

Universidad San Pablo – CEU
CEINDO – CEU Escuela Internacional de
Doctorado

PROGRAMA en CIENCIA Y TECNOLOGÍA DE LA SALUD



CEU

*Escuela Internacional
de Doctorado*

**Desafíos analíticos en estudios
metabolómicos de cáncer:
búsqueda de biomarcadores**

TESIS DOCTORAL

Presentada por:

Ángeles López López

Dirigida por:

Dra. Coral Barbas Arribas

Dra. M^a Ángeles López González

MADRID
2022

*“Debemos tener constancia y,
sobre todo, confianza en
nosotros mismos. Debemos creer
que estamos dotados para algo.”*

Marie Skłodowska-Curie

AGRADECIMIENTOS

El trabajo llevado a cabo y que ha dado lugar a la presentación de los resultados en esta Tesis Doctoral se ha realizado en el Centro de Metabolómica y Bioanálisis (CEMBIO) de la Universidad San Pablo CEU de Madrid (España) financiado por la Comunidad de Madrid y la Escuela Internacional de Doctorado CEU (CEINDO).

Los proyectos realizados surgen de diferentes colaboraciones con varios centros de investigación, incluidos el Instituto de Investigación de la Unidad Gastrointestinal del Hospital Universitario 12 de Octubre de Madrid (imas12), el Centro de Investigación Español de Cáncer (CNIO), el Centro Nacional de Biotecnología (CNB) y el Centro de Investigación Clínica de la Universidad Médica de Białystok en Polonia.

En estas líneas me gustaría poder agradecer a todas esas personas quienes de manera directa o indirecta han contribuido a la realización de este trabajo, desde los coautores y colaboradores como a mis compañeros del CEMBIO y mis supervisoras.

En primer lugar, me gustaría agradecer a mis directoras de tesis, Coral Barbas y M^a Angeles López González, por darme la oportunidad de incorporarme a su grupo de investigación. Gracias por toda vuestra ayuda, guía, motivación y paciencia, y en el ejemplo que he recibido por vuestra parte, tanto a nivel científico como en valores como el compromiso, la constancia, el esfuerzo y la responsabilidad que me han devuelto esta pasión por la investigación.

Me gustaría continuar agradeciendo de forma muy especial a mis compañeras, pero antes amigas, Vanesa Alonso y Carolina González, siempre me han acompañado en todo este camino, escuchándome, aconsejándome, ayudándome y apoyándome tanto en lo personal como en lo profesional. Gracias de corazón.

También me gustaría dar las gracias a Fernanda Rey-Stolle, Ana Gradillas, Alma Villaseñor, M^a Paz Martínez Alcázar, M^a Paz Lorenzo, José Miguel Cárdenas Rebollo, Santiago Angulo, Pilar Díaz Carrasco, David Rojo, Gloria Anemone, Javier Rupérez y Antonia García por todo el conocimiento y buenos momentos que han compartido conmigo durante estos años.

Durante este periodo he coincidido con mucha gente en el laboratorio del CEMBIO, gracias a todos ellos por el soporte, las discusiones científicas y los momentos divertidos. Con especial cariño a Paula, Carolina, David O., Borja, Cecilia y Elena por todo lo vivido junto a vosotros. Gracias por el apoyo científico, pero sobre todo por haber estado como amigos. Habéis hecho esta experiencia única.

A Joanna Godzien, gracias por todo lo que me has enseñado de la ciencia, del laboratorio, de los lípidos, de las oxidaciones, de Madrid, de Polonia, de convivencia y por encima de todo, de amistad. No quería pasar la oportunidad de agradecer al Prof. Michał Ciborowski y su grupo de investigación, permitirme la oportunidad de realizar una estancia internacional. Gracias por todo lo que aprendí y el apoyo a nivel personal que recibí. (*I wouldn't like to miss this opportunity to thank Prof. Michał Ciborowski and his research group, for allowing me the opportunity to carry out an international pre-doctoral scientific stay. Thank you for everything I learned and the support on a personal level*).

Han sido unos años con temporadas muy duras, y me siento tremendamente afortunada y agradecida a mi familia y amigos que siempre he contado con ellos y han valorado en todo momento mi trabajo. En especial a mis padres y mi hermana, gracias por animarme a comenzar esta etapa de mi vida, ayudándome y apoyándome en todo momento, no dejando que tirase la toalla y tendiéndome la mano siempre para que no cayese. Gracias Macarena por toda tu confianza y apoyo, incondicional e ilimitado, que me has ofrecido cada día sin excepción y con tanto amor. Gracias a Noelia, mi hermana mayor, que me ha regalado un sinfín de buenos momentos, siempre dándome ánimos a continuar; tú y yo sabemos que este trabajo también va por ella. A mis abuelos, mis tíos y primos. Gracias Teresa por recordarme siempre lo valiente que soy. A Gloria Curiel, mi profesora de química, por enseñarme la belleza de esta ciencia. Quiero agradecer a mis amigas, con especial cariño a Lara, Vir, Candi e Itziar, que han tenido siempre fe ciega en mí, gracias por vuestras palabras de aliento y los maravillosos momentos de desconexión. A Rafa, por tanta comprensión y tanto cariño, consiguiendo minimizar los problemas haciéndome sonreír, y ayudándome con las revisiones.

Gracias a todos ellos, por vuestra inmensa paciencia y vuestro gran apoyo.

ÍNDICE

Abreviaturas y acrónimos.....	15
Resumen.....	21
Objetivos Generales.....	27
Introducción	
- Introducción General – Metabolómica para la búsqueda de biomarcadores validados en estudios de cáncer.....	31
Capítulo 1	
- Perfil metabolómico plasmático de pacientes con tumores neuroendocrinos avanzados (NET). Diagnóstico y relevancia biológica.....	81
Capítulo 2	
- Desafíos analíticos e implicaciones biológicas de los lípidos oxidados en el perfil metabólico de tumores neuroendocrinos y cáncer de pulmón.....	147
Capítulo 3	
- Enfoque analítico para el estudio del perfil metabolómico de bilis e interpretación de las alteraciones biológicas en Colangiocarcinoma.....	201
Capítulo 4	
- Desarrollo de un método de análisis para compuestos aniónicos en CE–MS. Aplicación a un estudio metabolómico de cáncer de pulmón.....	255
Conclusiones.....	295
Lista de publicaciones científicas.....	301

Abreviaturas y acrónimos

A-CEHC , Alfa-carboxietilhidrocromano	CA , Antígeno carbohidrato	EAC , Adenocarcinoma esofágico
(d)MRM , Monitoreo de reacciones múltiples (dinámico)	CAR , Carnitina	EC , Cáncer de endometrio
5-HIA , Ácido 5-hidroxiindolacético	CC , Cáncer de cuello uterino	ECOG , Grupo de Oncología Cooperativa del Este
ACN , Acetonitrilo	CCA , Colangiocarcinoma	EFA , Ácido graso esterificado
ADC , Adenocarcinoma	CDKN1B , Inhibidor 1B de quinasa dependiente de ciclina	EHBC , Cáncer del conducto biliar extrahepático
ADN , Ácido desoxirribonucleico	CE , Electroforesis capilar	EI , Impacto electrónico
ADP , Adenosina difosfato	CEA , Antígeno carcinoembrionario	EIC , Cromatograma/electroferograma de iones extraídos
AEDT , Ácido etilendiaminotetraacético	CE-MS , Electroforesis capilar acoplado a la espectrometría de masas	EOC , Cáncer epitelial de ovario
AHT , Antihipertensivo	Cer , Ceramida	EPOC , Enfermedad pulmonar obstructiva crónica
AICAR , Ribonucleótido de amino imidazol carboxamida	CgA , Cromogranina A	ESCC , Carcinoma de células escamosas esofágicas
Ala , Alanina	CID , Disociación inducida por colisión	ESD , Displasia de células escamosas
AMP , Adenosina monofosfato	CMF , Fragmentación por migración de carga	ESI , Ionización electrospray
APAP , Paracetamol	CML , Leucemia mieloide crónica	EtOH , Etanol
APT , Antiplaquetario	CMM , Ceu Mass Mediator	F1,6BP , Fructosa 1,6-difosfato
Arg , Arginina	COX , Ciclooxygenasa	FA , Ácido graso
ARNm , Ácido ribonucleótido mensajero	CR , Respuesta completa	FAME , Ésteres metílicos de ácidos grasos
ARNt o tRNA , Ácido ribonucleótido de transferencia	CRC , Cáncer colorrectal	FbF , Filtrar compuestos por fórmula
ASAP , Análisis de sólidos a presión atmosférica	CRE , Creatinina	Fbl , Filtrado por ión
Asp , Asparagina	CRF , Fragmentación por retención de carga	FC , <i>del inglés fold-change</i> , razón entre medias
ASS1 , Argininosucinato sintasa 1	CV , Coeficiente de variación	FDR , Tasa de descubrimientos de falsos positivos
ATP , Adenosín trifosfato	Cys , Cisteína	FFA , Ácido graso libre
AUC , Área bajo la curva	CZE , Electroforesis de zona capilar	FIA , Análisis de inyección directa
BC , Cáncer de vejiga	DA , Ácidos dicarboxílicos	FTICR , Espectrometría de resonancia ciclotrónica por transformada de Fourier
BCAA , Aminoácidos de cadena ramificada	DB , Base de datos	G1 , Grado histopatológico 1
BD , Donador de sangre, control	DBS , <i>del inglés dried blood spot</i> , micromuestreo de sangre seca	G2 , Grado histopatológico 2
BGE , Electrolito de fondo	DC , Enfermedad-control.	G6P , Glucosa 6-fosfato
BRC , Cáncer de pecho	DG , Diacilglicérido	GCa , Cáncer gástrico
BSTFA , O-bis (trimetilsilil) trifluoroacetamida	DHAP , Dihidroxiacetona fosfato	GC , Cromatografía de gases
BTT , Tumor de tiroides benigno	DIA , Análisis de datos independiente	GCA , Ácido glicocólico
BTN , Nódulo tiroideo benigno	DIU , Diurético	
	DMA , Dietilamina	
	E4P , Eritrosa 4-fosfato	

GCDC , Ácido glucoquenodesoxicólico	LC , Cromatografía de líquidos	MLF , Fragmento específico de lípidos moleculares
GC-MS , Cromatografía de gases acoplado a la espectrometría de masas	LCC , Carcinoma de células grandes	MPA , Análisis de la ruta del metabolito
GDCA , Ácido glicodesoxicólico	LCF , Fragmento selectivo de clase lipídica	MS , Espectrometría de masas
Glu , Glutamato	LCMD , Base de datos del metaboloma de cáncer de pulmón	MS/MS , Espectrometría de masas en tándem
GLUC , Glucemia	LC-MS , Cromatografía de líquidos acoplado a la espectrometría de masas	MSEA , Análisis de enriquecimiento del conjunto de metabolitos
Gly , Glicina	LC-QqQ/MS , Cromatografía líquida acoplado con espectrometría de masas de triple cuadrupolo	MT , Tiempo de migración
GNM , Cáncer de glioblastoma multiforme	Leu , Leucina	MTBE , Metil-terc-butiletter
GPL , Glicerofosfolípidos	LG , Grado bajo	MTC , Cáncer de tiroides maligno
GSH , Glutación	LIT , Trampa lineal de iones	MVA , Análisis estadístico multivariante
GSSG , Glutación disulfuro	LLA , Leucemia linfoblástica aguda	MW , Masa molecular
HC , Controles sanos	LMSD , Base de datos de LIPID MAPS®	NA , No aplica
HCC , Carcinoma hepatocelular	LOX , Lipooxigenasa	NEN , Neoplasias neuroendocrinas
HETE , Ácido hidroxieicosatetraenoico	LPA , Ácido lisoglicerofosfatídico	NET , Tumores de origen neuroendocrino
HG , Grado alto	LPC , Lisoglicerofosfatidilcolina	NF1 , Gen neurofibromatosis tipo 1
HILIC , Cromatografía de interacción hidrófila	LPE , Lisoglicerofosfatidiletanolamina	NIST , Instituto Nacional de Normas y Tecnología
HMDB , Base de datos del metaboloma humano	LPI , Lisoglicerofosfatidilinositol	NMR , Resonancia magnética nuclear
HPV , Virus del papiloma humano	LPS , Lisoglicerofosfatidilserina	NP , No procede
HR-MAS , Espectroscopía de giro de ángulo mágico de alta resolución	LRM , Modelos de regresión logística	NS , No especificado
Hyp , Hidroxiprolina	Lys , Lisina	NSAIDs , Antiinflamatorio no esteroideo.
IC , Intervalo de confianza	m/z , Relación masa-carga	NSCLC , Cáncer de pulmón de células no pequeñas
ID , Identificación	MEN , Neoplasia endocrina múltiple	OC , Cáncer de ovario
IDP , Inosina difosfato	MeOH , Metanol	ONJ , Osteonecrosis de la mandíbula
IP , Vía de administración intraperitoneal	MES , Ácido 2-(N-morfolino) etanosulfónico	OPLS-DA , Análisis discriminante ortogonal por mínimos cuadrados parciales
iPrOH , Isopropanol	MetS , Metionina sulfona	Orn , Ornitina
IS , Patrón interno	MFE , Extracción de características moleculares	OSCC , Carcinoma oral de células escamosas
ISF , Fragmentación en fuente de ionización	MG , Monoacilglicérido	oxFA , Ácidos grasos oxidados
JK , Intervalo de confianza Jackknife		oxPC(s) , Glicerofosfatidilcolina(s) oxidada(s)
KC , Cáncer de riñón		
kNN , <i>del inglés, k-nearest neighbors</i> , "vecinos más cercanos"		
LCa , Cáncer de pulmón		

PaC , Cáncer de páncreas	r² , Coeficientes de correlación	TC , Cáncer de tiroides
PBM , Malfuncionamiento pancreaticobiliar	RAN , Ranitidina	TCA , Ciclo del citrato o ciclo de los ácidos tricarboxílicos
PC , Glicerofosfatidilcolina	RCC , Carcinoma de células renales	TCAc , Ácido taurocólico
PCA , Análisis de componentes principales	RFE , Extracción de características recursiva	TCDCa , Ácido tauroquenodesoxicólico
PCa , Cáncer de próstata	RI , Índice de retención	TDCA , Ácido taurodesoxicólico
PCr , Pancreatitis crónica	rMT , Tiempo de migración relativo	TFA , Ácido trifluoroacético
PD , Progresión de la enfermedad	RNS , Especies reactivas de nitrógeno	TFE , Extracción de características dirigida o <i>target</i>
PDAC , Adenocarcinoma ductal pancreático	ROC , Características operativas del receptor	THF , Tetrahydrofolato
PE , Glicerofosfatidiletanolamina	ROS , Especies reactivas de oxígeno	Thr , Treonina
PG , Glicerofosfatidilglicerol	RSD , Desviación estándar relativa	TIC , Cromatograma/electroferograma de iones totales
PGD , Disfunción primaria del injerto	RSM , Método de superficie de respuesta	TKI , Inhibidor de tirosina quinasa
Phe , Fenilalanina	RT , Tiempo de retención	TMCS , Trimetilclorosilano
PL , Fosfolípidos	RTL , <i>del inglés retención time looking</i> , bloqueo del tiempo de retención	TOF , Analizador de tiempo de vuelo
PLS-DA , Análisis discriminante por mínimos cuadrados parciales	Ru5P , Ribulosa 5-fosfato	TSC , Proteína de esclerosis tuberosa
PPI , Inhibidor de la bomba de protones	S/N , Relación señal-ruido	Tyr , Tirosina
PR , Respuesta parcial	S1P , Esfingosina 1-fosfato	UDC , Ácido ursodesoxicólico
Pro , Prolina	SCC , Carcinoma de células escamosas	UHPLC , Cromatografía líquida de ultra alta eficacia
PS , Glicerofosfatidilserina	SCLC , Cáncer de pulmón de células pequeñas	ULN , Límite superior de normalidad
PTC , Carcinoma papilar de tiroides	SD , Enfermedad estable	UVA , Análisis estadístico univariante
PVA , Alcohol polivinílico	SEOM , Sociedad española de oncología médica	Val , Valina
Q , Cuadrupolo	Ser , Serina	VCC , Compuesto carbonílico volátil
Q² , Bondad de predicción del modelo	SL , Líquido auxiliar	VEGF , Factor de crecimiento endotelial vascular
QA , Garantía de calidad	SM , Esfingomielina	VIP , Importancia de la variable en la proyección
QC , Control de calidad	SMART , Análisis del subconjunto de puntos de tiempo repetidos coincidentes	VOC , Compuesto orgánico volátil
QC-SVRC , Control de calidad y corrección de regresión de vectores de soporte	SMPD , Base de datos de rutas de moléculas pequeñas	
QIT , Cuadrupolo-trampa de iones	SPB , Base esfingoidea	
QqQ , Analizador de masas de triple cuadrupolo	SQ , Cuadrupolo sencillo	
QTOF , Analizador de masas cuadrupolo-tiempo de vuelo	STAT , Estatinas	
R² , Bondad de ajuste del modelo	STD , Desviación estándar	
	STE , Esteroides	

RESUMEN

El término cáncer engloba un numeroso grupo de enfermedades que se localiza dentro de las enfermedades humanas más devastadora que se conoce, responsable de un gran número de muertes cada año. Nos encontramos ante una enfermedad de la que se conocen más de 100 tipos y se definen, la mayor parte, por el nombre del órgano donde comienza el cáncer, aunque también se pueden identificar por el tipo de célula que los forma. Según los datos de incidencia, prevalencia y mortalidad de cáncer en nuestro país recogidos en el último informe anual por la Sociedad Española de Oncología Médica (SEOM, <https://seom.org/>) se estima una previsión de 27 millones de casos para el año 2040, siendo los tumores más frecuentes a nivel mundial los de pulmón, mama, colorrectal, próstata, estómago e hígado.

La ciencia encargada del estudio del metaboloma de un sistema, la metabolómica, última pieza de las principales plataformas ómicas, ofrece una visión instantánea del estado biológico. Los cambios en el metaboloma generan un rastro único de los procesos celulares que están teniendo lugar tanto en fluidos como en tejidos biológicos, por lo que aplicado a una enfermedad como la planteada en la presente tesis doctoral, permite informar sobre las bases moleculares del cáncer. Además, el enfoque intensivo que ofrece la metabolómica permite su uso para la identificación de características metabólicas específicas como potenciales biomarcadores mejorando el diagnóstico o la posible respuesta a un tratamiento, definiendo el término “biomarcador” como la característica medible que actúa como indicador de procesos biológicos, patogénicos o farmacológicos.

En la presente tesis se ha perseguido el desarrollo de nuevas metodologías o mejora de los métodos ya existentes para el análisis de distintas matrices biológicas que permitan obtener métodos analíticos optimizados, a través de las plataformas de electroforesis capilar (CE), cromatografía de líquidos (LC) y cromatografía de gases (GC) acopladas todas ellas a la espectrometría de masas (MS). Esta metodología se ha aplicado a diferentes estudios de cáncer abarcando una gran variedad de metabolitos que nos informan del estado alterado en esta enfermedad y nos han permitido establecer y validar, en algunas ocasiones, un potencial panel biomarcador.

En la “**Introducción General**” se recoge una descripción general de los diferentes estudios metabolómicos llevados a cabo en muestras biológicas basados en la búsqueda de biomarcadores validados en cáncer en los últimos, casi, diez años (2013 a enero 2022), resultado de una amplia revisión bibliográfica. La información obtenida ha sido clasificada en función del tipo de cáncer u órgano afectado, las cohortes de pacientes utilizadas tanto en la fase de descubrimiento como de validación, el tipo de

muestra biológica o la/s plataforma/s de análisis empleadas. A la hora de reportar estos resultados, debemos asegurarnos de que se cumplan unos criterios mínimos de validez, por lo que utilizar cohortes relativamente grandes e independientes o plataformas de análisis distintas para la validación de los resultados obtenidos ha sido uno de los criterios de selección. Por ello, en el siguiente capítulo, **capítulo 1**, se ha llevado a cabo el estudio del perfil metabolómico plasmático de pacientes con **tumores neuroendocrinos avanzados** (NET). Con el objetivo de comprender mejor la desregulación metabólica en estos tumores e identificar biomarcadores para un posible uso clínico, se realizaron los análisis multiplataforma (GC-MS, LC-MS y CE-MS) no dirigidos de una cohorte extensa de pacientes con NET e individuos sin cáncer (control). Además, se evaluó el potencial diagnóstico y la relevancia biológica de los metabolitos diferenciales encontrados con la exploración de las vías desreguladas. Finalmente se realizó el análisis dirigido de los metabolitos discriminantes de NET seleccionados a través de una plataforma de análisis diferente a la empleada en los análisis iniciales. Los resultados han revelado la significancia funcional de los metabolitos seleccionados con varias vías alteradas por oncogénesis, y algunas, más concretamente, en el desarrollo específico de NET como la angiogénesis, la vía mTOR, el metabolismo del triptófano, el estrés oxidativo o el ciclo de la urea, entre otros.

La correcta asignación de la identidad de esas posibles características que discriminen individuos caso de control es una etapa fundamental en el flujo de trabajo planteado en la mayoría de los estudios no dirigidos de metabolómica. Es por ello por lo que el **capítulo 2**, recoge los ecos de la existente preocupación por una anotación satisfactoria para la multitud de “nuevos” compuestos derivados de otros ya conocidos, muchos de ellos por ligeras modificaciones y, ofrece un flujo de trabajo optimizado para la asignación de **identidades correctas de lípidos oxidados** que no dispone de una identidad correcta en las bases de datos disponibles actualmente. Con la aplicación de este flujo ha sido posible la identificación de diferentes formas isoméricas de derivados hidroxilados del ácido araquidónico (HETE) y de tres formas oxidadas de lisoglicerofosfatidilcolinas (LPC) significativas en muestras de plasma de **pacientes con NET**. Asimismo, se ha trabajado en el desarrollo de una metodología aplicada al análisis de los diferentes productos de oxidación de estándares de lípidos comerciales que permitan caracterizar cada forma con su tiempo de retención y fragmentos característicos procedentes del análisis MS/MS. La información recolectada ha permitido la creación de una base de datos que ayude al investigador en la identificación de estas modificaciones lipídicas. Complementariamente, se ha explorado las

consecuencias biológicas de la aparición de diferentes formas oxidadas producto de las principales glicerofosfatidilcolinas (PC) en datos procedentes de un estudio de **cáncer de pulmón** empleando muestras de plasma y tejido de pacientes pertenecientes a distintos tipos y distintos estadios de la enfermedad. De los subtipos de cáncer de pulmón de células no pequeñas (NSCLC), ha sido posible discriminar adenocarcinoma (ADC) y cáncer de pulmón de células escamosas (SCC) de individuos control, respectivamente, a partir de glicerofosfatidilcolinas oxidadas (oxPC) encontradas tanto en plasma como en tejido. Se ha podido observar tendencias preliminares para derivados oxidados de cadena larga en plasma, y derivados oxidados de cadena truncada en tejido pulmonar, además de una marcada disminución en la abundancia de estos derivados oxidados a medida que aumenta la enfermedad, posiblemente explicado por la apoptosis generada en estadios avanzados de la misma.

La selección del tipo de muestra empleada para el estudio de la consecuente cascada de alteraciones metabólicas en cáncer es otro de los puntos críticos en la metabolómica no dirigida. Las matrices biológicas empleadas en estudios de metabolómica son muy diversas y van desde fluidos fácilmente accesibles como la orina, sangre o heces hasta órganos, tejidos o células. A la hora de elegir el tipo de matriz biológica empleada para el estudio de cualquier patología es importante conocer el tipo de información que vamos a obtener de su análisis. El **capítulo 3** aborda el empleo de **muestras de bilis** para el estudio de un tipo de cáncer que surge del epitelio de los conductos biliares, el **colangiocarcinoma (CCA)**, y que, a pesar del carácter más invasivo de este tipo de muestra, los resultados han demostrado que las bilis son una rica fuente de biomarcadores específicos de cáncer de hígado, vías biliares, vesicular biliar y páncreas. Como nos encontramos ante una matriz biológica menos explorada, ha sido necesario el desarrollo de un protocolo de análisis a través de CE-MS que ha permitido caracterizar el perfil metabólico de la bilis a través de esta plataforma y que, junto con el análisis de las muestras por LC-MS, ha permitido establecer la importancia del metabolismo alterado de ácidos biliares, fosfatidilcolinas y aminoácidos en individuos con CCA, y la representación del estado redox celular a través de la semicuantificación de varios de los compuestos responsables.

En la última década, ha sido ampliamente utilizada la CE-MS para el estudio metabolómico de compuestos polares catiónicos, siendo minoritarios los estudios que reportan compuestos polares aniónicos. La experiencia nos ha demostrado que los capilares empleados hasta la fecha con este fin presentan algunas desventajas, por lo que consideramos que hay, aún, cabida para implantar un método optimizado para el

análisis de compuestos aniónicos. Es así como, en el **capítulo 4**, se recoge el trabajo realizado durante estos años de investigación para el desarrollo e implementación de un método **CE-MS para compuestos aniónicos** en matrices biológicas. En él se estudia un nuevo capilar de polivinil alcohol (PVA) con el fin de ampliar las posibilidades de esta técnica y se compara con otros capilares más comunes. La optimización y desarrollo de esta metodología ha permitido la confección de una base de datos de más de doscientos metabolitos que se empleó para la caracterización del **plasma sanguíneo de NSCLC** y, el estudio comparativo de los dos subtipos más frecuentes, ADC y SCC *versus* controles. Los resultados obtenidos indican que los grupos estudiados pueden ser discriminados en función del perfil aniónico y se han encontrado cambios significativos en nuevos metabolitos relacionados con el cáncer que no se conocían previamente, como ciertos derivados de aminoácidos y ácidos dicarboxílicos.

Consideramos que tanto las metodologías desarrolladas en la presente tesis como su aplicación en diferentes estudios de cáncer han contribuido al avance del conocimiento en los campos analítico y bioquímico.

OBJETIVOS GENERALES

Durante la última década, los avances en metabolómica se han focalizado en la identificación de las características específicas de una enfermedad. La identificación y cuantificación de estos potenciales biomarcadores es una poderosa herramienta para el diagnóstico precoz de enfermedades y/o indicadores de la respuesta ante un tratamiento. Validar un potencial panel de biomarcadores puede implicar el análisis no dirigido, con todo el diseño que este conlleva, hasta la obtención de un escenario discriminante entre los grupos de análisis, seguido de la evaluación de ese panel, estableciendo los márgenes que ofrecen resultados de mayor eficacia, sensibilidad, especificidad y reproducibilidad. Esto podría simplificarse en validación analítica, validación clínica y validación con ensayos clínicos. La simple combinación de una lista corta de candidatos tiene como objetivo ofrecer un resultado predictivo efectivo y, debe estudiarse su utilidad clínica para asegurar que el ensayo tenga como objetivo final la mejora de resultados para los pacientes.

Sin embargo, este es un proceso largo, complejo y costoso, que debido a la gran diversidad fisicoquímica ante la que nos encontramos en el análisis en el caso de los metabolitos, a menudo es necesario mejorar las herramientas con las que nos aproximamos al estudio de las distintas muestras.

Los objetivos globales de esta tesis son el desarrollo y aplicación de diferentes herramientas analíticas para el estudio metabolómico de distintos tipos de cáncer, utilizando diversas matrices biológicas y empleando tres plataformas analíticas acopladas a la espectrometría de masas: CE-MS, LC-MS y GC-MS, con el objetivo de establecer posibles candidatos a biomarcadores.

Para el desarrollo del objetivo general de esta tesis, se han propuesto diferentes objetivos específicos:

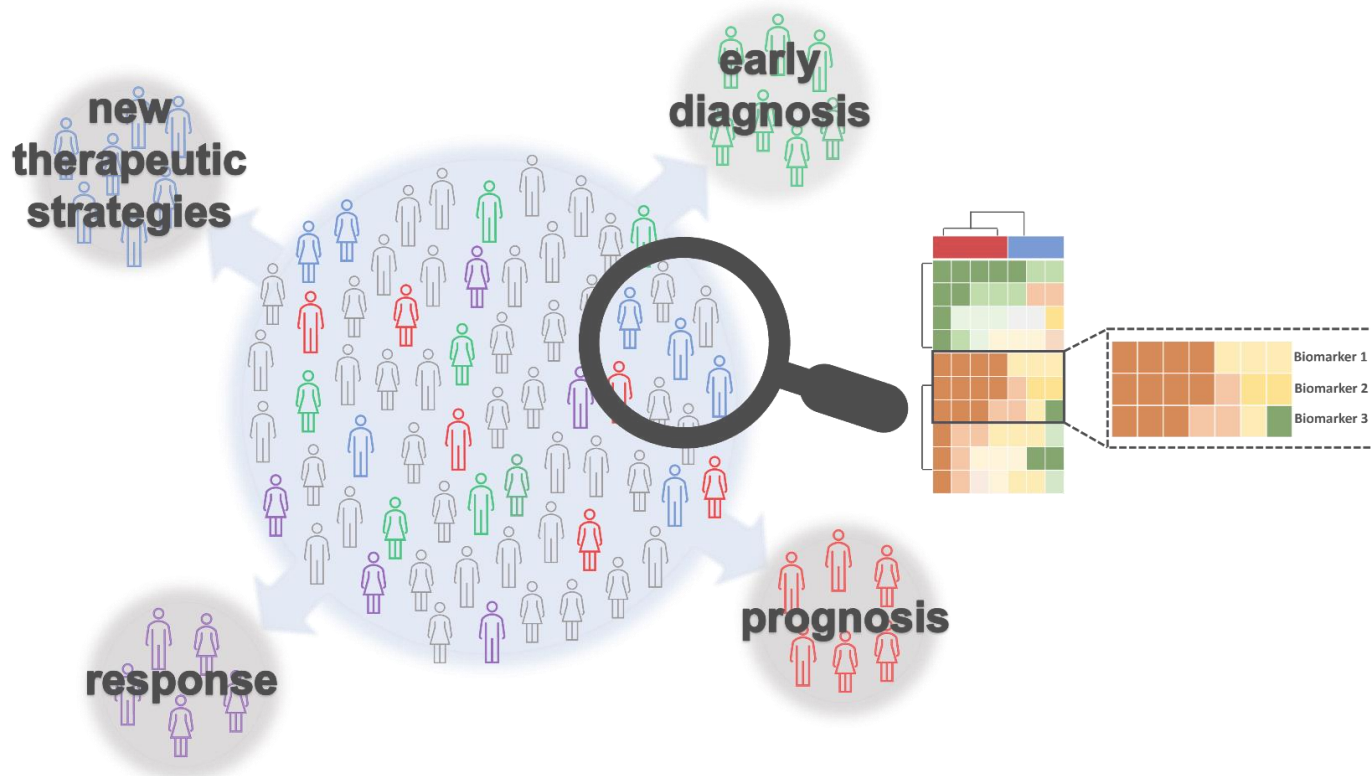
1. Optimizar las condiciones de trabajo para estudios con muestras de bilis y evaluar la huella metabólica del colangiocarcinoma. Se establecerán las condiciones adecuadas para el estudio metabolómico no dirigido incluyendo las etapas del método analítico y se aplicarán en el estudio metabolómico basado en CE-MS y LC-MS. A partir de los datos obtenidos, se tratará de esclarecer los procesos metabólicos alterados intentando obtener el máximo de información posible para, en futuros estudios, continuar en la búsqueda de un potencial panel biomarcador que ayude en la detección precoz de este tipo de cáncer.

2. Realizar el estudio metabolómico de tumores de origen neuroendocrino y la búsqueda de potenciales biomarcadores. Se realizará el análisis no dirigido multiplataforma para el estudio de la huella metabólica en muestras de plasma de

pacientes con tumores de origen neuroendocrino (NET). Se realizará la interpretación biológica de los resultados obtenidos y se intentará establecer un panel de biomarcadores de la enfermedad, realizando su validación analítica a través de otra técnica.

3. Establecer un flujo de trabajo para la identificación de lípidos oxidados sin correspondencia en las bases de datos disponibles actualmente y aplicarlo en diversos estudios de muestras de cáncer. Dada la creciente importancia de estos compuestos en relación con el diagnóstico y/o estratificación del cáncer, se realizará el estudio exhaustivo de las modificaciones lipídicas oxidadas por LC-MS para su identificación a través de coincidencia espectral, elucidación estructural y comparación con sus correspondientes formas nativas y, se optimizará un método de análisis de los productos de oxidación de los principales fosfolípidos. También, se creará una librería con sus tiempos de retención y fragmentos característicos procedentes del análisis MS/MS, aplicando esta metodología en el reprocesado de datos de los estudios de cáncer llevados a cabo. Finalmente se llevará a cabo el estudio de los productos de oxidación de las principales glicerofosfolípidos en datos procedentes del análisis de muestras de plasma y tejido de cáncer de pulmón.

4. Desarrollar un método de CE-MS para la determinación de compuestos aniónicos en muestras biológicas. El acoplamiento de la electroforesis capilar con el *electrospray* en condiciones que permitan medir compuestos aniónicos, con frecuencia de gran relevancia en el metabolismo del cáncer, ha probado tener diversas dificultades. Para resolverlas se estudiarán las condiciones de análisis que influyen tanto en la separación electroforética como en la espectrometría de masas. Además, se aplicará la metodología a un estudio de cáncer de pulmón con muestras de plasma. Este punto tiene el objetivo fundamental de establecer un método de análisis optimizado para compuestos aniónicos mediante CE-MS que puede ser aplicado a los futuros estudios metabolómicos multiplataforma que se realicen.



Introducción General: Metabolómica para la búsqueda de biomarcadores validados en estudios de cáncer

1. Introducción

La metabolómica nos ofrece un enfoque intensivo que permite la identificación de características metabólicas específicas de una enfermedad como biomarcadores potenciales, debido a que un cambio menor en los niveles de expresión de genes o proteínas genera un cambio significativo a nivel de metabolitos. La identificación y cuantificación de estos metabolitos son poderosas herramientas para el diagnóstico precoz de enfermedades, y estos metabolitos pueden actuar como indicadores de la respuesta a un tratamiento.

En esta introducción se describen los principales aspectos relacionados con la metabolómica para la búsqueda de biomarcadores enfocando los resultados en la enfermedad que engloba el objetivo de este trabajo de investigación, el cáncer. En ella se describen las estrategias metodológicas y hallazgos, repasando los principales artículos publicados en los últimos diez años, acotando los resultados a estudios que reporten biomarcadores validados para esta patología. Toda esta recopilación de información resultará de gran utilidad para los investigadores centrados en los cambios metabólicos originados por el cáncer, restringiendo a aquellos metabolitos que hayan resultado con mayor peso y hayan sido evaluados para ofrecer el mayor rendimiento, además de sensibilidad, especificidad y reproducibilidad.

Las principales metodologías utilizadas en metabolómica incluyen: i) análisis no dirigido: la medición de la más amplia gama de metabolitos extraídos de una muestra biológica sin una hipótesis *a priori* y ii) metabolómica dirigida, centrada en una ruta bioquímica o un conjunto específico de metabolitos que proporciona una mayor sensibilidad y selectividad, pero requiere información inicial para desarrollar y optimizar los métodos selectivos de análisis. Esta metodología es clave en la validación de biomarcadores para su traducción a la clínica (1). Ambos tipos de análisis se realizan siempre de modo diferencial, definiendo un grupo de prueba de la condición investigada y un grupo control con individuos que no presentan la condición investigada. La selección adecuada de las muestras a comparar es uno de los puntos críticos para el éxito de todo estudio.

En el área de investigación biomédica, la metabolómica no dirigida se centra en la búsqueda de biomarcadores para mejorar el diagnóstico de enfermedades. Sin embargo, los hallazgos pueden utilizarse para comprender sus mecanismos fisiopatológicos. Entre las plataformas analíticas empleadas en metabolómica destacan la espectroscopia de resonancia magnética nuclear (NMR) y la espectrometría de masas

(MS). La NMR es una técnica rápida, no invasiva y no destructiva para la identificación de metabolitos en una muestra, que requiere una preparación mínima de la muestra para obtener resultados altamente reproducibles. La NMR genera información sobre la identidad y cantidad de un alto número de metabolitos en un solo experimento, pero tiene la desventaja, inherente a su falta de sensibilidad, de perder información relacionada con los metabolitos presentes a bajas concentraciones, además de una importante superposición de señales en algunas regiones que impiden la identificación de ciertos metabolitos.

La espectrometría de masas, por su lado, es una técnica analítica con alta sensibilidad, pero con menor robustez. Para mejorar la información obtenida, se suele combinar con una técnica de separación como la cromatografía de líquidos (LC), la cromatografía de gases (GC) y/o la electroforesis capilar (CE). En general, el número y tipo de metabolitos que se pueden detectar difiere según la plataforma elegida (2).

En cualquier estudio metabólico no dirigido, la búsqueda de biomarcadores diagnóstico consta de una serie de etapas que deben diseñarse en función del objetivo del estudio, el cual debe estar claramente definido desde el inicio. Estas son: la selección de grupos adecuados para comparar; el tipo de muestra biológica y tratamiento adecuado; selección de la técnica y metodología analítica; adquisición de los datos; su tratamiento, incluyendo análisis estadísticos e identificación de metabolitos; e interpretación biológica de los resultados. Actualmente, uno de los cuellos de botella en este proceso es la identificación de metabolitos; por lo tanto, muchos artículos publican listas de metabolitos con solo una identificación tentativa, que muy a menudo no se puede confirmar. Este hecho, junto con el uso de pequeños grupos donde las correlaciones pueden aparecer por casualidad, debilita la credibilidad de las herramientas de metabolómica. Por lo tanto, la revisión bibliográfica se ha llevado a cabo solo con estudios que incluyan "biomarcadores validados". Además, en el capítulo 2 nos centraremos en el gran desafío que supone para el investigador la identificación correcta de las señales significativas, más concretamente de los lípidos oxidados, un grupo grande y heterogéneo de metabolitos modificados que ha despertado creciente interés en la última década por su relación con el estrés oxidativo, estrechamente ligado a patologías como el cáncer, y la limitada disponibilidad de información sobre ellos.

1.1. Biomarcadores y validación en un estudio no dirigido

La definición más genérica de biomarcador es "característica que se mide y evalúa objetivamente como un indicador de procesos biológicos normales, procesos patogénicos o respuestas farmacológicas a una intervención terapéutica" (3).

El gran crecimiento experimentado en el campo de la metabolómica se ha centrado en la mejora del diagnóstico y la prevención de enfermedades humanas con el descubrimiento de numerosos potenciales biomarcadores. Con la identificación de estos biomarcadores, y su posterior estudio y validación, se debería mejorar la clasificación de enfermedades con alta especificidad y sensibilidad. Por lo tanto, existe la preocupación de que, sin una atención cuidadosa al proceso completo de validación de un biomarcador, estos avances no puedan llegar a convertirse en pronósticos o tratamientos efectivos para los pacientes. Sin embargo, este escenario es el estado actual, con obstáculos que superar. Aunque el exitoso avance en el desarrollo de biomarcadores y criterios de valoración sustitutos, tiene el potencial de acercar la innovación científica a la aplicación clínica más deseada, todavía hay poco acuerdo sobre los pasos necesarios para validar estas nuevas entidades y conseguir la evaluación clínica adecuada (4).

La validación implica el proceso para evaluar un determinado biomarcador y establecer el margen de las condiciones bajo las cuales ofrece resultados con el mayor rendimiento, sensibilidad, especificidad y reproducibilidad (3). Este procedimiento conlleva mucho trabajo en el laboratorio y requiere un número elevado de participantes en el estudio, idealmente con diferente origen, aunque la investigación se debe comenzar con la necesidad de validar los marcadores en una cohorte externa o bajo otra plataforma analítica, distinta a la empleada para su descubrimiento, y no solo con la diferencia entre pacientes y controles en un grupo de muestra pequeño.

Una vez seleccionado un biomarcador potencial en un estudio de metabolómica, el proceso de validación incluye lo siguiente: i) validación analítica, ii) validación clínica y iii) validación en ensayos clínicos que proporcionan evidencia de que los métodos utilizados para un analito en particular respaldan el uso del marcador con poder estadístico suficiente para ser aprobado antes de ingresar al mercado. Dentro de cada etapa principal existen diferentes fases de evaluación (5, 6).

1.2. Validación analítica

Clásicamente, la validación analítica se define como la capacidad de la prueba para medir con precisión y fiabilidad los metabolitos seleccionados en muestras representativas de la población de interés.

Primero, se debe validar la metodología para el estudio de metabolómica. Para una visión completa y exhaustiva de la metodología a seguir para la validación analítica, revisión de D. Dudzik y C. Barbas-Bernardos (7). En segundo lugar, cuando el metabolito o conjunto de metabolitos se haya seleccionado en un experimento de metabolómica no dirigido, la validación analítica incluirá la confirmación de la identidad y la concentración del marcador potencial con una técnica diferente.

La metodología utilizada en la fase de descubrimiento (generalmente con un gradiente amplio a través LC-QTOF/MS) no es necesariamente la que se utilizará para confirmar la identidad, donde un método selectivo en LC-QQQ/MS es la opción más común y también, debe ser validado.

1.3. Validación clínica y validación de utilidad clínica

Después de la validación analítica, se debe evaluar el rendimiento del análisis tanto para la capacidad de predicción clínica como para la utilidad clínica basada en la mejora de los resultados del paciente.

Un primer paso fundamental debe ser la selección de posibles biomarcadores, que representa solo el punto inicial desde el descubrimiento en el laboratorio hasta el diagnóstico clínico. La mayoría de las publicaciones informan sobre el desarrollo de un ensayo basado en una lista corta, de no más de 10 candidatos, ya que es más probable que genere un modelo mucho más sólido que tratar de ajustar el modelo con cientos de características moleculares.

Esta selección se basa en la búsqueda de una combinación simple de metabolitos que ofrezcan un resultado predictivo efectivo, tratando de reducir el número de falsos positivos, evaluando la precisión de los modelos multivariantes y, optimizando individualmente cada candidato en dichos modelos estructurales validados (8).

La validación de un biomarcador está íntimamente ligada a su uso propuesto; por tanto, los mayores esfuerzos se centran en garantizar que dichos biomarcadores sean adecuados para el fin (*fit-for-purpose*).

Para llegar a la etapa final en el desarrollo de un biomarcador para su validación clínica, descubrimiento y prevalidación, los modelos se fijan con conjuntos de datos de entrenamiento y validación interna.

Como se mencionó anteriormente, un punto importante y limitante en la validación de un biomarcador corresponde a la temprana fase preanalítica en relación con la muestra. Otro factor limitante, el número de muestras en cada conjunto, suele ser muy pequeño y, por lo tanto, aumenta el efecto de las diferencias sutiles, como la edad, el sexo o la dieta. Otra práctica habitual que limita esta etapa es la de no trabajar con escenarios completamente independientes. Por lo tanto, para la validación completa de los biomarcadores debidamente identificados y prevalidados, se deben involucrar grandes cohortes de muestras de diferentes grupos de pacientes (9). La Figura 1 muestra un potencial diagrama de flujo para este proceso.

Finalmente, mientras que la validez clínica se enfoca en definir el desempeño del biomarcador con precisión para predecir el objetivo clínico, la utilidad clínica está relacionada con que el ensayo tenga como objetivo mejorar los resultados de los pacientes. Por lo que, respaldar la toma de decisiones y mejorar los resultados de los pacientes son los principales propósitos de la utilidad clínica.

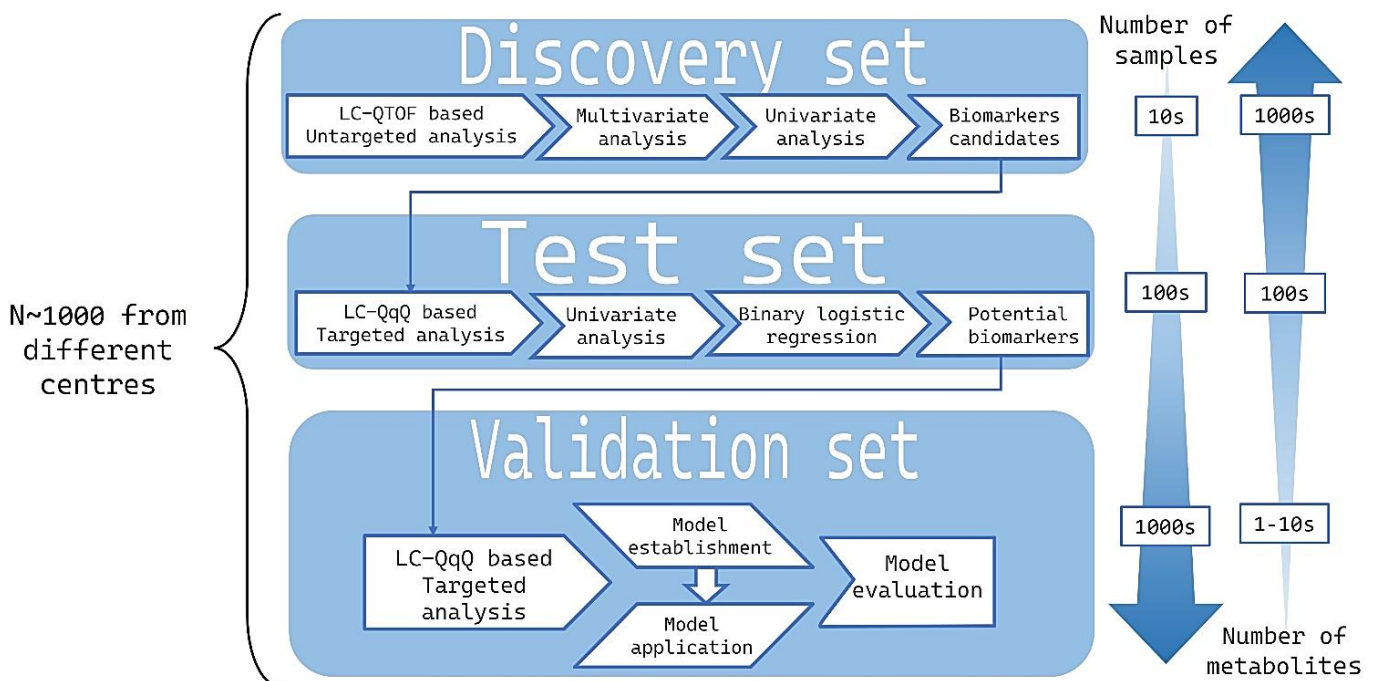


Figura 1. Diseño de las etapas de validación de un biomarcador con aplicación diagnóstica en un estudio metabolómico. Se indica el tamaño aproximado de muestras y metabolitos.

1.4. Cáncer

El cáncer es una enfermedad metabólica, y un diagnóstico temprano es crucial para la supervivencia. Las tecnologías de metabolómica han permitido una visión más profunda del metabolismo del cáncer. Actualmente, la investigación en metabolómica se está utilizando para descubrir biomarcadores de cáncer de diagnóstico en la clínica para comprender mejor su compleja naturaleza heterogénea y estadificación, descubrir vías involucradas en el cáncer que podrían usarse para nuevos objetivos y monitorear biomarcadores metabólicos durante la intervención terapéutica (10).

El desafío que engloba este trabajo de investigación es la aplicación de la metabolómica para la búsqueda de biomarcadores en cáncer, una de las principales causas de morbi-mortalidad del mundo. Por ello, esta introducción recoge una revisión bibliográfica que se centrará en resumir los estudios de metabolómica en cáncer para la búsqueda de biomarcadores en humanos que han aplicado un cierto grado de validación más allá de la validación matemática del modelo multivariante, publicados entre 2013 y enero del 2022.

2. Materiales y métodos

En la literatura se han publicado varios artículos para identificar biomarcadores en cáncer. Sin embargo, al considerar estudios con validación externa, este número disminuye considerablemente. El criterio de búsqueda utilizado consistió en utilizar las palabras clave 'metabolomic*' OR 'metabonomic*' AND 'biomarker' AND 'validat*' AND 'cancer' en las bases de datos Web of Science (<https://www.webofscience.com/>) y Pubmed (<https://pubmed.ncbi.nlm.nih.gov/>) donde se encontraron 574 artículos.

A continuación, se realizó un filtrado manual, artículo por artículo, para confirmar el resultado con el objetivo propuesto: estudios validados para el descubrimiento de biomarcadores en cáncer aplicados a muestras humanas, aunque la fase de descubrimiento y establecimiento del modelo podría incluir muestras de otras fuentes (11).

Se excluyeron las publicaciones donde el término validación se refería únicamente a los modelos matemáticos del análisis multivariante. Una parte importante del cribado se centró en encontrar la cohorte de individuos utilizada para la validación externa y, posteriormente, los artículos fueron clasificados, según la aplicación de los biomarcadores o tipo de cáncer u órgano afectado. La Figura 2 resume el número de artículos totales encontrados y aquellos que se incluyeron en la revisión.

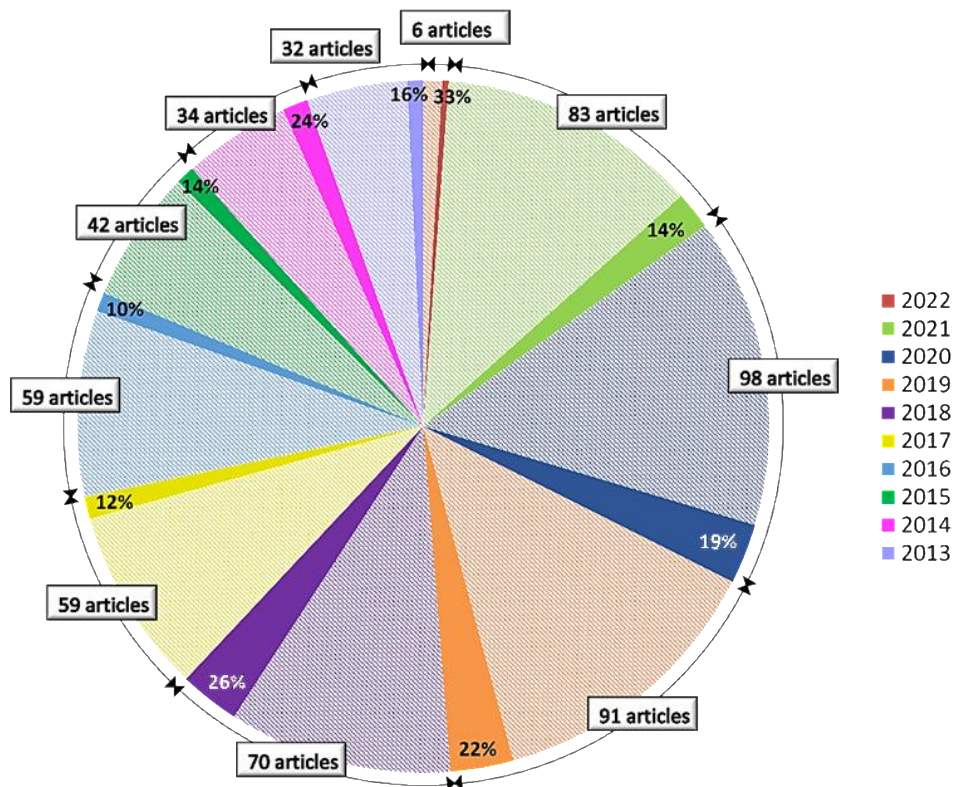


Figura 2. Representación gráfica del número de artículos publicados por año que cumplían nuestros requisitos, y de ellos, en color sólido, el porcentaje de artículos que, tras el cribado manual, se han utilizado para esta revisión. La leyenda indica el año asignado a cada color.

3. Resultados

De los 574 artículos que cumplían los criterios de inclusión establecidos para la búsqueda de publicaciones en las bases de datos, la revisión manual de todos ellos, vió reducido el número de publicaciones validas a 108. De esta revisión manual fueron excluidas todas aquellas publicaciones que consistían en: la revisión bibliográfica sobre el tema; aquellos que no incluían muestras humanas, al menos para validar los resultados; todos aquellos que su objetivo no fuese la búsqueda y posterior validación externa de biomarcadores de cáncer y, por último, se descartaron aquellos que únicamente se centaban en la validación del modelo multivariante. Tras esto, los artículos fueron agrupados por tipo de cáncer u órgano afectado por el tumor, y a continuación, ordenados por tipo de muestra empleada para el análisis.

La Tabla 1 reúne la información recogida de la revisión llevada a cabo desde 2013 a enero del 2022. Las publicaciones, como ya se ha comentado, están agrupadas

por tipo de cáncer y ordenadas por tipo de muestra. La información recogida en la tabla resume los puntos fundamentales del estudio realizado. Entre ellos se encuentra el tamaño de la cohorte de estudio o descubrimiento y el de la cohorte de validación, detallando el número de muestras por grupo siempre que fuese reportado por los investigadores. Además, se incluye el resultado de la validación, resumiendo el panel de biomarcadores empleados y el tipo de instrumentación utilizada, especificando la técnica de separación, si la hubiese, y la instrumentación aplicada para la detección (10). También se recoge si se ha empleado más de una plataforma.

Durante los últimos casi 10 años (2013 a enero 2022) se han identificado 108 artículos que han descrito estudios de búsqueda de biomarcadores de diferentes tipos de cáncer con cierto nivel de validación. Este proceso de validación utilizó principalmente dos conjuntos diferentes de muestras, uno para la detección y otro para la confirmación, o diferentes técnicas analíticas, o ambas estrategias.

Cáncer de vejiga

BC del inglés bladder cancer.

Este tipo de cáncer, el noveno tipo de cáncer más común en todo el mundo se forma en las células de la vejiga, y el más frecuente es el carcinoma de células de transición o de células transaccionales, con un 90% de los casos según la SEOM (<https://seom.org/>) (Sociedad Española de Oncología Médica).

Actualmente, la cistoscopia y la citología son los procedimientos estándar para la detección de BC. La cistoscopia es un procedimiento invasivo y costoso, mientras que la citología es un método no invasivo, que a menudo sirve como complemento de un estudio de cistoscopia negativo.

Aunque la citología tiene una alta especificidad, su sensibilidad no es satisfactoria, particularmente para un tumor de grado bajo (12). De los once estudios de BC seleccionados, ocho utilizan orina (13-20) que parece ser el tipo muestra ideal, por el contacto directo con el tumor y, por ser mínimamente invasiva, aunque sujeta a una alta variabilidad.

Además, tres estudios utilizaron muestras de plasma o suero (21-23). Bansal y colaboradores (21) y Tan y col. (23) se centraron en el estudio diferencial de BC de bajo grado (LG) y alto grado (HG), aunque encontraron que podían también diferenciar los controles sanos. Los marcadores potenciales encontrados en ambos estudios, resumidos en la Tabla 1, fueron diferentes, probablemente debido al uso de técnicas analíticas que detectan diferentes compuestos (^1H NMR y LC-QTOF/MS). El otro de los

estudios llevado a cabo con muestras de plasma (22), se centra en el descubrimiento de un panel de biomarcadores que permita la diferenciación entre BC y el carcinoma de células renales (RCC *del inglés Renal Cell Carcinoma*). A la vista de los resultados, y teniendo en cuenta la plataforma empleada, LC-MS, no es de extrañar que el panel obtenido se base fundamentalmente en lípidos, más concretamente en ácidos grasos (FA) y fosfolípidos (PL). Liu y *col.*(22) reportaron un panel para la detección del cáncer, al establecer diferencias con los controles sanos (HC) en FA y sus derivados oxidados (oxFA), y otro que diferencia BC de RCC a través de fosfolípidos como PS P-38:0 y esfingolípidos como SPB 18:2;O2, además de FA y oxFA.

En la diferenciación entre BC y RCC, Wang y *col.* (20) utilizaron una plataforma similar, pero en este caso con muestras de orina. Los resultados procedentes del estudio de la enfermedad remarcan biomarcadores implicados en el metabolismo de lípidos y purinas, sin embargo, para la discriminación entre BC y RCC, destacan desórdenes en metabolismo de purinas y aminoácidos fundamentalmente, con muestras que presentaban hematuria.

Metástasis ósea y caquexia

La metástasis ósea se produce cuando las células tumorales se propagan desde su lugar de origen al hueso. La mayoría de los tipos de cáncer pueden propagarse de esta forma, pero, es cierto que, existen algunos tipos de cáncer que son más propensos a comportarse así, como el cáncer de mama o el de próstata. Apenas carente de excepciones, el cáncer que se propaga a los huesos no tiene cura y los tratamientos se centran en ayudar a disminuir el dolor y otros síntomas de la metástasis ósea.

La osteonecrosis de la mandíbula (*del inglés ONJ*) es una interrupción del suministro vascular o necrosis avascular que conlleva la exposición de los huesos de la mandíbula. Esto puede ser debido al tratamiento administrado a un paciente con metástasis ósea a través de la radiación, la terapia en dosis elevadas con esteroides y los medicamentos que interrumpen el suministro vascular o el recambio óseo en la mandíbula. Yatsuoka y *col.* (24), a través de muestras de saliva, validaron niveles elevados de hipotaurina como biomarcador clave para la detección precoz de la osteonecrosis mandibular debida a la medicación suministrada. Hasta la fecha, existen escasos estudios que hayan investigado marcadores precoces para la detección de esta patología y, aun menos que utilicen saliva para llevarlo a cabo, debido a que la recolección de este tipo de muestra en pacientes con cáncer ha sido reportada como

una ardua tarea por el hecho de que el dolor causado por la patología probablemente disminuye la producción de saliva (25).

La caquexia es una complicación tardía generada en cáncer avanzado, entre otros. El diagnóstico de la caquexia debida a cáncer depende de medidas objetivas de síntomas clínicos y antecedentes de pérdida de peso, que no siguen la progresión de la enfermedad y tienen una utilidad limitada para su diagnóstico precoz y el tratamiento correspondiente. Quan-Jun Yang y *col.* (26) obtuvieron una ecuación diagnóstica que incluye carnosina, leucina y acetato de fenilo, y que podría ser utilizada como una herramienta diagnóstica factible para identificar poblaciones en riesgo a través de la detección de metabolitos séricos.

Cáncer de mama

BRC del inglés breast cancer.

El cáncer de mama es el tipo de cáncer más comúnmente diagnosticado en mujeres (más del 99% según la SEOM) y consiste en una acelerada e incontrolada proliferación de las células del epitelio glandular. Habitualmente, se usan la mamografía y la ecografía para la detección de BRC; sin embargo, están asociados con inconvenientes tales como la exposición a la radiación y la dependencia del nivel de habilidad de los operadores. Es por ello por lo que, la existencia de una prueba mínimamente invasiva, a través de plasma o suero, que pueda identificar con precisión a los pacientes con BRC, aumentaría la tasa de detección y provocaría como resultado la detección en una etapa más temprana.

En la tabla se recogen cinco estudios de BRC, de los cuales uno utiliza muestras de suero (27) y el resto, plasma (28-31). Cabe destacar que todos ellos son posteriores a 2018, lo que avanza una preocupación creciente en la búsqueda de biomarcadores para BRC en los últimos años. Los resultados obtenidos de los distintos estudios son heterogéneos, pero más focalizados en aminoácidos y sus derivados, y en lípidos como las carnitinas (CAR), FA y PL.

Con una gran cohorte de muestras, pero desproporcionada en cuanto al número reclutado, Zhao y *col.* (29) comparan cohortes de casos y controles muy alejadas de una posible semejanza a la población, debido a que el número de casos es muy elevado en comparación con el de controles, y los resultados obtenidos podrían verse sesgados por esta diferencia.

Por su parte, el estudio llevado a cabo por Satomi Tomida y *col.* (31) utiliza una cohorte de validación que incluye muestras recogidas tras la extirpación quirúrgica del

tumor. Esto provoca un efecto selectivo de restauración de las especies con cadenas 18:2 tanto en glicerofosfoetanolaminas (PE) como en plasmalógenos de etanolamina. Sin embargo, los niveles de especies poliinsaturadas FA 20:4 y FA 22:6 no se vieron afectadas o se redujeron tras la cirugía.

Cáncer de cuello uterino (CC)

CC del inglés cervical cancer.

El cáncer de cuello uterino es una enfermedad maligna común en las mujeres y su principal causa es la infección persistente por el virus del papiloma humano (HPV) oncogénico. A menudo, en las primeras etapas, los pacientes diagnosticados con cáncer de cuello uterino no muestran signos o síntomas evidentes, que puede llevar a un diagnóstico erróneo y, por lo tanto, una diagnosis en etapa avanzada. Hasta la fecha, se utiliza el biomarcador antígeno del carcinoma de células escamosas (SCC-Ag) que aparece elevado en los distintos estadios de la enfermedad y, aunque también son conocidos otros anticuerpos circulantes, bien es cierto que su precisión diagnóstica y rendimiento predictivo siguen siendo inciertos.

Los niveles de los metabolitos de un panel seleccionado, tras un primer estudio de huella dactilar, en una cohorte de pacientes externa, fueron determinados mediante UHPLC-QqQ/MS (32). Destacan una elevación significativa de AMP en el grupo de cáncer de cuello uterino en comparación con la neoplasia intraepitelial de grado 1 y los controles. El resto del panel lo completan metabolitos de pequeño tamaño como ácidos carboxílicos, aminoácidos y sus derivados, como son el ácido aspártico y glutámico, hipoxantina, lactato, prolina y piroglutamato. Una rápida proliferación celular y un metabolismo aumentado de las células tumorales requiere de una mayor cantidad de aminoácidos y moléculas de pequeño tamaño para la síntesis de proteínas y ácidos nucleicos, lo que concuerda con estos resultados.

Colangiocarcinoma (CCA)

Nos encontramos ante un cáncer del conducto biliar que es relativamente común en los países del Este, pudiendo ser detectado por 21-desoxicortisol y bilirrubina con alta precisión (33).

Macias y *col.* (34) establecieron un panel serológico de metabolitos para el diagnóstico diferencial de adenocarcinomas pancreáticos, colangiocarcinoma distal y adenocarcinoma ductal pancreático (PDAC), compuesto de varios lípidos, en concreto, CAR 16:0, Cer 42:1;O2, LPC 20:0, PC O-16:0/20:3, LPE P-18:2, SM 40:2;O2 y SM

41:2;O2. De esta forma se sobrepone la falta de biomarcadores no invasivos para el diagnóstico precoz de PDAC y dCCA, contribuyendo a mejorar el pésimo pronóstico de este tipo de tumores.

Leucemia

La leucemia mieloide crónica (CML) es un tipo de cáncer que se origina en determinadas células productoras de sangre de la médula ósea. En la tabla se recogen dos estudios sobre la leucemia, uno de CML (35) y otro de leucemia linfoblástica aguda (LLA) (36). En el primero de los estudios (35), se consideraron marcadores no solo para el diagnóstico sino también para evaluar la respuesta al tratamiento con inhibidores de la tirosina quinasa, y se encontró que el glicerol y el ácido mirístico son los más relevantes para monitorizar la eficacia terapéutica. Por su parte, a través de líquido cefalorraquídeo, se estableció el perfil metabolómico global para proporcionar información sobre los mecanismos alterados en la fatiga en pacientes pediátricos con LLA (36). Los biomarcadores validados asparagina y γ -glutamylglutamina proporcionan información sobre futuras terapias dirigidas a reducir la fatiga relacionada con el cáncer asociado a LLA.

Cáncer colorrectal (CRC)

Se produce cuando los tumores se forman en el revestimiento del intestino grueso y se han recogido nueve estudios para la búsqueda de biomarcadores de cáncer colorrectal a través de diferentes técnicas y distintos tipos de muestra. Orina a través de CE-MS se utilizó en dos estudios complementarios, puesto que Sakurai y *colaboradores* (37) utilizaron los resultados obtenidos de las primeras investigaciones del grupo, para llevar a cabo la validación de los biomarcadores (38). En él, se establecen varios paneles discriminativos según CRC, adenoma o CRC + adenoma, todos ellos con N₈-Acetilespermidina y N₁,N₈-Diacetilespermidina en común.

Los niveles plasmáticos de metabolitos para el diagnóstico en estadio temprano (39) y el análisis de la respuesta a un posible tratamiento fútil (40) fueron llevados a cabo estableciendo paneles de 16 y 21 biomarcadores, respectivamente. En el trabajo de Rattner y *col.* (40), se ofrece un novedoso enfoque con la obtención de biomarcadores para la evaluación temprana de la eficacia de un tratamiento. Continuando con plasma, Crotti y *col.* (41) utilizando GC-MS, encontraron que los ácidos grasos de cadena media entre 6 y 12 átomos de carbono (C6-C12) eran la clase

de lípidos que mostraba los cambios más marcados en el desarrollo de CRC, aunque los pacientes no fueron estratificados.

Haihong Zha y *col.* (42) utilizaron una técnica diferente, SWATH MRM-QqQ/MS, que aumenta la sensibilidad al medir entre 1000 y 2000 metabolitos. Los resultados procedentes del estudio a nivel tisular fueron validados en el plasma de pacientes con CRC para evaluar el pronóstico tras la resección quirúrgica del tumor. Igualmente, en un estudio con muestras de tejido y heces, se validó acetato fecal como biomarcador diagnóstico de CRC a través de ^1H NMR (43).

Por último, cabe destacar las investigaciones realizadas en el suero de pacientes con CRC (44, 45). Por su parte, Zha y *col.* (45) se basan en la derivatización para obtener el perfil global del metaboloma de ácidos carboxílicos a través del análisis de sus fragmentos característicos.

Cáncer de endometrio (EC) y epitelial de ovario (EOC)

El cáncer de endometrio (EC) es el cáncer más común del aparato reproductor femenino en los países desarrollados y, actualmente, no se dispone de ningún sistema de detección eficaz. Dos estudios (46, 47) concluyeron que la firma metabolómica sérica del cáncer de endometrio difiere de la de los controles sanos y de la de la enfermedad endometrial benigna y otros cánceres ginecológicos (como el cáncer de ovario). Sin embargo, el 3-hidroxibutirato fue el único marcador común en ambos estudios.

Por otra parte, los metabolitos identificados y validados, fosfocolina, asparagina y ácido málico del líquido cervicovaginal se consideran prometedores biomarcadores metabolómicos para la detección de EC mediante espectroscopia ^1H NMR (48).

Con respecto al cáncer de ovario epitelial (EOC) (49, 50), la metabolómica del suero basada en GC-MS permitió la construcción de dos modelos para la predicción de EOC en etapa temprana. Uno de ellos proporciona información útil sobre las alteraciones metabólicas de los ácidos grasos, demostrando que biomarcadores de ácidos grasos saturados son útiles para la detección y caracterización de pacientes con EOC (50). Por su parte, Wang y *col.* (49) establecieron un modelo predictivo de EOC a través del metabolismo de aminoácidos fundamentales, incluyendo metionina, glutamina, asparagina, ácido glutámico y ácido glicólico.

Asimismo, se estableció un modelo de 12 biomarcadores asociados a la β -oxidación anormal de FA y el metabolismo de PL y ácidos biliares en cáncer de ovario (OC) analizando suero con UHPLC-QTOF/MS; de los cuales, 2-piperidinona y LPE 17:0

fueron validados con una cohorte externa, proporcionando una herramienta aplicable al diagnóstico clínico y tratamiento de OC (51).

Cáncer de esófago

Siete artículos describieron marcadores de cáncer de esófago, el octavo cáncer más común en todo el mundo y uno de los más letales. Hay dos tipos histológicos principales de este tipo de cáncer, el adenocarcinoma esofágico (EAC) (52, 53) y el carcinoma de células escamosas esofágicas (ESCC) (54-58), que sigue siendo el tipo más prevalente de cáncer esofágico en el mundo. Debido a que el pronóstico y la supervivencia de los pacientes con ambos tipos de cáncer siguen siendo deficientes, se necesita más investigación para lograr cualquier herramienta de detección temprana útil a partir de las muestras no invasivas analizadas por metabolómica.

La displasia de células escamosas (ESD) es una lesión histológica confinada al epitelio y que es considerada una lesión premaligna del ESCC. A través del suero de pacientes con ESCC y ESD se establecieron diferentes paneles que, no solo permiten la discriminación de ESCC, sino que facilitan el reconocimiento de lesiones premalignas para un diagnóstico más temprano, lo que mejora la supervivencia de los pacientes (55).

Xu y *colaboradores* (57) fueron un paso más allá, utilizando HPLC-QTOF/MS para validar un panel de biomarcadores cuyos niveles están estrechamente relacionados con el efecto del tratamiento, empleando el plasma de pacientes antes, durante y después de la quimiorradioterapia (QRT), que es útil para el diagnóstico así como para el control de las respuestas terapéuticas y predicción de resultados de ESCC.

Cáncer gástrico (GCa)

Yinan Chen y *col.* (59) estudiaron el diagnóstico temprano de GCa en orina mediante GC-MS y encontraron, además de un panel de marcadores, un conjunto de ellos (prolina, p-cresol y ácido 4-hidroxibenzoico) para el valor de predicción de resultados mediante el análisis de supervivencia. En un estudio más amplio, Jeeyoun Jung y *col.* (60) adquirieron perfiles metabólicos urinarios y tisulares utilizando RMN y espectroscopia de giro de ángulo mágico de alta resolución (HR-MAS) respectivamente. Esta es una herramienta que combina las ventajas de las técnicas de RMN para muestras sólidas y líquidas, y ofrece la oportunidad de medir tejidos intactos. Además, se realizó el perfil metabólico de la orina de pacientes con GCa después de la cirugía curativa. Los niveles de 4-hidroxifenilacetato, alanina, fenilacetilglicina, manitol, glicolato y arginina se correlacionaron significativamente con el estadio del cáncer y, junto con el

nivel de hipoxantina, mostraron una tendencia de recuperación hacia los niveles de los controles sanos en las muestras postoperatorias en comparación con las muestras preoperatorias.

Recientemente, el estudio de plasma a través de UHPLC-QTOF/MS (61) encontró diferencias significativas que fueron posteriormente validadas en el perfil lipídico de pacientes con lesión gástrica precancerígenas frente a GCa. El panel compuesto por dos lisofosfatidilcolinas, LPC 18:3 y LPC 20:3, y cuatro ácidos grasos, FA 20:4, FA 18:3, FA 18:2, y FA 16:0, permite la discriminación de una posible lesión premaligna. Con el mismo objetivo, Corona y *col.* (62) utilizaron los niveles séricos de SM 22:1;OH y CAR 16:0 para diferenciar familiares de primer grado con mayor riesgo de desarrollar GCa.

Glioma

Según Hua Zhao y *col.* (63) los biomarcadores moleculares basados en tumores han redefinido la clasificación de los gliomas. Sin embargo, aún no se ha explorado la asociación de la metabolómica sistémica con el fenotipo de glioma. Sus hallazgos identificaron metabolitos y vías metabólicas que diferenciaron los fenotipos tumorales y que pueden ser útiles como candidatos a biomarcadores del huésped para ayudar aún más a la clasificación molecular del glioma. Con similar objetivo se ha utilizado el análisis de subconjuntos de puntos de tiempo repetidos coincidentes (SMART), en un estudio de metabolómica de muestras de sangre repetidas de futuros pacientes con glioma y sus controles emparejados (64). Este estudio destaca los beneficios del análisis de patrones de progresión y proporciona una herramienta para el descubrimiento de marcadores tempranos de enfermedad.

Carcinoma hepatocelular (HCC)

Once artículos describen la búsqueda de marcadores de hepatocarcinoma (HCC), nueve de ellos en suero (65-70), tres combinan los resultados con tejido (9, 71, 72), uno en orina (73) y otro con muestras de plasma (74). Entre ellos, Ping Lou y *col.* (9) realizaron un experimento modelo con un estudio multicentro, con 1448 sujetos, de los cuales la mayoría eran controles normales y pacientes con infección crónica por el virus de la hepatitis B, cirrosis hepática y HCC. Ciento ocho se utilizaron como prueba de descubrimiento, 648 como conjunto de prueba y 572 para evaluar el rendimiento del panel. Finalmente, se encontró que fenilalanil-triptófano y, a partir de otros estudios,

diferentes acilcarnitinas estaban potencialmente alterados, lo que concuerda con la ya conocida relación de la β -oxidación de ácidos grasos en cáncer.

Existen otros estudios que reportan niveles alterados de acilcarnitinas en HCC (66, 68, 71, 73), como el llevado a cabo en muestras de orina, que combina CAR 4:0 y ácido hidantoína-5-propiónico para distinguir HCC de cirrosis (73).

Cáncer de pulmón

LCa del inglés lung cancer

En los numerosos estudios de cáncer de pulmón encontrados se puede justificar la complejidad que este presenta ya que comprende no solo el cáncer de pulmón de células pequeñas (SCLC) y no pequeñas (NSCLC), sino también el adenocarcinoma de pulmón y el carcinoma de pulmón de células escamosas, dos subtipos principales de cáncer de pulmón de células no pequeñas.

SCLC y NSCLC presentan un crecimiento diferente y, por lo tanto, su tratamiento se lleva a cabo de distinta manera. El diagnóstico diferencial de SCLC frente a NSCLC es, por tanto, esencial para garantizar la selección del tratamiento adecuado. El micromuestreo de sangre, con su posterior secado y extracción, (DBS) ha permitido establecer dos modelos discriminantes multianálisis, especificados en género, para distinguir SCLC de NSCLC y controles (75).

Feng Peng y *col.* (76) centraron su investigación en los marcadores de respuesta basados en el tratamiento con platino. Al combinar el análisis estadístico multivariante y el análisis de vías alteradas con el análisis de correlación, 8 biomarcadores potenciales se asociaron significativamente con la respuesta a la quimioterapia con platino. De manera similar, Yanhua Tian (77) y Xiaoqun Nie (78) encontraron siete y dos metabolitos séricos, respectivamente, asociados con la respuesta al tratamiento oncológico.

El origen de los biomarcadores validados para este tipo de cáncer se ha centrado mayoritariamente en el uso de muestras mínimamente invasivas (suero, plasma, orina y DBS), que facilita su diagnóstico y seguimiento, evitando las posibles complicaciones y molestias que supone para el paciente una biopsia tisular.

Carcinoma oral de células escamosas (OSCC)

Un pequeño estudio abordó el conflictivo tema del uso de la quimioterapia personalizada de inducción en OSCC (79). Los autores encontraron ácido láctico, ácido glutámico y ácido aspártico como predictores serológicos de la respuesta.

Una alternativa novedosa con potencial para la detección precisa de metástasis ganglionar en pacientes con OSCC es el estudio del perfil metabólico de muestras de saliva (80). Además, establece un modelo conformado por cuatro aminoácidos, (glicina, prolina, citrulina y ornitina) asociados a OSCC en estadio temprano.

Cáncer de páncreas (PaC) y Adenocarcinoma ductal pancreático (PDAC)

Trece artículos describen marcadores de cáncer de páncreas, una enfermedad agresiva con altas tasas de mortalidad, pero de la que no se dispone de un análisis de sangre para su detección y diagnóstico tempranos, por lo que es imposible detectarlo en etapas iniciales, que es cuando es más curable.

Existen varios tipos de cáncer de páncreas, pero el más común comienza en las células que recubren los conductos que transportan las enzimas digestivas fuera del páncreas (adenocarcinoma ductal pancreático, PDAC). Sahni y *col.* (81) y, posteriormente, Hipperson (82) utilizaron el análisis de orina a través de ^1H NMR para validar trigonelina, ácido hipúrico y mioinositol. Por su parte, empleando UHPLC-QTOF/MS, se establecieron diferencias discriminativas según el estadio oncológico de PDAC en muestras de plasma (83) y, evidencias sustanciales a nivel metabólico tras la resección quirúrgica en suero para este tipo de cáncer pancreático (84).

Lindahl (85) y Mayerle (86) tuvieron como objetivo diferenciar el adenocarcinoma ductal pancreático y la pancreatitis crónica utilizando ambas muestras de plasma y suero, pero empleando técnicas de análisis diferentes, de ahí que no encontremos resultados compartidos.

La histidina se validó en cuatro de los estudios (86-89) mientras que xilitol y 1,5-anhidro-D-glucitol en dos de ellos (87, 89).

Cáncer de tiroides (TC)

El cáncer de la glándula tiroidea ocurre en las células de dicha glándula. Existen varios tipos de este, algunos de crecimiento lento y otros muy agresivos, pero en casi todos los casos es posible la curación del paciente con tratamiento. En un estudio sobre TC (90), se propuso el descubrimiento de nuevos biomarcadores séricos no invasivos para diferenciar los nódulos tiroideos benignos de los malignos para el diagnóstico del cáncer de tiroides. Así, los autores encontraron un biomarcador en suero, la fosfatidilcolina PC (34:1), con excelente valor diagnóstico para discriminar el cáncer de tiroides maligno (MTC) del tumor benigno (BTT).

Un estudio multicentro por HPLC-QTOF/MS llevado a cabo en muestras de suero, plasma y tejido, validó un panel de biomarcadores tisular para la discriminación conjunta de carcinoma papilar de tiroides (PTC) y nódulo tiroideo benigno (BTN) (91).

Cáncer de próstata (PCa)

Las técnicas de detección actuales para la detección de PCa se basan en la medición de los niveles séricos de antígeno prostático específico (PSA) y el examen rectal. Sin embargo, cada vez hay más pruebas en contra del valor de la detección basada en PSA por lo que son necesarios biomarcadores no invasivos para el diagnóstico (92). En los últimos 10 años se han encontrado siete estudios de validación de biomarcadores de cáncer de próstata en muestras no invasivas. De ellos, dos investigaciones emplearon suero como muestra de estudio (93, 94), otros dos plasma (95, 96) y tres orina (97-99).

A través de LC-MS se identificó citrato de 2-isopropilo, citidina y D-asparagina en suero. Estos biomarcadores candidatos se validaron en una cohorte de pacientes independientes de gran tamaño (n = 889). Potencialmente, el presente estudio proporciona la herramienta de diagnóstico para PCa en su etapa inicial con ese panel de biomarcadores (96).

En los estudios centrados en la búsqueda de biomarcadores úricos, aunque todos ellos emplearon GC-MS como técnica de análisis, dos de ellos (97, 99) se centraron en los compuestos orgánicos y carbonílicos volátiles (VOCs y VCCs). Por su parte, en el estudio llevado a cabo por Lima y *col.* (98) se validó un panel compuesto de 21 metabolitos utilizando el conjunto de resultados obtenidos del análisis de orina de pacientes con PCa a través de GC-MS y ¹H NMR. De los 21 metabolitos, trece fueron reportados por primera vez como asociados a la enfermedad de PCa, los ocho restantes corroboraron lo reportado en estudios anteriores.

Carcinoma de células renales (RCC)

Conocido también como cáncer de riñón o adenocarcinoma renal, se trata del tipo más común de cáncer en adultos y se presenta con mayor frecuencia en hombres de 60 a 70 años según la SEOM.

Como ya se comentó anteriormente en el apartado de cáncer de vejiga, Wang y *col.* (20) establecieron a través de los niveles en orina de 4-etoximetilfenol, prostaglandina F_{2B}, tromboxano B₃, CAR 4:0;O, 3-hidroxifloretina y N'-formilquinurenina, el panel discriminante entre BC vs RCC.

Además, en esta revisión se han incluido dos estudios con el objetivo común de la búsqueda de biomarcadores de RCC. Maria Elena Knott y *col.* (100) construyeron un modelo celular *in vitro* que permitió la obtención de metabolitos discriminantes, y que posteriormente, fue validado en suero de pacientes con estadio avanzado de RCC. En sus resultados remarcan, por primera vez, la asociación entre N-lactoil-leucina y RCC. Por otra parte, el análisis de orina estableció dos líneas diferenciadoras de RCC frente a tumor benigno y RCC frente a control (101).

Estudios multicáncer

Así pues, se han incluido también dos estudios llevados a cabo con el plasma de pacientes que presentaban distintos tipos de cáncer. El primero de ellos, identifica en una cohorte inicial, y posteriormente valida en una cohorte externa, un ácido dicarboxílico de cadena larga que se encuentra regulado a la baja en CRC, PC y cáncer de riñón (KC). En BC y cáncer de glioblastoma multiforme (GNM) no se presentó alteración en los niveles de este ácido dicarboxílico (102).

Según ha sido reportado en los últimos años, los cambios dinámicos en el perfil amínico están estrechamente relacionados con el inicio y la progresión del cáncer. El submetaboloma de aminas incluye aminoácidos y aminas biogénicas, y el empleo de una técnica más sensible para su detección, ayudaría en el descubrimiento de posibles biomarcadores, herramienta significativa para la detección del cáncer. Zhang y *col.* (103) llevaron a cabo un estudio en plasma de pacientes con LCa, BRC, CRC, GCa para discriminar de los controles y distinguir los tipos de cáncer entre sí. Por último, se validó un panel amínico, incluyendo sarcosina, ácido glutámico, citrulina, triptófano, glutamina e histidina, en pacientes con LCa anterior y posteriormente al tratamiento. Estos resultados indicaban el desarrollo de un método adecuado para el estudio de biomarcadores en cáncer que podría ser aplicado al diagnóstico precoz.

Tabla 1. Principales características de los artículos revisados de biomarcadores validados en cáncer.

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(21)	Bladder cancer (BC).	Low-grade BC (LGBC; n= 36), high-grade BC (HGBC; n= 31) and healthy controls (HC; n= 32).	Internal (n= 99) External validation: (n= 106).	6 biomarkers (dimethylamine, malonate, lactate, glutamine, histidine, and valine) of BC cases when compared to HC. 3 biomarkers (DMA, glutamine, and malonate) of cases of LG were able to differentiate from HG.	Plasma	¹ H NMR
(23)	Bladder cancer (BC).	LGBC (n = 34) and HGBC (n = 34) BC and HC (n = 30).	LGBC (n = 26) and HGBC (n = 26) and HC (n = 22).	3 serum metabolites (inosine, acetyl-N-formyl-5-methoxykynuren-amine, and LPS O-18:0 for HGBC when compared to LGBC and for LGBC from HC.	Serum	UHPLC-QTOF/MS
(13)	Bladder cancer (BC).	BC (n=38) and non-BC (n=61).	Model validation: BC (n = 7) and non-BC (n = 10) External validation: 32 BC (n = 32) and non-BC (n = 31).	External validation shows that levels of tryptophan were lower in BC compared to non-BC controls and levels of kynurenic acid were higher in BC compared to controls.	Urine	GCxGC-TOF/MS UHPLC-QqQ/MS (MRM)
(14)	Bladder cancer (BC).	BC (n = 92) and HC (n = 81).	BC (n = 46) and HC (n = 40).	Distinction between the BC and control with 12 differential metabolites including carnitine and acylcarnitines, the last steps of glycolysis, or the TCA cycle.	Urine	HPLC-QTOF/MS
(15)	Bladder cancer (BC).	Two centers: BC (n = 66) and non-BC (n = 266).	BC (n = 29) and non-BC (n = 79).	Palmitoyl sphingomyelin, lactate, adenosine, and succinate provide the strongest predictive power for differentiating cancer from non-cancer urines.	Urine	Training: UHPLC-LIT/MS and GC-MS Validation: UHPLC-Orbitrap/MS
(16)	Bladder cancer (BC).	BC (n = 50) and HC (n = 35).	BC (n = 59) and HC (n = 37).	A four-biomarker panel (5-hydroxyvaleric acid, cholesterol, 3-phosphoglyceric acid, and glycolic acid) that covered extensive metabolic characteristics was defined as a combinatorial biomarker for the discrimination of BCs vs HCs.	Urine	GC-MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(17)	Bladder cancer (BC).	LGBC (Low-grade BC; n =18), high-grade BC (HGBC; n= 15) and healthy controls (HC; n= 44).	BC (n = 20), controls (benign, n = 6) and healthy controls (HC; n= 20).	BC panel: trans-2-dodecenoylcarnitine, serinyl-valine, feruloyl-2-hydroxyputrescine, and CAR 9:0;O. HGBC vs LGBC panel: indolylacryloylglycine, N ² -galacturonyl-L-lysine, and aspartyl-glutamate.	Urine	UHPLC-LIT Orbitrap/MS
(18)	Bladder cancer (BC).	NMIBC (non-muscle invasive BC, n = 54), controls (n = 78).	NMIBC (n = 26), controls (n = 39).	NMIBC panel: dopamine 4-sulfate, MG 18:4, aspartyl-histidine, and tyrosyl-methionine.	Urine	UHPLC-LIT Orbitrap/MS
(19)	Early-stage bladder cancer (BC).	BC (n = 51) and hernia (n = 49).	BC (n = 12) and hernia (n = 12).	Desaminotyrosine, erythritol, d-ribose, ribitol, d-fructose, d-mannose, and d-galactose.	Urine	GC-MS
(20)	Differentiate bladder cancer (BC) from renal cell carcinoma (RCC).	Without hematuria: BC (n = 53), RCC (n = 64) and HC (n = 98). With hematuria: BC (n = 69) and RCC (n = 21).	Without hematuria: BC (n = 24), RCC (n = 30) and HC (n = 44).	Cancer panel: α-CEHC, β-cortolone, deoxyinosine, flunisolid, 11b, 17a, 21-trihydroxypreg-nenolone and glycerol tripropanoate. BC vs RCC: 4-ethoxymethylphenol, prostaglandin F2b, thromboxane B3, CAR 4:0;O, 3-hydroxyphloretin and N'-formylkynurenine.	Urine	UHPLC-QIT/MS
(22)	Discriminate Bladder cancer (BC) and Renal cell carcinoma (RCC).	BC (n = 42), RCC (n = 53) and HC (n = 95).	BC (n = 22), RCC (n = 21) and HC (n = 46).	Cancer panel: FA 18:1;O3(9,10,13), FA 20:1(11Z), FA 18:1;O2(12,13), 6Z-Heneicosen-9-one, FA 18:3, FA 22:0, and FA 16:0;16OH,10O. BC vs RCC panel: 7,8-Dihydropteroic acid, PS P-38:0, FA 18:1;O3(9,10,13), FA 18:2;O, 3,4-Dimethyl-5-pentyl-2-furanundecanoic acid and SPB 18:2;O2.	Plasma	UHPLC-LIT Orbitrap/MS
(24)	Bone metastasis. MRONJ.	MRONJ stage≥1 (n = 9) and control (n = 9).	MRONJ stage≥1 (n = 8) and control (n = 9).	Hypotaurine concentrations were significantly higher in the MRONJ.	Saliva	CE-TOF/MS
(26)	Cachexia.	Cachexia (n = 46), pre-cachectic (n = 22), cancer (n = 68) and HC (n = 28).	Cachexia (n = 38), pre-cachectic (n = 11), cancer (n = 37) and HC (n = 26).	Creation a distinct diagnostic model based on three metabolites: carnosine, leucine and phenylacetate.	Serum and urine	¹ H NMR

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(27)	Breast cancer (BRC).	BRC (n = 44), benign tumor (BE) (n = 38) and controls (n = 34).	BRC (n = 30), BE (n = 30) and controls (n = 30).	BRC panel: a higher level of taurine and glutamic acid, and the lower ethylmalonic acid.	Serum	UHPLC-QqQ/MS and GC-QIT/MS
(28)	Breast cancer (BRC).	BRC (n = 40) and HC (n = 30).	BRC (n = 30) and HC (n = 16).	CAR 8:0, 5-oxo-proline, hypoxanthine, and FA 22:6.	Plasma	UHPLC-QTOF/MS
(29)	Breast cancer (BRC).	BRC (n = 134) and HC (n = 57).	BRC (n = 224) and HC (n = 81).	Glucuronate, FA 18:0, FA 8:0;OH, LPA 16:0, 1-(1-enyl-stearoyl)-2-dihomo-linolenoyl-GPE, SM, bilirubin, urea, and 3-methylxanthine.	Plasma	UHPLC-LIT Orbitrap/MS and GC-MS
(30)	Breast cancer (BRC).	BRC (n = 80) and HC (n = 100).	BRC (n = 109) and HC (n = 50).	Glutamine, ornithine, threonine, tryptophan, methionine sulfoxide, CAR 2:0, CAR 3:0.	Plasma	UHPLC-MS/MS and FIA-MS/MS
(31)	Breast cancer (BRC).	Pre-surgery BRC (n= 69) and controls (n = 48).	Pre-surgery BRC (n= 62), post-surgery (n = 31) and controls (n = 121).	FA 18:2/FA 18:1 ratio, total ethanolamine plasmalogen, and total lysophosphatidylcholine.	Plasma	FIA-QIT Orbitrap/MS
(32)	Cervical cancer (CC).	CC (n = 54), CIN1 (Cervical intraepithelial neoplasia grade 1, n = 54), CIN2/3 (n = 27) and control (n = 70).	CC (n = 60), CIN1 (n = 55), CIN2/3 (n = 42) and control (n = 69).	AMP, aspartate, glutamate, hypoxanthine, lactate, proline, and pyroglutamate.	Plasma	UHPLC-QTOF/MS and UHPLC-QqQ/MS
(33)	Cholangiocarcinoma (CCA).	CCA (n = 176, 88 male and 88 female) and HC (n = 85).	CCA (n = 225) an HC (n = 101).	External validation showed that the combination of these four metabolites (21-deoxycortisol, bilirubin, LPC 14:0, and LPC 15:0) was effective for CCA diagnosis.	Serum	UHPLC-QTOF/MS
(34)	Distal cholangiocarcinoma (dCCA) and pancreatic ductal adenocarcinoma (PDAC).	dCCA (n = 16), PDAC (n = 19), BPD (benign pancreatic disease, n = 22) and control (n = 12).	dCCA (n = 18), PDAC (n = 19), BPD (n = 20) and control (n = 13).	dCCA vs PDAC panel: CAR 16:0, Cer 42:1;O2, LPC 20:0, PC O-16:0/20:3, LPE P-18:2, SM 40:2;O2 and SM 41:2;O2.	Serum	UHPLC-TOF/MS and UHPLC-SQ/MS
(36)	Cancer-related fatigue reported during pediatric acute lymphoblastic leukemia (ALL) therapy.	Discovery cohort: n = 86 Replication cohort: n = 85. (NS).	N = 48 from replication cohort. (NS).	Asparagine and gamma-glutamylglutamine.	Cerebrospinal fluid	UHPLC-MS/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(35)	Chronic myeloid leukemia (CML).	SCML (sensitive to tyrosine kinase inhibitors, n = 26), RCML (resistant to TKI treatment n = 26) and HC (n = 26).	SCML (sensitive to tyrosine kinase inhibitors, n = 112), RCML (resistant to TKI treatment n = 82) and HC (n = 26).	Myristic acid and glycerol for TKI therapeutic.	Plasma	GC-MS
(38)	Colorectal cancer (CRC).	n = 124 (NS).	n = 123 (NS).	Butyrate, 3-hydroxy-3-methylglutarate, and carnosine.	Urine	CE-TOF/MS
(37)	Colorectal cancer (CRC).	(38)	New samples added to before cohort. Total: CRC (n = 214), AD (adenoma, n = 34) and HC (n = 36).	AD vs HC: N ⁸ -Acetylspermidine, N ¹ ,N ⁸ -Diacetylspermidine and 2-Oxoglutarate. CRC vs AD: N ⁸ -Acetylspermidine, N ¹ ,N ⁸ -Diacetylspermidine, citrate and citrulline. AD + CRC vs HC: N ⁸ -Acetylspermidine, N ¹ ,N ⁸ -Diacetylspermidine, citrate.	Urine	CE-TOF/MS
(39)	Colorectal cancer (CRC).	CRC _{stage1} (n = 50), CRC _{stage2} (n = 50), CRC _{stage3} (n = 50) and CE (Chronic enteritis, n = 50).	CRC _{stage1} (n = 62), CRC _{stage2} (n = 209), CRC _{stage3} (n = 128) and CE (n = 23).	The optimal model includes 16 biomarkers.	Plasma	UHPLC-QTOF/MS
(40)	Colorectal cancer (CRC). Treatment futility.	PR (partial response, n = 22), PD (progressive disease, n = 24) and SD (stable disease, n = 22).	PR (partial response, n = 6), PD (progressive disease, n = 33) and SD (stable disease, n = 73).	PD vs PR: stable model of 21 metabolites.	Plasma	GC-MS
(41)	Colorectal Cancer (CRC).	CRC (n = 15) and HC (n = 15).	CRC I-II stage (n = 20), CRC III-IV stage (n = 28), HC (n = 20), BC (n = 14), ulcerative colitis (UC) (n = 20) and high-grade dysplasia adenoma (HGDA) (n = 15).	C10 fatty acid could be a valuable early diagnostic biomarker of CRC.	Plasma	GC-MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(44)	Colorectal cancer (CRC).	CRC (n = 32, stage I (n = 5), stage II (n = 5), stage III (n = 5) and stage IVa (n = 17)) and HC (n = 21).	An independent set of CRC (n = 28, stage I (n = 3), stage II (n = 3), stage III (n = 3) and stage IVa (n = 19)) and HC (n = 20).	The metabolomic profile that distinguished CRC from controls consisted of 48 metabolites. Selecting the most important metabolites, they generated a biomarker pattern consisting of 18 compounds.	Serum	FIA-QqQ/MS (MRM)
(45)	Colorectal cancer (CRC).	CRC (n = 43) and HC (n = 32).	CRC (n = 15) and HC (n = 15).	Serine, glycine, and methionine.	Serum	UHPLC-QTOF/MS
(42)	Colorectal cancer (CRC).	Tissue: CRC (n = 18) and HC (n = 18) paired.	Tissue: CRC (n = 42) and HC (n = 42) paired. Plasma: CRC pre-surgery (with tumor, n = 34) and CRC post-surgery (without tumor, n = 34) paired.	Panel of 20 metabolites from tissue panel. 17 metabolites were validated in post operative plasma samples.	Tissue and plasma	UHPLC-QTOF/MS and UHPLC-QqQ/MS
(43)	Colorectal cancer (CRC).	Tissue: CRC (n = 50) and HC (n = 50) paired. Feces: CRC (pre- and postoperative, n = 50) and HC (n = 50).	Feces: CRC (n = 20) and HC (n = 20).	Acetate fecal.	Tissue and feces.	¹ H NMR
(46)	Endometrial cancer (EC).	EC (n = 33) and HC (n = 36).	An independent validation group: EC (n = 23) and HC (n = 24).	3-hydroxybutyrate, CAR 14:2 and CAR 6:0 appeared to be the best predictors of EC.	Serum	¹ H NMR and HPLC-QqQ/MS
(47)	Endometrial cancer (EC).	EC (n = 88) and HC (n = 80).	EC (n = 30), ovarian cancer (n = 30), benign endometrial disease (n = 10) and HC (n = 50).	The most important metabolites were lactic acid, progesterone, homocysteine, 3-hydroxybutyrate, linoleic acid, stearic acid, myristic acid, threonine and valine.	Serum	GC-MS
(48)	Endometrial cancer (EC).	EC (n = 17) and control (n = 28).	EC (n = 4) and control (n = 5).	Phosphocholine, asparagine, and malate.	Cervicovaginal fluid	¹ H NMR
(50)	Epithelial ovarian cancer (EOC).	EOC (n = 40) and HC (n = 35).	EOC (n = 30) and HC (n = 30).	Esterified fatty acid (EFA) (C16:0), EFA (C18:0) and free fatty acid (FFA) (C16:0) were identified as the potential biomarkers of EOC patients.	Serum	GC-MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(49)	Epithelial ovarian cancer (EOC).	EOC (n = 39) and HC (n = 31).	EOC (n = 29) and HC (n = 28).	Methionine, glutamine, asparagine, glutamic acid and glycolic acid.	Serum	GC-QIT/MS
(51)	Ovarian cancer (OC).	OC (n = 30) and HC (n = 30).	OC (n = 17) and HC (n = 18).	2-piperidinone and LPE 17:0.	Serum	UHPLC-QTOF/MS
(52)	Esophageal adenocarcinoma (EAC).	EAC (n = 30) and HC (n = 30).	EAC (n = 321) and HC (n = 331).	The combined effects of the three metabolites (amino acid L-proline, ketone body 3-hydroxybutyrate, and carbohydrate D-mannose) were investigated by calculating a metabolite risk score (MRS).	Serum	LC-MS/MS
(53)	Esophageal adenocarcinoma (EAC).	Tissue: EAC (n = 50) and control (n = 50) paired. Serum: preoperative EC (n = 50) and control (n = 50).	Serum: EAC (n = 20) and control (n = 20).	Serum: acetate and pyruvate.	Tissue and serum	¹ H NMR
(54)	Esophageal squamous cell carcinoma (ESCC).	ESCC (n = 77) and HC (n = 84).	ESCC (n = 20) and HC (n = 21).	FA 12:0, LPA 18:1, and LPC 14:0 may potentially be used for the prediction of cancer stage and progression.	Serum	UHPLC-QTOF/MS
(55)	Esophageal squamous cell carcinoma (ESCC).	ESCC (n = 40), ESD (esophageal squamous dysplasia, n = 33) and HC (n = 37).	ESCC (n = 34), ESD (n = 39) and HC (n = 38).	ESCC/ESD panel: propanoic acid, linoleic acid, glycerol-3-phosphate, and L-glutamine. ESD panel: propanoic acid, L-leucine, and hydroxyproline. ESCC panel: hypoxanthine, 2-ketoisocaproic acid, L-glutamate, and L-aspartate.	Serum	GC-MS
(56)	Esophageal squamous cell carcinoma (ESCC).	ESCC (n = 80) and HC (n = 80).	ESCC (n = 60) and HC (n = 90).	Hypoxanthine, proline, betaine, indoleacrylic acid, inosine, CAR 10:1, FA 24:6, LPE 20:4, and LPC 20:5.	Plasma	UHPLC-QTOF/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(57)	Esophageal squamous cell carcinoma (ESCC).	ESCC (n = 44); mid- and post-treatment, ESCC-M and ESCC-P (n = 32); different responses to treatment, ESCC overall responders, OR (n = 21) and non-overall, non-OR (n = 11) and HC (n = 24).	ESCC (n = 80); ESCC-M and ESCC-P (n = 40); OR (n = 28) and non-OR (n = 12) and HC (n = 80).	Eleven of the resulting differential metabolites were identified as a potential therapeutic biomarker: two amino acids, five acylcarnitines, and four lysophatidylcholines, LPCs.	Plasma	HPLC-QTOF/MS
(58)	Esophageal squamous cell carcinoma (ESCC).	ESCC (n = 105) and control (n = 45).	ESCC (n = 36) and control (n = 25).	Uridine-diphosphate-n-acetylgalactosamine and maleylacetoacetic.	Tissue	UHPLC-Q-Orbitrap/MS
(62)	Gastric cancer (GCa).	GCa (n = 49) and FDR (first-degree relatives at higher risk of GC development, n = 37).	GCa (n = 22) and FDR (n = 17).	SM 22:1;OH and CAR 16:0.	Serum	HPLC-QqQ/MS
(104)	Gastric cancer (GaC). Response to chemotherapy.	Chemo-sensitive group (n = 16) and chemoresistant group (n = 15).	Chemo-sensitive group (n = 8) and chemoresistant group (n = 8).	Deoxyribose 1-phosphate, S-lactoylglutathione, LPC 16:0, and O-arachidonoyl ethanolamine.	Serum	HPLC-QTOF/MS
(61)	Gastric cancer (GCa).	GCa (n = 30) and precancerous gastric lesions (n = 170).	GCa (n = 48) and precancerous gastric lesions (n = 152).	LPC 18:3, LPC 20:3, FA 20:4, FA 18:3, FA 18:2, and FA 16:0.	Plasma	UHPLC-QTOF/MS
(105)	Gastric cancer (GCa).	GCa (n = 43) and HC (n = 40).	GCa (n = 41) and HC (n = 40).	D-Ile, D-Ser and β -(Pyrazol-1-yl)-L-alanine.	Urine	HPLC-QqQ/MS and HPLC-LIT TOF/MS
(59)	Gastric cancer (GCa).	GCa (n = 47) and HC (n = 47).	Gastric cancer (n = 112, 37 of early gastric cancer) and HC (n = 87).	14 metabolites show diagnostic value. Importantly, proline, p-cresol and 4-hydroxybenzoic acid disclose outcome-prediction value by means of survival analysis.	Urine	GC-MS
(60)	Gastric Cancer (GCa).	GCa (n = 50) and HC (n = 50); and matched tumor (n = 30) and normal stomach tissues (n = 30).	Case (n = 23, 14 with early GCa) and HC (n = 31).	A total of 17 urinary metabolites were tentatively identified as potential biomarkers for the diagnosis of GC, and seven showed correlations with cancer T stage.	Urine and tissues.	¹ H NMR

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(63)	Glioma.	Glioma (n = 42), 25 low-grade glioma (Grade I: n = 5; Grade II: n = 20) and 17 high-grade glioma (Grade III: n = 13; Grade IV GBM: n = 4).	Glioma (n = 45), 17 low-grade glioma (Grade I: n = 4; Grade II: n = 13) and 28 high-grade glioma (Grade III: n = 14; Grade IV GBM: n = 14).	Five metabolites (uracil, arginine, lactate, cystamine, and ornithine) significantly differed between high- and low-grade glioma patients in both cohorts.	Plasma	HPLC-QqQ/MS
(64)	Glioma.	Case (n = 64) and control (n = 64).	Case (n = 68) and control (n = 68).	Myo-inositol, cysteine, uric acid, N-acetylglucosamine, creatinine, glycine and erythronic acid.	Plasma	GC-MS
(65)	Hepatocellular carcinoma (HCC).	HCC (n = 26), hepatitis B virus (HBV) (n = 32), liver cirrhosis (n = 34) and HC (n = 40).	HCC (n = 13), hepatitis B virus (HBV) (n = 17), liver cirrhosis (n = 18) and HC (n = 21).	Phenylalanine, malic acid, and 5-methoxytryptamine for HBV vs NC, palmitic acid for cirrhosis vs HBV, and asparagine and β -glutamate for HCC vs cirrhosis were screened as the liver disease-specific potential biomarkers with an excellent discriminant performance.	Serum	GC-TOF/MS
(66)	Hepatocellular carcinoma (HCC).	HCC (early, n = 53), cirrhosis (n = 47) and controls (n = 50).	HCC (early, n = 82) and cirrhosis (n = 80).	Methionine, proline, ornithine, pimelylcarnitine and CAR 8:0.	Serum	GC-TOF/MS, FIA-QTOF/MS and UHPLC-QTOF/MS
(69)	Hepatocellular Carcinoma (HCC).	DEN-induced rat hepatocarcinogenesis model (LC-MS).	Two groups: group I, cirrhosis (n = 25) and HCC (n = 22); group II cirrhosis (n = 25) and HCC (n = 50, 20 small HCC and 30 general HCC).	Two metabolites (betaine and creatine) were defined as biomarker candidates.	Serum	CE-TOF/MS
(68)	Hepatocellular carcinoma (HCC).	HCC (n = 50) and distal noncancerous tissues (DNT) (n = 50).	HCC (n = 18), cirrhosis (n = 20) and HC (n = 20).	CAR 2:0 could serve as a marker for monitoring the development of hepatocellular carcinoma.	Serum	UHPLC-QTOF/MS and GC-QTOF/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(67)	Hepatocellular carcinoma (HCC).	HCC (n = 22), cirrhosis (n = 30) and HC (n = 30).	HCC (n = 50, 20 small HCC and 30 general HCC), cirrhosis (n = 25) and HC (n = 31).	"Serum biomarker model" including tryptophan, glutamine, and 2-hydroxybutyric acid was established for diagnosing HCC from non-HCC (health and cirrhosis). This model enabled the discrimination of small HCC from pre-cancer cirrhosis highlighting the potential of early diagnosis.	Serum	CE-TOF/MS.
(70)	Hepatocellular carcinoma (HCC).	HCC (n = 529), ICC (intrahepatic CCA, n = 55), HCA (Hepatocellular adenoma, n = 75) and HC (n = 102).	HCC (n = 95) and HC (n = 42).	Three ratios: GCA/TCA, GCDCA/TCDCA, and GDCA/TDCA.	Serum	UHPLC-QqQ/MS
(9)	Hepatocellular carcinoma (HCC).	Discovery set: HCC (n = 36), cirrhosis (n = 41) and HC (n = 31). Test set: small-HCC (S-HCC)/HCC (n = 92/325), cirrhosis (n = 126), gastric cancer (GCa) (n = 73) and HC (n = 160).	Cohort I: (S-HCC)/HCC (n = 37/155), cirrhosis (n = 143), intrahepatic cholangiocarcinoma (ICC) (n = 25), chronic hepatitis B virus infection (CHB) (n = 150) and HC (n = 99). Cohort II: HCC (n = 42) and HC (n = 42).	The combination of phenylalanyl-tryptophan (Phe-Trp) and glycocholate (GCA) was defined as the ideal biomarker panel to distinguish patients with HCC from subjects without HCC.	Serum and tissue	UHPLC-QqQ/MS
(71)	Hepatocellular carcinoma (HCC).	HCC (n = 139), chronic hepatitis (n = 81) and cirrhosis (n = 78).	HCC (n = 22) and cirrhosis (n = 25).	Combination biomarker, betaine, and CAR 3:0, is useful for the diagnosis of hepatocellular carcinoma.	Serum and tissue	UHPLC-QIT Orbitrap and CE-TOF/MS
(72)	Hepatocellular carcinoma (HCC).	Tissue: HCC (n = 156) and DNT (distal non-cancerous tissue, n = 156) paired.	Tissue: HCC (n = 80), cirrhosis (n = 20) and control (n = 20). Serum: HCC (n = 80), cirrhosis (n = 20) and control (n = 20) paired with tissue.	Retinol metabolism panel: retinol, retinal, retinoic acid, retinyl palmitate and β -carotene.	Serum and tissue	UHPLC-QTOF/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(73)	Hepatocellular carcinoma (HCC).	HCC (n = 33), cirrhosis (n = 27) and HC (n = 26).	HCC (n = 33), cirrhosis (n = 21) and HC (n = 26).	CAR 4:0 and hydantoin-5-propionic acid were defined as combinational markers to distinguish HCC from cirrhosis.	Urine	Nontargeted: UHPLC-QqQ/MS Pseudotargeted: UHPLC-QTOF/MS.
(74)	Hepatocellular carcinoma (HCC).	HCC _{AFP+} (α -fetoprotein+, n = 15), HCC _{AFP-} (α -fetoprotein-, n = 13), HBG (hepatitis B group, n = 16) and HC (n = 14).	HCC _{AFP+} (n = 15), HCC _{AFP-} (n = 27), HBG (n = 14) and HC (n = 16).	HCC vs HBG: LPC (24:0). HCC panel: LPC 17:0, LPC 20:4, LPC 22:0, LPC 24:0, PE P-16:0/22:4, SM 18:1/22:1;O2, creatinine, and L-isoleucine.	Plasma	HPLC-Q Orbitrap/MS
(76)	Lung cancer (LCa) clinical evaluation platinum efficacy.	PR (partial response, n = 12), PD (progressive disease, n = 16) and SD (stable disease, n = 15).	PR (n = 14) and PD (n = 12).	Phenylalanine, pyroglutamic acid, tyrosine, tryptophan, citric/isocitric acid, 2-hydroxyglutaric acid, succinic acid, α -ketoglutaric acid.	Plasma	UHPLC-QIT/MS
(106)	Lung Cancer (LCa).	Phase I: HC (n = 20), early-stage (n = 20), and late-stage lung cancer cases (n = 20) (after referred to as "trio").	50 trios (phase II) and 123 trios (phase III).	Identified and validated serum bilirubin as a risk predictor for lung cancer incidence as well as mortality in male smokers.	Serum	HPLC-QqQ/MS
(77)	Lung cancer (LCa): Pem-Cis, pemetrexed-cisplatin; pem-Carbo, pemetrexed-carboplatin.	Pem-cis: PR (partial response, n = 51), PD (progressive disease, n = 29) and SD (stable disease, n = 81). Pem-carbo: PR (n = 31), PD (n = 24) and SD (n = 35).	Pem-cis: PR (n = 22), PD (n = 12) and SD (n = 29). Pem-carbo: PR (n = 14), PD (n = 11) and SD (n = 15).	Hypotaurine, uridine, CAR 12:0, choline, dimethylglycine, niacinamide, CAR 16:0.	Serum	HPLC-Q Orbitrap/MS
(107)	Lung Cancer (LCa).	LCa (n = 23) and HC (n = 23).	LCa (n = 9) and HC (n = 9).	Choline and linoleic acid were defined as one combinational biomarker using binary logistic regression.	Serum	UHPLC-QTOF/MS
(108)	Lung cancer (LCa).	LCa (n = 30) and HC (n = 30).	LCa (n = 42) and HC (n = 38).	1-Methylhistidine, xanthosine, 8-hydroxynevirapine, tyrosyl-tyrosine, indoxyl.	Urine	UHPLC-QTOF/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(109)	Non-small cell lung cancer (NSCLC).	NS.	NS.	DC vs PD: pyruvate, methionine, salicylic acid, creatinine, and serine. CR/PR vs SD/PD: uric acid, fumarate, malonic acid, tryptophan, methionine sulfoxide, and creatinine.	Tissue	HPLC-Q Orbitrap/MS
(110)	Non-small cell lung cancer (NSCLC).	Non-small cell lung cancer (NSCLC) (n = 100) and HC (n = 199).	NSCLC (n = 108) and HC (n = 216).	They demonstrated significant additive performance of the combination of N ₁ ,N ₁₂ -diacetylspermine (DAS), and pro-surfactant protein B (pro-SFTPB) compared with either biomarker alone.	Serum	HPLC-QTOF/MS
(111)	Non-small cell lung cancer (NSCLC).	NSCLC (n = 90) and HC (n = 90).	NSCLC (n = 30) and HC (n = 30).	Increased saturated and monounsaturated PCs and decreased polyunsaturated PCs.	Serum	UHPLC-QTOF/MS, GC-TOF/MS
(112)	Non-small cell lung cancer (NSCLC).	LCa (lung cancer, n = 30), BC (benign tumor control, n = 6) and HC (n = 30).	LCa (n = 35) and HC (n = 35).	Cysteine, serine, and MG 18:1.	Serum	GC-MS
(113)	Non-small cell lung cancer (NSCLC).	NSCLC (n = 99) and HC (n = 109).	NSCLC (n = 43) and HC (n = 50).	Hypoxanthine, inosine, L-tryptophan, indoleacrylic acid, CAR 10:1, and LPC 18:2.	Serum	UHPLC-QTOF/MS
(78)	Non-small-cell lung cancer (NSCLC). Cell death protein 1 (PD-1) therapy.	Advanced NSCLC treated with nivolumab (n = 43), and controls (n = 43) paired, after 2 weeks of treatment.	Set 1: advanced NSCLC treated with tislelizumab (n = 21), and controls (n = 10) paired, after 3 weeks of treatment. Set 2: NSCLC treated with nivolumab (n = 10), and controls (n = 10) paired, after 2 weeks of treatment.	Hypoxanthine and histidine.	Serum	UHPLC-Q Orbitrap/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(114)	Non-small cell lung cancer (NSCLC).	Good survival (n = 55) and with poor survival (n = 55).	Good survival (n = 55) and with poor survival (n = 55).	They found levels of 4 metabolites, caffeine, paraxanthine, stachydrine, and methyl glucopyranoside (alpha +beta), differed significantly between NSCLC patients with poor and good survival.	Serum and plasma	HPLC-QqQ/MS
(115)	Non-small cell lung cancer (NSCLC), early-stage.	ADC _{stageI} (n = 32), SCC _{stageI} (n = 15), ADC _{stageII} (n = 29), SCC _{stageII} (n = 11), ADC _{stageIII/IV} (n = 20), SCC _{stageIII/IV} (n = 6) and HC (n = 40).	ADC _{stageI} (n = 18), SCC _{stageI} (n = 5), ADC _{stageII} (n = 11), SCC _{stageII} (n = 9) and HC (n = 20).	PC O-40:6, PC 38:0, PC 40:2, LPC 20:3, LPC 18:2, β-hydroxybutyric acid, CAR, tryptophan, methionine sulfoxide, citrate and fumaric acid.	Plasma	HPLC-LIT/MS FIA-LIT/MS
(116)	Non-small cell lung cancer (NSCLC).	ADC (adenocarcinoma, n = 28), SCC (squamous cell carcinoma, n = 25) and controls (n = 56).	ADC (adenocarcinoma, n = 24), SCC (squamous cell carcinoma, n = 22) and controls (n = 56).	NSCLC clinical diagnosis purposes: cortisol, cortisone, and 4-methoxyphenylacetic acid.	Plasma	UHPLC-LIT FTICR/MS
(75)	Small cell lung cancer (SCLC) vs non-small cell lung cancer (NSCLC).	SCLC (n = 37), NSCLC (n = 40), and controls (n = 37).	SCLC (n = 31), NSCLC (n = 20), and controls (n = 27).	SCLC diagnosis in males including PI 18:0/18:0, Cer 18:1/22:0;2O, MG 20:4, inosine monophosphate and Cholic acid. SCLC in females: PE 18:1/20:4, 5-methyltetrahydrofolic acid, Desmosterol, 4,5-Dihydroorotic acid and FA 20:4;O(9).	Dried blood spot (DBS)	HPLC-QqQ/MS
(80)	Oral cavity squamous cell carcinoma (OSCC) and oropharyngeal squamous cell carcinoma (OPC).	OSCC (n = 79): early stage (n = 40) and late stage (n = 39). Controls (n = 20). OSCC N+ (node positive, n = 45) and OSCC N- (node negative, n = 45).	OSCC (n = 22): early stage (n = 13) and late stage (n = 9). OPC (n = 58) and controls (n = 20).	Biomarker panel OCC: proline and glycine Biomarker panel early stage: citrulline, glycine, ornithine, and proline.	Saliva	¹ H NMR HPLC-QqQ/MS and HPLC-QTOF/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(79)	Oral squamous cell carcinoma (OSCC).	Pre/Post chemotherapy. With/without significant efficacy. (S/NS). (SPre) (n = 11) and (SPost) (n = 11); (NSPre) (n = 21) and (NSPost) (n = 21).	SPre (n = 4) and NSPre (n = 14).	Applicability of lactic acid, glutamic acid, and aspartic acid as predictors of the response to induction chemotherapy.	Serum	GC-MS
(117)	Pancreatic Cancer (PaC).	PaC (n = 30) and control (n = 30).	PaC (n = 18) and control (n = 18).	Acadesine, aspartic acid, and 1-methyl-L-tryptophan.	Serum	HPLC-QqQ/MS
(87)	Pancreatic cancer (PaC).	PaC (n = 43) and HC (n = 42).	PaC (n = 42), chronic pancreatitis (n = 23) and HC (n = 41).	4 metabolites were selected: xylitol, 1,5-anhydro-D-glucitol, inositol, and histidine.	Serum	GC-MS
(118)	Pancreatic cancer (PaC).	PaC (n = 64) and HC (n = 128).	PaC (n = 51) and HC (n = 102).	Aspartate/alanine, androstenediol monosulfate and glycylvaline.	Serum	UHPLC-Q Orbitrap/MS
(119)	Pancreatic cancer (PaC).	Exosome: PB (peripheral blood, n = 22), PCB (Para cancerous blood, n = 22) paired samples. HC (n = 17).	Exosome and corresponding serum: PB (n = 24), PCB (n = 24) and HC (n = 40).	LPC 22:0, PC P-14:0/22:2 and PE 16:0/18:1.	Exosome and serum	UHPLC-QTOF/MS and HPLC-LIT/MS (MRM)
(88)	Pancreatic cancer (PaC).	PaC (n = 120) and HC (n = 600).	PaC (n = 240), CP (n = 28, chronic pancreatitis), and HC (n = 7772).	PFAA, plasma free amino acids: serine, asparagine, isoleucine, alanine, histidine, and tryptophan, profile of PC was significantly different from HC.	Plasma	HPLC-MS
(89)	Pancreatic cancer (PaC).	Early PaC (n = 55) and HC (n = 58).	Early PaC (n = 16) and HC (n = 16).	To validate the diagnostic performance of their model, they constructed a diagnostic model for PaC based on the results of the 4 metabolites: 1,5-Anhydroglucitol, histidine, inositol, and xylitol.	Plasma	GC-MS
(120)	Pancreatic cancer (PaC).	Plasma: PaC (n = 60) and HC (n = 60).	Tissue: PaC (n = 6) and ANT (adjacent noncancerous tissue n = 6). Paired.	Creatine, inosine, beta-sitosterol, SPB 18:0;O2 and glycocholic acid.	Plasma and tissue.	UHPLC-QqQ/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(81)	Pancreatic ductal adenocarcinoma (PDAC).	PDAC (n = 56) and HC (n = 56).	PDAC (n = 36) and HC (n = 56).	Trigonelline, glycolate, hippurate, creatine, myoinositol and hydroxyacetone.	Urine	¹ H NMR
(82)	Pancreatic ductal adenocarcinoma (PDAC).	(81)	PDAC (n = 88) and HC (n = 56).	Trigonelline, hippurate and myoinositol.	Urine	¹ H NMR
(83)	Pancreatic ductal adenocarcinoma (PDAC).	1. PDAC (early-stage, n = 10; late-stage, n = 10), chronic pancreatitis (CP, n = 10) and HC (n = 70). 2. PDAC (early-stage, n = 5; late-stage, n = 4), low grade pancreatic cystic lesion (n = 51).	PDAC (early-stage, n = 39) and HC (n = 82).	Acetylspermidine, diacetylspermine indole-derivative, LPC 18:0 and LPC 20:3.	Plasma	UHPLC-QTOF/MS
(84)	Pancreatic ductal adenocarcinoma (PDAC).	Pre-PDAC (preoperative, n = 185), post-PDAC (postoperative paired with pre-PDAC, n = 87) and controls (n = 146).	Pre-PDAC (preoperative, n = 50), post-PDAC (postoperative paired with pre-PDAC, n = 32) and controls (n = 146).	Docosahexaenoic acid, LPC 14:0 and Histidinyl-Lysine.	Serum	UHPLC-QTOF/MS
(86)	Discrimination of pancreatic ductal Adenocarcinoma (PDAC) from chronic pancreatitis (CP).	Exploratory study: plasma of PDAC (n = 34), CP (n = 43), liver cirrhosis (n = 20) and BD (blood donors as controls; n = 104) Identification study: Serum of PDAC (n = 80), CP (n = 79), liver cirrhosis (n = 80) and BD (n = 77). Plasma of PDAC (n = 78), CP (n = 80).	Plasma of PDAC (n = 78), CP (n = 80), BD (n = 80).	Biomarker signature comprising nine metabolites and CA19-9.	Plasma and serum	GC-MS and HPLC-QqQ/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(85)	Discrimination of PDAC from chronic pancreatitis (CP).	PDAC (n = 44) and Chronic pancreatitis CP (n = 23).	PDAC (n = 20) and CP (n = 31).	Five-metabolite panel discriminating PDAC and CP: glycocholic acid, N-palmitoyl glutamic acid, CAR 6:0, phenylacetylglutamine and chenodeoxyglycocholate.	Plasma and serum	Discovery study: HPLC-QIT Orbitrap/MS Validation study: UHPLC-QTOF/MS
(121)	Papillary thyroid cancer (PTC).	Female, ♀: PTC (stage I/II, n = 29), BTN (benign thyroid nodule, n = 14) and HC (n = 15).	♀: PTC (stage I/II, n = 62; stage III/IV, n = 4), BTN (n = 36) and HC (n = 64). ♂: PTC (stage I/II, n = 28; stage III/IV, n = 2), BTN (n = 2) and HC (n = 37). Double-blind: thyroid nodules untreated (n = 31).	Serum: β-hydroxybutyrate, docosahexaenoic acid (DHA), 1-methyladenosine, pregnanediol-3-glucuronide. Urine: nicotinic acid mononucleotide and xanthosine. Corrected with creatinine.	Serum and urine	UHPLC-QTOF/MS and GC-MS
(90)	Malignant thyroid cancer (MTC) and benign thyroid tumor (BTT).	MTC (n = 42), BTT (n = 15) and HC (n = 40).	MTC (n = 82), BTT (n = 28) and HC (n = 82).	They have located a biomarker (PC 34:1 and a panel of biomarkers (PA 36:3, and SM 34:1;O2, sphingomyelin) with excellent diagnostic ability to differentiate both MTC and BTT patients from normal individuals.	Serum	MALDI-FTICR/MS
(91)	Papillary thyroid carcinoma (PTC) and benign thyroid nodule (BTN).	Center 1: PTC (n = 141), BTN (n = 93) and HC (n = 100).	Independent centers. Center 1: HC (n = 68). Center 2: PTC (n = 58), BTN (n = 94). Center 3: PTC (n = 19), BTN (n = 96). Center 4: PTC (n = 43), BTN (n = 12). Center 5: PTC (n = 9), BTN (n = 37). Center 6: tissue samples PTC (n = 31), BTN (n = 26).	BTN + PTC tissue panel: myo-inositol, α-N-phenylacetyl-L-glutamine, proline-betaine, L-glutamic acid, LPC 18:0 and LPC 18:1.	Serum, plasma, and tissue	HPLC-QTOF/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(93)	Prostate cancer (PCa).	PCa (n = 39), BPH (Benign Prostatic Hyperplasia, n = 46) and controls (n = 45).	PCa (n = 39) and controls (n = 50).	PE-NMe2 18:0/18:2, PC 16:0/20:2, PS 15:0/18:2, SM 16:0/24:1;O2, CAR 14:0.	Serum	GC-MS and UHPLC-Q Orbitrap/MS
(94)	Prostate cancer (PCa).	BCR (biochemical recurrence, n = 50) and non-BCR (n = 217).	BCR (n = 22) and non-BCR (n = 93).	1-methyladenosine and PA 18:0/22:0.	Serum	UHPLC-QTOF/MS and FIA-QTOF/MS
(95)	Prostate cancer (PCa).	PCa (n = 102) and controls (n = 190).	PCa (n = 44) and controls (n = 82).	FA 20:2, ethyl oleate, phosphate, CAR 5:1, SPB 18:0;O2, L-glutamic acid, serotonin, 2-hydroxyadenine, ST 27:2;O2, SPB 18:1;O2.	Plasma	UHPLC-Q Orbitrap/MS
(96)	Prostate cancer (PCa).	PCa (n = 560) and HC (n = 500).	PCa (n = 451) and HC (n = 438).	The present study provides the diagnosis tool for PCa in its early stage with a panel of biomarkers: 2-isopropyl citrate, cytidine, and D-asparagine.	Plasma	UHPLC-QTOF/MS
(97)	Prostate cancer (PCa).	PCa (n = 40) and control (n = 42).	PCa (n = 18) and control (n = 18).	Hexanal, 2,5-dimethylbenzaldehyde, 4-methylhexan-3-one, dihydroedulan IA, methylglyoxal and 3-phenylpropionaldehyde.	Urine	GC-MS
(99)	Prostate cancer (PCa).	PCa (n = 55) and control (n = 53).	PCa (n = 53) and control (n = 22).	Ethyl- α -hydroxymyristate, estradiol, 2,6-di-t-butyl-4-hydroxymethylene-2,3,5,6-tetrahydrocyclohexanone and 3,5-di-t-butyl-4-hydroxybenzaldehyde.	Urine	GC-MS
(98)	Prostate cancer (PCa).	PCa (n = 41) and controls (n = 42).	PCa (n = 18) and controls (n = 18).	Panel of 21 metabolites.	Urine	GC-MS and ¹ H NMR
(100)	Renal Cell Carcinoma (RCC).	<i>In vitro</i> model of human renal cell lines: RCC (two cell lines, n = 66) and control cells (n = 22).	Serum: RCC (stage IV, n = 10) and HC (n = 10).	Isoleucine/leucine, phenylalanine, N-lactoyl-leucine, N-acetyl-phenylalanine, cysteinylglutathione disulfide and hydroxypropyl-valine.	Cell lines and serum	UHPLC-QTOF/MS
(101)	Renal Cell Carcinoma (RCC).	RCC (n = 67), benign tumor (n = 34) and HC (n = 96).	RCC (n = 33), benign tumor (n = 7) and HC (n = 33).	RCC vs benign: L-3-hydroxykynurenine, 1,7-dimethylguanosine, and tetrahydroaldosterone-3-glucuronide. RCC diagnosis: N-formylkynurenine.	Urine	UHPLC-QIT Orbitrap/MS

Ref.	Tipo cancer /tratamiento	Cohorte de descubrimiento	Cohorte de validación	Biomarcadores/resultado de la validación	Tipo de muestra	Técnica/Equipo
(102)	Colorectal (CRC), pancreatic (PCa), kidney (KD), breast (BRC) and glioblastoma multiforme cancer (GBM).	KD (n = 7), BRC (n = 7), CRC (n = 9), GBM (n = 12) and control (n = 15).	KD (n = 9), CRC (n = 26), PCa (n = 45) and control (n = 16).	Very-long-chain dicarboxylic acids VLCDCAs 28:4 decreased levels in CRC, PC and KD. No alteration levels in BC and GNM.	Plasma	FIA-Q Orbitrap/MS
(103)	Several cancers: LCa (lung cancer), BRC (breast cancer), CRC (colorectal cancer), GCa (gastric cancer).	LCa (n = 28), BRC (n = 18), CRC (n = 21), GC (n = 11) and HC (n = 38).	Before treatment LCa (n = 8) and after treatment LCa (n = 8). Paired.	Sarcosine, glutamic acid, citrulline, tryptophan, glutamine, and histidine for LCa.	Plasma	HPLC-LIT/MS (MRM)

DMA, dimetilamina; LPS, lisoglicerofosfatidilserina; TCA *cyc/e*, vía de los ácidos tricarbónicos; CAR, carnitina; MG, monoglicérido; FA, ácido graso; SPB, base esfingóide; LPA, ácido lisofosfatídico; PA, ácido glicerofosfatídico; AMP, adenosina monofosfato; LPC, lisoglicerofosfatidilcolina, PC, glicerofosfatidilcolina, LPE, lisoglicerofosfatidiletanolamina; PE, glicerofosfatidiletanolamina; PE-NMe2, N,N-dimetilglicerofosfatidiletanolamina; Cer, ceramida; SM, esfingomielina; NS, no especificado; MRONJ, Osteonecrosis de la mandíbula relacionada con medicamentos; GCA, ácido glicocólico; GCDCA, ácido glucoquenodesoxicólico; GDCA, ácido glicodesoxicólico; TCA, ácido taurocólico; TCDCA, ácido tauroquenodesoxicólico; TDCA, ácido taurodesoxicólico; TKI, inhibidor de tirosina quinasa; EFA, ácido graso esterificado; FFA, ácido graso libre; PR, respuesta parcial; PD, enfermedad progresiva; SD, enfermedad estable; CR, respuesta completa; DC, enfermedad-control.

4. Conclusión

En la detección de compuestos cercanos al fenotipo y que cambian rápidamente según la situación fisiopatológica, la metabolómica es la herramienta perfecta para el descubrimiento de biomarcadores de diagnóstico. La selección de los grupos de estudio es un punto crítico para los resultados, aunque a veces las dificultades técnicas propias del reclutamiento de la cohorte son inherentes al investigador. Bien es cierto que, la aplicación de los perfiles metabolómicos en grandes cohortes permite obtener conclusiones más sólidas y hallazgos reproducibles externamente, lo que generaría marcadores clínicos reales.

Para la mejora del diagnóstico, la estadificación, pronóstico y/o terapia, la identificación de biomarcadores metabolómicos en la mayoría de las enfermedades es vital, siempre y cuando se consiga traspasar la barrera de la investigación a la práctica clínica. El hecho de elegir la metabolómica para esta búsqueda está fundamentado en los cambios tempranos y significativos que experimentan los metabolitos frente a los genes o proteínas. Además, estos cambios en la concentración de metabolitos se pueden asignar con las distintas vías metabólicas implicadas, lo que nos permite, en muchos casos, la explicación biológica, fortaleciendo su potencial.

Uno de los principales retos a superar es la falta de una plataforma analítica universal para el análisis de estos. Debido a la gran diversidad fisicoquímica ante la que nos encontramos, este desafío será difícil de superar, y a menudo, estos biomarcadores solo se identifican de forma tentativa, por lo que es fundamental establecer protocolos de consenso estableciéndose reglas claras en la comunidad científica con respecto a la identificación de metabolitos. Por otra parte, existe también la limitación que implica validar paneles diagnósticos utilizando cohortes de estudio más grandes. Desarrollar un experimento de descubrimiento con decenas de muestras es relativamente fácil; sin embargo, la obtención de muestras de diferentes biobancos u hospitales y su medición con los estándares adecuados, incluidos estándares marcados isotópicamente, solo se puede realizar bajo el paraguas de un proyecto internacional bien financiado. La búsqueda de biomarcadores basada en conjuntos de muestras pequeños puede verse alterada por diferentes factores de confusión lo que afecta directamente a los resultados del ensayo. Esto se demuestra al analizar los resultados obtenidos de estudios desarrollados para la misma enfermedad, en el mismo tipo de muestra y la misma instrumentación analítica que frecuentemente son dispares.

La aplicación de biomarcadores metabolómicos a la práctica clínica depende de la validación de estos, lo que demostraría su utilidad clínica, desarrollando pruebas de análisis clínicos adecuados que permitan una medición simple y confiable. Además, es importante resolver otro punto en relación con esto, la protección de la propiedad intelectual e industrial. Los biomarcadores derivados de genes, proteínas o metabolitos existentes de forma natural no son patentables. Los métodos de análisis de metabolitos pueden ser patentados; sin embargo, no existe una forma real de protegerlos, especialmente si se trata de un método cromatográfico.

Con esta revisión se establece una idea general sobre el estado del arte (*state-of-art*) que permite poner en contexto el tema central de esta tesis doctoral. El potencial de los metabolitos como marcadores de diagnóstico para llegar a la clínica y marcar una diferencia real en la salud del paciente es sustancial, pero quedan desafíos que deben explorarse.

5. Referencias

1. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. *Nature Reviews Molecular Cell Biology*. 2016;17(7):451-9.
2. Naz S, Vallejo M, García A, Barbas C. Method validation strategies involved in non-targeted metabolomics. *Journal of Chromatography A*. 2014;1353:99-105.
3. Wagner JA. Overview of biomarkers and surrogate endpoints in drug development. *Disease markers*. 2002;18(2):41-6.
4. Drucker E, Krapfenbauer K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *EPMA journal*. 2013;4(1):7.
5. Dobbin KK, Cesano A, Alvarez J, Hawtin R, Janetzki S, Kirsch I, et al. Validation of biomarkers to predict response to immunotherapy in cancer: Volume II—clinical validation and regulatory considerations. *Journal for immunotherapy of cancer*. 2016;4(1):77.
6. Masucci GV, Cesano A, Hawtin R, Janetzki S, Zhang J, Kirsch I, et al. Validation of biomarkers to predict response to immunotherapy in cancer: volume I—pre-analytical and analytical validation. *Journal for immunotherapy of cancer*. 2016;4(1):76.
7. Dudzik D, Barbas-Bernardos C, García A, Barbas C. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. *Journal of pharmaceutical and biomedical analysis*. 2017;147:149-73.
8. Xia JG, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics*. 2013;9(2):280-99.
9. Luo P, Yin P, Hua R, Tan Y, Li Z, Qiu G, et al. A Large-scale, multicenter serum metabolite biomarker identification study for the early detection of hepatocellular carcinoma. *Hepatology (Baltimore, Md)*. 2018;67(2):662-75.
10. Beger RD. A review of applications of metabolomics in cancer. *Metabolites*. 2013;3(3):552-74.
11. Zhang ZH, Chen H, Vaziri ND, Mao JR, Zhang L, Bai X, et al. Metabolomic Signatures of Chronic Kidney Disease of Diverse Etiologies in the Rats and Humans. *Journal of Proteome Research*. 2016;15(10):3802-12.
12. Konety BR. Molecular markers in bladder cancer: a critical appraisal. *Urol Oncol*. 2006;24(4):326-37.
13. Pasikanti KK, Esuvaranathan K, Hong Y, Ho PC, Mahendran R, Raman Nee Mani L, et al. Urinary metabotyping of bladder cancer using two-dimensional gas chromatography time-of-flight mass spectrometry. *Journal of proteome research*. 2013;12(9):3865-73.
14. Jin X, Yun SJ, Jeong P, Kim IY, Kim W-J, Park S. Diagnosis of bladder cancer and prediction of survival by urinary metabolomics. *Oncotarget*. 2014;5(6):1635.
15. Wittmann BM, Stirdivant SM, Mitchell MW, Wulff JE, McDunn JE, Li Z, et al. Bladder cancer biomarker discovery using global metabolomic profiling of urine. *PLoS One*. 2014;9(12):e115870.
16. Zhou Y, Song R, Ma C, Zhou L, Liu X, Yin P, et al. Discovery and validation of potential urinary biomarkers for bladder cancer diagnosis using a pseudotargeted GC-MS metabolomics method. *Oncotarget*. 2017;8(13):20719.

17. Liu X, Cheng X, He L, Zhang W, Wang Y, Sun W, et al. Investigation of the urinary metabolic variations and the application in bladder cancer biomarker discovery. *Int J Cancer*. 2018;143(2):408-18.
18. Cheng X, Liu X, Guo Z, Sun H, Zhang M, Ji Z, et al. Metabolomics of Non-muscle Invasive Bladder Cancer: Biomarkers for Early Detection of Bladder Cancer. *Front Oncol*. 2018;8:494.
19. Lin J-Y, Juo B-R, Yeh Y-H, Fu S-H, Chen Y-T, Chen C-L, et al. Putative markers for the detection of early-stage bladder cancer selected by urine metabolomics. *BMC Bioinformatics*. 2021;22(1).
20. Wang Z, Liu X, Sun H, Guo Z, Zheng G, Zhang Y, et al. UPLC-MS based urine untargeted metabolomic analyses to differentiate bladder cancer from renal cell carcinoma. *BMC Cancer*. 2019;19(1):1195.
21. Bansal N, Gupta A, Mitash N, Shakya PS, Mandhani A, Mahdi AA, et al. Low- and High-Grade Bladder Cancer Determination via Human Serum-Based Metabolomics Approach. *Journal of Proteome Research*. 2013;12(12):5839-50.
22. Liu X, Zhang M, Cheng X, Sun H, Guo Z, Li J, et al. LC-MS-Based Plasma Metabolomics and Lipidomics Analyses for Differential Diagnosis of Bladder Cancer and Renal Cell Carcinoma. *Front Oncol*. 2020;10:717.
23. Tan GG, Wang HB, Yuan JL, Qin WJ, Dong X, Wu H, et al. Three serum metabolite signatures for diagnosing low-grade and highgrade bladder cancer. *Scientific Reports*. 2017;7:46176.
24. Yatsuoka W, Ueno T, Miyano K, Uezono Y, Enomoto A, Kaneko M, et al. Metabolomic profiling reveals salivary hypotaurine as a potential early detection marker for medication-related osteonecrosis of the jaw. *PLoS One*. 2019;14(8):e0220712.
25. Lalla RV, Bowen J, Barasch A, Elting L, Epstein J, Keefe DM, et al. MASCC/ISOO clinical practice guidelines for the management of mucositis secondary to cancer therapy. *Cancer*. 2014;120(10):1453-61.
26. Yang QJ, Zhao JR, Hao J, Li B, Huo Y, Han YL, et al. Serum and urine metabolomics study reveals a distinct diagnostic model for cancer cachexia. *Journal of cachexia, sarcopenia and muscle*. 2018;9:171-85.
27. Wang X, Zhao X, Chou J, Yu J, Yang T, Liu L, et al. Taurine, glutamic acid and ethylmalonic acid as important metabolites for detecting human breast cancer based on the targeted metabolomics. *Cancer Biomark*. 2018;23(2):255-68.
28. Park J, Shin Y, Kim TH, Kim DH, Lee A. Plasma metabolites as possible biomarkers for diagnosis of breast cancer. *PLoS One*. 2019;14(12):e0225129.
29. Zhao H, Shen J, Moore SC, Ye Y, Wu X, Esteva FJ, et al. Breast cancer risk in relation to plasma metabolites among Hispanic and African American women. *Breast Cancer Res Treat*. 2019;176(3):687-96.
30. Yuan B, Schaffner S, Tang Q, Scheffler M, Nees J, Heil J, et al. A plasma metabolite panel as biomarkers for early primary breast cancer detection. *Int J Cancer*. 2019;144(11):2833-42.
31. Tomida S, Goodenowe DB, Koyama T, Ozaki E, Kuriyama N, Morita M, et al. Plasmalogen Deficiency and Overactive Fatty Acid Elongation Biomarkers in Serum of Breast Cancer Patients Pre- and Post-Surgery-New Insights on Diagnosis, Risk Assessment, and Disease Mechanisms. *Cancers (Basel)*. 2021;13(16) :4170.

32. Khan I, Nam M, Kwon M, Seo SS, Jung S, Han JS, et al. LC/MS-Based Polar Metabolite Profiling Identified Unique Biomarker Signatures for Cervical Cancer and Cervical Intraepithelial Neoplasia Using Global and Targeted Metabolomics. *Cancers (Basel)*. 2019;11(4):511.
33. Liang Q, Liu H, Zhang TY, Jiang Y, Xing HT, Zhang H. Serum metabolomics uncovering specific metabolite signatures of intra- and extrahepatic cholangiocarcinoma. *Molecular Biosystems*. 2016;12(2):334-40.
34. Macias RIR, Muñoz-Bellvís L, Sánchez-Martín A, Arretxe E, Martínez-Arranz I, Lapitz A, et al. A Novel Serum Metabolomic Profile for the Differential Diagnosis of Distal Cholangiocarcinoma and Pancreatic Ductal Adenocarcinoma. *Cancers (Basel)*. 2020;12(6):1433.
35. Yang B, Wang C, Xie Y, Xu L, Wu X, Wu D. Monitoring tyrosine kinase inhibitor therapeutic responses with a panel of metabolic biomarkers in chronic myeloid leukemia patients. *Cancer Sci*. 2018;109(3):777-84.
36. Brown AL, Sok P, Taylor O, Woodhouse JP, Bernhardt MB, Raghubar KP, et al. Cerebrospinal Fluid Metabolomic Profiles Associated With Fatigue During Treatment for Pediatric Acute Lymphoblastic Leukemia. *J Pain Symptom Manage*. 2021;61(3):464-73.
37. Sakurai T, Katsumata K, Udo R, Tago T, Kasahara K, Mazaki J, et al. Validation of Urinary Charged Metabolite Profiles in Colorectal Cancer Using Capillary Electrophoresis-Mass Spectrometry. *Metabolites*. 2022;12(1):59.
38. Udo R, Katsumata K, Kuwabara H, Enomoto M, Ishizaki T, Sunamura M, et al. Urinary charged metabolite profiling of colorectal cancer using capillary electrophoresis-mass spectrometry. *Sci Rep*. 2020;10(1):21057.
39. Zhang F, Li C, Deng K, Wang Z, Zhao W, Yang K, et al. Metabolic phenotyping to monitor chronic enteritis canceration. *Metabolomics*. 2020;16(3):29.
40. Rattner JI, Kopciuk KA, Vogel HJ, Tang PA, Shapiro JD, Tu D, et al. Early detection of treatment futility in patients with metastatic colorectal cancer. *Oncotarget*. 2022;13:61-72.
41. Crotti S, Agnoletto E, Cancemi G, Di Marco V, Traldi P, Pucciarelli S, et al. Altered plasma levels of decanoic acid in colorectal cancer as a new diagnostic biomarker. *Analytical and bioanalytical chemistry*. 2016;408(23):6321-8.
42. Zha H, Cai Y, Yin Y, Wang Z, Li K, Zhu ZJ. SWATHtoMRM: Development of High-Coverage Targeted Metabolomics Method Using SWATH Technology for Biomarker Discovery. *Anal Chem*. 2018;90(6):4062-70.
43. Lin Y, Ma C, Bezabeh T, Wang Z, Liang J, Huang Y, et al. (1) H NMR-based metabolomics reveal overlapping discriminatory metabolites and metabolic pathway disturbances between colorectal tumor tissues and fecal samples. *Int J Cancer*. 2019;145(6):1679-89.
44. Farshidfar F, Kopciuk KA, Hilsden R, McGregor SE, Mazurak VC, Buie WD, et al. A quantitative multimodal metabolomic assay for colorectal cancer. *Bmc Cancer*. 2018;18(1):26.
45. Bian X, Sun B, Zheng P, Li N, Wu JL. Derivatization enhanced separation and sensitivity of long chain-free fatty acids: Application to asthma using targeted and non-targeted liquid chromatography-mass spectrometry approach. *Anal Chim Acta*. 2017;989:59-70.

46. Bahado-Singh RO, Lugade A, Field J, Al-Wahab Z, Han B, Mandal R, et al. Metabolomic prediction of endometrial cancer. *Metabolomics*. 2017;14(1):6.
47. Troisi J, Sarno L, Landolfi A, Scala G, Martinelli P, Venturella R, et al. A metabolomic signature of endometrial cancer. *Journal of proteome research*. 2018;17(2):804-812
48. Cheng SC, Chen K, Chiu CY, Lu KY, Lu HY, Chiang MH, et al. Metabolomic biomarkers in cervicovaginal fluid for detecting endometrial cancer through nuclear magnetic resonance spectroscopy. *Metabolomics*. 2019;15(11):146.
49. Wang X, Zhao X, Zhao J, Yang T, Zhang F, Liu L. Serum metabolite signatures of epithelial ovarian cancer based on targeted metabolomics. *Clin Chim Acta*. 2021;518:59-69.
50. Yin R, Yang T, Su H, Ying L, Liu L, Sun C. Saturated fatty acids as possible important metabolites for epithelial ovarian cancer based on the free and esterified fatty acid profiles determined by GC-MS analysis. *Cancer Biomarkers*. 2016;17(3):259-69.
51. Yang W, Mu T, Jiang J, Sun Q, Hou X, Sun Y, et al. Identification of Potential Biomarkers and Metabolic Profiling of Serum in Ovarian Cancer Patients Using UPLC/Q-TOF MS. *Cell Physiol Biochem*. 2018;51(3):1134-48.
52. Sanchez-Espiridon B, Liang D, Ajani JA, Liang S, Ye Y, Hildebrandt MA, et al. Identification of serum markers of esophageal adenocarcinoma by global and targeted metabolic profiling. *Clinical Gastroenterology and Hepatology*. 2015;13(10):1730-7. e9.
53. Ye W, Lin Y, Bezabeh T, Ma C, Liang J, Zhao J, et al. (1) H NMR-based metabolomics of paired esophageal tumor tissues and serum samples identifies specific serum biomarkers for esophageal cancer. *NMR Biomed*. 2021;34(6):e4505.
54. Wang JL, Zhang T, Shen XT, Liu J, Zhao DL, Sun YW, et al. Serum metabolomics for early diagnosis of esophageal squamous cell carcinoma by UHPLC-QTOF/MS. *Metabolomics*. 2016;12(7):1-10.
55. Zhang S, Lu X, Hu C, Li Y, Yang H, Yan H, et al. Serum Metabolomics for Biomarker Screening of Esophageal Squamous Cell Carcinoma and Esophageal Squamous Dysplasia Using Gas Chromatography-Mass Spectrometry. *ACS Omega*. 2020;5(41):26402-12.
56. Zhu ZJ, Qi Z, Zhang J, Xue WH, Li LF, Shen ZB, et al. Untargeted Metabolomics Analysis of Esophageal Squamous Cell Carcinoma Discovers Dysregulated Metabolic Pathways and Potential Diagnostic Biomarkers. *J Cancer*. 2020;11(13):3944-54.
57. Xu J, Chen YH, Zhang RP, Song YM, Cao JZ, Bi N, et al. Global and Targeted Metabolomics of Esophageal Squamous Cell Carcinoma Discovers Potential Diagnostic and Therapeutic Biomarkers. *Molecular & Cellular Proteomics*. 2013;12(5):1306-18.
58. Chen Z, Gao Y, Huang X, Yao Y, Chen K, Zeng S, et al. Tissue-based metabolomics reveals metabolic biomarkers and potential therapeutic targets for esophageal squamous cell carcinoma. *J Pharm Biomed Anal*. 2021;197:113937.
59. Chen YN, Zhang J, Guo L, Liu L, Wen JR, Xu L, et al. A characteristic biosignature for discrimination of gastric cancer from healthy population by high throughput GC-MS analysis. *Oncotarget*. 2016;7(52):87496-510.
60. Jung J, Jung Y, Bang EJ, Cho SI, Jang YJ, Kwak JM, et al. Noninvasive Diagnosis and Evaluation of Curative Surgery for Gastric Cancer by Using NMR-based Metabolomic Profiling. *Annals of Surgical Oncology*. 2014;21:S736-S42.

61. Huang S, Guo Y, Li Z-W, Shui G, Tian H, Li B-W, et al. Identification and Validation of Plasma Metabolomic Signatures in Precancerous Gastric Lesions That Progress to Cancer. *JAMA Network Open*. 2021;4(6):e2114186.
62. Corona G, Cannizzaro R, Miolo G, Caggiari L, De Zorzi M, Repetto O, et al. Use of Metabolomics as a Complementary Omic Approach to Implement Risk Criteria for First-Degree Relatives of Gastric Cancer Patients. *Int J Mol Sci*. 2018;19(3):750.
63. Zhao H, Heimberger AB, Lu ZM, Wu XF, Hodges TR, Song RD, et al. Metabolomics profiling in plasma samples from glioma patients correlates with tumor phenotypes. *Oncotarget*. 2016;7(15):20486-95.
64. Jonsson P, Antti H, Späth F, Melin B, Björkblom B. Identification of Pre-Diagnostic Metabolic Patterns for Glioma Using Subset Analysis of Matched Repeated Time Points. *Cancers (Basel)*. 2020;12(11).
65. Gao R, Cheng J, Fan C, Shi X, Cao Y, Sun B, et al. Serum metabolomics to identify the liver disease-specific biomarkers for the progression of hepatitis to hepatocellular carcinoma. *Scientific reports*. 2015;5:18175.
66. Kim DJ, Cho EJ, Yu KS, Jang IJ, Yoon JH, Park T, et al. Comprehensive Metabolomic Search for Biomarkers to Differentiate Early Stage Hepatocellular Carcinoma from Cirrhosis. *Cancers (Basel)*. 2019;11(10):1497.
67. Zeng J, Yin PY, Tan YX, Dong LW, Hu CX, Huang Q, et al. Metabolomics Study of Hepatocellular Carcinoma: Discovery and Validation of Serum Potential Biomarkers by Using Capillary Electrophoresis-Mass Spectrometry. *Journal of Proteome Research*. 2014;13(7):3420-31.
68. Lu YH, Li N, Gao L, Xu YJ, Huang C, Yu KK, et al. Acetylcarnitine Is a Candidate Diagnostic and Prognostic Biomarker of Hepatocellular Carcinoma. *Cancer Research*. 2016;76(10):2912-20.
69. Zeng J, Huang X, Zhou LN, Tan YX, Hu CX, Wang XM, et al. Metabolomics Identifies Biomarker Pattern for Early Diagnosis of Hepatocellular Carcinoma: from Diethylnitrosamine Treated Rats to Patients. *Scientific Reports*. 2015;5:16101.
70. Chen T, Zhou K, Sun T, Sang C, Jia W, Xie G. Altered bile acid glycine : taurine ratio in the progression of chronic liver disease. *Journal of Gastroenterology and Hepatology*. 2022;37(1):208-15.
71. Huang Q, Tan YX, Yin PY, Ye GZ, Gao P, Lu X, et al. Metabolic Characterization of Hepatocellular Carcinoma Using Nontargeted Tissue Metabolomics. *Cancer Research*. 2013;73(16):4992-5002.
72. Han J, Han ML, Xing H, Li ZL, Yuan DY, Wu H, et al. Tissue and serum metabolomic phenotyping for diagnosis and prognosis of hepatocellular carcinoma. *Int J Cancer*. 2020;146(6):1741-53.
73. Shao Y, Zhu B, Zheng R, Zhao X, Yin P, Lu X, et al. Development of urinary pseudotargeted LC-MS-based metabolomics method and its application in hepatocellular carcinoma biomarker discovery. *Journal of proteome research*. 2014;14(2):906-16.
74. Sun J, Zhao Y, Qin L, Li K, Sun H, Zhang T, et al. Metabolomic Profiles for HBV Related Hepatocellular Carcinoma Including Alpha-Fetoproteins Positive and Negative Subtypes. *Front Oncol*. 2019;9:1069.

75. Yu L, Li K, Li X, Guan C, Sun T, Zhang X. Metabolomic profiling of dried blood spots reveals gender-specific discriminant models for the diagnosis of small cell lung cancer. *Aging (Albany NY)*. 2020;12(1):978-95.
76. Peng F, Liu Y, He C, Kong Y, Ouyang Q, Xie X, et al. Prediction of platinum-based chemotherapy efficacy in lung cancer based on LC-MS metabolomics approach. *J Pharm Biomed Anal*. 2018;154:95-101.
77. Tian Y, Wang Z, Liu X, Duan J, Feng G, Yin Y, et al. Prediction of Chemotherapeutic Efficacy in Non-Small Cell Lung Cancer by Serum Metabolomic Profiling. *Clin Cancer Res*. 2018;24(9):2100-9.
78. Nie X, Xia L, Gao F, Liu L, Yang Y, Chen Y, et al. Serum Metabolite Biomarkers Predictive of Response to PD-1 Blockade Therapy in Non-Small Cell Lung Cancer. *Front Mol Biosci*. 2021;8:678753.
79. Ye GZ, Liu Y, Yin PY, Zeng ZD, Huang Q, Kong HW, et al. Study of Induction Chemotherapy Efficacy in Oral Squamous Cell Carcinoma Using Pseudotargeted Metabolomics. *Journal of Proteome Research*. 2014;13(4):1994-2004.
80. Lohavanichbutr P, Zhang Y, Wang P, Gu H, Nagana Gowda GA, Djukovic D, et al. Salivary metabolite profiling distinguishes patients with oral cavity squamous cell carcinoma from normal controls. *PLoS One*. 2018;13(9):e0204249.
81. Sahni S, Pandya AR, Hadden WJ, Nahm CB, Maloney S, Cook V, et al. A unique urinary metabolomic signature for the detection of pancreatic ductal adenocarcinoma. *Int J Cancer*. 2021;148(6):1508-18.
82. Hipperson L, Hadden WJ, Nahm CB, Gill AJ, Samra JS, Dona A, et al. Urinary metabolite prognostic biomarker panel for pancreatic ductal adenocarcinomas. *Biochim Biophys Acta Gen Subj*. 2021;1865(11):129966.
83. Fahrman JF, Bantis LE, Capello M, Scelo G, Dennison JB, Patel N, et al. A Plasma-Derived Protein-Metabolite Multiplexed Panel for Early-Stage Pancreatic Cancer. *J Natl Cancer Inst*. 2019;111(4):372-9.
84. Zhang X, Shi X, Lu X, Li Y, Zhan C, Akhtar ML, et al. Novel Metabolomics Serum Biomarkers for Pancreatic Ductal Adenocarcinoma by the Comparison of Pre-, Postoperative and Normal Samples. *Journal of Cancer*. 2020;11(16):4641-51.
85. Lindahl A, Heuchel R, Forshed J, Lehtio J, Lohr M, Nordstrom A. Discrimination of pancreatic cancer and pancreatitis by LC-MS metabolomics. *Metabolomics*. 2017;13(5):61.
86. Mayerle J, Kalthoff H, Reszka R, Kamlage B, Peter E, Schniewind B, et al. Metabolic biomarker signature to differentiate pancreatic ductal adenocarcinoma from chronic pancreatitis. *Gut*. 2018;67(1):128-37.
87. Kobayashi T, Nishiumi S, Ikeda A, Yoshie T, Sakai A, Matsubara A, et al. A Novel Serum Metabolomics-Based Diagnostic Approach to Pancreatic Cancer. *Cancer Epidemiology Biomarkers & Prevention*. 2013;22(4):571-9.
88. Fukutake N, Ueno M, Hiraoka N, Shimada K, Shiraishi K, Saruki N, et al. A Novel Multivariate Index for Pancreatic Cancer Detection Based On the Plasma Free Amino Acid Profile. *Plos One*. 2015;10(7):e0132223.
89. Hirata Y, Kobayashi T, Nishiumi S, Yamanaka K, Nakagawa T, Fujigaki S, et al. Identification of highly sensitive biomarkers that can aid the early detection of pancreatic cancer using GC/MS/MS-based targeted metabolomics. *Clinica Chimica Acta*. 2017;468:98-104.

90. Guo S, Qiu L, Wang YM, Qin XZ, Liu H, He MW, et al. Tissue imaging and serum lipidomic profiling for screening potential biomarkers of thyroid tumors by matrix-assisted laser desorption/ionization-Fourier transform ion cyclotron resonance mass spectrometry. *Analytical and Bioanalytical Chemistry*. 2014;406(18):4357-70.
91. Huang FQ, Li J, Jiang L, Wang FX, Alolga RN, Wang MJ, et al. Serum-plasma matched metabolomics for comprehensive characterization of benign thyroid nodule and papillary thyroid carcinoma. *Int J Cancer*. 2019;144(4):868-76.
92. Lima AR, de Lourdes Bastos M, Carvalho M, de Pinho PG. Biomarker discovery in human prostate cancer: an update in metabolomics studies. *Translational oncology*. 2016;9(4):357-70.
93. Xu H, Chen J, He J, Ji J, Cao Z, Chen X, et al. Serum Metabolic Profiling Identifies a Biomarker Panel for Improvement of Prostate Cancer Diagnosis. *Front Oncol*. 2021;11:666320.
94. Kiebish MA, Cullen J, Mishra P, Ali A, Milliman E, Rodrigues LO, et al. Multi-omic serum biomarkers for prognosis of disease progression in prostate cancer. *J Transl Med*. 2020;18(1):10.
95. Lin X, Lécuyer L, Liu X, Triba MN, Deschasaux-Tanguy M, Demidem A, et al. Plasma Metabolomics for Discovery of Early Metabolic Markers of Prostate Cancer Based on Ultra-High-Performance Liquid Chromatography-High Resolution Mass Spectrometry. *Cancers (Basel)*. 2021;13(13):3140.
96. Li YF, Qiu S, Zhang AH. High-throughput metabolomics to identify metabolites to serve as diagnostic biomarkers of prostate cancer. *Analytical Methods*. 2016;8(16):3284-90.
97. Lima AR, Pinto J, Azevedo AI, Barros-Silva D, Jerónimo C, Henrique R, et al. Identification of a biomarker panel for improvement of prostate cancer diagnosis by volatile metabolic profiling of urine. *Br J Cancer*. 2019;121(10):857-68.
98. Lima AR, Pinto J, Barros-Silva D, Jerónimo C, Henrique R, Bastos ML, et al. New findings on urinary prostate cancer metabolome through combined GC-MS and (1)H NMR analytical platforms. *Metabolomics*. 2020;16(6):70.
99. Gao Q, Su X, Annabi MH, Schreiter BR, Prince T, Ackerman A, et al. Application of Urinary Volatile Organic Compounds (VOCs) for the Diagnosis of Prostate Cancer. *Clin Genitourin Cancer*. 2019;17(3):183-90.
100. Knott ME, Manzi M, Zabalegui N, Salazar MO, Puricelli LI, Monge ME. Metabolic Footprinting of a Clear Cell Renal Cell Carcinoma in Vitro Model for Human Kidney Cancer Detection. *J Proteome Res*. 2018;17(11):3877-88.
101. Liu X, Zhang M, Sun H, Guo Z, Tang X, Wang Z, et al. Urine Metabolomics for Renal Cell Carcinoma (RCC) Prediction: Tryptophan Metabolism as an Important Pathway in RCC. *Front Oncol*. 2019;9:663.
102. Wood PL. Endogenous Anti-Inflammatory Very-Long-Chain Dicarboxylic Acids: Potential Chemopreventive Lipids. *Metabolites*. 2018;8(4).
103. Zhang Q, Xu H, Liu R, Gao P, Yang X, Li P, et al. Highly Sensitive Quantification Method for Amine Submetabolome Based on AQC-Labeled-LC-Tandem-MS and Multiple Statistical Data Mining: A Potential Cancer Screening Approach. *Anal Chem*. 2018;90(20):11941-8.
104. Wang D, Li W, Yin L, Du Y, Zhang S, Suo J. Association of serum levels of deoxyribose 1-phosphate and S-lactoylglutathione with neoadjuvant chemotherapy

sensitivity in patients with gastric cancer: A metabolomics study. *Oncol Lett.* 2020;19(3):2231-42.

105. Huang R, Shen K, He Q, Hu Y, Sun C, Guo C, et al. Metabolic Profiling of Urinary Chiral Amino-Containing Biomarkers for Gastric Cancer Using a Sensitive Chiral Chlorine-Labeled Probe by HPLC-MS/MS. *J Proteome Res.* 2021;20(8):3952-62.

106. Wen C-P, Zhang F, Liang D, Wen C, Gu J, Skinner H, et al. The Ability of Bilirubin in Identifying Smokers with Higher Risk of Lung Cancer: A Large Cohort Study in Conjunction with Global Metabolomic Profiling. *Clinical Cancer Research.* 2015;21(1):193-200.

107. Li YJ, Song X, Zhao XJ, Zou LJ, Xu GW. Serum metabolic profiling study of lung cancer using ultra high performance liquid chromatography/quadrupole time-of-flight mass spectrometry. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences.* 2014;966:147-53.

108. Zhao C, Kong X, Han S, Li X, Wu T, Zhou J, et al. Analysis of differential metabolites in lung cancer patients based on metabolomics and bioinformatics. *Future Oncol.* 2020;16(18):1269-87.

109. Miller HA, Yin X, Smith SA, Hu X, Zhang X, Yan J, et al. Evaluation of disease staging and chemotherapeutic response in non-small cell lung cancer from patient tumor-derived metabolomic data. *Lung Cancer.* 2021;156:20-30.

110. Wikoff WR, Hanash S, DeFelice B, Miyamoto S, Barnett M, Zhao Y, et al. Diacetylspermine is a novel prediagnostic serum biomarker for non-small-cell lung cancer and has additive performance with pro-surfactant protein B. *Journal of clinical oncology.* 2015;33(33):3880-6.

111. Chen Y, Ma Z, Zhong J, Li L, Min L, Xu L, et al. Simultaneous quantification of serum monounsaturated and polyunsaturated phosphatidylcholines as potential biomarkers for diagnosing non-small cell lung cancer. *Sci Rep.* 2018;8(1):7137.

112. Mu Y, Zhou Y, Wang Y, Li W, Zhou L, Lu X, et al. Serum Metabolomics Study of Nonsmoking Female Patients with Non-Small Cell Lung Cancer Using Gas Chromatography-Mass Spectrometry. *J Proteome Res.* 2019;18(5):2175-84.

113. Ruiying C, Zeyun L, Yongliang Y, Zijia Z, Ji Z, Xin T, et al. A comprehensive analysis of metabolomics and transcriptomics in non-small cell lung cancer. *PLoS One.* 2020;15(5):e0232272.

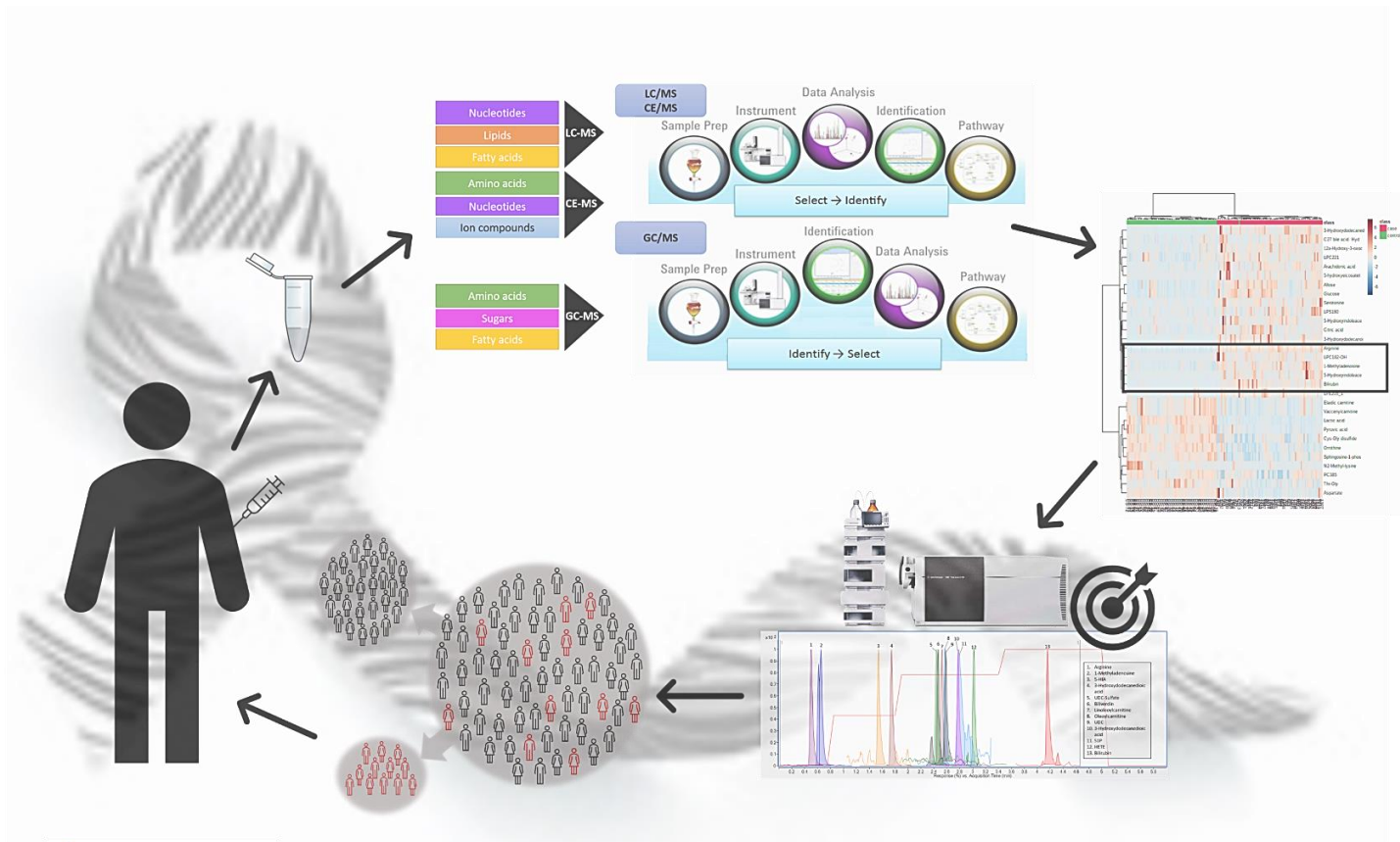
114. Shen J, Ye YQ, Chang DW, Huang MS, Heymach JV, Roth JA, et al. Circulating metabolite profiles to predict overall survival in advanced non-small cell lung cancer patients receiving first-line chemotherapy. *Lung Cancer.* 2017;114:70-8.

115. Zhang L, Zheng J, Ahmed R, Huang G, Reid J, Mandal R, et al. A High-Performing Plasma Metabolite Panel for Early-Stage Lung Cancer Detection. *Cancers (Basel).* 2020;12(3).

116. Xiang C, Jin S, Zhang J, Chen M, Xia Y, Shu Y, et al. Cortisol, cortisone, and 4-methoxyphenylacetic acid as potential plasma biomarkers for early detection of non-small cell lung cancer. *Int J Biol Markers.* 2018;33(3):314-20.

117. Jiao L, Maity S, Coarfa C, Rajapakshe K, Chen L, Jin F, et al. A Prospective Targeted Serum Metabolomics Study of Pancreatic Cancer in Postmenopausal Women. *Cancer Prev Res (Phila).* 2019;12(4):237-46.

118. Wang S, Li M, Yan L, He M, Lin H, Xu Y, et al. Metabolomics study reveals systematic metabolic dysregulation and early detection markers associated with incident pancreatic cancer. *International Journal of Cancer*. 2022;150(7):1091-100.
119. Tao L, Zhou J, Yuan C, Zhang L, Li D, Si D, et al. Metabolomics identifies serum and exosomes metabolite markers of pancreatic cancer. *Metabolomics*. 2019;15(6):86.
120. Luo X, Liu J, Wang H, Lu H. Metabolomics identified new biomarkers for the precise diagnosis of pancreatic cancer and associated tissue metastasis. *Pharmacol Res*. 2020;156:104805.
121. Chen J, Hu Q, Hou H, Wang S, Zhang Y, Luo Y, et al. Metabolite analysis-aided diagnosis of papillary thyroid cancer. *Endocr Relat Cancer*. 2019;26(12):829-41.



Capítulo 1. Perfil metabólico plasmático de pacientes con tumores neuroendocrinos avanzados (NET). Diagnóstico y relevancia biológica

1. Introducción

El metabolismo reprogramado engloba la capacidad de las células para responder o adaptar su señalización metabólica y, así, apoyar y permitir la supervivencia celular en condiciones desfavorables u hostiles. Esta capacidad aumenta en las células cancerosas para mejorar su fenotipo adaptativo y mantener tanto la viabilidad como la proliferación incontrolada. La flexibilidad metabólica es, por tanto, una de las principales características del cáncer (1), aunque quedan por dilucidar las vías implicadas en la maleabilidad metabólica de cada tipo de cáncer. Los metabolitos son los productos finales de esta adaptación, que reflejan los cambios aberrantes en la variabilidad genómica, transcriptómica y proteómica de los tumores y, por tanto, proporcionan información biológica y clínica útil sobre el inicio y la progresión del cáncer (2-4). Por todo esto, actualmente se está intentando comprender los cambios metabólicos asociados específicamente a un tipo de tumor.

Cada vez se encuentran más estudios del metabolismo de los tumores, y se están redescubriendo nuevas rutas metabólicas y nuevas formas de regular las rutas conocidas. El conocimiento de las regulaciones alostéricas de enzimas metabólicas está siendo complementado con el estudio de la regulación de su expresión y de modificaciones post-transcripción regidas por oncogenes y señales extracelulares. Modificaciones como la fosforilación, glicosilación o acetilación de metabolitos se recogen en el término conocido como epimetabolómica y, gracias a la mejora en las técnicas de análisis, se está avanzando en el conocimiento de los flujos metabólicos y, con ello, en el estudio de los cambios en el metabolismo asociado a la formación de tumores.

Por esa razón, aunque aún queda mucho por descubrir en el metabolismo del cáncer, se puede plantear lo siguiente; si no es suficiente con atacar las moléculas causantes del tumor (oncogenes), quizá inhibiendo ciertas rutas metabólicas responsables de mantener ese tumor se pueda controlar o parar su acción devastadora. Esto, junto con el hecho de que la metabolómica se puede realizar fácilmente en muestras biológicas de fácil acceso (plasma, orina), hace que el perfil metabólico de los pacientes con cáncer sea una herramienta prometedora para caracterizar el fenotipo tumoral e identificar nuevos biomarcadores de posible uso clínico. Los enfoques de la medicina de sistemas que integran tecnologías "-ómicas" de alto rendimiento en plataformas de diagnóstico han permitido, de hecho, el análisis detallado de las redes metabólicas (conocidas como metabolómica) en varios cánceres de alta incidencia, prevalencia y mortalidad (5-8), pero estos no incluyen las neoplasias neuroendocrinas

(NEN). La transcriptómica, ciencia “ómica”, perfila la expresión del ARNm; La proteómica identifica proteínas codificadas por estos ARNm; sin embargo, debido a que la actividad de un gen dado puede regularse en un nivel postraduccional, la transcriptómica y la proteómica no son marcadores directos de la actividad de una proteína determinada o de una vía metabólica y/o de señalización celular. Con la metabolómica se puede caracterizar la suma de metabolitos e intermediarios, de bajo peso molecular, definiendo así el estado fisiológico o patológico de un sistema biológico (9). La elaboración de perfiles metabolómicos ofrece resultados ventajosos sobre la transcripción y la elaboración de perfiles de proteínas. Para empezar, como los metabolitos se encuentran aguas abajo de las transcripciones y/o de las proteínas, los metabolitos ofrecen un mejor indicador de la actividad enzimática dentro de una célula, tejido u organismo (10). Además de que, ligeras perturbaciones en los niveles de enzimas dentro de una vía pueden generar grandes cambios en las reservas de metabolitos dentro de esa misma ruta (11, 12). Mientras que las secuencias de transcripciones y proteínas no son idénticas de un organismo a otro, los metabolitos pueden mostrar gran superposición entre especies (azúcares, aminoácidos, intermedios de la ruta de los nucleótidos, etc.). Por lo tanto, las mismas herramientas y técnicas pueden aplicarse a muchos sistemas biológicos (13).

El conocimiento actual define el sistema neuroendocrino como un sistema celular difuso que, junto con las neuronas locales y los ganglios, regulan la homeostasis local y del organismo completo vía aminos y péptidos secretados (14). Las células neuroendocrinas tienen rasgos tanto de células endocrinas, productoras de hormonas, como de células nerviosas. Se encuentran en todos los órganos del cuerpo y ayudan a controlar muchas de las funciones del cuerpo. Desde su descubrimiento inicial, las células neuroendocrinas han sido descritas y estudiadas en múltiples sistemas orgánicos, incluyendo el sistema digestivo, respiratorio y urogenital (15-17). La composición molecular de estas células, así como su contribución a la homeostasis de los órganos, sigue sin estar completamente definida.

Las neoplasias neuroendocrinas (NEN) comprenden una familia heterogénea de tumores raros de incidencia creciente y manejo clínico desafiante que surge de las células neuroendocrinas, e incluyen tumores carcinoides, gastroenteropancreáticos no carcinoides, tumores secretores de catecolaminas, carcinoma medular de tiroides, tumores hipofisarios cromófobos, cáncer de pulmón microcítico y tumores de células de Merkel (18). Aunque pueden surgir en prácticamente cualquier órgano, las localizaciones de tumores primarios más comunes son los pulmones (25%) y el tracto

digestivo (~ 65%). Los tumores neuroendocrinos (NET) bien diferenciados representan aproximadamente el 80% de todas las NEN, tienen un comportamiento clínico bastante indolente, en comparación con sus homólogos exocrinos, y se asocian con un pronóstico de bueno a moderado según esté localizado el tumor primario, el índice de proliferación (Ki67 o índice mitótico) y el estadio tumoral. Aproximadamente el 20% de los NET también tienen la capacidad única de producir y secretar aminas u hormonas peptídicas al torrente sanguíneo, los denominados "tumores funcionales", que producen síndromes endocrinos específicos (es decir, síndrome carcinoide) y que pueden afectar gravemente a la calidad de vida de los pacientes y su pronóstico (19). El síndrome carcinoide es un signo de NET "funcional" en el que se libera serotonina que puede causar diarrea y enrojecimiento facial entre otros síntomas. Los NET no funcionales no liberan sustancias o no las suficientes para causar síntomas. Aproximadamente el 60% de los NET son no-funcionales.

Aunque la supervivencia ha mejorado con el tiempo para todos los NET, probablemente reflejado en un diagnóstico más temprano y mejoras en la terapia (20, 21), a una proporción significativa de pacientes todavía se les diagnostica estadios avanzados de la enfermedad, lo que destaca la necesidad de identificar nuevos biomarcadores específicos que puedan contribuir a una detección más temprana y una mayor probabilidad de curación.

El estudio de los síndromes genéticos hereditarios asociados con una mayor predisposición a desarrollar NET (~ 5%) ha contribuido a dilucidar parcialmente algunos mecanismos implicados en su carcinogénesis (22-27). Las mutaciones de la línea germinal en MEN1, RET, CDKN1B, VHL, NF1 y TSC1/2 son las alteraciones moleculares más frecuentemente detectadas en las NEN hereditarias. Aunque algunas de estas mutaciones tienen una representación relevante en NET esporádicos (es decir, MEN1, TSC1/2), también se han implicado otras alteraciones moleculares implicadas en la regulación epigenética, reparación del ADN, regulación de telómeros y reordenamientos cromosómicos (28-30). Sin embargo, a pesar de estos avances recientes, los mecanismos moleculares de la génesis y la progresión de los NET siguen sin resolverse en gran medida. Además, pocos autores han explorado los NET desde una perspectiva metabolómica. Un estudio piloto prospectivo analizó muestras de orina de 28 pacientes con NET gastroenteropancreáticos mediante espectroscopía de resonancia magnética nuclear (NMR) y mostró distintos fenotipos metabolómicos según el sitio del tumor primario (intestino delgado *versus* NEN pancreático) y su función (31). Un segundo trabajo reciente describió la huella metabolómica de 46 muestras de tejido

NET del intestino delgado analizados por espectroscopía de NMR, sugiriendo la existencia de vías metabólicas complejas en los NET, que posiblemente influyen en el desarrollo y la evolución del tumor y, por lo tanto, en el resultado clínico (32). Con la excepción de estos dos pequeños estudios, el perfil metabolómico de los pacientes con NET no se ha estudiado en profundidad hasta la fecha.

En este contexto, el objetivo del presente estudio fue realizar un perfil metabólico integral de los NET para comprender mejor la desregulación metabólica en estos tumores e identificar nuevos biomarcadores de posible uso clínico. Con este objetivo, se realizaron análisis metabolómicos multiplataforma no dirigidos en muestras de plasma de 77 pacientes con NET gastrointestinales y pulmonares avanzados, y 68 individuos sin cáncer (controles). Se evaluó el potencial diagnóstico y la relevancia biológica de los metabolitos diferenciales identificados, y se exploraron las vías desreguladas para proporcionar más información sobre los mecanismos moleculares involucrados en el desarrollo y la progresión de los NET.

2. Materiales y métodos

2.1. Cohorte de estudio

La población de estudio estaba formada por un total de 145 muestras de plasma, de las cuales, 77 correspondían a pacientes diagnosticados con tumores neuroendocrinos no pancreáticos incluidos en un ensayo clínico doble ciego de fase II-III (AXINET) que aleatorizó a los pacientes para recibir un inhibidor selectivo de los receptores VEGF (*del inglés Vascular Endothelial Growth Factor*, factor de crecimiento endotelial vascular), Axitinib, *versus* placebo, ambos en combinación con octreotida LAR (EUDRACT: 2011-001550-29). Se establecieron una serie de criterios de inclusión que debían cumplir los pacientes: 1) diagnóstico histológico de tumor neuroendocrino no pancreático bien diferenciado; 2) enfermedad no extirpable avanzada o metastásica; 3) tumor de grado 1 o 2 (Ki-67 < 20%); 4) edad > 18 años; 5) función hematológica, renal y hepática adecuada; 6) estado de rendimiento del ECOG (*Eastern Cooperative Oncology Group*) < 1; 7) sin tratamiento anti angiogénico previo. Los pacientes dieron su consentimiento informado antes de ingresar al estudio, incluido un consentimiento específico para estudios traslacionales, y el protocolo fue aprobado por el comité de ética de todos los centros participantes. Las muestras de sangre iniciales fueron recolectadas y se procesaron antes de la administración del tratamiento del estudio. Se recogió sangre periférica en tubos con AEDT sódico, de acuerdo con procedimientos

estándar, y se centrifugó a 1000 x g durante 5 min. Se recuperó la capa de plasma en criotubos estériles, se congeló y se almacenó hasta su uso a -80 °C. Además, se usaron muestras de sangre de 68 individuos no cancerosos como cohorte de control. La distribución del género, la edad y el índice de masa corporal (IMC) fue similar en las cohortes NET y control (Tabla 1).

Tabla 1. Distribución de la edad, del género y el índice de masa corporal (IMC) en cohortes de casos (pacientes con NET) y de control (individuos sin cáncer).

	Pacientes NET		Individuos sin cáncer	
	n = 77 (100%)		n = 68 (100%)	
Sexo	n	%	n	%
Mujer	35	45.5	41	60.3
Hombre	42	54.5	27	39.7
Edad [años]	media	rango	media	rango
	59.1	37-83	61.7	38-83
IMC	media	rango	media	rango
	27.3	17.2- 52.5	26.8	20.7–35.9

IMC: Índice de Masa Corporal

Además, estos datos se complementaron con información extraída de la historia clínica de los pacientes de estudio como es la ingesta de fármacos (Tabla 2).

Tabla 2. Ingesta de fármacos en pacientes NET y sin cáncer.

Tipo de fármaco	Pacientes NET, n (%)	Individuos sin cáncer, n (%)
Alpha-adrenergic Blockers	2 (2.56%)	2 (2.9%)
Analgesics	20 (25.97%)	11 (16.2%)
Anaesthetics	1 (1.28%)	0 (0%)
Anticholinergics	2 (2.56%)	0 (0%)
Anticoagulants/Antiaggregants	12 (15.58%)	2 (2.9%)
Antidiabetics	5 (6.49%)	5 (7.4%)
Antidiarrheals	4 (5.13%)	0 (0%)
Antifungals	1 (1.28%)	0 (0%)
Antigout agents	3 (3.85%)	1 (1.4%)
Antihistaminics	3 (3.85%)	2 (2.9%)
Antihypertensive agents	42 (54.54%)	20 (29.4%)
Antivirals	2 (2.56%)	0 (0%)
Bile Acids Sequestrants	1 (1.28%)	1 (1.4%)
Bronchodilators	7 (8.97%)	4 (5.8%)
Diuretics	14 (17.95%)	4 (5.8%)
H2 Receptor Blockers	8 (10.26%)	0 (0%)
Laxative agents	2 (2.56%)	0 (0%)
Lipid-Lowering Medications	16 (20.77%)	19 (27.9%)
Medicinal Herbs	1 (1.28%)	0 (0%)
Nitrates	1 (1.28%)	0 (0%)
Propulsives	2 (2.56%)	0 (0%)
Proton Pump Inhibitors	15 (19.23%)	9 (13.2%)
Psychoactive drugs	19 (24.67%)	11 (16.2%)
Steroids	4 (5.13%)	0 (0%)
Thyroid agents	6 (7.79%)	6 (8.8%)
Vitamins and dietary supplements	24 (31.16%)	4 (5.8%)

2.2. Análisis multiplataforma de la huella dactilar metabolómica

Con el objetivo de abarcar el espectro completo del metaboloma de estudio, se llevó a cabo el enfoque no dirigido multiplataforma. Para ello, las muestras fueron analizadas a través de diferentes técnicas de separación acopladas, todas ellas, a la espectrometría de masas: electroforesis capilar (CE-MS), cromatografía de líquidos (LC-MS) y cromatografía de gases (GC-MS). Estos procedimientos se realizaron siguiendo protocolos de trabajo estandarizados establecidos por el centro (33-35).

- *Análisis no dirigido de muestras de plasma por CE-MS*

Para el tratamiento de la muestra, a 100 µL de plasma se le añadieron 100 µL de una mezcla acuosa de ácido fórmico en concentración 0.2 M, 5% de acetonitrilo y 0.4 mM de metionina sulfona, usada como patrón interno (IS). La mezcla se agitó en un vortex enérgicamente durante 1 min y, posteriormente, fue transferida a un filtro Centrifree Milipore (30 kDa) y se centrifugó a 2000 x g durante 70 min a 4 °C. La solución filtrada se transfirió directamente al vial para el análisis (33).

El equipo utilizado para el análisis CE-MS fue una electroforesis capilar 7100 (Agilent Technologies, Wilmington, USA) acoplado a un detector de masa exacta con un sistema analizador de tiempo de vuelo TOF (Agilent Technologies, Wilmington, USA). Este acoplamiento cuenta con una fuente de ionización por *electrospray* (ESI) como interfaz.

Inicialmente, un nuevo capilar de sílice fundida de Agilent Technologies (50 μm diámetro interno x 100 cm de longitud total) fue acondicionado con el flujo en tres etapas sucesivas de 30 min cada una de NaOH 1.0 M, agua MiliQ® y electrolito de fondo (BGE, ácido fórmico 1.0 M en MeOH al 10%). Previo a cada análisis el vial BGE fue automáticamente vaciado y llenado. A continuación, el capilar se enjuagó por 5 min (950 mbar) con BGE, aplicando un potencial de 30 kV durante 10 s con el objetivo de desplazar los iones del tampón. La inyección de la muestra se realizó durante 50 s con una presión de 50 mbar. Tras cada inyección de muestras se inyectó BGE durante 20 s a 100 mbar de presión. Las condiciones de separación del análisis incluyeron una presión de 25 mbar y 30 kV de potencial. Los datos de la fuente de ionización por ESI se adquirieron en modo de ion positivo a una velocidad de barrido de exploración de 1.0 espectro/s y con un rango de escaneo de masas de 74 a 1000 m/z . El líquido auxiliar (SL) que consiste en una mezcla 50:50 de metanol/agua y 10 μL de los estándares de masas de referencia (0.25 μM purina, (m/z 121.0509) + 0.25 μM HP-0921 (m/z 922.0098) fue usado con un flujo de 0.6 mL/min (división 1:100). La temperatura del gas de secado se mantuvo a 200 °C con una tasa de flujo de 10 L/min; la presión del nebulizador se fijó a 10 psig, el voltaje a 3500 V, el voltaje del fragmentor a 125 V y el *skimmer* a 65 V (33).

Los iones producto que nos permitieron la confirmación de la anotación tentativa del metabolito fueron obtenidos fragmentando las moléculas en la fuente de ionización (ISF) (36, 37). Las muestras se analizaron nuevamente utilizando las mismas condiciones de análisis salvo el voltaje del fragmentor que fue incrementado hasta los 200 V permitiendo la fragmentación de las moléculas en la fuente ESI. Esto otorgó información adicional muy útil a la hora de elucidar la estructura de las moléculas con la información proporcionada por un analizador de masas sencillo (38).

- *Análisis no dirigido de muestras de plasma por LC-MS*

La precipitación de proteínas y la extracción de metabolitos se realizó mezclando 100 μL de plasma con 300 μL de una mezcla fría (-20 °C) de metanol y etanol (1:1, v/v). La mezcla se agitó en un vortex durante 1 min, se dejó reposar durante 5 min en hielo y

se centrifugó a 16000 x g por 20 min a 4 °C. El sobrenadante se transfirió directamente a un vial de LC para su análisis.

El análisis LC-MS se realizó utilizando un sistema UHPLC serie 1290 Infinity II (Agilent Technologies, Waldbronn, Germany) acoplado a un analizador cuadrupolo-tiempo de vuelo, QTOF/MS (6545), adquiriendo en modo de ion positivo y negativo, respectivamente. El análisis se llevó a cabo como ha sido descrito anteriormente (34). Para la separación de los metabolitos, se inyectó 0.5 µL de muestra en una columna Zorbax Extended-C18 Rapid Resolution (Agilent Technologies, 2.1 mm x 50 mm, 1.8 µm) manteniendo la temperatura a 60 °C. El flujo se fijó en 0.6 mL/min con una fase móvil compuesta por (A) 0.1% de ácido fórmico en agua y (B) 0.1% de ácido fórmico en acetonitrilo. El gradiente de elución cromatográfico comenzó con un 5% de B durante el primer minuto, incrementado hasta el 80% de B en el minuto 6.0. Posteriormente se alcanzó el 100% de B en el minuto 11.5 volviendo a las condiciones iniciales en 0.5 min y permitiendo el reequilibrio hasta el minuto 15.0 (34).

Los datos de la fuente de ionización ESI se adquirieron en modo de ion positivo y negativo, en análisis independientes, con el escaneo completo de masas en la detección por MS en el rango de m/z de 100 a 1000, y de m/z de 40 a 1000 para el análisis MS y el MS/MS, respectivamente.

El voltaje del capilar se estableció en 3000 V con una velocidad de barrido de exploración de 1.5 espectros/s para ESI(+) y 4000 V con 1.0 espectros/s para ESI(-). El caudal del gas de secado fue de 12 L/min a 250 °C y la presión del nebulizador de gas 52.0 psi. Para mejorar la ionización se aplicó un calentamiento adicional usando gas auxiliar calentado hasta los 370 °C con un flujo de 11 L/min.

El análisis de MS/MS fue llevado a cabo manteniendo las mismas condiciones cromatográficas y espectrométricas empleadas en el análisis inicial. De acuerdo con la masa exacta y el tiempo de retención determinados *a priori*, los iones de interés fueron seleccionados según su m/z (iones precursores) y fragmentados por el método de disociación inducida por colisión (CID) con una ventana de error estrecha (aprox. 1.3 Da).

Dos masas de referencia fueron analizadas continuamente en el sistema durante todo el tiempo de análisis para permitir la corrección constante de la masa. Estos iones de referencia fueron la purina protonada ($m/z = 121.0509$) y HP-0921 ($m/z = 922.0097$) para ESI(+), mientras que para ESI(-) fueron el anión de TFA con la pérdida de un protón ($m/z = 112.9856$) y el aducto de formiato del HP-0921 ($m/z = 966.0007$).

- *Análisis no dirigido de muestras de plasma por GC-MS*

Para el tratamiento de muestra de esta plataforma se procedió inicialmente con la precipitación proteica a través del tratamiento con acetonitrilo frío en proporción (1:3), agitado-mezclado durante 2 min y se dejaron reposar en hielo durante 5 min. Las muestras fueron centrifugadas ($15400 \times g$, 10 min, 4 °C) lo que permitió aislar la fracción de metabolitos. Del sobrenadante resultante, 100 μL fueron transferidos a un vial de GC y evaporados completamente a sequedad (concentrador SpeedVac, Thermo Fisher Scientific, Waltham, MA, EE. UU.) para añadir a continuación a cada vial 10 μL de clorhidrato de O-metoxiamina en piridina (15 mg/mL) y agitar vigorosamente durante 1 min. La reacción de metoximación se llevó a cabo a temperatura ambiente bajo oscuridad durante 16 h. Posteriormente se añadió 10 μL de N, O-bis(trimetilsilil)trifluoroacetamida (BSTFA) con 1% de trimetilclorosilano (TMCS) como catalizador adicional y se mezclaron utilizando vortex durante 5 min. Los viales encapsulados fueron calentados en un horno a 70 °C durante 1 h para completar la siliación. Finalmente, se agregaron 100 μL de heptano que contenía 10 ppm del ester metílico C18:0 utilizado como patrón interno (IS) y se mezclaron en un vortex antes de su análisis (35).

El análisis de GC-MS se realizó en un sistema GC (Agilent Technologies 7890A) acoplado a un espectrómetro de masas de cuadrupolo sencillo con un detector de triple eje (5975C, Agilent Technologies). Dos microlitros (2 μL) de la muestra de plasma derivatizada fueron automáticamente inyectados a través del inyector (Agilent 7693) en modo *split* (radio de división 1:10) en una columna GC DB5-MS (30 m de longitud, 0.25 mm de diámetro interno, 0.25 μm de película de 95% dimetil / 5% difenilpolisiloxano) con una precolumna integrada (10 m J&W, Agilent). La división del flujo se realizó en el interior de una guía de lana de vidrio desactivada (Restek 20782, Bellefonte, PA, EE. UU). El flujo del gas portador (He) se fijó tras realizar el bloqueo en el tiempo de retención (*Retention Time Locking*, RTL) a aproximadamente 1 mL/min, con una temperatura del inyector y de la línea de transferencia de 250 °C y 280 °C, respectivamente. El gradiente de temperatura se programó de la siguiente manera: la temperatura inicial del horno fue fijada a 60 °C y se mantuvo durante 1 min, aumentada hasta los 325 °C a razón de 10 °C/min y mantenida a esta temperatura durante 10 min antes de comenzar el enfriado para la próxima inyección. La detección de MS se realizó en el modo de ionización de impacto electrónico (EI) con una energía de -70 eV y 230 °C en la fuente de filamento. Los espectros de masas se adquirieron en el rango de m/z 50 – 600 a una velocidad de 2 espectros/s. Al inicio del análisis se hicieron varias

inyecciones de IS, una mezcla estándar de ésteres metílicos de ácidos grasos (FAME C8-C30) y una mezcla de alcanos (C8-C28) disuelta en hexano para la determinación del índice de retención (RI). Para la adquisición de los datos se utilizó Agilent MSD ChemStation Software (Agilent Technologies).

El análisis para cada plataforma empleada se llevó a cabo aleatorizando el orden de inyección de las muestras. También fueron incluidas varias réplicas de una mezcla homogénea que contiene una pequeña cantidad de todas las muestras (controles de calidad, QC). Estos QCs siguieron el mismo procedimiento aplicado para las muestras experimentales según la plataforma de análisis y fueron analizados en un conjunto inicial de 10 inyecciones para equilibrar el sistema y, cada 10 muestras de estudio proporcionando una medida de la estabilidad del análisis, y del rendimiento y la reproducibilidad del procedimiento del tratamiento de muestra (39). En el caso de CE-MS y LC-MS, se utilizó una única muestra de QC, mientras que para GC-MS se prepararon muestras de QC independientes.

2.3 Procesado de los datos

Los datos adquiridos se procesaron según la plataforma empleada para el análisis. Así, los datos obtenidos a través de CE-MS fueron procesados para obtener la estructura y el formato apropiado utilizando el software MassHunter Profinder (B.08.00, Agilent Technologies). Inicialmente se llevó a cabo el análisis de extracción de características moleculares (MFE) para limpiar el ruido de fondo y los iones no relacionados. Este algoritmo permitió encontrar los aductos coeluyentes (+H⁺, +Na⁺ y la pérdida neutra de agua) de cada entidad química con el alineamiento de las muestras usando la masa y el tiempo de retención creando el espectro conjunto final para esa entidad química. En el siguiente paso se procedió con la reextracción de las características moleculares por conjunto completo (RFE) a través de la búsqueda por ion (FBI) con una lista final de compuestos de mejor calidad. Los compuestos encontrados en las diferentes muestras se alinearon utilizando una ventana de tiempo de retención de 10% ± 2 min y una ventana de masa de 20 ppm ± 2.0 mDa. La altura mínima se fijó en 600 cuentas con el filtrado de las características presentes en al menos el 75% de las muestras de al menos un grupo. Esto generó una lista final de grupos de compuestos que fue inspeccionada minuciosamente en todas las muestras para verificar la correcta integración de los iones extraídos.

Los datos procedentes del análisis por LC-MS se reprocesaron por medio del software MassHunter Qualitative (B.06.00, Agilent Software) a través del algoritmo de

extracción de características moleculares (MFE) creando una lista final de posibles componentes. El reprocesado de datos fuera de línea por DA Reprocessor Offline Utilities B.05.00 (Agilent) permitió encontrar los aductos coeluyentes de una misma característica, iones como $[M]^+$, $[M+H]^+$ y $[M+Na]^+$ en modo de ionización positiva; $[M-H]^-$, $[M+HCOOH-H]^-$, $[M+Cl]^-$ en modo negativo y, pérdida neutra de agua en ambas polaridades. Por último, utilizando el software Mass Profiler Professional (B.14.9 Agilent Software) los datos se alinearon y filtraron. Los compuestos encontrados en las diferentes muestras se alinearon utilizando una ventana de tiempo de retención de $0.5\% \pm 0.15$ min y una ventana de masa de 15 ppm ± 2.0 mDa y se mantuvieron los compuestos comunes encontrados en al menos el 75% de las muestras QCs.

El tratamiento de datos de GC-MS comenzó con la inspección exhaustiva de todos los cromatogramas de iones totales (TIC) obtenidos durante el análisis completo para examinar la calidad general del rendimiento analítico. Los datos sin procesar se convirtieron al formato apropiado para el análisis cuantitativo utilizando el software MassHunter Workstation GC-MS Translator (B.04.01). Posteriormente, se empleó Agilent MassHunter Unknowns Analysis Tool B.07.00 para la deconvolución espectral de masas automática y la identificación de los compuestos que coeluyen de acuerdo con sus índices (RI) y tiempos de retención (RT). Estos tiempos de retención de las muestras se obtuvieron del ajuste realizado con los datos de calibración procedentes del análisis de la solución estándar FAME C8-C30. La identificación se realizó mediante la búsqueda en dos bibliotecas específicas: Fiehn, versión 2008, y la biblioteca interna de espectros en muestras de plasma del CEMBio basada en las librerías de Fiehn y NIST (*National Institute of Standards and Technology*, versión 2014), usando los propios estándares. Para lograr la asignación de esta identidad se compararon sus índices de retención y tiempos de retención, además del espectro experimental extraído tras la deconvolución con el teórico incluido en las bibliotecas. Los compuestos a los que no fue asignada una identidad se intentaron identificar comparando sus patrones de fragmentación de masa con la biblioteca NIST 2014 a través del software MassHunter Qualitative (B.06.00, Agilent Software).

Tras esto, los datos se alinearon con MassProfiler Professional (B.14.9 Agilent Software) y se exportaron a Agilent MassHunter Quantitative Analysis (B.07.00) para la asignación de los iones objetivo y, obtener la abundancia de los compuestos tras la inspección de la integración correcta de los picos. Finalmente se generó la matriz de datos con la abundancia final de cada metabolito.

Para garantizar mediciones válidas, las matrices de datos obtenidas tras el reprocesamiento para cada una de las plataformas se filtraron de acuerdo con la variación en la abundancia de los QCs, expresada como desviación estándar relativa (RSD), y solo se mantuvieron aquellos con $RSD < 30\%$.

2.4 Análisis estadístico

Para el análisis estadístico de los datos, la sustitución de los valores que faltan en las matrices de CE-MS y LC-MS fue lo primero que se realizó. Durante el reprocesado de los datos se generan valores perdidos debido a los algoritmos MFE y Fbl y, para reemplazarlos, se utilizó una función matemática llamada “vecinos más cercanos” o kNN (*del inglés, k-nearest neighbors*). Este algoritmo permite estimar el valor perdido comparando la similitud del perfil con las muestras cercanas o vecinos del grupo al que pertenecen (40).

Inicialmente para todas las plataformas, con el análisis multivariante no supervisado de los datos (PCA) se observaron tendencias basadas en el orden de inyección que fueron corregidas con el algoritmo QC-SVRC (*Quality Control and Support Vector Regression Correction*) (41). Este algoritmo corrige la deriva instrumental ocurrida durante el análisis de una secuencia de muestras utilizando los datos de los QCs. Estas variaciones en la respuesta instrumental son inevitables en el análisis de secuencias de muestras de gran tamaño y su ajuste requiere de estrategias de corrección de la señal. Los tres hiperparámetros que determinan la precisión de la eliminación de los efectos de secuencia fueron: i) el umbral de tolerancia (ϵ) 5% de la mediana de distribución de los valores de los QC; ii) el término de penalización (C), mediana de la distribución de muestras QC y iii) el ancho de kernel (γ), logspace (0.3), 20 valores. Tras la corrección de la deriva, los datos se sometieron al procedimiento de garantía de calidad (QA), estudio de la distribución normal de los datos por las pruebas de Kolmogórov-Smirnov-Lilliefors y Shapiro-Wilk y la prueba de Levene para la relación de varianzas, para cada plataforma. Para determinar la significación estadística de cada metabolito por separado, se evaluaron las diferencias entre los individuos NET y del grupo no-cáncer aplicando la prueba t de Student (valor de $p \leq 0.05$) utilizando MATLAB (R2015a, MathWorks). Además, se aplicó el método de corrección múltiple de Benjamini-Hochberg a todos los valores de p para controlar la tasa de falsos positivos en el nivel $\alpha = 0.05$.

El análisis multivariante se realizó mediante el software SIMCA 15.0 (Sartorius Stedim Biotech) que permite la visualización de la matriz de datos completa y,

comprobar el rendimiento y la calidad de los datos obtenidos a partir de cada plataforma. El método no supervisado, PCA, se utilizó para observar el agrupamiento espontáneo de las muestras y la detección de muestras atípicas, “*outliers*”. Métodos supervisados como son el análisis discriminante por mínimos cuadrados parciales (PLS-DA) y el ortogonal por mínimos cuadrados parciales (OPLS-DA) fueron construidos usando este software. Tras obtener el modelo OPLS-DA se construyeron los intervalos de confianza del diagrama S (*S-plot*), herramienta útil para seleccionar grupos de metabolitos diferenciadores. De igual manera, los diagramas Volcano (*Volcano plots*) permiten su aplicación para la selección de variables con la representación combinada de la importancia de la variable en la proyección (VIP) y el coeficiente de correlación, $\rho(\text{corr})$ (42).

La calidad de los modelos construidos viene proporcionada por los valores de R^2 (bondad de explicación del modelo) y Q^2 (bondad de predicción del modelo) que proporciona el software. Los modelos OPLS-DA para cada plataforma (CE-MS, LC-MS/ESI(+), LC-MS/ESI(-) y GC-MS) fueron validados mediante 999 permutaciones con la respuesta de la prueba de permutaciones. Esta prueba evalúa si la clasificación específica de los individuos en los dos grupos diseñados es significativamente mejor que cualquiera otra clasificación aleatoria en dos grupos arbitrarios, comprobando la validez y el grado de sobreajuste de los modelos. Cuando los datos del “eje y” se permutan aleatoriamente, se compara la distribución de R^2 y Q^2 con los valores obtenidos de los datos originales (no perturbados), al mismo tiempo que los datos de “x” permanecen intactos. El gráfico de validación resultante representa en el “eje y” los valores de R^2 y Q^2 de los modelos originales y permutados mientras que el “eje x” representa los valores de los coeficientes de correlación entre ambos modelos. Los puntos se ajustaron a través de una recta de regresión y, la validez del modelo implica la distribución de R^2 y Q^2 en el modelo original con valores más altos que los resultantes de los modelos permutados (43).

Por otra parte, calculando el porcentaje de cambio se inspeccionó la distribución de cada variable entre los grupos. Esta variación se halló calculando el promedio de cada compuesto por grupo de muestras de la siguiente forma: (Promedio del grupo de casos – Promedio del grupo control) / Promedio del grupo control, en %.

2.5 Anotación e identificación de los metabolitos

En el caso de los resultados para CE-MS y LC-MS, se comenzó con una anotación tentativa de las señales mediante la herramienta CEU Mass Mediator (44),

basada en el m/z de los compuestos que presentaban diferencias significativas en la separación por grupos. Esta anotación incluyó, además de la coincidencia de masa precisa con el error de masa establecido en 10 ppm para LC-MS y 20 ppm para CE-MS, la determinación de la distribución isotópica y la verificación manual de los posibles iones y aductos. Para CE-MS, se complementó la anotación comprobando el tiempo de migración relativo, relación entre el tiempo de retención del metabolito y el del IS, y los fragmentos, dímeros o artefactos del conjunto de datos fueron eliminados (36).

Para confirmar esas anotaciones se utilizó la información proporcionada de espectros de fragmentación en fuente y de disociación inducida por colisión (espectrometría de masas en tándem), para CE-MS y LC-MS respectivamente.

Los espectros obtenidos con un elevado valor de voltaje del fragmentor (200 V) fueron estudiados para identificar patrones característicos de fragmentación de la fuente de iones en CE-MS, con la ayuda de la base de datos interna creada por el CEMBIO. Esta base de datos incluye 515 fragmentos junto con sus iones, aductos y multímeros, disponible en CEU Mass Mediator (<http://ceumass.eps.uspceu.es/>) (44, 45), además de una biblioteca de tiempos de migración relativa.

El análisis LC-MS/MS se realizó repitiendo el experimento con las mismas condiciones cromatográficas y espectrométricas, dirigiendo la fragmentación con los valores de m/z y de RT del compuesto de interés que se desea fragmentar para su identificación (ion precursor) y, usando una ventana de aislamiento estrecha en el cuadrupolo, Q (aprox. 1.3 Da). Se aplicaron diferentes energías de colisión fija (20 y 40 eV) a los iones precursores en análisis independientes para obtener espectros específicos de MS/MS. Con estos resultados se procedió a la identificación de cada metabolito con la interpretación manual de los espectros de fragmentación. La comparación de los patrones de fragmentación en base al mecanismo de fragmentación característico y las diferencias de masa, se realizó, cuando fue posible, con la información disponible de espectros MS/MS de estándares analizados por el grupo o, a través de la correspondencia espectral con los espectros disponibles en la base de datos de Metlin (<https://metlin.scripps.edu/>) (46) y la fragmentación *in silico* para la identificación asistida en MetFrag (<https://msbi.ipb-halle.de/MetFrag/>) (47).

2.6 Interpretación biológica

Finalmente, con los datos de los metabolitos identificados con un nivel 2 de confianza se estudió la interpretación funcional de estos resultados mediante la plataforma online MetaboAnalyst 4.0 (<https://www.metaboanalyst.ca/>). Esta herramienta

expone a través de distintos análisis y gráficos, los cambios más representativos de la distinción entre grupos o las posibles rutas metabólicas implicadas, entre otros (48). Aunque esto, no siempre es suficiente y, se debe recurrir a la literatura publicada para poder dar una explicación a los cambios observados y a la relación de esas rutas alteradas con la patología.

Con este propósito se construyó un “*heatmap*”, análisis que convierte una matriz de datos numérica en el correspondiente mapa de color en 2D proporcionando un resumen intuitivo de los datos obtenidos. Asimismo, se realizó el análisis de las rutas metabólicas, que proporcionó una visión del metaboloma destacando que vías metabólicas se encuentran significativamente modificadas o, que metabolitos tienen mayor impacto en la ruta.

Los metabolitos discriminantes se clasificaron por clases bioquímicas y por su relación con vías moleculares específicas (apoptosis, ciclo celular, angiogénesis, vía mTOR, efecto Warburg, estrés oxidativo, metabolismo del triptófano, metabolismo del colágeno, metabolismo de la carnitina, ciclo de la metionina, metabolismo del ácido araquidónico, ciclo de la urea, poliaminas y metabolismo del grupo hemo). Con el fin de refinar la identificación de las vías moleculares aberrantes en pacientes NET, se analizaron los datos mediante el análisis de la ruta del metabolito (MPA) y el análisis de enriquecimiento del conjunto de metabolitos (MSEA) utilizando la plataforma MetaboAnalyst 4.0. Las bases de datos de referencia empleadas fueron KEGG homo sapiens (<https://www.genome.jp/kegg/>)(octubre de 2019) y SMPD (<https://smpdb.ca/>) (49, 50).

2.7 Análisis de los datos clínicos y moleculares

Para evaluar el potencial diagnóstico de los metabolitos, se construyeron y estudiaron las curvas de las características operativas del receptor (ROC) y el área bajo la curva (AUC), y se calcularon los valores de sensibilidad y especificidad (según el índice de Youden) (51). Las asociaciones de las características clínicas relevantes (edad, sexo, IMC, glucemia y niveles plasmáticos de creatinina) y los medicamentos concomitantes comunes seleccionados por su supuesta influencia en la metabolómica (Tabla 2) se evaluaron para obtener los metabolitos con un valor de AUC > 0.85, utilizando la prueba exacta de Fisher, prueba chi-cuadrado o de correlación de Pearson, según corresponda ($p < 0.05$ se consideraron significativos). A continuación, se construyeron los modelos de regresión logística (LRM) para cada metabolito ajustando por edad, sexo, glucemia, niveles plasmáticos de creatinina y medicación significativa

del paciente, seleccionados a partir del análisis de asociación. Los metabolitos con AUC > 0.85 se consideraron como posibles biomarcadores. Además, metabolitos diferenciales de los modelos OPLS-DA con VIP > 1.0 y $|p(\text{corr})| > 0.5$ también se consideraron como posibles biomarcadores.

Adicionalmente, se realizaron *heatmaps* con el valor log₁₀ de los niveles de cada metabolito en muestras de plasma. Se llevó a cabo una agrupación jerárquica no supervisada para metabolitos y muestras utilizando la correlación de Pearson y el promedio como método de vinculación. Ambos se realizaron utilizando el software Morpheus (<https://software.broadinstitute.org/morpheus/>) (Broad Institute, consultado el 18 de diciembre de 2020).

2.8 Análisis dirigido para la validación analítica de los biomarcadores candidatos

Con el objetivo de confirmar la identidad y cuantificar los potenciales biomarcadores, se realizó un análisis dirigido mediante LC-QqQ/MS.

- *Preparación de estándares*

Los estándares comerciales utilizados se resumen en la Tabla 3. Cada compuesto se pesó independientemente y el polvo correspondiente se disolvió en el volumen requerido del disolvente elegido para dar una solución madre de 100 µg/ml que se almacenó a -20 °C. Las soluciones de trabajo se prepararon diluyendo las soluciones madre de cada compuesto a 10 µg/ml en el disolvente correspondiente. Además, se prepararon dos mezclas de compuestos, de concentración 10 µg/mL de cada uno de ellos, según su solubilidad en solvente acuoso o en solvente orgánico (MeOH/EtOH (1:1)). Para preparar la serie de diluciones de calibración, se diluyó la mezcla acuosa que contenía arginina, metiladenosina, biliverdina y bilirrubina a 6 niveles de concentración diferentes (desde 1 ng/ml a 1 µg/ml). De la misma forma, se utilizó la mezcla orgánica que contenía ácido 5-hidroxiindolacético, ácido 3-hidroxidodecanodioico, linoleoilcarnitina, oleoilcarnitina, esfingosina-1-fosfato, ácido 3-hidroxidodecanoico, HETE, ácido ursodesoxicólico y ácido ursodesoxicólico 3-sulfato para la preparación de las diluciones secuenciales (desde 1 ng/ml a 1 µg/ml).

Tabla 3. Estándares con su pureza, la marca y la casa comercial correspondiente que lo distribuye utilizados para el análisis dirigido de posibles biomarcadores en NET.

Compuesto	Marca	Casa comercial	Riqueza
Arginine	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>98%
Biliverdin	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	≥97%
Linoleoylcarnitine	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	≥95%
5-Hydroxyindoleacetic acid	Acros Organics	Fisher Scientific One Reagent Lane Fair Lawn, NJ 07410	≥99%
3-Hydroxydodecanoic acid	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	≥99%
Bilirubin	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	≥95%
1-Methyladenosine	Cayman Chemical Company	Cayman Chemical Company 1180 E. Ellsworth Rd. Ann Arbor, MI 48108	≥98%
15-HETE	Avanti Polar Lipids, INC	700 Industrial Park Drive Alabaster, Al 35007 United States of America	≥99%
Sphingosine-1-phosphate	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	≥98%
3-Hydroxydodecanedioic acid	Toronto Research Chemicals	Toronto Research Chemicals 2 Brisbane Road Toronto, ON M3J 2J8 CANADA	93%
Ursodeoxycholic acid 3-sulfate	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	≥98%
Oleoylcarnitine	Larodan AB	Larodan AB Karolinska Institutet Science Park Retzius väg 8 SE-171 65 SOLNA Sweden	≥97%
Ursodeoxycholic acid	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	≥99%

- *Preparación de la muestra*

Las muestras de plasma de 76 casos y 68 controles se prepararon utilizando el mismo protocolo utilizado para el análisis no dirigido mediante LC-MS (34).

- *Optimización del método HPLC QqQ-MS/MS*

El análisis de validación se realizó en un sistema Agilent 1290 Infinity UHPLC (Agilent Technologies, Waldbronn, Alemania) junto con un espectrómetro de masas de triple cuadrupolo QqQ Agilent 6460 con una fuente de ionización por *electrospray* (ESI). En el método final, la separación cromatográfica de los compuestos se logró con una columna Zorbax C8 Eclipse Plus (Agilent Technologies, 2.1 × 150 mm, 1.8 μm) termostaticada a 55 °C.

Para garantizar el mejor rendimiento del sistema LC-MS para metabolitos individuales, se llevó a cabo la optimización con respecto a las condiciones de LC y los parámetros relacionados con MS. Las condiciones dependientes de LC se optimizaron probando múltiples gradientes de elución de fase móvil, velocidad de flujo y adición de aditivos a la fase móvil orgánica.

Para los analitos individuales, los parámetros de MS fueron ajustados por Agilent MassHunter Optimizer (versión de software B.07.00), además de por la información recogida de la revisión bibliográfica, utilizando estándares individuales como referencia. MassHunter Optimizer optimizó automáticamente los parámetros de adquisición de datos para el modo MRM (monitoreo de reacciones múltiples). Se automatizó la selección de los mejores iones precursores, la optimización del voltaje del fragmento para cada ion precursor, la selección de los mejores iones producto y la optimización de los valores de energía de colisión de cada transición para una lista de compuestos especificados.

Siempre que fue posible, se identificaron dos de las transiciones iónicas precursor-producto más intensas, la transición con mayor respuesta se seleccionó como ion cuantificador, mientras que la otra fue seleccionada como ion cualificador.

- *Condiciones de análisis LC-MS optimizadas e identificación de compuestos*

Tras la optimización completa del sistema LC-MS, la fase móvil binaria seleccionada consistió en la fase A (0.1% ácido fórmico en agua) y la fase B (0.1% ácido fórmico y 10% iPrOH en ACN). El flujo utilizado se ajustó a 0.6 mL/min y el volumen de inyección fue de 1.00 μL para todos los estándares y muestras estudiadas. El gradiente

de LC se ajustó finalmente para obtener la separación cromatográfica más eficaz reduciendo el tiempo de ejecución total.

El gradiente cromatográfico comenzó con un 3% de fase B, aumentando hasta un 4% de B en el minuto 0.75. A continuación, el gradiente creció hasta un 45% de B en el minuto 0.85 y se mantuvo hasta el minuto 1.80. Se aplicó otro incremento al 79% B en el minuto 1.90, manteniendo este porcentaje hasta el minuto 3.40. En 0.10 minutos, subió a 100% B y se mantuvo hasta el minuto 5.00. Finalmente, se volvió a la condición de partida en 0.10 min permitiendo el reequilibrio hasta el minuto 6.00. La adquisición de datos espectrales de masas se realizó en polaridad positiva (ESI+) y negativa (ESI-) según la transición elegida para cada metabolito, y el modo de monitoreo dinámico de reacciones múltiples (dMRM). El voltaje del capilar se fijó en 4000 V para el modo de ionización positivo y 3500 V para el modo de ionización negativo; el caudal de gas de secado fue de 8 L/min a 200 °C, la presión del gas del nebulizador a 50 psi, el flujo de gas auxiliar en 8 L/min y la temperatura a 250 °C.

El sistema se optimizó en modo MRM dinámico; el ion precursor de cada compuesto se aisló mediante el primer cuadrupolo, se fragmentó en la celda de colisión, y el tercer cuadrupolo seleccionó los fragmentos más característicos. Se monitorizaron las transiciones específicas para cada analito ajustando la energía de colisión, el voltaje del fragmento, el voltaje del acelerador celular y el delta de tiempo de retención durante 1 min con un tiempo de permanencia de 100 ms para todos ellos.

La identificación de los metabolitos se logró comparando sus transiciones de pares de iones precursor-producto MRM (tanto cuantificador como cualificador) y los tiempos de retención con los de los estándares individuales analizados. La cuantificación se llevó a cabo con curvas de calibración establecidas utilizando el área del pico del analito de los iones cuantificadores.

- *Validación y calibración de métodos*

La secuencia analítica consistió en el conjunto de calibración de ambas mezclas estándar (acuosa y orgánica) al comienzo y al final de la secuencia. Durante el experimento también se inyectaron diferentes concentraciones de la mezcla estándar para monitorizar el análisis. Las muestras biológicas fueron analizadas aleatoriamente. Las mezclas de calibración se prepararon con 6 niveles de concentración diferentes con un amplio rango debido a los distintos niveles de concentración de analitos individuales en las diferentes muestras biológicas. Las curvas de calibración se construyeron mediante regresión lineal y los orígenes no se forzaron a pasar por cero, siempre que

fue posible. Para evaluar el ajuste de las curvas, se calcularon los coeficientes de correlación (r).

- *Tratamiento de datos y análisis estadístico*

El control del sistema durante el análisis y la revisión inicial del cromatograma se realizó con el software Agilent MassHunter Qualitative (versión B.08.00). El reprocesado de datos se llevó a cabo utilizando Agilent MassHunter Quantitative QqQ (versión B.09.00). Los metabolitos fueron semicuantificados según el factor de respuesta de la curva de calibración respectiva. La normalidad de los datos de la matriz generada se probó mediante las pruebas de Kolmogórov-Smirnov-Lilliefors y Shapiro-Wilk. Tras el procesamiento de los datos se realizó un análisis estadístico para evaluar las diferencias en la comparación de casos NET frente a individuos sin cáncer mediante la aplicación de la prueba t de Student (valor de $p \leq 0.05$) utilizando MATLAB (R2015a, MathWorks).

3. Resultados

En lo referente a las muestras empleadas en el estudio se recogieron datos de los pacientes procedentes tanto de sus historias clínicas como variables histopatológicas determinadas en el grupo con cáncer (Tabla 4). Si nos fijamos en la localización del tumor, el intestino delgado fue la categoría predominante y el 66% de los pacientes presentaron unos niveles plasmáticos de cromogranina A dos veces mayores al límite superior de normalidad (ULN). En torno al 30% de los NETs fueron clasificados como “tumores funcionantes” los cuales producen síndrome carcinoide y un 66% presentó grado histopatológico 2. Como se comentó anteriormente, las células neuroendocrinas han sido descritas y estudiadas en múltiples sistemas orgánicos, pero presentan mayor frecuencia en órganos como el tracto gastrointestinal (~65%) y los pulmones (~25%). La habilidad de una parte de este tipo de tumores para producir y secretar aminas u hormonas peptídicas provoca el síndrome carcinoide, pudiendo perjudicar gravemente en la calidad de vida y pronóstico de los pacientes. Cuando el tumor es clasificado como no funcionante, sin sintomatología de tipo productor, los niveles plasmáticos de CgA son útiles para el diagnóstico y seguimiento de tumores neuroendocrinos gastroenteropancreáticos (páncreas y tubo digestivo). La clasificación propuesta según el grado histopatológico se basa en las tasas de proliferación y presenta rangos numéricos específicos del recuento mitótico y el índice de proliferación Ki67 (52).

Tabla 4. Características clínicas, bioquímicas y patológicas de la población con NET (n=77).

Síndrome carcinoide (tumor funcionante)	
<i>positivo</i>	24 (31%)
<i>negativo</i>	51 (66%)
<i>no dato</i>	2 (3%)
Localización tumor 1ario	
<i>Colorectal</i>	8 (10%)
<i>Intestino delgado</i>	43 (56%)
<i>Pulmón</i>	18 (24%)
<i>Otros</i>	7 (9%)
<i>no dato</i>	1 (1%)
Grado histopatológico	
<i>G1</i>	26 (34%)
<i>G2</i>	51 (66%)
Cromogranina A basal valor > 2 xULN	
<i>positivo</i>	51 (66%)
<i>negativo</i>	26 (34%)
5-Hidroxi-indolacético (orina basal) valor > 2xULN	
<i>positivo</i>	39 (51%)
<i>negativo</i>	38 (49%)

ULN: del inglés *Upper Limit of Normal*, límite superior normal; G: grado

Con el fin de asegurar el cumplimiento de la hipótesis en un estudio metabólico, donde los niveles de los metabolitos obtenidos evidencien el estado biológico del sistema bajo estudio, se siguieron una serie de estrategias que ofrecían la calidad requerida y nos permitieron obtener resultados minimizando cualquier fuente de variación que no estuviese relacionada con la naturaleza biológica de las muestras. Esto fue posible con la inclusión de muestras “control de calidad”, QC, que permitieron el acondicionamiento inicial del equipo, además de blancos de muestras y estrategias de limpieza entre inyecciones, así como el análisis de estas mezclas de muestras, QC, durante la secuencia.

Asimismo, se incorporaron diferentes IS a las muestras, QC y blancos, en diferentes etapas del procesamiento de las muestras, por ejemplo, durante la homogenización y extracción de los metabolitos, durante la preparación de las muestras para el análisis CE-MS, o durante la preparación y derivatización de las muestras en

GC-MS. Estos IS proporcionan información crucial para el seguimiento del proceso de extracción del metaboloma durante la preparación de muestras, seguimiento de la reacción de derivatización en GC-MS, o la supervisión de la robustez en el análisis de separación (tiempos de retención y forma del pico) y detección (área del pico y exactitud de masa).

Los blancos de muestras, procesados de igual forma que las muestras, pero sin contenido biológico, permiten detectar cualquier contaminante proveniente del tratamiento de muestra o del análisis y, así, eliminarlo. Por su parte, con los QC, nos aseguramos un equilibrio inicial del sistema y el correcto funcionamiento del equipo durante el seguimiento del análisis a lo largo de la secuencia.

Los datos obtenidos tras el alineamiento y el filtrado se usaron para el análisis multivariante no supervisado (PCA) que permitió verificar la distribución de los QCs en cada técnica. El agrupamiento espontáneo de este grupo de muestras se puede ver en la Figura 1, lo que reflejó la estabilidad del sistema y, el rendimiento y reproducibilidad de los procedimientos empleados en el tratamiento de muestra.

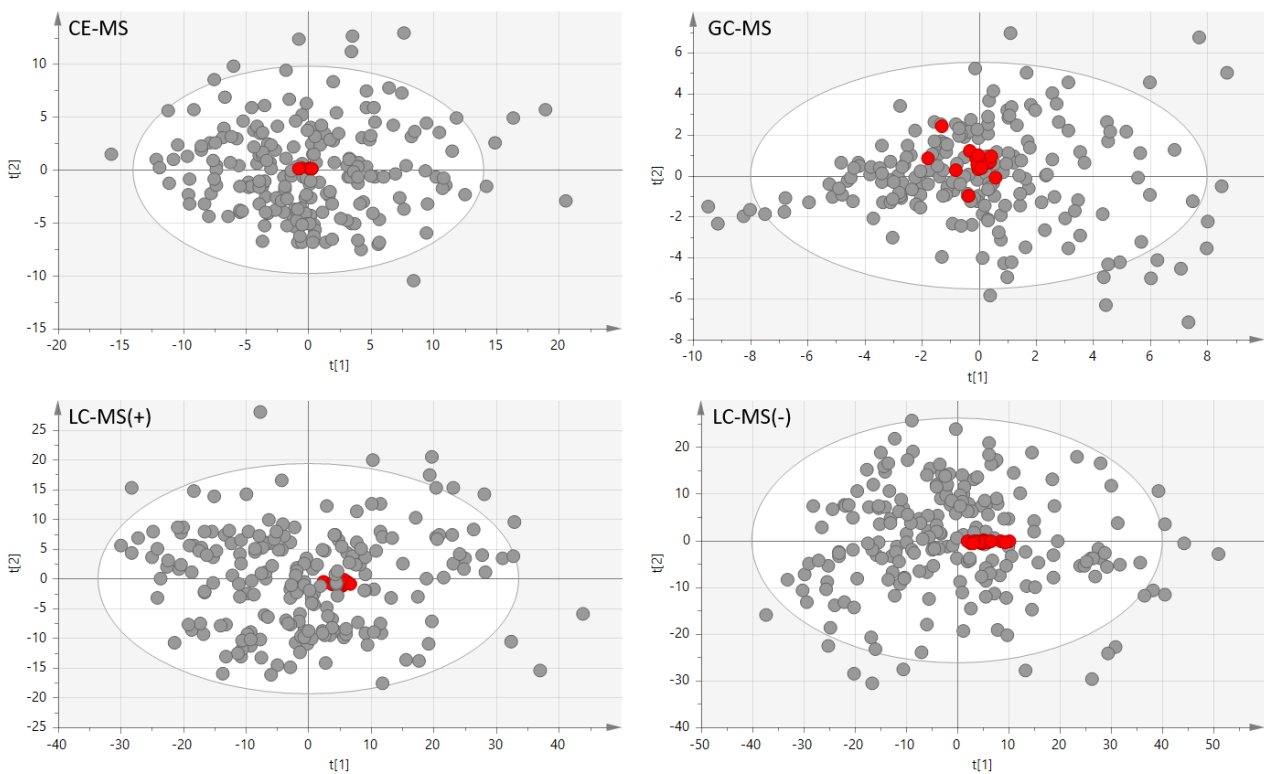


Figura 1. Análisis multivariante no supervisado con los modelos PCA-X generados del análisis de los datos normalizados de los QC y las muestras de estudio (círculo rojo, muestras QC), según la plataforma empleada. Valores de R^2 para $t[1]$ y $t[2]$ de 0.53 para CE-MS, $R^2 = 0.35$ para LC-MS/ESI(+), $R^2 = 0.46$ para LC-MS/ESI(-), y $R^2 = 0.414$ para GC-MS

Posteriormente se fue completando el análisis multivariante para cada plataforma. Se construyeron modelos no supervisados para observar el agrupamiento natural de los grupos casos y controles (Figura 2) junto con los modelos supervisados PLS-DA y OPLS-DA. Los modelos supervisados PLS-DA de la Figura 3 mostraron una clara separación entre grupos.

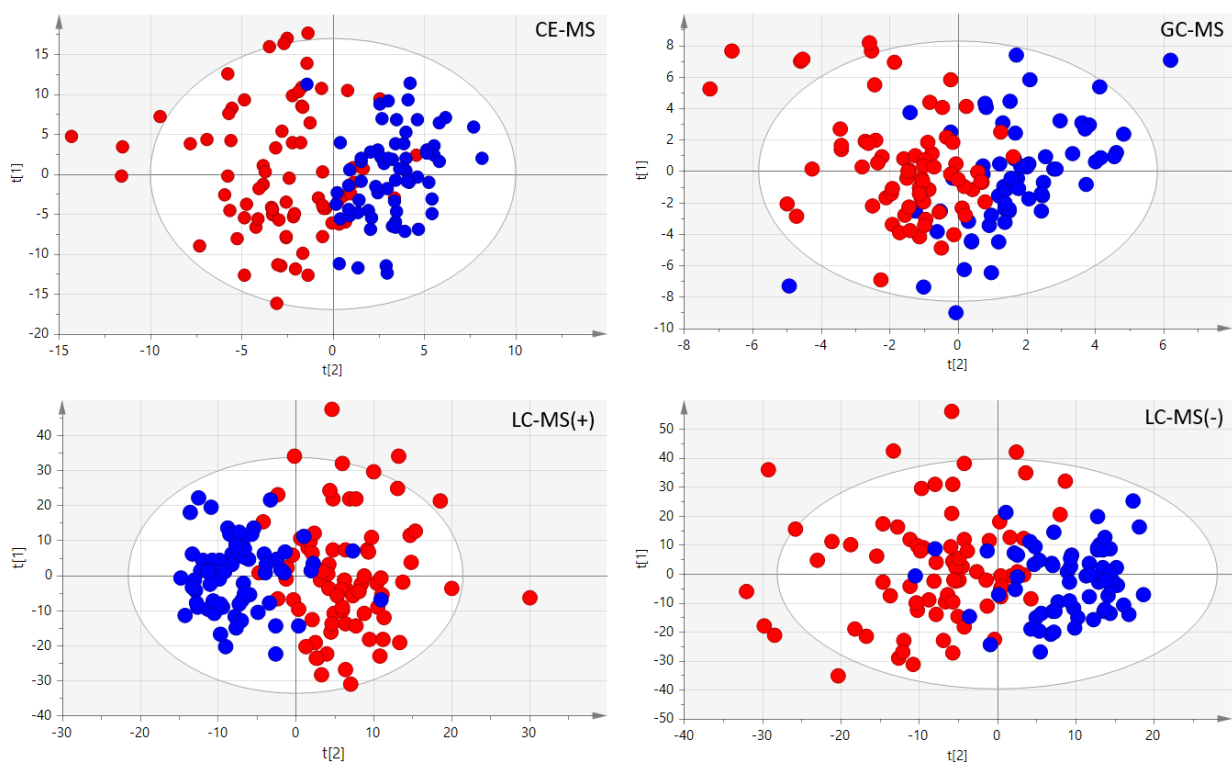


Figura 2. Modelo no supervisado PCA-X para el análisis multiplataforma de los dos grupos de estudio (los círculos rojos representan los individuos NET ($n = 77$) y, los círculos azules los individuos control ($n = 68$)) con una calidad de los valores que explican y predicen el modelo para CE-MS ($R^2 = 0.449$, $Q^2 = 0.298$), LC-MS/ESI(+) ($R^2 = 0.393$, $Q^2 = 0.194$), LC-MS/ESI(-) ($R^2 = 0.386$, $Q^2 = 0.219$) y GC-MS ($R^2 = 0.611$, $Q^2 = 0.352$).

Los modelos OPLS-DA se usaron para modelar aquellas diferencias entre grupos y se validaron usando la prueba de permutación (Figura 4). En ellos se pudo observar una clara separación entre NET y controles para todas las técnicas analíticas empleadas. Todos los modelos presentaron una diferencia en los valores de R^2 y Q^2 menor de 0.3, y unos valores de Q^2 elevados lo que refleja la alta calidad de los modelos. Estos modelos fueron validados mediante la respuesta a la prueba de permutaciones, a través de 999 permutaciones. Las gráficas de validación mostraron valores de Q^2 en Y por debajo de 0, además que los del modelo permutado era menores que los del modelo original. Esto indicó que el modelo original no ocurre por coincidencia.

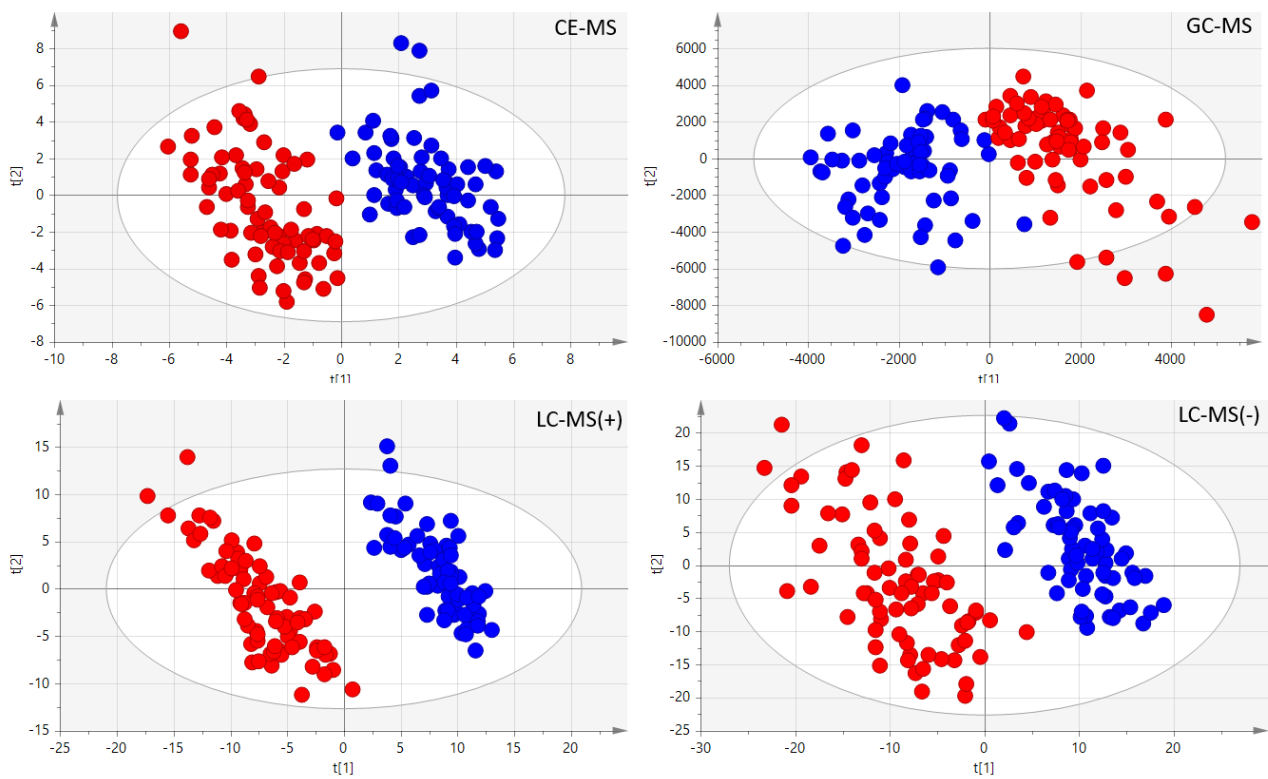


Figura 3. Modelos supervisados PLS-DA construidos del análisis de los dos grupos de estudio (NET ($n = 77$), círculos rojos, versus controles ($n = 68$), círculos azules). Valores de los modelos representados para los datos de CE-MS ($R^2 = 0.872$, $Q^2 = 0.831$), datos de LC-MS/ESI(+) ($R^2 = 0.954$, $Q^2 = 0.858$), datos de LC-MS/ESI(-) ($R^2 = 0.885$, $Q^2 = 0.775$) y datos GC-MS ($R^2 = 0.781$, $Q^2 = 0.732$).

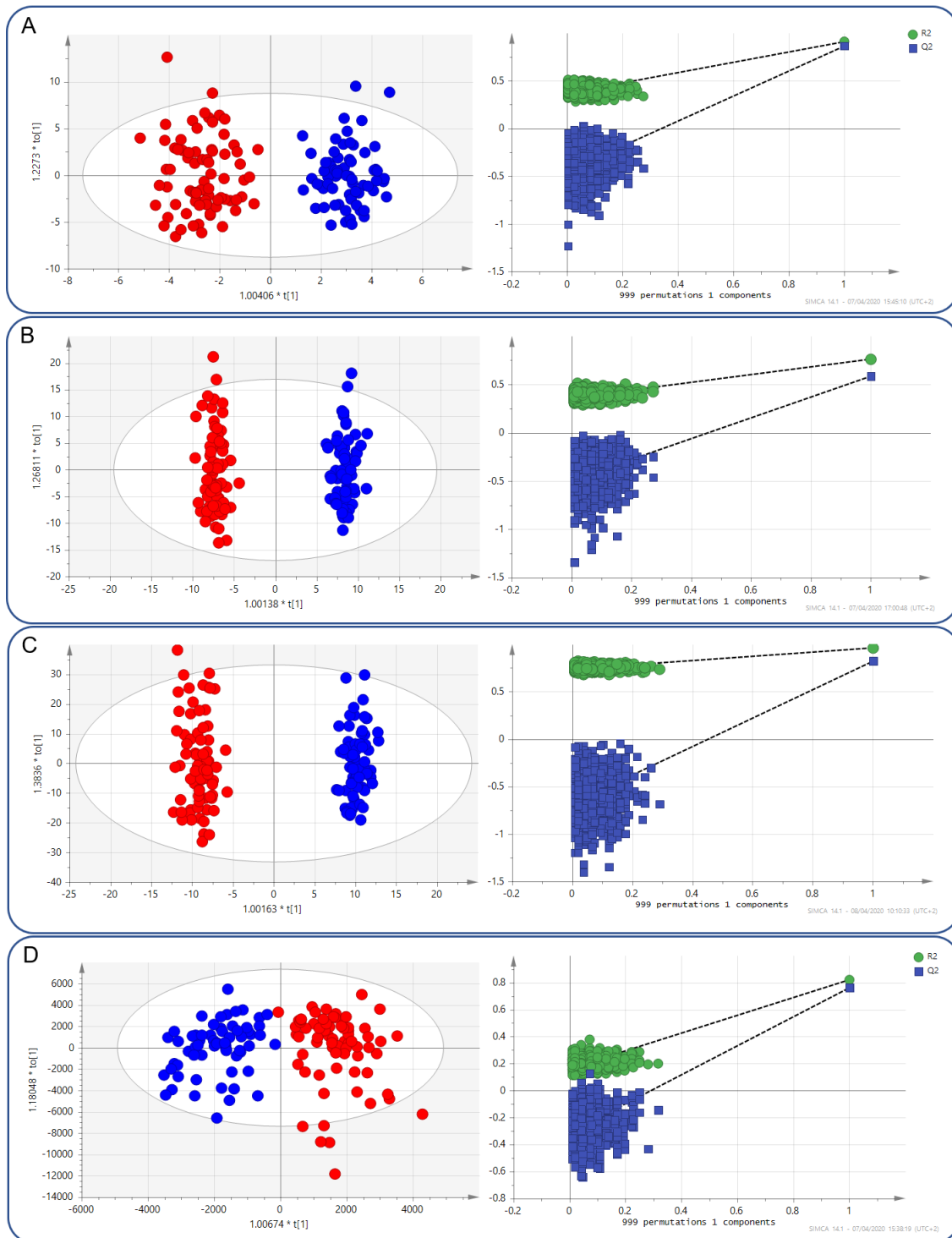


Figura 4. En los paneles se muestra los modelos OPLS-DA (círculos rojos, individuos NET ($n = 77$); círculos azules, individuos control ($n = 68$)) y los gráficos generados de la prueba de permutaciones de estos modelos para cada plataforma tras 999 permutaciones. El panel A pertenece a los datos de CE-MS ($R^2 = 0.872$, $Q^2 = 0.843$); panel B, datos de LC-MS/ESI(+) ($R^2 = 0.954$, $Q^2 = 0.871$); panel C, datos de LC-MS/ESI(-) ($R^2 = 0.885$, $Q^2 = 0.788$) y panel D, datos de GC-MS ($R^2 = 0.781$, $Q^2 = 0.744$). Los modelos indican que los individuos control pudieron ser claramente diferenciados de la mayoría de los individuos NET.

Como se ha explicado, el estudio de la huella metabólica se realizó mediante el análisis realizado por las plataformas CE-MS, LC-MS y GC-MS. Tras el reprocesado de datos para la extracción de las señales, la inspección de estas y el filtrado, el conjunto de datos de CE-MS se redujo a 219 potenciales compuestos. Para el análisis de LC-MS se obtuvieron tras el reprocesado *fuera de línea* 1410 potenciales compuestos en polaridad positiva y 2149 en polaridad negativa. Estos potenciales compuestos de LC-MS, se inspeccionaron y filtraron de acuerdo con el RSD de los QCs y el porcentaje de cambio, obteniéndose 219 y 508, respectivamente. El conjunto de datos de GC-MS mostró 60 compuestos tras el agrupamiento de los distintos derivados provenientes del mismo metabolito. El análisis estadístico univariante reveló la significancia estadística individual de cada compuesto, que se compararon en cada técnica aplicada, obteniendo 75 potenciales compuestos significativos en CE-MS, 150 en LC-MS/ESI(+), 296 en LC-MS/ESI(-) y 19 en GC-MS. Estas variables fueron anotadas o identificadas según lo descrito en el apartado “*Anotación e identificación de los metabolitos*”, y se muestran en la Tabla 5, entre otras características, el cálculo del error de masa medido experimentalmente comparado con la masa monoisotópica de las bases de datos (excepto para los obtenidos de GC-MS) o la clase a la cual pertenece. La tabla resumió aquellos compuestos anotados por la información proporcionada por MS1 (nivel 3) y aquellos identificados por fragmentación en fuente en CE-MS, MS/MS en LC-MS o índice de retención y librería espectral de plasma en GC-MS (nivel 2), de acuerdo con lo establecido en la iniciativa estandarizada de metabolómica (53), además de tiempo de retención, fórmula química, masa, % de cambio, p valor, $p(\text{corr})$ o VIP.

Tabla 5. Resumen de los metabolitos anotados estadísticamente significativos discriminantes entre los perfiles plasmáticos de pacientes NET (n = 77) y no-cáncer (n = 68), con sus características estadísticas procedentes tanto de UVA como de MVA (porcentaje de cambio, p valor, p(corr) y VIP) y descriptores analíticos (masa experimental y su desviación de la teórica, tiempo de retención experimental, plataforma analítica en la que se ha detectado, fuente de anotación, donde DB corresponde al resultado de la base de datos, nivel de confianza para la anotación según la iniciativa de estándares de metabolómica y su correspondiente código de la base de datos del metaboloma humano, HMDB (<https://hmdb.ca/>)).

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
Amines												
Triethylamine	C ₆ H ₁₅ N	101.1204	10.28	-53	0.0009			0	CE-MS	DB	3	HMDB0032539
Amino acids, peptides, and analogues												
Arginine	C ₆ H ₁₄ N ₄ O ₂	174.1131	9.98	243	4.58x10 ⁻³⁵	0.80	2.33	15	CE-MS LC-MS(+)	DB	2	HMDB0000517
Arginium(1+)	C ₆ H ₁₅ N ₄ O ₂	175.1184	11.60	33	0.0496			6	CE-MS	DB	3	HMDB0062762
Arg-Val	C ₁₁ H ₂₃ N ₅ O ₃	273.1822	9.95	66	0.0004			8	CE-MS	DB	3	HMDB0028722
Aspartate	C ₄ H ₇ NO ₄	133.0375	13.80	-32	4.98x10 ⁻¹³			2	CE-MS	DB	2	HMDB0000191
Cys-Gly	C ₅ H ₁₀ N ₂ O ₃ S	178.0412	12.19	-29	8.00x10 ⁻⁰⁹			0	CE-MS	DB	3	HMDB0000078
Cys-Gly disulfide	C ₈ H ₁₅ N ₃ O ₅ S ₂	297.0455	12.19	-41	1.39x10 ⁻⁰⁸	-0.49	1.63	1	CE-MS	DB	3	HMDB0000709
Cysteineglutathione disulphide	C ₁₃ H ₂₂ N ₄ O ₈ S ₂	426.0912	13.95	-37	7.40x10 ⁻⁰⁶			8	CE-MS	DB	3	
Dimethyl-Arginine	C ₈ H ₁₈ N ₄ O ₂	202.1428	10.67	25	0.0006			1	CE-MS	DB	2	HMDB0001539
GalactosylhydroxyLys	C ₁₂ H ₂₄ N ₂ O ₈	324.1558	11.49	42	0.0014			8	CE-MS	DB	3	HMDB0000600
gamma-Glu-orn *	C ₁₀ H ₁₉ N ₃ O ₅	261.1335	11.30	34	0.0103			4	CE-MS	DB	3	HMDB0002248
Glu-Ala*	C ₈ H ₁₄ N ₂ O ₅	218.0904	14.40	135	4.38x10 ⁻¹⁴			1	CE-MS	DB	3	HMDB0006248
Glu-Arg	C ₁₁ H ₂₁ N ₅ O ₅	303.1554	11.49	87	6.28x10 ⁻¹⁴	0.52	1.55	4	CE-MS	DB	3	HMDB0028813
Glu-Asp	C ₉ H ₁₄ N ₂ O ₇	262.0801	3.78	54	0.0036			7	LC-MS(-)	MS/MS	2	HMDB0030419
Glu-hyp	C ₁₀ H ₁₆ N ₂ O ₆	260.0993	12.98	37	5.12x10 ⁻⁰⁷			6	CE-MS	DB	3	HMDB0011161
Glu-Lys/ε-Glu-Lys	C ₁₁ H ₂₁ N ₃ O ₅	275.1481	12.32	67	1.20x10 ⁻⁰⁹			2	CE-MS	DB	3	HMDB0029154

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
Glu-Lys/ε-Glu-Lys	C ₁₁ H ₂₁ N ₃ O ₅	275.1481	11.37	58	1.55x10 ⁻⁰⁷			2	CE-MS	DB	3	HMDB0029155
Glutamine	C ₅ H ₁₀ N ₂ O ₃	146.0691	1.34	69	0.0006			1	LC-MS(+) LC-MS(-)	MS/MS	2	HMDB0000641
Glu-Val	C ₁₀ H ₁₈ N ₂ O ₅	246.1217	14.68	48	0.0009			1	CE-MS LC-MS(+)	DB	3	HMDB0028832
Glycine	C ₂ H ₅ NO ₂	75.0320	9.75	32	0.0202			-	GC-MS	Fiehn	2	HMDB0000123
Gly-Pro	C ₇ H ₁₂ N ₂ O ₃	172.0831	15.89	-20	0.0275			10	CE-MS	DB	3	HMDB0000721
Homocitrulline	C ₇ H ₁₅ N ₃ O ₃	189.1118	13.52	44	0.0027			3	CE-MS	DB	3	HMDB0000679
Iminodiacetic acid	C ₄ H ₇ NO ₄	133.0375	12.87	-37	0.0158			-	GC-MS CE-MS	Fiehn	2	HMDB0011753
Indoleacetyl glutamine	C ₁₅ H ₁₇ N ₃ O ₄	303.1219	1.70	203	1.43x10 ⁻⁰⁶	0.58	2.05	2	LC-MS(-) LC-MS(+)	DB	3	HMDB0013240
Leu-hyp	C ₁₁ H ₂₀ N ₂ O ₄	244.1423	2.00	87	0.0018			2	LC-MS(-)	DB	3	HMDB0028867
Leu-Phe	C ₁₅ H ₂₂ N ₂ O ₃	278.1630	2.14	Presented in cancer group	3.54x10 ⁻⁰⁵	0.60	2.45	0	LC-MS(-)	DB	3	HMDB0013243
Lys-Asp*	C ₁₀ H ₁₉ N ₃ O ₅	261.1335	11.30	34	0.0103			4	CE-MS	DB	3	HMDB0028947
Methionine S-oxide	C ₅ H ₁₁ NO ₃ S	165.0478	14.06	110	3.30x10 ⁻⁰⁹			11	CE-MS	DB	2	HMDB0002005
N2-Methyl-lysine	C ₇ H ₁₆ N ₂ O ₂	160.1208	10.67	-77	6.97x10 ⁻⁰⁹			2	CE-MS	DB	2	HMDB0002038
N2-Methylproline	C ₆ H ₁₁ NO ₂	129.0791	14.59	-43	0.0031			1	CE-MS	DB	2	HMDB0059649
N6-Acetyl-hydroxy-lysine*	C ₈ H ₁₆ N ₂ O ₄	204.1103	12.78	-55	0.0002			3	CE-MS	DB	3	HMDB0033891
N-acetyl-lysine	C ₈ H ₁₆ N ₂ O ₃	188.1155	13.70	20	0.0002			3	CE-MS	DB	2	HMDB0000206
Ornithine	C ₅ H ₁₂ N ₂ O ₂	132.0906	9.66	-39	5.58x10 ⁻¹⁹	-0.65	1.65	6	CE-MS	DB	2	HMDB0000214
Phenylalanine	C ₉ H ₁₁ NO ₂	165.0770	13.88	-58	0.0050			12	CE-MS GC-MS	DB	2	HMDB0000159
Pipecolic acid	C ₆ H ₁₁ NO ₂	129.0790	10.14	49	0.0307			-	GC-MS	Fiehn	2	HMDB0000716
Pyroglutamine	C ₅ H ₈ N ₂ O ₂	128.0583	11.48	68	1.41x10 ⁻⁰⁵			2	CE-MS	DB	3	HMDB0062558
Ser-Ala*	C ₆ H ₁₂ N ₂ O ₄	176.0789	13.07	-38	1.37x10 ⁻⁰⁸			4	CE-MS	DB	3	HMDB0029032
Ser-hyp*	C ₈ H ₁₄ N ₂ O ₅	218.0904	14.40	135	4.38x10 ⁻¹⁴			1	CE-MS	DB	3	HMDB0029040
Ser-Val*	C ₈ H ₁₆ N ₂ O ₄	204.1103	12.78	-55	0.0002			3	CE-MS	DB	3	HMDB0029052

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
Stearoyl-tyrosine*	C ₂₇ H ₄₅ NO ₄	447.3349	5.15	-66	9.67x10 ⁻¹⁷	-0.75	3.74	1	LC-MS(+)	DB	3	HMDB0062343
Suberylglycine	C ₁₀ H ₁₇ NO ₅	231.1106	1.19	Presented in cancer group	0.0001	0.73	3.82	1	LC-MS(-)	DB	3	HMDB0000953
Thr-Ala	C ₇ H ₁₄ N ₂ O ₄	190.0954	13.83	90	5.09x10 ⁻¹¹			0	CE-MS	DB	3	HMDB0029054
Thr-Gly*	C ₆ H ₁₂ N ₂ O ₄	176.0789	13.07	-38	1.37x10 ⁻⁰⁸			4	CE-MS	DB	3	HMDB0029061
Trp-Phe	C ₂₀ H ₂₁ N ₃ O ₃	351.1582	2.32	74	3.73x10 ⁻¹¹	0.63	1.56	1	LC-MS(-)	DB	3	HMDB0029090
Val-Leu	C ₁₁ H ₂₂ N ₂ O ₃	230.1616	12.83	51	0.0012			6	CE-MS	DB	3	HMDB0029131
<i>Benzene and substituted derivatives</i>												
Mandelic acid	C ₈ H ₈ O ₃	152.0473	2.28	335	0.0037			3	LC-MS(-)	DB	3	HMDB0000703
3-phenylprop-2-en-1-yloxysulfonic acid	C ₉ H ₁₀ O ₄ S	214.0299	2.77	118	0.0005			2	LC-MS(-)	DB	3	HMDB0135284
<i>Carbohydrates and carbohydrate conjugates</i>												
Allose	C ₆ H ₁₂ O ₆	180.0633	17.10	139	1.32x10 ⁻¹⁴	0.62	1.44	-	GC-MS	Fiehn	2	HMDB0001151
Glucose	C ₆ H ₁₂ O ₆	180.0633	17.45	174	1.98x10 ⁻¹⁵	0.62	1.64	-	GC-MS	Fiehn	2	HMDB0000122
Glycerol	C ₃ H ₈ O ₃	92.0473	9.28	45	0.0000			-	GC-MS	Fiehn	2	HMDB0000131
Mannitol	C ₆ H ₁₄ O ₆	182.0790	17.59	196	0.0017			-	GC-MS	Fiehn	2	HMDB0000765
Phenylglucuronide	C ₁₂ H ₁₄ O ₇	270.0739	0.93	308	0.0052	0.52	2.37	2	LC-MS(-)	DB	3	HMDB0060014
PYRANOSE (glucose/altrose /galactose /talose)	C ₆ H ₁₂ O ₆	180.0633	17.24	194	1.61x10 ⁻¹⁶	0.63	1.71	-	GC-MS	Fiehn	2	
<i>Carboximide acids and derivatives</i>												
Acetylspermidine	C ₉ H ₂₁ N ₃ O	187.1671	9.15	38	6.27x10 ⁻⁰⁶			7	CE-MS	DB	3	HMDB0001276
<i>Carboxylic acids and derivatives</i>												
1-Aminocyclohexanecarboxylic acid	C ₇ H ₁₃ NO ₂	143.0944	13.89	-32	0.0366			1	CE-MS	DB	3	HMDB0038249
di-Hydroxymelatonin*	C ₁₃ H ₁₆ N ₂ O ₄	264.1110	1.34	46	0.0268			2	LC-MS(-)	DB	3	HMDB0061136
Edetic Acid	C ₁₀ H ₁₆ N ₂ O ₈	292.0906	0.24	35	0.0001			1	LC-MS(+)	DB	3	HMDB0015109
Isocitric acid Citric acid	C ₆ H ₈ O ₇	192.0270	0.23	64	5.83x10 ⁻⁰⁵			0	LC-MS(-)	MS/MS	2	HMDB0000193 HMDB0000094

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
Diazines/ Pyrimidines and pyrimidine derivatives												
5,6-Dihydrothymine	C ₅ H ₈ N ₂ O ₂	128.0570	12.22	56	0.0169			12	CE-MS	DB	3	HMDB0000079
Fatty Acyls												
3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF)	C ₁₂ H ₁₆ O ₅	240.0998	3.78	69	0.004			2	LC-MS(+)	MS/MS	2	HMDB0061112
8-amino-7-oxo-nonanoic acid*	C ₉ H ₁₇ NO ₃	187.1208	2.06	418	0.0220			2	LC-MS(-)	DB	3	
Arachidonic acid	C ₂₀ H ₃₂ O ₂	304.2402	7.13	62	1.59x10 ⁻⁰⁵			1	LC-MS(-)	MS/MS	2	HMDB0001043
beta-Phenylalanoyl-CoA*	C ₃₀ H ₄₅ N ₈ O ₁₇ P ₃ S	914.1836	2.83	73	7.07x10 ⁻⁰³			1	LC-MS(-)	DB	3	
beta-Phenylalanoyl-CoA*	C ₃₀ H ₄₅ N ₈ O ₁₇ P ₃ S	914.1836	3.62	65	8.78x10 ⁻⁰³			0	LC-MS(-)	DB	3	
DG(31:0)	C ₃₄ H ₆₆ O ₅	554.4910	8.18	-37	0.0008			0	LC-MS(+)	DB	3	HMDB0093505
Docosapentaenoic acid	C ₂₂ H ₃₄ O ₂	330.2558	7.25	75	6.87x10 ⁻⁰⁹	0.55	1.51	1	LC-MS(-)	DB	3	HMDB0006528
Dodecenedioic acid	C ₁₂ H ₂₀ O ₄	228.1362	3.50	191	9.87x10 ⁻⁰⁸	0.60	2.24	2	LC-MS(-)	DB	3	HMDB0000933
Eicosapentaenoic acid	C ₂₀ H ₃₀ O ₂	302.2246	6.77	65	3.88x10 ⁻⁰⁴			2	LC-MS(-)	DB	3	HMDB0001999
Eicosatrienoic acid	C ₂₀ H ₃₄ O ₂	306.2558	7.39	82	1.85x10 ⁻¹⁰			1	LC-MS(-)	DB	3	HMDB0010378
Eicosenoic acid	C ₂₀ H ₃₈ O ₂	310.2871	8.34	79	2.75x10 ⁻¹⁰	0.51	1.50	1	LC-MS(-)	DB	3	HMDB0002231
Glucosylgalactosylhydroxylysine	C ₁₈ H ₃₄ N ₂ O ₁₃	486.2093	12.38	46	1.73x10 ⁻⁰⁵			7	CE-MS	DB	3	HMDB0000585
HETE	C ₂₀ H ₃₂ O ₃	320.2351	5.92	Presented in cancer group	1.24x10 ⁻⁰⁵	0.63	2.64	0	LC-MS(-)	MS/MS	2	HMDB0060101
MG 18:2	C ₂₁ H ₃₈ O ₄	354.2770	6.81	116	0.0007			2	LC-MS(+)	MS/MS	2	HMDB0011538
MG 20:0	C ₂₃ H ₄₆ O ₄	386.3396	7.79	-86	1.30x10 ⁻¹²			5	LC-MS(+)	DB	3	HMDB0072859
N-palmitoyl glutamic acid*	C ₂₁ H ₃₉ NO ₅	385.2828	4.13	62	0.004			0	LC-MS(+)	DB	3	
Oleic acid	C ₁₈ H ₃₄ O ₂	282.2559	20.47	67	0.0013			-	GC-MS LC-MS(+)	Fiehn	2	HMDB0000207
Vaccenic acid	C ₁₈ H ₃₄ O ₂	282.2559	20.56	23	0.0158			-	GC-MS	Fiehn	2	HMDB0041480
3-hydroxy-5-octenoylcarnitine	C ₁₅ H ₂₇ NO ₅	301.1889	4.23	Presented in cancer group	3.57x10 ⁻⁰²	0.72	3.96	3	LC-MS(-)	DB	3	
3-Hydroxy-5-tetradecenoylcarnitine*	C ₂₁ H ₃₉ NO ₅	385.2828	4.13	62	0.004			0	LC-MS(+)	DB	3	HMDB0013330

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
9-Decenoylcarnitine	C ₁₇ H ₃₁ NO ₄	313.2232	12.99	-15	0.0325			7	CE-MS	DB	3	HMDB0013205
Arachidonoylcarnitine*	C ₂₇ H ₄₅ NO ₄	447.3349	5.15	-66	9.67x10 ⁻¹⁷	-0.75	3.74	1	LC-MS(+)	DB	3	HMDB0062343
Linolenyl carnitine	C ₂₅ H ₄₃ NO ₄	421.3192	4.91	-41	1.79x10 ⁻⁰⁸	-0.53	2.26	0	LC-MS(+)	DB	3	HMDB0006319
Lysoleilcarnitine	C ₂₅ H ₄₅ NO ₄	423.3349	5.14	-59	4.03x10 ⁻¹⁶	-0.73	3.40	0	LC-MS(+)	DB	3	HMDB0006469
Vaccenylcarnitine Elaidic carnitine	C ₂₅ H ₄₇ NO ₄	425.3504	5.43	-41	1.41x10 ⁻⁰⁸	-0.51	2.40	1	LC-MS(+)	MS/MS	2	HMDB0006351 HMDB0006464
Flavonoids												
Anthraniloyl-CoA	C ₂₈ H ₄₁ N ₈ O ₁₇ P ₃ S	886.1523	0.23	186	3.47x10 ⁻²¹	0.69	2.39	3	LC-MS(-)	DB	3	
Glycerolipids												
11-Oxo-androsterone glucuronide	C ₂₅ H ₃₆ O ₉	480.2359	3.43	-45	8.06x10 ⁻⁰⁶	-0.53	2.06	1	LC-MS(-)	DB	3	HMDB0010338
Glycerophospholipids												
LPC 16:0;OH	C ₂₄ H ₅₀ NO ₈ P	511.3274	4.12	25	0.008			0	LC-MS(+) LC-MS(-)	MS/MS	2	
LPC 16:0;OH	C ₂₄ H ₅₀ NO ₈ P	511.3274	4.21	37	0.0005			0	LC-MS(+) LC-MS(-)	MS/MS	2	
LPC 18:0;OH	C ₂₆ H ₅₄ NO ₈ P	539.3587	4.71	27	0.0100			0	LC-MS(+) LC-MS(-)	MS/MS	2	
LPC 18:2;OH	C ₂₆ H ₅₀ NO ₈ P	535.3274	4.41	Presented in cancer group	2.10x10E ⁻⁰⁶			0	LC-MS(+) LC-MS(-)	MS/MS	2	
LPA 13:0	C ₁₆ H ₃₃ O ₇ P	368.1963	5.23	-57	4.95x10 ⁻⁰³	-0.56	2.29	7	LC-MS(-)	DB	3	HMDB0114760
LPC 22:1	C ₃₀ H ₆₀ NO ₇ P	577.4107	6.99	66	2.14x10 ⁻⁰⁵			2	LC-MS(-)	MS/MS	2	HMDB0010399
LPE 16:0	C ₂₁ H ₄₄ NO ₇ P	453.2855	5.65	25	0.008			0	LC-MS(+)	MS/MS	2	HMDB0011473
LPE 20:5	C ₂₅ H ₄₂ NO ₇ P	499.2699	5.08	179	2.80x10 ⁻⁰⁶			0	LC-MS(+)	MS/MS	2	HMDB0011489
LPE 20:5	C ₂₅ H ₄₂ NO ₇ P	499.2699	5.16	55	0.001			0	LC-MS(+) LC-MS(-)	MS/MS	2	HMDB0011489
LPE 22:6	C ₂₇ H ₄₄ NO ₇ P	525.2855	5.37	28	0.001			0	LC-MS(+)	MS/MS	2	HMDB0011526
LPE 22:6	C ₂₇ H ₄₄ NO ₇ P	525.2855	5.45	39	0.0005			0	LC-MS(+)	MS/MS	2	HMDB0011496
LPE(P-16:0)	C ₂₁ H ₄₄ NO ₆ P	437.2906	5.83	36	0.002			0	LC-MS(+)	DB	3	HMDB0011152
LPI 16:1	C ₂₅ H ₄₇ O ₁₂ P	570.2805	5.51	97	2.63x10 ⁻⁰³			1	LC-MS(-)	MS/MS	2	

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
LPS 18:0	C ₂₄ H ₄₈ NO ₉ P	525.3066	6.71	104	9.62x10 ⁻⁰⁶			2	LC-MS(-)	MS/MS	2	
PC 32:0	C ₄₀ H ₈₀ NO ₈ P	733.5622	10.47	51	3.91x10 ⁻⁰⁷			2	LC-MS(+)	DB	3	HMDB0007871
PC 38:2	C ₄₆ H ₈₈ NO ₈ P	813.6247	11.94	39	0.0004			0	LC-MS(+)	MS/MS	2	HMDB0007987
PC 38:5	C ₄₆ H ₈₂ NO ₈ P	807.5778	9.82	-28	5.05x10 ⁻⁰⁵			2	LC-MS(+)	MS/MS	2	HMDB0008156
PE 34:2 PE(O-34:3)	C ₃₉ H ₇₄ NO ₇ P	699.5203	10.25	-33	0.001			1	LC-MS(+)	MS/MS	2	HMDB0011343
PE 38:6	C ₄₃ H ₇₄ NO ₈ P	763.5151	9.31	-31	0.0087			1	LC-MS(+)	MS/MS	2	HMDB0009294
PG 20:2	C ₂₆ H ₄₉ O ₉ P	536.3114	7.41	-58	3.44x10 ⁻¹¹	-0.59	2.33	4	LC-MS(-)	DB	3	
PG 28:0	C ₃₄ H ₆₇ O ₁₀ P	666.4471	7.64	115	8.69x10 ⁻¹⁵	0.61	2.75	5	LC-MS(+)	DB	3	HMDB0116681
PS 39:5	C ₄₅ H ₇₈ NO ₁₀ P	823.5363	10.46	25	1.60x10 ⁻⁰⁷			5	LC-MS(+)	DB	3	
Hydroxy acids and derivatives												
3-Hydroxydodecanedioic acid	C ₁₂ H ₂₂ O ₅	246.1467	2.79	Presented in cancer group	2.42x10 ⁻⁰⁶	0.68	2.97	0	LC-MS(-)	MS/MS	2	HMDB0000413
3-Hydroxydodecanoic acid	C ₁₂ H ₂₄ O ₃	216.1725	4.62	Presented in cancer group	2.32x10 ⁻¹¹	0.72	2.79	2	LC-MS(-)	MS/MS	2	HMDB0000387
Imidazoles												
Methylimidazole	C ₄ H ₆ N ₂	82.0538	11.48	64	6.41x10 ⁻⁰⁶			9	CE-MS	DB	3	
Urocanate Nicotinamide N-oxide	C ₆ H ₆ N ₂ O ₂	138.0435	11.02	24	0.001			4	CE-MS	DB	3	HMDB0034174 HMDB0002730
Imidazopyrimidines / Purines and purine derivatives												
Hypoxanthine	C ₅ H ₄ N ₄ O	136.0402	14.19	-34	8.20x10 ⁻⁰⁶			13	CE-MS	DB	2	HMDB0000157
Indoles and derivatives												
3-Indoleacetic acid	C ₁₀ H ₉ NO ₂	175.0633	17.94	131	2.62x10 ⁻⁰⁷			-	GC-MS	Fiehn	2	HMDB0000197
5-Hydroxyindole	C ₈ H ₇ NO	133.0527	0.76	37	3.82x10 ⁻⁰²			1	LC-MS(-)	DB	3	HMDB0059805
5-Hydroxyindoleacetaldehyde	C ₁₀ H ₉ NO ₂	175.0633	2.51	150	2.63x10 ⁻⁰⁵	0.51	2.55	1	LC-MS(+)	MS/MS	2	HMDB0004073

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
5-Hydroxyindoleacetic acid	C ₁₀ H ₉ NO ₃	191.0582	0.80	Presented in cancer group	3.28x10 ⁻⁰⁸			0	LC-MS(+)	MS/MS	2	HMDB0000763
Lactones												
N-(4-Coumaroyl)-homoserine lactone	C ₁₃ H ₁₃ NO ₄	247.0845	1.34	55	0.0006			2	LC-MS(+)	DB	3	
Organic acids and derivatives												
(Homo)2-aconitate*	C ₈ H ₁₀ O ₆	202.0477	0.26	-75	7.87x10 ⁻⁵⁹	-0.84	2.98	11	LC-MS(-)	DB	3	
Lactic acid	C ₃ H ₆ O ₃	90.0317	6.06	-77	4.72x10 ⁻³⁶	-0.92	2.58	-	GC-MS	Fiehn	2	HMDB0001311
Pyruvic acid	C ₃ H ₄ O ₃	88.0160	5.89	-58	1.06x10 ⁻¹⁴	-0.69	1.78	-	GC-MS	Fiehn	2	HMDB0000243
Succinylacetoacetate*	C ₈ H ₁₀ O ₆	202.0477	0.26	-75	7.87x10 ⁻⁵⁹	-0.84	2.98	11	LC-MS(-)	DB	3	HMDB0240258
Organic sulfuric acids and derivatives												
Indoxylsulfuric acid	C ₈ H ₇ NO ₄ S	213.0095	1.00	42	3.47x10 ⁻⁰²			1	LC-MS(-)	MS/MS	2	HMDB0000682
p-Phenolsulfonic acid	C ₆ H ₆ O ₄ S	173.9986	0.65	174	3.45x10 ⁻⁰³			1	LC-MS(-)	MS/MS	2	HMDB0060015
Organonitrogen compounds												
Phosphocholine	C ₅ H ₁₄ NO ₄ P	183.0660	5.37	32	0.003			1	LC-MS(+)	MS/MS	2	
Organooxygen compounds												
4-Hydroxycyclohexylcarboxylic acid	C ₇ H ₁₂ O ₃	144.0786	0.76	789	5.08x10 ⁻¹⁰	0.71	2.99	1	LC-MS(-)	DB	3	HMDB0001988
Acetyl-N-formyl-5-methoxykynurenamine (AFMK)*	C ₁₃ H ₁₆ N ₂ O ₄	264.1110	1.34	46	0.0268			2	LC-MS(-)	DB	3	HMDB0004259
Phenols												
4-Methylcatechol	C ₇ H ₈ O ₂	124.0524	1.23	-43	1.68x10 ⁻⁰²			2	LC-MS(-)	DB	3	HMDB0000873
Prenol lipids												
Retinol	C ₂₀ H ₃₀ O	286.2296	7.04	34	0.0003			2	LC-MS(+)	MS/MS	2	HMDB0000305
Purine nucleosides												
1-Methyladenosine	C ₁₁ H ₁₅ N ₅ O ₄	281.1132	12.49	41	3.68x10 ⁻¹⁶	0.58	1.20	3	CE-MS	DB	2	HMDB0003331
Pyridines and derivatives												
Norcotinine	C ₉ H ₁₀ N ₂ O	162.0790	11.48	34	0.0039			2	CE-MS	DB	3	HMDB0001297

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
Piperidine	C ₅ H ₉ N	83.0739	13.77	28	0.0239			5	CE-MS	DB	3	
Serotonine	C ₁₀ H ₁₂ N ₂ O	176.0954	11.40	228	0.0007			3	CE-MS	DB	2	HMDB0001046
Quinolines and derivatives												
8-Hydroxycarteolol	C ₁₆ H ₂₄ N ₂ O ₄	308.1732	13.90	-31	0.0246			1	CE-MS	DB	3	HMDB0060990
Quinoline	C ₉ H ₇ N	129.0578	2.51	99	0.0008			1	LC-MS(+)	DB	3	HMDB0033731
Sphingolipids												
Cer(35:0)	C ₃₅ H ₆₉ NO ₃	551.5277	11.39	58	4.24x10 ⁻⁰⁸			3	LC-MS(-)	DB	3	
Cer(36:1)	C ₃₆ H ₇₁ NO ₃	565.5434	11.74	42	0.0003			1	LC-MS(+)	DB	3	HMDB0004950
SM(36:0)	C ₄₁ H ₈₅ N ₂ O ₆ P	732.6145	10.14	41	0.0014			1	LC-MS(+)	MS/MS	2	HMDB0012087
Sphingosine-1-phosphate	C ₁₈ H ₃₈ NO ₅ P	379.2488	5.02	-30	2.17x10 ⁻⁰⁸	-0.50	2.06	0	LC-MS(+)	MS/MS	2	HMDB0000277
Steroids and steroid derivatives												
12a-Hydroxy-3-oxocholeadienic acid	C ₂₄ H ₃₄ O ₄	386.2457	4.27	144	3.66x10 ⁻⁰⁶	0.54	2.12	0	LC-MS(-)	MS/MS	2	HMDB0000385
Biliverdin	C ₃₃ H ₃₄ N ₄ O ₆	582.2478	4.73	300	1.47x10 ⁻¹⁶	0.72	2.62	0	LC-MS(-)	DB	3	HMDB0001008
Hydroxy-3-oxo-4-cholestenoate	C ₂₇ H ₄₂ O ₄	430.3083	5.53	54	2.78x10 ⁻⁰⁶				LC-MS(+) LC-MS(-)	MS/MS	2	
Calcitric acid	C ₂₃ H ₃₄ O ₄	374.2457	7.39	84	6.25x10 ⁻¹¹	0.54	1.76	6	LC-MS(-)	DB	3	HMDB0006472
Chenodeoxycholic acid 3-glucuronide*	C ₃₀ H ₄₈ O ₁₀	568.3248	4.33	94	2.10x10 ⁻⁰³			1	LC-MS(-)	DB	3	
Cholestane-3,7,12,24,25-pentol	C ₂₇ H ₄₈ O ₅	452.3501	7.77	-56	3.81x10 ⁻¹⁰	-0.56	2.03	2	LC-MS(-)	DB	3	HMDB0000483
Cholestane-3,7,12,25-tetrol-3-glucuronide	C ₃₃ H ₅₆ O ₁₀	612.3873	4.71	97	4.89x10 ⁻⁰⁵			0	LC-MS(-)	DB	3	HMDB0010355
Cortisone acetate	C ₂₃ H ₃₀ O ₆	402.2042	3.86	63	5.89x10 ⁻⁰⁴			9	LC-MS(-)	DB	3	HMDB0015459
Dehydroepiandrosterone 3-glucuronide	C ₂₅ H ₃₆ O ₈	464.2410	3.91	-51	1.01x10 ⁻⁰⁵			2	LC-MS(-)	DB	3	HMDB0010348
Dehydroisoandrosterone 3-glucuronide	C ₂₅ H ₃₆ O ₈	464.2410	3.91	-51	1.01x10 ⁻⁰⁵			2	LC-MS(-)	DB	3	HMDB0010327
Deoxycholic acid 3-glucuronide*	C ₃₀ H ₄₈ O ₁₀	568.3248	4.33	94	2.10x10 ⁻⁰³			1	LC-MS(-)	DB	3	HMDB0002596
ecdysone 25-O-D-glucopyranoside	C ₃₃ H ₅₄ O ₁₁	626.3666	4.41	88	1.02x10 ⁻⁰⁴			1	LC-MS(-)	DB	3	
Pregnanediol	C ₂₁ H ₃₆ O ₂	320.2715	6.35	-51	4.21x10 ⁻⁰⁵			0	LC-MS(-)	DB	3	HMDB0004025
Pregnanolone sulfate	C ₂₁ H ₃₄ O ₅ S	398.2127	3.64	55	5.27x10 ⁻⁰³			0	LC-MS(-)	DB	3	HMDB0240591

Compuesto	Formula	Masa (Da)	RT (min)	% de cambio	p valor	p(corr)	VIP	Error de masa (ppm)	Plataforma analítica	Fuente de anotación	Nivel de confianza	HMDB código
Ursodeoxycholic acid	C ₂₄ H ₄₀ O ₄	392.2927	4.34	188	3.93x10 ⁻⁰²	0.51	2.35	1	LC-MS(-)	MS/MS	2	HMDB0000946
Ursodeoxycholic acid 3-sulfate	C ₂₄ H ₄₀ O ₇ S	472.2494	3.75	130	4.38x10 ⁻⁰³	0.54	2.36	1	LC-MS(-)	DB	3	HMDB0002642
Sterol Lipids												
24-Hydroxygemivitamin D3	C ₃₂ H ₅₄ O ₅	518.3971	6.79	-37	2.82x10 ⁻⁰⁷			1	LC-MS(+)	DB	3	
Tetrapyrroles and derivatives												
Bilirubin	C ₃₃ H ₃₆ N ₄ O ₆	584.2635	3.85	674	3.79x10 ⁻⁰⁸	0.68	2.63	4	LC-MS(-) LC-MS(+)	MS/MS	2	HMDB0000054

* = Múltiples opciones de identificación

Metabolitos como arginina, glutamina, ácido iminodiacético, fenilalanina o ácido oleico se mostraron en la Tabla 5 con anotaciones que proceden de más de una plataforma de análisis. La integración de los datos adquiridos por diferentes plataformas analíticas dio como resultado 155 metabolitos identificados. Además, algunas de estas identificaciones procedentes de más de una plataforma de análisis incrementaron el nivel de confianza en la asignación de la identidad.

Los resultados mostrados en la Figura 5 A ilustran, a través del “heatmap” generado, los 30 principales metabolitos clasificados según pruebas t para retener los patrones con mayor contraste responsables de la separación entre grupos. Esta representación se construyó con la abundancia de cada metabolito en cada una de las muestras de los dos grupos, y la escala de color va desde los aumentados (rojo) a los que se encuentran disminuidos (azul). En él se unió la utilización de agrupamientos jerárquicos en clúster de las variables, método no supervisado, que agrupa en un conglomerado aquellas variables que tengan mayor similitud (o menor distancia). A continuación, y siguiendo con la misma estrategia, se van agrupando las variables en grupos cada vez más grandes y heterogéneos hasta que al final del proceso todas las variables tratadas están englobadas en un mismo conglomerado. El análisis de las rutas metabólicas (MPA) de los 155 metabolitos diferenciales identificó varias vías moleculares desreguladas en los NET (Figura 5B). MPA mostró que la biosíntesis de la arginina, el metabolismo de la arginina y la prolina, la biosíntesis de animoacil-tRNA, el ciclo del citrato y del piruvato, el metabolismo del glutatión, alanina, aspartato y del glutamato eran las vías más comúnmente desreguladas en los pacientes con NET. La información proporcionada por el impacto en las rutas alteradas nos ayuda a visualizar aquellas vías que juegan un papel clave en la regulación de los procesos cancerígenos en pacientes con NET.

Los metabolitos discriminantes resumidos en la Tabla 5 se clasificaron según la clase bioquímica a la que pertenecen (Figura 6). Las principales categorías predominantes identificadas fueron los aminoácidos, péptidos y sus derivados (27.7%), ácidos grasos (16.1%), glicerofosfolípidos (14.2%), esteroides y derivados (9.7%) y carbohidratos y sus conjugados (3.9%). (Figura 6A). Setenta y dos por ciento de todos los metabolitos diferenciales se regularon positivamente en pacientes con NET. La proporción de metabolitos regulados hacia arriba o hacia abajo por clase bioquímica se resume en la Figura 6B. Se debe destacar el elevado número de aminoácidos y sus derivados alterados. Entre ellos, derivados como metil-lisina, metil-prolina o estearoil-tirosina mostraron una disminución en sus niveles frente a la línea base. Esto también

pudo observarse en las estructuras dipeptídicas que predominan en este grupo. Del análisis por LC-MS, los resultados fueron distribuidos en diferentes tipos de lípidos. Siguiendo la clasificación oficial de estos (54), predominaron las categorías principales de glicerofosfolípidos (GPL), ácidos grasos (FA) y esteroides. Los GPL representaron el grupo de lípidos alterado más grande encontrado. Las formas lisas se encontraron fundamentalmente incrementadas al contrario que lo encontrado en glicerofosfolípidos y glicerofosfoetanolaminas que se encontraban bajo-reguladas. Los lisoglicerofosfolípidos oxidados encontrados presentaban elevadas abundancias en el grupo NET, y hasta la fecha, existen muy pocos estudios que los reporten. Los ácidos grasos y sus derivados representaron aproximadamente el 30% de los lípidos encontrados y, debe ser remarcado que, en general, los niveles en carnitinas se presentaron disminuidos. Asimismo, ha sido posible la identificación de niveles incrementados en derivados oxidados del ácido araquidónico (HETE). En la identificación de todas estas modificaciones oxidadas se siguió un flujo de trabajo establecido que permitió la anotación de posibles compuestos desconocidos sustentada por la información estructural de sus espectros de fragmentación, así como por el mecanismo de fragmentación involucrado, que incluye datos de LC-MS/MS en ambas polaridades y con diferentes energías de colisión (55).

Otro de los grupos predominantes, los esteroides, está formado principalmente por ácidos biliares que presentaron un contenido elevado en el grupo NET. También se pudo observar niveles elevados en azúcares y ácido cítrico, y disminuidos en ácidos láctico y pirúvico, todos ellos relacionados mediante la glucólisis y el ciclo del ácido tricarboxílico. Estos resultados incluyen un destacado incremento en serotonina, hormona secretada en elevadas cantidades en pacientes NET que presentaron síndrome carcinoide y su principal producto, el ácido 5-hidroxiindolacético.

Adicionalmente, también se detectaron cambios en dos metabolitos de pigmentos hemo no conjugados, bilirrubina y biliverdina. Ambos se encontraron sobre-regulados con importantes diferencias entre los grupos. En las células de mamíferos, la bilirrubina es el producto final del catabolismo hemo. El grupo hemo es degradado a biliverdina por la hemooxigenasa, que posteriormente se reduce a bilirrubina por la enzima biliverdina reductasa. Esta es una proteína citoprotectora y promotora del crecimiento celular, directamente relacionada con la transcripción de genes involucrados en la respuesta al estrés oxidativo (56, 57). Junto con un aumento en la expresión en las células tumorales, la biliverdina reductasa se identificó como un promotor tumoral (58).

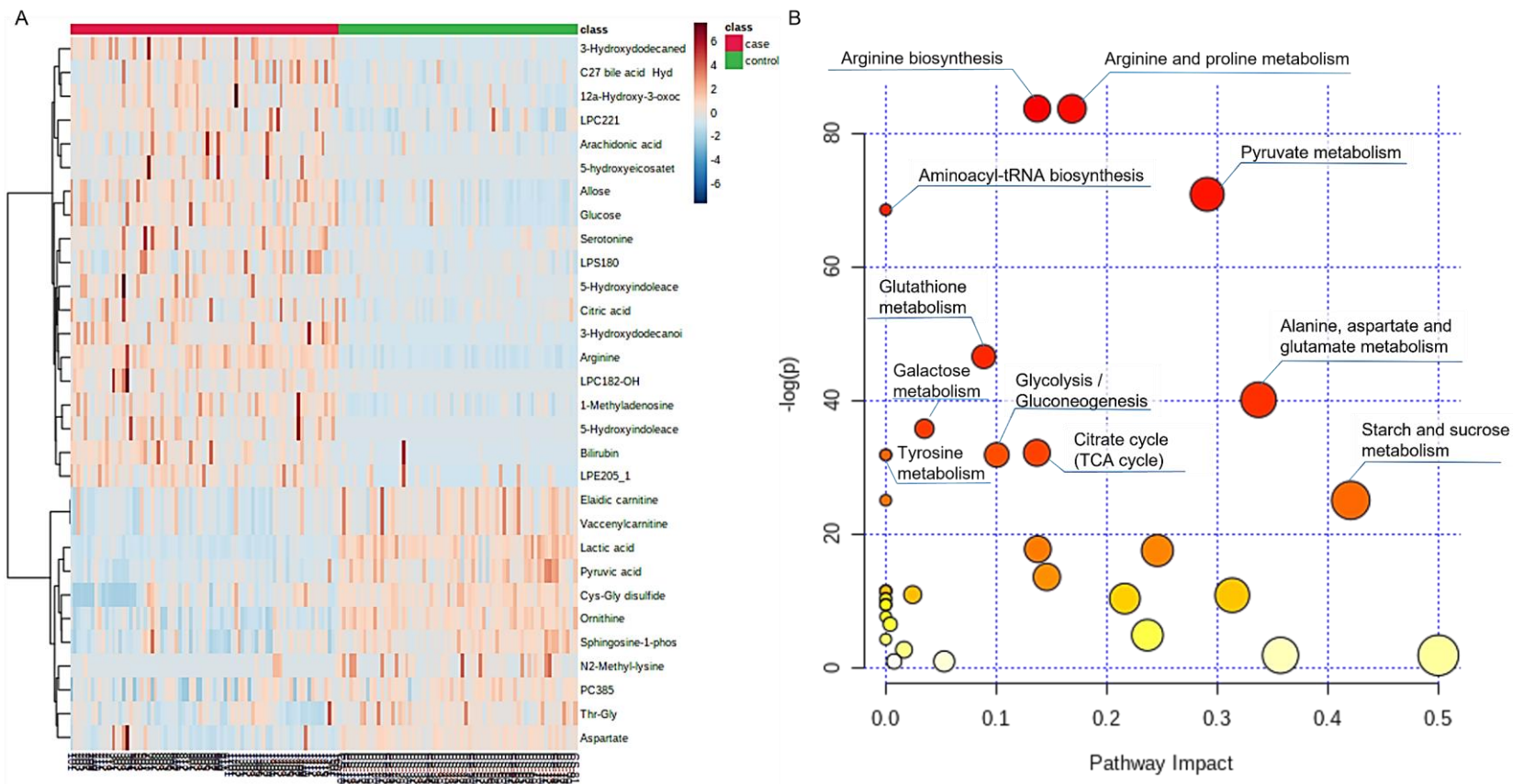
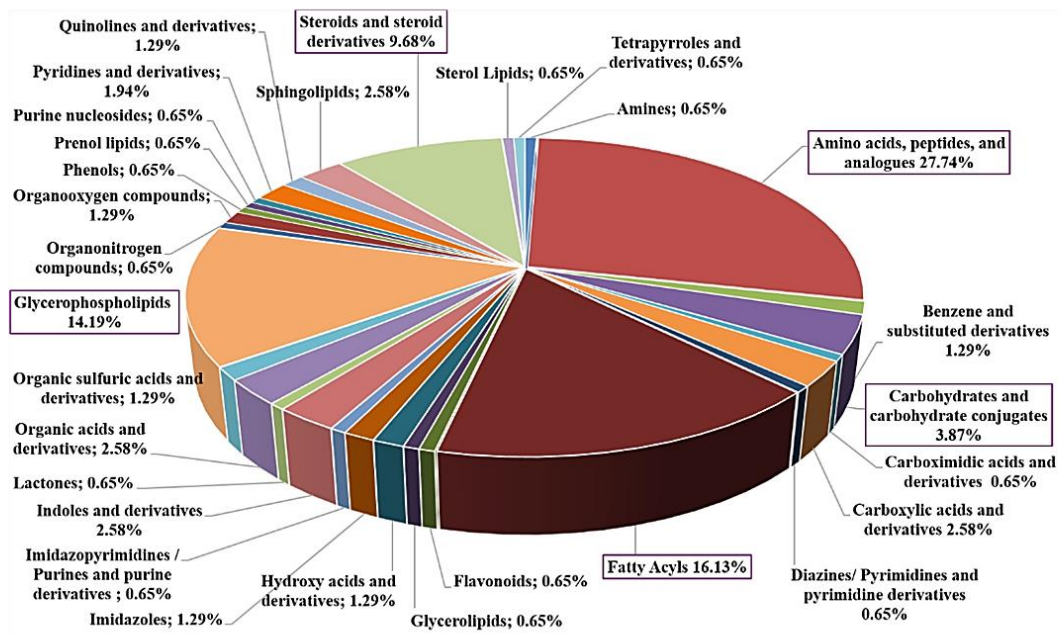


Figura 5. Resultados del estudio multiplataforma de diferencias metabólicas en los perfiles plasmáticos entre pacientes con NET ($n = 77$) e individuos control ($n = 68$). **A.** Mapa de calor o “heatmap” de los 30 principales metabolitos identificados clasificados por pruebas t . Las muestras se representan en columnas y las variables en filas. El espectro de color varía de azul oscuro a rojo oscuro para indicar valores de datos que cambian de muy bajo a extremadamente alto respectivamente. La agrupación jerárquica se realizó solo en los compuestos, mientras que las muestras se ordenan según su pertenencia a un grupo. **B.** Análisis de rutas metabólicas (MPA). Las vías metabólicas alteradas mostraron cambios en el metabolismo en la comparación NET versus control. El color y el tamaño de cada círculo se basan en su valor p y el valor de impacto de la ruta, respectivamente. Como resultado, las vías ubicadas en la esquina superior derecha del gráfico son las más afectadas por un tumor neuroendocrino.

A



B

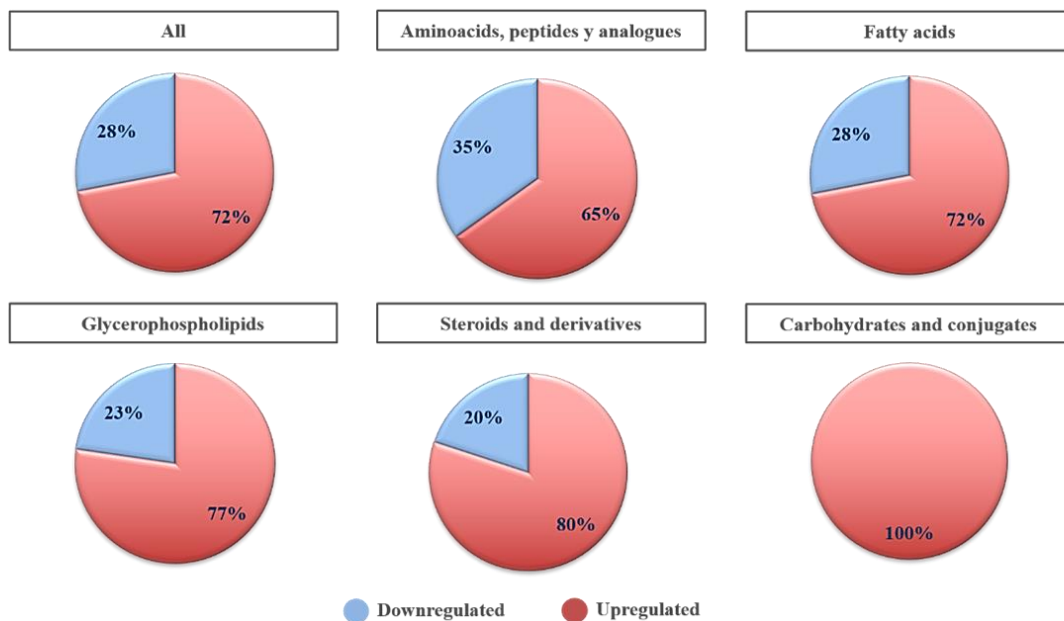


Figura 6. **A.** Gráfico circular de la distribución de compuestos discriminantes y anotados clasificados en las categorías bioquímicas principales. Este gráfico resume el porcentaje ($n = 155$) de cada clase de compuesto que se consideró biológicamente relevante mediante el análisis multiplataforma en la comparación de NET ($n = 77$) frente a individuos sin cáncer ($n = 68$). Los metabolitos se clasificaron en 28 clases según su naturaleza bioquímica. Este análisis reveló

aminoácidos, péptidos y análogos (n = 43; 27.74%), acilos grasos (n = 26; 16.13%), glicerofosfolípidos (n = 22; 14.19%), esteroides y derivados esteroides (n = 16; 9.68%) y carbohidratos y conjugados de carbohidratos (n = 6; 3.87%) como las clases bioquímicas más representadas. B. Gráficos circulares que resumen el porcentaje de metabolitos plasmáticos regulados al alza y a la baja en las principales clases bioquímicas que son biológicamente relevantes en el estudio comparativo. En general, la sobre-regulación (72%) prevaleció sobre la regulación a la baja (28%) en el conjunto de metabolitos completo (n = 155). Esta tendencia de metabolitos regulados al alza versus regulados a la baja se mantuvo constante para todas las clases bioquímicas principales: aminoácidos, péptidos y análogos (65% frente a 35%), acilos grasos (72% frente a 28%), glicerofosfolípidos (77% vs 23%), esteroides y derivados de esteroides (80% vs 20%) y carbohidratos y conjugados de carbohidratos (100%).

Por otra parte, considerando la relevancia funcional de los metabolitos seleccionados documentados en la literatura, encontramos varias vías enriquecidas relacionadas con la oncogénesis, algunas específicamente involucradas en el desarrollo de NET, como la angiogénesis, la vía mTOR, el metabolismo del triptófano, el efecto Warburg, el estrés oxidativo y el ciclo de la urea, entre otros (Figura 7).

De los resultados hasta ahora recogidos, es evidente la capacidad relevante para discriminar a los pacientes con NET de los individuos sin cáncer según los niveles de sus metabolitos, por lo que se aplicó el análisis ROC y se calculó el área bajo la curva (AUC) de cada metabolito identificado para determinar su desempeño individual como biomarcador diagnóstico de NET. También fue calculada la importancia de la variable en la puntuación de proyección (VIP) del modelo OPLS-DA. El valor de VIP estima la importancia de cada metabolito en el modelo y, por lo tanto, su capacidad para diferenciar los NET de los pacientes sin cáncer. La Tabla 6 resume los resultados obtenidos del análisis ROC y OPLS-DA de los 155 metabolitos plasmáticos diferenciales identificados entre NET y controles. En la tabla aparecen aquellos con un AUC > 0.85, un VIP > 1.0 y un $|p(\text{corr})| > 0.5$ o combinaciones de estos que se consideraron metabolitos con potencial biomarcador. Veintisiete metabolitos tenían tanto un AUC > 0.85 como un VIP > 1.0 y un $|p(\text{corr})| > 0.5$; 17 metabolitos tenían sólo un VIP > 1.0 y un $|p(\text{corr})| > 0.5$ y 5 metabolitos tenían solo un AUC > 0.85. Por lo tanto, se identificaron 49 metabolitos con potencial diagnóstico significativo.

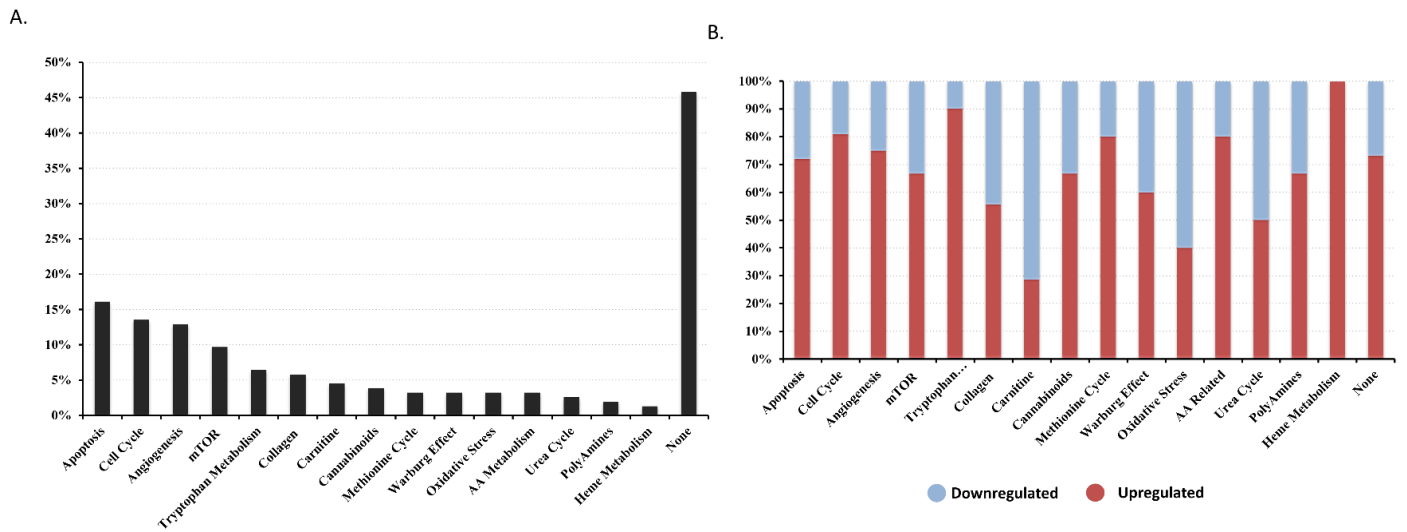


Figura 7. Clasificación de las vías de los metabolitos plasmáticos diferenciales en pacientes con NET. **A.** Gráfico de barras que representa la distribución porcentual de metabolitos diferenciales ($n = 155$) según las vías moleculares en pacientes con NET. Los metabolitos se clasificaron en 16 clases según su relevancia funcional. Las vías enriquecidas con mayor frecuencia incluyeron las vías clásicas del cáncer como la apoptosis ($n = 25$; 16.13%) y el ciclo celular ($n = 21$; 13.55%), y vías clave en la patogenia de los NET como la angiogénesis ($n = 20$; 12.9%), la vía mTOR ($n = 15$; 9.68%) y el metabolismo del triptófano ($n = 10$; 6.45%). **B.** Gráfico de barras que resume el porcentaje de metabolitos plasmáticos regulados al alza y a la baja en las vías moleculares relevantes en los NET ($n = 77$). Nuevamente, la tendencia de metabolitos regulados hacia arriba o hacia abajo que se encuentran en las clases bioquímicas se mantuvo constante en las vías: apoptosis (72% frente a 28%) y ciclo celular (80.95% frente a 19.05%), angiogénesis (75% frente a 25%), la vía mTOR (66.67% vs 33.33%) y el metabolismo del triptófano (90% vs 10%).

Tabla 6. Metabolitos con potencial biomarcador en pacientes NET. Esta tabla resume los resultados del análisis ROC y OPLS-DA de los 155 metabolitos plasmáticos diferenciales identificados entre pacientes con NET (n = 77) y pacientes sin cáncer (n = 68). Se muestran el AUC, la sensibilidad y la especificidad (según el índice de Youden) para la selección de los 49 metabolitos, además de p(corr) y VIP (parámetros OPLS-DA) para los metabolitos significativos. Aquellos con un AUC > 0.85, un VIP > 1.0 y un |p(corr)| > 0.5 o la combinación de estos, se consideraron metabolitos con potencial biomarcador. 27 metabolitos tenían ambos, un AUC > 0.85, un VIP > 1.0 y un |p(corr)| > 0.5; 17 metabolitos tenían un VIP > 1.0 y un |p(corr)| > 0.5 y, 5 metabolitos tenían solo un AUC > 0.85. Los criterios de selección los cumplieron 49 metabolitos y, fueron seleccionados para el estudio de potenciales biomarcadores. (* = múltiples opciones de identificación).

Metabolito	AUC	Especificidad	Sensibilidad	p(corr)	VIP
Lactic acid	0.999	0.985	1	-0.918	2.576
(Homo)2-aconitate*	0.998	1	0.987	-0.843	2.984
Succinylacetoacetate*	0.998	1	0.987	-0.843	2.984
Arginine	0.982	0.956	0.961	0.795	2.326
Biliverdín	0.962	0.941	0.896	0.724	2.62
Anthraniloyl-CoA	0.954	0.941	0.896	0.689	2.387
Stearoyl-tyrosine*	0.938	0.794	0.948	-0.755	3.742
Arachidonoylcarnitine*	0.938	0.794	0.948	-0.755	3.742
3-hydroxy-5-octenoylcarnitine	0.937	0.971	0.883	0.723	3.964
Lynoleilcarnitine	0.937	0.941	0.805	-0.732	3.395
Pyruvic acid	0.904	0.892	0.795	-0.686	1.784
Ornithine	0.899	0.926	0.779	-0.646	1.654
PYRANOSE (glucose/altrose /galactose /talose)	0.899	0.831	0.849	0.633	1.713
5-Hydroxyindoleacetic acid	0.891	0.985	0.792		
Suberylglycine	0.89	1	0.779	0.734	3.819
4-Hydroxycyclohexylcarboxylic acid	0.888	0.794	0.909	0.713	2.992
3-Hydroxydodecanoic acid	0.887	0.809	0.922	0.725	2.793
Glucose	0.887	0.862	0.808	0.62	1.636
Trp-Phe	0.887	0.765	0.896	0.626	1.555
Allose	0.88	0.815	0.863	0.618	1.437
Glu-Arg	0.879	0.882	0.779	0.524	1.548
3-Hydroxydodecanedioic acid	0.876	0.809	0.87	0.679	2.975
MG(20:0)	0.873	0.868	0.896		
Aspartate	0.872	0.912	0.766		
PG(28:0)	0.871	0.824	0.779	0.605	2.75
Bilirubin	0.861	0.926	0.779	0.679	2.628
Cholestane-3,7,12,24,25-pentol	0.858	0.794	0.818	-0.555	2.034
1-Methyladenosine	0.857	0.779	0.831	0.58	1.198
Glu-Ala / Ser-hyp*	0.853	0.926	0.753		
Dodecenedioic acid	0.852	0.809	0.779	0.6	2.237

Metabolito	AUC	Especificidad	Sensibilidad	$p(\text{corr})$	VIP
Indoleacetyl glutamine	0.852	0.824	0.766	0.58	2.047
PG(20:2)	0.846	0.985	0.636	-0.587	2.332
HETE	0.846	0.809	0.805	0.63	2.642
Cys-Gly disulfide	0.837	0.794	0.792	-0.493	1.63
Linolenyl carnitine	0.834	0.824	0.831	-0.534	2.261
Calcitroic acid	0.83	0.632	0.896	0.536	1.759
Oleoylcarnitine/Elaidic carnitine	0.828	0.765	0.818	-0.511	2.402
Docosapentaenoic acid	0.824	0.824	0.74	0.55	1.513
Eicosenoic acid	0.822	0.662	0.831	0.511	1.495
Leu-Phe	0.82	0.882	0.74	0.601	2.449
5-Hydroxyindoleacetaldehyde	0.817	0.632	0.87	0.512	2.553
Ursodeoxycholic acid 3-sulfate	0.807	0.662	0.922	0.544	2.357
Methionine S-oxide	0.806	0.838	0.688		
12a-Hydroxy-3-oxocholadienic acid	0.804	0.721	0.844	0.537	2.119
Sphingosine-1-phosphate	0.803	0.838	0.675	-0.502	2.058
Phenylglucuronide	0.794	0.853	0.74	0.524	2.375
Ursodeoxycholic acid	0.779	0.765	0.818	0.51	2.349
LPA(13:0)	0.779	0.971	0.714	-0.556	2.286
11-Oxo-androsterone glucuronide	0.753	0.956	0.584	-0.533	2.064

MG: monoacilglicerol, PG: glicerofosfoglicerol, LPA: lisoglicerofosfato

A continuación, para determinar qué moléculas estaban directamente relacionadas con los NET, independientemente de otros factores clínicos, se realizó un modelo de regresión logística (LRM) para cada metabolito seleccionado, ajustado por sexo, edad, glucemia, niveles de creatinina y fármacos concomitantes seleccionados como posibles variables de confusión, antes de la inclusión en el modelo. Solo se incluyeron hasta dos asociaciones de fármacos significativas ($p < 0.05$) en el LRM de cada metabolito para evitar el sobreajuste debido a características explicativas excesivas. De los 49 metabolitos evaluados, solo un metabolito, la suberilglicina, no contribuyó significativamente a explicar el resultado de la clasificación en su modelo (Tabla 7), lo que indica una capacidad de diagnóstico más pobre. Por lo tanto, 48 metabolitos contribuyeron significativamente a explicar el modelo de clasificación (NET frente a pacientes sin cáncer), independientemente de otros factores de confusión.

Tabla 7. Análisis de regresión logística (LRM) de metabolitos con potencial biomarcador en el plasma de pacientes NET. Se construyó un modelo LRM para cada metabolito con potencial biomarcador. Cada modelo se ajustó con covariables clínicas (sexo, edad, niveles de glucemia y niveles de creatinina) y con hasta 2 covariables de ingesta de fármaco significativamente asociadas con el fin de evaluar la capacidad diagnóstico independiente de cada metabolito. Las covariables de ingesta clínica y de fármaco incluidas en cada modelo se indican en el ajuste del modelo. Se muestra Exp (B) y, su IC del 95% y el valor p de cada metabolito en su modelo. Los valores de $p < 0.05$ se consideraron significativos. Todos los metabolitos, excepto la suberilglicina, discriminaron independientemente a los pacientes con NET de los no oncológicos.

Ajuste del modelo	Metabolito	
	Exp (B) (95% IC)	p-value
1-methyladenosine, AGE, SEX, GLUC, CRE, DIU, APAP	2.63·10 ⁷ (4.49·10 ⁴ – 1.55·10 ¹⁰)	1.49·10 ⁻⁷
11-Oxo-androsterone glucuronide, AGE, SEX, GLUC, CRE	0.06 (0.01 – 0.59)	1.49·10 ⁻²
12a-Hydroxy-3-oxocholadienic acid, AGE, SEX, GLUC, CRE, RAN	22.45 (4.67 – 108.01)	1.04·10 ⁻⁴
3-hydroxy-5-octenoylcarnitine, AGE, SEX, GLUC, CRE, RAN, STE	3.13 (2.15 – 4.55)	2.13·10 ⁻⁹
3-Hydroxydodecanedioic acid, AGE, SEX, GLUC, CRE	2.31 (1.82 – 2.94)	8.84·10 ⁻¹²
3-Hydroxydodecanoic acid, AGE, SEX, GLUC, CRE	3.04 (1.56 – 1.82·10 ³)	3.07·10 ⁻¹²
4-Hydroxycyclohexylcarboxylic acid, AGE, SEX, GLUC, CRE	2.53 (1.93 – 3.30)	1.25·10 ⁻¹¹
5-Hydroxyindoleacetaldehyde, AGE, SEX, GLUC, CRE, RAN, STE	77.24 (11.80 – 505.37)	6.00·10 ⁻⁶
5-Hydroxyindoleacetic acid, AGE, SEX, GLUC, CRE, RAN, APAP	3.83 (2.13 – 6.87)	7.00·10 ⁻⁶
Allose, AGE, SEX, GLUC, CRE, APT	11.78 (3.15 – 44.09)	2.50·10 ⁻⁴
Anthraniloyl-CoA, AGE, SEX, GLUC, CRE	2.51·10 ⁸ (1.39·10 ⁵ – 4.54·10 ¹¹)	4.32·10 ⁻⁷
Arachidonoylcarnitine*, AGE, SEX, GLUC, CRE, APT, AHT	1.10·10 ⁻⁵ (1.79·10 ⁻⁷ - 7.02·10 ⁻⁴)	6.66·10 ⁻⁸
Aspartate, AGE, SEX, GLUC, CRE, RAN, STAT	4.31·10 ⁻⁴ (2.10·10 ⁻⁵ - 9.07·10 ⁻³)	6.17·10 ⁻⁷
Arginine, AGE, SEX, GLUC, CRE, RAN	2.70·10 ⁷ (1.80·10 ⁴ - 4.05·10 ¹⁰)	4.00·10 ⁻⁶
Bilirubin, AGE, SEX, GLUC, CRE, STE	2.30 (1.78 – 2.97)	2.21·10 ⁻¹⁰
Biliverdin, AGE, SEX, GLUC, CRE, RAN, AHT	3.27·10 ⁵ (2.39·10 ³ – 4.49·10 ⁷)	4.25·10 ⁻⁷
Calcitroic acid, AGE, SEX, GLUC, CRE, NSAIDs, AHT	5.39·10 ³ (139.60 – 2.08·10 ⁵)	4.00·10 ⁻⁶
Cholestane-3,7,12,24,25-pentol, AGE, SEX, GLUC, CRE	3,04·10 ⁻³ (3.32·10 ⁻³ – 0.28)	2.91·10 ⁻⁷
Cys-Gly disulfide, AGE, SEX, GLUC, CRE	1.70·10 ⁻⁵ (2.22·10 ⁻⁷ - 1.29·10 ⁻³)	6.65·10 ⁻⁷
Docosapentaenoic acid, AGE, SEX, GLUC, CRE, STE, AHT	1.13·10 ⁴ (340.9 - 3.73·10 ⁵)	1.73·10 ⁻⁷
Dodecenedioic acid, AGE, SEX, GLUC, CRE	150.65 (18.62 – 1.21·10 ³)	3.00·10 ⁻⁶
Eicosenoic acid, AGE, SEX, GLUC, CRE, RAN	2.59·10 ³ (111.99 – 5.97·10 ⁴)	9.30·10 ⁻⁷
Glu-Ala*, AGE, SEX, GLUC, CRE, RAN, STAT	5.60·10 ³ (1.92·10 ² - 1.64·10 ⁵)	5.35·10 ⁻⁷
Glu-Arg, AGE, SEX, GLUC, CRE	1.52·10 ⁴ (5.22·10 ² - 4.43·10 ⁵)	2.19·10 ⁻⁸
Glucose, AGE, SEX, GLUC, CRE, APT, RAN	10.75 (3.16 – 36.52)	1.42·10 ⁻⁴
HETE, AGE, SEX, GLUC, CRE, PPI, AHT	2.00 (1.60 – 2.50)	2.22·10 ⁻⁹
Homo 2-aconitate*, AGE, SEX, GLUC, CRE, RAN, AHT	1.33·10 ⁻¹⁷ (0 - 0.52)	4.62·10 ⁻²
Indoleacetyl glutamine, AGE, SEX, GLUC, CRE, RAN	61.80 (9.72 - 392.56)	1.20·10 ⁻⁵
Lactate, AGE, SEX, GLUC, CRE, RAN, AHT	4.50·10 ⁻²⁵ (0 - 6.33)	5.78·10 ⁻²
Leu-Phe, AGE, SEX, GLUC, CRE, RAN, AHT	2.44 (1.81 – 3.27)	3.69·10 ⁻⁹
Linolenyl carnitine, AGE, SEX, GLUC, CRE, STE	3.50·10 ⁻⁴ (1.71·10 ⁻³ – 7.19·10 ⁻⁵)	2.45·10 ⁻⁷
Linoleyl carnitine, AGE, SEX, GLUC, CRE, STE, RAN	4.00·10 ⁻⁶ (4.57·10 ⁻⁸ – 3.45·10 ⁻⁴)	4.80·10 ⁻⁸
LPA (13:0), AGE, SEX, GLUC, CRE, RAN	0.29 (0.17 – 0.48)	2.00·10 ⁻⁶

MG (20:00), AGE, SEX, GLUC, CRE, RAN	0.39 (0.30 – 0.51)	$5.47 \cdot 10^{-12}$
Ornithine, AGE, SEX, GLUC, CRE	$1.73 \cdot 10^{-7}$ ($1.09 \cdot 10^{-9}$ - $2.74 \cdot 10^{-5}$)	$1.71 \cdot 10^{-9}$
PG (20:2), AGE, SEX, GLUC, CRE	$1.00 \cdot 10^{-3}$ ($5.50 \cdot 10^{-5}$ – 0.02)	$3.00 \cdot 10^{-6}$
PG (28:0), AGE, SEX, GLUC, CRE, APT, DIU	$2.36 \cdot 10^3$ ($1.28 \cdot 10^2$ – $4.37 \cdot 10^4$)	$1.82 \cdot 10^{-7}$
Phenylglucuronide, AGE, SEX, GLUC, CRE, STE, DIU	1.94 (1.54 - 2.44)	$2.72 \cdot 10^{-8}$
Pyranose, AGE, SEX, GLUC, CRE, APT, RAN	13.40 (3.80 – 47.18)	$5.30 \cdot 10^{-5}$
Pyruvate, AGE, SEX, GLUC, CRE, RAN, AHT	$1.02 \cdot 10^{-4}$ ($3.35 \cdot 10^{-6}$ - $3.13 \cdot 10^{-3}$)	$1.41 \cdot 10^{-7}$
Ser-Hyp*, AGE, SEX, GLUC, CRE, RAN, STAT	$5.60 \cdot 10^3$ ($1.92 \cdot 10^2$ - $1.64 \cdot 10^5$)	$5.35 \cdot 10^{-7}$
Sphingosine-1-phosphate, AGE, SEX, GLUC, CRE, STE, AHT	$7.20 \cdot 10^{-5}$ ($1.39 \cdot 10^{-6}$ – $3.71 \cdot 10^{-3}$)	$2.00 \cdot 10^{-7}$
Stearoyl-tyrosine*, AGE, SEX, GLUC, CRE, APT, AHT	$1.10 \cdot 10^{-5}$ ($1.79 \cdot 10^{-7}$ - $7.02 \cdot 10^{-4}$)	$6.66 \cdot 10^{-8}$
Suberylglycine, AGE, SEX, GLUC, CRE, STE, STAT	179.24 (0 – $1.79 \cdot 10^{308}$)	0.996
Succinylacetoacetate*, AGE, SEX, GLUC, CRE, RAN, AHT	1.3310^{-17} (0 - 0.52)	$4.62 \cdot 10^{-2}$
Trp-Phe, AGE, SEX, GLUC, CRE, RAN	$8.70E \cdot 10^4$ ($1.51 \cdot 10^3$ - $4.99 \cdot 10^6$)	$3.72 \cdot 10^{-8}$
Ursodeoxycholic acid 3-sulfate, AGE, SEX, GLUC, CRE, STAT, RAN	6.92 (2.21 – 21.69)	$9.11 \cdot 10^{-4}$
Ursodeoxycholic acid, AGE, SEX, GLUC, CRE, STAT	1.94 (1.55 – 2.41)	$4.58 \cdot 10^{-9}$
Vaccenyl/Elaidic carnitine, AGE, SEX, GLUC, CRE, RAN	$1.88 \cdot 10^{-3}$ ($1.44 \cdot 10^{-4}$ - 0.03)	$2.00 \cdot 10^{-6}$

IC: intervalo de confianza; GLUC: glucemia; CRE: Creatinina; RAN: Ranitidina; STAT: estatinas; STE: esteroides; AHT: antihipertensivo; APT: antiplaquetario; PPI: inhibidor de la bomba de protones; DIU: diurético; APAP: paracetamol; NSAIDs: Antiinflamatorio no esteroideo.

Además, se realizaron ambos análisis, MPA y MSEA, con los 48 metabolitos que mostraban mayor potencial diagnóstico y confirmaron que la biosíntesis de arginina, el metabolismo de arginina y prolina, la biosíntesis de aminoacil-tRNA, el ciclo del citrato, el metabolismo del piruvato y glutatión y, el ciclo de la urea fueron las vías desreguladas más relevantes en los pacientes con NET.

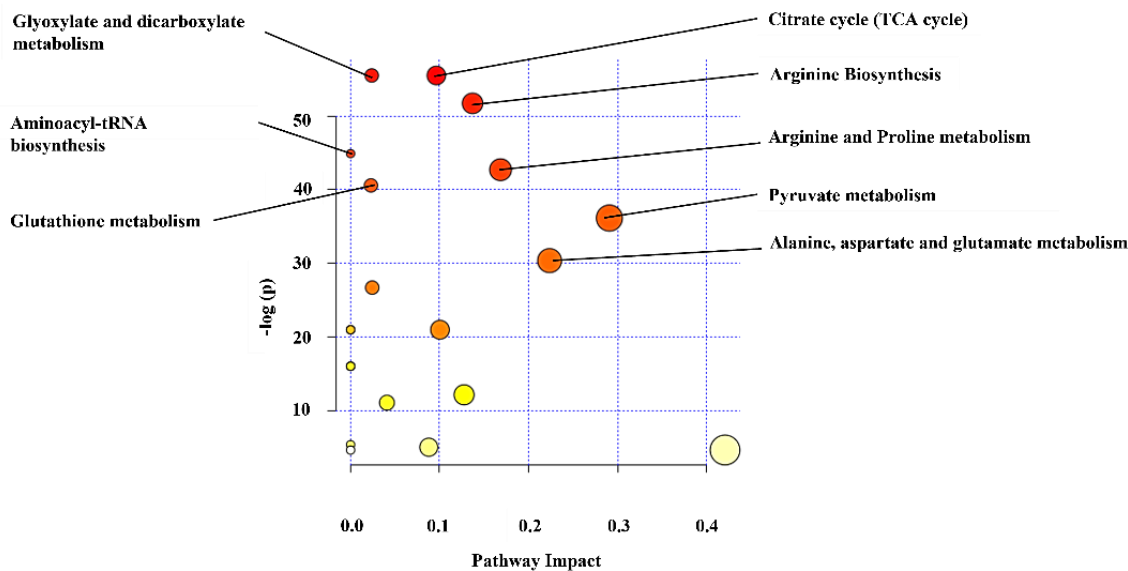
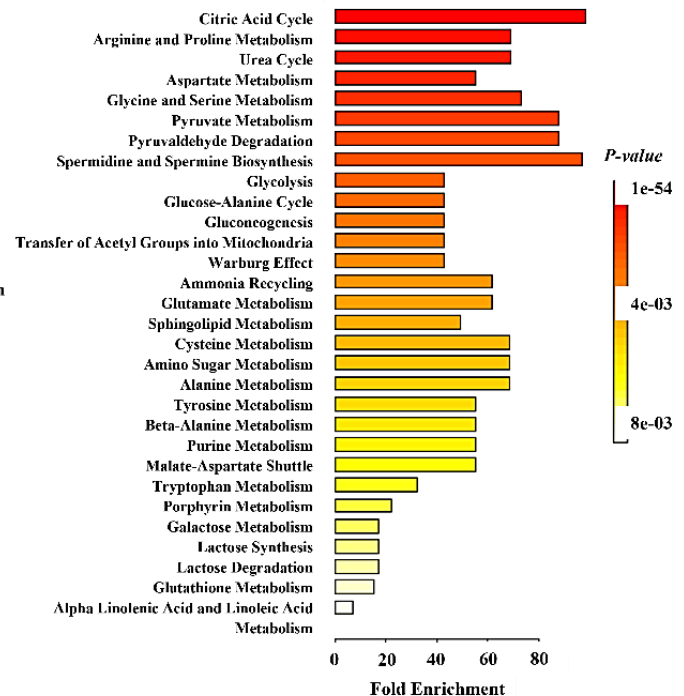
A**B.**

Figura 8. Análisis de las rutas metabólicas implicadas de los 48 metabolitos con potencial diagnóstico. **A.** Análisis de las rutas metabólicas (MPA) enriquecidas significativamente ($FDR < 0.05$) por la disponibilidad de los metabolitos seleccionados ($n = 48$) en el plasma de pacientes con NET. Las vías más enriquecidas entre las 23 significativas fueron: ciclo del citrato (ciclo TCA) ($FDR = 3.8 \cdot 10^{-55}$); metabolismo de glioxilato y dicarboxilato ($FDR = 3.8 \cdot 10^{-55}$); biosíntesis de arginina ($FDR = 1.53 \cdot 10^{-51}$); biosíntesis de aminoacyl-tRNA ($FDR = 7.46 \cdot 10^{-45}$); metabolismo de arginina y prolina ($FDR = 9.23 \cdot 10^{-43}$); metabolismo del glutatión ($FDR = 1.02 \cdot 10^{-40}$); metabolismo del piruvato ($FDR = 2.37 \cdot 10^{-36}$); y el metabolismo de alanina, aspartato y glutamato ($FDR = 1.24 \cdot 10^{-30}$). **B.** Análisis de enriquecimiento del conjunto de metabolitos (MSEA) (seleccionados, $n = 48$) en el plasma de pacientes NET. El eje X representa el enriquecimiento de cada conjunto de metabolitos y el color de las barras indica el p valor. Los resultados muestran que aparecían enriquecidos significativamente 30 conjuntos de metabolitos ($FDR < 0.05$). El ciclo del ácido cítrico ($Q = 68.138$), el metabolismo de la arginina y prolina ($Q = 47.931$), el ciclo de la urea ($Q = 47.808$), el metabolismo del aspartato ($Q = 38.216$) y el metabolismo de la glicina y la serina ($Q = 50.931$) fueron los conjuntos metabólicos más enriquecidos.

4. Validación

Con el objetivo de identificar aquellos metabolitos con mayor potencial diagnóstico, se procedió con la validación de los resultados procedentes del estudio metabolómico no dirigido multiplataforma de las muestras de plasma. Una vez seleccionado un conjunto de metabolitos resultantes de los análisis no dirigidos, la validación analítica confirmaría la identidad y la concentración de los potenciales biomarcadores a través de una técnica diferente.

De la figura 5A (Heatmap) se puede observar la capacidad de ciertos metabolitos para discriminar a los pacientes con NET de los individuos sin cáncer, por lo que se aplicó el análisis de la curva ROC, ampliamente considerado como el método más objetivo y estadísticamente válido para la evaluación del desempeño de posibles biomarcadores. Una curva ROC muestra cómo cambian la sensibilidad y la especificidad a medida que varía el límite de decisión de clasificación en el rango de puntuaciones de biomarcadores disponibles. En la interpretación de las curvas ROC se utiliza una métrica conocida como área bajo la curva (AUC) que se puede interpretar como la probabilidad de que el biomarcador clasifique un caso positivo elegido al azar más alto que uno negativo elegido al azar (59).

En cuanto a los criterios de selección se utilizaron dos de las herramientas principales para este propósito: $VIP > 1$ en el modelo OPLS-DA y $AUC > 0.85$ en el análisis ROC que proporcionó una lista de 46 candidatos para la validación. Estos fueron estudiados de acuerdo con su estabilidad, naturaleza, posibilidad de ser analizadas y cuantificadas a través de LC-QqQ-MS/MS, disponibilidad de estándar comercial, etc. Como resultado, se propuso el análisis de los siguientes 13 candidatos biomarcador (Tabla 8).

Tabla 8. Resumen de los posibles candidatos biomarcador elegidos para la validación.

Metabolito	Nivel de confianza en la anotación	Plataforma analítica	% de cambio	Adj AUC	Modelo	
					exp (B)	p valor
Arginine	2	CE-MS/LC-MS(+)	243	0.986	$2.70 \times 10^{+07}$	4.00×10^{-06}
Biliverdin	3	LC-MS(-)	300	0.975	$3.27 \times 10^{+05}$	4.25×10^{-07}
Linoleoylcarnitine	3	LC-MS(+)	-59	0.962	4.00×10^{-06}	4.80×10^{-08}
5-Hydroxyindoleacetic acid	2	LC-MS(+)	Cancer group	0.935	3.83	7.00×10^{-06}
Hydroxydodecanoic acid	2	LC-MS(-)	Cancer group	0.922	3.04	3.07×10^{-12}
Bilirubin	2	LC-MS(-)/LC-MS(+)	674	0.896	2.30	2.21×10^{-10}
Hydroxydodecanedioic acid	2	LC-MS(-)	Cancer group	0.873	2.31	8.84×10^{-12}
1-Methyladenosine	2	CE-MS	41	0.872	$2.63 \times 10^{+07}$	1.49×10^{-07}
HETE	2	LC-MS(-)	Cancer group	0.868	2.00	2.22×10^{-09}
Sphingosine-1-phosphate	2	LC-MS(+)	-30	0.856	7.20×10^{-05}	2.00×10^{-06}
Ursodeoxycholic acid 3-sulfate	3	LC-MS(-)	130	0.853	6.92	9.11×10^{-04}
Oleoylcarnitine	2	LC-MS(+)	-41	0.852	1.88×10^{-03}	2.00×10^{-06}
Ursodeoxycholic acid	2	LC-MS(-)	188	0.851	1.94	4.58×10^{-09}

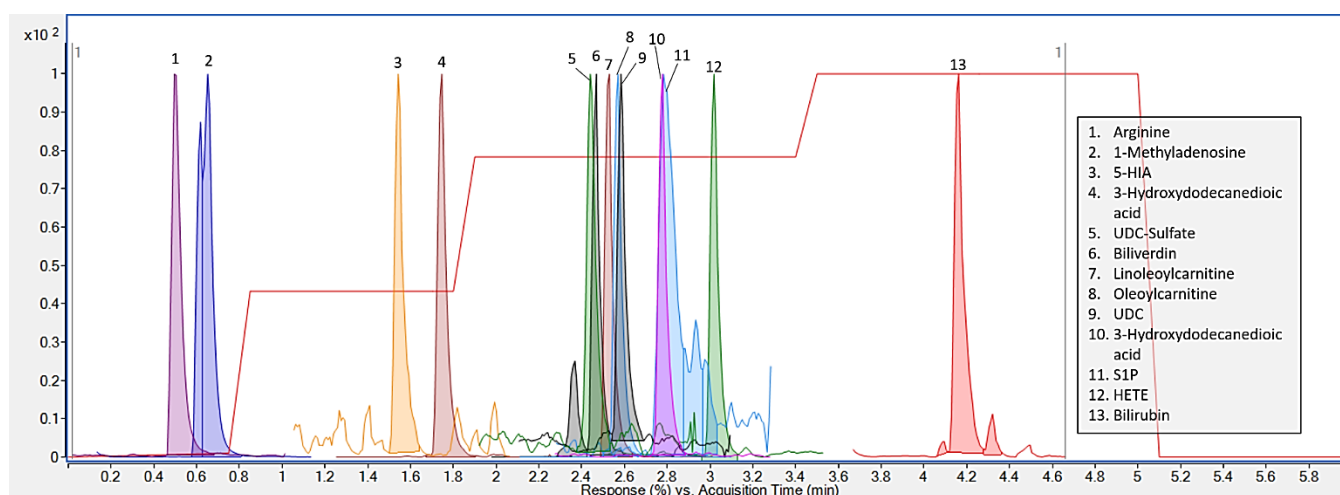


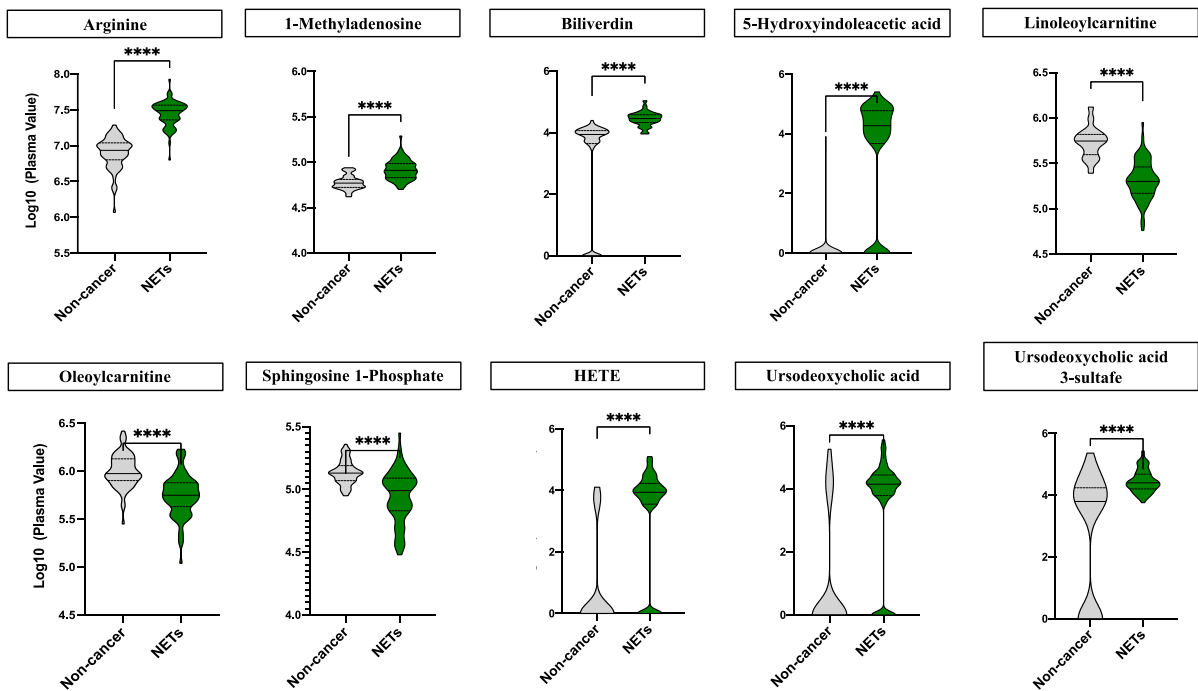
Figura 9. Cromatograma de los iones extraídos correspondientes a las transiciones fijadas para cada metabolito. La línea roja representa el gradiente de fase B diseñado para obtener la separación cromatográfica.

Los estándares elegidos se pidieron a distintas casas comerciales. Con ellos se diseñó el método más adecuado de separación cromatográfica y, junto con una extensa revisión bibliográfica y la ayuda de Optimizer software de Agilent se establecieron las transiciones más adecuadas para cada metabolito, además de, las energías de colisión y el voltaje del fragmentor requerido. Dado que los metabolitos individualmente tienen iones precursores y productos específicos, conocidos como transición, se asegura la identidad del metabolito, especialmente cuando se combina con un tiempo de retención cromatográfico conocido. Sin embargo, estas transiciones específicas, los tiempos de retención, el rango de concentración dinámica y los parámetros de colisión para cada metabolito deben definirse previamente mediante el análisis experimental. El método de separación cromatográfica finalmente optimizado se representa en la Figura 9, donde se puede apreciar que finalmente se logró llegar a un método corto, apenas 6 min de análisis, con la mejor separación posible.

Finalmente, se seleccionó un enfoque de MRM dinámico (dMRM) que ofrece varias ventajas sobre el método de MRM convencional, generando segmentos de tiempo para muestras complejas con niveles bajos de los analitos de interés. El algoritmo utilizado por el dMRM utiliza el tiempo de elución para cada analito y permite que los segmentos de transición de MRM se construyan de forma dinámica a lo largo del análisis de LC-MS. Con este enfoque se monitoriza cada transición exclusivamente en torno a su tiempo de elución de LC esperado, en lugar de adquirir transiciones en un segmento de tiempo como en los métodos de MRM convencionales. De esta forma se maximiza el número de transiciones iónicas monitorizadas en un solo método sin comprometer el tiempo de permanencia (*Dwell*) y la sensibilidad. Datos experimentales como el tiempo de retención o la transición utilizada para la cuantificación y, las concentraciones obtenidas por grupo junto con el *p* valor, el % de cambio o el resultado de r^2 del ajuste de la regresión lineal para cada metabolito de este nuevo análisis se resumen en la Tabla 9.

Arginina, 1-metiladenosina, biliverdina, ácido 5-hidroxiindolacético, linoleoilcarnitina, oleoilcarnitina, esfingosina-1-fosfato, ácido 15-hidroxi-eicosatetraenoico, ácido ursodesoxicólico y ácido ursodesoxicólico 3-sulfato fueron confirmados como posibles biomarcadores de NET, mientras que no fue posible validar bilirrubina, ácido 3-hidroxidodecanodioico y el 3-hidroxidodecanoico, probablemente debido a su inestabilidad. La distribución de la abundancia de cada metabolito en el plasma y su correspondiente curva ROC de los metabolitos con potencial diagnóstico entre pacientes NETs y no-cáncer se muestra en la Figura 10.

A



B

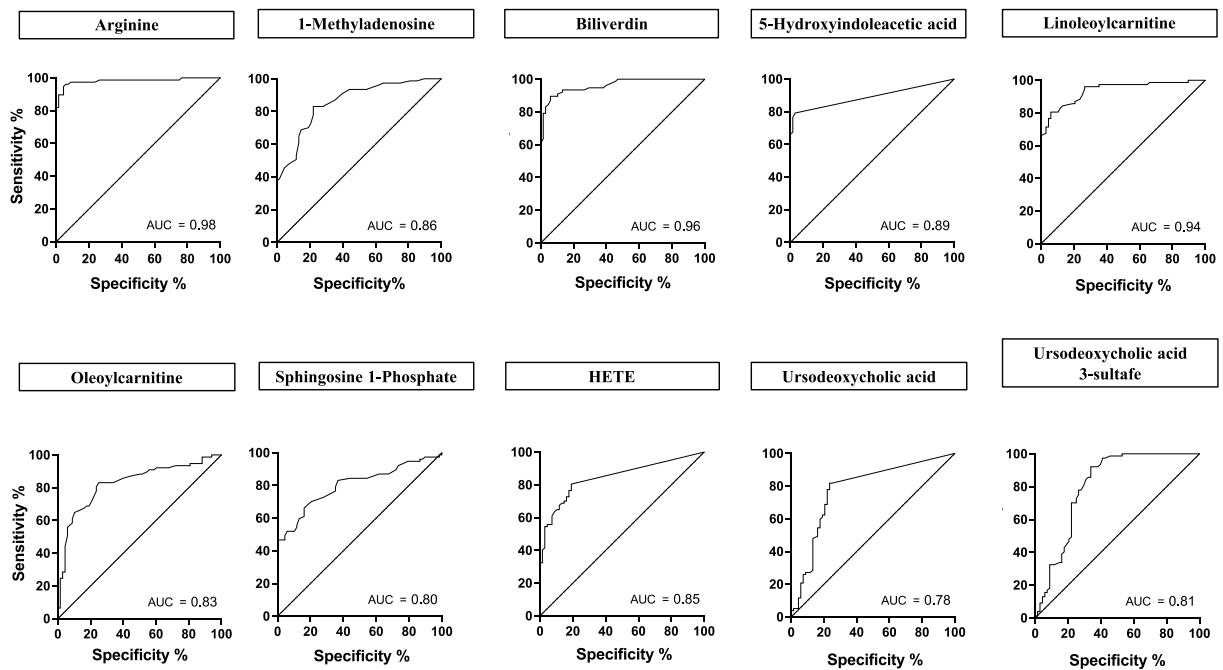


Figura 10. **A.** Abundancia plasmática de los 10 metabolitos validados con potencial diagnóstico en NET. Se realizaron diagramas de violín y la prueba de la U de Mann-Whitney, no paramétrica, para evaluar la abundancia y distribución de los 10 metabolitos validados en pacientes con NET ($n = 76$) y en individuos sin cáncer ($n = 68$). Los valores plasmáticos logarítmicos en pacientes con NET se muestran en verde, mientras que los individuos sin cáncer se representan en gris. Las líneas continuas corresponden a los valores medios y las líneas punteadas se relacionan

con los cuartiles. En general, todos los metabolitos validados mostraron diferencias significativas (****; $p < 0.0001$) entre individuos con NET y sin cáncer. **B.** Curvas ROC de los 10 metabolitos validados con potencial biomarcador en NET. Las curvas se construyeron basándose en el análisis del área bajo la curva (AUC) de pacientes ($n = 76$) y de individuos sin cáncer ($n = 68$). Los puntos de corte óptimos se seleccionaron de acuerdo con la maximización del índice de Youden. En general, todos los metabolitos validados mostraron un alto potencial de biomarcadores con $AUC > 0.75$.

Tabla 9. Validación de posibles biomarcadores de diagnóstico de NET. Esta tabla resume los datos experimentales del análisis LC-QqQ-MS/MS de los 13 metabolitos diferenciales entre pacientes con NET (n = 76) y pacientes sin cáncer (n = 68). El RT, tiempo de retención y los datos de transición utilizados para la cuantificación de cada metabolito, así como la concentración para cada grupo, el valor de p (prueba t de Student) y el % de cambio se muestran para cada metabolito. En total se validaron 10 metabolitos como biomarcadores de diagnóstico de NET.

Compuesto	Formula Molecular	RT (min)	Transición	Polaridad	p valor	% de cambio	Concentración NETs (ng mL ⁻¹)	Concentración no-cancer (ng mL ⁻¹)	Recta de calibrado	r
Arginine	C ₆ H ₁₄ N ₄ O ₂	0.521	175.1 → 70.1	+	9.67x10 ⁻⁰⁷	341	268 ± 76	60.8 ± 6.8	y = 145.7x + 48.9	0.9993
1-Methyladenosine	C ₁₁ H ₁₅ N ₅ O ₄	0.602	282.1 → 150.0	+	1.01x10 ⁻¹⁹	24	6.73 ± 0.18	5.44 ± 0.16	y = 2228.6x - 1063.7	0.9991
Bilirubin	C ₃₃ H ₃₆ N ₄ O ₆	4.162	585.3 → 299.1	+	NP	NP	NP	NP	y = 25.801x + 526.49	0.96
Biliverdin	C ₃₃ H ₃₄ N ₄ O ₆	2.376	583.3 → 297.1	+	5.12x10 ⁻¹⁴	179	97 ± 12	34.8 ± 5.6	y = 8.6391x + 2.013	0.999
5-Hydroxyindoleacetic acid	C ₁₀ H ₉ NO ₃	1.540	190.1 → 144.0	-	1.40x10 ⁻⁰⁶	407	152 ± 44	30.0 ± 8.0	y = 2.1188x - 9.8133	0.9998
3-Hydroxydodecanedioic acid	C ₁₂ H ₂₂ O ₅	1.746	245.1 → 185.2	-	NP	NP	NP	NP	y = 30.17635x + 0.23980	0.994
Linoleoylcarnitine	C ₂₅ H ₄₅ NO ₄	2.529	424.3 → 85.0	+	2.00x10 ⁻¹⁶	-53	81.2 ± 7.6	173 ± 16	y = 784.09x - 3404.6	0.9999
Oleoylcarnitine	C ₂₅ H ₄₇ NO ₄	2.570	426.4 → 85.0	+	1.07x10 ⁻⁰⁹	-50	120 ± 15	238 ± 31	y = 1003.7x - 6622.4	0.9998
Sphingosine-1-phosphate	C ₁₈ H ₃₈ NO ₅ P	2.773	380.3 → 82.1	+	4.70x10 ⁻¹³	-27	353 ± 25	482 ± 18	y = 107.95x - 1072.3	0.9994
3-Hydroxydodecanoic acid	C ₁₂ H ₂₄ O ₃	2.778	215.2 → 59.1	-	NP	NP	NP	NP	y = 70.765x + 120.68	0.9999
HETE	C ₂₀ H ₃₂ O ₃	3.018	319.2 → 301.2	-	2.13x10 ⁻⁰²	1319	21 ± 16	1.48 ± 0.64	y = 25.441x	0.9997
Ursodeoxycholic acid	C ₂₄ H ₄₀ O ₄	2.583	391.2 → 391.2	-	4.42x10 ⁻⁰²	90	91 ± 36	48 ± 18	y = 117.01x + 837.82	0.9998
Ursodeoxycholic acid 3-sulfate	C ₂₄ H ₄₀ O ₇ S	2.448	475.1 → 475.1	-	1.86x10 ⁻⁰⁵	27	33.2 ± 2.3	26.2 ± 2.0	y = 129.62x - 342.57	0.9999

NP: no procede

5. Discusión

El objetivo fundamental de este estudio fue proporcionar, por primera vez, un perfil metabólico completo de los NET, evaluado a través del análisis multiplataforma no dirigido de muestras de plasma de una cohorte grande de pacientes. Además, se ha identificado una firma metabólica distintiva que supone un importante potencial clínico.

Con la integración de los datos adquiridos a través de las 3 plataformas, CE-MS, LC-MS y GC-MS, 155 compuestos anotados presentaron diferencias significativas entre NET y pacientes sin cáncer. El análisis ROC y OPLS-DA reveló 49 metabolitos específicos con potencial diagnóstico, 48 de los cuales contribuyeron significativamente al modelo tras el ajuste con posibles variables de confusión, entre ellas, el sexo, la edad, niveles de glucosa en sangre o el uso concomitante de otros medicamentos.

Actualmente, en la práctica clínica, se lleva a cabo la evaluación bioquímica de varios dipéptidos como diagnóstico o seguimiento de pacientes con NET. Existen, además, otros marcadores tumorales como la cromogranina A (CgA) o la enolasa neuronal específica, los cuales no son recomendables para el cribado debido a que estos pueden verse alterados por otras condiciones, sean o no de origen oncológico (60). A tal efecto, los metabolitos identificados en nuestro estudio, con alta predicción diagnóstica, pueden proporcionar una herramienta muy valiosa para la detección específica de NET. Algunos de ellos están relacionados con vías habituales de sobreexpresión en cáncer, como el ciclo celular y la apoptosis, y otras más especializadas en la señalización de NET, como el metabolismo del triptófano, la angiogénesis o la vía mTOR. Asimismo, se han identificado 32 vías metabólicas que presentan algún tipo de alteración en NET relacionadas con el ciclo del TCA y el metabolismo de arginina, glutatión o piruvato, con distintas implicaciones en la oncogénesis.

Como se citó anteriormente, en lo reportado hasta la fecha, existen dos estudios pequeños que han explorado parcialmente el perfil metabólico de los pacientes con NET, sin embargo, ninguno de ellos empleó, como fuente principal de análisis, muestras de plasma. Kinross *y col.* (61) realizaron un estudio piloto de 28 muestras de orina de pacientes NET gastroenteropáticos utilizando H-NMR. Se consiguió identificar distintos fenotipos metabólicos, según la localización del tumor primario (intestino delgado *versus* NEN pancreático), por importantes alteraciones en el metabolismo del ácido hipúrico. Las limitaciones fundamentales de este estudio fueron el tamaño de la cohorte, el rango de edad de los individuos empleados como tumorales y como control, sin tener en cuenta otros factores que pueden generar confusión como el sexo, la función renal o la

concomitancia de fármacos. Imperiale y *col.* (32) evaluaron, por su parte, la huella metabólica a través H¹ NMR de NET de tumores primarios del intestino delgado (n = 46) y de metástasis en el hígado (n = 18). Treinta muestras normales de intestino delgado e hígado completaron la cohorte. Los resultados sugirieron alteraciones en vías metabólicas cruciales como el ciclo del ácido tricarboxílico (TCA). En relación con esto, esta vía también se encontró alterada en nuestros datos con la elevada disponibilidad en el plasma NET de compuestos citrato/isocitrato (+64%). Además, esta hipótesis se encuentra respaldada por los altos niveles de glucosa (+174%), glutamina (+69%) y ácidos grasos, protagonistas también del ciclo de TCA.

Otros estudios realizados en tumores dependientes de hormonas similares a los NET, como el cáncer de próstata, recalcan la alteración significativa del metabolismo intermediario en la disfunción maligna. Específicamente se ha destacado la anomalía de las células productoras de citrato que durante la oncogénesis se transforman en células malignas que oxidan ese citrato. El proceso bioquímico de oxidación del citrato en el TCA generando ATP está próximamente relacionado con vías celulares bioenergéticas, crecimiento celular, apoptosis, lipogénesis y angiogénesis (62). Es por tanto que, nos encontramos ante una vía metabólica que es pieza central de la respiración celular y, que está estrictamente regulada para cumplir con los requisitos de bioenergía celular, biosíntesis y equilibrio redox; y aunque varios tipos de tumores se caracterizan por una marcada desregulación de las enzimas TCA (63), su participación en el metabolismo del cáncer sigue sin comprenderse completamente.

Los niveles de citrato/isocitrato elevados en plasma también podrían ser una explicación de una alteración en la lipogénesis (64). En nuestros datos, el lipidoma característico de los pacientes NET está representado por el enriquecimiento de glicerofosfolípidos, ácidos grasos y esteroides. Además, se han encontrado en particular niveles aumentados de lisoglicerofosfolípidos oxidados (oxLPC) en NET, lo que indica fuerte estrés oxidativo, junto con altos niveles de derivados oxidados del ácido araquidónico (HETEs) y una importante disminución en los niveles de carnitinas (55). La L-carnitina es un metabolito crítico en el control del metabolismo de las células cancerosas para el transporte bidireccional entre el citosol y las mitocondrias del acilo graso de cadena larga y la acil CoA (65, 66). Estudios recientes señalan el sistema de carnitina como parte fundamental para la adaptación metabólica de las células cancerosas, obteniendo energía a través de la β -oxidación de lípidos. Por lo tanto, los bajos niveles encontrados en el plasma de pacientes NET podría tener su explicación por una sobreexpresión de este sistema en las mitocondrias tumorales y una regulación

positiva de la β -oxidación. Los HETEs constituyen una amplia variedad de mediadores lipídicos generados a partir del metabolismo del ácido araquidónico por tres vías enzimáticas diferentes (ciclooxigenasa, (COX), lipooxigenasa (LOX) y citocromo P450) y que se encuentran implicados en múltiples procesos fisiológicos y fisiopatológicos. Además de los elevados niveles de estos derivados, se han encontrado también un aumento importante en derivados de los eicosanoides, estrechamente relacionados con las respuestas antiinflamatorias (67).

El colesterol es otro de los compuestos protagonistas en la progresión tumoral. Las células cancerosas adaptan las elevadas necesidades de colesterol intracelular a través de distintos mecanismos que incluyen la producción endógena de colesterol y ácidos grasos, la reducción de su transporte o el aumento en la captación de partículas lipídicas de baja densidad (68). Igualmente, se ha observado un elevado contenido de esteroides en el plasma de los pacientes con NET, principalmente ácidos biliares y derivados del colesterol como la vitamina D, la biliverdina y la bilirrubina. Esto sugiere una importante dependencia de los NET con el metabolismo de lípidos y esteroides.

La vascularización y la angiogénesis tienen una relevancia particular en el desarrollo y la progresión de los NET (69), y hemos identificado varios metabolitos significativos en nuestro estudio que pueden contribuir al cambio de la angiogénesis. Como ejemplo, se encontró que la arginina presentaba niveles aumentados en los NET, y es la principal fuente de óxido nítrico a nivel endógeno. El NO produce efectos tanto anti- como pro-tumorales, y está profundamente involucrado en la regulación de la angiogénesis, apoptosis, ciclo celular, invasión y metástasis (70). De igual manera, el ácido lisofosfatídico, los HETEs o la biliverdina inducen la angiogénesis por regulación positiva de los factores de crecimiento endotelial vascular A, C (VEGFA, VEGFC), e interleuquinas 1 y 8 (IL-1, IL-8) (71-73). En general, estos hallazgos respaldan, aún más, el papel relevante que desempeña la angiogénesis en la patogenia de los NET.

Centrándonos en las vías más críticas en los NET, destaca la vía mTOR (74); además, se ha demostrado la actividad antiproliferativa de un inhibidor específico de mTOR, everolimus, en estas neoplasias aprobado para el tratamiento de NETs avanzados gastrointestinales, pancreáticos y pulmonares (75, 76). Nuestros resultados muestran niveles aumentados de arginina (+243%) y glutamina (+69%) debido a la patología, que son, junto con otro aminoácido, leucina, estimuladores de mTOR a través del complejo de regulación. Del mismo modo, se han obtenido en nuestros análisis, mayores niveles de fosfatidilcolina, cuya síntesis es promovida por mTOR C1. Curiosamente, la vía mTOR se ha asociado con el cáncer a través de su papel en la

regulación de la dinámica de las poliaminas (77). Los niveles de poliaminas se asocian con una reducción de la apoptosis y, un aumento de la proliferación de células cancerosas y la expresión de genes relacionados con la metástasis, aunque los mecanismos subyacentes a estos efectos no han sido bien definidos (78). La acetilpermidina es una poliamina que fue detectada en nuestro estudio con un aumento ~40%, lo que ilustra la importancia del metabolismo de las poliaminas en NET. Recientemente, Chalisehar y *col.* (79) observaron que el cáncer de pulmón de células escamosas pequeñas (NSCLC), impulsado por la familia de genes MYC, depende preferentemente de vías reguladas por arginina, incluida la biosíntesis de poliaminas y la activación de la vía mTOR. Argininosuccinato sintasa 1 (ASS1) genera arginina en el ciclo de la urea de manera indirecta y, en varios casos de cáncer, es eliminada mediante el silenciamiento epigenético, incluido el NSCLC. Además, este colapso de ASS1 provoca un aumento de la actividad mTOR y la auxotrofia de arginina. Por lo que, sería interesante plantear como estrategia terapéutica, la privación de arginina para los tumores que dependen de esta para su supervivencia (3).

Por último, es importante citar los niveles diferenciales en hipoxantina encontrados (-34%). Niveles bajos de esta purina se han observado en muestras de orina y plasma de pacientes con cáncer, especialmente aquellos que se encuentran en etapas avanzadas de la enfermedad ya que los tejidos hiperproliferativos requieren mayor síntesis de ADN (80). Y, aunque la causa y el efecto de la desregulación de la hipoxantina en los NET no está clara, es posible que se generen alteraciones en el metabolismo de las purinas durante la progresión del tumor. Esta hipótesis podría responder a la correlación inversa encontrada en nuestros datos entre los niveles plasmáticos de hipoxantina y la tasa de proliferación tumoral o índice Ki67 ($r^2 = -0.243$, $p = 0.033$)

Una de las características con mayor potencia del estudio aquí presentado, es que se ha llevado a cabo con una cohorte amplia y homogénea de muestras de plasma, recogidas y analizadas de forma uniforme y prospectiva. Además, cuenta con un conjunto de metabolitos validados en un análisis dirigido, en una plataforma analítica diferente, utilizando la misma cohorte. El objetivo futuro será continuar con el estudio de los resultados, utilizando una cohorte independiente de pacientes con NET y comparar los perfiles metabolómicos de pacientes con tumores exocrinos de origen tisular similar, para poder confirmar el potencial diagnóstico de la firma metabólica seleccionada en los NET. Así mismo, estos podrían complementarse con estudios de genómica, proteómica y/o transcriptómica.

6. Referencias

1. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. *Nature Reviews Cancer*. 2011;11(2):85-95.
2. Suhre K, Shin SY, Petersen AK, Mohny RP, Meredith D, Wägele B, et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature*. 2011;477(7362):54-60.
3. Sullivan LB, Gui DY, Heiden MG. Altered metabolite levels in cancer: implications for tumour biology and cancer therapy. *Nature Reviews Cancer*. 2016;16(11):680-93.
4. Newgard CB. Metabolomics and metabolic diseases: where do we stand? *Cell metabolism*. 2017;25(1):43-56.
5. Oakman C, Tenori L, Claudino WM, Cappadona S, Nepi S, Battaglia A, et al. Identification of a serum-detectable metabolomic fingerprint potentially correlated with the presence of micrometastatic disease in early breast cancer patients at varying risks of disease relapse by traditional prognostic methods. *Ann Oncol*. 2011;22(6):1295-301.
6. Martín-Blázquez A, Díaz C, González-Flores E, Franco-Rivas D, Jiménez-Luna C, Melguizo C, et al. Untargeted LC-HRMS-based metabolomics to identify novel biomarkers of metastatic colorectal cancer. *Scientific Reports*. 2019;9(1):20198.
7. Bertini I, Cacciatore S, Jensen BV, Schou JV, Johansen JS, Kruhøffer M, et al. Metabolomic NMR fingerprinting to identify and predict survival of patients with metastatic colorectal cancer. *Cancer research*. 2012;72(1):356-64.
8. Deja S, Porebska I, Kowal A, Zabek A, Barg W, Pawelczyk K, et al. Metabolomics provide new insights on lung cancer staging and discrimination from chronic obstructive pulmonary disease. *Journal of pharmaceutical and biomedical analysis*. 2014;100:369-80.
9. Kell DB. Metabolomics and systems biology: making sense of the soup. *Current opinion in microbiology*. 2004;7(3):296-307.
10. ter Kuile BH, Westerhoff HV. Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Letters*. 2001;500(3):169-71.
11. Fell D, Cornish-Bowden A. *Understanding the control of metabolism*: Portland press London; 1997.
12. Mendes P, Kell DB, Westerhoff HV. Why and when channelling can decrease pool size at constant net flux in a simple dynamic channel. *Biochimica et Biophysica Acta (BBA)-General Subjects*. 1996;1289(2):175-86.
13. Griffin JL, Shockcor JP. Metabolic profiles of cancer cells. *Nature Reviews Cancer*. 2004;4(7):551-61.
14. Delcore R, Friesen SR. Embryologic concepts in the APUD system. *Semin Surg Oncol*. 1993;9(5):349-61.
15. Andrew A, Kramer B, Rawdon BB. The origin of gut and pancreatic neuroendocrine (APUD) cells—the last word? *The Journal of Pathology*. 1998;186(2):117-8.

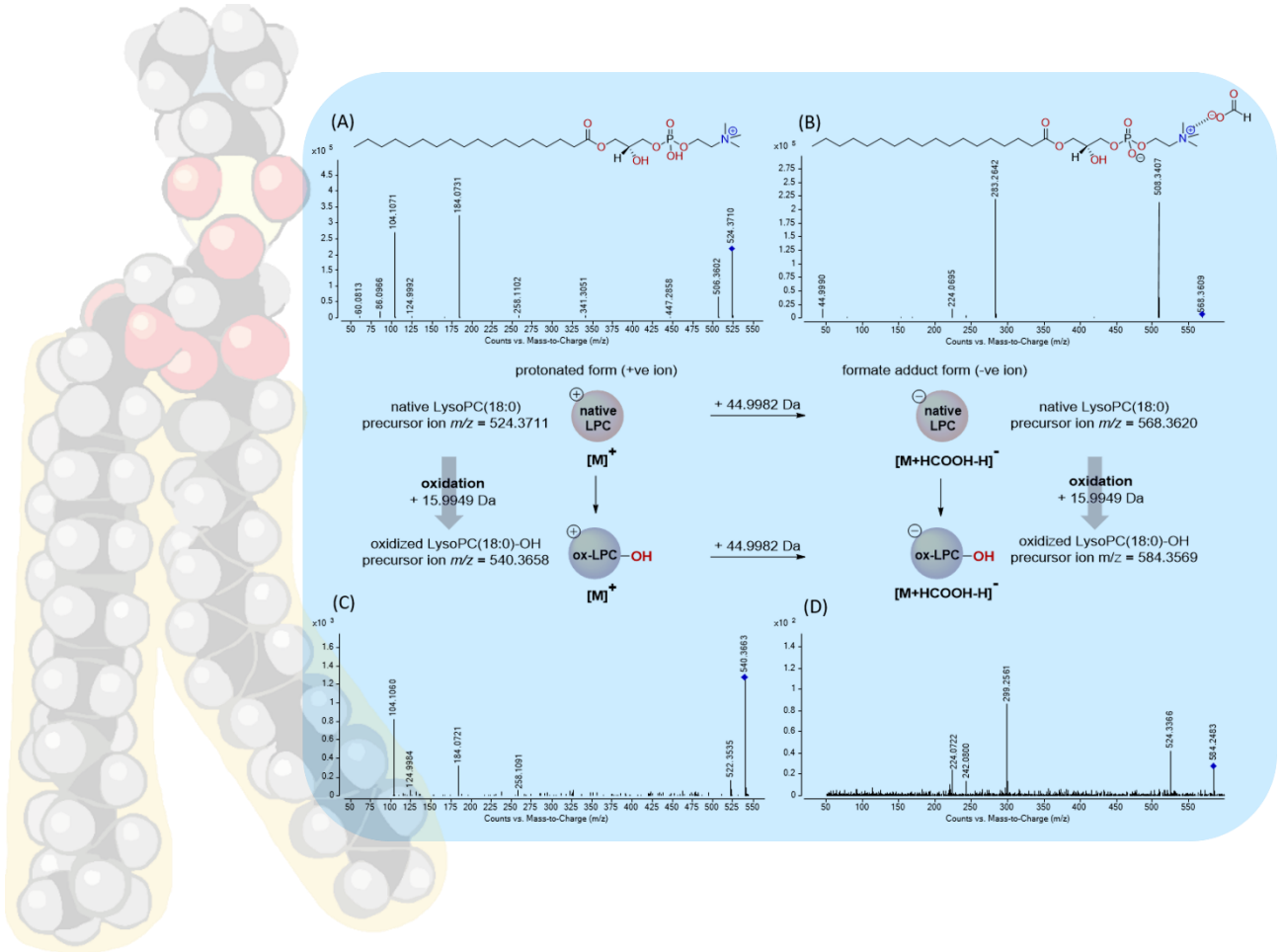
16. Johnson DE, Georgieff MK. Pulmonary neuroendocrine cells. Their secretory products and their potential roles in health and chronic lung disease in infancy. *Am Rev Respir Dis*. 1989;140(6):1807-12.
17. Abrahamsson PA. Neuroendocrine cells in tumour growth of the prostate. *Endocr Relat Cancer*. 1999;6(4):503-19.
18. González-Flores E, Serrano R, Sevilla I, Viúdez A, Barriuso J, Benavent M, et al. SEOM clinical guidelines for the diagnosis and treatment of gastroenteropancreatic and bronchial neuroendocrine neoplasms (NENs)(2018). *Clinical and Translational Oncology*. 2019;21(1):55-63.
19. Hofland J, Kaltsas G, de Herder WW. Advances in the Diagnosis and Management of Well-Differentiated Neuroendocrine Neoplasms. *Endocr Rev*. 2020;41(2):371-403.
20. Dasari A, Shen C, Halperin D, Zhao B, Zhou S, Xu Y, et al. Trends in the incidence, prevalence, and survival outcomes in patients with neuroendocrine tumors in the United States. *JAMA Oncology*. 2017;3(10):1335-42.
21. Nuñez-Valdovinos B, Carmona-Bayonas A, Jimenez-Fonseca P, Capdevila J, Castaño-Pascual Á, Benavent M, et al. Neuroendocrine tumor heterogeneity adds uncertainty to the World Health Organization 2010 Classification: Real-world data from the Spanish Tumor Registry (R-GETNE). *The oncologist*. 2018;23(4):422.
22. Stevenson M, Lines KE, Thakker RV. Molecular genetic studies of pancreatic neuroendocrine tumors: New therapeutic approaches. *Endocrinology and Metabolism Clinics*. 2018;47(3):525-48.
23. Brandi ML, Gagel RF, Angeli A, Bilezikian JP, Beck-Peccoz P, Bordi C, et al. Consensus: guidelines for diagnosis and therapy of MEN type 1 and type 2. *The Journal of Clinical Endocrinology & Metabolism*. 2001;86(12):5658-71.
24. Thakker RV. Multiple endocrine neoplasia type 1 (MEN1) and type 4 (MEN4). *Molecular and cellular endocrinology*. 2014;386(1-2):2-15.
25. Rednam SP, Erez A, Druker H, Janeway KA, Kamihara J, Kohlmann WK, et al. Von Hippel–Lindau and hereditary pheochromocytoma/paraganglioma syndromes: clinical features, genetics, and surveillance recommendations in childhood. *Clinical Cancer Research*. 2017;23(12):e68-e75.
26. Jett K, Friedman JM. Clinical and genetic aspects of neurofibromatosis 1. *Genet Med*. 2010;12(1):1-11.
27. Larson A, Hedgire S, Deshpande V, Stemmer-Rachamimov A, Harisinghani M, Ferrone C, et al. Pancreatic neuroendocrine tumors in patients with tuberous sclerosis complex. *Clinical Genetics*. 2012;82(6):558-63.
28. Scarpa A, Chang DK, Nones K, Corbo V, Patch A-M, Bailey P, et al. Whole-genome landscape of pancreatic neuroendocrine tumours. *Nature*. 2017;543(7643):65-71.
29. Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, et al. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science*. 2011;331(6021):1199-203.
30. Karpathakis A, Dibra H, Pipinikas C, Feber A, Morris T, Francis J, et al. Prognostic Impact of Novel Molecular Subtypes of Small Intestinal Neuroendocrine Tumor. *Clin Cancer Res*. 2016;22(1):250-8.

31. Kinross JM, Drymoussis P, Jiménez B, Frilling A. Metabonomic profiling: a novel approach in neuroendocrine neoplasias. *Surgery*. 2013;154(6):1185-92; discussion 92-3.
32. Imperiale A, Poncet G, Addeo P, Ruhland E, Roche C, Battini S, et al. Metabolomics of Small Intestine Neuroendocrine Tumors and Related Hepatic Metastases. *Metabolites*. 2019;9(12):300.
33. Naz S, Garcia A, Rusak M, Barbas C. Method development and validation for rat serum fingerprinting with CE-MS: application to ventilator-induced-lung-injury study. *Analytical and Bioanalytical Chemistry*. 2013;405(14):4849-58.
34. Godzien J, Kalaska B, Adamska-Patruno E, Siroka J, Ciborowski M, Kretowski A, et al. Oxidized glycerophosphatidylcholines in diabetes through non-targeted metabolomics: Their annotation and biological meaning. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2019;1120:62-70.
35. Dudzik D, Zorawski M, Skotnicki M, Zarzycki W, Kozłowska G, Bibik-Malinowska K, et al. Metabolic fingerprint of Gestational Diabetes Mellitus. *Journal of Proteomics*. 2014;103:57-71.
36. Godzien J, Armitage EG, Angulo S, Martinez-Alcazar MP, Alonso-Herranz V, Otero A, et al. In-source fragmentation and correlation analysis as tools for metabolite identification exemplified with CE-TOF untargeted metabolomics. *Electrophoresis*. 2015;36(18):2188-95.
37. Parcher JF, Wang M, Chittiboyina AG, Khan IA. In-source collision-induced dissociation (IS-CID): Applications, issues and structure elucidation with single-stage mass analyzers. *Drug testing and analysis*. 2018;10(1):28-36.
38. Mamani-Huanca M, Gradillas A, Gil de la Fuente A, López-González Á, Barbas C. Unveiling the Fragmentation Mechanisms of Modified Amino Acids as the Key for Their Targeted Identification. *Analytical Chemistry*. 2020;92(7):4848-4857.
39. Dudzik D, Barbas-Bernardos C, García A, Barbas C. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. *Journal of pharmaceutical and biomedical analysis*. 2017;147:149-73.
40. Armitage EG, Godzien J, Alonso-Herranz V, López-González Á, Barbas C. Missing value imputation strategies for metabolomics data. *Electrophoresis*. 2015;36(24):3050-60.
41. Kuligowski J, Sanchez-Illana A, Sanjuan-Herraez D, Vento M, Quintas G. Intra-batch effect correction in liquid chromatography-mass spectrometry using quality control samples and support vector regression (QC-SVRC). *Analyst*. 2015;140(22):7810-7.
42. Wheelock AM, Wheelock CE. Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Molecular BioSystems*. 2013;9(11):2589-96.
43. Barnes RJ, Dhanoa MS, Lister SJ. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*. 1989;43(5):772-7.
44. Gil de la Fuente A, Godzien J, Fernández López M, Rupérez FJ, Barbas C, Otero A. Knowledge-based metabolite annotation tool: CEU Mass Mediator. *J Pharm Biomed Anal*. 2018;154:138-49.

45. Gil-De-La Fuente A, Godzien J, Saugar S, García Carmona R, Badran H, Wishart DS, et al. CEU Mass Mediator 3.0: a metabolite annotation tool. *J Proteome Res.* 2019;18(2):797-802
46. Montenegro-Burke JR, Guijas C, Siuzdak G. METLIN: A Tandem Mass Spectral Library of Standards. *Methods Mol Biol.* 2020;2104:149-63.
47. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform.* 2016;8:3.
48. Chong J, Wishart DS, Xia J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Current Protocols in Bioinformatics.* 2019;68(1):e86.
49. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* 2014;42(Database issue):D199-205.
50. Jewison T, Su Y, Disfany FM, Liang Y, Knox C, Maciejewski A, et al. SMPDB 2.0: big improvements to the Small Molecule Pathway Database. *Nucleic Acids Res.* 2014;42(Database issue):D478-84.
51. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochem Med (Zagreb).* 2016;26(3):297-307.
52. Hofland J, Kaltsas G, de Herder WW. Advances in the Diagnosis and Management of Well-Differentiated Neuroendocrine Neoplasms. *Endocrine Reviews.* 2020;41(2):371-403.
53. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *Journal of The American Society for Mass Spectrometry.* 2016;27(12):1897-905.
54. Fahy E, Subramaniam S, Murphy RC, Nishijima M, Raetz CR, Shimizu T, et al. Update of the LIPID MAPS comprehensive classification system for lipids. *J Lipid Res.* 2009;50 Suppl:S9-14.
55. López-López Á, Godzien J, Soldevilla B, Gradillas A, López-González Á, Lens-Pardo A, et al. Oxidized lipids in the metabolic profiling of neuroendocrine tumors – Analytical challenges and biological implications. *Journal of Chromatography A.* 2020;1625:461233.
56. Gonzalez-Sanchez E, Perez MJ, Nytofte NS, Briz O, Monte MJ, Lozano E, et al. Protective role of biliverdin against bile acid-induced oxidative stress in liver cells. *Free Radical Biology and Medicine.* 2016;97:466-77.
57. Kapitulnik J, Maines M. The Role of Bile Pigments in Health and Disease: Effects on Cell Signaling, Cytotoxicity, and Cytoprotection. *Frontiers in Pharmacology.* 2012;3:136.
58. Stocker R, McDonagh AF, Glazer AN, Ames BN. [31] Antioxidant activities of bile pigments: Biliverdin and bilirubin. *Methods in Enzymology.* 186: Academic Press; 1990. p. 301-9.
59. Xia JG, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics.* 2013;9(2):280-99.

60. Sansone A, Lauretta R, Vottari S, Chiefari A, Barnabei A, Romanelli F, et al. Specific and Non-Specific Biomarkers in Neuroendocrine Gastroenteropancreatic Tumors. *Cancers*. 2019;11(8):1113.
61. Kinross JM, Drymoussis P, Jiménez B, Frilling A. Metabonomic profiling: a novel approach in neuroendocrine neoplasias. *Surgery*. 2013;154(6):1185-92; discussion 92-3.
62. Costello LC, Franklin RB. The Intermediary Metabolism of the Prostate: A Key to Understanding the Pathogenesis and Progression of Prostate Malignancy. *Oncology*. 2000;59(4):269-82.
63. Anderson NM, Mucka P, Kern JG, Feng H. The emerging role and targetability of the TCA cycle in cancer metabolism. *Protein & Cell*. 2018;9(2):216-37.
64. Jiang L, Boufersaoui A, Yang C, Ko B, Rakheja D, Guevara G, et al. Quantitative metabolic flux analysis reveals an unconventional pathway of fatty acid synthesis in cancer cells deficient for the mitochondrial citrate transport protein. *Metabolic Engineering*. 2017;43:198-207.
65. Reuter SE, Evans AM. Carnitine and Acylcarnitines. *Clinical Pharmacokinetics*. 2012;51(9):553-72.
66. Gao B, Lue HW, Podolak J, Fan S, Zhang Y, Serawat A, et al. Multi-Omics Analyses Detail Metabolic Reprogramming in Lipids, Carnitines, and Use of Glycolytic Intermediates between Prostate Small Cell Neuroendocrine Carcinoma and Prostate Adenocarcinoma. *Metabolites*. 2019;9(5):82.
67. Powell WS, Rokach J. Biosynthesis, biological effects, and receptors of hydroxyeicosatetraenoic acids (HETEs) and oxoeicosatetraenoic acids (oxo-ETEs) derived from arachidonic acid. *Biochim Biophys Acta*. 2015;1851(4):340-55.
68. Huang B, Song BL, Xu C. Cholesterol metabolism in cancer: mechanisms and therapeutic opportunities. *Nat Metab*. 2020;2(2):132-41.
69. Ia Rosa S, Uccella S, Finzi G, Albarello L, Sessa F, Capella C. Localization of vascular endothelial growth factor and its receptors in digestive endocrine tumors: Correlation with microvessel density and clinicopathologic features. *Human Pathology*. 2003;34(1):18-27.
70. Keshet R, Erez A. Arginine and the metabolic regulation of nitric oxide synthesis in cancer. *Disease Models & Mechanisms*. 2018;11(8 8):dmm033332.
71. Xu Y. Targeting Lysophosphatidic Acid in Cancer: The Issues in Moving from Bench to Bedside. *Cancers*. 2019;11(10):1523.
72. Panagiotopoulos AA, Kalyvianaki K, Castanas E, Kampa M. Eicosanoids in prostate cancer. *Cancer and Metastasis Reviews*. 2018;37(2):237-43.
73. Chau L-Y. Heme oxygenase-1: emerging target of cancer therapy. *Journal of Biomedical Science*. 2015;22(1):22.
74. Lamberti G, Brighi N, Maggio I, Manuzzi L, Peterle C, Ambrosini V, et al. The Role of mTOR in Neuroendocrine Tumors: Future Cornerstone of a Winning Strategy? *International Journal of Molecular Sciences*. 2018;19(3):747.
75. Yao JC, Shah MH, Ito T, Bohas CL, Wolin EM, Van Cutsem E, et al. Everolimus for Advanced Pancreatic Neuroendocrine Tumors. *New England Journal of Medicine*. 2011;364(6):514-23.

76. Yao JC, Fazio N, Singh S, Buzzoni R, Carnaghi C, Wolin E, et al. Everolimus for the treatment of advanced, non-functional neuroendocrine tumours of the lung or gastrointestinal tract (RADIANT-4): a randomised, placebo-controlled, phase 3 study. *The Lancet*. 2016;387(10022):968-77.
77. Zabala-Letona A, Arruabarrena-Aristorena A, Martín-Martín N, Fernandez-Ruiz S, Sutherland JD, Clasquin M, et al. mTORC1-dependent AMD1 regulation sustains polyamine metabolism in prostate cancer. *Nature*. 2017;547(7661):109-13.
78. Gerner EW, Meyskens FL. Polyamines and cancer: old molecules, new understanding. *Nature Reviews Cancer*. 2004;4(10):781-92.
79. Chalise MD, Wait SJ, Huang F, Ireland AS, Mukhopadhyay A, Lee Y, et al. MYC-Driven Small-Cell Lung Cancer is Metabolically Distinct and Vulnerable to Arginine Depletion. *Clinical Cancer Research*. 2019;25(16):5107.
80. Kim K, Yeo S-G, Yoo BC. Identification of Hypoxanthine and Phosphoenolpyruvic Acid as Serum Markers of Chemoradiotherapy Response in Locally Advanced Rectal Cancer. *Cancer Res Treat*. 2015;47(1):78-89.



Capítulo 2. Desafíos analíticos e implicaciones biológicas de los lípidos oxidados en el perfil metabólico de tumores neuroendocrinos y cáncer de pulmón

1. Introducción

Durante más de dos décadas, la metabolómica no dirigida ha sido aplicada con éxito en muchas áreas de investigación (1). Su potencial surge de su capacidad para nuevos descubrimientos, ya sean nuevos metabolitos o roles para compuestos ya conocidos. Sin embargo, esta ventaja a menudo se convierte en una gran limitación durante el proceso de identificación, cuando las señales deben asignarse a metabolitos específicos. El flujo de trabajo llevado a cabo por la mayoría de los estudios no dirigidos de metabolómica incluye como primer paso para la anotación, la coincidencia precisa de masa, estableciendo unos límites de error para la búsqueda en bases de datos. En muchas ocasiones, este es el primer y, único método, utilizado para la anotación de metabolitos, y produce diversos candidatos o anotaciones poco fundamentadas. Por tanto, la necesidad de una anotación mejorada es fundamental y, se refleja a través de las herramientas de software y los enfoques que tienen como objetivo ayudar en este proceso (2).

Una anotación tentativa puede tener mucho éxito en términos de metabolitos bien conocidos y catalogados, pero obviamente, falla en el caso de moléculas recién detectadas, especialmente cuando nos encontramos ante modificaciones de estos metabolitos. Por esta razón, se necesita más información estructural sobre los metabolitos anotados. Es por esto por lo que, el análisis MS/MS se convierte en una pieza indispensable, que permite generar los fragmentos clave característicos de una especie o, evidencias claras para corregir una anotación errónea. Los espectros de fragmentación generados que permitan la coincidencia espectral con los espectros disponibles en las bibliotecas proporcionarán una fácil determinación y una rápida anotación de la molécula. Esto es posible, siempre y cuando, la biblioteca disponga de la información espectral sobre el metabolito buscado, y aunque el número de espectros de fragmentación disponibles en estas bases de datos aumenta continuamente, se estima que solo entre el 5 y 10% de todos los metabolitos conocidos tienen información espectral (3).

Hasta la fecha, el análisis de los estándares de referencia disponibles comercialmente continúa siendo la mejor manera de confirmar la identificación de un compuesto, sin embargo, este hecho no siempre es posible, debido a la dificultad de acceso a los estándares de referencia de los cientos o miles de sustancias de interés que se pueden encontrar en una muestra compleja (4). Por lo tanto, para las sustancias sin estándares de referencia o que no están en las bibliotecas espectrales, las fragmentaciones asistidas *in-silico* son una gran ayuda para respaldar el proceso de

anotación. La fragmentación obtenida experimentalmente se enfrenta a la fragmentación generada computacionalmente para una estructura candidata, aplicando diferentes enfoques de fragmentación como métodos heurísticos, combinatorios, de aprendizaje automático y de química cuántica.

Sin embargo, ninguno de los métodos mencionados anteriormente (ni los enfoques basados en MS¹ ni los basados en MS/MS) serán efectivos, si no se encuentra el candidato buscado; la solidez del proceso de anotación está limitada por los candidatos reportados y, no será nunca efectiva y fiable, si no se ha proporcionado una identificación verdadera. Además, los enfoques computacionales a veces no consideran reglas más elaboradas de estabilidades iónicas que nos conducirían a identificaciones erróneas (1).

El término lipidoma fue definido por primera vez en 2001, como el espectro completo de lípidos que se encuentran en un tejido, orgánulo o membrana (5). En los sistemas biológicos, se define como lipidómica, el estudio a gran escala de las vías y redes de lípidos y nos encontramos con una de las llamadas ciencias “ómicas” que ha presentado mayor crecimiento en la última década. No solo nos encontramos ante un crecimiento en número de especies, sino también, al análisis y descubrimiento del papel de estos lípidos en estudios biomédicos. En la base de datos de estructuras lipídicas LIPID MAPS® (LMSD, <https://www.lipidmaps.org/>), a día de hoy (enero 2022), se recogen más de 47100 estructuras (6). Esto es debido en gran parte a los avances en espectrometría de masas (MS) con alta sensibilidad, precisión y resolución, combinada con técnicas de separación como la cromatografía de líquidos (LC).

Los lípidos desempeñan funciones diversas esenciales a nivel estructural, de señalización y/o producción energética y, como otras biomoléculas, están expuestos a sufrir modificaciones químicas que amplían su funcionalidad. La oxidación es una de las modificaciones que más ha ido ganando atención, cuyo interés es debido fundamentalmente a los avances ocurridos en la metodología, que han permitido la detección de estas especies que se encuentran en abundancias muy pequeñas, y al creciente número de referencias en la literatura, por su relativo reciente descubrimiento, que los implica en determinados procesos biológicos y roles patológicos. Estos son compuestos bioactivos, producidos en rutas biosintéticas específicas y, generalmente, como respuesta a un estímulo extracelular (7).

Los lípidos oxidados pueden derivar de la inclusión de ácidos grasos hidroxilados en la estructura lipídica (8-10) o de modificaciones del tipo enzimático o no-enzimático (11-13). Lo que ha llamado la atención de los investigadores es la estrecha relación

intrínseca entre estas modificaciones y el estrés oxidativo, y que se presentan desreguladas en muchas patologías como el Alzheimer, la diabetes mellitus, enfermedad cardiovascular o el cáncer, entre otras (14-19).

El estrés oxidativo que ocurre en los organismos vivos se refiere al desequilibrio entre las especies reactivas de oxígeno (ROS) y su eliminación por determinados mecanismos. El estado de un organismo se rige por la actividad de las células que lo constituyen las cuales equilibran reacciones bioquímicas para así mantener la homeostasis. La “homeostasis redox” es uno de los equilibrios cruciales consistente en la regulación del metabolismo oxidativo y reductor. Las últimas investigaciones en este campo se han centrado en el estudio del papel que desempeñan los lípidos oxidados en los procesos endógenos y, como están íntimamente relacionados con el estrés oxidativo (20). Los lípidos oxidados están involucrados en procesos muy importantes como la producción de energía a través de la β -oxidación, señalización a través de eicosanoides o la degradación oxidativa incontrolada provocada por los radicales libres. Por lo tanto, no solo los lípidos intactos sino también sus formas oxidadas representan unos de los metabolitos más importantes de la bioquímica de los mamíferos (21).

Es de sobra conocido que el estrés oxidativo, y con ello incluimos la oxidación de lípidos, son muy importantes en la patología de enfermedades como el cáncer. La proliferación celular, migración o el crecimiento metastásico juegan un papel crucial en la aparición y progresión del cáncer y, estos eventos, están regulados por diferentes factores, entre los que se encuentran los productos lipídicos. Productos derivados por oxidación de ácidos grasos y lípidos, y las enzimas responsables de sus producción, lipooxigenasas y ciclooxigenasas, se encuentran sobreexpresadas durante el desarrollo y crecimiento tumoral (22, 23).

Los productos generados del ataque de ROS sobre los lípidos dan como resultado un número elevado de especies con una alta diversidad estructural, que incluye desde especies con un simple hidroxilo en la cadena hidrofóbica, a reacciones de fragmentación o ciclación y, como consecuencia, una mayor complejidad. A esto se une la aparición de formas isoméricas, su escasa abundancia, su inestabilidad química y la diversidad, aún desconocida, en muestras biológicas, lo que los convierte en un verdadero reto para los investigadores.

Pese a que, al igual que el número de estudios que incluyen lípidos oxidados ha aumentado considerablemente, la información recogida en las bases de datos sobre estas especies es elevada pero no abarca su totalidad. Por ello, la interpretación exhaustiva de la información proporcionada a nivel de MS¹ y MS/MS es definitiva para

la correcta anotación. La anotación de metabolitos oxidados se ve facilitada por el hecho de que, generalmente, se construyen parcialmente a partir del esqueleto de moléculas conocidas y, por lo tanto, deben reconocerse al menos algunos iones producto. La búsqueda de patrones de fragmentación conocidos podría confirmar la presencia de un grupo químico particular y, por lo tanto, proporcionar información valiosa sobre la clase química (24).

Durante la movilidad internacional realizada en *Mass Spectrometry Laboratory of Clinical Research Center* de *Medical University of Białystok* en Polonia bajo la tutela del profesor asociado Michał Ciborowski se llevó a cabo la evaluación del estado de los lípidos oxidados en distintos tipos de cáncer.

El objetivo fundamental de la estancia fue ampliar capacidades y conocimientos relacionados con el uso de la metabolómica para la búsqueda de biomarcadores en cáncer, tema de la tesis doctoral. Como ya se ha comentado, poder encontrar biomarcadores específicos es crítico para obtener un diagnóstico temprano y proporcionar un tratamiento efectivo.

Por este motivo el trabajo se dividió en dos partes: una primera parte práctica, de trabajo en el laboratorio, para la optimización de un método de análisis de los diferentes productos de oxidación de los principales fosfolípidos y, una segunda parte de reprocesado de datos para la profundización en la identificación de lípidos oxidados en cáncer, con dos aplicaciones, cáncer de pulmón y tumores de origen neuroendocrino.

Como hasta ahora la información sobre estos metabolitos es todavía limitada, en los datos procedentes del análisis de muestras de cáncer de pulmón, se realizó una búsqueda *target* (dirigida) de las principales formas oxidadas de un tipo de glicerofosfolípido, glicerofosfocolina (PC), siguiendo un procedimiento similar al empleado por Godzien y *colaboradores* (25) en muestras de suero de pacientes diabéticos. A grandes rasgos conocemos, hasta ahora, dos tipos de oxidaciones posibles de estos lípidos, de cadena larga y de cadena corta, y como estas modificaciones afectan considerablemente la naturaleza hidrofóbica de estos metabolitos (20). Una vez identificadas las distintas modificaciones de las principales PC (oxPCs) se ha realizado un reprocesado *target* de los datos procedentes tanto de plasma como de tejido. Con esto se observó cómo están afectados los niveles de las formas oxidadas en comparación de sus formas nativas; oxPCs son significativamente menos abundantes que sus precursores no oxidados lo cual concuerda con el conocimiento general sobre la concentración de las formas oxidadas en muestras biológicas (26). Esto a su vez limita el trabajo y la cantidad de modificaciones que somos

capaces de localizar, lo que lo convierte en una ardua tarea. Esta limitación sería fácilmente superada con la caracterización del máximo número de modificaciones lipídicas, lo que nos ayudaría en la identificación de estas.

Con el objetivo de construir una base de datos que nos proporcione la mayor información para poder identificar estos metabolitos, se procedió al análisis de estándares modificados y su posterior caracterización. Partiendo de seis estándares de fosfolípidos, de los más abundantes en muestras biológicas, se sometieron a un tratamiento de oxidación para su posterior determinación mediante LC-MS y caracterización por MS/MS. La metodología empleada para la oxidación fue optimizada a partir del tratamiento empleado por el grupo de la Dra. Federova (2) y, la cual fue seguida con la toma de alícuotas a distintos tiempos para poder estudiar el desarrollo de la reacción. Todas estas formas oxidadas y sus correspondientes formas nativas se analizaron utilizando un método previamente establecido en el centro de investigación (25) y con los datos obtenidos se realizó la caracterización. Posteriormente se trabajó en la optimización del método de análisis donde se fueron modificando los parámetros relacionados con la separación cromatográfica y la espectrometría de masas para obtener los mejores resultados. Tras este análisis, los datos podrían ser reprocesados, utilizando un software libre desarrollado por el mismo grupo de trabajo.

Por lo tanto, este capítulo expone, por una parte, los experimentos llevados a cabo en el laboratorio donde se realizó la estancia, relacionados con el tratamiento de estándares de lípidos para su oxidación y, posterior análisis con una metodología optimizada. Esto se completó con el reprocesado de los datos con la intención de crear una librería de estos productos de oxidación con sus tiempos de retención y los fragmentos característicos procedentes del análisis MS/MS por disociación inducida por colisión (CID). Partiendo de una exhaustiva revisión bibliográfica de lo publicado en los últimos años en relación con el tratamiento de muestras y el método de análisis de estándares lipídicos oxidados, se decidió seguir con el procedimiento aplicado por Ni y colaboradores (2) que ofrecía resultados prometedores y, el consiguiente uso de la herramienta LPPTiger (<https://home.uni-leipzig.de/fedorova/software/lpptiger/>), de uso libre, desarrollada por el mismo grupo, para la identificación de los fosfolípidos oxidados. Esta revisión se centró no solo en el tratamiento de estándares para su oxidación, teniendo en cuenta toxicidad de los reactivos utilizados, atmosferas de trabajo y tiempo requerido, entre otros, sino también en los métodos de análisis, fundamentalmente mediante LC-MS en fase inversa, y en los resultados obtenidos, rendimiento y la posible aplicación del método de análisis a muestras biológicas.

Por otra parte, el reprocesado de datos para la identificación de lípidos oxidados se realizó con datos procedentes de dos estudios metabolómicos no dirigidos a través de la plataforma LC-MS. Uno de ellos consta de datos del análisis LC-MS de muestras de plasma y tejido de pacientes con cáncer de pulmón de células no pequeñas (NSCLC) llevado a cabo por el grupo donde se realizó la estancia. El otro conjunto de datos proviene del análisis no dirigido multiplataforma de muestras de plasma de pacientes con tumores neuroendocrinos (NET) (Capítulo 1) llevado a cabo como parte de la investigación de esta tesis doctoral. Para ello, se utilizaron los datos procedentes del análisis LC-MS y, se estableció un flujo de trabajo para la anotación de lípidos oxidados a nivel de MS/MS cuando la asignación de su anotación no es posible mediante la información de las bases de datos disponibles y las herramientas existentes.

El nivel MS/MS cubre el emparejamiento espectral de metabolitos conocidos y la elucidación estructural de moléculas desconocidas, aplicado a especies lipídicas oxidadas como los derivados del ácido araquidónico, conocidos como HETEs y las lisoglicerofosfolinas oxidadas (oxLPCs). Como ya hemos comentado, la oxidación es una de las reacciones más comunes que ocurren en las células vivas y se ha investigado durante décadas, especialmente en términos de oxidación de lípidos. Sin embargo, se desconocen muchos productos de oxidación. La investigación sobre los mecanismos de oxidación de los fosfolípidos y su papel en enfermedades específicas aún está en curso, y muchas preguntas siguen sin respuesta (12). Múltiples actividades biológicas de glicerofosfolípidos oxidados que causan lesiones oxidativas patógenas se han relacionado con cáncer (27-29), diabetes (25, 30, 31), enfermedades neurodegenerativas (32, 33) o enfermedades cardiovasculares (34, 35). Se sabe aún menos acerca de los oxLPCs, ya sea a nivel analítico o biológico.

Los estudios de reprocesado de datos reunieron información recogida de muestras de plasma de pacientes afectados por NET, y de plasma y tejido de pacientes afectados con NSCLC. Los NET son una familia heterogénea de neoplasias en lo que respecta a su biología molecular, comportamiento clínico, pronóstico y respuesta a la terapia, cuya incidencia ha aumentado sustancialmente en los últimos años (34). Aunque se han logrado mejoras significativas en el tratamiento de NET en los últimos años, existen todavía desafíos con respecto a la estratificación del paciente y al seguimiento del tratamiento. Nuevos biomarcadores son, por lo tanto, necesarios para ayudar en la toma de decisiones clínicas y, en última instancia, mejorar los resultados del paciente. Aunque las localizaciones principales son el páncreas, los pulmones y el tracto digestivo, estos tumores pueden aparecer en cualquier sitio anatómico ya que las

células neuroendocrinas se distribuyen por todo el cuerpo durante la etapa embrionaria (36). En comparación con otros tipos de cáncer, el número de investigaciones centradas en los NET es relativamente bajo. Sin embargo, la heterogeneidad antes mencionada de los NET y una incidencia creciente exige que ampliemos nuestro conocimiento actual.

El cáncer de pulmón del tipo de células no pequeñas, NSCLC, es uno de los más comunes y representa alrededor del 85% de los casos. Aunque el cáncer de pulmón está entre los más diagnosticados en el mundo, sigue siendo uno de los más agresivos y la principal causa de muerte por cáncer. EL NSCLC comienza en las células epiteliales que conforman mayoritariamente los pulmones y, según el tipo de célula epitelial donde comience, se originan los distintos tipos. El adenocarcinoma (ADC) comienza en las células productoras de mucosidad, mientras que el carcinoma de células escamosas (SCC) se origina en las células que recubren las vías respiratorias. Ambos representan el 85% de todos los casos (37). Por último, el carcinoma de células grandes (LCC) nace en células que no pertenecen a ninguno de los dos tipos descritos anteriormente. Esta parte es muy importante, delimitar en qué tipo de células se ha originado el tumor, ya que a nivel clínico determina las distintas opciones de tratamiento. Actualmente no existen disponibles métodos no invasivos para el diagnóstico temprano y el tipificado de NSCLC, lo que convierte en un aspecto crucial del tratamiento y la supervivencia del paciente, el diagnóstico de la enfermedad en el estado más precoz posible. Aunque se prefiere el uso de muestras de plasma y suero para el diagnóstico de cáncer y la monitorización de los efectos de un tratamiento, la información recopilada a nivel del tejido es vital para estudiar la bioquímica del cáncer. Por eso, este estudio se ha complementado con el análisis de muestras de plasma y tejido para intentar solventar las dificultades propias de la búsqueda de biomarcadores de cáncer en plasma o suero debido a la elevada homeostasis que se produce en la sangre, además de que el tejido es el centro de proliferación de reacciones y donde ocurre la secreción de los metabolitos a la sangre.

2. Materiales y métodos

2.1. Estándares de lípidos

1-palmitoil-2-araquidinoil-sn-glicerol-3-fosfocolina (PC 16:0/20:4 o PAPC), ácido 1-palmitoil-2-araquidinoil-sn-glicerol-3-fosfatídico (PA 16:0/20:4), 1-palmitoil-2-araquidonoil-sn-glicero-3-fosfo-L-serina (PS 16:0/20:4), 1-palmitoil-2-docosaheptaenoil-sn-glicerol-3-fosfocolina (PC 16:0/22:6), ácido 1-palmitoil-2-docosaheptaenoil-sn-glicero-

3-fosfatídico (PA 16:0/22:6) y 1-palmitoil-2-docosahexaenoil-sn-glicero-3-fosfo-L-serina (PS 16:0/22:6) de AVANTI Polar Lipids fueron los estándares comerciales empleados para la optimización del método de oxidación. Las características de cada estándar se resumen en la Tabla 1.

Tabla 1. Estándares comerciales empleados para el método de oxidación. Además del nombre completo y la abreviatura, se indica el CAS, la riqueza, el disolvente y su concentración.

Estándar	Abreviatura	CAS	Riqueza	Disolución	Concentración	Vendedor
1-palmitoil-2-arachidonoyl-sn-glycero-3-phosphocholine	PC 16:0/20:4	35418-58-7	>99%	CHCl ₃	10 mg/mL	AVANTI Polar Lipids
1-palmitoil-2-arachidonoyl-sn-glycero-3-phosphate	PA 16:0/20:4	322647-60-9	>99%	CHCl ₃	10 mg/mL	AVANTI Polar Lipids
1-palmitoil-2-arachidonoyl-sn-glycero-3-phospho-L-serine	PS 16:0/20:4	474943-16-3	>99%	CHCl ₃	10 mg/mL	AVANTI Polar Lipids
1-palmitoil-2-docosahexaenoyl-sn-glycero-3-phosphocholine	PC 16:0/22:6	59403-54-2	>99%	CHCl ₃	10 mg/mL	AVANTI Polar Lipids
1-palmitoil-2-docosahexaenoyl-sn-glycero-3-phosphate	PA 16:0/22:6	474943-28-7	>99%	CHCl ₃	10 mg/mL	AVANTI Polar Lipids
1-palmitoil-2-docosahexaenoyl-sn-glycero-3-phospho-L-serine	PS 16:0/22:6	474943-17-4	>99%	CHCl ₃	10 mg/mL	AVANTI Polar Lipids

2.2. Preparación y análisis de estándares oxidados

Brevemente, antes de comenzar con la reacción de oxidación, se deben formar los liposomas a partir de los estándares de lípidos debido a que la reacción de Fenton no funciona bien según el solvente orgánico que se use. Para ello se debe secar el estándar comercial, que generalmente viene en disolución con cloroformo (CHCl₃). Para ello se tomaron 65 µL de cada estándar comercial, se llevó a sequedad utilizando un sistema de nitrógeno gas (N₂) y se rehidrató con un tampón acuoso de NH₄HCO₃ de concentración 3 mM. La mezcla se agitó en un vortex y se sometió a ultrasonidos durante 15 min. De esta forma se consiguieron formar los liposomas (Figura 1).

Una vez obtenidas las vesículas, se continuó con la reacción propiamente de oxidación (Figura 1). A 125 µL de las vesículas formadas se le agregó una disolución de CuSO₄ 0.75 mM y 50 µL de una disolución de 1.5 mM de ácido ascórbico. El ascorbato debe estar en la disolución para reducir el Cu (II) (del sulfato de cobre) a Cu (I). Solo Cu(I) puede realizar la reacción de Fenton con H₂O₂. Por último, la mezcla fue incubada

a 37 °C durante 72 h y se fueron tomando alícuotas cada 24 horas para hacer el seguimiento de la reacción.

Adicionalmente se estudió un método de extracción, debido a posibles interferencias con el cobre y el ácido ascórbico. Para ello, una vez concluidos los distintos tiempos de incubación, se añadió a la mezcla 2 volúmenes de metil-terc-butileter y se agitó durante 30 min para, finalmente, recoger la fase superior para su análisis.

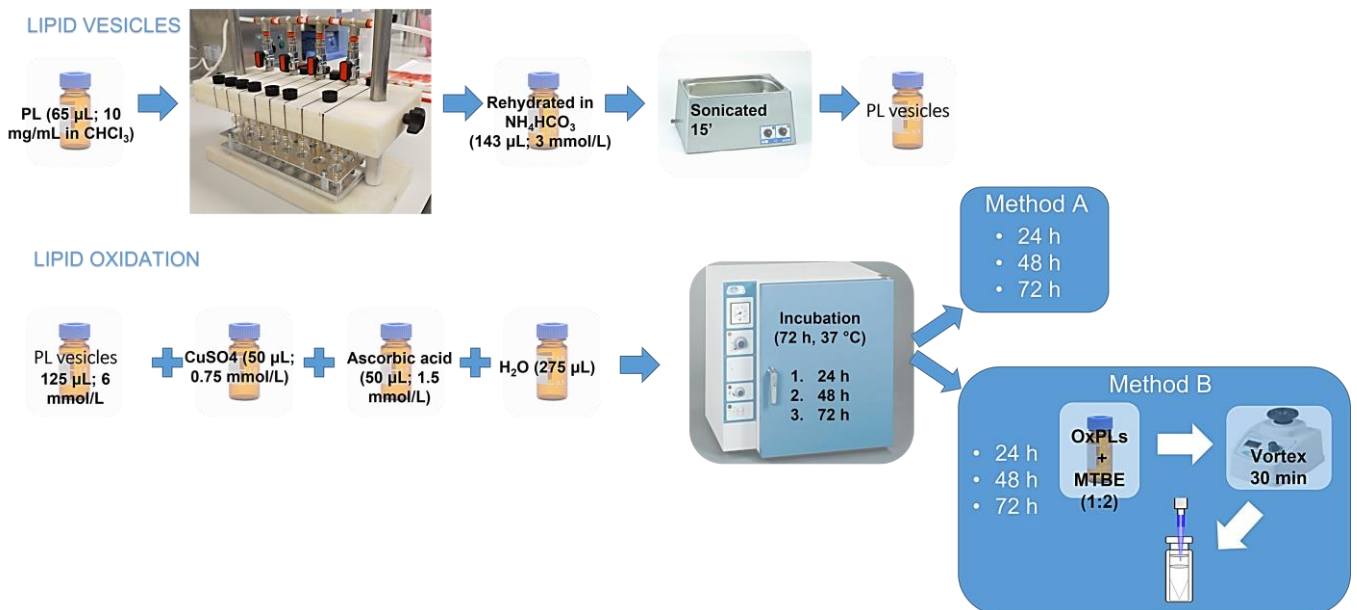


Figura 1. Esquema del tratamiento de oxidación empleado para los estándares comerciales. El proceso se divide en una primera etapa fundamental para formar los liposomas, seguido de la etapa de oxidación. Por último, se siguieron dos métodos antes del análisis, uno sin filtración (A) y otro, extrayendo con MTBE (B). Ambas metodologías fueron caracterizadas a distintos tiempos de incubación (24, 48 y 72 horas).

2.3. Cohortes de estudio

- *Cohorte de estudio de pacientes con NET*

La población de este estudio incluyó 77 muestras de plasma de pacientes diagnosticados con tumores neuroendocrinos no pancreáticos inscritos en un ensayo clínico de fase II-III doble ciego (AXINET) que aleatorizó a los pacientes para recibir un inhibidor selectivo de los receptores del factor de crecimiento endotelial vascular (VEGF), Axitinib, *versus* placebo, ambos en combinación con octreotida LAR (EUDRACT: 2011-001550-29). Los pacientes debían cumplir una serie de criterios de inclusión que se detallaron en el capítulo 1. Además, el estudio se completó con 68 muestras de plasma de pacientes no-cancerosos, utilizadas como control.

- *Cohorte de estudio de los pacientes con NSCLC*

Las muestras de tejido pulmonar se obtuvieron de pacientes sometidos a tratamiento quirúrgico por NSCLC primario en el Departamento de Cirugía Torácica del Hospital Clínico de la Universidad Médica de Białystok (Polonia). Las muestras se recolectaron siguiendo los más altos estándares de biobanco establecidos por la universidad (38). El estudio incluyó un total de 76 muestras de tejido de pacientes con NSCLC. Estas muestras, además de tejido control adyacente sin cambios morfológicos, se recolectaron y, fueron revisados histológicamente y clasificadas. Las muestras se congelaron y almacenaron a $-80\text{ }^{\circ}\text{C}$ hasta su análisis.

Las 33 muestras de plasma analizadas pertenecían al grupo clasificado como pacientes con NSCLC en etapa avanzada y, de los cuales, se disponía también, de muestras de tejido pulmonar que fueron analizadas en este estudio. Los pacientes fueron clasificados según el tipo de célula de origen como: 11 con ADC, 15 con SCC y 7 con LCC. Además, fueron incluidos en el estudio, plasma de 26 pacientes con enfermedad pulmonar obstructiva crónica (EPOC), clasificados como grupo control.

Las muestras de sangre se recogieron en tubos de 9 mL con sistema de vacío con K_2EDTA como anticoagulante. Tras la mezcla suave, se centrifugó la mezcla a $1300\text{ } \times g$ durante 20 minutos a temperatura ambiente y, la fracción del plasma fue recogida en tubos Eppendorf y almacenada hasta su análisis a $-80\text{ }^{\circ}\text{C}$. Este estudio fue aprobado por el Comité de Ética de la Universidad Médica de Białystok. Además, antes de recolectar las muestras, se obtuvo el consentimiento informado por escrito para la recolección de muestras de todos los participantes.

2.4. Preparación de las muestras

Las muestras de plasma de NET se sometieron a desproteización y extracción de los metabolitos según el protocolo establecido en el centro para el análisis LC-MS. Para ello, brevemente, se mezclaron 100 μl de plasma con 300 μl de MeOH:EtOH frío ($-20\text{ }^{\circ}\text{C}$) (1:1, v/v). Las muestras se mezclaron con vórtex durante 1 min, se incubaron en hielo durante 5 min y se centrifugaron a $16000\text{ } \times g$ durante 20 min a $4\text{ }^{\circ}\text{C}$. Por último, el sobrenadante resultante se transfirió directamente a un vial para su análisis.

El plasma de NSCLC se sometió al tratamiento de muestra optimizado por el centro donde se realizó el análisis (39), que difirió del protocolo empleado y descrito en el capítulo 1, en el filtrado a través de un filtro de nailon de $0.22\text{ }\mu\text{m}$ directamente al vial de análisis.

Por su parte, la preparación de las muestras de tejido se realizó siguiendo protocolo publicado por el centro (40) donde, brevemente, el tejido pulmonar fue homogeneizado con 50% de metanol frío utilizando un homogeneizador de molino de perlas (Tissue Lyser LT; Qiagen Hilden, Alemania) durante 8 min (30 Hz). Para extraer los metabolitos se añadió ACN frío que contenía 1 mg/L de Zomepirac (patrón interno, IS) y, las muestras fueron centrifugadas y filtradas antes de su análisis. El protocolo detallado de tratamiento de muestras de NSCLC fue descrito por Kowalczyk y colaboradores (41).

2.5. Análisis UHPLC–MS

Las muestras de NET se analizaron en un sistema de cromatografía de líquidos, UHPLC 1290 Infinity II (Agilent Technologies, Waldbronn, Alemania) que incluye una bomba binaria y un muestreador múltiple automático termostatzado, acoplado a un analizador cuadrupolo–tiempo de vuelo, QTOF, de masa precisa Agilent 6545. Para la separación, como ya se explicó en el capítulo 1, se inyectaron 0.5 μ l de muestra extraída en una columna de resolución rápida Zorbax Extended-C18 (Agilent Technologies, 2.1 x 50 mm, 1.8 μ m) termostatzada a 60 °C. El caudal fue de 0.6 mL/min con una fase móvil compuesta por agua con ácido fórmico al 0.1% para A y ACN con ácido fórmico al 0.1% para B. El gradiente cromatográfico partió del 5% B durante el primer minuto, aumentando a 80% B en 6.0 min, luego al 100% hasta el minuto 11.5 para volver a la condición inicial en 0.5 min permitiendo el reequilibrio hasta el minuto 15.0 (25).

Los datos se recopilaron en ambos modos de ionización ESI, positivo y negativo, en análisis separados y, operaron en el rango de m/z 100 a 1000 y m/z 40 a 1000 para análisis MS y análisis MS/MS, respectivamente.

Se estableció un voltaje adicional de 1000 V en la fuente de ionización *Jet Stream* (*nozzle voltage*) y el voltaje capilar fue de 3000 V con una velocidad de exploración de 1.5 espectros/s en modo positivo y –4000 V con una velocidad de exploración de 1.0 espectro/s en modo negativo. La temperatura del gas de secado se fijó en 250 °C y se hizo fluir a una velocidad de 12 L/min, a una presión de 52.0 psi. Adicionalmente, se utilizó un gas auxiliar para aumentar la temperatura hasta los 370 °C con una velocidad de 11 L/min y mejorar la ionización.

Para la corrección de masa durante la adquisición de datos, se infundió continuamente en el sistema dos compuestos de referencia durante el transcurso de todo el análisis: m/z 121.0509 (purina protonada) y m/z 922.0098 hexakis (1H, 1H, 3H-tetrafluoropropoxi) fosfazina protonada (HP-921) para el modo de ionización positivo; y

m/z 112.9856 (anión con pérdida de H del ácido trifluoroacético (TFA)) y m/z 966.0007 (aducto de formiato del HP-921) para el modo negativo.

Las muestras se analizaron de forma aleatoria durante la secuencia. Para controlar el rendimiento del sistema y determinar la reproducibilidad, se analizaron varias réplicas de una mezcla homogénea que contenía una pequeña alícuota de todas las muestras (controles de calidad, QC). Estos controles de calidad se trataron como el resto de las muestras y se inyectaron al comienzo de la secuencia (10 inyecciones) para equilibrar el sistema y, cada diez muestras para controlar aún más la estabilidad del análisis (42).

El análisis de los estándares oxidados y la toma de huellas dactilares metabólicas para las muestras de plasma y tejido de NSCLC se realizó utilizando los mismos métodos LC-MS descritos anteriormente (39, 40). Los productos de oxidación y las muestras extraídas se analizaron mediante un equipamiento LC-MS, sistema 1290 Infinity UHPLC (Agilent, Santa Clara, California, EE. UU.), similar al empleado en el análisis de las muestras NET, con la diferencia que este se combinaba con un espectrómetro de masas QTOF con tecnología *iFunnel* 6550 como detector (Agilent, Santa Clara, California, EE. UU.). Los análisis se realizaron en modo ESI(+) y ESI(-) tanto para los estándares como para las muestras de plasma y tejido de NSCLC, y en los estándares oxidados se llevó a cabo el análisis MS/MS utilizando la herramienta AUTO MS/MS. Esta tiene la capacidad de realizar un análisis MS/MS automatizado, lo que permite capturar la mejor información de fragmentación de la selección de iones. Este método fue optimizado para un rango de adquisición MS de m/z de 400 a 1000 y de MS/MS de m/z de 40 a 1000 utilizando las mismas condiciones de separación cromatográfica que en el análisis MS. Utilizando una ventana estrecha de aislamiento, los iones precursores fueron fragmentados con una energía de colisión de 20 eV. La selección de los iones precursores se optimizó con un umbral de 25000 cuentas, excluyendo los m/z pertenecientes a los compuestos empleados como masas de referencia. La información sobre el análisis LC-MS de NSCLC realizado se describió detalladamente por Kowalczyk y *col.* (41).

2.6. Procesado de datos

- *Procesado de datos de estándares oxidados*

Los cromatogramas, espectros de masas y de fragmentación fueron inspeccionados manualmente utilizando Mass Hunter Qualitative (B.06.00, Agilent). Los lípidos nativos, como sus diferentes productos de oxidación, se extrajeron utilizando su

correspondiente $[M + H]^+$, comprobando, también, su aducto $[M + Na]^+$ en ionización positiva y $[M - H]^-$, $[M + HCOOH-H]^-$, $[M + Cl]^-$ en el modo de ionización negativa con una tolerancia de masa de máximo 10 ppm.

- *Procesado de datos NET*

Los datos en crudo se procesaron mediante el algoritmo de extracción de características moleculares (MFE) en Mass Hunter Qualitative (B.06.00, Agilent). El algoritmo MFE permite borrar los datos del ruido de fondo y los iones no relacionados, y crea una lista de características. Para encontrar iones coeluyentes y co-relacionados, se consideraron los aductos $[M]^+$, $[M + H]^+$, $[M + Na]^+$ en ionización positiva y $[M - H]^-$, $[M + HCOOH-H]^-$, $[M + Cl]^-$ en el modo de ionización negativa. También se incluyó la pérdida de agua neutra para ambas polaridades. El método de reprocesamiento establecido en el software Mass Hunter Qualitative se utilizó en la herramienta DA Reprocessor Offline Utilities (B.05.00 Agilent) para el procesamiento automatizado fuera de línea de todo el conjunto de datos. Con el software Mass Profiler Professional (B.14.9 Agilent) se llevó a cabo la alineación de los datos reprocesados y el filtrado de los mismos, utilizando una ventana de tiempo de retención de $0.5\% \pm 0.15$ min y una de masa de 15 ppm ± 2.0 mDa, conservando los compuestos que estuviesen presentes en, al menos, el 75% de los QCs.

- *Procesado de datos NSCLC*

Los datos crudos fueron reprocesados de forma dirigida a través del software MassHunter Profinder (B.08.00, Agilent). A través de su algoritmo de extracción de características de forma *target*, dirigida, (TFE), utilizando una base de datos que contiene las fórmulas moleculares, masa y/o tiempo de retención, busca y extrae las características de los datos que cumplan esos criterios mediante un proceso denominado “buscar compuestos por fórmula” (FbF).

Para construir esa base de datos necesaria para el procesamiento dirigido, utilizando MassHunter Qualitative (B.07.00, Agilent), se realizó la búsqueda manual de todos los productos de oxidación posibles derivados de las principales glicerofosfatidilcolinas (PC) encontradas en muestras de humanos siguiendo la información publicada (25). Estos productos y sus formas nativas fueron caracterizados a través de los tiempos de retención de todas sus posibles formas isoméricas utilizando los datos de ESI positivo y negativo. Todas las características encontradas se inspeccionaron y, con esto, se construyó la base de datos necesaria para el reprocesado dirigido, asignando cada especie con su fórmula, masa molecular y tiempos de retención.

Cuando las características moleculares de entrada se caracterizan con su masa y tiempo de retención, en lugar de solo con la fórmula molecular, el algoritmo TFE calcula patrones isotópicos razonables y los utiliza para encontrar las características objetivo en los archivos de datos con una tolerancia de tiempo de retención. La inspección manual de todas las características encontradas, incluyendo la reintegración, completaron el procesado.

2.7. Análisis estadístico

- *Análisis estadístico NET*

El reprocesado de datos y el análisis estadístico llevado a cabo para el estudio LC-MS de NETs se explicó detalladamente en el Capítulo 1. Brevemente, para la corrección del efecto dentro de la secuencia se utilizó el algoritmo de muestras de control de calidad y regresión de vectores de soporte (QC-SVRC) con los parámetros detallados anteriormente (43). Tras la corrección, los datos se sometieron a un procedimiento de garantía de calidad (QA) para evaluar la varianza analítica de cada característica medida. Los datos se filtraron conservando las características con desviaciones estándar relativas (RSD) < 30%.

El análisis univariante (UVA) se llevó a cabo utilizando MATLAB (R2015a, MathWorks) y, se comprobó la normalidad aplicando la prueba de Kolmogórov-Smirnov (K-S). Por último, se evaluaron las diferencias entre casos NET y los individuos sin cáncer aplicando la prueba t de Student (valor de $p \leq 0.05$) con post-corrección múltiple de Benjamini-Hochberg.

El análisis multivariado (MVA) se realizó en SIMCA 15.0 (Sartorius Stedim Biotech) para visualizar los datos. Se utilizó un método no supervisado, el análisis de componentes principales (PCA), para examinar la calidad de los datos y para probar la agrupación espontánea de las muestras. Para evaluar la fuerza y la calidad de las diferencias entre los grupos se utilizó un método supervisado, el análisis discriminante de mínimos cuadrados ortogonales parciales (OPLS-DA). Por último, la calidad de los modelos construidos se evaluó a través de la varianza explicada (R^2) y la predicha (Q^2), y se realizó una prueba de permutaciones para evaluar cualquier sobreajuste de los modelos (999 permutaciones) (44).

Para inspeccionar las variaciones entre los grupos se calculó el porcentaje de cambio. La abundancia de cada compuesto se promedió en cada muestra entre dos grupos y la variación se calculó de la siguiente manera: (Promedio para el grupo de casos - Promedio para el grupo sin cáncer) / Promedio para el grupo sin cáncer, en %.

- *Análisis estadístico NSCLC*

Los datos resultantes del reprocesado dirigido fueron importados en Mass Profiler Profesional (MPP 12.6.1) (Agilent, Santa Clara, California, USA), donde se realizaron todos los análisis estadísticos y las diferentes comparaciones. Los análisis incluyeron pruebas paramétricas y no paramétricas sin corrección de prueba múltiple, debido a la pequeña cantidad de especies probadas. Los estudios comparativos de plasma fueron no pareados, mientras que las comparaciones con los datos de tejido fueron pareadas.

Las comparaciones entre grupos fueron representadas con diagramas de Venn, mientras que, la distribución de cada especie lipídica significativa se inspeccionó utilizando el gráfico de caja y bigotes, y un gráfico del perfil, donde las abundancias en cada muestra individual se representaron mediante líneas.

Para los datos de plasma no se utilizó normalización alguna, mientras que los datos de tejido fueron normalizados según el patrón interno. Se omitió cualquier otra normalización debido al peso estandarizado del tejido y el solvente de extracción.

El número de muestras utilizado para cada comparación, y la prueba estadística aplicada se resume en la Tabla 2.

Tabla 2. Resumen de los análisis estadísticos llevados a cabo con los datos procedentes de muestras de plasma y tejido de NSCLC, que incluye el número de muestras comparadas y los resultados significativos obtenidos, lípidos nativos y formas oxidadas.

PLASMA						
Comparación	Lípidos significativos	Formas nativas	Lípidos oxidados	N° muestras	Análisis	Tipo comparación
cáncer vs control	16	12	4	33 vs 26	prueba t	Enfermedad
ADC vs control	11	7	4	11 vs 26	Mann-Whitney	Subtipo
LCC vs control	6	6	0	7 vs 26		
SCC vs control	14	11	3	15 vs 26		
estadio 3 vs control	8	5	3	7 vs 26	Mann-Whitney	Estadio
estadio 4 vs control	16	10	6	19 vs 26		
estadio 5 vs control	7	4	3	7 vs 26		
TEJIDO						
Comparación	Lípidos significativos	Formas nativas	Lípidos oxidados	N° muestras	Análisis	Tipo comparación
cáncer vs control	22	12	10	76 vs 76	prueba t	Enfermedad
ADC _{tumor} vs ADC _{control}	21	18	3	25 vs 25	Mann-Whitney pareado	Subtipo
LCC _{tumor} vs LCC _{control}	7	6	1	15 vs 15		
SCC _{tumor} vs SCC _{control}	24	19	5	36 vs 36		
estadio 3 vs control	25	18	7	26 vs 26	Mann-Whitney pareado	Estadio
estadio 4 vs control	23	19	4	25 vs 25		
estadio 5 vs control	18	14	4	25 vs 25		

2.8. Anotación de los metabolitos

Para los datos procedentes del estudio NSCLC, al ser un estudio dirigido, los compuestos significativos se encontraban anotados.

El proceso de anotación de NETs fue más complejo y, cubrió dos etapas: primero, una asignación tentativa basada en los datos de MS¹ y, en segundo lugar, coincidencia espectral y elucidación estructural basada en datos de MS/MS. El flujo de trabajo establecido para las anotaciones se resume en la Figura 2.

- Nivel MS¹

Se realizó una identificación tentativa inicial de las características que discriminan entre grupos en función de la coincidencia de masa exacta utilizando CEU Mass Mediator (CMM). CMM es una herramienta de anotación, que utiliza información relacionada con la ionización, la formación de aductos y el orden de elución en una clasificación de candidatos tentativos recuperados a través de varias bases de datos (45).

La justificación de la asignación se representa mediante una puntuación numérica (46) o mediante explicaciones en lenguaje natural (47). Este proceso comenzó

con la asignación tentativa de las masas experimentales a los resultados candidatos ofrecidos por la base de datos y cubrió la coincidencia de masa exacta, la determinación de la distribución isotópica y la verificación de los posibles iones y aductos.

- Nivel MS/MS

Para confirmar la anotación de los compuestos, se realizó el análisis LC-MS/MS, repitiendo el análisis. El análisis de datos independiente (DIA) se realizó con las mismas condiciones cromatográficas y espectrométricas utilizadas para el análisis inicial no dirigido. De acuerdo con la masa exacta y el tiempo de retención determinados previamente, los iones de interés se seleccionaron para su fragmentación por disociación inducida por colisión (CID), utilizando una ventana de masa aislada estrecha (aproximadamente 1.3 Da). Los iones precursores se sometieron a dos energías de colisión fija, 20 y 40 eV, en análisis separados para obtener sus espectros MS/MS específicos de la energía de colisión.

Como la oxidación de un lípido para formar hidroperóxido o productos de hidróxido implica la adición de oxígeno a la estructura nativa, la masa de la molécula cambia y la modificación resultante puede detectarse fácilmente mediante MS. Por ello, se realizó la interpretación manual precisa y exhaustiva de los espectros MS/MS comprobando los múltiples productos de oxidación de lípidos que se pueden generar. Con el espectro MS/MS del compuesto estándar, la forma nativa, siempre y cuando estuviese disponible en la librería del centro o en la base de datos METLIN (45) o MEtFrag (46), de fragmentación *in-silico* por sistemas asistidos por computadora, se realizó la comparación del patrón de fragmentación, basado en el mecanismo de dicho proceso, aplicando las diferencias de masa para comprobar la posible identidad de un producto de oxidación lipídico. La nomenclatura utilizada en cualquier caso, incluida la anotación de los iones de fragmentos observados, se basó en las recomendaciones de la literatura (47). Brevemente, un tipo de fragmento de lípidos, definido como fragmento específico de especies de lípidos moleculares (MLF), proporciona información sobre la composición química de la cadena de acilo. Sin embargo, la detección de MLF no respalda la inferencia de restos ácidos grasos (FA) en los glicerofosfolípidos en posiciones sn-1 / sn-2 / sn-3. Por otro lado, otro tipo de fragmento lipídico definido como fragmento selectivo de clase lipídica (LCF) proporciona información sobre todas las moléculas que pertenecen a la misma clase lipídica y, es común a todas ellas.

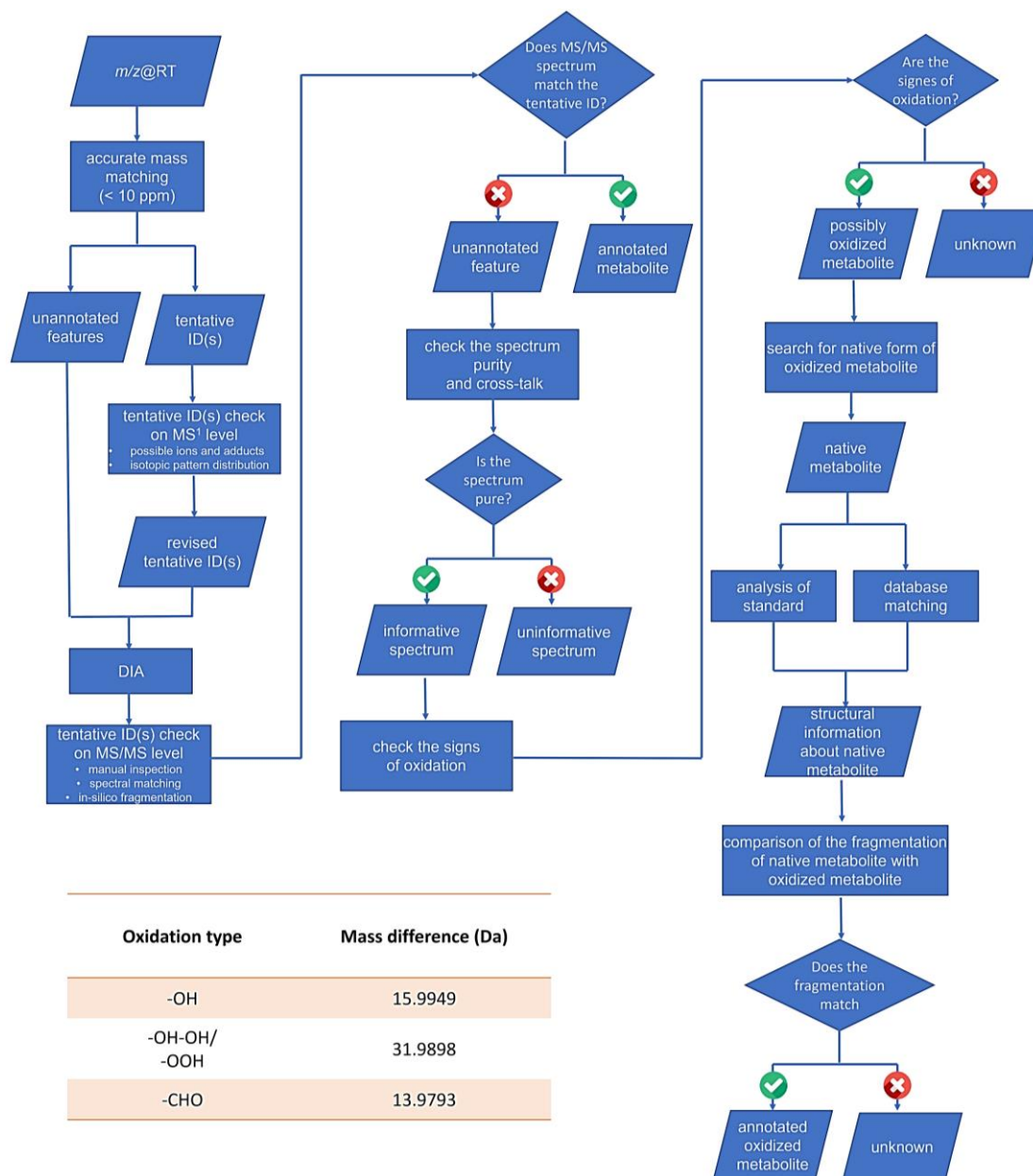


Figura 2. Flujo de trabajo de anotación para metabolitos oxidados. Esquema de los pasos realizados para la anotación de especies oxidadas en muestras biológicas. En el nivel MS¹, los candidatos asignados tentativamente de CMM se verificaron manualmente. Para la confirmación de las ID tentativas y la anotación de señales sin ningún candidato de las bases de datos, se realizó DIA. Los espectros de fragmentación obtenidos se compararon con la estructura de los candidatos obtenidos, lo que condujo a la conformación de los ID tentativos o su rechazo. Los espectros con anotaciones incorrectas se inspeccionaron primero verificando su calidad y, luego, buscando posibles signos de oxidación (por ejemplo, pérdidas neutras, cambios en la masa de acilos grasos). Con base a esta inspección, se determinaron y compararon formas nativas de lípidos oxidados a través de CMM. Los espectros obtenidos se inspeccionaron manualmente y se compararon con los espectros de las formas nativas obtenidas mediante el análisis del estándar, si se dispone, o de bases de datos. Por último, se realizó una inspección final de los espectros MS/MS en ambos modos de ionización para confirmar la oxidación comparando el patrón de fragmentación con la forma precursora.

3. Resultados

3.1. Estudio de los productos de oxidación

A partir de los estándares utilizados, se estudiaron los principales productos derivados de la reacción de oxidación. Para ello se trabajó con dos metodologías que se combinaron. Por una parte, se estudió el uso o no de un método de extracción y, por otro lado, se tomaron alícuotas a distintos tiempos de reacción de oxidación, concretamente a las 24, 48 y 72 horas, y se analizaron. Todos estos procedimientos se aplicaron de forma combinada y, se realizaron por triplicado para asegurarnos de los resultados obtenidos.

La Figura 3 muestra los cromatogramas obtenidos del análisis de los productos de oxidación de PAPC, PC 16:0/20:4. El panel A muestra los cromatogramas de iones totales (TIC) obtenidos de oxPAPC con el método de extracción planteado a los distintos tiempos de seguimiento establecidos. El panel B muestra los TIC de los productos de oxidación de PAPC sin método de extracción final, superponiendo de igual manera las alícuotas tomadas a distintos tiempos de reacción. Lo que se puede observar inicialmente, es la diferencia en el perfil que nos proporciona uno u otro método según llevemos a cabo la extracción o no. Aunque podríamos pensar en un perfil más pobre derivado del tratamiento de muestra con extracción, las señales obtenidas en el método B podrían ser derivadas de productos secundarios indeseados y/o artefactos. Por lo tanto, para tomar una decisión que nos permitiese seguir con el estudio, se decidió utilizar los datos MS/MS y seguir una estrategia como en este caso, de búsqueda del número de posibles precursores que presentaban el fragmento característico de la cabeza de glicerofosocolina, $m/z = 184.07$, en ionización positiva. Esto generó como resultados, 260 precursores para el método B, sin extracción, y solo 2 para el método A, con extracción, lo que nos ayudó a concluir que con el método de extracción perdíamos la mayor parte de los productos buscados frente al método sin extracción.

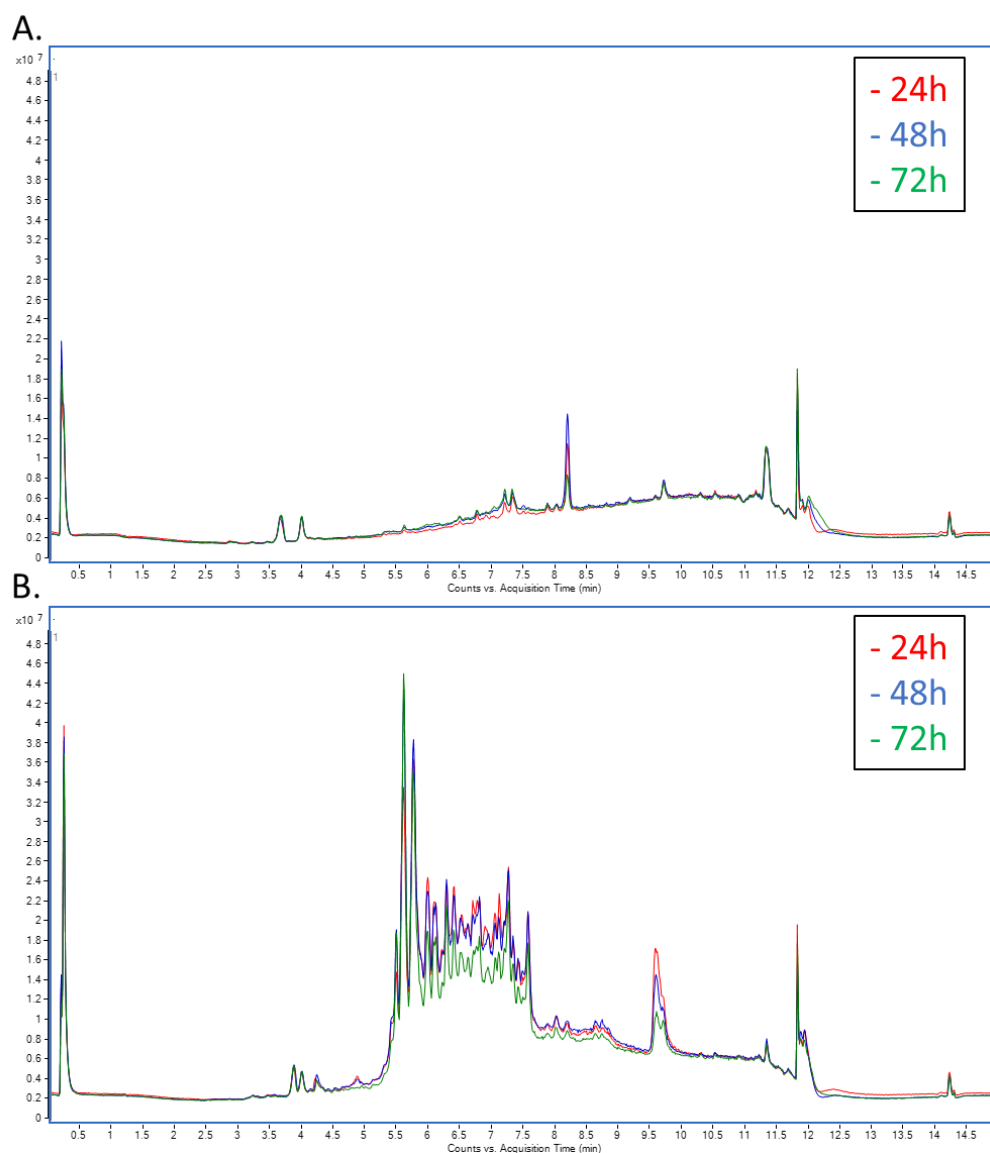


Figura 3. Cromatograma de iones totales (TIC) resultado de la oxidación de PAPC (PC 16:0/20:4) para los distintos métodos empleados. **A.** TICs resultantes de la oxidación de PAPC que incluye la extracción con MTBE. **B.** TICs resultantes para oxPAPC con tratamiento de muestra sin extracción. Los métodos completos se resumen en el apartado “Materiales y métodos”. Ambos métodos muestran superpuestos los TIC para el seguimiento de la reacción de oxidación tomando alícuotas cada 24 (rojo), 48 (azul) y 72 horas (verde).

En cuanto al seguimiento del tiempo de oxidación, también se pudo observar que no se destacaban remarcadas diferencias entre los 3 tiempos. No éramos capaces de distinguir zonas en las que vayan apareciendo progresivamente más señales y desapareciendo otras. Bien es cierto, que se pudo observar una ligera diferencia de intensidad según los tiempos de incubación de la oxidación, pero de forma global, aparecen, prácticamente, las mismas señales.

Este trabajo ha permitido identificar gran parte de los productos de oxidación que se conocen, derivados de los principales mecanismos de reacción para cada estándar comercial utilizado. La Figura 4 muestra el EIC de los productos de oxidación derivados de una de las principales glicerofosocolinas, PAPC, encontradas en muestras humanas. La caracterización de cada señal nos informa de cómo, a medida que aumentamos el grado de oxidación, la especie presenta menor tiempo de retención que su correspondiente especie nativa; y como los productos de reacción de cadena corta se representan en la zona intermedia (5-6 min).

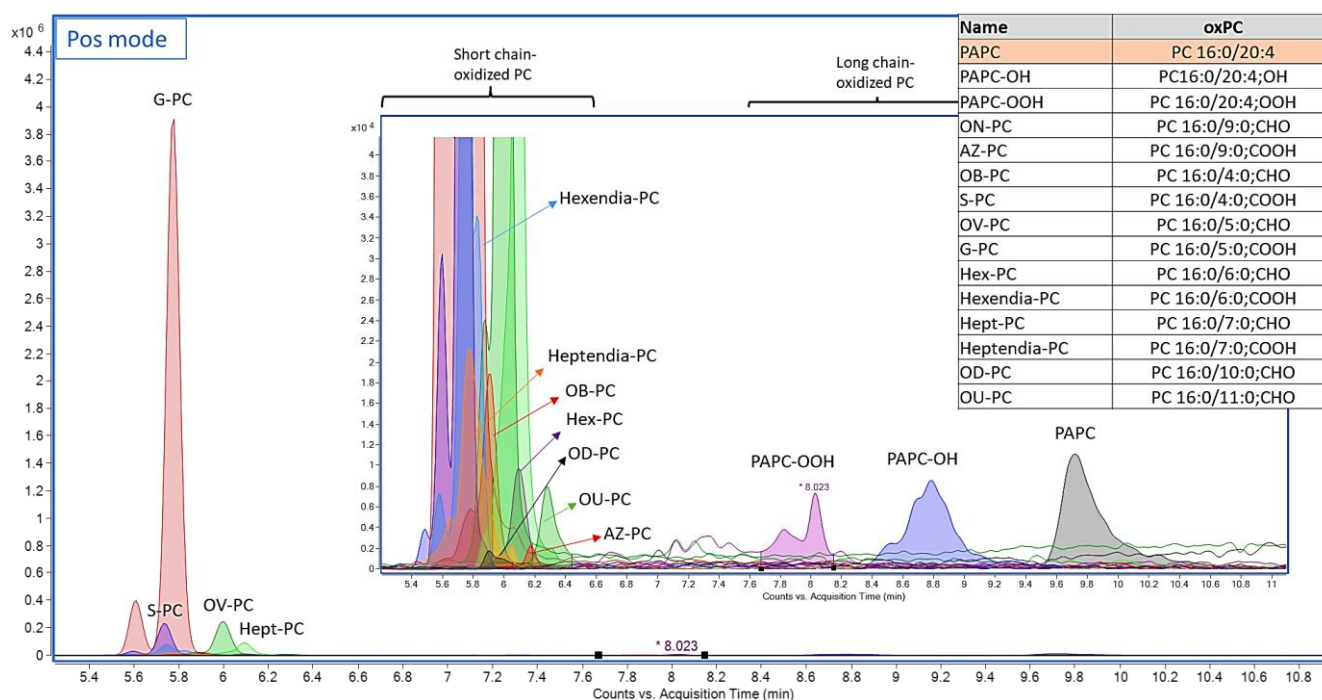


Figura 4. Cromatograma de iones extraídos (EIC) para los productos de oxidación de PC 16:0/20:4 en modo de ionización positivo. La tabla resume las abreviaturas comúnmente empleadas para esas especies. El cromatograma ampliado nos proporciona una imagen de la abundancia y de los tiempos de retención de la forma nativa, y los productos de oxidación derivados de cadena larga y de cadena corta.

3.2. Perfil metabólico en NET de muestras de plasma

Utilizando LC-QTOF/MS se llevó a cabo el estudio metabólico no dirigido de 145 muestras de plasma de individuos sin cáncer y pacientes con NET. El procedimiento de QA se realizó tras la reducción del efecto intra-secuencia con QC-SVRC, limitando la matriz a las características medidas de manera reproducible. La normalidad de los datos se demostró con la prueba de Kolmogórov-Smirnov (K-S) (el 73% de las características en modo positivo y el 85% en modo negativo tenían una distribución normal). El análisis univariante condujo a la selección de 150 y 296 metabolitos

discriminantes en modo positivo y negativo, respectivamente. La anotación provisional permitió asignar 58 compuestos en positivo y 74 en modo negativo, de las cuales 31 y 25 anotaciones se confirmaron en base a la información espectral DIA.

Los perfiles metabólicos de los pacientes con NET fueron claramente diferentes de los individuos sin cáncer. Esto fue apoyado por un gran número de masas estadísticamente significativas y una robusta separación de las muestras incluso en el nivel de PCA (Fig. 5A). Esta fuerte separación se confirmó aún más con el modelo OPLS-DA y sus parámetros de calidad: R^2 : 0.964 y Q^2 : 0.921 (Fig. 5B). La robustez de la clasificación obtenida se confirmó mediante la prueba de permutaciones que descartó la posibilidad de sobreajuste del modelo. Todas las señales discriminatorias y anotadas pertenecen a lípidos. Siguiendo el sistema de clasificación oficial de lípidos (48) se clasificaron en 6 categorías principales, predominando los glicerofosfolípidos, ácidos grasos y esteroides (ácido hidroxioxicólico, ácido ursodesoxicólico y ácido glicocólico). La Fig. 5C muestra el panel lipídico de las alteraciones metabólicas encontradas en muestras de plasma de pacientes con NET. Como se puede observar fácilmente, se detectó una elevación general del contenido de lípidos en el grupo NET. La abundancia de la gran mayoría de los lípidos medidos fue elevada excepto para esfingolípidos, carnitinas, glicerofosfolinas y glicerofosfoetanolaminas. Estos datos muestran las identificaciones finales; sin embargo, es importante discutir la forma en que estos lípidos fueron anotados.

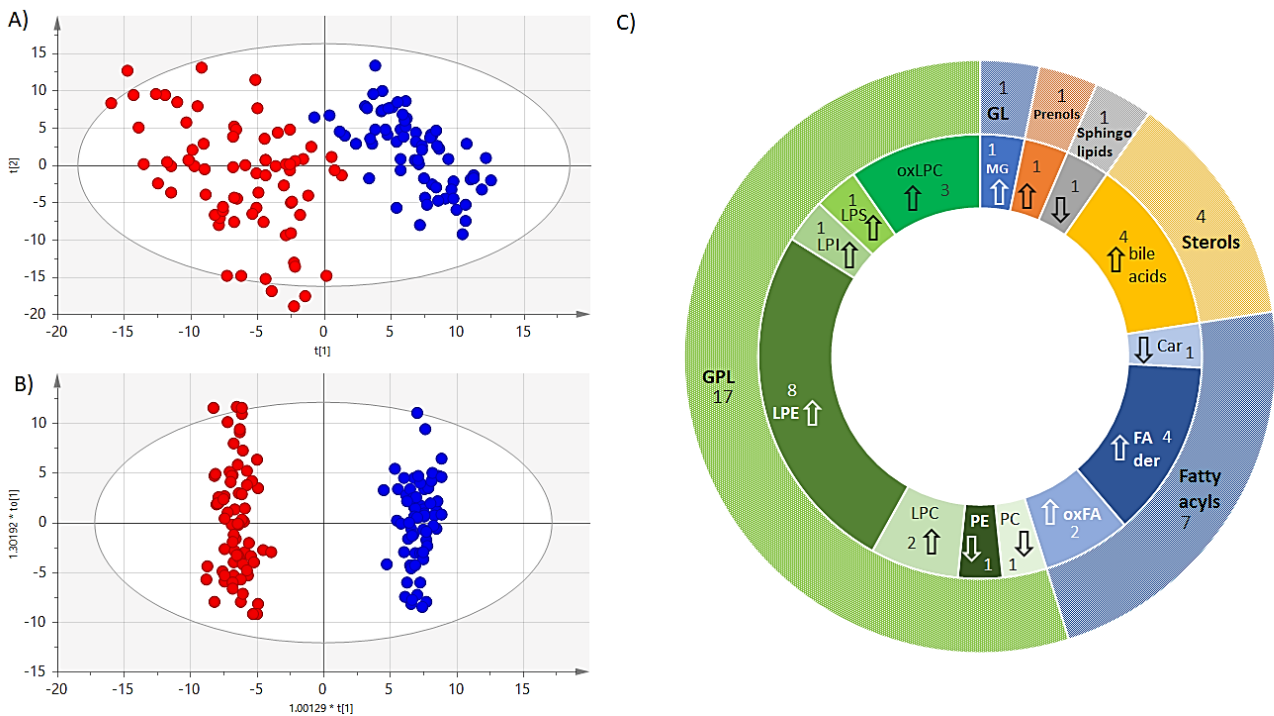


Figura 5. Diferencias metabólicas en los perfiles plasmáticos entre pacientes con NET ($n = 77$) e individuos sin cáncer ($n = 68$). **A.** Modelo PCA construido a partir de datos adquiridos en modo de ionización negativo con escalado UV y transformación logarítmica. Los puntos azules indican pacientes con NET y los rojos representan los controles. Parámetros del modelo: $R^2 = 0.186$ y $Q^2 = 0.156$; **B.** Modelo OPLS-DA construido a partir de datos adquiridos en modo de ionización negativo con escalado pareto y transformación logarítmica. Los puntos azules indican pacientes con NET y los rojos representan los controles. Parámetros del modelo: $R^2 = 0.964$ y $Q^2 = 0.921$; **C.** Gráfico circular de la distribución de lípidos discriminantes y anotados clasificados en las principales categorías (círculo externo) y clases (círculo interno). Las flechas indican los cambios entre los grupos comparados: hacia arriba, un aumento en NET en comparación con los controles; hacia abajo, una disminución en NET en comparación con los controles. Los números indican el número de especies identificadas para cada categoría de lípidos.

Durante el proceso de anotación, se establecieron diferentes enfoques para abordar mejor los desafíos encontrados en señales específicas. La combinación de metodologías, utilizando diferentes métodos y enfoques, aumentan la probabilidad de éxito de una asignación de ID a través del mejor uso de la información disponible, ya sea a nivel MS^1 o MS/MS . De las 46 identificaciones confirmadas en total para ambas polaridades, 31 eran lípidos, y tres de ellas no pudieron verificarse con la coincidencia espectral de la anotación tentativa y fueron identificadas por elucidación estructural. Por lo tanto, a continuación, se presentan los diferentes enfoques utilizados para la

anotación de lípidos oxidados, incluida la coincidencia espectral para metabolitos conocidos y la elucidación estructural de metabolitos modificados.

3.3. Anotación MS/MS de HETEs a través de coincidencia espectral

La anotación de ácidos grasos oxidados se ilustra con el ejemplo de los ácidos hidroieicosatetraenoicos, conocidos como HETE. Este es un metabolito del ácido araquidónico con un grupo hidroxilo. La diferencia en la posición del grupo hidroxilo da como resultado la formación de diferentes isoformas de HETE, por ejemplo, 5-HETE, 12-HETE o 15-HETE. Los HETE, como metabolitos del ácido araquidónico, se han investigado en profundidad y son bien conocidos. Sin embargo, la diversidad de isómeros, la multitud de vías químicas que conducen a su formación y el efecto fisiológico dependiente de la concentración, implican la necesidad de continuar con su investigación. Como ya se mencionó, la investigación sobre HETE tiene una larga tradición y existen ya, varios métodos robustos dedicados exclusivamente a ellos (49-55). Sin embargo, la capacidad de los enfoques no dirigidos para analizar estos lípidos brinda oportunidades únicas para su exploración y la de otros lípidos oxidados, en muchos casos fisiológicos y fisiopatológicos. En este estudio, el método de análisis no dirigido para el estudio del perfil de plasma permitió separar y detectar 5 isoformas de HETE: 5-HETE, 8-HETE, 11-HETE, 12-HETE y 15-HETE (Fig. 6A) en ambas polaridades. Teniendo en cuenta que contábamos con un gradiente muy pronunciado, se obtuvo una buena resolución cromatográfica. La inestabilidad de los lípidos oxidados es uno de los principales retos en el análisis de este grupo de metabolitos. La co-elución se confirmó mediante el análisis MS/MS, donde los espectros obtenidos contenían diferentes fragmentos procedentes de diferentes isoformas juntas. Tal caso se ejemplifica claramente en la Fig. 7B: el espectro presentado contiene fragmentos que se originan a partir de tres isoformas de HETE, que incluyen 12-HETE, 11-HETE y 8-HETE. Al comparar esta información con el orden de elución de HETE (Fig. 6A), tal co-elución puede justificarse fácilmente. La importancia de ambos modos de polaridad en el proceso de anotación es bien conocida. En el caso de los HETE, este punto no es tan crucial. Esto se debe a que se ionizan fácilmente en modo de ionización negativo (Fig. 6B, EIC azul), pero en modo positivo la eficiencia de ionización es muy pequeña (Fig. 6B, EIC rojo).

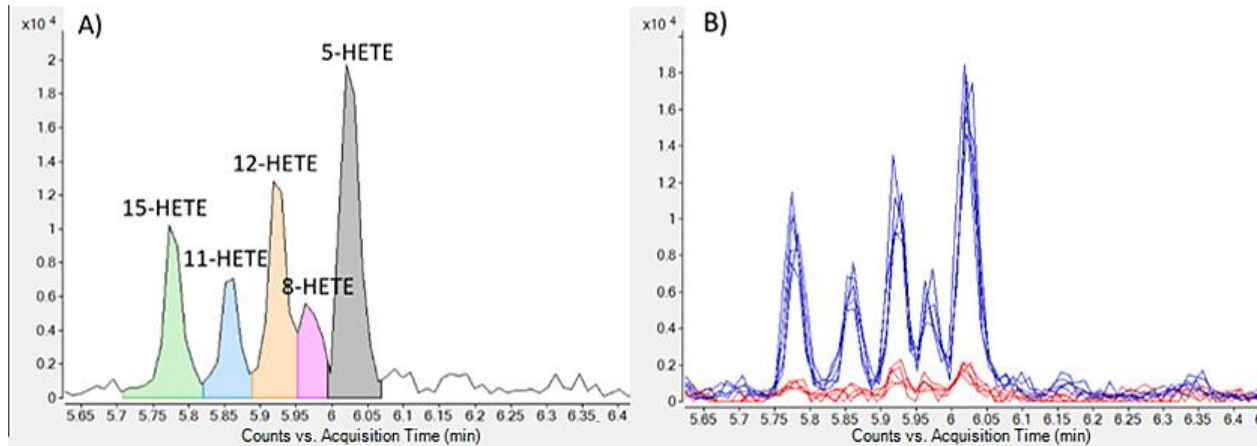


Figura 6. Cromatogramas de iones extraídos (EIC) de HETE en muestras de plasma de NET. **A.** EIC de 5 isoformas de HETE resueltas cromatográficamente en modo de ionización negativa. **B.** EIC de HETE en modo de ionización positivo para $[M + H]^+$ (color rojo) y en modo negativo (azul) para $[M - H]^-$.

Además, en según qué condiciones de análisis, se observa predominantemente la pérdida neutra de agua o los aductos de sodio frente a una señal imperceptible de los iones protonados. Sin embargo, la falta de señal en un modo iónico particular podría ser tan informativa como la presencia de un cierto aducto. Como se mencionó anteriormente, los HETE son un grupo de lípidos bastante conocido, por lo que es fácil encontrarlos en las bases de datos. Además, la mayoría de las isoformas de HETE tienen información de fragmentación que se puede utilizar para la coincidencia espectral. La Fig. 7A muestra un ejemplo de 5-HETE.

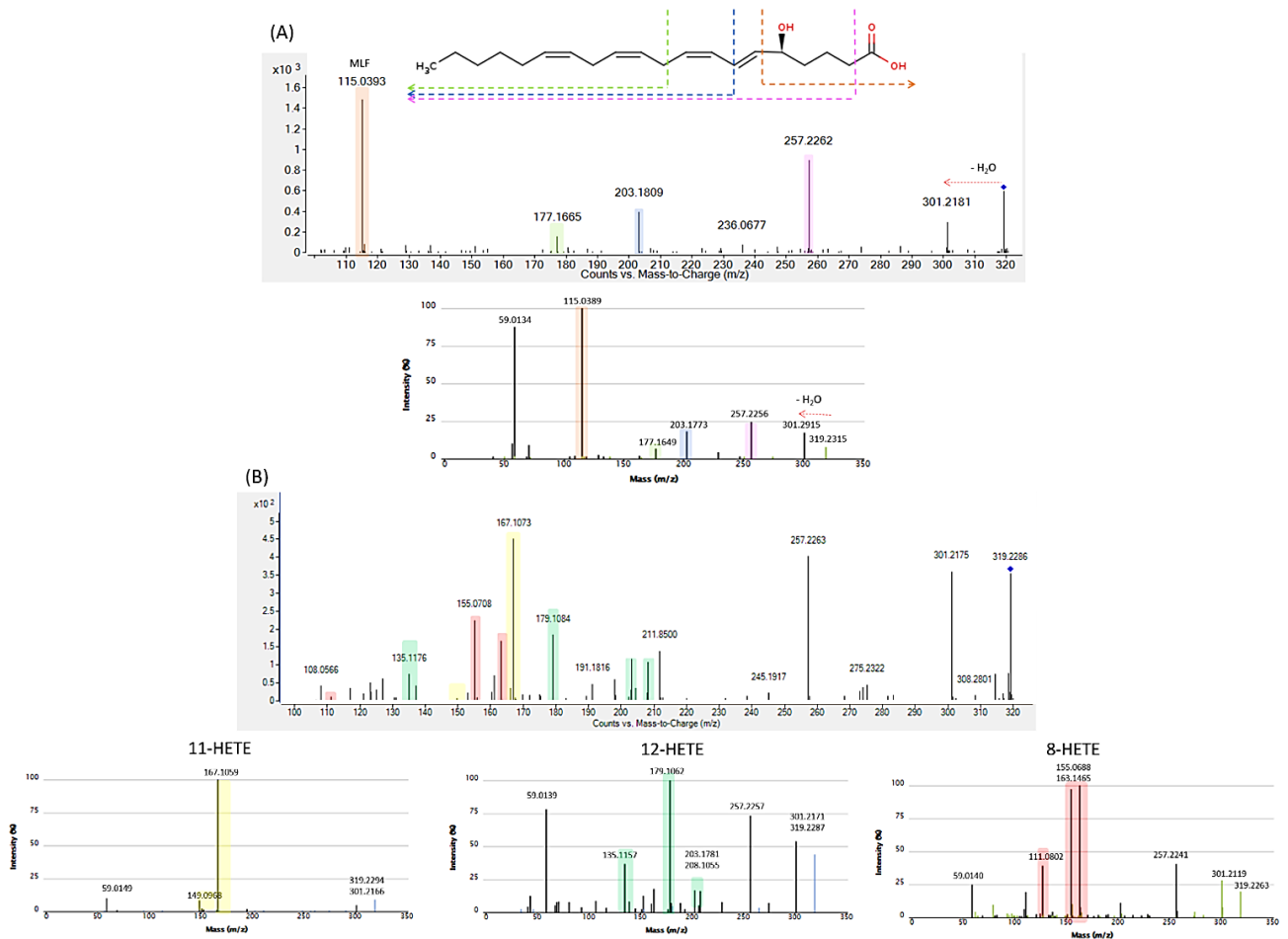


Figura 7. El panel (A) muestra el espectro experimental MS/MS de 5-HETE y el panel inferior presenta los espectros de la base de datos METLIN MS/MS; flechas indican fragmentos y sus colores corresponden a las marcas en los espectros. (B) Ejemplo de un espectro MS/MS de co-elución de tres isoformas de HETE: 12-HETE, 11-HETE y 8-HETE. La parte superior muestra el espectro experimental, mientras que la parte inferior muestra tres espectros de fragmentación de la base de datos METLIN MS/MS. Todos los espectros se adquirieron en modo negativo aplicando 20 eV. MLF: fragmento específico de especie de lípidos moleculares.

Las siguientes Figuras (Fig. 8-11) muestran las anotaciones restantes para los isómeros HETE en los datos experimentales de MS/MS en plasma. La similitud estructural de los HETE da como resultado una analogía espectral entre diferentes isoformas. Comúnmente, hay iones producto que se forman en diferentes HETE, por ejemplo: m/z 301.21 formado a través de la pérdida de agua de su precursor m/z 257.22, generado, a su vez, por la pérdida de un grupo carboxílico. A pesar de estas similitudes, hay muchos fragmentos de MLF que ayudan en el proceso de anotación diferenciando los isómeros. Los fragmentos de MLF que diferencian entre las isoformas de HETE incluyen: m/z 115.03 para 5- HETE, m/z 155.06 y 163.14 para 8-HETE, m/z 167.10 para 11-HETE, m/z 179.10 para 12-HETE y m/z 219.13 para 15-HETE. La mayoría de los HETE producen espectros similares con la excepción del 11-HETE que produce muy pocos fragmentos y, por lo tanto, tiene un espectro bastante diferente. Los HETE, como grupo, proporcionan espectros más informativos bajo energías de colisión pequeñas (entre 10 y 20 eV) porque las energías altas conducen a una sobrefragmentación.

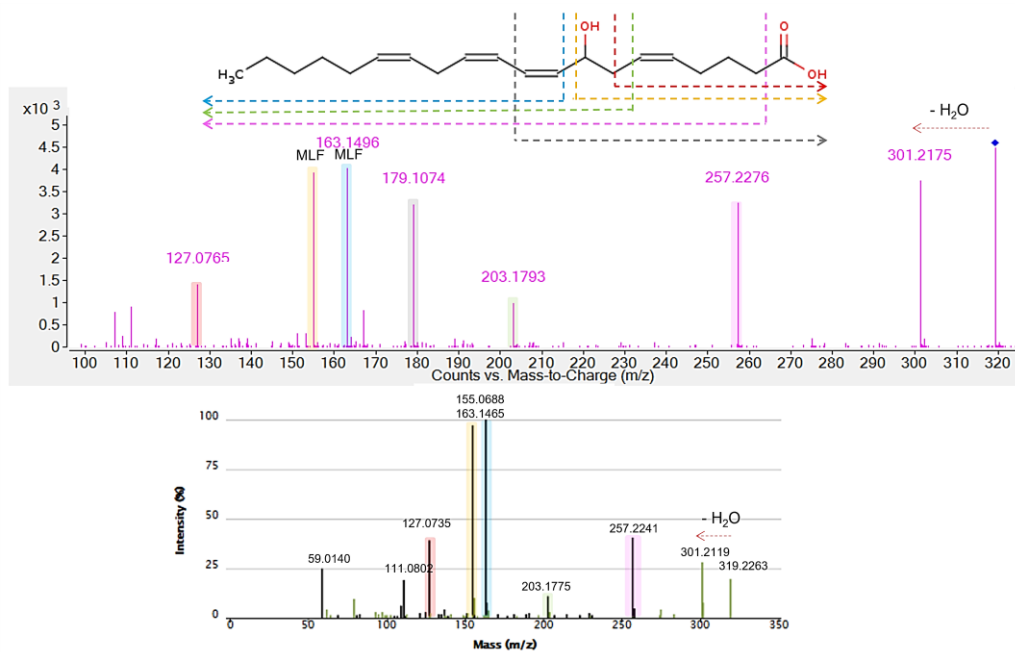


Figura 8. Panel superior: espectro de fragmentación experimental de 8-HETE presente en las muestras de plasma. Panel inferior: espectro de la base de datos METLIN MS/MS. Las flechas indican fragmentos y sus colores corresponden a las marcas en los espectros. Todos los espectros se adquirieron en modo negativo aplicando 20 eV. MLF: fragmento específico de especie de lípidos moleculares.

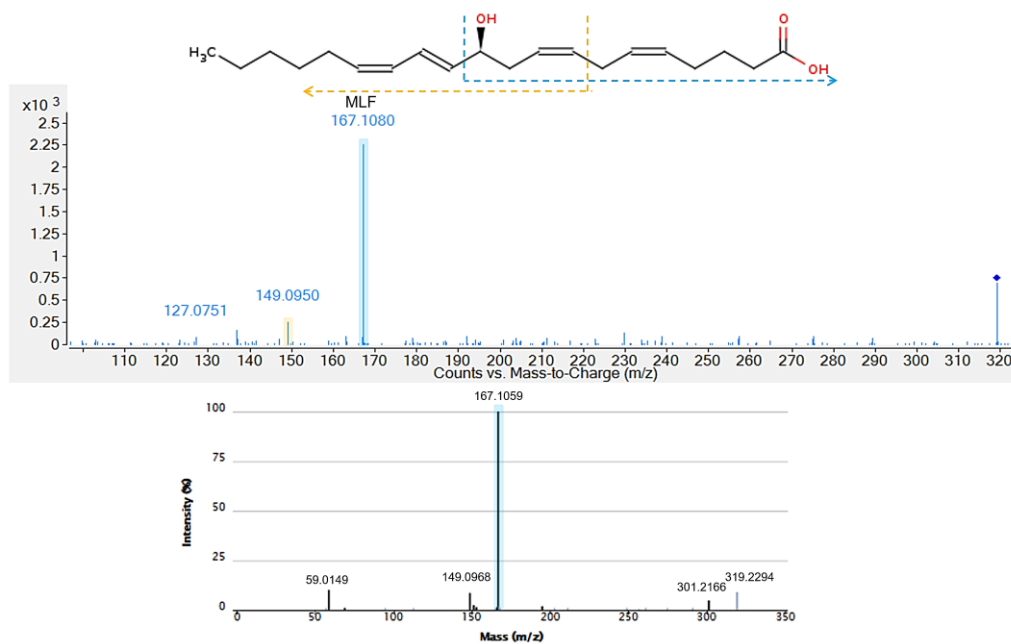


Figura 9. Panel superior: espectro de fragmentación experimental de 11-HETE presente en las muestras de plasma. Panel inferior: espectro de la base de datos METLIN MS/MS. Las flechas indican fragmentos y sus colores corresponden a las marcas en los espectros. Todos los espectros se adquirieron en modo negativo aplicando 20 eV. MLF: fragmento específico de especie de lípidos moleculares.

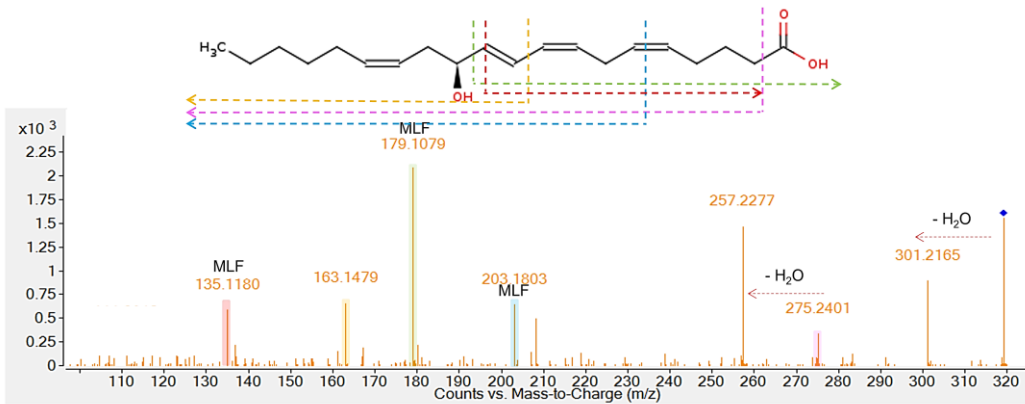


Figura 10. Panel superior: espectro de fragmentación experimental de 12-HETE presente en las muestras de plasma. Panel inferior: espectro de la base de datos METLIN MS/MS. Las flechas indican fragmentos y sus colores corresponden a las marcas en los espectros. Todos los espectros se adquirieron en modo negativo aplicando 20 eV. MLF: fragmento específico de especie de lípidos moleculares.

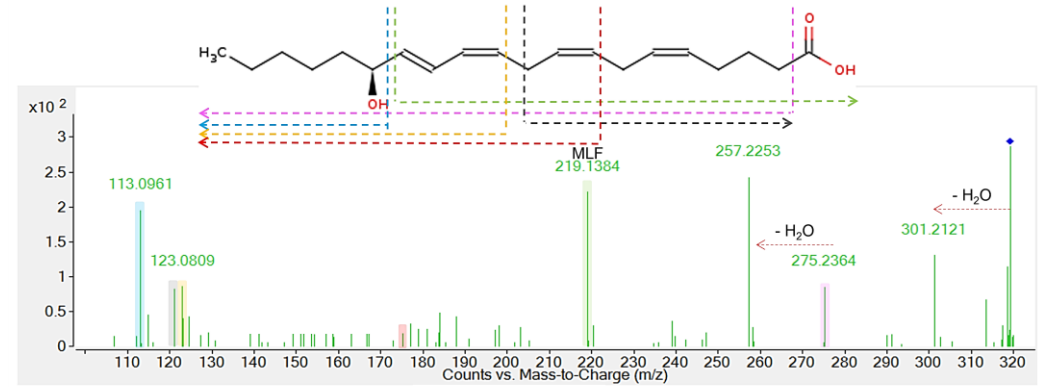
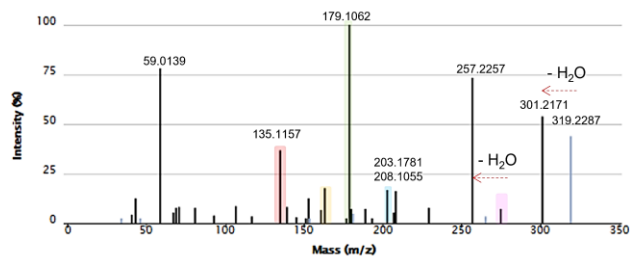
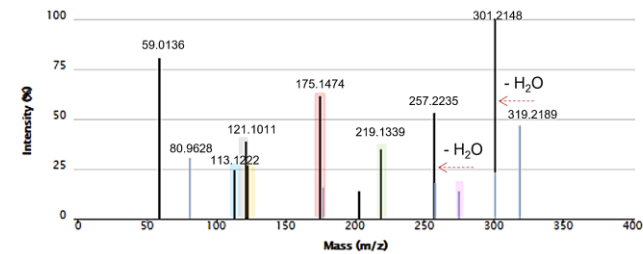


Figura 11. Panel superior: espectro de fragmentación experimental de 15-HETE presente en las muestras de plasma. Panel inferior: espectro de la base de datos METLIN MS/MS. Las flechas indican fragmentos y sus colores corresponden a las marcas en los espectros. Todos los espectros se adquirieron en modo negativo aplicando 20 eV. MLF: fragmento específico de especie de lípidos moleculares.



3.4. Elucidación estructural MS/MS de lisoglicerofosfolinas oxidadas (oxLPC)

Todavía existían tres anotaciones no confirmadas por la coincidencia espectral que requerían un análisis profundo y elucidación de los espectros MS/MS. El estudio de sus correspondientes espectros de fragmentación concluyó que estas incógnitas se referían a glicerofosfolípidos oxidados. En los últimos años se ha observado un mayor interés por las glicerofosfolinas oxidadas, como primeros productos de oxidación, lo que ha supuesto un enorme aumento del conocimiento adquirido sobre estos lípidos, tanto desde una perspectiva analítica como biológica. En marzo de 2018 se reportaron 26 especies oxidadas catalogadas en la base de datos LIPID MAPS (38), 260 especies en febrero de 2022. Sin embargo, se sabe muy poco sobre las monoacil- o monoalquil-glicerofosfolinas oxidadas (oxLPC). La falta de oxLPC en las bases de datos impide que incluso los enfoques más robustos para la anotación provisional logren obtener una asignación correcta. Como se mencionó anteriormente, tres lípidos se anotaron erróneamente basándose únicamente en la información de MS¹: $m/z = 512.3354$ Da a RT 4.21 min, $m/z = 540.3658$ Da a RT 4.70 min y $m/z = 536.3344$ Da a RT 4.41 min. Todas estas características se obtuvieron del análisis en modo de ionización positivos y, todas ellas fueron anotadas como fosfatidilserinas por CMM a través de información de MS¹: LPS O-18:0, LPS O-20:0 con pérdida de agua y LPS O-20:0, respectivamente, de acuerdo con el flujo de trabajo explicado en la Fig.2.

Se obtuvo un posible candidato adicional de PC 16:1/2:0 para $m/z = 536.3344$ Da. Una inspección de los espectros de fragmentación rechazó rápidamente la anotación tentativa y permitió clasificar los tres iones como lisoglicerofosfolípidos oxidados: LPC 16:0;OH, LPC 18:0;OH y LPC 18:2;OH. La evidencia y el razonamiento que nos llevaron a estas anotaciones se describen a continuación.

- *LPC 16:0;OH: emparejamiento del espectro experimental MS/MS con el espectro MS/MS de la forma no oxidada del estándar de referencia puro*

El m/z estadísticamente significativo, 512.3354 Da y RT 4.21 min se asignó provisionalmente como LPS O-18:0 a través de CMM. Una vez que se obtuvo el espectro experimental MS/MS, se pudo comprobar que esa asignación tentativa inicial era incorrecta, debido a la falta de iones producto característicos de las glicerofososerinas. Por su parte, si se pudo observar la presencia de iones producto característicos de las glicerofosfolinas en el espectro (Fig. 12). Tras la exclusión de posibles coeluciones, se consideró la presencia de una posible oxidación. De esta

forma, las posibles oxidaciones se calcularon con la diferencia de masa entre las formas oxidadas y nativa (no oxidada): 15.9949 Da para $-\text{OH}$, 31.9898 Da para $-(\text{OH})_2$ o $-\text{OOH}$ (20). Las masas recalculadas de formas no oxidadas para $-(\text{OH})_2$ / $-\text{OOH}$ y $-\text{OH}$ condujeron a la obtención, tras la búsqueda en CMM, de LPC O-16:1 y LPC 16:0 como posibles candidatos por $[\text{M}]^+$ respectivamente. Si bien solo se observó un candidato con pérdida de agua, y se sabe que $-(\text{OH})_2$ presenta una doble pérdida de agua (18.0108 y 36.0216 Da), no se pudo descartar esta opción, pues si bien esta pérdida neutra probablemente proviene de la forma liso, la segunda pérdida sólo es posible observarla con energías de colisión altas. Para confirmar el producto de oxidación, los espectros MS/MS para ambas polaridades se compararon con los espectros del estándar nativo (Fig. 12).

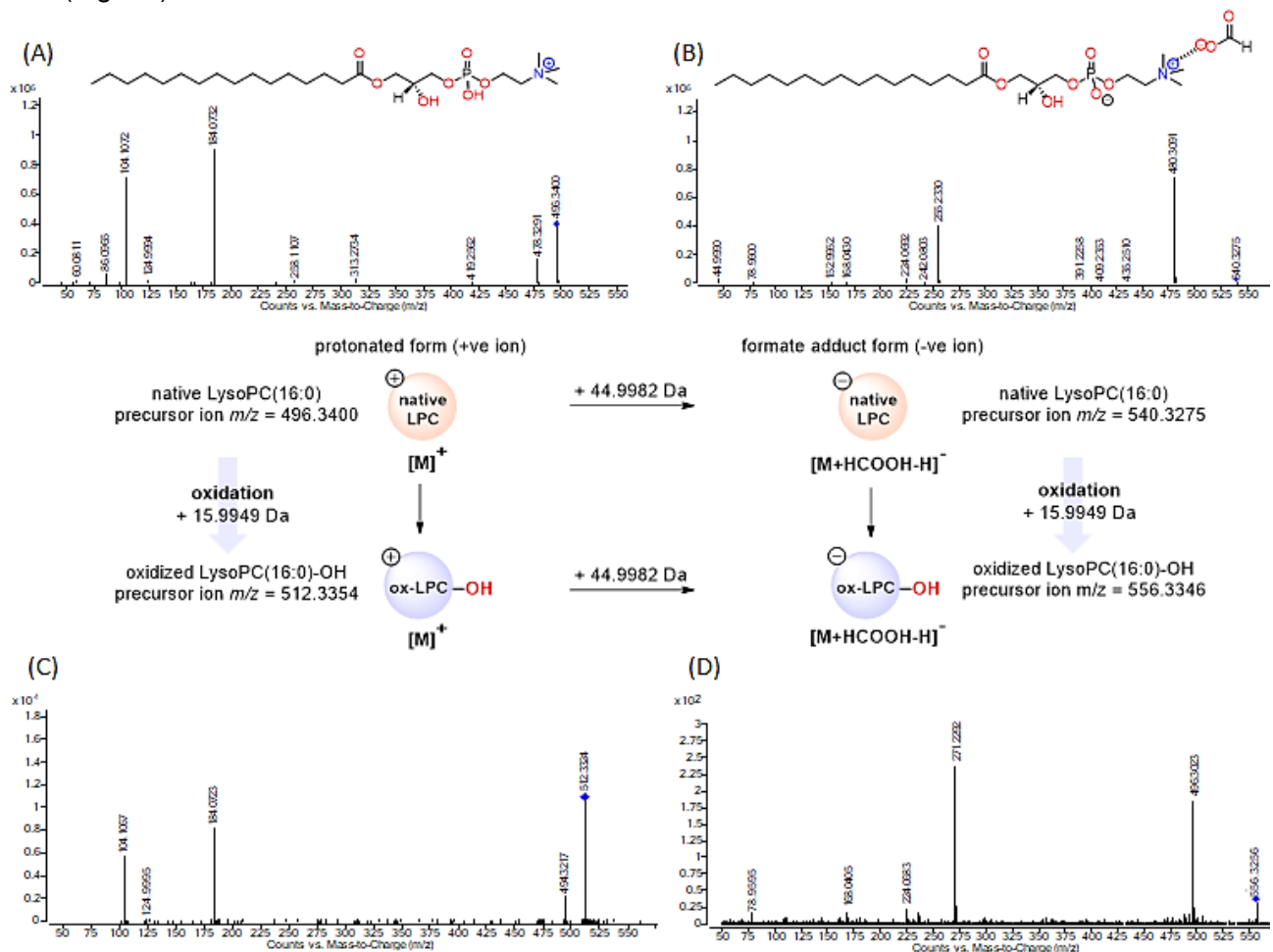


Figura 12. Comparación del patrón de fragmentación basado en el mecanismo de fragmentación (LCF y MLF) y las diferencias de masa en MLF. Los espectros superiores muestran los espectros MS/MS resultantes del LPC en estado nativo (16:0) a 20 eV adquirido en modo de ionización positivo (A) y en modo de ionización negativo (B). Los espectros inferiores ilustran los espectros experimentales de MS/MS de la forma oxidada LPC 16:0;OH presente en las muestras de plasma a la misma energía de colisión (20 eV) en modo de ionización positivo (C) y negativo (D). LCF: fragmento selectivo de la clase de lípidos; MLF: fragmento específico de especie de lípido molecular.

Los fragmentos característicos de estos espectros fueron relevantes para la identificación de las cadenas de acilo graso del glicerofosfolípido. En la región de masa media, el fragmento característico de la cadena de acilo graso se puede observar en modo negativo, y el resultado de la base de datos para $m/z = 271.2291$ (MLF) fue ácido hidroxihexadecanoico. Además, las MLF características pueden compararse entre los espectros de LPC nativos y oxidados. Se mostró una diferencia de masa de 15.9949 Da, la diferencia entre nuestro precursor y la forma nativa y, se correspondía con la variación de masa causada por la presencia de un grupo $-OH$. Todos los demás LCFs permanecen sin cambios, como ya se mencionó, ya que estos solo brindan información sobre la clase de lípidos.

Al unir la información proporcionada por los datos de MS/MS en ambos modos de ionización, se identificó LPC 16:0;OH como la característica significativa con $[M]^+ = 512.3354$ a RT 4.21 min.

Se han llevado a cabo estudios mecanísticos de las vías de fragmentación de lípidos para una amplia gama de estas moléculas, lo que demuestra que la comprensión de las reacciones de fragmentación del ion precursor es esencial para la caracterización estructural (56-58). Estos estudios han dedicado especial atención a estandarizar las reacciones de lípidos en condiciones de disociación inducida por colisión (CID) y han demostrado que la CID de estas moléculas se produce a través de dos mecanismos predominantes, fragmentación por migración de carga (CMF), un mecanismo que implica la carga del ion precursor, y fragmentación por retención de carga (CRF), mecanismo que tiene lugar físicamente lejos del sitio de carga. Estos mecanismos de fragmentación producen iones de fragmentos comunes y predecibles a partir de moléculas de lípidos que tienen diversas estructuras químicas (59).

La caracterización estructural ha sido respaldada con rutas de fragmentación propuestas basadas en estudios previos para los compuestos objetivo, como se muestra en las siguientes figuras (Fig. 13-15), donde se representan los iones característicos y las pérdidas neutras del LPC estándar 16:0 en modo ESI negativo y positivo.

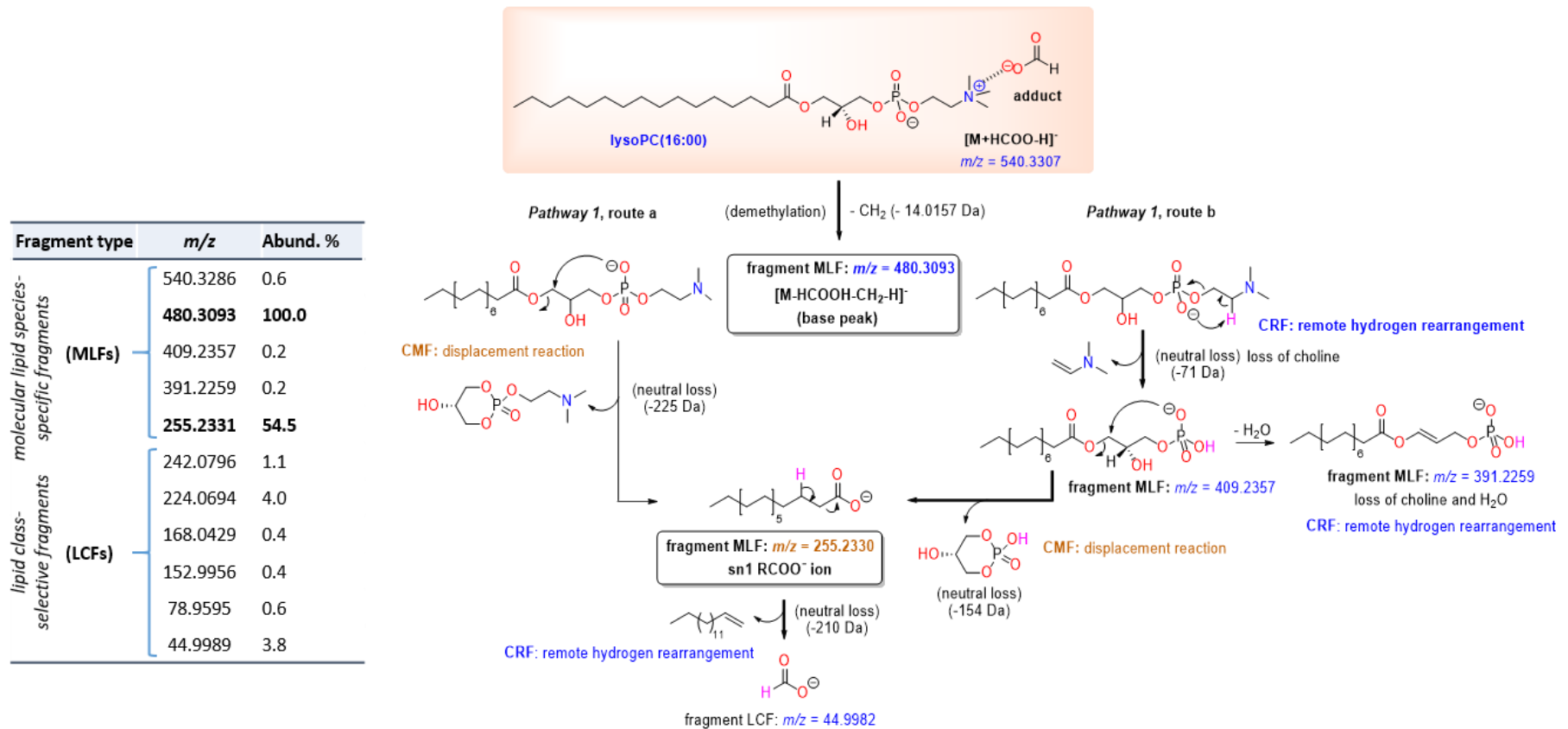
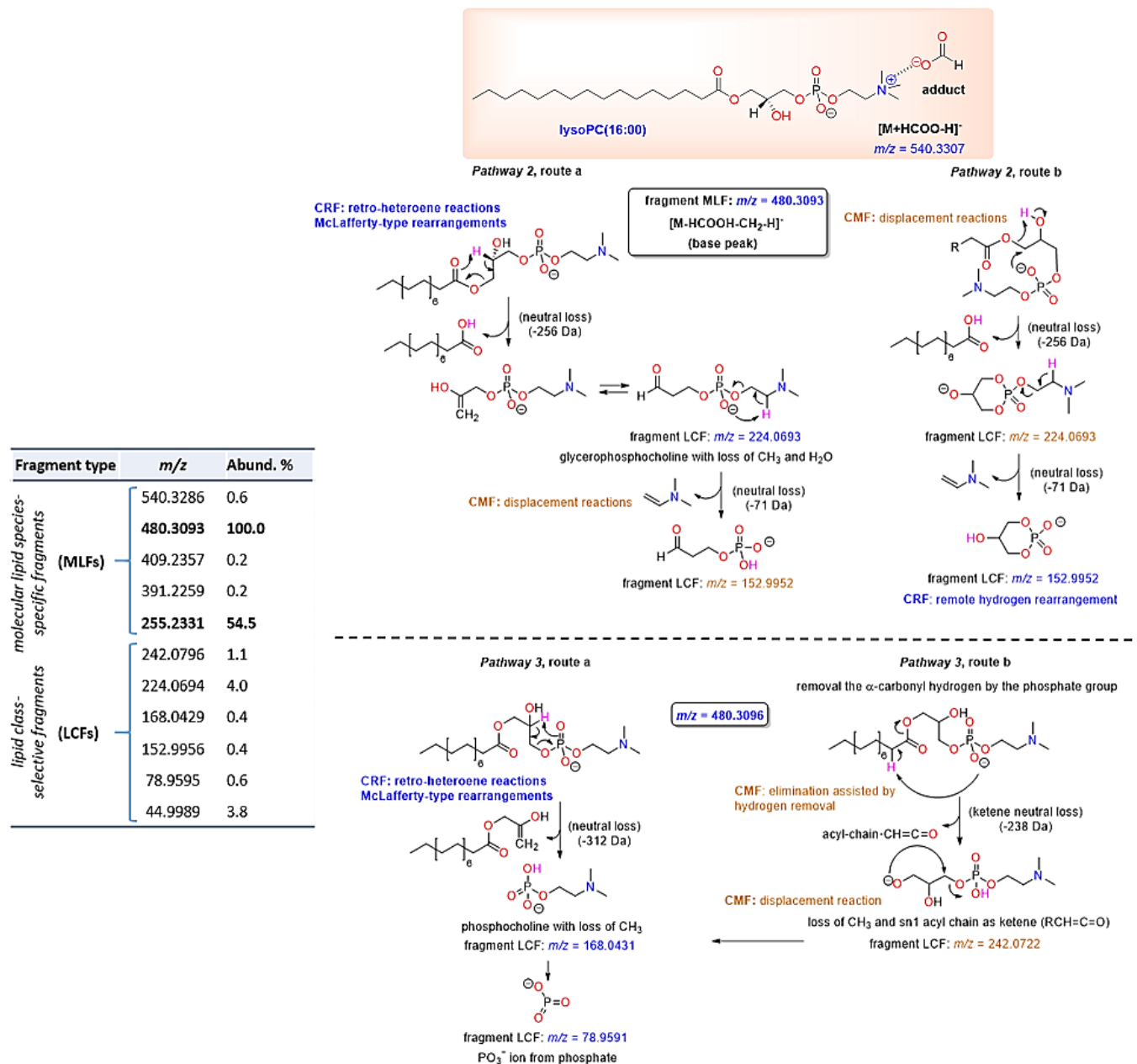
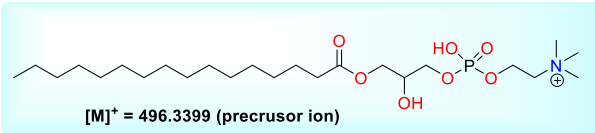


Figura 13. Ruta de fragmentación mayoritaria 1 propuesta en ionización negativa de LPC 16:0. Competencia entre reordenamiento de tipo McLafferty (*ruta a*) y una reacción de desplazamiento (*ruta b*) para producir fragmentos específicos de especies de lípidos moleculares (MLF) característicos.

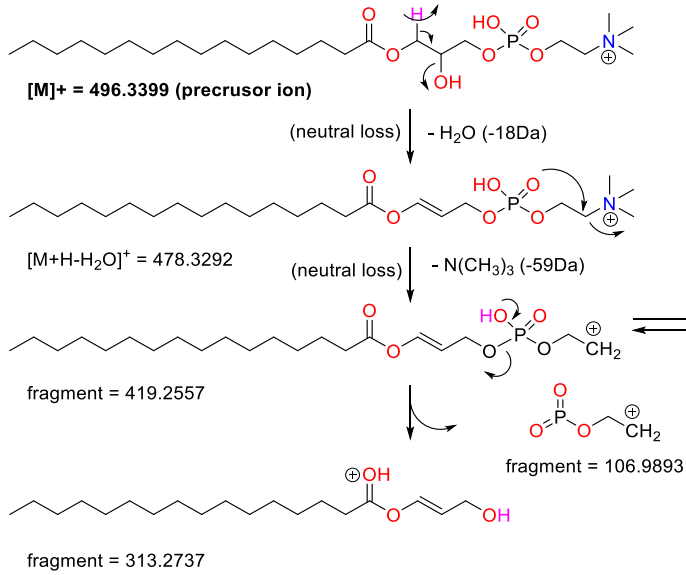


Fragment type	<i>m/z</i>	Abund. %
molecular lipid species-specific fragments (MLFs)	540.3286	0.6
	480.3093	100.0
	409.2357	0.2
	391.2259	0.2
	255.2331	54.5
lipid class-selective fragments (LCFs)	242.0796	1.1
	224.0694	4.0
	168.0429	0.4
	152.9956	0.4
	78.9595	0.6
	44.9989	3.8

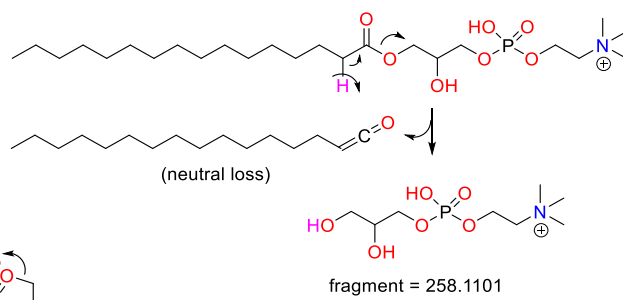
Figura 14. Vías de fragmentación secundarias 2 y 3 propuestas para ionización negativa de LPC 16:0. Competencia entre reordenamiento de tipo McLafferty (*ruta a*) y la reacción de desplazamiento o eliminación asistida por supresión de hidrógeno (*ruta b*) para producir LCF comunes (fragmentos selectivos de la clase de lípidos).



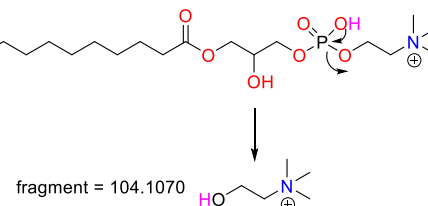
Pathway a



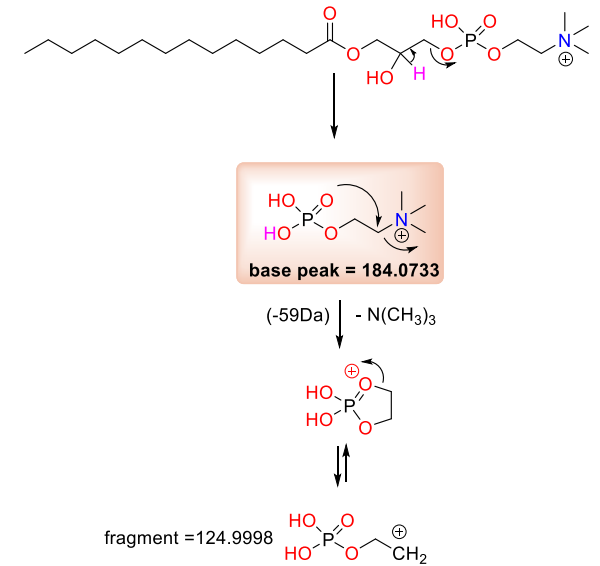
Pathway b



Pathway e



Pathway c



Pathway d

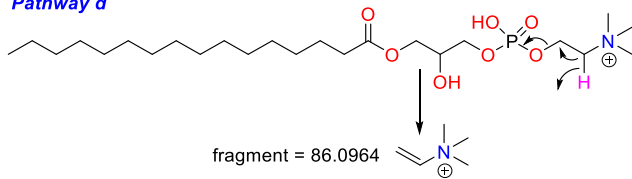


Figura 15. Propuesta de rutas de fragmentación en ionización positiva para LPC 16:0

- *LPC 18:0;OH*; emparejamiento del espectro experimental MS/MS con el espectro MS/MS de la forma no oxidada del estándar de referencia puro

Como en el caso anterior, la característica con $m/z = 540.3658$ Da y RT 4.70 min, asignada de forma tentativa como LPS O-20:0 por $[M]^+$, fue estudiada en detalle mediante la inspección de datos MS/MS experimentales. Debido a que los fragmentos característicos de las glicerofosocolinas dominaban el espectro MS/MS en modo positivo, se estudiaron posibles modificaciones del precursor. Extrayendo las posibles oxidaciones e inspeccionando los fragmentos obtenidos en modo negativo, se pudo establecer con relativa facilidad la composición exacta de los ácidos grasos, en base al espectro MS/MS adquirido en modo de ionización negativa. A través del espectro MS/MS del estándar nativo, los fragmentos característicos podrían compararse entre la forma oxidada y la no oxidada (Fig. 16).

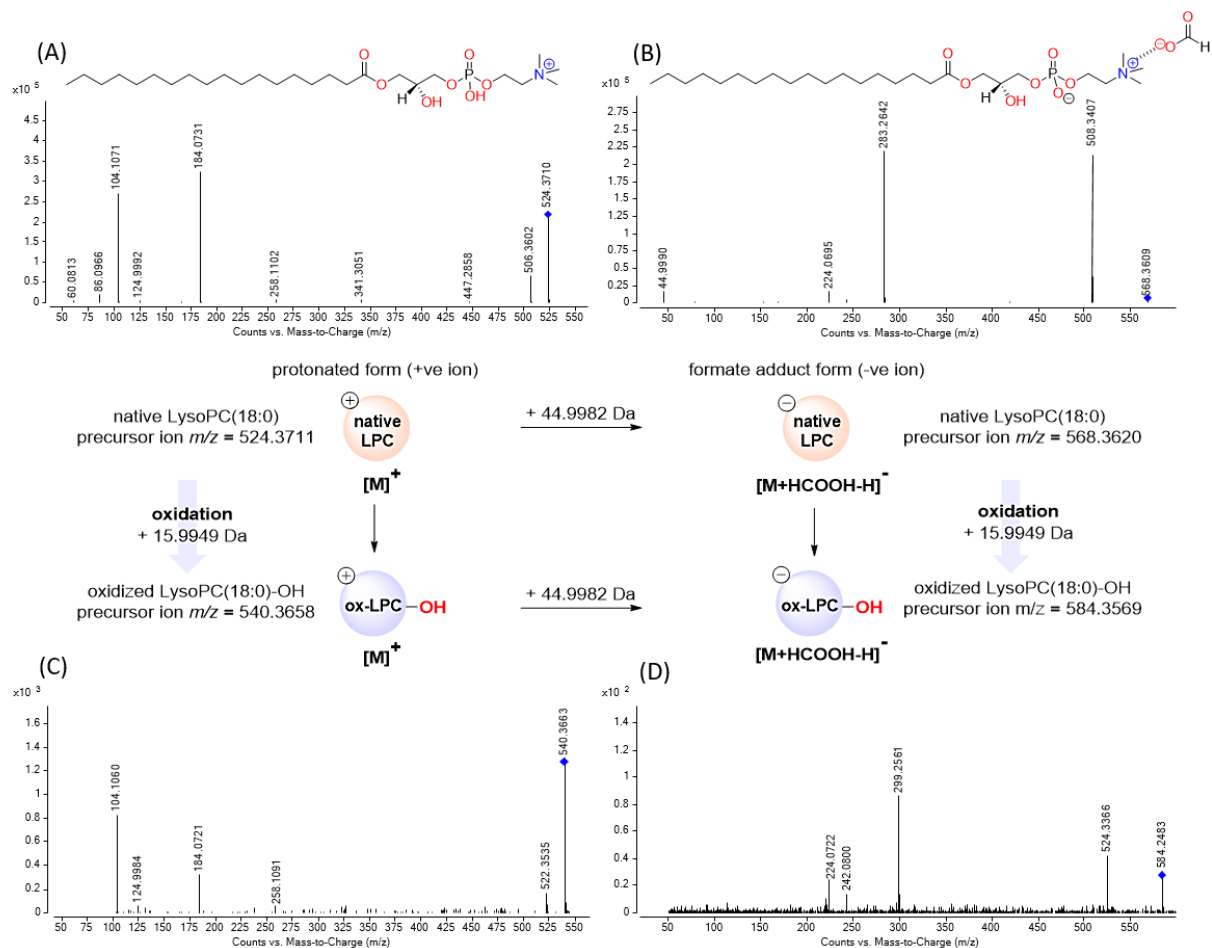


Figura 16. Comparación del patrón de fragmentación basado en el mecanismo de fragmentación (LCF y MLF) y las diferencias de masa en MLF. Los espectros superiores muestran los espectros MS/MS resultantes del estándar nativo LPC 18:0 a 20 eV, adquiridos en modo positivo (A) y en modo negativo (B). Los espectros inferiores ilustran los espectros experimentales MS/MS de la forma oxidada LPC 18:0;OH presente en las muestras de plasma con la misma energía de colisión (20 eV) en modo de ionización positivo (C) y negativo (D).

Las señales correspondientes al ácido graso desprotonado o a la forma liso desmetilada se pudieron observar en la región de masa media; por lo que, con la señal de $m/z = 299.2596$ (MLF) correspondiente al ácido hidroxiocetadecanoico, se identificó la molécula como LPC 18:0;OH.

- *LPC 18:2;OH: comparación del espectro experimental MS/MS con los disponibles en las bases de datos*

Otro caso fue el de la característica con $m/z = 536.3344$ Da y RT 4.41 min. La anotación tentativa se realizó de acuerdo con el flujo de trabajo explicado en la Fig. 2 y, así, se le asignaron los candidatos PC16:1/2:0 con $[M]^+$ o LPS 20:0 por $[M + H - H_2O]^+$. Aunque PS se descartó por completo porque se identificaron fragmentos característicos del grupo de cabeza de fosfocolina en el espectro MS/MS (modo de ionización positivo), también se pudieron observar dos pérdidas de agua neutra en la región de masa alta (Fig. 17A). Como el espectro no coincidía con la asignación tentativa y, se podían observar dos pérdidas de agua en modo positivo, se consideró la oxidación de ácidos grasos. De la misma manera que en el caso anterior, se calcularon y buscaron posibles precursores, y se confirmó LPC 18:2;OH, con el fragmento relacionado con el ácido graso en la región de masa media en el espectro MS/MS experimental en modo negativo (Fig. 17C). En este caso, se utilizó el espectro MS/MS de la forma nativa en modo de ionización negativo procedente de una base de datos en línea, para comparar sus fragmentos característicos (60) (Fig. 17B). En el modo de ionización positivo, se pudo asignar las señales pertenecientes a las pérdidas de agua a la forma liso, y otra al grupo -OH. Además, la pérdida relacionada con el grupo principal/aducto, identificada como pérdida de formiato de metilo, es una pérdida relevante para el reconocimiento de aductos en datos MS/MS en modo negativo. También se puede observar que las estructuras oxidadas con dobles enlaces conjugados pueden perder fácilmente agua como pérdida neutra.

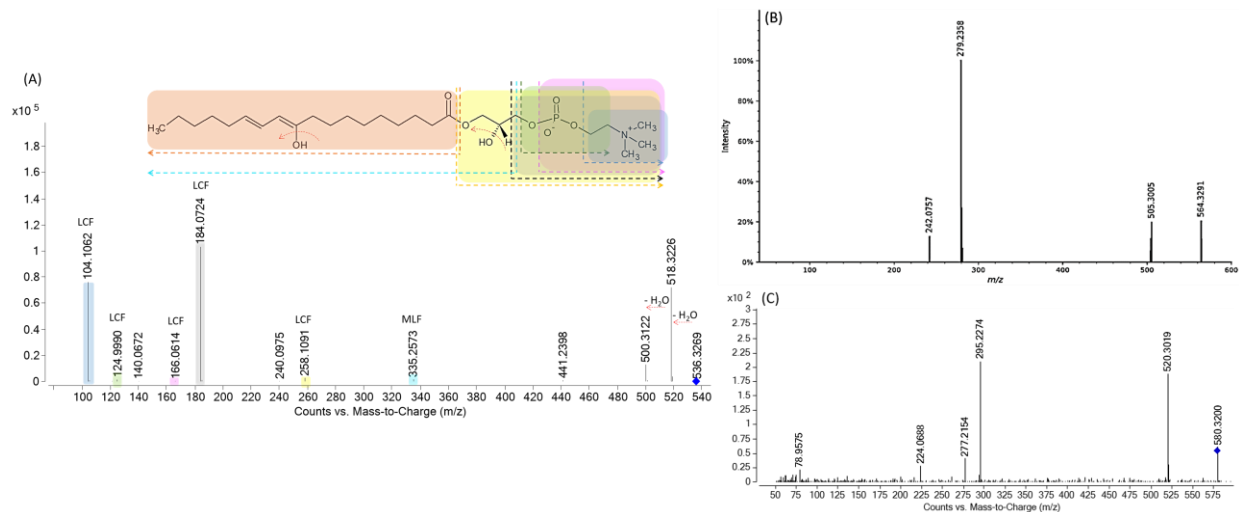


Figura 17. **A.** Espectro MS/MS experimental de LPC 18:2;OH de muestras de plasma adquiridas en modo positivo a 25 eV. **B y C.** Comparación del espectro MS/MS experimental de la forma oxidada LPC 18:2;OH con el patrón de fragmentación basado en el espectro nativo en negativo de la base de datos MassBank. El espectro superior muestra el espectro de la base de datos MS/MS del precursor nativo LPC 18:2 adquirido en modo de ionización negativo (**B**). El espectro inferior ilustra el espectro MS/MS experimental de la forma oxidada LPC 18:2;OH de una muestra de plasma a 20 eV en modo negativo (**C**).

3.5. Implicaciones biológicas de la oxidación lipídica en pacientes con NET

Aunque la relación patológica del estrés oxidativo con el cáncer ha sido bien establecida, la caracterización de los productos de oxidación de lípidos generados bajo esta condición no está bien desarrollada hasta la fecha. Hay estudios en los que se ha utilizado la espectrometría de masas para perfilar el lipidoma oxidado en muestras biológicas de pacientes con cáncer (61-66). Sin embargo, muchas publicaciones suelen proporcionar resultados contradictorios que podrían simplemente confirmar la compleja naturaleza de los lípidos oxidados (67, 68). Se ha demostrado que diferentes isoformas de la misma forma nativa pueden desempeñar diferentes funciones fisiológicas. Además, un solo lípido oxidado puede exhibir propiedades muy diferentes o incluso opuestas dependiendo de su concentración. Muy a menudo, las bajas concentraciones dan como resultado propiedades protectoras, mientras que un aumento en la concentración implica efectos pro-enfermedad. Esto demuestra claramente la necesidad de explorar más a fondo el papel de los lípidos oxidados.

Los derivados oxidados del ácido araquidónico se relacionaron con la modulación de los efectos de los factores de crecimiento en las células cancerosas, los efectos co-mitogénicos, la angiogénesis, la adhesión, la migración y la invasión de las células tumorales. Los procesos bioquímicos de los mecanismos de propagación y

migración están relacionados con estos productos de oxidación y son esenciales para la metástasis. El papel y la actividad de los HETE han sido investigados en diferentes tipos de cáncer, incluido el cáncer de próstata (69, 70), el adenocarcinoma esofágico humano (EAC) (71), el cáncer de mama (72), el cáncer gástrico (73) y el cáncer de pulmón (74), pero nunca ha sido estudiado en pacientes con NET. Independientemente de la isoforma, se ha encontrado que todos los HETE están significativamente elevados en pacientes con NET en comparación con individuos sin cáncer (Fig. 18A). También, se ha observado la misma tendencia para otros lípidos oxidados, como las oxLPC (Fig. 18B).

La formación de fosfolípidos oxidados es bien conocida y abarca la oxidación enzimática altamente controlada y los eventos no enzimáticos no controlados. La producción de lisofosfolípidos oxidados aún no se ha explorado profundamente, pero algunas fuentes brindan información sobre tres vías principales: a) fragmentación oxidativa de glicerofosfolípidos a través de la pérdida de un grupo acilo de un fosfolípido oxidado, b) esterificación de ácidos grasos oxidados a glicerol derivados (75) o c) enzimáticamente por LOX, COX o citocromo P450. Estos productos de oxidación están involucrados en la regulación de los procesos de inflamación, trombosis, angiogénesis, inmunotolerancia y funcionamiento de la barrera endotelial, ya que se alteran las propiedades físicas de fluidez y empaquetamiento de acilo. Si bien se ha estudiado el papel de HETE en el cáncer, no se sabe mucho sobre los oxLPC. La correlación estadística entre los oxLPC anotados con los HETE podría sugerir una relación biológica entre ambos tipos de productos de oxidación. Esto también implica la formación de dichos oxilípidos y, podría indicar su origen enzimático. Sin embargo, para confirmar esta relación y su naturaleza, se necesita más investigación. Los lípidos oxidados son diferentes entre el grupo sin cáncer y los pacientes con NET; sin embargo, también se observaron algunas tendencias dentro del grupo NET. Esta heterogeneidad se ilustra en dos parámetros clínicos, como son la actividad proliferativa tumoral y la cantidad de cromogranina A. Los NET se estratifican por actividad proliferativa en grado 1 y grado 2, donde el grado 1 significa actividad proliferativa baja (índice proliferativo Ki67 < 3%) y actividad proliferativa intermedia de grado 2 (índice proliferativo Ki67 de 3-20%). La Fig. 18C ilustra las diferencias en el nivel de HETE entre los grados. Como puede verse, con una mayor actividad de proliferación, el nivel plasmático de HETE está disminuyendo. Se sabe que los HETE promueven la proliferación de células tumorales (76, 77), por lo tanto, podrían acumularse localmente en el tejido tumoral, lo que a su vez da como resultado la disminución de su nivel en la sangre. La cromogranina A (CgA)

es la principal proteína soluble que se libera junto con hormonas peptídicas, neurotransmisores o aminas en respuesta a diversos estímulos fisiológicos y farmacológicos. Ubicada en las vesículas secretoras de las neuronas y células endocrinas, la CgA y sus diversos fragmentos bioactivos ejercen un amplio espectro de actividades reguladoras al influir en los sistemas endocrino, neurológico, cardiovascular e inmunológico, y afecta a la homeostasis de la glucosa y el calcio. La CgA y sus péptidos derivados se utilizan como marcadores de diagnóstico y pronóstico o para controlar la respuesta a la intervención farmacológica no solo en tumores neuroendocrinos, sino también en enfermedades inflamatorias, cardiovasculares y neuropsiquiátricas. En la Figura 18D se muestra una disminución de HETE en el grupo 2 donde se superó el límite superior del nivel de referencia de cromogranina A, con un aumento particularmente grande en el grupo 1 de 5-HETE. Todos estos hallazgos brindan una nueva e importante perspectiva sobre el metabolismo de los lípidos en los NET; sin embargo, se requiere más investigación. Las diferencias capturadas en el nivel de lípidos oxidados en plasma son muy prometedoras en el contexto de la investigación sobre los tumores NET. Además, para comprender completamente la función de estos productos de oxidación, se requiere un análisis funcional a través del análisis combinado de lípidos y proteínas.

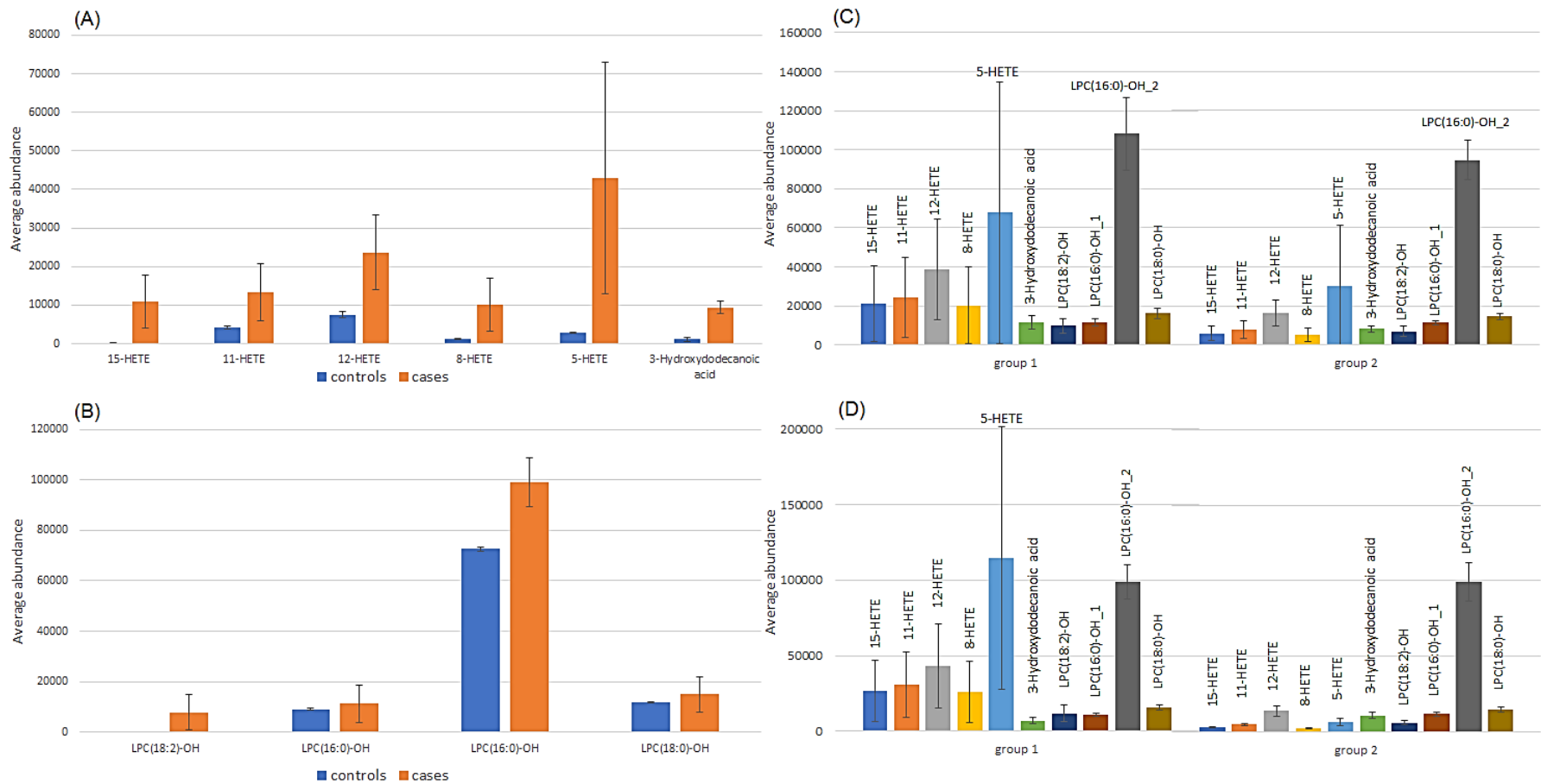


Figura 18. Gráficos de barras que ilustran la dirección y la magnitud del cambio de ácidos grasos oxidados (A) y oxLPC (B). Las barras azules corresponden al grupo sin cáncer (n = 68) mientras que las barras naranjas representan pacientes con NET (n = 77). Las barras de error representan el intervalo de confianza con el 95%. Los paneles C y D muestran la distribución de los lípidos oxidados en pacientes con NET estratificados según la actividad proliferativa del tumor (C: el grupo 1 significa grado 1 (n = 26) y el grupo 2 corresponde al grado 2 (n = 51)) y niveles de referencia de cromogranina A (D: el grupo 1 significa por debajo del doble del límite superior normal (n = 26) y el grupo 2 el doble por encima del límite superior normal (n = 51)).

3.6. Estudio lípidos oxidados en NSCLC

En el análisis dirigido para la búsqueda de diferencias significativas entre los distintos grupos comparados en lípidos oxidados presentes en muestras de plasma y tejido de NSCLC, se han encontrado importantes resultados preliminares que continúan siendo analizados para intentar ofrecer la mayor información posible.

Los análisis de plasma mostraron diferencias significativas en lípidos como PC 16:0/20:4;OOH, PC 16:0/20:3;OH o PC 18:0/18:2;OH en las comparativas ADC vs control y SCC vs control. El estudio comparativo LCC vs control no reportó diferencias significativas en cuanto a lípidos oxidados. Se ha podido observar con los resultados de estos primeros análisis preliminares que, las especies mayoritarias que se encuentran alteradas en los datos procedentes de plasma son lípidos oxidados de cadena larga, y que las formas hidroperóxido se encuentran disminuidas en el grupo cáncer, sin importar el subtipo, mientras que el resto, presentan niveles aumentados (Figura 19).

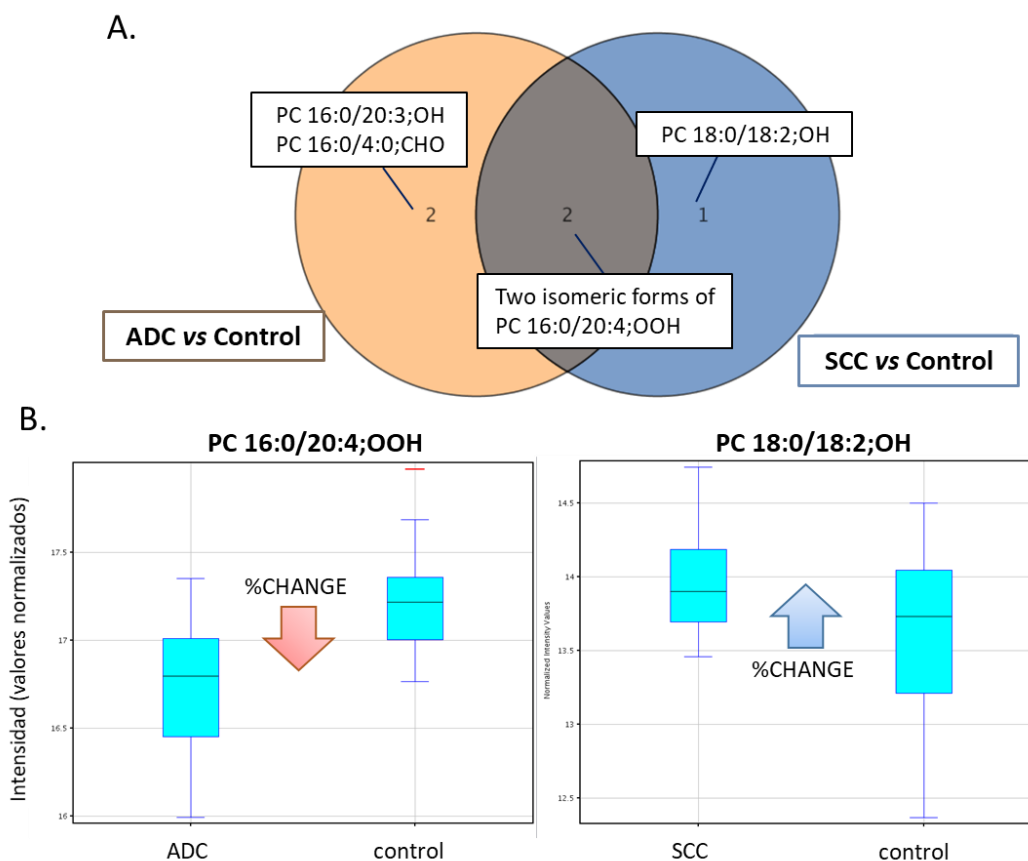


Figura 19. Resultados obtenidos del análisis de lípidos oxidados en plasma de NSCLC. **A.** Diagrama de Venn con la comparación de los resultados obtenidos del estudio estadístico según el subtipo de cáncer de pulmón (ADC vs control, y SCC vs control). El estudio comparativo LCC vs control no proporcionó resultados significativos en cuanto a lípidos oxidados **B.** Diagrama de

caja y bigotes para las abundancias obtenidas de los derivados oxidados PC 16:0/20:4;OOH y PC 18:0/18:2;OH en ambas comparativas

En cuanto a los resultados obtenidos del estudio comparativo con los datos procedentes del análisis de muestras de tejido se concluyó que estos pertenecían en su mayoría a lípidos oxidados de cadena corta. El estudio comparativo según el subtipo de cáncer de pulmón se llevó a cabo comparando los resultados encontrados en muestras de tejido tumorales y de tejido adyacente asignado como control. Además, se realizaron diferentes análisis para observar la tendencia y el comportamiento de las especies lipídicas oxidadas con el progreso de la enfermedad en el estudio comparativo de los estadios de NSCLC (Figura 20). En todos los casos se pudo observar que los niveles de oxPC se encuentran disminuidos en el tejido tumoral frente al control y que esos niveles se van reduciendo a medida que aumenta el estadio. El diagrama de Venn para los resultados de tejido en diferentes subtipos de cáncer se realizó solo con lípidos oxidados.

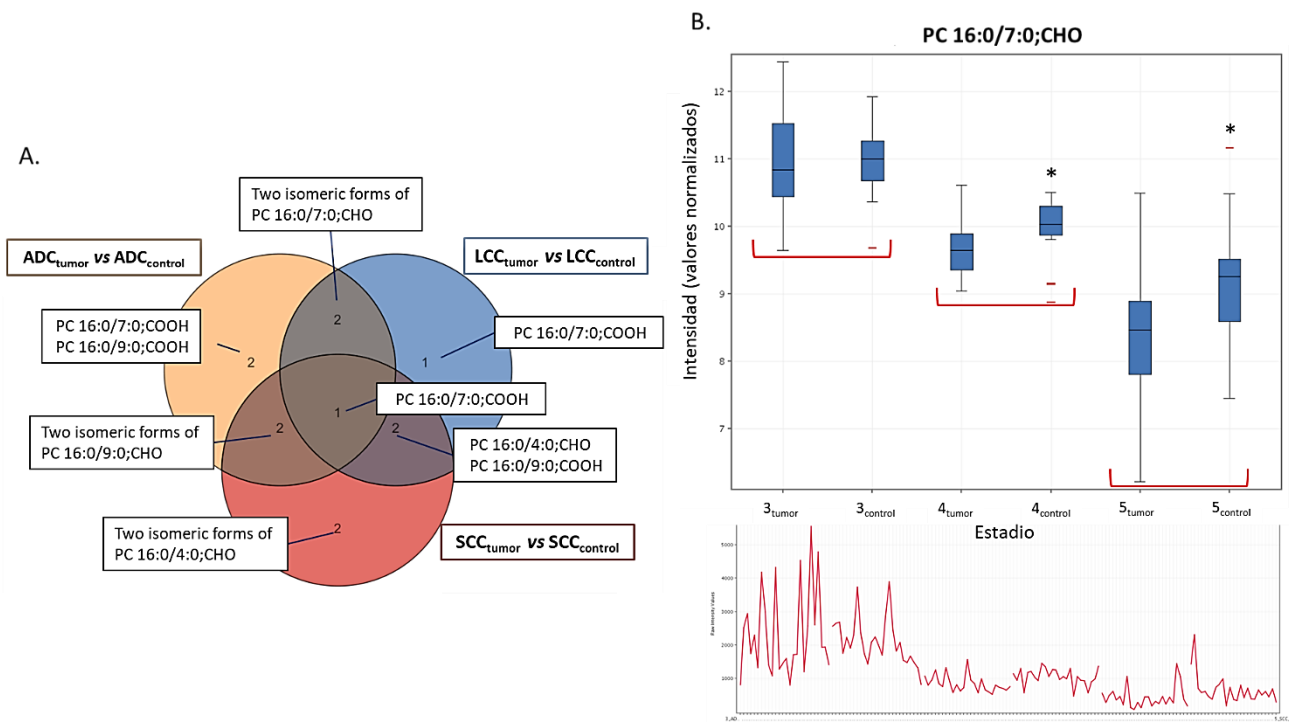


Figura 20. Resultados obtenidos del análisis de lípidos oxidados en tejido de NSCLC. **A.** Diagrama de Venn con la comparación de los resultados obtenidos del estudio estadístico según el subtipo de cáncer de pulmón (ADC_{tumor} vs ADC_{control}, LCC_{tumor} vs LCC_{control} y SCC_{tumor} vs SCC_{control}). **B.** Diagrama de caja y bigotes para las abundancias obtenidas del derivado oxidado PC 16:0/7:0;CHO según los distintos estadios. Este incluye una representación gráfica de la intensidad de esta especie y su distribución en las muestras. El asterisco representa resultados significativos.

Como se observa en el gráfico, 3 lípidos son exclusivos en las comparaciones, 6 lípidos se comparten entre 2 tipos de cáncer (2 para cada combinación) y 1 se comparte entre todos los tipos de cáncer. Los datos del lípido que dio los mejores resultados, PC 16:0/7:0;CHO, se analizó con un diagrama de caja y bigotes según los diferentes estadios y se representaron los diagramas de perfil que muestran los cambios de forma más evidente.

Estos resultados continúan aún en estudio y se están completando con el análisis de otra cohorte independiente de pacientes. Con esos resultados, junto con los colaboradores donde se realizó la movilidad, se continúa trabajando en el desarrollo de una publicación científica que reúna la información más relevante posible.

4. Conclusión

En este trabajo se realizó, por una parte, el perfilado metabólico de muestras de plasma de pacientes con NET y controles que condujo a la selección de lípidos desregulados. Entre los lípidos seleccionados y anotados, se prestó especial atención a los lípidos oxidados, incluida la serie de HETE y los oxLPC anotados. Debido a su importancia biológica, se afrontaron varios desafíos analíticos, especialmente durante el proceso de anotación. Estos desafíos se pueden dividir en dos categorías principales según el tipo de información que se utilizó en la identificación. La anotación MS/MS a través de la coincidencia espectral se realizó para HETE, y la elucidación estructural de MS/MS se llevó a cabo para los oxLPC haciendo la comparación con el espectro del estándar nativo, siempre que fuese posible. Además, en nuestro estudio, se utilizó una combinación de ionización positiva y negativa, para obtener una mejor anotación de confianza, aunque los datos del modo negativo nos proporcionaron más información sobre la cadena de acilo. La anotación tentativa puede ser interesante, pero la información estructural es obligatoria para una anotación inequívoca. Nuestro proceso de elucidación de la estructura para LPC estándar (16:0), basado en datos de modo negativo ESI-MS/MS, mostró las principales reacciones y mecanismos derivados de la participación del grupo de cabeza glicerofosfolina en el proceso de fragmentación. También se proporcionó evidencia de que varios productos de oxidación tienen diferencias significativas entre pacientes con cáncer y personas sin cáncer, y establecimos un flujo de trabajo para identificar dichas modificaciones. Se reportó la regulación positiva de ácidos grasos oxidados y oxLPC en pacientes con NET y, hasta donde sabemos, este es el primer intento de determinar HETE en NET. Además, hay

muy pocos estudios en los que se miden los oxLPC, por lo que los resultados obtenidos proporcionan información importante sobre la oxidación de lípidos en los NET. Sin embargo, las funciones fisiológicas de los lisofosfolípidos oxidados aún deben establecerse y requieren más investigación.

Por otro lado, se llevó a cabo la búsqueda de especies oxidadas derivadas de las principales glicerofosfolinas humanas en un estudio de datos procedentes del análisis de muestras de plasma y tejido de cáncer de pulmón. Los resultados preliminares mostraron la existencia de especies oxidadas discriminantes según el tipo de subtipo cancerígeno, más concretamente para diferencias ADC o SCC frente a control. La comparativa LCC vs control no reveló diferencias en cuanto a oxPCs en plasma. Este estudio forma parte de un estudio preliminar. Además, también se pudo establecer diferencias en cuanto al estadio con muestras con enfermedad avanzada (estadio 3, 4 y 5). Una observación importante es el hecho de que aparecen alteradas, fundamentalmente, oxPC de cadena larga en datos plasmáticos, con formas hidroperóxido que disminuyen en casos, y formas hidróxido aumentadas, y oxPC de cadena corta en los datos procedentes de tejido pulmonar. Estos resultados nos han permitido concluir, también, una marcada disminución en la abundancia de la forma oxidada, a medida que avanza la enfermedad, posiblemente debido a la apoptosis provocada en estadios tan avanzados.

Por último, el análisis de estándares oxidados permite la construcción de una librería que caracteriza cada producto de oxidación con sus tiempos de retención y fragmentos característicos. Todos estos datos podrán reunirse en herramientas como CMM, facilitando al investigador el complicado proceso de anotación, debido al reciente interés que han mostrado las especies oxidadas por su importante rol patológico.

Como resumen, con todo este trabajo se podrá continuar con el desarrollo de una biblioteca que reúna la mayor cantidad de especies identificadas, con sus espectros de fragmentación correspondientes, que simplifique el cuello de botella que supone la identificación de lípidos y, más recientemente, de especies lipídicas oxidadas. Además, se ofrece un flujo de trabajo establecido para la identificación de especies oxidadas sin resultado en las bases de datos disponibles.

La búsqueda de especies oxidadas en distintos tipos de muestras procedentes de estudios de cáncer demuestra la conocida relación entre el estrés oxidativo y esta patología. Estos análisis deben validarse y apoyarse en ensayos complementarios con nuevas cohortes de pacientes que validen los resultados.

5. Referencias

1. González-Riano C, Dudzik D, Garcia A, Gil-de-la-Fuente A, Gradillas A, Godzien J, et al. Recent Developments along the Analytical Process for Metabolomics Workflows. *Anal Chem.* 2020;92(1):203-26.
2. Ni Z, Angelidou G, Hoffmann R, Fedorova M. LPPtiger software for lipidome-specific prediction and identification of oxidized phospholipids from LC-MS datasets. *Scientific reports.* 2017;7(1):1-14.
3. Wishart DS, Guo A, Oler E, Wang F, Anjum A, Peters H, et al. HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Res.* 2022;50(D1):D622-d31.
4. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic acids research.* 2018;46(D1):D608-D17.
5. Kishimoto K, Urade R, Ogawa T, Moriyama T. Nondestructive Quantification of Neutral Lipids by Thin-Layer Chromatography and Laser-Fluorescent Scanning: Suitable Methods for "Lipidome" Analysis. *Biochemical and Biophysical Research Communications.* 2001;281(3):657-62.
6. Fahy E, Sud M, Cotter D, Subramaniam S. LIPID MAPS online tools for lipid research. *Nucleic Acids Research.* 2007;35(suppl_2):W606-W12.
7. Kodali ST, Kauffman P, Kotha SR, Yenigalla A, Veeraraghavan R, Pannu SR, et al. Oxidative Lipidomics: Analysis of Oxidized Lipids and Lipid Peroxidation in Biological Systems with Relevance to Health and Disease. In: Berliner LJ, Parinandi NL, editors. *Measuring Oxidants and Oxidative Stress in Biological Systems.* Cham: Springer International Publishing; 2020. p. 61-92.
8. Hama H. Fatty acid 2-Hydroxylation in mammalian sphingolipid biology. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids.* 2010;1801(4):405-14.
9. Hardwick J, Song B-J, Huberman E, Gonzalez F. Isolation, complementary DNA sequence, and regulation of rat hepatic lauric acid omega-hydroxylase (cytochrome P-450LA omega). Identification of a new cytochrome P-450 gene family. *Journal of Biological Chemistry.* 1987;262(2):801-10.
10. Miura Y. The biological significance of ω -oxidation of fatty acids. *Proc Jpn Acad Ser B Phys Biol Sci.* 2013;89(8):370-82.
11. Yin H, Xu L, Porter NA. Free Radical Lipid Peroxidation: Mechanisms and Analysis. *Chemical Reviews.* 2011;111(10):5944-72.
12. Reis Pereira A, Spickett C. Chemistry of phospholipid oxidation. *Biochim Biophys Acta.* 2012 Oct;1818(10):2374-87.
13. Bochkov VN, Oskolkova OV, Birukov KG, Levonen AL, Binder CJ, Stöckl J. Generation and biological activities of oxidized phospholipids. *Antioxid Redox Signal.* 2010;12(8):1009-59.
14. Zhong S, Li L, Shen X, Li Q, Xu W, Wang X, et al. An update on lipid oxidation and inflammation in cardiovascular diseases. *Free Radical Biology and Medicine.* 2019;144:266-78.
15. Levitan I, Volkov S, Subbaiah PV. Oxidized LDL: Diversity, Patterns of Recognition, and Pathophysiology. *Antioxidants & Redox Signaling.* 2009;13(1):39-75.

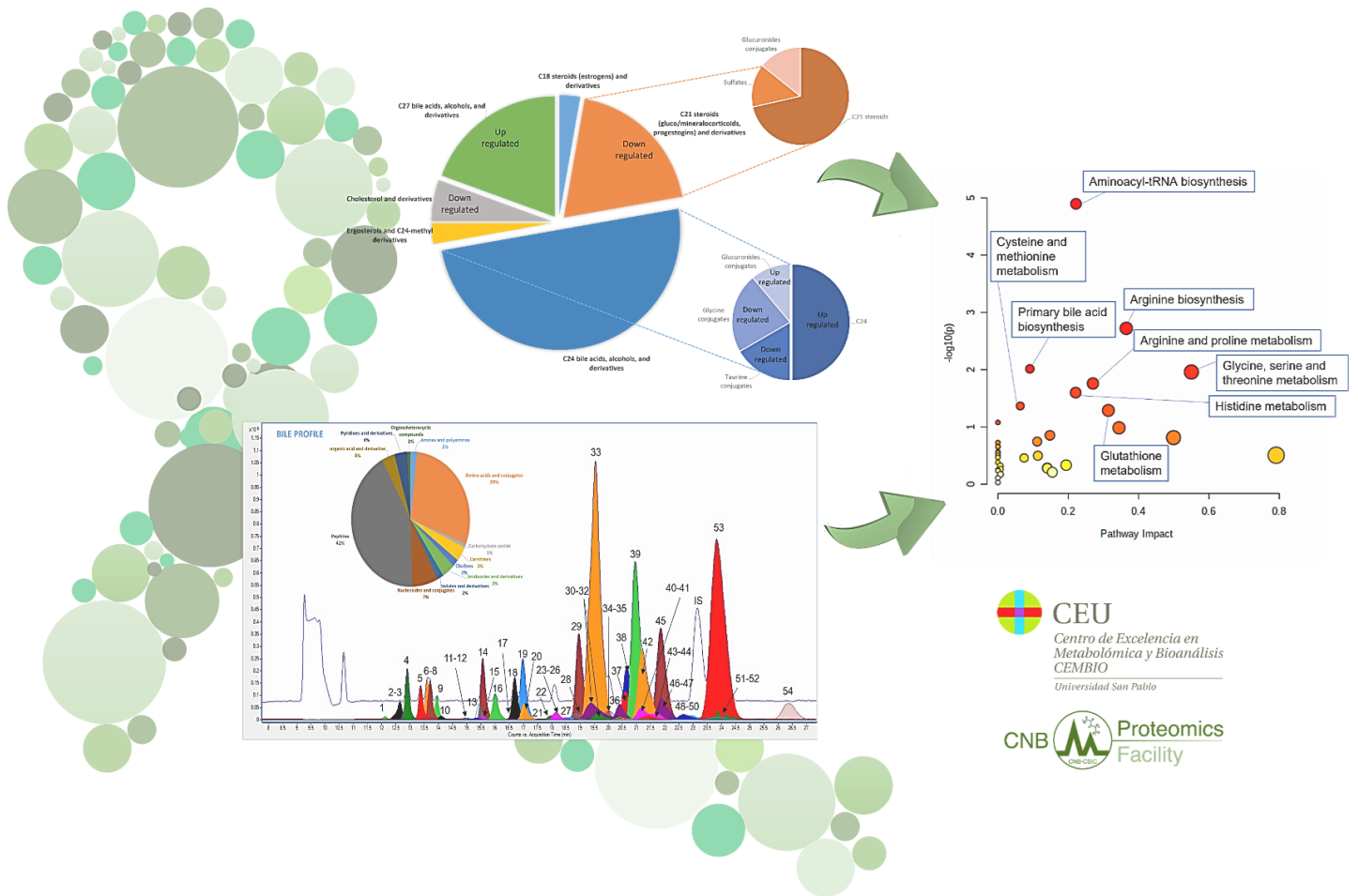
16. Sottero B, Rossin D, Staurengi E, Gamba P, Poli G, Testa G. Omics analysis of oxysterols to better understand their pathophysiological role. *Free Radical Biology and Medicine*. 2019;144:55-71.
17. Lee JY, Byeon SK, Moon MH. Profiling of oxidized phospholipids in lipoproteins from patients with coronary artery disease by hollow fiber flow field-flow fractionation and nanoflow liquid chromatography-tandem mass spectrometry. *Anal Chem*. 2015;87(2):1266-73.
18. Ademowo O, Dias H, Milic I, Devitt A, Moran R, Mulcahy R, et al. Phospholipid oxidation and carotenoid supplementation in Alzheimer's disease patients. *Free Radical Biology and Medicine*. 2017;108:77-85.
19. Pickens CA, Petritis K. High resolution mass spectrometry newborn screening applications for quantitative analysis of amino acids and acylcarnitines from dried blood spots. *Analytica Chimica Acta*. 2020;1120:85-96.
20. Gil de la Fuente A, Traldi F, Siroka J, Kretowski A, Ciborowski M, Otero A, et al. Characterization and annotation of oxidized glycerophosphocholines for non-targeted metabolomics with LC-QTOF-MS data. *Anal Chim Acta*. 2018;1037:358-68.
21. Salomon RG, Hong L, Hollyfield JG. Discovery of carboxyethylpyrroles (CEPs): critical insights into AMD, autism, cancer, and wound healing from basic research on the chemistry of oxidized phospholipids. *Chem Res Toxicol*. 2011;24(11):1803-16.
22. Kagan VE, Tyurina YY, Sun WY, Vlasova II, Dar H, Tyurin VA, et al. Redox phospholipidomics of enzymatically generated oxygenated phospholipids as specific signals of programmed cell death. *Free Radic Biol Med*. 2020;147:231-41.
23. Ghosh J, Myers CE. Inhibition of arachidonate 5-lipoxygenase triggers massive apoptosis in human prostate cancer cells. *Proceedings of the National Academy of Sciences*. 1998;95(22):13182.
24. Godzien J, Gil de la Fuente A, Otero A, Barbas C. Chapter Fifteen - Metabolite Annotation and Identification. In: Jaumot J, Bedia C, Tauler R, editors. *Comprehensive Analytical Chemistry*. 82: Elsevier; 2018. p. 415-45.
25. Godzien J, Kalaska B, Adamska-Patruno E, Siroka J, Ciborowski M, Kretowski A, et al. Oxidized glycerophosphatidylcholines in diabetes through non-targeted metabolomics: Their annotation and biological meaning. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2019;1120:62-70.
26. Reis A, Domingues P, Domingues M. Structural motifs in primary oxidation products of palmitoyl-arachidonoyl-phosphatidylcholines by LC-MS/MS. *Journal of Mass Spectrometry*. 2013;48(11):1207-16.
27. Reuter S, Gupta SC, Chaturvedi MM, Aggarwal BB. Oxidative stress, inflammation, and cancer: how are they linked? *Free Radic Biol Med*. 2010;49(11):1603-16.
28. Gupta RK, Patel AK, Shah N, Chaudhary AK, Jha UK, Yadav UC, et al. Oxidative stress and antioxidants in disease and cancer: a review. *Asian Pac J Cancer Prev*. 2014;15(11):4405-9.
29. Mantovani G, Madeddu C, Macciò A. Cachexia and oxidative stress in cancer: an innovative therapeutic management. *Curr Pharm Des*. 2012;18(31):4813-8.
30. Nie Q, Xing M, Chen H, Hu J, Nie S. Metabolomics and Lipidomics Profiling Reveals Hypocholesterolemic and Hypolipidemic Effects of Arabinoxylan on Type 2 Diabetic Rats. *J Agric Food Chem*. 2019;67(38):10614-23.

31. Chen Z, Zang L, Wu Y, Nakayama H, Shimada Y, Shrestha R, et al. Lipidomic Profiling on Oxidized Phospholipids in Type 2 Diabetes Mellitus Model Zebrafish. *Anal Sci.* 2018;34(10):1201-8.
32. Okuzumi A, Hatano T, Ueno SI, Ogawa T, Saiki S, Mori A, et al. Metabolomics-based identification of metabolic alterations in PARK2. *Ann Clin Transl Neurol.* 2019;6(3):525-36.
33. Tyurina YY, Polimova AM, Maciel E, Tyurin VA, Kapralova VI, Winnica DE, et al. LC/MS analysis of cardiolipins in substantia nigra and plasma of rotenone-treated rats: Implication for mitochondrial dysfunction in Parkinson's disease. *Free Radical Research.* 2015;49(5):681-91.
34. Paynter NP, Balasubramanian R, Giulianini F, Wang DD, Tinker LF, Gopal S, et al. Metabolic Predictors of Incident Coronary Heart Disease in Women. *Circulation.* 2018;137(8):841-53.
35. Tsimikas S, Brilakis ES, Miller ER, McConnell JP, Lennon RJ, Kornman KS, et al. Oxidized phospholipids, Lp(a) lipoprotein, and coronary artery disease. *N Engl J Med.* 2005;353(1):46-57.
36. Barriuso J, Custodio A, Afonso R, Alonso V, Astudillo A, Capdevila J, et al. Prognostic and predictive biomarkers for somatostatin analogs, peptide receptor radionuclide therapy and serotonin pathway targets in neuroendocrine tumours. *Cancer Treat Rev.*70(1532-1967 (Electronic)):209-22.
37. Barta JA, Powell CA, Wisnivesky JP. Global Epidemiology of Lung Cancer. *Annals of global health.* 2019;85(1):8.
38. Niklinski J, Kretowski A, Moniuszko M, Reszec J, Michalska-Falkowska A, Niemira M, et al. Systematic biobanking, novel imaging techniques, and advanced molecular analysis for precise tumor diagnosis and therapy: The Polish MOBIT project. *Advances in medical sciences.* 2017;62(2):405-13.
39. Daniluk U, Daniluk J, Kucharski R, Kowalczyk T, Pietrowska K, Samczuk P, et al. Untargeted Metabolomics and Inflammatory Markers Profiling in Children With Crohn's Disease and Ulcerative Colitis-A Preliminary Study. *Inflamm Bowel Dis.* 2019;25(7):1120-8.
40. Ciborowski M, Kisluk J, Pietrowska K, Samczuk P, Parfieniuk E, Kowalczyk T, et al. Development of LC-QTOF-MS method for human lung tissue fingerprinting. A preliminary application to nonsmall cell lung cancer. *ELECTROPHORESIS.* 2017;38(18):2304-12.
41. Kowalczyk T, Kisluk J, Pietrowska K, Godzien J, Kozłowski M, Reszeć J, et al. The Ability of Metabolomics to Discriminate Non-Small-Cell Lung Cancer Subtypes Depends on the Stage of the Disease and the Type of Material Studied. *Cancers (Basel).* 2021;13(13).
42. Dudzik D, Barbas-Bernardos C, García A, Barbas C. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. *Journal of pharmaceutical and biomedical analysis.* 2017;147:149-73.
43. Kuligowski J, Sanchez-Illana A, Sanjuan-Herraez D, Vento M, Quintas G. Intra-batch effect correction in liquid chromatography-mass spectrometry using quality control samples and support vector regression (QC-SVRC). *Analyst.* 2015;140(22):7810-7.

44. Barnes RJ, Dhanoa MS, Lister SJ. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*. 1989;43(5):772-7.
45. Montenegro-Burke JR, Guijas C, Siuzdak G. METLIN: A Tandem Mass Spectral Library of Standards. *Methods Mol Biol*. 2020;2104:149-63.
46. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform*. 2016;8:3.
47. Pauling JK, Hermansson M, Hartler J, Christiansen K, Gallego SF, Peng B, et al. Proposal for a common nomenclature for fragment ions in mass spectra of lipids. *PLoS One*. 2017;12(11):e0188394.
48. Fahy E, Subramaniam S, Murphy RC, Nishijima M, Raetz CR, Shimizu T, et al. Update of the LIPID MAPS comprehensive classification system for lipids. *J Lipid Res*. 2009;50 Suppl:S9-14.
49. Cai C, Bian X, Xue M, Liu X, Hu H, Wang J, et al. Eicosanoids metabolized through LOX distinguish asthma-COPD overlap from COPD by metabolomics study. *Int J Chron Obstruct Pulmon Dis*. 2019;14:1769-78.
50. Hu DY, Luo Y, Li CB, Zhou CY, Li XH, Peng A, et al. Oxylipin profiling of human plasma reflects the renal dysfunction in uremic patients. *Metabolomics*. 2018;14(8):104.
51. Banaszczak M, Maciejewska D, Drozd A, Ryterska K, Milc DJ, Raszeja-Wyszomirska J, et al. 5-Lipoxygenase Derivatives as Serum Biomarkers of a Successful Dietary Intervention in Patients with NonAlcoholic Fatty Liver Disease. *Medicina (Kaunas)*. 2020;56(2).
52. Wang N, Zhao X, Huai J, Li Y, Cheng C, Bi K, et al. Arachidonic acid metabonomics study for understanding therapeutic mechanism of Huo Luo Xiao Ling Dan on rat model of rheumatoid arthritis. *J Ethnopharmacol*. 2018;217:205-11.
53. Bian X, Sun B, Zheng P, Li N, Wu JL. Derivatization enhanced separation and sensitivity of long chain-free fatty acids: Application to asthma using targeted and non-targeted liquid chromatography-mass spectrometry approach. *Anal Chim Acta*. 2017;989:59-70.
54. Honn KV, Tang DG, Gao X, Butovich IA, Liu B, Timar J, et al. 12-lipoxygenases and 12(S)-HETE: role in cancer metastasis. *Cancer Metastasis Rev*. 1994;13(3-4):365-96.
55. Wang N, Zhao X, Wang W, Peng Y, Bi K, Dai R. Targeted profiling of arachidonic acid and eicosanoids in rat tissue by UFLC-MS/MS: Application to identify potential markers for rheumatoid arthritis. *Talanta*. 2017;162:479-87.
56. Cheng C, Gross ML. Applications and mechanisms of charge-remote fragmentation. *Mass Spectrometry Reviews*. 2000;19(6):398-420.
57. Hsu F-F, Turk J. Characterization of phosphatidylinositol, phosphatidylinositol-4-phosphate, and phosphatidylinositol-4,5-bisphosphate by electrospray ionization tandem mass spectrometry: a mechanistic study. *Journal of the American Society for Mass Spectrometry*. 2000;11(11):986-99.
58. Griffiths WJ. Tandem mass spectrometry in the study of fatty acids, bile acids, and steroids. *Mass Spectrom Rev*. 2003;22(2):81-152.

59. Demarque DP, Crotti AE, Vessecchi R, Lopes JL, Lopes NP. Fragmentation reactions using electrospray ionization mass spectrometry: an important tool for the structural elucidation and characterization of synthetic and natural products. *Nat Prod Rep.* 2016;33(3):432-55.
60. Tsugawa H, Nakabayashi R, Mori T, Yamada Y, Takahashi M, Rai A, et al. A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms. *Nature Methods.* 2019;16(4):295-8.
61. Ramakrishnan R, Tyurin VA, Tuyrin VA, Veglia F, Condamine T, Amoscato A, et al. Oxidized lipids block antigen cross-presentation by dendritic cells in cancer. *J Immunol.* 2014;192(6):2920-31.
62. Hammad LA, Wu G, Saleh MM, Klouckova I, Dobrolecki LE, Hickey RJ, et al. Elevated levels of hydroxylated phosphocholine lipids in the blood serum of breast cancer patients. *Rapid Commun Mass Spectrom.* 2009;23(6):863-76.
63. Goto T, Terada N, Inoue T, Kobayashi T, Nakayama K, Okada Y, et al. Decreased expression of lysophosphatidylcholine (16:0/OH) in high resolution imaging mass spectrometry independently predicts biochemical recurrence after surgical treatment for prostate cancer. *Prostate.* 2015;75(16):1821-30.
64. Bürger F, Krieg P, Kinzig A, Schurich B, Marks F, Fürstenberger G. Constitutive expression of 8-lipoxygenase in papillomas and clastogenic effects of lipoxygenase-derived arachidonic acid metabolites in keratinocytes. *Mol Carcinog.* 1999;24(2):108-17.
65. Ding XZ, Hennig R, Adrian TE. Lipoxygenase and cyclooxygenase metabolism: new insights in treatment and chemoprevention of pancreatic cancer. *Mol Cancer.* 2003;2:10.
66. Ikeda K, Mutoh M, Teraoka N, Nakanishi H, Wakabayashi K, Taguchi R. Increase of oxidant-related triglycerides and phosphatidylcholines in serum and small intestinal mucosa during development of intestinal polyp formation in Min mice. *Cancer Sci.* 2011;102(1):79-87.
67. Bochkov VN. Inflammatory profile of oxidized phospholipids. *Thromb Haemost.* 2007;97(3):348-54.
68. Bochkov V, Gesslbauer B, Mauerhofer C, Philippova M, Erne P, Oskolkova OV. Pleiotropic effects of oxidized phospholipids. *Free Radic Biol Med.* 2017;111:6-24.
69. Kelavkar UP, Nixon JB, Cohen C, Dillehay D, Eling TE, Badr KF. Overexpression of 15-lipoxygenase-1 in PC-3 human prostate cancer cells increases tumorigenesis. *Carcinogenesis.* 2001;22(11):1765-73.
70. Hassan S, Carraway RE. Involvement of arachidonic acid metabolism and EGF receptor in neurotensin-induced prostate cancer PC3 cell growth. *Regul Pept.* 2006;133(1-3):105-14.
71. Chen X, Li N, Wang S, Hong J, Fang M, Youselfson J, et al. Aberrant arachidonic acid metabolism in esophageal adenocarcinogenesis, and the effects of sulindac, nordihydroguaiaretic acid, and alpha-difluoromethylornithine on tumorigenesis in a rat surgical model. *Carcinogenesis.* 2002;23(12):2095-102.
72. Avis I, Hong SH, Martinez A, Moody T, Choi YH, Trepel J, et al. Five-lipoxygenase inhibitors can mediate apoptosis in human breast cancer cell lines through complex eicosanoid interactions. *Faseb j.* 2001;15(11):2007-9.

73. Chen FL, Wang XZ, Li JY, Yu JP, Huang CY, Chen ZX. 12-lipoxygenase induces apoptosis of human gastric cancer AGS cells via the ERK1/2 signal pathway. *Dig Dis Sci.* 2008;53(1):181-7.
74. Avis IM, Jett M, Boyle T, Vos MD, Moody T, Treston AM, et al. Growth control of lung cancer by interruption of 5-lipoxygenase-mediated growth factor signaling. *J Clin Invest.* 1996;97(3):806-13.
75. Karara A, Dishman E, Falck JR, Capdevila JH. Endogenous epoxyeicosatrienoyl-phospholipids. A novel class of cellular glycerolipids containing epoxidized arachidonate moieties. *J Biol Chem.* 1991;266(12):7561-9.
76. Powell WS, Rokach J. Biosynthesis, biological effects, and receptors of hydroxyeicosatetraenoic acids (HETEs) and oxoeicosatetraenoic acids (oxo-ETEs) derived from arachidonic acid. *Biochim Biophys Acta.* 2015;1851(4):340-55.
77. Koundouros N, Pouligiannis G. Reprogramming of fatty acid metabolism in cancer. *British Journal of Cancer.* 2020;122(1):4-22.



Capítulo 3. Enfoque analítico para el estudio del perfil metabolómico de bilis e interpretación de las alteraciones biológicas en Colangiocarcinoma

1. Introducción

El colangiocarcinoma (CCA) es un tipo de tumor devastador, altamente maligno, que surge del epitelio de los conductos biliares. Nos encontramos ante el segundo tumor hepático primario más común de todo el mundo y la forma más común de malignidad del árbol biliar (1). Aunque se trata de un cáncer raro, su incidencia ha aumentado considerablemente en las últimas décadas y su elevada tasa de mortalidad acumulada representa un problema de salud global (2).

Los pacientes con CCA aparecen con una presentación clínica tardía debido a la ausencia de síntomas específicos en las neoplasias malignas tempranas, por lo que es difícil de diagnosticar en estadios no avanzados de la enfermedad y, está asociado a un pronóstico precario (3). Hasta la fecha, la confirmación diagnóstica de CCA está basada en los niveles séricos de dos marcadores, el antígeno carcinoembrionario (CEA) y el antígeno carbohidrato 19-9 (CA 19-9), y técnicas de imágenes hepatobiliares, con alto índice de sospecha clínica previa. El principal problema es que estos biomarcadores no son sensibles ni específicos de CCA y, con el diagnóstico por imagen, no es posible diferenciar con precisión la enfermedad biliar benigna de la maligna (4, 5). Como consecuencia, la mayoría de los casos se encuentran en estados muy avanzados para la resección quirúrgica con éxito y, el tratamiento con radioterapia y quimioterapia conlleva a una tasa de supervivencia pobre (6, 7).

La falta de biomarcadores adecuados para el diagnóstico, pronóstico y tratamiento del CCA ha supuesto un desafío importante en los últimos años en la investigación en cáncer y, el estudio a nivel metabólico permitirá mejorar la comprensión de la etiopatogenia de la enfermedad.

Las células cancerosas muestran marcadas alteraciones metabólicas y la medición de metabolitos en distintas muestras biológicas representa una oportunidad alentadora para la búsqueda de biomarcadores de diagnóstico y/o pronóstico. Hoy en día, existen diversos estudios de búsqueda de biomarcadores en fluidos ampliamente utilizados en metabolómica como suero, plasma u orina para el diagnóstico de CCA (8-13). También se han llevado a cabo estudios con muestras de tejido (14, 15), pero el origen invasivo de la muestra ha planteado nuevos horizontes de estudio.

El muestreo de bilis con fines diagnóstico se ha convertido en una práctica clínica común y, su flujo continuo a través del tracto hepatobiliar la hace una muestra potencialmente ideal para la búsqueda de biomarcadores de CCA, siendo una potente fuente de información para enfermedades malignas del sistema hepatobiliar.

La bilis es un líquido complejo que el hígado produce constantemente y que es segregado hasta la vesícula biliar a través de los conductos intrahepáticos, donde se almacena y concentra hasta su liberación hacia el duodeno. Está compuesta en su mayor parte de agua (~ 95%) en la que se encuentran disueltos varios componentes, incluidos ácidos biliares, colesterol, glutatión, bilirrubina, fosfolípidos (principalmente glicerofosfatidilcolina, PC), aminoácidos, esteroides, enzimas, porfirinas, vitaminas y metales pesados, así como productos exógenos procedentes de fármacos, xenobióticos y toxinas ambientales (16, 17). Cuenta con una amplia gama de importantes funciones fisiológicas entre las que se encuentran, la digestión y absorción de las grasas, proteínas y vitaminas liposolubles por parte de los ácidos biliares, el mantenimiento de la homeostasis del colesterol y, la excreción de productos como la bilirrubina y proteínas envejecidas, además de ser vehículo para la eliminación de fármacos y toxinas endógenas y exógenas.

A pesar de la naturaleza más invasiva de la muestra si la comparamos con muestras de plasma u orina, se ha podido comprobar que la bilis está emergiendo como una rica fuente de biomarcadores específicamente para cánceres de hígado, vías biliares, vesícula biliar y páncreas (18). Por tanto, se podrían encontrar biomarcadores específicos en condiciones metabólicas alteradas. Sin embargo, se han invertido más esfuerzos en el perfilado de los ácidos biliares (19, 20), que en el estudio del perfil global de la bilis con el correspondiente potencial para descubrir nuevos marcadores aún sin explorar.

Se espera que la metabolómica, el estudio de todos los metabolitos del cuerpo, pueda proporcionar información más precisa y significativa sobre los sistemas biológicos implicados, si lo comparamos con la información proporcionada hasta ahora por la genómica, la proteómica y la transcriptómica (21). En estudios anteriores, mediante espectroscopía de resonancia magnética de protones a 800 MHz (H^1 NMR) se analizó un extracto de bilis (22). Cuarenta compuestos fueron asignados, entre los que se incluían fosfatidilcolinas, catorce aminoácidos, diez ácidos orgánicos, cuatro carbohidratos y polioles, colina, fosfocolina, betaína, trimetilamina-N-óxido y otras moléculas pequeñas. Existen también estudios utilizando la espectrometría de masas, aunque son menos numerosos. Empleando un espectrómetro de masas híbrido ortogonal de cuadrupolo con tiempo de vuelo (QTOF) equipado con una sonda de análisis de sólidos a presión atmosférica (ASAP-MS), Plumb y colaboradores (23) elaboraron el caracterizado de los perfiles de bilis de ratas y perros. Sin embargo, los autores concluyeron que los perfiles son menos completos que los resultados

procedentes de LC o GC, pero lo suficiente para producir una clasificación de la muestra en un entorno clínico. Por otro lado, el análisis de bilis con LC y GC conlleva resolver varios desafíos importantes; entre los más significativos destaca que, debido a la complejidad del perfil metabólico de este tipo de muestras conteniendo muchos ácidos biliares isobáricos e isoméricos, su separación e identificación se ve dificultada. Además, en los métodos basados en GC-MS, se requiere un prolongado pretratamiento de la muestra, lo que aumenta el riesgo de interferencias indeseables por la manipulación y el tiempo empleado para el análisis (24).

Previamente, en nuestro grupo, se aplicó con éxito un método simple de CE con diodo array para la obtención del perfil de huellas dactilares de 26 muestras de bilis tomadas en animales de experimentación, antes y después del trasplante de hígado de donantes y receptores, respectivamente, con el fin de desentrañar las complejas vías metabólicas involucradas en el trasplante (25). Los electroferogramas se compararon directamente para agrupar las muestras, sin embargo, solo se encontró alteración en dos ácidos biliares, ácido taurocólico (TCAc) y taurodesoxicólico (TDCA), que se encontraban aumentados al comparar los resultados obtenidos en muestras de donantes de injertos normales con dos de las 15 muestras que desarrollaron disfunción primaria (PGD).

Aunque la CE se ha empleado principalmente para la toma de huellas dactilares de proteínas, en la última década han aumentado los estudios metabólicos que reportan esta técnica analítica para la toma de huellas dactilares de biofluidos (26, 27). Con esta técnica, los compuestos se separan en relación con su masa/carga, por lo que es adecuada para el análisis de compuestos polares/ionizables. Entre sus principales ventajas destaca el pequeño volumen (nanolitros) tanto de muestra como de reactivos, mínimo pretratamiento de muestra, corto tiempo de análisis y alta eficiencia (28-31).

De manera que, con base en los estudios previos sobre bilis en enfermedades hepatobiliares, lo que este trabajo quiere presentar es una descripción de la caracterización del perfil metabólico de bilis de rata por CE-MS junto con un enfoque metabólico no dirigido en un grupo con CCA frente a un grupo control. Para completar la información obtenida, en el estudio se llevó a cabo el análisis no dirigido utilizando, además, HPLC-QTOF/MS.

2. Materiales y métodos

2.1. Cohorte de estudio

La población de estudio se basó en 38 muestras de bilis procedentes de ratas macho Sprague-Dawley de ocho semanas (~ 250 g) (Envigo, Barcelona, España). Para la extracción de la muestra, los animales fueron anestesiados con pentobarbital sódico (50 mg/kg, ip), se realizó laparotomía abdominal y, se canuló el conducto biliar para la recolección de bilis como describe Rodríguez-Ortigosa y *col.* (32). Los animales se sacrificaron mediante dislocación cervical. El protocolo experimental fue aprobado y realizado según las directrices del Comité de Atención Animal de la Universidad de Navarra.

En total, fueron analizadas 10 muestras de bilis de control para la caracterización del perfil, mientras que el estudio comparativo se llevó a cabo con muestras de bilis de rata control de 16 semanas, C_16, y 30 semanas, C_30, (n = 12, en cada grupo) y las muestras de un modelo de patología de colangiocarcinoma de 30 semanas, T_30, (n = 14) y se analizaron utilizando el protocolo de tratamiento de muestras optimizado para CE-MS como prueba de concepto de la utilidad del perfil para la caracterización de enfermedades. Todas las muestras se almacenaron a -80 °C, y el día del análisis se descongelaron.

Debido a la limitación de muestra y, con la vista puesta en los estudios previos realizados por el grupo colaborador en análisis de proteómica, se decidió complementar el estudio metabólico utilizando la cromatografía de líquidos, además de la electroforesis capilar, acopladas ambas plataformas analíticas a la espectrometría de masas.

2.2. Análisis no dirigido por CE-TOF/MS

Como hasta la fecha, no existía metodología establecida para el tratamiento de este tipo de muestra en el centro, se comenzó con la optimización para la precipitación de las proteínas y el grado de dilución de la muestra.

Para ello se llevaron a cabo diferentes ensayos con muestra de bilis almacenada para este tipo de pruebas. La bilis se descongeló en hielo y fue mezclada con diferentes volúmenes de reactivo para la obtención de la mejor respuesta posible. Con relación a lo recogido en la bibliografía se siguieron dos tipos de tratamiento posibles. El método modificado del utilizado para plasma en el CEMBIO fue uno de los que se probó, en el que, a un volumen de bilis se le añadieron diferentes volúmenes (1, 5 y 100 volúmenes)

de agua MiliQ® que contenía un 5% de ACN, ácido fórmico 0.1 M y metionina sulfona 0.2 mM como patrón interno (IS). La mezcla se agitó en un vortex durante 1 min y, posteriormente, se transfirió a un filtro de ultracentrifugación Centrifree (30 kDa) (Millipore Irlanda, Eire) para su desprotenización (2000 x g, 80 min, 4 °C). El ultrafiltrado se transfirió a un vial directamente para su análisis. Por otra parte, el segundo método de tratamiento de muestra consistió en la precipitación proteica con fase orgánica. La mezcla de un volumen de muestra con tres volúmenes de ACN frío (-20 °C) se centrifugó durante 10 min a 4 °C a la velocidad de 2000 x g. El sobrenadante se evaporó hasta sequedad y fue reconstituido, en el mismo volumen que el inicial de bilis, con ácido fórmico 0.1 M que contenía metionina sulfona 0.2 mM como IS. La figura 1 resume el proceso realizado.

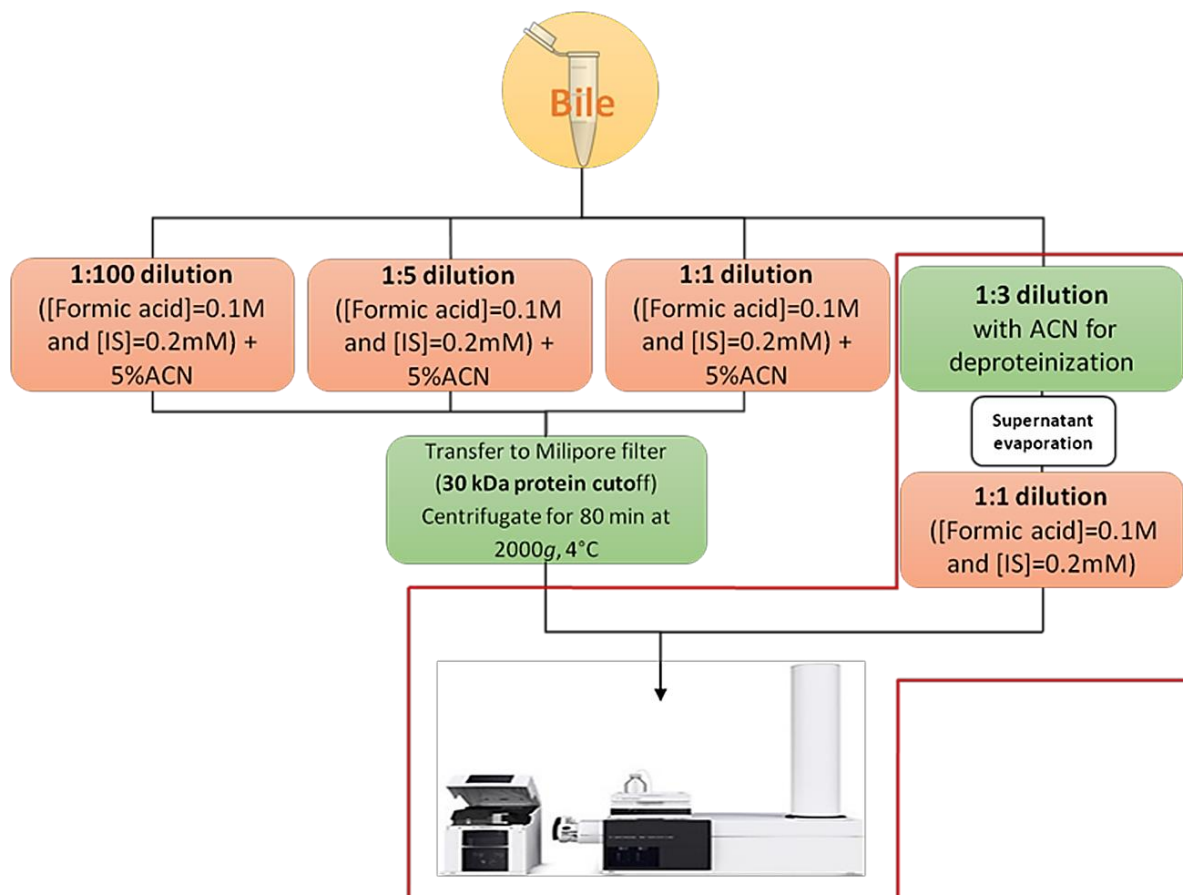


Figura 1. Etapas del procedimiento aplicado en la optimización del tratamiento de muestras de bilis para CE-MS.

Para verificar el rendimiento y la reproducibilidad del equipo se utilizaron controles de calidad. Para ello, se analizaron varias réplicas de este tipo de muestra que contenían una pequeña cantidad igual de todas las muestras de la cohorte de estudio ("Controles de calidad", QC). Estos controles de calidad se trataron como el resto de las

muestras y, se inyectaron al comienzo de la secuencia de análisis (10 inyecciones) para acondicionar o equilibrar el sistema y, posteriormente, cada siete muestras para probar la estabilidad durante el análisis.

Los análisis de CE-MS se realizaron utilizando una electroforesis capilar 7100 (Agilent Technologies, Wilmington, EE. UU.) acoplada a un detector de masa exacta con un sistema analizador de tiempo de vuelo TOF/MS 6224 (Agilent Technologies, Wilmington, EE. UU.). El acoplamiento está equipado con una fuente de ionización por *electrospray* (ESI).

De la misma forma que se expone en el capítulo 1, se preparó un nuevo capilar de sílice fundida de Agilent Technologies (50 μm id x 100 cm de longitud total) con el acondicionamiento en tres etapas de 30 minutos cada una: NaOH 1.0 M, seguido de agua MiliQ® y electrolito de fondo (BGE) (ácido fórmico 1.0 M en metanol al 10%). El análisis se realizó utilizando un método modificado desarrollado en el CEMBIO (33). Antes de cada análisis, el BGE se reemplaza automáticamente del vial. A continuación, se enjuagó el capilar durante 5 min (950 mbar) con BGE y se aplicó un potencial de 30 kV durante 10 s para desplazar los iones tampón. Las inyecciones de muestra se realizaron durante 50 s con una presión de 50 mbar. Después de cada inyección de muestra, se inyectó BGE durante 20 s a 100 mbar de presión. Las condiciones de separación incluyeron 25 mbar de presión y 30 kV de voltaje. Los datos se adquirieron en modo ESI positivo con una velocidad de barrido de 1 espectro/s y un rango de masa de m/z de 74 a 1000. El líquido auxiliar (SL) consistía en 50% de metanol, 50% de agua y los estándares de referencia (20 μL de purina 5 mmol/L, (m/z 121.0509) + 20 μL de HP-0921 2.5 mmol/L (m/z 922.0098)) usando un caudal de 0.6 mL/min (división 1:100). La temperatura del gas de secado se mantuvo a 200 °C con un flujo de 10 L/min, la presión del nebulizador se ajustó a 10 psig, el voltaje a 3500 V, el voltaje del fragmentor a 125 V y el *skimmer* a 65 V.

2.3. Análisis no dirigido por LC-QTOF/MS

A la vista de lo recogido en la bibliografía (17, 34, 35), para la precipitación proteica y la extracción de los metabolitos, a un volumen de muestra de bilis se le añadió 3 volúmenes de una mezcla fría (-20 °C) de etanol:metanol (1:1, v/v). La mezcla se mantuvo en un vortex durante un 1 min, se dejó reposar durante 5 min en hielo y se centrifugó a 16000 x g durante 20 min a 4 °C. El sobrenadante se recogió en un vial para su análisis.

La separación cromatográfica y la detección por espectrometría de masas se llevó a cabo con el sistema UHPLC serie 1290 Infinity II (Agilent Technologies, Waldbronn, Germany) acoplado a un analizador QTOF/MS 6545, adquiriendo en modo de ion positivo y negativo, respectivamente. El sistema de UHPLC consta de un desgasificador, dos bombas binarias, un inyector automático termostatzado y un horno de columna. Durante todos los análisis, se infusieron continuamente en el sistema dos masas de referencia para permitir una corrección de masa constante: m/z 121.0509 ($C_5H_4N_4$) y m/z 922.0098 ($C_{18}H_{18}O_6N_3P_3F_{24}$) para el modo de ionización positiva y m/z 112.9856 ($C_2O_2F_3(NH_4)$) y m/z 1033.9881 ($C_{18}H_{18}O_6N_3P_3F_{24}$) para modo negativo. Las condiciones de análisis utilizadas fueron similares a las empleadas en el capítulo 1, optimizadas en nuestro laboratorio siguiendo el desarrollo de un método para muestras de plasma (36) y, los análisis preliminares realizados sobre muestras de bilis de prueba otorgaron resultados satisfactorios. Brevemente, 0.5 μ L y 1.0 μ L de muestra, para positivo y negativo respectivamente, fueron inyectados en una columna Zorbax Extended-C18 Rapid Resolution (Agilent Technologies, 2.1 mm \times 50 mm, 1.8 μ m) controlando la temperatura a 60 °C. La fase móvil compuesta por (A) 0.1% de ácido fórmico en agua y (B) 0.1% de ácido fórmico en acetonitrilo se utilizó con un flujo de 0.6 mL/min. Las condiciones iniciales del gradiente cromatográfico fue un 5% de B durante el primer minuto, con posterior incremento hasta el 80% de B en el minuto 6. Por último, se alcanzó el 100% de B en el minuto 11.5 volviendo a las condiciones iniciales en 0.5 min, seguido de 3 min de reequilibrado.

Los datos se adquirieron en modo de ionización ESI positivo y negativo, en análisis independientes, con el escaneo completo de masas en el rango de m/z de 100 a 1000, y de m/z de 40 a 1000 para el análisis MS y MS/MS, respectivamente. Las condiciones de la fuente de espectrometría de masas consistieron en un voltaje del capilar de 3000 V con una velocidad de barrido de exploración de 2.0 espectros/s (modo positivo) o 4000 V y 1.0 espectro/s (modo negativo), un caudal de gas nebulizador de 12.0 L/min, una temperatura de la fuente de 250 °C y una presión de la fuente de 52 psig. Para mejorar la ionización se aplicó un calentamiento adicional usando gas auxiliar calentado hasta los 370 °C con un flujo de 11 L/min.

Para el análisis de MS/MS se utilizó la herramienta de adquisición iterativa dependiente de datos (DDA), que consiste en un modo de adquisición automático con la exclusión de iones precursores en cada réplica. Por ello, es necesario llevar a cabo el re-análisis de algunas muestras por grupo manteniendo las mismas condiciones cromatográficas y de la espectrometría de masas empleadas en el análisis inicial. Este

análisis se puede adquirir en solo unas pocas muestras representativas, lo que ahorra tiempo de adquisición y procesamiento. Con una ventana de error estrecha (aprox. 1.3 Da), las energías de colisión fueron determinadas mediante la rampa para todas las cargas, basada en m/z y la carga del precursor: 3.6 fue la pendiente y el factor de corrección de 4.8 según la ecuación: Energía de colisión = [(pendiente x m/z del ion precursor) / 100.0] + factor de corrección. Cada muestra fue analizada 5 veces utilizando esta herramienta.

2.4. Tratamiento de datos

Los datos adquiridos para ambas plataformas fueron procesados, de una forma similar al explicado en el capítulo 1, utilizando el software MassHunter Profinder (B.08.00, Agilent Technologies) para obtener los datos estructurados y en un formato apropiado para su posterior filtrado de calidad y análisis estadístico. El primer paso es la extracción de características moleculares (MFE) eliminando el ruido de fondo y los iones no relacionados. Con ello se agrupan los aductos co-eluyentes (+H⁺, +Na⁺ y la pérdida neutra de agua para CE-MS y LC-MS(+)) y -H⁺, -Cl⁻, +COO⁻ y la pérdida neutra de agua para LC-MS(-)) de cada entidad química. Esto, junto con el alineado de las muestras usando la masa y el tiempo de retención crea el espectro final para esa entidad química. En el siguiente paso se procedió con la reextracción de las características moleculares por conjunto completo (RFE) a través de la búsqueda por ion (FBI) con una lista final de compuestos de mejor calidad. Los compuestos encontrados en las diferentes muestras se alinearon utilizando una ventana de tiempo de retención de 0.0% ± 0.5 min y una ventana de masa de 10 ppm ± 2.0 mDa para LC-MS (positivo y negativo) y 13% ± 3 min y una ventana de masa de 20 ppm ± 2.0 mDa para CE-MS. La altura mínima se estableció en 1000 cuentas para LC-MS y 600 cuentas para CE-MS y, el filtrado de las características presentes en al menos el 70% de las muestras de algún grupo. Con esto se generó una lista final de grupos de compuestos que, inspeccionada minuciosamente en cada muestra, se comprobó la correcta integración de los iones extraídos.

Una vez obtenida la matriz de datos se realizó el filtrado de calidad de los mismos de acuerdo con la variación en la abundancia de los QCs, manteniendo exclusivamente aquellos con RSD < 30%.

2.5. Análisis estadístico

Previo al análisis estadístico se llevó a cabo la sustitución de los valores que faltaban, valores que se pierden durante el procesado de datos debido a la complejidad del trabajo realizado por los algoritmos MFE y Fbl. Para reemplazarlos se utilizó una función matemática conocida como kNN (*k-nearest neighbors*) que permite sustituir el valor perdido por la estimación de su valor comparando la similitud del perfil con otras muestras o vecinos cercanos pertenecientes al mismo grupo (37).

Una vez reemplazados estos valores, los datos de ambas técnicas fueron analizados utilizando estadística multivariante y univariante. Primero se realizó un análisis multivariante no supervisado de los datos (PCA) que nos permite una visión simplificada inicial de posibles tendencias en los datos. Para LC-MS en polaridad negativa se observó una tendencia clara basada en el orden de inyección que fue corregida con el algoritmo QC-SVRC (*Quality Control and Support Vector Regression Correction*) explicado anteriormente (capítulo 1) (38).

Los datos de CE-MS y LC-MS para ambas polaridades fueron, entonces, sometidos al siguiente filtro de garantía de calidad (QA) donde se estudió su distribución normal por las pruebas de Kolmogórov-Smirnov-Lilliefors y Shapiro-Wilk y la prueba de Levene para la relación de varianza. Posteriormente se continuó con el análisis univariante (UVA) utilizando Matlab R2015a (The MathWorks, Inc) para evaluar las diferencias entre los individuos con CCA *versus* el grupo control, aplicando la prueba t de Student (valor de $p \leq 0.05$). Posteriormente estos valores fueron corregidos con el método de Benjamini-Hochberg, controlando de esta forma la tasa de falsos positivos en el nivel $\alpha = 0.05$.

El análisis multivariante (MVA) se llevó a cabo utilizando el software SIMCA 15.0 (Sartorius Stedim Biotech) que nos permite estudiar el conjunto de datos simplificando su dimensión. Con algoritmos de enfoque no supervisado y supervisados seremos capaces de visualizar la matriz de datos completa y comprobar el rendimiento y la calidad de los mismos. Con el Análisis de Componentes Principales (PCA) se comprobó el agrupamiento espontáneo de las muestras y, nos permitió identificar aquellas que más se alejaban del resto, "outliers". Métodos supervisados como el análisis discriminante por mínimos cuadrados parciales (PLS-DA) y el ortogonal por mínimos cuadrados parciales (OPLS-DA) fueron construidos también. A partir de este último se estudiaron los intervalos de confianza del diagrama S (*S-plot*) y los diagramas Volcano (*Volcano plots*) que a partir de los valores de importancia de la variable en la proyección (VIP) y el coeficiente de correlación, $\rho(\text{corr})$ nos permite la selección de variables

diferenciadoras (39). Estos gráficos son, generalmente, muy empleados en metabolómica para proporcionar una buena descripción general de los datos y del modelo.

La capacidad de los modelos para la representación de datos reales o bondad de ajuste está representada por R^2 . Por otra parte, la bondad de predicción del modelo viene representada por Q^2 y, ambos valores, son prácticos a la hora de decidir el número de variables latentes a utilizar. Es por ello por lo que, para estudiar la calidad de los modelos obtenidos se han utilizado R^2 y Q^2 . Los modelos OPLS-DA generados para cada técnica fueron validados utilizando la prueba de validación cruzada (*cross validation*), dejando fuera 1/3 de las muestras y, construyendo los modelos con las muestras restantes hasta que todas las muestras excluidas fueran predichas por el nuevo modelo al menos una vez. El porcentaje de muestras correctamente clasificadas para cada predicción fue calculado.

Por último, se estudió la distribución de cada variable entre los grupos calculando el porcentaje de cambio (Promedio del grupo de casos – Promedio del grupo control) / Promedio del grupo control, en %.

2.6. Anotación e identificación

Los resultados con diferencias estadísticamente significativas para cada plataforma fueron anotados de forma tentativa según lo detallado anteriormente en el capítulo 1. Utilizando la herramienta CEU Mass Mediator (<http://ceumass.eps.uspceu.es/>) con los m/z de los compuestos. Para ello, se tuvieron en cuenta, la coincidencia de masa exacta con un error menor de 10 ppm para LC-MS y de 20 ppm para CE-MS, la distribución isotópica y la verificación manual de los posibles iones y aductos. La anotación para CE-MS se complementó con el tiempo de migración relativo, eliminando fragmentos, dímeros o artefactos (40).

Para confirmar estas anotaciones tentativas de LC-MS se utilizaron los datos procedentes del análisis MS/MS de exclusión iterativa. Con estos datos se procedió de dos formas. Por una parte, fueron importados a la herramienta Agilent MassHunter Lipid Annotator Version 1.0 empleada para analizar rápidamente grandes conjuntos de datos de lipidómica y, mejorar la precisión de la anotación. Se utilizaron los parámetros del método predeterminado, excepto que solo se consideraron los precursores $[M + H]^+$ y $[M + Na]^+$ para el análisis en modo de iones positivos, y los precursores $[M - H]^-$ y $[M + COOH-H]^-$ para resultados de análisis en modo negativo. Esta herramienta utiliza una combinación de puntuación bayesiana, un algoritmo de densidad de probabilidad y el

ajuste por mínimos cuadrados no negativos para buscar en la biblioteca de lípidos generados según información teórica (LipidBlast) y anotar los espectros MS/MS (41). Utilizando algoritmos propios para realizar la deconvolución de las mezclas de espectros MS/MS y, empleando el teorema bayesiano, aproxima la posible identificación de los lípidos candidatos analizando las distribuciones de probabilidad derivadas de los errores aleatorios en el rendimiento de la medición (m/z , fidelidad isotópica, artefactos y distribuciones de las señales de fragmentación). Además, los métodos bayesianos pueden proporcionar estimaciones más precisas mediante la capacidad de incorporar conocimientos previos y mediciones experimentales directas en distribuciones de error (42). Las anotaciones reportadas por Lipid Annotator fueron comprobadas analizando manualmente los espectros de fragmentación. Por otra parte, la información contenida en los análisis MS/MS iterativos fue utilizada, también, para intentar anotar otros posibles compuestos significativos, no reportados por el software, con la búsqueda e inspección manual de las señales y los espectros de fragmentación. El estudio de los patrones de fragmentación se complementó con la información disponible de espectros MS/MS de estándares analizados por el grupo, o a través de los espectros disponibles en la base de datos de Metlin (<https://metlin.scripps.edu/>) (43) y la fragmentación “in silico” para la identificación asistida en MetFrag (<https://msbi.ipb-halle.de/MetFrag/>) (44).

2.7. Interpretación biológica

Finalmente, se utilizó la plataforma online MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca/>) para el estudio funcional de los resultados obtenidos (45). Esta herramienta sirve de guía para plantear una posible hipótesis, pero debe complementarse con el estudio de la literatura publicada para poder generar una explicación de los cambios observados y relacionarlos con las rutas alteradas para la patología.

Para ello, se llevó a cabo el análisis de los resultados con el mapa de color en 2D que reporta un resumen intuitivo de los datos recogidos en la matriz numérica, “*heatmap*” que nos facilitó la visualización de tendencias dentro de cada grupo experimental. Asimismo, se realizó el análisis para identificar las vías más relevantes según el impacto de los metabolitos previamente seleccionados dentro de la ruta metabólica. Esto nos ofrece una visión completa del metaboloma remarcando aquellas vías metabólicas con mayor impacto en la ruta de los metabolitos involucrados alterados.

La figura 2 es una representación esquemática de los pasos llevados a cabo en el análisis de muestras de bilis mediante CE-MS.

2.8. Análisis semicuantitativo a través de CE-MS

- *Tratamiento de muestra*

Para las muestras de bilis se utilizó el tratamiento de muestra optimizado que consistió, brevemente, en mezclar un volumen de bilis con tres volúmenes de acetonitrilo (ACN) frío (-20 ° C). La mezcla se centrifugó durante 10 minutos a 4 °C con 2000 x g y el sobrenadante se evaporó hasta sequedad y se reconstituyó en el mismo volumen de bilis con ácido fórmico 0.1 M que contenía metionina sulfona 0.2 mM como IS.

- *Análisis de bilis por CE-TOF/MS*

Los análisis de CE-MS se realizaron utilizando el mismo equipo descrito previamente, mediante el método modificado desarrollado en el CEMBIO (33).

- *Semi-cuantificación de metabolitos*

Se seleccionaron los metabolitos glicina, serina, ornitina, lisina, arginina, citrulina, cistina, ácido oftálmico y glutatión oxidado como posibles candidatos de la respuesta redox a nivel celular (tabla 1). A partir de disoluciones madre preparadas de 1000 mg/L se construyeron las diferentes diluciones de las concentraciones requeridas.

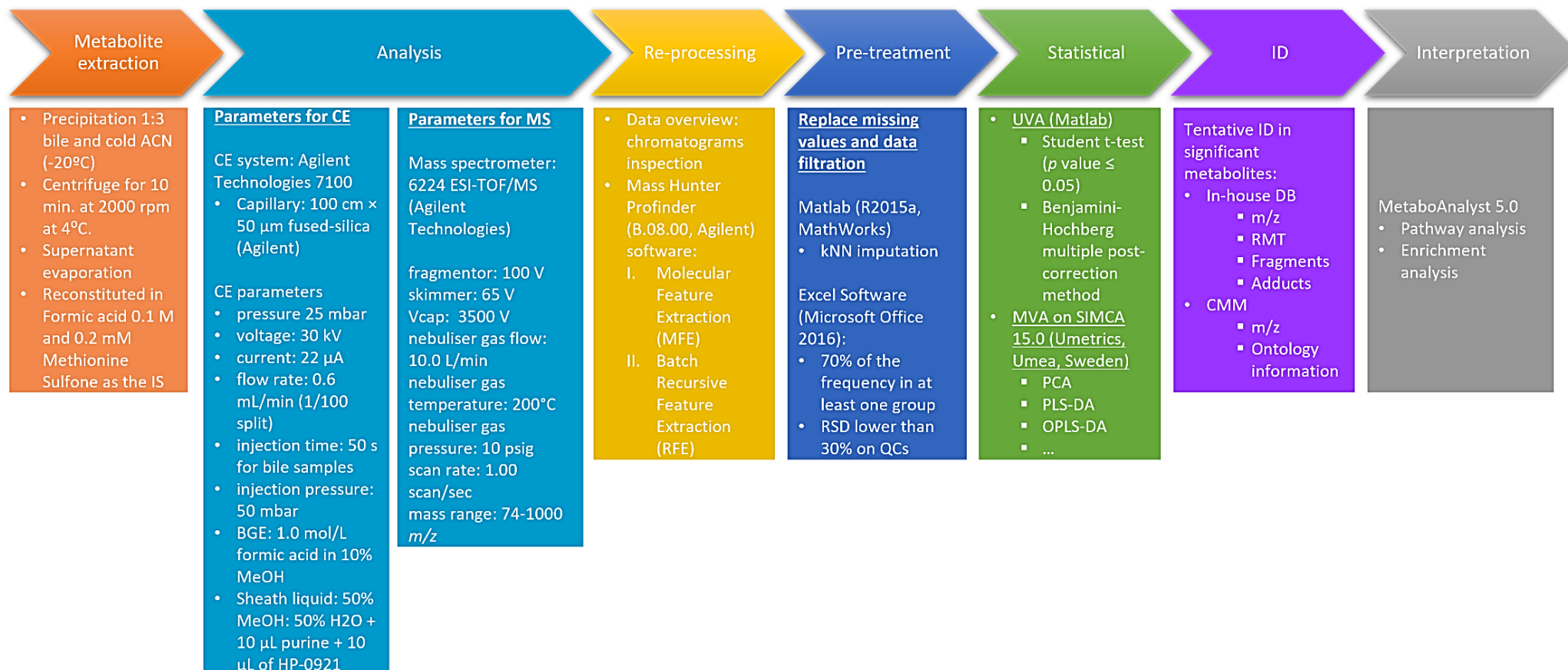


Figura 2. Flujo de trabajo completo para el análisis de muestras de bilis de rata a través de CE-MS. Representación esquemática de los pasos fundamentales y los softwares utilizados que se detallan en el apartado “Materiales y métodos”.

Tabla 1. Lista de los estándares usados en la semicuantificación a través de CE-MS.

Comuesto	Marca	Casa comercial	Riqueza
Glycine	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>99%
L-Serine	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>99%
L-Ornithine hydrochloride	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>99%
L-Lysine	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>98%
L-Arginine	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>98%
DL-Citrulline	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>98%
L-Cystine	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>99%
Ophthalmic acid	BACHEM	Bachem AG, Hauptstrasse 144, CH-4416 Bubendorf, Switzerland	>98%
Glutathione disulfide	SIGMA ALDRICH	Merck Life Science S.L. Calle Maria de Molina 40 E-28006 MADRID	>98%

Para obtener los valores de concentración de estos compuestos en bilis, se realizó la semicuantificación de cada uno de ellos con relación al pico del patrón interno ($A_{\text{metabolito}} / A_{\text{IS}}$) tanto en muestras como en patrones. La semicuantificación se realizó interpolando el valor integrado $A_{\text{metabolite}} / A_{\text{IS}}$ en las curvas de calibración construidas a partir del análisis de tres réplicas de cada estándar comercial relativizado con el estándar interno.

3. Resultados y Discusión

El estudio del perfil metabólico de muestras de bilis era un análisis novedoso para el centro, por lo que uno de los objetivos primero, era la optimización del tratamiento de muestra y las condiciones de análisis de este tipo de muestras. Una vez se llegó un protocolo sencillo y de fácil aplicación, que generaba buenos resultados, se aplicó para el caracterizado del perfil metabólico de bilis de rata vía CE-MS y, el estudio de

huella dactilar de este tipo de muestra para la búsqueda de posibles biomarcadores en la patología de colangiocarcinoma a través de CE-MS y LC-MS. El estudio del perfil metabolómico de la bilis a través de estas técnicas proporciona información sobre la patogénesis de las enfermedades de hígado, conductos biliares, vesícula biliar y páncreas, y permite la identificación de biomarcadores bioquímicos de la patología (18).

3.1. Optimización del tratamiento de muestra para CE-MS

Hasta la fecha, existían muy pocos estudios de metabolómica en muestras de bilis (20, 23, 34, 46, 47), la mayoría de ellos empleando NMR (46, 48-50), y mucho menos utilizando CE-MS. Papaspyridonos y *col.* (25) siguieron un protocolo de tratamiento de muestras reportado previamente por Mims y Hercules para muestras de plasma (51) y orina (52) mediante extracción en fase sólida (SPE) al vacío. Los ácidos biliares eluían con metanol y posteriormente se evaporaba para finalmente diluir las muestras en una mezcla de MeOH/tampón.

A la vista de lo reportado, iniciamos la optimización siguiendo una modificación de nuestro protocolo para el tratamiento CE-MS de plasma (53). Este consistía en añadir a un volumen de bilis, diferentes volúmenes de agua MiliQ que contenían un 5% de ACN, ácido fórmico 0.1M y el IS en 0.2mM, en concentración final. Se estudiaron diferentes ratios frente al volumen de bilis (1:1, 1:5 y 1:100) y la mezcla fue ultrafiltrada utilizando Centrifree (30 kDa) para su desproteinización. Con este método, los resultados no fueron del todo satisfactorios. Durante el análisis, se observaron caídas repetidas de la corriente del capilar y rotura de este. Esto fue, posiblemente, por obstrucción del capilar, por lo que se probó otro medio para precipitar las proteínas y eliminarlas de la muestra. Además, la mezcla 1:100 presentó una señal muy baja, apenas diferenciable del ruido de fondo, por lo que esta proporción fue descartada.

Para intentar obtener una muestra más “limpia” para el análisis, el segundo método de tratamiento de muestra fue la precipitación con fase orgánica, mezclando un volumen de bilis con tres de ACN frío, centrifugando y recogiendo sobrenadante. Este sobrenadante se secó y la muestra se reconstituyó en el mismo volumen que el inicial de bilis con ácido fórmico y IS, 0.1 M y 0.2mM de concentración final, respectivamente.

La Figura 3 resume los resultados obtenidos de estos procedimientos. Para la proporción 1:1, el área total de los electroferogramas de iones extraídos (EIC) para los metabolitos seleccionados fue ligeramente mayor sin ultrafiltración $((276.4 \pm 1.8) \times 10^6$, media \pm desviación estándar, STD) que con ultrafiltración $((256.76 \pm 0.25) \times 10^6$, media

± STD). De otra forma, se analizaron las diferencias relacionadas con el posible bloqueo del capilar para asegurar que no hubiera problemas de movilidad o cambios de los tiempos de migración, como ya había sido reportado (53). Utilizando el método de precipitación proteica con tres volúmenes de ACN, se analizaron 10 réplicas de una muestra y, los resultados obtenidos fueron analizados en términos de sensibilidad y reproducibilidad. El área del patrón interno se mantuvo en $148.5 \pm 3.7 \times 10^5$, media ± STD, con un RSD de 2.52% y un tiempo de migración de 22.87 ± 0.17 min, media ± STD, con un RSD de 0.74%. Como se pudo comprobar, ni la sensibilidad ni la reproducibilidad se vieron comprometidas, además de la estabilidad durante el análisis, sin pérdidas de corriente y obstrucción del capilar. Si bien la diferencia en cuanto a los resultados de sensibilidad en ambos métodos es mínima, con el método de precipitación proteica sin ultrafiltración se estableció un protocolo para el tratamiento de bilis más sencillo y corto, que presentó menor variabilidad y mejor sensibilidad.

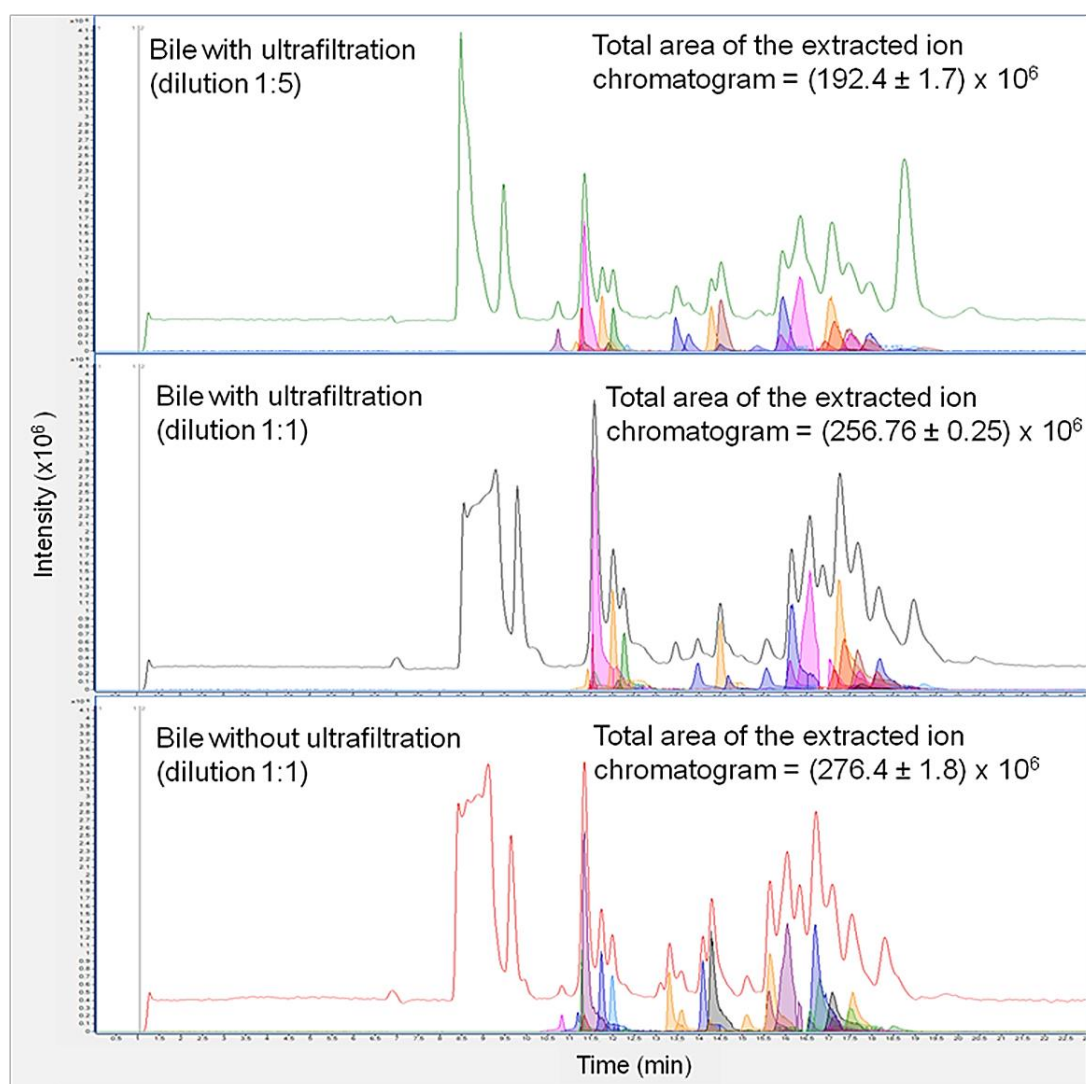


Figura 3. Electroferograma de iones totales (TIC) y Electroferograma de iones extraídos (EIC) comparativos de muestra de bilis bajo tres protocolos de tratamiento para su optimización. El área total del EIC para el mismo panel de metabolitos tentativos seleccionados se calculó con media \pm STD ($n = 10$).

3.2. Caracterizado del perfil metabolómico de bilis a través de CE-MS

El protocolo de tratamiento de muestra optimizado se aplicó en muestras procedentes de ratas y, ayudó a caracterizar el perfil de bilis a través de CE-MS (Figura 4). Esta aplicación permitió caracterizar el perfil metabolómico de bilis de ratas de individuos sanos donde se pudieron observar más de doscientos picos. Tras el filtrado de 219 características, 213 presentaban un RSD por debajo del 30% en los controles de calidad (QC) y se anotaron 101 señales a través de la plataforma CEU Mass Mediator (54). Ciento un compuestos se encontraban repartidos a lo largo del electroferograma representados en la Figura 4, donde los picos más representativos, anotados de forma tentativa, se designaron con sus números correspondientes en la Tabla 2. Los compuestos identificados son, en su mayoría, aminoácidos y sus derivados, ácidos carboxílicos y compuestos orgánicos, nucleobases y carnitinas. También fueron anotados algunos dipéptidos. Este análisis ha permitido mejorar los resultados de los aminoácidos y sus conjugados obtenidos previamente en un perfil de bilis humana por Duarte y *col.* (46) mediante espectroscopia de NMR.

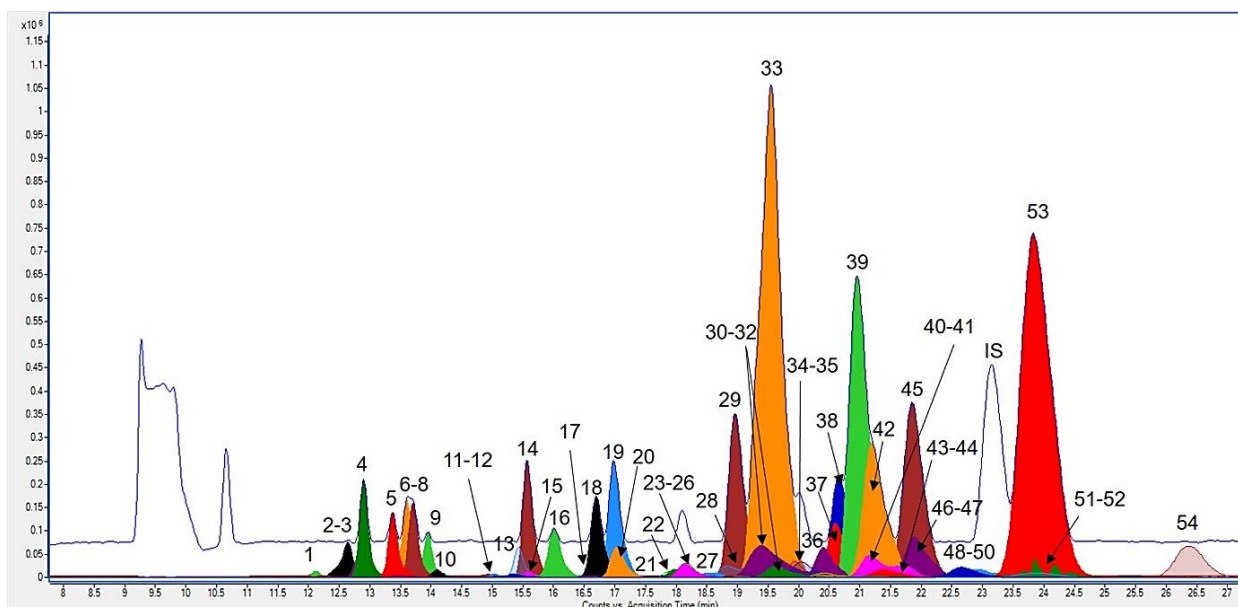


Figura 4. Electroferograma de iones totales (TIC) y electroferograma de iones extraídos (EIC) para los metabolitos encontrados en el perfil de muestra de bilis analizada en polaridad normal y modo de iones positivos mediante CE-TOF/MS. Los números asignados a cada pico se resumen en la Tabla 2.

Recientemente Mori y col. (55) publicaban un trabajo que aúna los resultados biliares de CE-MS con LC-MS en la comparativa que incluye una cohorte limitada de pacientes con: malfuncionamiento pancreaticobiliar sin cáncer (PBM), cáncer del conducto biliar extrahepático (EHBC) y controles con enfermedad benigna. Nuestros resultados, además de la caracterización biliar por CE-MS, reportan un perfil de aminoácidos y derivados más completo, además de unos resultados a nivel de lípidos y ácidos biliares más concretos y, con importante implicación biológica en la carcinogénesis.

Con los resultados obtenidos se analizaron las rutas metabólicas implicadas según los metabolitos que pueden ser caracterizados en el perfil de muestras de bilis a través de CE-MS (Figura 5). Entre ellas destacan las relacionadas con el metabolismo de los aminoácidos y sus derivados, como el metabolismo del glutatión, que tiene un papel importante en la defensa antioxidante, el metabolismo de los nutrientes y la regulación de eventos celulares entre los que se incluyen la expresión génica, síntesis de proteínas y ADN, proliferación celular y apoptosis, transducción de señales o la producción de citoquinas (56).

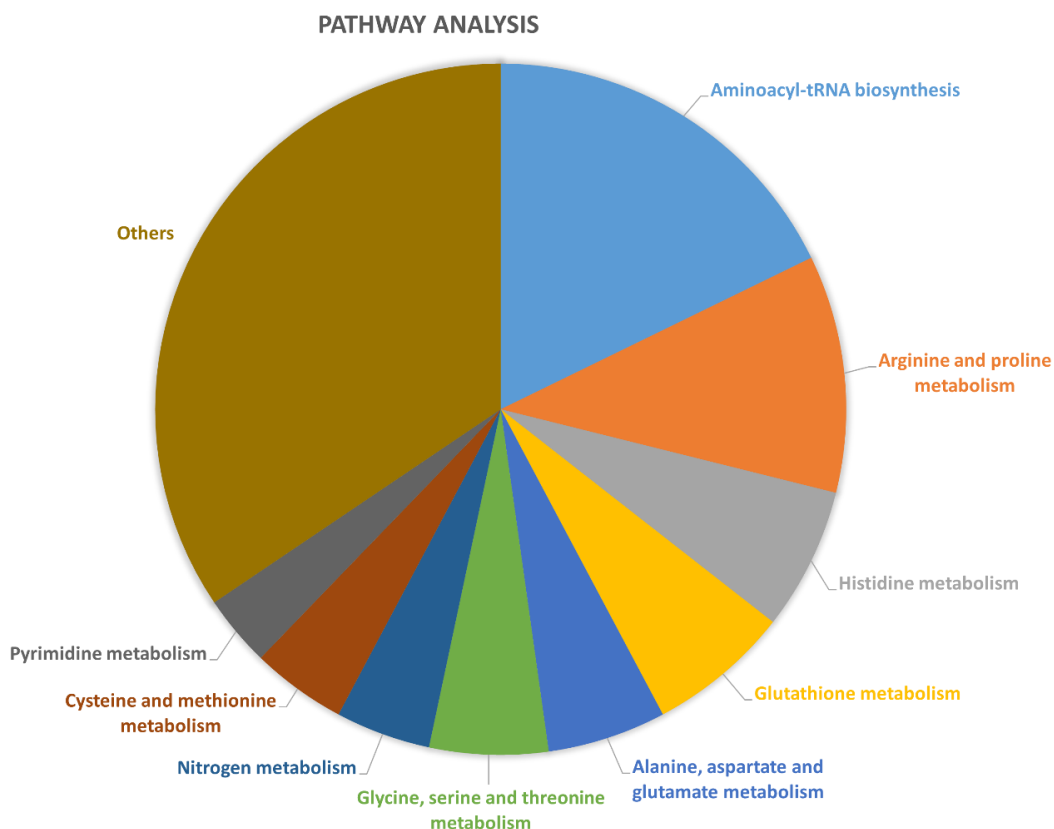


Figura 5. Análisis de las rutas metabólicas más representativas con los resultados obtenidos del caracterizado del perfil de bilis. Los metabolitos encontrados nos permitirán, en el estudio metabólico de bilis de ratas con CCA, encontrar las posibles alteraciones de estas vías.

También, existen pequeñas moléculas polares relacionadas con la biosíntesis de aminoacil-tRNA o el metabolismo de nucleótidos, vía crítica en la replicación del ADN, síntesis de ARN y bioenergía celular. El desorden del metabolismo de los nucleótidos genera moléculas de purina y pirimidina que favorecen la proliferación celular, sello distintivo del cáncer (57, 58).

3.3. Estudio metabolómico no dirigido a través de CE-MS y LC-MS

Con el protocolo de tratamiento de muestra propuesto para CE-MS se analizaron las 38 muestras para el estudio comparativo. Complementariamente, utilizando el protocolo propuesto por el grupo para el análisis de muestras de plasma a través de LC-MS se analizaron las mismas muestras de bilis por esta plataforma (36).

Tras el reprocesado de los datos un análisis de PCA inicial se generó para cada plataforma que verificó la agrupación estrecha de las muestras QC para CE-MS y LC-MS(+), lo que indica la estabilidad y reproducibilidad del sistema. En el caso de LC-MS(-) se observó una tendencia por orden de inyección en los QC. Para corregir esta deriva de la respuesta instrumental dentro de la secuencia se utilizó el algoritmo QC-SVRC (38). Tras el análisis y el tratamiento previo de los datos, se obtuvieron 221 señales, 212 de ellas con un RSD por debajo del 30% para CE-MS. El análisis estadístico univariante se llevó a cabo comparando el grupo CCA y el grupo control, ambos para tiempo 30 semanas (T_30 y C_30, respectivamente), y proporcionó 101 señales químicas con diferencias estadísticamente significativas para la anotación tentativa. Para LC-MS, los conjuntos de datos tras el filtrado y el análisis estadístico consistieron en 703 señales químicas en polaridad positiva y 117 para polaridad negativa.

Tras el estudio de los compuestos para su anotación a través de la plataforma online CMM, siguiendo los criterios explicados en el apartado “*Anotación e identificación*” y, la identificación manual y los resultados reportados por el software Lipid Annotator de los espectros de fragmentación para LC-MS, se encontraron 151 metabolitos que explicarían diferencias metabolómicas en la patología de estudio y que son recogidos en la Tabla 3. Esta tabla proporciona, además de la identidad química, otros parámetros como el tiempo de migración o el de retención, la masa monoisotópica, la fórmula molecular, su categoría bioquímica, el valor de p o el porcentaje de cambio, entre otros. Además, esta recoge la fuente de la que provienen y, si está anotado por más de una plataforma, lo que nos proporciona un nivel superior de confianza en el resultado obtenido.

El análisis de componentes principales (PCA) mostró una tendencia espontánea a la agrupación de muestras según su pertenencia a un grupo (Figura 6). Además, también se aplicaron modelos supervisados como el análisis discriminante de mínimos cuadrados parciales (PLS-DA) y el análisis PLS-DA ortogonal (OPLS-DA) para visualizar la separación entre los grupos de muestras T_30 y C_30, y la fortaleza del modelo. Con el análisis PLS-DA se evaluó la clara tendencia de las muestras a encontrarse en el grupo al que pertenecen lo que indica que los niveles de metabolitos cambian debido a la presencia de la patología. Estas diferencias fueron estudiadas mediante un modelo OPLS-DA y, además, para evitar el riesgo de sobreajuste, el modelo se validó mediante una herramienta de validación cruzada, dejando fuera 1/3 de las muestras por grupo y prediciendo su clasificación. El porcentaje de muestras clasificadas correctamente para el modelo OPLS-DA en cada plataforma se recogen en la Figura 7 incluyendo el valor de R^2 y Q^2 para cada modelo. Todas las gráficas presentaron una diferencia entre R^2 y Q^2 menor a 0.3, lo que confirma la potencia de los modelos. Los modelos presentaron, por tanto, una buena calidad, lo que nos permitió identificar las entidades responsables de esta separación. Los resultados tanto del MVA como del UVA fueron incluidos en la Tabla 3.

A partir de los modelos OPLS-DA se construyeron diferentes gráficos como el de intervalos de confianza del diagrama S (*S-plot*) y, los diagramas con forma de “V” conocidos como “volcán” (*Volcano plots*). En la Figura 8 se muestran los *Volcano plots* para el estudio realizado. En ellos están señaladas las zonas que incluyen los metabolitos más discriminatorios con valores de $VIP \geq 1$ y de $|p(\text{corr})| \geq 0.5$. Los valores de $p(\text{corr})$ negativos (izquierda) indican una disminución de las concentraciones de metabolitos en los genotipos caso frente a control, mientras que los valores de $p(\text{corr})$ positivos (derecha) indican un aumento de las concentraciones de metabolitos en los genotipos casos vs control. Estos resultados están recogidos en la Tabla 3.

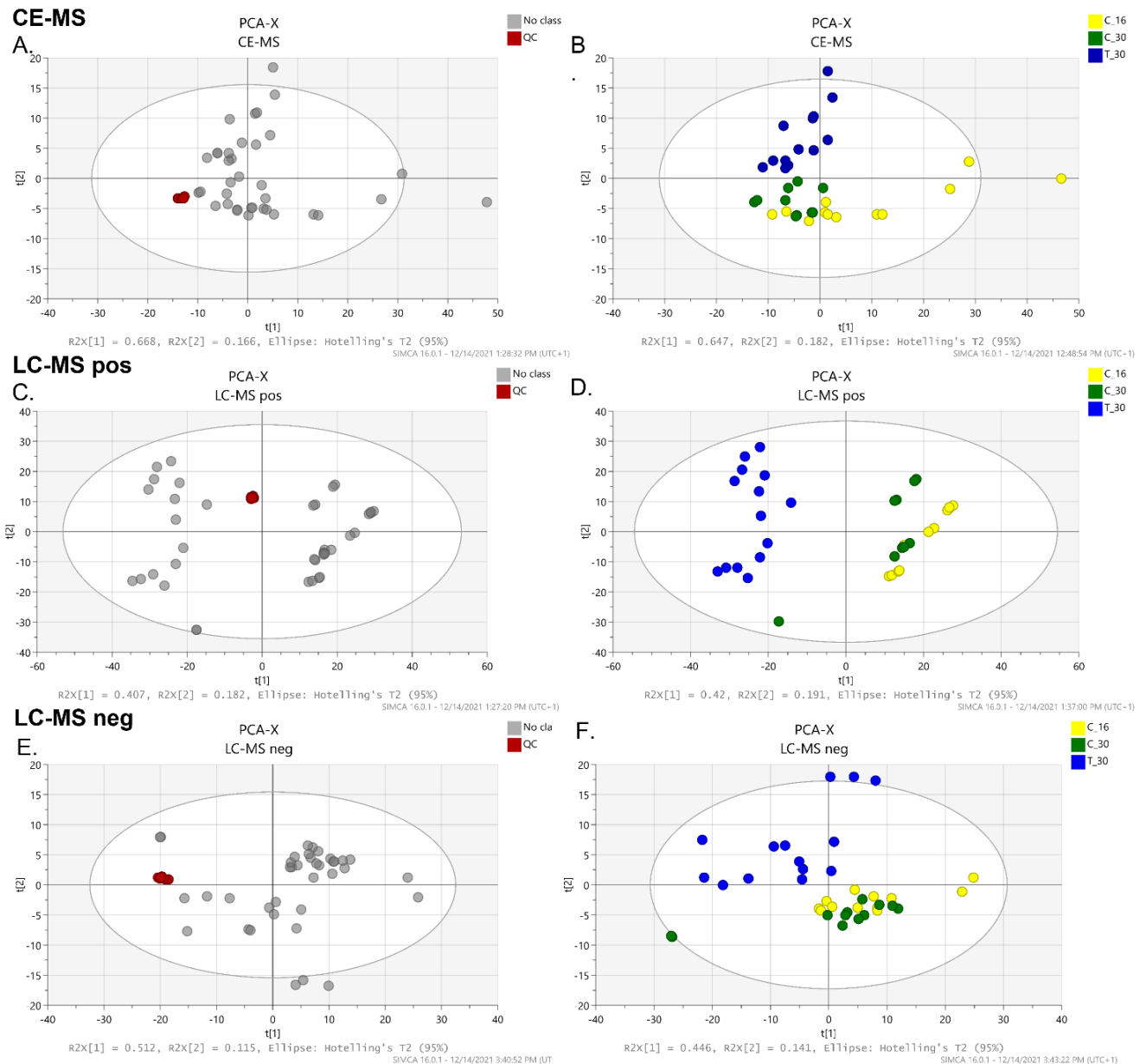


Figura 6. Modelos PCA que muestran la calidad de los datos en las 3 plataformas de análisis empleadas. Gráficos A, C y E: modelos PCA-X para las muestras y los QCs (controles de calidad (rojo), muestras (gris)) para CE-MS, LC-MS(+) y LC-MS(-), respectivamente. **Modelo A**, CE-MS ($R^2 = 0.833$; $Q^2 = 0.802$), **Modelo C**, LC-MS(+) ($R^2 = 0.589$; $Q^2 = 0.539$), **Modelo E**, LC-MS(-) ($R^2 = 0.627$; $Q^2 = 0.569$). Gráficos B, D y F: modelos PCA-X para las muestras de análisis (C_16, controles a tiempo 16 semanas (amarillo); C_30, controles a tiempo 30 semanas (verde); T_30, casos a tiempo 30 semanas (azul)) para CE-MS, LC-MS(+) y LC-MS(-), respectivamente. **Modelo B**, CE-MS ($R^2 = 0.829$; $Q^2 = 0.795$), **Modelo D**, LC-MS(+) ($R^2 = 0.611$; $Q^2 = 0.554$), **Modelo F**, LC-MS(-) ($R^2 = 0.587$; $Q^2 = 0.507$).

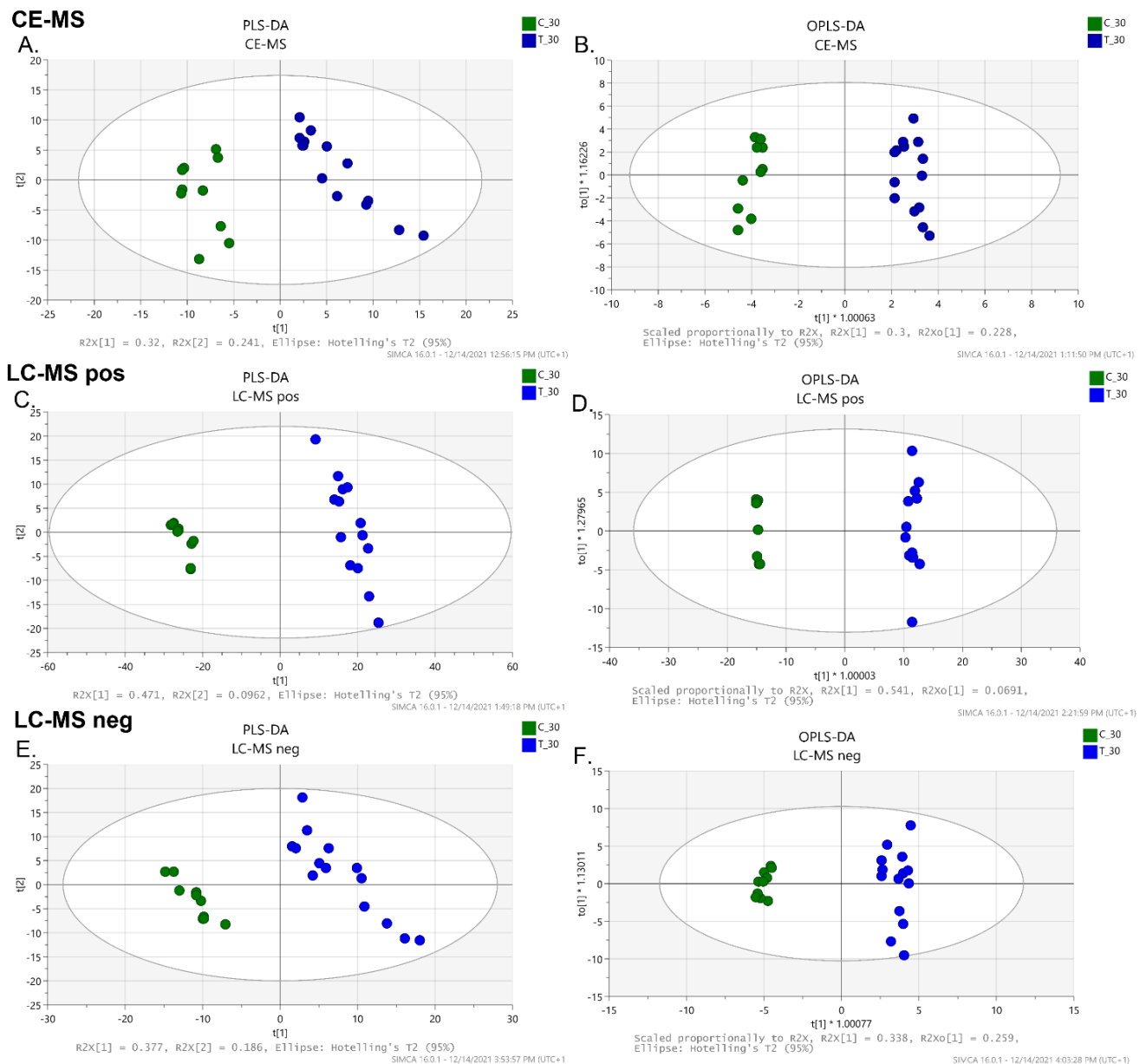


Figura 7. Modelos supervisados PLS-DA y OPLS-DA para el estudio comparativo casos (T_{30} : círculos azules) versus controles (C_{30} : círculos verdes) ambos a tiempo 30 semanas para las tres plataformas de análisis empleadas. Gráficos A, C y E: modelos PLS-DA para CE-MS (modelo A, $R^2 = 0.925, Q^2 = 0.880$), LC-MS(+), (modelo C, $R^2 = 0.993, Q^2 = 0.986$) y LC-MS(-), (modelo E, $R^2 = 0.973, Q^2 = 0.950$). Gráficos B, D y F: modelos OPLS-DA. **Gráfico B:** modelo OPLS-DA para CE-MS con calidad de la varianza explicada y la varianza predicha ($R^2 = 0.969, Q^2 = 0.957$) y el porcentaje de muestras correctamente clasificadas (media \pm desviación estándar) $97.3\% \pm 1.9$ SD. **Gráfico D:** modelo OPLS-DA para LC-MS(+) con calidad de la varianza explicada y la varianza predicha ($R^2 = 0.925, Q^2 = 0.880$) y el porcentaje de muestras correctamente clasificadas $98.7\% \pm 1.9$ SD. **Gráfico F:** modelo OPLS-DA para LC-MS(-) con calidad de la varianza explicada y la varianza predicha ($R^2 = 0.925, Q^2 = 0.880$) y el porcentaje de muestras correctamente clasificadas $95.5\% \pm 1.9$ SD.

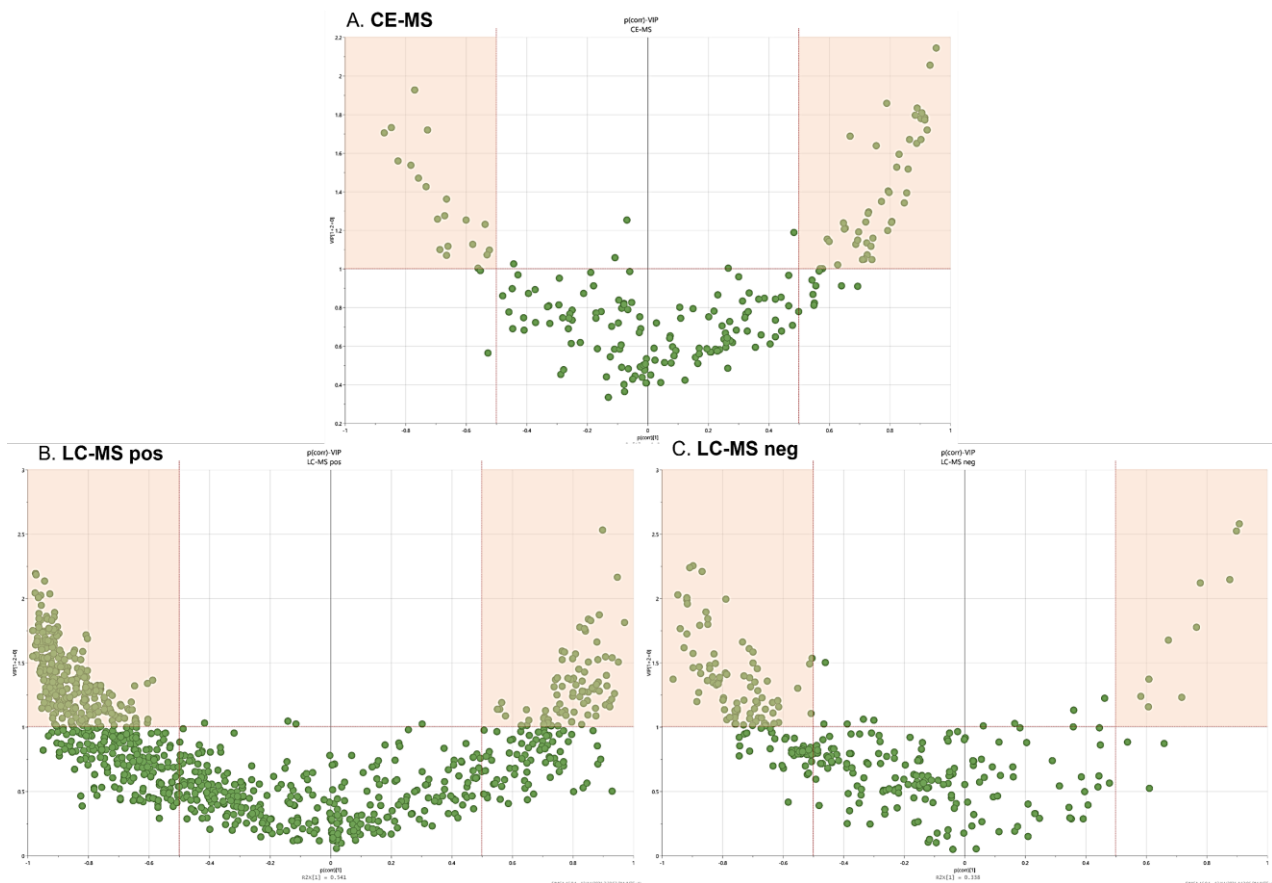


Figura 8. Gráficos Volcano ($p(\text{corr})$ frente a VIP) derivados de los modelos OPLS-DA que se muestran en la figura 7 para cada plataforma. Gráfico A, CE-MS. Gráfico B, LC-MS(+). Gráfico C, LC-MS(-). Las zonas sombreadas representan los metabolitos más representativos para la discriminación caso versus control con valores de VIP ≥ 1 y de $|p(\text{corr})| \geq 0.5$. Los valores de los metabolitos significativos que cumplen estos criterios están resumidos en la Tabla 3.

En cuanto a los resultados obtenidos se observa el gran impacto que presentan diferentes aminoácidos y sus derivados. Los resultados de este trabajo revelaron alteraciones en vías metabólicas donde están implicados ciertos aminoácidos, destacando el metabolismo de la arginina y prolina, el metabolismo de la glicina y serina y el metabolismo del glutatión. Se puede observar la alteración de 17 aminoácidos y sus derivados como ornitina, lisina, arginina, creatina, betaína o tirosina. La mayoría de ellos han sido reportados como posibles validos biomarcadores en enfermedades como el carcinoma hepatocelular (HCC) (10, 59-61), el cáncer de páncreas (PaC) (62-64) o la discriminación entre cáncer de páncreas y pancreatitis (65) y, adenocarcinoma ductal pancreático (PDAC) y pancreatitis crónica (PCr) (66). Zeng y colaboradores (61) definieron dos biomarcadores candidatos para discriminar entre cirrosis y HCC utilizando una muestra de suero. En nuestro estudio, estos dos metabolitos, betaína y

creatina, se han visto alterados con un cambio muy significativo en la creatina. La histidina es otro metabolito reportado en varios estudios sobre PaC (62-64) en muestras de suero y plasma. Fukutake y col.(63) construyeron un panel de aminoácidos libres en plasma para la detección del cáncer de páncreas con concentraciones plasmáticas de histidina y triptófano particularmente disminuidas. En contraste, nuestros resultados muestran en los aminoácidos estadísticamente significativos una tendencia positiva, notablemente aumentados ornitina, arginina, N-N dimetilglicina y citrulina. Esta tendencia positiva de los aminoácidos corrobora su metabolismo anormal en este tipo de enfermedades (67).

Además, utilizando la herramienta online "*Metaboanalyst*", se ha conseguido obtener una visión global de los resultados obtenidos, que nos ayudará con la interpretación de los mismos. La Figura 9 representa las vías metabólicas que presentan mayor impacto por la patología de estudio. Este análisis de rutas remarca como vías principales más alteradas la biosíntesis de aminoacil-tRNA y de ácidos biliares primarios, resultados que se observan claramente al estudiar los metabolitos afectados provenientes del análisis de LC-MS, fundamentalmente. Por otra parte, muchas de las vías relacionadas con el metabolismo y biosíntesis de aminoácidos y sus derivados obtuvieron también puntuaciones altas. El metabolismo de cisteína y metionina, arginina y prolina, glicina, serina y prolina, histidina y la biosíntesis de arginina destacan entre las rutas de mayor impacto, resultados que están claramente relacionados con los datos obtenidos del análisis por CE-MS. Estos resultados son satisfactorios y claramente nos ayudan a comprender mejor los mecanismos subyacentes al CCA.

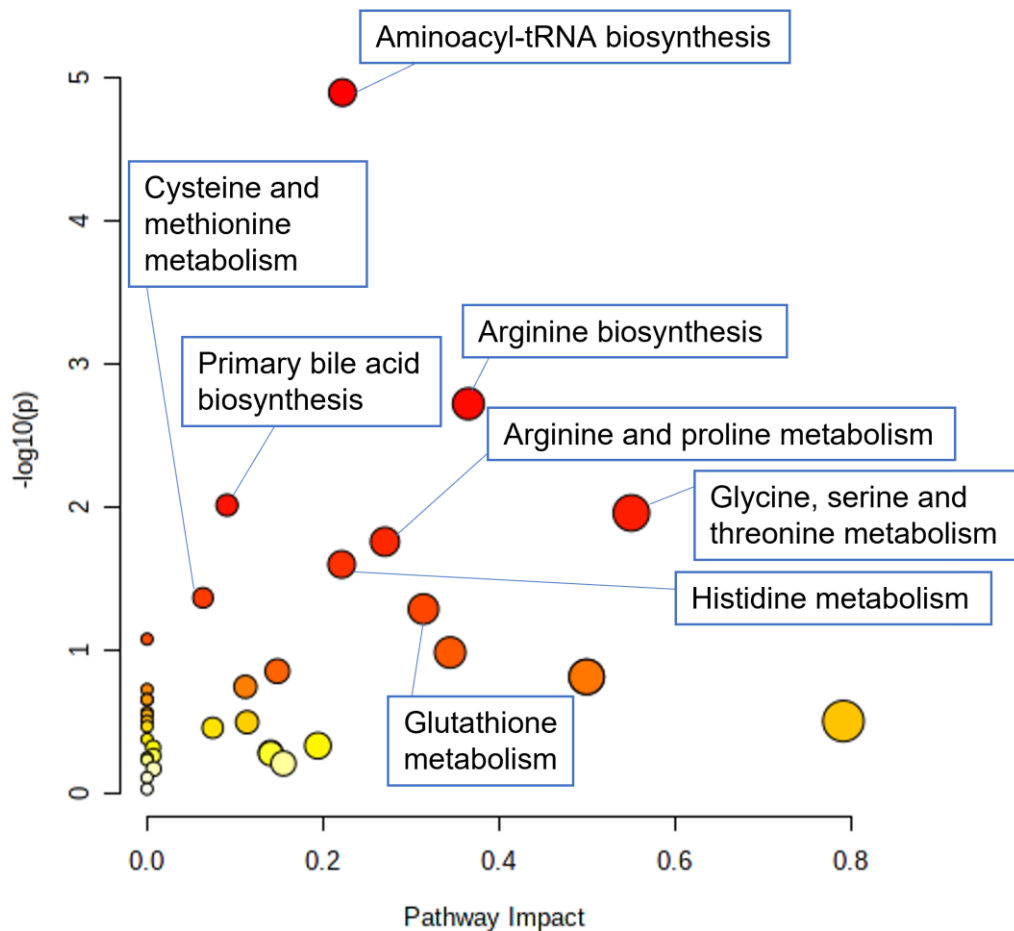


Figura 9. MetPA (del inglés *Metabolim Pathway Analysis*) es el análisis de las vías metabólicas alteradas por la representación de los metabolitos con diferencias significativas entre casos y controles a 30 semanas obtenidos en los análisis de CE-MS y LC-MS para ambas polaridades. En el gráfico, el tamaño y color de los círculos varía según la importancia de la vía y su p valor respectivamente. Así, las vías ubicadas en la esquina superior derecha son las que más afectan a la patología de colangiocarcinoma.

Para complementar esta visión global de los resultados se representó un “*heatmap*” a través de la misma plataforma. Este es una simplificación gráfica de los datos donde se representa el nivel de abundancia de cada metabolito en cada una de las muestras utilizando una escala de colores. La escala de colores va del rojo (aumentados) al azul (disminuidos) y, utiliza agrupamientos jerárquicos, método no supervisado que permite agrupar tanto muestras como variables en conglomerados de mayor tamaño cada vez. La Figura 10 recoge este tipo de gráfico con los 60 metabolitos más representativos de la diferencia entre los grupos. En ella se pueden observar patrones de comportamiento claros. Fijándonos en la mayoría de glicerofosfolípidos, estos se encuentran disminuidos en el grupo caso. Sin embargo, si observamos los

aminoácidos y sus derivados y la mayoría de los esteroides, estos presentan una mayor abundancia debido a la patología.

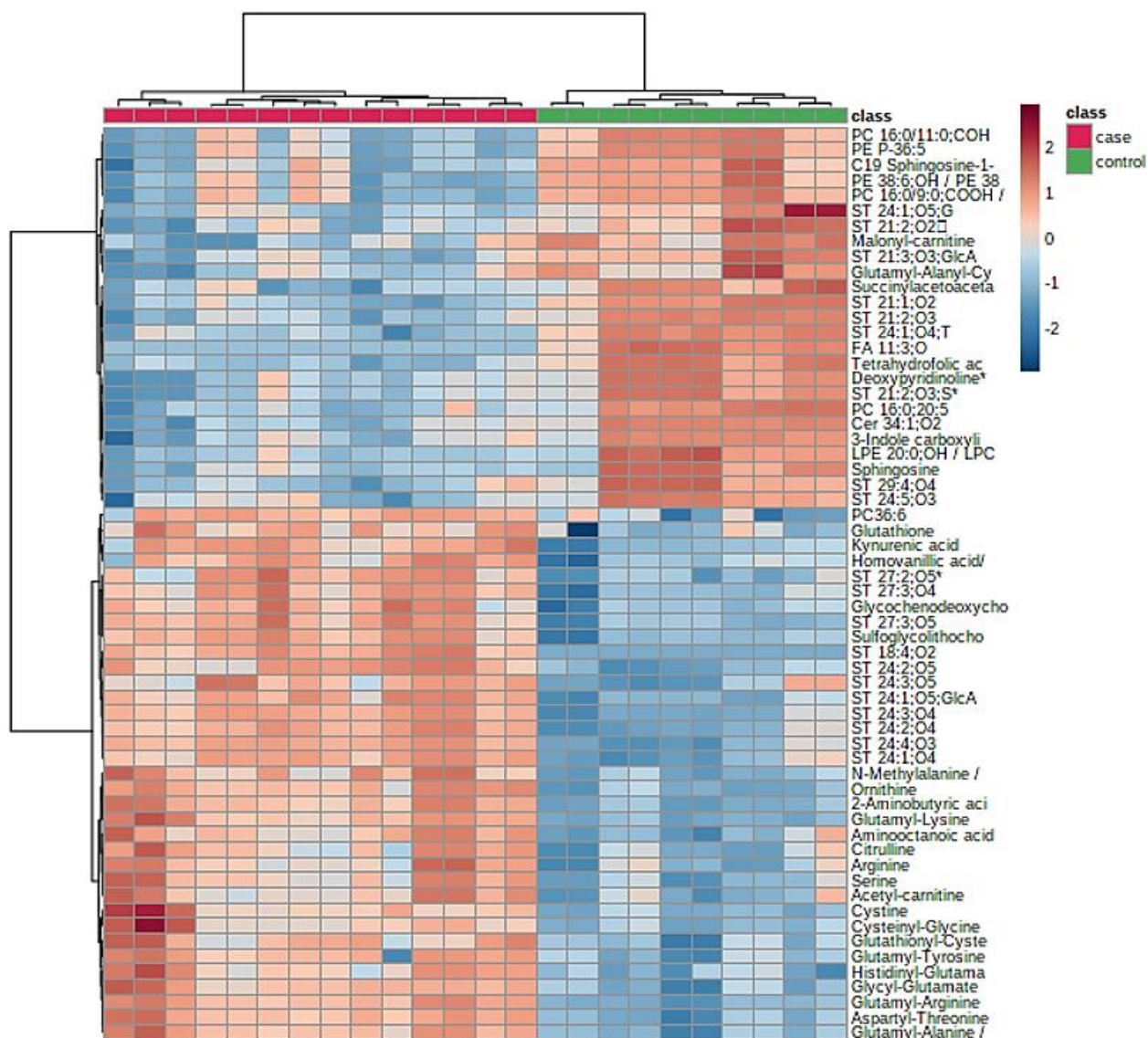


Figura 10. "Heatmap" o mapa de calor de los 60 metabolitos más representativos detectados en la bilis de rata por CE-MS y LC-MS, significativamente alterados por el estado caso/control, o lo que es lo mismo, CCA/no-cáncer. El espectro de color va del rojo al azul e indica el rango de intensidad de las señales respectivamente para cada muestra y para cada metabolito. Los resultados están agrupados por pertenencia a un grupo (rojo, caso; verde, control). Este gráfico también muestra la agrupación jerárquica, tanto de observables cómo de variables.

Con los metabolitos significativos anotados resultado del estudio comparativo, se ha obtenido un gráfico *volcano* conjunto (Figura 11) que permite representar en el mismo gráfico aquellos compuestos que están por encima y por debajo de los umbrales de *fold change* (2 y -2) y *p* valor ≥ 0.05 para todas las plataformas utilizadas. En el eje de abscisas se representa las diferencias obtenidas de los metabolitos con el $\log_2(\text{FC})$, mientras que en el de ordenadas se representa el $-\log$ del *p* valor. Esto permiten la rápida identificación de los resultados más significativos entre grandes cantidades de datos. En la esquina superior derecha, marcados en rojo, se encuentran lo metabolitos aumentados por la patología, entre los que destacan gran parte de los aminoácidos y sus derivados y los ácidos biliares de C24 y C27 sin modificaciones. Por su parte la esquina superior izquierda representa, en color azul, los metabolitos significativos disminuidos en CCA entre los que destacan los ácidos grasos, las fosfatidilcolinas, los ácidos biliares conjugados con taurina y glicina, ceramidas o el THF (tetrahydrofolato) entre otros.

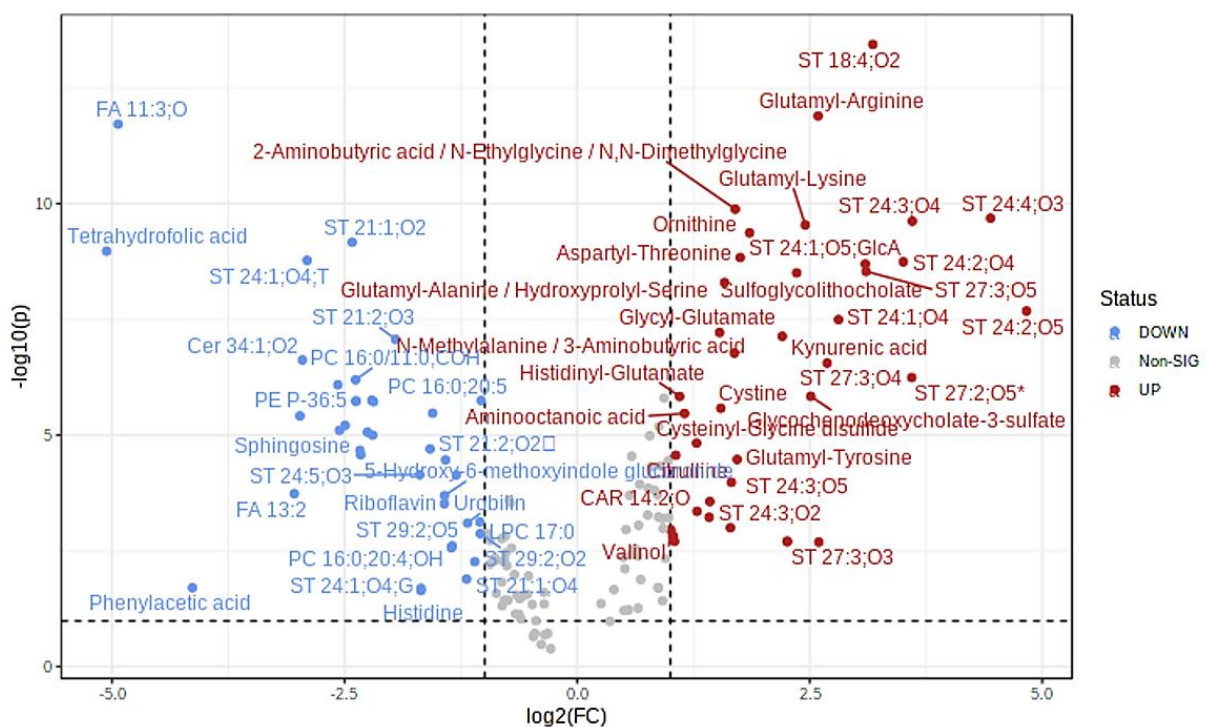


Figura 11. Metabolitos importantes seleccionadas por el gráfico Volcano con umbral de cambio $\log_2(\text{FC})$ y umbral de prueba *t* ($-\log(p)$ valor)). Los círculos rojos representan características por encima del umbral. Cuanto más alejada esté su posición de (0,0), más significativa será la característica.

Del análisis exhaustivo de los resultados obtenidos respecto a los esteroides y sus derivados es importante remarcar como se han detectado alterados los principales grupos como los esteroides de 18 y 21 carbonos (C18 y C21, respectivamente) y los ácidos biliares, alcoholes y sus derivados de 24 y 27 carbonos (correspondientes a C24 y C27). En todos estos aparece una distribución característica donde los derivados del colesterol y de los esteroides de C21 se encuentran disminuidos en los casos. Dentro de la categoría C21 se han obtenido derivados sulfato y glucurónido conjugados, todos ellos con una regulación a la baja. Por otra parte, los derivados de ácidos biliares de C27 se mostraron aumentados con respecto a los individuos no-cáncer. El grupo más destacado es el de los ácidos biliares y derivados de C24 entre los que podemos encontrar los ácidos biliares primarios, secundarios y los derivados conjugados de taurina y glicina fundamentalmente. En estos, la distribución es, también, característica, y se resume en unos niveles aumentados en los ácidos biliares primarios, secundarios y los glucurónido conjugados; los niveles de derivados de C24 conjugados con glicina y taurina se encuentran disminuidos en la comparación CCA/no-cáncer. Todos estos resultados se encuentran resumidos en la Figura 12.

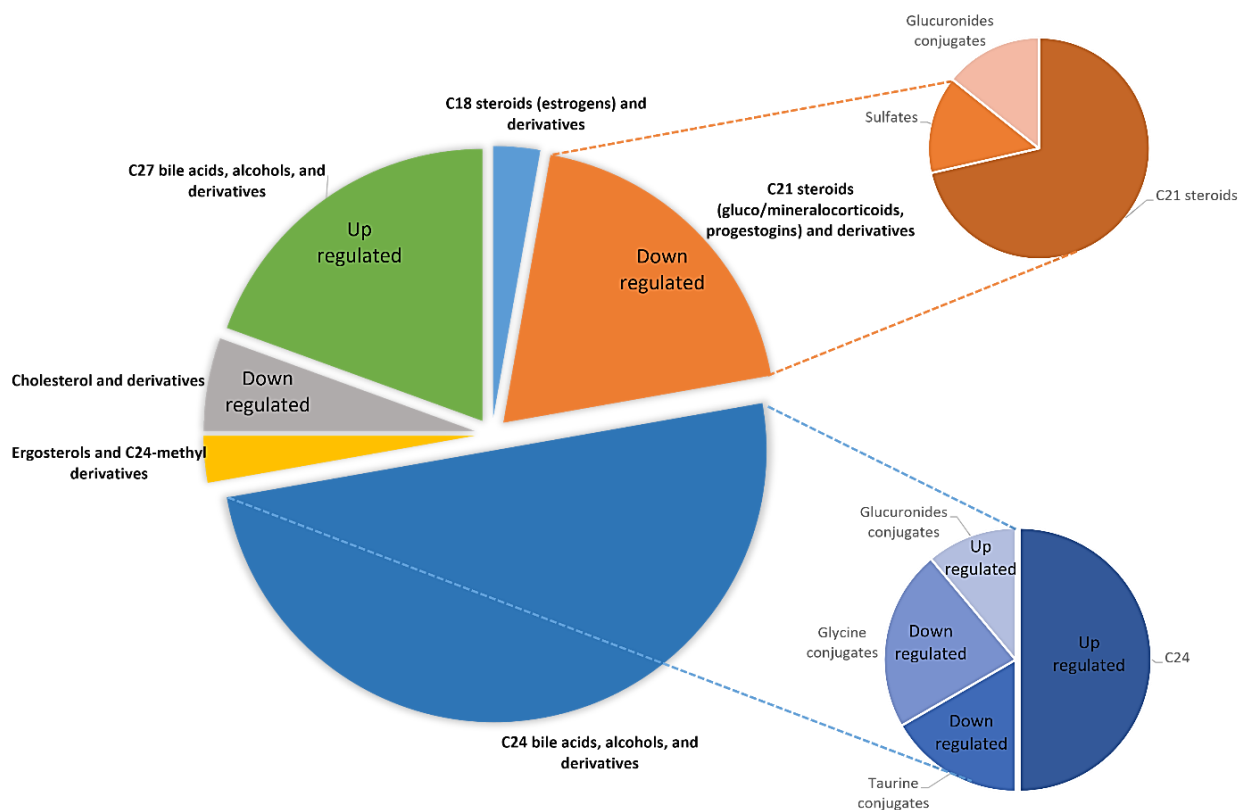


Figura 12. Clasificación de los esteroides y derivados resultantes del análisis comparativo en muestras de bilis de rata para la patología de CCA a través de, fundamentalmente, la plataforma LC-MS. La clasificación de los derivados de C21 y C24 se ha dividido a su vez en formas conjugadas, sulfato o nativas (sin modificación).

En lo reportado hasta la fecha sobre CCA se han recogido diferencias sustanciales en la composición molecular de las fosfolípidos hepáticas y biliares, por lo que se podría estar hablando de un grupo de fosfatidilcolina (PC) dedicadas a la secreción biliar (68). En concreto, Macias y *col.* (69) reportaron niveles disminuidos de PC y, varias LPC fueron establecidas como firma metabólica relevante para el CCA e investigadas por su potencial diagnóstico (8). Nuestros resultados revelaron niveles significativamente disminuidos en varias LPC, entre ellas LPC 17:0, 18:0 o 20:5, y la mayor parte de las PC, lo que concuerda con lo reportado. Las PC son fosfolípidos biliares dominantes con efecto citoprotector. La alteración y, por lo tanto, la ausencia de transporte de estos hacia la bilis conduce a una exposición prolongada del epitelio biliar con la bilis "tóxica" lo que favorece el desarrollo de la patología (70, 71). Algo similar ocurre con los esfingolípidos significativos obtenidos, entre los que se encuentran Cer 34:1;O₂, SM 34:2;O₂ o SM 18:1;16:0;O₂, y que de acuerdo con hallazgos previos (11, 22, 72, 73), se observa una reducción general de la concentración de PC y esfingolípidos en la bilis de CCA.

El estudio de la firma metabólica también señaló la bilirrubina como potencial biomarcador (8). Esta biomolécula es el pigmento color amarillo biliar que resulta de la degradación de los eritrocitos y que se almacena en la vesícula biliar. La bilirrubina, por lo tanto, puede ser respuesta de inflamación hepatobiliar pero no es sensible de CCA, ya que se ha demostrado que sus niveles pueden aumentar drásticamente en estenosis benignas como la colangitis esclerosante primaria o los cálculos biliares (74).

Otro hallazgo significativo de este estudio es la disminución de fosfolípidos oxidados; al igual que las PC en CCA, sus derivados oxidados presentan la misma tendencia. Destacan derivados hidroxilados de PAPC (PC 16:0;20:4) y de PLPC (PC 16:0;18:2) o de cadena trunca como PC 16:0/9:0;COOH o PC 16:0/11:0;COH. Para identificar estas especies se siguió el flujo de trabajo propuesto en el capítulo 2, comparando los fragmentos adquiridos de los espectros MS/MS, con los disponibles en las bases de datos, si fuese posible, o bien, comparándolo con el espectro de fragmentación de la especie nativa en ionización positiva y negativa.

Aunque es indiscutible la relación patológica entre el cáncer y el estrés oxidativo, las implicaciones fisiológicas que juegan las distintas iso-formas o incluso los niveles de concentración de las distintas especies, son condiciones que no están bien definidas hasta la fecha. A menudo, niveles disminuidos de estos metabolitos, están relacionados con propiedades protectoras (75, 76), pero el efecto específico que provoca la

disminución de estas especies, o la causa de esta misma, son preguntas que deben seguir investigándose para esclarecerlas.

El estado redox celular está representado, además, por compuestos como el ácido oftálmico, glutatión (GSH) y glutatión disulfuro (GSSG) y, que se incluyeron en el análisis para la semicuantificación a través de CE-MS. GSH es un tripéptido endógeno formado por γ -glutamato, cisteína y glicina, siendo la cisteína el aminoácido limitante. Cuando la cisteína se reemplaza por ácido 2-aminobutírico se forma el ácido oftálmico en el hígado y, este es conducido al sistema circulatorio. Estudios metabólicos han revelado el papel del ácido oftálmico como biomarcador de estrés oxidativo (77, 78) y la relación que le une a GSH y GSSG.

El tetrahidrofolato o THF es esencial para la función biológica normal actuando como co-factor enzimático en la biosíntesis de moléculas con un importante papel biológico como los nucleótidos de purina y pirimidina (componentes básicos del ADN y ARN) y, numerosos aminoácidos como la metionina o la glicina (79). Por lo tanto, como presentan los resultados de este estudio, su deficiencia puede conducir a la inhibición de la síntesis de ADN, ARN y proteínas, fundamental para el crecimiento y supervivencia celular.

La importancia del metabolismo de ácidos biliares, fosfatidilcolinas y aminoácidos han sido resaltados en este estudio con individuos con CCA. Estos hallazgos sugieren alteraciones importantes en las vías metabólicas moleculares del CCA y requieren de una validación adicional.

3.4. Análisis semicuantitativo a través de CE-MS

A la vista de los resultados obtenidos y, teniendo presente los estudios preliminares realizados por el grupo colaborador (80), se llevó a cabo el análisis semicuantitativo de los metabolitos significativos con mayor relevancia para los resultados de proteómica obtenidos. Los trastornos de las vías hepatobiliares presentan como rasgo común, la remodelación metabólica, donde las células cancerígenas adaptan su metabolismo a los requerimientos proliferativos inherentes, descrito por primera vez por Warburg (81). ROS y RNS, especies reactivas de oxígeno y nitrógeno respectivamente, dan como resultado modificaciones del estado oxidativo celular. Estas especies presentan una alta reactividad con distintas biomoléculas provocando daños oxidativos permanentes irreversibles. Los organismos vivos eliminan estos radicales libres a través de antioxidantes endógenos los cuales reaccionan con ellos antes de que estos interaccionen con las distintas biomoléculas induciendo el daño celular. Glicina,

serina, ornitina, lisina, arginina, citrulina, cistina, ácido oftálmico y glutatión oxidado fueron seleccionados para el análisis como posibles candidatos de la respuesta redox a nivel celular.

Para obtener los valores de concentración de estos compuestos en bilis, se realizó la semicuantificación de cada uno de ellos con relación al pico del patrón interno ($A_{\text{metabolito}} / A_{\text{IS}}$) tanto en muestras como en patrones. La semicuantificación se realizó interpolando el valor integrado $A_{\text{metabolite}} / A_{\text{IS}}$ en las curvas de calibración construidas a partir del análisis de tres réplicas de cada estándar comercial relacionado con el estándar interno (Tabla 4).

Tabla 2. Compuestos anotados en la caracterización de muestras de bilis mediante CE-MS. La columna 1 muestra los metabolitos extraídos en el perfil biliar de la figura 4.

Número Fig. 4	Compuesto	Formula	Masa Monoisotópica (Da)	Error (ppm)	MT (min)	RMT (min)	HMDB ID	Clase bioquímica
1	Trimethylamine N-oxide	C ₃ H ₉ NO	75.0684	4.0	12.05	0.53	HMDB00925	Amines and polyamines
2	Choline	C ₅ H ₁₃ NO	103.0997	0.1	12.57	0.55	HMDB00097	Alcohols and polyols
3	Ornithine	C ₅ H ₁₂ N ₂ O ₂	132.0899	0.2	12.72	0.56	HMDB00214	Amino acids and conjugates
4	Lysine	C ₆ H ₁₄ N ₂ O ₂	146.1055	0.2	12.8	0.56	HMDB00182	Amino acids and conjugates
5	Arginine	C ₆ H ₁₄ N ₄ O ₂	174.1117	0.1	13.27	0.58	HMDB00517	Amino acids and conjugates
	Prolyl-Histidine	C ₁₁ H ₁₆ N ₄ O ₃	252.1222	3.6	13.38	0.59	HMDB0029019	Peptides
6	Creatinine	C ₄ N ₃ H ₇ O	113.0589	0.8	13.5	0.59	HMDB00562	Imidazoles and derivatives
7	N-Methylnicotinamide	C ₇ H ₈ N ₂ O	136.0637	4.4	13.54	0.60	HMDB0003152	Pyridines and derivatives
8	Histidine	C ₆ H ₉ N ₃ O ₂	155.0695	1.1	13.62	0.60	HMDB00177	Amino acids and conjugates
9	Niacinamide	C ₆ H ₆ N ₂ O	122.0480	4.1	13.85	0.61	HMDB01406	Pyridines and derivatives
10	3-Methylhistidine	C ₇ H ₁₁ N ₃ O ₂	169.0851	1.0	13.99	0.61	HMDB0000479	Amino acids and conjugates
	3-Methylindole	C ₉ H ₉ N	131.0735	3.8	14.91	0.66	HMDB0000466	Indoles and derivatives
11	Imidazole 4-acetate	C ₅ H ₆ N ₂ O ₂	126.0429	11.9	14.92	0.66	HMDB02024	Imidazoles and derivatives
12	Acetylcholine	C ₇ H ₁₅ NO ₂	145.1103	10.5	14.99	0.66	HMDB0000895	Cholines
13	Cysteinyl-glycine	C ₅ H ₁₀ N ₂ O ₃ S	178.0412	0.6	15.2	0.67	HMDB00078	Peptides
	Methylimidazoleacetic acid	C ₆ H ₈ N ₂ O ₃	140.0586	0.6	15.28	0.67	HMDB02820	Imidazoles and derivatives
	Triethanolamine	C ₆ H ₁₅ NO ₃	149.1052	3.4	15.29	0.67	HMDB0032538	Amine
14	Glycine	C ₂ H ₅ NO ₂	75.0320	7.6	15.41	0.68	HMDB00123	Amino acids and conjugates
15	4-Guanidinobutanoic acid	C ₅ H ₁₁ N ₃ O ₂	145.0851	1.4	15.43	0.68	HMDB03464	Organic acids and derivatives
	Imidazolelactic acid	C ₆ H ₈ N ₂ O ₃	156.0534	14	15.45	0.68	HMDB0002320	Imidazoles and derivatives

Número Fig. 4	Compuesto	Formula	Masa Monoisotopica (Da)	Error (ppm)	MT (min)	RMT (min)	HMDB ID	Clase bioquímica
	2-Amino-octanoic acid	C ₈ H ₁₇ NO ₂	159.1259	3.8	15.59	0.69	HMDB0000991	Amino acids and conjugates
16	Carnitine	C ₇ H ₁₅ NO ₃	161.1052	1.9	15.85	0.70	HMDB00062	Amino acids and conjugates
	Glutamyl-Lysine	C ₁₁ H ₂₁ N ₃ O ₅	275.1481	0.1	16.08	0.71	HMDB03869	Peptides
	Indole-3-propionic acid	C ₁₁ H ₁₁ NO ₂	189.0790	6.5	16.35	0.72	HMDB02302	Indoles and derivatives
	Histidinyl-Glutamate	C ₁₁ H ₁₆ N ₄ O ₅	284.1120	0.0	16.35	0.72	HMDB002884	Peptides
	Glutamyl-Arginine	C ₁₁ H ₂₁ N ₅ O ₅	303.1543	0.3	16.39	0.72	HMDB0028813	Peptides
17	Pyridoxal	C ₈ H ₉ NO ₃	167.0582	2.4	16.4	0.72	HMDB01545	Pyridines and derivatives
18	Creatine	C ₄ H ₉ N ₃ O ₂	131.0695	0.6	16.52	0.73	HMDB00064	Amino acids and conjugates
19	L-Alanine	C ₃ H ₇ NO ₂	89.0477	5.9	16.8	0.74	HMDB00161	Peptides
20	Acetylcarnitine (CAR 2:0)	C ₉ H ₁₇ NO ₄	203.1158	0.3	16.86	0.74	HMDB00201	Carnitines
	Glycylproline	C ₇ H ₁₂ N ₂ O ₃	172.0848	1.8	16.94	0.74	HMDB00721	Peptides
21	Propionyl-L-carnitine (CAR 3:0)	C ₁₀ H ₁₉ NO ₄	217.1314	7.4	17.42	0.77	HMDB00824	Carnitines
	Alanyl-Proline	C ₈ H ₁₄ N ₂ O ₃	186.1004	0.2	17.61	0.77	HMDB0028695	Peptides
	L-prolyl-L-proline	C ₁₀ H ₁₆ N ₂ O ₃	212.1161	0.0	17.72	0.78	HMDB0011180	Peptides
22	2'-Deoxycytidine	C ₉ H ₁₃ N ₃ O ₄	227.0906	0.4	17.78	0.78	HMDB00014	Nucleosides and conjugates
	1-(beta-D-Ribofuranosyl)-1,4-dihydronicotinamide	C ₁₁ H ₁₆ N ₂ O ₅	256.1059	3.9	17.9	0.79	HMDB0011648	Carbohydrate amide
	gamma-L-Glutamyl-L-pipecolic acid	C ₁₁ H ₁₈ N ₂ O ₅	258.1216	3.5	17.9	0.79	HMDB0038614	Peptides
23	Isobutyryl-L-carnitine	C ₁₁ H ₂₁ NO ₄	231.1470	4.3	17.91	0.79	HMDB00736	Carnitines
24	L-Cysteinylglycine disulfide	C ₈ H ₁₅ N ₃ O ₅ S ₂	297.0453	1.7	17.95	0.79	HMDB00709	Peptides
25	2-Aminobutyric acid	C ₄ H ₉ NO ₂	103.0633	10.7	17.98	0.79	HMDB00452	Amino acids and conjugates

Número Fig. 4	Compuesto	Formula	Masa Monoisotopica (Da)	Error (ppm)	MT (min)	RMT (min)	HMDB ID	Clase bioquímica
	Cysteinyl-Alanine	C ₆ H ₁₂ N ₂ O ₃ S	192.0569	0.5	17.98	0.79	HMDB0028684	Peptides
	Cysteinyl-Serine	C ₆ H ₁₂ N ₂ O ₄ S	208.0518	8.2	18.19	0.80	HMDB0028784	Peptides
	Alanyl-Hydroxyproline	C ₈ H ₁₄ N ₂ O ₄	202.0954	12.1	18.25	0.80	HMDB0028688	Peptides
	Prolyl-Serine	C ₈ H ₁₄ N ₂ O ₄	202.0954				HMDB0029026	Peptides
26	Cytidine	C ₉ H ₁₃ N ₃ O ₅	243.0865	6.2	18.29	0.80	HMDB00089	Nucleosides and conjugates
27	D-Alanyl-D-Valine	C ₈ H ₁₆ N ₂ O ₃	188.1161	1.1	18.33	0.81	HMDB0028700	Peptides
	Glycyl-L-leucine	C ₈ H ₁₆ N ₂ O ₃	188.1161				HMDB00759	Peptides
	N6-Acetyl-L-Lysine	C ₈ H ₁₆ N ₂ O ₃	188.1161				HMDB00206	Amino acids and conjugates
	5-Methyldeoxycytidine	C ₁₀ H ₁₅ N ₃ O ₄	241.1063	2.5	18.45	0.81	HMDB02224	Nucleosides and conjugates
	Valyl-Proline	C ₁₀ H ₁₈ N ₂ O ₃	214.1317	0.0	18.58	0.82	HMDB0029135	Peptides
	Cysteinyl-Cysteine	C ₆ H ₁₂ N ₂ O ₃ S ₂	224.0289	3.1	18.63	0.82	HMDB0028772	Peptides
	Asparaginylyl-Proline	C ₉ H ₁₅ N ₃ O ₄	229.1062	7.4	18.69	0.82	HMDB0028739	Peptides
28	Serine	C ₃ H ₇ NO ₃	105.0426	1.0	18.7	0.82	HMDB00187	Amino acids and conjugates
29	Valine	C ₅ H ₁₁ NO ₂	117.0790	2.7	18.73	0.82	HMDB00883	Amino acids and conjugates
	5-Methylcytidine	C ₁₀ H ₁₅ N ₃ O ₅	257.1028	3.5	18.87	0.83	HMDB00982	Nucleosides and conjugates
	Prolyl-Methionine	C ₁₀ H ₁₈ N ₂ O ₃ S	246.1038	6.5	19.01	0.84	HMDB0029023	Peptides
30	5'-Methylthioadenosine	C ₁₁ H ₁₅ N ₅ O ₃ S	297.0896	1.0	19.14	0.84	HMDB01173	Nucleosides and conjugates
31	Leucyl-Proline	C ₁₁ H ₂₀ N ₂ O ₃	228.1474	3.0	19.15	0.84	HMDB11175	Peptides
	Isoleucyl-Proline	C ₁₁ H ₂₀ N ₂ O ₃	228.1474				HMDB0011174	Peptides
32	L-Pipecolic acid	C ₆ H ₁₁ NO ₂	129.0790	10.7	19.23	0.85	HMDB0000716	Amino acids and conjugates
33	Isoleucine	C ₆ H ₁₃ NO ₂	131.0946	4.4	19.3	0.85	HMDB00172	Amino acids and conjugates

Número Fig. 4	Compuesto	Formula	Masa Monoisotopica (Da)	Error (ppm)	MT (min)	RMT (min)	HMDB ID	Clase bioquímica
	Leucine	C ₆ H ₁₃ NO ₂	131.0946				HMDB00687	Amino acids and conjugates
34	Asparagine	C ₄ H ₈ N ₂ O ₃	132.0535	11.4	19.68	0.87	HMDB00168	Amino acids and conjugates
35	4-Aminobenzoic acid	C ₇ H ₇ NO ₂	137.0477	5.8	19.76	0.87	HMDB01392	Organic acids and derivatives
	Leucyl-Valine	C ₁₁ H ₂₂ N ₂ O ₃	230.1630	3.5	19.76	0.87	HMDB0028942	Peptides
	Isoleucyl-Valine	C ₁₁ H ₂₂ N ₂ O ₃	230.1630				HMDB0028920	Peptides
	Threoninyl-Leucine	C ₁₀ H ₂₀ N ₂ O ₄	232.1423	0.9	19.77	0.87	HMDB0029065	Peptides
	Threoninyl-IsoLeucine	C ₁₀ H ₂₀ N ₂ O ₄	232.1423				HMDB0029064	Peptides
	Aspartyl-L-proline	C ₉ H ₁₄ N ₂ O ₅	230.0903	2.2	19.93	0.88	HMDB0002335	Peptides
	Leucyl-Leucine	C ₁₂ H ₂₄ N ₂ O ₃	244.1787	2.5	20.15	0.89	HMDB0028933	Peptides
	Isoleucyl-Leucine	C ₁₂ H ₂₄ N ₂ O ₃	244.1787				HMDB0028911	Peptides
36	Phenylalanyl-Threonine	C ₁₃ H ₁₈ N ₂ O ₄	266.1267	9.4	20.16	0.89	HMDB0029005	Peptides
37	Methionine	C ₅ H ₁₁ NO ₂ S	149.0511	1.0	20.17	0.89	HMDB00696	Amino acids and conjugates
	Hydroxy glutamate	C ₅ H ₉ NO ₅	163.0480	4	20.19	0.89	HMDB00002273	Amino acids and conjugates
38	L-Glutamine	C ₅ H ₁₀ N ₂ O ₃	146.0691	0.4	20.31	0.89	HMDB00641	Amino acids and conjugates
	L-Proline	C ₅ H ₉ NO ₂	115.0633	4.1	20.39	0.90	HMDB00162	Amino acids and conjugates
	6-Keto-decanoylcarnitine	C ₁₇ H ₃₃ NO ₄	329.2202	0.3	20.43	0.90	HMDB0013202	Organic acids and derivatives
	Leucyl-Aspartate	C ₁₀ H ₁₈ N ₂ O ₅	246.1216	5.3	20.61	0.91	HMDB0028925	Peptides
	Isoleucyl-Aspartate	C ₁₀ H ₁₈ N ₂ O ₅	246.1216				HMDB0028903	Peptides
	1-Aminocyclopropane-1-carboxylate	C ₄ H ₇ NO ₂	101.0477	9.0	20.68	0.91	HMDB0036458	Amino acids and conjugates
39	Glutamic acid	C ₅ H ₉ NO ₄	147.0532	2.3	20.68	0.91	HMDB00148	Amino acids and conjugates
40	N-N dimetilglicine	C ₄ H ₉ NO ₂	103.0633	8.7	20.69	0.91	HMDB00092	Amino acids and conjugates

Número Fig. 4	Compuesto	Formula	Masa Monoisotopica (Da)	Error (ppm)	MT (min)	RMT (min)	HMDB ID	Clase bioquímica
41	Tryptophan	C ₁₁ H ₁₂ N ₂ O ₂	204.0899	0.9	20.88	0.92	HMDB00929	Amino acids and conjugates
42	Phenylalanine	C ₉ H ₁₁ NO ₂	165.0790	8.6	20.92	0.92	HMDB00159	Amino acids and conjugates
43	L-Cystine	C ₆ H ₁₂ N ₂ O ₄ S ₂	240.0239	2.3	20.94	0.92	HMDB00192	Amino acids and conjugates
44	DL-Citrulline	C ₆ H ₁₃ N ₃ O ₃	175.0957	0.5	20.96	0.92	HMDB00904	Amino acids and conjugates
45	Betaine	C ₅ H ₁₁ NO ₂	117.0790	5.3	21.52	0.95	HMDB00043	Amino acids and conjugates
46	Tyrosine	C ₉ H ₁₁ NO ₃	181.0739	0.6	21.56	0.95	HMDB00158	Amino acids and conjugates
47	Aspartic acid	C ₄ H ₇ NO ₄	133.0375	2.3	21.8	0.96	HMDB00191	Amino acids and conjugates
48	Methylcysteine	C ₄ H ₉ NO ₂ S	135.0354	3.7	21.87	0.97	HMDB02108	Amino acids and conjugates
49	Glutathione	C ₁₀ H ₁₇ N ₃ O ₆ S	307.0838	1.6	22.29	0.98	HMDB00125	Peptides
50	Cysteineglutathione disulfide	C ₁₃ H ₂₂ N ₄ O ₈ S ₂	426.0879	1.6	22.29	0.98	HMDB0000656	Oligopeptides
51	Methionine sulfoxide	C ₅ H ₁₁ NO ₃ S	165.0460	0.8	22.57	0.99	HMDB02005	Amino acids and conjugates
	Glycyl-Glutamate	C ₇ H ₁₂ N ₂ O ₅	204.0746	0.1	23.03	1.01	HMDB0028840	Peptides
52	Trans-4-Hydroxyproline	C ₅ H ₉ NO ₃	131.0582	3.1	23.07	1.01	HMDB00725	Amino acids and conjugates
	Methionyl-Serine	C ₈ H ₁₆ N ₂ O ₄ S	236.0831	4.2	23.15	1.02	HMDB0028982	Peptides
	AICAR	C ₉ H ₁₅ N ₄ O ₈ P	338.0627	2.4	23.4	1.03	HMDB0001517	Nucleosides and conjugates
53	Glutathione disulfide	C ₂₀ H ₃₂ N ₆ O ₁₂ S ₂	612.1520	1.4	23.44	1.03	HMDB03337	Peptides
	Hydroxypropyl-Serine	C ₈ H ₁₄ N ₂ O ₅	218.0902	0.0	23.54	1.03	HMDB0028872	Peptides
	Glutamyl-Alanine	C ₈ H ₁₄ N ₂ O ₅	218.0903				HMDB0003764	Peptides
	1-methylguanosine	C ₁₁ H ₁₅ N ₅ O ₅	297.1073	1.3	23.96	1.05	HMDB01563	Nucleoside
	Aspartyl-Threonine	C ₈ H ₁₄ N ₂ O ₆	234.0851	3.0	24.14	1.06	HMDB0028763	Peptides
	Methionyl-Alanine	C ₈ H ₁₆ N ₂ O ₃ S	220.0882	11.8	24.17	1.06	HMDB0028966	Peptides

Número Fig. 4	Compuesto	Formula	Masa Monoisotopica (Da)	Error (ppm)	MT (min)	RMT (min)	HMDB ID	Clase bioquímica
	Valyl-Cysteine	C ₈ H ₁₆ N ₂ O ₃ S	220.0882				HMDB0029124	Peptides
	Glutaryl-Valine	C ₁₀ H ₁₈ N ₂ O ₅	246.1216	0.4	24.32	1.07	HMDB0028832	Peptides
	Aspartyl-Leucine	C ₁₀ H ₁₈ N ₂ O ₅	246.1216	0.4	24.32	1.07	HMDB0028757	Peptides
	Threoninyl-Glutamate	C ₉ H ₁₆ N ₂ O ₆	248.1008	12.1	24.44	1.07	HMDB0029060	Peptides
	FAD	C ₂₇ H ₃₃ N ₉ O ₁₅ P ₂	785.1571	3.6	24.8	1.09	HMDB01248	Nucleosides and conjugates
	Glutaryl-Methionine	C ₁₀ H ₁₈ N ₂ O ₅ S	278.0936	3.2	24.89	1.09	HMDB0028825	Peptides
	Glutaryl-phenylalanine	C ₁₄ H ₁₈ N ₂ O ₅	294.1216	1.7	25.08	1.10	HMDB0000594	Peptides
	5-hydroxy-L-tryptophan	C ₁₁ H ₁₂ N ₂ O ₃	220.0848	3.2	25.17	1.11	HMDB00472	Amino acids and conjugates
	Ophthalmic acid	C ₁₁ H ₁₉ N ₃ O ₆	289.1273	1.7	25.52	1.12	HMDB05765	organic acid and derivative
	Glutaryl-tyrosine	C ₁₄ H ₁₈ N ₂ O ₆	310.1165	1.0	25.58	1.12	HMDB0028831	Peptides
54	L-L-Homoglutathione	C ₁₁ H ₁₉ N ₃ O ₆ S	321.0995	0.6	25.77	1.13	HMDB0029395	Hybrid peptides
	L-Urobilin	C ₃₃ H ₄₆ N ₄ O ₆	594.3417	2.6	25.88	1.14	HMDB04159	Organoheterocyclic compounds
	Glutaryl-Cysteinyl-Serine	C ₁₁ H ₁₉ N ₃ O ₇ S ₃	337.0943	1	26.15	1.15	HMDB0039733	Peptides
	Hydroxymethyl-Glutathione	C ₁₁ H ₁₉ N ₃ O ₇ S ₃	337.0943				HMDB0004662	Peptides
	Malonylcarnitine (CAR 3:1;O2)	C ₁₀ H ₁₇ NO ₆	247.1056	4.5	27.22	1.20	HMDB02095	Carnitines

Error (ppm), Error de masa expresado en ppm; **MT**, (del inglés Migration time) tiempo de migración; **RMT**, tiempo de migración relativo frente al IS; **HMDB ID**, código de HMDB (<https://hmdb.ca/>);

AICAR, Ribonucleótido de aminoimidazol carboxamida

Tabla 3. Metabolitos estadísticamente significativos resultantes del estudio comparativo llevado a cabo en muestras de bilis de ratas con colangiocarcinoma (cases, T_30, n = 12) versus controles (controls, C_30, n=14) a través de CE-MS y LC-MS en modo de ionización positivo y negativo.

Nombre	Formula	MW (Da)	Error (ppm)	RT (min)	rMT MetS	CV (%)	Plataforma Analítica	Aducto	HMDB	Fuente de anotación	%cambio	p valor	p(corr)	VIP
Amino acids, peptides, and analogues														
Glycine	C2H5NO2	75.0320	8	15.41	0.68	7.8	CEMS	[M+H] ⁺	HMDB0000123	DB	44	1.73E-03	---	---
Alanine	C3H7NO2	89.0477	6	16.80	0.74	8.9	CEMS	[M+H] ⁺	HMDB0000161	DB	58	0.003	0.63	1.02
Serine	C3H7NO3	105.0426	1	18.70	0.82	6.4	CEMS	[M+H] ⁺	HMDB0000187	DB	91	1.25E-04	0.81	1.25
Creatinine	C4N3H7O	113.0589	1	13.50	0.59	9.5	CEMS	[M+H] ⁺	HMDB0000562	DB	32	0.033	---	---
Proline	C5H9NO2	115.0633	4	20.39	0.90	10.0	CEMS/LCMS(+)	[M+H] ⁺	HMDB0000162	DB	41	0.011	---	---
Pipecolic acid	C6H11NO2	129.0789	2	19.23	0.85	9.1	CEMS	[M+H] ⁺	HMDB0000716	DB	42	0.050	---	---
Creatine	C4H9N3O2	131.0695	1	16.52	0.73	7.9	CEMS/LCMS(+)	[M+H] ⁺	HMDB0000064	MS/MS	56	7.72E-03	0.58	1.00
Ornithine	C5H12N2O2	132.0899	0	12.72	0.56	6.0	CEMS	[M+H] ⁺	HMDB0000214	DB	Cases	4.71E-05	0.90	1.78
Glutamine	C5H10N2O3	146.0691	0	20.31	0.89	6.1	CEMS	[M+H] ⁺	HMDB0000641	DB	48	7.60E-04	---	---
Lysine	C6H14N2O2	146.1055	0	12.80	0.56	7.1	CEMS	[M+H] ⁺	HMDB0000182	DB	89	3.11E-03	0.65	1.24
Histidine	C6H9N3O2	155.0695	0	0.25	---	11.9	LCMS (+)	[M+H] ⁺	HMDB0000177	DB	-68	0.033	-0.65	1.05
Arginine	C6H14N4O2	174.1117	0	13.27	0.58	6.8	CEMS	[M+H] ⁺	HMDB0000517	DB	97	3.96E-04	0.72	1.24
Tyrosine	C9H11NO3	181.0739	1	21.56	0.95	5.9	CEMS/LCMS(+)	[M+H] ⁺	HMDB0000158	DB	59	4.94E-04	0.71	1.05
Citrulline	C6H13N3O3	175.0957	1	20.96	0.92	6.1	CEMS	[M+H] ⁺	HMDB0000904	DB	108	4.94E-04	0.73	1.30
Cystine	C6H12N2O4S2	240.0238	2	20.94	0.92	7.0	CEMS	[M+H] ⁺	HMDB0000192	DB	192	4.71E-05	0.82	1.53
Valinol	C5H13NO	103.0997	4	0.21	---	2.1	LCMS (+)	[M+H] ⁺	---	DB	92	2.21E-03	---	---
1-Aminocyclopropane-1-carboxylate	C4H7NO2	101.0477	9	20.68	0.91	4.2	CEMS	[M+H] ⁺	HMDB0036458	DB	20	0.050	---	---
N,N-Dimethylglycine	C4H9NO2	103.0633	8	20.69	0.91	8.3	CEMS	[M+H] ⁺	HMDB0000092	DB	Cases	4.71E-05	0.92	1.72
2--Aminobutyric acid	C4H9NO2	103.0633	8	17.98	0.79	11.6	CEMS	[M+H] ⁺	HMDB0031654	DB	Cases	4.71E-05	0.87	1.67
Hydroxy glutamate	C5H9NO5	163.0480	4	20.19	0.89	12.0	CEMS	[M+H] ⁺	HMDB00002273	DB	-41	0.015	-0.52	1.10
3-Methylhistidine	C7H11N3O2	169.0851	1	13.99	0.61	5.6	CEMS	[M+H] ⁺	HMDB0000479	DB	54	7.72E-03	---	---
Aminooctanoic acid	C8H17NO2	159.1259	4	15.59	0.69	9.8	CEMS	[M+H] ⁺	HMDB0000991	DB	123	3.17E-04	0.77	1.35
Cysteinyl-glycine	C5H10N2O3S	178.0412	1	15.20	0.67	13.5	CEMS	[M+H] ⁺	HMDB0028775	DB	57	0.050	---	---
Cysteinyl-Glycine disulfide	C8H15N3O5S2	297.0453	2	17.95	0.79	6.0	CEMS	[M+H] ⁺	HMDB0000709	DB	144	4.71E-05	0.79	1.41
Cysteinyl-Alanine	C6H12N2O3S	192.0569	0	17.98	0.79	11.9	CEMS	[M+H] ⁺	HMDB0028768	DB	-39	3.75E-03	---	---
Cysteinyl-Serine	C6H12N2O4S	208.0518	8	18.19	0.80	27.6	CEMS	[M+H] ⁺	HMDB0028784	DB	-46	7.72E-03	-0.53	1.07
Glycyl-Glutamate	C7H12N2O5	204.0746	0	23.03	1.01	5.8	CEMS	[M+H] ⁺	HMDB0028840	DB	189	4.71E-05	0.89	1.65
Valyl-Proline	C10H18N2O3	214.1317	2	0.38	---	2.2	LCMS (+)	[M+H] ⁺	HMDB0029135	DB	-33	0.033	---	---
Aspartyl-Threonine	C8H14N2O6	234.0852	3	24.14	1.06	7.9	CEMS	[M+H] ⁺	HMDB0028763	DB	Cases	4.71E-05	0.91	1.77
Glutamyl-Alanine / Hydroxypropyl-Serine	C8H14N2O5	218.0902	0	23.54	1.03	6.1	CEMS	[M+H] ⁺	HMDB0003764/ HMDB0028872	DB	Cases	4.71E-05	0.90	1.67

Nombre	Formula	MW (Da)	Error (ppm)	RT (min)	rMT MetS	CV (%)	Plataforma Analítica	Aducto	HMDB	Fuente de anotación	%cambio	p valor	p(corr)	VIP
Glutamyl-Valine / Aspartyl-Leucine	C10H18N2O5	246.1216	0	24.32	1.07	8.2	CEMS	[M+H] ⁺	HMDB0028832/ HMDB0028757	DB	43	0.028	---	---
Glutamyl-Lysine	C11H21N3O5	275.1481	0	16.08	0.71	8.3	CEMS	[M+H] ⁺	HMDB0003869	DB	Cases	4.71E-05	0.93	2.06
Glutamyl-Methionine	C10H18N2O5S	278.0936	3	24.89	1.09	0.0	CEMS/LCMS(+)	[M+H] ⁺	HMDB0028825	DB	70	1.73E-03	0.74	1.12
Histidinyl-Glutamate	C11H16N4O5	284.1120	0	16.35	0.72	12.8	CEMS	[M+H] ⁺	HMDB002884	DB	115	7.74E-05	0.85	1.34
Glutamylphenylalanine	C14H18N2O5	294.1216	2	25.08	1.10	6.2	CEMS	[M+H] ⁺	HMDB0000594	DB	82	7.60E-04	0.73	1.29
Glutamyl-Arginine	C11H21N5O5	303.1543	0	16.39	0.72	6.2	CEMS	[M+H] ⁺	HMDB0028813	DB	Cases	4.71E-05	0.95	2.15
Glutamyl-Tyrosine	C14H18N2O6	310.1165	1	25.58	1.12	4.4	CEMS	[M+H] ⁺	HMDB0011741	DB	Cases	3.96E-04	0.76	1.64
Glutathione	C10H17N3O6S	307.0838	2	22.29	0.98	3.4	CEMS	[M+H] ⁺	HMDB0000125	DB	86	2.53E-04	0.74	1.16
Glutathione disulfide	C20H32N6O12S2	612.1520	1	23.44	1.03	1.4	CEMS	[M+2H] ⁺	HMDB0062697	DB	69	1.73E-03	0.69	1.15
Ophthalmic acid	C11H19N3O6	289.1274	2	25.52	1.12	6.6	CEMS	[M+H] ⁺	HMDB0005765	DB	167	0.012	0.67	1.69
Cysteineglutathione disulfide	C13H22N4O8S2	426.0879	2	22.29	0.98	4.7	CEMS	[M+H] ⁺	HMDB0000656	DB	84	6.05E-05	0.79	1.20
Glutamyl-Alanyl-Cysteine-Serine / Aspartyl-Glycyl-Methionine	C11H19N3O6S	321.0995	1	0.27	---	1.9	LCMS (+)	[M+H] ⁺	---	DB	-60	1.17E-05	---	---
Glutamyl-Cysteinyl-Serine / Hydroxymethyl-Glutathione	C11H19N3O7S3	337.0943	1	26.15	1.15	18.5	CEMS	[M+H] ⁺	HMDB0039733/ HMDB0004662	DB	-53	6.47E-03	-0.54	1.23
Pyridinoline	C18H28N4O8	428.1907	7	2.57	---	3.7	LCMS (+)	[M+H] ⁺	HMDB0000851	DB	-40	3.43E-03	---	---
Deoxy pyridinoline*	C18H28N4O7	412.1958	11	2.83	---	6.8	LCMS (+)	[M+H] ⁺	HMDB0000569	DB	-81	3.28E-05	-0.90	1.37
Benzene and substituted derivatives														
Phenylacetic acid	C8H8O2	136.0524	1	2.28	---	6.6	LCMS (+)	[M+H] ⁺	HMDB0000209	DB	Controls	NA	---	---
Homovanillic acid/ Hydroxyphenyllactic acid	C9H10O4	182.0579	2	0.27	---	3.4	LCMS (+)	[M+H-H2O] ⁺	HMDB0000118/ HMDB0000755	DB	46	2.35E-04	---	---
Fatty Acyls														
FA 11:3;O	C11H16O3	196.1099	2	2.13	---	9.3	LCMS (+)	[M+H] ⁺	---	DB	Controls	NA	---	---
FA 13:2	C13H22O2	210.1620	0	2.78	---	7.5	LCMS (+)	[M+H-H2O] ⁺	---	DB	-88	1.16E-03	-0.89	1.47
FA 12:2;O	C12H20O3	212.1412	3	2.79	---	9.8	LCMS (+)	[M+H-H2O] ⁺	---	DB	-48	4.31E-04	---	---
CAR 2:0	C9H17NO4	203.1158	0	16.86	0.74	9.1	CEMS	[M+H] ⁺	HMDB0000201	DB	74	7.60E-04	0.72	1.13
CAR 3:0	C10H19NO4	217.1314	7	17.42	0.77	8.4	CEMS	[M+H] ⁺	HMDB0000824	DB	106	2.57E-03	0.60	1.14
CAR 3:1;O2	C10H17NO6	247.1056	5	27.22	1.20	11.4	CEMS	[M+H] ⁺	HMDB0002095	DB	-78	1.98E-04	-0.77	1.93
CAR 14:1	C21H39NO4	369.2879	3	4.48	---	5.4	LCMS (+)	[M+H] ⁺	HMDB0002014	DB	88	7.78E-03	---	---
CAR 14:2;O	C21H37NO5	383.2671	1	3.85	---	7.9	LCMS (+)	[M+H] ⁺	HMDB0013332	DB	141	5.77E-04	---	---
CAR 16:1;O*	C23H43NO5	413.3141	1	4.68	---	4.7	LCMS (+)	[M+H] ⁺	HMDB0013333	DB	76	0.025	---	---
CAR 20:4;O	C27H45NO5	463.3297	0	4.89	---	10.1	LCMS (+)	[M+H] ⁺	---	DB	-44	5.96E-03	---	---
N-stearoyl glutamic acid*	C23H43NO5	413.3141	1	4.68	---	4.7	LCMS (+)	[M+H] ⁺	---	DB	76	0.025	---	---

Nombre	Formula	MW (Da)	Error (ppm)	RT (min)	rMT MetS	CV (%)	Plataforma Analítica	Aducto	HMDB	Fuente de anotación	%cambio	p valor	p(corr)	VIP
3-Indole carboxylic acid glucuronide	C15H15NO8	337.0798	0	1.33	---	3.2	LCMS (+)	[M+H] ⁺	HMDB0013189	DB	-68	2.35E-04	-0.89	1.14
N-palmitoyl-threonine	C20H39NO4	357.2879	1	4.45	---	5.3	LCMS (+)	[M+H] ⁺	---	DB	84	0.001		
Glycerophospholipids														
LPS 16:1	C22H42NO9P	495.2597	12	2.78	---	2.9	LCMS (+)	[M+H] ⁺	---	DB	54	0.025	---	---
LPE 18:2	C23H44NO7P	477.2855	1	5.25	---	22.6	LCMS (+)/LCMS (-)	[M-H] ⁻	HMDB0011477	MS/MS	81	4.31E-03	---	---
LPE 20:4	C25H44NO7P	501.2855	4	5.34	---	2.9	LCMS (+)	[M+H] ⁺	HMDB0011517	MS/MS	50	5.06E-03	---	---
LPE 20:0;OH / LPC 17:1;OOH / LPE 20:1;OOH	C25H50NO9P	539.3223	9	4.90	---	4.0	LCMS (+)	[M+H] ⁺	---	DB	-77	5.09E-05	-0.94	1.30
LPC 17:0	C25H52NO7P	509.3481	0	5.90	---	2.9	LCMS (+)	[M+H] ⁺	HMDB0012108	MS/MS	-55	1.09E-03	---	---
LPC 18:0	C26H54NO7P	523.3638	1	6.16	---	2.5	LCMS (+)	[M+H] ⁺	HMDB0010384	MS/MS	-47	3.10E-03	---	---
LPC 20:5	C28H48NO7P	541.3168	2	5.08	---	2.2	LCMS (+)	[M+H] ⁺	HMDB0010397	MS/MS	-33	9.39E-03	---	---
LPC 20:1	C28H56NO7P	549.3794	1	6.34	---	18.0	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	HMDB0010391	MS/MS	-45	4.31E-03	-0.85	1.36
PC 16:0/9:0;COOH / PS 27:0	C33H64NO10P	665.4267	13	6.31	---	1.4	LCMS (+)	[M+H-H2O] ⁺	---	DB	-78	5.40E-05	-0.78	1.32
PC 16:0/11:0;COH	C35H68NO9P	677.4632	12	6.35	---	1.2	LCMS (+)	[M+H] ⁺	---	DB	-80	6.71E-05	-0.84	1.32
PC 16:0;18:2;OH	C42H80NO9P	773.5571	4	8.65	---	23.9	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	---	MS/MS	-42	0.013	-0.68	1.28
PC 16:0;20:4;OH	C44H80NO9P	797.5571	0	8.60	---	24.8	LCMS (-)	[M+FA-H] ⁻	---	MS/MS	-64	5.54E-04	-0.73	1.66
PC 14:0;16:0	C38H76NO8P	705.5309	0	9.77	---	4.0	LCMS (+)	[M+H] ⁺	HMDB0007869	MS/MS	-35	6.93E-03	---	---
PC 14:0;18:3	C40H74NO8P	727.5152	1	8.65	---	20.6	LCMS (-)	[M+FA-H] ⁻	HMDB0007876	MS/MS	-22	0.033	-0.66	1.01
PC 15:1;18:2	C41H76NO8P	741.5309	1	9.05	---	16.9	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	---	MS/MS	-31	2.21E-03	-0.75	1.13
PC 14:0;20:4	C42H76NO8P	753.5309	1	8.96	---	23.4	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	HMDB0008459	MS/MS	-51	2.21E-03	-0.84	1.48
PC 16:2;18:2	C42H76NO8P	753.5309	1	8.66	---	25.0	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	---	MS/MS	-28	6.93E-03	-0.71	1.10
PC 15:0;20:4	C43H78NO8P	767.5465	0	9.41	---	2.5	LCMS (+)	[M+H] ⁺	HMDB0007949	MS/MS	-45	1.09E-03	---	---
PC 17:1;18:2	C43H80NO8P	769.5622	1	9.74	---	21.3	LCMS (-)	[M+FA-H] ⁻	---	MS/MS	-22	5.93E-03	---	---
PC 17:0;18:2	C43H82NO8P	771.5778	2	10.91	---	32.5	LCMS (-)	[M+FA-H] ⁻	---	MS/MS	-28	0.019	-0.62	1.16
PC36:6	C44H76NO8P	777.5309	2	9.35	---	4.9	LCMS (+)	[M+H+HCOONa] ⁺	---	MS/MS	81	5.40E-05	---	---
PC 16:0;20:5	C44H78NO8P	779.5465	1	9.30	---	17.9	LCMS (+)	[M+H] ⁺	HMDB0008624	MS/MS	-53	6.71E-05	---	---
PC 18:2;18:3	C44H78NO8P	779.5465	0	9.08	---	22.8	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	HMDB0008140	MS/MS	-45	8.65E-04	-0.83	1.35
PC 16:1;20:4	C44H78NO8P	779.5465	0	9.24	---	18.0	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	HMDB0008463	MS/MS	-41	1.31E-03	-0.88	1.27
PC 16:0;20:4	C44H80NO8P	781.5622	1	9.82	---	15.7	LCMS (-)	[M+FA-H] ⁻	HMDB0008429	MS/MS	-24	3.66E-03	---	---
PC 18:2;18:2	C44H80NO8P	781.5622	1	9.40	---	21.7	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	HMDB0008138	MS/MS	-24	4.31E-03	-0.66	1.02
PC 18:2;18:1	C44H82NO8P	783.5778	1	10.15	---	5.2	LCMS (+)	[M+H] ⁺	---	MS/MS	-39	5.93E-03	---	---
PC 17:0;20:4	C45H82NO8P	795.5778	1	10.45	---	25.4	LCMS (-)	[M+FA-H] ⁻	---	MS/MS	-27	8.08E-03	-0.66	1.03
PC 20:4;18:3	C46H78NO8P	803.5465	0	8.73	---	23.9	LCMS (-)	[M+FA-H] ⁻	HMDB0008213	MS/MS	-49	3.10E-03	-0.84	1.45
PC 18:2;20:4	C46H80NO8P	805.5622	0	9.26	---	22.4	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	HMDB0008147	MS/MS	-32	3.66E-03	-0.76	1.10

Nombre	Formula	MW (Da)	Error (ppm)	RT (min)	rMT MetS	CV (%)	Plataforma Analítica	Aducto	HMDB	Fuente de anotación	%cambio	p valor	p(corr)	VIP
PC 20:4;20:4	C48H80NO8P	829.5622	0	9.09	---	24.5	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	HMDB0008443	MS/MS	-43	3.10E-03	-0.82	1.33
PC 42:9	C50H82NO8P	855.5778	1	11.94	---	7.0	LCMS (+)	[M+H] ⁺	HMDB0008515	DB	50	3.93E-03	---	---
PE 38:6;OH / PE 38:5;O / PE P-38:6;OOH	C43H74NO9P	779.5101	8	6.32	---	2.7	LCMS (+)	[M+H] ⁺	---	DB	-81	1.03E-04	-0.76	1.31
PE P-36:5	C41H72NO7P	721.5046	10	6.35	---	1.3	LCMS (+)	[M+H] ⁺	HMDB0011354	DB	-82	4.33E-05	-0.83	1.39
Imidazoles														
Imidazolepropionic acid / Methylimidazoleacetic acid*	C6H8N2O2	140.0586	1	0.22	---	11.3	LCMS (+)	[M+H-H2O] ⁺	HMDB0002271/ HMDB0002820	DB	-42	1.86E-03	---	---
Imidazolelactic acid	C6H8N2O3	156.0534	14	15.45	0.68	6.5	CEMS	[M+H-H2O] ⁺	HMDB0002320	DB	78	2.01E-04	0.69	1.13
Imidazopyrimidines / Purines and purine derivatives														
Tetrahydrofolic acid	C19H23N7O6	445.1710	7	2.79	---	5.8	LCMS (+)	[M+H] ⁺	HMDB0001846	DB	-97	1.74E-05	-0.96	2.02
Nucleosides, nucleotides, and analogues														
Cytidine	C9H13N3O5	243.0855	10	18.29	0.80	14.8	CEMS	[M+H] ⁺	HMDB0000089	DB	50	6.47E-03	---	---
Organic acids and derivatives														
Succinylacetoacetate / Ethyl aconitate	C8H10O6	202.0477	12	0.22	---	11.6	LCMS (+)	[M+H] ⁺	HMDB0240258/ HMDB0040275	DB	-68	3.44E-05	-0.87	1.10
Organonitrogen compounds														
Trimethylamine N-oxide	C3H9NO	75.0684	4	12.05	0.53	8.2	CEMS	[M+H] ⁺	HMDB0000925	DB	-22	0.024	---	---
Acetylcholine	C7H15NO2	146.2074	0	14.99	0.66	26.9	CEMS	[M+H] ⁺	HMDB0000895	DB	54	6.47E-03	0.57	1.00
Carnitine	C7H15NO3	161.1051	2	15.85	0.70	8.5	CEMS	[M+H] ⁺	HMDB0000062	DB	28	0.038	---	---
Sphingosine	C18H37NO2	299.2824	0	4.77	---	2.2	LCMS (+)	[M+H] ⁺	HMDB0000252	DB	-81	1.61E-04	-0.92	1.37
Riboflavin	C17H20N4O6	376.1382	0	1.94	---	3.7	LCMS (+)	[M+H] ⁺	HMDB0000244	DB	-66	4.40E-04	-0.78	1.10
Organoxygen compounds														
Pantothenate	C9H17NO5	219.1107	5	0.46	---	3.1	LCMS (+)	[M+H] ⁺	HMDB0000210	DB	103	4.26E-04	---	---
5-Hydroxy-6-methoxyindole glucuronide	C15H17NO8	339.0954	0	1.76	---	2.7	LCMS (+)	[M+H] ⁺	HMDB0010363	DB	-64	2.35E-04	-0.85	1.13
Pyridines and derivatives														
Nicotinamide	C6H6N2O	122.0480	0	13.85	0.61	11.0	CEMS/LCMS(+)	[M+H] ⁺	HMDB0001406	DB	-40	1.16E-03	-0.69	1.10
Quinolines and derivates														
Kynurenic acid	C10H7NO3	189.0426	0	0.75	---	5.0	LCMS (+)	[M+H] ⁺	HMDB0000715	DB	Cases	3.12E-06	0.91	1.17
Sphingolipids														
C19 Sphingosine-1-phosphate	C19H40NO5P	393.2644	1	2.38	---	6.8	LCMS (+)	[M+H] ⁺	---	DB	-81	4.33E-05	-0.77	1.29
Cer 34:1;O2	C34H67NO3	537.5121	1	10.94	---	8.1	LCMS (+)	[M+H] ⁺	HMDB0004949	DB	-88	2.77E-05	-0.92	1.56
SM 34:2;O2	C39H77N2O6P	700.5519	1	8.32	---	28.8	LCMS (-)	[M+FA-H] ⁻	---	MS/MS	-34	0.017	-0.69	1.20
SM 18:1;16:0;O2	C39H79N2O6P	702.5676	1	9.05	---	21.8	LCMS (+)/LCMS (-)	[M+FA-H] ⁻	HMDB0061712	MS/MS	-36	3.10E-03	-0.69	1.19

Nombre	Formula	MW (Da)	Error (ppm)	RT (min)	rMT MetS	CV (%)	Plataforma Analítica	Aducto	HMDB	Fuente de anotación	%cambio	p valor	p(corr)	VIP
Steroids and steroid derivatives														
ST 18:4;O2	C18H22O2	270.1619	2	4.01	---	2.3	LCMS (+)	[M+H] ⁺	HMDB0000145	DB	Cases	NA	---	---
ST 21:1;O2	C21H34O2	318.2559	1	4.33	---	5.1	LCMS (+)	[M+H] ⁺	HMDB0062782	DB	-82	1.75E-05	-0.95	1.41
ST 21:2;O2	C21H32O2	316.2402	0	3.66	---	2.8	LCMS (+)	[M+H] ⁺	HMDB0000253	DB	-66	1.03E-04	-0.81	1.07
ST 21:1;O3	C21H34O3	334.2508	1	3.86	---	9.5	LCMS (+)	[M+H-H2O] ⁺	HMDB0000879	DB	-44	1.56E-03	---	---
ST 21:1;O4	C21H34O4	350.2457	1	3.24	---	2.5	LCMS (+)	[M+H-H2O] ⁺	HMDB0000449	DB	-58	4.31E-03	---	---
ST 21:2;O3	C21H32O3	332.2351	0	3.79	---	2.2	LCMS (+)	[M+H-H2O] ⁺	HMDB0001830	DB	-75	2.77E-05	-0.94	1.28
ST 21:2;O3;S*	C21H32O6S	412.1919	2	2.83	---	6.8	LCMS (+)	[M+H] ⁺	HMDB0000416	DB	-81	3.28E-05	-0.90	1.37
ST 21:3;O3;GlcA	C27H38O9	506.2516	1	3.66	---	5.3	LCMS (+)	[M+H] ⁺	HMDB0010364	DB	-80	1.75E-05	-0.84	1.34
ST 24:1;O4	C24H40O4	392.2927	6	3.72	---	5.2	LCMS (+)	[M+H-H2O] ⁺	HMDB0000518	DB	Cases	4.33E-05	0.79	1.37
ST 24:2;O4	C24H38O4	390.2770	1	3.71	---	2.3	LCMS (+)	[M+H-H2O] ⁺	HMDB0000328	MS/MS	Cases	1.75E-05	0.85	1.51
ST 24:2;O5	C24H38O5	406.2719	1	3.36	---	6.1	LCMS (+)	[M+H-H2O] ⁺	HMDB0000391	DB	Cases	1.75E-05	0.84	1.77
ST 24:3;O2	C24H36O2	356.2715	2	4.45	---	5.9	LCMS (+)	[M+H] ⁺	---	DB	105	2.98E-03	---	---
ST 24:3;O4	C24H36O4	388.2614	1	4.02	---	2.8	LCMS (+)	[M+H] ⁺	HMDB0062744	MS/MS	Cases	1.75E-05	0.88	1.52
ST 24:3;O5	C24H36O5	404.2563	0	4.21	---	2.9	LCMS (+)	[M+H] ⁺	HMDB0000447	DB	121	0.011	---	---
ST 24:4;O3	C24H34O3	370.2507	1	4.02	---	4.1	LCMS (+)	[M+H] ⁺	HMDB0000476	MS/MS	Cases	1.75E-05	0.86	1.83
ST 24:4;O4	C24H34O4	386.2457	0	3.59	---	3.9	LCMS (+)	[M+H] ⁺	HMDB0000385	DB	71	1.92E-03	---	---
ST 24:5;O3	C24H32O3	368.2351	1	3.72	---	16.4	LCMS (+)	[M+H-H2O] ⁺	---	DB	-68	1.23E-04	-0.84	1.15
ST 24:1;O3;T	C26H45NO5S	483.3018	0	4.87	---	1.3	LCMS (+)	[M+H-H2O] ⁺	---	DB	-55	0.025	---	---
ST 24:1;O4;T	C26H45NO6S	499.2967	2	3.75	---	5.0	LCMS (+)/LCMS (-)	[M+H] ⁺	HMDB0000951	MS/MS	-87	1.75E-05	-0.95	1.54
ST 24:1;O4;G	C26H43NO5	449.3141	3	4.54	---	1.9	LCMS (+)	[M+H] ⁺	HMDB0006898	MS/MS	-78	0.011	-0.65	1.07
ST 24:1;O4;GlcA	C30H48O10	568.3248	2	4.38	---	60.1	LCMS (+)/LCMS (-)	[M-H] ⁻	HMDB0002596	MS/MS	Cases	NA	0.78	2.12
ST 24:1;O5;GlcA	C30H48O11	584.3197	7	3.59	---	1.8	LCMS (+)	[M+H] ⁺	HMDB0002577	DB	Cases	3.91E-09	0.88	1.35
ST 24:1;O5;G	C26H43NO6	465.3090	2	3.51	---	5.1	LCMS (+)/LCMS (-)	[M+H-H2O] ⁺	HMDB0000138	MS/MS	-85	5.40E-05	-0.81	1.31
ST 24:1;O5;T	C26H45NO7S	515.2916	1	3.01	---	1.9	LCMS (+)/LCMS (-)	[M+H-H2O] ⁺	HMDB0000036	MS/MS	-45	3.62E-03	---	---
ST 27:1;O4*	C27H46O4	434.3396	1	5.22	---	2.0	LCMS (+)	[M+H-H2O] ⁺	---	DB	Cases	3.66E-03	---	---
ST 27:2;O3*	C27H44O3	416.3290	1	5.22	---	2.0	LCMS (+)	[M+H] ⁺	---	DB	Cases	3.66E-03	---	---
ST 27:2;O4	C27H44O4	432.3240	0	5.23	---	12.2	LCMS (+)	[M+H] ⁺	HMDB0062712	DB	136	8.08E-03	---	---
ST 27:2;O5*	C27H44O5	448.3188	2	4.45	---	5.9	LCMS (+)	[M+H-H2O] ⁺	---	DB	Cases	2.10E-06	0.76	1.32
ST 27:3;O3	C27H42O3	414.3134	2	4.28	---	13.4	LCMS (+)	[M+H] ⁺	---	DB	Cases	4.00E-03	---	---
ST 27:3;O4	C27H42O4	430.3083	0	4.67	---	11.7	LCMS (+)	[M+H] ⁺	HMDB0012458	DB	Cases	8.47E-07	0.84	1.18
ST 27:3;O5	C27H42O5	446.3032	2	3.74	---	13.8	LCMS (+)	[M+H-H2O] ⁺	---	DB	Cases	4.07E-05	0.91	1.40
ST 28:2;O4	C28H46O4	446.3396	1	7.36	---	2.2	LCMS (+)	[M+H] ⁺	HMDB0041527	DB	-42	9.39E-03	---	---
ST 29:2;O2	C29H48O2	428.3654	2	6.49	---	3.6	LCMS (+)	[M+H] ⁺	---	DB	-53	4.26E-04	---	---
ST 29:2;O5	C29H48O5	476.3500	0	5.07	---	11.6	LCMS (+)	[M+H] ⁺	---	DB	-59	1.31E-03	-0.77	1.04
ST 29:4;O4	C29H44O4	456.3239	1	5.32	---	6.3	LCMS (+)	[M+H] ⁺	---	DB	-86	4.33E-05	-0.84	1.53
Sulfoglycolithocholate	C26H43NO7S	513.2760	1	3.22	---	2.0	LCMS (+)	[M+H] ⁺	HMDB0002639	DB	Cases	4.44E-07	0.92	1.19

Nombre	Formula	MW (Da)	Error (ppm)	RT (min)	rMT MetS	CV (%)	Plataforma Analítica	Aducto	HMDB	Fuente de anotación	%cambio	p valor	p(corr)	VIP
Glycochenodeoxycholate-3-sulfate	C26H43NO8S	529.2709	3	2.63	---	3.2	LCMS (+)	[M+H] ⁺	HMDB0002497	DB	Cases	2.77E-05	0.80	1.12
Tetrapyrroles and derivatives														
Biliverdin	C33H34N4O6	582.2478	1	4.47	---	6.7	LCMS (+)	[M+H] ⁺	HMDB0001008	DB	-49	4.31E-03	---	---
Urobilin	C33H46N4O6	594.3417	3	25.88	1.14	0.0	CEMS/LCMS(+)	[M+H] ⁺	HMDB0004159	DB	-56	2.57E-03	-0.67	1.28

MW, Masa molecular (del inglés Molecular Weight); **Error** (ppm), Error de masa expresado en ppm; **RT**, (del inglés Retention time) tiempo de retención o migración según corresponda; **rMT_{MetS}**, tiempo de migración relativo frente al IS para los datos de CE-MS; **CV**, coeficiente de variación de los metabolitos en los QCs expresados en %; **Plataforma analítica**, plataforma analítica de la que proviene la anotación; **Aducto**, aducto mayoritario; **HMDB ID**, código de HMDB (<https://hmdb.ca/>); **Fuente de anotación**, fuente de identificación, donde DB (del inglés Data base) corresponde a la base de datos CMM y, MS/MS, a la identificación confirmada con fragmentación; **% cambio**, variación de la abundancia en el estudio comparativo expresada como % (ver más detalles en el apartado “Análisis estadístico”); **cases**, compuesto presente solo en el grupo de casos; **controls**, compuesto no presente en el grupo casos; **p valor**, p valor global obtenido de la prueba t de Student; **p(corr)** y **VIP**, selección de variables usando una combinación del coeficiente de correlación ($|p(corr)| \geq 0.5$) y la importancia de la variable en la proyección ($VIP \geq 1$) (Figura 8).

Tabla 4. Cuantificación de metabolitos relacionados con el estado redox celular en muestras de bilis. El tiempo de migración relativo (rMT) se calculó en relación con el MT del IS. [control]_{16semanas}, concentración de cada compuesto en 12 controles de dieciséis semanas; [control]_{30semanas}, concentración en 12 controles de treinta semanas; [case]_{30semanas}, concentración en 14 casos de treinta semanas; SD: desviación estándar; p value, p valor obtenido de la prueba t de Student del estudio comparativo casos/controles de 30 semanas; %cambio, porcentaje de cambio obtenido de $([casos]_{30weeks} - [control]_{30weeks}) / [control]_{30weeks}$.

Metabolito	Formula	Masa Monoisotópica (Da)	rMT	Curva de calibración	r ²	[control] _{16semanas} mg/l	SD	[control] _{30semanas} mg/l	SD	[case] _{30semanas} mg/l	SD	p valor	%cambio
Glycine	C2H5NO2	75.0326	0.68	y = 24599x - 22557	0.996	47.8	2.1	44.5	2.1	55.2	5.4	2.51E-06	24
L-Serine	C3H7NO3	105.0427	0.82	y = 25377x - 2235	0.997	31.6	2.3	39.3	1.0	65.6	4.9	8.36E-12	67
L-Ornithine	C5H12N2O2	132.0899	0.56	y = 25001x - 255	0.992	4.3	1.1	7.1	1.5	22.3	2.5	9.54E-15	213
L-Lysine	C6H14N2O2	146.1055	0.56	y = 36551x - 8041	0.997	1.202	0.083	1.086	0.085	1.45	0.14	8.86E-08	33
L-Arginine	C6H14N4O2	174.1117	0.58	y = 44603x - 19052	0.995	15.0	1.3	18.5	3.7	30.7	4.2	1.81E-07	66
DL-Citrulline	C6H13N3O3	175.0956	0.92	y = 31312x + 1834	0.998	4.98	0.90	6.5	1.5	12.1	2.0	9.02E-08	87
L-Cystine	C6H12N2O4S2	240.0244	0.92	y = 25351x + 416	0.997	0.768	0.091	1.00	0.15	2.6	1.1	9.37E-05	159
Ophthalmic acid	C11H19N3O6	289.1268	1.12	y = 22239x + 255	0.999	0.57	0.23	1.6	2.1	3.8	2.6	3.74E-02	132
Glutathione disulfide	C20H32N6O12S2	612.152	0.93	y = 12524x - 1595	0.993	1.74	0.35	1.86	0.54	2.71	0.54	1.08E-03	46

4. Referencias

1. Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. *Nature Reviews Gastroenterology & Hepatology*. 2020;17(9):557-88.
2. Banales JM, Cardinale V, Carpino G, Marzioni M, Andersen JB, Invernizzi P, et al. Expert consensus document: Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). *Nature reviews Gastroenterology & hepatology*. 2016;13(5):261-80.
3. Strijker M, Belkouz A, van der Geest LG, van Gulik TM, van Hooft JE, de Meijer VE, et al. Treatment and survival of resected and unresected distal cholangiocarcinoma: a nationwide study. *Acta Oncologica*. 2019;58(7):1048-55.
4. Loosen SH, Vucur M, Trautwein C, Roderburg C, Luedde T. Circulating Biomarkers for Cholangiocarcinoma. *Digestive Diseases*. 2018:281-8.
5. Dos Santos A, Court M, Thiers V, Sar S, Guettier C, Samuel D, et al. Identification of cellular targets in human intrahepatic cholangiocarcinoma using laser microdissection and accurate mass and time tag proteomics. *Mol Cell Proteomics*. 2010;9(9):1991-2004.
6. Nakagohri T, Kinoshita T, Konishi M, Takahashi S, Gotohda N. Surgical Outcome and Prognostic Factors in Intrahepatic Cholangiocarcinoma. *World Journal of Surgery*. 2008;32(12):2675-80.
7. Rosen CB, Heimbach JK, Gores GJ. Liver transplantation for cholangiocarcinoma. *Transplant International*. 2010;23(7):692-7.
8. Liang Q, Liu H, Zhang TY, Jiang Y, Xing HT, Zhang H. Serum metabolomics uncovering specific metabolite signatures of intra- and extrahepatic cholangiocarcinoma. *Molecular Biosystems*. 2016;12(2):334-40.
9. Banales JM, Iñarrairaegui M, Arbelaiz A, Milkiewicz P, Muntané J, Muñoz-Bellvis L, et al. Serum metabolites as diagnostic biomarkers for cholangiocarcinoma, hepatocellular carcinoma, and primary sclerosing cholangitis. *Hepatology*. 2019;70(2):547-62.
10. Gao R, Cheng J, Fan C, Shi X, Cao Y, Sun B, et al. Serum metabolomics to identify the liver disease-specific biomarkers for the progression of hepatitis to hepatocellular carcinoma. *Scientific reports*. 2015;5:18175.
11. Alsaleh M, Leftley Z, Barbera TA, Koomson LK, Zabron A, Crossey MME, et al. Characterisation of the Serum Metabolic Signature of Cholangiocarcinoma in a United Kingdom Cohort. *J Clin Exp Hepatol*. 2020;10(1):17-29.
12. Alsaleh M, Sithithaworn P, Khuntikeo N, Loilome W, Yongvanit P, Chamadol N, et al. Characterisation of the Urinary Metabolic Profile of Liver Fluke-Associated Cholangiocarcinoma. *J Clin Exp Hepatol*. 2019;9(6):657-75.
13. Winter H, Kaisaki PJ, Harvey J, Giacobuzzi E, Ferla MP, Pentony MM, et al. Identification of Circulating Genomic and Metabolic Biomarkers in Intrahepatic Cholangiocarcinoma. *Cancers*. 2019;11(12).
14. Murakami Y, Kubo S, Tamori A, Itami S, Kawamura E, Iwaisako K, et al. Comprehensive analysis of transcriptome and metabolome analysis in Intrahepatic Cholangiocarcinoma and Hepatocellular Carcinoma. *Scientific Reports*. 2015;5(1):16294.

15. Kotawong K, Chaijaroenkul W, Roytrakul S, Phaonakrop N, Na-Bangchang K. Screening of Molecular Targets of Action of Atractylodin in Cholangiocarcinoma by Applying Proteomic and Metabolomic Approaches. *Metabolites*. 2019;9(11):260.
16. Barbhuiya MA, Sahasrabuddhe NA, Pinto SM, Muthusamy B, Singh TD, Nanjappa V, et al. Comprehensive proteomic analysis of human bile. *Proteomics*. 2011;11(23):4443-53.
17. Farina A, Dumonceau JM, Lescuyer P. Proteomic analysis of human bile and potential applications for cancer diagnosis. *Expert Rev Proteomics*. 2009;6(3):285-301.
18. Gowda GA. Human bile as a rich source of biomarkers for hepatopancreatobiliary cancers. *Biomark Med*. 2010;4(2):299-314.
19. Ijare OB, Somashekar B, Jadegoud Y, Nagana Gowda G. ¹H and ¹³C NMR characterization and stereochemical assignments of bile acids in aqueous media. *Lipids*. 2005;40(10):1031.
20. Lee G, Lee H, Hong J, Lee SH, Jung BH. Quantitative profiling of bile acids in rat bile using ultrahigh-performance liquid chromatography-orbitrap mass spectrometry: Alteration of the bile acid composition with aging. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2016;1031:37-49.
21. Pearson H. Meet the human metabolome. *Nature*. 2007;446(7131):8-.
22. Hashim AbdAlla MS, Taylor-Robinson SD, Sharif AW, Williams HR, Crossey MM, Badra GA, et al. Differences in phosphatidylcholine and bile acids in bile from Egyptian and UK patients with and without cholangiocarcinoma. *HPB*. 2011;13(6):385-90.
23. Plumb RS, Rainville PD, Potts WB, Johnson KA, Gika E, Wilson ID. Application of ultra performance liquid chromatography-mass spectrometry to profiling rat and dog bile. *J Proteome Res*. 2009;8(5):2495-500.
24. Wang Y, Griffiths WJ. Modern methods of bile acid analysis by mass spectrometry: a view into the metabolome. *Current Analytical Chemistry*. 2007;3(2):103-26.
25. Papaspyridonos K, Garcia-Perez I, Angulo S, Domann PJ, Vilca-Melendez H, Heaton N, et al. Fingerprinting of human bile during liver transplantation by capillary electrophoresis. *J Sep Sci*. 2008;31(16-17):3058-64.
26. Zhang W, Ramautar R. CE-MS for metabolomics: Developments and applications in the period 2018-2020. *Electrophoresis*. 2021;42(4):381-401.
27. Ramautar R, Somsen GW, de Jong GJ. CE-MS for metabolomics: Developments and applications in the period 2016-2018. *Electrophoresis*. 2019;40(1):165-79.
28. Ramautar R. Capillary Electrophoresis-Mass Spectrometry for Clinical Metabolomics. *Adv Clin Chem*. 2016;74:1-34.
29. Drouin N, Ramautar R. Capillary Electrophoresis-Mass Spectrometry for Metabolomics: Possibilities and Perspectives. *Adv Exp Med Biol*. 2021;1336:159-78.
30. Garcia A, Godzien J, Lopez-Gonzalvez A, Barbas C. Capillary electrophoresis mass spectrometry as a tool for untargeted metabolomics. *Bioanalysis*. 2017;9(1):99-130.
31. López-González Á, Godzien J, García A, Barbas C. Capillary Electrophoresis Mass Spectrometry as a Tool for Untargeted Metabolomics. *Methods Mol Biol*. 2019;1978:55-77.

32. Rodríguez-Ortigosa CM, Celay J, Olivas I, Juanarena N, Arcelus S, Uriarte I, et al. A GAPDH-mediated trans-nitrosylation pathway is required for feedback inhibition of bile salt synthesis in rat liver. *Gastroenterology*. 2014;147(5):1084-93.
33. Naz S, Garcia A, Rusak M, Barbas C. Method development and validation for rat serum fingerprinting with CE-MS: application to ventilator-induced-lung-injury study. *Anal Bioanal Chem*. 2013;405(14):4849-58.
34. Su L, Mao J, Hao M, Lu T, Mao C, Ji D, et al. Integrated Plasma and Bile Metabolomics Based on an UHPLC-Q/TOF-MS and Network Pharmacology Approach to Explore the Potential Mechanism of Schisandra chinensis-Protection From Acute Alcoholic Liver Injury. *Frontiers in Pharmacology*. 2020;10(1543).
35. Attard JA, Dunn WB, Mergental H, Mirza DF, Afford SC, Perera MTPR. Systematic Review: Clinical Metabolomics to Forecast Outcomes in Liver Transplantation Surgery. *OMICS*. 2019;23(10):463-76.
36. Godzien J, Kalaska B, Adamska-Patruno E, Siroka J, Ciborowski M, Kretowski A, et al. Oxidized glycerophosphatidylcholines in diabetes through non-targeted metabolomics: Their annotation and biological meaning. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2019;1120:62-70.
37. Armitage EG, Godzien J, Alonso-Herranz V, López-González Á, Barbas C. Missing value imputation strategies for metabolomics data. *Electrophoresis*. 2015;36(24):3050-60.
38. Kuligowski J, Sanchez-Illana A, Sanjuan-Herraez D, Vento M, Quintas G. Intra-batch effect correction in liquid chromatography-mass spectrometry using quality control samples and support vector regression (QC-SVRC). *Analyst*. 2015;140(22):7810-7.
39. Wheelock AM, Wheelock CE. Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Molecular BioSystems*. 2013;9(11):2589-96.
40. Godzien J, Armitage EG, Angulo S, Martinez-Alcazar MP, Alonso-Herranz V, Otero A, et al. In-source fragmentation and correlation analysis as tools for metabolite identification exemplified with CE-TOF untargeted metabolomics. *Electrophoresis*. 2015;36(18):2188-95.
41. Kind T, Liu K-H, Lee DY, DeFelice B, Meissen JK, Fiehn O. LipidBlast in silico tandem mass spectrometry database for lipid identification. *Nature Methods*. 2013;10(8):755-8.
42. Hernández B, Pennington SR, Parnell AC. Bayesian methods for proteomic biomarker development. *EuPA Open Proteomics*. 2015;9:54-64.
43. Montenegro-Burke JR, Guijas C, Siuzdak G. METLIN: A Tandem Mass Spectral Library of Standards. *Methods Mol Biol*. 2020;2104:149-63.
44. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform*. 2016;8:3.
45. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research*. 2021;49(W1):W388-W96.
46. Duarte IF, Legido-Quigley C, Parker DA, Swann JR, Spraul M, Braumann U, et al. Identification of metabolites in human hepatic bile using 800 MHz ¹H NMR spectroscopy, HPLC-NMR/MS and UPLC-MS. *Mol Biosyst*. 2009;5(2):180-90.

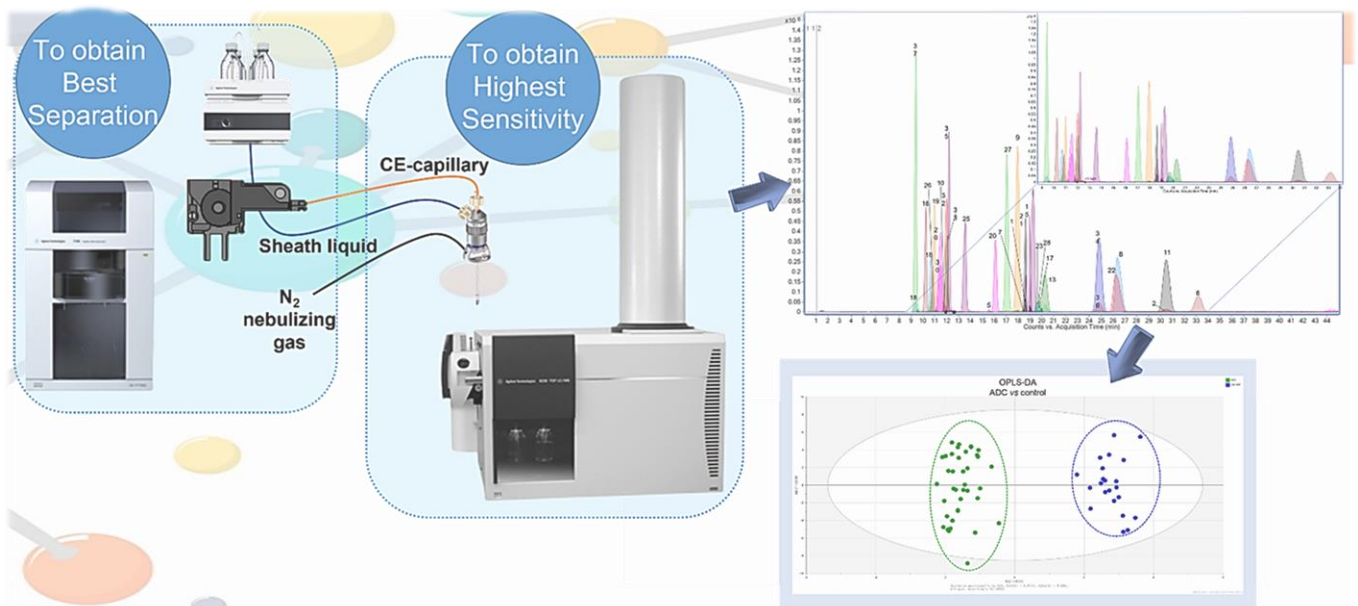
47. Petrov VA, Fernández-Peralbo MA, Derks R, Knyazeva EM, Merzlikin NV, Sazonov AE, et al. Biliary Microbiota and Bile Acid Composition in Cholelithiasis. *BioMed Research International*. 2020;2020:1242364.
48. Gowda GA, Ijare OB, Somashekar BS, Sharma A, Kapoor VK, Khetrapal CL. Single-step analysis of individual conjugated bile acids in human bile using ¹H NMR spectroscopy. *Lipids*. 2006;41(6):591-603.
49. Khan SA, Cox IJ, Thillainayagam AV, Bansil DS, Thomas HC, Taylor-Robinson SD. Proton and phosphorus-31 nuclear magnetic resonance spectroscopy of human bile in hepatopancreaticobiliary cancer. *Eur J Gastroenterol Hepatol*. 2005;17(7):733-8.
50. Nagana Gowda GA, Shanaiah N, Cooper A, Maluccio M, Raftery D. Visualization of bile homeostasis using (1)H-NMR spectroscopy as a route for assessing liver cancer. *Lipids*. 2009;44(1):27-35.
51. Mims D, Hercules D. Quantification of bile acids directly from plasma by MALDI-TOF-MS. *Anal Bioanal Chem*. 2004;378(5):1322-6.
52. Mims D, Hercules D. Quantification of bile acids directly from urine by MALDI-TOF-MS. *Anal Bioanal Chem*. 2003;375(5):609-16.
53. Naz S, Garcia A, Barbas C. Multiplatform Analytical Methodology for Metabolic Fingerprinting of Lung Tissue. *Analytical Chemistry*. 2013;85(22):10941-8.
54. Gil-De-La Fuente A, Godzien J, Saugar S, García Carmona R, Badran H, Wishart DS, et al. CEU Mass Mediator 3.0: a metabolite annotation tool. *J Proteome Res*. 2018.
55. Mori H, Morine Y, Mawatari K, Chiba A, Yamada S, Saito YU, et al. Bile Metabolites and Risk of Carcinogenesis in Patients With Pancreaticobiliary Maljunction: A Pilot Study. *Anticancer Res*. 2021;41(1):327-34.
56. Townsend DM, Tew KD, Tapiero H. The importance of glutathione in human disease. *Biomed Pharmacother*. 2003;57(3-4):145-55.
57. Villa E, Ali ES, Sahu U, Ben-Sahra I. Cancer cells tune the signaling pathways to empower de novo synthesis of nucleotides. *Cancers*. 2019;11(5):688.
58. Siddiqui A, Ceppi P. A non-proliferative role of pyrimidine metabolism in cancer. *Molecular metabolism*. 2020;35:100962-.
59. Lu YH, Li N, Gao L, Xu YJ, Huang C, Yu KK, et al. Acetylcarnitine Is a Candidate Diagnostic and Prognostic Biomarker of Hepatocellular Carcinoma. *Cancer Research*. 2016;76(10):2912-20.
60. Zeng J, Huang X, Zhou LN, Tan YX, Hu CX, Wang XM, et al. Metabolomics Identifies Biomarker Pattern for Early Diagnosis of Hepatocellular Carcinoma: from Diethylnitrosamine Treated Rats to Patients. *Scientific Reports*. 2015;5.
61. Zeng J, Yin PY, Tan YX, Dong LW, Hu CX, Huang Q, et al. Metabolomics Study of Hepatocellular Carcinoma: Discovery and Validation of Serum Potential Biomarkers by Using Capillary Electrophoresis-Mass Spectrometry. *Journal of Proteome Research*. 2014;13(7):3420-31.
62. Kobayashi T, Nishiumi S, Ikeda A, Yoshie T, Sakai A, Matsubara A, et al. A Novel Serum Metabolomics-Based Diagnostic Approach to Pancreatic Cancer. *Cancer Epidemiology Biomarkers & Prevention*. 2013;22(4):571-9.
63. Fukutake N, Ueno M, Hiraoka N, Shimada K, Shiraishi K, Saruki N, et al. A Novel Multivariate Index for Pancreatic Cancer Detection Based On the Plasma Free Amino Acid Profile. *Plos One*. 2015;10(7).

64. Hirata Y, Kobayashi T, Nishiumi S, Yamanaka K, Nakagawa T, Fujigaki S, et al. Identification of highly sensitive biomarkers that can aid the early detection of pancreatic cancer using GC/MS/MS-based targeted metabolomics. *Clinica Chimica Acta*. 2017;468:98-104.
65. Lindahl A, Heuchel R, Forshed J, Lehtio J, Lohr M, Nordstrom A. Discrimination of pancreatic cancer and pancreatitis by LC-MS metabolomics. *Metabolomics*. 2017;13(5).
66. Mayerle J, Kalthoff H, Reszka R, Kamlage B, Peter E, Schniewind B, et al. Metabolic biomarker signature to differentiate pancreatic ductal adenocarcinoma from chronic pancreatitis. *Gut*. 2017:gutjnl-2016-312432.
67. Zhang L, Jin H, Guo X, Yang Z, Zhao L, Tang S, et al. Distinguishing pancreatic cancer from chronic pancreatitis and healthy individuals by ¹H nuclear magnetic resonance-based metabolomic profiles. *Clin Biochem*. 2012;45(13-14):1064-9.
68. Alvaro D, Cantafora A, Attili AF, Ginanni Corradini S, De Luca C, Minervini G, et al. Relationships between bile salts hydrophilicity and phospholipid composition in bile of various animal species. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*. 1986;83(3):551-4.
69. Macias RIR, Banales JM, Sangro B, Muntané J, Avila MA, Lozano E, et al. The search for novel diagnostic and prognostic biomarkers in cholangiocarcinoma. *Biochim Biophys Acta Mol Basis Dis*. 2018;1864(4 Pt B):1468-77.
70. Komichi D, Tazuma S, Nishioka T, Hyogo H, Chayama K. Glycochenodeoxycholate plays a carcinogenic role in immortalized mouse cholangiocytes via oxidative DNA damage. *Free Radic Biol Med*. 2005;39(11):1418-27.
71. Khan SA, Thomas HC, Davidson BR, Taylor-Robinson SD. Cholangiocarcinoma. *The Lancet*. 2005;366(9493):1303-14.
72. Bala L, Tripathi P, Bhatt G, Das K, Roy R, Choudhuri G, et al. ¹H and ³¹P NMR studies indicate reduced bile constituents in patients with biliary obstruction and infection. *NMR in Biomedicine*. 2009;22(2):220-8.
73. Urman JM, Herranz JM, Uriarte I, Rullán M, Oyón D, González B, et al. Pilot Multi-Omic Analysis of Human Bile from Benign and Malignant Biliary Strictures: A Machine-Learning Approach. *Cancers*. 2020;12(6):1644.
74. Khan SA, Davidson BR, Goldin RD, Heaton N, Karani J, Pereira SP, et al. Guidelines for the diagnosis and treatment of cholangiocarcinoma: an update. *Gut*. 2012;61(12):1657-69.
75. Nie J, Yang J, Wei Y, Wei X. The role of oxidized phospholipids in the development of disease. *Molecular Aspects of Medicine*. 2020;76:100909.
76. Fruhwirth GO, Loidl A, Hermetter A. Oxidized phospholipids: from molecular properties to disease. *Biochim Biophys Acta*. 2007;1772(7):718-36.
77. Soga T, Baran R, Suematsu M, Ueno Y, Ikeda S, Sakurakawa T, et al. Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption. *J Biol Chem*. 2006;281(24):16768-76.
78. Kaur G, Leslie EM, Tillman H, Lee WM, Swanlund DP, Karvellas CJ, et al. Detection of Ophthalmic Acid in Serum from Acetaminophen-Induced Acute Liver Failure Patients Is More Frequent in Non-Survivors. *PLoS One*. 2015;10(9):e0139299.

79. Hitchings GH. Functions of Tetrahydrofolate and the Role of Dihydrofolate Reductase in Cellular Metabolism. In: Hitchings GH, editor. *Inhibition of Folate Metabolism in Chemotherapy: The Origins and Uses of Co-trimoxazole*. Berlin, Heidelberg: Springer Berlin Heidelberg; 1983. p. 11-23.

80. Mora MI, Molina M, Odriozola L, Elortza F, Mato JM, Sitek B, et al. Prioritizing Popular Proteins in Liver Cancer: Remodelling One-Carbon Metabolism. *J Proteome Res*. 2017;16(12):4506-14.

81. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. *Nature Reviews Cancer*. 2011;11(2):85-95.



Capítulo 4. Desarrollo de un método de análisis para compuestos aniónicos en CE-MS. Aplicación a un estudio metabolómico de cáncer de pulmón

1. Introducción

La determinación de todos los metabolitos presentes en un sistema biológico mediante un único método analítico no es posible ya que las propiedades físicas y químicas de los metabolitos son diversas, por lo que generalmente se emplean varios métodos analíticos que se complementen. Los métodos basados en espectrometría de masas (MS), como GC-MS y LC-MS han sido ampliamente empleados en metabolómica (1). Entre estos métodos, la cromatografía de interacción hidrofílica (HILIC) acoplada a la MS es el método más utilizado para la separación de metabolitos polares. Sin embargo, los azúcares fosforilados múltiples y los ácidos orgánicos a menudo no se separan bien. La cromatografía líquida de par iónico también es una técnica de separación útil para metabolitos cargados, sin embargo, los reactivos añadidos a la fase móvil para promover la formación de pares iónicos con los analitos de la muestra en ocasiones pueden causar supresión de iones en la MS y reducir la precisión de la cuantificación (2). A pesar de los importantes avances en la tecnología y la metodología de cromatografía líquida, el análisis selectivo y eficiente de metabolitos altamente polares y cargados se ha determinado en muchos estudios metabolómicos mediante la electroforesis de zona capilar (CZE o CE). Además, como el mecanismo de separación de la CE es fundamentalmente diferente de las técnicas de separación basadas en cromatografía, proporciona una visión complementaria sobre la composición de los metabolitos presentes en una muestra biológica determinada (3). En CE-MS convencional, los compuestos se separan de acuerdo con su relación carga/tamaño; por lo tanto, esta técnica es naturalmente adecuada para el análisis de compuestos muy polares/ionizables. También se sabe que esta técnica es apta para trabajar con una gran variedad de matrices biológicas.

Durante la última década, la CE-MS se ha empleado ampliamente en el estudio metabolómico de compuestos polares catiónicos y, solo en algunos casos, se ha complementado con el estudio de los compuestos polares aniónicos. Actualmente, esto está cambiando y como se puede observar en la revisión de esta técnica realizada por Zhang y Raumatar en el periodo de 2019-2020, el 60% de los estudios se centran tanto en los compuestos polares positivos como negativos (4). Aun así, el análisis de los metabolitos aniónicos sigue suponiendo un esfuerzo desafiante para el investigador.

En los últimos años se han empleado básicamente 3 tipos de capilares (COSMOS, SMILE y sílice fundida) con el fin de determinar metabolitos aniónicos en estudios metabolómicos. Aunque se estén empleando ciertos métodos en este campo específico, llegamos a la conclusión de que todavía hay espacio para el diseño de un

método CE-MS confiable para el perfil aniónico-metabólico. Un método aniónico de CE-MS y la disponibilidad de una base de datos que sirva para identificar los metabolitos sería muy favorable para hacer de la CE-MS una técnica analítica alternativa y viable para la metabolómica.

Bajo nuestra experiencia los capilares propuestos hasta la fecha para el análisis aniónico presentan algunas desventajas. El capilar de COSMOS es una columna capilar recubierta de un polímero catiónico y, aunque ha sido ampliamente desarrollado por el grupo de Soga (5), las principales ventajas que nos ofrece la CE son una alta eficacia de separación, un tiempo de análisis corto y un bajo coste operacional. Sin embargo, a la hora de elegir un capilar con el que trabajar, se deben considerar otros aspectos a parte de la eficacia de separación. El coste del capilar es un importante factor a examinar, siendo mucho más costoso un capilar COSMO(+) que un capilar de alcohol polivinílico (PVA), por ejemplo. A esto se suma el hecho de que el capilar COSMO(+) es considerablemente más complicado en su puesta a punto para el análisis y de mantener su correcto funcionamiento. A nivel del capilar, no hemos podido asegurar análisis reproducibles por pérdidas de señal. Las principales causas de pérdida de señal pueden ser la obstrucción o la rotura. Capilares con tendencia a obstruirse y romperse no son recomendables puesto que, dado que la CE está principalmente enfocada al análisis de un amplio rango de muestras biológicas, generarían problemas a la hora de analizar un elevado número de muestras. Por otra parte, el capilar SMILE, se trata de un capilar comercial de Nacalai Tesque (Kioto, Japón) del que muy poco se conoce de su recubrimiento. Muchas de las aplicaciones descritas utilizan un tampón comercial del que no se conoce la composición, además de que, ha sido puesto de manifiesto la poca reproducibilidad de este capilar generando aumentos graduales en los tiempos de migración y disminución de las alturas de los picos en secuencias consecutivas (6).

En este capítulo se recoge el trabajo realizado durante estos años de investigación para el desarrollo e implementación de un método de análisis de compuestos aniónicos a través de CE-MS. En él, ofrecemos un nuevo capilar de PVA (polivinilalcohol) con el fin de ampliar las posibilidades de la electroforesis capilar-espectrometría de masas (CE-MS) para el estudio del perfil aniónico-metabólico y, lo comparamos con el empleo de otros capilares utilizados con este fin. Además, se ha realizado la optimización y comparación de dos capilares (PVA y sílice fundida) empleando dos electrolitos de fondo (BGE), acetato amónico 50 mM a pH 8.5 y ácido fórmico 0.1 M, con el fin de obtener un método que reúna las mejores condiciones de separación y de detección. Al mismo tiempo, se ha confeccionado una

base de datos de 240 metabolitos, hasta la fecha y que continúa en construcción, con los datos relativos a los tiempos de migración relativos obtenidos frente a metionina sulfona (MetS) y ácido 2-(N-morfolino) etanosulfónico (MES) empleados como patrones internos (IS). Finalmente el método optimizado se ha empleado en la caracterización del perfil metabólico del plasma sanguíneo de cáncer de pulmón de células no pequeñas (NSCLC) y, del estudio comparativo NSCLC frente a control, y de los dos subtipos histológicos más frecuentes, adenocarcinoma (ADC) (n = 39) y carcinoma de células escamosas (SCC) (n = 57) *versus* plasma no-canceroso (n = 22).

2. Materiales y métodos

2.1. Instrumentación

Los análisis de CE-ESI-MS fueron realizados con el sistema de electroforesis capilar 7100 de Agilent acoplado a un espectrofotómetro de masas con analizador de tiempo de vuelo (TOF/MS) Agilent 6230 (Agilent Technologies, Waldbronn, Germany). El acoplamiento de ambos sistemas consta de una fuente de *electrospray* G1607 de Agilent Technologies. Cuando se aplica un potencial negativo, se necesita una aguja de platino para evitar la corrosión (5). El flujo de líquido auxiliar se optimizó utilizando una bomba isocrática Agilent 1200 Infinity I y, estaba compuesto de una mezcla metanol/agua 50:50 (v/v) y dos estándares de referencia, purina ($C_5H_4N_4$, m/z 121.0509) y HP-0921 ($C_{18}H_{18}O_6N_3P_3F_{24}$, m/z 922.0098) para el control de la precisión de masa.

La separación se llevó a cabo utilizando un capilar de sílice fundida (Agilent Technologies) (dimensiones 50 μ m ID x 92 cm longitud total) en polaridad normal y un capilar recubierto de alcohol polivinílico (PVA) (50 μ m ID x 100 cm longitud total) en polaridad inversa. El impacto en términos de resolución y eficiencia de la separación para metabolitos aniónicos se comparó con dos buffer o electrolito de fondo (BGE): ácido fórmico de 0.05 a 0.20 M y, por otro lado, acetato de amonio de 25 a 100 mM a pH = 8.5, en polaridad inversa y normal, respectivamente. Los rangos experimentales para las condiciones de optimización se resumen en la Tabla 1.

Tabla 1. Rangos experimentales de las condiciones de optimización para desarrollar una nueva metodología aniónica mediante el análisis CE-MS.

	Capilar	Silice	PVA
CE condiciones	BGE condition	Formic acid 0.05 to 0.2M or NH ₄ Ac 25 to 100 mM	
	Voltage (kV)	10 to 30	-10 to -30
	BGE pressure (mbar)	10 to 100	
	Injection time(s)/(mbar)	35 and 50/50	
MS parámetros (ionización negativa)	Iso.Pump flow (mL/min)	0.2 to 1.0	
	Drying gas temperature (°C)	200 to 275	
	Gas flow rate (L/min)	2 to 10	
	Capillary voltage (V)	2000 to 4500	
	Nebulizer gas pressure (psig)	4 to 10	

Ambos capilares deben ser acondicionados previo al primer uso. Un nuevo capilar de sílice se acondicionó tras tres etapas sucesivas de 30 minutos cada una: NaOH 0.1M, seguido de agua MiliQ® y BGE. Por el contrario, el capilar de PVA solo se acondicionó con BGE. Antes de cada análisis, el capilar se equilibra con un lavado durante 5 min con BGE. Posteriormente, la muestra se inyecta en el capilar aplicando una presión determinada durante 35 s y, seguidamente, se introduce BGE durante 10 s a 100 mbar.

Los datos se adquirieron utilizando ESI(-) con una velocidad de escaneo de 1.00 espectro/s y el rango de masa de m/z 74 a 1000. Las condiciones de ESI se optimizaron en los siguientes rangos: temperatura del gas de secado = 200 °C - 275 °C y caudal = 2 L/min - 10 L/min; V_{cap} = 2000 V - 4500 V; presión de gas del nebulizador = 4 psig - 10 psig.

Además, se ha llevado a cabo la fragmentación en fuente (ISF, *del inglés in-source fragmentation*) para fragmentar las moléculas y obtener iones producto que nos permita identificar el metabolito. Para ello, tanto las soluciones individuales de los estándares, la mezcla de estándares, como las muestras de plasma se analizaron en las mismas condiciones, pero aumentando el voltaje del fragmentor usando una rampa de 125 a 250 V (7, 8).

Para intentar abarcar el mayor número de compuestos aniónicos durante el análisis a través de un método robusto y fiable, se estudió el conjunto de parámetros específicos elegidos de forma cuidadosa y optimizados con el fin de obtener la mayor resolución y sensibilidad posibles. Sin embargo, parámetros que pueden parecer

óptimos para la separación en el capilar podrían tener efectos adversos sobre la sensibilidad en el detector MS. Para la optimización es fundamental el estudio de los parámetros de la interfaz, ya que en ella se produce el acoplamiento entre la electroforesis capilar y MS. De esta manera, los parámetros relativos a la optimización del método de análisis de compuestos aniónicos se dividieron en tres partes. Sobre la CE, se estudiaron los capilares de sílice y PVA con diferentes electrolitos de fondo, además de parámetros que afectan a la separación de los metabolitos en el capilar. Entre ellos se encuentran el potencial aplicado, el tiempo de inyección o la presión aplicada. A continuación, se estudiaron los parámetros de *electrospray*, o lo que es lo mismo, los parámetros óptimos para formar las partículas cargadas desolvatadas a partir de la muestra inicial en disolución, entre los que se incluyen: el flujo del líquido auxiliar, la temperatura y flujo del gas de secado, el nebulizador y el voltaje del capilar. Finalmente, solo se estudió el voltaje del fragmentor en lo que respecta a MS, siempre trabajando en condiciones de ionización negativas. La Tabla 2 recoge las condiciones óptimas finales.

Tabla 2. Condiciones óptimas finales para cada capilar y BGE utilizado.

Capilar	BGE condition	Condiciones CE			Parámetros de <i>Electrospray</i>					Parámetros MS (ionización negativa)
		Voltage (kV)	BEG pressure (mbar)	Injection time (s)/(mBar)	Iso.Pump flow (mL/min)	Drying gas temperature (°C)	Gas flow rate (L/min)	Capillary voltage (V)	Nebulizer gas pressure (psig)	Fragmentor voltage (V)
Silice	Formic Acid 0.1M	20	75	35/50	1.00	250	10	4000	10	120
	NH4Ac 50 mM	30	50	35/50	1.00	250	10	4000	10	120
PVA	Formic Acid 0.1M	-30	50	35/50	0.60	275	10	2000	10	125
	NH4Ac 50 mM	-30	50	35/50	1.00	275	10	2000	10	125

2.2. Patrones y soluciones de trabajo

Una mezcla de 37 metabolitos diferentes (MIX 37) en concentración 25 ppm se preparó por dilución de las correspondientes disoluciones madre (1000 ppm) y, se empleó para optimizar las condiciones de análisis y obtener la mejor separación posible y la mayor sensibilidad de detección. Esta mezcla incluía ácidos carboxílicos, azúcares y derivados fosfato, metabolitos característicos de la glucólisis, ciclo del ácido tricarboxílico (TCA), pentosa fosfato y 0.2 mM de metionina sulfona (MetS), 0.25 mM de ácido 2-(N-morfolino) etano sulfónico (MES) y 1 mM de paracetamol como IS (Tabla 3).

Tabla 3. Lista de los 37 compuestos presente en la mezcla utilizada para optimizar las condiciones de análisis del método para compuestos aniónicos a través de CE-MS.

Num.	Metabolito	Formula	CAS	Masa	[M-H] ⁻
1	Lactic acid	C3H6O3	10326-41-7	90.0321	89.0248
2	2-Aminobutyric acid	C4H9NO2	1492-24-6	103.0633	102.056
3	Creatinine	C4H7N3O	60-27-5	113.0589	112.0516
4	Guanidineacetic acid	C3H7N3O2	352-97-6	117.0538	116.0465
5	Succinic acid	C4H6O4	110-15-6	118.0266	117.0193
6	Pipecolinic acid	C6H11NO2	3105-95-1	129.0789	128.0716
7	Glutaric acid	C5H8O4	110-94-1	132.0423	131.035
8	Aspartic acid	C4H7NO4	1783-96-6	133.0375	132.0302
9	Malic acid	C4H6O5	6915-15-7	134.0215	133.0142
10	Aminobenzoic acid	C7H7NO2	118-92-3	137.0477	136.0404
11	Glutamic acid	C5H9NO4	11070-68-1	147.0532	146.0459
12	Tartaric acid	C4H6O6	147-71-7	150.0164	149.0092
13	Xanthine	C5H4N4O2	69-89-6	152.0334	151.0261
14	2-Aminoadipic acid	C6H11NO4	542-32-5	161.0688	160.0615
15	Phenyllactic acid	C9H10O3	828-01-3	166.0619	165.0546
16	Phosphoenolpyruvate	C3H5O6P	138-08-9	167.9824	166.9751
17	Uric acid	C5H4N4O3	69-93-2	168.0283	167.021
18	DHAP	C3H7O6P	57-04-5	169.998	168.9907
19	Glycerol-1-phosphate	C3H9O6P	17989-41-2	172.0137	171.0064
20	Aconitic acid	C6H6O6	585-84-2	174.0164	173.0091
21	Hippuric acid	C9H9NO3	495-69-2	179.0582	178.0509
22	Methionine sulfone	C5H11NO4S	7314-32-1	181.0409	180.0336
23	Mannitol	C6H14O6	69-65-8	182.0777	181.0704
24	Phosphocholine	C5H15NO4P	3616-04-4	184.1507	182.0592
25	Phospho-L-serine	C3H8NO6P	407-41-0	185.0091	184.0018
26	Glycerate 3-phosphate	C3H7O7P	820-11-1	185.9929	184.9857
27	Isocitric acid	C6H8O7	320-77-4	192.027	191.0197
28	MES	C6H13NO4S	4432-31-9	195.0565	194.0501
29	E4P	C4H9O7P	7183-41-7	200.0086	201.0159
30	Ru5P	C5H11O8P	4151-19-3	230.0192	229.0119
31	Cytidine	C9H13N3O5	65-46-3	243.0855	242.0782
32	G6P	C6H13O9P	299-31-0	260.0297	259.0224
33	6-phosphogluconate	C6H13O10P	921-62-0	276.0246	275.0174
34	Ophthalmic acid	C11H19N3O6	495-27-2	289.1274	288.1201
35	Sedoheptulose-7-phosphate	C7H15O10P	2646-35-7	290.0403	289.033
36	Argininosuccinic acid	C10H18N4O6	2387-71-5	290.1226	289.1154
37	F1,6BP	C6H14O12P2	488-69-7	339.996	338.9888

DHAP: (*del inglés dihydroxyacetone phosphate*) dihidroxiacetona fosfato, E4P: (*del inglés eritrose-4-phosphate*) eritrosa 4-fosfato, Ru5P: (*del inglés ribulose-5-phosphate*) ribulosa 5-fosfato, G6P: (*del inglés glucose-6-phosphate*) glucosa 6-fosfato, F1,6BP: (*del inglés fructose-1,6-biphosphate*) fructosa 1,6-difosfato.

Por otra parte, se prepararon 240 disoluciones individuales de 25 ppm a partir de disoluciones madre de 1000 ppm que se utilizaron para la caracterización de compuestos aniónicos (RMT y fragmentación) y, posterior construcción de una base de datos. Todas estas disoluciones individuales tenían, además, como patrones internos (IS) MetS en concentración 0.2 mM y MES en 0.25 mM.

2.3. Perfil de plasma de NSCLC y estudio comparativo

Las muestras de plasma procedentes de un estudio de cáncer de pulmón de células no pequeñas (NSCLC), ADC (n = 39), SCC (n = 57) y plasma no canceroso (n = 22), fueron descongeladas el día del análisis y se homogeneizaron mediante agitación en vortex durante 1 min. El tratamiento de la muestra de plasma se realizó según el protocolo establecido por el grupo para este tipo de muestra (4). Sin embargo, en este caso, se agregaron tres patrones internos a la solución: metionina sulfona con una concentración final de 0.2 mM, MES en una concentración final de 0.25 mM y paracetamol en una concentración final de 1 mM. Brevemente, a 100 μ L de plasma se le añadieron 100 μ L de una mezcla acuosa de ácido fórmico en concentración 0.2 M, 5% de acetonitrilo y los patrones internos. Durante 1 min, la mezcla se agitó en vortex y, posteriormente, se transfirió a un filtro Centrifree Milipore (30 kDa) para su centrifugación a 2000 x g durante 80 min a 4 °C. La solución filtrada se transfirió directamente al vial para el análisis.

Durante el análisis, periódicamente, se analizaron muestras conocidas como controles de calidad (QCs) que permiten el control de la estabilidad y el rendimiento del sistema, y la reproducibilidad del procedimiento de tratamiento de muestras (9). Estos se prepararon combinando volúmenes iguales de cada una de las muestras y fueron tratadas como el resto de las muestras.

2.4. Manejo de los datos

El método de trabajo y la adquisición de datos se supervisaron mediante el software Agilent MassHunter Workstation Data Acquisition (B.06.00).

Los electroferogramas de los iones extraídos para cada molécula objetivo de cada estándar individual y para la mezcla 37 fueron obtenidos utilizando el software MassHunter Qualitative (B.10.00, Agilent). Cada pico se caracterizó con su RMT frente a MetS y MES, y se inspeccionaron los espectros de fragmentación para caracterizar los iones producto correspondientes.

Para el procesamiento de los datos procedentes del análisis no dirigido de las muestras de plasma se empleó MassHunter Profinder versión B.10.00 (Agilent Technologies, Santa Clara, CA, EE. UU.) Como ya se ha descrito en detalle previamente (capítulos 1 y 3), este tratamiento consta de la extracción de las características moleculares (MFE) para eliminar el ruido de fondo y los iones no relacionados, la búsqueda de los aductos que coeluyen ($-H^+$, $-COO^-$ y pérdida neutra de agua) y posterior alineación de todas las características (iones) a través de las muestras usando masa y tiempo de migración (MT). Con esto se construyó un espectro final para cada grupo de compuestos para pasar al siguiente paso, la reextracción de los archivos por lotes. Brevemente, la extracción de características recursivas por lotes (RFE) refina la calidad de la lista de objetivos para la función Find by Ion (Fbi) al mejorar la calidad de la lista final de compuestos. Mediante RFE, utilizando tanto la masa como el MT de los resultados de MFE anteriores, se generó una lista final de grupos de compuestos, que fue inspeccionada manualmente para asegurar la correcta integración del electroferograma iónico extraído para cada característica en todas las muestras. El estado de carga máximo permitido fue el doble y se aplicó como filtro 600 cuentas como límite para el ruido de fondo. La alineación se realizó basándose en similitudes entre las muestras de m/z y MT. Los parámetros aplicados fueron 15% para la ventana MT y 20 ppm para tolerancia de masa.

Una vez obtenida la matriz de datos tras el reprocesado, los valores faltantes o perdidos se reemplazaron con los valores apropiados utilizando el script integrado y construido internamente para MATLAB (R2015a, MathWorks) y, se filtró según la variación en la abundancia de los compuestos en las muestras de QC, expresada como una desviación estándar relativa, donde solo se mantuvieron aquellas con $RSD < 30\%$.

2.5. Análisis estadístico

A través del análisis no supervisado PCA, se observaron tendencias basadas en orden de inyección y, por lo tanto, se aplicó el algoritmo de Control de Calidad y Corrección de Regresión de Vector de Soporte (QC-SVRC) (10) para la normalización de la matriz ajustando este efecto dentro de la secuencia de análisis. Con este algoritmo, como se explicó en el capítulo 1, se corrigió la deriva instrumental que se produce durante el análisis utilizando los datos adquiridos de las muestras de control de calidad. Los tres parámetros que determinan la precisión en la eliminación de esta deriva se establecieron en un 5% de la distribución mediana de los valores QC para el umbral de tolerancia (ϵ) y el término de penalización (C) y el ancho del kernel (γ) con 0 y 3

respectivamente. Tras la corrección de los datos, se estudió la normalidad utilizando las pruebas de Kolmogórov-Smirnov-Lilliefors y Shapiro-Wilk y la prueba de Levene para la relación de varianza. Por último, se determinó la significancia estadística de cada metabolito por separado, y las diferencias caso-control y, entre los casos SCC y ADC y los individuos sin cáncer, respectivamente, se evaluaron de forma independiente, aplicando la prueba t de Student (valor de $p \leq 0.05$) utilizando MATLAB (R2015a, MathWorks). Se aplicó el método de post-corrección múltiple de Benjamini-Hochberg a todos los valores de p para controlar la tasa de falsos positivos en el nivel $\alpha = 0.05$.

Con el análisis multivariado (MVA) realizado en SIMCA 15.0 (Sartorius Stedim Biotech), se visualizó la matriz de datos y se verificó el rendimiento y la calidad del análisis. Para la observación del agrupamiento natural se utilizó el método no supervisado PCA y, en este software, también, se construyeron métodos supervisados, como el análisis discriminante de mínimos cuadrados parciales (PLS-DA) y el análisis discriminante de mínimos cuadrados ortogonales (OPLS-DA). Se utilizó el modelo OPLS-DA para la construcción y evaluación de la gráfica *S-plot* y, para la selección de variables, los gráficos “*Volcano*” que representan la importancia de la variable en la proyección (VIP) frente al valor $p(\text{corr})$ (11). La calidad de los modelos construidos se evaluó a través del valor de la varianza (R^2) y la varianza predicha (Q^2), proporcionada por el software. Los resultados del OPLS-DA también incluyen el intervalo de confianza de Jackknife de los modelos con valor diferente de cero. Los modelos se validaron utilizando la prueba de permutación a través de 999 permutaciones, prueba que muestra la validez y el grado de sobreajuste del modelo (12).

Para inspeccionar las variaciones entre los grupos, el porcentaje de cambio se calculó con el promedio de la abundancia para cada compuesto en cada muestra de cada grupo, y se calculó de la siguiente manera: (Promedio para el grupo de casos - Promedio para el grupo sin cáncer) / Promedio para el grupo sin cáncer, en %.

Finalmente, los metabolitos anotados se analizaron en MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca/>) (13) y se construyó un “mapa de calor” o *heatmap*. Además, se llevó a cabo un análisis de las vías metabólicas, proporcionando una visión del metaboloma, destacando aquellas rutas metabólicas que se modifican significativamente o tienen un mayor impacto debido a los metabolitos alterados involucrados.

2.6. Anotación e identificación de los metabolitos

Bajo las condiciones optimizadas para cada capilar, cada uno de los más de 200 estándares analizados individualmente fueron caracterizados con sus tiempos de migración relativo (RMT) frente a los IS, junto con los iones producto característicos obtenidos de los espectros de fragmentación en fuente a través de su extracción manual.

Con el análisis de las muestras mediante metodología de ionización negativa, se caracterizó el perfil de metabolitos aniónicos en muestras de cáncer de pulmón. La identificación tentativa de características basadas en el m/z de los compuestos se realizó mediante la herramienta CEU Mass Mediator (<http://ceumass.eps.uspceu.es/>) (14). La anotación provisional cubría, además de la coincidencia exacta de la masa con el error de masa establecido en 20 ppm, la determinación de la distribución isotópica y la verificación manual de los posibles iones y aductos. Siempre que fue posible, el RMT con cada IS se comparó con los resultados de la base de datos interna construida. Se utilizaron los espectros de fragmentación (250 V) para confirmar las anotaciones y se estudiaron los patrones característicos de la fragmentación en fuente comparándolos con los de la base de datos interna.

3. Resultados

3.1. Condiciones de optimización para el método de análisis

En la búsqueda de las mejores condiciones de análisis se tuvieron en cuenta varios factores. La separación de los compuestos está influenciada principalmente por los parámetros de CE generando unos picos más definidos en el electroferograma. El objetivo era obtener un método relativamente rápido con picos lo más resueltos posible, buscando el compromiso entre cantidad y calidad de la información obtenida.

La optimización del potencial aplicado en la CE se realizó utilizando la relación señal/ruido (S/N) y la resolución de los picos. Tanto las áreas como la relación señal/ruido permitieron estimar la sensibilidad de los métodos. La presión del BGE y el potencial aplicado al capilar CE influyen en la separación y están correlacionados. La Figura 1 muestra el método de superficie de respuesta (RSM) diseñado para ayudar a cuantificar las relaciones entre la presión y el potencial de CE y, el valor de S/N para cada combinación. El potencial de CE aplicado en el capilar de PVA con tampón ácido fue de -10 a -30 kV y la presión de 25 a 100 mbar. Con este gráfico fue posible optimizar los valores de presión y voltaje CE para obtener la mejor relación señal/ruido (S/N). Esto

se llevó a cabo con el estudio de la señal para el compuesto de fosfo-L-serina ($m/z = 184.0018$), entre otros, y se obtuvo el ruido de la señal obtenida (por el método *peak-to-peak definition*). La figura muestra una zona roja entre -20 a -30 kV y de 50 a 80 mbar, correspondiente a la relación señal-ruido máxima. Esta fluctuación asegura la robustez del método, ligeras variaciones en estos parámetros no afectan significativamente los resultados obtenidos. Para obtener la mejor relación S/N, sin perder la resolución de los picos en el electroferograma en el menor tiempo posible, las condiciones óptimas se establecieron en -30 kV y 50 mbar.

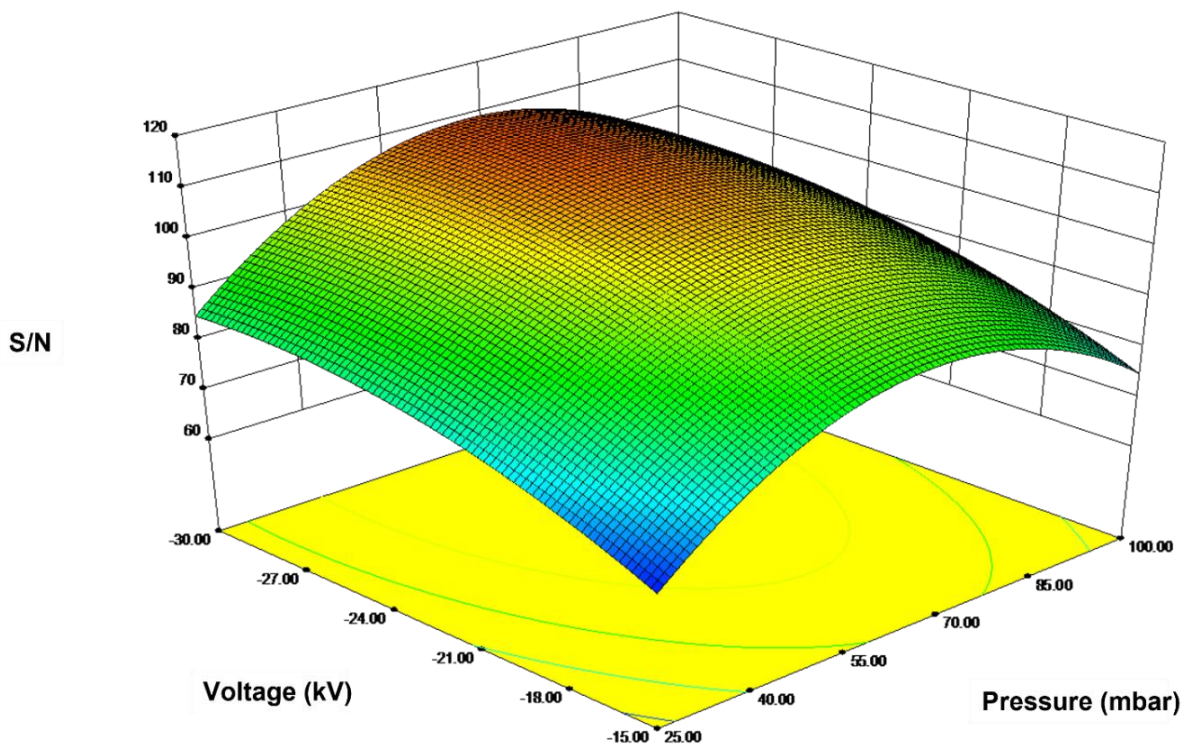


Figura 1. Gráfico RSM para la optimización del potencial y la presión aplicada en CE. Los datos corresponden al capilar PVA en condiciones ácidas y, el rango de voltaje va de -10 a -30 kV y la presión de 25 a 100 mbar. La variable de respuesta recopila los valores de S/N para el compuesto fosfo-L-serina ($m/z = 184.0018$) en cada situación. El valor máximo se localiza a -30 kV y 50 mbar de presión.

Por otro lado, la sensibilidad máxima del método se estableció optimizando los parámetros de la espectrometría de masas: voltaje del capilar, flujo SL, temperatura del gas de secado, caudal del gas, presión del gas del nebulizador y voltaje del fragmentor. Esta optimización se ha realizado mediante el estudio de las señales obtenidas en atención a su intensidad, área, o S/N, entre otros, frente a la variación del parámetro correspondiente.

La Figura 2A muestra cómo varía la relación S/N para los diferentes metabolitos elegidos a medida que aumenta el potencial del capilar. Para el capilar de PVA en condiciones ácidas, aunque la señal aumenta, también fluctúan el ruido y el valor de S/N. Debido a los resultados, si bien 2000 V tuvo una señal más baja, también ofreció menos ruido y, se pudo observar que, solo a 2000 V, no hubo pérdida de rango de masa debido al fenómeno de supresión iónica que se observó en compuestos que tienen tiempos de migración similares. Este fenómeno, conocido como supresión de iones, describe el efecto adverso en la respuesta del detector debido a una reducción en la eficiencia de ionización de los analitos diana. Como los compuestos utilizados como masas de referencia se introducen junto con el SL cuando llegan a la fuente de ionización lo hacen junto con el resto de los metabolitos. En este caso, tres compuestos (ácido láctico, ácido feniláctico y ácido hipúrico) que presentan tiempos de migración muy similares, cuando llegan a la fuente, compiten por la ionización, y se pierde la exactitud de masa que se corrige con las masas de referencia, obteniendo una variación a masa monoisotópica de los compuestos, dando como resultado esos picos desdoblados (Figura 2B).

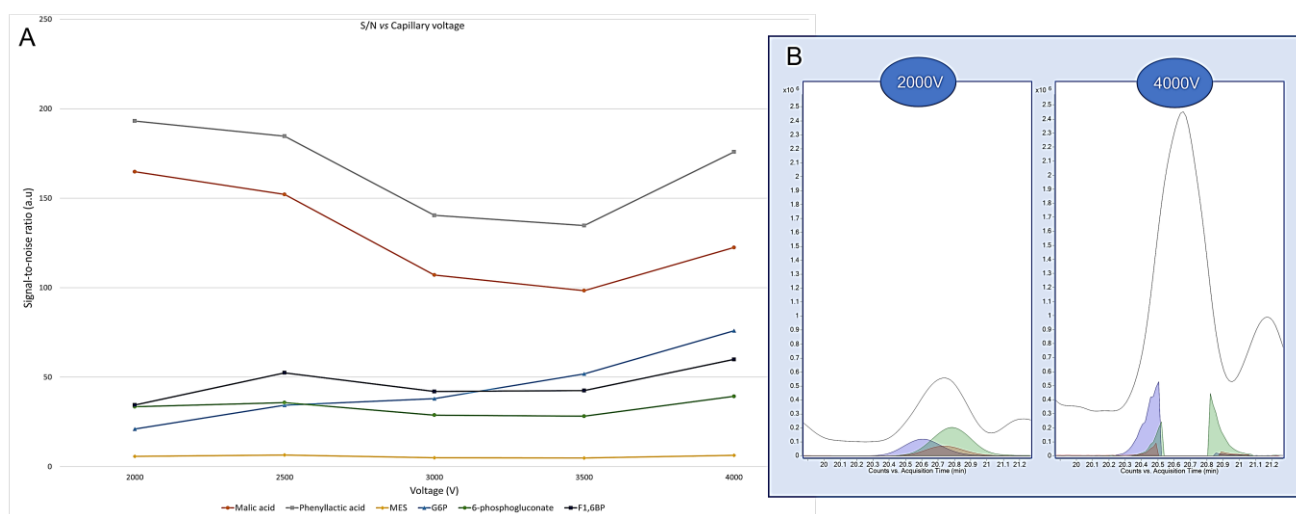


Figura 2. **A.** Efecto del potencial MS sobre la sensibilidad (S/N) de metabolitos aniónicos. Condiciones experimentales: capilar PVA, BGE de ácido fórmico 0.1 M. Metabolitos seleccionados: ácido málico en color naranja ($m/z = 133.0142$), ácido feniláctico en color gris ($m/z = 165.0546$), MES en color amarillo (ácido 2-morfolinoetanosulfónico) ($m/z = 194.0501$), G6P en color azul claro (glucosa-6-fosfato) ($m/z = 259.0224$), 6-fosfogluconato en color verde ($m/z = 275.0174$) y F1,6BP en color azul oscuro (fructosa-1,6-bifosfato) ($m/z = 338.9888$). **B.** Electroferograma de iones totales (TIC) y de iones extraídos (EIC) de los metabolitos seleccionados para 2000 y 4000V. Ácido láctico ($m/z = 89.0248$), color rojo; ácido feniláctico ($m/z = 165.0546$), color verde; ácido hipúrico ($m/z = 178.0509$), color azul.

También es importante tener en cuenta que, al aumentar el voltaje, también podría aumentar la fragmentación en la fuente y, causar una disminución en la sensibilidad. Esto se ha podido comprobar con los resultados obtenidos a valores superiores de 4000 V.

El voltaje MS tanto para el capilar de PVA como de sílice se optimizaron utilizando el área de los iones extraídos de varios metabolitos (ácido málico, ácido feniláctico, MES, G-6-P, 6-fosfogluconato y F-1,6-BP). En todos los casos se aplicó un rango de potenciales de 2000 V a 4000 V y, los mejores resultados se obtuvieron con el potencial de 2000 V en el capilar PVA. Para el capilar de sílice, se eligió 4000 V.

Otro parámetro optimizado fue el flujo de líquido auxiliar (SL), que se modificó de 0.2 a 1.0 mL/min en ambos capilares. La función principal es proporcionar una formación de *electrospray* estable, manteniendo la ionización, independientemente del volumen de líquido que emerja del capilar, y así, evitar cualquier alteración durante el análisis. Además, el SL resultó ventajoso para impulsar la ionización hacia la formación de cationes o aniones independientemente del pH del BGE (14). Al ir comprobando los resultados, se rechazó el flujo de 0.2 mL/min porque apenas se obtuvo señal distinguible del ruido, y, por ejemplo, para el capilar PVA en condiciones ácidas, los valores S/N obtenidos para fosfo-L-serina ($m/z = 184.0018$) fueron S/N (0.4 mL/min) = 206; S/N (0.6 mL/min) = 160.8; S/N (0.8 mL/min) = 164.9; S/N (1.0 mL/min) = 99.4. La mejor S/N se obtuvo con un flujo de 0.4 mL/min. Sin embargo, al extraer los iones del electroferograma, se observó que aparecen los tres compuestos, señalados anteriormente, desdoblados ya que tienen un tiempo de migración muy similar y su señal se superpone. En algunas condiciones, estos picos no tenían una forma definida y comenzaron a desdoblarse. La supresión de iones, descrita con anterioridad, está detrás de este fenómeno, y que no solo se vio afectado por el flujo del SL, sino que también debe considerarse el potencial de ionización. Por esta razón, se comparó la variación de flujo con modificaciones en el potencial de ionización (Figura 3). Las mejores condiciones S/N, sin problemas de pérdida de masa exacta, se dieron para 0.6 mL/min de SL y 2000 V.

En el siguiente paso, el flujo de gas de nebulización se varió entre 2.0 y 10.0 L/min. Las señales en el electroferograma a 2.0 L/min eran apenas visibles debido a que la pulverización no se podía realizar con efectividad. Un incremento del flujo de gas de nebulización mejoró la sensibilidad y las señales presentaron un área más significativa (Figura 4).

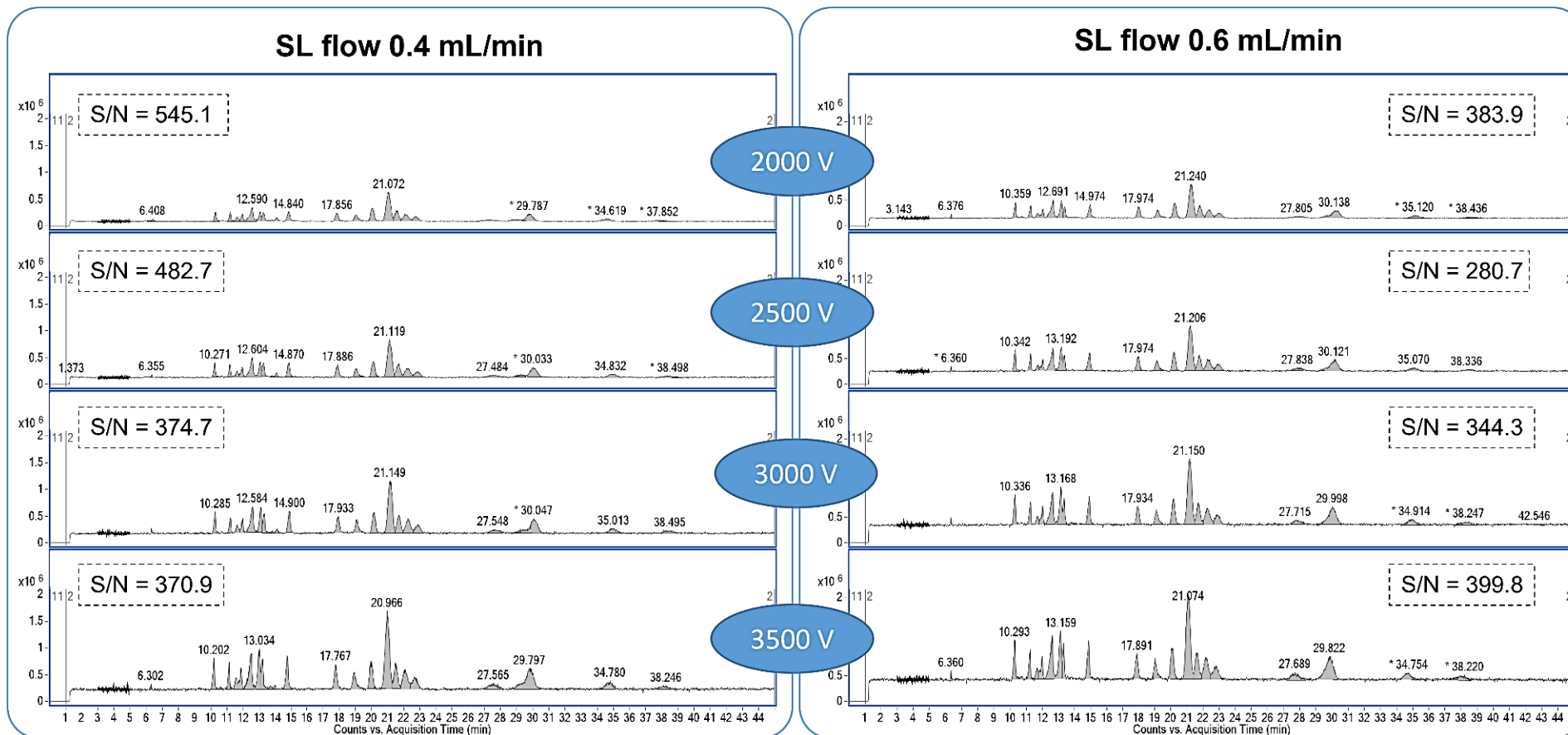


Figura 3. TIC de la mezcla de 37 estándares para los flujos de SL de 0.4 y 0.6 mL/min variando el potencial de ionización, en el capilar de PVA con ácido fórmico 0.1 M como BGE. El valor de S/N se calculó con el área total del electroferograma y el ruido para la región de tiempo entre el minuto 3 y el 5.

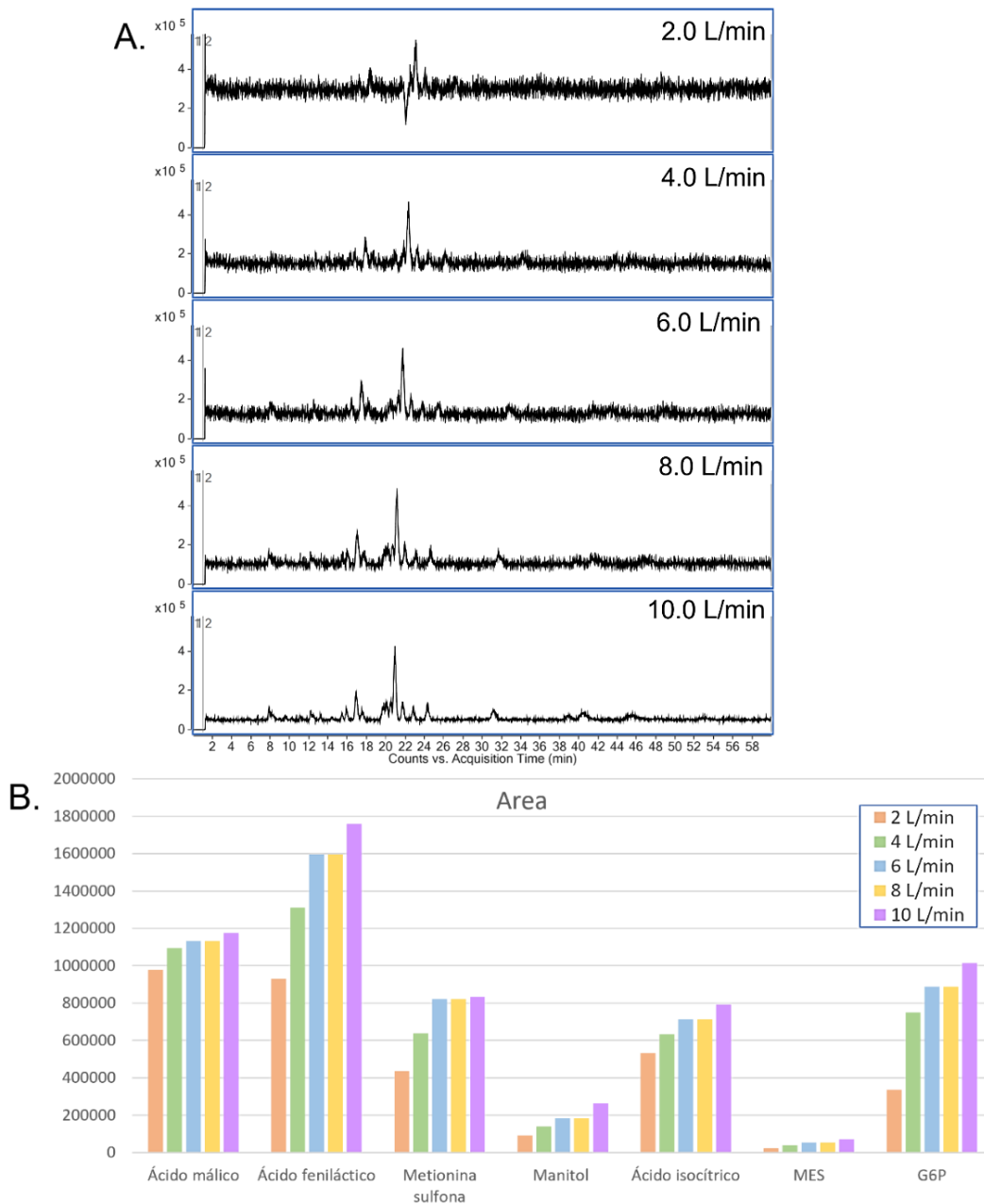


Figura 4. **A.** TIC resultado de los distintos flujos de gas de nebulización estudiados mediante el capilar PVA en condiciones ácidas. **B.** Representación gráfica de las áreas de metabolitos seleccionados (u.a) en el eje abcisas (ácido málico, ácido feniláctico, metionina sulfona, manitol, ácido isocítrico, MES y G6P) para la optimización de los caudales de gas de nebulización estudiados en las mismas condiciones.

Una vez optimizado el caudal, se pasó a probar diferentes presiones de pulverización (Figura 5). El rango estudiado fue de 4 a 10 psig, ya que a 2 psig la presión era insuficiente y el electroferograma obtenido carecía completamente de señal. De la misma forma, se pudo observar cómo este parámetro influye en la sensibilidad de la

detección y cómo, al aumentar la presión, mejora considerablemente la intensidad de la señal.

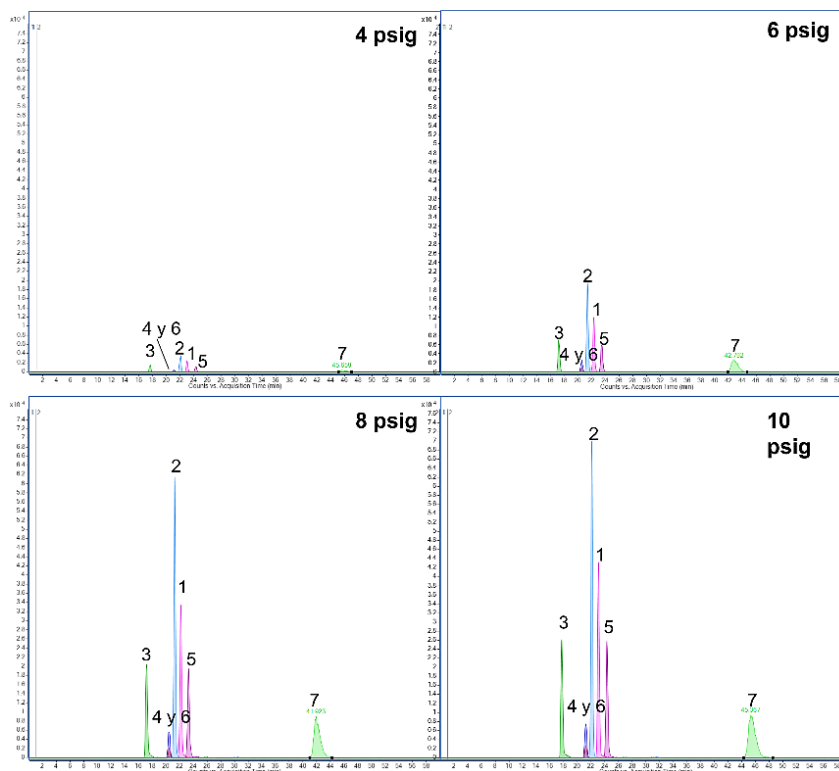


Figura 5. EIC de metabolitos seleccionados para la optimización de la presión del gas de nebulización. Estos resultados corresponden a la optimización del método de análisis para el capilar de PVA con ácido fórmico 0.1 M. 1. Ácido málico, 2. Ácido feniláctico, 3. Metionina sulfona, 4. Manitol, 5. Ácido isocítrico, 6. MES, 7. G6P, todos ellos extraídos con su ión monovalente $[M-H]^-$.

Para estudiar la reproducibilidad y sensibilidad de los 4 métodos en las condiciones óptimas de cada capilar y BGE, se analizaron 6 réplicas de la mezcla de 37 estándares (mix 37). Todos los compuestos se detectaron con sus iones moleculares desprotonados monovalentes $[M-H]^-$. La desviación estándar relativa (RSD) de los RMT con respecto al MES para las especies aniónicas se calcularon y representaron en la Figura 6. Estos resultados estaban comprendidos entre 0.05 y 1.50%. En esta figura se puede observar que, tanto para el capilar de sílice como para el capilar de PVA, la elección de NH_4Ac como tampón proporcionó valores de RSD más altos que para el ácido fórmico. Además, se concluyó que los mejores resultados se obtuvieron para el capilar PVA con BGE ácido. La reproducibilidad también se estudió para las áreas de los picos y, se obtuvieron resultados por debajo del 10% a excepción de: xantina en capilar PVA y BGE de ácido fórmico; ácido glutámico, ácido úrico y dihidroxiacetona

fosfato (DHAP) en el capilar PVA y BGE acetato amónico; ácido aconítico, manitol y ribulosa 5-fosfato en capilar de sílice y BGE acetato de amonio.

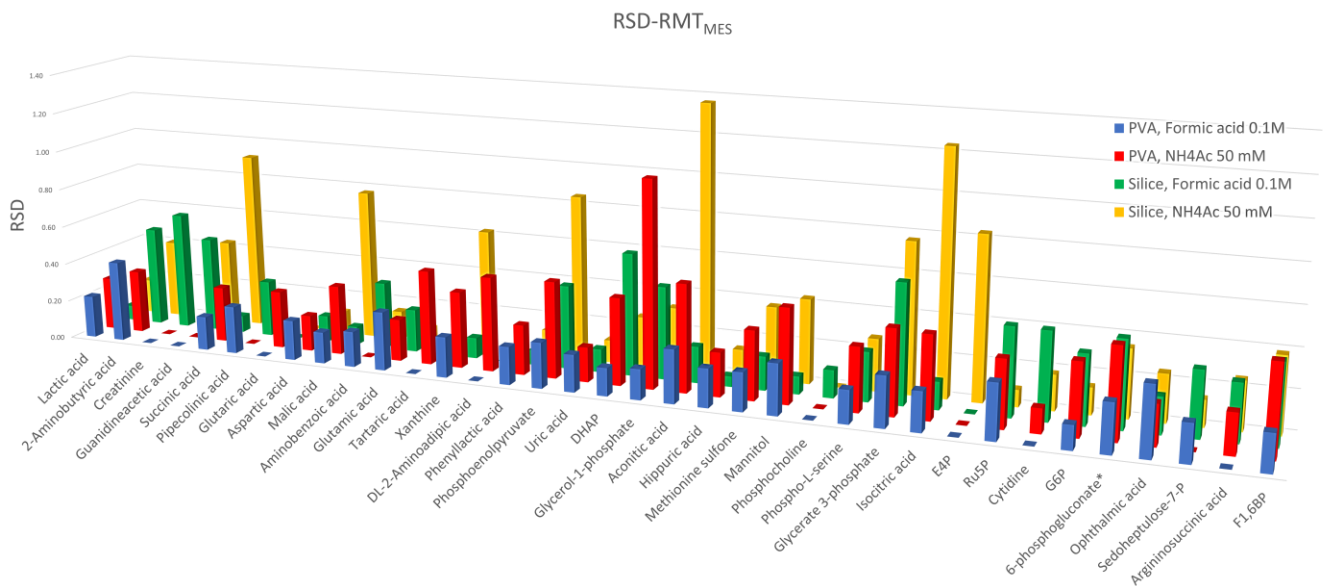


Figura 6. Representación de los valores de RSD calculados para el RMT versus MES de los metabolitos presentes en la mezcla de estándares (mix 37) para el análisis en las condiciones óptimas para cada capilar y BGE estudiado. Las condiciones experimentales se resumieron en la Tabla 2.

Por último, se fijó el voltaje del fragmentor a 125 V para todas las metodologías y, para la fragmentación en fuente (ISF), se trabajó con segmentos de tiempo en el modo de adquisición. Este modo de adquisición permite obtener diferentes espectros de masas para un segmento de tiempo muy corto. En este caso, se mantuvieron los parámetros optimizados y se varió el rango de m/z adquirido, disminuyéndolo hasta 40 Da y, el voltaje del fragmentor fue modificado en cada segmento desde 125 a 250 V. De esta forma se pudo caracterizar cada señal con los iones fragmentos adquiridos experimentalmente, proporcionando un nivel de confianza superior en la anotación del compuesto. A medida que el voltaje fue incrementado, disminuyó la intensidad del ion precursor y fue aumentando la intensidad de los iones fragmentos característicos del metabolito. La Figura 7 resume lo explicado, donde se puede observar en los espectros de fragmentación de la adenosina-5'-monofosfato, como el ion precursor correspondiente al ion mayoritario $[M-H]^- = 346.055$ va disminuyendo, y como consecuencia, aumentaron los iones fragmento característicos con $m/z = 78.95, 96.96, 134.04, 150.98, 192.99$ y 211.00 .

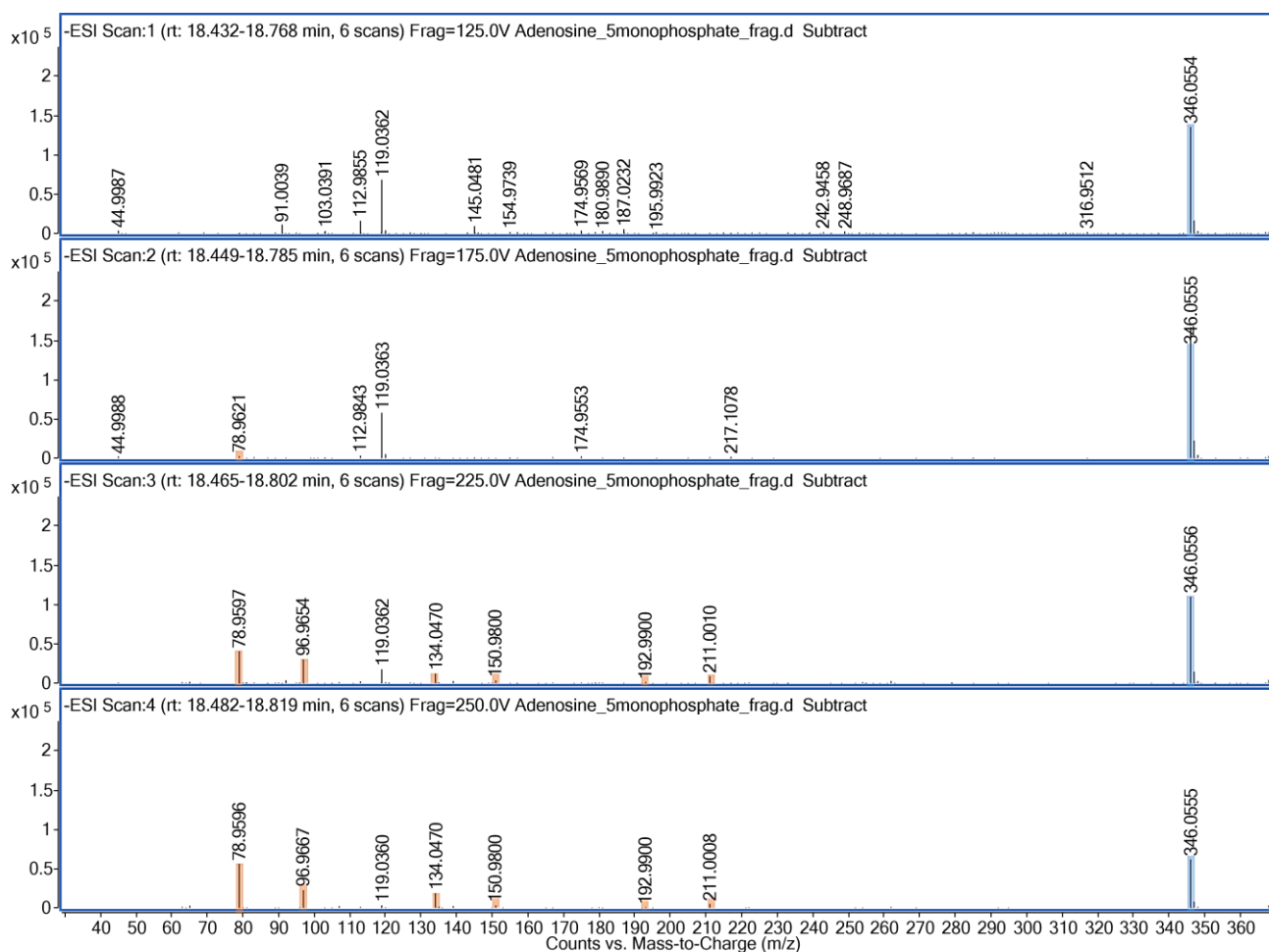


Figura 7. Espectros de fragmentación en fuente con distintos voltajes de fragmentor (125, 175, 225, 250 V) para la adenosina-5´ monofosfato. Ion precursor marcado en color azul, iones fragmento en color naranja.

Finalmente, para cada capilar y en ambas condiciones de BGE, se caracterizó el perfil de la mezcla 37 utilizando el método optimizado (Tabla 2). Los EIC de cada metabolito se extrajeron y nombraron de acuerdo con la tabla de identificación (Tabla 3) y se resumieron en la Figura 8. Considerando los resultados en términos de repetibilidad, reproducibilidad, estabilidad del capilar durante el análisis y el número de metabolitos obtenidos, se seleccionó el capilar PVA con BGE de ácido fórmico 0.1 M. El número de metabolitos detectados en la mezcla 37, la intensidad de los picos, la resolución de estos, el RSD en los RMT y el tiempo de vida del capilar fueron algunos de los resultados que nos hicieron decidir por esta metodología.

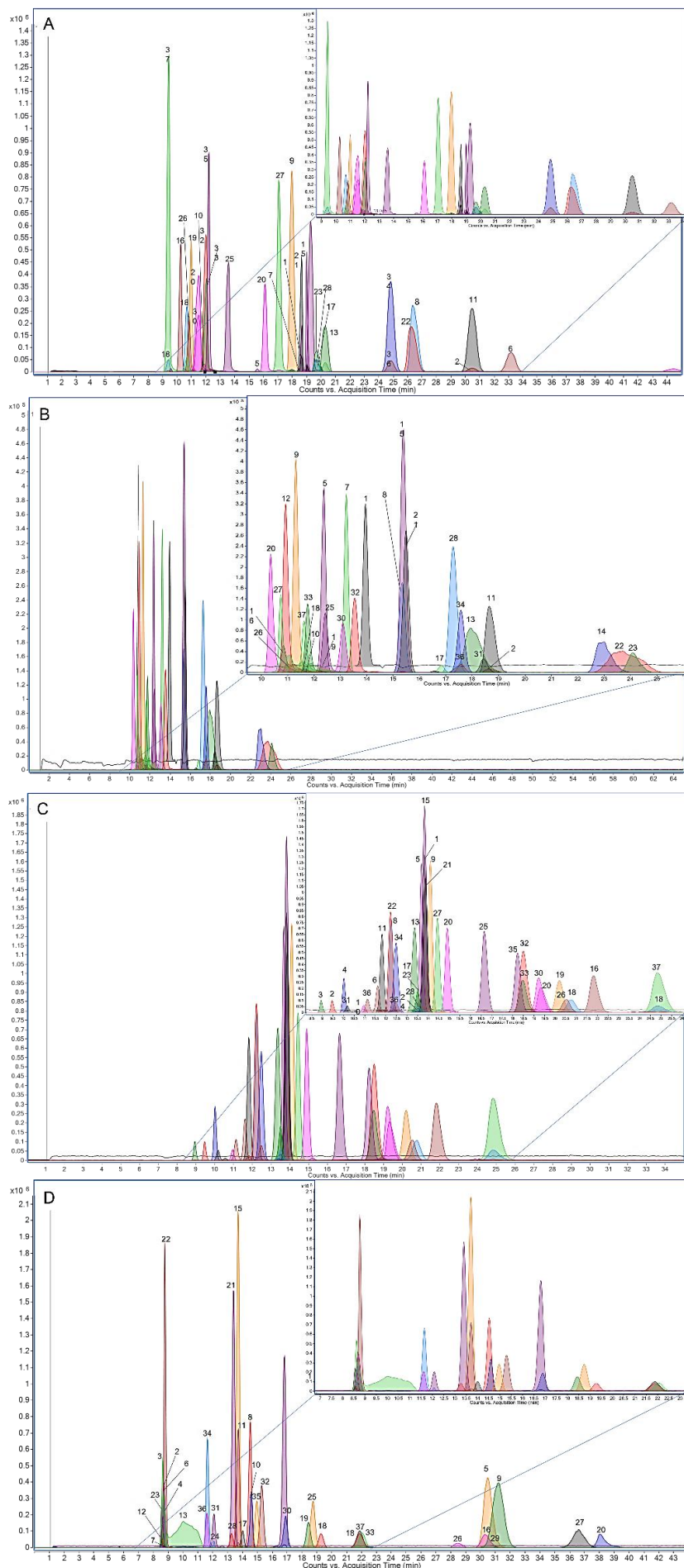


Figura 8. EIC de la mezcla 37 para cada capilar en ambas condiciones de BGE. Las condiciones óptimas para cada análisis se enumeraron en la Tabla 2. **A.** Capilar de PVA con BGE de ácido fórmico 0.1 M, **B.** Capilar de PVA con BGE de NH₄Ac 50 mM, **C.** Capilar de sílice con BGE de ácido fórmico de 0.1 M, **D.** Capilar de sílice con BGE de NH₄Ac de 50 mM. La anotación numérica se resume en la tabla 3.

Durante esta investigación, también se trabajó con otros capilares de los cuales no se obtuvieron resultados reproducibles y/o el capilar era bastante inestable, solo permitía unos pocos análisis antes de romperse. Es el caso de la columna DB-WAX (J&W DB-WAX GC Column, Agilent Technologies, Santa Clara, CA, EE. UU.), con fase estacionaria de polietilenglicol, propia de GC-MS, y el capilar recubierto catiónico COSMO(+) (50 μm d.i.) (Nacalai Tesque, Kyoto, Japón), ampliamente estudiado por el grupo de Soga (5).

3.2. Base de datos de CE-MS aniónico

Los datos adquiridos del análisis individual de 240 estándares fueron procesados y caracterizados con sus RMT frente a MetS y MES. Todos los metabolitos se caracterizaron a través de su ion monovalente $[M-H]^-$ como aducto mayoritario. Asimismo, se evaluaron los espectros ISF adquiridos experimentalmente y, la lista de iones fragmentos característicos del metabolito se recogió en la base de datos. Estos datos fueron recopilados del análisis de los estándares utilizando cuatro segmentos de tiempo en el modo de adquisición con potenciales de fragmentación desde los 125 a los 250 V. Con estos datos se pudo observar claramente lo comentado con anterioridad; a medida que aumentaba el fragmentor, la intensidad del ion precursor disminuía y a su vez, aumentaban los iones fragmento. La recolección de estos datos incluye la sustracción del fondo y la comparación del patrón de fragmentación observado para cada compuesto, con los resultados registrados en la base de datos MS/MS METLIN (<https://metlin.scripps.edu/>) a 10 eV.

Nuestra base de datos está formada por diferentes ácidos carboxílicos y derivados, con diferente longitud y rigidez de la cadena, derivados fosforilados, azúcares y sus derivados, aminoácidos y ácidos biliares, y se continúa trabajando en su curado y ampliación para la integración en la plataforma CEU Mass Mediator (CMM).

3.3. Perfil aniónico de NSCLC

El plasma sanguíneo ($n = 96$) procedente de un estudio de cáncer de pulmón de células no pequeñas (NSCLC) fue utilizado para evaluar el método optimizado para el análisis de compuestos aniónicos por CE-MS. Este tipo de muestra es uno de los biofluidos más accesibles para el estudio metabolómico y, permite evaluar los mecanismos implicados en la patología para identificar nuevos biomarcadores de la enfermedad. Para el estudio del perfil aniónico de NSCLC, las muestras se analizaron utilizando el método elegido y se realizó la deconvolución siguiendo el proceso explicado

en el apartado “*Manejo de los datos*”. Posteriormente, se filtró la lista de características obtenidas y, los datos recolectados se utilizaron para una primera anotación tentativa utilizando la plataforma CMM y la base de datos creada. Siempre que fue posible, estas anotaciones se confirmaron utilizando los RMT y fragmentos de iones de ISF recopilados en la base de datos. Se identificaron 66 metabolitos en plasma de NSCLC gracias al método establecido para analizar compuestos aniónicos mediante CE-MS. Es necesario tener en cuenta que la incorrecta identificación de los compuestos podría llevar a una conclusión biológica falsa.

Tras el filtrado, 329 de 343 características tenían RSD por debajo del 30% en las muestras de QC y, tras eliminar dímeros, fragmentos y artefactos, se anotaron 66 metabolitos, resumidos en la Tabla 4. La tabla resume, entre otros, aquellos compuestos anotados por la información proporcionada por MS¹ (nivel 3), aquellos identificados por RMT del análisis de estándares incluidos en nuestra base de datos interna (L2*) y aquellos que también incluyen fragmentación en fuente (L2), siguiendo las disposiciones establecidas de la iniciativa metabolómica estandarizada (15). Los compuestos identificados son en su mayoría aminoácidos y sus derivados, ácidos carboxílicos y compuestos orgánicos del ciclo del TCA, azúcares y sus fosfoderivados.

3.4. Estudio comparativo

Tras el alineado y procesamiento de los datos por RFE a través de MassHunter Profinder, se obtuvieron 615 características que, con el filtrado manual de duplicados, artefactos, etc., se adquirió una matriz compuesta por 343 características. Se realizaron análisis no supervisados, PCA, para el análisis inicial del conjunto de datos y verificar el agrupamiento espontáneo de los QCs. Posteriormente, se completó el análisis multivariado con modelos supervisados PLS-DA y OPLS-DA. Los modelos OPLS-DA supervisados de la Figura 9 muestran una clara separación entre ambos grupos y controles.

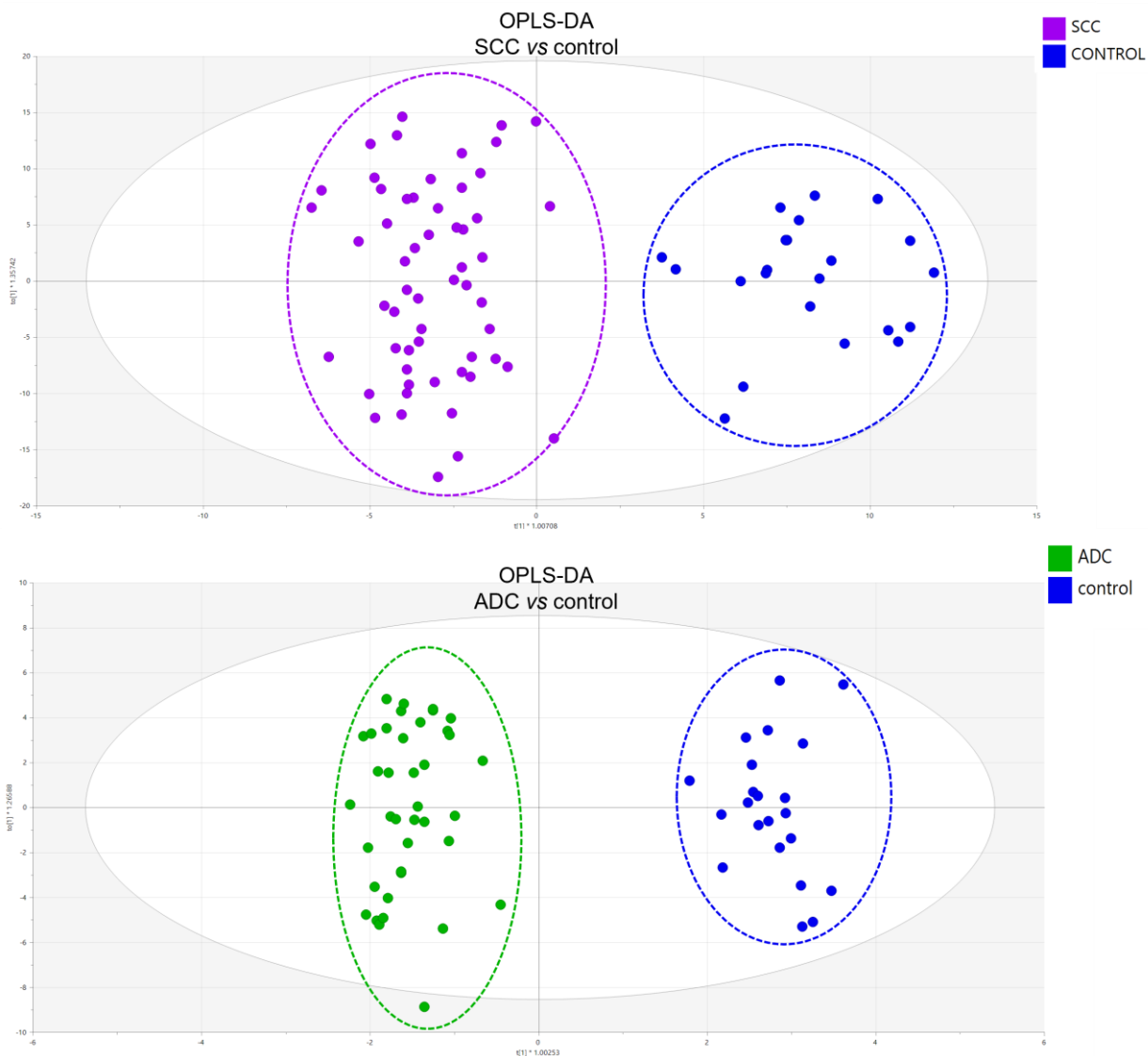


Figura 9. El panel muestra los gráficos de OPLS-DA. A. Modelo OPLS-DA para SCC vs control (puntos morados, pacientes con SCC ($n = 57$); puntos azules, pacientes sin cáncer ($n = 22$)) ($R^2 = 0.704$, $Q^2 = 0.592$); B. Modelo OPLS-DA para ADC vs control (puntos verdes, pacientes con ADC ($n = 39$); puntos azules, pacientes sin cáncer ($n = 22$)) ($R^2 = 0.769$, $Q^2 = 0.653$). Los gráficos indican que los pacientes no-cáncer se pueden separar claramente de la mayoría de los individuos con NSCLC.

Tabla 4. Compuestos identificados en la caracterización de muestra de plasma procedentes de pacientes NSCLC (n = 96) mediante el método optimizado para CE-MS aniónico. Cada identificación se caracterizó por los niveles de confianza según la fuente de información: L3, compuestos anotados por la información brindada por MS¹ a través de CMM; L2*, compuestos anotados por m/z y tiempo de migración relativo del análisis de estándares incluidos en nuestra base de datos interna; L2, identificación de compuestos que también incluyen fragmentación en fuente. Todos los metabolitos fueron caracterizados con diferentes parámetros: MT, tiempo de migración obtenido mediante análisis aniónico CE-MS en minutos; tiempo de migración relativo rMT sobre metionina sulfona (MetS) y ácido 2- (N-Morfolino) -etanosulfónico (MES); HMDB, código de base de datos del metaboloma humano; Nivel, nivel de confianza de la anotación; RSD, desviación estándar relativa de los metabolitos en las muestras QC.

Metabolito	Formula	Masa Molecular (Da)	Error (ppm)	Aducto	MT (min)	rMT MetS	rMT MES	HMDB	Nivel	Fuente de anotación	RSD
Lactic acid	C3H6O3	90.0317	9	M-H	14.72	0.78	0.97	HMDB0001311	L2*	In-house DB	5
Phosphoric acid	H3PO4	97.9769	2	M-H	10.35	0.55	0.68	HMDB0000973	L3	CMM	4
N,N-Dimethylglycine/N-methylalanine	C4H9NO2	103.0633	1	M-H	20.69	1.10	1.36	HMDB0000092	L2*	in-house DB	5
2-hydroxy-butyrac acid	C4H8O3	104.0473	1	M-H	14.81	0.79	0.97	HMDB0000710	L2	In-house DB	5
3-hydroxy-butyrac acid	C4H8O3	104.0473	5	M-H	15.02	0.80	0.99	HMDB0000442	L2*	In-house DB	4
Serine	C3H7NO3	105.0426	0	M-H	22.45	1.19	1.48	HMDB0003406	L2*	In-house DB	5
Proline	C5H9NO2	115.0633	0	M-H	19.73	1.05	1.30	HMDB0000162	L2*	In-house DB	8
ketovaleric acid	C5H8O3	116.0473	1	M-H	10.09	0.54	0.66	HMDB0001865	L2*	In-house DB	4
amino-pentanoic acid/valine/norvaline/betaine	C5H11NO2	117.0790	1	M-H	22.90	1.22	1.51	HMDB0003355	L3	CMM	8
Succinic acid	C4H6O4	118.0266	4	M-H	15.01	0.80	0.99	HMDB0000254	L2*	In-house DB	4
Taurine	C2H7NO3S	125.0147	6	M-H	15.25	0.81	1.00	HMDB0000251	L2	In-house DB	3
Pyroglutamic acid	C5H7NO3	129.0426	2	M-H	13.95	0.74	0.92	HMDB0000267	L2*	In-house DB	4
Ketoleucine 2-/3-/5-keto-n-caproic acid	C6H10O3	130.0630	2	M-H	9.92	0.53	0.65	HMDB0000695 HMDB0001864	L3	CMM	4
hydroxyproline	C5H9NO3	131.0582	3	M-H	18.33	0.97	1.21	HMDB0000725	L3	CMM	8
Creatine	C4H9N3O2	131.0695	2	M-H	31.28	1.66	2.06	HMDB0000064	L3	CMM	14
Isoleucine/Leucine/Norleucine	C6H13NO2	131.0946	2	M-H	22.67	1.21	1.49	HMDB0000172	L2*	In-house DB	8

Glutaric acid	C5H8O4	132.0423	14	M-H	15.09	0.80	0.99	HMDB0000661	L2*	In-house DB	4
Asparagine	C4H8N2O3	132.0535	2	M-H	21.36	1.14	1.41	HMDB0000168	L2	CMM	4
Aspartic acid	C4H7NO4	133.0375	3	M-H	18.91	1.01	1.24	HMDB00006483	L2	In-house DB	5
Malic acid	C4H6O5	134.0215	2	M-H	14.29	0.76	0.94	HMDB0000156	L2*	In-house DB	4
Threonic Acid	C4H8O5	136.0372	0	M-H	14.61	0.78	0.96	HMDB0062620	L2	In-house DB	4
Adipic acid	C6H10O4	146.0579	13	M-H	15.15	0.81	1.00	HMDB0000448	L2*	in-house DB	20
Glutamine	C5H10N2O3	146.0691	3	M-H	21.01	1.12	1.38	HMDB0000641	L2*	In-house DB	4
Glutamic acid	C5H9NO4	147.0532	9	M-H	20.70	1.10	1.36	HMDB0000148	L2	In-house DB	6
D-Ribose / D-Ribulose / D-Arabinose / D-Xylulose / D-Lyxose	C5H10O5	150.0528	11	M-H+FA	14.65	0.78	0.96	HMDB0000621	L3	CMM	5
N-Formyl-L-aspartate	C5H7NO5	161.0324	5	M-H	12.59	0.67	0.83	HMDB0060495	L3	in-house DB	7
Methionine sulfoxide	C5H11NO3S	165.0460	16	M-H	15.44	0.82	1.02	HMDB0002005	L3	CMM	4
Phenylalanine	C9H11NO2	165.0790	7	M-H	20.56	1.09	1.35	HMDB0000159	L2*	In-house DB	6
Ribonic acid	C5H10O6	166.0477	10	M-H	14.64	0.78	0.96	HMDB0000867	L3	CMM	5
Pyridoxal	C8H9NO3	167.0582	2	M-H	15.39	0.82	1.01	HMDB0001545	L3	CMM	4
Uric acid	C5H4N4O3	168.0283	1	M-H	15.23	0.81	1.00	HMDB0000289	L2*	In-house DB	4
Cysteate	C3H7NO5S	169.0045	0	M-H	19.60	1.04	1.29	HMDB0002757	L3	CMM	7
2,3-Dihydrodipicolinate	C7H7NO4	169.0375	5	M-H	20.66	1.10	1.36	HMDB0012247	L3	CMM	4
Glycerol 1-phosphate	C3H9O6P	172.0137	5	M-H	9.83	0.52	0.65	HMDB0000126	L2*	In-house DB	6
Acetylleucine	C8H15NO3	173.1052	6	M-H	15.17	0.81	1.00	HMDB0011756	L2*	in-house DB	4
Aconitic acid (pico 1)	C6H6O6	174.0164	3	M-H	10.20	0.54	0.67	HMDB0000072	L2*	In-house DB	4
Aconitic acid (pico 2)	C6H6O6	174.0164	0	M-H	13.79	0.73	0.91	HMDB0000072	L2*	In-house DB	5
Suberic acid	C8H14O4	174.0892	6	M-H+FA	15.16	0.81	1.00	HMDB0000893	L2*	in-house DB	4
Citrulline	C6H13N3O3	175.0957	2	M-H	20.99	1.12	1.38	HMDB0000904	L2*	In-house DB	4
Fructose/glucose/galactose	C6H12O6	180.0634	2	M-H	14.73	0.78	0.97	HMDB0000660	L2*	In-house DB	5
Tyrosine	C9H11NO3	181.0739	6	M-H	20.46	1.09	1.35	HMDB0000158	L3	CMM	4
2-/3-Phosphoglyceric acid	C3H7O7P	185.9929	7	M-H	9.75	0.52	0.64	HMDB0000807	L2	In-house DB	5
N-Heptanoylglycine	C9H17NO3	187.1208	5	M-H	15.20	0.81	1.00	HMDB0013010	L2*	in-house DB	4

Isocitric acid Citric acid	C6H8O7	192.0270	3	M-H	13.77	0.73	0.91	HMDB0000193 HMDB0000094	L2	In-house DB	4
Quinic Acid	C7H12O6	192.0634	7	M-H	14.58	0.78	0.96	HMDB0003072	L2*	In-house DB	4
Glucuronic acid	C6H10O7	194.0427	10	M-H	14.06	0.75	0.93	HMDB0000127	L3	CMM	4
Gluconic Acid	C6H12O7	196.0583	9	M-H	14.65	0.78	0.96	HMDB0000625	L2*	In-house DB	5
Cysteine-S-sulfate	C3H7NO5S2	200.9766	3	M-H	10.52	0.56	0.69	HMDB0000731	L3	CMM	4
Capryloylglycine	C10H19NO3	201.1365	4	M-H	15.21	0.81	1.00	HMDB0000832	L3	CMM	4
Tryptophan	C11H12N2O2	204.0899	0	M-H	21.57	1.15	1.42	HMDB0000929	L2	In-house DB	5
Sedoheptulose	C7H14O7	210.0740	10	M-H	15.40	0.82	1.01	HMDB0003219	L3	CMM	5
Galactaric acid Glucaric acid	C6H10O8	210.0376	3	M-H	10.68	0.57	0.70	HMDB0000639 HMDB0000663	L3	CMM	4
Phosphocreatine	C4H10N3O5P	211.0358	3	M-H	13.20	0.70	0.87	HMDB0001511	L3	CMM	5
Cystine	C6H12N2O4S2	240.0238	7	M-H	19.60	1.04	1.29	HMDB0000192	L2*	In-house DB	4
Uridine	C9H12N2O6	244.0695	1	M-H	15.37	0.82	1.01	HMDB0000296	L2*	In-house DB	5
Glucosa-/Fructose 6-phosphate	C6H13O9P	260.0297	6	M-H	10.46	0.56	0.69	HMDB0000124	L2*	In-house DB	4
2,3-Diphosphoglyceric acid	C3H8O10P2	265.9593	7	M-H	8.37	0.44	0.55	HMDB0001294	L2*	In-house DB	5
Norphthalmic acid	C10H17N3O6	275.1117	3	M-H	17.96	0.95	1.18	HMDB0005766	L3	CMM	5
3-hydroxy-tetradecanedioic acid	C14H26O5	274.1780	4	M-H+FA	15.19	0.81	1.00	HMDB0000394	L3	CMM	4
Cysteinylglycine disulfide	C8H15N3O5S2	297.0453	4	M+FA-H	13.89	0.74	0.91	HMDB0000709	L3	CMM	24
Phenethylamine glucuronide	C14H19NO6	297.1212	7	M-H	10.35	0.55	0.68	HMDB0010323	L3	CMM	4
N-Acetylglucosamine 6-sulfate	C8H15NO9S	301.0468	25	M+FA-H	13.82	0.73	0.91	HMDB0000814	L3	CMM	9
Acetylneuraminic acid	C11H19NO9	309.1060	4	M-H	13.41	0.71	0.88	HMDB0000230	L2*	In-house DB	8
S-(Hydroxymethyl)glutathione	C11H19N3O7S	337.0944	11	M-H	20.77	1.10	1.37	HMDB0004662	L3	CMM	6
Sedoheptulose 1,7-bisphosphate	C7H16O13P2	370.0066	7	M-H	9.73	0.52	0.64	HMDB0060274	L3	CMM	5
ADP	C10H15N5O10P2	427.0294	3	M-H	11.24	0.60	0.74	HMDB0001341	L2*	In-house DB	7
IDP	C10H14N4O11P2	428.0134	7	M+Na-2H	13.74	0.73	0.90	HMDB0003335	L3	CMM	4

MT, tiempo de migración; rMT, tiempo de migración relativo; MetS, metionina sulfona; MES, 2-(N-Morpholino)-ethanesulfonic acid; HMDB, (<https://hmdb.ca/>) código identificativo de Human Metabolome Database; DB, base de datos; CMM, base de datos de CEU Mass Mediator (http://ceumass.eps.uspceu.es/mediator/index_cesearch.shtml); RSD, desviación estándar relativa; ADP, adenosina difosfato; IDP, inosina difosfato.

Los OPLS-DA fueron validados mediante prueba de permutación ($n = 999$) y los modelos presentaron valores de R^2 y Q^2 con diferencia menor a 0.3, y valores altos de Q^2 , lo que refleja la alta calidad de los modelos.

A continuación, los datos se refinaron mediante el coeficiente de variación de los QC ($RSD < 30\%$), obteniéndose 329 características, y, a partir de ahí, se realizó el estudio comparativo. Por un lado, el estudio de SCC frente a control generó 114 características significativas (valor de $p < 0.05$ y % de cambio $> 20\%$). ADC vs control resultó en 69 características significativas. Finalmente, se pudieron anotar 31 compuestos para la comparación SCC y 24 para la ADC. Estos metabolitos significativos se anotaron como se describe en la sección "*Anotación e identificación de compuestos*" y, se incluyeron en la Tabla 5, para ambos análisis comparativos, SCC vs control y ADC vs control, entre otras características, error de masa medido experimentalmente en comparación con la masa monoisotópica de las bases de datos o el tiempo relativo de migración. La tabla resumió la fuente de anotación utilizada, explicada anteriormente, y que los agrupa en nivel 3, nivel 2* y nivel 2 (15), además de la fórmula química, masa, % de cambio, valor de p , $p(\text{corr})$ o VIP.

De ambas comparaciones fue posible observar alteraciones en la vía de glucólisis y gluconeogénesis, el metabolismo de aminoácidos como taurina o glutamina/glutamato y, la biosíntesis de fenilalanina, tirosina y triptófano (Figura 10). Destacan metabolitos como los azúcares y derivados fosforilados, aldosa, ribosa, sedoheptulosa, ácido difosfoglicérico, sedoheptulosa 1,7-bisfosfato y ácidos carboxílicos como el ácido láctico, aspártico, glutárico, adípico o piroglutámico.

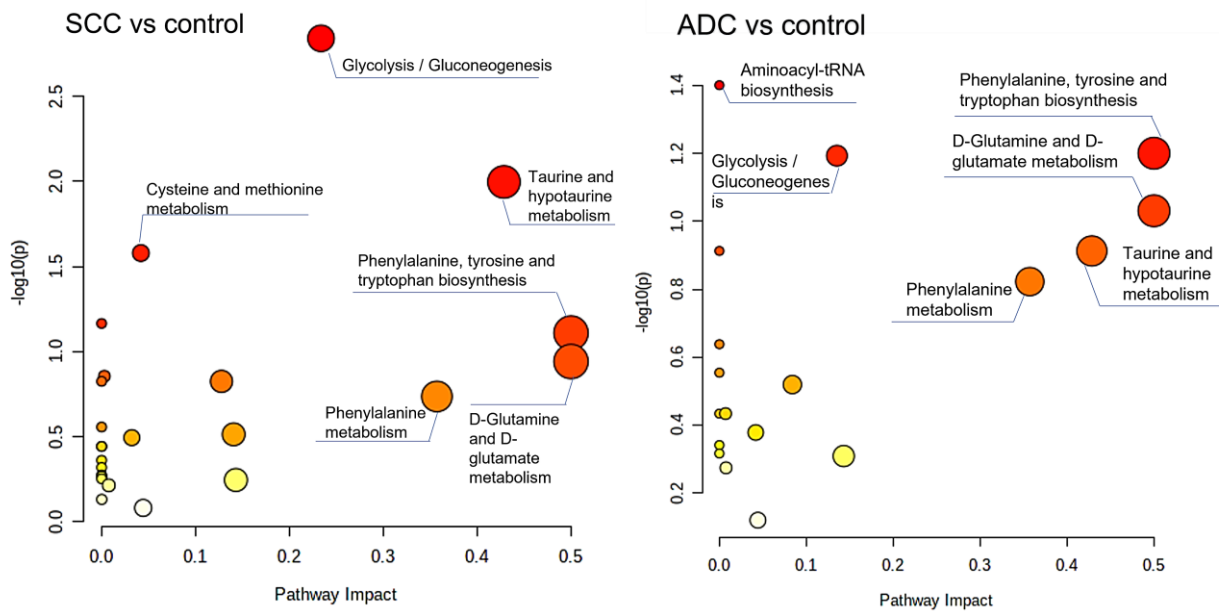


Figura 10. Análisis de rutas metabólicas. Las vías metabólicas alteradas mostraron varios cambios en el metabolismo en la comparación de NSCLC versus control. El color y el tamaño de cada círculo se basan en su valor p y el valor del impacto de la vía, respectivamente. Como resultado, las vías ubicadas en la esquina superior derecha del gráfico son las más afectadas por NSCLC.

Finalmente, la construcción de un mapa de calor o *heatmap*, representa la abundancia de cada metabolito en cada muestra de ambas comparaciones estudiadas, SCC vs control y ADC vs control, y nos permite una visión completa de la distribución de metabolitos según el subtipo de cáncer. Este gráfico se combinó con el uso de agrupaciones jerárquicas, agrupando las variables mediante un método no supervisado (Figura 11).

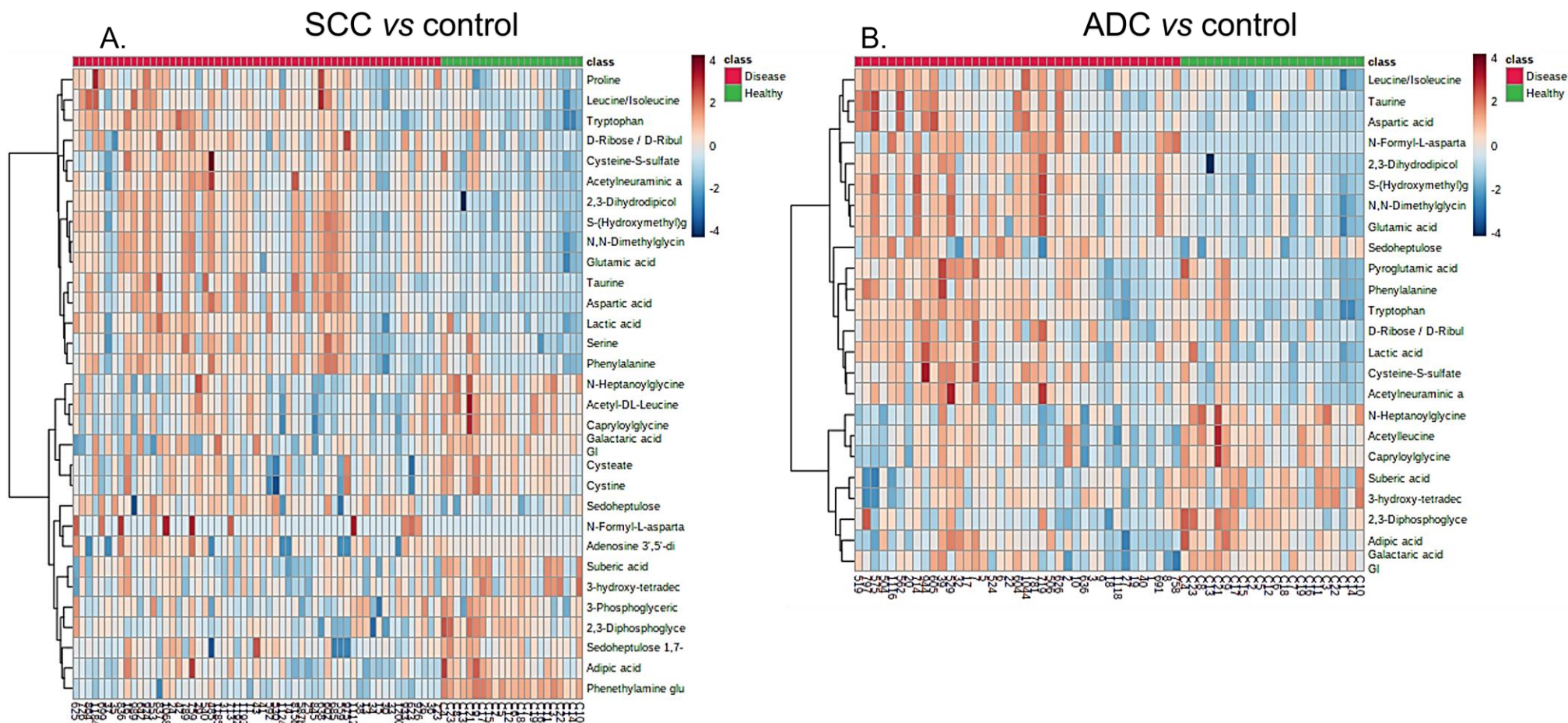


Figura 11. Mapa de calor o “heatmap” de los metabolitos significativos detectados mediante el análisis de muestras de plasma de NSCLC a través del método optimizado para CE-MS aniónico. **A.** Resultados del estudio comparativo SCC versus control y **B.** ADC versus control. El espectro de color que va del rojo al azul indica el rango de intensidades de señales altas a bajas, respectivamente, para cada metabolito.

4. Discusión y conclusión

A la vista de los resultados obtenidos, se ha podido concluir la utilización del capilar de PVA para el análisis aniónico por CE-MS. La optimización y comparación de dos capilares (PVA y sílice fundida) utilizando dos electrolitos de fondo (BGE), acetato de amonio 50 mM a pH 8,5 y ácido fórmico 0,1 M, ha permitido obtener un método que reúne las mejores condiciones de separación y detección para este tipo. de metabolitos.

La reprogramación metabólica, incluida el consumo y metabolismo de los aminoácidos y las alteraciones del metabolismo del carbono central han sido ampliamente observados en cáncer (16, 17). Como esos son metabolitos identificados en el plasma usando el método CE-MS optimizado para compuestos aniónicos, decidimos probar la utilidad clínica de este método aplicándolo al estudio de muestras de plasma de pacientes con cáncer de pulmón, más concretamente, cáncer de pulmón de células no pequeñas (NSCLC). El cáncer de pulmón es la principal causa de mortalidad relacionada con el cáncer en el mundo, y el cáncer de pulmón de células no pequeñas (NSCLC) representa aproximadamente el 85% de los casos de cáncer de pulmón (18).

Una base de datos del metaboloma del cáncer de pulmón (LCMD, http://cosbi4.ee.ncku.edu.tw/z70180/lung_home), ha sido publicada recientemente lo que constituye una valiosa fuente de información sobre los metabolitos relacionados con el cáncer de pulmón (19). La información recopilada en esta base de datos muestra que los datos sobre la regulación de metabolitos específicos relacionados con NSCLC no siempre son consistentes. Esto puede deberse a la diversidad de subtipos de NSCLC y/o los diferentes estadios de la enfermedad estudiada o del grupo control seleccionado (20-22). De manera similar a nuestros resultados, varios estudios también informaron niveles elevados de isoleucina/leucina, fenilalanina, ácido aspártico, serina, prolina, triptófano o ácido glutámico en pacientes con cáncer de pulmón (23-27). La regulación al alza de aminoácidos como leucina/isoleucina o fenilalanina está asociado con mayores demandas metabólicas de crecimiento y proliferación tumoral (28). La leucina, al igual que otros aminoácidos de cadena ramificada (BCAA), se puede utilizar como una fuente alternativa de moléculas orgánicas para alimentar el ciclo TCA (17). Del mismo modo, el aspartato es un precursor de varios intermedios del ciclo TCA y se consume activamente en la proliferación de tejidos tumorales de pulmón (29), mientras que el ácido glutámico puede ser una fuente de nitrógeno y carbono en reacciones anapleróticas del ciclo TCA (30). La rápida proliferación tumoral también depende de la disponibilidad de serina extracelular, que puede ser importada a la célula por el

transportador SLC1A4 dependiente de Na⁺, que está sobreexpresado en el cáncer de pulmón (16). La prolina desempeña un papel importante en el desarrollo y la progresión del cáncer, y su disponibilidad influye en la síntesis y maduración del colágeno, lo que puede influir en el crecimiento, la plasticidad y la heterogeneidad de las células cancerosas (31). Cascino y *col.* (27), de manera similar a nuestros resultados, observaron un aumento en los niveles plasmáticos de triptófano en pacientes con cáncer de pulmón. Se ha sugerido que el nivel de triptófano puede estar asociado con el desarrollo de resistencia a los medicamentos en las células de cáncer de pulmón (32). Se han podido encontrar datos inconsistentes con respecto a los cambios en los niveles séricos de taurina en relación con NSCLC. Algunos estudios reportaron una disminución del nivel de taurina en pacientes con NSCLC (33-35), mientras que no se encontraron diferencias significativas en el estudio de Klupczynska y *col.* (36), y en nuestros resultados se encontró un aumento de taurina en el plasma de pacientes con NSCLC. Curiosamente, el nivel de taurina puede diferir entre los pacientes con NSCLC, ya que se encontró que los pacientes sensibles a la quimioterapia combinada basada en platino tienen niveles más bajos de taurina que los no-sensibles (34). No solo hemos encontrado diferencias en los aminoácidos estándar, sino que hemos observado diferencias en los derivados de aminoácidos, p. ej. N-formil-L-aspartato, cisteína-S-sulfato, N-heptanoilglicina o acetilleucina, siendo N-formil-L-aspartato detectado solo en las muestras de pacientes con NSCLC. Este es el primer estudio en el que se informa de cambios en el nivel de estos metabolitos en relación con el cáncer de pulmón, y la información sobre otros tipos de cáncer en relación con estos derivados es, también escasa. Se ha reportado un aumento del nivel de N-heptanoilglicina en la orina de pacientes con cáncer colorrectal (37), mientras que la acetilleucina se asoció con el cáncer de próstata (38).

Otro grupo de compuestos que discrimina significativamente a los pacientes con NSCLC de los controles son los ácidos dicarboxílicos (DA). Cuatro DA (adípico, pimérico, subérico, galactárico/glucárico) se encontraron disminuidos en el plasma de pacientes con NSCLC. Por un lado, la oxidación beta de DA de cadena impar da como resultado malonil-CoA, que se utiliza posteriormente para la síntesis de lípidos de novo, mientras que los DA de cadena par da lugar a succinil-CoA, que se utiliza en el ciclo del TCA (39). Por lo tanto, un nivel disminuido de DA puede deberse a su utilización para el crecimiento de células cancerosas, ya que el transportador de dicarboxilato mitocondrial se encuentra sobreexpresado en muchos cánceres (40). Hasta la fecha, ha sido

reportado la reducción del nivel plasmático de ácido dicarboxílico 28:4 en pacientes con cáncer colorrectal (41).

Entre otros metabolitos significativos, se pueden mencionar el piroglutámico, en la comparativa ADC vs Control, y el ácido láctico (ambos aumentados en pacientes con NSCLC en este estudio). La mayoría de los estudios recopilados en LCMD informaron de un aumento del ácido láctico y una disminución del nivel de ácido piroglutámico en pacientes con cáncer de pulmón. El exceso de ácido láctico es fundamentalmente responsable de la acidosis metabólica, que se encuentra comúnmente en los tumores denominados sólidos. Incluso en condiciones altamente aeróbicas, la principal fuente de energía para las células cancerosas es la glucosa, siendo el ácido láctico el producto final. Este fenómeno, descubierto por Otto Warburg, se conoce como efecto Warburg (42). Por otra parte, se ha observado una disminución del nivel y la actividad de 5-oxoprolinasa, una enzima que cataliza la conversión de ácido piroglutámico en glutamato, en tejido tumoral de cáncer de pulmón (43), lo que podría dar una explicación a nuestros niveles aumentado de piroglutamato.

Este trabajo recoge la información clave y las conclusiones alcanzadas tras años de investigación para el desarrollo e implementación de un método para el análisis de compuestos aniónicos mediante CE-MS. La implementación de este método para el análisis de un gran número de estándares ha permitido construir una base de datos de metabolitos aniónicos proporcionando un incremento de información sobre los metabolitos involucrados en gran parte de los procesos relacionados con muchas patologías, entre ellas, el cáncer. El método desarrollado se utilizó para analizar muestras de plasma recolectadas de pacientes con NSCLC y un grupo de control. Los resultados obtenidos indican que los grupos estudiados pueden discriminarse en función del perfil metabólico aniónico. No solo hemos confirmado cambios en varios metabolitos previamente informados (p. ej., aminoácidos o ácido láctico), sino que también hemos encontrado nuevos metabolitos relacionados con el cáncer que no se conocían previamente (p. ej., derivados de aminoácidos y DA). Especialmente interesantes son los resultados obtenidos para N-formil-L-aspartato, que se detectó solo en las muestras recolectadas de pacientes con NSCLC. Los nuevos metabolitos relacionados con el cáncer descubiertos con el método desarrollado requieren una mayor validación en conjuntos de muestras independientes y con el uso de un método analítico específico, pero esto está fuera del objetivo de este trabajo.

Tabla 5. Metabolitos estadísticamente significativos para ambas comparaciones realizadas (SCC (n = 57) vs control ((n = 22) y ADC (n = 39) vs control). Cada anotación se caracterizó por los niveles de confianza según la fuente de identificación: L3, compuestos anotados por la información proporcionada por MS1 a través de CMM; L2*, compuestos anotados por m/z y tiempo de migración relativo del análisis de los estándares incluido en nuestra base de datos interna; L2, identificación de compuestos que también incluyen fragmentación en fuente. Diferentes parámetros se han utilizado para identificar todos los metabolitos: error de masa, expresado en ppm, MT, tiempo de migración obtenido a través del análisis aniónico CE-MS en minutos, tiempo de migración relativo rMT sobre metionina sulfona (MetS) y ácido 2-(N-morfolino)-etanosulfónico (MES); HMDB, código de la base de datos del metaboloma humano; Nivel, nivel de confianza de la anotación; RSD, desviación estándar relativa de los metabolitos en las muestras de control de calidad; % de cambio, (Promedio para el grupo de casos - Promedio para el grupo sin cáncer) / Promedio del grupo sin cáncer, en %; valor p, obtenido después de la prueba t de Student; JK, intervalo de confianza Jackknife y selección de variables usando la combinación de VIP y p(corr) ($|p(\text{corr})| > 0.5$ y $VIP \geq 1$).

SCC vs control																
Metabolito	Formula	Masa Molecular	Error (ppm)	Aducto	MT (min)	rMT MetS	rMT MES	HMDB	Nivel	Fuente de anotación	RSD	% cambio	p valor	JK	p(corr)	VIP
Lactic acid	C3H6O3	90.0317	9	M-H	14.72	0.78	0.97	HMDB0000190	L2*	in-house DB	5	21	3.57E-04	JK		7.23
N,N-Dimethylglycine/N-methylalanine	C4H9NO2	103.0633	1	M-H	20.69	1.10	1.36	HMDB0000092	L2*	in-house DB	5	158	2.15E-08	JK		
Serine	C3H7NO3	105.0426	0	M-H	22.45	1.19	1.48	HMDB0003406	L2	in-house DB	5	25	0.033			
Proline	C5H9NO2	115.0633	0	M-H	19.73	1.05	1.30	HMDB0003411	L2*	in-house DB	8	24	0.021	JK		
Taurine	C2H7NO3S	125.0147	6	M-H	15.25	0.81	1.00	HMDB0000251	L2	in-house DB	3	88	3.20E-03	JK		2.05
Leucine/Isoleucine/ Norleucine	C6H13NO2	131.0946	2	M-H	22.67	1.21	1.49	HMDB0000687	L2*	in-house DB	8	36	1.60E-05	JK		1.52
Aspartic acid	C4H7NO4	133.0375	3	M-H	18.91	1.01	1.24	HMDB0006483	L2	in-house DB	5	591	2.24E-06	JK		1.17
Adipic acid	C6H10O4	146.0579	13	M-H	15.15	0.81	1.00	HMDB0000448	L2*	in-house DB	20	-32	2.49E-04			
Glutamic acid	C5H9NO4	147.0532	9	M-H	20.70	1.10	1.36	HMDB0003339	L2	in-house DB	6	182	7.26E-08	JK	0.55	2.30
D-Ribose / D-Ribulose / D-Arabinose / D-Xylulose / D-Lyxose	C5H10O5	150.0528	11	M-H+FA	14.65	0.78	0.96	HMDB0000621	L3	CMM	5	27	0.0020	JK		
N-Formyl-L-aspartate	C5H7NO5	161.0324	5	M-H	12.59	0.67	0.83	HMDB0060495	L3	in-house DB	7	SCC group	NA			
Phenylalanine	C9H11NO2	165.0790	7	M-H	20.56	1.09	1.35	HMDB0000159	L2*	in-house DB	6	24	0.010	JK		

Cysteate	C3H7NO5S	169.0045	0	M-H	19.60	1.04	1.29	HMDB0002757	L3	CMM	7	-23	3.74E-03	JK		
2,3-Dihydrodipicolinate	C7H7NO4	169.0375	5	M-H	20.66	1.10	1.36	HMDB0012247	L3	CMM	4	161	1.99E-07	JK	0.57	1.14
Acetylucine	C8H15NO3	173.1052	6	M-H	15.17	0.81	1.00	HMDB0011756	L2*	in-house DB	4	-29	1.13E-04	JK		
Suberic acid	C8H14O4	174.0892	6	M-H+FA	15.16	0.81	1.00	HMDB0000893	L2*	in-house DB	4	-39	2.76E-04	JK		
3-Phosphoglyceric acid 2-Phosphoglyceric acid	C3H7O7P	185.9929	7	M-H	9.75	0.52	0.64	HMDB0000807 HMDB0003391	L2	in-house DB	5	-31	0.021			
N-Heptanoylglycine	C9H17NO3	187.1208	5	M-H	15.20	0.81	1.00	HMDB0013010	L2*	in-house DB	4	-38	8.19E-05	JK	-0.54	1.53
Cysteine-S-sulfate	C3H7NO5S2	200.9766	3	M-H	10.52	0.56	0.69	HMDB0000731	L3	CMM	4	36	0.0047			
Capryloylglycine	C10H19NO3	201.1365	4	M-H	15.21	0.81	1.00	HMDB0000832	L3	CMM	4	-26	1.97E-04	JK		
Tryptophan	C11H12N2O2	204.0899	0	M-H	21.57	1.15	1.42	HMDB0000929	L2*	in-house DB	5	37	1.77E-04	JK		
Galactaric acid Glucaric acid	C6H10O8	210.0376	3	M-H	10.76	0.57	0.71	HMDB0000639 HMDB0000663	L3	CMM	6	-23	0.0006			
Sedoheptulose	C7H14O7	210.0740	10	M-H	15.40	0.82	1.01	HMDB0003219	L3	CMM	5	22	0.0028	JK		
Cystine	C6H12N2O4S2	240.0238	7	M-H	19.60	1.04	1.29	HMDB0000192	L2*	in-house DB	4	-23	1.03E-02			
2,3-Diphosphoglyceric acid	C3H8O10P2	265.9593	7	M-H	8.37	0.44	0.55	HMDB0001294	L2*	in-house DB	5	-60	8.02E-06	JK		
3-hydroxy-tetradecanedioic acid	C14H26O5	274.1780	4	M-H+FA	15.19	0.81	1.00	HMDB0000394	L3	CMM	4	-37	0.015	JK	-0.57	2.16
Phenethylamine glucuronide	C14H19NO6	297.1212	7	M-H	10.35	0.55	0.68	HMDB0010323	L3	CMM	4	-69	1.08E-19	JK		
Acetylneuraminic acid	C11H19NO9	309.1060	4	M-H	13.41	0.71	0.88	HMDB0000230	L2*	in-house DB	8	46	3.71E-04	JK		
S-(Hydroxymethyl)glutathione	C11H19N3O7S	337.0944	11	M-H	20.77	1.10	1.37	HMDB0004662	L3	CMM	6	77	7.84E-08	JK		
Sedoheptulose 1,7-bisphosphate	C7H16O13P2	370.0066	7	M-H	9.73	0.52	0.64	HMDB0060274	L3	CMM	5	-20	1.13E-02	JK		
Adenosine 3',5'-diphosphate	C10H15N5O10P2	427.0294	3	M-H	11.24	0.60	0.74	HMDB0000061	L2*	in-house DB	7	81	3.84E-02	JK		

ADC vs control

Metabolito	Formula	Masa Molecular	Error (ppm)	Aducto	MT (min)	rMT MetS	rMT MES	HMDB	Nivel	Fuente de anotación	RSD	% cambio	p valor	JK	p(corr)	VIP
Lactic acid	C3H6O3	90.0317	9	M-H	14.72	0.78	0.97	HMDB0001311	L2*	in-house DB	5	21	0.0007	JK		1.01
N,N-Dimethylglycine/N-methylalanine	C4H9NO2	103.0633	1	M-H	20.69	1.10	1.36	HMDB0000092	L2*	in-house DB	5	146	1.74E-06	JK	0.59	2.21

Taurine	C2H7NO3S	125.0147	6	M-H	15.25	0.81	1.00	HMDB0000251	L2	in-house DB	3	70	0.0020	JK		1.42
Pyroglutamic acid	C5H7NO3	129.0426	2	M-H	13.95	0.74	0.92	HMDB0000267	L2*	in-house DB	4	34	0.023	JK		1.42
Leucine/Isoleucine/ Norleucine	C6H13NO2	131.0946	2	M-H	22.67	1.21	1.49	HMDB0000687	L2*	in-house DB	8	31	6.09E-06	JK	0.54	1.26
Aspartic acid	C4H7NO4	133.0375	3	M-H	18.91	1.01	1.24	HMDB0006483	L2	in-house DB	5	367	1.48E-06	JK	0.53	2.43
Adipic acid	C6H10O4	146.0579	13	M-H	15.15	0.81	1.00	HMDB0000448	L2*	in-house DB	20	-24	0.015	JK		1.25
Glutamic acid	C5H9NO4	147.0532	9	M-H	20.70	1.10	1.36	HMDB0003339	L2*	in-house DB	6	165	4.49E-06	JK	0.57	2.28
D-Ribose / D-Ribulose / D-Arabinose / D-Xylulose / D-Lyxose	C5H10O5	150.0528	11	M+FA-H	14.65	0.78	0.96	HMDB0000621	L3	CMM	5	32	0.0014	JK		1.06
N-Formyl-L-aspartate	C5H7NO5	161.0324	5	M-H	12.59	0.67	0.83	HMDB0060495	L2*	in-house DB	7	ADC group	NA			
Phenylalanine	C9H11NO2	165.0790	7	M-H	20.56	1.09	1.35	HMDB0000159	L2	in-house DB	6	24	0.0092	JK		1.03
2,3-Dihydrodipicolinate	C7H7NO4	169.0375	5	M-H	20.66	1.10	1.36	HMDB0012247	L3	CMM	4	158	2.69E-07			
Acetylleucine	C8H15NO3	173.1052	6	M-H	15.17	0.81	1.00	HMDB0011756	L2*	in-house DB	4	-28	0.0009	JK		1.36
Suberic acid	C8H14O4	174.0892	6	M+FA-H	15.16	0.81	1.00	HMDB0000893	L2*	in-house DB	4	-31	0.0070	JK		1.20
N-Heptanoylglycine	C9H17NO3	187.1208	5	M-H	15.20	0.81	1.00	HMDB0013010	L2*	in-house DB	4	-40	4.42E-05	JK	-0.54	1.67
Cysteine-S-sulfate	C3H7NO5S2	200.9766	3	M-H	10.52	0.56	0.69	HMDB0000731	L3	CMM	4	73	0.0011	JK		1.55
Capryloylglycine	C10H19NO3	201.1365	4	M-H	15.21	0.81	1.00	HMDB0000832	L3	CMM	4	-26	0.0005	JK		1.32
Tryptophan	C11H12N2O2	204.0899	0	M-H	21.57	1.15	1.42	HMDB0000929	L2	in-house DB	5	42	2.27E-04	JK		1.38
Galactaric acid Glucaric acid	C6H10O8	210.0376	3	M-H	10.76	0.57	0.71	HMDB0000639 HMDB0000663	L3	CMM	6	-28	0.0010	JK		1.55
Sedoheptulose	C7H14O7	210.0740	10	M-H	15.40	0.82	1.01	HMDB0003219	L3	CMM	5	23	0.0013			
2,3-Diphosphoglyceric acid	C3H8O10P2	265.9593	7	M-H	8.37	0.44	0.55	HMDB0001294	L2*	in-house DB	5	-45	0.0017	JK		1.66
3-hydroxy-tetradecanedioic acid	C14H26O5	274.1780	4	M-H+FA	15.19	0.81	1.00	HMDB0000394	L3	CMM	4	-32	0.0248			1.10
Acetylneuraminic acid	C11H19NO9	309.1060	4	M-H	13.41	0.71	0.88	HMDB0000230	L2*	in-house DB	8	66	0.0006	JK		1.49
S-(Hydroxymethyl)glutathione	C11H19N3O7S	337.0944	11	M-H	20.77	1.10	1.37	HMDB0004662	L3	CMM	6	67	3.70E-06	JK	0.51	1.58

SCC, carcinoma de células escamosas; ADC, adenocarcinoma; MT, tiempo de migración (min); rMT, tiempo de migración relativo; HMDB, *The Human Metabolome Database* (<https://hmdb.ca/>); código; DB, base de datos; NA, no aplica; CMM, *CEU Mass Mediator database* (http://ceumass.eps.uspceu.es/index_cesearch.xhtml); RSD, desviación estándar relativa; JK, Jackknife.

5. Referencias

1. González-Riano C, Dudzik D, Garcia A, Gil-de-la-Fuente A, Gradillas A, Godzien J, et al. Recent Developments along the Analytical Process for Metabolomics Workflows. *Anal Chem.* 2020;92(1):203-26.
2. Hirayama A, Tabata S, Kudo R, Hasebe M, Suzuki K, Tomita M, et al. The use of a double coaxial electrospray ionization sprayer improves the peak resolutions of anionic metabolites in capillary ion chromatography-mass spectrometry. *J Chromatogr A.* 2020;1619:460914.
3. Zhang W, Hankemeier T, Ramautar R. Next-generation capillary electrophoresis-mass spectrometry approaches in metabolomics. *Curr Opin Biotechnol.* 2017;43:1-7.
4. Zhang W, Ramautar R. CE-MS for metabolomics: Developments and applications in the period 2018-2020. *Electrophoresis.* 2021;42(4):381-401.
5. Soga T, Igarashi K, Ito C, Mizobuchi K, Zimmermann HP, Tomita M. Metabolomic profiling of anionic metabolites by capillary electrophoresis mass spectrometry. *Anal Chem.* 2009;81(15):6165-74.
6. Soga T, Ueno Y, Naraoka H, Matsuda K, Tomita M, Nishioka T. Pressure-assisted capillary electrophoresis electrospray ionization mass spectrometry for analysis of multivalent anions. *Anal Chem.* 2002;74(24):6224-9.
7. Godzien J, Armitage EG, Angulo S, Martinez-Alcazar MP, Alonso-Herranz V, Otero A, et al. In-source fragmentation and correlation analysis as tools for metabolite identification exemplified with CE-TOF untargeted metabolomics. *Electrophoresis.* 2015;36(18):2188-95.
8. Mamani-Huanca M, Gradillas A, Gil de la Fuente A, López-González Á, Barbas C. Unveiling the Fragmentation Mechanisms of Modified Amino Acids as the Key for Their Targeted Identification. *Analytical Chemistry.* 2020; 92(7):4848–4857
9. Dudzik D, Barbas-Bernardos C, García A, Barbas C. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. *Journal of pharmaceutical and biomedical analysis.* 2017;147:149-73.
10. Kuligowski J, Sanchez-Illana A, Sanjuan-Herraez D, Vento M, Quintas G. Intra-batch effect correction in liquid chromatography-mass spectrometry using quality control samples and support vector regression (QC-SVRC). *Analyst.* 2015;140(22):7810-7.
11. Wheelock AM, Wheelock CE. Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Molecular BioSystems.* 2013;9(11):2589-96.
12. Barnes RJ, Dhanoa MS, Lister SJ. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy.* 1989;43(5):772-7.
13. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research.* 2021;49(W1):W388-W96.
14. Lauer HH, Rozing GP. *High Performance Capillary Electrophoresis* 2010.
15. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *Journal of The American Society for Mass Spectrometry.* 2016;27(12):1897-905.

16. Wei Z, Liu X, Cheng C, Yu W, Yi P. Metabolism of Amino Acids in Cancer. *Frontiers in Cell and Developmental Biology*. 2021;8:603837.
17. Lieu EL, Nguyen T, Rhyne S, Kim J. Amino acids in cancer. *Experimental & Molecular Medicine*. 2020;52(1):15-30.
18. Wang M, Herbst RS, Boshoff C. Toward personalized treatment approaches for non-small-cell lung cancer. *Nature Medicine*. 2021;27(8):1345-56.
19. Wu W-S, Wu H-Y, Wang P-H, Chen T-Y, Chen K-R, Chang C-W, et al. LCMD: Lung Cancer Metabolome Database. *Computational and Structural Biotechnology Journal*. 2022;20:65-78.
20. Hori S, Nishiumi S, Kobayashi K, Shinohara M, Hatakeyama Y, Kotani Y, et al. A metabolomic approach to lung cancer. *Lung Cancer*. 2011;74(2):284-92.
21. Klupczynska A, Plewa S, Dyszkiewicz W, Kasprzyk M, Sytek N, Kokot ZJ. Determination of low-molecular-weight organic acids in non-small cell lung cancer with a new liquid chromatography-tandem mass spectrometry method. *J Pharm Biomed Anal*. 2016;129:299-309.
22. Kowalczyk T, Kisluk J, Pietrowska K, Godzien J, Kozłowski M, Reszeć J, et al. The Ability of Metabolomics to Discriminate Non-Small-Cell Lung Cancer Subtypes Depends on the Stage of the Disease and the Type of Material Studied. *Cancers (Basel)*. 2021;13(13):3314.
23. Puchades-Carrasco L, Jantus-Lewintre E, Pérez-Rambla C, García-García F, Lucas R, Calabuig S, et al. Serum metabolomic profiling facilitates the non-invasive identification of metabolic biomarkers associated with the onset and progression of non-small cell lung cancer. *Oncotarget*. 2016;7(11):12904-16.
24. Louis E, Adriaensens P, Guedens W, Vanhove K, Vandeurzen K, Darquennes K, et al. Metabolic phenotyping of human blood plasma: a powerful tool to discriminate between cancer types? *Ann Oncol*. 2016;27(1):178-84.
25. Maeda J, Higashiyama M, Imaizumi A, Nakayama T, Yamamoto H, Daimon T, et al. Possibility of multivariate function composed of plasma amino acid profiles as a novel screening index for non-small cell lung cancer: a case control study. *BMC Cancer*. 2010;10:690.
26. Miyamoto S, Taylor SL, Barupal DK, Taguchi A, Wohlgemuth G, Wikoff WR, et al. Systemic Metabolomic Changes in Blood Samples of Lung Cancer Patients Identified by Gas Chromatography Time-of-Flight Mass Spectrometry. *Metabolites*. 2015;5(2):192-210.
27. Cascino A, Muscaritoli M, Cangiano C, Conversano L, Laviano A, Ariemma S, et al. Plasma amino acid imbalance in patients with lung and breast cancer. *Anticancer research*. 1995;15(2):507-10.
28. Bamji-Stocke S, van Berkel V, Miller DM, Frieboes HB. A review of metabolism-associated biomarkers in lung cancer diagnosis and treatment. *Metabolomics*. 2018;14(6):81.
29. Kami K, Fujimori T, Sato H, Sato M, Yamamoto H, Ohashi Y, et al. Metabolomic profiling of lung and prostate tumor tissues by capillary electrophoresis time-of-flight mass spectrometry. *Metabolomics*. 2013;9(2):444-53.
30. Kodama M, Oshikawa K, Shimizu H, Yoshioka S, Takahashi M, Izumi Y, et al. A shift in glutamine nitrogen metabolism contributes to the malignant progression of cancer. *Nature Communications*. 2020;11(1):1320.

31. D'Aniello C, Patriarca EJ, Phang JM, Minchiotti G. Proline Metabolism in Tumor Growth and Metastatic Progression. *Frontiers in Oncology*. 2020;10:776.
32. Li C, Zhao H. Tryptophan and Its Metabolites in Lung Cancer: Basic Functions and Clinical Significance. *Frontiers in Oncology*. 2021;11:707277.
33. Hu J-M, Sun H-T. Serum proton NMR metabolomics analysis of human lung cancer following microwave ablation. *Radiation Oncology*. 2018;13(1):40.
34. Xu S, Zhou Y, Geng H, Song D, Tang J, Zhu X, et al. Serum Metabolic Profile Alteration Reveals Response to Platinum-Based Combination Chemotherapy for Lung Cancer: Sensitive Patients Distinguished from Insensitive ones. *Scientific Reports*. 2017;7(1):17524.
35. Ruiying C, Zeyun L, Yongliang Y, Zijia Z, Ji Z, Xin T, et al. A comprehensive analysis of metabolomics and transcriptomics in non-small cell lung cancer. *PLoS One*. 2020;15(5):e0232272.
36. Klupczynska A, Dereziński P, Dyszkiewicz W, Pawlak K, Kasprzyk M, Kokot ZJ. Evaluation of serum amino acid profiles' utility in non-small cell lung cancer detection in Polish population. *Lung Cancer*. 2016;100:71-6.
37. Deng Y, Yao H, Chen W, Wei H, Li X, Zhang F, et al. Profiling of polar urine metabolite extracts from Chinese colorectal cancer patients to screen for potential diagnostic and adverse-effect biomarkers. *Journal of Cancer*. 2020;11(23):6925-38.
38. Mondul AM, Moore SC, Weinstein SJ, Karoly ED, Sampson JN, Albanes D. Metabolomic analysis of prostate cancer risk in a prospective cohort: The alpha-tocopherol, beta-carotene cancer prevention (ATBC) study. *Int J Cancer*. 2015;137(9):2124-32.
39. Grego AV, Mingrone G. Dicarboxylic acids, an alternate fuel substrate in parenteral nutrition: an update. *Clin Nutr*. 1995;14(3):143-8.
40. Lytovchenko O, Kunji ERS. Expression and putative role of mitochondrial transport proteins in cancer. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 2017;1858(8):641-54.
41. Wood PL, Donohue MM, Cebak JE, Beckmann TG, Messias MCF, Credidio L, et al. Reduced Plasma Levels of Very-Long-Chain Dicarboxylic Acid 28:4 in Italian and Brazilian Colorectal Cancer Patient Cohorts. *Metabolites*. 2018;8(4):91.
42. Choi SY, Collins CC, Gout PW, Wang Y. Cancer-generated lactic acid: a regulatory, immunosuppressive metabolite? *J Pathol*. 2013;230(4):350-5.
43. Chen X, Schechter RL, Griffith OW, Hayward MA, Alpert LC, Batist G. Characterization of 5-oxo-L-prolinase in normal and tumor tissues of humans and rats: a potential new target for biochemical modulation of glutathione. *Clin Cancer Res*. 1998;4(1):131-8.

CONCLUSIONES

La búsqueda de biomarcadores a nivel de metabolitos en cáncer tiene como objetivo identificar un *target*, rastrear el desarrollo y la progresión de la enfermedad o su respuesta a la terapia. Para llevar a cabo este objetivo es clave conocer las alteraciones metabólicas que se producen según el tumor ante el que nos encontremos. A este respecto, la metabolómica nos ofrece una herramienta particularmente útil para la búsqueda de biomarcadores en cáncer debido a su carácter intensivo, estudiando los importantes cambios a nivel de metabolitos que se producen por alteraciones menores en los niveles de genes o proteínas. El destino de los biomarcadores de cáncer sería en última instancia su aplicación en la práctica clínica, por lo que, el análisis de su utilidad y confiabilidad para identificar la patología es crucial para su último fin: marcar una diferencia real en la salud del paciente.

A través de las estrategias de análisis desarrolladas y empleadas, esta tesis doctoral ha logrado sus principales objetivos. En primer lugar, el desarrollo, optimización y aplicación de diferentes herramientas analíticas para el estudio metabolómico de tres tipos distintos de cáncer (tumores de origen neuroendocrino, colangiocarcinoma y cáncer de pulmón de células no pequeñas) empleando las matrices biológicas de plasma, bilis y tejido a través de las plataformas CE-MS, LC-MS y GC-MS. En consecuencia, con lo anterior, la implementación de estos métodos ha posibilitado dilucidar los mecanismos biológicos involucrados en la iniciación y progresión de los procesos tumorales y, el establecimiento de posibles biomarcadores candidatos.

Además, es posible destacar importantes aspectos que complementan la contribución de esta investigación al campo de la metabolómica:

- Una exhaustiva revisión bibliográfica de los estudios metabolómicos sobre cáncer, que incluyen cierto grado de validación, publicados en la última década, ha permitido obtener el mapa de los potenciales biomarcadores propuestos a nivel de metabolitos en cada tipo de cáncer, así como la muestra utilizada y las herramientas analíticas empleadas. Esta ha sido la referencia para conocer el estado del arte (*state-of-art*) a lo largo de la presente tesis.
- La aplicación de una estrategia multiplataforma, que en algunos casos ha supuesto el desarrollo de una metodología *ad hoc* para cierto tipo de muestras como las bilis, proporcionando información clave sobre las vías metabólicas alteradas en los distintos tipos de cáncer estudiados: tumores de origen neuroendocrino, colangiocarcinoma y cáncer de pulmón de células no pequeñas.

- La validación de los resultados de estudios multiplataforma no dirigidos en tumores de origen neuroendocrino ha dado lugar a un conjunto de biomarcadores que pueden distinguir el estado patológico de la población. Este es un avance importante para establecer estrategias de práctica clínica que mejoren la capacidad de diagnosticar la enfermedad y establecer su pronóstico.
- La implantación de un flujo de trabajo optimizado para la anotación de modificaciones oxidadas de lípidos ha permitido establecer una herramienta que soluciona uno de los principales obstáculos en el empleo de la metabolómica no dirigida para la búsqueda de biomarcadores, la identificación de compuestos sin correspondencia resultante en las bases de datos disponibles gracias a la coincidencia espectral, elucidación estructural y cotejo con sus correspondientes formas nativas. Esta estrategia ha sido aplicada satisfactoriamente a los estudios con LC-MS de cáncer llevados a cabo, permitiéndonos identificar distintos productos de oxidación. Asimismo, el método planteado para el análisis a través de LC-MS de los distintos productos de oxidación obtenidos por oxidación *in vitro* de estándares comerciales ha hecho posible la creación de una base de datos para la caracterización de lípidos oxidados por su tiempo de retención y fragmentos característicos procedentes del análisis MS/MS.
- La búsqueda dirigida de productos de oxidación de las principales glicerofosfolinas ha mostrado un patrón preliminar discriminativo de cáncer de pulmón entre las formas predominantemente alteradas en plasma y tejido. Con esto ha sido posible proponer un posible panel candidato biomarcador para el diagnóstico de SCC o ADC *versus* control en ambos tipos de muestra. Conjuntamente se pudo concluir que la apoptosis celular estaría detrás de la disminución en la abundancia de las formas oxidadas a medida que la enfermedad va avanzando.
- La relación entre el cáncer, el estrés oxidativo y los procesos pre y proinflamatorios ha sido estudiada en detalle gracias a la identificación de un gran número de lípidos oxidados en los diferentes estudios oncológicos llevados a cabo. Estos hallazgos han sido relacionados con la implicación de los derivados oxidados del ácido araquidónico en la modulación de los factores de crecimiento en células cancerosas, efectos co-mitogénicos, angiogénesis, y la adhesión, migración e invasión de las células tumorales. Por su parte los productos de oxidación de lisofosfolípidos han sido involucrados en la regulación de procesos de inflamación, trombosis y, en general, alteración del funcionamiento de la barrera endotelial, ya que se modifica el empaquetamiento de acilo y, con ello, las propiedades físicas de fluidez.

- El desarrollo y optimización de una metodología para el análisis de compuestos aniónicos en CE-MS y la construcción de una base de datos de estos metabolitos ha proporcionado un aumento de información sobre los metabolitos implicados en gran parte de los procesos relacionados con el cáncer. Este método ha sido aplicado con éxito a muestras de plasma de cáncer de pulmón lo que ha permitido extender el amplio abanico metabolómico que el análisis multiplataforma no dirigido puede ofrecer. Con este método de análisis se han reportado, por primera vez, algunos derivados de aminoácidos y ácidos dicarboxílicos relacionados con el cáncer.

Como continuación del trabajo planteado en esta tesis doctoral y, como en cualquier otro proyecto de investigación, se encuentran líneas de investigación que permanecen abiertas y sobre las que merece continuar trabajando. Durante este trabajo de investigación surgieron varias líneas de futuro, algunas de ellas relacionadas de manera directa con los resultados obtenidos o los problemas encontrados, con el objetivo final de seguir avanzando en el conocimiento sobre la búsqueda de biomarcadores en cáncer o la aplicación a cualquier otra patología.

Entre los posibles futuros trabajos destacan aquellos que continúen con el desarrollo de la investigación aquí planteada, superando el alcance de esta tesis, y se sugieren desarrollos específicos que apoyen y mejoren los modelos, metodologías y resultados propuestos. Cabe destacar:

- Aplicación de sucesivas etapas de validación de los resultados obtenidos en los estudios de cáncer llevados a cabo. Esto incluye: validación del panel de biomarcadores planteado para plasma de tumores neuroendocrinos en una cohorte externa y de gran tamaño de pacientes; estudio de los resultados obtenidos para la discriminación de colangiocarcinoma en muestras de bilis que incluiría el establecimiento de un panel discriminante, validación de los resultados mediante otra técnica y en otra cohorte de muestras y, aplicación de los resultados en muestras de origen humano; validación de los resultados preliminares obtenidos para distintos tipos de muestra de cáncer de pulmón en relación tanto a las oxidaciones lipídicas encontradas como a los derivados de aminoácidos y ácidos dicarboxílicos de los análisis LC-MS y CE-MS aniónico, respectivamente.
- Ampliación y desarrollo de las bases de datos propuestas tanto para la caracterización de las múltiples formas oxidadas derivadas de lípidos como del

- análisis CE-MS aniónico con la metodología optimizada. Aplicación a estudios metabolómicos tanto en cáncer como en otras patologías.
- Aplicación de la espectrometría de movilidad iónica para separar e identificar especies isoméricas, principalmente en lípidos oxidados. Esta técnica analítica proporciona una dimensión adicional que permite la separación de iones en fase gaseosa bajo la influencia de un campo eléctrico. En definitiva, la velocidad resultante de los iones en el tubo de deriva dependerá de sus secciones transversales de colisión función de su tamaño, forma, carga eléctrica y masa. Con la información obtenida sería posible aumentar el nivel de identificación, lo que permitiría localizar la posición del derivado de oxidación ajustando de forma más precisa el mecanismo de oxidación que lo ha producido y las consecuencias a nivel biológico que eso supone.
 - *Single cell metabolomics* (SCM) o el estudio metabolómico a nivel celular ofrece una instantánea de las moléculas pequeñas, productos intermedios y los productos finales del metabolismo celular dentro de un sistema biológico. Esto permitiría abordar la heterogeneidad de los tejidos cancerígenos y estudiar los cambios a nivel celular. Por lo tanto, SCM sería especialmente útil en estudios de oncología, beneficiando en varias facetas de la investigación, incluida la evaluación de fármacos, el diagnóstico de enfermedades y el diseño de fármacos de precisión.

Actualmente se está concluyendo con la redacción de dos publicaciones científicas derivadas de los trabajos llevados a cabo resumidos en los capítulos 2 y 4, en colaboración ambos con el grupo del profesor Michał Ciborowski de la Universidad Médica de Białystok (Polonia). Además, se ha planteado la redacción de una revisión bibliográfica con el trabajo presentado en el capítulo “*Introducción General*”, como continuación de la revisión ya publicada en 2017, focalizando el tema exclusivamente en cáncer.

Lista de publicaciones científicas

Listado de publicaciones científicas derivadas de esta tesis:

1. **Ángeles López-López**, Ángeles López-González, Tomás Clive Barker-Tejeda & Coral Barbas (2018) A review of validated biomarkers obtained through metabolomics, *Expert Review of Molecular Diagnostics*, 18:6, 557-575, DOI: 10.1080/14737159.2018.1481391
 - Factor de impacto: 5.225 , Q1 (PATHOLOGY – SCIE)
 - Ranking: 17/77
2. Soldevilla, B*.; **López-López, A***.; Lens-Pardo, A.; Carretero-Puche, C.; Lopez-Gonzalez, A.; La Salvia, A.; Gil-Calderon, B.; Riesco-Martinez, M.C.; Espinosa-Olarte, P.; Sarmentero, J.; et al. Comprehensive Plasma Metabolomic Profile of Patients with Advanced Neuroendocrine Tumors (NETs). Diagnostic and Biological Relevance. *Cancers* 2021, 13, 2634
 - Factor de Impacto: 6.639, Q1 (ONCOLOGY - SCIE)
 - Ranking: 51/242
3. **Ángeles López-López**, Joanna Godzien, Beatriz Soldevilla, Ana Gradillas, Ángeles López-González, Alberto Lens-Pardo, Anna La Salvia, María del Carmen Riesco-Martínez, Rocío García-Carbonero, Coral Barbas, Oxidized lipids in the metabolic profiling of neuroendocrine tumors – Analytical challenges and biological implications, *Journal of Chromatography A*, Volume 1625, 2020, 461233
 - Factor de Impacto: 4.759, Q1 (CHEMISTRY, ANALYTICAL - SCIE)
 - Ranking: 13/78

Otras publicaciones científicas:

1. Tomás Clive Barker-Tejeda, Alma Villaseñor, Carolina Gonzalez-Riano, **Ángeles López-López**, Ana Gradillas, Coral Barbas, In vitro generation of oxidized standards for lipidomics. Application to major membrane lipid components, *Journal of Chromatography A*, Volume 1651, 2021, 462254
 - Factor de Impacto: 4.759, Q1 (CHEMISTRY, ANALYTICAL - SCIE)
 - Ranking: 13/78
2. Alma Villaseñor, Joanna Godzien, Tomás Clive Barker-Tejeda, Carolina Gonzalez-Riano, **Ángeles López-López**, Danuta Dudzik, Ana Gradillas, Coral Barbas, Analytical approaches for studying oxygenated lipids in the search of potential biomarkers by LC-MS, *TrAC Trends in Analytical Chemistry*, Volume 143, 2021, 116367
 - Factor de Impacto: 12.296, Q1 (CHEMISTRY, ANALYTICAL - SCIE)
 - Ranking: 1/87

