



# Instituto Politécnico Nacional

---

## CENTRO DE INVESTIGACIÓN EN CIENCIA APLICADA Y TECNOLOGÍA AVANZADA

Unidad Querétaro

"Mecanismo molecular del efecto antidiabético de bebidas  
funcionales elaboradas de plantas medicinales, a partir de datos  
metabolómicos de animales diabéticos "

Tesis que presenta

Ing. Yuvika Reyes Albarracín

Para obtener el Grado de:

Maestría en Tecnología Avanzada

Directores de Tesis:

Dr. Luis Miguel Salgado Rodríguez

Dra. Ma. Guadalupe del Carmen Méndez Montealvo

1 de julio de 2021





# INSTITUTO POLITÉCNICO NACIONAL SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

## ACTA DE REGISTRO DE TEMA DE TESIS Y DESIGNACIÓN DE DIRECTOR DE TESIS

Ciudad de México, a 21 de junio del 2021

El Colegio de Profesores de Posgrado de **CICATA Unidad Querétaro** en su Sesión **extraordinaria** No. **210618** celebrada el día **18** del mes **junio** de **2021**, conoció la solicitud presentada por el alumno:

Apellido Paterno:	<b>Reyes</b>	Apellido Materno:	<b>Albarracín</b>	Nombre (s):	<b>Yuvika</b>
-------------------	--------------	-------------------	-------------------	-------------	---------------

Número de registro: 

A	1	9	0	8	1	4
---	---	---	---	---	---	---

del Programa Académico de Posgrado: 

Maestría en Tecnología Avanzada
---------------------------------

Referente al registro de su tema de tesis; acordando lo siguiente:

1.- Se designa al aspirante el tema de tesis titulado:

Mecanismo molecular del efecto antidiabético de bebidas funcionales elaboradas de plantas medicinales, a partir de datos metabolómicos de animales diabéticos.
--

Objetivo general del trabajo de tesis:

Proponer un mecanismo molecular de acción para el efecto hipoglicemiantre de bebidas funcionales a base de flor de Jamaica ( <i>Hibicus sabdaiffa</i> ) y de té verde ( <i>camellia sinensis</i> ).
---

2.- Se designa como Directores de Tesis a los profesores:

Director: 

Dr. Luis Miguel Salgado Rodríguez
-----------------------------------

2° Director: 

Dra. Ma. Guadalupe del Carmen Méndez Montealvo
--

No aplica:

3.- El Trabajo de investigación base para el desarrollo de la tesis será elaborado por el alumno en:

CICATA-Unidad Querétaro
-------------------------

que cuenta con los recursos e infraestructura necesarios.

4.- El interesado deberá asistir a los seminarios desarrollados en el área de adscripción del trabajo desde la fecha en que se suscribe la presente, hasta la aprobación de la versión completa de la tesis por parte de la Comisión Revisora correspondiente.

Director de Tesis

Dr. Luis Miguel Salgado Rodríguez

2° Director de Tesis

Dra. Ma. Guadalupe del Carmen Méndez Montealvo

Aspirante

Yuvika Reyes Albarracín

Presidente del Colegio

Dr. Juan Bautista Ramos  
DIRECCIÓN



# INSTITUTO POLITÉCNICO NACIONAL

## SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

### ACTA DE REVISIÓN DE TESIS

En la Ciudad de  siendo las  horas del día  del mes de  del  se reunieron los miembros de la Comisión Revisora de la Tesis, designada por el Colegio de Profesores de Posgrado de:  para examinar la tesis titulada:

del (la) alumno (a):

Apellido Paterno:	Reyes	Apellido Materno:	Albarracín	Nombre (s):	Yuvika
-------------------	-------	-------------------	------------	-------------	--------

Número de registro:

Aspirante del Programa Académico de Posgrado:

Una vez que se realizó un análisis de similitud de texto, utilizando el software antiplagio, se encontró que el trabajo de tesis tiene 13 % de similitud. **Se adjunta reporte de software utilizado.**

Después que esta Comisión revisó exhaustivamente el contenido, estructura, intención y ubicación de los textos de la tesis identificados como coincidentes con otros documentos, concluyó que en el presente trabajo **SI**  **NO**  **SE CONSTITUYE UN POSIBLE PLAGIO.**

**JUSTIFICACIÓN DE LA CONCLUSIÓN:** *(Por ejemplo, el % de similitud se localiza en metodologías adecuadamente referidas a fuente original)*

El porcentaje de similitud reportado por el análisis fue menor al 1% en todos los casos y corresponde, generalmente, a grupos de palabras utilizadas de forma conjunta en investigaciones correspondientes al área de estudio.

**\*\*Es responsabilidad del alumno como autor de la tesis la verificación antiplagio, y del Director o Directores de tesis el análisis del % de similitud para establecer el riesgo o la existencia de un posible plagio.**

Finalmente y posterior a la lectura, revisión individual, así como el análisis e intercambio de opiniones, los miembros de la Comisión manifestaron **APROBAR**  **SUSPENDER**  **NO APROBAR**  la tesis por **UNANIMIDAD**  o **MAYORÍA**  en virtud de los motivos siguientes:

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

### COMISIÓN REVISORA DE TESIS

Director de Tesis  
Dr. Luis Miguel Salgado Rodríguez

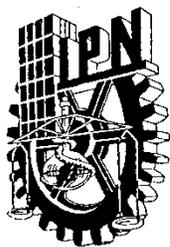
Dr. Gonzalo Velázquez de la Cruz

Dra. Regina Hernández Gama

2° Director de Tesis  
Dra. Ma. Guadalupe del Carmen Méndez Montealvo

Dra. Rosalía Reynoso Camacho

Dr. Juan Bautista Hernández  
**PRESIDENTE DEL COLEGIO DE PROFESORES**



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA DE INVESTIGACIÓN Y POSGRADO**

*CARTA CESIÓN DE DERECHOS*

En la Ciudad de México, D.F. el día 29 del mes de junio del año 2021, la que suscribe Yuvika Reyes Albarracín, alumna del Programa de Maestría en Tecnología Avanzada, con número de registro A190814, adscrito(a) al **Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada (CICATA)**, manifiesto(a) que es el (la) autor(a) intelectual del presente trabajo de Tesis bajo la dirección del Dr. Luis Miguel Salgado Rodríguez y la Dra. Ma. Guadalupe del Carmen Méndez Montealvo y ceden los derechos del trabajo titulado “Mecanismo molecular del efecto antidiabético de bebidas funcionales elaboradas de plantas medicinales, a partir de datos metabolómicos de animales diabéticos”, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso de la autora y/o directores del trabajo. Este puede ser obtenido escribiendo a las siguientes direcciones [yreyesa1900@alumno.ipn.mx](mailto:yreyesa1900@alumno.ipn.mx). Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

---

Yuvika Reyes Albarracín

# Dedicatoria

*Este trabajo se lo dedico a mis padres y familia. Mi fuente de energía para seguir adelante y quienes me dieron su apoyo a pesar de la distancia.*

*A mi padre Walter, un hombre fuerte y sabio. Espero un día poder parecerme a ti, papá.*

*A mi madre Dana. La mujer más constante y alegre que tengo la dicha de llamar “mamá”.*

*También mi hermano Bratzo y a mi cuñada Yazmín. Porque llevaron alegría y risas a un hogar que me tocó abandonar momentáneamente.*

*Y no me olvido de mis dos ángeles en el cielo, mis abuelos José y “Peta”, porque desde arriba nos cuidan.*

# Agradecimientos

Gracias a CONACYT por haberme brindado el apoyo económico para poder realizar mis estudios de maestría.

Gracias a mi director de tesis, el Dr. Luis Miguel Salgado Rodríguez por ser mi mentor y guiar este trabajo a pesar de todos los contratiempos que se presentaron en el camino.

Agradezco también el apoyo académico del resto de mi comité tutorial: a la subdirectora de tesis, la Dra. Ma. Guadalupe Méndez Montealvo, al Dr. Gonzalo Velázquez de la Cruz, a la Dra. Regina Hernández Gama y a la Dra. Rosalía Reynoso Camacho; por los constantes consejos y observaciones que fueron necesarios para la realización de este trabajo.

Me gustaría agradecer a Fernando, por acompañarme y apoyarme durante este proceso, en el cual crecí en muchos aspectos.

De igual forma, agradecer a Alejandra Rangel y a Karen Castellanos quienes fueron un soporte académico y emocional constante.

Finalmente, me gustaría agradecer al Instituto Politécnico Nacional por haberme acogido en sus instalaciones y por formar una parte muy importante en mi formación académica.

# Contenido

<b>Dedicatoria .....</b>	<b>3</b>
<b>Agradecimientos .....</b>	<b>4</b>
<b>Contenido .....</b>	<b>5</b>
<b>Índice de Figuras .....</b>	<b>8</b>
<b>Índice de Tablas .....</b>	<b>9</b>
<b>Glosario .....</b>	<b>10</b>
<b>Resumen .....</b>	<b>11</b>
<b>Abstract .....</b>	<b>12</b>
<b>Capítulo 1 .....</b>	<b>13</b>
<b>Introducción .....</b>	<b>13</b>
<b>Capítulo 2.....</b>	<b>15</b>
<b>Antecedentes .....</b>	<b>15</b>
2.1. Bioinformática .....	15
2.1.1. Biología de sistemas.....	15
2.1.2. Ciencias ómicas .....	16

2.1.3. Metabolómica .....	18
2.2. Diabetes .....	24
2.2.1 Tipos de diabetes.....	25
2.2.2 Epidemiología .....	26
2.2.3 Cuadro clínico .....	26
2.2.4 Mecanismo .....	27
2.2.5 Proteína quinasa activada por AMP .....	28
2.3. Modelo animal de diabetes .....	29
<b>Capítulo 3 .....</b>	<b>31</b>
<b>Justificación.....</b>	<b>31</b>
<b>Hipótesis .....</b>	<b>32</b>
<b>Objetivo general.....</b>	<b>33</b>
<b>Objetivos específicos.....</b>	<b>33</b>
<b>Capítulo 4 .....</b>	<b>34</b>
<b>Metodología.....</b>	<b>34</b>
4.1 Equipo de laboratorio y detalles computacionales .....	34
4.1.1 Equipo:.....	34
4.1.2 Software: .....	34
4.1.3 Servidores: .....	35
4.2 Métodos .....	35
4.2.1 Exportación de datos .....	35
4.2.2 Conversión de formato .....	35
4.2.3 Cambio de MS2 por MS1 .....	36

4.2.4	Preprocesamiento de datos.....	36
4.2.5	Análisis estadístico.....	37
4.2.6	Identificación de metabolitos.....	38
4.2.7	Construcción de redes metabólicas.....	39
<b>Capítulo 5</b>	.....	<b>40</b>
<b>Resultados y Discusión</b>	.....	<b>40</b>
5.1.	Exportación de datos.....	40
5.2.	Cambio de formato.....	40
5.3.	Preprocesamiento de datos.....	41
5.4.	Análisis estadístico.....	43
5.5.	Identificación de metabolitos.....	52
5.6.	Construcción de redes metabólicas.....	56
<b>Conclusiones</b>	.....	<b>63</b>
<b>Recomendaciones</b>	.....	<b>64</b>
<b>Referencias Bibliográficas</b>	.....	<b>65</b>
<b>Anexos</b>	.....	<b>72</b>
<b>Anexo A</b>	.....	<b>73</b>

# Índice de Figuras

Figura 1. Relaciones interdisciplinarias que componen la Biología de sistemas (Cambiaghi, Ferrario, & Masseroli, 2017).....	16
Figura 2. Primeras tecnologías ómicas desarrolladas (Hasin et al., 2017).....	18
Figura 3. Factores relacionados con el metaboloma. ....	19
Figura 4. Tipos de estrategias metabolómicas (Hasin et al., 2017).....	21
Figura 5. Pasos del preprocesamiento de datos (Bauer, Cramer, & Schuchhardt, 2011).....	22
Figura 6. Metabolismo de una persona sana y personas con diabetes tipo I y II. ....	25
Figura 7. Diseño experimental de animales sanos y alimentados con una dieta alta en grasa y fructosa y tratados con bebidas. ....	29
Figura 8. Diagrama de espectrometría de masas (MS/MS).....	40
Figura 9. Cromatograma de iones totales.....	42
Figura 10. Diagrama PCA de set de datos filtrados. ....	44
Figura 11. Diagrama PLSDA de set de datos filtrados. ....	45
Figura 12. Test de validación de set de datos filtrados. ....	46
Figura 13. Diagramas O-PLSDA correspondientes a los dos grupos validados .....	48
Figura 14. Validación por permutación de grupos validados.....	50
Figura 15. Features y valores VIP de grupos estudiados. ....	51
Figura 16. Análisis de enriquecimiento de rutas metabólicas. ....	57
Figura 17. Gráfico de impacto en las rutas metabólicas.....	59

# Índice de Tablas

Tabla 1. Composición de bebidas funcionales. ....	29
Tabla 2. Grupos formados para PLSDA - ortogonal. ....	47
Tabla 3. Metabolitos identificados. ....	53
Tabla 4. Rutas metabólicas relacionadas con los metabolitos identificados. ....	55
Tabla 5. Compuestos detectados por MS/MS relacionados con Diabetes tipo II. ....	73
Tabla 6. Abundancias relativas de los compuestos identificados por grupo. ....	76

# Glosario

**DM:** Diabetes mellitus.

**ARNm:** Ácido ribonucleico mensajero.

**VNTR:** *Variable number of tandem repeats.*

**kDa:** Kilo Daltons.

**Target:** Célula diana.

**RNM:** Resonancia magnética nuclear.

**MS:** Espectrometría de masas.

**Features:** Posibles compuestos.

**OMS:** Organización mundial de la salud.

**IRS:** Sustrato receptor de insulina.

**ATP:** Adenosin tri fosfato

**ADP:** Adenosin di fosfato

**HFD:** *High fructose diet*

**Peak picking:** Selección de picos

# Resumen

La metabolómica es la manera más adecuada de abordar el estudio de la diabetes mellitus, debido a que es una enfermedad metabólica. Para esto se requiere de técnicas y metodologías robustas que permitan la detección de metabolitos y compuestos de interés mediante programas accesibles y rutinas de trabajo sencillas. Sin embargo, el uso de métodos complejos dificulta la visión integral de los datos, impidiendo localizar factores causales de ruido. En este trabajo se llevó a cabo un análisis metabolómico usando programas comerciales y plataformas libres para la identificación de metabolitos y redes metabólicas. Se utilizaron datos provenientes de ratas macho tipo *Wistar* en un modelo de resistencia a la insulina. Estas fueron tratadas con bebidas funcionales con un contenido mayoritario de tres plantas usadas en herbolaria: jamaica roja, jamaica blanca y té verde. Se realizó un preprocesamiento de datos adicional, que constó de control de calidad y limpieza de valores perdidos por tratamiento. Permitiendo así, obtener resultados más robustos en los análisis estadísticos. Este paso fue crucial para la identificación de compuestos. Los datos fueron validados por un modelo ortogonal de PLSDA con un *p value* menor a 0.005 en cuatro tratamientos. Como resultado, se logró identificar siete metabolitos; sin embargo, acetoacetyl-CoA y ácido lisofosfatídico (18:1) mostraron diferencias en su abundancia relativa entre el control y en animales obesos que fueron tratados con la bebida blanca y roja. Dichos compuestos guardan relación en el metabolismo del piruvato, la oxidación de ácidos grasos, metabolismo de glicerofosfolípidos y adición de sustratos gluconeogénicos.

**Palabras clave:** Diabetes mellitus tipo II, metabolómica, biología de sistemas.

# Abstract

Metabolomics is the most appropriate way to approach the study of diabetes mellitus because it is a metabolic disease. For this robust techniques and methodologies are required that allow the detection of metabolites and compounds of interest through accessible programs and simple work routines. However, the use of complex methods difficulties a holistic view of the data, preventing noise locate causative factors. In this work, a metabolomic analysis was perform using commercial programs and free platforms for the identification of metabolites and metabolic networks. Metabolomic data from male Wistar-type rats were used in an insulin resistance model. These were treated with functional drinks with a majority content of three plants used in herbalism: red hibiscus, white hibiscus and green tea. Additional data pre-processing was performed, which consisted of quality control and cleaning of lost values by treatment. Thus, allowing for more robust results in statistical analyzes. It was one of the main reasons why this step was crucial for the identification of compounds. The data were validated by an orthogonal PLSDA model with a *p value* less than 0.005 in four treatments. As a result, it was possible to identify seven metabolites; however, acetoacetyl-CoA and lysophosphatidic acid (18: 1) showed differences in relative abundance between the control group and in obese animals that were treated with the white hibiscus drink. Such compounds keep relationship among pyruvate metabolism, fatty acid oxidation, glycerophospholipid metabolism and addition of gluconeogenic substrates.

**Key words:** Type II diabetes mellitus, metabolomics, system biology.

# Capítulo 1

## Introducción

La biología de sistemas se originó junto a la genómica; sin embargo, la popularidad de la primera surge después del auge de la segunda. Desde el surgimiento de otras ciencias ómicas, como la proteómica y metabolómica, se generaron grandes cantidades de datos y la capacidad de análisis e interpretación se vio mermada. Así, se generaron nuevas tecnologías y técnicas para optimizar los análisis y aumentar la sensibilidad de estos (Phelan, 2020). Sin embargo, esto causó la generación de una gran variedad de metodologías; causando la falta de estandarización. La metabolómica ha sido una de las más afectadas pues el análisis de los resultados y su interpretación se ha vuelto cada vez más compleja. Para resolver esto, los laboratorios han desarrollado técnicas y rutinas, tanto comerciales como libres (Lederer & La Manno, 2020). Estas son capaces de evaluar y facilitar la interpretación de mega datos. La metabolómica es la herramienta más adecuada para abordar el estudio de enfermedades con un origen metabólico, como la diabetes (Hasin, Seldin, & Lusic, 2017; S. Park, Sadanala, & Kim, 2015).

La diabetes mellitus (DM) es un trastorno crónico que implica la desregulación del metabolismo de carbohidratos, lípidos y proteínas; su principal característica es la presencia de un índice glucémico elevado (World Health Organization, 2019). Esto, además de otras complicaciones ha hecho de la diabetes una enfermedad peligrosa que ha cobrado importancia debido al índice de mortalidad que tiene. Se ha convertido en la tercera causa de muerte en México y a nivel mundial, por debajo de las causadas por COVID-19 (Inegi, 2021). Las enfermedades cardiovasculares siguen ocupando el primer lugar (Inegi, 2021). Según la Federación Mexicana de Diabetes, de cada 10 mexicanos entre los 19 y 60 años, aproximadamente 4 tienen algún tipo de diabetes. Existen varios tipos de la enfermedad, sin embargo, la DM tipo II es la más importante. La principal característica de ésta es la resistencia a la insulina (World Health Organization, 2019).

Estados metabólicos como la obesidad, malos hábitos alimenticios y tendencia al sedentarismo, son factores que contribuyen a generar resistencia a la insulina y que empeoran el cuadro clínico. A pesar de tratarse de un desorden metabólico irreversible,

la terapia farmacológica puede mejorar la calidad de vida del paciente. El principal tratamiento para la diabetes son drogas hipoglicemiantes; sin embargo, una parte de la población recurre a la herbolaria por diferentes causas.

Con este propósito, se estudiaron técnicas metabolómicas con la capacidad de identificar metabolitos endógenos y redes metabólicas que puedan detectar un cambio significativo en un modelo animal de resistencia a la insulina tratado con decocciones herbales. También, se analizaron las capacidades de diferentes programas para ensamblar una rutina que permita elaborar un correcto preprocesamiento de datos, dado que este paso es crucial para obtener resultados confiables.

# Capítulo 2

## Antecedentes

### 2.1. Bioinformática

#### 2.1.1. Biología de sistemas

La biología tradicional se basó en el direccionamiento del entendimiento entre genes, proteínas y su función biológica en respuesta a una perturbación; dando como resultado, una cantidad grande de conocimiento. Sin embargo, no ha proporcionado información sobre la integración de todas las células y compuestos; ni tampoco cómo se relaciona una respuesta frente a la de todo un organismo (Pucker, Schilbert, & Schumacher, 2019).

A medida que los avances científicos y tecnológicos dieron como resultado información más compleja y en mayor proporción, se requirió describir e identificar sus componentes y mecanismos. Así, se inició la búsqueda de la integración de toda la información en una teoría unificadora que pueda explicar el comportamiento sistemático de un organismo frente a un estímulo.

La biología de sistemas es un enfoque holístico de la biología tradicional. Busca comprender cómo los fenómenos biológicos son impulsados por un comportamiento interconectado entre las diferentes moléculas de un organismo en particular, mientras este se encuentra en funcionamiento. Su principal objetivo es la comprensión en red de los organismos vivos como un todo; entendiendo a su vez, que este tipo de concepto es más acertado que la suma de sus partes. Requiere la integración y colaboración de disciplinas científicas para pronosticar cómo un sistema puede cambiar con respecto al tiempo y bajo diferentes condiciones. Planteando así, el desarrollo de soluciones a perturbaciones que provoquen caos en las redes (Abascal, 2020; Pucker et al., 2019).

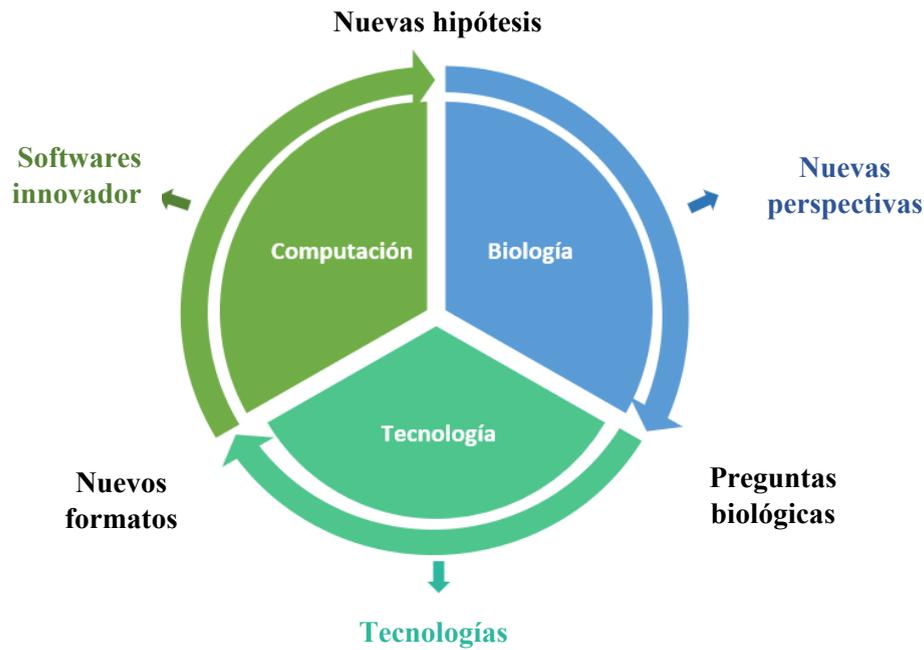


Figura 1. Relaciones interdisciplinarias que componen la Biología de sistemas (Cambiaghi, Ferrario, & Masseroli, 2017).

Se basa en la teoría de sistemas, que es el estudio interdisciplinario de todo sistema de forma general. Se investigan fundamentos y estrategias que puedan ser aplicados a cualquier nivel en un sistema en estudio. Estos deben descubrir y comprender de forma sistemática y dinámica; cuáles son todos los componentes del sistema, cómo interactúan y luego cómo esa interacción da lugar a un fenómeno emergente como comparación a dos fenotipos distintos. (Aloy, 2017)

Su principal justificación es la relación de la biología, la computación y la tecnología como una recurrente necesidad para el desarrollo de técnicas y softwares innovadores (ver Figura 1). Generando así, el acceso a nuevos tipos de datos que permitan la producción de constantes preguntas, perspectivas biológicas y respuestas. Esto, se convierte en un círculo virtuoso donde sólo un entorno interdisciplinario puede reunir a médicos, ingenieros, informáticos, biólogos y otros, capaces de abordar grandes desafíos. Su principal tarea es la correcta integración, modelización e interpretación de los datos recabados en las aproximaciones de las ciencias ómicas, tales como la genómica, transcriptómica, proteómica, metabolómica, entre otras (Abascal, 2020; Aloy, 2017).

### 2.1.2. Ciencias ómicas

El término “ómica” se refiere al estudio de grandes cantidades de moléculas en un

organismo en particular. Fue empleado con mayor frecuencia luego del proceso de mapeo y secuenciación del genoma humano, permitiendo la obtención de grandes cantidades de datos genéticos de una célula o tejido, brindando una especie de fotografía instantánea a diferentes niveles biológicos dentro de un organismo (Hasin et al., 2017).

Se diferenciaron varios tipos de ciencias ómicas; sin embargo, la pionera fue la genómica, con la cual se logró estudiar todos los genes. Con el tiempo, surgió la transcriptómica que aborda la investigación de los transcritos (ARNm principalmente), es decir de cómo se transcriben los genes (Hasin et al., 2017).

Luego de las dos antes mencionadas, las siguientes fueron implementadas rápidamente. Por ejemplo, la proteómica estudia todas las proteínas en conjunto, dando información de cómo se tradujo el gen. Por otro lado, se encuentra la metabolómica que estudia los metabolitos o sus derivados, especificando cómo funcionó el gen, acercándonos más aún al fenotipo. También, se describe la epigenómica que estudia el conjunto de procesos por los cuales se regula la transcripción de los genes sin afectar la secuencia de ADN. Encontraremos también a ómicas específicas, como el caso de la nutrigenómica o la lipidómica. La primera es la herramienta que permite conocer a los cambios en la expresión de genes en respuesta a una exposición crónica de alimentos. Por otro lado, la segunda estudia el conjunto de moléculas con carácter lipídico, en una célula. Para una descripción gráfica, observar Figura 2 (Lederer & La Manno, 2020).

La integración e interpretación holística de los datos recabados de las metodologías mencionadas ha sido el objetivo perseguido en los últimos años. Este brinda un panorama general del estado fisiológico del organismo, a varios niveles. Así, se ha logrado identificar biomarcadores que abrieron las puertas a la detección temprana de patologías y su posible prevención.

Un ejemplo muy claro es la identificación de marcadores genéticos de diabetes mellitus tipo II. A pesar de ser clasificada como una enfermedad poligénica, se han identificado algunos polimorfismos como rs2291166 del gen *TJPI* o el *VNTR* (*Variable Number of Tandem Repeats*) CAG(n) del gen *ATXN2*. También se reportan marcadores bioquímicos como aminoácidos de cadena ramificada o ciertos tipos de azúcares (ej. glucosa, sacarosa, fructuosa) detectados por métodos metabolómicos de alto rendimiento. Estos han brindado información sobre la desregulación metabólica implícita de la enfermedad. (Gan, Ramachandran, Lim, & Koh, 2020)

### 2.1.3. Metabolómica

Los metabolitos son moléculas endógenas y exógenas de bajo peso molecular (<1 kDa) que pueden jugar el rol de sustratos, productos o intermediarios en una reacción bioquímica llevada a cabo dentro de la célula. Con esta premisa, se puede definir la metabolómica como el conjunto de técnicas de alto rendimiento que tienen como fin la detección, identificación y cuantificación de metabolitos presentes en una muestra fenotípicamente deseable, evaluando de forma integral su estado metabólico. El conjunto de metabolitos detectados se le llama metaboloma. (Liu & Locasale, 2017; S. Park et al., 2015)

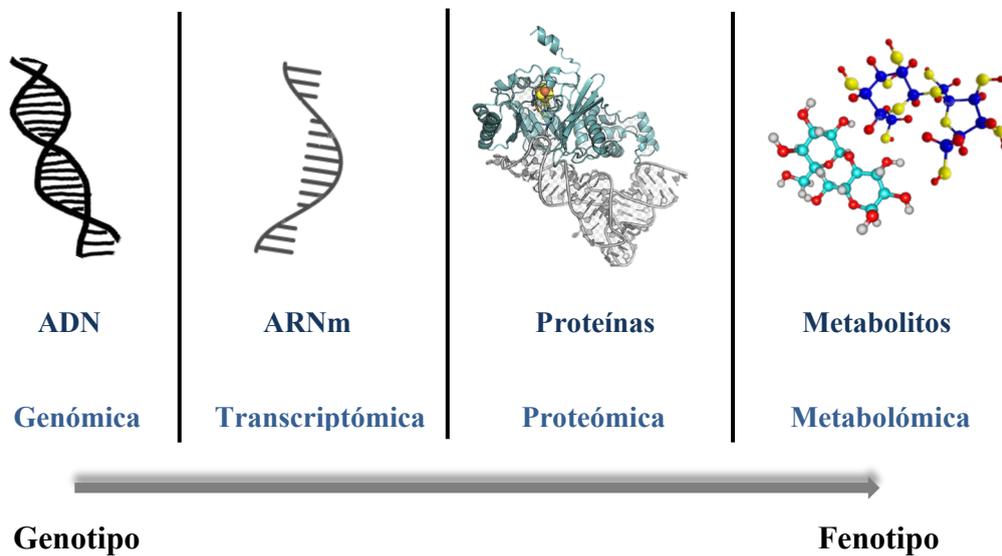


Figura 2. Primeras tecnologías ómicas desarrolladas (Hasin et al., 2017).

La composición y concentración de estos metabolitos se ve afectada por factores genéticos y por factores ambientales, como el tipo de dieta, el estilo vida, enfermedades o la medicación. Los cambios que presentan estos metabolitos dan información sobre las vías metabólicas de un organismo. Al ser analizados de forma integral, son capaces de exhibir una enorme especificidad tisular y dinámica temporal (Duncan, Fyrestam, & Lanekoff, 2019).

Una aplicación importante de la metabolómica está en el área de medicina molecular y personalizada; como control de trasplantes y detección de enfermedades en recién nacidos, con el fin de fortalecer la prevención. La información extraída del análisis del metaboloma puede ser tan compleja, que proporciona valiosa información acerca de

la expresión de genes y proteínas; así como la influencia del entorno. Los metabolitos detectados hasta el momento superan los 14 mil compuestos, mientras que las reacciones metabólicas suman aproximadamente 18000, según *The Human Metabolome Database (HMDB)*, reportados hasta el 2018 (Wishart et al., 2018).

Dichos compuestos son el producto final de cambios y expresiones genéticas, interacciones de proteínas y compuestos que se reflejan finalmente en el metaboloma del organismo. Por esta razón se dice que la metabolómica es la ciencia ómica que más se aproxima al fenotipo (ver Figura 3) (Liu & Locasale, 2017).

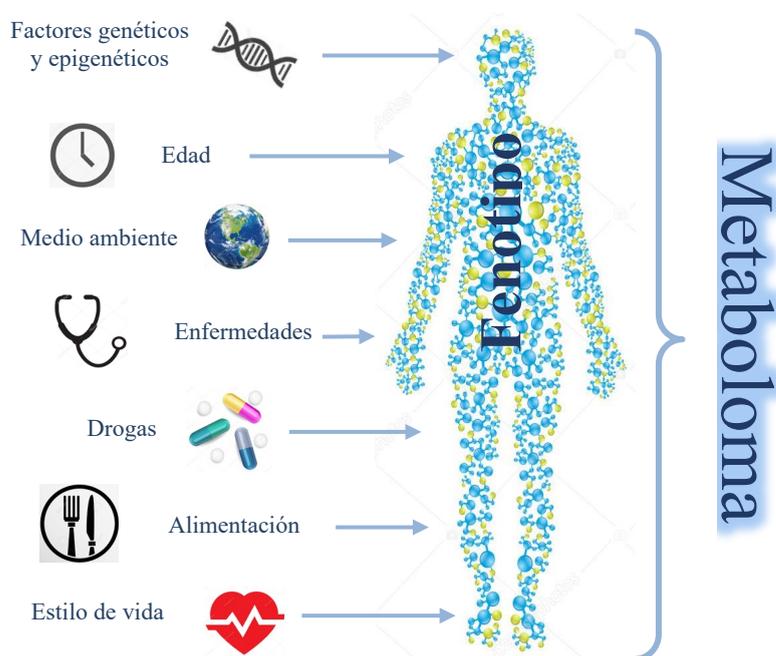


Figura 3. Factores relacionados con el metaboloma.

A pesar de la utilidad y la variedad de aplicaciones, la complejidad del metaboloma genera inherentemente desafíos y limitaciones debido a la variedad química de sus blancos moleculares. Es decir, no existen aún metodologías que permitan la identificación integral de todas las clases de metabolitos presentes en el organismo, como lo son: aminoácidos, ácidos orgánicos, lípidos, carbohidratos, esteroides, nucleótidos, entre otros. Esto se debe, principalmente, a la diversidad de propiedades fisicoquímicas que posee cada grupo (Jacob et al., 2018).

Generalmente se separa el metaboloma en subgrupos de metabolitos, según su afinidad a una propiedad medible, como los grupos funcionales o la polaridad. Luego, estos son analizados por separado a través de otras técnicas. Por esta razón, analizar el

metaboloma es una tarea compleja, ya que debe ser estudiado por partes y luego, integrado en formatos diferentes (Ren, Hinzman, Kang, Szczesniak, & Lu, 2015). Debido a esta problemática, se han confeccionado múltiples procedimientos y metodologías para el análisis de datos que requieren de un proceso de refinamiento bioestadístico. Esto abrió paso a que laboratorios y empresas dedicadas al rubro desarrollaran técnicas individuales patentadas, por lo que para su uso se requiere la adquisición de licencias. La creación de una técnica metabolómica estándar, unificando las demás, es menester para los profesionales de esta área. De esta forma, se incrementa el grado de sensibilidad y permite que los resultados muestren un panorama real (Majumder et al., 2021).

Para la detección y cuantificación, en metabolómica se trabaja con dos técnicas principalmente. Estas son la resonancia magnética nuclear (RNM) y la espectrometría de masas (MS). Ambas pueden cuantificar e identificar una gran variedad de compuestos; sin embargo, ambas son utilizadas en campos distintos. Por un lado, los resultados de RNM son más reproducibles, pero menos sensibles. Además, se requiere una mayor cantidad de muestra para llevar a cabo el experimento. Estas características no son deseables en áreas médicas ni en la investigación, ya que la mayoría de las veces la muestra es limitada. Por esta razón, se trabaja principalmente con espectrometría de masas precedida de un tipo de cromatografía. Esta técnica se vale, principalmente, de dos estrategias que han guiado la metabolómica desde sus inicios (Majumder et al., 2021).

### ***2.1.3.1. Estrategias metabolómicas***

Planteado lo anterior, las estrategias metabolómicas se pueden dividir en dos grandes grupos: análisis dirigido (*targeted metabolomics*) y no dirigido (*untargeted metabolomics*), según el objetivo del análisis. De igual forma, se elige entre el endometaboloma (metabolitos que forman parte de metabolismo interno del organismo) y el exometaboloma (incluye compuestos provenientes de la dieta u otras fuentes) y se puede optar por una metodología semi dirigida (*semi-targeted metabolomics*) (Phelan, 2020).

El análisis no dirigido es el primero en realizarse cuando se desconocen detalles del metaboloma. Pues, por la cantidad de información que es integrada de forma holística, nos permite generar hipótesis de los supuestos fenómenos que puedan causar la disrupción armónica del metabolismo. También se pueden identificar patrones característicos de los fenotipos de interés; generando así, las huellas metabólicas. El análisis realizado en este

caso es semicuantitativo ya que las características detectadas (*features*) requieren de una comprobación de componentes en el siguiente paso. Estas son clasificadas y cribadas de forma integral para una correcta identificación (Begou, Gika, Theodoridis, & Wilson, 2018).

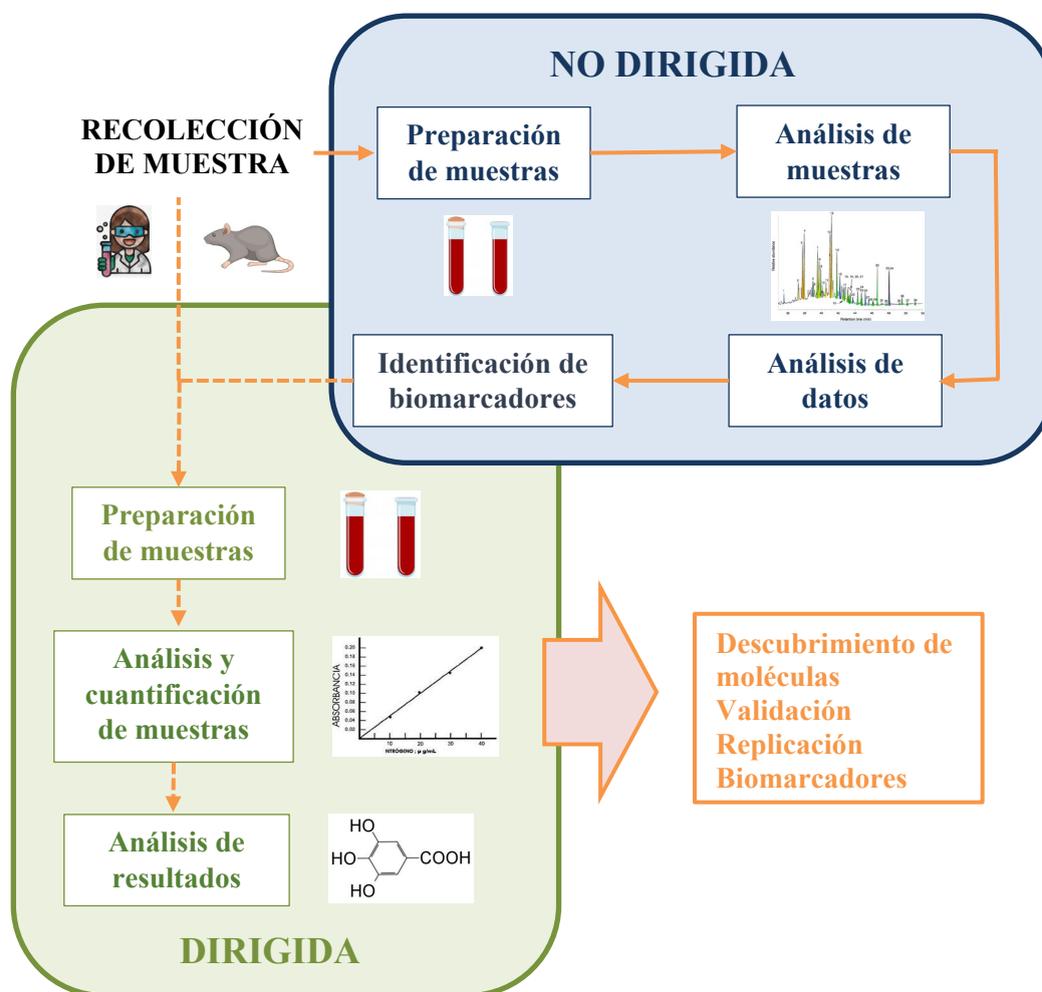


Figura 4. Tipos de estrategias metabolómicas (Hasin et al., 2017).

Por otro lado, las metodologías de análisis dirigido permiten dar una respuesta a las hipótesis generadas en el estudio no dirigido. De esta forma se puede validar el análisis que le precede. Los estudios dirigidos permiten la verificación y cuantificación de moléculas específicas en las rutas metabólicas afectadas, mediante el uso de estándares. Estos ayudan a asegurar la presencia y a realizar el diagnóstico completo de un posible compuesto validado sin la necesidad de recurrir a la comparación subjetiva de patrones de fragmentación. También ayudan a precisar sus interacciones en el metaboloma como

biomarcador (S. Park et al., 2015).

Sobre la preparación de las muestras, ambos siguen metodologías similares. Los análisis no dirigidos suelen constar de una extracción global de metabolitos de forma hidrofílica e hidrofóbica. También, se trabaja con un estándar interno; de manera que este ayuda a llevar un control de calidad en la detección de posibles compuestos. Por otro lado, en la metodología dirigida, se debe tener claro qué metabolitos se desea hallar para orientar la estrategia de identificación hacia esos componentes (Phelan, 2020).

En ambos casos se realizan dos pasos, el primero es un método de separación que siempre es un tipo de cromatografía (líquida o de gases). El segundo paso se lleva a cabo por una ionización (positiva y negativa) de los compuestos, donde se detectan las cargas y los números de masas respectivas. En el caso de realizar una técnica dirigida, se debe realizar la búsqueda múltiple sólo en iones de interés para la investigación que se lleve a cabo, con el fin de obtener resultados más finos (Majumder et al., 2021).

## Preprocesamiento

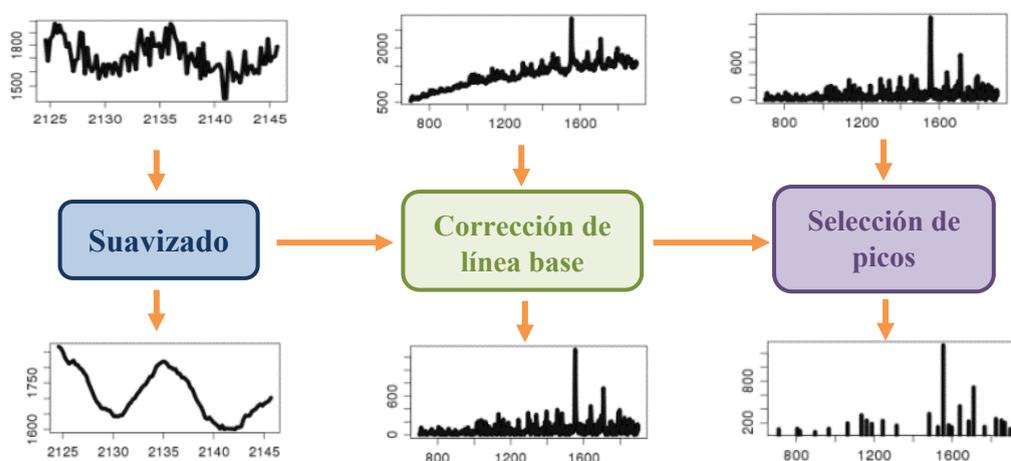


Figura 5. Pasos del preprocesamiento de datos (Bauer, Cramer, & Schuchhardt, 2011).

El preprocesamiento de datos es un paso crucial para el análisis de datos metabolómicos (ver Figura 5). Este está formado del suavizado de datos, con lo que se logra moderar el ruido de alta frecuencia (los picos muy altos serán eliminados y los similares, unidos) usando principalmente tres tipos de algoritmos: filtro Savitzky Golay, filtros Gaussianos, y técnicas basadas en ondículas. Dichos picos serán eliminados ya que los algoritmos tienen como propósito, disminuir las dimensiones exacerbadas de las

señales. El siguiente paso importante es la corrección de la línea base, que elimina el ruido de baja frecuencia e inherentemente la correlación de características cercanas. Este aplica filtros de espacio de Fourier, filtros derivados de Loess para poder calcular la línea base. Se continúa con la selección de picos que puede usar algoritmos muy comunes como la relación señal/ruido, la transformación de ondas continuas o funciones de modelo usando funciones Gaussianas. Esta etapa corresponde al filtrado de ruido y es importante para la normalización de datos. Se prosigue con la corrección del tiempo de retención, alineamiento de cromatogramas e identificación de compuestos (Bauer et al., 2011; Phelan, 2020; Spicer, Salek, Moreno, Cañueto, & Steinbeck, 2017).

Es importante, también, revisar la integridad de los datos, ya que cada equipo tiene distintas formas de exportar los archivos de salida. Sin embargo, estos deben ser compatibles entre los siguientes programas a usar. El programa usado debe dar los resultados con la corrección del estándar interno; de caso contrario, esta debe realizarse manualmente. En esta etapa, el último paso es el análisis estadístico multivariado que disminuye el número de variables en el estudio. También, da valores de importancia a los posibles compuestos, indicando si el modelo es predecible y sometiendo a los datos a diferentes validaciones estadísticas. (Phelan, 2020; Spicer et al., 2017)

Por otro lado, la ruta de trabajo en metabolómica dirigida consta de tres pasos: el preprocesamiento de datos, donde se limpia el ruido y se despeja la zona de interés, el análisis estadístico multivariado, y finalmente la cuantificación absoluta de las concentraciones de los metabolitos. A partir de este punto, los pasos a seguir son los mismos hasta llegar a la formación de redes metabólicas (Bauer et al., 2011).

Trabajos así requieren de la intervención de programas y plataformas especializadas en análisis metabolómicos que usan distintas estrategias para un objetivo específico. Algunos son amigables con el usuario y otros, sólo se manejan mejor por medio de comandos. Estos son interdisciplinarios, por lo que tienen un campo grande de aplicaciones.

### ***2.1.3.2. Aplicaciones***

Hoy en día la medicina moderna se enfoca, como se mencionó en párrafos anteriores, en la prevención efectiva y en terapias personalizadas. La metabolómica y determinadas ciencias ómicas son aplicables para este fin (Dąbek, Wojtala, Pirola, & Balcerczyk, 2020).

La medicina de precisión o personalizada es la rama encargada de abordar la inherente heterogeneidad de las enfermedades, pues un mismo desorden puede reflejarse de forma distinta en las personas. Este punto, precisamente, es intervenido por la metabolómica. Evalúa la suma significativa del genotipo, efectos de las interacciones ambientales y fisiológicas. De esta forma se buscarán biomarcadores para la detección anticipada de enfermedades y para la selección correcta del tratamiento, mostrando un antes y un después en el fenotipo (Hasin et al., 2017).

Es precisamente el diagnóstico temprano la fórmula del éxito para una terapia satisfactoria. Esto se hace posible con el descubrimiento de nuevos biomarcadores ya que tienen un panorama abierto para aplicaciones clínicas. Algunas de las investigaciones claves en el desarrollo de biomarcadores ha girado en torno a varios tipos de cáncer (ej. de mama, próstata, colon, estómago), enfermedades del sistema nervioso como Alzheimer y enfermedades metabólicas como diabetes (Klasic et al., 2018).

La diabetes ha sido ampliamente investigada y a través de herramientas metabolómicas pues es la mejor forma de abordar este tema. Se ha logrado identificar biomarcadores de factores de riesgo que lograron mejorar la predicción de la enfermedad. Además, ha sido gracias a estas herramientas que se pudo conocer un poco más de una enfermedad multifactorial que afecta aproximadamente a un 10% de mexicanos, según la Federación Internacional de Diabetes (FID) (Basto-Abreu, Barrientos-Guitérrez, Rojas Martínez, & Al., 2020).

## **2.2. Diabetes**

La diabetes mellitus (DM) es un trastorno metabólico caracterizado por un índice de glucosa en sangre elevado debido a diferentes causas; entre las que destaca la secreción anormal de insulina, la falla en los receptores celulares u otro tipo de factores como la obesidad, el sedentarismo y algunos estados clínicos (World Health Organization, 2019).

Hay varios tipos de DM, destacando las de tipo 1 y 2. Sin embargo, se describen otras que no son muy frecuentes, como diabetes gestacional y tipo M.O.D.Y. (*Maturity Onset Diabetes of the Young*, por sus siglas en inglés). El trastorno que antecede al cuadro patológico, donde la concentración de glucosa en sangre es elevada, pero no lo suficiente para un diagnóstico concreto, es conocido como prediabetes. Personas que se mantienen en este estado de forma persistente, tienen más probabilidades de desarrollar enfermedades cardiovasculares, así como terminar de desarrollar la enfermedad. (World

## 2.2.1 Tipos de diabetes

### *Tipo I*

También llamada “insulino-dependiente”, debido a una disminución de las células  $\beta$  de los islotes de Langerhans del páncreas, o su destrucción por procesos autoinmunes. Una vez que esto ocurre, la concentración de insulina se ve afectada ya que son las encargadas de su producción. En este caso, los pacientes presentan nula o muy poca concentración de insulina (Figura 6) (Janghorbani, Dam, Willett, & Hu, 2007).

### *Tipo II*

La DM tipo II, es producto de la interacción de factores genéticos, ambientales y del comportamiento. Es considerada la forma más común de este desorden debido a que más del 90% de los casos pertenecen a esta clasificación. La mayoría de la población en riesgo tienen condiciones médicas, incluyendo obesidad, hipertensión, hipercolesterolemia, hiperlipidemia y el síndrome metabólico (S. Park et al., 2015).

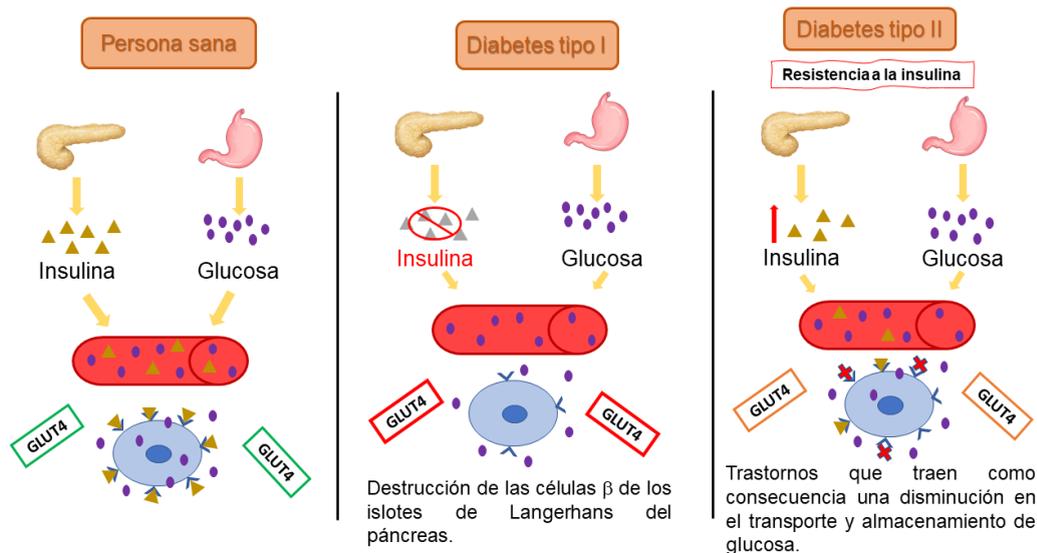


Figura 6. Metabolismo de una persona sana y personas con diabetes tipo I y II.

La desregulación en la respuesta metabólica de la insulina es la consecuencia de

fallas a nivel de la cascada de señalización. Así, se genera un ineficiente transporte de glucosa hacia hepatocitos, miocitos y adipocitos creando un estado de resistencia a la insulina (Figura 6) (S. Park et al., 2015).

En un estado de resistencia, el cuerpo se encuentra en un ayuno prolongado a pesar de haber ingerido alimentos. Esto se debe a que las células no pueden introducir la cantidad necesaria de glucosa debido a las fallas de su captación. De esta forma, los niveles de glucagón se elevan (hormona presente en estados de ayuno), provocando que el hígado comience a producir glucosa hepática a partir de fuentes gluconeogénicas como ciertos aminoácidos. Por otro lado, el páncreas al detectar un exceso de glucosa en circulación seguirá produciendo insulina. Y como esta concentración no disminuye, ocasionará una sobreproducción de insulina. Consecuentemente se generará la presencia de los característicos picos en la concentración de la hormona. Este panorama deteriorará la capacidad de producción del páncreas, y con el tiempo el órgano se atrofiará como consecuencia (Czech, 2017).

### **2.2.2 Epidemiología**

La diabetes es considerada una epidemia, lo que plantea una amenaza para la salud mundial. Según la Organización Mundial de la Salud (OMS), en el 2019 alrededor de 802 millones de adultos en todo el mundo tenían diabetes. Sin embargo, su prevalencia se ha duplicado desde ese año en el mundo. Aun así, las estadísticas reportadas se encuentran por debajo de la verdadera cifra global ya que, en las regiones de transiciones veloces de epidemiología, los números varían cada vez más rápido. Esto se debe principalmente a que el sedentarismo en niños y adultos es cada vez más difícil de erradicar (Hasin et al., 2017; World Health Organization, 2019).

México no es la excepción para las cifras internacionales, hoy en día la diabetes es la tercera causa de muerte, por debajo de las defunciones a causa de la pandemia por COVID-19 (Inegi, 2021).

### **2.2.3 Cuadro clínico**

La DM no presenta una sintomatología característica de la enfermedad. Es probable que el cuadro pase desapercibido los primeros años sin una revisión médica. Entre los principales síntomas destacan: sensación intensa de sed y hambre, necesidad de miccionar frecuentemente, pérdida de peso sin causa aparente, cansancio extremo, visión

borrosa, hormigueo en las extremidades, malestares estomacales y vómitos. El especialista de la salud debe conocer los síntomas y hábitos para sospechar y realizar el despistaje de diabetes mellitus. Un factor importante es la edad en la que se presentan todos los síntomas antes mencionados, para poder realizar un despistaje correcto de un tipo de diabetes mellitus (tipo I, tipo II, gestacional, M.O.D.Y.) (World Health Organization, 2019).

#### **2.2.4 Mecanismo**

La glucosa es el principal sustrato energético de la célula y para su ingreso se requieren proteínas transportadoras en la membrana celular. Esto se debe a que por su tamaño molecular y polaridad no pueden atravesar la membrana por difusión simple. Se han descrito dos sistemas de transporte de glucosa y de otros monosacáridos: los transportadores de sodio y glucosa llamados SGLT (transportadores sodio-glucosa) y los transportadores de glucosa llamados GLUT (transportadores de glucosa) (Lotfy, Adeghate, Kalasz, Singh, & Adeghate., 2017).

##### ***Transportadores SGLT***

Son proteínas de 14 hélices transmembranales que poseen ambos extremos en el espacio extracelular. Estas llevan a cabo un transporte acoplado, ya que también ingresa sodio y glucosa (o galactosa en algunos casos).

Se localizan en la membrana luminal de las células epiteliales encargadas de la absorción (intestino delgado) y la reabsorción (túbulo proximal) de nutrientes. Se aprovecha el ingreso de sodio a favor del gradiente electroquímico, entre el exterior e interior de la célula, para transportar la glucosa en contra de un gradiente químico (Lotfy et al., 2017).

##### ***Transportadores GLUT***

Los transportadores GLUT están encargados del ingreso de los monosacáridos a todas las células del organismo mediante difusión facilitada. Se han identificado catorce de ellos, enumerados desde GLUT 1 hasta GLUT 14. Presentan una conformación proteica similar, glicoproteínas de 45 a 55 kDa. Además, presentan un sitio de

glicosilación en la región externa de la membrana. Las isoformas de los GLUT tienen ubicación y características cinéticas propias, adaptadas a las necesidades metabólicas de los distintos tejidos del organismo. Estos transportadores se pueden agrupar en tres clases dependiendo de su secuencia homóloga y del asa de glicosilación (Vargas & Carrillo Sepulveda, 2019).

La ubicación de los transportadores depende de la necesidad de insulina de los tejidos para poder ingresar glucosa a la célula. Los GLUT1, se localizan en los eritrocitos; GLUT2, en el hígado; GLUT3 en cerebro. Estos últimos no dependen de la hormona para poder efectuar el transporte de la glucosa dentro de la célula. Los GLUT4 se encuentra exclusivamente en tejido adiposo y músculo, y estos responden a insulina (Vargas & Carrillo Sepulveda, 2019).

### **2.2.5 Proteína quinasa activada por AMP**

La proteína quinasa activada por AMP (AMPK) es un sensor del estado de energía, altamente conservado en células eucariotas. Consta de tres subunidades ( $\alpha$ ,  $\beta$ ,  $\gamma$ ), que poseen múltiples isoformas. Tienen una expresión de una isoforma específica según el tejido célula (Hardie, 2014).

Cuando las células se encuentran estables, la concentración de AMP es muy baja debido a que la alta relación de ATP:ADP favorece la reacción reversible de adenilato quinasa hacia la síntesis de ADP (Ecuación 1). Cuando las células entran en estrés metabólico (debido a la escasez de glucosa, oxígeno, o debido al consumo acelerado de ATP, se genera un aumento en la relación en la ADP:ATP, por lo que la reacción antes mencionada favorecería a la síntesis de AMP. Estas son las señales que activan a AMPK y sus funciones (Hardie, 2014).

*Ecuación 1*



El AMP es un intermediario regulador en varios procesos metabólicos, ya que es capaz de fosforilar proteínas para activarlas o desactivarlas. También está relacionado con los transportadores GLUT4, por lo que se dice que tiene un papel imprescindible en la regulación de utilización de glucosa (Hardie, 2014).

## 2.3. Modelo animal de diabetes

Este trabajo retoma el estudio sobre el impacto de bebidas funcionales en el metabolismo de ratas obesas (Rubio-Rodríguez, Reynoso-Camacho, Rocha-Guzmán, & Salgado, 2021).

En la investigación citada, se emplearon 5 tipos de plantas con propiedades antiobesogénicas y antidiabéticas en diferente concentración para la elaboración de las bebidas funcionales. Se crearon tres tratamientos que fueron resumidos en la Tabla 1. Adicionalmente, todas las bebidas contenían: acesulfamo, sucralosa, ácido cítrico y sorbato de potasio.

Tabla 1. Composición de bebidas funcionales.

ID	Jamaica roja *	Jamaica blanca *	Té verde**	Canela**	Tamarindo de malabar**	Menta*
Roja	2	-	1.125	0.75	0.8	-
Blanca	-	2	1.125	1.0	1.125	-
Verde	-	-	1.5	-	1.5	0.5

\*Valores en porcentaje.

\*\*Valores en g/L.

Para el tratamiento se usaron ratas macho tipo *Wistar* conforme al protocolo de cuidado y uso de Animales de la Universidad Autónoma de Querétaro, como se recomienda en la norma (NOM-062-ZOO-1999).

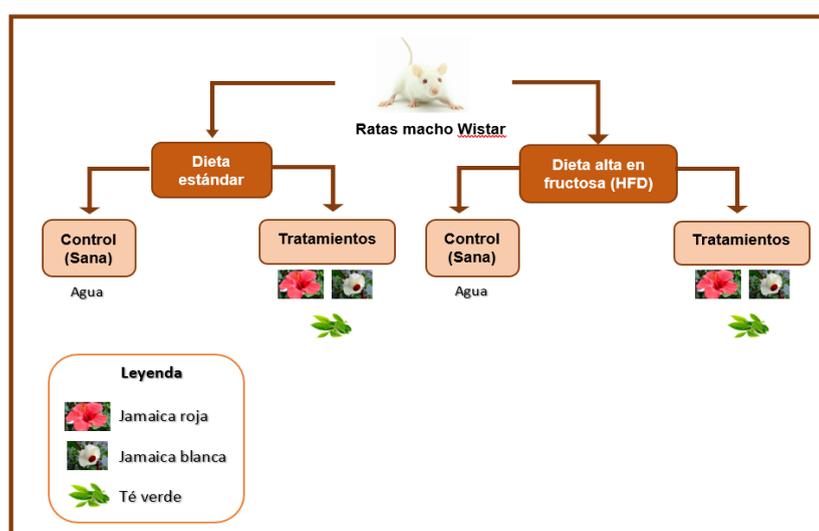


Figura 7. Diseño experimental de animales sanos y alimentados con una dieta alta en grasa y fructosa y tratados con bebidas.

Se trabajaron en grupos de 8 animales, donde el grupo control negativo consumió alimento estándar y al grupo control positivo se le administró una dieta alta en grasa y fructosa durante 19 semanas y 3 de estos grupos fueron tratados con las bebidas. De esta forma, se produjo obesidad en los animales de experimentación (Figura 7).

Luego del tiempo mencionado, se recolectó la sangre por punción cardiaca, la cual fue tratada por separación mecánica y química para poder ser inyectada al equipo. De esta forma, fue analizada como se menciona a partir de la sección 4.1.1 más adelante.

# Capítulo 3

## Justificación

La bioinformática tomó mayor importancia gracias a la genómica, pues vino acompañada de una constante innovación en métodos y técnicas de laboratorio. Esto ocurrió con el fin de mejorar la capacidad y sensibilidad de los análisis. Años más tarde, con la implementación de la biología de sistemas y las ciencias ómicas, se hizo más grande la brecha entre la interpretación de metadatos y las metodologías existentes. La rama de la biología de sistemas que más ha presentado una falta de estándares ha sido la metabolómica.

La metabolómica es una ciencia ómica que estudia todos los metabolitos de un organismo, un tejido o una célula, como respuesta a un estímulo. De esta forma se describe y analiza las características implícitas de un determinado fenotipo. Usualmente, se genera una gran cantidad de datos, por lo cual su análisis e interpretación es una tarea compleja. Para esto, se desarrollaron técnicas y metodologías que los laboratorios usan de forma exclusiva o combinada, causando una falta de estandarización.

Las rutinas que hacen uso de softwares que requieren la adquisición de licencias; y también de plataformas o servidores de uso abierto. Se desea analizar qué tan precisos son este tipo análisis en un estudio metabolómico de un modelo animal de obesidad y resistencia a la insulina. Esta enfermedad, al ser un desorden metabólico importante, generará variabilidad en el metabolismo. Se considera que la metabolómica es la mejor vía para el estudio de este trastorno, por lo que sería posible la detección tanto de los metabolitos afectados, así como de las rutas metabólicas que presenten una variación estadísticamente significativa. Así mismo, es importante determinar la capacidad de las bebidas para regular dichas vías metabólicas y con ello identificar y proponer posibles mecanismos de acción.

# Hipótesis

Dado los reportes sobre el efecto hipoglicemiante de plantas terapéuticas, entre las que destacan la flor de Jamaica (*Hibiscus sabdariffa*) y el té verde (*Camellia sinensis*), es posible determinar el mecanismo de acción molecular de bebidas funcionales con el uso de metodologías bioinformáticas y metabolómicas.

# Objetivo general

Proponer un mecanismo molecular de acción para el efecto hipoglicemiante de bebidas funcionales hechas a base de flor de Jamaica (*Hibiscus sabdariffa*) o de té verde (*Camellia sinensis*).

## Objetivos específicos

- ✓ Recolectar datos metabolómicos de los diferentes grupos de animales.
- ✓ Analizar y comparar los datos recolectados de los metabolomas estudiados.
- ✓ Integrar, en un solo modelo, la información de las diferentes rutas metabólicas afectadas.
- ✓ Proponer un mecanismo de acción a partir de los nodos regulatorios identificado.

# Capítulo 4

## Metodología

### 4.1 Equipo de laboratorio y detalles computacionales

#### 4.1.1 Equipo:

Las muestras de suero de 56 inyecciones fueron analizadas en un equipo de cromatografía líquida de ultra resolución (UPLC) modelo *Acquity* clase I con detector PDA acoplado a un espectrómetro de masas tándem (MS-MS) (Waters, Vion IMS QTOF). Se usó una columna BEH C8 (2.1 X 100 mm, 1.7  $\mu\text{m}$  partícula)

#### 4.1.2 Software:

**UNIFI (WATERS).** Plataforma de análisis que fusiona datos de cromatografía líquida (LC) y espectrometría de masas (MS). Este programa es el encargado de realizar la corrección de masas en base a un estándar interno. Tiene código privado, por lo que se requiere licencia para su uso.

**PROGENESI QI (WATERS).** Programa de análisis de datos LC-MS que realiza estudios estadísticos, cuantifica e identifica compuestos pequeños en base a rutinas cargadas predeterminadamente en sus bases de datos. También tiene código privado y requiere licencia.

**OPENMS 2.0.** Es un programa multiplataforma robusto y de código abierto (no requiere de licencia) para el análisis de datos de espectrometría de masas. Fue programado en Python y C++ (Röst et al., 2016).

**PROTEOWIZARD 3.0.** Convertidor de una amplia gama de formatos. De código libre con distribución Linux y para Windows. Se puede elegir entre la versión gráfica y la manejada por medio de comandos (Chambers et al., 2012).

### 4.1.3 Servidores:

**METABOANALYST 5.0.** Es un servidor online de procesamiento de datos metabolómicos. Se pueden realizar diferentes tipos de análisis estadísticos, identificación de metabolitos, mapeo de rutas, agrupaciones jerárquicas, entre otras (Xia, Psychogios, Young, & Wishart, 2009).

**METLIN.** Es una base de datos online donde se encuentran anotaciones de espectrometría de masas de una lista muy grande de compuestos (Smith et al., 2005).

**HMDB 4.0.** Es una base de datos libre que alberga información detallada de metabolitos que fueron hallados en el cuerpo humano, tanto endógenos como exógenos (Wishart et al., 2018).

## 4.2 Métodos

### 4.2.1 Exportación de datos

Se exportaron los datos metabolómicos del equipo mencionado en la sección 4.1.1, Con la ayuda de UNIFI (*Waters*), pues este es el primer software con el que tiene contacto en la metodología.

Este momento es donde se realizó la corrección de masas con respecto al estándar interno del equipo. Se inyectaron 10  $\mu\text{L}/\text{min}$  de leucina encefalina a una concentración de 200  $\text{pg}/\mu\text{L}$ . Se exportaron los datos correspondientes a la ionización negativa en dos formatos: el que es nativo de la empresa Waters (*.UEP*) y uno abierto o comercial (*.mzml*, escrito en minúsculas, indicando un error; pues el correcto es *.mzML*).

### 4.2.2 Conversión de formato

1. Se realizó en el sistema operativo Linux para comodidad del uso de comandos y visualización de datos. Se usó el software Proteowizard 3.0 para convertir los datos abiertos (*.mzml*) a su formato correcto (*.mzML*). Para esto, primero se visualizaron los datos en el editor de textos, se comprobó un encabezado y una señal de finalización. Seguido, se realizó la conversión en el programa mencionado usando el siguiente comando:

```
msconvert input.mzml --mzML -o output_directory
```

Donde: *msconvert* es la forma en la que se da la orden para llamar al programa en distribución Linux, *input.mzml* es el archivo de entrada, *--mzML* es la extensión en la que queremos que se obtengan los archivos resultantes, y *output\_directory* es donde queremos que se guarden los resultados.

2. Seguido a esto, se comprobó la integridad de los datos en el editor de textos ofrecido por LINUX, pues nos permitió comparar el archivo inicial y el archivo final fácilmente mediante el uso de comandos. Sin embargo, al intentar abrir los archivos en un visor de metadatos, estos no pudieron ser abiertos ni visualizados.

### 4.2.3 Cambio de MS2 por MS1

1. Se analizaron los niveles de masas (MS1, MS2) y al no hallar MS1 en los archivos, se los introdujo en cada archivo. Se filtró las zonas donde se hallaba lo siguiente:

```
<Param accesión="MS:1000511" cvRef="MS" name="ms level" value="2"/>
```

Y se sustituyó con lo siguiente:

```
<Param accesión="MS:1000511" cvRef="MS" name="ms level" value="1"/>
```

Seguidamente, se procedió a verificar la integridad en visores para datos metabolómicos para asegurar que dichos archivos contaran con las partes básicas requeridas por los programas para poder ser leídos. Si dichos archivos no cuentan con una señal de apertura o de finalización, un programa o plataforma no podrá leerlo, dando un mensaje de "ERROR".

### 4.2.4 Preprocesamiento de datos

1. Esta etapa fue realizada en PROGENESIS QI. El primer paso fue importar los datos desde el programa UNIFI para realizar la revisión de la alineación. Este es un paso importante en el análisis estadístico de los datos ya que permite que la selección de picos o *Peak picking* sea un proceso óptimo y rápido. Además, abre paso a un análisis estadístico robusto y confiable. Una vez que finalizó esta etapa, se dio click en "*Section complete*"
2. El siguiente paso realizado fue *Peack picking*, por el cual PROGENESIS QI localizó posibles compuestos o *features* en las muestras. Finalizado el proceso, se dio click en "*Section complete*"

3. Luego, se continuó con el proceso de deconvolución. En este paso los iones hallados en el paso anterior y los aductos se agrupan en compuestos. Se analizó la integridad de los cromatogramas y espectros para seleccionar cada señal o *feature* en: Pico (cromatograma y espectro correctos), No pico (espectro correcto, cromatograma con fallas), No patrón (cromatograma correcto, espectro fallado) y No pico/No patrón (cromatograma y patrón con fallas). Este paso sirvió como un filtro de los datos, pues de esta forma las señales situadas en el último grupo fueron clasificadas como ruido. Al finalizar este proceso se dio click en “*Section complete*”.

### 4.2.5 Análisis estadístico

Los análisis estadísticos se realizaron en la plataforma de análisis metabolómicos METABOANALYST 5.0, donde se llevaron a cabo estudios multivariables.

1. Se exportaron los datos resultantes de revisión de la deconvolución desde PROGENESIS QI, en formato *.csv*. Luego, se editó para obtener un archivo en el que los tratamientos se situaron en las filas y la relación carga masa en las columnas.
2. Se realizó una limpieza extra de los datos, pues se observó una gran cantidad de valores perdidos (NA). Se obtuvo el promedio de NA's por cada tratamiento y se eliminaron los que superaban este porcentaje. Adicionalmente, se observó una gran cantidad de señales que no mostraron valores en las columnas por lo que se decidió eliminarlas. De esta forma, se logró un análisis más estricto y limpio.
3. Seguido, se ingresó al portal de METABOANALYST 5.0 y se seleccionó la pestaña correspondiente a “Análisis estadístico”. Se seleccionó la opción “Concentración” en tipo de datos, en “Formato” se eligió las muestras en filas y no pareado. Luego se cargó el archivo en formato *.csv*. Inmediatamente, se da “*Submit*” para enviar el trabajo a servidor.
4. A continuación, se mostraron especificaciones sobre la integridad de los datos, donde se nos indicó la cantidad de valores perdidos, porcentaje y demás. En esta pantalla, se dio click en “*Missing value*” para realizar un tratamiento a estos valores.
5. Se retiró el 31% de valores perdidos (NA) y se excluyeron todas las variables con estos mismos valores. Luego, se dio click en “*Process*”.
6. El siguiente paso que es el filtrado de datos, donde se seleccionó la opción “Ninguna” o “*None*” ya que por el número de *features* que manejamos no es recomendable filtrar datos. Este número se obtiene del total de señales en el archivo

- .csv. Se dio click en “*Proceed*”.
7. Se prosiguió con la normalización de datos. Se seleccionó la opción de normalizar los datos por su suma (*By sum*), en la transformación de datos se eligió un proceso logarítmico (*Log transformation*) y en el escalado de datos se seleccionó la opción de Pareto (*Pareto scalling*). Se dio click en “*Normalize*”. Seguido a esto se da en “*Proceed*”.
  8. Se obtuvieron los resultados estadísticos del método multivariable de PCA y al observar que los tratamientos no se separaron en base a estos, se analizó el diagrama PLSDA (*Partial least squares-discriminant análisis*). Dada la naturaleza de los datos, se decidió realizar una validación de los datos según el modelo PLSDA ortogonal. Con este fin, se formaron grupos de parejas entre los tratamientos en archivos .csv y se inició el procedimiento nuevamente.
  9. A continuación, se validaron los datos en el modelo seleccionado por permutación, teniendo en cuenta que el *p value* obtenido sea menor a 0.005.
  10. De los grupos que pudieron ser validados, se procedió a buscar los posibles *features* a través del puntaje VIP. El corte se realizó en 3.0 con el fin de ser más restrictivos en el análisis estadístico.
  11. Se identificaron los *features* que cumplieron el valor de corte VIP, ubicándolos en la tabla de formato .cvs para identificar su relación carga / masa y tiempo de retención. Esta relación fue anotada y guardada para continuar con la identificación.

#### 4.2.6 Identificación de metabolitos

1. Este procedimiento fue llevado a cabo en PROGENESIS QI. Esto fue después que los *features* fueran localizados junto a sus valores de carga / masa y tiempo de retención (RT). En la sección de identificación de metabolitos, se buscó y se marcó estos *features* con un nuevo identificador. Seguido, fueron filtrados en base a este para trabajar sólo con ellos.
2. Luego se seleccionó a “*Chemspider*” como el método de identificación, pues este compara su base de datos con más de 32 millones de estructuras. Además, alberga información de otras bases de datos como HMDB.
3. En la sección donde se indican los filtros, se cambió de 10 ppm a 5 ppm; con el fin de hacer la identificación más sensible. También se seleccionó el set llamado “*Metabolomics*”, pues en esta rutina se encuentran algunos parámetros ya definidos

para los trabajos de laboratorio del grupo de investigación. Seguido a esto, se dio click en “*Search for identifications*”.

4. Una vez obtenidos los resultados, se mostró una lista con posibles opciones por *feature*, con un score y un puntaje de fragmentación. Se eligieron los compuestos con un score mayor a 40 y un puntaje de fragmentación alto, principalmente.
5. Los compuestos seleccionados fueron buscados en las diferentes bases de datos (ChEMBL, HMDB, Kegg Database, etc) para confirmar sus valores de tiempo de retención, peso molecular, relación m/z y también se recurrió a observar su espectro de masas para confirmarlo.
6. Se anotaron los códigos de acceso de cada compuesto, tanto el principal como el secundario.

#### **4.2.7 Construcción de redes metabólicas.**

Esta parte de la metodología se llevó a cabo en la plataforma de METABOANALYST 5.0.

1. Se accedió a la plataforma de forma online y se seleccionó la opción de “*Click to start*” para empezar y en el menú principal se seleccionó la opción de “*Pathway analysis*”.
2. Se seleccionó el tipo de entrada y por cuestiones de practicidad, se eligió “*Kegg ID*”. Luego se introdujo manualmente los códigos ID de cada compuesto de esta base de datos en una columna y se dio click en “*Submit*”.
3. En el paso siguiente, se seleccionó en la opción de “*Pathway library*” en mamíferos, la opción de “*Rattus norvegicus*”. Finalmente se subió el trabajo a la plataforma dando click en “*Submit*”.

Finalmente, se prosiguió con el análisis e interpretación de los resultados obtenidos hasta el momento.

# Capítulo 5

## Resultados y Discusión

### 5.1. Exportación de datos

Se logró exportar datos de un total de 56 inyecciones, correspondiente a la ionización negativa (ESI -) en dos tipos de formato: *.UEP*, el cual es una extensión nativa de Waters; y el formato libre o abierto *.mzml* por el programa UNIFI.

### 5.2. Cambio de formato

Se modificó exitosamente el tipo de extensión de los archivos con formato *.mzml* a *.mzML* usando el convertidor online PROTEOWIZARD, sin embargo, estos no pudieron ser analizados ni visualizados por otras herramientas bioinformáticas debido a incompatibilidades entre plataformas.

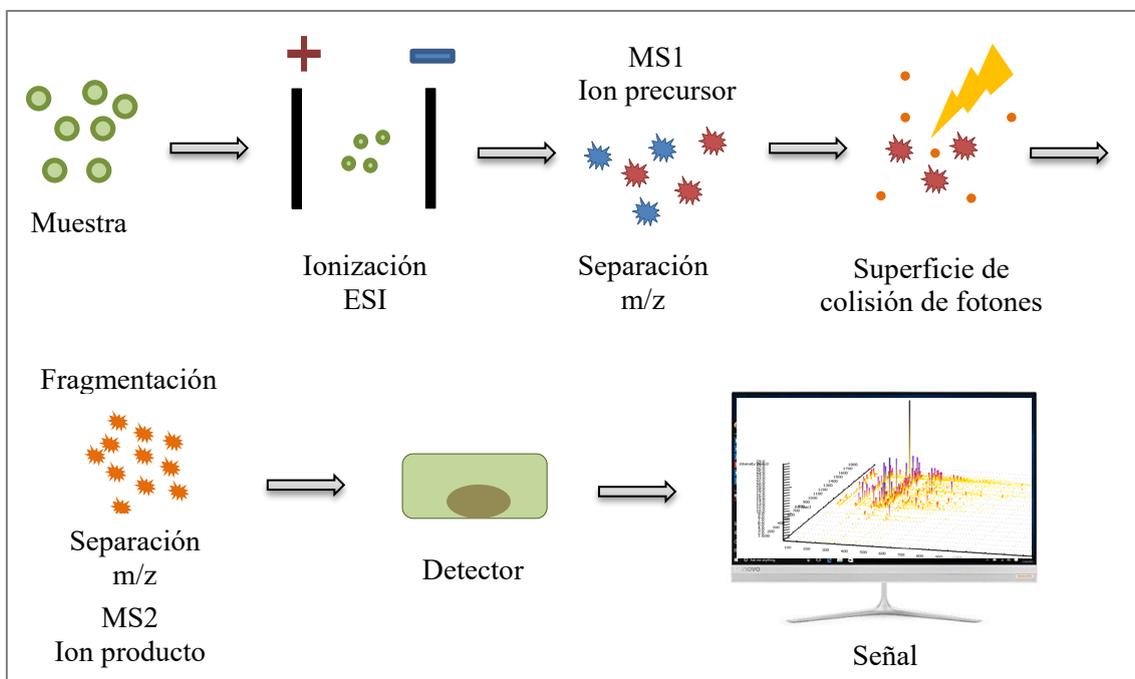


Figura 8. Diagrama de espectrometría de masas (MS/MS)

Al analizar la estructura de cada archivo se notó que estos no contenían valores MS1, los cuales son imprescindibles para un análisis metabolómico ya que brindan

información sobre los primeros precursores del compuesto padre. Por esta razón, no fue posible visualizarlos a través de programas especializados como los mencionados en la sección de metodología, ya que nuestros archivos sólo contenían valores de MS2. Este fenómeno se explica desde el funcionamiento del equipo, pues el UPLC/MS-MS; luego de la primera ionización, genera iones que son precursores inmediatos de una relación carga masa ( $m/z$ ), es decir MS1. Estos son seleccionados y fragmentados nuevamente, generando otros iones que son detectados como MS2. De la misma forma el proceso puede alargarse como MS3, MS4; dependiendo de la metodología que se siga (Figura 8).

Al investigar la rutina de trabajo del equipo con el que se trabajó, se halló la metodología que genera MS1 y MS2. Fue así como se decidió editar, de forma deliberada, todos los archivos para poder brindarle al programa un valor MS1 que nos permitiera abrirlos y visualizarlos. La visualización se realizó en el programa OPENMS 2.0, pues nos permitió observar y analizar aspectos sobre los espectros de masas de cada inyección. Sin embargo, los estadísticos iniciales no pudieron ser llevados cabo por incompatibilidades de versiones y formatos de UNIFI (Röst et al., 2016).

A pesar de que se reportan plataformas que permiten trabajar con conjuntos de datos que incluyan sólo valores MS2, no es este el caso. Esto se debe a que los valores MS1 están destinados a realizar identificaciones putativas, principalmente debido al tipo de información que aportan al estudio (Majumder et al., 2021).

Los datos con información de MS1 y MS2 son más complejos, pero generan más datos. Con los primeros obtendremos información detallada acerca del ion precursor ( $m/z$ , RT, tiempo de vuelo), mientras que con los segundos se analizarán principalmente las propiedades físicas de los iones producto que sean detectables por el equipo. De esta forma, el contar sólo con un tipo de datos, limita en gran parte el análisis.

Perez-Riverol & Moreno en su publicación del 2020 hablan sobre el escalamiento de datos en metodologías de metabolómica, proteómica y su condicionamiento implícito. En dicho estudio, se hace mención sobre la recurrente limitación de conversión e intercambio de formatos donde se explica las consecuencias de conversiones entre formatos privados y abiertos: ralentización del desarrollo, demanda de mantenimiento, introducción de errores o pérdida de datos (Perez-Riverol & Moreno, 2020).

### **5.3. Preprocesamiento de datos**

Se exportaron un total de 17510 señales desde UNIFI hacia PROGENESIS QI

provenientes de las 56 inyecciones que fueron alineados con respecto a una muestra representativa. Luego se revisó la alineación, pues como se mencionó en el Capítulo 2, es en este paso donde se lleva a cabo tres procesos principales que permiten un posterior estadístico robusto.

Para visualizar las señales detectadas antes y después de la alineación, se muestra la Figura 9 correspondiente al cromatograma de iones totales. En esta figura se observa en color verde el cromatograma correspondiente a la alineación con respecto a la muestra representativa, la que fue elegida en forma aleatoria. Usualmente esta alineación se realiza en base al QC (*Quality control*), el cual es una muestra formada a partir de una mezcla de una muestra por grupo de tratamiento. De esta forma se asegura que el análisis contará con todos los metabolitos presentes en los controles y en todos los tratamientos, jugando el papel de estándar interno. Sin embargo, en este estudio no se contó con la presencia de uno. A pesar de esto, se observan bajas fluctuaciones entre picos (Begou et al., 2018).

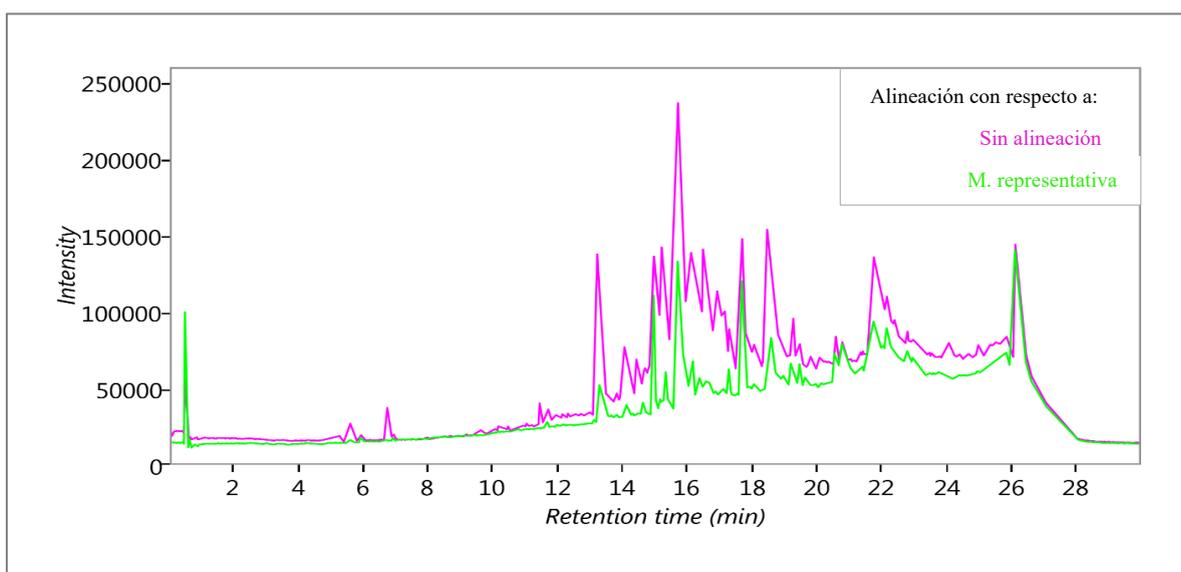


Figura 9. Cromatograma de iones totales.

En color magenta se ven las señales de todas las inyecciones del estudio, sin alineación alguna. La calidad del alineamiento nos da una idea de cuán confiable y robusto será nuestro análisis estadístico. Al comparar ambas señales en dicha figura, se ve claramente una disminución en la intensidad. Dicho fenómeno ocurre gracias al alineamiento que el programa realiza respecto a la muestra elegida, ayudando a disminuir el ruido de fondo. En este paso se realizan los principales pasos del preprocesamiento de

los datos: suavizado y corrección de la línea base, ya que la selección de picos es un paso especial en PROGENESIS QI (ver Figura 5).

En el proceso de selección de picos o *Peak picking*, se hallaron un total de 14810 *features* lo cual nos indicó la eliminación de un total de 2700 señales que fueron clasificadas como ruido. Explicando el por qué en la Figura 9 se observan picos distintos entre la muestra de referencia entre el minuto 13 y el minuto 20. Adicionalmente, se visualizan señales antes del minuto 0.7 y después del minuto 26; mismas que fueron descartadas proceso pues corresponden a señales del equipo correspondientes al lavado de la columna.

En el momento de revisar la deconvolución, se reportó un número final de 2440 *features* aptos para identificar como posibles metabolitos. Para este proceso, los aductos son los responsables de brindar información teórica acerca del compuesto o señal detectada.

La disminución del número de *features* se debe a que luego de haber captado las señales, el programa los agrupa en compuestos usando la información experimental de aductos que se ingresó al estudio. Para que una señal y un aducto sean considerados parte de un mismo compuesto, deben presentar el mismo tiempo de retención (RT) y la masa de sus iones pueden diferir en una cantidad igual a la diferencia de dos aductos incluidos en la rutina experimental. Entonces, habrá muchas señales de aductos detectadas por el equipo de forma repetida, por lo que el proceso de deconvolución los detecta y si forman parte de un compuesto los conserva. De lo contrario, son eliminados y clasificados como ruido. Esta es la principal razón por la que el número de *features* disminuye drásticamente en este paso (Erngren et al., 2019).

## 5.4. Análisis estadístico

Se analizaron los datos luego de haber atravesado por un control de calidad para evaluar integridad y eliminar el ruido de fondo, el cual será explicado a continuación. Las principales causas que en un inicio se hiciera imposible analizarlos fue la presencia de un número significativo de datos vacíos y remanentes del ruido de fondo.

Al realizar una revisión de la integridad del archivo en Excel, se halló un promedio de 35% de valores perdidos que causaron que procedimientos como la normalización y escalado no fueran llevados a cabo con éxito. Los tratamientos que superaron este valor fueron separados del estudio ya que, al tener una cantidad grande de éstos, se añaden

falsos positivos al estudio. Finalmente, el ensayo se llevó a cabo con 31 inyecciones de 56. La principal causa de la merma en el grupo se debe a que ratas obesas que consumieron la bebida funcional “Verde” fue eliminado de forma íntegra debido a que todas las inyecciones superaron con gran ventaja el porcentaje promedio establecido de valores perdidos. De haber continuado con el grupo de estudio completo, los pasos posteriores no hubieran sido posibles, pues cada procedimiento debe ser validado estadísticamente para seguir trabajando con determinados datos.

Como se hace mención en el artículo de Wei en el 2018, el proceso elegido para la imputación de valores perdidos es clave del desarrollo del proceso estadísticos. Sin embargo, se menciona también que la cantidad obtenida de estos habla sobre la calidad del análisis. Ya que, con un número creciente de valores perdidos, la posterior imputación, los análisis estadísticos y la identificación se verán afectados. Tomando en cuenta esta premisa, de no haber filtrado un porcentaje de valores perdidos, el estudio se hubiera visto afectado (Wei et al., 2018).

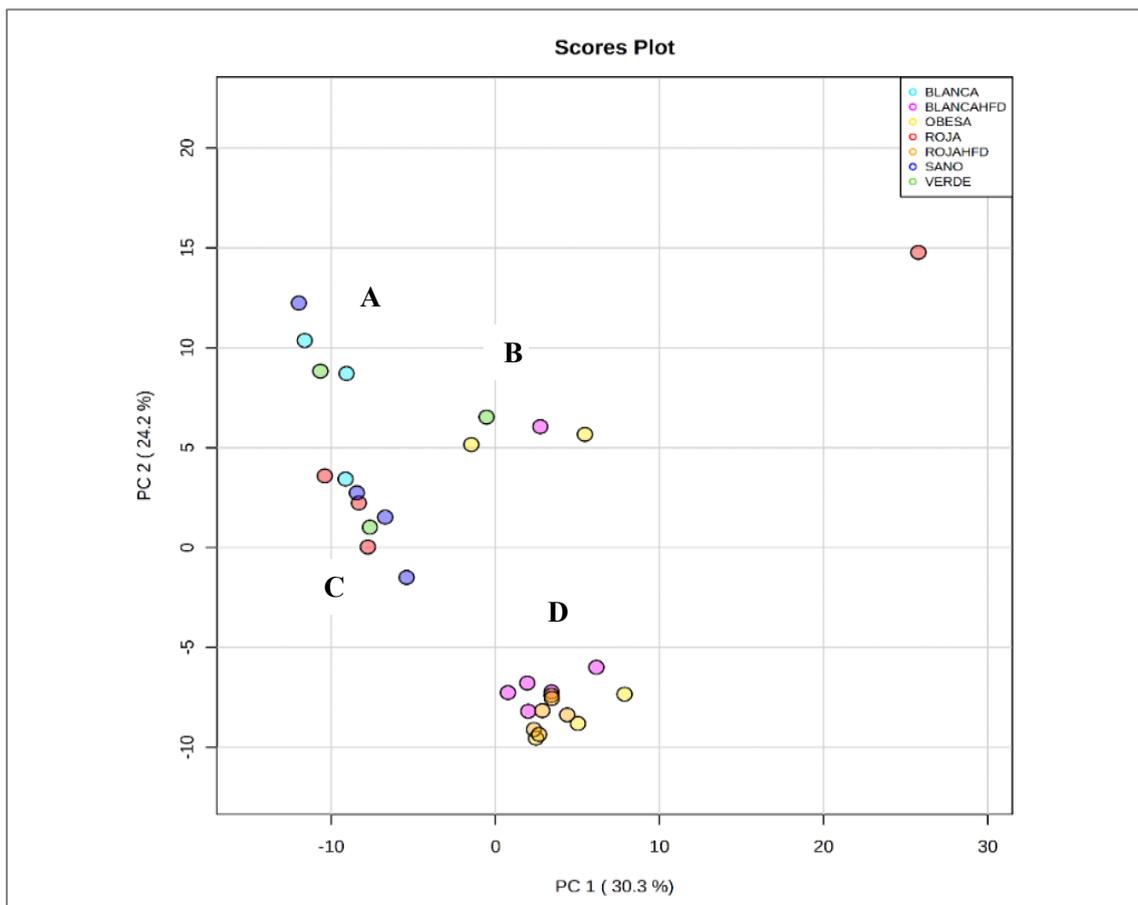


Figura 10. Diagrama PCA de set de datos filtrados.

El análisis de compuestos principales (PCA) del ensayo general se muestra en la

Figura 10. Es un método exploratorio y es capaz de reducir las dimensiones de las variables de respuestas. Es decir, por su naturaleza multivariable tiene la capacidad de normalizar las intensidades que saltan de los máximos establecidos. De esta forma, fue factible observar el efecto de los tratamientos.

Se formaron 4 grupos principalmente: A, B, C y D; donde las dos últimas fueron las más pobladas (Figura 10). A pesar de no haberse agrupado por tratamiento estrictamente, estos mostraron cierta semejanza. Brindando la idea que los tratamientos se agrupaban por similitud al estado metabólico. Sin embargo, esto no ocurrió en el grupo “B”, ya que hallamos animales que consumieron dieta estándar y dieta alta en grasa y en fructosa; descartando la teoría. También se observó la presencia de una muestra perteneciente al grupo “roja” que se encontró muy alejada de todos los demás; considerándole como un dato (o conjuntos de datos) que escapa(n) de la media del estudio permaneciendo hasta un cuadrante distantes de los grupos de datos (Li, Shao, & Fu, 2018).

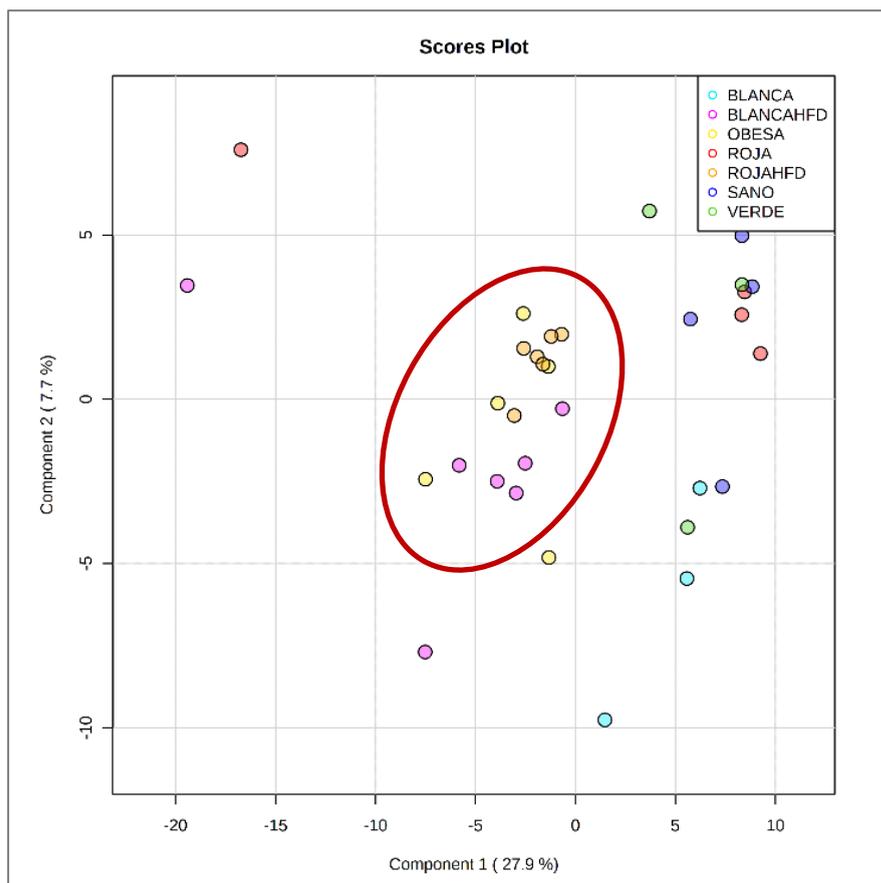


Figura 11. Diagrama PLSDA de set de datos filtrados.

Al analizar de forma integral la Figura 10, se muestra en un inicio que el estudio

los datos no mostró una diferencia estadísticamente significativa, ya que se esperaba que los grupos se separen en base al tratamiento. Esto se confirma al observar el comportamiento de los grupos control (obeso y sano), pues no se separan de forma contundente.

Este fenómeno indicaría la presencia de un ruido de fondo significativo, falta de sensibilidad en la metodología experimental o errores en los procesos de normalización y estandarización, lo que arrastraría consigo ruido constante a lo largo de todo el estudio.

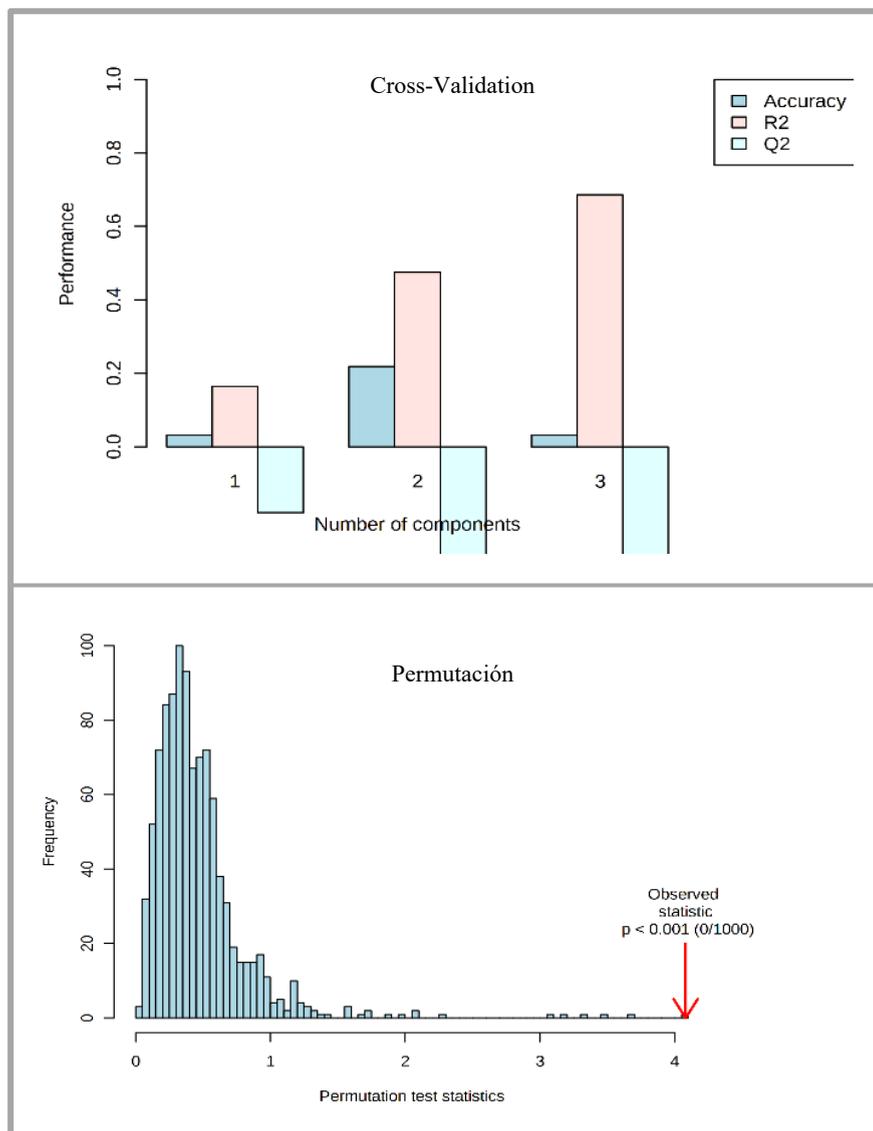


Figura 12. Test de validación de set de datos filtrados.

Sin embargo, estas especulaciones se descartan al observar las siguientes pruebas estadísticas mostradas a continuación. Puede ocurrir que un modelo estadístico cargado predeterminadamente en plataformas como METABOANALIST 5.0 no se ajusten muy

bien al modelo biológico que se maneja en el presente estudio.

Como lo mencionó Antonelli y colaboradores en su artículo referido a rutinas estadísticas de trabajo para metadatos, todo resultado biológico sometido a distintos modelos estadísticos varía según lo que sea de interés investigar de esos datos en específico. Siendo así, que algunos no generarán la respuesta deseada frente a un mismo modelo a pesar de modificar los parámetros de normalización y escalado (Antonelli et al., 2019).

Plataformas bioestadísticas en su versión gráfica pueden estar limitadas en cuanto a la flexibilidad de las rutinas que se llevan a cabo. Esto se debe principalmente a que ya vienen programadas y en ocasiones se requieren pasos adicionales. Dicha limitación quedó expuesta desde el momento en el que se requirió que los datos atravesaran por un control de calidad extra, ya que el programa que eliminó el ruido de fondo no tomó en cuenta los valores perdidos por tratamiento.

*Tabla 2. Grupos formados para PLSDA - ortogonal.*

<b>GRUPOS FORMADOS</b>		<b>VALOR “p value”</b>
<i>Blanca + HFD</i>	<i>Blanca</i>	<i>0.005</i>
<i>Blanca + HFD</i>	<i>Obesa</i>	<i>0.733</i>
<i>Obesa</i>	<i>Sana</i>	<i>0.015</i>
<i>Obesa</i>	<i>Verde</i>	<i>0.15</i>
<i>Roja + HFD</i>	<i>Blanca + HFD</i>	<i>0.154</i>
<i>Roja + HFD</i>	<i>Obesa</i>	<i>0.224</i>
<i>Roja + HFD</i>	<i>Roja</i>	<i>0.003</i>
<i>Sana</i>	<i>Roja</i>	<i>0.918</i>
<i>Sana</i>	<i>Verde</i>	<i>0.609</i>
<i>Sana</i>	<i>Blanca</i>	<i>0.878</i>

Por otro lado, en el análisis PLSDA (Figura 11) se muestra una agrupación conformada por blanca + HFD, roja + HFD y obesas; aunque de forma dispersa. Los

grupos observados en la Figura 10 se mostraron disociadas en la Figura 11, confirmando la presencia de ruido. Esto indica que la fuerza de atracción que las mantenía unidas en el diagrama PCA no es lo suficientemente fuerte como para mantenerse en el correspondiente al PLSDA. Así se resta importancia a la hipótesis sobre la acción del tratamiento y al estado metabólico, a lo cual se le atribuía la agrupación de grupos. Al ser más restrictivos en la eliminación de ruido y en la limpieza de tratamientos y de señales, se obtuvo el mismo resultado.

Al validar los datos por los métodos ofrecidos en la plataforma de METABOANALYST, se obtuvo un *p value* de 0.001 a través de la prueba de permutación. Esta se valida obteniendo un valor menor a 0.005.

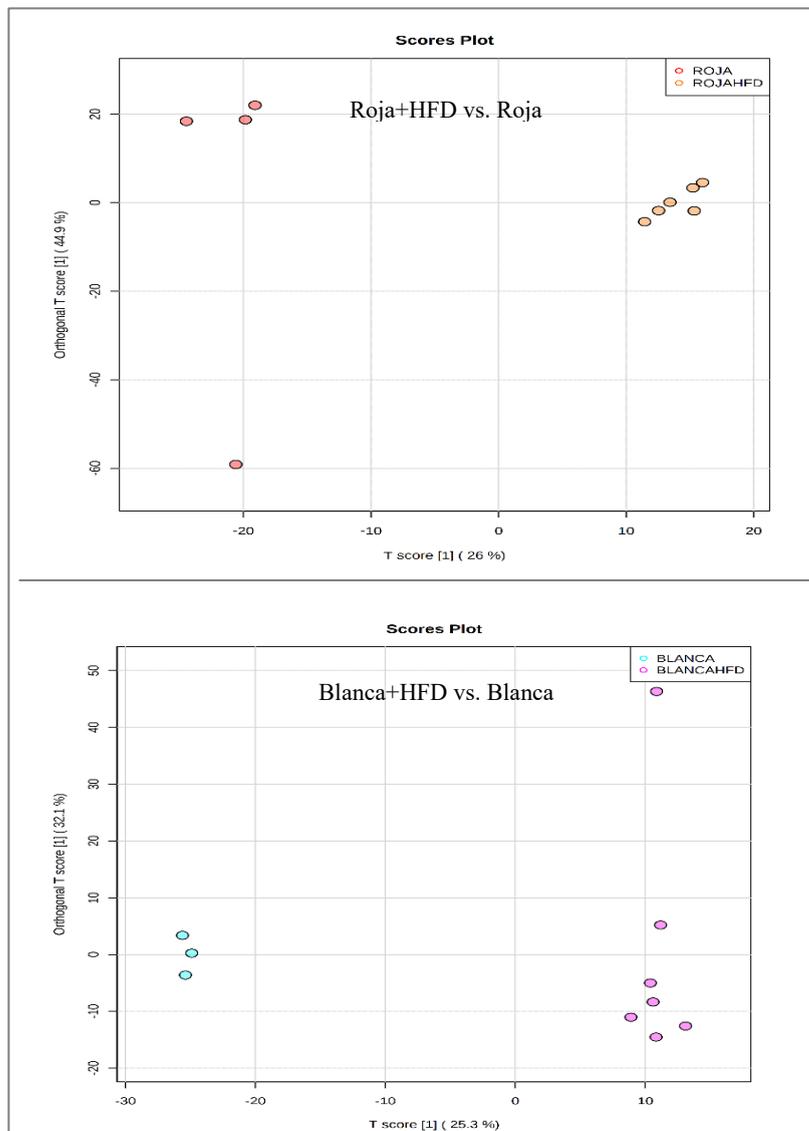


Figura 13. Diagramas O-PLSDA correspondientes a los dos grupos validados

Por otro lado, en la prueba *Cross validation* se generó una  $Q^2$  negativa. La cual, para validar el análisis, debe ser positiva. Se indicaría así que el modelo no tiene predictibilidad, ya que el valor mínimo debe ser de 0.7 en esta prueba (Xia et al., 2009).

Se observa en la parte superior de la Figura 12 la prueba por *Cross validation*, y se nota una inconsistencia entre los valores de  $R^2$  y  $Q^2$ , pues usualmente ambos dan como resultado un mínimo de 0.7. Por otro lado, en la prueba de permutación (Figura 12, parte inferior), se generó un *p value* deseable, siendo este menor a 0.001. Sin embargo, al trabajar con el modelo PLSDA, es necesario validar los datos por ambos test, de caso contrario no es posible usar este modelo estadístico.

Como es mencionado por Barupal y colaboradores en su artículo sobre las recientes aproximaciones de las estrategias para una correcta interpretación e integración de datos metabólicos, el probar con diferentes modelos estadísticos es muy importante. Esto se realiza para analizar y descubrir el comportamiento de dichos datos a determinados arreglos matemáticos (Barupal, Fan, & Fiehn, 2018).

El modelo ortogonal del análisis discriminante de mínimos cuadrados parciales (O-PLSDA) al usar filtro matemático, elimina la varianza de los datos que no está relacionada con su clase. Dicho fenómeno sucede gracias al arreglo matemático que tiene como base a la matriz de Hessian, ocasionando la exacerbación de las variables de respuestas, incrementando su magnitud como un evento aislado. Dando como resultado una separación entre dos grupos comparados en los diagramas (Allen, Payne, & Cole, 2018; Xia et al., 2009).

Este modelo estadístico presentó mejores probabilidades y resultados para el análisis dado su fundamento, ya que finalmente los datos pudieron ser validados por el test de permutación. Para esto, fue necesario formar parejas entre los tratamientos de diferente y el mismo estado metabólico y el tratamiento, mismos que están detallados en la Tabla 2.

Los diagramas de O-PLSDA se observan en la Figura 13, donde se observa que los tratamientos se separan en cada extremo del diagrama. Lo mismo ocurre de la misma manera tanto en la comparación de Roja+HFD vs. Roja (superior) y Blanca+HFD vs. Blanca (inferior). La notoria separación de los grupos indica que los datos responden de forma satisfactoria a este modelo estadístico, permitiendo su validación para poder continuar con el estudio a pesar de que no sea obligatorio que todos los grupos con los que se trabaja sean validados.

Como se mencionó, en ocasiones los modelos matemáticos programados en las

plataformas en versión gráfica pueden tener limitaciones debido a su flexibilidad y la de los datos con los que se trabaja. Como se mencionó anteriormente, los datos recabados requirieron un tratamiento filtrante adicional, así como el uso de un modelo estadístico en específico (O-PLSDA en este caso). Sólo de esta forma fue posible la validación de los datos, esto puede deberse a que los datos presentaron un alto ruido y variables internas de amplia magnitud que impidieron ver la respuesta hacía los tratamientos que se esperaba observar.

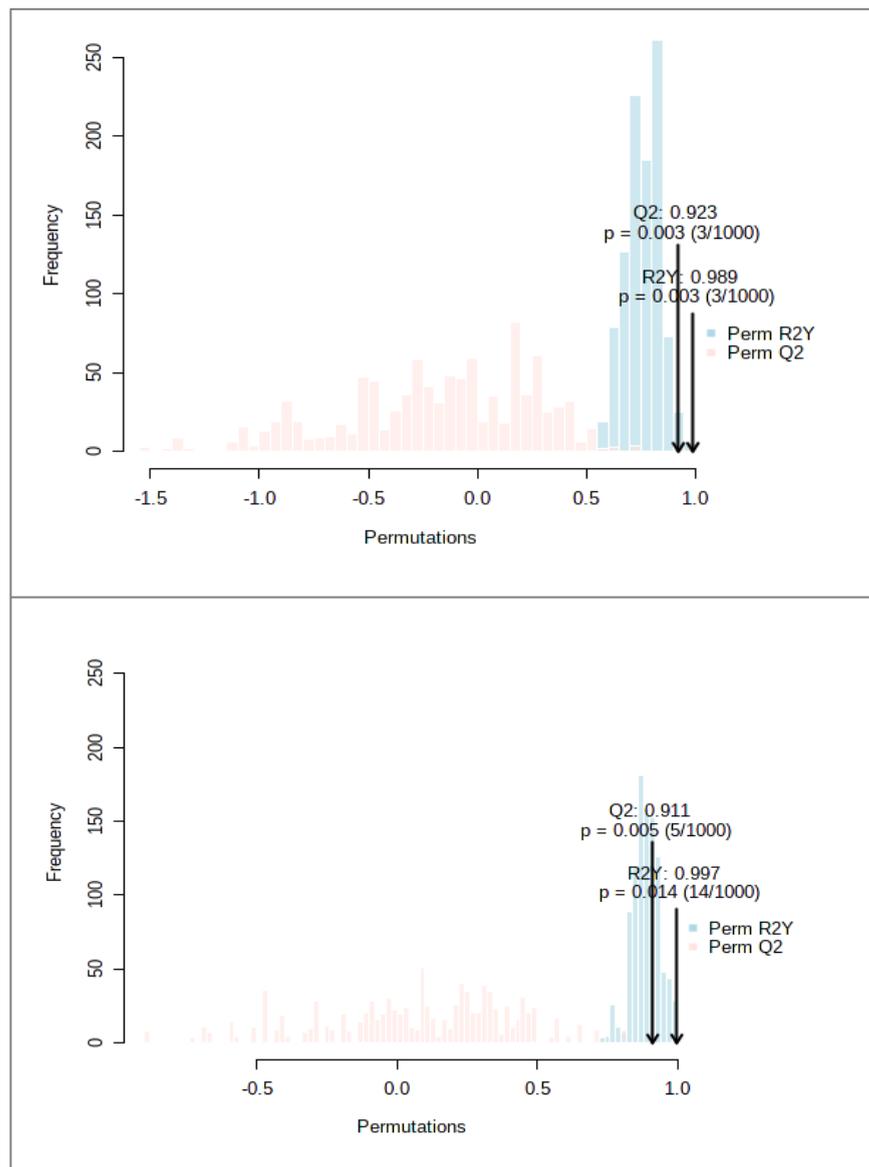


Figura 14. Validación por permutación de grupos validados.

También puede ocurrir que las plataformas presenten limitaciones para trabajar con datos biológicos. Esto ocurre debido a la variabilidad implícita de un organismo vivo.

Esta es la principal razón por la que se manejan modelos para datos biológicos, los cuales pueden ser flexibles o restrictivos. Sin embargo, se requiere amplia experiencia en el campo de la bioestadística para entender a fondo cómo es que estos funcionan (Xia et al., 2009).

Luego de someter a pruebas de validación estadística los grupos de estudio, sólo 2 fueron aptos para continuar el proceso. Los 8 grupos restantes no obtuvieron *p values* menores a 0.005 en la prueba de permutación. Los grupos que no lograron ser validados, fueron eliminados del estudio, pues de continuar con estos, se obtendrían falsos positivos. La comparación del grupo Blanca + HFD y Blanca obtuvo un *p value* de 0.005 y el grupo Roja + HFD con Roja uno de 0.003. Estos se muestran en la Figura 14 en la zona inferior y superior; respectivamente. Los valores los podemos encontrar en la Tabla 2.

Una vez validados los datos, se recolectaron un total de 22 *features* con un valor VIP (*variable importance parameter*) mayor a 3, pues esta fue la línea de corte que se decidió usar para hacer más estricto el análisis. En la Figura 15 se observa que este valor es el punto de inflexión donde, si subimos un poco se obtiene menos señales, pero más selectas y con más fuerza.

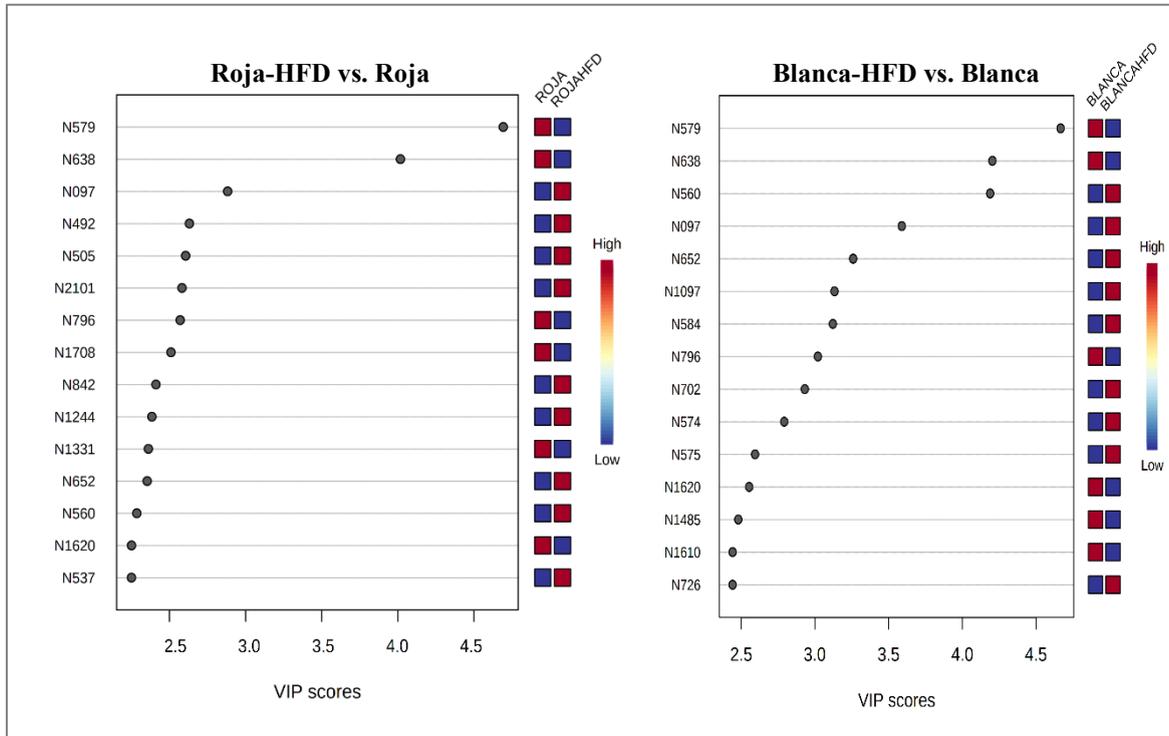


Figura 15. Features y valores VIP de grupos estudiados.

Dichas señales, fueron localizadas e integradas en un archivo de lectura. Se

obtuvieron los respectivos tiempos de retención (RT) y su relación m/z y se reservaron para continuar con el proceso de identificación. En el momento de iniciar los análisis estadísticos se detectaron 2440 señales, las cuales disminuyeron a 22 luego de atravesar por controles de calidad, tratamientos estadísticos y demás; para poder continuar con el proceso de identificación.

Esta es la principal ventaja del por qué se decidió llevar a cabo primero el análisis multivariado y luego la identificación. De esta forma, se disminuye considerablemente el número de variables y posibles compuestos, dándole importancia a los que presenten una diferencia estadísticamente significativa con respecto a los demás. Evitando, así, identificar compuestos que no sean de relevancia en el estudio (Michel et al., 2020).

Por otro lado, al realizarse el proceso de identificación en un grupo de datos muy grande llega a ser contraproducente, ya que esto indicaría que los procesos estadísticos a los que fueron sometidos no son lo suficientemente robustos.

No se afirma que ninguna de las metodologías sea mejor que la otra, sin embargo, una es claramente más práctica que la otra. Eliminar todo el ruido de fondo antes de identificar compuestos es de gran ayuda cuando se buscan nodos reguladores o biomarcadores.

Por otro lado, identificar antes de disminuir las variables, habla de un ensayo de amplio espectro que estaría aplicado para mapeos metabólicos. Son formas distintas de procesar los datos, para aplicaciones diferentes. Sin embargo, al ser rutinas estadísticas queda implícito que están sujetas a muchas variables de por sí y no es funcional agregar más variabilidad al estudio (Michel et al., 2020; Zhang et al., 2020).

La rutina implementada permitió comprobar información ya reportada, dando validez al procedimiento llevado a cabo.

## **5.5. Identificación de metabolitos**

Los 22 *features* seleccionados en base al valor VIP fueron mapeados o localizados en la relación de señales filtradas y detectadas por el equipo. A su vez, fueron marcados con un identificador especial para poder rastrearlos en la lista de todas las señales captadas y de esta manera poder ser ubicados con facilidad dentro de PROGENESIS QI. De 22 señales con puntaje VIP mayor a 3, se hallaron 17 coincidencias con otros compuestos de bases de datos reconocidas. Sin embargo, sólo se logró identificar 7 compuestos (Tabla 3). A pesar de haber usado una metodología y un equipo sumamente sensibles, el resto de

las señales detectadas no mostraron una abundancia relativa alta en los análisis quimiométricos.

Fueron detectados en una concentración mínima de 5 ppm para darle más sensibilidad a la identificación, la cual está detallada en la metodología ya establecida en PROGENESIS QI. Los compuestos identificados fueron hallados en bases de datos como *Kegg*, *Human Metabolome Data Base (HMDB)*, *Chemspider*, *ChMBL-EBI Data base* y otros más. Dentro de estas, podremos hallar compuestos que son producto del metabolismo endógeno y metabolitos exógenos.

Se trabajó principalmente con la base de datos de *Kegg* debido a que el proceso de identificación arrojó identificadores de esta. Se espera que todos los identificadores sean mapeados en HMDB ya que se busca compuestos endógenos detectados como producto del metabolismo interno del ser humano, pues es una base de datos especializada en esta tarea. Pero, al no ser posible, se ubicó todos los compuestos en la base de datos de *Kegg*. Esta es más general, pues dentro de esta se ubican toda clase de metabolitos; de mamíferos, plantas, bacterias, etc.

*Tabla 3. Metabolitos identificados.*

<b>Nombre</b>	<b>HMDB ID</b>	<b>KEGG ID</b>	<b>Score</b>	<b>Fragmentación</b>
Molybdopterina Z (precursor)	HMDB0011683	C05924	44.9	37.3
5-O-Mycaminosyprotylonolide	-	C12001	51.3	61.7
(17R(18S)-EpETE)	HMDB0010212	C13843	40.1	3.93
LysoPA(0:0/18:1(9z))	HMDB0007851	C00416	44	26.1
Lolitriol	-	C20548	48	56.8
Norrubrofusarin-6-beta-gentiobiosido	HMDB0034570	C17671	46.8	50.3
Acetoacetyl-CoA	HMDB0001484	C00332	40	0.88

La plataforma de METABOANALYST brinda la herramienta para trabajar con datos obtenidos a partir de esta. Como lo describe Chong y Xia, este servidor web es ampliamente usado por las tareas que es capaz de realizar y por su accesibilidad. Puede realizar desde análisis estadísticos multivariados, identificación de compuestos, genes y compilación de redes metabólicas (Chong & Xia, 2020). En contraste, otros autores como Leenders y colaboradores hacen estricta mención sobre la rigidez de plataformas

metabolómicas y proteómicas, inclinándose y aconsejando el uso de rutinas de estadísticas y de identificación mediante código R y Python (Leenders et al., 2020).

Se continuó la rutina filtrando y seleccionando los compuestos en base a un score mayor a 40 (Tabla 3), ya que este parámetro es el mínimo establecido por PROGENESIS para hacer de elegible un compuesto. Esta es una puntuación que se enumera de 0 a 100, considerando 5 parámetros de similitud: masiva, de isótopos, de tiempo de retención, de librería CCS y puntuación de fragmentación. Un promedio de estos debe dar como mínimo 40 para poder ser elegible (Röst et al., 2016).

De esta forma, el porcentaje de fragmentación puede llegar a ser ambiguo, ya que lo recomendado es que este valor sea lo más alto posible para asegurarse de forma rápida y sencilla que el compuesto elegido es realmente el que se identificó. Sin embargo, en ocasiones como esta no se logra conseguir un porcentaje alto. El siguiente paso es observar y analizar los espectros de cada compuesto y se los compara con su respectiva fragmentación para asegurarse que el compuesto identificado tiene el mismo patrón de fragmentación que el detectado por el equipo (Pucker et al., 2019).

El primer compuesto identificado fue Molybdopterina Z ( $C_{10}H_{12}N_5O_7P$ ), este es un precursor de molybdopterina, un tipo de oxotransferasa que usa un átomo de molibdeno como cofactor y que contiene pterina. Este tipo de enzimas generalmente catalizan la transferencia de átomos de oxígeno desde o hacia un sustrato usando el Mo como aceptor o donador de electrones (Wishart et al., 2018). Este compuesto ha sido identificado en sangre y en orina, como lo muestra la investigación realizada en el 2018 por Suganuma encontró que el complejo de molibdopterina previene el daño oxidativo causado por un exceso de aminoácidos azufrados, promoviendo así, la acelerada oxidación de ácidos grasos y formación de acetyl-CoA (Suganuma et al., 2018).

El segundo compuesto de la Tabla 3, 5-O-Mycaminosyprotylonolido ( $C_{31}H_{53}NO_8$ ) también se le conoce como 5-O-beta-D-Mycaminosyltylactona y es un antibiótico que pertenece a la clase de macrólidos, se deriva de un monosacárido y de una enona. No es un metabolito que se produzca del metabolismo endógeno. Sin embargo, haber detectado su presencia puede deberse al metabolismo hepático de determinados compuestos de la planta (Adeva-Andany et al., 2014; Wishart et al., 2018).

El tercer compuesto identificado es 17R(18S)-EpETE ( $C_{20}H_{30}O_3$ ) o ácido 17,18-epoxyeicosatetraenoico es un compuesto orgánico clasificado como ácido graso de cadena larga. Tiene una cola alifática de 13 - 21 átomos de carbono, siendo una molécula altamente hidrófoba, de carácter eicosanoide (originadas luego de la oxidación de ácidos

grasos esenciales de 20 carbonos, como omega-3 y omega-6) (Arneth, Arneth, & Shams, 2019; Wishart et al., 2018). Park menciona que la presencia en suero de este ácido grasos habla sobre una disminución de los mediadores de la respuesta inflamatoria así como de las especies reactivas de oxígeno (S. J. Park & Park, 2020).

Tabla 4. Rutas metabólicas relacionadas con los metabolitos identificados.

Ruta metabólica*	Total	Hits	Impacto
Síntesis y degradación de cuerpos cetónicos	5	1	0.13
Metabolismo de butanoato	15	1	0.26
Metabolismo de glicerofosfolípidos	16	1	0.01
Biosíntesis de terpenoides	18	1	0.00
Metabolismo de piruvato	22	1	0.00
Metabolismo de propanoato	23	1	0.00
Degradación de lisina	25	1	0.05
Biosíntesis de folato	27	1	0.05
Sistema de señalización de Fosfatidilinositol	28	1	0.00
Metabolismo de glioxilato y dicarboxilato	32	1	0.00
Metabolismo de glicerofosfolípidos	36	1	0.14
Degradación de ácidos grasos	39	1	0.02
Degradación de valina, leucina e isoleucina	40	1	0.00
Metabolismo de triptofano	41	1	0.00

\*Dichas rutas provenientes del análisis de rutas metabólicas.

También se logró identificar LysoPA(0:0/18:1(9z)) o ácido lisofosfatídico (0:018:1) (C<sub>21</sub>H<sub>41</sub>O<sub>7</sub>P). Se describe como un glicerofosfolípido con un resto fosfato en la posición del glicerol. Este compuesto es mediador lipídico, tiene tareas similares al factor de crecimiento que se libera de forma instantánea al detectar plaquetas activadas (Wishart et al., 2018). Estudios llevados a cabo por Fan y colaboradores han demostrado que la producción excesiva de este metabolito puede deteriorar la tolerancia a la glucosa, siendo un marcador de inflamación (Fan et al., 2021).

Lolitriol (C<sub>37</sub>H<sub>49</sub>NO<sub>7</sub>) es un compuesto identificado en semillas de una especie perenne infectada con ciertos tipos de hongos, como *Neotyphodium lolii*. Es precursor activo de Lolitrem B, neurotoxina responsable de causar el “escalofrío de raigrás”, la cual es una enfermedad en animales de granja que se caracteriza por la presencia de temblores

y descoordinaciones motoras (Wishart et al., 2018). Es probable que se hace detectado su presencia debido al consumo de granos infectados con el patógeno.

El penúltimo compuesto orgánico de la tabla lleva como nombre norrubrofusarin-6-beta-gentiobiosido ( $C_{26}H_{30}O_{15}$ ). Son conocidos como glucósidos de naftopiranona, y tienen un carbohidrato y un enlace glicosídico a la naftopiranona. Es un compuesto prácticamente neutro, y considerado una base muy débil. Ha sido detectado en el cuerpo, y fuera de él en algunos alimentos como el café y derivados (Wishart et al., 2018).

Por último, el compuesto detectado con un menor porcentaje de fragmentación, Acetoacetil-CoA ( $C_{25}H_{40}N_7O_{18}P_3S$ ) es un compuesto conocido como aminopiperidinas. Es un sustrato para varias enzimas, sobre todo para succinil-CoA que forma parte del Ciclo de Krebs. También es esencial para síntesis de colesterol ya que es un precursor de HMG-CoA. Es crucial en la ruta del mevalonato y de los cuerpos cetónicos (Wishart et al., 2018). Fukao y colegas realizaron una revisión, en la cual hace mención a este cuerpo cetónico que, en abundancia, brinda información sobre una desregulación del metabolismo de lípidos (Fukao et al., 2014).

Haber identificado los compuestos mencionados brinda una idea panorámica de cómo podría funcionar la dieta y confirmar que las bebidas causaron un efecto en los animales.

## 5.6. Construcción de redes metabólicas.

La integración de las rutas metabólicas fue gracias a la herramienta de la plataforma METABOANALYST 5.0. Se tomaron los identificadores de los metabolitos detectados en el paso anterior, corroborándose en la base de datos de KEGG y fueron introducidos en la sección de “*Join pathway*” en la plataforma metabolómica.

Realizar esta tarea, se refiere a la simulación mediante un conjunto de técnicas que recompila información relevante de un organismo, traduciéndola y compilándola en un modelo matemático. Luego de manejarlo, este se refina a través de datos ya conocidos y se puede analizar de manera holística e interactiva (Majumder et al., 2021).

De esta forma, es posible basarse en la conservación de la masa de los metabolitos para expresar la red mediante una matriz estequiométrica, equilibrando la ecuación pues se considera que no existe la acumulación interna de metabolitos. Así lo mencionan Ren y colegas, quienes en 2015 modificaron el código en R de la rutina implementada por el equipo liderado por Fernández en 2014 de la paquetería MAIT para la identificación

automática de compuestos (Fernández-Albert, Llorach, Andrés-Lacueva, & Perera, 2014; Ren et al., 2015).

De este análisis holístico de las rutas metabólicas, se encontró que los metabolitos identificados tienen presencia en un total de 14 rutas, detalladas en la Tabla 4. Este es resultado fue obtenido a través de la rutina predeterminada del servidor web METABOANALYST.

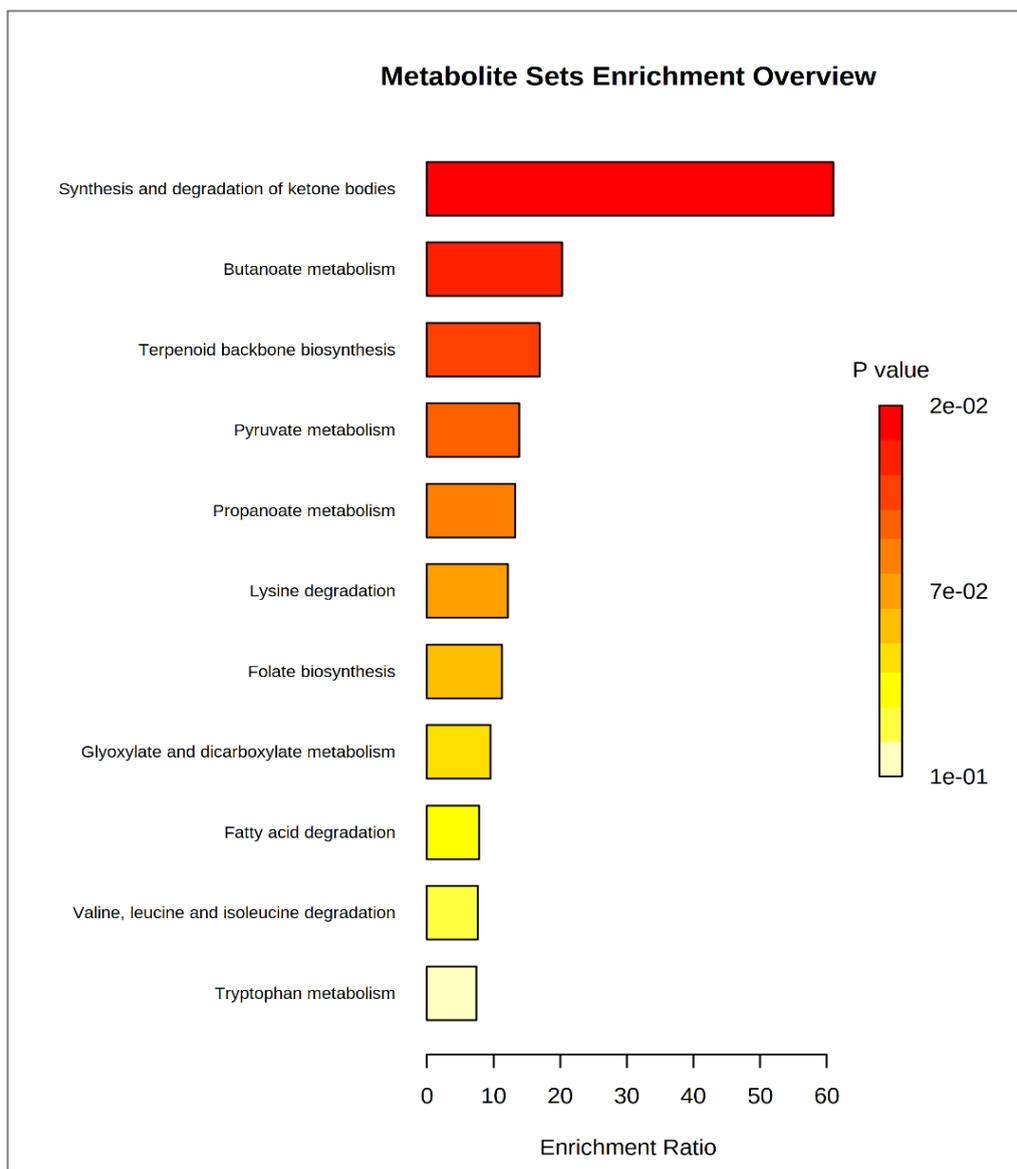


Figura 16. Análisis de enriquecimiento de rutas metabólicas.

En dicha tabla, “total” es el número total de compuestos involucrados en la ruta metabólica, “hits” se refiere al número de los coincidentes cargados para realizar el análisis. El impacto se da en un valor calculado a partir de la topología de la vía (S. Park et al., 2015).

El estudio de enriquecimiento de rutas metabólicas mostró principalmente que dos vías tienen un cambio significativo en sus concentraciones: síntesis y degradación de cuerpos cetónicos y el metabolismo de butanoato. A pesar de que hay otras rutas metabólicas involucradas, estas no tienen un impacto grande con respecto a las ya mencionadas, como se muestra en la Figura 16. Estas vías se encuentran de un color rojo intenso debido al *p value* que obtuvieron, el cual fue menor a 0.002. La primera mencionada con un valor de impacto entre 0.014 y 0.015 y la segunda con uno mayor a 0.025.

De la misma forma, se guarda relación con la gráfica de impacto de las rutas metabólicas obtenida como resultado de la plataforma usada (Figura 17), donde se observa las mismas vías con un alto impacto. Así, las dos siguientes vías más afectadas son el metabolismo de glicerofosfolípidos y de glicerolípidos. Esta información guarda relación, con la serie de reportes publicados por D'Souza los cuales los puentes que existen entre la alimentación y los cambios bioquímicos y metabólicos que ocurren en un organismo (D'Souza, Paramel, & Kienesberger, 2018).

Al realizarse una comparación mediante análisis quimiométricos en base a la abundancia relativa, se observaron cambios sutiles entre los compuestos mencionados. Es decir, variaciones ligeras entre ellos. Dichas abundancias se encuentran en la Tabla 6 en la sección de Anexo A.

La varianza más notable la mostró el ácido lisofosfatídico, que se encuentra naturalmente elevado en estados de obesidad y resistencia a la insulina (Lee, Kim, Oh, & Jun, 2019). Este compuesto se encontró incrementado en el grupo control "Obesas". Con respecto a este, se observó que en los grupos de "Blanca+HFD" y "Roja+HFD" se encuentra ligeramente disminuido. Sin embargo, no todos los animales obesos del estudio mostraron una disminución del compuesto mencionado. Dicho evento, guarda relación con lo reportado por Rubio-Rodríguez, investigador principal del proyecto quien menciona que la bebida que nosotros llamamos "Blanca", tuvo un efecto hepatoprotector al reducir lípidos en sangre (Rubio-Rodríguez et al., 2021).

Entre ambos grupos, los animales obesos que presentaron una mayor disminución de dicho compuesto fueron los pertenecientes al grupo que consumió la bebida funcional nombrada como "Blanca", en promedio un 27.84%. Con esta premisa, podríamos tener un indicio que el tratamiento con esta bebida podría haber sido el responsable de disminuir la concentración de este metabolito en el grupo de estudio (ver Tabla 6) (Rubio-Rodríguez et al., 2021; Tangvarasittichai, 2015)

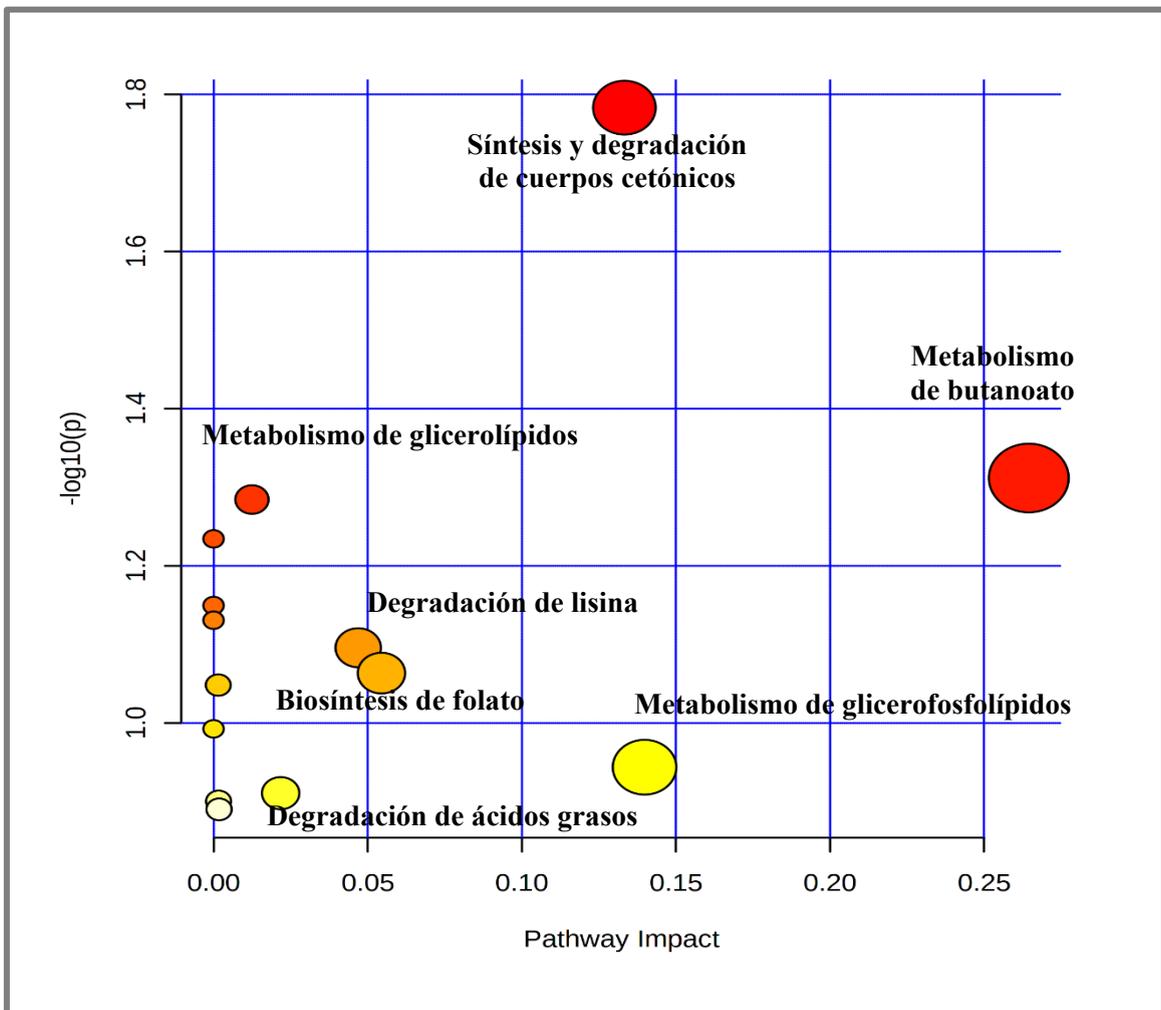


Figura 17. Gráfico de impacto en las rutas metabólicas.

Por otro lado, acetoacetyl-CoA también mostró una diferencia muy sutil al comparar la abundancia relativa obtenida con respecto al control. A pesar de que estos están disminuidos en ambos tratamientos comparados y validados, sólo uno demostró mayor diferencia con respecto al control. Nuevamente fueron los animales obesos tratados con la bebida “Blanca” los que presentaron una menor abundancia con respecto al control, aproximadamente un 5.61% (ver Tabla 6).

Acetoacetyl-CoA tiene un papel biológico de importancia, pues participa en una lista grande de rutas metabólicas. Esto se menciona en el artículo de Qiu y colaboradores donde se habla sobre la cetoacidosis en ratas alimentadas con una dieta alta en grasa y fructosa que desarrollaron un estado de diabetes (Qiu, Novikov, & Vallon, 2017). Al observar los resultados obtenidos en el presente estudio, como la gráfica de impacto de los compuestos en las rutas metabólicas y la gráfica de enriquecimiento, se mostró que el metabolito tuvo incidencia principalmente en el metabolismo de cuerpos cetónicos, el

metabolismo de butanoato, en la degradación de lisina y de ácidos grasos; confirmando la importancia de su papel metabólico.

En hepatocitos, dicho metabolito se encuentra abundancia junto con acetil-CoA, lo que indica una beta-oxidación, según Fukao y colaboradores. Dicho proceso, se encuentra desregulado en un estado de resistencia a la insulina, debido a la incapacidad del organismo de aprovechar la glucosa exógena. En el artículo mencionado, se expone sobre cuerpos cetónicos y sus efectos en el cuerpo; explicando la relevancia de los metabolitos que dan lugar a 3-hidroxiacetil-CoA y acetoacetato; ya que son vectores importantes para el aprovechamiento de energía en hígado y tejidos extrahepáticos (Fukao et al., 2014). El poco o nulo aprovechamiento de insulina y un inherente aumento del glucagón en pacientes diabéticos conduce a una concentración alta de ácidos libres en suero y en hígado. Los cuales se metabolizan de manera acelerada a dióxido de carbono y agua, al mismo tiempo que acetil-CoA se convierte en acetoacetil-CoA y finalmente en ácido acetoacético e hidroxibutirato (Chandrasegaran et al., 2018).

Como lo confirma Chandrasegaran, un estado metabólico afectado por la diabetes es capaz de producir cuerpos cetónicos como acetoacetil-CoA. Dichos metabolitos se caracterizan por ser fuente energética alternativa de ciertos órganos como el corazón y el cerebro en estados donde el cuerpo no pueda ingresar la glucosa a las células. Evidencias muestran que la captación de acetoacetil-CoA en los animales de experimentación usados es indicador de un estado de lipólisis debido a la dieta alta en grasa y fructosa (Chandrasegaran et al., 2018; Qiu et al., 2017).

Por otro lado, nuestro estudio arrojó que lysoPA(0:0/18:1(9z)) o ácido lisofosfatídico se encuentra involucrado en el metabolismo de glicerofosfolípidos y de glicerolípidos. Interviene en procesos de señalización, proliferación y migración como lo menciona Arneth y colaboradores en el artículo donde se plasma la diferencia metabólica de entre individuos que padecen diabetes tipo I y tipo II. Este compuesto en concentraciones elevadas puede producir daño celular a través de la superposición de rutas metabólicas produciendo especies reactivas de oxígeno (Arneth et al., 2019).

Aunque es un compuesto simple, es clasificado como una molécula altamente importante en procesos de señalización celular por su estrecha relación a los receptores de la proteína G. Además, las alteraciones de este metabolito se pueden detectar de manera directa e indirecta gracias a las alteraciones en expresión y producción de enzima autotaxina, pues esta es la encargada de producirlo (D'Souza et al., 2018).

D'Souza y colaboradores demostraron que las concentraciones circulantes en

sangre de autotaxina y ácido lisofosfatídico son reguladas por la dieta y el ayuno, comprobando su participación en procesos de desregulación metabólica como la diabetes y obesidad (D'Souza et al., 2017). A pesar que el mecanismo por el cual la dieta regula sus concentraciones no está del todo claro, este mismo autor indica que se encontró una sobreexpresión de mRNA y proteínas de autotaxina en ratones macho que fueron alimentados con una dieta alta en grasa y fructosa (D'Souza et al., 2018). Por otro lado, dicho compuesto fue detectado en sangre de pacientes que desarrollaron nefropatía diabética (Lee et al., 2019).

A pesar de que el estudio llevado a cabo por Nishimura y colegas contradice lo mencionado, la discrepancia entre ambos autores no está clara (Sun et al., 2017). Por el momento, nuestro estudio corrobora los hallazgos de D'Souza y guarda relación con el modelo animal que se diseñó para estudiar las variaciones de respuesta en animales con resistencia a la insulina.

Los metabolitos identificados y las rutas mapeadas analizadas de forma holística forman una pieza interesante en los trastornos metabólicos relacionados con la resistencia a la insulina, estado predominante en la diabetes mellitus tipo II (Dąbek et al., 2020; Lee et al., 2019). Evidencias muestran que acetoacetyl-CoA es metabolito importante, ya que interconecta el metabolismo de la síntesis y degradación de cuerpos cetónicos a través de hidroxibutanoato y según su abundancia en el organismo, ser indicador de la desregulación a nivel de lípidos (Qiu et al., 2017).

Dicho fenómeno estaría relacionado con una elevada oxidación de ácidos grasos en el hígado, para su producción y aprovechamiento (Oguntibeju, 2019). Lo encontrado en este estudio, reafirma lo reportado por numerosos autores. Dicho fenómeno ocurre debido a la imposibilidad de aprovechar la glucosa proveniente de la dieta, estado ocasionado por el desarrollo de una resistencia a la insulina. Esto traería como consecuencia un falso estado de inanición, debido a que el organismo posee glucosa exógena que es incapaz de ingresar a las células (Alsahli & Gerich, 2017). De esta forma se cree que la metodología implementada dio frutos al contrastar los resultados con lo obtenido por Fukao en su trabajo donde habla de los efectos de la cetoacidosis en pacientes diabéticos. (Fukao et al., 2014).

Por otro lado, lysoPA(0:0/18:1(9z)) o ácido lisofosfatídico, juega el rol de conector y regulador de algunas rutas metabólicas que están relacionadas con el aporte de sustratos gluconeogénicos (Lee et al., 2019). Relacionando principalmente en el metabolismo de los aminoácidos glicina, serina y treonina, además de aportar materia

para el metabolismo del piruvato y de la glucosa (Alsahli & Gerich, 2017). Diversos estudios reportan la presencia de este compuesto y de la enzima productora de este de forma directamente proporcional al nivel de resistencia a la insulina que se genera al detectar una abundancia elevada de estos compuestos (D'Souza et al., 2017).

La dieta ha demostrado ser un factor importante para el mantenimiento de la homeostasis, por lo que, si esta presenta alteraciones, el equilibrio metabólico interno se verá afectado. El reflejo medible del estado metabólico interno son los metabolitos, responsables de distintas tareas y reacciones dentro del organismo, las cuales van desde el aprovechamiento de la glucosa hasta llevar a cabo procesos de inflamación. Diversos autores han identificado y cuantificado metabolitos involucrados en enfermedades complejas y en diferentes estados metabólicos como inflamación y resistencia a la insulina,

Los metabolitos y rutas halladas guardan relación con trastornos metabólicos, esto sugiere la presencia de un desorden del metabolismo de carbohidratos y ácidos grasos, principalmente (Arneith et al., 2019; Qiu et al., 2017). Esta afirmación concuerda con lo establecido por D'Souza y Fukao quienes trabajaron con animales obesos.

Algunos de los compuestos detectados se ven involucrados de distinta manera con la diabetes y la resistencia a la insulina se presentan en la Tabla 5 de la sección Anexo A, donde hallaremos el piruvato (Vanderperre, Bender, Kunji, & Martinou, 2015). Si bien en el presente estudio no se identificó este metabolito, tienen relación con acetoacetil-CoA y con el ácido lisofosfatídico (18:1) (Dąbek et al., 2020; Lee et al., 2019).

A pesar de que dos compuestos presentaron diferencias entre tratamientos con respecto al control, la información recabada hasta el momento no es suficiente para predecir su comportamiento mediante un mecanismo molecular. Esto principalmente se debe a que dos compuestos con una diferencia sutil con respecto a su control no aportan información relevante al estudio.

Sin embargo, se puede agregar que la bebida nombrada como “Blanca” en este estudio fue capaz de reducir, de forma discreta, la concentración de acetoacetil-CoA y del ácido lisofosfatídico. Así también, podemos mencionar este estudio brinda un panorama general de qué rutinas mejorar para realizar análisis metabólicos más efectivos y que brinden más información.

# Conclusiones

La migración hacia un formato abierto fue posible, así como la visualización de espectros de masas por inyección; no por componente.

Sin la presencia de valores de la primera fragmentación (MS1), los análisis posteriores a la rutina establecida no fueron posibles.

El pretratamiento y una limpieza adicional a la que realizan los programas convencionalmente hizo posible la validación y la posterior detección de los metabolitos.

En lo que respecta a la identificación de metabolitos relevantes en un estado de resistencia a la insulina, se identificaron un total de 7 compuestos. De los cuales, 2 son de importancia en estados de diabetes y resistencia a la insulina: Acetoacetil-CoA y ácido lisofosfatídico.

Acetoacetil-CoA y ácido lisofosfatídico (18:1) se encontraron aumentados con respecto al control, esto es característico en un estado de obesidad y resistencia a la insulina.

El tratamiento de la bebida “Blanca” en animales obesos, disminuyó de forma discreta dichos compuestos, generando un efecto protector hacia una dieta desequilibrada.

No haber encontrado una diferencia significativa entre tratamientos, dificulta predecir un mecanismo de acción de los tratamientos empleados.

# Recomendaciones

Es necesario promover el uso de técnicas alternativas que no requieran el uso de softwares de licencia, esto facilita en gran medida la capacidad de análisis en el laboratorio. Sin embargo, al no contar con un formato libre con valores de la primera fragmentación (MS1) esta tarea se hace difícil.

Como alternativa se recomienda la implementación de metodologías que sólo usen datos de la segunda fragmentación (MS2) y solicitar información al fabricante sobre cómo rescatar los valores (MS1) del equipo, sin causar daños.

# Referencias Bibliográficas

- Abascal, F. (2020, April 6). Spotlight on Bioinformatics. *Nature*.  
<https://doi.org/10.1038/nj0478>
- Adeva-Andany, M., López-Ojén, M., Funcasta-Calderón, R., Ameneiros-Rodríguez, E., Donapetry-García, C., Vila-Altesor, M., & Rodríguez-Seijas, J. (2014). Comprehensive review on lactate metabolism in human health. *Mitochondrion*, Vol. 17, pp. 76–100.  
<https://doi.org/10.1016/j.mito.2014.05.007>
- Allen, A. E. A., Payne, M. C., & Cole, D. J. (2018). Harmonic Force Constants for Molecular Mechanics Force Fields via Hessian Matrix Projection. *Journal of Chemical Theory and Computation*, 14(1), 274–281. <https://doi.org/10.1021/acs.jctc.7b00785>
- Aloy, P. (2017). *A NETWORK BIOLOGY APPROACH TO NOVEL THERAPEUTIC STRATEGIES*. Retrieved from <https://riuma.uma.es/xmlui/handle/10630/13772>
- Alsahli, M., & Gerich, J. E. (2017, November 1). Renal glucose metabolism in normal physiological conditions and in diabetes. *Diabetes Research and Clinical Practice*, Vol. 133, pp. 1–9. <https://doi.org/10.1016/j.diabres.2017.07.033>
- Antonelli, J., Claggett, B. L., Henglin, M., Kim, A., Ovsak, G., Kim, N., ... Cheng, S. (2019, July 1). Statistical workflow for feature selection in human metabolomics data. *Metabolites*, Vol. 9, p. 143. <https://doi.org/10.3390/metabo9070143>
- Arneth, B., Arneth, R., & Shams, M. (2019, May 2). Metabolomics of type 1 and type 2 diabetes. *International Journal of Molecular Sciences*, Vol. 20, p. 2467.  
<https://doi.org/10.3390/ijms20102467>
- Barupal, D. K., Fan, S., & Fiehn, O. (2018, December 1). Integrating bioinformatics approaches for a comprehensive interpretation of metabolomics datasets. *Current Opinion in Biotechnology*, Vol. 54, pp. 1–9.  
<https://doi.org/10.1016/j.copbio.2018.01.010>
- Basto-Abreu, A., Barrientos-Guitérrez, T., Rojas Martínez, R., & Al., E. (2020). Prevalencia de diabetes y descontrol glucémico e México: resultados de la Ensanut 2016. *Salud Pública Mex.*, 62(1), 50–59.
- Bauer, C., Cramer, R., & Schuchhardt, J. (2011). Evaluation of peak-picking algorithms for

- protein mass spectrometry. *Methods in Molecular Biology (Clifton, N.J.)*, 696, 341–352. [https://doi.org/10.1007/978-1-60761-987-1\\_22](https://doi.org/10.1007/978-1-60761-987-1_22)
- Begou, O., Gika, H. G., Theodoridis, G. A., & Wilson, I. D. (2018). Quality control and validation issues in LC-MS metabolomics. In *Methods in Molecular Biology* (Vol. 1738, pp. 15–26). [https://doi.org/10.1007/978-1-4939-7643-0\\_2](https://doi.org/10.1007/978-1-4939-7643-0_2)
- Cambiaghi, A., Ferrario, M., & Masseroli, M. (2017). Analysis of metabolomic data: Tools, current strategies and future challenges for omics data integration. *Briefings in Bioinformatics*, 18(3), 498–510. <https://doi.org/10.1093/bib/bbw031>
- Chambers, M. C., MacLean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., ... Mallick, P. (2012, October 10). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, Vol. 30, pp. 918–920. <https://doi.org/10.1038/nbt.2377>
- Chandirasegaran, G., Ghosh, K., Manivel, G., Hassan, T., Elanchezhyan, C., & Manivel, G. (2018). Diabetes Mellitus Induced Oxidative Stress, Inflammation and Apoptosis: A Concise Review Cronicon EC DIABETES AND METABOLIC RESEARCH Diabetes Mellitus Induced Oxidative Stress, Inflammation and Apoptosis: A Concise Review. In *Metabolic Research* (Vol. 2). Retrieved from <https://www.researchgate.net/publication/331199644>
- Chong, J., & Xia, J. (2020). Using MetaboAnalyst 4.0 for Metabolomics Data Analysis, Interpretation, and Integration with Other Omics Data. In *Methods in Molecular Biology* (Vol. 2104, pp. 337–360). [https://doi.org/10.1007/978-1-0716-0239-3\\_17](https://doi.org/10.1007/978-1-0716-0239-3_17)
- Czech, M. P. (2017). Insulin action and resistance in obesity and type 2 diabetes. *Nature Medicine*, 23(7), 804–814. <https://doi.org/10.1038/nm.4350>
- D’Souza, K., Kane, D. A., Touaibia, M., Kershaw, E. E., Pulinilkunnil, T., & Kienesberger, P. C. (2017). Autotaxin is regulated by glucose and insulin in adipocytes. *Endocrinology*, 158(4), 791–803. <https://doi.org/10.1210/en.2017-00035>
- D’Souza, K., Paramel, G. V., & Kienesberger, P. C. (2018, April 1). Lysophosphatidic acid signaling in obesity and insulin resistance. *Nutrients*, Vol. 10, p. 399. <https://doi.org/10.3390/nu10040399>
- Dąbek, A., Wojtala, M., Pirola, L., & Balcerczyk, A. (2020, March 1). Modulation of cellular biochemistry, epigenetics and metabolomics by ketone bodies. Implications of

- the ketogenic diet in the physiology of the organism and pathological states. *Nutrients*, Vol. 12, p. 788. <https://doi.org/10.3390/nu12030788>
- Duncan, K. D., Fyrestam, J., & Lanekoff, I. (2019, February 7). Advances in mass spectrometry based single-cell metabolomics. *Analyst*, Vol. 144, pp. 782–793. <https://doi.org/10.1039/c8an01581c>
- Erngren, I., Haglöf, J., Engskog, M. K. R., Nestor, M., Hedeland, M., Arvidsson, T., & Pettersson, C. (2019). Adduct formation in electrospray ionisation-mass spectrometry with hydrophilic interaction liquid chromatography is strongly affected by the inorganic ion concentration of the samples. *Journal of Chromatography A*, 1600, 174–182. <https://doi.org/10.1016/j.chroma.2019.04.049>
- Fan, H., Li, Y., Wang, J., Shao, J., Tang, T., Lai, T., ... Lai, S. (2021). *Untargeted Metabolomic Characteristics of Skeletal Muscle Dysfunction in Rabbits Induced by a High Fat Diet*. <https://doi.org/10.21203/rs.3.rs-159084/v1>
- Fernández-Albert, F., Llorach, R., Andrés-Lacueva, C., & Perera, A. (2014). An R package to analyse LC/MS metabolomic data: MAIT (Metabolite Automatic Identification Toolkit). *Bioinformatics*, 30(13), 1937–1939. <https://doi.org/10.1093/bioinformatics/btu136>
- Fukao, T., Mitchell, G., Sass, J. O., Hori, T., Orii, K., & Aoyama, Y. (2014). Ketone body metabolism and its defects. *Journal of Inherited Metabolic Disease*, 37(4), 541–551. <https://doi.org/10.1007/s10545-014-9704-9>
- Gan, W. Z., Ramachandran, V., Lim, C. S. Y., & Koh, R. Y. (2020, March 1). Omics-based biomarkers in the diagnosis of diabetes. *Journal of Basic and Clinical Physiology and Pharmacology*, Vol. 31. <https://doi.org/10.1515/jbcpp-2019-0120>
- Hardie, D. G. (2014). Review AMPK — Sensing Energy while Talking to Other Signaling Pathways. *Cell Metabolism*, 20(6), 939–952. <https://doi.org/10.1016/j.cmet.2014.09.013>
- Hasin, Y., Seldin, M., & Lusis, A. (2017, May 5). Multi-omics approaches to disease. *Genome Biology*, Vol. 18, pp. 1–15. <https://doi.org/10.1186/s13059-017-1215-1>
- Inegi. (2021). *COMUNICADO DE PRENSA NÚM. 61/21 27 DE ENERO DE 2021 PÁGINA 2/4 COMUNICACIÓN SOCIAL*. Retrieved from <https://www.paho.org/hq/dmdocuments/2016/2016-cha-epidemiological-calendar.pdf>

- Jacob, M., Malkawi, A., Albast, N., Al Bougha, S., Lopata, A., Dasouki, M., & Abdel Rahman, A. M. (2018). A targeted metabolomics approach for clinical diagnosis of inborn errors of metabolism. *Analytica Chimica Acta*, *1025*, 141–153.  
<https://doi.org/10.1016/j.aca.2018.03.058>
- Janghorbani, M., Dam, R. M. Van, Willett, W. C., & Hu, F. B. (2007). Article Navigation Systematic Review of Type 1 and Type 2 Diabetes Mellitus and Risk of Fracture. *American Journal of Epidemiology*, *166*, 495–505.
- Klasic, A., Isakovic, A., Kocic, G., Kavacic, N., Jovanovic, M., Zvrko, E., ... Ninic, A. (2018). Relationship between Oxidative Stress, Inflammation and Dyslipidemia with Fatty Liver Index in Patients with Type 2 Diabetes Mellitus. *Experimental and Clinical Endocrinology and Diabetes*, *126*(6), 371–378. <https://doi.org/10.1055/s-0043-118667>
- Lederer, A. R., & La Manno, G. (2020, June 1). The emergence and promise of single-cell temporal-omics approaches. *Current Opinion in Biotechnology*, Vol. 63, pp. 70–78.  
<https://doi.org/10.1016/j.copbio.2019.12.005>
- Lee, J. H., Kim, D., Oh, Y. S., & Jun, H. S. (2019, June 1). Lysophosphatidic acid signaling in diabetic nephropathy. *International Journal of Molecular Sciences*, Vol. 20, p. 2850.  
<https://doi.org/10.3390/ijms20112850>
- Leenders, J., Grootveld, M., Percival, B., Gibson, M., Casanova, F., & Wilson, P. B. (2020). Benchtop low-frequency 60 MHz NMR analysis of urine: A comparative metabolomics investigation. *Metabolites*, *10*(4), 155. <https://doi.org/10.3390/metabo10040155>
- Li, S., Shao, M., & Fu, Y. (2018). Multi-view low-rank analysis with applications to outlier detection. *ACM Transactions on Knowledge Discovery from Data*, *12*(3).  
<https://doi.org/10.1145/3168363>
- Liu, X., & Locasale, J. W. (2017, April 1). Metabolomics: A Primer. *Trends in Biochemical Sciences*, Vol. 42, pp. 274–284. <https://doi.org/10.1016/j.tibs.2017.01.004>
- Lotfy, M., Adeghate, J., Kalasz, H., Singh, J., & Adeghate, E. (2017). Chronic Complications of Diabetes Mellitus: A Mini Review. *Current Diabetes Reviews*, *13*, 3–10.
- Majumder, E. L. W., Billings, E. M., Benton, H. P., Martin, R. L., Palermo, A., Guijas, C., ... Siuzdak, G. (2021). Cognitive analysis of metabolomics data for systems biology. *Nature Protocols*, *16*(3), 1376–1418. <https://doi.org/10.1038/s41596-020-00455-4>

- Michel, M., Dubowy, K. O., Zlamy, M., Karall, D., Adam, M. G., Entenmann, A., ... Laser, K. T. (2020). Targeted metabolomic analysis of serum phospholipid and acylcarnitine in the adult Fontan patient with a dominant left ventricle. *Therapeutic Advances in Chronic Disease*, 11. <https://doi.org/10.1177/2040622320916031>
- Oguntibeju, O. O. (2019). Type 2 diabetes mellitus, oxidative stress and inflammation: examining the links. *International Journal of Physiology, Pathophysiology and Pharmacology*, 11(3), 45–63. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/31333808>
- Park, S. J., & Park, D. H. (2020). Revisiting lipids in retinal diseases: A focused review on age-related macular degeneration and diabetic retinopathy. *Journal of Lipid and Atherosclerosis*, 9(3), 406–418. <https://doi.org/10.12997/jla.2020.9.3.406>
- Park, S., Sadanala, K. C., & Kim, E. (2015). A Metabolomic Approach to Understanding the Metabolic Link between Obesity and Diabetes. *Molecules and Cells*, 38(7), 587–596.
- Perez-Riverol, Y., & Moreno, P. (2020). Scalable Data Analysis in Proteomics and Metabolomics Using BioContainers and Workflows Engines. *PROTEOMICS*, 20(9), 1900147. <https://doi.org/10.1002/PMIC.201900147>
- Phelan, V. V. (2020). Feature-Based Molecular Networking for Metabolite Annotation. In *Methods in Molecular Biology* (Vol. 2104). [https://doi.org/10.1007/978-1-0716-0239-3\\_13](https://doi.org/10.1007/978-1-0716-0239-3_13)
- Pucker, B., Schilbert, H. M., & Schumacher, S. F. (2019). Integrating Molecular Biology and Bioinformatics Education. *Journal of Integrative Bioinformatics*, 16(3). <https://doi.org/10.1515/jib-2019-0005>
- Qiu, H., Novikov, A., & Vallon, V. (2017, July 1). Ketosis and diabetic ketoacidosis in response to SGLT2 inhibitors: Basic mechanisms and therapeutic perspectives. *Diabetes/Metabolism Research and Reviews*, Vol. 33, p. e2886. <https://doi.org/10.1002/dmrr.2886>
- Ren, S., Hinzman, A. A., Kang, E. L., Szczesniak, R. D., & Lu, L. J. (2015, December 1). Computational and statistical analysis of metabolomics data. *Metabolomics*, Vol. 11, pp. 1492–1513. <https://doi.org/10.1007/s11306-015-0823-6>
- Röst, H. L., Sachsenberg, T., Aiche, S., Bielow, C., Weisser, H., Aicheler, F., ... Kohlbacher, O. (2016, August 30). OpenMS: A flexible open-source software platform

for mass spectrometry data analysis. *Nature Methods*, Vol. 13, pp. 741–748.

<https://doi.org/10.1038/nmeth.3959>

- Rubio-Rodríguez, J. C., Reynoso-Camacho, R., Rocha-Guzmán, N., & Salgado, L. M. (2021). Functional beverages improve insulin resistance and hepatic steatosis modulating lysophospholipids in diet-induced obese rats. *Food Science and Nutrition*, 9(4), 1961–1971. <https://doi.org/10.1002/fsn3.2162>
- Smith, C. A., O'Maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., ... Siuzdak, G. (2005). METLIN: A metabolite mass spectral database. *Therapeutic Drug Monitoring*, 27(6), 747–751. <https://doi.org/10.1097/01.ftd.0000179845.53213.39>
- Spicer, R., Salek, R. M., Moreno, P., Cañueto, D., & Steinbeck, C. (2017, September 1). Navigating freely-available software tools for metabolomics analysis. *Metabolomics*, Vol. 13, p. 106. <https://doi.org/10.1007/s11306-017-1242-7>
- Suganuma, T., Swanson, S. K., Gogol, M., Garrett, T. J., Conkright-Fincham, J., Florens, L., ... Workman, J. L. (2018). MPTAC Determines APP Fragmentation via Sensing Sulfur Amino Acid Catabolism. *Cell Reports*, 24(6), 1585–1596. <https://doi.org/10.1016/j.celrep.2018.07.013>
- Sun, S., Wang, R., Song, J., Guan, M., Li, N., Zhang, X., ... Zhang, J. (2017). Blocking gp130 signaling suppresses autotoxin expression in adipocytes and improves insulin sensitivity in diet-induced obesity. *Journal of Lipid Research*, 58(11), 2102–2113. <https://doi.org/10.1194/jlr.M075655>
- Tangvarasittichai, S. (2015). Oxidative stress, insulin resistance, dyslipidemia and type 2 diabetes mellitus. *World Journal of Diabetes*, 6, 456–480.
- Vanderperre, B., Bender, T., Kunji, E. R. S., & Martinou, J. C. (2015, April 1). Mitochondrial pyruvate import and its effects on homeostasis. *Current Opinion in Cell Biology*, Vol. 33, pp. 35–41. <https://doi.org/10.1016/j.ceb.2014.10.008>
- Vargas, E., & Carrillo Sepulveda, M. A. (2019). Physiology, Glucose Transporter Type 4 (GLUT4). In *StatPearls*. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/30726007>
- Wei, R., Wang, J., Su, M., Jia, E., Chen, S., Chen, T., & Ni, Y. (2018). Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Scientific Reports*, 8(1), 663. <https://doi.org/10.1038/s41598-017-19120-0>

- Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., ... Scalbert, A. (2018). HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Research*, 46(D1), D608–D617. <https://doi.org/10.1093/nar/gkx1089>
- World Health Organization. (2019). Classification of diabetes mellitus. In *Classification of diabetes mellitus*. [https://doi.org/10.5005/jp/books/12855\\_84](https://doi.org/10.5005/jp/books/12855_84)
- Xia, J., Psychogios, N., Young, N., & Wishart, D. S. (2009). MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Research*, 37(SUPPL. 2). <https://doi.org/10.1093/nar/gkp356>
- Zhang, T., Huang, F., Li, B., Huang, C., Xu, C., Lin, K., & Lin, D. (2020). NMR-based metabolomic analysis for the effects of Huiyang Shengji extract on rat diabetic skin ulcers. *Journal of Ethnopharmacology*, 261, 112978. <https://doi.org/10.1016/j.jep.2020.112978>

# **Anexos**

# Anexo A

Tabla 5. Compuestos detectados por MS/MS relacionados con Diabetes tipo II.

Compound	m/z	RT	Frag1	Frag2	Formula	Sample	Reference	Comments
Aspartic acid	132.0304	0.88	NR	NR	C <sub>4</sub> H <sub>7</sub> NO <sub>4</sub>	Liver	Jiang P. (2019).	Aminoacyl-tRNA biosynthesis
Proline	116.0700	0.91	NR	NR	C <sub>5</sub> H <sub>9</sub> NO <sub>2</sub>	Liver	Jiang P. (2019).	Arginine and proline metabolism
Tyrosine	182.0802	0.95	NR	NR	C <sub>9</sub> H <sub>11</sub> NO <sub>3</sub>	Liver	Jiang P. (2019).	Tyrosine metabolism
Pyroglutamic acid	128.0356	0.99	NR	NR	C <sub>5</sub> H <sub>7</sub> NO <sub>3</sub>	Liver	Jiang P. (2019).	Glutathione metabolism
Taurine	126.0210	1.01	NR	NR	C <sub>2</sub> H <sub>7</sub> NO <sub>3</sub> S	Liver	Jiang P. (2019).	Bile acid biosynthesis
Glutathione	308.0894	1.03	NR	NR	C <sub>10</sub> H <sub>17</sub> N <sub>3</sub> O <sub>6</sub> S	Liver	Jiang P. (2019).	Glutathione metabolism
L-Carnitine	162.1116	1.11	NR	NR	C <sub>7</sub> H <sub>15</sub> NO <sub>3</sub>	Liver	Jiang P. (2019).	Fatty acid metabolism
L-Methionine	150.0573	1.17	NR	NR	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub> S	Liver	Jiang P. (2019).	Cysteine and methionine metabolism; aminoacyl-tRNA
L-Acetylcarnitine	204.1219	1.17	NR	NR	C <sub>9</sub> H <sub>17</sub> NO <sub>4</sub>	Liver	Jiang P. (2019).	Fatty acid metabolism
3-Hydroxybutyric acid	103.0402	1.23	NR	NR	C <sub>4</sub> H <sub>8</sub> O <sub>3</sub>	Liver	Jiang P. (2019).	Synthesis and degradation of ketone bodies
Betaine	118.0852	1.34	NR	NR	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub>	Liver	Jiang P. (2019).	Betaine metabolism
Alanine	90.0543	1.45	NR	NR	C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub>	Liver	Jiang P. (2019).	Pyruvate metabolism; amino acids metabolism
Taurocholic acid	514.2832	3.85	NR	NR	C <sub>26</sub> H <sub>45</sub> NO <sub>7</sub> S	Liver	Jiang P. (2019).	Primary bile acid biosynthesis; bile acid metabolism
Tauroursodeoxycholic acid	498.2886	4.25	NR	NR	C <sub>26</sub> H <sub>45</sub> NO <sub>6</sub> S	Liver	Jiang P. (2019).	Primary bile acid biosynthesis; bile acid metabolism
Glycocholic acid	464.3009	5.02	NR	NR	C <sub>26</sub> H <sub>43</sub> NO <sub>6</sub>	Liver	Jiang P. (2019).	Bile acid biosynthesis; bile acid metabolism
Glycoursodeoxycholic acid	448.3063	5.04	NR	NR	C <sub>26</sub> H <sub>43</sub> NO <sub>5</sub>	Liver	Jiang P. (2019).	Primary bile acid biosynthesis; bile acid metabolism
Cholic acid	407.2793	6.06	NR	NR	C <sub>24</sub> H <sub>40</sub> O <sub>5</sub>	Liver	Jiang P. (2019).	Bile acid biosynthesis; bile acid metabolism
Piruvic acid	88.4301	7.43	NR	NR	C <sub>3</sub> H <sub>4</sub> O <sub>3</sub>	Liver	Jiang P. (2019).	Bile acid biosynthesis; bile acid metabolism
LysoPE(18:2)	478.2903	9.26	NR	NR	C <sub>23</sub> H <sub>44</sub> NO <sub>7</sub> P	Liver	Jiang P. (2019).	Glycerophospholipid metabolism
Lactic acid	90.3977	10.37	NR	NR	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub>	Liver	Jiang P. (2019).	Glycerophospholipid metabolism
Sphingosine	300.2883	9.52	NR	NR	C <sub>18</sub> H <sub>37</sub> NO <sub>2</sub>	Liver	Jiang P. (2019).	Phospholipid metabolism
LysoPC(18:1)	522.353	11.07	NR	NR	C <sub>26</sub> H <sub>52</sub> NO <sub>7</sub> P	Liver	Jiang P. (2019).	Glycerophospholipid metabolism
L-Palmitoylcarnitine	400.3398	11.85	NR	NR	C <sub>23</sub> H <sub>45</sub> NO <sub>4</sub>	Liver	Jiang P. (2019).	Fatty acid metabolism

Compound	m/z	RT	Frag1	Frag2	Formula	Sample	Reference	Comments
LysoPC(15:0)	482.3217	12.65	NR	NR	C <sub>23</sub> H <sub>48</sub> NO <sub>7</sub> P	Liver	Jiang P. (2019).	Glycerophospholipid metabolism
LysoPC(18:0)	524.367	12.71	NR	NR	C <sub>26</sub> H <sub>54</sub> NO <sub>7</sub> P	Liver	Jiang P. (2019).	Glycerophospholipid metabolism
Stearoylcarnitine	428.3708	13.59	NR	NR	C <sub>25</sub> H <sub>49</sub> NO <sub>4</sub>	Liver	Jiang P. (2019).	Fatty acid metabolism
Docosahexaenoic acid	327.2324	13.95	NR	NR	C <sub>22</sub> H <sub>32</sub> O <sub>2</sub>	Liver	Jiang P. (2019).	Fatty acid metabolism
PC 1-acyl 32:1	732.560	22.687	NR	NR	C <sub>40</sub> H <sub>80</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
LPE 1-acyl 22:5	528.314	12.318	NR	NR	C <sub>27</sub> H <sub>46</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
Vaccenylcarnitine	426.375	12.689	NR	NR	C <sub>25</sub> H <sub>47</sub> NO <sub>4</sub>	Liver	Huang, Q. (2011).	Fatty acid oxidation
LPC 1-acyl 16:1	494.339	10.943	NR	NR	C <sub>24</sub> H <sub>48</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
LPC 1-acyl 22:5	570.35	12.084	NR	NR	C <sub>30</sub> H <sub>52</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
LPE 1-acyl 18:2	478.308	11.258	NR	NR	C <sub>22</sub> H <sub>46</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
LPE 1-acyl 20:4	502.303	11.807	NR	NR	C <sub>25</sub> H <sub>44</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
PC 1-acyl 38:5	808.605	23.216	NR	NR	C <sub>46</sub> H <sub>82</sub> NO <sub>8</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
LPC 1-acyl 18:2	520.342	11.34	NR	NR	C <sub>26</sub> H <sub>50</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
SM 2-amido 16:0	703.571	21.194	NR	NR	C <sub>39</sub> H <sub>81</sub> N <sub>2</sub> O <sub>6</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
SM 2-amido 16:1	701.557	20.157	NR	NR	C <sub>39</sub> H <sub>77</sub> N <sub>2</sub> O <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
LPE 1-acyl 20:0	510.365	13.123	NR	NR	C <sub>25</sub> H <sub>52</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
PE 1-acyl 36:2	744.559	24.083	NR	NR	C <sub>42</sub> H <sub>80</sub> NO <sub>8</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
LPE 1-acyl 22:0	538.391	15.032	NR	NR	C <sub>27</sub> H <sub>56</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
PC 1-acyl 33:2	744.557	22.407	NR	NR	C <sub>26</sub> H <sub>52</sub> NO <sub>8</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
PC 1-acyl 31:0	720.551	22.934	NR	NR	C <sub>39</sub> H <sub>78</sub> NO <sub>8</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
PC 1-alk 40:6	820.607	23.269	NR	NR	C <sub>48</sub> H <sub>86</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
LPC 1-acyl 15:0	482.34	11.195	NR	NR	C <sub>23</sub> H <sub>48</sub> NO <sub>7</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
PC 1-acyl 38:7	804.572	21.676	NR	NR	C <sub>46</sub> H <sub>78</sub> NO <sub>8</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
Leucine	131.8429	22.934	NR	NR	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	Liver	Huang, Q. (2011).	Tyrosine metabolism
Orotic acid	156.0482	26.676	NR	NR	C <sub>5</sub> H <sub>4</sub> N <sub>2</sub> O <sub>4</sub>		Huang, Q. (2011).	Pirimidine metabolism
AMP	348.077	1.105	NR	NR	C <sub>10</sub> H <sub>14</sub> N <sub>5</sub> O <sub>7</sub> P	Liver	Huang, Q. (2011).	Purine metabolism
PC 1-acyl 36:1	772.6	23.518	NR	NR	C <sub>44</sub> H <sub>86</sub> NO <sub>8</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
PC 1-alk 36:2	772.625	24.391	NR	NR	C <sub>44</sub> H <sub>84</sub> NO <sub>8</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
Inositol*	C <sub>6</sub> H <sub>12</sub> O <sub>6</sub>	180.631	NR	NR	C <sub>6</sub> H <sub>12</sub> O <sub>6</sub>		Huang, Q. (2011).	Inositol phosphate metabolism
PC 1-alk 38:6	792.573	22.191	NR	NR	C <sub>46</sub> H <sub>80</sub> NO <sub>8</sub> P	Liver	Huang, Q. (2011).	Lipid metabolism
GCA	466.333	6.299	NR	NR	C <sub>26</sub> H <sub>43</sub> NO <sub>6</sub>	Liver	Huang, Q. (2011).	Bile acid biosynthesis

<b>Compound</b>	<b>m/z</b>	<b>RT</b>	<b>Frag1</b>	<b>Frag2</b>	<b>Formula</b>	<b>Sample</b>	<b>Reference</b>	<b>Comments</b>
Glutamine	146.0336	12.299	NR	NR	C <sub>5</sub> H <sub>10</sub> N <sub>2</sub> O <sub>3</sub>	Liver	Huang, Q. (2011).	Bile acid biosynthesis
Serine	105.035	22.3734	NR	NR	C <sub>3</sub> H <sub>7</sub> NO <sub>3</sub>	Liver	Huang, Q. (2011).	Lipid metabolism
FFA C22:5	329.258	17.588	NR	NR	C <sub>22</sub> H <sub>34</sub> O <sub>2</sub>	Liver	Huang, Q. (2011).	Fatty acid biosynthesis
FFA C22:4	331.272	18.128	NR	NR	C <sub>22</sub> H <sub>36</sub> O <sub>2</sub>	Liver	Huang, Q. (2011).	Fatty acid biosynthesis
Ascorbic acid	175.015	0.788	NR	NR	C <sub>6</sub> H <sub>8</sub> O <sub>6</sub>	Liver	Huang, Q. (2011).	Acorbate and aldarate metabolism
FFA C18:0	283.238	16.504	NR	NR	C <sub>18</sub> H <sub>35</sub> O <sub>2</sub>	Liver	Huang, Q. (2011).	Fatty acid biosynthesis
AMP	346.063	1.102	NR	NR	C <sub>10</sub> H <sub>14</sub> N <sub>5</sub> O <sub>7</sub> P	Liver	Huang, Q. (2011).	Purine metabolism

Tabla 6. Abundancias relativas de los compuestos identificados por grupo.

<b>COMPOUNDS</b>							
<b>ID</b>	<b>LysoPA(0:0/18:1(9z))</b>	<b>Acetoacetyl-CoA</b>	<b>Norrubrofusarin-6-beta-gentiobiosido</b>	<b>Molybdopterina Z</b>	<b>(17R(18S)-EpETE)</b>	<b>5-O-Mycaminosyprotylonolide</b>	<b>Lolitriol</b>
<b>B4</b>	5.90695371	0.709309999	2.340062117	0.194694132	47.3397765	106.6060973	95.587695
<b>B5</b>	6.008679338	0.016490671	0.891117226	0.069050169	126.971722	286.7763937	288.510595
<b>B7</b>	4.884816513	0.011697308	NA	0.017309748	13.0137775	86.56502973	86.6205985
<b>BF1</b>	354.5132554	19.02985549	0.334693485	7.79526007	NA	2.098005652	1.22741699
<b>BF2</b>	479.1587012	8.638327087	0.138767398	2.120992068	0.01516726	4.525117054	12.8932526
<b>BF3</b>	584.3078704	18.46934338	0.038013412	6.182462856	18.0086991	5.175409106	29.9107812
<b>BF4</b>	672.115488	23.28725666	0.479002972	11.30255741	17.8746643	83.77132048	66.6514782
<b>BF5</b>	845.1685727	20.19118691	2.077464265	10.07930175	8.6055309	139.5254936	118.458191
<b>BF6</b>	663.8947063	23.82074961	2.643018641	5.495691572	NA	3.886474609	3.04345703
<b>BF7</b>	781.3179889	28.15423623	2.610918985	4.019509557	3.09646608	5.01113163	4.60405224
<b>O1</b>	600.8969883	32.16999532	0.220878677	30.62240468	3.69654848	3.212646484	2.07723999
<b>O2</b>	695.6763914	29.1258422	0.076069263	10.63640214	21.42305	106.3645383	86.893446
<b>O3</b>	1196.543178	27.95722277	1.985718914	16.23282175	0.95928955	19.57098607	14.9649264
<b>O4</b>	114.9696584	0.874931976	0.814397611	3.249487208	1.11453247	8.891621252	5.18600816
<b>O5</b>	6.406551793	2.27232156	NA	0.173756403	3.69503784	93.16404864	61.3922048
<b>R1</b>	454.0146715	26.64888352	0.056159945	29.48462799	17.1536043	40.12346292	25.2985214
<b>R5</b>	7.508907983	0.176593422	0.003221689	0.166701667	122.542226	488.3500579	409.644551
<b>R7</b>	234.4835966	0.081517536	0.145843416	1.980267712	30.4091872	149.8573782	123.738651
<b>R8</b>	5.102766972	0.31782219	1.35569786	0.303415288	21.410824	107.2969216	101.585812
<b>RF1</b>	889.4382199	24.98521336	0.399538362	5.725668106	5.86462405	16.75906566	11.6098328
<b>RF2</b>	1225.9682	28.16466347	1.200260676	26.66715236	2.00927734	12.1604632	9.76382446
<b>RF3</b>	750.5649984	22.39637202	2.032470687	7.181526375	1.53617859	3.08102417	2.19689941
<b>RF4</b>	688.2776211	19.75325819	0.56269852	28.15188764	4.87161255	25.16238403	22.5290863

<b>RF5</b>	777.1439447	30.81803126	2.034503034	53.79525872	25.239141	186.0280991	153.707988
<b>RF6</b>	1080.697102	26.24775589	0.891222561	16.01734866	0.28330994	1.984497049	2.05780035
<b>S2</b>	28.57259781	0.125877715	1.870197115	0.199183666	14.6150135	8.599726263	5.78518677
<b>S4</b>	6.301583146	0.244989327	0.769917679	0.589290385	2.82531739	44.50828901	39.2955516
<b>S5</b>	5.428957377	0.078390882	0.648834673	0.016666221	5.22061546	17.38539748	13.9413065
<b>S6</b>	4.533474584	NA	NA	0.059749157	21.145252	106.5631136	104.146096
<b>V3</b>	11.82817844	0.168360383	1.032933726	0.324568529	30.0918463	133.5923907	113.611574
<b>V4</b>	7.514504661	0.178766461	0.180392393	0.049768249	34.2714344	66.61064919	53.2914596
<b>V7</b>	5.438091314	0.391719411	NA	0.033832841	13.4014818	86.74529933	89.8123116

---