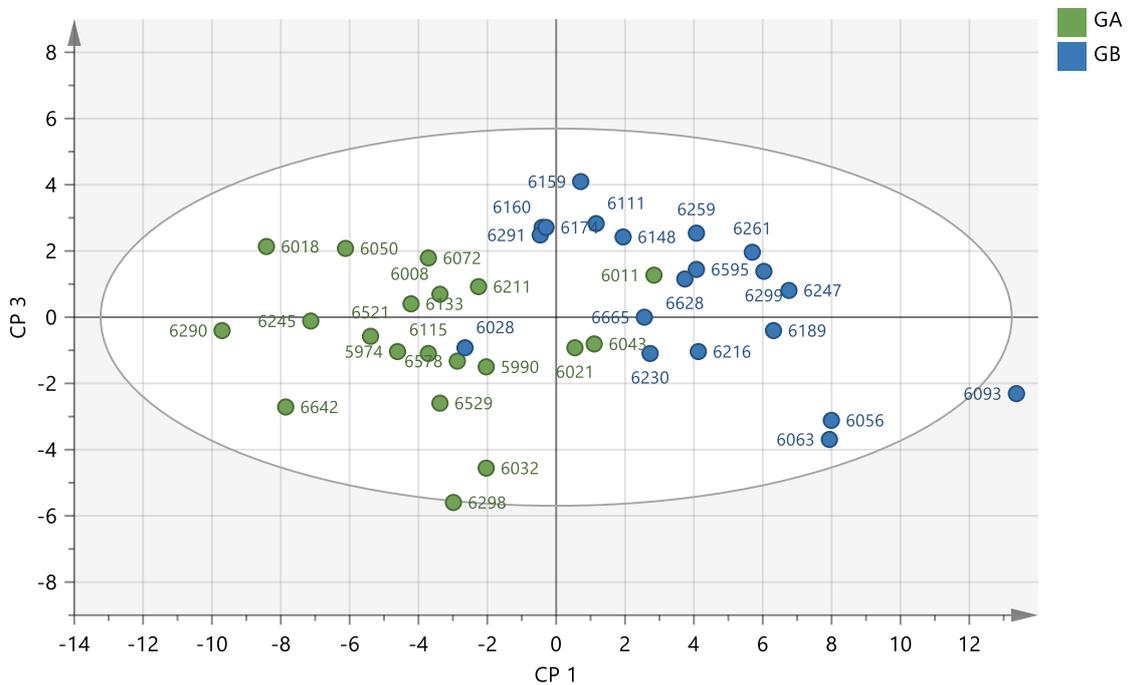


Figura 17. Gráfico de dispersión de los *scores* de la 1ª con la 3ª componente del PLS-DA para los resultados de LC-MS. Elipse T² de Hotelling (95%). Cada punto corresponde a un individuo. CP: componente principal; GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

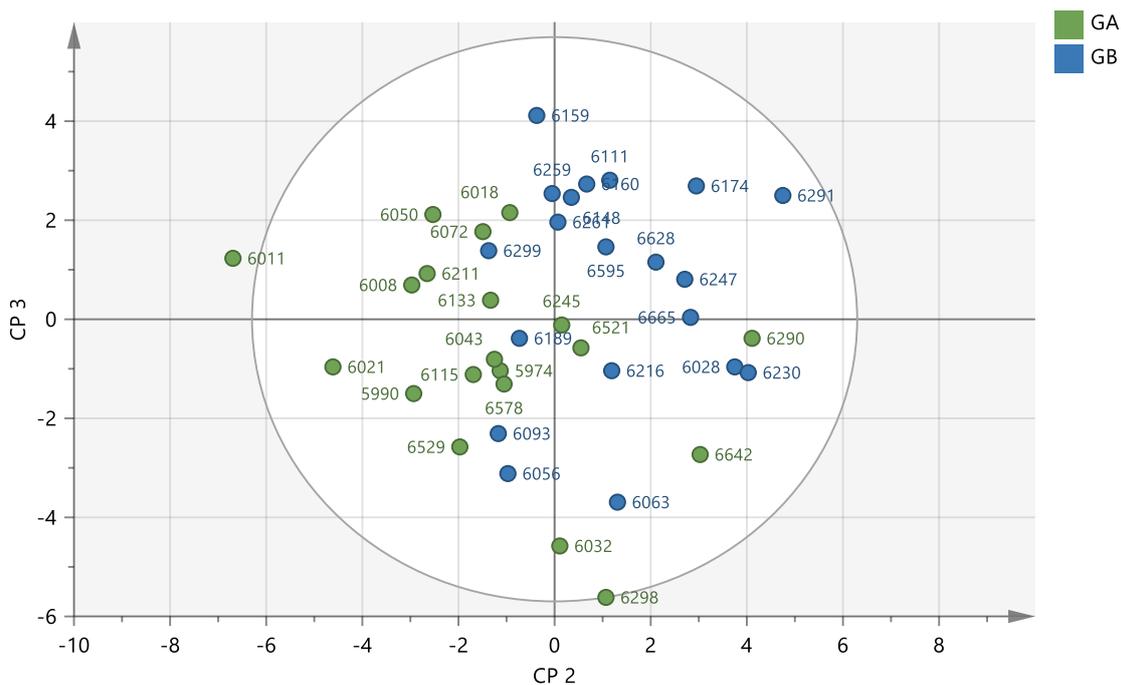


Figura 18. Gráfico de dispersión de los *scores* de la 2ª con la 3ª componente del PLS-DA para los resultados de LC-MS. Elipse T² de Hotelling (95%). Cada punto corresponde a un individuo. CP: componente principal; GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

En el análisis PLS-DA realizado para los resultados de la CE-MS se utilizaron las primeras tres componentes del PLS (R^2_Y : 0.839, Q^2 : 0.602). En las figuras 19, 20 y 21 se muestran los gráficos de dispersión de los *scores*, con su correspondiente T² de Hotelling al 95%. El gráfico de dispersión de la 1ª con la 2ª componente (figura 19) muestra la diferenciación entre grupos de alta y baja GIM, pero al igual que en los análisis realizados para los datos de LC-MS, no se puede recoger la información en una única componente, ambos ejes están interviniendo. En cambio, en los gráficos de dispersión de la 1ª con la 3ª componente (figura 20) y de la 2ª con la 3ª (figura 21) no se aprecia una clara separación entre los grupos.

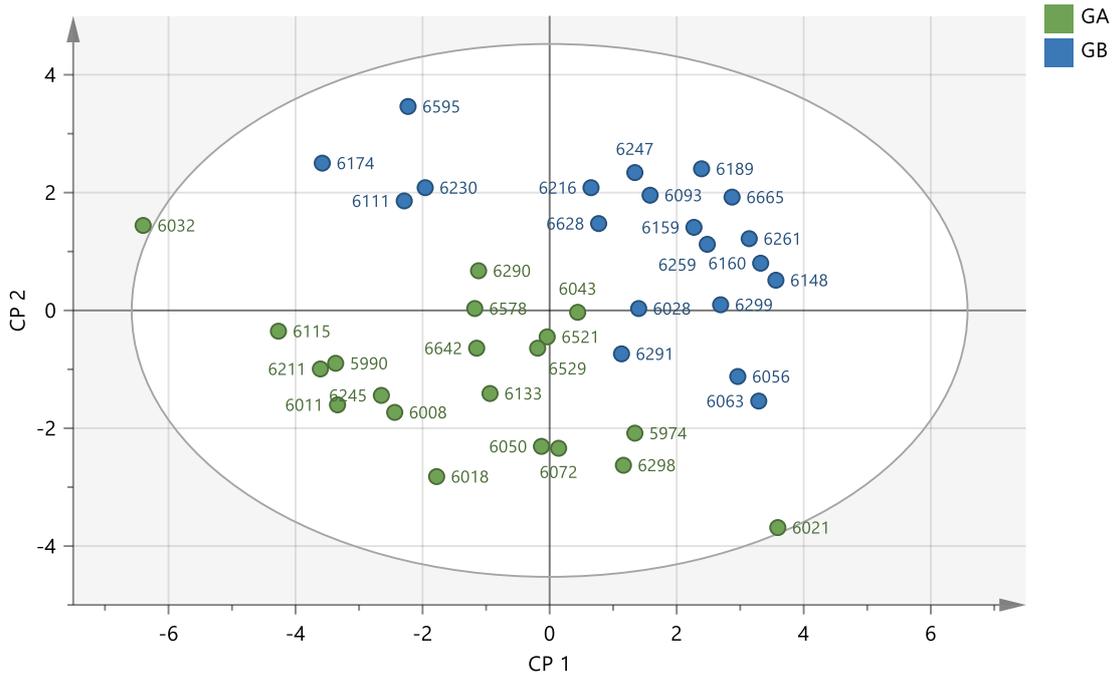


Figura 19. Gráfico de dispersión de los *scores* de la 1ª con la 2ª componente del PLS-DA para los resultados de CE-MS. Elipse T^2 de Hotelling (95%). Cada punto corresponde a un individuo. CP: componente principal; GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

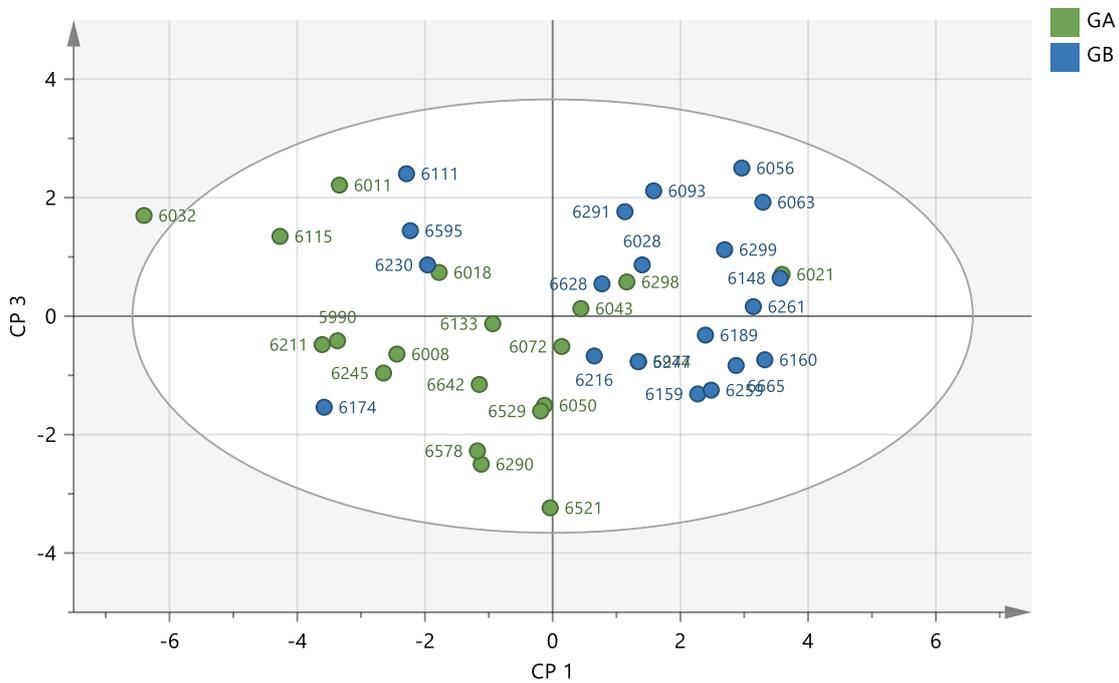


Figura 20. Gráfico de dispersión de los *scores* de la 1ª con la 3ª componente del PLS-DA para los resultados de CE-MS. Elipse T^2 de Hotelling (95%). Cada punto corresponde a un individuo. CP: componente principal; GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

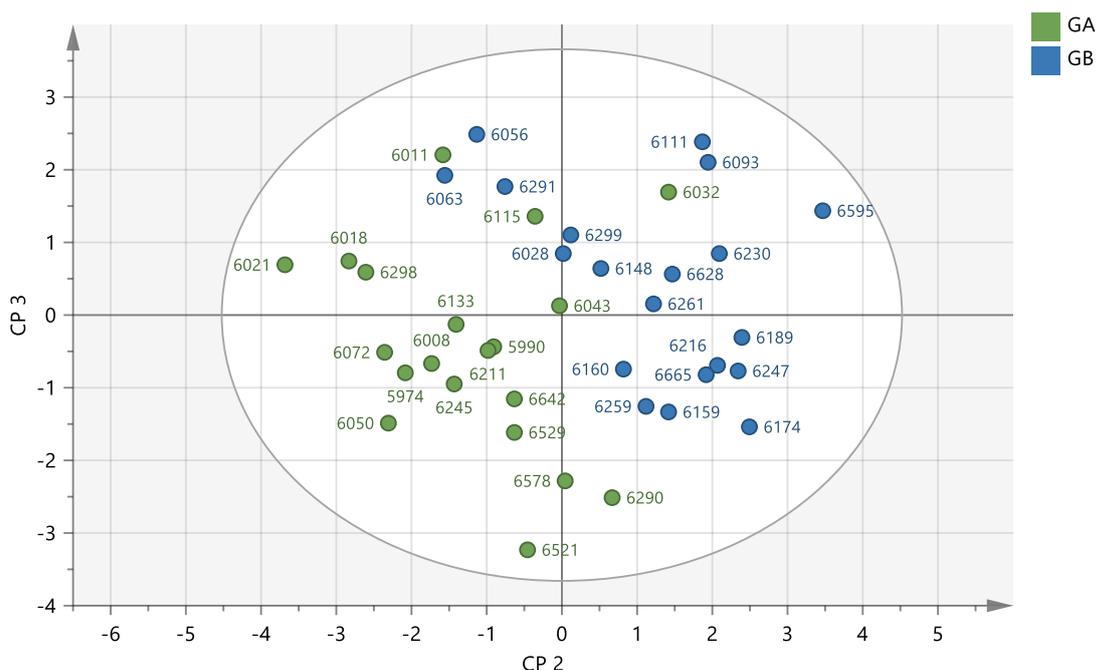


Figura 21. Gráfico de dispersión de los *scores* de la 2ª con la 3ª componente del PLS-DA para los resultados de CE-MS. Elipse T² de Hotelling (95%). Cada punto corresponde a un individuo. CP: componente principal; GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

4.3.2 Proyección ortogonal sobre estructuras latentes- análisis discriminante (OPLS-DA)

Tal como se mencionó en la introducción, el OPLS-DA sigue la misma metodología que el PLS-DA, por lo que tiene la misma capacidad predictiva, pero su interpretación es más sencilla; por este motivo, se conservó el mismo número de componentes que en el PLS-DA. Como ya se explicó, el OPLS-DA divide la matriz **X** en una parte correlacionada y otra parte no correlacionada a la variable **y**; de este modo la información correlacionada con la variable **y** se recoge toda en una componente (denominada predictiva), mientras que las siguientes componentes contendrán la información no correlacionada (denominadas ortogonales) (Trygg and Wold, 2002).

En la LC-MS, al igual que para el PLS-DA, se conservaron las primeras tres componentes (una componente predictiva y dos componentes ortogonales). En las figuras 22, 23 y 24 se muestran los gráficos de dispersión de los *scores*, con sus correspondientes elipses T² de Hotelling. En las figuras 22 y 23, que representan la componente predictiva frente a la 1ª y la 2ª componente ortogonal, respectivamente, se puede ver con claridad que toda la variabilidad correlacionada con la variable **y** se agrupa en la componente predictiva, mientras que las componentes ortogonales recogen la variabilidad no relacionada con la

clasificación. La figura 24 representa, como es de esperar, que el plano formado por la 1ª y 2ª componente ortogonal no es capaz de clasificar a los individuos por grupos.

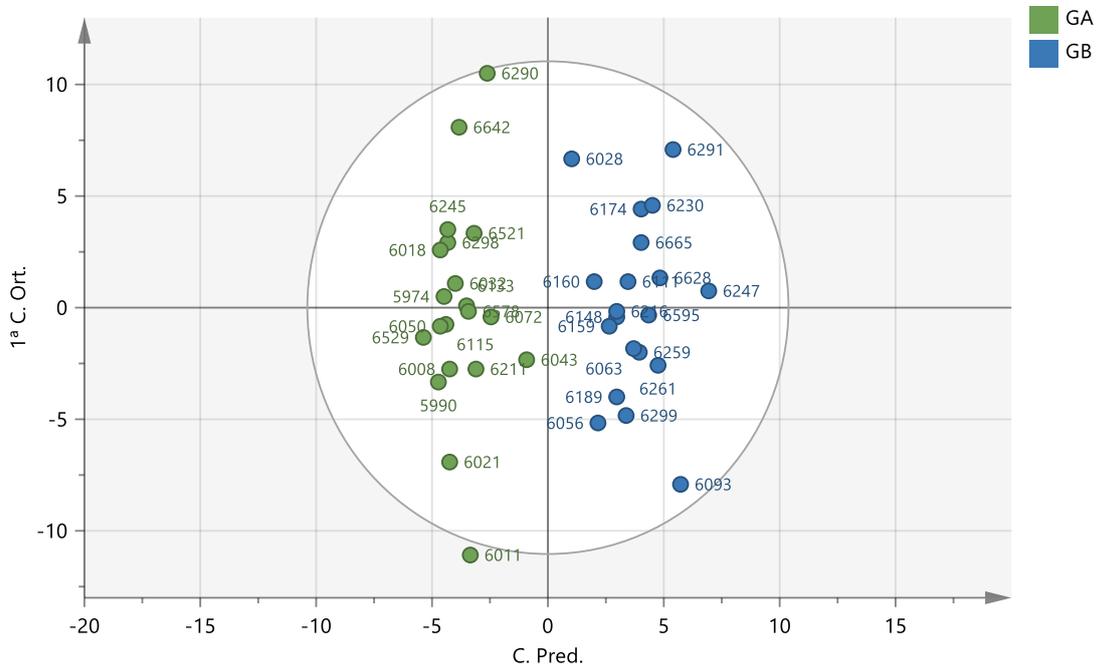


Figura 22. Gráfico de dispersión de los *scores* de la componente predictiva con la 1ª componente ortogonal del OPLS-DA para los resultados de LC-MS. Elipse T^2 de Hotelling (95%). Cada punto corresponde a un individuo. C. Pred.: componente predictiva; C. Ort.: componente ortogonal GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

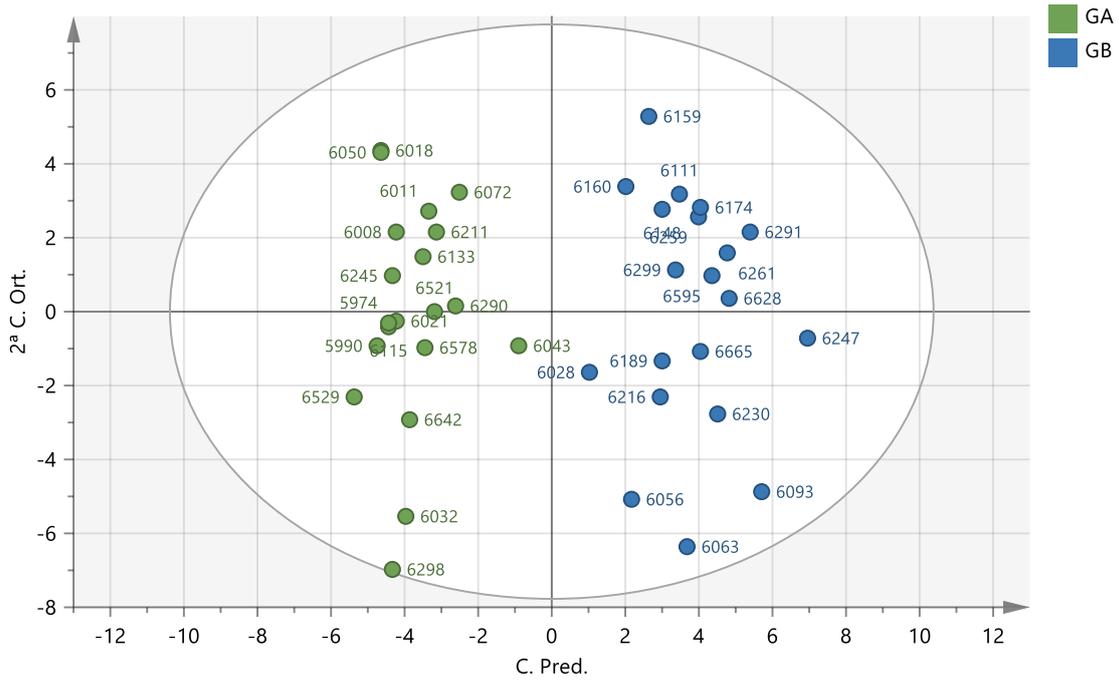


Figura 23. Gráfico de dispersión de los *scores* de la componente predictiva con la 2ª componente ortogonal del OPLS-DA para los resultados obtenidos de LC-MS. Elipse T² de Hotelling (95%). Cada punto corresponde a un individuo. C. Pred.: componente predictiva; C. Ort.: componente ortogonal GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

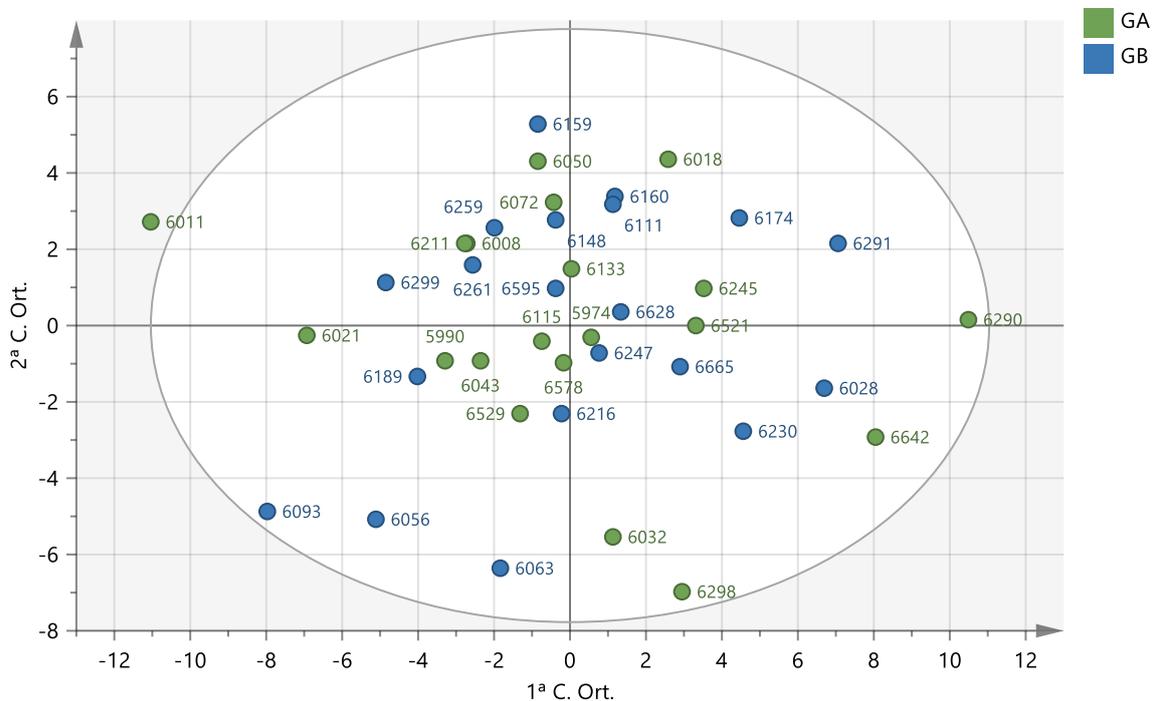


Figura 24. Gráfico de dispersión de los *scores* de la 1ª vs la 2ª componente ortogonal del OPLS-DA para los resultados obtenidos de LC-MS. Elipse T² de Hotelling (95%). Cada punto corresponde a un individuo. C. Pred.: componente predictiva; C. Ort.: componente ortogonal GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

Por el mismo motivo que se discutió anteriormente, en la CE-MS se conservaron las primeras 3 componentes (la componente predictiva y las dos primeras componentes ortogonales). En las figuras 25, 26 y 27 se muestran los gráficos de dispersión de los *scores*, con sus correspondientes elipses T^2 de Hotelling. Las figuras 25 y 26, que representan la componente predictiva frente a la 1ª y la 2ª componente ortogonal, respectivamente, exponen la separación entre los grupos dada por la componente predictiva. La forma alargada observada en la figura 25, distinto a lo observado hasta ahora, pone en evidencia que la variabilidad en \mathbf{X} correlacionada con la variable \mathbf{y} es mucho menor a la variabilidad no correlacionada. Esto concuerda con las conclusiones sacadas luego de analizar el PCA. La figura 27 representa, como es de esperar, que el plano formado por la 1ª y 2ª componente ortogonal no es capaz de clasificar a los individuos por grupos.

El OPLS-DA no sólo es más claro a la hora de interpretar los resultados gráficamente, sino que también permite analizar la información recogida en la componente predictiva de manera independiente a las componentes ortogonales. Es decir, permite separar la variabilidad responsable de la discriminación, del ruido que genera la variabilidad no correlacionada. Por este motivo, el OPLS-DA fue el método escogido para la selección de variables consideradas relevantes en la discriminación de grupos.

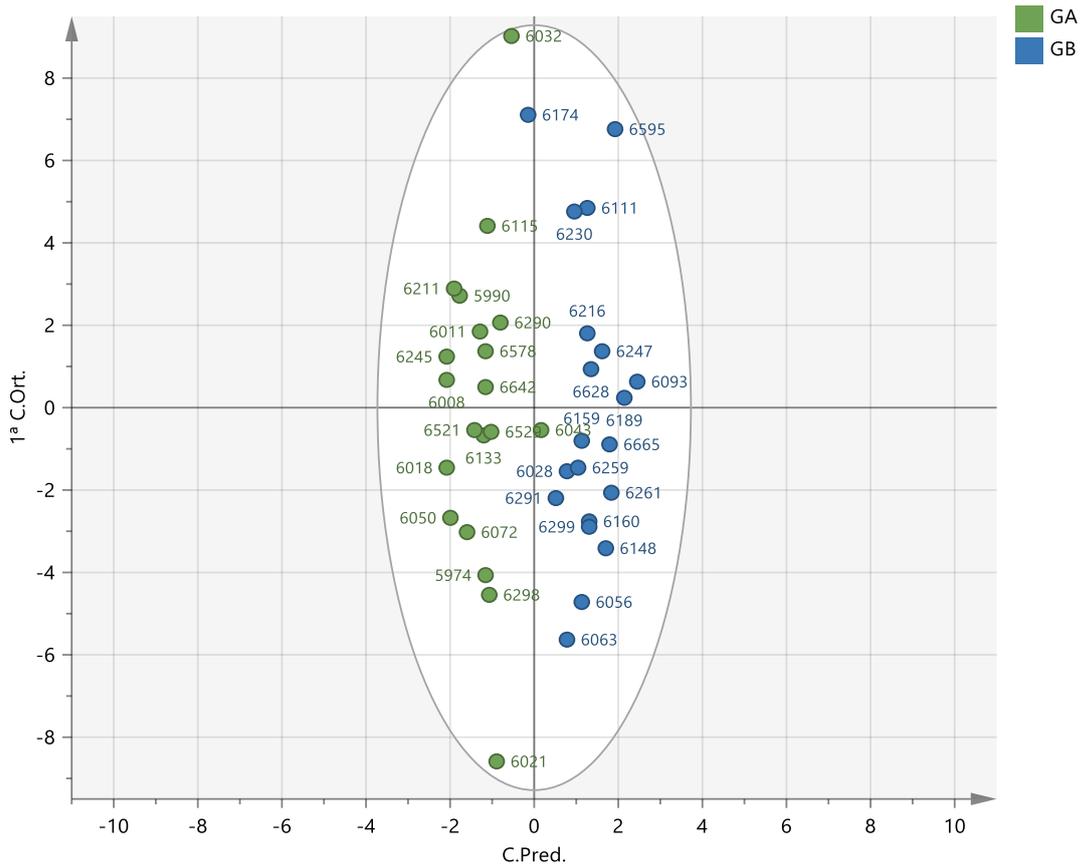


Figura 25. Gráfico de dispersión de los *scores* de la componente predictiva con la 1ª componente ortogonal del OPLS-DA para los resultados de CE-MS. Elipse T^2 de Hotelling (95%). Cada punto corresponde a un individuo. C. Pred.: componente predictiva; C. Ort.: componente ortogonal GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

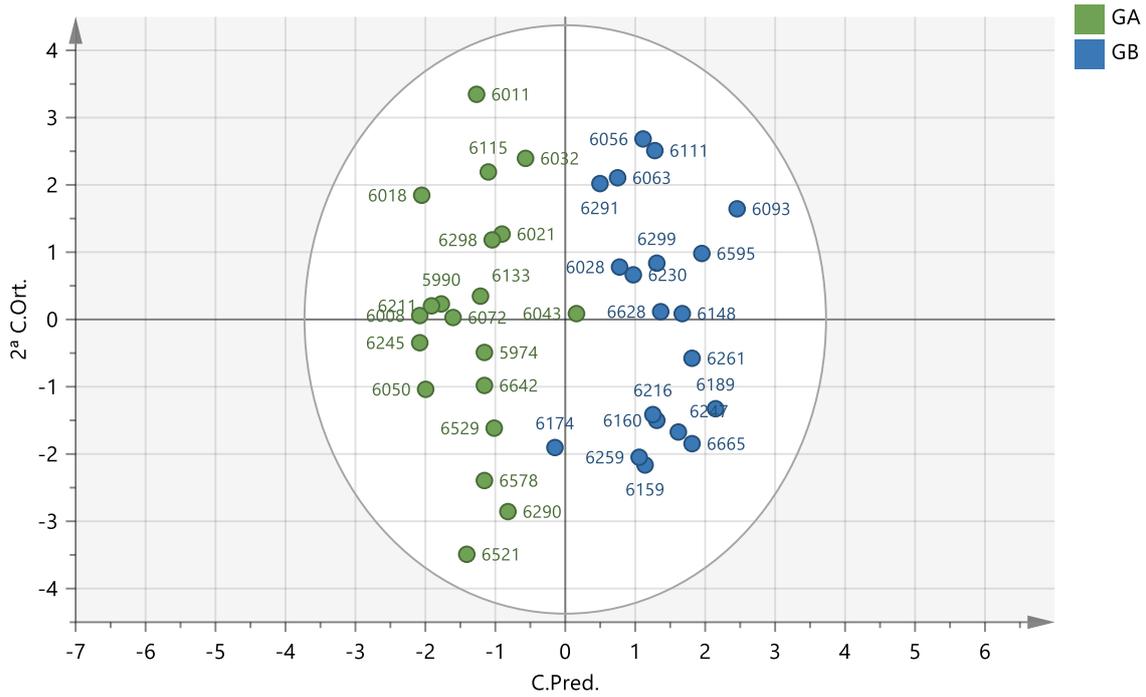


Figura 26. Gráfico de dispersión de los *scores* de la componente predictiva con la 2ª componente ortogonal del OPLS-DA para los resultados de CE-MS. Elipse T² de Hotelling (95%). Cada punto corresponde a un individuo. C. Pred.: componente predictiva; C. Ort.: componente ortogonal GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

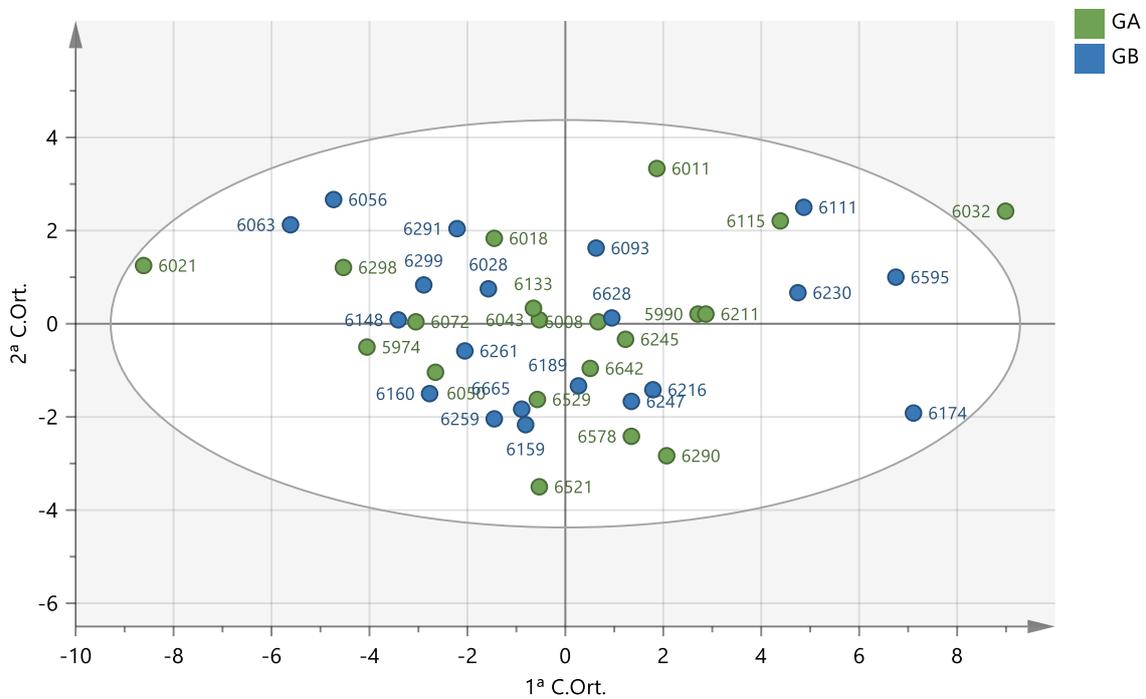


Figura 27. Gráfico de dispersión de los *scores* de la 1ª con la 2ª componente ortogonal del OPLS-DA para los resultados de CE-MS. Elipse T² de Hotelling (95%). Cada punto corresponde a un individuo. C. Pred.: componente predictiva; C. Ort.: componente ortogonal GA, GB: línea de alta y baja grasa intramuscular, respectivamente.

4.4 Selección e identificación de variables

Tras realizar los análisis multivariantes, es necesario seleccionar un número manejable de datos que me permitan orientar análisis posteriores. La elección de los criterios de selección puede variar según la cantidad de variables de las que se parte, la estructura de los datos o los análisis que se quieran realizar a continuación.

Como ya se mencionó, en la LC-MS, 456 variables fueron analizadas tras del control de calidad. A partir de los análisis multivariantes y univariantes, se seleccionaron 148 variables (tabla 2 de anexos). Los criterios de selección utilizados se describen a continuación: a partir del análisis multivariante se seleccionaron las variables que tuvieran un VIP (*variable importance in projection*) mayor a 1 (Chong and Jun, 2005), cuyo intervalo de confianza no incluyera el 0. El VIP es un parámetro que se utiliza para calcular la influencia de una variable en el modelo (Galindo-Prieto et al., 2014). Además, otra herramienta comúnmente utilizada para encontrar las variables que contribuyen a la discriminación entre grupos, teniendo en cuenta su relación señal/ruido, es el gráfico-S, que se muestra en la figura 28 (H.J. Kim et al., 2010). Después de algunas pruebas se decidió que las variables seleccionadas fueron aquellas cuyo $p(\text{corr})$ fue mayor a 0.3 o menor a -0.3. El gráfico-S es una forma de medir la fiabilidad en la estimación del peso de una variable sobre la componente predictiva. Al dividir el peso de cada variable (covarianza, representado en el eje X) por las desviaciones típicas (correlación, representada en el eje Y) se está realizando una tipificación según la magnitud de cada variable. Las zonas extremas son las que tienen menor riesgo de correlaciones al azar (Wiklund et al., 2008). El análisis univariante realizado fue el test-t, con corrección de Benjamini-Hochberg y se seleccionaron las variables con un $p\text{-valor} < 0.05$. La corrección de Benjamini-Hochberg es una corrección equivalente a la de Bonferroni, pero menos estricta. Por último, se consideraron solamente aquellas variables que tuvieran un porcentaje de cambio (FC%) entre los grupos del $\pm 15\%$, según la ecuación 3 descrita en Materiales y Métodos.

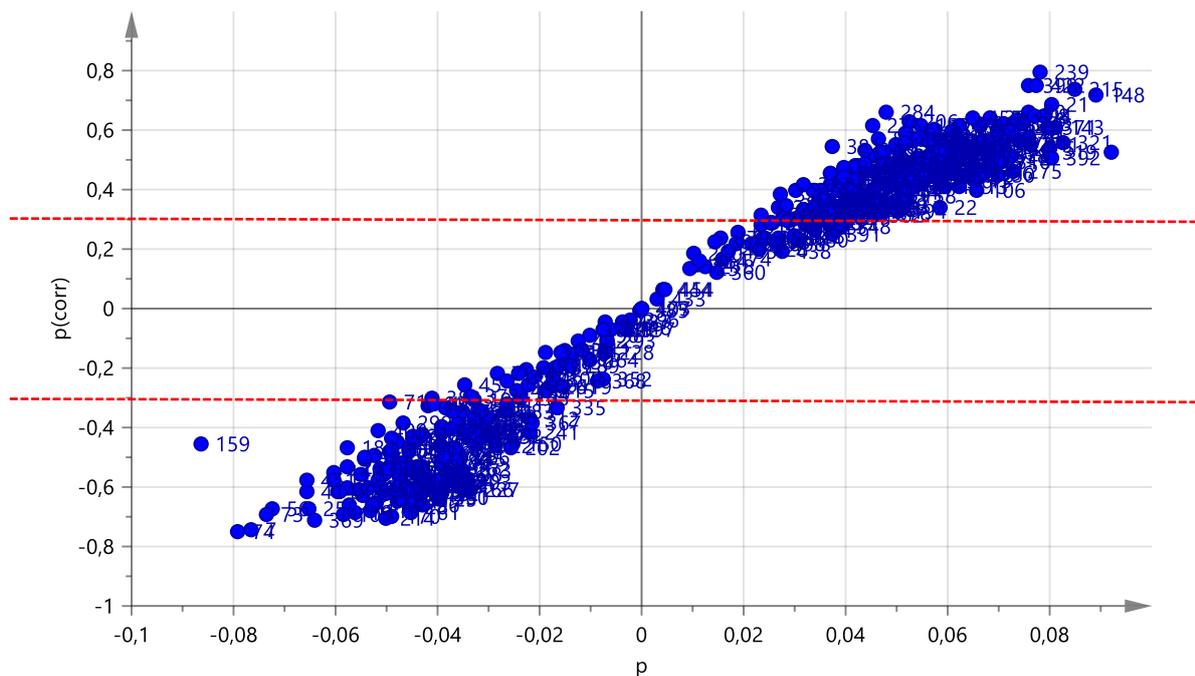


Figura 28. Gráfico-S del OPLS-DA para los resultados de LC-MS. Las líneas rojas delimitan el umbral a partir del cual las variables fueron preseleccionadas ($-0.3 > p(\text{corr}) > 0.3$). Cada punto se corresponde a una variable.

Los datos de masa experimental y tiempo de retención de las 149 variables seleccionadas fueron ingresados al software CEU Mass-Mediator para la identificación de metabolitos. Este software cruza la información de varias bases de datos, las utilizadas para este análisis fueron Human Metabolome Database (HMDB), LipidMaps, KEGG y Metlin. La identificación de estas variables (tabla 2 de anexos) son resultados aún preliminares. Posteriormente se seleccionaron aquellos metabolitos que podían referenciarse a una única fórmula molecular y no se tuvieron en cuenta para este trabajo aquellos a los que no se les pudo asignar una identificación. Tras este filtrado quedaron 16 metabolitos, que se exponen en la tabla 3.

Tabla 3. Selección de compuestos identificados mediante LC-MS diferencialmente expresados en las líneas seleccionadas por alta y baja GIM.

Masa experimental	FC(%)	Peso molecular	Identificación
552,3102	-30,25	553,3168	LysoPE(0:0/24:6(6Z,9Z,12Z,15Z,18Z,21Z))
		553,3168	LysoPE(24:6(6Z,9Z,12Z,15Z,18Z,21Z)/0:0)
973,4802	-31,27	938,4922	PIP(20:4(8Z,11Z,14Z,17Z)/16:0)
		938,4922	PIP(20:4(5Z,8Z,11Z,14Z)/16:0)
		938,4922	PIP(18:2(9Z,12Z)/18:2(9Z,12Z))
		938,4922	PIP(16:0/20:4(8Z,11Z,14Z,17Z))
		938,4922	PIP(16:0/20:4(5Z,8Z,11Z,14Z))
554,3436	-30,78	519,3689	PC(P-19:1(12Z)/0:0)
477,28	-31,45	496,3036	(3b,16b,20R)-Pregn-5-ene-3,16,20-triol 3-glucoside
		496,3036	Pregnanediol-3-glucuronide
		496,3036	3-alpha,20-alpha-dihydroxy-5-beta-pregnane 3-glucuronide
918,5777	-32,16	883,6091	PC(22:4(7Z,10Z,13Z,16Z)/22:5(7Z,10Z,13Z,16Z,19Z))
		883,6091	PC(22:5(4Z,7Z,10Z,13Z,16Z)/22:4(7Z,10Z,13Z,16Z))
		883,6091	PC(22:4(7Z,10Z,13Z,16Z)/22:5(4Z,7Z,10Z,13Z,16Z))
		883,6091	PC(22:5(7Z,10Z,13Z,16Z,19Z)/22:4(7Z,10Z,13Z,16Z))
662,3639	-33,19	663,3747	1-Palmitoyl-2-(5-keto-6-octenediyl)-sn-glycero-3-phosphocholine
486,3297	-34,15	505,3532	PC(P-18:1(9Z)/0:0)
		505,3532	CerP(d18:1/8:0)
		505,3532	C-8 Ceramide-1-phosphate
942,5399	-39,85	907,5575	PS(DiMe(11,3)/MonoMe(13,5))
850,6292	-40,97	851,6404	PE(44:4)
625,3458	-45,27	626,3431	PI(20:1(11Z)/0:0)
458,2641	-45,36	477,2855	PE(18:2(9Z,12Z)/0:0)
		477,2855	LysoPE(0:0/18:2(9Z,12Z))
258,1829	-46,63	259,1896	Lysyl-Leucine
		259,1896	Leucyl-Lysine
		259,1896	Isoleucyl-Lysine
		259,1896	Lysyl-Isoleucine
884,6801	-48,44	903,6928	PS(22:0/22:0)
568,3235	-47,83	533,3481	1-linoleoyl-sn-glycero-3-phosphocholine
		533,3481	LysoPE(0:0/22:2(13Z,16Z))
		533,3481	PE(22:2(13Z,16Z)/0:0)
390,2764	-47,56	355,3086	Pristanoylglycine
		355,3086	N-palmitoyl valine
		355,3086	N-stearoyl alanine
		355,3086	N-Hexadecyl-L-hydroxyproline

FC: porcentaje de cambio.

Por su parte, en la CE-MS se obtuvieron 208 variables en el análisis, de las cuales tras los análisis multivariantes y univariantes fueron seleccionadas 29 (tabla 3 en anexos). Los criterios de selección para el análisis multivariante incluyeron, al igual que para LC-MS, un VIP mayor a 1, cuyo intervalo de confianza no incluyera el 0 y un valor de p(corr) mayor a 0.3 o menor a -0.3 (figura 29). El análisis univariante realizado fue el test no paramétrico de Kruskal-Wallis, asumiendo un umbral para el p-valor de 0.05. De la misma manera que

Tabla 4. Selección de compuestos identificados mediante CE-MS diferencialmente expresados en las líneas seleccionadas por alta y baja GIM.

Masa experimental	FC(%)	Peso molecular	Identificación
328,2229	108,31	305,2355	Desconocido 1
158,1175	45,83	175,1195	Desconocido 2
205,0828	43,43	204,0746	Desconocido 3
217,1906	34,58	234,1984	Desconocido 4
306,085	30,82	283,0917	Desconocido 5
323,1945	30,62	322,1852	Desconocido 6
112,037	26,94	89,0477	Desconocido 7
134,0444	26,53	133,0375	Ácido aspártico
123,0557	25,42	122,048	Nicotinamida
126,0217	24,86	125,0147	Taurina
425,1405	21,36	424,1383	Desconocido 8
162,1124	20,08	161,1052	L-Carnitina
939,2723	19,40	916,2849	Desconocido 9
247,0149	16,93	246,0107	Desconocido 10
163,0605	15,82	162,0528	Desconocido 11
325,1132	15,82	342,1162	Desconocido 12
203,1496	-20,08	404,2927	Dimetil-L-arginina
205,0828	-27,55	204,0746	Desconocido 13
151,0711	-28,37	128,0837	Desconocido 14

Los metabolitos resaltados son aquellos con identificación confirmada. FC%: porcentaje de cambio.

4.5 Metabolitos diferencialmente expresados entre las líneas de alta y baja GIM

Se utilizó el software MetaboAnalyst 4.0 (Chong et al., 2018) para identificar las rutas metabólicas más relevantes de acuerdo a los metabolitos diferencialmente expresados en las líneas de alta y baja GIM. Para obtener información sobre los metabolitos y las rutas metabólicas implicadas se utilizaron las bases de datos KEGG (*Kyoto encyclopedia of genes and genomes*, Kanehisa et al., 2012), SMPDB (*Small molecule pathway database*, Jewison et al., 2014) y HMDB (*Human metabolome database*, Wishart et al., 2018)

En la tabla 3 se muestran los metabolitos diferencialmente expresados entre líneas obtenidos mediante LC-MS. Estos metabolitos están implicados fundamentalmente en la ruta del metabolismo de los glicerofosfolípidos. En concreto, los metabolitos identificados fueron fosfocolinas (PC), fosfoetanolaminas (PE) y lisofosfocolinas (LysoPC), los cuales se encontraban todos sobreexpresados en la línea de baja grasa intramuscular. Los glicerofosfolípidos son un tipo de fosfolípidos que tienen una función estructural en la célula, ya que forman parte de la membrana lipídica. Los fosfolípidos tienen mayor

proporción de PUFA, comparados con los triglicéridos (De Smet et al., 2004). Estos resultados concuerdan con la respuesta correlacionada a la selección por GIM en la composición de ácidos grasos obtenida previamente en estas líneas de conejos por Martínez-Álvarez et al., 2017. En este estudio, se observó una correlación genética alta y positiva entre la GIM y los MUFA (entre 0.88 y 0.97) y alta y negativa entre la GIM y los PUFA (entre -0.83 y -0.91). Esto tuvo como consecuencia un mayor porcentaje de MUFA y menor porcentaje de PUFA en la línea de alta GIM con respecto a la de baja GIM (Martínez-Álvarez et al., 2017), debido a la mayor proporción de triglicéridos respecto a los fosfolípidos (localizados en la membrana plasmática) (De Smet et al., 2004). Otros autores han documentado el incremento diferencial de los distintos tipos de ácidos grasos al aumentar el contenido de grasa (De Smet et al., 2004; Wood et al., 2008). Por otra parte, resultados preliminares de los análisis genómicos realizados por Sosa Madrid et al., 2018 en la generación 9 de estas líneas detectaron regiones candidatas donde se encontraron genes relacionados con el transporte de los fosfolípidos en la membrana, entre ellos el *Stard10* (específicamente para el transporte de fosfatidilcolina y fosfatidiletanolamina, Olayioye et al., 2005), *Apold1* (Rosell et al., 2014) y *Arap1* (Moisá et al., 2015). También se encontró el gen *Plbd1*, que codifica para una proteína con débil actividad fosfolipasa para varios fosfolípidos, entre ellos la fosfatidilcolina, fosfatidilinositol, fosfatidiletanolamina y lisofosfolípidos (Zhou et al., 2016). Es necesario seguir investigando para tratar de dilucidar la relación entre los metabolitos detectados en este estudio y los genes anteriormente mencionados.

Los metabolitos diferencialmente expresados en las líneas de alta y baja GIM detectados mediante la técnica de CE-MS se muestran en la tabla 4. Hay varios metabolitos que han sido asignados como desconocidos; esto se debe a que son compuestos cuya identificación no ha podido ser confirmada por el momento. Dentro de estos metabolitos, destaca el desconocido 1 cuya expresión es más de 100 veces superior en la línea de alta GIM que en la de baja. De confirmarse su importancia, este metabolito podría ser utilizado como un biomarcador, aún sin conocer su identidad; aunque si este es el caso, se podrían realizar otros análisis para intentar identificarlo.

Se encontró mayor contenido de ácido aspártico, nicotinamida, taurina y L-carnitina en la línea de alta GIM.

La nicotinamida es precursor de la nicotinamida adenina dinucleótido fosfato (NADP⁺), una coenzima indispensable para el funcionamiento de muchas enzimas, entre las que se encuentra la enzima lipogénica glucosa 6-fostato deshidrogenasa. La glucosa 6-fostato deshidrogenasa (G6PDH) está involucrada en la ruta metabólica de la pentosa fosfato. Cataliza la oxidación de la glucosa 6-fostato a 6-fosfogluconato, produciendo NADPH a partir del NADP⁺. En tejidos con alta actividad de síntesis de ácidos grasos o de colesterol y hormonas esteroides, la ruta de la pentosa fosfato es indispensable para proveer el NADPH que requieren esas rutas (Nelson and Cox, 2013). Estudios previos en la actividad de enzimas lipogénicos de las líneas de alta y baja GIM, mostraron una mayor actividad de la G6PDH en el hígado (Martinez-Alvaro et al., 2018), en el músculo y la grasa perirrenal (Martinez-Alvaro et al., 2017). La mayor actividad de la G6PDH en la línea de alta GIM podría estar relacionada con la sobreexpresión de la nicotinamida. El ácido aspártico también se encontró en mayor cantidad en la línea de alta GIM. Este metabolito se encuentra implicado en muchas rutas metabólicas, entre las que también se encuentra el metabolismo de la nicotinamida.

La L-carnitina se encuentra en mayor cantidad en la línea de alta GIM. La L-carnitina es una amina cuaternaria sintetizada en el hígado y los riñones a partir de dos aminoácidos esenciales, la lisina y la metionina. La L-carnitina está directamente relacionado con la degradación de los ácidos grasos ya que cumple un rol importante en el transporte de los ácidos grasos de cadena larga al interior de la mitocondria para la β oxidación. De acuerdo a la bibliografía, no hay una relación clara entre el nivel de L-carnitina en el plasma y la deposición lipídica. En bovino, un análisis realizado para estudiar las vías metabólicas afectadas por dos mutaciones con efecto divergente en el crecimiento y la deposición de lípidos (Weikard et al., 2010) demostró que los niveles de L-carnitina en plasma se veían disminuidos en animales que tenían mayor crecimiento corporal, pero menor cantidad de grasa (mutación *Q204X* del gen *GDF8*). Sin embargo, un estudio realizado en humanos donde se evaluaron las diferencias metabólicas entre hombres obesos y de peso normal (J.Y. Kim et al., 2010), detectó niveles más bajos de L-carnitina en plasma en la población obesa. La taurina es una amina biogénica que se encuentra en mayor cantidad en la línea de alta GIM. La taurina está involucrada en numerosas rutas metabólicas, entre ellas la biosíntesis de los ácidos biliares primarios. Los ácidos biliares son un grupo de compuestos derivados de esteroides que actúan como detergente en el intestino para facilitar la digestión y absorción de las grasas. Los ácidos biliares se conjugan con taurina o glicina para que puedan

interactuar tanto con compuestos hidrofóbicos como hidrofílicos (Agellon, 2002). Como ya se mencionó, los resultados genómicos de Sosa Madrid et al., 2018 identificaron como gen de interés asociado al carácter GIM el gen *Stard10*. Un estudio en ratones knockout para dicho gen demostró que en estos ratones la secreción de ácidos biliares del hígado y los niveles de ácidos biliares conjugados con taurina en la bilis eran mayores que en los ratones *wild type* (Ito et al., 2013). En este sentido, tanto los resultados metabolómicos como los genómicos señalan esta ruta como potencial involucrada en la regulación del fenotipo de interés.

Por otra parte, la dimetil-L-arginina se encontró más expresada en la línea de baja GIM. Este metabolito interfiere en la producción del óxido nítrico. En humanos, la ruta dimetil-L-arginina/óxido nítrico ha sido muy estudiada debido a su relación con el desarrollo de varias enfermedades, entre ellas la hipertensión (Surdacki et al., 1999) y enfermedades del riñón (Fleck et al., 2003; Wang et al., 2008). Sin embargo, no se ha encontrado aún una relación clara entre la expresión diferencial de este metabolito y la expresión del carácter GIM en las líneas de selección.

Es necesario seguir investigando la interacción de los principales metabolitos implicados a través de una integración multiómica para entender cómo actúan en el fenotipo expresado. Estudios posteriores incluirán análisis metabolómicos con la plataforma de espectrometría de masas acoplada a cromatografía gaseosa y espectrometría de masas acoplada a cromatografía líquida en polaridad positiva. Además, la integración multiómica entre los resultados genómicos y metabolómicos podrán clarificar los mecanismos implicados en la deposición de la grasa intramuscular

5. CONCLUSIONES

Los métodos de análisis multivariantes utilizados han permitido abordar el análisis del gran conjunto de datos obtenido en un estudio metabolómico no dirigido de las líneas de conejos de alta y baja GIM. La reducción de la dimensionalidad ha permitido ver mejor la estructura de los datos con mínima pérdida de información, facilitando la interpretación de los resultados.

Los resultados preliminares del análisis metabolómico han permitido encontrar diversos metabolitos diferencialmente expresados entre las líneas. Mediante LC-MS se han encontrado en mayor cantidad en la línea de baja grasa intramuscular, diversos metabolitos implicados en la ruta del metabolismo de los glicerofosfolípidos: fosfocolinas (PC), fosfoetanolaminas (PE) y lisofosfocolinas (LysoPC). Mediante CE-MS se encontró mayor contenido de ácido aspártico, nicotinamida, taurina y L-carnitina en la línea de alta GIM y menor de Dimetil-L-arginina.

Se han detectado algunas rutas metabólicas potencialmente implicadas en la diferenciación entre líneas del carácter grasa intramuscular. La ruta del metabolismo de los glicerofosfolípidos, la ruta de la pentosa 5-fosfato, la biosíntesis de ácidos biliares y la β oxidación, han sido las rutas más destacadas en la deposición diferencial de GIM entre las líneas.

6. REFERENCIAS.

Abdi, H., and L. J. Williams. 2010. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* 2:433–459. doi:10.1002/wics.101.

Agellon, L. 2002. Metabolism and function of bile acids. *Biochem. lipids, lipoproteins Membr.* 432–448.

Akiyama, K., E. Chikayama, H. Yuasa, Y. Shimada, T. Tohge, K. Shinozaki, M. Y. Hirai, T. Sakurai, J. Kikuchi, and K. Saito. 2008. PRIME: a Web site that assembles tools for metabolomics and transcriptomics. *In Silico Biol.* 8:339–345.

Alonso, A., S. Marsal, and A. Juliá. 2015. Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. *Front. Bioeng. Biotechnol.* 3:1–20. doi:10.3389/fbioe.2015.00023.

Artati, A., C. Prehn, G. Möller, and J. Adamski. 2012. Assay Tools for Metabolomics. In: K. Suhre, editor. *Genetics meets metabolomics. From experiments to system biology.* Springer. p. 13–38.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57:289–300. doi:10.2307/2346101.

van den Berg, R. A., H. C. J. Hoefsloot, J. A. Westerhuis, A. K. Smilde, and M. J. van der Werf. 2006. Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics.* 7:1–15. doi:10.1186/1471-2164-7-142.

Bingol, K., D.-W. Li, L. Bruschiweiler-Li, O. A. Cabrera, T. Megraw, F. Zhang, and R. Brüschiweiler. 2015. Unified and Isomer-Specific NMR Metabolomics Database for the Accurate Analysis of ¹³C–¹H HSQC Spectra. *ACS Chem. Biol.* 10:452–459. doi:10.1021/cb5006382.

Blasco, A., M. Martínez-Álvarez, M. L. García, N. Ibáñez-Escriche, and M. J. Argente. 2017. Selection for environmental variance of litter size in rabbits. *Genet. Sel. Evol.* 49:1–8. doi:10.1186/s12711-017-0323-4.

Bylesjo, M., M. Rantalainen, O. Cloarec, J. K. Nicholson, E. Holmes, and J. Trygg. 2006. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J. Chemom.* 20:341–351. doi:10.1002/cem.1006.

Canuto, G. A. B., J. L. Da Costa, P. L. R. Da Cruz, A. R. L. De Souza, A. T. Faccio, A. Klassen, K. T. Rodrigues, and M. F. M. Tavares. 2018. Metabolômica: Definições, Estado-Da-Arte E Aplicações Representativas. *Quim. Nova.* 41:75–91. doi:10.21577/0100-4042.20170134.

Chapinal, N., M. E. Carson, S. J. LeBlanc, K. E. Leslie, S. Godden, M. Capel, J. E. P. Santos, M. W. Overton, and T. F. Duffield. 2012. The association of serum metabolites in the transition period with milk production and early-lactation reproductive performance. *J. Dairy Sci.* 95:1301–1309. doi:10.3168/jds.2011-4724.

Chong, I. G., and C. H. Jun. 2005. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* 78:103–112. doi:10.1016/j.chemolab.2004.12.011.

Chong, J., O. Soufan, C. Li, I. Caraus, S. Li, G. Bourque, D. S. Wishart, and J. Xia. 2018. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* 46:W486–W494. doi:10.1093/nar/gky310.

Ciborowski, M., F. Javier Rupérez, M. P. Martínez-Alcázar, S. Angulo, P. Radziwon, R. Olszanski, J. Kloczko, and C. Barbas. 2010. Metabolomic approach with LC-MS reveals significant effect of pressure on diver's plasma. *J. Proteome Res.* 9:4131–4137. doi:10.1021/pr100331j.

Ciobanu, D. C., S. M. Lonergan, and E. J. Huff-Lonergan. 2011. Genetics of Meat Quality and Carcass Traits. In: Rothschild, MF and Ruvinsky, A, editor. *GENETICS OF THE PIG, 2ND EDITION.* CABI PUBLISHING-C A B INT, CABI PUBLISHING, WALLINGFORD OX10 8DE, OXON, ENGLAND. p. 355–389.

Cui, Q., I. A. Lewis, A. D. Hegeman, M. E. Anderson, J. Li, C. F. Schulte, W. M. Westler, H. R. Eghbalnia, M. R. Sussman, and J. L. Markley. 2008. Metabolite identification via the Madison Metabolomics Consortium Database. *Nat. Biotechnol.* 26:162. Available from: <http://dx.doi.org/10.1038/nbt0208-162>

Curso “Análisis, monitorización y diagnóstico de procesos multivariantes”. Ferrer Riquelme, Alberto J.

Curso “Multivariate Data Analysis for Omics”, Umetrics.

Dunn, W. B., A. Erban, R. J. M. Weber, D. J. Creek, M. Brown, R. Breitling, T. Hankemeier, R. Goodacre, S. Neumann, J. Kopka, and M. R. Viant. 2013. Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics. *Metabolomics*. 9:44–66. doi:10.1007/s11306-012-0434-4.

Fiehn, O. 2002. Metabolomics--the link between genotypes and phenotypes. *Plant Mol. Biol.* 48:155–71. doi:082/11.

Fleck, C., F. Schweitzer, E. Karge, M. Busch, and G. Stein. 2003. Serum concentrations of asymmetric (ADMA) and symmetric (SDMA) dimethylarginine in patients with chronic kidney diseases. *Clin. Chim. Acta.* 336:1–12.

Fontanesi, L. 2016. Metabolomics and livestock genomics: Insights into a phenotyping frontier and its applications in animal breeding. *Anim. Front.* 6:73. doi:10.2527/af.2016-0011.

Fuhrer, T., and N. Zamboni. 2015. High-throughput discovery metabolomics. *Curr. Opin. Biotechnol.* 31:73–78. doi:10.1016/j.copbio.2014.08.006.

Galindo-Prieto, B., L. Eriksson, and J. Trygg. 2014. Variable influence on projection (VIP) for orthogonal projections to latent structures (OPLS). *J. Chemom.* 28:623–632. doi:10.1002/cem.2627.

García, A., J. Godzien, Á. López-González, and C. Barbas. 2017. Capillary electrophoresis mass spectrometry as a tool for untargeted metabolomics. *Bioanalysis*. 9:99–130. doi:10.4155/bio-2016-0216.

Gil de la Fuente, A., J. Godzien, M. Fernández López, F. J. Rupérez, C. Barbas, and A. Otero. 2018. Knowledge-based metabolite annotation tool: CEU Mass Mediator. *J. Pharm. Biomed. Anal.* 154:138–149. doi:10.1016/j.jpba.2018.02.046.

Goldansaz, S. A., A. C. Guo, T. Sajed, M. A. Steele, G. S. Plastow, and D. S. Wishart. 2017. Livestock metabolomics and the livestock metabolome: A systematic review. *PLoS One*. 12:1–26. doi:10.1371/journal.pone.0177675.

Goodacre, R. 2005. Metabolomics - The way forward. *Metabolomics*. 1:1–2. doi:10.1007/s11306-005-1111-7.

Gromski, P. S., H. Muhamadali, D. I. Ellis, Y. Xu, E. Correa, M. L. Turner, and R.

Goodacre. 2015. A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. *Anal. Chim. Acta.* 879:10–23. doi:10.1016/j.aca.2015.02.012.

Hailemariam, D., R. Mandal, F. Saleem, S. M. Dunn, D. S. Wishart, and B. N. Ametaj. 2014. Identification of predictive biomarkers of disease state in transition dairy cows. *J. Dairy Sci.* 97:2680–2693. doi:10.3168/jds.2013-6803.

Hernández, P., M. Pla, and A. Blasco. 1998. Carcass characteristics and meat quality of rabbit lines selected for different objectives: II. Relationships between meat characteristics. *Livest. Prod. Sci.* 54:125–131. doi:10.1016/S0301-6226(97)00178-4.

Hirayama, A., M. Wakayama, and T. Soga. 2014. Metabolome analysis based on capillary electrophoresis-mass spectrometry. *TrAC - Trends Anal. Chem.* 61:215–222. doi:10.1016/j.trac.2014.05.005.

Horai, H., M. Arita, S. Kanaya, Y. Nihei, T. Ikeda, K. Suwa, Y. Ojima, K. Tanaka, S. Tanaka, K. Aoshima, Y. Oda, Y. Kakazu, M. Kusano, T. Tohge, F. Matsuda, Y. Sawada, M. Y. Hirai, H. Nakanishi, K. Ikeda, N. Akimoto, T. Maoka, H. Takahashi, T. Ara, N. Sakurai, H. Suzuki, D. Shibata, S. Neumann, T. Iida, K. Tanaka, K. Funatsu, F. Matsuura, T. Soga, R. Taguchi, K. Saito, and T. Nishioka. 2010. MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* 45:703–714. doi:10.1002/jms.1777.

<https://www.agilent.com/>. Agilent Technologies.

Hummel, J., J. Selbig, D. Walther, and J. Kopka. 2007. The Golm Metabolome Database: a database for GC-MS based metabolite profiling. In: J. Nielsen and M. C. Jewett, editors. *Metabolomics: A Powerful Tool in Systems Biology*. Springer Berlin Heidelberg, Berlin, Heidelberg. p. 75–95.

Ito, M., Y. Yamanashi, Y. Toyoda, H. Izumi-Nakaseko, S. Oda, A. Sugiyama, M. Kuroda, H. Suzuki, T. Takada, and S. Adachi-Akahane. 2013. Disruption of Stard10 gene alters the PPAR α -mediated bile acid homeostasis. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids.* 1831:459–468. doi:10.1016/j.bbalip.2012.11.008.

Jewison, T., Y. Su, F. M. Disfany, Y. Liang, C. Knox, A. Maciejewski, J. Poelzer, J. Huynh, Y. Zhou, D. Arndt, Y. Djoumbou, Y. Liu, L. Deng, A. C. Guo, B. Han, A. Pon, M. Wilson,

S. Rafatnia, P. Liu, and D. S. Wishart. 2014. SMPDB 2.0: big improvements to the Small Molecule Pathway Database. *Nucleic Acids Res.* 42:D478-84. doi:10.1093/nar/gkt1067.

Kanehisa, M., S. Goto, Y. Sato, M. Furumichi, and M. Tanabe. 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40:D109-14. doi:10.1093/nar/gkr988.

Karaman, I. 2017. Preprocessing and Pretreatment of Metabolomics Data for Statistical Analysis. In: *Metabolomics: From Fundamentals to Clinical Applications*. United States. p. 145–161.

Karisa, B. K., J. Thomson, Z. Wang, C. Li, Y. R. Montanholi, S. P. Miller, S. S. Moore, and G. S. Plastow. 2014. Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livest. Sci.* 165:200–211. doi:10.1016/j.livsci.2014.03.002.

Kim, H. J., J. H. Kim, S. Noh, H. J. Hur, M. J. Sung, J. T. Hwang, J. H. Park, H. J. Yang, M. S. Kim, D. Y. Kwon, and S. H. Yoon. 2010. Metabolomic Analysis of Livers and Serum from High Fat Diet Induced Obese Mice. *J. Proteome Res.* 10:722–731. doi:10.1021/pr100892r.

Kim, J. Y., J. Y. Park, O. Y. Kim, B. M. Ham, H.-J. Kim, D. Y. Kwon, Y. Jang, and J. H. Lee. 2010. Metabolic profiling of plasma in overweight/obese and lean men using ultra performance liquid chromatography and Q-TOF mass spectrometry (UPLC– Q-TOF MS). *J. Proteome Res.* 9:4368–4375. doi:10.1021/pr100101p.

Krastanov, A. 2010. Metabolomics - The state of art. *Biotechnol. Biotechnol. Equip.* 24:1537–1543. doi:10.2478/V10133-010-0001-y.

Kühn, C., R. Weikard, and P. Widmann. 2014. Metabolomics: a pathway for improved understanding of genetic modulation of mammalian growth and tissue deposition. *Proc. World Congr. Genet. Appl. to Livest. Prod. Genetics* o:134.

LeBlanc, S. J., K. E. Leslie, and T. F. Duffield. 2005. Metabolic Predictors of Displaced Abomasum in Dairy Cattle. *J. Dairy Sci.* 88:159–170. doi:10.3168/jds.S0022-0302(05)72674-6.

Ludwig, C., J. M. Easton, A. Lodi, S. Tiziani, S. E. Manzoor, A. D. Southam, J. J. Byrne,

L. M. Bishop, S. He, T. N. Arvanitis, U. L. Günther, and M. R. Viant. 2012. Birmingham Metabolite Library: A publicly accessible database of 1-D1H and 2-D1H J-resolved NMR spectra of authentic metabolite standards (BML-NMR). *Metabolomics*. 8:8–18. doi:10.1007/s11306-011-0347-7.

Martín Gómez, M. C., and M. Ballesteros González. 2010. Espectrometría de masas y análisis de biomarcadores. In: *Monografías de la Real Academia Nacional de Farmacia*. p. 113–168.

Martinez-Alvaro, M., S. Agha, A. Blasco, and P. Hernandez. 2017. Muscle lipid metabolism in two rabbit lines divergently selected for intramuscular fat. *J. Anim. Sci.* 95:2576–2584. doi:10.2527/jas.2017.1371.

Martínez-Álvaro, M., A. Blasco, and P. Hernández. 2017. Effect of selection for intramuscular fat on the fatty acid composition of rabbit meat. *Animal*. 1–7. doi:10.1017/S1751731117003494.

Martínez-Álvaro, M., P. Hernández, and A. Blasco. 2016. Divergent selection on intramuscular fat in rabbits: Responses to selection and genetic parameters. *J. Anim. Sci.* 94:4993–5003. doi:10.2527/jas.2016-0590.

Martinez-Alvaro, M., Y. Paucar, K. Satue, A. Blasco, and P. Hernandez. 2018. Liver metabolism traits in two rabbit lines divergently selected for intramuscular fat. *Animal*. 12:1217–1223. doi:10.1017/S1751731117002695.

Mastrangelo, A., A. Ferrarini, F. Rey-Stolle, A. García, and C. Barbas. 2015. From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. *Anal. Chim. Acta*. 900:21–35. doi:10.1016/j.aca.2015.10.001.

Mateescu, R. G. 2015. Genetics of meat quality. *The genetics of cattle*. 2nd ed. (CAB International, editor.). Oxfordshire, UK.

Melzer, N., D. Wittenburg, S. Hartwig, S. Jakubowski, U. Kesting, L. Willmitzer, J. Lisec, N. Reinsch, and D. Repsilber. 2013. Investigating associations between milk metabolite profiles and milk traits of Holstein cows. *J. Dairy Sci.* 96:1521–1534. doi:10.3168/jds.2012-5743.

Melzer, N., D. Wittenburg, and D. Repsilber. 2013. Integrating Milk Metabolite Profile

Information for the Prediction of Traditional Milk Traits Based on SNP Information for Holstein Cows. *PLoS One*. 8. doi:10.1371/journal.pone.0070256.

Moisá, S. J., D. W. Shike, L. Shoup, S. L. Rodriguez-Zas, and J. J. Loor. 2015. Maternal plane of nutrition during late gestation and weaning age alter Angus × Simmental offspring longissimus muscle transcriptome and intramuscular fat. *PLoS One*. 10:1–20. doi:10.1371/journal.pone.0131478.

Naz, S., A. Garcia, M. Rusak, and C. Barbas. 2013. Method development and validation for rat serum fingerprinting with CE-MS: Application to ventilator-induced-lung-injury study. *Anal. Bioanal. Chem.* 405:4849–4858. doi:10.1007/s00216-013-6882-5.

Naz, S., D. C. M. dos Santos, A. García, and C. Barbas. 2014a. Analytical protocols based on LC – MS , GC – MS and CE – MS for nontargeted metabolomics of biological tissues. *Bioanalysis*. 6:1657–1677. doi:10.4155/bio.14.119.

Naz, S., M. Vallejo, A. García, and C. Barbas. 2014b. Method validation strategies involved in non-targeted metabolomics. *J. Chromatogr. A*. 1353:99–105. doi:10.1016/j.chroma.2014.04.071.

Nelson, D. L., and M. M. Cox. 2013. *Lehninger Principles of Biochemistry*. 6th ed.

Olayioye, M. A., S. Vehring, P. Müller, A. Herrmann, J. Schiller, C. Thiele, G. J. Lindeman, J. E. Visvader, and T. Pomorski. 2005. StarD10, a START domain protein overexpressed in breast cancer, functions as a phospholipid transfer protein. *J. Biol. Chem.* 280:27436–27442. doi:10.1074/jbc.M413330200.

Patti, G. J., O. Yanes, and G. Siuzdak. 2013. Metabolomics: the apogee or the omic triloggy. *Nature Rev. Molecular Cell Biol.* 13:263–269. doi:10.1038/nrm3314. Metabolomics.

Putri, S. P., S. Yamamoto, H. Tsugawa, and E. Fukusaki. 2013a. Current metabolomics: Technological advances. *J. Biosci. Bioeng.* 116:9–16. doi:10.1016/j.jbiosc.2013.01.004. Available from: <http://dx.doi.org/10.1016/j.jbiosc.2013.01.004>

Putri, S. P., S. Yamamoto, H. Tsugawa, and E. Fukusaki. 2013b. Current metabolomics: Technological advances. *J. Biosci. Bioeng.* 116:9–16. doi:10.1016/j.jbiosc.2013.01.004.

Robinette, S. L., F. Zhang, L. Bruschweiler-Li, and R. Bruschweiler. 2008. Web server based complex mixture analysis by NMR. *Anal. Chem.* 80:3606–3611.

doi:10.1021/ac702530t.

Rosell, M., M. Kaforou, A. Frontini, A. Okolo, Y.-W. Chan, E. Nikolopoulou, S. Millership, M. E. Fenech, D. MacIntyre, J. O. Turner, J. D. Moore, E. Blackburn, W. J. Gullick, S. Cinti, G. Montana, M. G. Parker, and M. Christian. 2014. Brown and white adipose tissues: intrinsic differences in gene expression and response to cold exposure in mice. *AJP Endocrinol. Metab.* 306:E945–E964. doi:10.1152/ajpendo.00473.2013.

Sakurai, T., Y. Yamada, Y. Sawada, F. Matsuda, K. Akiyama, K. Shinozaki, M. Y. Hirai, and K. Saito. 2013. PRIME Update: Innovative content for plant metabolomics and integration of gene expression and metabolite accumulation. *Plant Cell Physiol.* 54:1–8. doi:10.1093/pcp/pcs184.

Sapp, R. L., J. K. Bertrand, T. D. Pringle, and D. E. Wilson. 2002. Effects of selection for ultrasound intramuscular fat percentage in Angus bulls on carcass traits of progeny. *J. Anim. Sci.* 80:2017–2022.

Schwab, C. R., T. J. Baas, K. J. Stalder, and D. Nettleton. 2009. Results from six generations of selection for intramuscular fat in Duroc swine using real-time ultrasound. I. Direct and correlated phenotypic responses to selection. *J. Anim. Sci.* 87:2774–2780. doi:10.2527/jas.2008-1335.

De Smet, S., K. Raes, and D. Demeyer. 2004. Meat fatty acid composition as affected by fatness and genetic factors: a review. *Anim. Res.* 53:81–98. doi:10.1051/animres:2004003.

Soriano Lopes, A., E. Castañeda Santa Cruz, A. Sussulini, and A. Klassen. 2017. Metabolomic Strategies Involving Mass Spectrometry Combined with Liquid and Gas Chromatography. In: *Metabolomics: From Fundamentals to Clinical Applications*.

Sosa Madrid, B. S., P. Hernández, P. Navarro, C. Haley, L. Fontanesi, R. N. Pena, M. A. Santacreu, A. Blasco, and N. Ibañez-Escriche. 2018. Genomic regions influencing intramuscular fat in divergently selected rabbit populations.

Steinbeck, C., P. Conesa, K. Haug, T. Mahendraker, M. Williams, E. Maguire, P. Rocca-Serra, S. A. Sansone, R. M. Salek, and J. L. Griffin. 2012. MetaboLights: Towards a new COSMOS of metabolomics data management. *Metabolomics.* 8:757–760. doi:10.1007/s11306-012-0462-0.

Sud, M., E. Fahy, D. Cotter, A. Brown, E. A. Dennis, C. K. Glass, A. H. J. Merrill, R. C. Murphy, C. R. H. Raetz, D. W. Russell, and S. Subramaniam. 2007. LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* 35:D527-32. doi:10.1093/nar/gkl838.

Surdacki, A., M. Nowicki, J. Sandmann, D. Tsikas, R. H. Boeger, S. M. Bode-Boeger, O. Kruszelnicka-Kwiatkowska, F. Kokot, J. S. Dubiel, and J. C. Froelich. 1999. Reduced urinary excretion of nitric oxide metabolites and increased plasma levels of asymmetric dimethylarginine in men with essential hypertension. *J. Cardiovasc. Pharmacol.* 33:652-658.

Tautenhahn, R., K. Cho, W. Uritboonthai, Z. Zhu, G. J. Patti, and G. Siuzdak. 2012. An accelerated workflow for untargeted metabolomics using the METLIN database. *Nat. Biotechnol.* 30:826-828. doi:10.1038/nbt.2348.

Troyanskaya, O., M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, and R. B. Altman. 2001. Missing value estimation methods for DNA microarrays. *Bioinformatics.* 2001 Jun; 17:520-525. doi:10.1093/bioinformatics/17.6.520.

Trygg, J., and S. Wold. 2002. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* 16:119-128. doi:10.1002/cem.695.

Ulrich, E. L., H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. Kent Wenger, H. Yao, and J. L. Markley. 2008. BioMagResBank. *Nucleic Acids Res.* 36:D402-8. doi:10.1093/nar/gkm957.

Varmuza, K., and P. Filzmoser. 2009. *Introduction to Multivariate Statistical Analysis in Chemometrics.* 1st. CRC Press, Boca Raton.

Wang, D., S. Strandgaard, M. L. Borresen, Z. Luo, S. G. Connors, Q. Yan, and C. S. Wilcox. 2008. Asymmetric dimethylarginine and lipid peroxidation products in early autosomal dominant polycystic kidney disease. *Am. J. Kidney Dis.* 51:184-191. doi:10.1053/j.ajkd.2007.09.020.

Weikard, R., E. Altmaier, K. Suhre, K. M. Weinberger, H. M. Hammon, E. Albrecht, K. Setoguchi, A. Takasuga, and C. Kuhn. 2010. Metabolomic profiles indicate distinct physiological pathways affected by two loci with major divergent effect on *Bos taurus* growth and lipid deposition. *Physiol Genomics.* 42A:79-88.

doi:physiolgenomics.00120.2010 [pii]\r10.1152/physiolgenomics.00120.2010.

Wiklund, S., E. Johansson, L. Sjöström, E. J. Mellerowicz, U. Edlund, J. P. Shockcor, J. Gottfries, T. Moritz, and J. Trygg. 2008. Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Anal. Chem.* 80:115–122. doi:10.1021/ac0713510.

Wishart, D. S., Y. D. Feunang, A. Marcu, A. C. Guo, K. Liang, R. Vazquez-Fresno, T. Sajed, D. Johnson, C. Li, N. Karu, Z. Sayeeda, E. Lo, N. Assempour, M. Berjanskii, S. Singhal, D. Arndt, Y. Liang, H. Badran, J. Grant, A. Serra-Cayuela, Y. Liu, R. Mandal, V. Neveu, A. Pon, C. Knox, M. Wilson, C. Manach, and A. Scalbert. 2018. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46:D608–D617. doi:10.1093/nar/gkx1089.

Wishart, D. S., T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, E. Dong, S. Bouatra, I. Sinelnikov, D. Arndt, J. Xia, P. Liu, F. Yallou, T. Bjorndahl, R. Perez-Pineiro, R. Eisner, F. Allen, V. Neveu, R. Greiner, and A. Scalbert. 2013. HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res.* 41:D801-7. doi:10.1093/nar/gks1065.

Wold, H. 1966. Nonlinear estimation by iterative least squares procedures. F. N. David, editor. *Res. Pap. Stat.*

Wood, J. D., M. Enser, A. V. Fisher, G. R. Nute, P. R. Sheard, R. I. Richardson, S. I. Hughes, and F. M. Whittington. 2008. Fat deposition, fatty acid composition and meat quality: A review. *Meat Sci.* 78:343–358. doi:10.1016/j.meatsci.2007.07.019.

Wood, J. D., R. I. Richardson, G. R. Nute, A. V. Fisher, M. M. Campo, E. Kasapidou, P. R. Sheard, and M. Enser. 2003. Effects of fatty acids on meat quality: A review. *Meat Sci.* 66:21–32. doi:10.1016/S0309-1740(03)00022-6.

Worley, B., S. Halouska, and R. Powers. 2013. Utilities for quantifying separation in PCA/PLS-DA scores plots. 433:102–104. doi:10.1016/j.ab.2012.10.011.Utilities.

Worley, B., and R. Powers. 2012. Multivariate Analysis in Metabolomics. *Curr. Metabolomics.* 1:92–107. doi:10.2174/2213235X11301010092.

Zhang, A., H. Sun, P. Wang, Y. Han, and X. Wang. 2012. Modern analytical techniques in

metabolomics analysis. *Analyst*. 137:293–300. doi:10.1039/c1an15605e.

Zhao, G. P., J. L. Chen, M. Q. Zheng, J. Wen, and Y. Zhang. 2007. Correlated responses to selection for increased intramuscular fat in a Chinese quality chicken line. *Poult. Sci.* 86:2309–2314. doi:10.3382/ps.2007-00013.

Zhou, L., J. Ji, S. Peng, Z. Zhang, S. Fang, L. Li, Y. Zhu, L. Huang, C. Chen, and J. Ma. 2016. A GWA study reveals genetic loci for body conformation traits in Chinese Laiwu pigs and its implications for human BMI. *Mamm. Genome*. 27:610–621. doi:10.1007/s00335-016-9657-4.

Zomeño, C., P. Hernández, and A. Blasco. 2011. Use of near infrared spectroscopy for intramuscular fat selection in rabbits. *World Rabbit Sci.* 19:203–208. doi:10.4995/vrs.2011.939.

Zomeño, C., P. Hernández, and A. Blasco. 2013. Divergent selection for intramuscular fat content in rabbits. I. Direct response to selection1. *J. Anim. Sci.* 91:4526–4531. doi:10.2527/jas.2013-6361.

Zong, G., Y. Li, A. J. Wanders, M. Alsema, P. L. Zock, W. C. Willett, F. B. Hu, and Q. Sun. 2016. Intake of individual saturated fatty acids and risk of coronary heart disease in US men and women: two prospective longitudinal cohort studies. *Bmj*. i5796. doi:10.1136/bmj.i5796.

ANEXOS

Anexos. Tabla 1. Detalles de los individuos utilizados para el estudio metabólico								
Nº experimental	Fecha Sacrificio	Línea	Sexo	Peso al destete (g)	Peso al sacrificio (g)	Grasa escapular (g)	Grasa perirrenal (g)	Contenido GIM (g/100g)
6008	27/02/2017	GA	M	576	2.058	4,95	15,55	1,72
6021	27/02/2017	GA	F	516	1.884	4,97	16,21	1,41
6072	13/03/2017	GA	F	500	1.916	4,39	13,25	1,93
6018	27/02/2017	GA	M	508	2.008	7,75	13,54	1,44
6298	24/04/2017	GA	M	420	1.524	2,71	5,12	1,04
6050	06/03/2017	GA	M	528	2.018	5,47	13,92	1,32
6211	03/04/2017	GA	M	622	1.912	4,47	11,19	1,14
6529	05/06/2017	GA	F	558	1.984	7,35	17,60	1,29
5974	20/02/2017	GA	M	584	1.870	4,04	10,25	1,17
5990	20/02/2017	GA	M	396	1.742	4,90	8,59	1,67
6290	24/04/2017	GA	F	442	1.776	4,20	11,00	1,46
6642	19/06/2017	GA	M	562	1.844	5,15	11,88	1,25
6032	27/02/2017	GA	M	498	1.632	2,60	4,31	1,15
6043	06/03/2017	GA	F	474	1.848	3,28	5,33	1,19
6011	27/02/2017	GA	F	582	1.632	NA	11,48	1,15
6133	27/03/2017	GA	F	552	1.790	3,59	9,42	1,34
6578	12/06/2017	GA	F	418	1.770	4,57	15,29	1,37
6521	05/06/2017	GA	M	550	1.856	5,68	11,80	1,21
6115	20/03/2017	GA	M	516	1.930	3,96	7,32	1,45
6245	10/04/2017	GA	M	576	2.036	4,79	11,19	1,39
6093	13/03/2017	GB	F	442	1.702	1,80	1,54	0,84
6189	03/04/2017	GB	F	520	1.804	1,31	4,96	0,77
6111	20/03/2017	GB	F	504	1.658	2,85	4,22	0,75
6299	24/04/2017	GB	F	442	1.654	3,88	2,39	0,73
6063	06/03/2017	GB	F	610	1.726	2,75	5,45	0,67
6056	06/03/2017	GB	F	678	1.752	2,60	3,40	0,83
6247	10/04/2017	GB	M	566	1.888	3,97	5,74	0,81
6174	03/04/2017	GB	M	524	1.900	5,25	5,42	0,80
6261	24/04/2017	GB	F	466	1.852	4,78	6,24	0,75
6291	24/04/2017	GB	M	504	1.746	3,51	4,03	0,74
6628	19/06/2017	GB	M	460	1.618	1,79	3,01	0,82
6665	26/06/2017	GB	M	422	1.546	1,36	3,58	0,77
6028	27/02/2017	GB	F	736	2.044	6,59	15,88	1,08
6230	10/04/2017	GB	M	604	1.844	3,29	5,54	0,78
6148	27/03/2017	GB	F	574	1.888	2,93	5,30	0,83
6160	27/03/2017	GB	M	670	2.024	5,43	7,76	0,80
6259	24/04/2017	GB	M	460	1.906	4,42	5,94	0,75
6595	12/06/2017	GB	F	510	1.702	2,23	4,69	0,77
6159	27/03/2017	GB	M	726	2.076	5,71	11,60	0,95
6216	10/04/2017	GB	M	600	1.884	2,74	5,82	0,79

Anexos. Tabla 2. Variables seleccionadas para LC-MS tras aplicar los métodos de selección de variables.								
Masa experimental	VIP	p(corr)	FC(%)	p-valor	Aducto	Error ppm	Peso molecular	Identificación
392,2928	1,41	0,54	-37,5135	7,81E-07	M-H	2	393,2991	N-palmitoyl histidine
449,3134	1,55	0,46	-46,00	1,39E-06	M-H	4	450,3223	SM(d17:1/0:0)
					M-H	4	450,3223	C17 sphingosine-1-phosphocholine
					M+Cl	14	414,3498	28:5(10Z, 13Z, 16Z, 19Z, 22Z)
					M+Cl	14	414,3498	1alpha,25-dihydroxy-3alpha-methyl-3-deoxyvitamin D3 /1alpha,25-dihydroxy-3alpha-methyl-3-deoxycholecalciferol
					M+Cl	14	414,3498	1-Phenyl-1,3-docosanediene
					M+Cl	14	414,3498	4alpha-Formyl-5alpha-cholesta-8-en-3beta-ol
					M+Cl	14	414,3498	4Alpha-hydroxymethyl-5alpha-cholesta-8,24-dien-3beta-ol
					M-H-H2O	17	468,324	3,21-dioxoolean-18-en-28-oic acid
582,4615	1,31	0,34	-46,47	6,82E-05	M-H-H2O	10	601,474	1alpha,25-dihydroxy-11-(3-hydroxy-1-propynyl)-9,11-didehydrovitamin D3 / 1alpha,25-dihydroxy-11-(3-hydroxy-1-propynyl)-9,11-didehydrocholecalciferol
312,2296	1,33	0,44	-50,90	9,06E-05	No compound suitable			N-oleoyl tryptophan
806,5663	1,72	0,54	-48,00	9,38E-05	M-H	5	807,5778	PC(38:5)
					M-H	5	807,5778	PE(41:5)
					M-H-H2O	14	825,5731	PI-Cer(36:0)
					M-H-H2O	4	825,5884	PS(40:4)
					M-H-H2O	4	825,5884	PC(38:4)
					M-H-H2O	4	825,5884	PS(40:3)
593,3684	1,04	-0,43	29,45	3,82E-04	M-H-H2O	13	612,3791	PA(36:4)
					M-H-H2O	1	612,3873	Cholestane-3,7,12,25-tetrol-3-glucuronide
539,3216	1,46	-0,75	70,04	2,63E-04	M+Cl	14	504,3451	(3beta,19alpha)-3,19,23,24-Tetrahydroxy-12-oleanen-28-oic acid
					M+Cl	14	504,3451	(1beta,2alpha,3alpha)-1,2,3,24-Tetrahydroxy-12-oleanen-28-oic acid
390,2764	1,58	0,65	-47,56	2,63E-04	M+Cl	5	355,3086	Pristanoylglycine
					M+Cl	5	355,3086	N-palmitoyl valine
					M+Cl	5	355,3086	N-stearoyl alanine
					M+Cl	5	355,3086	N-Hexadecyl-L-hydroxyproline
488,2797	1,12	0,52	-25,16	5,22E-04	M+Cl	3	453,309	10-Deoxymethymycin
314,2454	1,29	0,34	-39,37	5,22E-04	No compound suitable			
460,2796	1,06	0,55	-23,02	5,22E-04	M-H	20	461,2777	17-phenyl trinor Prostaglandin F2a serinol amide
					M-H-H2O	7	479,3012	LysoPE(18:1(11Z)/0:0)
					M-H-H2O	7	479,3012	PE(18:1(9Z)/0:0)
					M-H-H2O	7	479,3012	Glycerophospho-N-Oleoyl Ethanolamine
					M-H-H2O	7	479,3012	PC(15:1(9Z)/0:0)
438,2974	1,26	0,51	-30,73	3,68E-04	M-H	4	439,3063	PC(O-12:0/O-1:0)
					M-H	4	439,3063	PE(O-16:0/O:0)
					M-H	9	439,3086	N-arachidonoyl vanillylamine
					M-H	9	439,3086	N-arachidonoyl dopamine
782,5682	1,57	0,43	-47,44	3,68E-04	M-H	3	783,5778	PC(36:3)
					M-H	3	783,5778	PE(39:3)
					M-H-H2O	2	801,5884	PS(38:2)
428,2684	1,40	0,52	-37,83	5,22E-04	M+Cl	0	393,2991	N-palmitoyl histidine
378,2762	1,24	0,48	-40,02	5,22E-04	M+Cl	5	343,3086	Ethanolamine Oleate
635,3406	1,09	-0,69	37,05	5,29E-04	M-H	11	636,3411	Ansatrienin A
					M+Cl	8	600,3662	Phorbol 12-tiglate 13-decanoate
568,3233	1,33	0,64	-43,50	5,75E-04	M+Cl	11	533,3481	1-linoleoyl-sn-glycero-3-phosphocholine
					M+Cl	11	533,3481	LysoPE(0:0/22:2(13Z, 16Z))
					M+Cl	11	533,3481	PE(22:2(13Z, 16Z)/0:0)
568,3235	1,50	0,69	-47,83	4,53E-04	M+Cl	11	533,3481	1-linoleoyl-sn-glycero-3-phosphocholine
					M+Cl	11	533,3481	LysoPE(0:0/22:2(13Z, 16Z))
					M+Cl	11	533,3481	PE(22:2(13Z, 16Z)/0:0)
884,6801	1,26	0,32	-48,44	5,01E-04	M-H-H2O	6	903,6928	PS(22:0/22:0)
310,2124	1,58	0,44	-67,44	6,10E-04	M-H	10	311,2228	13-HPODE(1-)
					M-H-H2O	14	329,2355	17beta-Acetamidoandroster-4-en-3-one
615,3545	1,02	-0,66	36,14	7,68E-04	M+Cl	11	580,3916	(3S,4S,3'R)-4-Hydroxyalloxanthin
449,3131	1,52	0,44	-52,79	7,68E-04	M-H	4	450,3223	SM(d17:1/0:0)
					M-H	4	450,3223	C17 sphingosine-1-phosphocholine
					M+Cl	15	414,3498	28:5(10Z, 13Z, 16Z, 19Z, 22Z)
					M+Cl	15	414,3498	1alpha,25-dihydroxy-3alpha-methyl-3-deoxyvitamin D3 /1alpha,25-dihydroxy-3alpha-methyl-3-deoxycholecalciferol
					M+Cl	15	414,3498	22alpha-Hydroxy-campesterol-4-en-3-one; (22S,24R)-22-Hydroxy-ergost-4-en-3-one
					M+Cl	15	414,3498	(22S, 24R)-22-hydroxyergost-4-en-3-one
					M-H-H2O	16	468,324	1alpha,25-dihydroxy-11-(3-hydroxy-1-propynyl)-9,11-didehydrovitamin D3 / 1alpha,25-dihydroxy-11-(3-hydroxy-1-propynyl)-9,11-didehydrocholecalciferol
					M-H-H2O	8	468,3352	N-oleoyl tryptophan

Continúa

436,2817	1,53	0,54	-44,41	7,68E-04	M-H	4	437,2906	PE(P-16:0/0:0)
					M+Cl	11	401,3079	Prostaglandin F2a Ethanolamide-d4
					M+Cl	12	401,3075	PGF2alpha-EA(d4)
					M-H-H2O	8	455,3036	15-HETE-DA
486,3334	1,15	0,47	-35,02	6,34E-04	M-H-H2O	3	505,3532	PC(O-18:2(9Z,12Z)/0:0)[U]
					M-H-H2O	3	505,3532	CerP(d18:1/8:0)
					M-H-H2O	3	505,3532	C-8 Ceramide-1-phosphate
528,2673	1,20	0,57	-29,00	7,68E-04	M-H	7	529,2709	Glycochenodeoxycholate-3-sulfate
					M-H	7	529,2709	glycochenodeoxycholic acid 7-sulfate
					M-H	7	529,2709	N-[(3a,5b,7b)-7-hydroxy-24-oxo-3-(sulfoxy)cholan-24-yl]-Glycine
					M-H-H2O	10	547,291	Glycine, N-[(3a,5b,7b)-3-hydroxy-24-oxo-7-(sulfoxy)cholan-24-yl]
					M-H-H2O	14	547,2934	PS(20:3(8Z,11Z,14Z)/0:0)
539,3213	1,09	-0,63	38,45	7,68E-04	M+Cl	14	504,3451	(3beta,19alpha)-3,19,23,24-Tetrahydroxy-12-oleanen-28-oic acid
					M+Cl	14	504,3451	(1beta,2alpha,3alpha)-1,2,3,24-Tetrahydroxy-12-oleanen-28-oic acid
838,5018	1,61	0,51	-44,21	7,68E-04	M+Cl	18	803,5465	PE(41:7)
					M+Cl	18	803,5465	PC(38:7)
					M+Cl	18	803,5465	2,3-Bis-O-(geranylgeranyl)-sn-glycero-1-phospho-L-serine
					M-H-H2O	1	857,5207	PS(44:9)
482,323	1,22	0,62	-29,77	6,68E-04	M+Cl	10	447,358	N-arachidonoyl-dopamine-d8
650,4483	1,05	0,34	-33,28	8,70E-04	No compound suitable			
456,323	1,13	0,39	-24,24	8,70E-04	M+Cl	15	421,3599	9,14,19,19,19-pentadeuterio-1alpha,25-dihydroxyprevitamin D3 / 9,14,19,19,19-pentadeuterio-1alpha,25-dihydroxyprecholecalciferol
286,2141	1,32	0,41	-49,40	7,77E-04	M-H-H2O	10	305,2355	17beta-Nitro-5alpha-androstane
214,0298	1,05	-0,33	57,08	8,63E-04	M+Cl	12	179,0582	Hippuric acid
					M+Cl	12	179,0582	3-Succinoylpyridine
					M+Cl	12	179,0582	Methyl n-formylanthranilate
					M-H-H2O	12	233,0453	Vitamin B6
258,1829	1,34	0,42	-46,63	1,10E-03	M-H	2	259,1896	Lysyl-Leucine
					M-H	2	259,1896	Leucyl-Lysine
					M-H	2	259,1896	Isoleucyl-Lysine
					M-H	2	259,1896	Lysyl-Isoleucine
593,3667	1,05	-0,44	30,39	9,61E-04	M-H-H2O	10	612,3791	PA(30:4)
					M-H-H2O	4	612,3873	Cholestane-3,7,12,25-tetrol-3-glucuronide
430,2668	1,40	0,53	-38,00	9,61E-04	M-H	16	431,2672	17-phenoxy trinor PGF2a ethyl amide
					M-H-H2O	1	449,2849	Gentamicin C1a
591,3501	1,38	-0,67	64,64	1,13E-03	M+Cl	8	556,3764	3beta,15alpha-Diacetoxylanosta-8,24-dien-26-oic acid
563,3212	1,06	-0,51	50,29	1,04E-03	No compound suitable			
334,2114	1,40	0,45	-52,17	1,15E-03	M-H	11	335,2222	9-peroxy-5Z,7E,11Z,14Z-eicosatetraenoate
					M-H	11	335,2222	11-peroxy-5Z,8Z,12E,14Z-eicosatetraenoate
					M-H	12	335,2228	5(S),15(R)-DiHETE(1-)
					M-H	12	335,2228	Prostaglandin C1(1-)
					M+Cl	14	299,246	Pentadecanoylglycine
					M-H-H2O	9	353,2331	C16 Sphinganine-1-phosphate
492,2173	1,02	0,46	-21,07	1,06E-03	M+Cl	4	457,2498	10,11-Dihydro-12R-hydroxy-leukotriene E4
718,4367	1,03	0,34	-33,80	1,23E-03	No compound suitable			
607,309	1,42	-0,69	66,55	1,46E-03	M-H	8	608,321	Boc-Phe(NMe)-Pro-Phe(NMe)-Gly-OMe
					M+Cl	14	572,3478	dolichyl-4-D-xylosyl phosphate
539,3211	1,49	-0,75	74,53	1,57E-03	M+Cl	13	504,3451	(3beta,19alpha)-3,19,23,24-Tetrahydroxy-12-oleanen-28-oic acid
486,267	1,58	0,56	-46,32	1,63E-03	M+Cl	19	451,3063	PC(O-14:1(1E)/0:0)
					M+Cl	19	451,3063	PC(P-14:0/0:0)
					M-H-H2O	13	505,2788	HMB-Val-Ser-Leu-VE
625,3458	1,36	0,59	-45,27	1,71E-03	M-H	16	626,3431	PI(20:1(11Z)/0:0)
554,3436	1,08	0,49	-30,78	1,70E-03	M+Cl	10	519,3689	PC(P-19:1(12Z)/0:0)
450,2511	1,24	0,56	-30,56	1,72E-03	No compound suitable			
517,3006	1,33	0,47	-36,47	1,80E-03	M-H-H2O	14	536,3114	PG(20:2(11Z,14Z)/0:0)
					M-H-H2O	10	536,3138	Dihydrocelastryl Diacetate
302,2086	1,47	0,66	-44,71	1,83E-03	No compound suitable			
422,3014	1,60	0,58	-47,55	1,90E-03	No compound suitable			
134,0728	1,07	-0,32	54,73	2,00E-03	M-H-H2O	6	153,0902	N-Acetylhistamine
					M-H-H2O	6	477,2855	PE(18:2(9Z,12Z)/0:0)
458,2641	1,56	0,65	-45,36	2,04E-03	M-H-H2O	6	477,2855	LysoPE(0:0/18:2(9Z,12Z))
					M-H-H2O	6	477,2855	LysoPE(0:0/18:2(9Z,12Z))
608,4195	1,08	0,43	-35,04	2,36E-03	M-H	12	609,4192	Heptanoyl thio-PC
472,3173	1,21	0,59	-37,34	2,81E-03	No compound suitable			

Continúa

479,3232	1,63	0,40	-44,32	3,04E-03	M-H	18	480,3216	PA(21:0/0:0)
					M+Cl	15	444,3603	11alpha-ethyl-1alpha,25-dihydroxyvitamin D3 / 11alpha-ethyl-1alpha,25-dihydroxycholecalciferol
					M+Cl	15	444,3603	3-beta-Hydroxy-4-beta-methyl-5-alpha-cholest-7-ene-4-alpha-carboxylate
					M+Cl	15	444,3603	4alpha-Carboxy-4beta-methyl-5alpha-cholesta-8-en-3beta-ol
					M-H-H2O	19	498,3322	PG(O-18:0/0:0)
365,2168	1,02	0,32	-33,26	2,99E-03	No compound suitable			
850,6292	1,46	0,54	-40,97	3,05E-03	M-H	5	851,6404	PE(44:4)
641,389	1,12	-0,69	39,06	3,09E-03	M+Cl	11	606,4261	PA(29:0)
703,3284	1,02	-0,68	31,64	3,09E-03	M-H	7	704,3408	25-Acetyl-6,7-didehydrofivicordin F 3-glucoside
608,3533	1,35	0,58	-41,54	3,23E-03	M-H	6	609,3642	PGPC
					M-H	6	609,3642	PC(16:0/5:0(COOH))
507,3329	1,03	-0,61	29,47	3,32E-03	M-H	0	508,34	3alpha,7alpha,12alpha-trihydroxy-27-carboxymethyl-5beta-cholestan-26-oic acid
					M+Cl	17	472,3553	26,27-diethyl-1alpha,25-dihydroxy-20,21-didehydro-23-oxavitamin D3 / 26,27-diethyl-1alpha,25-dihydroxy-20,21-didehydro-23-oxacholecalciferol
568,3232	1,13	0,56	-40,01	3,25E-03	M+Cl	11	533,3481	1-linoleoyl-sn-glycero-3-phosphocholine
					M+Cl	11	533,3481	LysoPE(0:0/22:2(13Z,16Z))
					M+Cl	11	533,3481	PE(22:2(13Z,16Z)/0:0)
850,434	1,03	0,57	-31,76	3,41E-03	No compound suitable			
942,5399	1,51	0,55	-39,85	3,41E-03	M+Cl	14	907,5575	PS(DiMe(11,3)/MonoMe(13,5))
328,2241	1,42	0,64	-44,45	3,73E-03	No compound suitable			
822,5438	1,54	0,52	-41,45	3,94E-03	M-H	18	823,5363	PE(DiMe(9,5)/DiMe(11,3))
					M-H	18	823,5363	PS(39:5)
					M+Cl	2	787,5727	PS(O-20:0/17:2(9Z,12Z))
					M+Cl	2	787,5727	PS(P-18:0/19:1(9Z))
					M-H-H2O	0	841,5622	PE(44:9)
					M-H-H2O	7	841,568	PI-Cer(t18:0/18:0(2OH))
477,28	1,22	0,56	-31,45	3,94E-03	M-H-H2O	11	496,3036	(3b,16b,20R)-Pregn-5-ene-3,16,20-triol 3-glucoside
					M-H-H2O	11	496,3036	Pregnanediol-3-glucuronide
					M-H-H2O	11	496,3036	3-alpha,20-alpha-dihydroxy-5-beta-pregnane 3-glucuronide
390,276	1,03	0,48	-34,13	4,20E-03	M-H	20	391,2756	N-stearoyl taurine
					M+Cl	6	355,3086	Pristanoylglycine
					M+Cl	6	355,3086	N-palmitoyl valine
					M+Cl	6	355,3086	N-stearoyl alanine
					M+Cl	6	355,3086	N-Hexadecyl-L-hydroxyproline
474,333	1,09	0,55	-37,35	4,22E-03	M-H-H2O	4	493,3532	PC(P-17:0/0:0)
					M-H-H2O	4	493,3532	PE(P-20:0/0:0)
224,1776	1,22	0,37	-42,21	4,69E-03	No compound suitable			
581,3682	1,04	-0,55	37,86	4,63E-03	No compound suitable			
376,2966	1,67	0,61	-50,09	4,73E-03	No compound suitable			
451,2701	1,25	-0,55	56,36	4,91E-03	M+Cl	19	416,2927	(24R,25S)-25,26-epoxy-1alpha,24-dihydroxy-27-norvitamin D3 / (24R,25S)-25,26-epoxy-1alpha,24-dihydroxy-27-norcholecalciferol
					M+Cl	19	416,2927	1alpha,25-Dihydroxy-16-ene-19-nor-24-oxovitamin D3
					M+Cl	19	416,2927	6a-Testosterone Enanthate
472,3173	1,31	0,60	-39,63	5,18E-03	No compound suitable			
512,2962	1,65	0,71	-55,96	5,03E-03	M+Cl	10	477,3219	PE(P-19:1(12Z)/0:0)
					M+Cl	10	477,3219	PC(O-16:2(9E,10E)/0:0)[U]
446,2985	1,04	-0,50	40,88	5,58E-03	M-H-H2O	17	465,309	Glycocholic acid
					M-H-H2O	17	465,309	3a,7b,12a-Trihydroxyoxocholanyl-Glycine
					M-H-H2O	11	465,3219	PC(P-15:0/0:0)
					M-H-H2O	11	465,3219	PE(O-18:1(9Z)/0:0)
664,2419	1,08	0,62	-25,54	5,61E-03	No compound suitable			
771,314	1,01	-0,65	32,28	5,73E-03	No compound suitable			
342,2764	1,27	0,62	-40,86	5,70E-03	M-H-H2O	9	361,2981	N-(2-methoxy-ethyl) arachidonoyl amine
					M-H-H2O	9	361,2981	N-methyl N-(2-hydroxy-ethyl) arachidonoyl amine
					M-H-H2O	9	361,2981	N-(1R-methyl-2-hydroxy-ethyl) arachidonoyl amine
					M-H-H2O	9	361,2981	D-erythro-MAPP
282,1817	1,28	0,38	-46,95	5,75E-03	No compound suitable			
572,3105	1,86	-0,47	273,14	5,77E-03	M-H	19	573,3067	PS(22:4)
					M+Cl	4	537,3431	PC(18:0)
					M+Cl	4	537,3431	PS(P-20:0/0:0)
					M+Cl	4	537,343	PE(21:0)
					M+Cl	4	537,343	1-O-Palmitoyl-2-O-acetyl-sn-glycero-3-phosphorylcholine

Continúa

486,3297	1,11	0,42	-34,15	6,16E-03	M-H-H2O	10	505,3532	PC(P-18:1(9Z)/0:0)
					M-H-H2O	10	505,3532	CerP(d18:1/8:0)
					M-H-H2O	10	505,3532	C-8 Ceramide-1-phosphate
522,2505	1,00	0,51	-21,15	6,91E-03	M-H	0	523,2577	Met-Arg-Phe-Ala acetate salt
585,2914	1,52	0,47	-45,16	6,97E-03	No compound suitable			
621,3254	1,09	-0,61	36,33	7,41E-03	M+Cl	13	586,3482	PI(O-18:0/0:0)
662,3639	1,04	0,35	-33,19	7,56E-03	M-H	5	663,3747	1-Palmitoyl-2-(5-keto-6-octenedioyl)-sn-glycero-3-phosphocholine
452,2516	1,24	0,53	-30,03	7,59E-03	M-H	9	453,2549	Leukotriene E4 methyl ester
406,2302	1,01	0,32	-29,71	8,14E-03	M-H	11	407,242	Leu-leu-tyr; LLY
					M-H-H2O	13	425,2542	PC(0:0/11:0)[U]
					M-H-H2O	13	425,2542	PE(14:0/0:0)
					M-H-H2O	13	425,2542	LysoPE(0:0/14:0)
438,2946	1,28	0,46	-38,11	8,31E-03	M-H	10	439,3063	PC(O-12:0/O-1:0)[U]
					M-H	10	439,3063	PE(O-16:0/0:0)
506,2873	1,32	0,52	-32,68	8,56E-03	M-H-H2O	2	525,3067	PS(18:0/0:0)
					M-H-H2O	2	525,3067	L-a-Lysophosphatidylserine
					M-H-H2O	2	525,3067	PS(18:0/0:0)
552,3102	1,22	0,49	-30,25	8,94E-03	M-H	1	553,3168	LysoPE(0:0/24:6(6Z,9Z,12Z,15Z,18Z,21Z))
					M-H	1	553,3168	LysoPE(24:6(6Z,9Z,12Z,15Z,18Z,21Z)/0:0)
500,312	1,03	0,42	-39,69	9,09E-03	M+Cl	7	465,3457	LysoSM(d18:1)
					M-H-H2O	4	519,3325	PC(18:2(9Z,12Z)/0:0)[U]
					M-H-H2O	4	519,3325	2-linoleoyl-sn-glycero-3-phosphocholine
534,3149	1,02	0,53	-21,27	9,09E-03	M-H-H2O	8	553,338	PS(20:0/0:0)
598,4324	1,27	0,32	-44,47	9,09E-03	No compound suitable			
582,3198	1,12	0,47	-34,94	9,30E-03	No compound suitable			
628,3808	1,25	0,55	-36,68	1,00E-02	M+Cl	10	593,4057	PE(12:0/13:0)
					M+Cl	10	593,4057	PC(11:0/11:0)
					M+Cl	10	593,4057	PE-NMe(12:0/12:0)
					M+Cl	10	593,4057	PE(13:0/12:0)
					M+Cl	10	593,4057	PC(16:0/6:0)
					M+Cl	10	593,4057	PC(4:0/18:0)
					M+Cl	10	593,4057	PC(10:0/12:0)[U]
					M+Cl	10	593,4057	PC(11:0/11:0)[U]
					M+Cl	10	593,4057	PC(12:0/10:0)[U]
					M+Cl	10	593,4057	PC(13:0/9:0)[U]
					M+Cl	10	593,4057	PC(16:0/6:0)[S]
					M+Cl	10	593,4057	PC(16:0/6:0)[U]
					M+Cl	10	593,4057	PC(18:0/4:0)[U]
					M+Cl	10	593,4057	PC(6:0/16:0)[U]
					M+Cl	10	593,4057	PC(9:0/13:0)[U]
					M+Cl	10	593,4057	PE(12:0/13:0)[U]
					M+Cl	10	593,4057	PE(11:0/14:0)[U]
M+Cl	10	593,4057	PE(14:0/11:0)[U]					
M+Cl	10	593,4057	PE(13:0/12:0)[U]					
M+Cl	10	593,4057	PE(10:0/15:0)[U]					
M+Cl	10	593,4057	PE-NMe(12:0/12:0)[U]					
650,3645	1,27	0,53	-51,97	1,02E-02	No compound suitable			
468,2714	1,10	0,42	-33,65	1,05E-02	No compound suitable			
478,3283	1,10	-0,54	46,36	1,11E-02	M-H	4	479,3376	PC(O-16:1(9E)/0:0)[U]
					M-H	4	479,3376	PC(O-16:1(9Z)/0:0)[U]
					M-H	4	479,3376	PC(O-16:1(1E)/0:0)[U]
590,3057	1,54	0,74	-48,81	1,14E-02	No compound suitable			
661,3595	1,01	-0,44	32,87	1,15E-02	M+Cl	7	626,3948	PA(13:0/18:4(6Z,9Z,12Z,15Z))
					M+Cl	4	626,3877	Sucrose monostearate
					M+Cl	6	626,3866	15-deoxy-?12,14-Prostaglandin J2-biotin
644,4299	1,04	0,31	-32,57	1,18E-02	M-H	0	645,437	PE(17:2(9Z,12Z)/12:0)
					M-H	0	645,437	PE(12:0/17:2(9Z,12Z))
					M-H	0	645,437	PE(14:1(9Z)/15:1(9Z))
					M-H	0	645,437	PE(15:1(9Z)/14:1(9Z))
M-H-H2O	1	663,4475	PS(P-16:0/12:0)					
973,4802	1,34	0,47	-31,27	1,20E-02	M+Cl	20	938,4922	PIP(20:4(8Z,11Z,14Z,17Z)/16:0)
					M+Cl	20	938,4922	PIP(20:4(5Z,8Z,11Z,14Z)/16:0)
					M+Cl	20	938,4922	PIP(18:2(9Z,12Z)/18:2(9Z,12Z))
					M+Cl	20	938,4922	PIP(16:0/20:4(8Z,11Z,14Z,17Z))
					M+Cl	20	938,4922	PIP(16:0/20:4(5Z,8Z,11Z,14Z))

Continúa

906,4924	1,33	0,49	-32,87	1,27E-02	No compound suitable			
593,3669	1,10	-0,45	42,59	1,29E-02	M-H-H2O	10	612,3791	PA(12:0/18:4(6Z,9Z,12Z,15Z))
					M-H-H2O	10	612,3791	PA(18:4(6Z,9Z,12Z,15Z)/12:0)
					M-H-H2O	3	612,3873	Cholestane-3,7,12,25-tetrol-3-glucuronide
					M-H-H2O	3	612,3874	Cholestane-3,7,12,25-tetrol-3-glucuronide
561,324	1,21	0,50	-34,61	1,31E-02	M+Cl	17	526,3635	PG(O-20:0/0:0)
					M-H-H2O	6	580,3459	alpha,alpha'-Trehalose 6-palmitate
563,3064	1,48	0,45	-50,48	1,32E-02	No compound suitable			
879,5087	1,47	0,51	-39,19	1,34E-02	M-H	7	880,5102	PI(20:5(5Z,8Z,11Z,14Z,17Z)/18:2(9Z,12Z))
					M-H	7	880,5102	PI(18:3(9Z,12Z,15Z)/20:4(5Z,8Z,11Z,14Z))
					M-H	7	880,5102	PI(18:4(6Z,9Z,12Z,15Z)/20:3(8Z,11Z,14Z))
					M-H	7	880,5102	PI(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/16:1(9Z))
					M-H	7	880,5102	PI(20:3(8Z,11Z,14Z)/18:4(6Z,9Z,12Z,15Z))
					M-H	7	880,5102	PI(18:2(9Z,12Z)/20:5(5Z,8Z,11Z,14Z,17Z))
					M-H	7	880,5102	PI(20:4(5Z,8Z,11Z,14Z)/18:3(6Z,9Z,12Z))
					M-H	7	880,5102	PI(20:4(5Z,8Z,11Z,14Z)/18:3(9Z,12Z,15Z))
					M-H	7	880,5102	PI(18:3(6Z,9Z,12Z)/20:4(5Z,8Z,11Z,14Z))
					M-H	7	880,5102	PI(16:1(9Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))
					M-H	8	880,5231	PGP(18:0/20:3(8Z,11Z,14Z))
					M-H	8	880,5231	PGP(18:0/20:3(5Z,8Z,11Z))
					M+Cl	9	844,5466	PI(O-18:0/18:4(6Z,9Z,12Z,15Z))
					M+Cl	9	844,5466	PI(O-16:0/20:4(5Z,8Z,11Z,14Z))
					M+Cl	9	844,5466	PI(P-16:0/20:3(8Z,11Z,14Z))
					M+Cl	9	844,5466	PI(P-18:0/18:3(6Z,9Z,12Z))
					M+Cl	9	844,5466	PI(P-18:0/18:3(9Z,12Z,15Z))
					M+Cl	16	844,5254	PG(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/20:3(8Z,11Z,14Z))
M+Cl	16	844,5254	PG(20:3(8Z,11Z,14Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))					
M+Cl	16	844,5254	PG(20:5(5Z,8Z,11Z,14Z,17Z)/22:4(7Z,10Z,13Z,16Z))					
M+Cl	16	844,5254	PG(22:4(7Z,10Z,13Z,16Z)/20:5(5Z,8Z,11Z,14Z,17Z))					
909,608	1,04	0,37	-22,86	1,36E-02	M-H-H2O	15	928,6405	PI(O-20:0/22:4(7Z,10Z,13Z,16Z))
880,447	1,38	0,79	-41,18	1,37E-02	No compound suitable			
196,146	1,29	0,33	-38,74	1,37E-02	No compound suitable			
272,199	1,25	0,45	-41,75	1,39E-02	M-H	11	273,2093	N-Isobutyl-2,4,8,10,12-tetradecapentaenamide
526,2511	1,53	0,62	-43,89	1,40E-02	M-H-H2O	7	545,2659	N-[(3a,5b,7a,12a)-3,7-dihydroxy-24-oxo-12-(sulfoxy)cholan-24-yl]-Glycine
					M-H-H2O	11	545,2754	PS(20:4(5Z,8Z,11Z,14Z)/0:0)
628,3806	1,14	0,50	-36,52	1,48E-02	M+Cl	9	593,4057	PE(12:0/13:0)
					M+Cl	9	593,4057	PC(11:0/11:0)
					M+Cl	9	593,4057	PE-NMe(12:0/12:0)
					M+Cl	9	593,4057	PE(13:0/12:0)
					M+Cl	9	593,4057	PC(16:0/6:0)
					M+Cl	9	593,4057	PC(4:0/18:0)
					M+Cl	9	593,4057	PC(10:0/12:0)[U]
					M+Cl	9	593,4057	PC(11:0/11:0)[U]
					M+Cl	9	593,4057	PC(12:0/10:0)[U]
					M+Cl	9	593,4057	PC(13:0/9:0)[U]
					M+Cl	9	593,4057	PC(16:0/6:0)[S]
					M+Cl	9	593,4057	PC(16:0/6:0)[U]
					M+Cl	9	593,4057	PC(18:0/4:0)[U]
					M+Cl	9	593,4057	PC(6:0/16:0)[U]
					M+Cl	9	593,4057	PC(9:0/13:0)[U]
					M+Cl	9	593,4057	PE(12:0/13:0)[U]
					M+Cl	9	593,4057	PE(11:0/14:0)[U]
					M+Cl	9	593,4057	PE(14:0/11:0)[U]
M+Cl	9	593,4057	PE(13:0/12:0)[U]					
M+Cl	9	593,4057	PE(10:0/15:0)[U]					
M+Cl	9	593,4057	PE-NMe(12:0/12:0)[U]					
330,2397	1,16	0,60	-34,32	1,52E-02	No compound suitable			
553,336	1,16	-0,60	49,65	1,54E-02	No compound suitable			
674,1976	1,13	0,58	-25,32	1,54E-02	No compound suitable			

Continúa

864,5412	1,43	0,55	-36,34	1,59E-02	M+Cl	14	829,5833	PS(19:1(9Z)/20:1(11Z))
					M+Cl	14	829,5833	PS(17:0/22:2(13Z,16Z))
					M+Cl	14	829,5833	PS(22:2(13Z,16Z)/17:0)
					M+Cl	14	829,5833	PS(17:1(9Z)/22:1(11Z))
					M+Cl	14	829,5833	PS(20:1(11Z)/19:1(9Z))
					M+Cl	14	829,5833	PS(17:2(9Z,12Z)/22:0)
					M+Cl	14	829,5833	PS(20:2(11Z,14Z)/19:0)
					M+Cl	14	829,5833	PS(18:2(9Z,12Z)/21:0)
					M+Cl	14	829,5833	PS(21:0/18:2(9Z,12Z))
					M+Cl	14	829,5833	PS(22:0/17:2(9Z,12Z))
					M+Cl	14	829,5833	PS(19:0/20:2(11Z,14Z))
					M+Cl	14	829,5833	PS(22:1(11Z)/17:1(9Z))
					M+Cl	12	829,5622	PC(20:4(5E,8E,11E,14E)/20:4(5E,8E,11E,14E))
					M+Cl	12	829,5622	PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:3(6Z,9Z,12Z))
					M+Cl	12	829,5622	PC(20:3(8Z,11Z,14Z)/20:5(5Z,8Z,11Z,14Z,17Z))
					M+Cl	12	829,5622	PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:3(6Z,9Z,12Z))
					M+Cl	12	829,5622	PC(18:2(9Z,12Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))
					M+Cl	12	829,5622	PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:3(8Z,11Z,14Z))
					M+Cl	12	829,5622	PC(20:4(8Z,11Z,14Z,17Z)/20:4(8Z,11Z,14Z,17Z))
					M+Cl	12	829,5622	PC(20:4(8Z,11Z,14Z,17Z)/20:4(5Z,8Z,11Z,14Z))
					M+Cl	12	829,5622	PC(18:2(9Z,12Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))[U]
					M+Cl	12	829,5622	PC(20:4(5E,8E,11E,14E)/20:4(5E,8E,11E,14E))[U]
					M+Cl	12	829,5622	PC(20:4(5Z,8Z,11Z,14Z)/20:4(5Z,8Z,11Z,14Z))[U]
					M+Cl	12	829,5622	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:2(9Z,12Z))[U]
					M+Cl	12	829,5622	PC(18:4(6Z,9Z,12Z,15Z)/22:4(7Z,10Z,13Z,16Z))
M+Cl	12	829,5622	PC(18:3(9Z,12Z,15Z)/22:5(7Z,10Z,13Z,16Z,19Z))					
M+Cl	12	829,5622	PC(20:4(5Z,8Z,11Z,14Z)/20:4(5Z,8Z,11Z,14Z))					
M+Cl	12	829,5622	PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:3(9Z,12Z,15Z))					
M+Cl	12	829,5622	PC(18:3(6Z,9Z,12Z)/22:5(7Z,10Z,13Z,16Z,19Z))					
M+Cl	12	829,5622	PC(18:3(9Z,12Z,15Z)/22:5(4Z,7Z,10Z,13Z,16Z))					
M+Cl	12	829,5622	PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:3(5Z,8Z,11Z))					
M+Cl	12	829,5622	PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:3(9Z,12Z,15Z))					
M+Cl	12	829,5622	PC(22:4(7Z,10Z,13Z,16Z)/18:4(6Z,9Z,12Z,15Z))					
M+Cl	12	829,5622	PC(20:4(5Z,8Z,11Z,14Z)/20:4(8Z,11Z,14Z,17Z))					
M+Cl	12	829,5622	PC(18:3(6Z,9Z,12Z)/22:5(4Z,7Z,10Z,13Z,16Z))					
M+Cl	12	829,5622	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:2(9Z,12Z))					
M+Cl	12	829,5622	PC(20:3(5Z,8Z,11Z)/20:5(5Z,8Z,11Z,14Z,17Z))					
613,3576	1,24	-0,68	52,34	1,59E-02	M+Cl	11	578,3948	PA(12:0/15:0)
					M+Cl	11	578,3948	PA(13:0/14:0)
					M+Cl	11	578,3948	PA(14:0/13:0)
					M+Cl	11	578,3948	PA(15:0/12:0)
					M+Cl	11	578,3819	1,25-Dihydroxyvitamin D3 3-glycoside
360,2866	1,29	0,56	-32,70	1,63E-02	M-H	12	361,2981	N-(2-methyl-2S-hydroxy-ethyl) arachidonoyl amine
					M-H	12	361,2981	L-erythro MAPP
					M-H-H2O	10	379,3086	N-oleoyl proline
					M-H-H2O	10	379,3086	N-linoleoyl valine
918,5777	1,31	0,62	-32,16	1,70E-02	M+Cl	1	883,6091	PC(22:4(7Z,10Z,13Z,16Z)/22:5(7Z,10Z,13Z,16Z,19Z))
					M+Cl	1	883,6091	PC(22:5(4Z,7Z,10Z,13Z,16Z)/22:4(7Z,10Z,13Z,16Z))
					M+Cl	1	883,6091	PC(22:4(7Z,10Z,13Z,16Z)/22:5(4Z,7Z,10Z,13Z,16Z))
					M+Cl	1	883,6091	PC(22:5(7Z,10Z,13Z,16Z,19Z)/22:4(7Z,10Z,13Z,16Z))
954,5003	1,35	0,44	-32,35	1,74E-02	No compound suitable			
587,3337	1,02	-0,60	30,81	1,83E-02	M-H-H2O	19	606,3404	(23S)-23,25-dihydroxy-24-oxovitamin D3 23-(beta-glucuronide) / (23S)-23,25-dihydroxy-24-oxocholecalciferol 23-(beta-glucuronide)
928,5355	1,41	0,57	-36,34	1,89E-02	No compound suitable			
448,2036	1,52	0,46	-57,60	1,92E-02	M-H-H2O	14	467,2284	PS(14:1(9Z)/0:0)
388,2632	1,06	0,49	-25,09	2,03E-02	M+Cl	2	353,293	N-palmitoyl proline
					M+Cl	2	353,293	N-oleoyl alanine
936,5182	1,39	0,50	-35,20	2,20E-02	No compound suitable			
771,5412	1,01	0,36	-22,54	2,22E-02	M-H	17	772,5618	PG(P-20:0/17:2(9Z,12Z))
					M+Cl	14	736,5618	PG(O-20:0/14:0)
					M+Cl	14	736,5618	PG(O-18:0/16:0)
					M+Cl	14	736,5618	PG(O-16:0/18:0)
					M-H-H2O	11	790,5513	PA(21:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))
					M-H-H2O	11	790,5513	PA(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/21:0)

Continúa

950,5129	1,27	0,53	-30,62	2,40E-02	No compound suitable			
400,2811	1,11	0,55	-36,74	2,52E-02	M+Cl	3	365,3128	Prostaglandin F1a-d9
					M-H-H2O	10	419,3036	N-palmitoyl tyrosine
533,3453	1,04	-0,42	27,05	2,63E-02	M-H	6	534,356	1-Linoleoylglycerophosphocholine
					M+Cl	10	498,3709	(22S)-1alpha,22,25-trihydroxy-26,27-dimethyl-23,23,24,24-tetradecahydro-24a,24b,24c-trihomovitamin D3 / (22S)-1alpha,22,25-trihydroxy-26,27-dimethyl-23,23,24,24-tetradecahydro-24a,24b,24c-trihomocholecalciferol
675,2939	1,05	-0,58	36,58	2,69E-02	No compound suitable			
503,3008	1,05	-0,56	35,94	2,86E-02	M+Cl	8	468,3352	N-oleoyl tryptophan
					M+Cl	16	468,324	1alpha,25-dihydroxy-11-(3-hydroxy-1-propynyl)-9,11-didehydrovitamin D3 / 1alpha,25-dihydroxy-11-(3-hydroxy-1-propynyl)-9,11-didehydrocholecalciferol
866,5295	1,55	0,49	-39,58	3,09E-02	No compound suitable			
438,2947	1,20	0,33	-36,65	3,20E-02	M-H	10	439,3063	PC(O-12:0/O-1:0)
					M-H	10	439,3063	PC(O-12:0/O-1:0)[U]
					M-H	10	439,3063	PE(O-16:0/O:0)
					M-H	15	439,3086	N-arachidonoyl vanillylamine
					M-H	15	439,3086	N-arachidonoyl dopamine
624,4188	1,02	0,57	-28,40	3,20E-02	M+Cl	4	589,4471	CerP(d18:1/14:0)
518,1642	2,04	0,53	-55,47	3,28E-02	No compound suitable			
444,2863	1,76	0,56	-51,96	3,42E-02	M-H-H2O	9	463,3006	Gentamicin C2b; Sagamicin
370,3073	1,01	0,60	-29,46	3,82E-02	M-H-H2O	9	389,3294	N-(2-isopropyl-5Z,8Z,11Z,14Z-eicosatetraenoyl)-ethanolamine
					M-H-H2O	9	389,3294	N-(5Z,8Z,11Z,14Z-tricosatetraenoyl)-ethanolamine
					M-H-H2O	9	389,3294	N-(17,17-dimethyl-5Z,8Z,11Z,14Z-heneicosatetraenoyl)-ethanolamine
					M-H-H2O	9	389,3294	N-(17-methyl-5Z,8Z,11Z,14Z-docosatetraenoyl)-ethanolamine
					M-H-H2O	9	389,3294	N-(5-hydroxy-pentyl) arachidonoyl amine
					M-H-H2O	9	389,3294	N-propyl N-(2-hydroxy-ethyl) arachidonoyl amine
					M-H-H2O	9	389,3294	(+)N-(2S-hydroxy-propyl) alpha, alpha-dimethylarachidonoyl amine
					M-H-H2O	9	389,3294	(-)N-(1R-methyl-2-hydroxy-ethyl) alpha, alpha-dimethylarachidonoyl amine
					M-H-H2O	9	389,3294	(+)N-(1S-methyl-2-hydroxy-ethyl) alpha, alpha-dimethylarachidonoyl amine
					M-H-H2O	9	389,3294	(-)N-(2R-hydroxy-propyl) alpha, alpha-dimethylarachidonoyl amine
					M-H-H2O	9	389,3294	N-(2-isopropyl-5Z,8Z,11Z,14Z-eicosatetraenoyl)-ethanolamine
					M-H-H2O	9	389,3294	N-(5Z,8Z,11Z,14Z-tricosatetraenoyl)-ethanolamine
					M-H-H2O	9	389,3294	N-(17,17-dimethyl-5Z,8Z,11Z,14Z-heneicosatetraenoyl)-ethanolamine
M-H-H2O	9	389,3294	N-(17-methyl-5Z,8Z,11Z,14Z-docosatetraenoyl)-ethanolamine					
904,5256	1,40	0,48	-37,80	4,01E-02	No compound suitable			
536,2075	1,48	0,60	-43,76	4,02E-02	No compound suitable			

Continúa

437,3126	1,24	0,51	-33,53	4,02E-02	M-H	15	438,3134	(17Z)-1alpha,25-dihydroxy-26,27-dimethyl-17,20,22,22,23,23-hexadehydrovitamin D3 / (17Z)-1alpha,25-dihydroxy-26,27-dimethyl-17,20,22,22,23,23-hexadehydrocholecalciferol
					M+Cl	16	402,3498	19-Hydroxy-10S,19-dihydrovitamin D3
					M+Cl	16	402,3498	12,14-Heptacosadiynoic acid
					M+Cl	16	402,3498	7alpha-Hydroxy-5beta-cholestan-3-one
					M+Cl	16	402,3498	5beta-cholest-24-ene-3alpha,7alpha-diol
					M+Cl	16	402,3498	5beta-cholest-25-ene-3alpha,7alpha-diol
					M+Cl	16	402,3498	12alpha-Hydroxy-5beta-cholestan-3-one
					M+Cl	16	402,3498	C27:4
					M+Cl	16	402,3498	7a-Hydroxycholesterol
					M+Cl	16	402,3498	5-(16-Heneicosenyl)-1,3-benzenediol
					M+Cl	16	402,3498	5-(12-Heneicosenyl)-1,3-benzenediol
					M+Cl	16	402,3498	eta-Tocopherol
					M+Cl	16	402,3498	2-Methyl-6-phytylhydroquinone
					M+Cl	16	402,3498	(24R)-Cholest-5-ene-3-beta,24-diol
					M+Cl	16	402,3498	Cholest-5-ene-3-beta,7-alpha-diol
					M+Cl	16	402,3498	3-Hydroxy-1-phenyl-1-heneicosanone
					M+Cl	16	402,3498	5-(14-Heneicosenyl)-1,3-benzenediol
					M+Cl	16	402,3498	(3beta,5alpha,6alpha)-Cholest-8-ene-3,6-diol
					M+Cl	16	402,3498	25-Hydroxycholesterol
					M+Cl	16	402,3498	7a-Hydroxy-5b-cholestan-3-one
					M+Cl	16	402,3498	delta-tocopherol
					M+Cl	16	402,3498	2-Methyl-6-phytylquinol; 2-Methyl-6-phytyl-1,4-hydroquinone; 2-Methyl-6-phytylbenzene-1,4-diol; 2-Methyl-6-phytyl-1,4-benzoquinol
					M+Cl	16	402,3498	7a-Hydroxy-5b-cholestan-3-one
					M+Cl	16	402,3498	7B-Hydroxycholesterol
					M+Cl	16	402,3498	7a-Hydroxy-cholesterol
					M+Cl	16	402,3498	25-Hydroxy-cholesterol
					M+Cl	16	402,3498	20-Hydroxycholesterol
					M+Cl	16	402,3498	27-Hydroxycholesterol
					M+Cl	16	402,3498	5,6beta-epoxy-5alpha-cholestan-3beta-ol
					M+Cl	16	402,3498	7a-Hydroxy-5B-cholestan-3-one
M+Cl	16	402,3498	5a-Cholestan-3B-OI-6-One					
M+Cl	16	402,3498	5a-Cholestan-3B-ol-7-one					
M+Cl	16	402,3498	5B-cholest-24-ene-3a,7a-diol					
M+Cl	16	402,3498	5B-cholest-25-ene-3a,7a-diol					
M+Cl	16	402,3498	12a-Hydroxy-5B-cholestan-3-one					
M+Cl	19	402,3354	Bis(2-methylundecan-2-yl) Disulfide					
M-H-H2O	15	456,324	Callystatin A					
472,2863	1,62	0,55	-49,33	4,05E-02	M-H-H2O	19	491,2955	N2'-Acetylgentamicin C1a
414,2731	1,07	0,59	-23,95	4,08E-02	M-H	20	415,2723	15(R)-17-phenyl trinor Prostaglandin F2a ethyl amide
					M-H	20	415,2723	5-trans-17-phenyl trinor Prostaglandin F2a ethyl amide
					M+Cl	13	379,3086	N-oleoyl proline
					M+Cl	13	379,3086	N-linoleoyl valine
436,2812	1,23	0,40	-36,21	4,45E-02	M-H	5	437,2906	PE(P-16:0/0:0)
					M+Cl	10	401,3079	Prostaglandin F2a Ethanolamide-d4
					M+Cl	11	401,3075	PGF2alpha-EA(d4)
					M-H-H2O	9	455,3036	15-HETE-DA
348,2638	1,43	0,47	-44,38	4,74E-02	No compound suitable			
574,2502	1,30	0,51	-33,23	4,89E-02	No compound suitable			
442,3061	1,23	0,43	-28,46	4,99E-02	No compound suitable			

Continúa

449,3123	1,56	0,45	-52,44	4,99E-02	M-H	19	450,311	PA(P-20:0/0:0)
					M-H	6	450,3223	SM(d17:1/0:0)
					M-H	6	450,3223	C17 sphingosine-1-phosphocholine
					M+Cl	17	414,3498	28:5(10Z,13Z,16Z,19Z,22Z)
					M+Cl	17	414,3498	1alpha,25-dihydroxy-3alpha-methyl-3-deoxyvitamin D3 / 1alpha,25-dihydroxy-3alpha-methyl-3-deoxycholecalciferol
					M+Cl	17	414,3498	22-Hydroxy Vitamin D4
					M+Cl	17	414,3498	28:5(13Z,16Z,19Z,22Z,25Z)
					M+Cl	17	414,3498	C28:5
					M+Cl	17	414,3498	1-Phenyl-1,3-docosanedione
					M+Cl	17	414,3498	Ergostane-3,6-dione
					M+Cl	17	414,3498	(3beta,5alpha,7alpha)-14-Methylcholesta-9(11),24-diene-3,7-diol
					M+Cl	17	414,3498	Stenocereol
					M+Cl	17	414,3498	4alpha-Formyl-5alpha-cholesta-8-en-3beta-ol
					M+Cl	17	414,3498	4Alpha-hydroxymethyl-5alpha-cholesta-8,24-dien-3beta-ol
					M+Cl	17	414,3498	(22Alpha)-hydroxy-campest-4-en-3-one
					M+Cl	17	414,3498	22alpha-Hydroxy-campest-4-en-3-one; (22S,24R)-22-Hydroxy-ergost-4 en-3-one
					M-H-H2O	14	468,324	1alpha,25-dihydroxy-11-(3-hydroxy-1-propynyl)-9,11-didehydrovitamin D3 / 1alpha,25-dihydroxy-11-(3-hydroxy-1-propynyl)-9,11-didehydrocholecalciferol
M-H-H2O	10	468,3352	N-oleoyl tryptophan					

Anexos. Tabla 3. Variables seleccionadas para CE-MS tras aplicar los métodos de selección de variables.								
Masa experimental	VIP	p(corr)	FC(%)	p-valor	Aducto	Error ppm	Peso molecular	Identificación
163,0605	1,10	-0,32	15,82	0,03	M+H	2	162,0528	2-hydroxyadipic acid
					M+H	2	162,0528	3,6-Anhydrogalactose; 3,6-Anhydro-D-galactose
					M+H-H2O	2	180,0634	Myoinositol
					M+H-H2O	2	180,0634	D-Fructose
					M+H-H2O	2	180,0634	D-Glucose
					M+H-H2O	2	180,0634	Allose
					M+H-H2O	2	180,0634	D-Galactose
144,1019	1,46	-0,45	23,21	0,03	M+H-H2O	1	161,1052	L-Carnitine
325,1132	1,10	-0,31	15,82	0,03	M+H-H2O	1	342,1162	alpha-Cellobiose; 1-beta-D-Glucopyranosyl-4-alpha-D-glucopyranose
					M+H-H2O	1	342,1162	Maltose
					M+H-H2O	1	342,1162	D-Lactose
					M+H-H2O	1	342,1162	Maltulose
					M+H-H2O	1	342,1162	Inulobiose
					M+H-H2O	1	342,1162	Mannobiose
					M+H-H2O	1	342,1162	Cellobiose
					M+H-H2O	1	342,1162	Trehalulose
					M+H-H2O	1	342,1162	Epimelibiose
					M+H-H2O	1	342,1162	Trehalose
					M+H-H2O	1	342,1162	Melibiose
					M+H-H2O	1	342,1162	Turanose
					M+H-H2O	1	342,1162	Isomaltose
					M+H-H2O	1	342,1162	Galactinol
					M+H-H2O	1	342,1162	Sakebiose
					M+H-H2O	1	342,1162	Allolactose
					M+H-H2O	1	342,1162	DEAE-cellulose
					M+H+HCOONa	6	256,1172	HistidinyL-Threonine
M+H+HCOONa	6	256,1172	ThreoninyL-Histidine					
158,1175	1,98	-0,65	45,83	0,00	M+H-H2O	7	175,1195	L-argininium(1+)/L-arginine
162,1124	1,20	-0,37	20,08	0,01	M+H	0	161,1052	L-Carnitine
151,0711	1,73	0,57	-28,37	0,00	M+Na	14	128,0837	methyl-hexenoic acid
981,0826	1,15	-0,35	16,16	0,03	NO compounds suitable.			
306,085	1,44	-0,44	30,82	0,01	M+Na	15	283,0917	Guanosine
					I+H+HCOON	7	237,092	Ketamine
981,0817	1,22	-0,33	20,54	0,03	NO compounds suitable.			
164,0652	1,10	-0,32	15,44	0,03	NO compounds suitable.			
134,0444	1,37	-0,42	26,53	0,01	M+H	3	133,0375	L-Aspartic acid
					M+H-H2O	12	151,0494	Guanine
328,2229	2,75	-0,71	108,31	0,00	M+Na	6	305,2355	17beta-Nitro-5alpha-androstane
205,0828	1,75	0,57	-27,55	0,00	M+H	4	204,0746	Alanyl-Aspartate
					M+H	4	204,0746	Aspartyl-Alanine
					M+H	4	204,0746	Glycyl-Glutamate
					M+H	4	204,0746	Glutamyl-Glycine
					M+H-H2O	14	222,0892	2E,4E,8E,10E-Dodecatetraenedioic acid
247,0149	1,09	-0,33	16,93	0,05	M+H	13	246,0107	Magnesium sulfate; Magnesium sulfate heptahydrate
					M+H	7	246,0058	Isopentenyl pyrophosphate
					M+Na	15	224,0289	Cysteinyl-Cysteine
123,0557	1,16	-0,41	25,42	0,03	M+H	3	122,048	Niacinamide
126,0217	1,29	-0,39	24,86	0,03	M+H	2	125,0147	Taurine
217,1906	1,69	-0,58	34,58	0,00	M+H-H2O	19	234,1984	4,6,11-hexadecatrienal

425,1405	1,25	-0,41	21,36	0,01	M+H	12	424,1383	His-Phe-OH
					M+H	12	424,1383	HoPhe-His-OH
					M+H	9	424,1369	(1'x,2S)-2-(1,2-Dihydroxy-1-methylethyl)-2,3-dihydro-7H-furo[3,2-g][1]benzopyran-7-one 2' glucoside
					M+H	15	424,127	Nap-Thr-OH
					M+H	15	424,127	Ser-Nap-OH
203,1496	1,40	0,51	-20,08	0,00	M+H	3	202,143	Dimetil-L-arginina
205,0828	1,53	-0,42	43,43	0,01	M+H	4	204,0746	Alanyl-Aspartate
					M+H	4	204,0746	Aspartyl-Alanine
					M+H	4	204,0746	Glycyl-Glutamate
					M+H	4	204,0746	Glutamyl-Glycine
					M+H-H2O	14	222,0892	2E,4E,8E,10E-Dodecatetraenedioic acid
232,9998	1,10	-0,33	16,96	0,04	NO compounds suitable.			
112,037	1,28	-0,38	26,94	0,03	M+Na	1	89,0477	L-Alanine
					M+H-H2O	19	129,0426	4-Oxoproline;4-Oxo-L-proline
					M+H-H2O	19	129,0426	Pyroglutamic acid
216,9874	1,09	-0,32	16,72	0,04	NO compounds suitable.			
939,2723	1,11	-0,32	19,40	0,03	M+Na	2	916,2849	Isorhamnetin 3-(2G-rhamnosylrutinoside)-7-rhamnoside
323,1945	1,56	-0,51	30,62	0,00	M+H	6	322,1852	Neamine; Neomycin A
					M+Na	12	300,2089	C20:6
					M+Na	12	300,2089	17alpha-Methyl-17beta-hydroxyandrosta-4,6-dien-3-one; 17alpha-Methyl-6,7-dehydrotestosterone
					M+Na	12	300,2089	Retinoic Acid
					M+H-H2O	18	340,2038	Piperochromenoic acid
184,9619	1,11	-0,33	16,43	0,05	NO compounds suitable.			
234,9981	1,08	-0,33	16,75	0,03	NO compounds suitable.			
346,9791	1,10	-0,32	17,37	0,04	NO compounds suitable.			
658,9331	1,22	-0,36	20,98	0,04	NO compounds suitable.			

