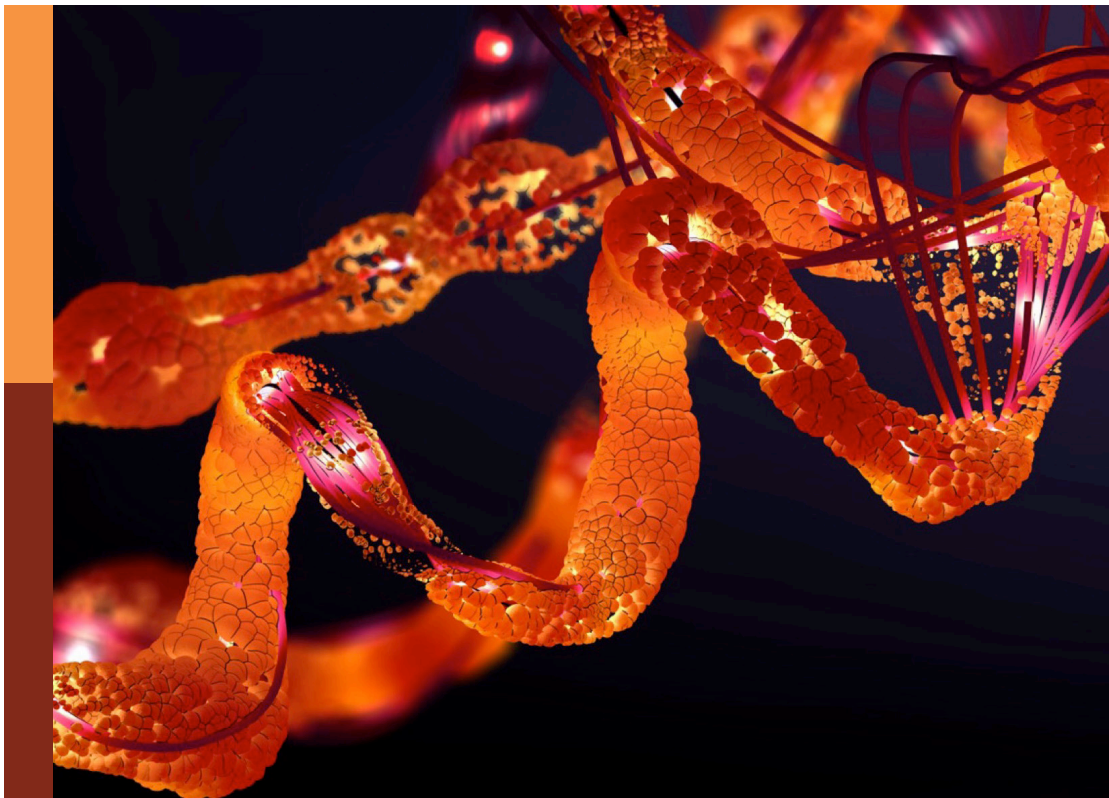
frontiers | Research Topics

The IV Latin American Metabolic Profiling Society (LAMPS) symposium 2022

Edited by
Guillermo Moyna, Pablo Hoijemberg and Martin Aran

Published in
Frontiers in Molecular Biosciences



ISSN 1664-8714
ISBN 978-2-8325-5104-2
DOI 10.3389/978-2-8325-5104-2

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

The IV Latin American Metabolic Profiling Society (LAMPS) symposium: 2022

Topic editors

Guillermo Moyna — Universidad de la República, Uruguay

Pablo Hoijemberg — Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

Martin Aran — IIBBA-CONICET Leloir Institute Foundation, Argentina

Citation

Moyna, G., Hoijemberg, P., Aran, M., eds. (2024). *The IV Latin American Metabolic Profiling Society (LAMPS) symposium: 2022*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-5104-2

Table of contents

frontiers | Frontiers in Molecular Biosciences

TYPE Editorial
PUBLISHED 21 June 2024
DOI 10.3389/fmolb.2024.1430789

OPEN ACCESS

EDITED AND REVIEWED BY
Ricardo Borges,
Federal University of Rio de Janeiro, Brazil

*CORRESPONDENCE
Martín Arán,
maran@leloir.org.ar
Pablo A. Hoijemberg,
pablo.hoijemberg@cibion.conicet.gov.ar
Guillermo Moyna,
gmoyna@fq.edu.uy

RECEIVED 10 May 2024
ACCEPTED 10 June 2024
PUBLISHED 21 June 2024

CITATION
Arán M, Hoijemberg PA and Moyna G (2024), Editorial: The IV Latin American Metabolic Profiling Society (LAMPS) symposium: 2022. *Front. Mol. Biosci.* 11:1430789. doi: 10.3389/fmolb.2024.1430789

Editorial: The IV Latin American Metabolic Profiling Society (LAMPS) symposium: 2022

Martín Arán[1]*, Pablo A. Hoijemberg[2]* and Guillermo Moyna[3]*

[1]Laboratorio de RMN, Fundación Instituto Leloir (IIBBA-CONICET), Buenos Aires, Argentina, [2]Centro de Investigaciones en Bionanociencias (CIBION-CONICET), Buenos Aires, Argentina, [3]Departamento de Química Del Litoral, CENUR Litoral Norte, Universidad de La República, Paysandú, Uruguay

KEYWORDS

Latin America, Latin American Metabolic Profiling Society, metabolomics, nuclear magnetic resonance, mass spectrometry

Editorial on the Research Topic
The IV Latin American Metabolic Profiling Society (LAMPS) symposium: 2022

In the past two decades, metabolomic analysis has evolved from being an important tool to becoming a standalone discipline. It has been applied to a wide variety of problems, including clinical diagnosis and biomarker identification, crop selection and protection, and food authentication, among others. Most of the advances in metabolomics have gone hand in hand with the standardization of methods and protocols, as well as with advances in analytical instrumentation and bioinformatics tools capable of handling large amounts of samples and data. The latter are inherently associated with considerable investments in resources, putting researchers in emerging regions such as Latin America at a clear disadvantage. However, the number of research groups, institutions, and dedicated facilities focusing on metabolomic analysis has increased modestly but steadily in the past years in Latin American countries. One of the catalysts of this growth has been the Latin American Metabolic Profiling Society (LAMPS), an organization founded initially by researchers in Argentina, Colombia, Brazil, Peru, and Uruguay in the mid 2010s. Since its early days, LAMPS has promoted collaborative research projects between members and fostered training opportunities for students and young investigators, and has also organized biennial meetings attended by researchers in the field from Latin America and the rest of the World. The first LAMPS meeting was held in Lima, Perú (2014), and gatherings in Rosario, Argentina (2016) and Rio de Janeiro, Brazil (2018) followed. The fourth LAMPS meeting planned for 2020 was postponed due to the COVID-19 pandemic, and was then celebrated in Cartagena de Indias, Colombia, in 2022. With 14 invited panelists and over 150 participants, this proved to be the largest LAMPS gathering so far.

To partly document this continuing growth, this Research Topic collects manuscripts from Latin American groups that participated in the IV LAMPS. As evidenced in the volume, the contributions describe the use of metabolomics analysis and related methods to tackle a wide gamut of problems which are briefly summarized here. For example, COVID-19 was the focus of two studies. In one of them, López-Hernandez et al. used untargeted MS-based metabolomics to investigate dysregulations in lipid pathways 2 years after recovery from the disease, adding important insights into our understanding of long COVID-19. In the other, the effects of SARS-Cov-2 on lung parenchyma were compared to those caused by

other severe pulmonary infections through analysis of the ^{1}H NMR profiles of tissue extracts, corroborating that distinct metabolic signatures associated with energy metabolism and inflammatory pathways differentiate COVID-19 from other respiratory infections (Hurtado et al.). Through a translational study supported largely by untargeted NMR-based metabolomic analysis and targeted GC-EI-MS data, a multinational team led by Argentine researchers reported the identification of nicotinamide as a potential biomarker for Alzheimer's disease (Dalmasso et al.). The application of untargeted metabolomics and lipidomics based on LC-QTOF-MS and GC-QTOF-MS found dysregulation of glycerolipid and sphingolipid metabolic pathways in the plasma of acute leukemia patients (Arévalo et al.). These differences are independent of lifestyle, race, or geographic location, providing valuable clues for the development of global therapies. NMR-based plasma metabolomics of individuals with differential responses to HIV-1 exposure and/or infection revealed that different pathways are affected in each group relative to controls (Gómez-Archila et al.). In particular, the study was the first to identify that HIV-1-exposed but seronegative (HESN) individuals have a specific metabolic fingerprint with significant alterations in LDL, glucose, lactate, and phosphocholine levels. Based on a systematic review of 26 independent metabolomics studies on systemic sclerosis, Morales-Gonzáles et al. identified 151 metabolites associated with the condition. Species linked to amino acid, lipid, and TCA cycle metabolic pathways are the most dysregulated. These confirm the impact of autoimmune inflammation, vascular damage, fibrosis, and gut dysbiosis in the progression of this disease, and also represent potential biomarkers for its early diagnosis and prognosis. The gut microbiome in a Colombian cohort of pregnant and lactating women was investigated by untargeted metabolomics based on LC-QTOF-MS coupled to molecular networking (Londoño-Osorio et al.). The report helps to identify metabolites with potential use in nutritional and physiological state assessments as well as personalized health and nutrition strategies. Using MS-based multiplatform metabolomics, Pardo-Rodriguez et al. investigated alterations in the metabolism of *Trypanosoma cruzi* after treatment with extracts of the Andean shrub *Clethra fimbriata*. More than 150 altered metabolites were identified in the treated parasites, with those related to energy metabolism pathways being the most affected. In addition, the authors found that triterpenes originating on the plant contributed to the disruption of essential processes in the parasite. Based on a study involving MS/MS data of beauvericins, depsipeptides present in *Fusarium* spp. Fungi, Selegato et al. demonstrated that the combination of feature-based molecular networking and MassQL is an effective strategy for accelerating the decoding of mass fragmentation pathways and identifying molecules with comparable fragmentation patterns. In combination with molecular networking analysis, MS-based untargeted metabolomics was applied to investigate the influence of altitudinal variations in the chemical composition of different bamboo species (Chivita et al.). The study uncovered 89 differential metabolites between the altitudinal ranges investigated, with an increase in the profile of flavonoids observed at high altitude and a boost in the levels of cinnamic acid derivatives registered at low altitude. Finally, Arrieta-Echeverri et al. carried out a characterization of the microbial populations and chemical space composition of a water kefir fermentation using culture-dependent methods, compositional metagenomics, and untargeted metabolomics based on LC-QTOF-MS. The work provides specific knowledge that could be easily applied to the rational development of novel probiotic and postbiotic ingredients for functional nutrition.

To conclude, it is worth mentioning that several of the manuscripts published in this Research Topic are the product of synergistic collaborations between Latin American research groups, and that those collaborations developed, in great part, from events fostered by LAMPS. We are already organizing the upcoming V LAMPS Meeting, which will include several workshops and will have the participation of invited experts from around the World. We are certain that this meeting, which will be held in Montevideo, Uruguay, from October 30th to 1 November 2024 (https://sites.google.com/unesp.br/v-lamps-2024/home), will confirm that the field of metabolomics, as well as its ancillary methodologies, continues to expand and mature in the region.

Author contributions

MA: Writing–review and editing, Writing–original draft. PAH: Writing–review and editing, Writing–original draft. GM: Writing–review and editing, Writing–original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. GM received financial support from the Programa de Desarrollo de las Ciencias Básicas (PEDECIBA).

Acknowledgments

We wish to thank the authors of the manuscripts published in this Research Topic for their valuable contributions and the reviewers for their rigorous evaluations. We also thank the editorial board and specialists of Frontiers in Molecular Biosciences for their support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 06 January 2023
DOI 10.3389/fmolb.2022.1067296

OPEN ACCESS

EDITED BY
Manuel Portero-Otin,
Universitat de Lleida, Spain

REVIEWED BY
James Adams,
Independent Researcher, United States
Maria Isabel Behrens,
University of Chile, Chile

*CORRESPONDENCE
Laura Morelli,
lmorelli@leloir.org.ar

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Metabolomics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 11 October 2022
ACCEPTED 16 December 2022
PUBLISHED 06 January 2023

CITATION
Dalmasso MC, Arán M, Galeano P, Perin S, Giavalisco P, Martino Adami PV, Novack GV, Castaño EM, Cuello AC, Scherer M, Maier W, Wagner M, Riedel-Heller S, Ramirez A and Morelli L (2023), Nicotinamide as potential biomarker for Alzheimer's disease: A translational study based on metabolomics.
Front. Mol. Biosci. 9:1067296.
doi: 10.3389/fmolb.2022.1067296

Nicotinamide as potential biomarker for Alzheimer's disease: A translational study based on metabolomics

María C. Dalmasso[1,2,3], Martín Arán[4], Pablo Galeano[1], Silvina Perin[5], Patrick Giavalisco[5], Pamela V. Martino Adami[2], Gisela V. Novack[1], Eduardo M. Castaño[1], A. Claudio Cuello[6], Martin Scherer[7], Wolfgang Maier[8,9], Michael Wagner[8,9], Steffi Riedel-Heller[10], Alfredo Ramirez[2,8,9,11,12†] and Laura Morelli[1*†]

[1]Laboratory of Brain Aging and Neurodegeneration-Fundación Instituto Leloir-IIBBA-National Scientific and Technical Research Council (CONICET). Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina, [2]Division of Neurogenetics and Molecular Psychiatry, Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany, [3]Studies in Neuroscience and Complex Systems Unit (ENyS-CONICET-HEC-UNAJ). Florencio Varela, Florencio Varela, Argentina, [4]Laboratory of NMR-Fundación Instituto Leloir-IIBBA-National Scientific and Technical Research Council (CONICET). Ciudad Autónoma de Buenos Aires, Cologne, Argentina, [5]Max Planck Institute for Biology of Ageing, Cologne, Germany, [6]Department of Pharmacology and Therapeutics, McGill University, Montreal, CA, Canada, [7]Department of Primary Medical Care, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany, [8]Department of Neurodegenerative and Geriatric Psychiatry, University Hospital Bonn, Medical Faculty, Bonn, Germany, [9]German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany, [10]Institute of Social Medicine, Occupational Health and Public Health, University of Leipzig, Leipzig, Germany, [11]Department of Psychiatry and Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, San Antonio, TX, United States, [12]Cluster of Excellence Cellular Stress Responses in Aging-associated Diseases (CECAD), University of Cologne, Cologne, Germany

Introduction: The metabolic routes altered in Alzheimer's disease (AD) brain are poorly understood. As the metabolic pathways are evolutionarily conserved, the metabolic profiles carried out in animal models of AD could be directly translated into human studies.

Methods: We performed untargeted Nuclear Magnetic Resonance metabolomics in hippocampus of McGill-R-Thy1-APP transgenic (Tg) rats, a model of AD-like cerebral amyloidosis and the translational potential of these findings was assessed by targeted Gas Chromatography-Electron Impact-Mass Spectrometry in plasma of participants in the German longitudinal cohort AgeCoDe.

Results: In rat hippocampus 26 metabolites were identified. Of these 26 metabolites, nine showed differences between rat genotypes that were nominally significant. Two of them presented partial least square-discriminant analysis (PLS-DA) loadings with the larger absolute weights and the highest Variable Importance in Projection (VIP) scores and were specifically assigned to nicotinamide adenine dinucleotide (NAD) and nicotinamide (Nam). NAD levels were significantly decreased in Tg rat brains as compared to controls. In agreement with these results, plasma of AD patients showed significantly reduced levels of Nam in respect to cognitively normal participants. In addition, high plasma levels of Nam showed a 27% risk reduction of progressing to AD dementia within the following 2.5 years, this hazard ratio is lost afterwards.

Discussion: To our knowledge, this is the first report showing that a decrease of Nam plasma levels is observed couple of years before conversion to AD, thereby suggesting its potential use as biomarker for AD progression.

KEYWORDS

NAD salvage pathway, vit B3, transgenic rats, alzheimer's disease, biomarkers, brain alterations, nicotinamide (NAM), case-control analysis

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative proteinopathy characterized by deposition of amyloid β (Aβ) and hyperphosphorylated tau protein in the brain of patients. The pathology observed in AD begins years, or even decades, before the appearance of clinical symptoms. Thus, identification of biomarkers reporting on pathways modulating AD pathology in asymptomatic individuals at-risk is of paramount importance to define target groups for early prevention strategies once these become available. This, however, has been proven to be a major challenge as several, partially unknown, pathways contribute to the pathology leading to neurodegeneration, cognitive decline and finally dementia (De Strooper and Karran, 2016). Unfortunately, current validated biomarkers inform on the neuropathological hallmarks of the disease following the amyloid cascade hypothesis leaving other pathways uncovered (Jack et al., 2018). This assumption receives further support from disappointing results from amyloid-specific therapies in AD.

Dementia stage in AD is the culmination of a series of events that begin with a complex interplay between genetic and environmental susceptibility factors years before cognitive symptoms become apparent. This interplay triggers a sequence of pathological changes which involves process altering Aβ homeostasis, as well as processes beyond amyloid such as vascular changes, neuroinflammation and age-related factors relevant for reserve and resilience of the brain (De Strooper and Karran, 2016). Given the difficulty linked to the search for biomarkers informing on these pathways in humans, research has turned into model organisms to identify and to characterize conserved pathogenic pathways and molecules that could serve as biomarkers for AD (Wang et al., 2021). Herein, a promising animal model is the McGill-R-Thy1-APP rat (Leon et al., 2010) expressing the human amyloid precursor protein (APP) with the Swedish and Indiana mutations responsible for familial AD in humans. The hemozygous Tg ± rats do not develop extracellular plaques, but show intraneuronal accumulation of Aβ in cortex and hippocampus (Leon et al., 2010; Iulita et al., 2014), a similar feature was described in the human brain at early stages of AD amyloid pathology (Christensen et al., 2010). Moreover, these animals show accumulation of SDS-resistant Aβ oligomers (~30 kDa) from 6 months onwards (Galeano et al., 2014); synaptosomal bioenergetic defects (Martino Adami et al., 2017a) and cognitive impairments in different hippocampal-dependent behavioral tasks (Leon et al., 2010; Galeano et al., 2014; Iulita et al., 2014; Martino Adami et al., 2017a; Martino Adami et al., 2017b; Habif et al., 2021) resulting and interesting model of early AD-amyloid pathology. The homozygous Tg+/+ rats show the full AD-like-amyloid pathology, accompanied by neuroinflammation and cognitive impairment, reflecting stages of late AD (Leon et al., 2010). While the Tg rat model has been extensively used to explore stages of AD pathology and validation of experimental therapeutic candidates, studies linking the metabolic profiles in hippocampus in association with the degree of amyloid pathology are still lacking. Furthermore, translational research is still needed to define whether findings made in the McGill-R-Thy1-APP rat can also be seen in AD patients.

Identification of novel biomarkers covering pathogenic pathways beyond classic amyloid cascade pathways will derive in better clinical diagnosis, particularly at preclinical stage of the disease. Recent developments in sensitivity and specificity of proteomics and metabolomics technologies have made it possible to identify different molecules targeting these additional pathological pathways. Thus, for example, cerebrospinal fluid (CSF) and blood levels of the neurofilament light chain (NfL) have been used as a sensitive biomarker for neuroaxonal damage that can monitor neurodegeneration and progression of Alzheimer's disease dementia, albeit not specific (Norgren et al., 2003; Gaiottino et al., 2013). While most reports have been done using data derived from mouse models for neurodegenerative diseases (Wilkins and Trushina, 2017), few reports have been focused on the McGill-R-Thy1-APP rat model (Nilsen et al., 2012; Nilsen et al., 2014b).

Consequently, this study aimed to characterize metabolic abnormalities in the hippocampus of homo- and hemizygous McGill-R-Thy1-APP rats by using Nuclear Magnetic Resonance (^{1}H-NMR) spectroscopy. Promising findings in the rat were followed up in human plasma samples by Gas Chromatography Electron Impact Mass Spectrometry (GC-EI-MS) to explore their potential utility as AD biomarkers.

Materials and methods

Rat model

Transgenic (Tg) McGill-R-Thy1-APP rats (Leon et al., 2010) were provided to Fundación Instituto Leloir (FIL) by The Royal Institution for the Advancement of Learning/McGill University, Montreal, Canada, and an in-house colony was established at FIL. Rats' genotypes were determined by real time qPCR as previously described (Galeano et al., 2014). To avoid the litter effect, groups were made up of pups from three to four different litters. Homozygous (Tg+/+), hemizygous (Tg+/-), and littermates' wild type (WT) control animals were maintained in polycarbonate cages in a temperature-controlled animal facility with a 12-h dark/light cycle and allowed to consume standard diet and water *ad libitum*. Only 9-month-old male rats were used for experiments to avoid any potential effects of female estrus cycle. All experimental procedures were performed in accordance with the guidelines of ARRIVE and OLAW–NIH. The protocol was approved by the local animal care committee (CICUAL # A5168-01).

Rat hippocampal tissue collection

Rats were anesthetized with ketamine (50 mg/kg) and xylacine (10 mg/kg), placed under a guillotine blade, decapitated and brains quickly removed. Sacrifices were carried out during the morning. Hippocampi were dissected and processed as described in the Supplementary Figure S1 minimizing the time between sacrifice and tissue freeze.

Human plasma samples

Samples were selected from the German study on Aging, Cognition and Dementia (AgeCoDe) biobank (Ramirez et al., 2015). The original study protocol was approved by the local ethics committees at the following German institutions: University of Bonn; University of Hamburg; University of Duesseldorf; University of Heidelberg/Mannheim; University of Leipzig and the Technical University of Munich. Written informed consent was obtained from all participants. The main assessment instrument at all visits included the Structured Interview for Diagnosis of Dementia of Alzheimer type, Multi-infarct Dementia and Dementia of other etiology according to DSM-IV and ICD-10 (SIDAM), and diagnosis of AD was established according to the NINCDS-ADRDA criteria for probable AD (McKhann et al., 1984; Zaudig et al., 1991).

This is a longitudinal study, where participants were recruited in primary care centers in six German cities. Inclusion criteria were to be at least 75 years old and cognitively healthy according to the general practitioner's judgment. Every ~18 months interval participants are followed-up with personal interviews and neuropsychological assessments. To date, nine follow-ups (FUs) were completed, but results from the last one are still in process. Blood samples were obtained at the third visit, processed and store at -80°C. For this study the third visit is considered the baseline. Controls (n = 189) remained cognitively unimpaired until the last FU, and were 83.6 ± 3.1 years old, 64.0% female and 20.6% Apolipoprotein E4 (*APOE4*) carriers. In this report, participants who converted to AD at baseline were denominated *incident* AD (n = 68), and participants with diagnosis of AD before the baseline, were denominated *prevalent* AD (n = 29). Participants with *incident* AD were 86.0 ± 3.6 years old, 64.7% female and 33.8% *APOE4* carriers; and those with *prevalent* AD, were 84.2 ± 3.1 years old, 75.8% female and 37.9% *APOE4* carriers. Subjects converting to AD in the next three visits following baseline (FU1, FU2 and FU3) were included in the analysis. At FU1 there were 25 participants with mean age of 84.8 ± 3.5 years old, 80% women, and 28% *APOE4* carriers; at FU2 there were 37 participants with mean age of 83.6 ± 2.6 years old, 67.6% women, and 32.4% *APOE4* carriers; and at FU3 there were 23 participants with mean age of 82.7 ± 2.6 years old, 60.9% women, and 21.7% *APOE4* carriers.

Expression of aβ isoforms in rat hippocampus

To quantify human Aβ 38/40/42 MSD® V-PLEX PLUS Aβ Peptide Panel one kit was used following the manufacturer's instructions. Methodology is described in the SI.

Untargeted nuclear magnetic resonance (NMR) spectroscopy

Frozen rat hemi-hippocampus were homogenized with a teflon-glass grinder in 2 ml ice-cold 80% methanol (Nagana Gowda et al., 2018) and centrifuged at 4°C for 10 min at 15000 xg. Supernatants were collected, dried in a Savant SpeedVac (Thermo Scientific) and solubilized in .5 ml sodium phosphate buffer (100 mM dissolved in D_2O, pH = 7.4), supplemented with 3-trimethylsilyl-[2,2,3,3,-2H4]-propionate (TSP, final concentration .33 mM) as chemical shift reference. Sample sizes for NMR experiments were chosen using an analysis based approach, MetSizeR (Nyamundanda et al., 2013). All NMR experiments were performed at 298 K on a Bruker Avance III spectrometer operating at a proton frequency of 600.3 MHz. ^{1}H-NMR 1D spectra were acquired using a standard Bruker 1D NOESY pulse program with pre-saturation during relaxation delay and mixing time, and spoil gradients (noesygppr1d). The following experimental parameters were used in all measurements: 256 scans, 1.85 s relaxation delay, 1.36 s acquisition time, 20 ppm spectral width, 10 m mixing time, and 32 K acquisition points. The NMR data were zero-filled, Fourier transformed, phase corrected using NMRPipe and converted to a Matlab-compatible format for further processing and analysis. All spectra were referenced to TSP (1H δ = 0 ppm) and submitted to water peak elimination, baseline correction, normalization, and scaling. The assignment was achieved using the freely available electronic databases HMDB and BMRB, and subsequently confirmed by 2D spectra including heteronuclear single quantum coherence (HSQC) and total correlation spectroscopy (TOCSY) (Supplementary Table S1). 2D ^{1}H–^{1}H TOCSY spectra were collected with N1 = 512 and N2 = 2048 complex data points. The spectral widths for the indirect and the direct dimensions were 9,615.4 and 9,604.9 Hz, respectively. The number of scans per t1 increment was set to 36. The transmitter frequency offset was 4.7 ppm in both ^{1}H dimensions. 2D ^{13}C-^{1}H HSQC spectra were collected with N1 = 512 and N2 = 2048 complex data points. The spectral widths for the indirect and direct dimensions were 24,906.9 and 12,019.2 Hz, respectively. The number of scans per t1 increment was set to 256. The transmitter frequency offset was 70 ppm in the ^{13}C dimension and 4.7 ppm in the ^{1}H dimension. The estimated detection limit for ^{1}H NMR (at 600 MHz) is dependent on the compound and varies between 1–10 μM.

Measurement of NAD+ and NADH in rat hippocampal tissues

NAD+/NADH levels were measured using NAD+/NADH assay kit from Abcam (ab65348) as described in SI.

Determination of enzymes transcript levels of NAD rate-limiting and NAD salvage pathway

MRNA levels of NAMPT (rate-limiting); NMNAT (NAD-generation) and CD38, PARP1, PARP2 and Sirt 3 (NAD-consuming) enzymes were assessed by qRT-PCR as described in the SI.

Targeted Gas Chromatography Electron Impact Mass Spectrometry (GC-EI-MS)

Human plasma samples were thawed on ice, and 100 ul were extracted with 900 ul of cold extraction buffer containing 40:40: 20 methanol:acetonitrile:water [v:v:v]. After 30 min in an orbital mixer at 4°C, samples were sonicated for 10 min in an ice-cooled bath-type sonicator and centrifuged for 10 min at 16000xg at 4°C. Supernatants were collected and dried in a SpeedVac until complete dryness. Standard curves of Nam were prepared with concentrations ranging from .005 to 50 ug/ml (expected limit of detection .1–.5 ng/ml). Standards were processed in the same way as samples. Dried down samples and standards were derivatized using methoxyamine and MSTFA/FAMEs solution (N-methyl-N-trimethylsilyl-trifluoracetamid/Fatty acid methyl esters) following standard procedures (Lisec et al., 2006; Caldana et al., 2013). After that, samples were analyzed in a GC-EI-MS (Q Exactive GC Orbitrap system, ThermoFisher) using a 30-m DB-35M capillary column. Representative fragments from the GC-EI-MS analysis of Nam were extracted using TraceFinder (Version 4.1, ThermoFischer) and quantified using the linear range of the obtained standard curve. All analysis were performed using peak areas, transformed into Z-scores, for easier comparison among experiments.

Statistical analysis

The normalized NMR spectral areas (AUC) of assigned metabolites were subjected to Pareto scaling and analyzed by multivariate analysis using MetaboAnalyst 5.0 (Chong et al., 2018) (Supplementary Figure S1). The statistical significance was assessed by one-way ANOVA, Fisher's LSD were performed for all post-hoc tests taking $p < 0.05$ as significant.

Data of NADH and NAD + levels were analyzed by one-way ANOVA tests followed by post-hoc Tukey's multiple comparisons tests.

Statistical analysis and plots of data from GC-EI-MS experiments performed with human plasma were done using R-project v. 4.0.0 (https://www.R-project.org) and R-studio v.1.2.5042 (http://www.rstudio.com/). Normal distribution was visualized using qqnorm plots, and outliers (defined as mean ± 3 standard deviation) were eliminated from the analysis ($n = 2$). For easier comparison among experiments in human samples a Z-score standardization was applied and subsequently data were analyzed by one-way ANOVA tests followed by post-hoc Tukey's multiple comparisons tests. In all cases, assumption of normality was examined using Kolmogorov-Smirnov or Shapiro–Wilk tests. A probability equal or less that 5% was considered as significant. All analyses were carried out using GraphPad Prism for Windows (version 7.0).

Linear regression models adjusted for sex, age and apoe4 were used to estimate the association of Nam levels in cases vs. controls in the *discovery* and *replication* experiments, as well as in FUs groups. Meta-analysis was performed using the R-package "metafor" (Viechtbauer, 2010) and visualized with the general function forest. The cox proportional hazards regression model, which relates time dependent variables, time dependent strata, and multiple events per subject, were performed with the R-package "survival" (Therneau, 2020) and "survminer" (Kassambara, 2021). Samples of paritcipants converting toAD at FU1-3 were included, time variable was time to conversion to AD in years, and the event per subject was conversion (no = 0, yes = 1). Proportional hazard assumption was tested by Schoenfeld's test, and consequently two cox regressions were performed with a split-time = 2.5 years.

Results

Comparison of ^{1}H-NMR metabolomics profiles of Tg and control rats

In this report we did not assess the cognitive status of Tg rats. However, it was previously reported by us and others that from 3 to 9 months of age, Tg rats show impairments in learning and spatial reference memory (Galeano et al., 2014; Wilson et al., 2017a), in long-term memory of inhibitory avoidance to a foot-shock, in novel object recognition memory and social approaching behavior (Habif et al., 2021), in cued fear-conditioning recall (Wilson et al., 2017a), and associative learning (Wilson et al., 2017b). Studies by Leon et al. (2010) and Iulita et al. (2014) established that 13-month-old Tg+/+ rats show marked cognitive impairments, while Tg ± rats perform intermediately between homozygous and WT genotypes. To determine Aβ-associated shifts in brain metabolites, we first performed a highly sensitive multiplex ELISA to quantify total Aβ levels within the hippocampus of a sub-set of Tg rats ($n = 3–7$). The median value of the concentration of Aβ40 showed nearly 4-fold increase in Tg+/+ [54.4 [pg/mg] (IQR: 31.8–55.7)] vs. Tg+/− [13.8 [pg/mg] (IQR: 9.8–31.1), $p = .048$]. For Aβ42, the increased was more than 20-fold [48.8 [pg/mg] (IQR: 16.7–153.4)] vs. Tg+/− [2.4 [pg/mg] (IQR: 2.0–7.5), $p < 0.035$]. These results confirmed the impact of the two copies of human mutant APP transgene on the accumulation of cerebral amyloid. To identify metabolic changes in pathways relevant for the hippocampus of the McGill-R-Thy1-APP rat, we carried out untargeted ^{1}H-NMR metabolomics on methanol-extracted samples from freshly isolated tissues. A total of 26 compounds were detected and identified (residual methanol was excluded from the analysis), including mainly amino acids, carboxylic acids, and nucleotides (Figure 1A; Supplementary Table S1).

Principal component analysis (PCA) and partial least square-discriminant analysis (PLS-DA) were performed in order to detect the differences among control and Tg rats (Supplementary Figure S1). Although the 95% confidence intervals generated for each group overlapped in the PCA score plots, a pronounced separation was observed for Tg+/+. Supervised PLS-DA was in line with PCA and gave a clearer separation between groups (Supplementary Figure S1). Nine metabolites were found to be significantly altered in Tg rats vs. control (Table 1).

Interestingly, most of them (i.e., taurine, glutathione, tyrosine, and glutamate) have been previously reported in several studies on metabolomics performed in biological samples from the CNS, both in animal models and in AD patients (Altine-Samey et al., 2021). However, the PLS-DA loadings with the larger absolute weights and the highest VIP scores were NAD and Nam (Supplementary Figure S1). These two metabolites showed significant differences between genotypes (Figures 1B, D). Standard runs analysis was performed in order to confirm de identity of NAD in contrast to its related metabolite Nicotinamide Adenine Dinucleotide Phosphate (NADP) (Figure 1C). As previously described for these key molecules in the

10.3389/fmolb.2022.1067296

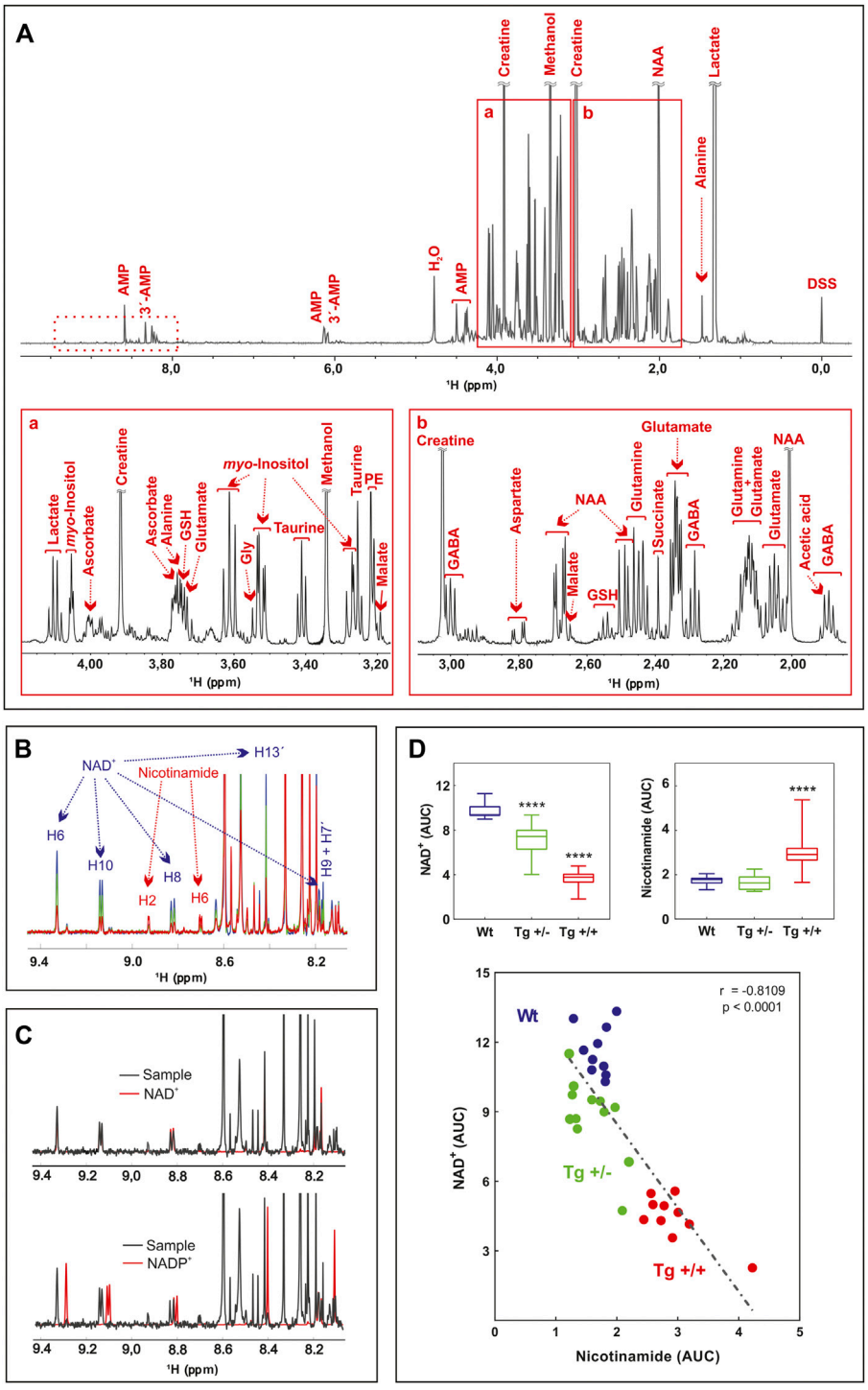


FIGURE 1
Untargeted ^{1}H-NMR metabolomics of hippocampus of AD-like amyloid pathology transgenic rats **(A)** Typical 600 MHz ^{1}H-NMR spectrum of WT rats, representative of all the registered spectra. Assigned resonances of specific metabolites are indicated in red. Expanded views of the spectrum between 3.1–4.4 ppm **(A)** and 1.7–3.2 ppm **(B)** are shown **(B)** Overlaid of averaged ^{1}H-NMR spectra of WT rats (blue) (n = 10), Tg +/- (green) (n = 12) and Tg+/+ (red) (n = 10) in the 9.5–8.0 ppm zone (dashed box in A). The resonances assigned to NAD and Nam protons are indicated **(C)** Overlaid of representative ^{1}H-NMR spectra of WT rats (black, sample), NAD standard (red, upper panel) and NADP standard (red, lower panel) **(D)** Lower panel: correlation between NAD and Nam levels of samples analyzed. The AUC of H6 of NAD and H2 of Nam were plotted (n = 32). The linear regression (dashed line), the Pearson's correlation coefficient and the p-value (two tailed) are shown. Upper panels: differences in the AUC of NAD (left) and Nam (right) among groups (WT, blue; Tg+/-, green and Tg+/+, red) were analyzed by one-way ANOVA. Significant differences are indicated accordingly to Fishers´s LSD test. ****$p < 0.0001$.

TABLE 1 Hippocampal metabolites detected by ^{1}H-NMR spectroscopy that showed significant differences between control and Tg rats.

Metabolite	FRD			
	Tg+/-	Tg+/+	Tg−/+	Tg+/+
NAD	2,50E-06	8,36E-11	↓	↓
Nam	>0.05	3,20E-05	=	↑
Taurine	2,39E-06	2,70E-05	↓	↓
Valine	1,71E-02	>0.05	↓	=
GSH	>0.05	2,08E-02	=	↑
Tyrosine	>0.05	2,47E-02	=	↑
NAA	3,55E-02	>0.05	↓	=
Creatine	3,55E-02	>0.05	↓	=
Glutamate	4,59E-02	4,80E-02	↓	↓

Differences in the AUC, of metabolites were analyzed by one-way ANOVA, with Fisher's LSD, post-hoc test; FDR, false discovery rate. Arrows indicate significant increase (up) or decrease (down) with respect to control rats. = , no change from control rats. NAD, nicotinamide adenine dinucleotide; Nam, nicotinamide; GSH, glutathione; NAA, N-acetylaspartate.

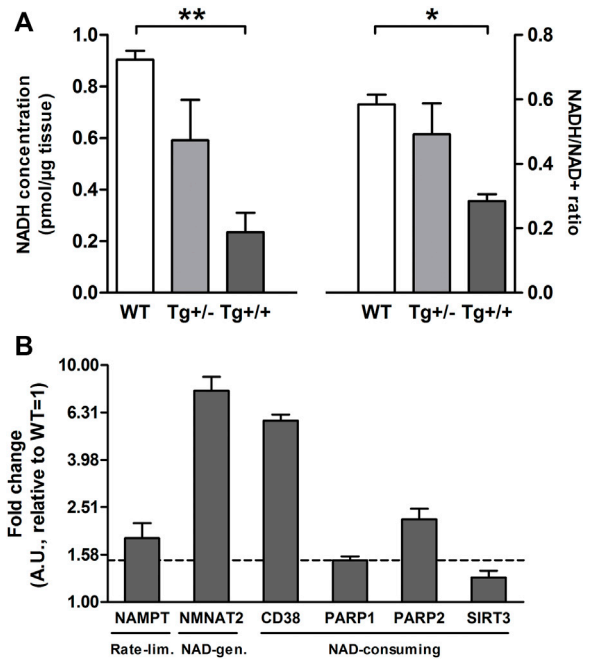


FIGURE 2
Hippocampal Aβ deposition alters the brain NAD + metabolism **(A)** Bars show mean ± SEM levels of NADH (left panel) and NADH/NAD + ratio (right panel) in hippocampal homogenates of control (WT; n = 3), hemizygous (Tg+/; n = 3) and homozygous (Tg+/+; n = 3) transgenic rats. One-way ANOVA tests and *post hoc* analyses revealed that Tg+/+ showed significantly lower levels of NADH ($F_{(2, 6)}$ = 10.76, p = 0.01; post-hoc: **p < 0.01) and NADH/NAD + ratio ($F_{(2, 6)}$ = 6.71, p = 0.02; post-hoc:*p < 0.05) compared with those observed in WT **(B)** Transcript levels of rate-limiting (NAMPT), NAD + -generation (NMNAT2) and NAD + consuming enzymes (CD38, PARP1, PARP2 and SIRT3) in hippocampal homogenates of Tg+/+ rats. Each bar represents the mean ± SEM of at least three independent experiments performed by triplicate for each sample normalized by GAPDH or Eukaryotic Translation Elongation Factor 1 Alpha 1 (EEF1A1). The mean ± SEM relative to WT (=1) is shown. Values above the dashed line (+1.5) were considered different from WT (=1).

NAD salvage-pathway, Nam and NAD levels showed an inverse relationship (Figure 1D; Supplementary Figure S1). Since NMR analysis cannot differentiate between NADH and NAD+, and taking into account that NADH levels decrease as a function of age and that the ability to regenerate NADH drops sharply in aged brain (Lautrup et al., 2019) we quantified NADH and NAD + by a colorimetric kit and found that homozygous rats (Tg+/+) showed significantly lower levels of NAD+ and NADH compared with those observed in WT. In contrast, hemizygous rats (Tg+/-) showed an intermediate level that did not reach significance neither with WT nor with Tg+/+ (Figure 2A). In this regard, NADH/NAD + ratio was significantly lower in the Tg+/+ as compared to WT and Tg +/- (Figure 2B) suggesting a clear alteration of the redox state in the brains of Tg+/+ rats, which is probably still incipient in the Tg ± ones. To explain alterations in Nam and NAD+/H levels observed in Tg+/+ brains we first evaluated transcript levels of Nicotinamide phosphoribosyltransferase (NAMPT) the rate-limiting component in the NAD + rescue pathway (Garten et al., 2015) and found a slight increase (1.85 ± 0.29) as compared with the control group (WT = 1) (Figure 2B). In addition, gene expression of NAD + -generating enzyme nicotinamide mononucleotide adenylyltransferase (NMNAT2) and NAD + -consuming (CD38, PARP1, PARP2 and SIRT3) enzymes (Okabe et al., 2019) were also assessed. We detected increments greater than 1.5 fold-change in NMNAT2 (7.8 ± 1.12), CD38 (member of the cyclic ADP-ribose synthase family) (5.83 ± 0.35) and PARP2 (member of the ADP-ribose transferases family) (2.23 ± 0.24). Whereas transcript levels of PARP1 and SIRT3 (sirtuin) were unaffected (1.5 ± 0.06 and 1.27 ± 0.08, respectively) (Figure 2B). Based on these results, expression of rate-limiting enzyme in Tg+/+ seems to be slight different from WT, while NAD + -consuming and the NAD + -generating pathways seem to be activated in Tg+/+ brain suggesting potential disturbance of the NAD + rescue pathway following the ongoing amyloid pathology. While central disturbance in NAD + metabolism in Tg rats was observed, its translation to peripheral tissue was unclear.

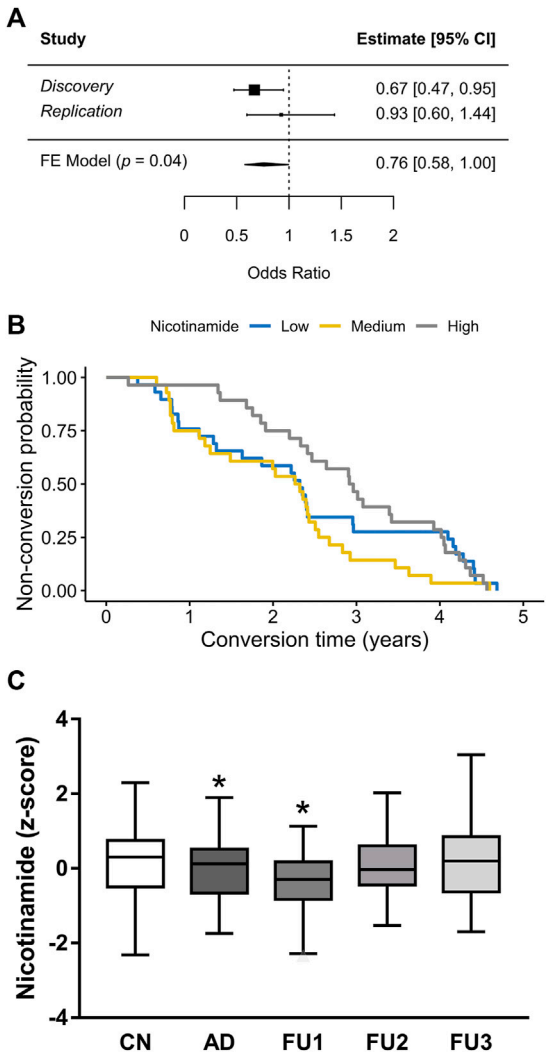


FIGURE 3
Plasma levels of Nam in association with AD **(A)** Meta-analysis forest plot of Nicotinamide plasma levels in human samples. Discovery experiment includes 68 cases and 93 controls. Replication experiment includes 29 cases and 93 controls. Estimates are in Odds Ratios; CI, confidence interval; FE Model, fixed effects meta-analysis results. **(B)** Kaplan-Meier conversion to AD survival of 85 participants after blood test for Nicotinamide, stratified in high, medium or low levels. High levels of Nicotinamide seem to be a predictor of dementia survival for 2.5 years (HR = 0.73, p = 0.04) **(C)** Box plots represent the normalized GC-EI-MS spectral areas of Nicotinamide in human plasma of CN (cognitive normal) subjects (n = 189); AD (n = 85) patients and FU1 (n = 25); FU2 (n = 37) and FU3 (n = 23) participants.*p < 0.05.

Plasma nam levels as a potential biomarker of AD

The results in the brain of the rat prompted us to explore whether these findings can be translated to humans. Herein, we focused our analysis on plasma because it might offer a promising alternative for biomarker in blood. Since measure of NAD in clinical practice is methodologically complicated because of its size (665 Da) and its stability in chromatography solvents, Nam levels were measured. To this aim, targeted detection of Nam was performed in human plasma samples from the longitudinal study AgeCoDe using GC-EI-MS. First, we compared whether Nam plasma levels of 68 participants with AD dementia (AD) showed statistical differences compared to 93 cognitively normal (CN) participants and found that Nam levels were significantly reduced in cases compared to CN (odd ratio (OR) = 0.67, p = 0.02, Figure 3A). In an independent replication sample drawn from AgeCoDe, including 96 CN and 29 AD, Nam showed the same trend (OR = 0.93) which, however, did not reach significance (p = 0.7, Figure 3A) probably due to the small number of cases analyzed. The meta-analysis of both samples confirmed the protective effect of plasma levels of Nam (OR = 0.76, p = 0.04, Figure 3A).

Nam plasma levels as a prospective biomarker of AD conversion

To better analyze our results we explored whether Nam plasma levels, measured at baseline, were associated with the time to conversion to AD. Consequently, participants were included in the analysis if they have available data on plasma levels of Nam and converted to AD at any of the next three follow-ups (FU) for which data was available [(FU1) = 0.94 ± 0.35 years after baseline; FU2 = 2.43 ± 0.38 years after baseline; FU3 = 4.13 ± 0.37 years after baseline]. For the analysis, the impact of Nam levels was stratified in tertiles (high, medium and low) and their effect on time to conversion was visualized by *Kaplan-Meier* survival curves (Figure 3B) and analyzed by Cox regression models. This analysis showed that only the higher levels of Nam are associated with a later conversion to dementia (hazard ratio (HR) 0.73, p = 0.04). However, we also observed that the HR is not proportional over time (curves intersect). Thus, while a person with Nam levels in plasma within the high tertile has 27% risk reduction of progressing to AD within the next 2.5 years, this HR is lost afterwards. Supporting this finding, we observed that only participants progressing to AD at FU1 showed significantly lower levels of Nam compared with CN (p = 0 .04, Figure 3C).

Discussion

Alzheimer's disease is a complex phenotype involving several pathogenic pathways leading to a metabolic imbalance already at early stages of the disease before symptom become apparent. Relevant pathways altered in AD include lipid and amino acid metabolism, as well as dysregulation of the glucose metabolism and mitochondrial dysfunction guiding to energetic imbalance and oxidative stress (Yan et al., 2020). Using untargeted ^{1}H-NMR spectroscopy, we observed in hippocampus of the McGill-R-Thy1-APP rat a significant reduction of NAD level in Tg rats compared to their healthy littermates at 9 months of age suggesting an energetic imbalance in the Tg rats.

NAD + and related metabolites are critical compounds essential to adaptive stress responses and cell survival. It was well established that PARP-1 (a NAD + consuming enzyme) functions as a DNA repair enzyme under intense DNA damage as is the case of late AD brain neuropatholgy. It was postulated that PARP-1 activity depletes neurons of NAD+ and ATP leading to neuronal death by a caspase-independent mechanism that shares characteristics of apoptosis and necrosis (known as Parthanatos), recently reviewed by Salech et al. (2020). Interestingly, Nam is a well-known inhibitor of PARP-1. It is of note that McGill-R-Thy1-APP rats used in this study lack neuropathology of late-AD brains and PARP-1 transcript levels in

Tg rat brains are similar to control animals, suggesting that Parthanatos is not operative in this animal model of brain amyloidosis. Considering the number of enzymes and transcription factors sensitive to the redox potential, NAD+/H redox state acquires pathophysiological relevance for aging and neurodegenerative diseases (Verdin, 2015; Fang et al., 2017). While several studies in mouse models for AD have shown the relevance of the NAD(P)+/NAD(P)H homeostasis in the brain, especially in hippocampus and cortex (Ghosh et al., 2012; Dong and Brewer, 2019; Dong et al., 2019), few reports have been published on the role of the NAD(P)+/NAD(P)H homeostasis in the McGill-R-Thy1-APP rats. A previous *in vivo* study using Magnetic Resonance Spectroscopy (^{1}H-MRS) identified in tissue derived from hippocampus and frontal cortex of McGill-R-Thy1-APP rat significant difference in levels of several metabolites compared to the WT littermates (Nilsen et al., 2012). Herein, the Tg+/+ rats, compared to WT rats, showed lower levels of glutamate, GABA, N-acetylaspartate (NAA) and elevated myo-inositol and taurine. These differences become apparent during the progression of amyloid pathology in Tg+/+ in time window of 6 months between three-to 9-months of age. Previously, the NAA and myo-inositol findings were reproduced in dorsal hippocampus tissue derived from this rat model, though only in males Tg+/+ rats (Nilsen et al., 2014b). Metabolites identified in these studies suggested brain damage and mitochondrial dysfunction that might be gender specific. In line with this report, we were able to replicate in part these previous differences using brain tissue from this rat model. Minor differences might derive from dissimilarities in the techniques used in previous studies compared to ours (*in vivo* ^{1}H-MRS vs *ex vivo* ^{1}H-NMR). By using *in vivo* ^{1}H-MRS the regional concentration of low molecular weight metabolites can be measured non-invasively. Conversely, *ex vivo* ^{1}H-NMR spectroscopy detects only hydrophilic metabolites extracted from tissue homogenates. Hence, both approaches might be complementary for the identification of neurochemical processes related to AD pathology and its progression over time. Our findings on NAD receive further supports from ^{1}H- and ^{13}C NMR spectroscopy and HP-LC experiments done in cingulate cortex derived from aged McGill-R-Thy1-APP (15-month-old) that showed decreased levels of NAD + in Tg rats compared to WT (Nilsen et al., 2014a).

In correlation with our findings, experimental evidence supports a protective effect of NAD + supplementation on cognitive deficits in AD models (Gong et al., 2013; Liu et al., 2013). It was previously reported (Xing et al., 2019) in hippocampal tissue of 6-month-old APPswe/PS1Δe9 transgenic mice decrements of NAD-generating enzyme (NAMPT) levels which were reverted by the administration of NAD, suggesting that increasing NAMPT expression levels may promote NAD production. Our results showed a slight increase of NAMPT transcript levels in Tg+/+ as compared to control rats (1,8 fold-change as compared to WT = 1). It is of note that NAMPT expression is induced by inflammatory signals and is considered a biomarker of chronic and acute inflammatory disease (Audrito et al., 2020). In this regard, hippocampal accumulation of Aβ in Tg+/+ rats may act as an alarmin triggering proinflammatroy cytokines (Wilson et al., 2018) and promoting increments of NAMPT transcript levels as reported here. Consequently, we postulate the possibility to use NAD + metabolites as peripheral biomarkers for AD. In line with this hypothesis, our study identified a significant lower level of plasma Nam in AD patients compared to healthy controls. This difference was also seen before the patients progressed to AD. Thus, Nam levels in plasma could serve as biomarker for progression to AD. However, risk reduction associated to high levels of Nam is lost after 2.5 years, meaning it is only valid in the close proximity to its assessment. The relevance of our observation is reinforced by a recent report showing that by untargeted metabolomics 308 CSF metabolites from 338 individuals were identified and associated using principal components (PCs) analysis with CSF total tau (t-tau), phosphorylated tau (p-tau), Aβ42, and Aβ42/40 ratio. Employing linear regression models 5 PCs were significantly associated with CSF p-tau and t-tau and 3 PCs with CSF Aβ42. Pathway analysis suggested that these PCS were enriched in six pathways, including metabolism of caffeine, nicotinate and Nam. (Dong et al., 2022).

In addition to the role of Nam as a potential biomarker for AD progression Nam may be also involved in AD onset. In this regard, a new mechanism of AD induction was recently postulated in which NAD depletion due to inadequate levels of Nam may have a relevant role in neuronal damage. On this point, the dietary habits in the aging characterized by low fruits and vegetable consumption and the presence of visceral fat which secretes visfatin, an inflammatory adipokine that deplets blood Nam, may explain why many people do develop AD due to lifestyle (Adams, 2021). However, research in humans has shown that plasma levels of NAD + decrease while levels of Nam increase significantly with age (Clement et al., 2019). These data have fueled several clinical trials of NAD + precursors, which still, produced inconsistent results (Rainer et al., 2000; Demarin et al., 2004).

The utility of Nam as a treatment for prevention of AD is still on debate. It was reported in animal models of AD that dietary supplements of Nam can increase the amount of (NAD) (+) in the brain, reduce the production of Aβ, and slow the decline of cognitive function. While Nam has shown promise in the treatment of AD, a Phase II Clinical Trial failed to demonstrate that Nam improves cognitive function in subjects with mild to moderate AD over 24 weeks. The lack of efficacy of Nam was explained due to several factors including a low sample size (n = 15); inclusion of subjects with moderate AD, and a relatively short treatment phase (Phelan et al., 2017). Currently few more human clinical trials are ongoing to evaluate the safety concerns of Nam supplementation however the outcomes are yet to be available (Nadeeshani et al., 2021).

We are aware that this study has also limitations. We only employed male Tg rats due to the well known effect that estrous cycle has on biochemical parameters, increasing intra-group variability. This is a clear impediment for generalizing results. However, in the experiments with human samples both genders were included. Finally, the replication analysis on human plasma did not reach significance. Since the effect goes in the same direction and the meta-analysis is still significant, this could be a minor problem. A replication in independent cohorts is necessary to validate the potential use of plasma levels of Nam as biomarker for AD progression.

Conclusion

In summary, our study provides additional supporting evidence indicating that hippocampal Aβ burden and/or hAPP processing is associated with the degree of NADH/NAD + shift in McGill-R-Thy1-

APP rat brain. Although this information cannot infer causal direction it offers a different perspective on the Aβ-mediating mechanisms involved in brain energy dysfunction observed in AD. Besides, our findings indicate that plasma Nam content has a potential role as short-term AD risk biomarker. Nevertheless, further studies in larger cohorts and independent populations of patients will be needed to confirm our results and the potential use of Nam as peripheral biomarker. To our knowledge, this is the first report showing a significant decrease of Nam plasma levels in people with AD that is observed couple of years before conversion, thereby suggesting its potential use as biomarker for progression.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics committees of the following German institutions: University of Bonn; University of Hamburg; University of Duesseldorf; University of Heidelberg/Mannheim; University of Leipzig and the Technical University of Munich. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by Leloir Institute animal care committee (CICUAL # A5168-01).

Author contributions

Conceptualization: MD, AR, and LM; Methodology: MD, MA, PG, PG, SP, GN, MS, WM, MW, and SR-H; Statistical Analysis: MD, MA, PG, and PM; Investigation: MD, AR, and LM; Resources: MD, AR, LM, and ACC; Writing—original draft: MD, MA, AR, and LM; Writing—review and editing: MD, MA, PG, PM, SP, PG, EC, ACC, MS, WM, MW, SR-H, AR, and LM; Supervision: AR and LM; Project administration: MD and LM; Funding acquisition: MD, ACC, AR, and LM. All authors read and approved the final manuscript.

Funding

This study was supported by funding from the Agencia Nacional de Promoción Científica y Tecnológica (PICT-2014-1,537 to MD, and PICT-2015-0285, PICT-2016-4647 and PIBT/09-2013 to LM.), International Society for Neurochemistry (CAEN Award 2015 to MCD), Alexander von Humboldt to MCD, Canadian Institutes of Health Research (201603PJT-364544 to ACC). The AgeCoDe cohort is part of the German Research Network on Degenerative. Dementia (KNDD, AgeCoDe study group), which is funded by the German Federal Ministry of Education and Research (BMBF) grants KNDD: 01GI1007A, 01GI0710, 01GI0711, 01GI0712, 01GI0713, 01GI0714, 01GI0715, 01GI0716, 01GI0717, 01ET1006B. MCD, PG, EMC and LM are members of the Research Career of CONICET. ACC is member of the Canadian Consortium of Neurodegeneration in Aging (CCNA) and holder of the McGill University Charles E. Frosst/Merck Chair in Pharmacology. PM Adami is supported by an Alexander von Humboldt Georg Forster Research Fellowship. Part of this investigation was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) grants RA1971/7-1 and RA1971/8-1 to AR.

Acknowledgments

We thank Luca Kleindeinman (University of Bonn and University of Cologne) and Rafael Campos (University of Cologne) for their advices on statistical analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.1067296/full#supplementary-material

References

Adams, J. D. (2021). Probable causes of Alzheimer's disease. *Sci.* 3, 16. doi:10.3390/sci3010016

Altine-Samey, R., Antier, D., Mavel, S., Dufour-Rainfray, D., Balageas, A. C., Beaufils, E., et al. (2021). The contributions of metabolomics in the discovery of new therapeutic targets in Alzheimer's disease. *Fundam. Clin. Pharmacol.* 35, 582–594. doi:10.1111/fcp.12654

Audrito, V., Messana, V. G., and Deaglio, S. (2020). NAMPT and NAPRT: Two metabolic enzymes with key roles in inflammation. *Front. Oncol.* 10, 358. doi:10.3389/fonc.2020.00358

Caldana, C., Li, Y., Leisse, A., Zhang, Y., Bartholomaeus, L., Fernie, A. R., et al. (2013). Systemic analysis of inducible target of rapamycin mutants reveal a general metabolic switch controlling growth in *Arabidopsis thaliana*. *Plant J.* 73, 897–909. doi:10.1111/tpj.12080

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

Christensen, D. Z., Schneider-Axmann, T., Lucassen, P. J., Bayer, T. A., and Wirths, O. (2010). Accumulation of intraneuronal Abeta correlates with ApoE4 genotype. *Acta Neuropathol.* 119, 555–566. doi:10.1007/s00401-010-0666-1

Clement, J., Wong, M., Poljak, A., Sachdev, P., and Braidy, N. (2019). The plasma NAD+ metabolome is dysregulated in "normal" aging. *Normal" Aging. Rejuvenation Res.* 22, 121–130. doi:10.1089/rej.2018.2077

De Strooper, B., and Karran, E. (2016). The cellular phase of alzheimer's disease. *Cell* 164, 603–615. doi:10.1016/j.cell.2015.12.056

Demarin, V., Podobnik, S. S., Storga-Tomic, D., and Kay, G. (2004). Treatment of alzheimer's disease with stabilized oral nicotinamide adenine dinucleotide: A randomized, double-blind study. *Drugs Exp. Clin. Res.* 30, 27–33.

Dong, R., Denier-Fields, D. N., Lu, Q., Suridjan, I., Kollmorgen, G., Wild, N., et al. (2022). Principal components from untargeted cerebrospinal fluid metabolomics associated with Alzheimer's disease biomarkers. *Neurobiol. Aging.* 117, 12–23. doi:10.1016/j.neurobiolaging.2022.04.009

Dong, Y., and Brewer, G. J. (2019). Global metabolic shifts in age and alzheimer's disease mouse brains pivot at NAD+/NADH redox sites. *J. Alzheimers Dis.* 71, 119–140. doi:10.3233/JAD-190408

Dong, Y., Digman, M. A., and Brewer, G. J. (2019). Age- and AD-related redox state of NADH in subcellular compartments by fluorescence lifetime imaging microscopy. *GeroScience* 41, 51–67. doi:10.1007/s11357-019-00052-8

Fang, E. F., Lautrup, S., Hou, Y., Demarest, T. G., Croteau, D. L., Mattson, M. P., et al. (2017). NAD(+) in aging: Molecular mechanisms and translational implications. *Trends Mol. Med.* 23, 899–916. doi:10.1016/j.molmed.2017.08.001

Gaiottino, J., Norgren, N., Dobson, R., Topping, J., Nissim, A., Malaspina, A., et al. (2013). Increased neurofilament light chain blood levels in neurodegenerative neurological diseases. *PLoS One* 8, e75091. doi:10.1371/journal.pone.0075091

Galeano, P., Martino Adami, P. V., Do Carmo, S., Blanco, E., Rotondaro, C., Capani, F., et al. (2014). Longitudinal analysis of the behavioral phenotype in a novel transgenic rat model of early stages of Alzheimer's disease. *Front. Behav. Neurosci.* 8, 321. doi:10.3389/fnbeh.2014.00321

Garten, A., Schuster, S., Penke, M., Gorski, T., de Giorgis, T., and Kiess, W. (2015). Physiological and pathophysiological roles of NAMPT and NAD metabolism. *Nat. Rev. Endocrinol.* 11, 535–546. doi:10.1038/nrendo.2015.117

Ghosh, D., LeVault, K. R., Barnett, A. J., and Brewer, G. J. (2012). A reversible early oxidized redox state that precedes macromolecular ROS damage in aging nontransgenic and 3xTg-AD mouse neurons. *J. Neurosci.* 32, 5821–5832. doi:10.1523/JNEUROSCI.6192-11.2012

Gong, B., Pan, Y., Vempati, P., Zhao, W., Knable, L., Ho, L., et al. (2013). Nicotinamide riboside restores cognition through an upregulation of proliferator-activated receptor-γ coactivator 1α regulated β-secretase 1 degradation and mitochondrial gene expression in Alzheimer's mouse models. *Neurobiol. Aging* 34, 1581–1588. doi:10.1016/j.neurobiolaging.2012.12.005

Habif, M., Do Carmo, S., Báez, M. V., Colettis, N. C., Cercato, M. C., Salas, D. A., et al. (2021). Early long-term memory impairment and changes in the expression of synaptic plasticity-associated genes, in the McGill-R-Thy1-APP rat model of alzheimer's-like brain amyloidosis. *Front. Aging Neurosci.* 12, 585873. doi:10.3389/fnagi.2020.585873

Iulita, M. F., Allard, S., Richter, L., Munter, L. M., Ducatenzeiler, A., Weise, C., et al. (2014). Intracellular Aβ pathology and early cognitive impairments in a transgenic rat overexpressing human amyloid precursor protein: A multidimensional study. *Acta Neuropathol. Commun.* 2, 61. doi:10.1186/2051-5960-2-61

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., et al. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement.* 14, 535–562. doi:10.1016/j.jalz.2018.02.018

Kassambara, A. (2021). *Survminer: Drawing survival curves using 'ggplot2*. R package version 0.4.8. Available at: https://CRAN.R-project.org/package=survminer.

Lautrup, S., Sinclair, D. A., Mattson, M. P., and Fang, E. F. (2019). NAD(+) in brain aging and neurodegenerative disorders. *Cell Metab.* 30, 630–655. doi:10.1016/j.cmet.2019.09.001

Leon, W. C., Canneva, F., Partridge, V., Allard, S., Ferretti, M. T., DeWilde, A., et al. (2010). A novel transgenic rat model with a full Alzheimer's-like amyloid pathology displays pre-plaque intracellular amyloid-beta-associated cognitive impairment. *J. Alzheimers Dis.* 20, 113–126. doi:10.3233/JAD-2010-1349

Lisec, J., Schauer, N., Kopka, J., Willmitzer, L., and Fernie, A. R. (2006). Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protoc.* 1, 387–396. doi:10.1038/nprot.2006.59

Liu, D., Pitta, M., Jiang, H., Lee, J. H., Zhang, G., Chen, X., et al. (2013). Nicotinamide forestalls pathology and cognitive decline in alzheimer mice: Evidence for improved neuronal bioenergetics and autophagy procession. *Neurobiol. Aging* 34, 1564–1580. doi:10.1016/j.neurobiolaging.2012.11.020

Martino Adami, P. V., Galeano, P., Wallinger, M. L., Quijano, C., Rabossi, A., Pagano, E. S., et al. (2017b). Worsening of memory deficit induced by energy-dense diet in a rat model of early-Alzheimer's disease is associated to neurotoxic Aβ species and independent of neuroinflammation. *Biochim. Biophys. Acta Mol. Basis Dis.* 1863, 731–743. doi:10.1016/j.bbadis.2016.12.014

Martino Adami, P. V., Quijano, C., Magnani, N., Galeano, P., Evelson, P., Cassina, A., et al. (2017a). Synaptosomal bioenergetic defects are associated with cognitive impairment in a transgenic rat model of early Alzheimer's disease. *J. Cereb. Blood Flow. Metab.* 37, 69–84. doi:10.1177/0271678X15615132

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on alzheimer's disease. *Neurology* 34, 939–944. doi:10.1212/wnl.34.7.939

Nadeeshani, H., Li, J., Ying, T., Zhang, B., and Lu, J. (2021). Nicotinamide mononucleotide (NMN) as an anti-aging health product - promises and safety concerns. *J. Adv. Res.* 37, 267–278. doi:10.1016/j.jare.2021.08.003

Nagana Gowda, G. A., Djukovic, D., Bettcher, L. F., Gu, H., and Raftery, D. (2018). NMR-guided mass spectrometry for absolute quantitation of human blood metabolites. *Anal. Chem.* 90, 2001–2009. doi:10.1021/acs.analchem.7b04089

Nilsen, L. H., Melo, T. M., Saether, O., Witter, M. P., and Sonnewald, U. (2012). Altered neurochemical profile in the McGill-R-Thy1-APP rat model of alzheimer's disease: A longitudinal *in vivo* 1 H MRS study. *J. Neurochem.* 123, 532–541. doi:10.1111/jnc.12003

Nilsen, L. H., Melo, T. M., Witter, M. P., and Sonnewald, U. (2014b). Early differences in dorsal hippocampal metabolite levels in males but not females in a transgenic rat model of Alzheimer's disease. *Neurochem. Res.* 39, 305–312. doi:10.1007/s11064-013-1222-x

Nilsen, L. H., Witter, M. P., and Sonnewald, U. (2014a). Neuronal and astrocytic metabolism in a transgenic rat model of Alzheimer's disease. *J. Cereb. Blood Flow. Metab.* 34, 906–914. doi:10.1038/jcbfm.2014.37

Norgren, N., Rosengren, L., and Stigbrand, T. (2003). Elevated neurofilament levels in neurological diseases. *Brain Res.* 987, 25–31. doi:10.1016/s0006-8993(03)03219-0

Nyamundanda, G., Gormley, I. C., Fan, Y., Gallagher, W. M., and Brennan, L. (2013). MetSizeR: Selecting the optimal sample size for metabolomic studies using an analysis based approach. *BMC Bioinforma.* 14, 338. doi:10.1186/1471-2105-14-338

Okabe, K., Yaku, K., Tobe, K., and Nakagawa, T. (2019). Implications of altered NAD metabolism in metabolic disorders. *J. Biomed. Sci.* 26, 34. doi:10.1186/s12929-019-0527-8

Phelan, M. J., Mulnard, R. A., Gillen, D. L., and Schreiber, S. S. (2017). Phase II clinical trial of nicotinamide for the treatment of mild to moderate alzheimer's disease. *J. Geriatr. Med. Gerontol.* 3, 021. doi:10.23937/2469-5858/1510021

Rainer, M., Kraxberger, E., Haushofer, M., Mucke, H. A., and Jellinger, K. A. (2000). No evidence for cognitive improvement from oral nicotinamide adenine dinucleotide (NADH) in dementia. *J. Neural Transm.* 107, 1475–1481. doi:10.1007/s007020070011

Ramirez, A., Wolfsgruber, S., Lange, C., Kaduszkiewicz, H., Weyerer, S., Werle, J., et al. (2015). Elevated HbA1c is associated with increased risk of incident dementia in primary care patients. *J. Alzheimers Dis.* 44, 1203–1212. doi:10.3233/JAD-141521

Salech, F., Ponce, D. P., Paula-Lima, A. C., SanMartin, C. D., and Behrens, M. I. (2020). Nicotinamide, a poly [ADP-Ribose] polymerase 1 (PARP-1) inhibitor, as an adjunctive therapy for the treatment of alzheimer's disease. *Front. Aging Neurosci.* 12, 255. doi:10.3389/fnagi.2020.00255

Therneau, T. M. (2020). *A package for survival analysis in R*. Available at: https://CRAN.R-project.org/package=survival.

Verdin, E. (2015). NAD⁺ in aging, metabolism, and neurodegeneration. *Science* 350, 1208–1213. doi:10.1126/science.aac4854

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–48. doi:10.18637/jss.v036.i03

Wang, H., Robinson, J. L., Kocabas, P., Gustafsson, J., Anton, M., Cholley, P. E., et al. (2021). Genome-scale metabolic network reconstruction of model animals as a platform for translational research. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2102344118. doi:10.1073/pnas.2102344118

Wilkins, J. M., and Trushina, E. (2017). Application of metabolomics in alzheimer's disease. *Front. Neurol.* 8, 719. doi:10.3389/fneur.2017.00719

Wilson, E. N., Abela, A. R., Do Carmo, S., Allard, S., Marks, A. R., Welikovitch, L. A., et al. (2017b). Intraneuronal amyloid beta accumulation disrupts hippocampal CRTC1-dependent gene expression and cognitive function in a rat model of alzheimer disease. *Cereb. Cortex* 27, 1501–1511. doi:10.1093/cercor/bhv332

Wilson, E. N., Do Carmo, S., Iulita, M. F., Hall, H., Austin, G. L., Jia, D. T., et al. (2018). Microdose lithium NP03 diminishes pre-plaque oxidative damage and neuroinflammation in a rat model of alzheimer's-like amyloidosis. *Curr. Alzheimer Res.* 15, 1220–1230. doi:10.2174/1567205015666180904154446

Wilson, E. N., Do Carmo, S., Iulita, M. F., Hall, H., Ducatenzeiler, A., Marks, A. R., et al. (2017a). BACE1 inhibition by microdose lithium formulation NP03 rescues memory loss and early stage amyloid neuropathology. *Transl. Psychiatry* 7, e1190. doi:10.1038/tp.2017.169

Xing, S., Hu, Y., Huang, X., Shen, D., and Chen, C. (2019). Nicotinamide phosphoribosyltransferaserelated signaling pathway in early Alzheimer's disease mouse models. *Mol. Med. Rep.* 20, 5163–5171. doi:10.3892/mmr.2019.10782

Yan, X., Hu, Y., Wang, B., Wang, S., and Zhang, X. (2020). Metabolic dysregulation contributes to the progression of alzheimer's disease. *Front. Neurosci.* 14, 530219. doi:10.3389/fnins.2020.530219

Zaudig, M., Mittelhammer, J., Hiller, W., Pauls, A., Thora, C., Morinigo, A., et al. (1991). SIDAM--A structured interview for the diagnosis of dementia of the Alzheimer type, multi-infarct dementia and dementias of other aetiology according to ICD-10 and DSM-III-R. *Psychol. Med.* 21, 225–236. doi:10.1017/s0033291700014811

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 03 March 2023
DOI 10.3389/fmolb.2023.1100486

OPEN ACCESS

EDITED BY
Guillermo Moyna,
Universidad de la República, Uruguay

REVIEWED BY
Tae Jin Lee,
Augusta University, United States
Monica Cala,
University of Los Andes, Colombia, Colombia

*CORRESPONDENCE
Yamilé López-Hernández,
ylopezher@conacyt.mx
Juan José Oropeza-Valdez,
lmcuazjjov@gmail.com

SPECIALTY SECTION
This article was submitted to Metabolomics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 16 November 2022
ACCEPTED 16 February 2023
PUBLISHED 03 March 2023

CITATION
López-Hernández Y, Oropeza-Valdez JJ, García Lopez DA, Borrego JC, Murgu M, Valdez J, López JA and Monárrez-Espino J (2023), Untargeted analysis in post-COVID-19 patients reveals dysregulated lipid pathways two years after recovery.
Front. Mol. Biosci. 10:1100486.
doi: 10.3389/fmolb.2023.1100486

Untargeted analysis in post-COVID-19 patients reveals dysregulated lipid pathways two years after recovery

Yamilé López-Hernández[1]*, Juan José Oropeza-Valdez[2]*, David Alejandro García Lopez[3], Juan Carlos Borrego[4], Michel Murgu[5], Jorge Valdez[6], Jesús Adrián López[7] and Joel Monárrez-Espino[8]

[1]CONACyT-Metabolomics and Proteomics Laboratory, Academic Unit of Biological Sciences, Autonomous University of Zacatecas, Zacatecas, Mexico, [2]Metabolomics and Proteomics Laboratory, Academic Unit of Biological Sciences, Autonomous University of Zacatecas, Zacatecas, Mexico, [3]Doctorado en Ciencias Básicas, Universidad Autónoma de Zacatecas, Zacatecas, Mexico, [4]Departamento de Epidemiología, Hospital General de Zona #1 "Emilio Varela Luján", Instituto Mexicano del Seguro Social, Centro, Zacatecas, Mexico, [5]Waters Technologies of Brazil, Alameda Tocantins, Barueri, Brazil, [6]Waters Corporation, Mexico City, Mexico, [7]MicroRNAs and Cancer Laboratory, Academic Unit of Biological Sciences, Autonomous University of Zacatecas, Zacatecas, Mexico, [8]Department of Health Research, Christus Muguerza del Parque Hospital Chihuahua, University of Monterrey, San Pedro Garza García, Mexico

Introduction: Similar to what it has been reported with preceding viral epidemics (such as MERS, SARS, or influenza), SARS-CoV-2 infection is also affecting the human immunometabolism with long-term consequences. Even with underreporting, an accumulated of almost 650 million people have been infected and 620 million recovered since the start of the pandemic; therefore, the impact of these long-term consequences in the world population could be significant. Recently, the World Health Organization recognized the post-COVID syndrome as a new entity, and guidelines are being established to manage and treat this new condition. However, there is still uncertainty about the molecular mechanisms behind the large number of symptoms reported worldwide.

Aims and Methods: In this study we aimed to evaluate the clinical and lipidomic profiles (using non-targeted lipidomics) of recovered patients who had a mild and severe COVID-19 infection (acute phase, first epidemic wave); the assessment was made two years after the initial infection.

Results: Fatigue (59%) and musculoskeletal (50%) symptoms as the most relevant and persistent. Functional analyses revealed that sterols, bile acids, isoprenoids, and fatty esters were the predicted metabolic pathways affected in both COVID-19 and post-COVID-19 patients. Principal Component Analysis showed differences between study groups. Several species of phosphatidylcholines and sphingomyelins were identified and expressed in higher levels in post-COVID-19 patients compared to controls. The paired analysis (comparing patients with an active infection and 2 years after recovery) show 170 dysregulated features. The relationship of such metabolic dysregulations with the clinical symptoms, point to the importance of developing diagnostic and therapeuthic markers based on cell signaling pathways.

KEYWORDS

long covid, metabolomics, metabolite, lipidomics, post-COVID-19

1 Introduction

Post-COVID-19 exists, and matters. The World Health Organization (WHO) has recently recognized the post-COVID-19 (or long COVID) condition (Delphi consensus), as one "that occurs in individuals with a previous history of probable or confirmed SARS-CoV-2 infection, usually three months after the onset, with symptoms lasting at least two months that cannot be explained by an alternative diagnosis" (Soriano et al., 2022). Studies show that 60% of COVID-19 survivors experience post-COVID symptoms (Fernández-de-Las-Peñas et al., 2021), and that these are associated with worse quality of life (Malik et al., 2022).

Why some patients experience long-term symptoms after COVID-19 infection remains uncertain (Yong, 2021). Genetic susceptibility, age, and viral load could be related to this syndrome (Collaborators, 2022). Organ damage due to an excessive inflammatory response caused by the virus, persistent reservoirs of SARS-CoV-2 in certain tissues triggering post-infection morbidity, pathogen reactivation as a result of immune dysregulation, host-microbiome alterations, coagulation problems, and autoimmunity due to molecular mimicry between SARS-CoV-2 and autoantibodies could be involved mechanisms (Collaborators, 2022). In addition, it has also been suggested that the prolonged symptoms of COVID-19 may not be a direct result of SARS-CoV-2 infection, but rather the consequence of Epstein-Barr virus reactivation induced by COVID-19 inflammation (Ortona and Malorni, 2022).

Multiple studies point to an immunometabolic dysregulation, highlighting certain metabolites, such as those involved in the tryptophan and kynurenine pathways, which have a very important role in the immune system. To date, the plethora of symptoms and protracted disorders documented suggest that various concurrent mechanisms might be involved, and that different therapeutic approaches need to be established when dealing with these patients.

It remains unknown whether SARS-CoV-2 can cause substantial tissue damage leading to a chronic form of the disease, such as the chronic convalescent lesions seen with other viral infections including the Human Immunodeficiency Virus (HIV), hepatitis C virus (HCV), hepatitis B virus (HBV) and some herpesviruses. Previous studies with SARS survivors have shown lung abnormalities months after infection (Lopez-Leon et al., 2021).

A recent comprehensive molecular investigation revealed extensive inflammation and degeneration in the brains of patients who died of COVID-19, even among those with no reported neurological symptoms. In that study, the authors report that SARS-CoV-2 virus induced vascular damage affecting endothelial cells and caused generalized neuroinflammation. Cytokines like interleukin-1 and interleukin-6 were highly elevated in patients with COVID-19, which are the ones driving neurodegeneration and Alzheimer's disease (Reiken et al., 2022).

Another study showed that between one and 12 months after infection, patients recovered from COVID-19 are at increased risk of incident cardiovascular disease, including cerebrovascular disorders, arrhythmias, ischemic and non-ischemic heart disease, pericarditis, myocarditis, heart failure, and thromboembolic events; these risks were documented even among people who were not hospitalized during the acute phase of the infection, and increased gradually depending on the care setting during the acute phase (non-hospitalized, hospitalized, and intensive care) (Xie et al., 2022).

In terms of mortality, a large study comprising nearly five million healthy controls and 90,000 COVID-19 patients revealed that the risk of death among COVID-19 survivors in the following 6 months after infection, increased by 60% (Al-Aly et al., 2021).

In the present work, we investigated the health status of patients who recovered from a mild, severe, and critical COVID-19 in 2020 (post-COVID-19 patients). Two years later, patients who provided informed consent were surveyed. Clinical symptomatology and blood samples for laboratory analyses were obtained. Untargeted lipidomic analysis was performed with plasma samples to assess potential dysregulation of lipid metabolism. These long-term alterations need to be deeply analyzed to find a possible connection with symptoms persistence and to find effective therapeutic alternatives to treat (or cure) these patients.

2 Materials and methods

2.1 Patients' recruitment

Symptomatic individuals aged 35–70 years who were RT-qPCR-tested for SARS-CoV-2 between March 15 and 1 November 2020, at the Zacatecas General Hospital's Respiratory Triage Unit of the Mexican Institute of Social Security (IMSS) and Christus Muguerza del Parque Hospital of Chihuahua city were included in this study. Negative controls were RT-qPCR negative patients. Inclusion criteria for negative controls, and mild, severe, and critically ill patients are shown in Supplementary Table S1.

The COVID-19 group were patients who were positive for SARS-CoV-2 in the first epidemic wave (2020). Blood specimens for plasma isolation were collected within two days after hospital admission on average. Baseline information including age, sex, comorbidities, clinical and laboratory data, and disease severity classification according to WHO guidelines (WHO, 2023).

Clinical data was obtained from the electronic medical records of each patient and stored by a password-protected database and provided in Table 1. Blood samples were collected in 2020 and stored at −80°C in the biobank at the Autonomous University of Zacatecas, Mexico.

Two years after hospital discharge and recovery, plasma samples were obtained from 22 COVID-19 patients (post-COVID-19 group). Chest computed tomography (CT) scans (in patients that had a baseline CT), basic blood biochemical markers (i.e., hemoglobin, platelets, leukocytes, lymphocytes, and creatinine), and a questionnaire to assess the persistence of clinical symptoms were used to evaluate their clinical recovery. Blood collection was done in fasting conditions. Only three patients had one reinfection 20 months on average after recovery from the initial infection.

The study was conducted in accordance with the Declaration of Helsinki (1976). It was also revised and approved by the Research and Ethics Committees of the Instituto Mexicano de Seguridad Social, with the registration number R-2022-3301-038 and Christus Muguerza del Parque Hospital (folio HCMP-CEI-28022022-A01 and HCMP-CEI-15042020-3). Informed consent was

TABLE 1 Clinical and demographic characteristics.

Variable	Negative controls (n = 15)	COVID-19 (n = 28)	Post-COVID-19 (n = 20)	*p*-value
Age, mean ± SD (years)	47.2 ± 8.4	56.3 ± 13	51.8 ± 11.6	0.0537
Male gender, n (%)	9 (60)	15 (53.5)	11 (55)	0.9198
Smoking, n (%)	3 (20)	2 (7.6)	3 (15)	0.4501
Comorbidities (self-reported), n (%)				
Diabetes	1 (6.66)	10 (35.7)	3 (15)	0.0592
Hypertension	4 (26.6)	8 (28.5)	9 (45)	0.4044
Obesity	1 (6.6)	8 (28.5)	0 (0)	**0.0128** *b
Symptomatology, n (%)				
Fever	0 (0)	15 (53.5)	0 (0)	**<0.0001** *a, b
Cough	0 (0)	24 (85.7)	3 (15)	**<0.0001** *a, b
Headache	10 (66)	18 (64)	6 (3)	**0.0334** *b, c
Dyspnea	2 (13.3)	23 (82.1)	6 (3)	**<0.0001** *a, b
Diarrhea	0 (0)	6 (21.4)	2 (10)	0.1201
Chest tightness	0 (0)	14 (50)	4 (20)	**0.0015** *a, b
Pharyngalgia	8 (53.3)	9 (32.1)	2 (10)	**0.0209** *c
Myalgia	8 (53.3)	18 (64.2)	9 (45)	0.4072
Arthralgias	5 (33.3)	18 (64.2)	10 (50)	0.1482
Anosmya	0 (0)	5 of 14 (35)	1 (5)	**0.006** *a, b
Laboratory data, median (Q1-Q3)				
Erythrocytes (million/mL)	5.3 (5–5.5)	4.8 (1.2–5.32)	5 (0.9–5.48)	0.0935
Hemoglobin (g/dL)	15.5 (14.9–16.4)	13.25 (9.78–15.2)	15.45 (14.63–16.55)	**0.001** *a, b
Platelets (thousands/mL)	277 (242–320)	233.5 (145.5–306.5)	237 (216–248)	0.1183
Leukocytes ($\times 10^3$)	6.6 (5.9–7.8)	8.3 (5.08–11.68)	7.23 (6.38–7.98)	0.4093
Lymphocytes (%)	33.3 (25.8–37)	7 (2.25–12.8)	34.95 (29.88–38.25)	**<0.0001** *a, b
Monocytes (%)	6.5 (5.4–8.3)	2.95 (0.4–4.78)	6.3 (1.16–7.1)	**0.0036** *a
Neutrophils (%)	58 (51.3–62.7)	81.45 (26.56–91.88)	51.1 (8.82–58.1)	**<0.0001** *a, b
Glucose (mg/dL)	91 (84–116)	123 (80–191.3)	100.3 (19.02–131.9)	0.1314
Creatinine (mg/dL)	1 (0.7–1.1)	0.67 (0.5–0.78)	0.795 (0.7–0.92)	**0.0083** *a

*a: Negative Controls vs. COVID-19.
*b: COVID-19, vs. post-COVID-19.
*c: Negative Controls vs. post-COVID-19.
Significant values ($p < 0.05$) are highlighted in bold.

obtained from all participants. All patients included in this study were informed in writing regarding the collection of their samples for research aims and were given the right to refuse participation.

2.2 Sample preparation

Blood collected in vacutainer tubes (EDTA) was centrifuged at 4°C and 3000 g for 15 min. Plasma was aliquoted and stored at −80°C until use. Plasma thawed in ice was extracted with pre-cooled isopropanol in a 1:3 ratio (LCMS grade, Honeywell, Charlotte, NC, United States) (Medina et al., 2020) vortexed for 1 min and incubated at −20°C overnight for protein precipitation. Subsequently, the extraction mix was centrifuged at 16000 *g* and 4°C for 15 min and supernatants were carefully collected. An additional step of centrifugation was done to remove any debris collected. For the analysis, each aliquot was transferred into certified LC vials (TruView LCMS, United States) and diluted to 1:20 ratio with a mixture of isopropanol/acetonitrile/water (2:1:1, v: v: v). Sample preparation order was randomized for sample picking to ensure no systematic biases were present during sample preparation.

2.3 Quality controls (QC) and quality assurance (QA)

These processes are referred to as the procedures applied in preparation for data acquisition (QA) and during/after data acquisition (QC) (Kirwan et al., 2022). As part of QA procedures, the equipment was subjected to a complete maintenance twice a year. This maintenance included both the chromatography system, the mass analyzer, and the nitrogen source. Sample cone and ion source cleaning were performed between every analytical batch. Calibration and manual tuning were also performed immediately before running samples. Temperature control, standardized protocols of operations and qualification of technical staff were also considered.

A pool of human plasma from all participants in the study served as a technical replicate throughout the dataset (pooled QCs). QCs were prepared identically as individual samples. Overall process variability was determined by calculating the median relative standard deviation (RSD) for all endogenous metabolites present in 100% of the pooled QCs samples.

Experimental samples were randomized across the platform run with 10 QCs samples at the beginning (for instrument and column equilibration) and one QC sample was acquired every ten samples injected.

2.4 Ultra-performance liquid chromatography (UPLC)-Mass spectrometry method for lipidomic analysis

The analysis was performed using an ACQUITY UPLC I-Class (Waters Corp., Milford, MA, United States) coupled to a XEVO-G2 XS quadrupole time-of-flight (ToF) mass spectrometer (Waters, Manchester, NH, United States) with an electrospray ionization source. The samples were analyzed in positive (ESI+) mode.

A UPLC CSH C18 column (2.1 × 100 mm, 1.7 μm) with a binary gradient elution of solvents was used for lipid separation. The mobile phase A was 10 mM ammonium formate with 0.1% formic acid in acetonitrile/water (60:40, v:v) and mobile phase B was 10 mM ammonium formate with 0.1% formic acid in isopropanol/acetonitrile (90:10, v:v). The mobile phases were delivered at a flow rate of 0.3 mL/min, initially with 60% A, followed by a linear gradient to 57% A over 2 min, and then the percentage of A was decreased to 50% within 0.5 min. Over the next 10 min, the gradient was ramped to 46% A, and the amount of A was then decreased to 30% in 0.5 min. Over 6 min, the amount of A decreased to 1%, and returned to initial conditions (60%) at the end of 25 min. The column temperature was adjusted to 55°C and the injection volume was five uL. Data was acquired using positive electrospray ionization mode with the capillary voltage set to 3.2 kV, the cone voltage to 40 eV and the source temperature to 130 °C. The desolvation gas was nitrogen, with a flow rate of 900 L/h, cone gas flow of 25 L/h and temperature of 550°C. Data was acquired in the m/z range of 50–1,200 in data independent analysis (DIA) mode in which the collision energy was alternated between low energy (6 eV) and high energy (ramped from 10–40 eV) in consecutive scans of 0.2 s generating high and low chromatograms and spectra. Lockmass correction was made by the acquisition of mass reference leucine enkephalin in intervals of 30 s.

2.4.1 Data analysis

Raw data were processed under default parameters as a UNIFI file (UNIFI 1.8.2, Waters Corp., Milford, United States), which was exported to Progenesis QI (version 3.0.7, Waters Corp., Milford, United States). For the alignment, retention times below 0.5 min and after 18 min were excluded. A width peak of 0.06 s was defined. Deconvolution was automatically performed, considering M + H, M + Na, M + H-H20, M + K, and M + NH4 as adducts. However, manual inspection was done, eliminating those features with incorrect alignment in chromatograms and neutral and m/z mass. An excel file was exported and a signal to noise (S/N) ratio was calculated for each sample based on the extraction blank. All features with a S/N < 5 in the 80% of samples were eliminated. Besides, RSD was calculated taking QCs as references. Features with RSD > 20% were also eliminated.

2.5 Lipid identification

A manual inspection about putative identification was done by searching accurate mass in HMDB (https://hmdb.ca), LipidBlast (https://fiehnlab.ucdavis.edu/projects/lipidblast) and METLIN (https://metlin.scripps.edu). Putative identification was assigned based on accurate mass, retention time, and fragmentation patterns [Progenesis QI (version 3.0.7, Waters Corp., Milford, United States)]. Confidence levels in annotation were as following: level 4 (molecular formula): molecular formula identification of features is completed *via* isotope abundance distribution, charge state and adduct ion determination. Level 3 (tentative structure): tentative structural identification includes a unique match of the parent ion (MS1) data searched through literature and/or libraries and databases. Level 2 (putative identification): putative identification reveals probable structure using fragmentation data from literature and/or libraries and databases (Schrimpe-Rutledge et al., 2016). For the significant features putatively identified, a MS/MS method was performed. Briefly, precursor ions were fragmented with collision energies 10 eV, 20 eV, and 40 eV. Mass spectra were analyzed and based on the fragmentation pattern; an identification (level 2–4 of confidence) was assigned.

2.6 Statistical analysis

Medians with interquartile ranges (IQRs) or means [with standard deviation (s.d.)] and frequencies (%) were used to describe baseline characteristics of non-COVID-19 subjects, and COVID-19 or post-COVID-19 patients for continuous and nominal data, respectively. Normality was assessed using the D'Agostino-Pearson normality test. Continuous variables were analyzed using Mann-Whitney U or Kruskal-Wallis tests. For nominal variables (e.g., sex, smoking, death, symptoms, and comorbidities) chi-square tests for trends were used. All *p*-values less than 0.05 considered statistically significant. Analyses were conducted using GraphPad Prism version 8.0.1 for Windows (GraphPad Software, La Jolla California United States).

For lipidomics data, functional and statistical analyses were performed with MetaboAnalyst 5.0 (https://www.metaboanalyst.ca). After filtering and eliminating possible artifacts or redundancies, data was normalized by sum (TIC), transformed by square, and scaled by range.

Mummichog pathway activity profile was done with the intention to reveal the most important metabolic pathway altered without dealing with identification of all features. Mass tolerance was set to 5 ppm, *p*-value: 10^{-4} cut off (default top 10% peaks), pathway library: *homo sapiens* (human) metabolite sets for lipids (main chemical classes and sub-classes) and metabolic pathways with at least three entries were considered. Metabolite sets were manually curated and originate from a number of sources (KEGG, BiGG, and Edinburgh Model).

Analysis of continuous and categorical data was performed by Mann-Whitney rank sum and Fisher's exact tests, respectively. Adjusted *p*-values (false discovery rate, FDR) < 0.05 were considered as significant. Univariate analysis of covariance (ANCOVA) was conducted in SPSS (version 29, SPSS Inc., United States) to examine the differences between post-COVID,

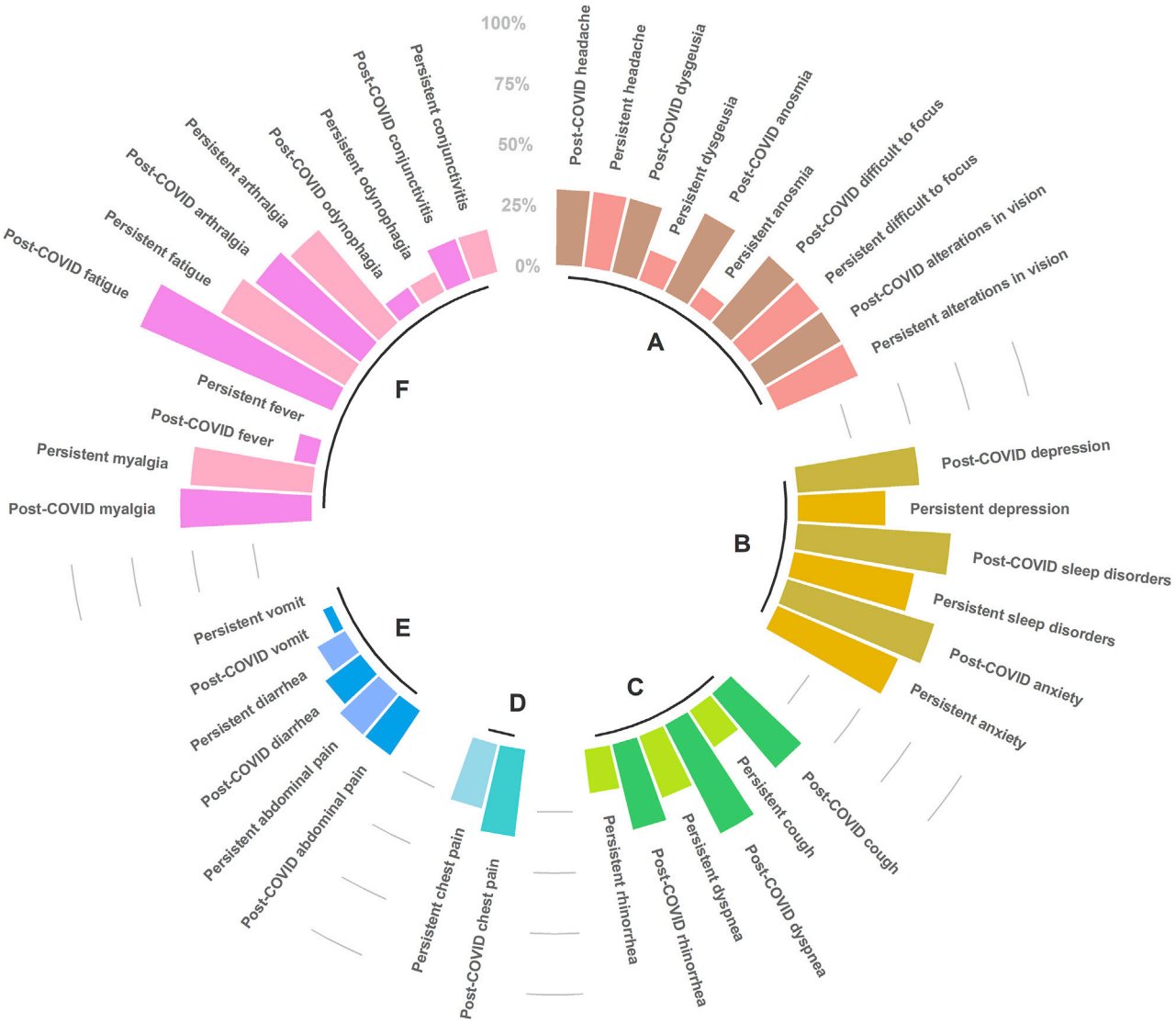


FIGURE 1
Post-COVID-19 (immediately after infection resolution), and persistent symptoms (until September 2022), clustered by etiology condition. **(A)** Neurological, **(B)** Psychiatric, **(C)** Respiratory, **(D)** Cardiac, **(E)** Digestive, **(F)** Systemic. The scale (0%–100%) represents the percentage of patients reporting each particular symptom. Figure was built with R package (ggplot2). Only symptoms that were non-significant (chi^2 test) between post-COVID-19 and persistent status were represented.

COVID-19 patients and negative controls, including all features from the lipidomic dataset adjusted by age and comorbidities that were reported as significant (presence or absence of diabetes and obesity).

Principal component analysis (PCA) and two-dimensional partial least squares discriminant analysis (2D PLS-DA) scores plots were used to compare plasma lipidomic data across and between study groups; 2000-fold permutation tests were used to minimize the possibility that the observed separation of the PLS-DA was due to chance. Discriminant Q^2 (DQ^2) is an improvement for the Q^2 value used in the validation of PLSDA models since it does not penalize class predictions beyond the class label value. DQ^2 estimations were performed using Matlab (version 2020B, The MathWorks, United States) using the DQ^2 Matlab routine written by Westerhuis et al. (Westerhuis et al., 2008) (http://www.bdagroup.nl/). Variable importance in projection (VIP) and heat maps were also plotted. Significant features were considered when having a VIP score >1.5 and a FDR <0.05.

3 Results

Table 1 shows the baseline characteristics of negative controls, COVID-19 patients (acute phase), and post-COVID-19 patients (2 years after recovery). At the time of the clinical examination, none of these patients had an active infection.

Among the 20 patients with post-COVID-19, 55% were male and the mean age was 51.8 ± 11.6 years. Four (27.3%) patients had

10.3389/fmolb.2023.1100486

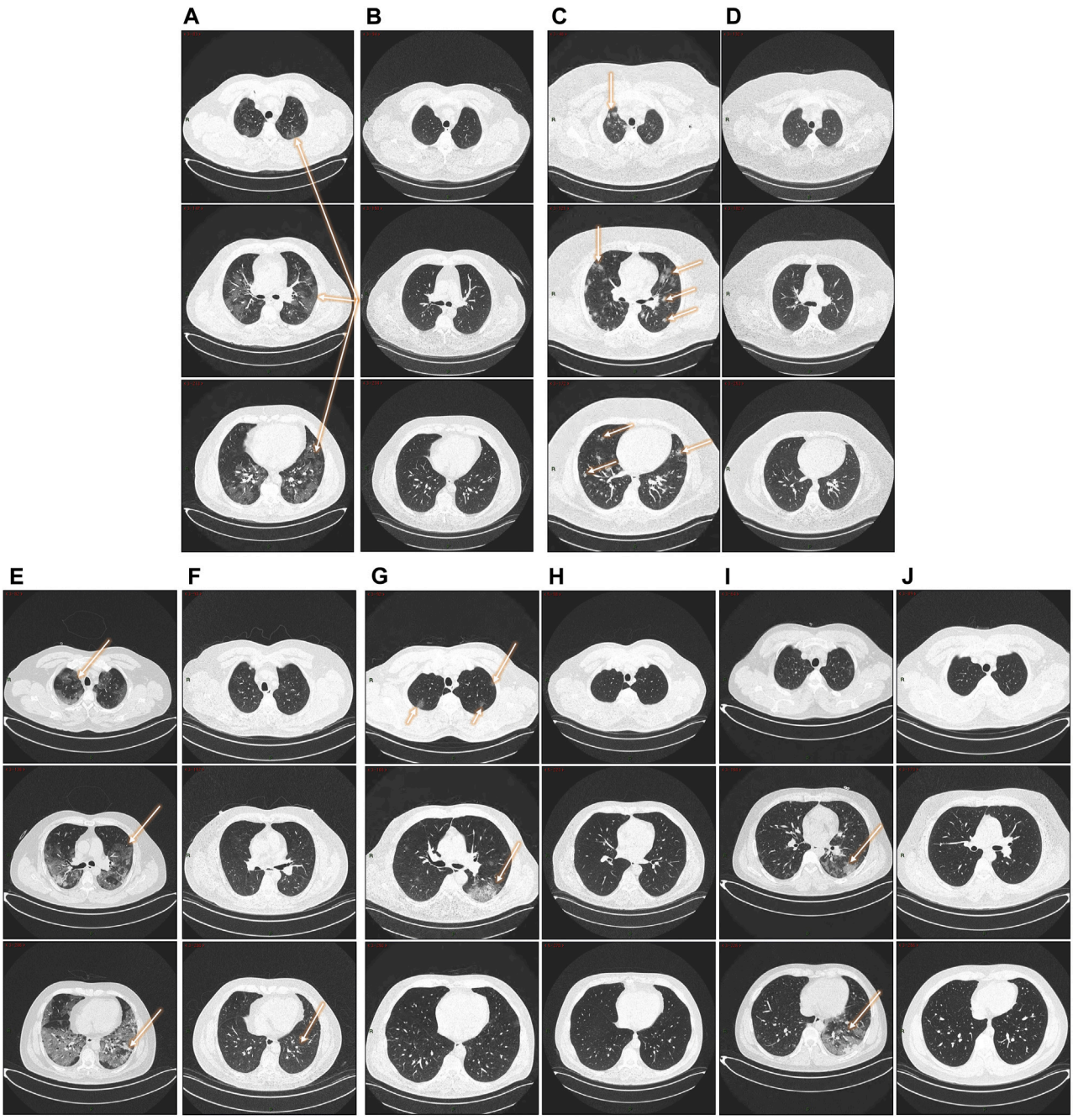


FIGURE 2
Follow-up lung CTs in patients with different degree of severity. **(A, B):** 47-year- old male patient (severe disease). **(A)** 2020. Pneumonia with presence of a ground glass pattern in the three levels of the parenchyma, with a predominant injury in the left basal region (arrows) with scattered areas of consolidation. **(B)** 2022. Complete resolution of radiological findings. **(C, D):** five- year-old male patient (severe disease). **(C)** 2020. Multiple nodular opacities in the right upper lobe (arrow) with ground glass images in the middle level of the parenchyma (coronal section) and predominance the left lobe. At the basal level, the affected areas are smaller with a diffuse nodular pattern. **(D)** Complete resolution of radiological findings. **(E, F)**: 51-year- old male patient (severe disease). **(E)** 2020. Pneumonia with the presence of a pattern in ground glass patches in the three levels of the parenchyma, with injury predominantly in the left basal region (arrows) with areas of consolidation and a pattern of bronchoalveolar distention. **(F)** 2022. Almost complete resolution of COVID-19 pneumonia with a mild residual interstitial pattern, predominantly at the basal level, coronal section (arrow). **(G, H):** 51-year-old male patient (severe disease). **(G)** 2020. Ground glass-type opacities are identified in the middle level of the parenchyma in coronal section, left lobe (arrow). At the apical region, a diffuse pattern of cotton-nodular appearance. **(H)** 2022. Complete resolution of the radiological findings. **(I, J):** 30-year-old male patient (severe disease). **(I)** 2020. Pneumonia with presence of a minimally affected pattern in the apical area with barely perceptible frosted glass smears, diffuse ground-glass patches at the mid-level of the parenchyma, with considerable involvement of left basal region (arrows). **(J)** 2022: Complete resolution of radiological findings.

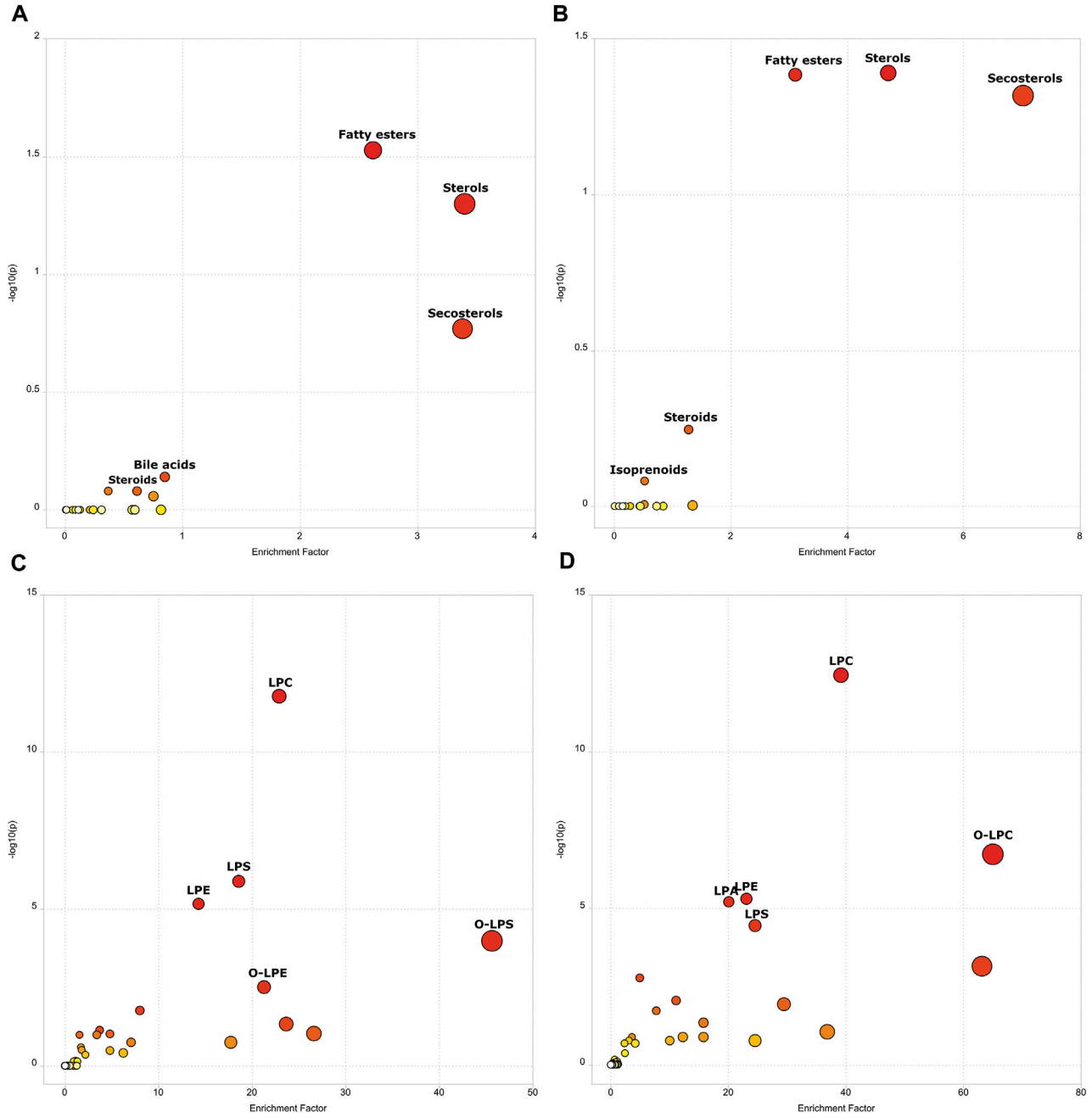


FIGURE 3
Scatter Plot representing Mummichog Pathway Activity Profile showing the enrichment factor. The enrichment factor is the ratio between the number of significant pathway hits and the expected number of hits within the pathway. **(A)** Negative controls vs COVID-19 (lipids main chemical class). **(B)** Negative controls vs Post-COVID-19 (lipids main chemical class). **(C)** Negative controls vs COVID-19 (lipids sub chemical class). **(D)** Negative controls vs Post-COVID-19 (lipids sub chemical class). Size and color increases as -log10(*p*) and enrichment factor increases. Figure was built with MetaboAnalyst software v 5.0 (https://www.metaboanalyst.ca).

developed a mild disease during the acute phase, 11 (50%) had severe disease, and five (22.7%) were critically ill requiring intubation. After two years, most of the laboratory parameters were normal. Lymphocytes, monocytes and neutrophils, which were altered parameters during the acute phase, showed statistical differences with the post-COVID phase. Various clinical symptoms persisted after two years of recovery; patients basically reported the same symptoms of the acute phase, except for vomiting and fever (Figure 1).

Patients with an abnormal chest CT in 2020 were taken a follow-up CT scan to assess pulmonary sequelae. Representative CT images of the two COVID-19 groups in 2020 and 2022 are shown in Figure 2. While some patients had a complete resolution of abnormal findings two years after the initial infection, others had

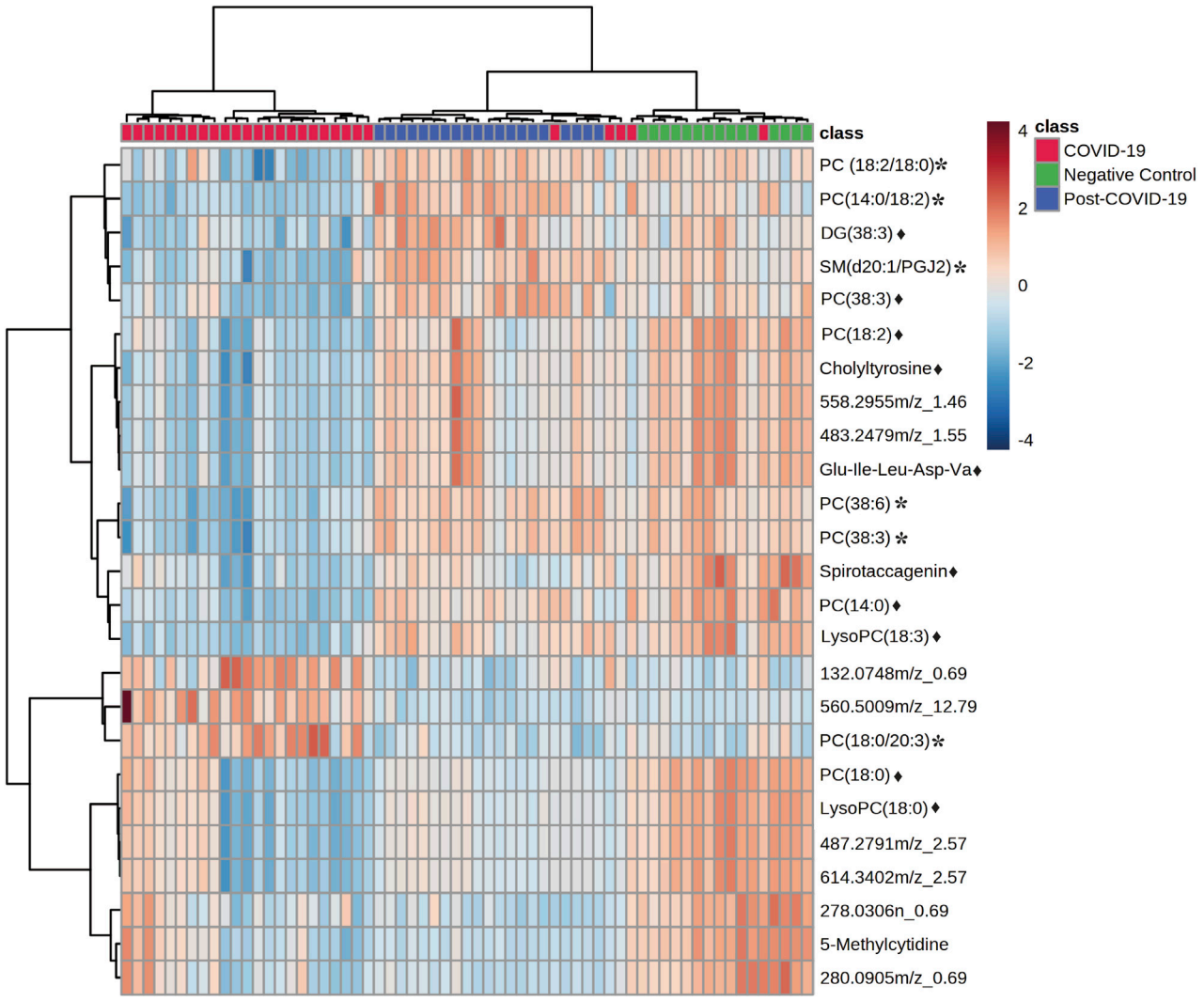


FIGURE 4
Heat map representing the most important features differentiating negative controls, post-COVID-19 patients, and COVID-19 patients. Data was normalized, distance measure: Euclidean, clustering method: Ward. Ranked by ANOVA test. Figure was built with MetaboAnalyst software v 5.0 (https://www.metaboanalyst.ca). * Represents features with confidence level 2 and ◆ represents features with level 3.

persistent lung abnormalities (interstitial thickening, ground glass opacity, and subpleural bands).

3.1 Functional analysis of untargeted lipidomics data generated from high-resolution mass spectrometry (HRMS)

With the goal to know the most important metabolic pathways dysregulated, both in the infection and recovery phases, a functional analysis (mummichog pathway activity profile) was done. A total of 401 features were detected after filtering as previously described. When the mummichog pathway activity profile was performed, putative dysregulated metabolic pathways were associated with both states (acute phase and post-COVID-19) with respect to negative controls (Figure 3). When analyzing active infection (COVID-19 group) and recovery (post-COVID-19 patients), fatty esters, sterols, secosterols and steroids were dysregulated. For active COVID-19 patients, bile acids pathway was dysregulated, and for post-COVID-19, it was the isoprenoids pathway. When subclasses of lipids were represented, both in the case of acute phase and post-COVID phase, phospholipids were the most important family of compounds dysregulated. Supplementary Tables S2, S3 show the specific details for these metabolic pathways, respectively.

3.2 Univariate and hierarchical clustering analysis

Significant statistical differences between the three study groups were observed. In total, 306 features were significantly dysregulated; 97 showed differences (FDR < 0.05) between controls and post-COVID-19, while 251 were different between post-COVID-19 and

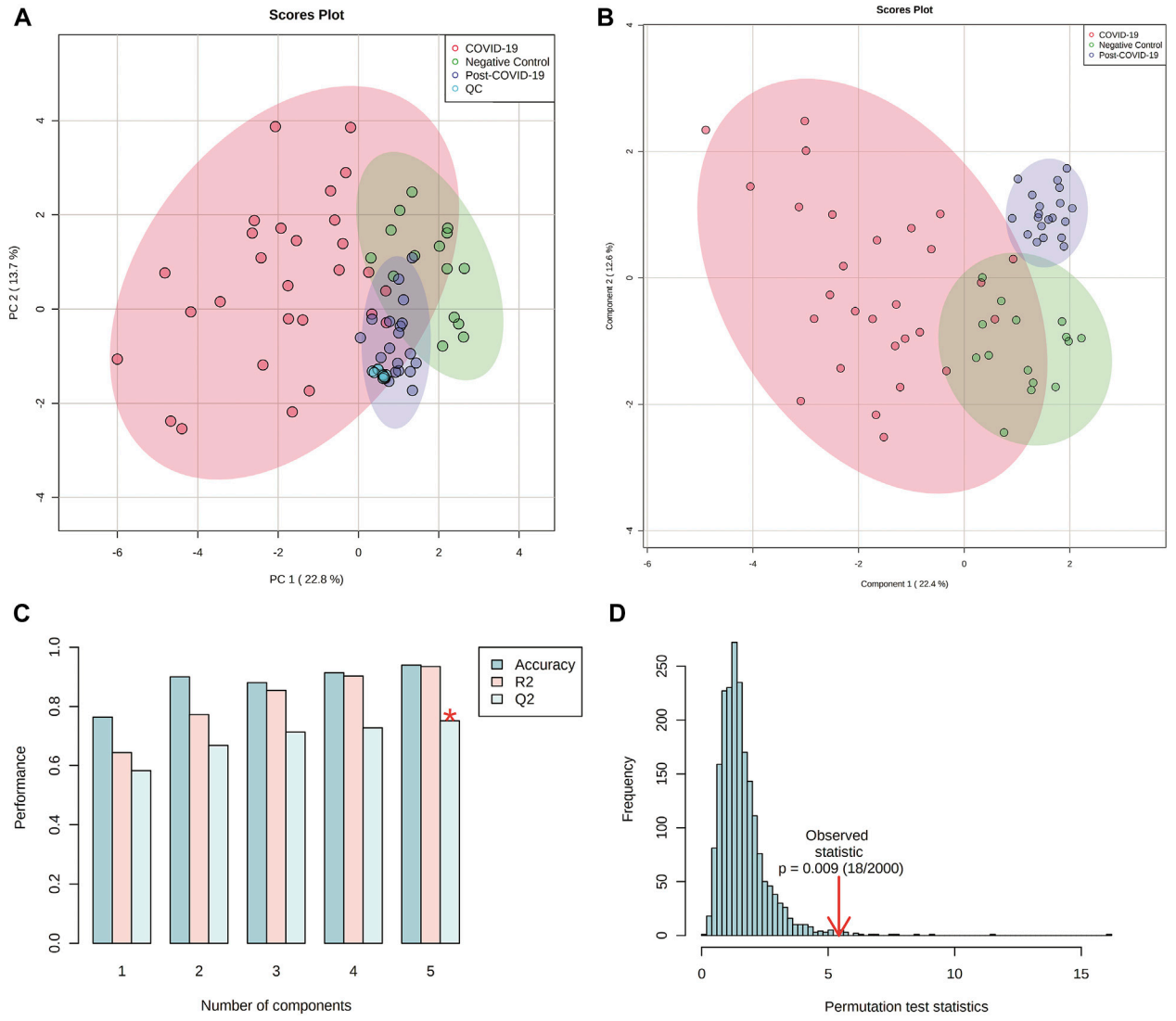


FIGURE 5
Multivariate analyses from plasma lipidomics profile of negative controls, COVID-19 patients, and post-COVID-19 patients. **(A)** Score scatter plot based on PCA models including quality controls (QCs) samples **(C)**. **(B)** Score scatter plot based on PLS-DA models. **(C)** Cross validation (10-foldCV). **(D)** permutation test (2000 permutations). Figures were produced in Metabo Analyst software v 5.0 (https://www.metaboanalyst.ca/).

COVID-19. With respect to controls, 239 features were different compared with COVID-19 patients. ANCOVA replicated the findings of ANOVA after adjustment by age, diabetes, or obesity for the putative features (Supplementary Table S4). The heatmap revealed differences between the COVID-19 group compared with the control and post-COVID-19 groups, but there were also differences between post-COVID-19 patients and the negative controls (Figure 4).

3.3 Multivariate analysis

The multivariate analysis showed a clear separation between the three groups. Principal Component Analysis (PCA) revealed a good clustering in QC samples indicating that technical reproducibility and stability of the system was achieved during the analysis. Once the QCs were inspected, they were eliminated from subsequent analysis. Partial Least Square Discriminant Analysis (PLS-DA) also showed good discrimination. The performance for this model was evaluated by 10-fold cross-validation and permutation test showing no overfitting (Figure 5). Double-check of the models was done as a validation resource of diagnostic statistics for PLS-DA. DQ^2 were as follow: 0.65, 0.79 and 0.73 for the comparisons between: controls vs COVID-19, controls vs Post-COVID, and post-COVID vs COVID-19, respectively.

VIP plots showed the most differentiated features with higher concentrations were found in the post-COVID-19 group (Figure 6). Details about m/z, retention time, adducts, and adjusted *p*-value are shown in Supplementary Table S4. Tandem analyses (MS/MS) were performed for the most important and abundant ions. Figure 7 and Supplementary Table S5 shows a representation of the

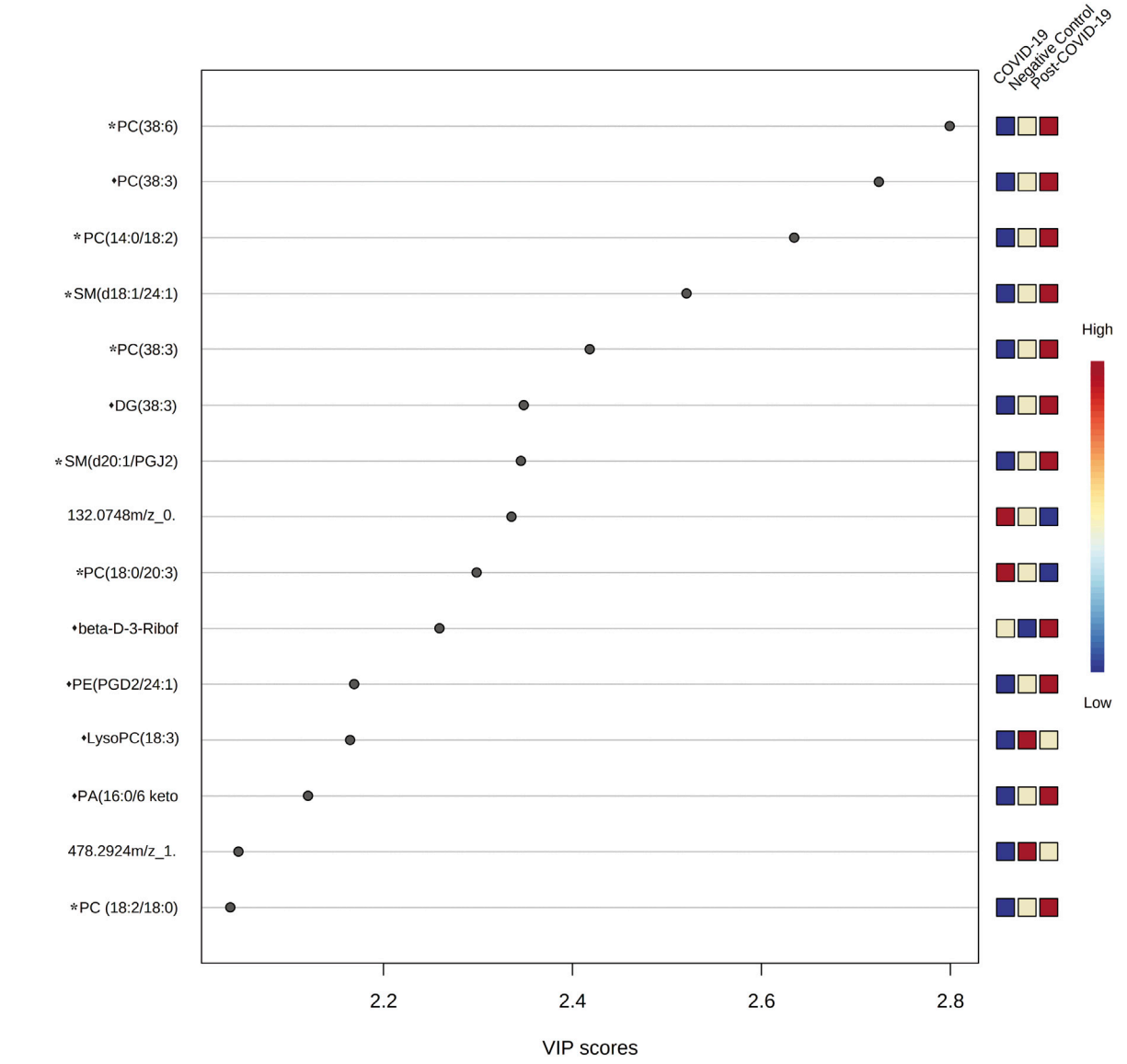


FIGURE 6
Variable importance in Projection (VIP) plot representing the rank of 15 features identified by PLS-DA according to VIP score on *x*-axis. The most discriminating metabolites are shown in descending score order. The color boxes indicate whether metabolite concentration was increased (red) or decreased (blue). Figures were produced in MetaboAnalyst software v 5.0 (https://www.metaboanalyst.ca/). * Represents features with confidence level 2 and ◆ represents features with level 3.

fragmentation pattern leading to the identification of glycerophospholipids and sphingomyelins.

3.4 Paired analysis

Thirteen patients with paired samples (acute phase and post-COVID-19) were included; 38.4% male, and mean age 54.6 ± 7.9 years. Three (23%) developed a mild disease during the acute phase, 8 (61%) had a severe disease, and two (15%) were critically ill requiring intubation (Table 2). After two years, most laboratory parameters became normal. Lympochytes counts, altered during acute phase, was statistically different compared with the post-COVID phase. After two years of recovery, patients reported the same symptoms experienced during the acute immediate recovery phase, except for vomit and fever.

Paired T-tests) showed 170 features dysregulated (FDR<0.05). The volcano plot revealed that 55 features were upregulated in post-COVID-19 patients, and 172 remained downregulated.

Figure 8 shows the multivariate analysis. PCA (Figure 8A) revealed a good clustering of samples. Partial Least Square Discriminant Analysis (PLS-DA) also showed good discrimination; the performance for this model was evaluated by 10-fold cross-validation and permutation test showing no overfitting (Figures 8B–D).

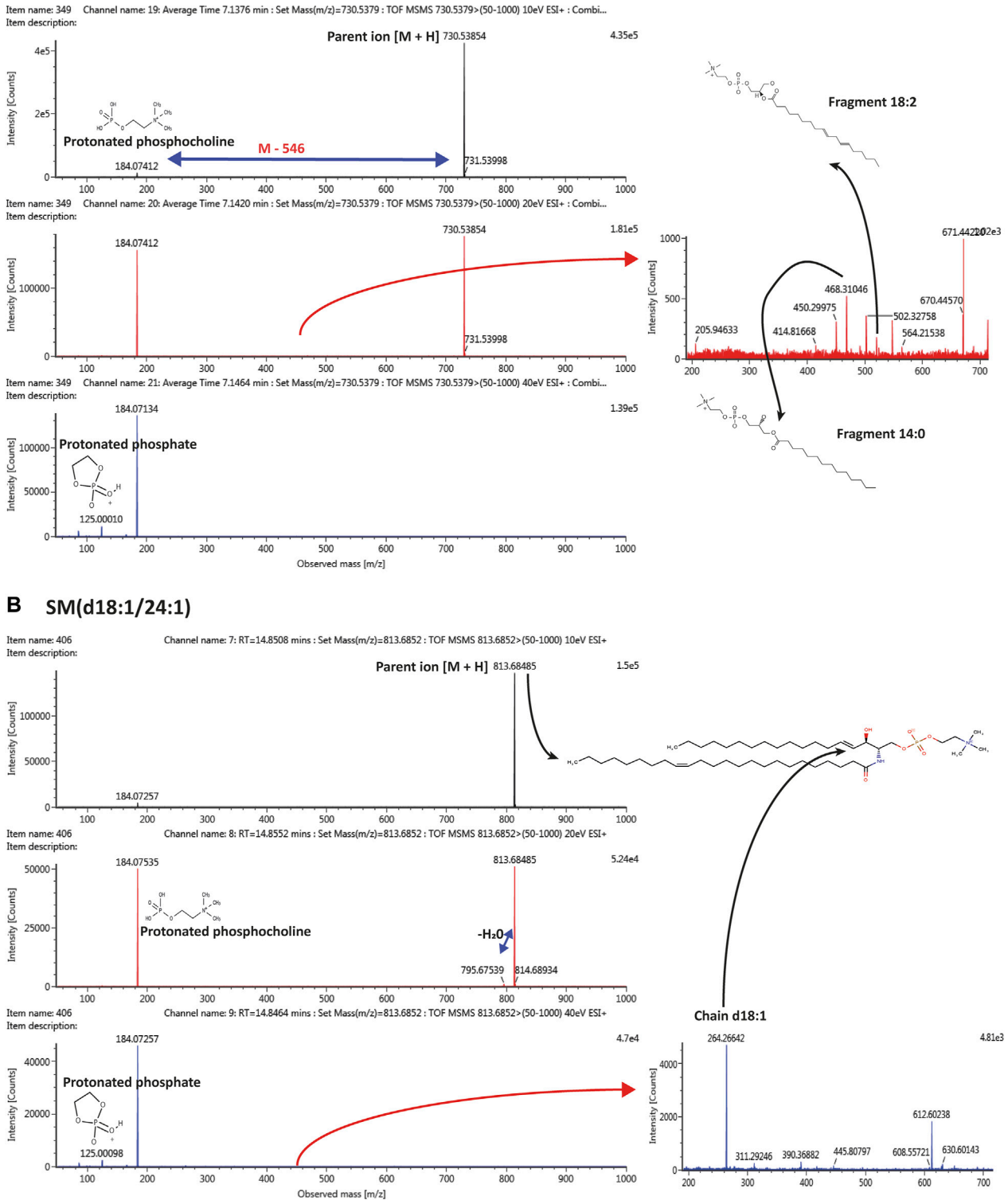


FIGURE 7
Tandem mass spectrometry (MS/MS) representing the pattern of fragmentation with 10, 20, 40 eV in ESI (+) mode, leading to the identification of **(A)** PC(14:0/18:2); **(B)** SM(d18:1/24:1).

TABLE 2 Clinical characteristics from patients with paired samples 2 years after recovery.

Variable	COVID-19 (N = 13)	Post-COVID-19 (N = 13)	*p*-value
Male gender, n (%)	**5 (38.4%)**		**NA**
Age, mean (± s.d)	**54.6 ± 7.9**		**NA**
Disease severity			
Mild, n (%)	**3 (23%)**		**NA**
Severe, n (%)	**8 (61%)**		**NA**
Critical, n (%)	**2 (15%)**		**NA**
Laboratory			
Hemoglobin (g/dL), median (Q1-Q2)	**14.90(13.80–16.40)**	**15.40(14.50–17.40)**	**0.08**
Platelets (×10^3/mL), median (Q1-Q2)	**238 (171.0–318.0)**	**238 (219.0–265.0)**	**0.96**
Leukocytes (×10^3), median (Q1-Q2)	**9.2 (5.6–11.8)**	**7.3 (6.3–8.1)**	**0.36**
Lymphocytes (%), median (Q1-Q2)	**13.4 (7.2–22.8)**	**36.4 (31.9–38.4)**	**<0.001**
Creatinine (mg/dL), median (Q1-Q2)	**0.7 (0.68–0.8)**	**0.80 (0.6–0.8)**	**0.96**

Significant values ($p < 0.05$) are highlighted in bold.

The VIP plot showed that for most differentiated features, higher concentrations were seen among in post-COVID-19 patients (Figure 8E).

4 Discussion

This study aimed at describing clinical and metabolic alterations persisting two years after patients had a SARS-CoV-2 infection of different severity. The metabolic pathways dysregulated during infection and after two years of recovery were identified. A functional analysis approach was used assuming that putative annotation at individual compound level can collectively predict changes at functional levels, as demonstrated by Li et al. (Li et al., 2013).

Lipid classes belonging to sterols, steroids, and fatty esters were dysregulated in both COVID-19 groups studied. Very recently, Guntur et al. (Guntur et al., 2022) found higher levels of poly and highly unsaturated fatty acids in patients with post-COVID-19 syndrome (more than 28 days after infection: recruitment phase done in a time-interval of two years). This finding was consistent with a reduced fatty acid oxidation at mitochondrial level. The accumulation of such molecules has been associated with erythrocyte dysfunction and impairment of oxygen transportation that could persist for months, thereby explaining symptoms such as fatigue and exercise intolerance.

Sterols are a subgroup of steroids. The most familiar type is cholesterol, which is vital for the membrane structure, and it is a precursor of fat-soluble vitamins and steroid hormones. A recent systematic review and meta-analysis demonstrated that lower concentrations of total HDL, and LDL-cholesterol were significantly associated with COVID-19 severity and mortality suggesting that cholesterol concentrations might be useful for risk stratification and monitoring (Zinellu et al., 2021). Ghini et al. (Ghini et al., 2022) recently demonstrated that the lipoproteome of recovered patients slowly reverted to the healthy state.

Corticosteroids such as dexamethasone belong to steroids. They have significant anti-inflammatory and anti-fibrotic effects, which may play a role reducing lung and systemic inflammation, especially in severe pneumonia and in advanced stages of COVID-19 (Leistner et al., 2022).

Among fatty esters, monoacylglycerols, diacylglycerols, and in particular, triacylglycerols, have also been associated with metabolic dysregulation in COVID-19 patients (Masana et al., 2021).

In COVID-19 patients bile acid was also found dysregulated. Bile acids are signaling molecules with immune, metabolic, and intestinal microbiota control actions (Wahlström et al., 2016). Bile acids pathways have been widely reported in COVID-19 due to the proven association between gut dysbiosis and inflammatory processes that lead to severe disease. However, anomalies in bile acids metabolism are also associated with liver injury, and affect the substance transport system (cholesterol transport), which is common in severe COVID-19 (Shen et al., 2020). A disordered metabolism of bile acids among recovered COVID-19 patients (three months after discharge) has been documented, suggesting that the intestinal equilibrium at mucosal level is delayed before is fully repaired (Zhang et al., 2021).

In post-COVID-19 patients, the isoprenoids pathway, also recognized as mevalonate pathway (MVP) or HMG-CoA reductase pathway, was also dysregulated. Isoprenoids are a highly diverse class of biomolecules, ranging from cholesterol, vitamin K, coenzyme Q10, all steroids hormones (Holstein and Hohl, 2004). The MVP limits the activation of inflammasomes and cytokine release, and for this reason, unbalanced signaling could be associated with the pathobiology of COVID-19. A recent *in silico* study revealed dysregulation of genes involved in the MVP in SARS-CoV-2 infection, but not with H3N2 influenza virus, H1N1 influenza virus, or respiratory syncytial virus (Gomez Marti et al., 2021). Finally, the use of statins, namely, HMG-CoA-reductase inhibitors, frequently used as therapeutic agents, reduce cholesterol levels lowering viral titers through immunomodulatory, anti-inflammatory and anti-thrombotic effects (Proto et al., 2021).

These previously described metabolic alterations could account for the plethora of symptoms reported in this study. In recovered patients, values of hemoglobin, lymphocytes, monocytes, neutrophils, and creatine were within normal range two years

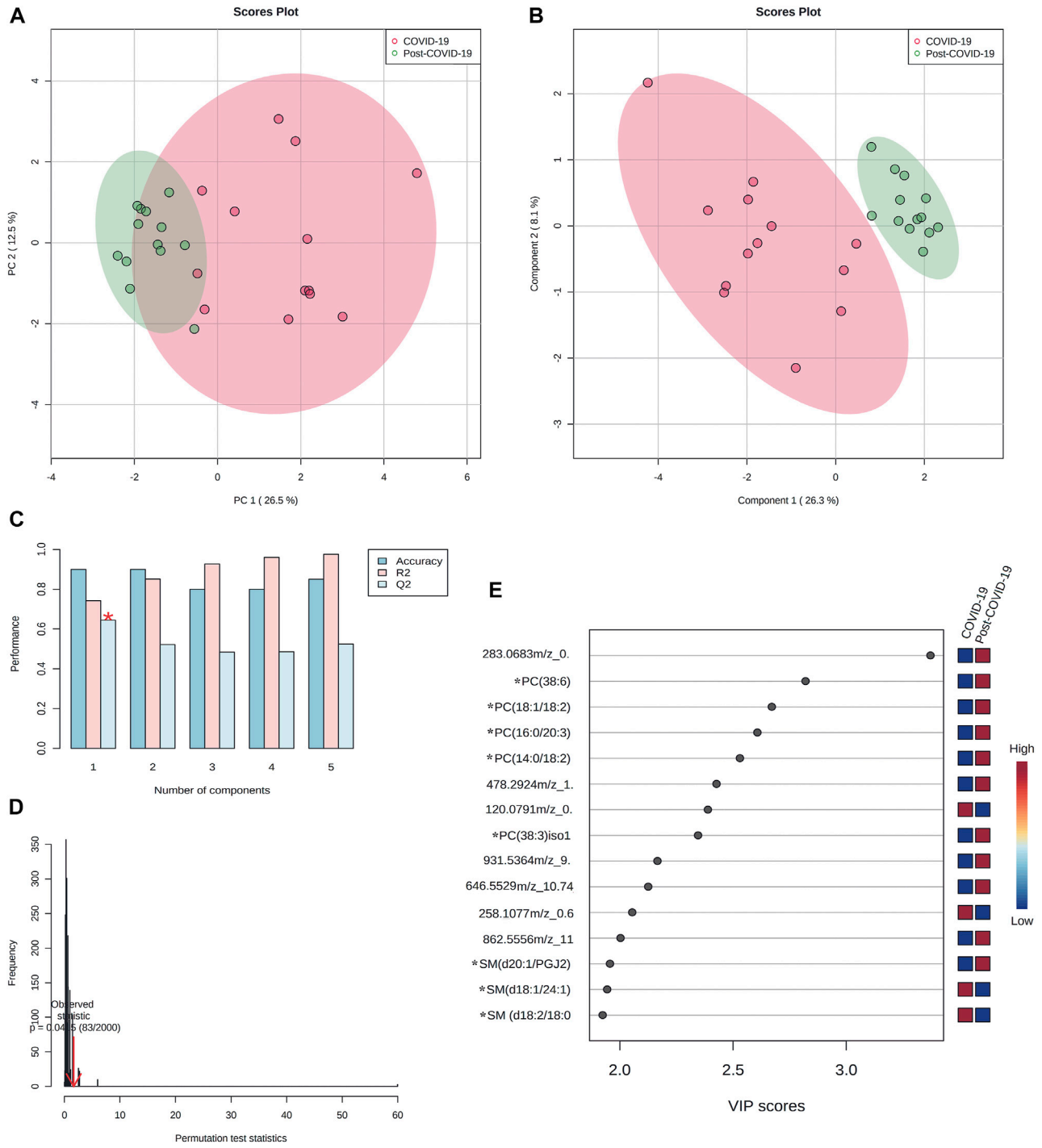


FIGURE 8
Multivariate paired analyses from plasma lipidomics profile of COVID-19, and post-COVID-19 patients. **(A)** Score scatter plot based on PCA models including paired samples. **(B)** Score scatter plot based on PLS-DA models. **(C)** Cross validation (10-foldCV). **(D)** permutation test (2000 permutations). **(E)** Variable importance in Projection (VIP) plot representing the rank of 15 features identified by PLS-DA according to VIP score on *x*-axis. The most discriminating metabolites are shown in descending score order. The color boxes indicate whether metabolite concentration was increased (red) or decreased (blue). * Represents features with confidence level 2. Figures were produced in MetaboAnalyst software v 5.0 (https://www.metaboanalyst.ca/).

after the acute infection. Only three out of 22 patients had one reinfection (with the Omicron variant, January 2022). By the moment of the follow-up laboratory tests, all patients tested negative for SARS-CoV-2. However, the heterogeneity of persistent symptomatology indicates that multiple organ systems were affected during the recovery phase. The etiologies of the

reported conditions are post-acute COVID-19 cardiovascular syndrome, post-acute COVID-19 neuropsychiatric syndrome, and multi-system syndrome.

In the cohort of patients with baseline and follow-up CT scans, interstitial thickening, ground glass opacity, and subpleural bands were the most frequent sequelae observed. Ground glass opacity, interstitial thickening, parenchymal bands, bronchiectasis, lymphadenopathy, and pleural effusion has been reported (Yu et al., 2020). A systematic review and meta-analysis of 15 studies including over 3000 patient's follow-up CT scans at 1–6 months after discharge showed residual CT changes in 55.7% of the cases (So et al., 2021). It could be possible that these anomalies could be reverted two years after infection in the absence of other lung diseases.

Remarkably, fatigue was the predominant alteration reported (59%), as well as musculoskeletal symptoms such as arthralgias and myalgias. Therefore, measuring the plasma lipid profile was relevant, as lipids plays an essential role in energy metabolism.

When the lipid profile was analyzed, multivariate analysis showed that after two years of recovery, post-COVID-19 patients cannot be grouped with negative controls neither clustered with COVID-19 patients, even though most basic laboratory parameters are normalized; yet, the presence of a wide spectrum of symptoms reflects that metabolic mediators are not reestablished at all.

Alterations in lipids have been found in recovered patients from SARS (2003). Wu et al. followed 25 recovered SARS patients 12 years after infection. The authors found increased levels of phosphatidylinositol and lysophosphatidylinositol (Wu et al., 2017).

In SARS-CoV-2, lipid metabolism has been reported altered in all the stages of the disease (Sun et al., 2020) and in the recovery phase. Our group has reported alteration in the levels of acylcarnitines and glycerophospholipids (phosphatidylcholines and lysophosphatidylcholines) upon admission in emergency rooms (early onset of symptoms) (López-Hernández et al., 2021). Chen and cols (Chen et al., 2022). reported that in COVID-19 patients with nucleic acid turning negative (still hospitalized), lipid metabolism was dysregulated. Acosta-Ampudia et al. also found that approximately two months after discharge, the phenotype of recovered patients did not return to a similar phenotype of pre-pandemic controls, and altered levels of unsaturated fatty acids, such as arachidonic and linoleic acid were seen (Acosta-Ampudia et al., 2021). Li et al. (Li et al., 2022) found that metabolic disturbance of lipids was associated with long-term chronic discomfort and immune dysregulation in COVID-19 survivors 6 months after discharge. The authors also reported dysregulated levels of TG, LTB4, PGE2, polyunsaturated fatty acids, including 5-hydroxyeicosatetraenoic acid (5-HETE), 12-hydroxyeicosatetraenoic acid (12-HETE), and 15-oxoeicosatetraenoic acid (15-oxoETE).

In this study, increased levels of several lipids (e.g., glycerophospholipids and sphingolipids) in the plasma of recovered patients were observed. Despite not identifying all the features dysregulated, we were able to identify (confidence level 2) the most important lipids contributing to the differentiation.

Phosphatidylcholines have been found altered in COVID-19 patients with mixed results. This is because the pattern of lipid regulation in COVID-19 patients depends upon the infection severity (asymptomatic, mild, or severe) (Hao et al., 2021). However, in most of the studies published so far, some lipid species (even within the same family) are upregulated and others downregulated, revealing a complex regulation in the context of various concomitant factors playing a role, such as the patient´s immune status and the presence of comorbidities. Here, some species of PCs were found to decrease during the active phase of the disease, and two years later these species increased in post-COVID-19 patients, even in comparison with negative controls. This could be explained as: 1) a compensatory mechanism, 2) the persistence of molecular mechanisms that are still dysregulating lipid homeostasis, and as 3) the cross-talking with the immune system and gut microbiota. In our research conducted in 2020 (Herrera-Van Oostdam et al., 2021), we observed a positive correlation between PCs and SMs with IL-12p70 and IFN-λ1. In line with this, a recent work has reported that patients with long COVID showed elevated expression of type I IFN (IFN-β) and type III IFN (IFN-λ1) that remained high after 8 months of infection (Phetsouphanh et al., 2022). On the other hand, the observed dysregulation could be due to lifestyle changes since it has been observed that mitochondrial dysfunction affect the mechanisms generating energy. Either abnormally high, or abnormally low, phospholipids can influence energy metabolism, and have large implications on general metabolic parameters (van der Veen et al., 2017).

Sphingolipids (SLs) also represent an important group of bioactive molecules involved in crucial processes such as inflammation, cellular differentiation, regeneration, aging, among others, particularly important in musculoskeletal cells (Meacci et al., 2022). The results of this study showed a dysregulation in sphingolipid metabolism, that could be associated with the reported symptoms: fatigue and muscular pain. It has been previously observed that sphingolipids impairment affect skeletal muscle cells (Danieli-Betto et al., 2005; Cowart, 2010). Accumulation of sphingolipids has been associated with inflammatory processes, and mass decrease of skeletal muscle cells of aged mice (Trayssac et al., 2018). Also, inactivity or disuse of musculoskeletal cells, as seen after disabled conditions, correlate well with remodeling of membranes enriched in SL and cholesterol along with changes in ceramide contents (Petrov et al., 2019). Ceramides and Hexocylceramides, which are derivatives of sphingomyelins, have been found increased in female patients with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) and among those with chronic hepatitis C infection and autoimmune disease (Zhang et al., 2016; Filippatou et al., 2021).

Increased sphingolipids levels have also been observed in metabolic syndrome (Chavez and Summers, 2012) and in the acute cell danger response (Naviaux, 2014). Regarding ME/CFS, a general agreement is that metabolic features are consistent with a hypometabolic state, characterized by a decrease in sphingolipids, glycosphingolipids, phospholipids, purines, microbiome aromatic amino acid, and branch chain amino acid. In this study, an increase in sphingolipids and phosphocholines was observed, so an underlying mechanism like ME/CFS unlikely explains the fatigue and muscular alterations seen.

Li et al. (Li et al., 2022) found that total levels of LysoPC, PA, PC, PE, PS and Cer were significantly downregulated in elderly survivors after a maximum of 9 months of a mild disease. This difference may be explained by the type of patients studied; in that study only mild

disease was included, and patients were stratified by age; and the time after the acute disease was shorter.

Summarizing, our results show that post-COVID-19 is a relevant entity that requires further research.

Results also show that after two years of SARS-CoV-2 infection, some metabolic pathways are not normalized. It was worth noting that some lipid species were downregulated in post-COVID patients, while others were upregulated even within the same lipid family. These lipid dysregulations could explain some of the persistent symptoms reported by patients, especially those related to musculoskeletal disorders. Targeted studies reporting absolute concentrations for these markers are needed to establish the precise molecular mechanisms involved, and most importantly, to eventually design potential therapeutic interventions.

Finally, some important limitations ought to be acknowledged. The sample size was small due to the exploratory nature of this study; this is the result of focus given to the recruitment of patients that previously participated in protocols approved in 2020 (first epidemic wave). As a result, after two years, nearly one-third of the patients had died at hospitals within the following months after the infection. Also, from the patients that agreed to participate, there was limited data on the type and dosage of the medications prescribed during the recovery phase to be considered when interpreting the results. There was also lack of data regarding the occurrence of new illnesses and/or the reactivation of latent ones that could have affected the lipidomic profile of the patients.

All participating patients included in this study were sent to specialists to receive medical assistance for treating persistent symptoms.

Data availability statement

The dataset presented in this study can be found in Mendeley repository (www.mendeley.com) with doi: 10.17632/47n7n25z5w.1. (https://data.mendeley.com/datasets/47n7n25z5w)

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the Comité Nacional de Investigación Científica del Instituto Mexicano de Seguridad Social, and the Ethics and Research Committees at Christus Muguerza del Parque Hospital in Chihuahua, Mexico. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JV and MM performed instrument suitability tests and MS/MS experiments, JM-E participated in patient recruitment and wrote part of the clinical database, DL participated in patient recruitment and performed statistical analysis, JL, performed manuscript writing and editing, JB, participated in patient recruitment and wrote part of the clinical database, JJO-V participated in manuscript writing, editing, performed part of the methodology and conceptualization, YL-H participated in manuscript writing, editing, performed part of the methodology, conceptualization, project administration and funding. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

CONACyT grants No. 311880, 316258 and 319503.

Conflict of interest

MM and JV are employed by Waters Corporation. The remaining authors declare that the research was conducted in the absence of any commercial of financial relationships that could be seen as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1100486/full#supplementary-material

References

Acosta-Ampudia, Y., Monsalve, D. M., Rojas, M., Rodríguez, Y., Gallo, J. E., Salazar-Uribe, J. C., et al. (2021). COVID-19 convalescent plasma composition and immunological effects in severe patients. *J. Autoimmun.* 118, 102598. doi:10.1016/j.jaut.2021.102598

Al-Aly, Z., Xie, Y., and Bowe, B. (2021). High-dimensional characterization of post-acute sequelae of COVID-19. *Nature* 594, 259–264. doi:10.1038/s41586-021-03553-9

Chavez, J. A., and Summers, S. A. (2012). A ceramide-centric view of insulin resistance. *Cell Metab.* 15, 585–594. doi:10.1016/j.cmet.2012.04.002

Chen, W., Yao, M., Chen, M., Ou, Z., Yang, Q., He, Y., et al. (2022). Using an untargeted metabolomics approach to analyze serum metabolites in COVID-19 patients with nucleic acid turning negative. *Front. Pharmacol.* 13, 964037. doi:10.3389/fphar.2022.964037

Collaborators, G. B. O. D. L. C., Wulf Hanson, S., Abbafati, C., Aerts, J. G., Al-Aly, Z., Ashbaugh, C., et al. (2022). Estimated global proportions of individuals with persistent fatigue, cognitive, and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. *JAMA* 328, 1604–1615. doi:10.1001/jama.2022.18931

Cowart, L. A. (2010). A novel role for sphingolipid metabolism in oxidant-mediated skeletal muscle fatigue. Focus on "Sphingomyelinase stimulates oxidant signaling to weaken skeletal muscle and promote fatigue". *Am. J. Physiology-Cell Physiology* 299, C549–C551. doi:10.1152/ajpcell.00236.2010

Danieli-Betto, D., Germinario, E., Esposito, A., Megighian, A., Midrio, M., Ravara, B., et al. (2005). Sphingosine 1-phosphate protects mouse extensor digitorum longus skeletal muscle during fatigue. *Am. J. Physiol. Cell Physiol.* 288, C1367–C1373. doi:10.1152/ajpcell.00246.2004

Declaration of Helsinki (1976). Recommendations guiding medical doctors in biomedical research involving human subjects. *Ugeskr. Laeger* 138, 399–400.

Fernández-De-Las-Peñas, C., Palacios-Ceña, D., Gómez-Mayordomo, V., Florencio, L. L., Cuadrado, M. L., Plaza-Manzano, G., et al. (2021). Prevalence of post-COVID-

19 symptoms in hospitalized and non-hospitalized COVID-19 survivors: A systematic review and meta-analysis. *Eur. J. Intern Med.* 92, 55–70. doi:10.1016/j.ejim.2021.06.009

Filippatou, A. G., Moniruzzaman, M., Sotirchos, E. S., Fitzgerald, K. C., Kalaitzidis, G., Lambe, J., et al. (2021). Serum ceramide levels are altered in multiple sclerosis. *Mult. Scler.* 27, 1506–1519. doi:10.1177/1352458520971816

Ghini, V., Meoni, G., Pelagatti, L., Celli, T., Veneziani, F., Petrucci, F., et al. (2022). Profiling metabolites and lipoproteins in COMETA, an Italian cohort of COVID-19 patients. *PLOS Pathog.* 18, e1010443. doi:10.1371/journal.ppat.1010443

Gomez Marti, J. L., Wells, A., and Brufsky, A. M. (2021). Dysregulation of the mevalonate pathway during SARS-CoV-2 infection: An *in silico* study. *J. Med. Virol.* 93, 2396–2405. doi:10.1002/jmv.26743

Guntur, V. P., Nemkov, T., De Boer, E., Mohning, M. P., Baraghoshi, D., Cendali, F. I., et al. (2022). Signatures of mitochondrial dysfunction and impaired fatty acid metabolism in plasma of patients with post-acute sequelae of COVID-19 (PASC). *Metabolites* 12, 1026. doi:10.3390/metabo12111026

Hao, Y., Zhang, Z., Feng, G., Chen, M., Wan, Q., Lin, J., et al. (2021). Distinct lipid metabolic dysregulation in asymptomatic COVID-19. *iScience* 24, 102974. doi:10.1016/j.isci.2021.102974

Herrera-Van Oostdam, A. S., Castañeda-Delgado, J. E., Oropeza-Valdez, J. J., Borrego, J. C., Monárrez-Espino, J., Zheng, J., et al. (2021). Immunometabolic signatures predict risk of progression to sepsis in COVID-19. *PLOS ONE* 16, e0256784. doi:10.1371/journal.pone.0256784

Holstein, S. A., and Hohl, R. J. (2004). Isoprenoids: Remarkable diversity of form and function. *Lipids* 39, 293–309. doi:10.1007/s11745-004-1233-3

Kirwan, J. A., Gika, H., Beger, R. D., Bearden, D., Dunn, W. B., Goodacre, R., et al. (2022). Quality assurance and quality control reporting in untargeted metabolic phenotyping: mQACC recommendations for analytical quality management. *Metabolomics* 18, 70. doi:10.1007/s11306-022-01926-3

Leistner, R., Schroeter, L., Adam, T., Poddubnyy, D., Stegemann, M., Siegmund, B., et al. (2022). Corticosteroids as risk factor for COVID-19-associated pulmonary aspergillosis in intensive care patients. *Crit. Care* 26, 30. doi:10.1186/s13054-022-03902-8

Li, F., Fu, L., Liu, X., Liu, X. A., Liang, Y., Lv, Y., et al. (2022). Serum metabolomic abnormalities in survivors of non-severe COVID-19. *Heliyon* 8, e10473. doi:10.1016/j.heliyon.2022.e10473

Li, S., Park, Y., Duraisingham, S., Strobel, F. H., Khan, N., Soltow, Q. A., et al. (2013). Predicting network activity from high throughput metabolomics. *PLOS Comput. Biol.* 9, e1003123. doi:10.1371/journal.pcbi.1003123

López-Hernández, Y., Monárrez-Espino, J., Oostdam, A. H., Delgado, J. E. C., Zhang, L., Zheng, J., et al. (2021). Targeted metabolomics identifies high performing diagnostic and prognostic biomarkers for COVID-19. *Sci. Rep.* 11, 14732. doi:10.1038/s41598-021-94171-y

Lopez-Leon, S., Wegman-Ostrosky, T., Perelman, C., Sepulveda, R., Rebolledo, P. A., Cuapio, A., et al. (2021). More than 50 long-term effects of COVID-19: A systematic review and meta-analysis. *Sci. Rep.* 11, 16144. doi:10.1038/s41598-021-95565-8

Malik, P., Patel, K., Pinto, C., Jaiswal, R., Tirupathi, R., Pillai, S., et al. (2022). Post-acute COVID-19 syndrome (PCS) and health-related quality of life (HRQoL)-A systematic review and meta-analysis. *J. Med. Virol.* 94, 253–262. doi:10.1002/jmv.27309

Masana, L., Correig, E., Ibarretxe, D., Anoro, E., Arroyo, J. A., Jericó, C., et al. (2021). Low HDL and high triglycerides predict COVID-19 severity. *Sci. Rep.* 11, 7217. doi:10.1038/s41598-021-86747-5

Meacci, E., Pierucci, F., and Garcia-Gil, M. (2022). Skeletal muscle and COVID-19: The potential involvement of bioactive sphingolipids. *Biomedicines* 10, 1068. doi:10.3390/biomedicines10051068

Medina, J., Van Der Velpen, V., Teav, T., Guitton, Y., Gallart-Ayala, H., and Ivanisevic, J. (2020). Single-step extraction coupled with targeted HILIC-MS/MS approach for comprehensive analysis of human plasma lipidome and polar metabolome. *Metabolites* 10, 495. doi:10.3390/metabo10120495

Naviaux, R. K. (2014). Metabolic features of the cell danger response. *Mitochondrion* 16, 7–17. doi:10.1016/j.mito.2013.08.006

Ortona, E., and Malorni, W. (2022). Long COVID: To investigate immunological mechanisms and sex/gender related aspects as fundamental steps for tailored therapy. *Eur. Respir. J.* 59, 2102245. doi:10.1183/13993003.02245-2021

Petrov, A. M., Shalagina, M. N., Protopopov, V. A., Sergeev, V. G., Ovechkin, S. V., Ovchinina, N. G., et al. (2019). Changes in membrane ceramide pools in rat soleus muscle in response to short-term disuse. *Int. J. Mol. Sci.* 20, 4860. doi:10.3390/ijms20194860

Phetsouphanh, C., Darley, D. R., Wilson, D. B., Howe, A., Munier, C. M. L., Patel, S. K., et al. (2022). Immunological dysfunction persists for 8 months following initial mild-to-moderate SARS-CoV-2 infection. *Nat. Immunol.* 23, 210–216. doi:10.1038/s41590-021-01113-x

Proto, M. C., Fiore, D., Piscopo, C., Pagano, C., Galgani, M., Bruzzaniti, S., et al. (2021). Lipid homeostasis and mevalonate pathway in COVID-19: Basic concepts and potential therapeutic targets. *Prog. Lipid Res.* 82, 101099. doi:10.1016/j.plipres.2021.101099

Reiken, S., Sittenfeld, L., Dridi, H., Liu, Y., Liu, X., and Marks, A. R. (2022). Alzheimer's-like signaling in brains of COVID-19 patients. *Alzheimers Dement.* 18, 955–965. doi:10.1002/alz.12558

Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., and Mclean, J. A. (2016). Untargeted metabolomics strategies—challenges and emerging directions. *J. Am. Soc. Mass Spectrom.* 27, 1897–1905. doi:10.1007/s13361-016-1469-y

Shen, B., Yi, X., Sun, Y., Bi, X., Du, J., Zhang, C., et al. (2020). Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell* 182, 59–72.e15. doi:10.1016/j.cell.2020.05.032

So, M., Kabata, H., Fukunaga, K., Takagi, H., and Kuno, T. (2021). Radiological and functional lung sequelae of COVID-19: A systematic review and meta-analysis. *BMC Pulm. Med.* 21, 97. doi:10.1186/s12890-021-01463-0

Soriano, J. B., Murthy, S., Marshall, J. C., Relan, P., and Diaz, J. V. (2022). A clinical case definition of post-COVID-19 condition by a Delphi consensus. *Lancet Infect. Dis.* 22, e102–e107. doi:10.1016/S1473-3099(21)00703-9

Sun, J. T., Chen, Z., Nie, P., Ge, H., Shen, L., Yang, F., et al. (2020). Lipid profile features and their associations with disease severity and mortality in patients with COVID-19. *Front. Cardiovasc Med.* 7, 584987. doi:10.3389/fcvm.2020.584987

Trayssac, M., Hannun, Y. A., and Obeid, L. M. (2018). Role of sphingolipids in senescence: Implication in aging and age-related diseases. *J. Clin. Invest.* 128, 2702–2712. doi:10.1172/JCI97949

Van Der Veen, J. N., Kennelly, J. P., Wan, S., Vance, J. E., Vance, D. E., and Jacobs, R. L. (2017). The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochimica Biophysica Acta (BBA) - Biomembr.* 1859, 1558–1572. doi:10.1016/j.bbamem.2017.04.006

Wahlström, A., Sayin, S. I., Marschall, H. U., and Bäckhed, F. (2016). Intestinal crosstalk between bile acids and microbiota and its impact on host metabolism. *Cell Metab.* 24, 41–50. doi:10.1016/j.cmet.2016.05.005

Westerhuis, J. A., Van Velzen, E. J. J., Hoefsloot, H. C. J., and Smilde, A. K. (2008). Discriminant Q2 (DQ2) for improved discrimination in PLSDA models. *Metabolomics* 4, 293–296. doi:10.1007/s11306-008-0126-2

Who (2023). *Clinical management of COVID-19: Living guideline*. Geneva).

Wu, Q., Zhou, L., Sun, X., Yan, Z., Hu, C., Wu, J., et al. (2017). Altered lipid metabolism in recovered SARS patients twelve years after infection. *Sci. Rep.* 7, 9110. doi:10.1038/s41598-017-09536-z

Xie, Y., Xu, E., Bowe, B., and Al-Aly, Z. (2022). Long-term cardiovascular outcomes of COVID-19. *Nat. Med.* 28, 583–590. doi:10.1038/s41591-022-01689-3

Yong, S. J. (2021). Long COVID or post-COVID-19 syndrome: Putative pathophysiology, risk factors, and treatments. *Infect. Dis. (Lond)* 53, 737–754. doi:10.1080/23744235.2021.1924397

Yu, M., Liu, Y., Xu, D., Zhang, R., Lan, L., and Xu, H. (2020). Prediction of the development of pulmonary fibrosis using serial thin-section CT and clinical features in patients discharged after treatment for COVID-19 pneumonia. *Korean J. Radiol.* 21, 746–755. doi:10.3348/kjr.2020.0215

Zhang, J. Y., Qu, F., Li, J. F., Liu, M., Ren, F., Zhang, J. Y., et al. (2016). Up-regulation of plasma hexosylceramide (d18: 1/18: 1) contributes to genotype 2 virus replication in chronic hepatitis C: A 20-year cohort study. *Med. Baltim.* 95, e3773. doi:10.1097/MD.0000000000003773

Zhang, S., Luo, P., Xu, J., Yang, L., Ma, P., Tan, X., et al. (2021). Plasma metabolomic profiles in recovered COVID-19 patients without previous underlying diseases 3 Months after discharge. *J. Inflamm. Res.* 14, 4485–4501. doi:10.2147/JIR.S325853

Zinellu, A., Paliogiannis, P., Fois, A. G., Solidoro, P., Carru, C., and Mangoni, A. A. (2021). Cholesterol and triglyceride concentrations, COVID-19 severity, and mortality: A systematic review and meta-analysis with meta-regression. *Front. Public Health* 9, 705916. doi:10.3389/fpubh.2021.705916

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 24 May 2023
DOI 10.3389/fmolb.2023.1192088

OPEN ACCESS

EDITED BY
Guillermo Moyna,
Universidad de la República, Uruguay

REVIEWED BY
Andrés Perez Parada,
Universidad de la República, Uruguay
Ying Liu,
Canadian Food Inspection Agency (CFIA), Canada
Wilton Ricardo Sala-Carvalho,
University of São Paulo, Brazil

*CORRESPONDENCE
Ian Castro-Gamboa,
ian.castro@unesp.br
Geison M. Costa,
modesticosta.g@javeriana.edu.co

†These authors have contributed equally to this work and share first authorship

RECEIVED 22 March 2023
ACCEPTED 10 May 2023
PUBLISHED 24 May 2023

CITATION
Chitiva LC, Lozano-Puentes HS, Londoño X, Leão TF, Cala MP, Ruiz-Sanchez E, Díaz-Ariza LA, Prieto-Rodríguez JA, Castro-Gamboa I and Costa GM (2023), Untargeted metabolomics approach and molecular networking analysis reveal changes in chemical composition under the influence of altitudinal variation in bamboo species.
Front. Mol. Biosci. 10:1192088.
doi: 10.3389/fmolb.2023.1192088

Untargeted metabolomics approach and molecular networking analysis reveal changes in chemical composition under the influence of altitudinal variation in bamboo species

Luis Carlos Chitiva[1,2†], Hair Santiago Lozano-Puentes[1,3†], Ximena Londoño[4], Tiago F. Leão[2], Mónica P. Cala[5], Eduardo Ruiz-Sanchez[6], Lucía Ana Díaz-Ariza[3], Juliet A. Prieto-Rodríguez[1], Ian Castro-Gamboa[2*] and Geison M. Costa[1*]

[1]Department of Chemistry, Faculty of Sciences, Pontificia Universidad Javeriana, Bogotá, Colombia, [2]Institute of Chemistry, São Paulo State University (UNESP), Araraquara, Brazil, [3]Department of Biology, Faculty of Sciences, Pontificia Universidad Javeriana, Bogotá, Colombia, [4]Faculty of Agricultural Sciences, Universidad Nacional de Colombia, Palmira, Colombia, [5]Metabolomics Core Facility-MetCore, Universidad de los Andes, Bogotá, Colombia, [6]Department of Botany and Zoology, Universidad de Guadalajara, Jalisco, México

Bamboo species have traditionally been used as building material and potential source of bioactive substances, as they produce a wide variety of phenolic compounds, including flavonoids and cinnamic acid derivatives that are considered biologically active. However, the effects of growth conditions such as location, altitude, climate, and soil on the metabolome of these species still need to be fully understood. This study aimed to evaluate variations in chemical composition induced by altitudinal gradient (0–3000 m) by utilizing an untargeted metabolomics approach and mapping chemical space using molecular networking analysis. We analyzed 111 samples from 12 bamboo species collected from different altitudinal ranges using liquid chromatography coupled with quadrupole time-of-flight mass spectrometry (LC-QTOF-MS). We used multivariate and univariate statistical analyses to identify the metabolites that showed significant differences in the altitude environments. Additionally, we used the Global Natural Products Social Molecular Networking (GNPS) web platform to perform chemical mapping by comparing the metabolome among the studied species and the reference spectra from its database. The results showed 89 differential metabolites between the altitudinal ranges investigated, wherein high altitude environments significantly increased the profile of flavonoids. While, low altitude environments significantly boosted the profile of cinnamic acid derivatives, particularly caffeoylquinic acids (CQAs). MolNetEnhancer networks confirmed the same differential molecular families already found, revealing metabolic diversity. Overall, this study provides the first report of variations induced by altitude in the chemical profile of bamboo species. The findings may possess fascinating active biological properties, thus offering an alternative use for bamboo.

KEYWORDS

bamboo, *Guadua*, altitudinal variation, flavonoids, cinnamic acid derivatives, untargeted metabolomics, GNPS, natural products

1 Introduction

Bamboo has gained immense value in recent times due to its versatile applications in construction, food, cosmetics, and medicine (Liese et al., 2015; Ming et al., 2017; Chongtham and Bisht, 2020). Bamboo is a rich source of active compounds such as flavonoids, phenolic acid derivatives, alkaloids, terpenes, and essential oils (Coffie et al., 2014; Gomez et al., 2021; Gagliano et al., 2022; Indira et al., 2022; Okido et al., 2022; Cheng et al., 2023) that are characterized by their antioxidant (Panche et al., 2016; Speisky et al., 2022), antimicrobial (Xie et al., 2015), antiviral (Badshah et al., 2021), and anti-inflammatory (Maleki et al., 2019) properties, among others. While most chemical and biological studies have focused on the Asian continent (Clark et al., 2015; Gagliano et al., 2022; Tamang et al., 2022), there is a pressing need to deepen research on Neotropical bamboos to identify alternative uses and create additional value for these species.

The metabolic profile of plants can be affected by several environmental factors, such as temperature, light, ultraviolet radiation levels, precipitation, humidity, nutrients, and altitude (Khalil et al., 2020; Kumari et al., 2022). Additionally, genetic factors, including the presence of genes that control metabolite biosynthesis and the participation of enzymes in different biosynthetic pathways, have been shown to contribute to this variation (Dhami and Mishra, 2015; Sampaio et al., 2016; Pant et al., 2021). However, few studies have investigated changes in the chemical composition of bamboo species due to environmental or genetic effects.

Recent studies have revealed that certain changes can significantly impact the biological potential of bamboo. Specifically, research has shown that seasonal and altitudinal variation in *Sasa argenteastriatus* (*Pleioblastus argenteostriatus* (Regel) Nakai) and *S. quelpaertensis* Nakai leaves is positively correlated with an increase in the content of phenolic and flavonoid compounds, with chlorogenic acid, isoorientin, and vitexin being the most notable compounds exhibiting significant changes (Ni et al., 2012; Ko et al., 2018). Another study on *Indocalamus latifolius* (Keng) McClure evaluated the impact of altitude on the chemical composition of flavonoids, phenols, and triterpenes, demonstrating that an increase in altitude led to the accumulation of metabolites and a subsequent increase in antioxidant potential (Ni et al., 2013). Based on these findings, we anticipate observing a similar correlation between the increase in phenolic compound content and the species in our study, providing valuable insights for improving crop production and obtaining biologically active metabolites.

Metabolomics is a valuable tool for evaluating metabolic changes in various biological matrices caused by environmental or genetic factors. To analyze large amounts of metabolites in a biological sample, different analytical platforms are currently used. When combined with multivariate analysis, these platforms can identify differentially expressed metabolites, helping to elucidate possible metabolic pathways affected (Verpoorte et al., 2010; Shen et al., 2023). To integrate other platforms and complement the global analysis of the metabolome in bamboo species, we used Global Natural Products Social Molecular Networking (GNPS), a novel platform that facilitates the creation of molecular networks to analyze mass spectrometry (MS/MS) data sets, providing a comprehensive visualization of the chemical space (Wang et al., 2016; Aron et al., 2020; Ramabulana et al., 2021). To our knowledge, this is the first report that applies the untargeted metabolomics approach and molecular networking analysis to the study of the chemical composition of bamboo species under the influence of altitude. This study aimed to evaluate the variations in chemical composition under the influence of an altitudinal gradient (0–3000 m) by utilizing an untargeted metabolomics approach and the mapping of the chemical diversity using molecular networking analysis of the global metabolome in bamboo species.

2 Materials and methods

2.1 Study design

A completely random sampling method was implemented for the metabolomic study by selecting different collection sites in Colombian locations [Cundinamarca (CU), Nariño (NA), Putumayo (PU), and Quindío (QU)] situated at different altitudes, ranging from (0–3000 m). The geographical distributions of the collected species are shown in Figure 1. All samples were collected from natural bamboo in the period 2020 to 2022 (more details are provided in Supplementary Table S1).

2.2 Plant material collection and identification

We collected a total of 111 leaf samples from the upper branches of 12 bamboo species, which were distributed across 40 different altitudes and collected under the same conditions. Harvested bamboo leaves were immediately air-dried at room temperature and then ground into powder for extraction. Based on altitude, we divided the 111 leaf samples into two groups: the low altitude group (~0–1500 m; n = 40) and high altitude group (~1500–3000 m; n = 71), as indicated in Table 1. At least three individuals (n = 3) were collected at each collection site, and voucher specimens were deposited in the (HPUJ) Herbarium of the Pontificia Universidad Javeriana.

2.3 Metabolite extraction and sample preparation

To extract the plant material (leaves), 300 mg of dried and ground samples were mixed with 10 mL of a solvent containing chloroform, methanol, and water (in a ratio of 5:2.5:2.5 *v*/*v*/*v*). The resulting mixture was then vortexed for 1 min, sonicated for 20 min, and centrifuged at 5000 rpm and 20°C for 10 min. The liquid

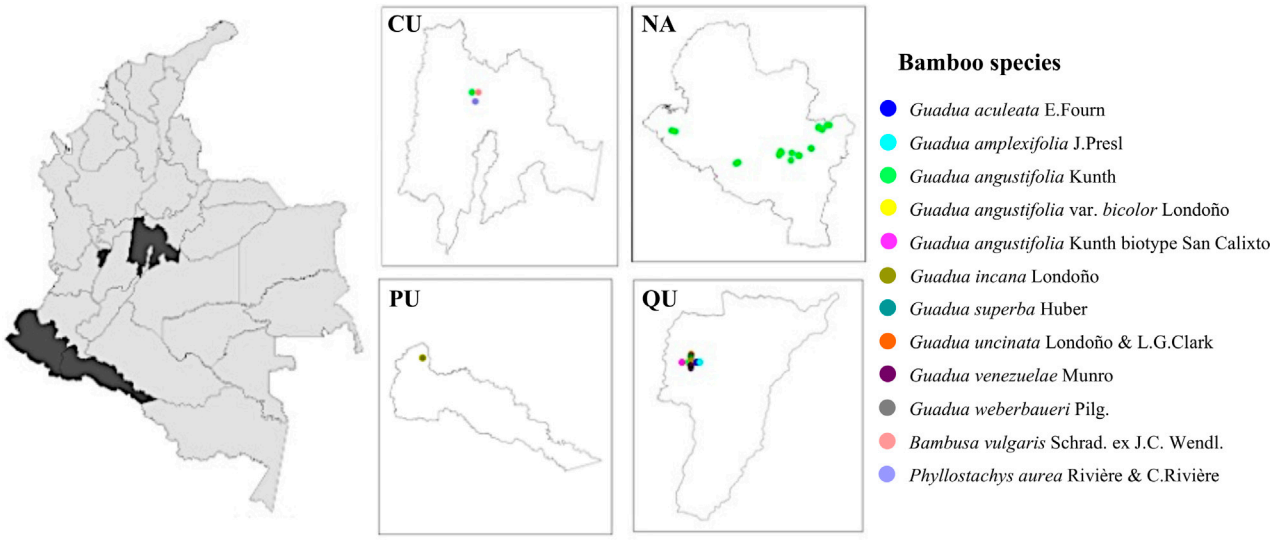


FIGURE 1
Geographical distribution map showing the bamboo species studied. Each mark corresponds to the place of collection of each specimen.

supernatant was filtered using PTFE syringe filters with a pore size of 0.22 µm (Thermo Scientific, Rockwood, TN), and stored at −80°C until used for metabolomics analysis.

2.4 Untargeted metabolomics using LC-QTOF-MS

Untargeted metabolomics was conducted using an Agilent Infinity 1260 HPLC system coupled to an Agilent 6545 quadrupole time-of-flight (QTOF) mass spectrometer equipped with electrospray ion source (Waldbronn, Germany). 5 µL of the extracts were injected on a C18 column (Kinetex 100 × 2.1 mm, 2.6 µm) at 30°C. The mobile phase employed in LC analyses was composed of 0.1% formic acid in Milli-Q water (*v/v*) (Solvent A) and acetonitrile (Solvent B) with a constant flow of 400 µL/min. The gradient was set as follows: 3% for 1 min; 3%–97% B in 15 min; 97% B for 2 min; the column was re-equilibrated for 6 min at the initial conditions. Full-scan MS1 and MS/MS spectra were acquired. Data mass spectra were acquired in negative ionization mode (ESI^-), in a mass range of *m/z* 80–1700 Da in data-dependent acquisition (DDA) mode. The QTOF instrument was operated in the 4 GHz (high resolution) mode. The parameters used for data acquisition were set as follows: nitrogen used as nebulizer gas with pressure at 52 psi, a capillary voltage of 3000 V, ion source temperature of 250°C, dry gas flow at 12 L/min, and acquisition rate of one spectrum per second. MS/MS fragmentation was performed using a collision-induced dissociation energy of 20 eV. Throughout the analysis, two reference masses were used for mass correction: *m/z* 112.9856 [$C_2O_2F_3(NH_4)$] and *m/z* 1033.9881($C_{18}H_{18}O_6N_3P_3F_{24}$).

2.5 Quality control samples

To evaluate system performance and reproducibility in sample analysis, multiple QC samples were created by pooling and mixing equal volumes of each extracted sample. To assess the instrument's robustness, pooled QC samples were injected before the sample analysis until system equilibration was achieved and after every ten randomized sample injections.

2.6 Metabolomics data processing

The LC-QTOF-MS raw data sets were processed using Agilent MassHunter Workstation Profinder software (B.10.0, Agilent Technologies) to extract molecular features for deconvolution, alignment, and integration. The data were manually inspected to eliminate noise and unrelated ions, and a presence filter was applied. For statistical analysis, features that were present in 100% of the samples for each altitudinal group and had a coefficient of variation (CV) of less than 20% in the QC were selected.

2.7 Statistical analysis

Univariate and multivariate analyses were conducted using MatLab (R2019b, MathWorks, Inc., Natick) and SIMCA 14.0 software (Umetrics, Umeå, Sweden), respectively. The multivariate analysis generated PCA and OPLS-DA models, which were validated using cross-validation less than 0.05 and evaluated based on R^2X (change in X explained by the model), R^2Y (the total of Y explained), and Q^2 (sum parameter in cross-validation). The significantly differential metabolites were identified by calculating the variable importance in the projection (VIP) greater than 1, with jackknife confident interval (JK) not including zero combined with FC > 2.0 or FC < 0.5. The annotated metabolites and their peak area were organized in a table (.csv), which was uploaded to MetaboAnalyst 5.0 software for statistical, functional, and integrative analysis of metabolomics data (https://www.metaboanalyst.ca/). The software was used for

10.3389/fmolb.2023.1192088

TABLE 1 Classification of bamboo species according to altitudinal range.

Plant scientific name	Sampling location	Altitude (m)	Group
G. aculeata E.Fourn.	Quindío, Montenegro	1256	Low
G. amplexifolia J.Presl	Quindío, Montenegro	1256	Low
G. angustifolia Kunth	Nariño, Tumaco	18	Low
	Nariño, Tumaco	21	Low
	Nariño, Ricaurte	1053	Low
	Nariño, Ricaurte	1089	Low
	Quindío, Montenegro	1256	Low
	Cundinamarca, Pacho	1343	Low
	Nariño, Samaniego	1478	Low
	Nariño, Samaniego	1565	High
	Nariño, Samaniego	1588	High
	Nariño, San Lorenzo	1595	High
	Nariño, La Unión	1598	High
	Nariño, Consacá	1606	High
	Nariño, Samaniego	1606	High
	Nariño, La Unión	1610	High
	Nariño, La Unión	1631	High
	Nariño, San Lorenzo	1713	High
	Nariño, Sandoná	1720	High
	Nariño, Sandoná	1744	High
	Nariño, San Lorenzo	1779	High
	Nariño, San Lorenzo	1808	High
	Nariño, San Lorenzo	1826	High
	Nariño, Chachagüí	1857	High
	Nariño, San Lorenzo	1876	High
	Nariño, San Lorenzo	1930	High
	Nariño, San Lorenzo	1970	High
	Nariño, La Florida	2089	High
	Nariño, La Florida	2122	High
	Nariño, La Florida	2137	High
G. angustifolia var. *bicolor* Londoño	Quindío, Montenegro	1256	Low
G. angustifolia Kunth biotype San Calixto	Quindío, Montenegro	1256	Low
G. incana Londoño	Putumayo, Mocoa	604	Low
	Quindío, Montenegro	1256	Low
G. superba Huber	Quindío, Montenegro	1256	Low
G. uncinata Londoño & L.G.Clark	Quindío, Montenegro	1256	Low
G. venezuelae Munro	Quindío, Montenegro	1256	Low
G. weberbaueri Pilg.	Quindío, Montenegro	1256	Low

(Continued on following page)

TABLE 1 (*Continued*) **Classification of bamboo species according to altitudinal range.**

Plant scientific name	Sampling location	Altitude (m)	Group
B. vulgaris Schrad. ex J.C. Wendl.	Cundinamarca, Pacho	1343	Low
P. aurea Rivière & C.Rivière	Quindío, Montenegro	1256	Low

visualization using heatmap clustering (Pang et al., 2022), normalized by Pareto scaling. Additionally, univariate analysis was performed to determine the *p*-value features.

2.8 Metabolite identification

Differential metabolite annotation was conducted by considering the precision of the mass (maximum error of mass 10 ppm), isotopic pattern distribution, and adduct formation, using different public online databases such as METLIN (https://metlin.scripps.edu/), KEGG (https://genome.jp/kegg), HMDB (https://hmdb.ca/), PubChem (https://pubchem.ncbi.nlm.nih.gov/) and ChEBI (https://www.ebi.ac.uk/chebi/) through the CEU Mass Mediator (https://ceumass.eps.uspceu.es/) tool. The identity of the metabolites was confirmed through MS/MS analysis, which included the use of MS-DIAL 4.80 (https://prime.psc.riken.jp/compms/msdial/main.html), MS-FINDER 3.52 (https://prime.psc.riken.jp/compms/msfinder/main.html), CFM-ID 4.0 (https://cfmid.wishartlab.com/) for *in silico* mass spectral fragmentation, GNPS web platform (https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp) and manual interpretation with the Agilent Mass Hunter Qualitative Analysis software (version 10.0). The metabolites were identified according to the metabolomics standards initiative (Schymanski et al., 2014).

2.9 Global natural products social molecular networking (GNPS) web platform workflow description

A molecular network was created using the online workflow (https://ccms-ucsd.github.io/GNPSDocumentation/) on the GNPS website (https://gnps.ucsd.edu/). The precursor ion mass tolerance was set to 0.02 Da and an MS/MS fragment ion tolerance of 0.02 Da. A network was then created where edges were filtered to have a cosine score above 0.6 and more than four matched peaks. Further, edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 0, and the lowest-scoring edges were removed from molecular families until the molecular family size was below this threshold. The spectra in the network were then searched against GNPS spectral libraries. All matches kept between network spectra and library spectra were required to have a score above 0.6 and at least four matched peaks (Wang et al., 2016). To enhance chemical structural information within the molecular network, information from *in silico* structure annotations from GNPS Library Search, Network Annotation Propagation, Dereplicator were incorporated into the network using the GNPS MolNetEnhancer workflow (https://ccms-ucsd.github.io/GNPSDocumentation/molnetenhancer/). Chemical class annotations were performed using the ClassyFire chemical ontology (Djoumbou Feunang et al., 2016; Mohimani et al., 2017; Da Silva et al., 2018; Ernst et al., 2019). The attribute table of the generated nodes was visualized in the Cytoscape software to analyze the molecular network. The data used for the analysis of molecular networks were deposited in the MassIVE Public GNPS database (http://massive.ucsd.edu) with the accession number MSV000090298. The workflow used in this study is summarized in Figure 2.

3 Results

3.1 Chemical variation of bamboo species presented between low and high altitudes

The quality and stability of the instrument were assessed by conducting a principal component analysis (PCA) on the quality control samples. Supplementary Figure S1 shows a clustering of the QC samples, which indicates the analytical platform's reliability and the data's validity. To account for altitudinal variation, we performed PCA and orthogonal partial least squares discriminant analysis (OPLS-DA) models on samples classified into two altitude ranges. Based on this classification, we developed an unsupervised PCA model to observe trends between the two altitude groups. As depicted in Figure 3A, the PCA model showed an initial exploration of the data set, evidencing certain trends between the variables of the altitude groups. Additionally, we constructed a supervised OPLS-DA model to differentiate between low and high altitude groups, with appropriate quality parameters, demonstrating complete separation of the groups. (Figure 3B). This allowed us to identify the significant variables that contributed to the separation between the groups, yielding a total of 89 significant variables with values (false discovery rate, FDR) < 0.05, VIP >1, and FC > 2.0 (or <0.5), as identified according to the metabolomics standards initiative (Schymanski et al., 2014). To validate the OPLS-DA model, we conducted permutation analysis, plotting R^2 and Q^2 of 200 permutation tests, as shown in Figure 3C.

Table 2 presents the differential metabolites (89) that belong to various families, such as flavonoids (48%), fatty acids (19%), cinnamic acid derivatives (11%), unknowns (6%), peptides (3%), glycosylated lignans (2%), alkaloids (2%), carboxylic acids (1%), phenols (1%), carbohydrates (1%), steroids (1%), glycosylated stilbenes (1%) and prenolipids (1%). The largest group of differentially expressed

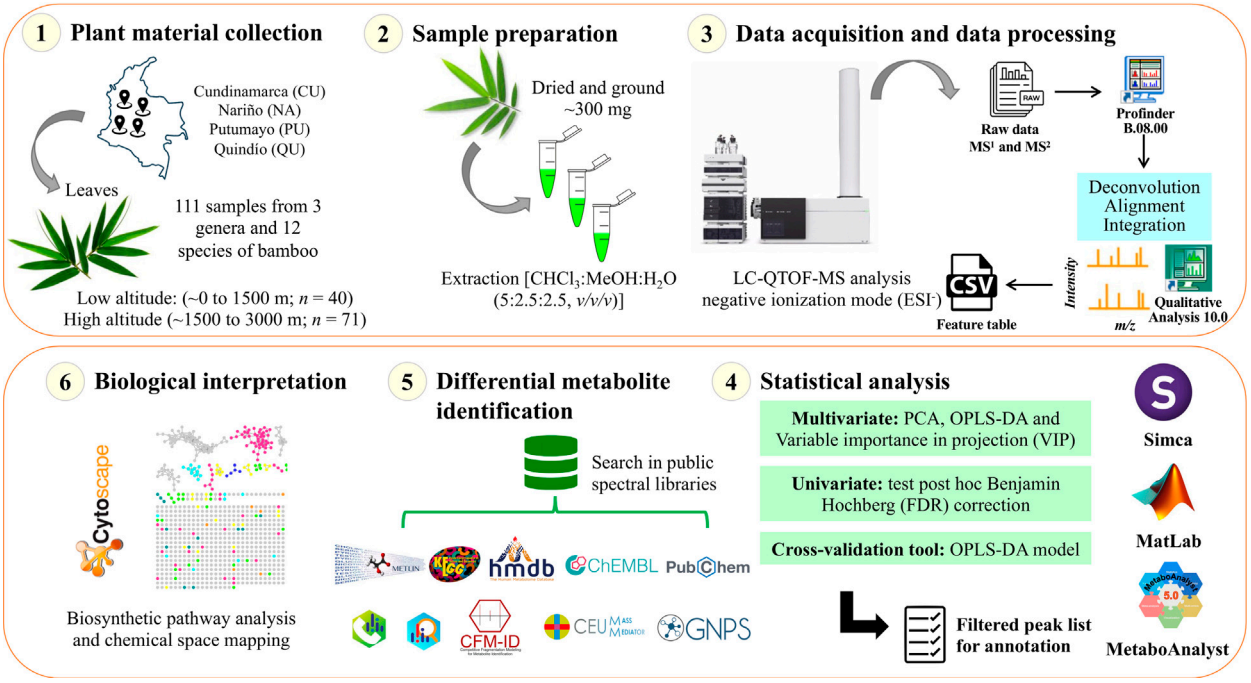


FIGURE 2
Schematic of the experiment analytical workflow used in this study.

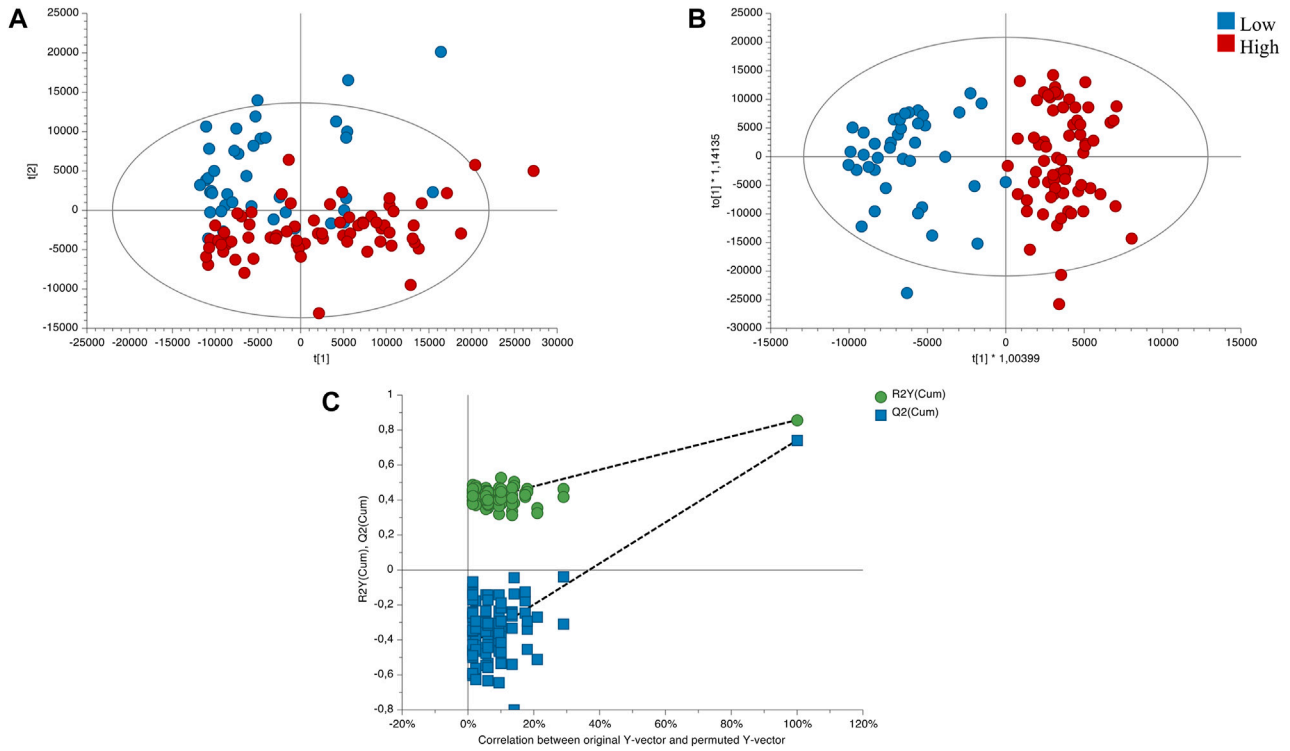


FIGURE 3
Score plots of the influence of direction compared low/high altitudes. **(A)** PCA score plot of total samples ($R^2X_{(cum)}$: 0.813; $Q^2_{(cum)}$: 0.54). **(B)** OPLS-DA score plot of the low vs. high altitudes sample ($R^2X_{(cum)}$: 0.494; $R^2Y_{(cum)}$:0.858; $Q^2_{(cum)}$:0.742; CV Anova: 9.61549e-27). **(C)** Cross-validation plot of the OPLS-DA model with 200 permutation test. Low altitude (~0–1500 m; *n* = 40) and high altitude (~1500–3000 m; *n* = 71).

TABLE 2 Significantly differential metabolites found the altitudinal variation.

Compound name	Molecular formula	Molecular weight (g/mol)	RT (min)	Mass error (ppm)	Adduct	ID level	[a]CV for QC (%)	Low vs. high altitude samples		
								[b]FC	[c]VIP	[d]*p*-value with FDR
Flavonoids										
Luteolin 6-*C*-glucoside 8-*C*-arabinoside	$C_{27}H_{30}O_{16}$	610.1534	5.17	1	$[M-H]^-$	2	0.64	0.67	3.26	2.41E-02
Quercetin 3,7-dirhamnoside	$C_{27}H_{30}O_{15}$	594.1585	5.39	1	$[M-H]^-$	2	1.73	0.65	3.71	2.81E-04
Kaempferol-3-*O*-rutinoside	$C_{27}H_{30}O_{15}$	594.1585	6.09	4	$[M-H]^-$	4	0.86	0.77	2.34	1.70E-02
Vitexin 6”-(3-hydroxy-3-methylglutarate)	$C_{27}H_{28}O_{14}$	576.1479	6.86	1	$[M-H-H_2O]^-$	3	3.42	0.45	3.03	6.10E-05
Maysin	$C_{27}H_{28}O_{14}$	576.1479	5.86	1	$[M-H]^-$	2	3.53	0.53	3.58	1.47E-04
Cassiaoccidentalin B	$C_{27}H_{28}O_{14}$	576.1479	6.72	2	$[M-H]^-$	2	1.76	0.58	2.70	2.91E-04
Nicotiflorin	$C_{27}H_{30}O_{15}$	594.1585	5.96	0	$[M-H]^-$	2	2.95	0.68	1.88	8.91E-03
Scutellarein 4′-methyl ether 7-glucuronide	$C_{22}H_{20}O_{12}$	476.0955	6.44	1	$[M-H]^-$	2	1.46	0.44	2.68	3.25E-06
Denticulaflavonol	$C_{35}H_{42}O_6$	558.2981	12.92	9	$[M+Cl]^-$	2	0.94	0.26	2.03	7.02E-12
Apigenin 7-[6”-(3-Hydroxy-3-methylglutaryl)glucoside]	$C_{27}H_{28}O_{14}$	576.1479	8.29	5	$[M-H-H_2O]^-$	4	1.00	0.61	2.21	1.23E-03
Quercetin 3-(2″,3″,4″-triacetylgalactoside)	$C_{27}H_{26}O_{15}$	590.1272	6.90	1	$[M-H-H_2O]^-$	3	0.74	0.35	2.33	5.17E-07
Hosloppin	$C_{22}H_{16}O_7$	392.0896	8.82	7	$[M+HCOO-H]^-$	2	0.62	0.68	1.54	1.25E-03
Vitexin 2″-*O*-rhamnoside*	$C_{27}H_{30}O_{14}$	578.1636	5.96	3	$[M-H]^-$	2	0.58	0.67	1.87	1.41E-03
Allivicin	$C_{27}H_{30}O_{16}$	610.1534	5.81	1	$[M-H]^-$	2	1.38	0.75	2.35	2.99E-02
Paniculatin	$C_{27}H_{30}O_{15}$	594.1585	6.86	0	$[M-H-H_2O]^-$	4	2.88	0.50	2.42	1.24E-03
6″-*O*-(3-Hydroxy-3-methylglutaroyl)astragalin	$C_{27}H_{28}O_{15}$	592.1428	6.21	1	$[M-H]^-$	2	1.60	0.58	1.31	3.73E-03
Baicalin	$C_{21}H_{18}O_{11}$	446.0849	6.18	3	$[M-H]^-$	4	2.31	0.47	1.59	2.05E-03
4′-*O*-Methylneobavaisoflavone 7-*O*-(2″-*p*-coumaroylglucoside)	$C_{36}H_{36}O_{11}$	644.2258	6.14	8	$[M-H]^-$	3	0.63	2.14	1.37	8.96E-03
Apigenin 7-[rhamnosyl-(1->2)-galacturonide]	$C_{27}H_{28}O_{15}$	592.1428	6.21	2	$[M-H]^-$	4	2.60	0.66	1.09	2.01E-03
Isorhamnetin 3-galactoside-7-rhamnoside	$C_{28}H_{32}O_{16}$	624.1690	5.52	2	$[M-H]^-$	2	0.71	0.67	1.47	8.10E-03
Violanthin	$C_{27}H_{30}O_{14}$	578.1636	6.51	4	$[M-H]^-$	4	1.05	0.56	1.27	2.10E-05
Eruberin B	$C_{30}H_{40}O_{15}$	640.2367	6.23	2	$[M-H]^-$	3	1.81	5.15	1.07	9.72E-04
Kaempferol 7-sophoroside	$C_{27}H_{30}O_{16}$	610.1534	5.17	6	$[M-H]^-$	2	0.26	0.74	2.44	4.45E-02
Vicenin 2*	$C_{27}H_{30}O_{15}$	594.1585	5.96	1	$[M-H]^-$	2	3.76	0.74	1.74	2.13E-02
Thonningianin B	$C_{35}H_{30}O_{17}$	722.1483	9.07	3	$[M+HCOO-H]^-$	2	0.86	0.58	1.49	1.25E-04
Bracteoside	$C_{22}H_{20}O_{12}$	476.0955	6.01	2	$[M-H]^-$	4	5.58	0.48	1.07	4.91E-05
Orientin 2″-rhamnoside	$C_{27}H_{30}O_{15}$	594.1585	6.09	1	$[M-H]^-$	2	0.77	0.74	2.47	6.33E-03
7,8,3′,4′-Tetrahydroxyflavanone 7-(2,4,6-triacetylglucoside)	$C_{27}H_{28}O_{14}$	576.1479	6.72	1	$[M-H]^-$	2	1.66	0.57	2.72	2.80E-04

(Continued on following page)

TABLE 2 (*Continued*) Significantly differential metabolites found the altitudinal variation.

Compound name	Molecular formula	Molecular weight (g/mol)	RT (min)	Mass error (ppm)	Adduct	ID level	[a]CV for QC (%)	Low vs. high altitude samples		
								[b]FC	[c]VIP	[d]*p*-value with FDR
Epigallocatechin 3-gallate	$C_{22}H_{18}O_{11}$	458.0849	7.93	2	$[M-H]^-$	2	2.83	0.45	1.15	3.29E-06
Glychalcone A	$C_{22}H_{22}O_5$	366.1467	1.34	4	$[M-H]^-$	4	1.85	0.48	1.29	6.72E-08
Quercetin 3-(3″,6″-diacetylgalactoside)	$C_{25}H_{24}O_{14}$	548.1166	6.94	6	$[M-H-H_2O]^-$	3	1.39	0.52	1.06	9.13E-05
2″,4″-Diacetylafzelin	$C_{25}H_{24}O_{12}$	516.1268	6.79	1	$[M-H]^-$	3	0.39	0.73	2.54	3.65E-03
Epigallocatechin 3-*O*-caffeate	$C_{24}H_{20}O_{10}$	468.1057	8.30	5	$[M+HCOO-H]^-$	3	1.06	0.48	1.16	1.05E-06
Cyanidin 3-rutinoside	$C_{27}H_{31}O_{15}$	595.1663	6.72	8	$[M-H-H_2O]^-$	4	1.72	0.73	1.29	1.24E-02
5,7,3′,4′-Tetrahydroxyflavanone 7-alpha-L-arabinofuranosyl-(1->6)-glucoside	$C_{26}H_{30}O_{15}$	582.1585	5.17	2	$[M+HCOO-H]^-$	4	0.24	0.63	2.60	2.76E-03
Isoorientin 2″-*O*-rhamnoside	$C_{27}H_{30}O_{15}$	594.1585	5.86	2	$[M-H]^-$	4	1.15	0.60	2.82	1.20E-03
Neosaponarin	$C_{27}H_{30}O_{15}$	594.1585	6.72	2	$[M-H]^-$	4	0.99	0.64	2.92	1.85E-03
Chamaemeloside	$C_{27}H_{28}O_{14}$	576.1479	6.08	1	$[M-H]^-$	2	1.39	0.59	1.68	3.16E-05
5′,5‴,8,8″-Tetrahydroxy-3′,3‴,4′,4‴,7′,7″-hexamethoxy-5,5″-biflavan	$C_{36}H_{38}O_{12}$	662.2363	5.90	8	$[M-H-H_2O]^-$	3	0.33	0.61	1.45	1.39E-05
Kaempferol 7-neohesperidoside	$C_{27}H_{30}O_{15}$	594.1585	6.09	1	$[M-H]^-$	2	0.85	0.72	2.56	4.44E-03
Saponarin*	$C_{27}H_{30}O_{15}$	594.1585	5.40	2	$[M-H]^-$	2	1.32	0.69	2.78	7.22E-04
Astragalin 7-rhamnoside	$C_{27}H_{30}O_{15}$	594.1585	6.54	1	$[M-H-H_2O]^-$	3	1.66	0.51	1.61	5.64E-05
Apigenin 7-glucuronide-4′-rhamnoside	$C_{27}H_{28}O_{15}$	592.1428	6.54	1	$[M-H-H_2O]^-$	3	1.95	0.50	1.64	5.20E-05
Fatty acids										
TriHODE	$C_{18}H_{32}O_5$	328.2250	8.07	1	$[M-H]^-$	4	0.36	0.65	3.90	2.34E-08
Sativic acid	$C_{18}H_{36}O_6$	348.2512	8.45	1	$[M-H-H_2O]^-$	4	0.26	0.61	3.96	4.91E-13
Coriolic acid	$C_{18}H_{32}O_3$	296.2351	11.70	1	$[M-H]^-$	2	0.75	0.35	2.40	7.02E-12
HoTrE	$C_{18}H_{30}O_3$	294.2195	11.20	1	$[M-H]^-$	2	0.76	0.33	2.44	1.84E-12
Dodecanedioic acid	$C_{12}H_{22}O_4$	230.1518	6.43	0	$[M-H]^-$	2	0.45	0.67	1.88	1.95E-04
TriHOME	$C_{18}H_{34}O_5$	330.2406	8.88	2	$[M-H]^-$	4	1.37	0.48	1.30	3.66E-07
Lauric acid	$C_{12}H_{22}O_3$	214.1569	8.39	1	$[M-H]^-$	2	0.64	0.43	1.14	2.13E-07
Undecylenic acid	$C_{11}H_{20}O_2$	184.1463	6.70	1	$[M+HCOO-H]^-$	4	0.61	0.68	1.26	1.89E-03
HpODE	$C_{18}H_{32}O_4$	312.2301	10.50	1	$[M-H]^-$	4	1.25	0.46	1.02	1.74E-07
HpOTrE	$C_{18}H_{30}O_4$	310.2144	10.09	1	$[M-H]^-$	4	1.74	0.41	1.03	2.56E-07
Undecenoic acid	$C_{11}H_{20}O_2$	184.1463	8.07	1	$[M+HCOO-H]^-$	4	0.87	0.65	1.23	3.88E-07
Cascarillic acid	$C_{11}H_{20}O_2$	184.1463	8.45	1	$[M+HCOO-H]^-$	4	0.88	0.53	1.43	5.49E-12
Sorbic acid	$C_6H_8O_2$	112.0524	3.33	1	$[M-H-H_2O]^-$	3	1.11	0.61	1.24	9.88E-07
Malyngic acid	$C_{18}H_{32}O_5$	328.2250	9.08	2	$[M-H]^-$	4	9.26	0.60	1.36	9.87E-07
Fulgidic acid	$C_{18}H_{32}O_5$	328.2250	8.55	1	$[M-H]^-$	4	1.53	0.56	1.01	1.06E-05

(Continued on following page)

TABLE 2 (*Continued*) Significantly differential metabolites found the altitudinal variation.

Compound name	Molecular formula	Molecular weight (g/mol)	RT (min)	Mass error (ppm)	Adduct	ID level	[a]CV for QC (%)	Low vs. high altitude samples		
								[b]FC	[c]VIP	[d]*p*-value with FDR
Hydroxyjasmonic acid	$C_{12}H_{18}O_4$	226.1205	7.50	1	$[M-H]^-$	2	0.83	0.61	1.04	3.25E-06
Norlinolenic acid	$C_{17}H_{28}O_2$	264.2089	9.08	2	$[M+HCOO-H]^-$	4	0.89	0.41	1.72	7.02E-12
Cinnamic acid derivatives										
O-Caffeoylquinic acid	$C_{25}H_{24}O_{12}$	516.1268	6.79	5	$[M-H]^-$	2	0.26	0.77	2.16	5.80E-03
p-Coumaroylquinic acid	$C_{16}H_{18}O_8$	338.1002	4.97	2	$[M-H]^-$	4	1.22	4.38	1.15	2.61E-06
O-Feruloyl-beta-D-glucose	$C_{16}H_{20}O_9$	356.1107	4.24	1	$[M-H-H_2O]^-$	4	1.39	2.39	1.14	2.12E-03
Dihydrocaffeic acid 3-*O*-glucuronide	$C_{15}H_{18}O_{10}$	358.0900	5.12	7	$[M+HCOO-H]^-$	2	0.60	0.73	1.37	1.12E-02
O-Feruloylgalactarate	$C_{16}H_{18}O_{11}$	386.0849	5.83	7	$[M+HCOO-H]^-$	2	0.45	0.62	2.09	1.80E-06
Caffeic acid 3-glucoside	$C_{15}H_{18}O_9$	342.0951	7.31	8	$[M+HCOO-H]^-$	4	1.05	0.66	1.49	3.44E-03
Dihydroferulic acid 4-*O*-glucuronide	$C_{16}H_{20}O_{10}$	372.1056	5.30	1	$[M-H]^-$	4	0.83	2.96	1.51	3.60E-07
1-Caffeoyl-4-deoxyquinic acid	$C_{16}H_{18}O_8$	338.1002	5.28	1	$[M-H]^-$	4	9.71	2.79	1.04	1.32E-02
Quinic acid	$C_7H_{12}O_6$	192.0634	4.73	2	$[M-H]^-$	2	0.91	5.05	1.42	3.14E-03
1-*O*-Sinapoylglucose	$C_{17}H_{22}O_{10}$	386.1213	5.48	1	$[M-H]^-$	4	0.88	6.16	1.24	8.80E-09
Unknowns										
Unknown 1 (396.036@4.25)	—	—	4.25	—	—	5	1.89	0.32	2.27	1.54E-07
Unknown 2 (586.0623@8.9)	—	—	8.90	—	—	5	1.77	0.67	1.49	4.69E-04
Unknown 3 (572.0832@8.44)	—	—	8.44	—	—	5	0.45	0.79	1.02	1.65E-02
Unknown 4 (255.989@8.9)	—	—	8.90	—	—	5	1.36	0.66	1.11	8.14E-05
Unknown 5	$C_{27}H_{36}O_{12}$	—	6.14	—	—	4	0.88	2.10	1.36	1.02E-02
Unknown 6	$C_{27}H_{36}O_{12}$	—	5.90	—	—	4	0.39	0.61	1.80	2.56E-05
Peptides										
Tripeptide 1	$C_{19}H_{25}N_3O_7$	407.1692	4.65	6	$[M-H]^-$	3	1.16	9.13	1.62	2.22E-09
Tripeptide 2	$C_{19}H_{25}N_3O_7$	407.1692	4.55	7	$[M-H]^-$	3	1.46	10.31	1.61	1.56E-10
Tripeptide 3	$C_{15}H_{20}N_4O_6$	352.1383	4.81	5	$[M-H]^-$	4	1.09	4.36	1.89	1.84E-12
Glycosylated lignans										
Prupaside	$C_{27}H_{36}O_{12}$	552.2207	5.90	1	$[M+HCOO-H]^-$	3	0.40	0.61	1.88	2.37E-05
Citrusin B	$C_{27}H_{36}O_{13}$	568.2156	6.27	2	$[M+HCOO-H]^-$	4	1.00	0.58	1.02	6.10E-05
Alkaloids										
2′-Norberbamunine	$C_{35}H_{38}N_2O_6$	582.2730	8.51	5	$[M+HCOO-H]^-$	4	1.30	0.39	1.07	4.43E-12
Vomilenine	$C_{21}H_{22}N_2O_3$	350.1630	4.88	8	$[M+Cl]^-$	4	0.50	1.36	1.44	1.04E-02
Carboxylic acids										
3,4,5-trihydroxy-6-(2-hydroxy-6-methoxyphenoxy)oxane-2-carboxylic acid	$C_{13}H_{16}O_9$	316.0794	2.06	2	$[M-H]^-$	4	0.88	2.14	1.54	5.09E-04

(Continued on following page)

TABLE 2 (*Continued*) Significantly differential metabolites found the altitudinal variation.

Compound name	Molecular formula	Molecular weight (g/mol)	RT (min)	Mass error (ppm)	Adduct	ID level	[a]CV for QC (%)	Low vs. high altitude samples		
								[b]FC	[c]VIP	[d]*p*-value with FDR
Phenols										
Phenol glucuronide	$C_{12}H_{14}O_7$	270.0740	2.06	0	$[M+HCOO-H]^-$	4	0.90	2.14	1.54	5.10E-04
Carbohydrates										
Ribulose	$C_5H_{10}O_5$	150.0528	0.60	0	$[M-H]^-$	4	0.76	0.75	1.14	1.21E-03
Steroids										
Physalin L	$C_{28}H_{32}O_{10}$	528.1995	1.54	3	$[M-H]^-$	2	1.11	0.40	1.08	1.07E-07
Glycosylated stilbenes										
Piceatannol 4′-galloylglucoside	$C_{27}H_{26}O_{13}$	558.1373	7.69	2	$[M-H]^-$	3	2.60	0.36	1.23	2.24E-05
Prenolipids										
Auxin b	$C_{18}H_{30}O_4$	310.2144	9.08	1	$[M-H]^-$	4	1.13	0.41	1.81	7.07E-12

[a]CV, coefficient of variation in the metabolites in the QC samples.
[b]FC, fold change in the comparison (average in low altitude/average in high altitude).
[c]VIP, variable importance in projection.
[d]*p*-value corresponding to the *p* values calculated by the Benjamini–Hochberg false discovery rate *post hoc* correction (FDR <0.05); *Metabolites annotated with GNPS. RT: retention time; Confidence levels in annotation were as following: Level 1: Confirmed structure, Level 2: Probable structure, Level 3: Tentative candidate(s), Level 4: Unequivocal molecular formula, Level 5: Exact mass (Schymanski et al., 2014).

metabolites comprises flavonoids which are the most frequently reported metabolites for these species. Many metabolites in this group showed an upward trend at higher altitudes, potentially contributing to the medicinal potential of these species.

3.2 Molecular networks and heatmap analyses reveal the chemical variation under the influence of altitudinal variation

To explore the leading chemical classes, we created a molecular network using the GNPS platform, which enabled us to visualize the chemical space of the metabolome in these species. Figure 4A illustrates the molecular network analyzed in the MolNetEnhancer platform, revealing the classes of MF present in the metabolome. We identified seven groups of molecular families, including phenylpropanoids and polyketides, lipids and lipid-like molecules, organic acid derivatives, lignans, neolignans and related compounds, alkaloids, and derivatives, organoheterocyclic compounds, and oxygenated organic compounds. However, due to the complexity of the metabolome in these species and the limitations of the spectral libraries, many of the nodes did not match any spectral reference. Therefore, it is crucial to continue exploring the chemical composition of these species, which is still limited in the literature. Notably, when we created a molecular network of the flavonoid cluster and compared the low vs. high altitude groups of the annotated compounds, we found that vitexin 2″-*O*-rhamnoside (FC = 1.494; *p*-value = 1.41E-03), saponarin (FC = 1.458; *p*-value = 7.22E-04), and vicenin 2 (FC = 1.360; *p*-value = 2.13E-02) (Supplementary Figures S2A–C) exhibited positive correlations at higher altitudes, confirming the statistical analysis performed previously (Figure 4B).

By conducting hierarchical clustering analysis and generating a heatmap based on the two most relevant groups of annotated metabolites that exhibited significant changes with altitude (flavonoids and cinnamic acid derivatives), we identified two distinct groups that were associated with altitude. We found that each group exhibited different patterns in terms of chemical composition and levels of presence and abundance. Figure 5 depicts the clustering of the two altitudinal ranges evaluated, with the first cluster corresponding to the high altitude samples that exhibited a high accumulation of flavonoids, while the low altitude samples exhibited a high accumulation of cinnamic acid derivatives.

3.3 Changes in the concentration and analysis of biosynthesis pathways of the most relevant metabolites in bamboo species

The box plot shows the differential metabolites for the two altitude groups. In Figure 6 we present the fold change of the main metabolites that exhibited a significant change. We found that cinnamic acid derivatives and flavonoids showed a significant change with the low and high altitude groups, respectively. Our study also highlighted the significance of quinic acid (QA) is a major

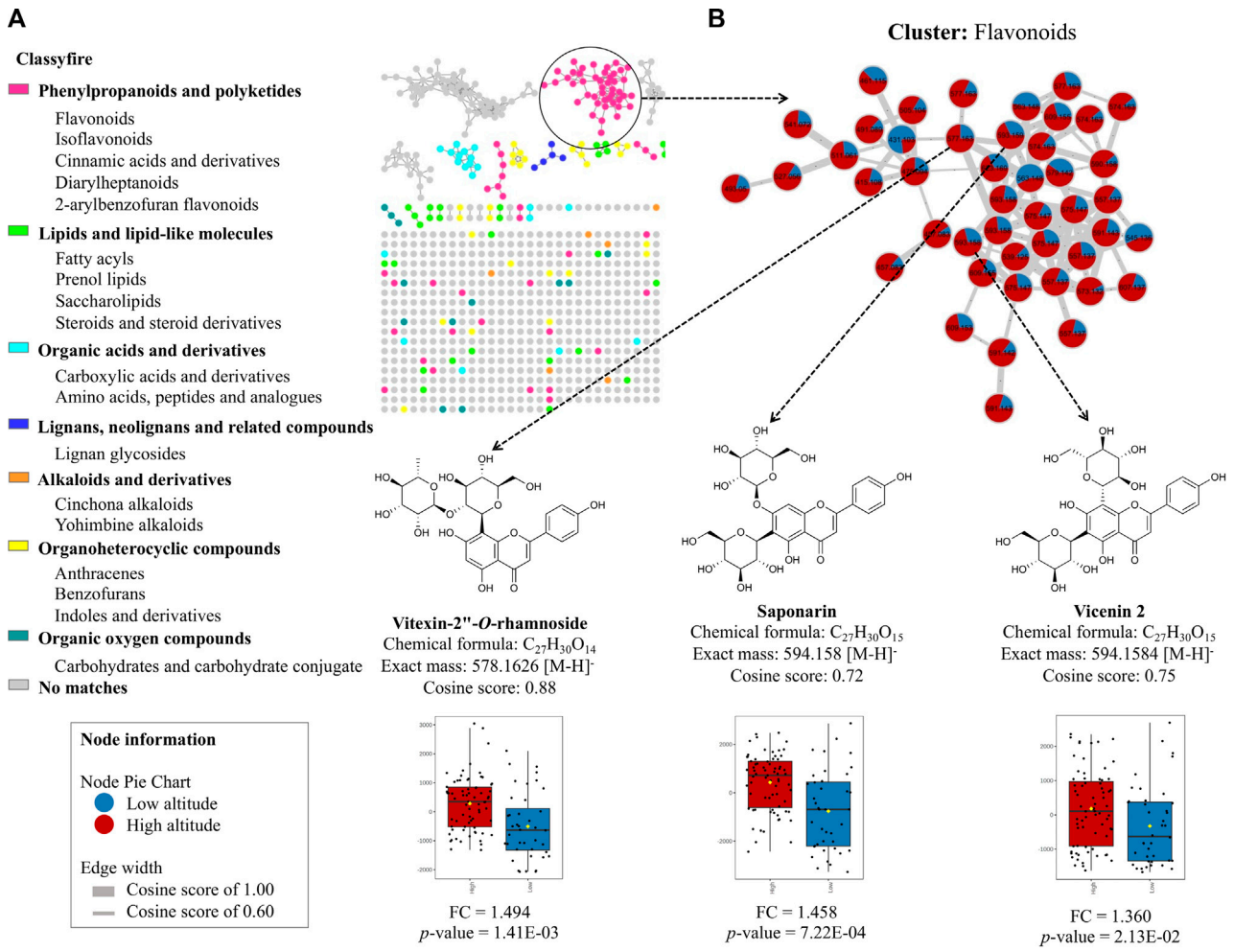


FIGURE 4
Molecular network analysis obtained from GNPS platform online. **(A)** Identification of major classes of chemical constituents of bamboo species using MolNetEnhancer technique and GNPS molecular networking. Color of the node is set according to the chemical class using "Classyfire". **(B)** Molecular network of the flavonoid cluster comparing low altitude (blue) and high altitude (red).

differential metabolite, which acts as a crucial precursor in the biosynthetic pathway of phenylpropanoids and is essential in the production of a diverse array of phenolic compounds.

3.4 Molecular networking for chemical space mapping of the metabolome of bamboo species by MolNetEnhancer

To complement the study on the chemical composition of the global metabolome of bamboo species, a molecular network was constructed using the GNPS platform to compare other types of flavonoids that were not statistically significant among the twelve species studied. Four flavonoid *C*-glycosides (isorhamnetin 7-rhamnoside, isovitexin, isoschaftoside, and rhoifolin) and one flavonoid *O*-glycoside (cyanidin 3-*O*-sophoroside) were annotated (Supplementary Figures S2D–H). It was observed that the *Guadua angustifolia* species was found to have an abundant profile of flavonoid *C*-glucosides, specifically of the compound isoschaftoside, which is functionally related to apigenin (Figure 7). In terms of flavonoid variation between species, isovitexin and isorhamnetin 7-rhamnoside were found in most species.

4 Discussion

Our findings indicate that flavonoids significantly increased at high altitudes, while cinnamic acid derivatives exhibited an increasing trend at low altitudes. Notably, flavonoids from *G. angustifolia* showed a positive correlation with the high altitude group, as did cinnamic acid derivatives from *G. aculeata*, *G. amplexifolia*, *G. angustifolia* var. *bicolor*, *G. angustifolia* biotype San Calixto, *G. incana*, *G. superba*, *G. uncinata*, *G. venezuelae*, *B. vulgaris*, and *P. aurea* with low altitude (Figure 8). However, further research is needed to determine the exact nature of this correlation. The effect of altitudinal variation on the chemical composition of the leaves of bamboo species is poorly documented. Previous studies have suggested that plants grown at high altitudes tend

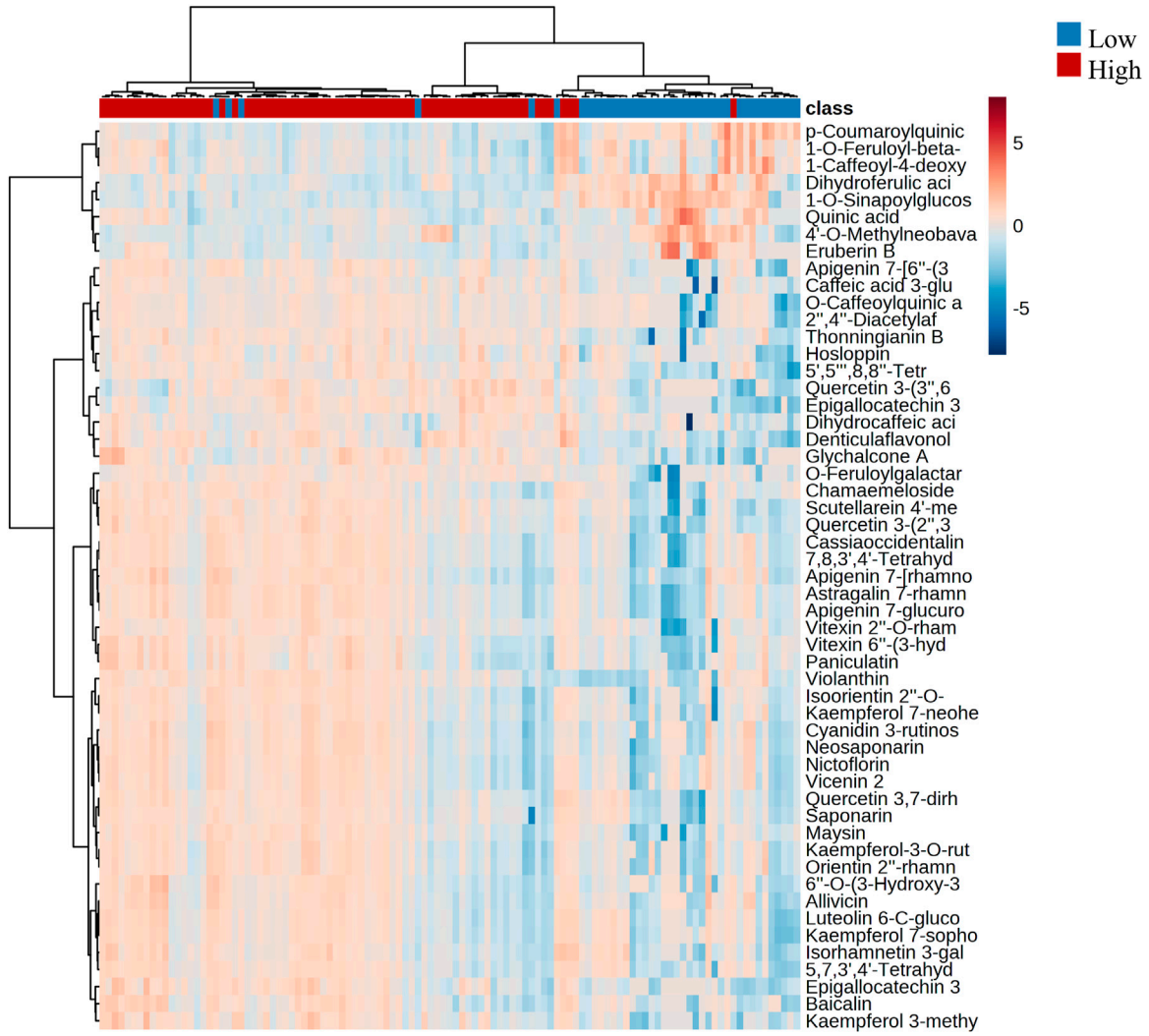


FIGURE 5
Hierarchical clustering with heatmap illustrating the differences in the metabolite abundance between low and high altitudes. The *x*-axis shows the clustering of all the samples, and the *y*-axis shows the clustering of the annotated flavonoids and cinnamic acid derivatives.

to have higher levels of flavonoids compared to those grown at low altitudes (Wang et al., 2020; Zhou et al., 2021). This phenomenon can be attributed to an increase of in ultraviolet radiation, illumination time and the delay of the phenophase of the plant, along with elevation, which is a protective mechanism that plants use against unfavorable environmental conditions (Ni et al., 2013). In a recent study, Zhou et al. (2021) investigated the influence of an altitudinal gradient on the variation of flavonoids in *Agriophyllum squarrosum* and found that these metabolites were enriched at high altitudes. The study also demonstrated a strong positive correlation between the contents of flavonoids, such as quercetin, tricine, and rutin, and environmental variables, such as latitude, longitude, and precipitation gradients. These findings corroborate our results and provide valuable information on the variation in the chemical profile of bamboo under various abiotic factors such as altitude, temperature, light, and soil, etc.

The metabolic differences expressed in bamboo species between high and low altitudes provide valuable information about the ideal growing conditions to promote the production of phenolic compounds with important biological properties in medicine. Moreover, the bamboo leaves have added value as are agro-industrial residues generated in the construction industry. Our results suggest that the accumulation of flavonoids at high altitudes is due to the adaptability of bamboo species to environmental conditions. These giant grasses biosynthesize flavonoid-type phenolic compounds that are used as defence and potential antioxidants (Falcone Ferreyra et al., 2012; Panche et al., 2016; Wang et al., 2018). Notably, the phenylpropanoid biosynthesis pathway, which is the starting point for producing many essential compounds such as flavonoids, coumarins, lignans, and hydroxycinnamic acid conjugates (Fraser and Chapple, 2011), is possibly the pathway that has been altered considering the results of this study. Flavonoid production occurs mainly through a diverse biosynthetic pathway involving the shikimate pathway and polyketide pathways (Mouradov and Spangenberg, 2014; Liu et al., 2021). Interestingly, the phenylpropanoid biosynthetic pathway showed a significant change in the presence of QA, with a fold change of 5.05, VIP of 1.42, and a

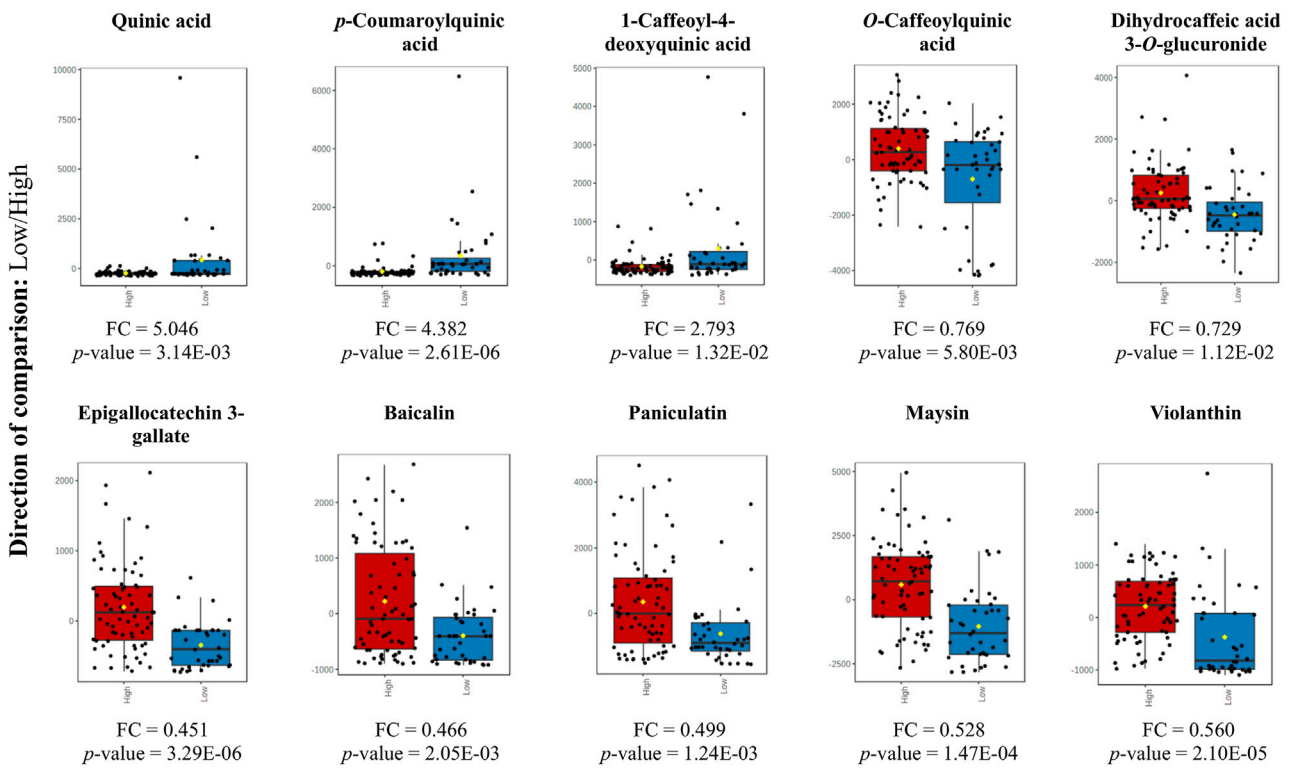


FIGURE 6
Box plot for the altered metabolites corresponding to cinnamic acid derivatives and flavonoids (p-value <0.05) at low altitude (blue) and high altitude (red). Y-axes are represented as relative units. The data were normalized with respect to the total spectral area. Bar charts show normalized values (mean ± one standard deviation). Boxes range from the 5% and 95% percentiles are indicated as error bars; individual data points are indicated by circles. The medians are indicated by horizontal lines within each box.

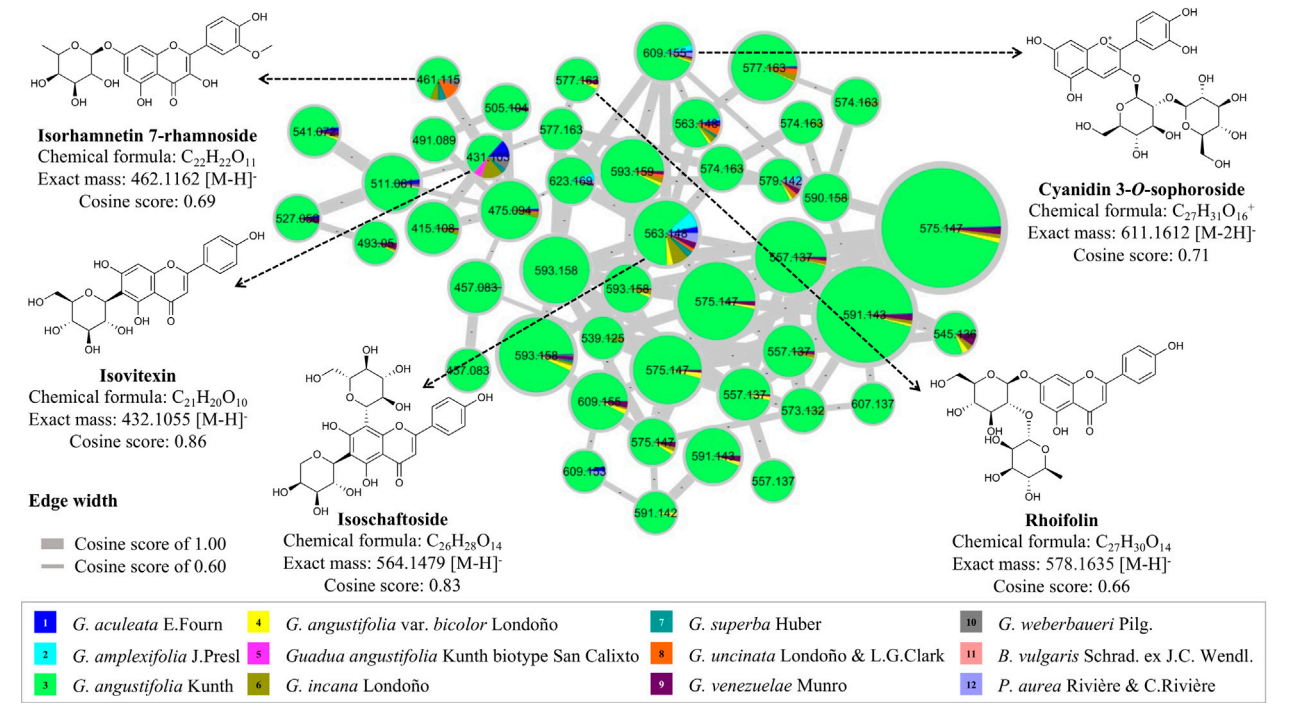


FIGURE 7
Molecular networking and dereplication of flavonoids comparing twelve bamboo species.

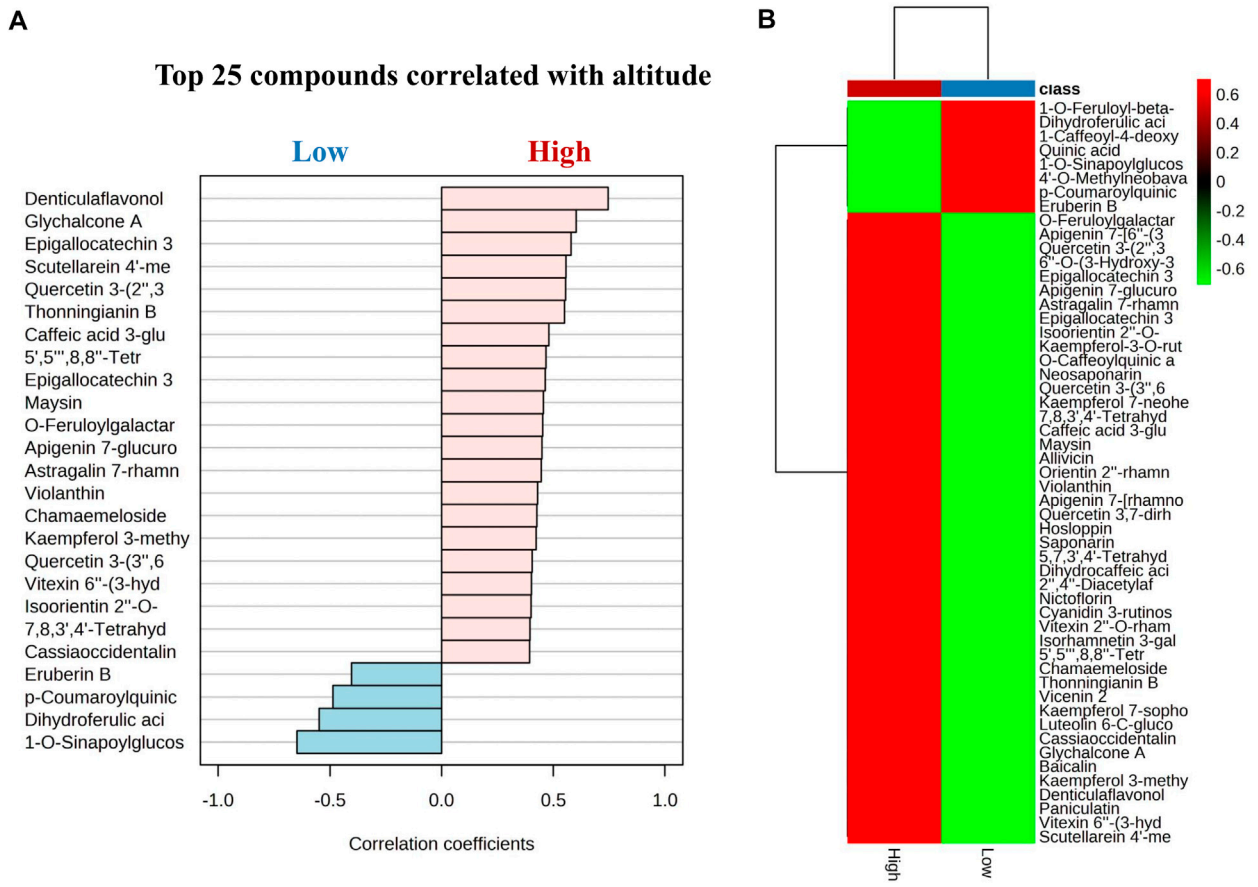


FIGURE 8
(A) Pearson's correlations using pattern search of the top 25 metabolites between low and high altitude. **(B)** Cluster heatmap based on correlation between flavonoids and acid cinnamic derivatives and the altitude variable. The analysis shows a positive (red) and negative (green) correlation with a *p*-value = 0.05 for all the metabolites.

p-value with FDR 3.14E-03. QA is a metabolite that is closely related to the biosynthesis pathway of caffeoylquinic acids (CQAs) and are specialized bioactive metabolites that are derived from the phenylpropanoid biosynthesis pathway. Consequently, QA is a critical intermediate in the biosynthesis of many flavonoids through cinnamic acid, which is a necessary precursor (Alcázar Magaña et al., 2021; Liu et al., 2021). Furthermore, the study revealed interesting chemical diversity in these species, with a predominance of groups of metabolites that are flavonoids and cinnamic acid derivatives. The mapping of the chemical space (including information about known reference spectra) allowed for the visualization of a large part of the global chemical composition of these species. However, the GNPS platform had a low annotation rate, resulting in many no-matches.

To complement the study of changes in chemical composition under the influence of altitude, the molecular network was used to analyze the specific cluster for flavonoids, comparing low and high altitudes. The study showed that the flavonoids such as vitexin 2″-*O*-rhamnoside, saponarin, and vicenin 2 had a positive tendency to increase at high altitudes. This finding suggests a close relationship between the concentration of phenolic compounds (flavonoids and cinnamic acid derivatives) in bamboo species. The heatmap shows a clear difference in the content of flavonoids and cinnamic acid derivatives between the groups compared under the effect of the altitudinal gradient, which could serve as marker compounds for chemical classification. The metabolites exposed to the variable altitude showed a significant difference, indicating that environmental factors have previously influenced the genetic and chemical diversity of plants (Pacheco-Hernández et al., 2021).

The metabolome of various bamboo species was compared, it was found that *G. angustifolia* had a rich profile of *C*-glycoside flavonoids. Similar metabolites have been reported for other bamboo species such as *P. nigra* var. *henonis* (Zhang et al., 2008; Ibrahim et al., 2021) *P. pubescens* (Tanaka et al., 2014) and *B. vulgaris* (Akhtar and Patowary, 2022). Some flavonoids were found to be shared among other species, such as *G. aculeata*., *G. angustifolia*, *G. angustifolia* biotype San Calixto, *G. incana*, *G. uncinata,* and *G. venezuelae*. Out of the twelve species analyzed, *B. vulgaris* and *P. aurea* were found to be the most studied species at the chemical level in Asia. However, considering the limited information available in the literature and the results obtained from the metabolomic

analysis, this study presents an opportunity to explore the metabolome of these species further, especially those belonging to the genus *Guadua* and distributed in the Neotropical region.

5 Conclusion

This study employed an untargeted metabolomics approach and molecular networking analysis to assess changes in the chemical composition of bamboo species due to variations in altitude. The study revealed that high altitude had a significant influence on the increase of flavonoid profiles, while low altitude led to an increase in cinnamic acid derivatives profiles. The molecular network analysis further demonstrated the diverse chemical composition of these species, including flavonoid glycosides, cinnamic acid derivatives, lignans, alkaloids, carbohydrates, and fatty acids. Conducting metabolomic studies on bamboo can provide a detailed understanding of its chemical composition and metabolic conservation, aiding in identifying patterns and trends in relation to environmental and cultivation factors. This information can help to enhance the production and quality of bamboo, identify bioactive compounds for natural health products, and improve sustainability, positively impacting the industry and the economy.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

LC and HL-P writing of the original manuscript; LC, HL-P, LD-A, and GC experimental design and activities; XL taxonomy, classification, and description of bamboo; XL supply of bamboo samples of Quindío; LC, HL-P, and MC data analysis; LC, HL-P, TL, and IC-G molecular networking analysis; XL, TL, MC, ER-S, LD-A, JP-R, IC-G, and GC writing, review and editing of the manuscript. GC for research supervision. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to acknowledge Pontificia Universidad Javeriana, al Ministerio de Ciencia, Tecnología e Innovación, al Ministerio de Educación Nacional, al Ministerio de Industria, Comercio y Turismo e ICETEX, 2ª Convocatoria Ecosistema Científico—Colombia Científica 792-2017, Programa "Generación de alternativas terapéuticas en cáncer a partir de plantas a través de procesos de investigación y desarrollo traslacional, articulados en sistemas de valor sostenibles ambiental y económicamente" (Contrato no. FP44842-221-2018) and the producers linked to the Federación Nacional de Cafeteros and Fundación Suyusama and El Paraiso del Bambú y la Guadua for providing the raw material from different *Guadua* species. We would also thank the Ministerio de Ambiente y Desarrollo Sostenible for allowing the use of genetic resources and products derived (Contract number 212/2018; Resolution 210/2020).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1192088/full#supplementary-material

References

Akhtar, J., and Patowary, L. (2022). *Bambusa vulgaris*: A comprehensive review of its traditional uses, phytochemicals and pharmacological activities. *Sci. Phytochemistry* 1, 11–21.

Alcázar Magaña, A., Kamimura, N., Soumyanath, A., Stevens, J. F., and Maier, C. S. (2021). Caffeoylquinic acids: Chemistry, biosynthesis, occurrence, analytical challenges, and bioactivity. *Plant J.* 107, 1399. doi:10.1111/TPJ.15390

Aron, A. T., Gentry, E. C., McPhail, K. L., Nothias, L. F., Nothias-Esposito, M., Bouslimani, A., et al. (2020). Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat. Protoc.* 15, 1954–1991. doi:10.1038/s41596-020-0317-5

Badshah, S. L., Faisal, S., Muhammad, A., Poulson, B. G., Emwas, A. H., and Jaremko, M. (2021). Antiviral activities of flavonoids. *Biomed. Pharmacother.* 140, 111596. doi:10.1016/J.BIOPHA.2021.111596

Cheng, Y., Wan, S., Yao, L., Lin, D., Wu, T., Chen, Y., et al. (2023). Bamboo leaf: A review of traditional medicinal property, phytochemistry, pharmacology, and purification technology. *J. Ethnopharmacol.* 306, 116166. doi:10.1016/j.jep.2023.116166

Chongtham, N., and Bisht, M. S. (2020). *Bamboo shoot: Superfood for nutrition, health and medicine*. Boca Raton; London: CRC Press.

Clark, L. G., Londoño, X., and Ruiz-Sanchez, E. (2015). "Bamboo taxonomy and habitat," in *Springer* (Springer International Publishing Switzerland), 1–30. doi:10.1007/978-3-319-14133-6_1

Coffie, G., Antwi-Boasiako, C., and Darkwa, N. (2014). Phytochemical constituents of the leaves of three bamboo (Poaceae) species in Ghana. *J. Pharmacogn. Phytochem.* 2, 34–38.

Da Silva, R. R., Wang, M., Nothias, L. F., van der Hooft, J. J. J., Caraballo-Rodríguez, A. M., Fox, E., et al. (2018). Propagating annotations of molecular networks using *in silico* fragmentation. *PLoS Comput. Biol.* 14, e1006089. doi:10.1371/journal.pcbi.1006089

Dhami, N., and Mishra, A. D. (2015). Phytochemical variation: How to resolve the quality controversies of herbal medicinal products? *J. Herb. Med.* 5, 118–127. doi:10.1016/j.hermed.2015.04.002

Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., et al. (2016). ClassyFire: Automated chemical classification with a comprehensive, computable taxonomy. *J. Cheminform* 8, 1–20. doi:10.1186/s13321-016-0174-y

Ernst, M., KangBinCaraballo-Rodríguez, K. A. M., Nothias, L. F., Wandy, J., Chen, C., et al. (2019). MolNetEnhancer: Enhanced molecular networks by integrating metabolome mining and annotation tools. *Metabolites* 9. doi:10.3390/METABO9070144

Falcone Ferreyra, M. L., Rius, S. P., and Casati, P. (2012). Flavonoids: Biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* 3, 222. doi:10.3389/FPLS.2012.00222/BIBTEX

Fraser, C. M., and Chapple, C. (2011). *The phenylpropanoid pathway in arabidopsis*. American Society of Plant Biologists. doi:10.1199/TAB.0152

Gagliano, J., Anselmo-Moreira, F., Sala-Carvalho, W. R., and Furlan, C. M. (2022). What is known about the medicinal potential of bamboo? *Adv. Traditional Med.* 22, 467–495. doi:10.1007/s13596-020-00536-5

Gomez, J. P., Velez, J. P. A., Pinzon, M. A., Arango, J. A. M., and Muriel, A. P. (2021). Chemical characterization and antiradical properties of pyroligneous acid from a preserved Bamboo, *Guadua angustifolia* Kunth. *Braz. Archives Biol. Technol.* 64, 1–13. doi:10.1590/1678-4324-2021190730

Ibrahim, A., Abdelhameed, R., Habib, E., Ahmed, S., and Badr, J. (2021). Biological activities of different species of the genus *Phyllostachys*. *Rec. Pharm. Biomed. Sci.* 5, 64–73. doi:10.21608/rpbs.2021.61559.1092

Indira, A., Santosh, O., Koul, A., and Nirmala, C. (2022). Comparative assessment of the antioxidant potential of bamboo leaves, along with some locally and commercially consumed beverages in India. *Adv. Bamboo Sci.* 1, 100007. doi:10.1016/j.bamboo.2022.100007

Khalil, N., El-Jalel, L., Yousif, M., and Gonaid, M. (2020). Altitude impact on the chemical profile and biological activities of *Satureja thymbra* L. essential oil. *BMC Complement. Med. Ther.* 20, 186. doi:10.1186/S12906-020-02982-9

Ko, H. C., Lee, J. Y., Jang, M. G., Song, H., and Kim, S. J. (2018). Seasonal variations in the phenolic compounds and antioxidant activity of *Sasa quelpaertensis*. *Ind. Crops Prod.* 122, 506–512. doi:10.1016/j.indcrop.2018.06.031

Kumari, A., Lakshmi, G. A., Krishna, G. K., Patni, B., Prakash, S., Bhattacharyya, M., et al. (2022). Climate change and its impact on crops: A comprehensive investigation for sustainable agriculture. *Agronomy* 12, 3008. doi:10.3390/agronomy12123008

Liese, W., Welling, J., and Tang, T. K. H. (2015). "Utilization of bamboo," in *Bamboo: The plant and its uses* (Hamburg; Germany: Springer International Publishing), 299–346. doi:10.1007/978-3-319-14133-6_10

Liu, W., Feng, Y., Yu, S., Fan, Z., Li, X., Li, J., et al. (2021). The flavonoid biosynthesis network in plants. *Int. J. Mol. Sci.* 22. doi:10.3390/IJMS222312824

Maleki, S. J., Crespo, J. F., and Cabanillas, B. (2019). Anti-inflammatory effects of flavonoids. *Food Chem.* 299, 125124. doi:10.1016/J.FOODCHEM.2019.125124

Ming, C., Jye, W., and Ahmad, H. (2017). Mechanical properties of bamboo and bamboo composites: A review. *J. Adv. Res. Mater. Sci.* 35, 7–26.

Mohimani, H., Gurevich, A., Mikheenko, A., Garg, N., Nothias, L. F., Ninomiya, A., et al. (2017). Dereplication of peptidic natural products through database search of mass spectra. *Nat. Chem. Biol.* 13, 30–37. doi:10.1038/nchembio.2219

Mouradov, A., and Spangenberg, G. (2014). Flavonoids: A metabolic network mediating plants adaptation to their real estate. *Front. Plant Sci.* 5, 1–16. doi:10.3389/FPLS.2014.00620

Ni, Q., Wang, Z., Xu, G., Gao, Q., Yang, D., Morimatsu, F., et al. (2013). Altitudinal variation of antioxidant components and capability in *Indocalamus latifolius* (Keng) McClure leaf. *J. Nutr. Sci. Vitaminol. (Tokyo)* 59, 336–342. doi:10.3177/jnsv.59.336

Ni, Q., Xu, G., Wang, Z., Gao, Q., Wang, S., and Zhang, Y. (2012). Seasonal variations of the antioxidant composition in ground bamboo *Sasa argenteastriatus* leaves. *Int. J. Mol. Sci.* 13, 2249–2262. doi:10.3390/ijms13022249

Okido, T., Ino, M., Kamada, I., Kihara, J., and Ueno, M. (2022). Biological activity of Kuma bamboo grass (*Sasa veitchii*) extract against the fungal causal agent (*Pyricularia oryzae*) of blast disease. *Adv. Bamboo Sci.* 1, 100004. doi:10.1016/j.bamboo.2022.100004

Pacheco-Hernández, Y., Villa-Ruano, N., Lozoya-Gloria, E., Barrales-Cortés, C. A., Jiménez-Montejo, F. E., and Cruz-López, M. D. C. (2021). Influence of environmental factors on the genetic and chemical diversity of *Brickellia veronicifolia* populations growing in fragmented shrublands from Mexico. *Plants* 10, 1–21. doi:10.3390/plants10020325

Panche, A. N., Diwan, A. D., and Chandra, S. R. (2016). Flavonoids: An overview. *J. Nutr. Sci.* 5, 1–15. doi:10.1017/jns.2016.41

Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N., et al. (2022). Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. *Nat. Protoc.* 17, 1735–1761. doi:10.1038/s41596-022-00710-w

Pant, P., Pandey, S., and Dall'Acqua, S. (2021). The influence of environmental conditions on secondary metabolites in medicinal plants: A literature review. *Chem. Biodivers.* 18, e2100345. doi:10.1002/CBDV.202100345

Ramabulana, A. T., Petras, D., Madala, N. E., and Tugizimana, F. (2021). Metabolomics and molecular networking to characterize the chemical space of four *Momordica* plant species. *Metabolites* 11, 763. doi:10.3390/METABO11110763/S1

Sampaio, B. L., Edrada-Ebel, R., and Da Costa, F. B. (2016). Effect of the environment on the secondary metabolic profile of *Tithonia diversifolia*: A model for environmental metabolomics of plants. *Sci. Rep.* 6, 1–11. doi:10.1038/srep29265

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., et al. (2014). Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environ. Sci. Technol.* 48, 2097–2098. doi:10.1021/es5002105

Shen, S., Zhan, C., Yang, C., Fernie, A. R., and Luo, J. (2023). Metabolomics-centered mining of plant metabolic diversity and function: Past decade and future perspectives. *Mol. Plant* 16, 43–63. doi:10.1016/j.molp.2022.09.007

Speisky, H., Shahidi, F., de Camargo, A. C., and Fuentes, J. (2022). Revisiting the oxidation of flavonoids: Loss, conservation or enhancement of their antioxidant properties. *Antioxidants* 11, 133. doi:10.3390/antiox11010133

Tamang, M., Nandy, S., Srinet, R., Das, A. K., and Padalia, H. (2022). Bamboo mapping using earth observation data: A systematic review. *J. Indian Soc. Remote Sens.* 50, 2055–2072. doi:10.1007/S12524-022-01600-0

Tanaka, A., Zhu, Q., Tan, H., Horiba, H., Ohnuki, K., Mori, Y., et al. (2014). Biological activities and phytochemical profiles of extracts from different parts of bamboo (*Phyllostachys pubescens*). *Molecules* 19, 8238–8260. doi:10.3390/molecules19068238

Verpoorte, R., Choi, Y. H., and Kim, H. K. (2010). Metabolomics: what's new? *Flavour Fragr. J.* 25, 128–131. doi:10.1002/ffj.1982

Wang, L., Yuan, S., Nie, Y., Zhao, J., Cao, X., Dai, Y., et al. (2020). Dietary flavonoids and the altitudinal preference of wild giant pandas in Foping National Nature Reserve, China. *Glob. Ecol. Conserv.* 22, e00981. doi:10.1016/j.gecco.2020.e00981

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* 34, 828–837. doi:10.1038/nbt.3597

Wang, T. Y., Li, Q., and Bi, K. S. (2018). Bioactive flavonoids in medicinal plants: Structure, activity and biological fate. *Asian J. Pharm. Sci.* 13, 12–23. doi:10.1016/j.ajps.2017.08.004

Xie, Y., Yang, W., Tang, F., Chen, X., and Ren, L. (2015). Antibacterial activities of flavonoids: Structure-activity relationship and mechanism. *Curr. Med. Chem.* 22, 132–149. doi:10.2174/0929867321666140916113443

Zhang, Y., Jiao, J., Liu, C., Wu, X., and Zhang, Y. (2008). Isolation and purification of four flavone *C*-glycosides from antioxidant of bamboo leaves by macroporous resin column chromatography and preparative high-performance liquid chromatography. *Food Chem.* 107, 1326–1336. doi:10.1016/j.foodchem.2007.09.037

Zhou, S., Yan, X., Yang, J., Qian, C., Yin, X., Fan, X., et al. (2021). Variations in flavonoid metabolites along altitudinal gradient in a desert medicinal plant *Agriophyllum squarrosum*. *Front. Plant Sci.* 12, 683265. doi:10.3389/fpls.2021.683265

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 29 June 2023
DOI 10.3389/fmolb.2023.1204273

OPEN ACCESS

EDITED BY
Martin Aran,
IIBBA-CONICET Leloir Institute Foundation, Argentina

REVIEWED BY
Sakshi Arora,
Children's Hospital of Philadelphia, United States
Gaia Meoni,
University of Florence, Italy

*CORRESPONDENCE
León Gabriel Gómez-Archila,
lgabriel.gomez@udea.edu.co
Martina Palomino-Schätzlein,
mpalomino@protoqsar.com

RECEIVED 12 April 2023
ACCEPTED 16 June 2023
PUBLISHED 29 June 2023

CITATION
Gómez-Archila LG, Palomino-Schätzlein M, Zapata-Builes W, Rugeles MT and Galeano E (2023), Plasma metabolomics by nuclear magnetic resonance reveals biomarkers and metabolic pathways associated with the control of HIV-1 infection/progression. *Front. Mol. Biosci.* 10:1204273. doi: 10.3389/fmolb.2023.1204273

Plasma metabolomics by nuclear magnetic resonance reveals biomarkers and metabolic pathways associated with the control of HIV-1 infection/progression

León Gabriel Gómez-Archila[1,2]*, Martina Palomino-Schätzlein[3]*, Wildeman Zapata-Builes[4,5], Maria T. Rugeles[4] and Elkin Galeano[1]

[1]Grupo de Investigación en Sustancias Bioactivas, Facultad de Ciencias Farmacéuticas y Alimentarias, Universidad de Antioquia (UdeA), Medellín, Colombia, [2]Grupo de Investigación en Ciencias Farmacéuticas ICIF-CES, Facultad de Ciencias y Biotecnología, Universidad CES, Medellín, Colombia, [3]Servicio de RMN, Centro de Investigación Príncipe Felipe, Valencia, España, [4]Grupo Inmunovirología, Facultad de Medicina, Universidad de Antioquia (UdeA), Medellín, Colombia, [5]Grupo Infettare, Facultad de Medicina, Universidad Cooperativa de Colombia, Medellín, Colombia

How the human body reacts to the exposure of HIV-1 is an important research goal. Frequently, HIV exposure leads to infection, but some individuals show natural resistance to this infection; they are known as HIV-1-exposed but seronegative (HESN). Others, although infected but without antiretroviral therapy, control HIV-1 replication and progression to AIDS; they are named controllers, maintaining low viral levels and an adequate count of $CD4^+$ T lymphocytes. Biological mechanisms explaining these phenomena are not precise. In this context, metabolomics emerges as a method to find metabolites in response to pathophysiological stimuli, which can help to establish mechanisms of natural resistance to HIV-1 infection and its progression. We conducted a cross-sectional study including 30 HESN, 14 HIV-1 progressors, 14 controllers and 30 healthy controls. Plasma samples (directly and deproteinized) were analyzed through Nuclear Magnetic Resonance (NMR) metabolomics to find biomarkers and altered metabolic pathways. The metabolic profile analysis of progressors, controllers and HESN demonstrated significant differences with healthy controls when a discriminant analysis (PLS-DA) was applied. In the discriminant models, 13 metabolites associated with HESN, 14 with progressors and 12 with controllers were identified, which presented statistically significant mean differences with healthy controls. In progressors, the metabolites were related to high energy expenditure (creatinine), mood disorders (tyrosine) and immune activation (lipoproteins), phenomena typical of the natural course of the infection. In controllers, they were related to an inflammation-modulating profile (glutamate and pyruvate) and a better adaptive immune system response (acetate) associated with resistance to progression. In the HESN group, with anti-inflammatory (lactate and phosphocholine) and virucidal (lactate) effects which constitute a protective profile in the sexual transmission of HIV. Concerning the significant metabolites of each group, we identified 24 genes involved in HIV-1 replication or virus proteins that were all altered in progressors but only partially in controllers and HESN. In summary, our results indicate that exposure to HIV-1 in

HESN, as well as infection in progressors and controllers, affects the metabolism of individuals and that this affectation can be determined using NMR metabolomics.

KEYWORDS

NMR, metabolomics, HIV-1, biomarkers, pathways

Introduction

Human immunodeficiency virus type 1 (HIV-1), the causal agent of the acquired immunodeficiency syndrome (AIDS) in humans (Barré-Sinoussi et al., 1983; Gallo et al., 1983), continues to be a serious public health problem, after nearly 40 years of research (ONUSIDA and Comunicaciones y Promoción Mundial, 2021), generating considerable mortality among those infected, and an excessive cost for the healthcare system (Roth et al., 2018; Frank et al., 2019).

The progression of infection from the acute phase to advanced infection or AIDS is a very complex process, which takes approximately 10 years in absence of treatment (Vergis and Mellors, 2000; Kumar, 2013). Some individuals can naturally control HIV-1 replication, maintaining low viral load (VL) levels and an adequate count of $CD4^+$ T lymphocytes, in the absence of antiretroviral therapy (ART) for at least 1 year (Cao et al., 1995; Baker et al., 2009). These individuals are known as controllers (elite or viremic) and exhibit specific resistance mechanisms to disease progression, including the presence of HLA alleles, HLA-B27 and HLA-B57 (Gonzalo-Gil et al., 2017).

Likewise, researchers have tried to characterize the natural resistance to HIV-1 infection among people exposed to the virus, who remain seronegative, known as HIV-exposed seronegative (HESN) individuals (Meyers and Fowke, 2010; Young et al., 2011). To date, only the homozygous *Δ32* mutation in the *CCR5* gene, the main entry coreceptor of the virus, has been consistently associated with host resistance to HIV-1 in less than 3% of resistant individuals (Huang et al., 1996; Ding et al., 2021). Other known genetic and immunologic factors involved in resistance to HIV-1 infection only partially explain this phenomenon (Lederman et al., 2010; Taborda-Vanegas et al., 2011), which means that further mechanisms remain unclear.

In this context, metabolomics understood as the objective identification and quantification of small molecules in biological fluids (Nicholson et al., 1999), might help to understand the biochemical state of an organism for discovering biomarkers. Through case-control studies of metabolites in plasma, urine, or cells, by quantitative measurement using nuclear magnetic resonance (NMR) spectroscopy, different pathophysiological states have been explained (Bertini et al., 2012; Li and Deng, 2016; Paris et al., 2018), suggesting that metabolomics could be a potential tool for prognosis, diagnosis, and monitoring the efficacy of treatment (Puchades-Carrasco and Pineda-Lucena, 2015), including HIV-1 infection.

Studies of the HIV effects on metabolism during *in vitro* replication and infection in animal and human models have provided new insights and targets for biomarker development and therapy. To date, little is known about the metabolic profiles that generate resistance to infection or a differential response to AIDS and its progression.

In the current study, we hypothesize that differences in the phenotype of infected individuals with high or low viral loads and seronegative individuals continuously exposed to HIV-1 will result in a dissimilar metabolomic plasma profile. Therefore, we collected and analyzed plasma samples of age and sex-matched groups of progressors, controllers, HESN and healthy controls by proton Nuclear Magnetic Resonance Spectroscopy (^{1}H NMR). We aim to identify a specific metabolic fingerprint of each group and to obtain biomarkers related to HIV-1 progression and natural resistance, providing valuable information on the pathogenesis of HIV-1 infection.

Materials and methods

Chemicals and materials

All solvents and reagents were analytical grade, sodium phosphate dibasic dihydrate, sodium azide, deuterium oxide, 3-(Trimethylsilyl) propionic-2,2,3,3-d_4 acid sodium salt (TSP-d_4) and 3-(Trimethylsilyl)-1-propanesulfonic acid-d_6 sodium salt (DSS-d_6) were supplied by Merck (Germany). The ultrapure water was obtained in a Milli-Q purification system of Merck Millipore. The Vivaspin® 500 3000 K MWCO Centrifugal Concentrators were provided by Sartorius.

Human subjects

A retrospective cross-sectional study was developed using a defined database of volunteers to build case-control relationships. Plasma samples from 88 volunteers were evaluated distributed as follows:

- HESN: thirty individuals, from serodiscordant couples (couples in which one partner is HIV-positive and the other HIV-negative). HESN reported multiple unprotected sexual episodes for >2 years at the time of enrollment, with at least five episodes of at-risk intercourse within 6 months before study entry with an HIV positive partner with a detectable viral load (Zapata et al., 2008). The median VL of the partner was 2,569 RNA copies/mL (interquartile range = 400–25,250 copies/mL) (Aguilar-Jiménez et al., 2013). From these individuals, 10 (35%) were ART-naïve [VL median (interquartile range)] [10,257 (718–23,188)]. Eight (25%) were ART-responders VL < 400. Finally, 12 (40%) were ART-non-responders [(Fisher et al., 1987), 806 (18,200–118,770)]. No Δ32-homozygous subjects were included.
- Controllers: fourteen, with 1 year of diagnosis of HIV-1 infection, and viral load less than 2000 copies/mL in the

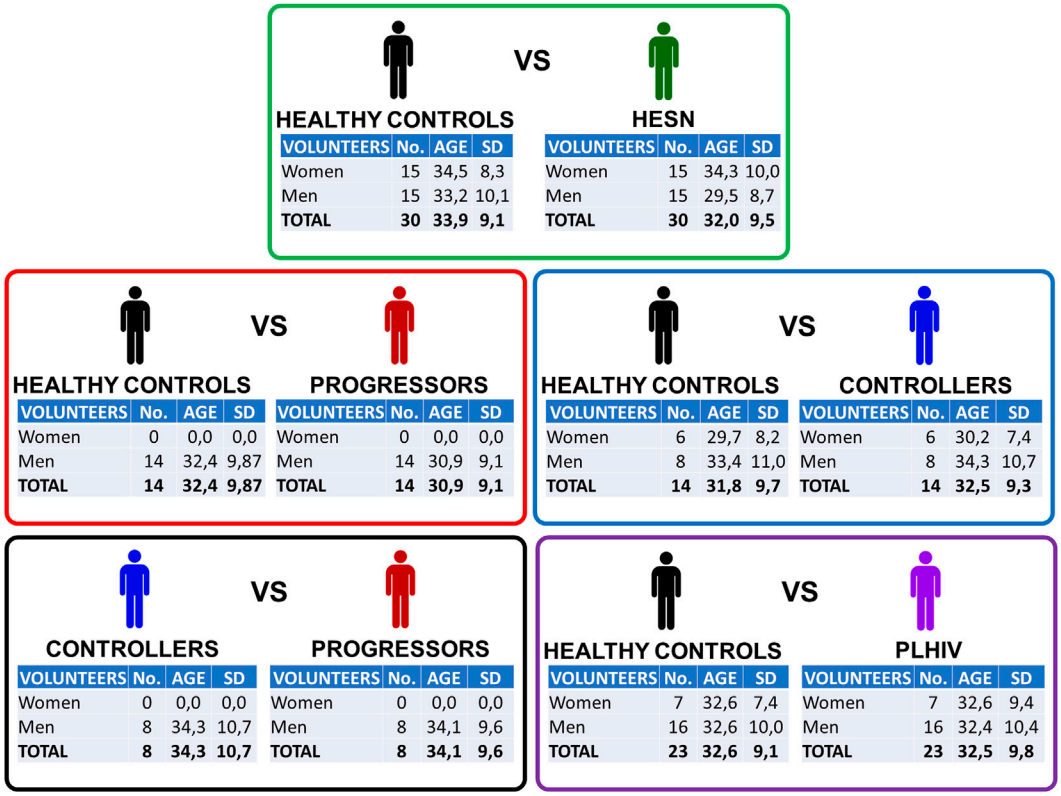


FIGURE 1
Case-control relationships. Detail of each of the five relationships built to analyze plasma samples in a direct and filtered way. To know the details of each of the volunteers who participated in the study, see Supplementary Table S1. Exclusion criteria: Individuals with hemoglobin ≤8.0 g/dl; neutrophil count ≤1,000/mm3; receiving some immunosuppressive treatment; pregnant or lactating women; cancer or with an active infection or disease requiring hospitalization. Age reported as the sample mean. SD: standard deviation.

absence of Antiretroviral therapy (ART) and normal $CD4^+$ T lymphocyte count (Pereyra et al., 2008). The median diagnosis time was 46 months (range 12–168). The median VL was 211 copies/mL (range 20–1885), and the median $CD4^+$ T cells count was 745 cells/uL (range 514–1,367). Only 2 (14%) controllers showed the HLA-B*27 allele, and 3 (21%) controllers showed the HLA-B*57 allele.

- Chronic progressors: fourteen, with a CD4 + T lymphocyte count> 350 cells/μL and a viral load between 10,000 and 100,000 copies/mL without receiving ART (Taborda et al., 2015). The median diagnosis time was 51 months (range 12–120). From these, ten individuals had between 1 and 5 years of infection, three reported 6–9 years, and one had 10 years of infection. The median VL was 31,552 copies/mL (range 11,206–160405) and the median $CD4^+$ T cells count was 443 cells/uL (range 267–819).
- PLHIV: people living with HIV. In this case it refers to a mix of controllers and progressors.
- Healthy controls: thirty, with negative serological tests for HIV-1 without risk behaviors.

This study was approved by the Bioethical Committee Universidad de Antioquia; and all the individuals signed informed consent prepared according to Colombian Legislation Resolution 008,430/1993.

Processing of blood samples

The Blood sample was collected from all participants by using potassium-EDTA collection tubes. Then, it was centrifuged at 1,000 x g for 5 min at 4°C, and 2 mL of plasma was stored at −80°C until processing. Two methodologies were established to process the biofluid: a direct analysis and a deproteinization analysis.

For the direct analysis 300 μL of Buffer pH 7.4 (Na_2HPO_4 75 mM DSS 2.3 mM and NaN_3 0.04%) was added to a microcentrifuge tube (1.5 or 2 mL) and reserved. Then, the plasma sample was thawed, homogenized, and 300 μL transferred to the previously mentioned vial with buffer. The resulting solution was mixed and 550 μL transferred to a 5 mm NMR tube for analysis. Tubes were degassed for 3 min before capping.

In the case of deproteinization analysis, 300 μL of Buffer pH 7.4 (Na_2HPO_4 75 mM TSP 2.3 mM and NaN_3 0.04%) was added to a microcentrifuge tube (1.5 or 2 mL) and reserved. Then, the vial containing the sample was thawed, the plasma was homogenized, and 500 μL of plasma were taken to a Vivaspin 500 centrifugal filter (Previously pre-washed 5 times with Buffer pH 7.4). The sample was centrifuged at 12,000 gravities at 4°C for 60 min 300 μL of the filtrate was brought to the microcentrifuge tube containing the 300 μL of Buffer pH 7.4 they were mixed with a micropipette. Finally, 550 μL of the solution were taken and transferred to a

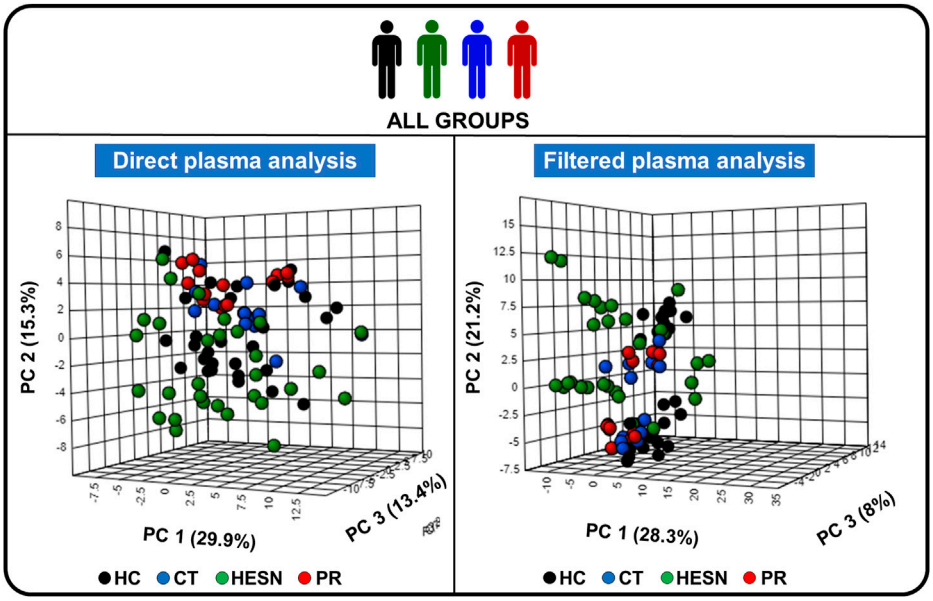


FIGURE 2
Principal component analysis (PCA) of the study groups. 3D score plot charts of the PCA analysis of all study volunteers, on the left side the analysis with direct plasma and on the right side with filtered plasma. PC: Principal component, HC: Healthy control, CT: controller, HESN: HIV-exposed seronegative, PR: Progressor. Each of the three axes of the graph represents a principal component. The values that each of the axes takes is related to the fact that there is a score value for each observation (row) in the data set; so, there is score values for the first component, another for the second component, and one for the third. The score value for an observation, say the first component, is the distance from the origin, along the direction (load vector) of the first component, to the point where that observation projects onto the direction vector.

5 mm NMR tube for analysis. The tubes were degassed for 3 min before being capped.

^{1}H-NMR experiments

The ^{1}H-NMR spectra of extracts were recorded at 300 K by a Bruker AVANCE III 600.13 MHz spectrometer equipped with 5 mm triple-resonance z-gradient cryoprobe (Prodigy TCI 1H-13C/15N-2H). TopSpin version 3.6.2 (Bruker GmbH Karlsruhe Germany) was used for spectrometer control purposes. Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence with water presaturation and spoil gradients (*cpmgpr1d* pulse sequence) for direct analysis (64 k data points, spectral width 12,019 Hz, dummy scans 8, 64 scans, loop for T2 filter 80, gain 80,6, delay time 4 s, fixed echo time 0.0007 s and the 90° pulse length was adjusted to about 10.40 μs). ^{1}H 1D Nuclear Overhauser Effect Spectroscopy (NOESY) NMR spectra with water presaturation and spoil gradients (*noesygppr1d* pulse sequence) was used for analysis of deproteinization samples Spectra were acquired with (Thomsen et al., 2011) scans, 64 k data points, spectral width of 7,211 Hz, and relaxation delay of 20 s (dummy scans 4, gain 203, and the 90° pulse length was adjusted to about 10.42 μs)

Total Correlation Spectroscopy (TOCSY) and multiplicity Heteronuclear Single Quantum Correlation (HSQC) were performed on representative samples with 256–512 t1 increments 32–96 transients and a relaxation delay of 1.5 s. The TOCSY spectra were recorded by a standard MLEV-17 pulse sequence with mixing times (spin-lock) of 65 ms.

Data analysis and statistics

NMR spectra processing: ^{1}H-NMR spectra were transformed with a 0.5 line-broadening and manually baseline- and phase-corrected with Topspin 3.6.4. NMR signals of DSS-d6 (for direct analysis) or TSP-d_4 (for deproteinization analysis) were referenced to 0.0 ppm. For metabolite identification purposes the ^{1}H and chemical shift values and multiplicity of signals were compared with the reference data from the Chenomx software (Chenomx NMR Suite 8.4 Chenomx Inc. Edmonton Canada) in combination with spectral databases Human Metabolome Database, and the Biological Magnetic Resonance Bank and several literature reports (Ulrich et al., 2007; Wishart et al., 2018). Optimal integration regions were defined for each metabolite to select signals without overlapping. Integration was performed with MestreNova 14 (Mestrelab Research SL Santiago de Compostela Spain) by manually integrating of the previously identified signals. With these regions an integration matrix (Integral Regions) was built which was later applied to the 88 acquired spectra and a matrix of integrals was built for all the spectra (Integral series). This matrix of integrals was normalized by the sum of the total signals of the spectrum using Excel (Microsoft United States of America).

Multivariate and univariate analysis of Metabolomic Profiles: The previously normalized matrix of integrals was processed using MetaboAnalyst 5.0. First a principal component analysis (PCA) was performed which allowed finding groups of samples with a similar metabolic pattern and/or segmenting those with a different metabolome.

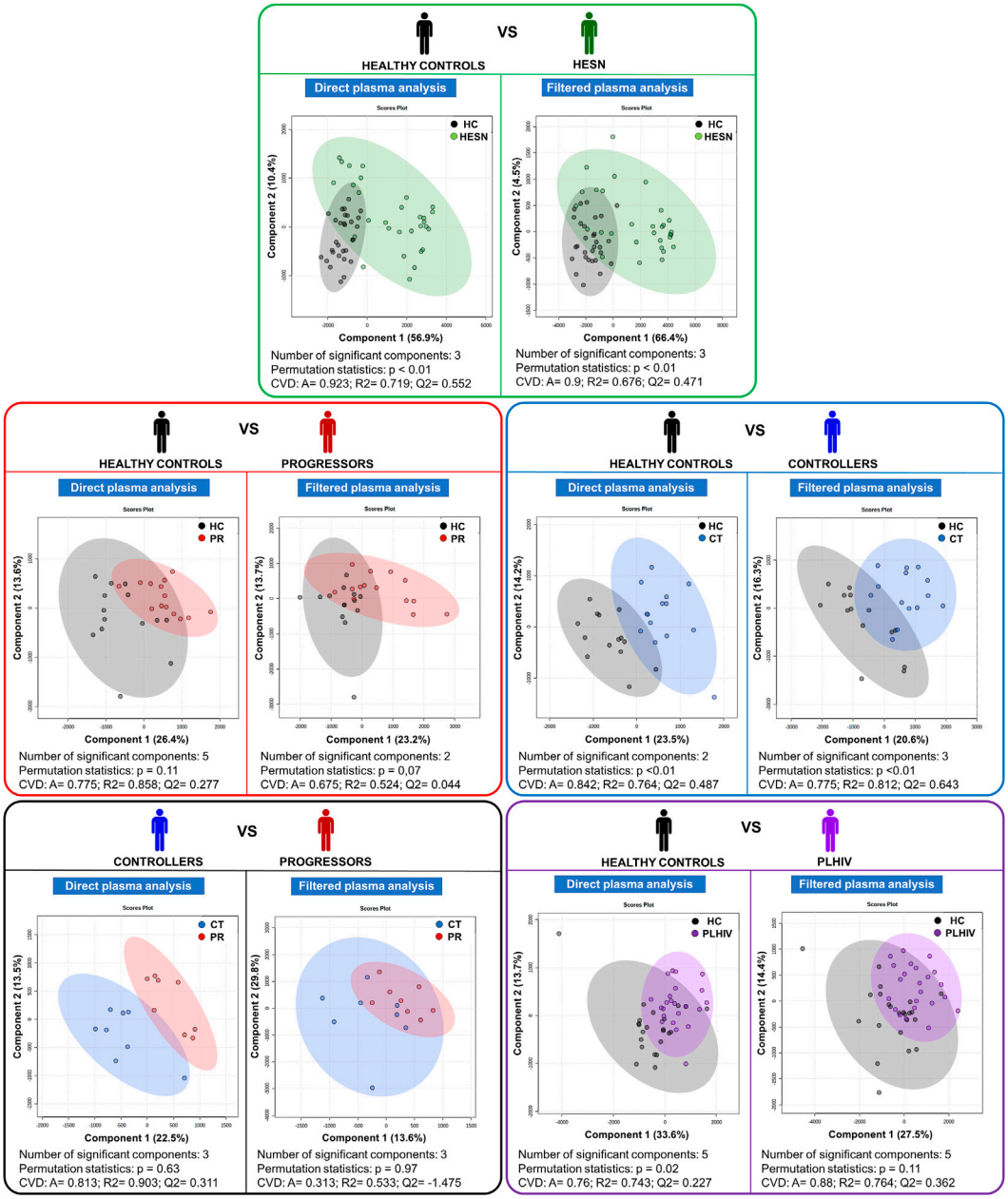


FIGURE 3
Partial Least Squares Discriminant Analysis (PLS-DA). Score plot charts of the PLS-DA of the five case-control relationships established; on the left side, the analysis with direct plasma and on the right side with filtered plasma (in each colored box). HC: Healthy control, CT: controller, HESN: HIV-exposed seronegative, PR: Progressor. CVD: Cross-validation details, A: Accuracy, R2: goodness of fit, Q2: predictive capacity. Healthy controls in black, HESN in green, Progressors in red, Controllers in blue and PLHIV in purple.

Then five case-control relationships were established: healthy controls *versus* progressors; healthy controls *versus* controller; healthy controls *versus* PLHIV (progressors and controllers); healthy controls *versus* HESN, and controllers *versus* progressors. These relationships were evaluated by Partial Least Squares Discriminant Analysis (PLS-DA) which links two data matrices and improves the separation between different groups of samples. The quality of the PLS-DA was evaluated, and a permutation test was carried out to calculate the goodness of fit (R2) and the predictive capacity (Q2) of the randomly generated models. An analysis of the results of the PLS-DA statistic (VIP scores) was performed and the metabolites that contributed significantly to the separation of the groups were identified. Variables with a VIP score greater than 1.0 were considered significant for the model.

Finally, the selected variables were subjected to a Univariate analysis using a difference of means test (Wilcoxon Test). For tests with a *p*-value less than 0.05 ($p < 0.05$) a statistically significant difference between the means (mean or median as appropriate) was assumed for the variable evaluated. In the case of obtaining more

10.3389/fmolb.2023.1204273

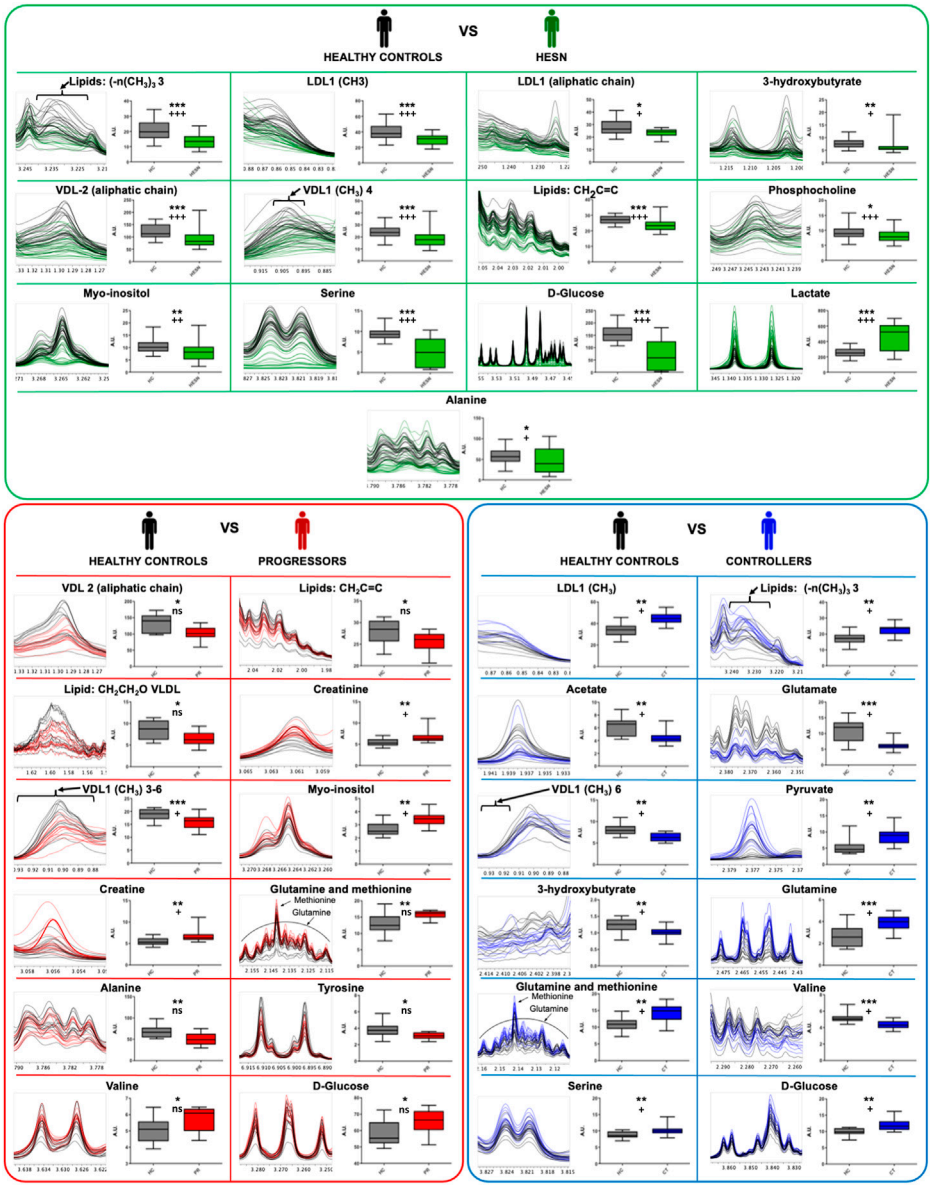


FIGURE 4
Metabolites that explain the difference between the groups. Results of the Wilcoxon test for the significant variables (metabolites) for each proposed PLS-DA model. In the panel **(A)** (main green box), the comparison of healthy controls and HESN, in the panel **(B)** (main red box) the comparison between healthy controls and progressives, and in the panel **(C)** (main blue box) the comparison between healthy controls and controllers. Within each small individual box: on the left the region of the spectrum (metabolite signal) 1H-NMR superimposed of all the samples analyzed in each comparison, on the right side the box-and-whisker plot for the normalized concentration and the statistical significance of each test. *: p-value < 0.05, **: p-value < 0.01 and ***: p-value < 0.001. +: False Discovery Rate (FDR) < 0.05, + +: FDR <0.01, + + +: FDR <0.001 and ns: FDR ˚0.05. Healthy controls in black, HESN in green, progressors in red, and the controllers in blue.

than one significant signal for a given metabolite, we selected the signal with less overlapping for graphical representation.

Gene analysis

Analysis of associated genes was carried out with the metabolites that were statistically significant after univariate analysis. For this, the web interface of MetaboAnalyst 5.0 [14] was used. Metabolites were introduced in the *Network Explorer* section of the platform and the *Metabolite-Gene-Disease Interaction Network analysis* was carried out, which provides a global view of potential functional relationships between metabolites, connected genes, and target diseases. The network integrates gene-metabolite, metabolite-disease, and gene-disease interaction networks.

The genes identified through the previous analysis were filtered through a comparison process with the National Center for Biotechnology Information (NCBI) gene database of the U.S. National Library of Medicine, excluding genes unrelated to HIV-

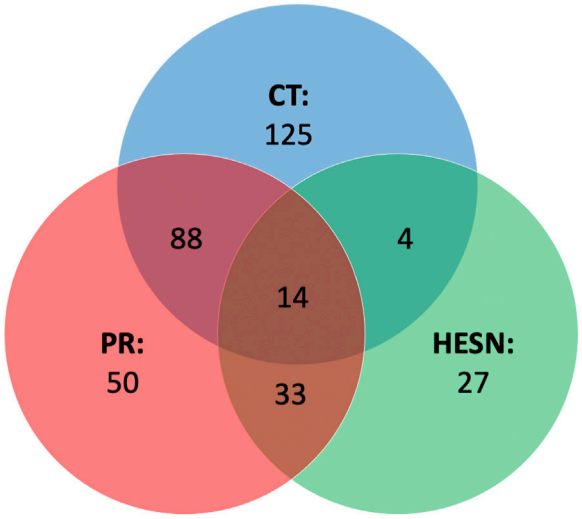


FIGURE 5
Venn diagram of metabolites that explain differences among groups. Set Analysis showing the metabolites related to differences similarities among the metabolomic profiles of controllers, progressors and HESN. In the main part of each set the metabolites related only to one study group and in the intercepts the metabolites related to two or more groups (central intercept). CT: controllers (In blue color set), HESN: HIV-exposed seronegative (In green color set), PR: Progressors (In red color set). HC: Healthy controls, 3-HB: 3-Hydroxybutyrate, ACE: Acetate, ALA: Alanine, CRE: Creatine, Cre: Creatinine, Glc: D-Glucose, GLN: Glutamine, GLU: Glutamate, LAC: Lactate, LDL1 (CH3): Low density lipoprotein (CH3), LDL (AC): Low density lipoprotein aliphatic chain, MET: Methionine, MYO: Myo-inositol, PHO: Phosphocholine, PYR: Pyruvate, SER: Serine, TYR: Tyrosine, UNK: Unknown, VAL: Valine, VDL (AC): VDL-2 (aliphatic chain), VDL1 (CH3): Very low-density lipoprotein (CH3), (-n (CH3)3: Lipids (-n (CH3)3, CH2CH2O VLDL: Lipid: CH2CH2O VLDL, CH2-C=C: Lipids CH2C = C.

1. Additionally, to perform a more specific analysis, genes that were not related to two or more groups and/or metabolites were excluded.

Then, we proceeded to perform an individual analysis of the selected genes through a review in the NCBI gene database. Looking specifically at the section on HIV-1 interactions, we filter further into the subcategories *Replication interactions* (human proteins shown to be required for HIV-1 infectivity and replication) and *Protein interactions* (proteins that have been shown to interact with proteins from HIV-1).

Results

Human subjects

Figure 1 shows the main characteristics of the five case-control relationships analyzed in the study. Out of the 88 available volunteers, a selection was made to generate groups balanced by gender and age. A summarized table of all individuals can be found in Supplementary Table S1

To compare HESN with healthy individuals, we achieved a completely gender-balanced comparison, with 15 women and 15 men in each group. Also, in the case of controllers, the number of women and men was similar in both groups. In the case of joining controllers and progressors (PLHIV), we built up groups with a higher number of men than women, but that was still gender-matched between patients and controls.

Analysis of metabolic plasma profiles

^{1}H NMR metabolomics analysis was performed on intact and filtered plasma samples, to identify the highest possible number of compounds. During the spectral analysis process, it was possible to identify fifty-five metabolites in the direct plasma samples: two alcohols, twenty-one amino acids, fourteen lipid-related signals, sixteen organic acids, one purine derivative, and one sugar. In the filtered plasma samples, forty-five metabolites were identified: two alcohols, twenty amino acids, two lipid-related signals, eighteen organic acids, two derived from purine and one from sugar. It should be noted that three metabolites that were not observed in direct plasma could be detected in filtered plasma samples: two organic acids (2-hydroxybutyrate and 3-Hydroxyisovalerate) and one purine derivative (Inosine). In Supplementary Figures S1, S2 a model NMR spectrum can be seen with the relative assignment for direct and filtered Plasma respectively. Furthermore, the quantification of small metabolites was more accurate in the filtered samples due to the absence of overlapping with broad lipoprotein samples. It should be noted that the lipoproteins evaluated in the unfiltered samples complement the metabolomic analysis of the plasma samples from the volunteers. To know the details of the metabolites identified in the plasma samples, see Supplementary Table S2.

After assignment, normalized integration tables were obtained of all spectra (Supplementary Table S3) and analyzed by multivariate analysis.

Multivariate analysis of plasma metabolomic profiles

Initially, a principal component analysis (PCA) was performed with all 88 samples to get a general overview. The result is shown in Figure 2, which corresponds to the PCA score plots of direct plasma samples and filtered plasma samples. In this unsupervised analysis, we detected a clustering between samples belonging to the same group (healthy control, controller, HESN or progressor). This grouping was more evident for direct plasma samples, which could indicate that lipoproteins play a key role in the differentiation of the groups. HESN samples seem to have the highest dispersion. PCAs were also performed for the five established case-control relationships, which can be seen in Supplementary Figure S3.

In addition, to observe specific metabolic differences between our five case-control relationships, a pair-wise Partial Least Squares Discriminant Analysis (PLS-DA) was performed (see Figure 3).

Statistically significant PLS-DA models were obtained comparing HESN and controllers with healthy controls. Models between healthy controls and progressors and healthy controls *versus* PLHIV were statistically significant. The analysis of the

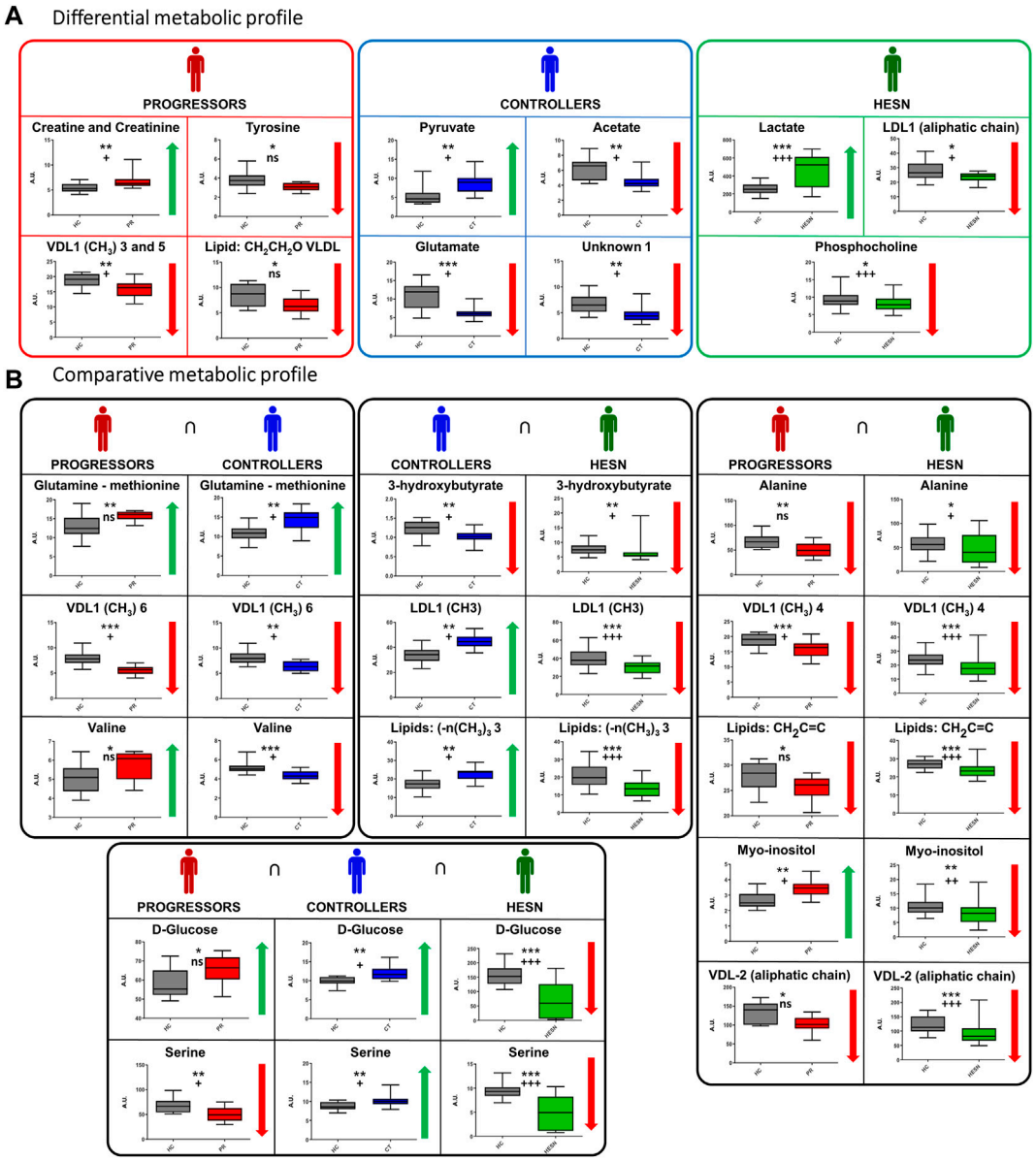


FIGURE 6
Specific and related metabolites in the study groups. **(A)** Differential metabolic profile: Metabolites that only exhibit a statistically significant mean difference compared to healthy controls in a study group, **(B)** Comparative metabolic profile: Metabolites exhibiting a statistically significant mean difference compared to healthy controls in two or more study groups. Variation: ↓ (Smaller area or relative concentration in the reference method) ↑ (Bigger area or relative concentration in the reference method). The box-and-whisker plot for the normalized concentration and the statistical significance of each test. *: p-value < 0.05, **: p-value < 0.01 and ***: p-value < 0.001. +: False Discovery Rate (FDR) < 0.05, + +: FDR <0.01, + + +: FDR <0.001 and ns: FDR ˃0.05. Healthy controls in gray, HESN in green, progressors in red, and the controllers in blue.

PLS-DA statistics through the VIP score allowed reducing the number of relevant variables (metabolites) to be analyzed as follows:

- Healthy controls *versus* HESN: 27 in direct plasma (DP) and 26 in filtered plasma (FP).
- Healthy controls *versus* progressors: 47 in DP and 33 in FP.
- Healthy controls *versus* controllers: 35 in DP and 38 in FP.
- Healthy controls *versus* PLHIV (progressors and controllers): 39 in DP and 36 in FP.
- Controllers *versus* progressors: 45 in DP and 39 in FP.

For the details of the PLS-DA statistics see Supplementary Table S4.

Univariate analysis of plasma metabolomic profiles

The metabolites relevant for PLS discriminant modeling were further submitted to univariate statistical analysis to identify significant changes in each case-control comparison. The detail

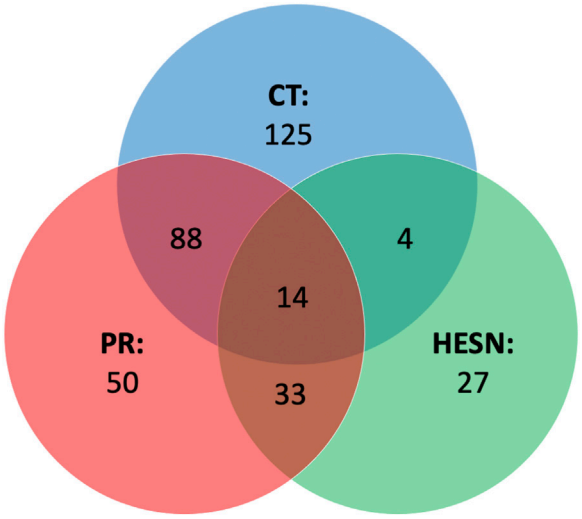


FIGURE 7
Venn diagram of HIV-related genes. Genes identified by Metabolite-Gene-Disease Interaction Network analysis and filtered through the NCBI gene database as related to HIV-1. In the main part of each set, the genes are related only to one study group and in the intercepts, the genes are related to two or more groups (central intercept). CT: controllers (In blue color set), HESN: HIV-exposed seronegative (In green color set), PR: Progressors (In red color set).

of the mean difference analysis (Wilcoxon Test) can be seen in Supplementary Table S5.

For the Healthy controls *versus* HESN comparison, 21 signals (variables) were identified with a significant variation, 13 associated with the FP. These signals are associated with 13 metabolites: six lipid-related signals (Low density lipoprotein aliphatic chain, Low density lipoprotein (CH3), very low-density lipoprotein (CH3), Lipids CH2C = C, lipids (-n (CH3)3 and VDL (aliphatic chain)), three amino acids (Serine, alanine and phosphocholine), two organic acids (3-Hydroxybutyrate and Lactate), one alcohol (Myo inositol) and one sugar (Glucose). Only the lactate was increased in HESN, the rest of the metabolites were decreased compared to healthy controls.

In healthy controls *versus* progressors comparison, 28 signals were identified, 19 associated with DP. These signals correspond to 14 metabolites: eight amino acids (Creatine, creatinine, Glutamine, Methionine, Serine, Alanine, Tyrosine and Valine), four lipid-related signals (Lipid: CH2CH2O, VLDL, Lipids CH2C = C, VDL-2 (aliphatic chain) and very low-density lipoprotein (CH3)), one alcohol (Myo inositol) and one sugar (Glucose). The metabolites of the lipid-related signals, Serine, Alanine and Tyrosine, are decreased.

While for the comparison healthy controls *versus* controllers 24 signals were identified (equal amount of each matrix), which were associated with 12 metabolites: five amino acids (Glutamine, Glutamate, Methionine, Serine and Valine), three lipid-related signals (Low density lipoprotein (CH3), very low-density lipoprotein (CH3) and lipids (-n (CH3)3), three organic acids (3-Hydroxybutyrate, Acetate and Pyruvate) and one sugar (Glucose).

Figure 4 shows a quantitative comparison of the most representative metabolites that change between healthy controls and HESN, progressor and controller groups.

The comparisons of healthy controls *versus* PLHIV and controllers *versus* progressors can be seen in Supplementary Figure S4.

Discussion

All three study groups show a specific metabolic profile

PCA and the PLS-DA analysis demonstrated significant differences between the controllers, progressors, and HESN study groups *versus* healthy controls. In contrast, only weak models were obtained comparing controllers and progressors, which may be due to the low sample number in this case (n = 8). Differences were related to specific metabolites present in different concentrations between the groups as demonstrated by univariate analysis (Wilcoxon Test). For each case-control comparison, a list of metabolites with altered levels were established. Figure 5 provides an overview of all the important metabolites in the different comparisons.

We observed that the disease altered the metabolomics blood profile of progressors, showing the highest number of relevant variables compared to healthy controls in the PLS-DA (47 in DP and 33 in FP), and the highest number of differentiated metabolites (Ding et al., 2021) in the univariate analysis. The impact on controllers was lower (35 variables in DP and 38 in FP, 12 relevant metabolites), while HESNs are the group with the lowest number of relevant variables in the PLS-DA (27 in DP and 26 in FP) and 13 differentiated metabolites. The progression of HIV induces the massive elimination of $CD4^+$ T lymphocytes and alterations in various components of the immune system (Brenchley et al., 2006), which would explain the difference in the metabolic profiles of the progressors compared to controllers and HESN. Since the controllers resist the progression to AIDS, maintaining low levels of viral load and an adequate count of $CD4^+$ T lymphocytes (Cao et al., 1995; Baker et al., 2009), it is directly reflected in a lesser impairment of their metabolism. Likewise, identifying of a differential metabolomic profile in HESNs allows us to affirm that the natural resistance of the host to HIV-1 is associated with a differential phenotype. Detailed analysis of the metabolomics changes (positive or negative variation) between the groups, as shown in Figure 6, allowed establishing differential (Figure 6A) and comparative profiles (Figure 6B) for each of the case-control comparisons.

Alterations associated to progressors: Progressors stand out for the variation of specific lipoproteins and creatine/creatinine and tyrosine, while changes in glutamate, pyruvate and acetate seem to be specific for controllers. On the other hand, alterations in lactate, phosphocholine and VDL are characteristic of HESN.

Concerning the comparative analysis (Figure 6B), it is worth mentioning that glucose and serine change in all three case-control comparisons, although these changes do not always have the same sign. It is also noteworthy that progressors and HESN have five metabolic changes in common, four of which have the same sign.

10.3389/fmolb.2023.1204273

TABLE 1 HIV-related genes that were associated with two or more study groups and/or metabolites.

Gene	PR	CT	HESN
GLUD2	ALA↓ GLN↑	GLU↓ GLN↑	ALA↓
DNAJB1 PABPN1 FOXP2	ALA↓ GLN↑	GLN↑	ALA↓
LDHA LDHB LDHC	CRE↑ Cre↑ ALA↓	GLU↓ PYR↑	ALA↓ LAC↑
MDH2	ALA↓	GLU↓ PYR↑	ALA↓
PKLR PKM	ALA↓ CRE↑	PYR↑	ALA↓ LAC↑
SLC38A2 SLC38A1	ALA↓ GLN↑ MET↑		ALA↓
GPT2	ALA↓ Cre↑	GLU↓ PYR↑	ALA↓
SLC16A10	ALA↓ GLN↑ MET↑ TYR↓ VAL↑		ALA↓
ARG2	CRE↑ VAL↑	GLU↓ VAL↓	N/A
SERPINC1 CAD CREBBP F13A1 MTOR HTT HIP1 DNAJA1 HSPA1A HSPA4 HSPB1 KCNN3 PML PPP2R2B MAPK8 PSMD2 ATXN2 TAF4 TGM4 TGM1 TGM3 UBA52 SUM O 1 UFD1 VCP VEGFA NCOA3 HAP1 TGM5 HDAC6 DNAJB6 STUB1 BAIAP2 TARDBP UBQLN2 ASRGL1 RBM17 TGM7 TMEM37 TGM6	GLN↑		
BDNF CREB1 GAPDH GART GRIN2B IMPDH2 JUN MAPT MSN PFAS ALDH18A1 QARS1 SPTBN TGM2 NME6	GLN↑	GLU↓ GLN↑	
CASP3	GLN↑ VAL↑	GLU↓ GLN↑ VAL↓	
CAT	CRE↑ MET↑ TYR↓	GLU↓ MET↑	
COMT	VAL↑	GLU↓ VAL↓	
DARS1 EPRS1 GLUL KARS1 RARS1 AIMP2 AIMP1 EEF1E1 LARS1	GLN↑ MET↑	GLU↓ GLN↑ MET↑	
DLD	VAL↑	GLU↓ PYR↑ VAL↓	
F2	Cre↑	GLU↓	
FM O 3 HBB MAT2A MSRA MTR MTRR MYH9 SMUG1	MET↑		
GAD1 SOD1	MET↑	GLU↓ MET↑	
IARS1	GLN↑ MET↑ VAL↑	GLU↓ GLN↑ MET↑ VAL↓	
IGF1 OAT PFKM	CRE↑	GLU↓	
MARS1	Cre↑ GLN↑ MET↑	GLU↓ GLN↑ MET↑	
TH	TYR↓	GLU↓	
MAP3K14	Cre↑ GLN↑	GLN↑	
ALB	Cre↑ TYR↓	N/A	PHO↓
BGLAP	Cre↑		LAC↑
FOXL2 FTL HMOX1 HOXA13 LYZ PIN1 PPIA SRSF1 SLC7A5 KYNU	ALA↓		ALA↓
CD79A CRP	Cre↑		PHO↓
DNMT1 FCN2 GCK B4GALT1 HK2 IFNB1 LGALS3 PYGL SGCB B4GALT2 SIGLEC5 H6PD SIGLEC7 CD207 GBA2 GXYLT1	Glc↑		Glc↓
F3	Cre↑ Glc↑		Glc↓
GYPA	CRE↑ Glc↑		Glc↓
FOS	N/A	GLU↓	LAC↑

(Continued on following page)

TABLE 1 (*Continued*) HIV-related genes that were associated with two or more study groups and/or metabolites.

Gene	PR	CT	HESN
OXCT1		3 HB↓	
PLA2G1B		GLU↓	PHO↓
LDHD		PYR↑	LAC↑
ALPI MPO	Cre↑ TYR↓	N/A	N/A
GAA CRNKL1 PIK3C2A	CRE↑ Cre↑		

CT: controllers, HESN: HIV-exposed seronegative, PR: progressors, N/A: Not Applicable. Variation: ↓ (Smaller area or relative concentration in the reference method) ↑ (Bigger area or relative concentration in the reference method). 3-HB: 3-Hydroxybutyrate, ALA: alanine, CRE: creatine, Cre: Creatinine, Glc: D-Glucose, GLN: glutamine, GLU: glutamate, LAC: lactate, MET: methionine, PHO: phosphocholine, PYR: pyruvate, TYR: tyrosine, VAL: valine.

Differentially, it was possible to identify an increase in the expression of creatine-creatinine in progressors; creatine is found in muscles (Kreider and Stout, 2021). Altered creatine-creatinine values have been previously found in HIV positive patients (Sitole et al., 2019) and are related to the prolonged period of high energy expenditure (Kosmiski, 2011) and cachexia (Von Roenn et al., 1992).

Likewise, tyrosine was decreased in the HIV progressors group. Tyrosine is a precursor of catecholamines (adrenaline, dopamine and noradrenaline) whose altered metabolism is related to mood disorders (Hasler et al., 2008). An increased phenylalanine/tyrosine ratio is common in patients with HIV-1 infection and is related to immune activation (Zangerle et al., 2010). Previous NMR studies identified tyrosine downregulation in untreated HIV-infected patients (Sitole et al., 2019). This change was not observed in HIV controllers, which allows us to state that low tyrosine levels are a biomarker of HIV infection progression.

One factor associated with HIV progression is the response of immune cells (Shi et al., 2022). Immune cells undergo energetic and structural remodeling following immune activation. It generates metabolic changes associated with increased energy and biosynthetic demands as viral load increases and the immune system responds (González Plaza et al., 2016). This metabolic changes, including lipid homeostasis, since mitochondria plays a key role in the biosynthesis of phospholipids for membranes, as well as in the catabolism of fatty acids (Tilokani et al., 2018). Unsurprisingly, phospholipid alterations are a common finding in the metabolic profiles of HIV-infected individuals.

A comparison of the commonly altered metabolites in controllers and progressors showed that valine is differentially regulated between both groups; it increased in progressors and decreased in controllers compared to healthy controls. A recent study show a significant increase in Valine levels in TEC before the loss of control compared to PEC, therefore, valine was defined as the main differentiating factor between the studied groups (Tarancon-Diez et al., 2019). That is, elevated valine levels could be a potential biomarker for the prediction of virological progression in controllers and progressors.

Alterations associated to controllers: L-glutamic acid and pyruvate are differentially altered in controllers; these metabolites modulate latent HIV reactivation and/or macrophage inflammation *in vitro* (Giron et al., 2021). A previous study demonstrated that glutamic acid was elevated in Persistent Elite Controllers (PEC) compared to Transient HIV Elite Controllers (TEC) (Tarancon-Diez et al., 2019), suggesting that glutamate metabolism is associated with a delay in the recovery time from HIV.

Likewise, in controllers, acetate is differentially decreased. This metabolite is transiently released into the circulation in response to systemic bacterial infection, as a resistance mechanism of the host's adaptive immune system. (Bose et al., 2019) (Balmer et al., 2016) (Vysochan et al., 2017). The virus-associated mechanism may be related to the group of HIV controllers, where downregulation of acetate concentration would slow down the lipogenesis.

Alterations associated to HESN: In all groups (controllers, progressors and HESN) evaluated, changes in the lipid profile were observed. However, HESNs showed variations in response compared to controllers and progressors; all significant changes were downward. It highlights the Low-density lipoproteins (LDL) signals that are decreased in HESN and increased in the other groups. LDLs are considered proinflammatory lipid species (Brennan et al., 2021) and associated with immune activation in HIV-infected persons (Funderburg and Mehta, 2016). Other NMR metabolomics studies support our findings on altered lipid metabolism in HIV-infected people (Hewer et al., 2006; Riddler et al., 2008; Philippeos et al., 2009; Swanson et al., 2009; Rodríguez-Gallego et al., 2018; Sitole et al., 2019). We did not identify previous studies that used NMR metabolomics to characterize HESNs as done in this study; however, a previous study that included 32 HESN individuals demonstrated that immune status secondary to HIV exposure influences the plasma efflux capacity of HDL cholesterol, which is buffered in HESN (Tort et al., 2018).

Likewise, glucose was altered in all groups. It is decreased in HESN and increased in the other groups. Early steps of virus replication are moderately affected by the ability of the target cell to perform glycolysis at the time of infection. Similarly, virion production in cultures containing galactose was reduced by 20%–60% compared to the amount produced in glucose-containing cultures (Hegedus et al., 2014). That suggests that high glucose availability in the body is associated with infection process and virus replication, which do not occur in HESN.

The HESN group showed higher lactate expression compared to healthy controls, lactate has anti-inflammatory effects modeling the production of interleukins and other proinflammatory molecules (Hearps et al., 2017) (Aldunate et al., 2013). This suggests a protective role of lactate in the sexual transmission of HIV.

Phosphocholine was found significantly decreased in the HESN group. Phosphocholine shown to be able to suppress

TABLE 2 List of genes differentially expressed in the study groups and associated with replication and protein interactions with HIV-1.

Description	Progressors	Controllers	HESN	Replication and/or protein interactions with HIV-1
DnaJ heat shock protein family (Hsp40) member B1	ALA↓ - GLN↑	GLN↑	ALA↓	Knockdown of DnaJ inhibits HIV-1 replication in HeLa-derived TZM-bl cells Brass et al., 2008, while an increase in gene expression is relevant for Tat recruitment in HIV-infected cells Dhamija et al., 2015
				Hsp40 protein is required for HIV-1 Nef-mediated enhancement of viral gene expression and replication Kumar and Mitra, (2005), and that members of this family of interferon-inducible proteins should be considered within its anti-HIV function Urano et al., 2013
Pyruvate kinase L/R	ALA↓ - CRE↑	PYR↑	ALA↓ - LAC↑	*PKLR* has shown a regulatory role in HIV replication in HeLa P4/R5 cells (Zhou et al., 2008)
solute carrier family 38 member 2	ALA↓ - GLN↑ - MET↑	GLU↓ - GLN↑ - MET↑	ALA↓	Down-regulation of *SLC38A1* and *SLC38A2* is associated with HIV interference with immunometabolism in activated primary human $CD4^+$ T cells Matheson et al., 2015
solute carrier family 38 member 1	ALA↓ - GLN↑ - MET↑	GLU↓ - GLN↑ - MET↑	ALA↓	This gene uses alanine as an endogenous substrate for T cell mitogenesis Matheson et al., 2015
Glutamic--pyruvic transaminase 2	ALA↓ Cre↑	GLU↓ PYR↑	ALA↓	Knockdown has been shown to inhibit early stages of HIV-1 replication in an *in vitro* model König et al., 2008
Caspase-3	GLN↑ VAL↑	GLU↓ GLN↑ VAL↓	N/A	Is related to HIV-associated dementia (HAD) Yndart et al., 2015 (104)
Coagulation factor II thrombin	Cre↑	GLU↓		Encodes the protein prothrombin Smolkin and Perrotta, (2017). Knockdown of *F2* has previously been suggested to have a regulatory role in HIV replication Zhou et al., 2008
				Thrombin was shown to activate gp120/gp41 of HIV-1, enhances virus-cell fusion Cheng et al., 2010, and enhance the gp160-mediated fusion of HIV-1 with R5 tropism Ling et al., 2004
Glyceraldehyde-3-phosphate dehydrogenase	GLN↑	GLU↓ GLN↑		Negatively regulates HIV-1 infection by directly interacting with *Gag* and *Gag-Pol* Kishimoto et al., 2012
Glutamate ionotropic receptor NMDA type subunit 2B	GLN↑	GLU↓ GLN↑		*GRIN2B* deletion inhibits HIV-1 replication in HeLa P4/R5 cells Zhou et al., 2008, this inhibition is related to HIV-gp120 and Tat upregulating *GRIN2B* Che et al., 2014; Xiong et al., 2014
Lysyl-trna synthetase 1	GLN↑ MET↑	GLU↓ GLN↑ MET↑		Knockdown inhibited the initial stages of HIV-1 replication *in vitro* König et al., 2008
Methionyl-trna synthetase 1	Cre↑ GLN↑ MET↑	GLU↓ GLN↑ MET↑		Knockdown inhibited HIV-1 replication Yeung et al., 2009
Moesin	GLN↑	GLU↓ GLN↑		Knockdown inhibited HIV-1 replication Yeung et al., 2009
Phosphofructokinase muscle	CRE↑	GLU↓		Knockdown inhibited the initial stages of HIV-1 replication *in vitro* König et al., 2008
Mitogen-activated protein kinase 14	Cre↑ GLN↑	GLN↑		MAP kinases (MAPK) have been associated with HIV proteins such as *gp120* Jin et al., 2016, *Nef* Hashimoto et al., 2014, *Tat* Planès et al., 2016 and *Vpr* Hoshino et al., 2010, which generate a strong activation of these enzymes
Aminoacyl trna synthetase complex interacting multifunctional protein 1	GLN↑ MET↑	GLU↓ GLN↑ MET↑		Knockdown inhibited HIV-1 replication Yeung et al., 2009
Leucyl-trna synthetase 1	GLN↑ MET↑	GLU↓ GLN↑ MET↑		Knockdown of DnaJ inhibits HIV-1 replication in HeLa-derived TZM-bl cells Brass et al., 2008
Ficolin 2	Glc↑	N/A	Glc↓	Ficolin-2 binds to HIV-1 *gp120* and blocks viral infection (Luo et al., 2016)
Glucokinase	Glc↑		Glc↓	Knockdown inhibited HIV-1 replication in HeLa-derived TZM-bl cells Brass et al., 2008
Interferon beta 1	Glc↑		Glc↓	Interferon-beta, encoded by *IFNB1* gene, has antiviral, antibacterial and anticancer properties Graber et al., 2014. HIV-1 replication upregulates the expression of *IFNB1* gene Czubala et al., 2016

(Continued on following page)

TABLE 2 (*Continued*) List of genes differentially expressed in the study groups and associated with replication and protein interactions with HIV-1.

Description	Progressors	Controllers	HESN	Replication and/or protein interactions with HIV-1
Galectin 3	Glc↑		Glc↓	In HIV infection, deletion of *LGALS3* by shRNA was shown to inhibit HIV-1 production *in vitro* Wang et al., 2014. Furthermore, it promotes HIV-1 budding through association with Alix and Gag p6 Wang et al., 2014
Sialic acid binding Ig like lectin 5	Glc↑		Glc↓	Siglec-5 associated with divergent outcomes of HIV-1 infection in human and chimpanzee CD4 T cells Soto et al., 2013
Phospholipase A2 group IB	N/A	GLU↓	PHO↓	This gene was involved in CD4 anergy and CD4 lymphopenia in HIV-infected patients Pothlichet et al., 2020
Myeloperoxidase	Cre↑ TYR↓	N/A	N/A	Knockdown of MPO inhibits HIV-1 replication in HeLa P4/R5 cells Zhou et al., 2008
Crooked neck pre-mrna splicing factor 1	CRE↑ Cre↑			CRNKL1 was identified as a highly Selective Regulator of Intron-Retaining HIV-1 and Cellular mRNAs Xiao et al., 2021

CT: controllers, HESN: HIV-exposed seronegative, PR: progressors, N/A: Not Applicable. Variation: ↓ (Smaller area or relative concentration in the reference method) ↑ (Bigger area or relative concentration in the reference method). 3-HB: 3-Hydroxybutyrate, ALA: alanine, CRE: creatine, Cre: Creatinine, Glc: D-Glucose, GLN: glutamine, GLU: glutamate, LAC: lactate, MET: methionine, PHO: phosphocholine, PYR: pyruvate, TYR: tyrosine, VAL: Valine.

immune response in human placenta (Lovell et al., 2007), initiate phagocytic immune recognition (Thompson et al., 1999) and is an intermediate in the synthesis of phosphatidylcholine in tissues (El-Bacha et al., 2016). Phosphatidylcholine has anti-inflammatory effects (Treede et al., 2007/07). The low concentration of phosphocholine in HESN could be related to the production of phosphatidylcholine that would reduce the inflammation of the colon and rectum that occurs in anal intercourse, a common means of HIV exposure in HESN.

It should be noted that cell activation and inflammation have been reported to enhance infection. (Masson et al., 2015; Liebenberg et al., 2017; Wall et al., 1994). The metabolites identified in HESN, and the metabolic and signaling pathways associated with these metabolites, may contribute to the reduction of inflammation and cell activation. Inflammation increases the risk of contracting HIV by causing the activation of HIV target cells ($CD4^+$ T cells), increasing their susceptibility to HIV infection (Koning et al., 2005). Inflammation also leads to increased recruitment of these activated target $CD4^+$ T cells at the site of HIV exposure (Arnold et al., 2016).

A specific comparison of progressors and HESN revealed that Myo-inositol was elevated in progressors and decreased in HESN. Myo-inositol is a marker of glial reactivity, gliosis and neuroinflammation (Bertran-Cobo et al., 2022). Previously, it was found to be elevated in different brain regions by Magnetic Resonance Spectroscopy (MRS) studies of HIV-infected patients with cognitive impairment (Chang et al., 2004; Cohen et al., 2010; Cassol et al., 2013; Cysique et al., 2013). A previous NMR study on CSF found that impairments in late recall and motor function were associated with higher levels of myo-inositol (Dickens et al., 2015). Alterations in myo-inositol levels were also identified in human mouthwashes (Ghannoum et al., 2013) and brain tissue of HIV-infected rodents (Epstein et al., 2013). All these findings, including those presented in our research, suggest that elevated levels of myo-inositol promote HIV progression in infected people and resistance to infection in HESN.

HIV-related genes associated with significant metabolites

In addition to the altered metabolic pathways, we also wanted to study the genes that were related to these pathways, whose expression could be altered. An analysis of the genes associated to HIV-1 that were related to the metabolites responsible for the difference between the study groups and the healthy controls (See Supplementary Table S6), revealed that patients living with HIV (Controllers and progressors) had the highest number of genes involved in their infectious status (231 and 185 genes *versus* 78 for HESN) (Figure 7). affected by the expression or differential regulation of the metabolites. A total of 341 genes.

We further obtained a more reduce list of genes by filtering only those found in two or more study groups or related to two or more metabolite (Table 1). From these, 21 genes were specifically related to replications interactions (Table 2). The table also summarizes how the alteration of these genes has been previously related to HIV.

Among the genes listed in Table 2, five are related to the three study groups (progressors, controllers and HESN): *DnaJ heat shock protein family (Hsp40) member B1,* Pyruvate kinase liver and red blood cell (*PKLR*) gene, Solute carrier gene family *SLC38A1* and *SLC38A2,* and glutamic-pyruvic transaminase 2 (*GPT2*).

High pyruvate expression in controllers may be associated with elevated PKLR function, whereas in HESN, high lactate expression that is associated with reduced pyruvate would explain a differential gene response in these two groups. *SLC38A1* and *SLC38A2* uses alanine as an endogenous substrate for T cell mitogenesis (Matheson et al., 2015). This metabolite is markedly down regulated in HESN and progressors, but not in controllers, which instead have a low concentration of GLU. In contrast to the HESN, progressors and controllers increase GLN and MET metabolites.

GPT2 encodes a mitochondrial alanine transaminase, a pyridoxal enzyme that catalyzes the reversible transamination between alanine and 2-oxoglutarate to generate pyruvate and

glutamate (Qing et al., 2016). The differential regulations of pyruvate and glutamate in Controllers may be related to a differential expression of this gene and low rates of virus replication.

On the other hand, when the association between HIV progressors and controllers was analyzed, eleven genes (*CASP3* gene, coagulation factor II thrombin gene (*F2*), Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), Glutamate ionotropic receptor NMDA type subunit 2B (*GRIN2B*), Mitogen-activated protein kinase 14 (*MAP3K14*) gene, Lysyl-trna synthetase 1 (*KARS*), Phosphofructokinase muscle (*PFKM*), Methionyl-tRNA synthetase 1 (*MARS*) Moesin (*MSN*), Aminoacyl tRNA synthetase complex interacting multifunctional protein 1 (*AIMP1*)) were identified expressed in both groups, but related to different metabolites or opposite metabolite concentration changes (See Table 2).

Between progressors and controllers there are two differences in terms of the associated metabolites: the expression of valine (elevated in progressors and decreased in controllers) and the expression of glutamate that is exclusively decreased in controllers. For this reason, glutamate is then the most important metabolite in the difference between progressors and controllers. Glutamate causes neuronal cell death by apoptosis at high concentrations (Froissard and Duval, 1994; Behl et al., 1995) and glutamate-induced apoptotic cell death was associated with caspase-3 gene regulation (Zhang and Bhavnani, 2006). It could then be stated that low concentrations of glutamate in Controllers may be related to a neuroprotective profile and regulation of apoptosis in HIV infection.

The relationship between prothrombin and/or thrombin with glutamate has been previously determined (Hoogland et al., 2005; Chinnaraj et al., 2018). Thus, the differential expression of glutamate in controllers may be related to an optimized cell proliferation mediated by *F2* that participates in resistance to HIV progression.

There is a relationship between glutamate and *GAPDH* shown previously (Ikemoto et al., 2003), so, it can be inferred that the differential levels observed in controllers can promote the downregulation of the infection.

Creatinine was only found to be altered in progressors. It has been demonstrated that dietary supplementation with creatinine generates a decrease in MAPK expression (Alves et al., 2012); this could occur in progressors with high creatinine levels. MAP kinases (MAPK) are involved in cellular processes such as development, proliferation, differentiation, and transcription regulation (Plotnikov et al., 2011).

Furthermore, in our study we were able to identify five genes that were associated with progressors and HESN: *Ficolin 2* (*FCN2*), *Glucokinase* (*GCK*), *Interferon-beta 1* (*IFNB1*), *Galectin 3* (*LGALS3*), and *Sialic acid binding Ig like lectin 5* (*SIGLEC5*) (See Table 2). All five genes are associated with D-Glucose regulation in study groups. In the progressors, D-Glucose is upregulated and downregulated in the HESNs group. Glucokinase is a type IV isozyme found exclusively in the liver. It is highly specific, only uses D-glucose as a substrate (Sanchez Caballero et al., 2021); it is encoded by the GCK gene, and its knockdown inhibited HIV-1 replication in HeLa-derived TZM-bl cells (Brass et al., 2008). The low concentration of D-Glucose in HESNs may be related to this phenomenon.

It was possible to identify a gene that is related to controllers and HESN, phospholipase A2 group IB gene (*PLA2G1B*). Recently, this gene was involved in CD4 anergy and CD4 lymphopenia in HIV-infected patients (Pothlichet et al., 2020). This gene could potentially be differentially expressed between controllers and HESN. It is enough to identify that HESN present a decreased concentration of PHO, a metabolite directly related to *PLA2G1B*.

Finally, within the replication interactions, two progressor-specific genes associated with increased expression of CRE and Cre were identified: Myeloperoxidase (MPO) and Crooked neck pre-mRNA splicing factor 1 (CRNKL1).

Conclusion

This study presents differential metabolic profile for controllers, progressors and HESN individuals. Thus, the resistance to HIV-1 progression is associated with changes in the individual's metabolome, represented in the form of metabolites, that could provide biomarkers of the infectious status of PLHIV, and it will be key to determine the factors that control the infection. Based on our results, we propose tyrosine, glutamate, and valine as biomarkers of progression in HIV infection in progressors and controllers.

It should be noted that the variation in the viral load during progression could affect the metabolomic profile of these individuals. Therefore, conducting a longitudinal study with this population help to resolve this issue. However, according to the HIV "test and treat" guidelines, it is challenging to recruit HIV-positive individuals without receiving ART; therefore, evaluations spanning continuous years of suppressive or not suppressive ART compared with our cohort can help to elucidate the impact of the ART in the metabolomic profile.

Likewise, our study visualized that natural resistance to HIV-1 infection in HESN individuals is associated with a specific metabolic fingerprint, described here for the first time according to our research. We consider LDL, glucose, lactate and Phosphocholine plausible biomarkers of natural resistance to HIV infection in HESN. However, additional studies should be carried out with other HESN groups: female sex workers (FSWs), children born to HIV infected mothers and men who have sex with men (MSM). These analyses allow us to compare and contrast our results, to determine if the metabolites are repeated in the different groups or if there are other phenotypes associated with HIV resistance.

The specific biomarkers for each group were associated with genes and proteins related to HIV-1, therefore, differential expression among groups could potentially explain the characteristics of each group. Finally, we consider that proteomic, transcriptomic and genomic analyzes should be carried out to have a more comprehensive look at the progression and the natural resistance to HIV-1 infection.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by the Bioethical Committee Universidad de Antioquia; and all the individuals signed informed consent prepared according to Colombian Legislation Resolution 008430/1993. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization: LG-A, MP-S, EG. Data curation: LG-A, MP-S, and EG. Formal analysis: LG-A and MP-S. Funding acquisition: MP-S, WZ-B, and EG. Investigation: LG-A. Methodology: LG-A, MP-S, and EG. Project administration: EG. Resources: WZ-B, EG, and MR. Supervision: EG. Validation: LG-A and MP-S. Visualization: LG-A and MP-S. Writing original draft preparation: LG-A, MP-S, WZ-B, and EG. Writing review and editing: LG-A, MP-S, WZ-B, and MR. All authors contributed to the article and approved the submitted version.

Funding

The authors acknowledge the Comité para el Desarrollo de la Investigación (CODI) of the University of Antioquia https://www.udea.edu.co for the financial support of this work. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1204273/full#supplementary-material

References

Aguilar-Jiménez, W., Zapata, W., Caruz, A., and Rugeles, M. T. (2013). High transcript levels of vitamin D receptor are correlated with higher mRNA expression of human beta defensins and IL-10 in mucosa of HIV-1-Exposed seronegative individuals. *PLoS One [Internet* 8 (12), e82717. Available from. doi:10.1371/journal.pone.0082717

Aldunate, M., Tyssen, D., Johnson, A., Zakir, T., Sonza, S., Moench, T., et al. (2013). Vaginal concentrations of lactic acid potently inactivate HIV. *J. Antimicrob. Chemother. [Internet]* 68 (9), 2015–2025. –25. Available from. doi:10.1093/jac/dkt156

Alves, C. R. R., Ferreira, J. C., de Siqueira-Filho, M. A., Carvalho, C. R., Lancha, A. H., and Gualano, B. Creatine-induced glucose uptake in type 2 diabetes: A role for AMPK-α? Amino acids *Amino Acids*, [Internet]. 2012;43(4):1803–1807. doi:10.1007/s00726-012-1246-6

Arnold, K. B., Burgener, A., Birse, K., Romas, L., Dunphy, L. J., Shahabi, K., et al. (2016). Increased levels of inflammatory cytokines in the female reproductive tract are associated with altered expression of proteases, mucosal barrier proteins, and an influx of HIV-susceptible target cells. *Mucosal Immunol.* 9 (1), 194–205. doi:10.1038/mi.2015.51

Baker, B. M., Block, B. L., Rothchild, A. C., and Walker, B. D. (2009). Elite control of HIV infection: implications for vaccine design. *Expert Opin. Biol. Ther.* 9 (1), 55–69. doi:10.1517/14712590802571928

Balmer, M. L., Ma, E. H., Bantug, G. R., Grählert, J., Pfister, S., Glatter, T., et al. (2016). Memory CD8(+) T cells require increased concentrations of acetate induced by stress for optimal function. *Immunity* 44 (6), 1312–1324. doi:10.1016/j.immuni.2016.03.016

Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., et al. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220 (4599), 868–871. doi:10.1126/science.6189183

Behl, C., Widmann, M., Trapp, T., and Holsboer, F. (1995). 17-beta estradiol protects neurons from oxidative stress-induced cell death *in vitro*. *Biochem. Biophys. Res. Commun.* 216 (2), 473–482. doi:10.1006/bbrc.1995.2647

Bertini, I., Cacciatore, S., Jensen, B. V., Schou, J. V., Johansen, J. S., Kruhøffer, M., et al. (2012). Metabolomic NMR fingerprinting to identify and predict survival of patients with metastatic colorectal cancer. *Cancer Res.* 72 (1), 356–364. Available at:. doi:10.1158/0008-5472.CAN-11-1543

Bertran-Cobo, C., Wedderburn, C. J., Robertson, F. C., Subramoney, S., Narr, K. L., Joshi, S. H., et al. (2022), A neurometabolic pattern of elevated myo-inositol in children who are HIV-exposed and uninfected: A South African birth cohort study *Front. Immunol*, [Internet]. Vol. 13, 800273, doi:10.3389/fimmu.2022.800273

Bose, S., Ramesh, V., and Locasale, J. W. Acetate metabolism in physiology, cancer, and beyond. *Trends Cell Biol.* [Internet]. 2019;29(9):695–703. doi:10.1016/j.tcb.2019.05.005

Brass, A. L., Dykxhoorn, D. M., Benita, Y., Yan, N., Engelman, A., Xavier, R. J., et al. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 319 (5865), 921–926. doi:10.1126/science.1152725

Brenchley, J. M., Price, D. A., and Douek, D. C. (2006). HIV disease: Fallout from a mucosal catastrophe? *Nat. Immunol.* 7 (3), 235–239. doi:10.1038/ni1316

Brennan, E., Kantharidis, P., Cooper, M. E., and Godson, C. (2021). Pro-resolving lipid mediators: Regulators of inflammation, metabolism and kidney function. *Nat. Rev. Nephrol. [Internet]* 17 (11), 725–739. Available from. doi:10.1038/s41581-021-00454-y

Cao, Y., Qin, L., Zhang, L., Safrit, J., and Ho, D. D. (1995). Virologic and immunologic characterization of long-term survivors of human immunodeficiency virus type 1 infection. *N. Engl. J. Med.* 332 (4), 201–208. doi:10.1056/NEJM199501263320401

Cassol, E., Misra, V., Holman, A., Kamat, A., Morgello, S., and Gabuzda, D. (2013). Plasma metabolomics identifies lipid abnormalities linked to markers of inflammation, microbial translocation, and hepatic function in HIV patients receiving protease inhibitors. *BMC Infect. Dis.* 13, 203. doi:10.1186/1471-2334-13-203

Chang, L., Lee, P. L., Yiannoutsos, C. T., Ernst, T., Marra, C. M., Richards, T., et al. (2004). A multicenter *in vivo* proton-MRS study of HIV-associated dementia and its relationship to age. *Neuroimage* 23 (4), 1336–1347. doi:10.1016/j.neuroimage.2004.07.067

Che, X., He, F., Deng, Y., Xu, S., Fan, X., Gu, P., et al. (2014). HIV-1 Tat-mediated apoptosis in human blood-retinal barrier-associated cells. *PLoS One* 9 (4), e95420. doi:10.1371/journal.pone.0095420

Cheng, D-C., Zhong, G-C., Su, J-X., Liu, Y-H., Li, Y., Wang, J-Y., et al. (2010). A sensitive HIV-1 envelope induced fusion assay identifies fusion enhancement of thrombin. *Biochem. Biophys. Res. Commun.* 391 (4), 1780–1784. doi:10.1016/j.bbrc.2009.12.155

Chinnaraj, M., Planer, W., and Pozzi, N. (2018). Structure of coagulation factor II: Molecular mechanism of thrombin generation and development of next-generation anticoagulants [internet]. *Front. Med.* 5, 1.

Cohen, R. A., Harezlak, J., Gongvatana, A., Buchthal, S., Schifitto, G., Clark, U., et al. (2010). Cerebral metabolite abnormalities in human immunodeficiency virus are associated with cortical and subcortical volumes. *J. Neurovirol* 16 (6), 435–444. doi:10.3109/13550284.2010.520817

Cysique, L. A., Moffat, K., Moore, D. M., Lane, T. A., Davies, N. W. S., Carr, A., et al. (2013). HIV, vascular and aging injuries in the brain of clinically stable HIV-infected adults: A (1)H MRS study. *PLoS One* 8 (4), e61738. doi:10.1371/journal.pone.0061738

Czubala, M. A., Finsterbusch, K., Ivory, M. O., Mitchell, J. P., Ahmed, Z., Shimauchi, T., et al. (2016). TGFβ induces a SAMHD1-independent post-entry restriction to HIV-1 infection of human epithelial langerhans cells. *J. Invest. Dermatol* 136 (10), 1981–1989. doi:10.1016/j.jid.2016.05.123

Dhamija, N., Choudhary, D., Ladha, J. S., Pillai, B., and Mitra, D. (2015). Tat predominantly associates with host promoter elements in HIV-1-infected T-cells - regulatory basis of transcriptional repression of c-Rel. *FEBS J.* 282 (3), 595–610. doi:10.1111/febs.13168

Dickens, A. M., Anthony, D. C., Deutsch, R., Mielke, M. M., Claridge, T. D. W., Grant, I., et al. (2015). Cerebrospinal fluid metabolomics implicate bioenergetic adaptation as a neural mechanism regulating shifts in cognitive states of HIV-infected patients. *Aids* 29 (5), 559–569. doi:10.1097/QAD.0000000000000580

Ding, J., Liu, Y., and Lai, Y. (2021). Knowledge from london and berlin: Finding threads to a functional HIV cure [internet]. 12, Available at: https://www.frontiersin.org/articles/10.3389/fimmu.2021.688747.

El-Bacha, T., and Torres, A. G. (2016). in *Phospholipids: Physiology*. Editors B. Caballero, P. M Finglas, and F. and H. Toldrá FBT-E of (Oxford: Academic Press), 352–359.

Epstein, A. A., Narayanasamy, P., Dash, P. K., High, R., Bathena, S. P. R., Gorantla, S., et al. (2013). Combinatorial assessments of brain tissue metabolomics and histopathology in rodent models of human immunodeficiency virus infection. *J. Neuroimmune Pharmacol.* 8 (5), 1224–1238. doi:10.1007/s11481-013-9461-9

Fisher, M. J., Dickson, A. J., and Pogson, C. I. (1987). The role of insulin in the modulation of glucagon-dependent control of phenylalanine hydroxylation in isolated liver cells. *Biochem. J. [Internet* 242 (3), 655–660. Available at:. doi:10.1042/bj2420655

Frank, T. D., Carter, A., Jahagirdar, D., Biehl, M. H., Douwes-Schultz, D., Larson, S. L., et al. (2019). Global, regional, and national incidence, prevalence, and mortality of HIV, 1980-2017, and forecasts to 2030, for 195 countries and territories: A systematic analysis for the global burden of diseases, injuries, and risk factors study 2017. *Lancet HIV* 6 (12), e831–e859. Available from. doi:10.1016/S2352-3018(19)30196-1

Froissard, P., and Duval, D. (1994). Cytotoxic effects of glutamic acid on PC12 cells. *Neurochem. Int.* 24 (5), 485–493. doi:10.1016/0197-0186(94)90096-5

Funderburg, N. T., and Mehta, N. N. (2016). Lipid abnormalities and inflammation in HIV inflection. *Curr. HIV/AIDS Rep. [Internet]* 13 (4), 218–225. Available from. doi:10.1007/s11904-016-0321-0

Gallo, R. C., Sarin, P. S., Gelmann, E. P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V. S., et al. (1983). Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science* 220 (4599), 865–867. doi:10.1126/science.6601823

Ghannoum, M. A., Mukherjee, P. K., Jurevic, R. J., Retuerto, M., Brown, R. E., Sikaroodi, M., et al. (2013). Metabolomics reveals differential levels of oral metabolites in HIV-infected patients: Toward novel diagnostic targets. *Omi A J. Integr. Biol.* 17 (1), 5–15. doi:10.1089/omi.2011.0035

Giron, L., Palmer, C., Liu, Q., Yin, X., Papasavvas, E., Sharaf, R., et al. (2021). Non-invasive plasma glycomic and metabolic biomarkers of post-treatment control of HIV. *Nat. Commun.* 12, 3922. doi:10.1038/s41467-021-24077-w

González Plaza, J. J., Hulak, N., Kausova, G., Zhumadilov, Z., and Akilzhanova, A. (2016). Role of metabolism during viral infections, and crosstalk with the innate immune system. *Intractable Rare Dis. Res.* 5 (2), 90–96. doi:10.5582/irdr.2016.01008

Gonzalo-Gil, E., Ikediobi, U., and Sutton, R. E. Mechanisms of virologic control and clinical characteristics of HIV+ elite/viremic controllers. *Yale J. Biol. Med.* [Internet]. 2017;90(2):245–259. Available at: https://pubmed.ncbi.nlm.nih.gov/28656011.

Graber, J. J., and Dhib-Jalbut, S. (2014). Interferons *Daroff RBBT-E of the NS* M. J. Aminoff Second E (Oxford: Academic Press), 718–723.

Hashimoto, M., Nasser, H., Chihara, T., and Suzu, S. (2014). Macropinocytosis and TAK1 mediate anti-inflammatory to pro-inflammatory macrophage differentiation by HIV-1 Nef. *Cell Death Dis.* 5 (5), e1267. doi:10.1038/cddis.2014.233

Hasler, G., Fromm, S., Carlson, P. J., Luckenbaugh, D. A., Waldeck, T., Geraci, M., et al. (2008). Neural response to catecholamine depletion in unmedicated subjects with major depressive disorder in remission and healthy subjects. *Arch. Gen. Psychiatry [Internet* 65 (5), 521–531. Available at:. doi:10.1001/archpsyc.65.5.521

Hearps, A. C., Tyssen, D., Srbinovski, D., Bayigga, L., Diaz, D. J. D., Aldunate, M., et al. (2017). Vaginal lactic acid elicits an anti-inflammatory response from human cervicovaginal epithelial cells and inhibits production of pro-inflammatory mediators associated with HIV acquisition. *Mucosal Immunol. [Internet]* 10 (6), 1480–1490. Available from. doi:10.1038/mi.2017.27

Hegedus, A., Kavanagh Williamson, M., and Huthoff, H. (2014). HIV-1 pathogenicity and virion production are dependent on the metabolic phenotype of activated CD4+ T cells. *Retrovirology* 11, 98. Availableat:. doi:10.1186/s12977-014-0098-4

Hewer, R., Vorster, J., Steffens, F. E., and Meyer, D. (2006). Applying biofluid 1H NMR-based metabonomic techniques to distinguish between HIV-1 positive/AIDS patients on antiretroviral treatment and HIV-1 negative individuals. *J. Pharm. Biomed. Anal.* 41 (4), 1442–1446. doi:10.1016/j.jpba.2006.03.006

Hoogland, G., Bos, I. W. M., Kupper, F., van Willigen, G., Spierenburg, H. A., van Nieuwenhuizen, O., et al. (2005). Thrombin-stimulated glutamate uptake in human platelets is predominantly mediated by the glial glutamate transporter EAAT2. *Neurochem. Int.* 47 (7), 499–506. Available at:. doi:10.1016/j.neuint.2005.06.006

Hoshino, S., Konishi, M., Mori, M., Shimura, M., Nishitani, C., Kuroki, Y., et al. (2010). HIV-1 Vpr induces TLR4/MyD88-mediated IL-6 production and reactivates viral production from latency. *J. Leukoc. Biol.* 87 (6), 1133–1143. doi:10.1189/jlb.0809547

Huang, Y., Paxton, W. A., Wolinsky, S. M., Neumann, A. U., Zhang, L., He, T., et al. (1996). The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat. Med.* 2 (11), 1240–1243. doi:10.1038/nm1196-1240

Ikemoto, A., Bole, D. G., and Ueda, T. (2003). Glycolysis and glutamate accumulation into synaptic vesicles: Role of glyceraldehyde phosphate dehydrogenase and 3-PHOSPHOGLYCERATE kinase. *J. Biol. Chem. [Internet]* 278 (8), 5929–5940. Available from. doi:10.1074/jbc.M211617200

Jin, C., Li, J., Cheng, L., Liu, F., and Wu, N. (2016). Gp120 binding with DC-SIGN induces reactivation of HIV-1 provirus via the NF-κB signaling pathway. *Acta Biochim. Biophys. Sin. (Shanghai).* 48 (3), 275–281. doi:10.1093/abbs/gmv138

Kishimoto, N., Onitsuka, A., Kido, K., Takamune, N., Shoji, S., and Misumi, S. (2012). Glyceraldehyde 3-phosphate dehydrogenase negatively regulates human immunodeficiency virus type 1 infection. *Retrovirology* 9, 107. doi:10.1186/1742-4690-9-107

König, R., Zhou, Y., Elleder, D., Diamond, T. L., Bonamy, G. M. C., Irelan, J. T., et al. (2008). Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 135 (1), 49–60. doi:10.1016/j.cell.2008.07.032

Koning, F. A., Otto, S. A., Hazenberg, M. D., Dekker, L., Prins, M., Miedema, F., et al. (2005). Low-level CD4+ T cell activation is associated with low susceptibility to HIV-1 infection. *J. Immunol.* 175 (9), 6117–6122. doi:10.4049/jimmunol.175.9.6117

Kosmiski, L. Energy expenditure in HIV infection. *Am. J. Clin. Nutr.* [Internet]. 2011;94(6):1677S–1682S. doi:10.3945/ajcn.111.012625

Kreider, R. B., and Stout, J. R. (2021). Creatine in health and disease. *Nutrients* 13, 447. doi:10.3390/nu13020447

Kumar, M., and Mitra, D. (2005). Heat shock protein 40 is necessary for human immunodeficiency virus-1 Nef-mediated enhancement of viral gene expression and replication. *J. Biol. Chem.* 280 (48), 40041–40050. doi:10.1074/jbc.M508904200

Kumar, P. (2013). Long term non-progressor (LTNP) HIV infection. *Indian J. Med. Res.* 138 (3), 291–293. Available at: https://pubmed.ncbi.nlm.nih.gov/24135172.

Lederman, M. M., Alter, G., Daskalakis, D. C., Rodriguez, B., Sieg, S. F., Hardy, G., et al. (2010). Determinants of protection among HIV-exposed seronegative persons: An overview. *J. Infect. Dis.* 202, S333–S338. doi:10.1086/655967

Li, T., and Deng, P. (2016). Nuclear Magnetic Resonance technique in tumor metabolism. *Genes Dis.* 4 (1), 28–36. Available at:. doi:10.1016/j.gendis.2016.12.001

Liebenberg, L. J. P., Masson, L., Arnold, K. B., Mckinnon, L. R., Werner, L., Proctor, E., et al. Genital-systemic chemokine gradients and the risk of HIV acquisition in women. *J. Acquir Immune Defic. Syndr.* [Internet]. 2017;74(3):318–325. doi:10.1097/QAI.0000000000001218

Ling, H., Xiao, P., Usami, O., and Hattori, T. (2004). Thrombin activates envelope glycoproteins of HIV type 1 and enhances fusion. *Microbes Infect.* 6 (5), 414–420. doi:10.1016/j.micinf.2004.01.010

Lovell, T. M., Woods, R. J., Butlin, D. J., Brayley, K. J., Manyonda, I. T., Jarvis, J., et al. (2007). Identification of a novel mammalian post-translational modification, phosphocholine, on placental secretory polypeptides. *J. Mol. Endocrinol. [Internet* 39 (3), 189–198. Available at:. doi:10.1677/JME-07-0007

Luo, F., Chen, T., Liu, J., Shen, X., Zhao, Y., Yang, R., et al. (2016). Ficolin-2 binds to HIV-1 gp120 and blocks viral infection. *Virol. Sin.* 31 (5), 406–414. doi:10.1007/s12250-016-3808-3

Masson, L., Passmore, J-A. S., Liebenberg, L. J., Werner, L., Baxter, C., Arnold, K. B., et al. (2015). Genital inflammation and the risk of HIV acquisition in women. *Clin. Infect. Dis. [Internet* 61 (2), 260–269. Available at:. doi:10.1093/cid/civ298

Matheson, N. J., Sumner, J., Wals, K., Rapiteanu, R., Weekes, M. P., Vigan, R., et al. Cell surface proteomic map of HIV infection reveals antagonism of amino acid metabolism by vpu and nef. *Cell Host Microbe* [Internet]. 2015;18(4):409–423. doi:10.1016/j.chom.2015.09.003

Meyers, A. F. A., and Fowke, K. R. (2010). *International symposium on natural immunity to HIV: A gathering of the HIV-exposed seronegative clan*, 202 Suppl. United States: The Journal of infectious diseases, S327–S328.

Nicholson, J. K., Lindon, J. C., and Holmes, E. (1999). Metabonomics: Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate

statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 29 (11), 1181–1189. doi:10.1080/004982599238047

ONUSIDA, Comunicaciones y Promoción Mundial (2021). HOJA INFORMATIVA 2021. Estadísticas mundiales sobre el VIH. Available at: https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_es.pdf.

Paris, D., Maniscalco, M., and Motta, A. (2018). Nuclear magnetic resonance-based metabolomics in respiratory medicine. *Eur. Respir. J. [Internet]* 52 (4), 1801107. Available at:. doi:10.1183/13993003.01107-2018

Pereyra, F., Addo, M. M., Kaufmann, D. E., Liu, Y., Miura, T., Rathod, A., et al. (2008). Genetic and immunologic heterogeneity among persons who control HIV infection in the absence of therapy. *J. Infect. Dis.* 197 (4), 563–571. doi:10.1086/526786

Philippeos, C., Steffens, F. E., and Meyer, D. (2009). Comparative 1H NMR-based metabonomic analysis of HIV-1 sera. *J. Biomol. NMR* 44 (3), 127–137. doi:10.1007/s10858-009-9329-8

Planès, R., Ben Haij, N., Leghmari, K., Serrero, M., BenMohamed, L., and Bahraoui, E. (2016). HIV-1 tat protein activates both the MyD88 and TRIF pathways to induce tumor necrosis factor alpha and interleukin-10 in human monocytes. *J. Virol.* 90 (13), 5886–5898. doi:10.1128/JVI.00262-16

Plotnikov, A., Zehorai, E., Procaccia, S., and Seger, R. (2011). The MAPK cascades: Signaling components, nuclear roles and mechanisms of nuclear translocation. *Biochim. Biophys. Acta - Mol. Cell Res. [Internet* 1813 (9), 1619–1633. Available from:. doi:10.1016/j.bbamcr.2010.12.012

Pothlichet, J., Rose, T., Bugault, F., Jeammet, L., Meola, A., Haouz, A., et al. (2020). PLA2G1B is involved in CD4 anergy and CD4 lymphopenia in HIV-infected patients. *J. Clin. Invest.* 130 (6), 2872–2887. doi:10.1172/JCI131842

Puchades-Carrasco, L., and Pineda-Lucena, A. (2015). Metabolomics in pharmaceutical research and development. *Curr. Opin. Biotechnol. [Internet]* 35, 73–77. Available from. doi:10.1016/j.copbio.2015.04.004

Qing, O., Tojo, N., Ozan, B., Chendong, Y., Michael, S., et al. (2016). Mutations in mitochondrial enzyme GPT2 cause metabolic dysfunction and neurological disease with developmental and progressive features. *Proc. Natl. Acad. Sci. [Internet]* 113 (38), E5598–E5607. Available from. doi:10.1073/pnas.1609221113

Riddler, S. A., Li, X., Otvos, J., Post, W., Palella, F., Kingsley, L., et al. (2008). Antiretroviral therapy is associated with an atherogenic lipoprotein phenotype among HIV-1-infected men in the multicenter AIDS cohort study. *J. Acquir Immune Defic. Syndr.* 48 (3), 281–288. doi:10.1097/QAI.0b013e31817bbbf0

Rodríguez-Gallego, E., Gómez, J., Domingo, P., Ferrando-Martínez, S., Peraire, J., Viladés, C., et al. (2018). Circulating metabolomic profile can predict dyslipidemia in HIV patients undergoing antiretroviral therapy. *Atherosclerosis* 273, 28–36. Available from:. doi:10.1016/j.atherosclerosis.2018.04.008

Roth, G. A., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., et al. (2018). Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: A systematic analysis for the global burden of disease study 2017. *Lancet [Internet]* 392 (10159), 1736–1788. Available from. doi:10.1016/S0140-6736(18)32203-7

Sanchez Caballero, L., Gorgogietas, V., Arroyo, M. N., and Igoillo-Esteve, M. (2021). "Chapter Three - molecular mechanisms of β-cell dysfunction and death in monogenic forms of diabetes," in *Pancreatic ß-cell biology in health and disease [internet]*. Editors I. Santin and C. Galluzzi LBT-IR of, (Oxford: Academic Press), 139–256. Available at: https://www.sciencedirect.com/science/article/pii/S1937644821000101.

Shi, Y., Su, J., Chen, R., Wei, W., Yuan, Z., Chen, X., et al. (2022). The role of innate immunity in natural elite controllers of HIV-1 infection. *Front. Immunol.*, 13. Available at: https://www.frontiersin.org/article/10.3389/fimmu.2022.780922.

Sitole, L. J., Tugizimana, F., and Meyer, D. (2019). Multi-platform metabonomics unravel amino acids as markers of HIV/combination antiretroviral therapy-induced oxidative stress. *J. Pharm. Biomed. Anal. [Internet]* 176, 112796. Available from. doi:10.1016/j.jpba.2019.112796

Smolkin, M. B., and Perrotta, P. L. (2017). "Chapter 18 - molecular diagnostics for coagulopathies," in *Coleman WB, tsongalis GJBT-DMP* (New York: Academic Press), 221–233. Available at: https://www.sciencedirect.com/science/article/pii/B9780128008867000182.

Soto, P. C., Karris, M. Y., Spina, C. A., Richman, D. D., and Varki, A. (2013). Cell-intrinsic mechanism involving Siglec-5 associated with divergent outcomes of HIV-1 infection in human and chimpanzee CD4 T cells. *J. Mol. Med. Berl.* 91 (2), 261–270. doi:10.1007/s00109-012-0951-7

Swanson, B., Sha, B. E., Keithley, J. K., Fogg, L., Nerad, J., Novak, R. M., et al. (2009). Lipoprotein particle profiles by nuclear magnetic resonance spectroscopy in medically underserved HIV-infected persons. *J. Clin. Lipidol.* 3 (6), 379–384. doi:10.1016/j.jacl.2009.10.005

Taborda, N. A., González, S. M., Alvarez, C. M., Correa, L. A., Montoya, C. J., and Rugeles, M. T. (2015). Higher frequency of NK and CD4+ T-cells in mucosa and potent cytotoxic response in HIV controllers. *PLoS One [Internet]* 10 (8), e0136292. Available from. doi:10.1371/journal.pone.0136292

Taborda-Vanegas, N., Zapata, W., and Rugeles, M. T. (2011). Genetic and immunological factors involved in natural resistance to HIV-1 infection. *Open Virol. J.* 5, 35–43. doi:10.2174/1874357901105010035

Tarancon-Diez, L., Rodríguez-Gallego, E., Rull, A., Peraire, J., Viladés, C., Portilla, I., et al. (2019). Immunometabolism is a key factor for the persistent spontaneous elite control of HIV-1 infection. *EBioMedicine* 42, 86–96. Available at:. doi:10.1016/j.ebiom.2019.03.004

Thompson, D., Pepys, M. B., and Wood, S. P. (1999). The physiological structure of human C-reactive protein and its complex with phosphocholine. *Structure* 7 (2), 169–177. Available from. doi:10.1016/S0969-2126(99)80023-9

Thomsen, T., Schlosser, A., Holmskov, U., and Sorensen, G. L. (2011). Ficolins and FIBCD1: Soluble and membrane bound pattern recognition molecules with acetyl group selectivity. *Mol. Immunol. [Internet* 48 (4), 369–381. Available at:. doi:10.1016/j.molimm.2010.09.019

Tilokani, L., Nagashima, S., Paupe, V., and Prudent, J. (2018). in *Mitochondrial dynamics: Overview of molecular mechanisms*. Editors C. Garone and M. Minczuk, 62, 341–360. Available from. doi:10.1042/EBC20170104

Tort, O., Escribà, T., Egaña-Gorroño, L., de Lazzari, E., Cofan, M., Fernandez, E., et al. Cholesterol efflux responds to viral load and CD4 counts in HIV+ patients and is dampened in HIV exposed. *J. Lipid Res. [Internet*]. 2018;59(11):2108, 2115. doi:10.1194/jlr.M088153

Treede, I., Braun, A., Sparla, R., Kühnel, M., Giese, T., Turner, J. R., et al. (2007/07/18. 2007). Anti-inflammatory effects of phosphatidylcholine. *J. Biol. Chem. [Internet* 282 (37), 27155–27164. Available at:. doi:10.1074/jbc.m704408200

Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., et al. (2007). BioMagResBank. *Nucleic Acids Res. [Internet]* D402, D402–D408. doi:10.1093/nar/gkm957

Urano, E., Morikawa, Y., and Komano, J. (2013). Novel role of HSP40/DNAJ in the regulation of HIV-1 replication. *J. Acquir Immune Defic. Syndr.* 64 (2), 154–162. doi:10.1097/QAI.0b013e31829a2ef8

Vergis, E. N., and Mellors, J. W. (2000). Natural history of HIV-1 infection. *Infect. Dis. Clin. North Am.* 14 (4), 809–825, v-vi. v–vi. doi:10.1016/s0891-5520(05)70135-5

Von Roenn, J. H., Roth, E. L., and Craig, R. (1992). HIV-related cachexia: Potential mechanisms and treatment. *Oncology* 49, 50–54. doi:10.1159/000227129

Vysochan, A., Sengupta, A., Weljie, A. M., Alwine, J. C., and Yu, Y. (2017). ACSS2-mediated acetyl-CoA synthesis from acetate is necessary for human cytomegalovirus infection. *Proc. Natl. Acad. Sci. U. S. A.* 114 (8), E1528–E1535. doi:10.1073/pnas.1614268114

Wall, K. M., Kilembe, W., Vwalika, B., Haddad, L. B., Hunter, E., Lakhi, S., et al. Risk of heterosexual HIV transmission attributable to sexually transmitted infections and non-specific genital inflammation in Zambian discordant couples, 1994. *Int. J. Epidemiol.* [Internet].;46(5):1593–1606. doi:10.1093/ije/dyx045

Wang, S-F., Tsao, C-H., Lin, Y-T., Hsu, D. K., Chiang, M-L., Lo, C-H., et al. (2014). Galectin-3 promotes HIV-1 budding via association with Alix and Gag p6. *Glycobiology* 24 (11), 1022–1035. doi:10.1093/glycob/cwu064

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., et al. (2018). Hmdb 4.0: The human metabolome database for 2018. *Nucleic Acids Res.* 46, D608. Available at: https://pubmed.ncbi.nlm.nih.gov/29140435.

Xiao, H., Wyler, E., Milek, M., Grewe, B., Kirchner, P., Ekici, A., et al. (2021). CRNKL1 is a highly selective regulator of intron-retaining HIV-1 and cellular mRNAs. *MBio* 12 (1), e02525. doi:10.1128/mBio.02525-20

Xiong, W., Xie, J., Liu, J., and Xiong, H. (2014). HIV-1gp120 injures neurons by alteration of neuronal expressions of NR2B and PSD-95]. Xi bao yu fen zi mian yi xue za zhi = Chinese. *J. Cell Mol. Immunol.* 30 (2), 139–142. doi:10.1111/jnc.14640

Yeung, M. L., Houzet, L., Yedavalli, V. S. R. K., and Jeang, K-T. (2009). A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. *J. Biol. Chem.* 284 (29), 19463–19473. doi:10.1074/jbc.M109.010033

Yndart, A., Kaushik, A., Agudelo, M., Raymond, A., Atluri, V. S., Saxena, S. K., et al. (2015). Investigation of neuropathogenesis in HIV-1 clade B and C infection associated with IL-33 and ST2 regulation. *ACS Chem. Neurosci.* 6 (9), 1600–1612. doi:10.1021/acschemneuro.5b00156

Young, J. M., Turpin, J. A., Musib, R., and Sharma, O. K. (2011). Outcomes of a national Institute of allergy and infectious diseases workshop on understanding HIV-exposed but seronegative individuals. *AIDS Res. Hum. retroviruses* 27, 737–743. doi:10.1089/AID.2010.0313

Zangerle, R., Kurz, K., Neurauter, G., Kitchen, M., Sarcletti, M., and Fuchs, D. Increased blood phenylalanine to tyrosine ratio in HIV-1 infection and correction following effective antiretroviral therapy. *Brain Behav. Immun.* [Internet]. 2010;24(3):403–408. doi:10.1016/j.bbi.2009.11.004

Zapata, W., Rodriguez, B., Weber, J., Estrada, H., Quiñones-Mateu, M. E., Zimermman, P. A., et al. (2008). Increased levels of human beta-defensins mRNA in sexually HIV-1 exposed but uninfected individuals. *Curr. HIV Res.* 6 (6), 531–538. doi:10.2174/157016208786501463

Zhang, Y., and Bhavnani, B. R. (2006). Glutamate-induced apoptosis in neuronal cells is mediated via caspase-dependent and independent mechanisms involving calpain and caspase-3 proteases as well as apoptosis inducing factor (AIF) and this process is inhibited by equine estrogens. *BMC Neurosci. [Internet]* 7 (1), 49. Available from. doi:10.1186/1471-2202-7-49

Zhou, H., Xu, M., Huang, Q., Gates, A. T., Zhang, X. D., Castle, J. C., et al. (2008). Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe* 4 (5), 495–504. doi:10.1016/j.chom.2008.10.004

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 01 August 2023
DOI 10.3389/fmolb.2023.1238475

OPEN ACCESS

EDITED BY
Guillermo Moyna,
Universidad de la República, Uruguay

REVIEWED BY
Andrés Perez Parada,
Universidad de la República, Uruguay
Kaifeng Hu,
Chengdu University of Traditional Chinese Medicine, China

*CORRESPONDENCE
Denise M. Selegato,
denise.selegato@embl.de

RECEIVED 11 June 2023
ACCEPTED 18 July 2023
PUBLISHED 01 August 2023

CITATION
Selegato DM, Zanatta AC, Pilon AC, Veloso JH and Castro-Gamboa I (2023), Application of feature-based molecular networking and MassQL for the MS/MS fragmentation study of depsipeptides. *Front. Mol. Biosci.* 10:1238475. doi: 10.3389/fmolb.2023.1238475

Application of feature-based molecular networking and MassQL for the MS/MS fragmentation study of depsipeptides

Denise M. Selegato[1]*, Ana C. Zanatta[2], Alan C. Pilon[1], Juvenal H. Veloso[1] and Ian Castro-Gamboa[1]

[1]Nucleus of Bioassays, Biosynthesis, and Ecophysiology of Natural Products (NuBBE), Institute of Chemistry, São Paulo State University (UNESP), Araraquara, Brazil, [2]Núcleo de Pesquisa em Produtos Naturais e Sintéticos (NPPNS), Faculdade de Ciências Farmacêuticas, São Paulo University (USP), São Paulo, Brazil

The Feature-based Molecular Networking (FBMN) is a well-known approach for mapping and identifying structures and analogues. However, in the absence of prior knowledge about the molecular class, assessing specific fragments and clusters requires time-consuming manual validation. This study demonstrates that combining FBMN and Mass Spec Query Language (MassQL) is an effective strategy for accelerating the decoding mass fragmentation pathways and identifying molecules with comparable fragmentation patterns, such as beauvericin and its analogues. To accomplish this objective, a spectral similarity network was built from ESI-MS/MS experiments of *Fusarium oxysporum* at various collision energies (CIDs) and paired with a MassQL search query for conserved beauvericin ions. FBMN analysis revealed that sodiated and protonated ions clustered differently, with sodiated adducts needing more collision energy and exhibiting a distinct fragmentation pattern. Based on this distinction, two sets of particular fragments were discovered for the identification of these hexadepsipeptides: ($[M + H]^+$) *m/z* 134, 244, 262, and 362 and ($[M + Na]^+$) *m/z* 266, 284 and 384. By using these fragments, MassQL accurately found other analogues of the same molecular class and annotated beauvericins that were not classified by FBMN alone. Furthermore, FBMN analysis of sodiated beauvericins at 70 eV revealed subclasses with distinct amino acid residues, allowing distinction between beauvericins (beauvericin and beauvericin D) and two previously unknown structural isomers with an unusual methionine sulfoxide residue. In summary, our integrated method revealed correlations between adduct types and fragmentation patterns, facilitated the detection of beauvericin clusters, including known and novel analogues, and allowed for the differentiation between structural isomers.

KEYWORDS

MS/MS fragmentation, beauvericin, feature-based molecular networking, MassQL, PCA

1 Introduction

LC-MS/MS-based metabolomics is a well-established technique for analyzing metabolites present in biological systems. However, the process of metabolite identification remains challenging. The structural complexity, and presence of similar fragments across different compounds, stereoisomers, adducts, and other spectral interferences make metabolite annotation a difficult

task, even for experienced spectroscopists. Consequently, a significant number of metabolites in biological systems remain unidentified (de Jonge et al., 2023).

Although statistical tools have been developed for analyzing large datasets, they are primarily designed for data exploration and analysis rather than the specific task of metabolite annotation. These tools often rely on clustering, classification, or regression approaches, which may not provide the level of detail required for accurate and comprehensive metabolite identification. Therefore, specialized approaches are needed to effectively address the challenges associated with metabolite annotation in LC-MS-based metabolomics (Lindon et al., 2000; Vuckovic, 2012; Johnson et al., 2016; Wishart, 2016).

In recent years, molecular networking has emerged as a powerful tool for large-scale annotation of LC-MS/MS data (Wang et al., 2016). Molecular networking groups the mass spectra (MS^2) of metabolites in a dataset based on fragmentation patterns, such as neutral losses, comparable losses, and ions present in the mass spectra. This approach offers a systematic way to organize and analyze vast amounts of LC-MS/MS data, allowing for the identification and classification of metabolites into distinct metabolic classes (Pilon et al., 2019). This network can also be enriched by annotating the experimental MS^2 spectra against MS^2 spectral libraries or compound databases, propagating annotations through the network edges to adjacent unknown nodes (Schmid et al., 2021).

Several investigations have demonstrated the effectiveness of molecular networking for metabolite annotation. For example, Naman et al. (2017) successfully identified a novel cytotoxic peptide from *Symploca* sp. cyanobacteria using molecular networks and the GNPS library (Naman et al., 2017). Similarly, Klein-Júnior et al. (2017) analyzed *Palicourea sessilis* using molecular networks and the DNP-ISDB tool, resulting in the identification of eight monoterpene alkaloids, three of which were novel and exhibited moderate anticholinesterase activity (Klein-Júnior et al., 2017). Olivon et al. (2017) employed molecular networks and bioactivity data against the Chikungunya virus to prioritize and isolate four esters of 12-deoxyforbol, two of which were novel, from 107 Euphorbiaceae species with anti-viral activity (Olivon et al., 2017).

Molecular networks have also been utilized to discover chalcones with antimicrobial properties against *Staphylococcus aureus* from *Angelica keiskei* (Caesar et al., 2018), as well as to characterize new indole alkaloids from *Geissospermum laeve* with potential antiparasitic and cytotoxic activity (Fox Ramos et al., 2017). Furthermore, Nothias et al. (2018) developed the bioactive molecular network approach for bio-guided studies, facilitating the dereplication process by associating bioactivity values with ion quantifications detected in LC-MS experiments (Nothias et al., 2018). This approach led to the isolation of two new substances, maridric acids A and B, from previously uncharacterized marine microorganisms.

The use of molecular networking can also assist in understanding metabolic fragmentation. This method involves analyzing metabolites with different collision energies, resulting in unique cleavage patterns. By examining protonation sites, cleavage types, fragment stability, and characteristic ions, researchers can accurately identify metabolites and their specific fragmentation patterns. Importantly, this approach would reveal the connections between structurally similar metabolites by highlighting subtle and intricate fragmentation patterns (De Souza et al., 2020; de Jonge et al., 2023).

Based on advancements in molecular networking, Jarmusch et al. (2022) have developed a complementary method called MassQL to further explore underutilized MS/MS data (Jarmusch et al., 2022). MassQL captures the unique characteristics of MS data, including isotopic patterns, diagnostic fragmentation, and neutral loss, and establishes a comprehensive MS terminology for searching MS patterns across datasets. This powerful tool formalizes terms for MS1 patterns (e.g., precursor ion *m/z*, isotopic patterns) and MS/MS fragmentation patterns, ensuring compatibility with all types of mass spectrometry data. By integrating molecular networking with MassQL, researchers can employ a comprehensive and powerful approach to analyze LC-MS/MS data and gain insights into metabolite fragmentation pathways.

In the present study, our aim is to demonstrate the potential of combining principal component analysis (PCA), spectral similarity networking, and MassQL as an effective approach for decoding mass fragmentation pathways of beauvericins and analogues. The beauvericin class a cyclic hexadepsipeptides composed of alternating *N*-methyl amino acid and hydroxy acid residues (Urbaniak et al., 2020). This class shares structural similarities with other enniantins but differs primarily in the types of amino acids and hydroxy acids that encompass them. Enniantins usually contain 2-hydroxyisovaleric acid, *N*-methylvaline, *N*-methylisoleucine, *N*-methylleucine (Sy-Cordero et al., 2012; Renaud et al., 2017; Li et al., 2020), while beauvericins commonly feature *N*-methylpheylalanine (MePhe), 2-hydroxyisovaleric acid (Hiv) and 2-hydroxy-3-methylpentanoic acid (Hamill et al., 1969; Gupta et al., 1995; Li et al., 2020; Urbaniak et al., 2020).

Due to similar fragmentation of peptides and the challenge to identify them, the beauvericins were selected as a model metabolic class for application of our strategy using FBMN and MassQL. Depsipeptides usually have the proton located at the *N*-terminus or a basic residue side chain and, during fragmentation, it can move along the backbone, breaking at different conserved sites (Pallerla et al., 2019). Thus, to achieve this objective, we employ these approaches in a combination of different collision energies from ESI-MS/MS experiments from a *Fusarium oxysporum* extract to determine both the mass fragmentation pathways and the identification of known and novel hexadepsipeptides. The selection of these compound classes is based on the ongoing debate surrounding the fragmentation profile of beauvericin, specifically the assignment of the ion at *m/z* 362 (Liuzzi et al., 2017; Tolosa et al., 2019). While this ion is commonly identified as diagnostic fragment ions for beauvericin and some derivatives, there is controversy regarding their presence in other cyclic peptides, lacking the required specificity for their automated discovery. Thus, our study aims to provide insights into the ambiguous fragmentation between these compound classes and demonstrate the potential of our approach in deciphering these complex fragmentation patterns.

2 Materials and methods

2.1 Fungi fermentation and metabolite extraction

Fusarium oxysporum (Nectriaceae) was cultured in Czapek-Broth ($NaNO_3$, 1.5 g L^{-1}; KH_2PO_4, 0.5 g L^{-1}; $MgSO_4$, 0.25 g L^{-1}; $FeSO_4.7H_2O$, 0.025 g L^{-1}; KCl, 2.5 g L^{-1}; and D-glucose,

30.0 g L^{-1}) in three biological replicates. Each fungal replicate was cultivated separately in eight Erlenmeyer flasks containing 300 mL of Czapek broth. The medium was first autoclaved at 121 °C for 20 min and, after sterilization, *F. oxysporum* was inoculated and incubated while stationary at 26°C for 28 days.

At the end of the incubation period, the flasks were vacuum filtered to remove mycelium and extracted with ethyl acetate (3 × 500 mL). The solvent was evaporated, and the extracts were then submitted to a clean-up process with solid phase extraction (SPE) in a cartridge filled with C-18 reversed phase silica after reconstitution in methanol HPLC grade (Strata X, C18), followed by filtration in 0.22 μM membrane.

2.2 MS/MS parameter optimization

Direct flow infusion of the samples was performed using a high-resolution micrOTOF-QII mass spectrometer (Bruker Daltonics, Bremen, Germany) for the optimization of the ionization and fragmentation parameters. This preliminary analysis used beauvericin as targeted precursor ions, both in its protonated (theoretical *m/z* 784.4173) and sodiated (*m/z* 806.3992) adducts, aiming to obtain MS/MS data with distinctive product ions characteristic. To optimize the collision energy (CE), fragmentation experiments were performed altering the collision-induced dissociation (CID) energy. For each experiment, these precursor ions were selected from the full-scan mass spectrum using a mass error of 10 ppm and fragmented at different CID energies (10, 20, 25, 30, 40, 50, 60 and 70 eV). Optimized CID energies were selected based on the ability to produce different MS/MS spectra in terms of type of product ions and their intensity.

The optimized ESI parameters were set to the following values: positive ionization mode, nebulizer (N_2) gas pressure of 4.5 Bar, dry gas flow rate of 9.0 L min^{-1}, ion source temperature of 200°C, capillary voltage of 4500 V, and voltage source of 5 kV. Full scan analysis was conducted in the *m/z* range of 50–1,500.

2.3 LC-MS/MS parameters

The LC-MS/MS analysis of the extracts of each replicate of *F. oxysporum* was performed using ultra-fast liquid chromatography (UFLC) Shimadzu system (Shimadzu Prominence UFLC, Shimadzu) equipped with two solvent pumps (LC-20AD), a degassing system (DGU-20A3), an autosampler (SIL-20AHT), a column oven (CTO-20A), a system controller (CBM-20A), and a diode array detector (SPD-M20AV, Shimadzu). The UFLC system was coupled to the previously described high-resolution micrOTOF-QII mass spectrometer.

To perform the analysis, the samples were suspended in MeOH/H_2O (8:2, v/v) at a concentration of 5.0 mg mL^{-1}, centrifuged, and the supernatant was transferred to a vial for subsequent analysis. LC separations were performed on a Kinetex 2.6 μm XB-C18 core-shell column (100 × 2.1 mm ID, Phenomenex, Torrance, United States). The injection volume was 2 μL, the flow rate was 250 μL min^{-1}, and the column oven temperature was set to 40°C. The samples were eluted using a mobile phase consisting of water (solvent A) and acetonitrile (solvent B), both acidified with 0.1% formic acid, in a linear gradient from 5% to 100% (B) over a period of 45 min.

ESI and MS/MS conditions were set based on the previously optimized parameters. The spectra acquisition rate was established at 1 Hz. The instrument provided a resolving power of 9,000 per FWHM for the beauvericin precursor ion *m/z* 784.4173. The mass spectra were externally calibrated in Enhanced Quadratic mode, using the exact masses of the sodium trifluoroacetate (NaTFA) clusters ions from a 500 ppm solution in methanol/water (1:1, v/v). After each run, the calibrant solution was consistently injected at a flow rate of 3 μL/min via a six-port divert valve. The MS/MS fragmentation step involved selecting the 5 ions with the highest intensity, and active exclusion was applied after 4 spectra with a release time of 30 s. The mass spectra were processed using the DataAnalysis software (version 4.3, Bruker). MS data collected from *F. oxysporum* replicates can be found in the MassIVE dataset under the register number MSV000091616.

2.4 LC-MS/MS data processing

The mass spectrometry data were centroided and converted from the proprietary format (.raw) to the *m/z* extensible markup language format (.mzML) using the peak picking algorithm of ProteoWizard (ver. 3.0.19, MSConvert tool) (Chambers et al., 2012). The mzML files were then processed with MZmine3 (Schmid et al., 2023). In short, feature detection and deconvolution were performed with the ADAP chromatogram builder and local minimum resolver algorithm. The isotopologues were regrouped and the features were aligned, and gap filled across samples.

The aligned peak list was filtered to contain only peaks with an associated fragmentation spectrum. Finally, the feature quantification table results (.CSV) and spectral information (.MGF) were exported with the GNPS module for feature-based molecular networking analysis on GNPS and with SIRIUS export modules. The MZmine3 project, the MZmine3 batch file (.XML format), and results files (.MGF and. CSV) are available in the MassIVE dataset under the register number MSV000091616. The MZmine3 batch file contains all the parameters used during the processing.

2.5 LC–MS/MS data annotation

The files exported from MZmine3 were uploaded to GNPS (http://gnps.ucsd.edu) (Wang et al., 2016) platform, in which spectral library matching was performed against public fragmentation spectra (MS2) spectral libraries. Annotation has also been performed with SIRIUS (v. 5.5.7) to systematically annotate the MS2 spectra (Böcker et al., 2009; Böcker and Dührkop, 2016; Dührkop et al., 2019; 2021). Molecular formulas were computed with the SIRIUS module by matching the experimental and predicted isotopic patterns and from fragmentation trees analysis of MS2. The parameters for SIRIUS tools were set as follows: molecular formula candidates retained, 10; maximum precursor ion *m/z* computed, 850; profile, Q-TOF; MS2 mass accuracy (ppm), 10; possible ionizations $[M + H]^+$ and $[M + Na]^+$.

2.6 Feature-based molecular networking

A molecular network was created with the FBMN workflow (Nothias et al., 2020) on GNPS (Wang et al., 2016). The precursor ion mass tolerance was set to 0.01 Da and the MS/MS fragment ion tolerance to 0.02 Da. A molecular network was then created where edges were filtered to have a cosine score above 0.7 and more than 4 matched peaks. Further, edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 100, and the lowest-scoring edges were removed from molecular families until the molecular family size was below this threshold.

The analogue search mode was used by searching against MS/MS spectra with a maximum difference of 100.0 in the precursor ion value. The DEREPLICATOR software was used to annotate MS/MS spectra (Mohimani et al., 2018). The molecular networks were visualized using the Cytoscape software (Shannon et al., 2003) and the links are available in the Supplementary Materials for public access.

2.7 Principal component analysis (PCA)

The.MGF file containing the MS/MS information was binned (bin size of *m/z* 0.1) and the resulting table was used to visualize the clustering of compounds based on the presence of conserved ions using Principal Component Analysis. SDV function from the R package (version 2.5.5) was applied to calculate the Euclidean dissimilarity matrix based on the metabolite levels. Subsequently, classical metric multidimensional scaling was carried out based on the Euclidean distance matrix to obtain different principal coordinates.

2.8 MassQL analysis

The MassQL tool version 31.4 (Jarmusch et al., 2022) was employed to search for specific MS/MS fragments that were identified as specific for the beauvericin molecular class. Thus, these analogs were grouped by searching for MS/MS diagnostic ions at (1) *m/z* 134, 244, 262, 362 or (2) *m/z* 384, 284, 266 with a 0.1 m/*z* tolerance and a minimum percentage intensity relative to the base peak of 10.0% for analysis at CID energy of 25 eV, and 100% for 50 and 70 eV. The MassQL jobs can be publicly accessed, and the links are available in the Supplementary Materials. The extracted data was re-analyzed to generate a molecular network, using the same parameters as described above.

3 Results

To assess the ability of MS/MS Molecular Networking to detect beauvericin analogues and distinguish their fragmentation patterns in different collision energies (25, 50 and 70 eV), a fragmentation study was established by the untargeted tandem mass spectrometry of *F. oxysporum* extract based on the fast data-independent acquisition (DIA) function of QToF mass spectrometer. This *Fusarium* extract has been previously reported to produce beauvericin (Selegato et al., 2016) and the DIA setting allowed the automatic analysis and detection of known and unknown analogs that range in structural features and concentration in these fungal cultures.

3.1 Principal component analysis (PCA) of MS/MS data

PCA analysis was performed on the MS/MS data using the SDV algorithm to evaluate the effect of collision energy on fragmentation data of beauvericin and analogues. Separate PCAs were conducted for each collision energy, showing a cumulative explained variance of 58.27% (25 eV), 70.42% (50 eV), and 96.32% (70 eV) for the first five principal components (PCs). For the higher collision energy, a strong correlation was observed between beauvericin in PCs 3–6 (for 50 eV) and PCs 1–3 (for 70 eV), respectively (Figure 1). However, this correlation was associated with the type of adduct formed during MS acquisition (protonated and sodiated ions) rather than the fragmentation pattern itself.

The protonated beauvericins $[M + H]^+$ were clustered on the positive sides of PC3 and PC1 (for 50 and 70 eV, respectively) with four characteristic fragments identified: *m/z* 362, 262, 244 and 134. The *m/z* 134 fragment exhibited higher abundance at 70 eV and corresponds to the *N*-methylphenylalanine (MePhe) residue, while the ions *m/z* 244, 262, and 362 showed higher intensities at 50 eV and are associated with the c2 and b2-fragments of MePhe-Hiv residues and the b3-fragment of the Hiv-MePhe-Hiv residue, respectively. In general, the depsipeptide fragmentation occurs mostly at the peptide bond (b-fragments), leading to the formation of stable amino acid immonium ions ($H_2N^+ = CHR_2$), or, at higher energies, in neighboring sites, generating the a-fragment alkyl carbonyl (CHR-CO) and the c-fragment aminoalkyl bond (NH-CHR) (Hohmann et al., 2008).

Contrarily, the positive sides of PC6 and PC3 (for 50 and 70 eV, respectively) were dominated by sodiated adducts $[M + Na]^+$ which are clustered due to the presence of the ions *m/z* 384, 284, and 266. These fragments are also originated from the b3-fragmentation of Hiv-MePhe-Hiv residue and the b2/c2-fragments of the MePhe-Hiv residue, respectively. Interestingly, the sodiated adducts exhibited a lower overall abundance of fragments compared to the protonated ions, indicating increased stability of the Na^+ adducts. Only the ion *m/z* 384 was significant at both 50 and 70 eV, while no fragments were observed for beauvericin at 25 eV. Lastly, a higher number of beauvericin ions were clustered at 70 eV, confirming that at higher energy levels, the conserved ions are more prevalent and explain the variance in the dataset better than other fragments and the precursor ion.

PCA analysis mainly showed the significant distinctions between sodiated and protonated beauvericin adducts. However, evaluation of the diagnostic fragments correlated with this adducts also reveals variations in the fragmentation pattern, specifically regarding the type and quantity of fragments generated through peptide cleavages. Indeed, the representative fragmentation

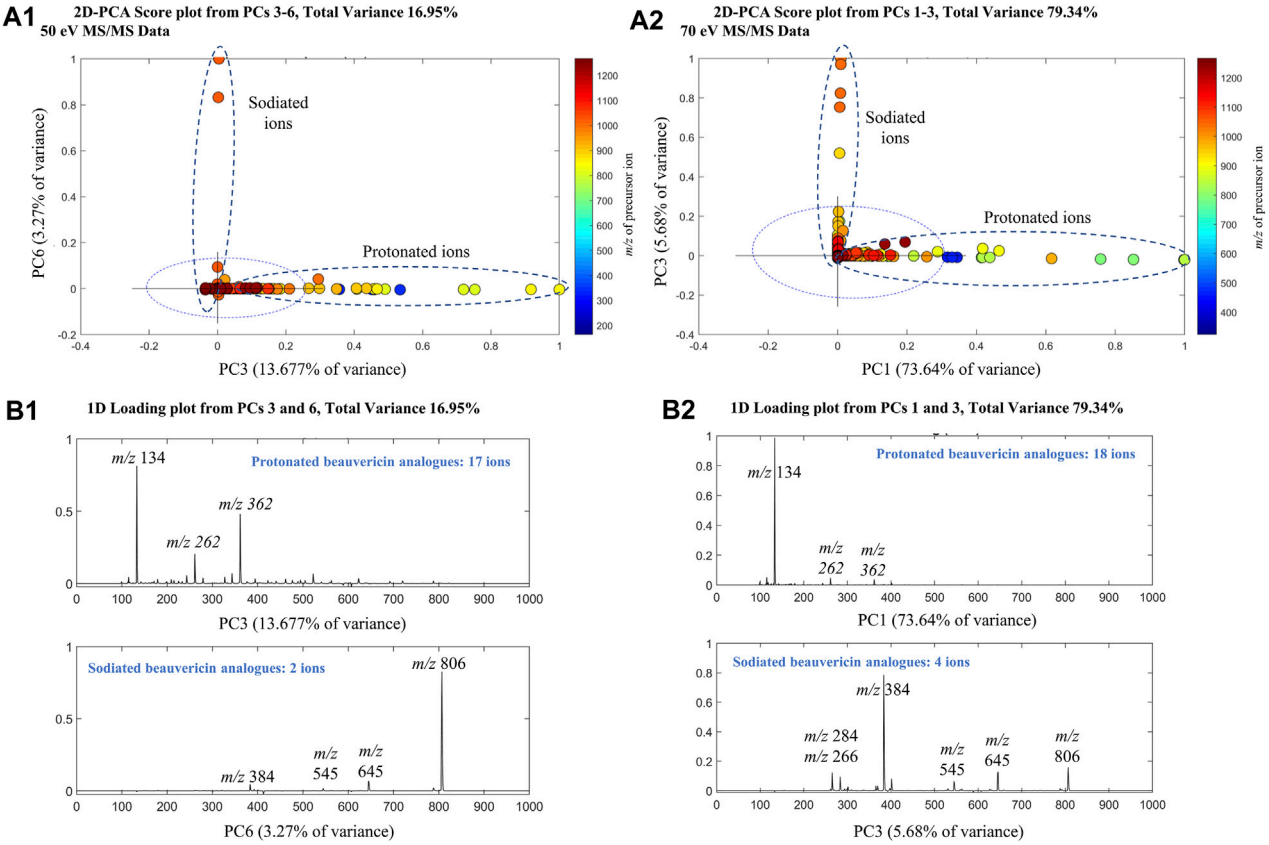


FIGURE 1
Principal Component Analysis (PCA) of the MS/MS beauvericin data at 50 eV and 70 eV. Analysis was performed using tandem MS data at different collision energies to obtain principal components (PCs) that explained the total beauvericin variation **(A)** 2D-Score plots of the PCs for 50 (PC3 and PC6) and 70 eV (PC1 and PC3). **(B)** 1D-Loading plots of the PCs that displayed a strong correlation was observed with beauvericin. For 50 eV, PC3 and PC6 show a combined variance of 16.95%, whereas for 70eV, PC1 and PC3 have 79.34% of explained variance.

spectra of beauvericin suggests that lower energies break the protonated ions into both b and c fragments, along with neutral loss of water (−18 Da) and methyl (−15 Da), whereas at higher energies, more stable b2-fragments have higher abundance in the spectra.

For protonated beauvericin, the most abundant fragment at 25 eV is the c4-fragment at *m/z* 541, as well as all five b-fragments (*m/z* 134, 262, 362, 523 and 623). As the collision energy increases, these b-fragments become more dominant over the c-fragments, in which the most significant ions are one, two and three-residue b-fragments (*m/z* 134, 262 and 362). At 70 eV, most fragments above *m/z* 400 have undergone fragmentation, and the *m/z* 262 and 362 appear at low intensities. Consequently, the single residues MePhe (*m/z* 134) and OH-Hiv (*m/z* 180) become the only representative fragments, with normalizes abundances of 100% and 20%, respectively (Figure 2). This fragmentation plurality is not observed in sodiated adducts, regardless of the collision energy. At 25 eV, no fragments are generated, and, at 50 and 70 eV, only b-fragments are observed. Interestingly, neither a nor c-fragments were formed at any of the collision energies, resulting in a cleaner spectrum compared to their respective protonated precursor.

3.2 Feature-based molecular networking (FBMN) of beauvericin analogues in different collision energies

The FBMN approach involves organizing MS/MS data (spectra) into a network based on the similarity of fragmentation patterns. Each node within the network represents a collection of spectra sharing the same precursor ion, and the relationships between nodes are determined by the degree of similarity among the spectra. Thus, to examine the variations in the fragmentation pattern of beauvericins under different CID energies, each dataset (25, 50, and 70 eV) was submitted to molecular networking approach.

Although all three networks contain a similar number of precursor ions, there is a disparity in the number of clusters. At higher collision energy (70 eV), there was a higher prevalence of singletons, which do not exhibit similar fragmentation patterns to other nodes, and the presence of smaller clusters with fewer than 10 nodes each. Conversely, at lower energies, a larger part of the nodes is grouped in big clusters, displaying almost three times less single nodes than at 70 eV. Moreover, at 25 eV, fewer clusters are present, but most of them contain many nodes (Figure 3).

10.3389/fmolb.2023.1238475

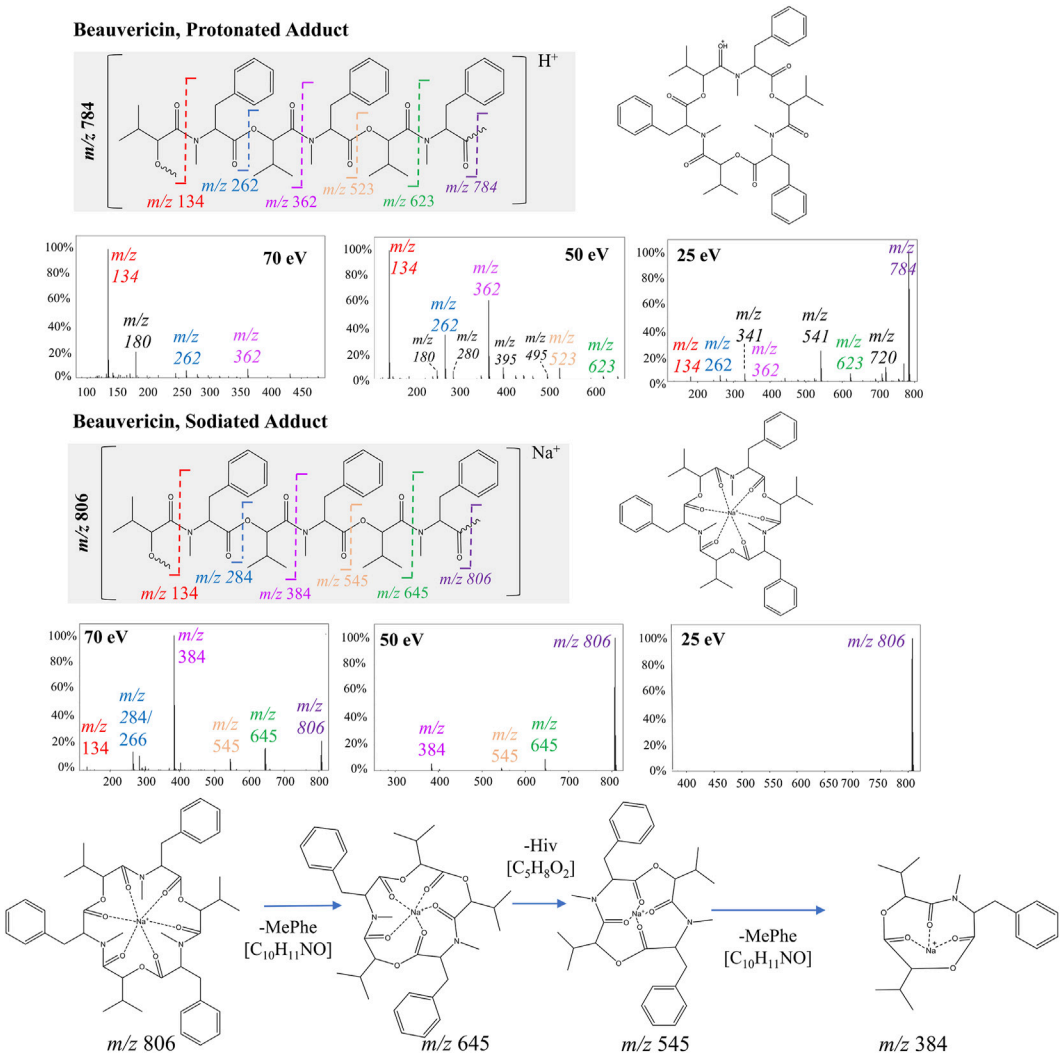


FIGURE 2
Fragmentation of beauvericin as a protonated ion and sodiated adduct at 25, 50, and 70 eV. The structure is shown both in its linear and cyclic form to facilitate visualization. Breaks in the peptide bond of each amino acid residue are shown in different colors and correspond to the b-fragments.

Protonated beauvericin ions were consistently clustered together, regardless of the collision energy applied. However, as the collision energy increases (50 and 70 eV), the number of ions grouped within the protonated clusters increases. In contrast, sodiated ions are mostly found as singletons (self-loop nodes) or scattered in random groups at 25 eV. At increasing energies, beauvericin nodes are grouped into numerous clusters of only a few nodes. For instance, at 50 eV, three highly specific beauvericin clusters are formed and MS/MS pattern is similar to the protonated cluster at 25 eV, as it balances the presence of the fragments with the precursor ion. In all three clusters, the identification of both the conserved ions (*m/z* 266, 284, and 384) and the repetitive loss of amino acid residues at their peptide bonds (b-fragments) are equally important to determining similarity. At energies above 50 eV, nodes that contain conserved beauvericin ions are grouped into different clusters and the presence of these fragments are no longer exclusively what determines the similarity.

FBMN analysis of sodiated beauvericins at 70 eV revealed subclasses with distinct amino acid residues, allowing distinction between known beauvericins (beauvericin and beauvericin D) and two previously unknown structural isomers (Figure 4). Both clusters contain the conserved ions at *m/z* 266, 284, 384. However, the MS/MS spectra of the unknown compounds, eluting in earlier retention time (~30 min), exhibit an additional neutral loss of 64 Da, which corresponds to the loss of methanesulfenic acid (CH_3SOH) from the methionine sulfoxide residue (Supplementary Figure S1). Putative annotation of the first novel metabolite (compound 1) showed a sodiated mass of *m/z* 806.3695, repetitive losses of 161–100–161–100 (from the MePhe-Hiv-MePhe-Hiv residues) and an additional loss of 97 Da, which corresponds to the *N*-methylated methionine residue (MeMe(O)) after the loss of methanesulfenic acid. The proposed molecular formula is $C_{41}H_{57}N_3O_{10}S$, displaying a mass error below 3 ppm, and retention time of 30.62 min. Compound (2) had the same

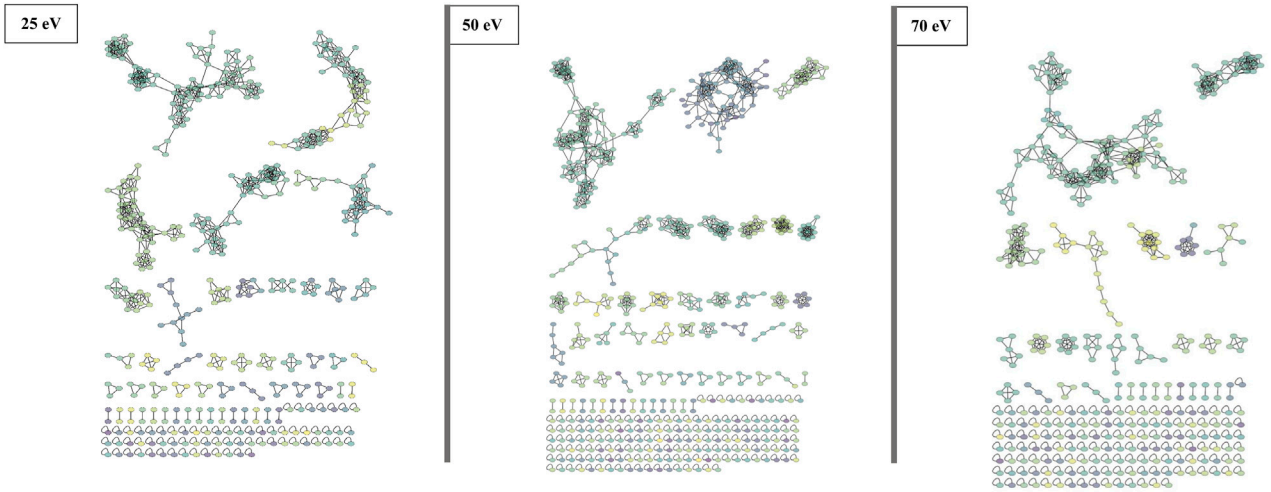


FIGURE 3
FBMN of beauvericin analogues at different collision energies (25, 50, and 70 eV). Each dataset was obtained from replicates of the same fungi extract and contains a similar number of MS/MS features. Nodes are colored according to the mass of the parent ions, in which yellowish colors represent lower masses (~*m/z* 250) and purple-like nodes constitute higher masses (~*m/z* 1,200).

fragmentation pattern as compound 1 (losses of 161–100–161–100 Da), expect for the loss of 83 Da, correlated to the non-methylated methionine sulfoxide residue (Me(O)) after the loss of methanesulfenic acid. This new depsipeptide has a *m/z* of 792.3536 and proposed molecular formula $C_{40}H_{55}N_3O_{10}S$ (mass error <3 ppm). Lastly, apart from the novel compounds, beauvericin D was also putatively annotated by MS/MS fragmentation. Although this molecule follows the exact same fragmentation pattern as beauvericin, its spectra also display a specific loss of 147 Da, which is characteristic of phenylalanine without any *N*-methylation. The putative MS/MS annotation of all four isomers are detailed in the Supplementary Materials (Supplementary Figures S2–5).

Overall, FBMN was successful in assisting the detection of beauvericin and other 13 analogues that share the same MS/MS fragmentation pattern. Among these molecules, some are known molecules that have been putatively identified, while others require further characterization and assessment of their unmatched fragment spectra. Table 1 shows the beauvericin ions identified for *F. oxysporum*, the proposed molecular formular and, when available, putative annotation.

3.3 Finding beauvericin analogues with MassQL analysis

The MassQL tool was used to search for spectra containing specific fragments related to beauvericin. For this, MassQL used the FBMN with different CEs to search for a query in a mass spectrometry-centric fashion, targeting to find its related ions without ambiguity. Three different strategies have been applied to search for this molecular class and a summary with their description and their main findings are available on Table 2. The complete results generated from each query and the links for their jobs on GNPS are available on Supplementary Material and Supplementary Table S2.

The initial query (query 1) aimed to search for MS/MS spectra that incorporated the product ion mass of at least one of the conserved ions previously identified in the beauvericin analogues. Thus, the search was conducted using the protonated fragment ions c_2, b_2 and b_2-H_2O at *m/z* 362, 262 and 244, and their sodium adducts at *m/z* 384, 284 and 266, respectively. The y ion at *m/z* 134 was searched for protonated ions only. This query resulted in 112, 167, and 156 scans for CID energies of 25, 50, and 70 eV, respectively. Furthermore, some ions identified by this query did not exhibit the characteristic fragmentation pattern of the hexadepsipeptide class, exhibiting *m/z* values below 600 or above 900. For purpose of illustration, at 50 eV, one of the most abundant ion was at *m/z* 625, and its fragmentation produced the major product ions at *m/z* 449, 405, 378, 322, and 181, which are not in agreement with the fragmentation pattern of beauvericin analogs.

To enhance the specificity of the search for beauvericin analogues and to define a precise sequence of *N*-MePhe and *D*-Hiv residues, the second query was employed (query 2) where conserved ions were systematically concatenated. This methodology aimed to refine the search parameters, thereby providing a more targeted examination of potential analogues. Upon implementing this query, CID energy levels of 25, 50, and 70 eV yielded 27, 54, and 65 scans, respectively. The significant reduction in the number of returned MS/MS spectra points out the increased specificity of this search strategy.

Given the potential modifications at the N- or C-termini of beauvericin analogues, identifying all derivatives via pseudoprecursor ion scanning and specific product ion formation becomes challenging without priorly conducting a fragmentation study, as demonstrated previously. To address this challenge, an alternative method was employed (query 3), which involved searching for a neutral loss sequence that included at least one

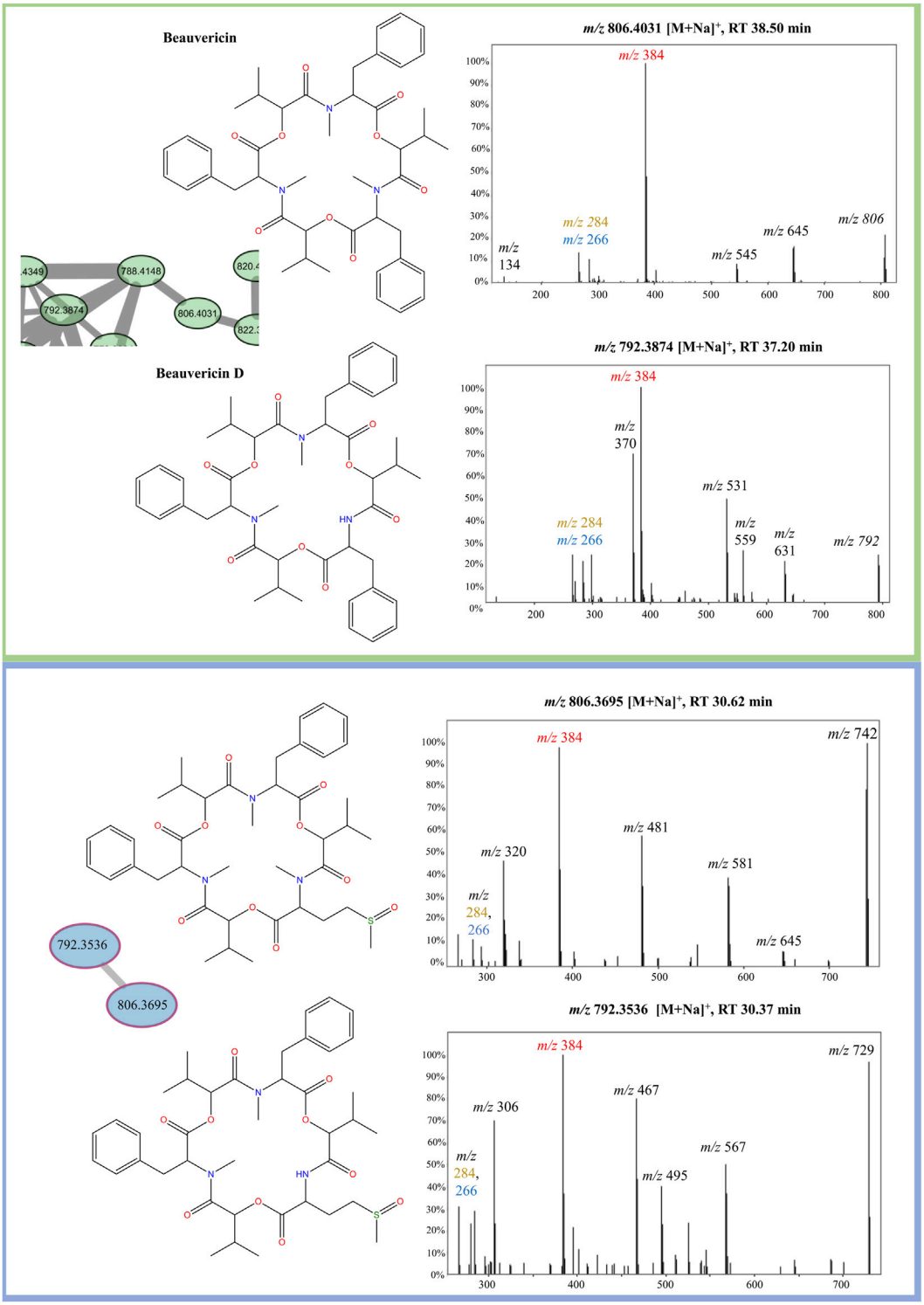


FIGURE 4
Elucidation of structural isomers from the beauvericin molecular class. Two sodiated ions at *m/z* 806 and *m/z* 792 have been clustered separately and contain different fragmentation patterns. The group colored in green belongs to the known compounds, beauvericin (*m/z* 806.4,031) and beauvericin D (*m/z* 792.3874) and elute at later retention time (~38 min). The other group, colored in blue, elutes at earlier retention times (~30 min) and belongs the to the novel beauvericin analogues (1, *m/z* 806.3695) and (2, *m/z* 792.3536). These metabolites contain an unusual methionine sulfoxide residue and a diagnostic loss of 64 Da, corresponding to the neutral loss of methanesulfenic acid. Both groups display the diagnostic beauvericin ions at *m/z* 266, 284, 384.

TABLE 1 Beauvericin analogues identified by FBMN and MassQL in *Fusarium oxysporum* extracts. All ions have been clustered with known beauvericin analogues and contain at least one of the diagnostic ions for the beauvericin molecular class. Molecules have been annotated by comparison with curated databases. Compound 1 and 2 have never been reported in literature and its putative annotation is described in the Supplementary Materials. Ions with (*) were identified by MassQL only. Compounds that have not been putatively annotated had their molecular formula proposed.

	Exp. *m/z*	Theoretical *m/z*	Error (ppm)	RT (min)	Putative annotation	Proposed	Most abundant fragments at 50 eV
						Molecular formula	
1	792.352 $[M + Na]^+$	792.3505	1.89	30.33	Compound 2	$C_{40}H_{55}N_3O_{10}S$	306, 384, 467, 567, 728, 792
2	806.3695 $[M + Na]^+$	806.3662	4.09	30.61	Compound 1	$C_{41}H_{57}N_3O_{10}S$	320, 384, 481, 581, 742, 806
3	808.378 $[M + Na]^+$	-	-	32.54	-	$C_{45}H_{59}N_3O_9$	284, 384, 547, 575, 647, 728, 808
4	822.396 $[M + Na]^+$	822.3941	2.31	32.7077	Beauvericin J	$C_{45}H_{57}N_3O_{10}$	384, 400, 545, 563, 645, 663, 822
5	820.411 $[M + Na]^+$	820.4149	4.75	33.74	Beauvericin A/F	$C_{46}H_{59}N_3O_9$	384, 398, 559, 659, 820
6	824.3755 $[M + Na]^+$	-	-	35.03	-	$C_{45}H_{59}N_3O_{10}$	384, 402, 563, 663, 824
7	825.464 $[M + Na]^+$	-	-	35.04	-	$C_{45}H_{58}N_2O_{11}$	402, 563, 663, 825
8	808.413 $[M + Na]^+$	-	-	35.78	-	$C_{45}H_{59}N_3O_9$	206, 386, 547, 647, 808
9	788.432 $[M + Na]^+$	-	-	36.24	-	$C_{48}H_{58}N_3O_7$	366, 384, 527, 627, 645, 788
10	822.3944 $[M + Na]^+$	822.3941	0.36	36.39	Beauvericin J	$C_{45}H_{57}N_3O_{10}$	300, 384, 400, 563, 663, 822
11	778.3641 $[M + Na]^+$	778.3679	4.88	36.85	Beauvericin G2	$C_{43}H_{53}N_3O_9$	356, 384, 517, 617, 778
12	758.3961 $[M + Na]^+$	758.3992	4.08	37.10	Beauvericin E	$C_{41}H_{57}N_3O_9$	758, 597, 525, 497, 384, 336
13	792.387 $[M + Na]^+$	792.3836	4.29	37.2159	Beauvericin D	$C_{44}H_{55}N_3O_9$	792, 631, 531, 384, 370
14	806.403 $[M + Na]^+$	806.3992	4.71	38.17	Beauvericin	$C_{45}H_{57}N_3O_9$	806, 645, 545, 384
15	722.3968*	-	-	36.2046	-	C40H55N3O9	433, 362, 333, 300, 262, 134
16	750.4325* $[M + H]^+$	750.4329	0.5	38.3355	Beauvericin K	$C_{42}H_{59}N_3O_9$	489, 362, 328, 262, 134
17	764.4355*	-	-	39.4984	-	$C_{43}H_{61}N_3O_9$	503, 603, 362, 342, 262, 134
18	780.4490*	-	-	34.5900	-	$C_{43}H_{61}N_3O_{10}$	537, 437, 262, 180

TABLE 2 MassQL strategies used for the search of beauvericin analogues. Queries 1 and 2 search for product ion formation, whereas the third uses neutral loss or delta mass.

Query N	Description	Targeted IONS	Main findings
QUERY 1	Search for conserved ions individually	Protonated: *m/z* 362 or 262 or 244 or 134; Sodiated: *m/z* 384 or 284 or 266	112 (25 eV), 167 (50eV), and 156 (70 eV) scans
QUERY 2	Search for conserved ions systematically concatenated	Protonated: *m/z* 362 and 262 and 244; Sodiated: *m/z* 384 and 284 and 266	27 (25 eV), 54 (50eV), and 65 (70 eV) scans
QUERY 3	Search for a neutral loss sequence of MePhe → Hiv → MePhe residues	Addition of a delta mass of 161 (X+161), followed by 100 (X+261) and then 161 (X+422)	150 (25 eV), 212 (50eV), and 81 (70 eV) scans

Hiv and two *N*-MePhe unit in the analogs. The MassQL query involved the addition of a delta mass of 161 (X+161), followed by 100 (X+261) and then 161 (X+422); this indicates the formation of a product ion at *m/z* = X, resulting from the sequential neutral loss of the MePhe→Hiv→MePhe residues. The query yielded 150, 212 and 81 scans for CID energies of 25, 50, and 70 eV, respectively. The higher number of identifications, compared to the results from the product ion scan (query 2), can be attributed to a greater sensitivity to mass deviation. This is because the query measures the relative mass loss between a precursor ion and a particular fragment ion. For example, in the dataset obtained at 50 eV, the query successfully pinpointed the beauvericin derivatives with precursor ions at *m/z* 784 $[M + H]^+$ (RT = 30.59 min), and *m/z* 784 $[M + H]^+$ (RT = 38.19 min), returning 12 scans and 2 scans for each respective MS/MS spectrum.

Given that the most promising results were obtained from the second and third queries, and to simplify the comparison of query results while avoiding redundancy, molecular networks were constructed using data derived from MassQL (Supplementary Figure S6). The dataset from each CID energy (25, 50, and 70 eV) contributed to the formation of distinctive clusters within the molecular network, displaying similar results as the preliminary FBMN. Upon visual inspection of the precursor ion clusters generated based on the proposed queries, it was observed that sodiated ions demonstrated improved clustering at higher fragmentation energies, with the most optimal clustering seen at a CID energy of 70 eV. On the contrary, protonated ions presented an opposite pattern. These findings are consistent with the FBMN results previously presented. In the context of sodiated ions clustering, query 03 appeared to be more effective, whereas query 02 identified two precursor ions at *m/z* 676 and 748 that do not fall within the beauvericin analogs. Furthermore, the identified sodiated precursor ions align with those manually analyzed in the molecular network generated by the FBMN approach.

With regard to protonated ions, query 02 demonstrated superior selectivity. Analyzing the molecular networks of the protonated ions at collision energies of 25 and 50 eV, it was possible to identify precursor ions that were not distinguished at the higher energy level or for the corresponding sodiated ions. The analysis of MS/MS mass spectra for the precursor ions identified in these clusters, specifically at *m/z* 722, m/z 750, m/z 764, and *m/z* 780, showed fragmentation patterns characteristic of beauvericin analogues (Supplementary Figure S7; Table 1). These precursor ions demonstrated similar fragmentations, due to the formation of the product ions at *m/z* 362, 262, and 134, along with consecutive neutral losses of 161 Da and 100 Da. As an illustration, the precursor ion at *m/z* 750 generated sequential neutral losses of 161, 100, and 127 Da; the neutral loss of 127 Da suggests that this ion contains an N-leucine/isoleucine unit in place of a phenylalanine unit (Xu et al., 2016). The annotation derived from structure-based propagation for the other ions implies that the modification occurs at the amino acid unit, as indicated by the observed mass differences of 128, 141 and 157 Da.

4 Discussion

4.1 Finding conserved ions by PCA-based molecular fingerprinting

Initially, we conducted PCA analysis on the MS/MS data, revealing a positive and correlation between collision energy and the percentage of variance explaining the presence of beauvericin. This correlation can be attributed to the promotion of successive cleavages at increasing energies, leading to the generation of building block structures ('ions") formed by single amino acid residues. This evidence is more prominent by analysis of PC1, which contains the highest explained variance of all PCs, embracing 21.22%, 26.70%, and 73.69% of the whole abundance variability.

PCA analysis for high collision energies data showed that beauvericin analogues were grouped based on the type of adduct (protonated or sodiated). Based on this distinction, two sets of particular fragments were discovery for the identification of these hexadepsipeptides: ($[M + H]^+$) *m/z* 134, 244, 262, and 362 and ($[M + Na]^+$) *m/z* 266, 284 and 384. To date, numerous LC-MS/MS methods have been developed to detect enniantins and beauvericins. However, most of these methods target the specific congeners at *m/z* 362, or semi-targeted relying on one or two product ions. Consequently, identifying a total of seven diagnostic fragments could assist in the identification of this molecular class, enabling a more automatic detection of analogues, reducing analysis time and minimizing compound overestimation.

Interestingly, sodiated adducts exhibited a lower overall abundance of the fragments, implying that these adducts are less prone to fragmentation than their protonated counterparts. This limited number of fragments can be attributed to the chelation effect of the carbonyl groups from the six amino acid residues with the sodium ion, further stabilizing the cycle and increasing the energy necessary for the cleavage of the peptides. On the one hand, this chelation effect makes the annotation of known beauvericin more straightforward given the spectra usually contains only b2-fragments, which preserve the entire amino acid residue. However, for unknown analogues, the information derived by other fragments in the protonated MS/MS spectra can aid in the determination of the right amino acid residues based on specific losses. The MePhe residue (*m/z* 134) was the only ion that appeared in both sodiated and protonated forms. This fragment thus represents the sole conserved ion presents in both adducts, making it a valuable fragment for this molecular class.

While this unsupervised analysis provided valuable insight into the presence of class-conserved ions, PCA alone falls short of clustering molecular features based on their fragmentation patterns, thus failing to identify compounds with similar MS/MS losses. Furthermore, the number of observed protonated and sodiated ions in the PCA is lower than the actual number of beauvericin analogues, and conserved ions often go unnoticed in lower intensity spectra. This limitation presents challenges, particularly for molecular classes that lack class-specific fragments. In such cases, a comprehensive assessment of the MS/MS data requires the integration of complementary tools that incorporate similarity outputs into the analysis process.

4.2 Exploration of beauvericin analogues by FBMN

Molecular networking has been widely used in natural product research to visualize and interpret untargeted MS/MS data. This algorithm is based on principle that structurally related molecules

tend to exhibit similar fragment patterns. As a result, MS/MS spectra can be mapped as molecular networks by comparing the spectral similarity between each pair of spectra. Given the structural relation of beauvericin analogues, these cyclic depsipeptides generate highly similar fragment patterns in tandem mass spectrometry. Consequently, these patterns can be organized into molecular families within the network, facilitation the annotation of known novel analogues. This approach has already been successfully applied to beauvericin and enniatin analogues, leading to the discovery of four enniatins and three beauvericin-producing fungi, as well as the identification of one new isomers of enniatin A and three new bassianolide analogs (Li et al., 2020).

Although MN has been extensively employed for compound mapping, only a few studies have used this approach to investigate fragmentation differences within a specific molecular class. Varying collision energies for the same molecule renders distinct MS/MS spectra, providing a deeper understanding of the chemical structure, fragment stability, and the accurate compound characterization.

To address the fragmentation pattern variations of beauvericin under different collision energies, we subjected each dataset (25, 50, and 70 eV) to feature-based molecular networking. This approach provided a visual representation of the clustering pattern and revealed differences between subclasses and adducts. For instance, the lower number of singletons at lower collision energies illustrates the reduced fragmentation at 25 eV, resulting in spectra formed mainly by the precursor ions and a weaker contribution of the fragments for overall fragmentation pattern. Moreover, at higher CID energies, there is a higher number of clusters that contain less than 10 nodes. This indicates that the increased number of fragments enhances specificity of the MS/MS spectra, differentiating fragmentation patterns even within the same molecular class.

Interesting, when we conducted a targeted search of the ions clustered on the PCA analysis for sodiated and protonated adducts, we observed a similar separation within MN. This finding was unexpected, considering that both types of ions theoretically have the same tendency of losing amino acid residues at the peptide bond. However, this discrepancy highlights the capacity of FBMN to distinguish classes not only based on their primary losses (such as b-fragments), but also on the more subtle spectral differences, as the presence of a and c-fragments, conserved ions, and losses specific to particular amino acid residues.

The protonated ions of beauvericin were consistently clustered together, irrespective of the collision energy applied (Figure 5). This suggests that the beauvericin analogues exhibit both diagnostic fragments and similar fragmentation patterns even at lower CID energies. However, a closer examination of the protonated cluster revealed that the highest number of ions were obtained at 50 eV. This improved accuracy in identifying this molecular class can be attributed to presence of MePhe, MePhe-Hiv, and Hiv-MePhe-Hiv fragment peaks at similar abundances (ranging from 50%–100%). At 70 eV, the cleavage of fragments containing two and three amino acid residues, increases the abundance and prevalence of the single-residue MePhe fragment. Consequently, other peptides that do not belong to the beauvericin analogs but also contain MePhe residues, such as pentacyclic peptides at *m/z* 522, 541, and 654, are clustered in the same group. This increases the specificity of peptides as opposed to specifically finding beauvericin hexadepsipeptides. Lastly, despite the protonated clusters at 25 eV contains the lowest number of ions grouped together, this ionization energy gives the most detailed chemical information about the amino acid residues, displaying the majority of the b-fragments and thereby facilitating the annotation of known and novel compounds (Figure 5).

The opposite scenario is encountered for sodiated beauvericin ions. At low CID energies, the fragmentation of these molecules is insufficient due to the chelation effect on the sodium ion. Consequently, beauvericin nodes either appear as singletons or clustered with nodes from other molecular classes that also do not fragment. At higher energies, MS/MS spectra contains a higher number of fragments. Therefore, its FBMN reveals the clustering of analogs based on their diagnostic fragmentation, resulting in smaller and more specific clusters (Figure 6). At 50 eV, the MS/MS spectra exhibited all the conserved ions (*m/z* 266, 284, and 384), leading to the most accurate detection of analogs. At energies above 50 eV, the specificity of these conserved ions diminishes in favor of providing information about amino acid residues other than MePhe and Hiv. This results in the formation of a larger number of small networks representing subclasses of beauvericin, which encompass the ions at *m/z* 266, 284, and 384 as well as residue-specific fragmentations.

The analysis of sodiated ions at high collision energies allowed the identification of two clusters that contain precursor masses of *m/z* 792 and 806. The first was annotated as beauvericin and its analogue beauvericin D. The latter has been previously isolated from *Beauveria* species (Fukuda et al., 2004), however, it has never been reported for *F. oxysporum*. The second group had their metabolites elucidated as the novel beauvericin analogues (MePhe-Hiv-MePhe-Hiv-MeMet(O)-Hiv and MePhe-Hiv-MePhe-Hiv-Met(O)) that contain an unusual methionine sulfoxide residue. Both compounds have never been reported for this species, although Gunasekera and co-workers have already shown the potential of bacteria to produce cyclic depsipeptides with this type of residue (Gunasekera et al., 2008). Putative annotation was mainly based on the neutral loss of 64 Da, which has been previously associated to the loss of methanesulfenic acid from precursors containing the oxidized methionine Met(O) residue (Pilo and McLuckey, 2014). This neutral loss is unique to Met(O) and was essential to differentiate between this residue and phenylalanine, given that both contain the same nominal mass.

Lastly, FBMN enabled the identification of beauvericin and other 13 analogues that share the same MS/MS fragmentation pattern. These include known analogues that have not been described for this species and unknown molecules, highlighting the potential of the molecular networking for the identification of novel secondary metabolites, as well as F. oxysporum potential to produce hexadepsipeptides.

Overall, FBMN provided valuable insights in several areas: (1) establishing the relationship between adducts and the fragmentation patterns, (2) identifying clusters specific to beauvericin, (3) characterizing both known and unknown analogues, and (4) differentiating between structural isomers. However, the extensive data generated by LC-MS/MS analysis, along with the occurrence of multiple fragmentation events during mass spectrometry analysis, can complicate the molecular

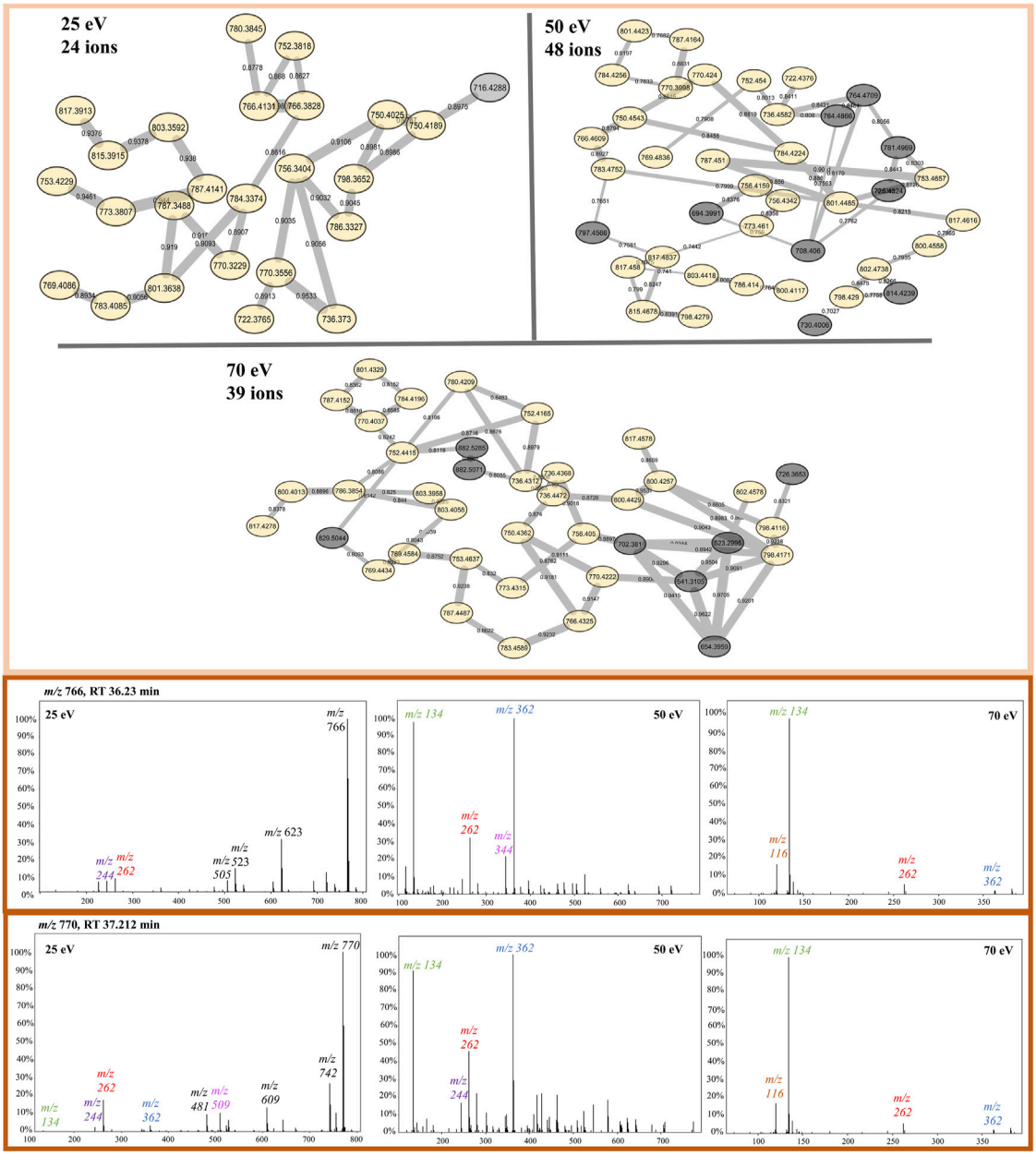


FIGURE 5
Protonated beauvericin clusters from the FBMN at 25, 50, and 70 eV. Nodes colored in yellow are found in all CIDs whereas nodes colored in gray are specific for only one CID. Example of the most abundant ions found on the protonated cluster. MS/MS spectra of the ions *m/z* 766 and *m/z* 770 are sequentially shown in all three collision energies. The diagnostic fragments for this molecular class are colored in the spectra as *m/z* 134, 244, 262, and 362.

networking process leading to inaccuracies in clustering and annotations. The fragmentation patterns, depending on the collision energy, may result in clustering of metabolites that are not exclusive to the beauvericin family, as observed for the sodiated ions at 25 eV (where no fragments were formed) and the protonated cluster at 70 eV (where only one abundant fragment was formed). Moreover, without prior knowledge derived from PCA regarding diagnostic fragments, determining the optimal collision energy for the FBMN and identifying its clusters become limiting, relying solely on manual validation and visual inspection of the MS/MS spectra.

4.3 Identification of beauvericin analogues by MassQL

Based on the analysis of the fragmentation pattern of beauvericin analogues and the insights derived from PCA and FBMN analysis, a

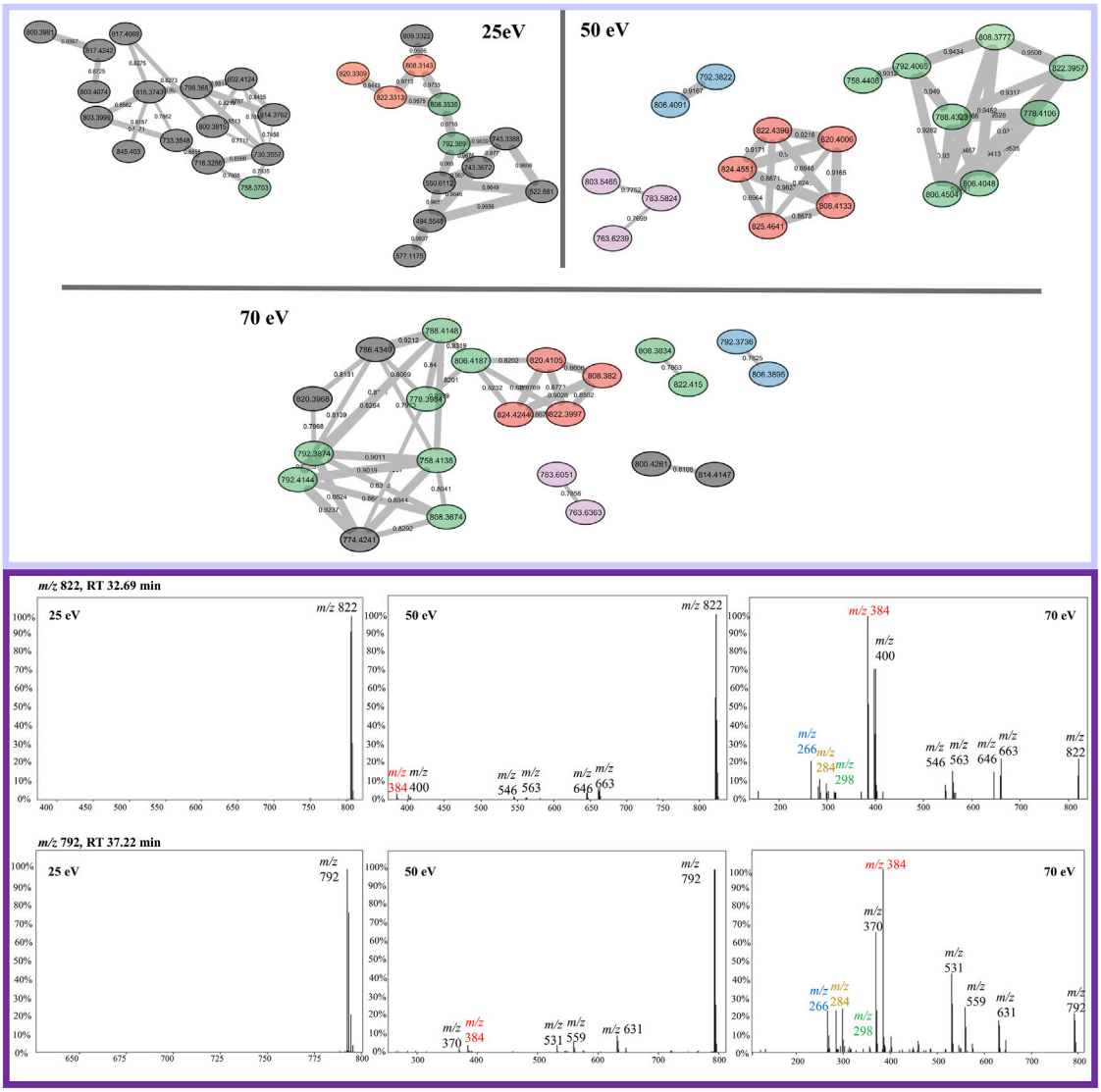


FIGURE 6
Sodiated beauvericin clusters from FBMN at 25, 50, and 70 eV. Nodes can be tracked by their color in all three collision energies, except for the nodes colored in gray which are specific for only one CID energy. Example of the most abundant ions found on the protonated cluster. MS/MS spectra of the ions *m/z* 822 and *m/z* 792 are sequentially shown in all three collision energies. The diagnostic fragments for this molecular class are colored in the spectra as *m/z* 266, 284, and 384.

series of investigations were conducted using the MassQL tool. The goal of this investigative process was to demonstrate the functionality of MassQL and to highlight its significant potential for mining beauvericin-like structures. The implementation of this tool was specifically designed to automate the search for spectra containing specific fragments associated with beauvericin. The automation facilitated by MassQL streamlines the identification process, enhancing the efficiency and accuracy of beauvericin analogue detection.

The application of combined multiple ion search proved to be a potent strategy to refine the search parameters, thereby increasing specificity and selectivity in the search for beauvericin analogues. Conserved ions serve as distinctive signatures in mass spectrometry studies, which enable efficient identification and grouping of compound analogues. This principle of structural similarity allows for a more targeted and efficient search within complex datasets, reducing false positives and enhancing the discovery of known and potentially novel analogues. However, the effectiveness of this method relies heavily on a thorough understanding of the fragmentation behavior of the compound class, and accurate identification of the conserved ions. Therefore, careful interpretation and validation of results are key to ensuring the correct grouping of analogues and the discovery of potential new ones.

The neutral loss process, on the other hand, focused on identifying specific mass differences between the precursor ion and its subsequent fragment ions, corresponding to the loss of a particular neutral moiety during fragmentation. This approach can

thus account for variability in fragmentation patterns. If a precursor ion can lose a specific group or moiety during fragmentation, the resulting mass difference can be tracked across various spectra, regardless of the exact identity of the fragment ions.

Neutral loss searches allowed for the identification of a broader range of compounds that contain the same functional group or moiety, even if their overall structures are different. This is especially important when studying a family of compounds, like beauvericin analogues, where subtle structural differences can lead to significant variations in fragmentation pattern.

These different strategies demonstrate the adaptability of mass spectrometry-based techniques for compound identification. Using an ion-focused approach to conserved products can provide high selectivity, which is beneficial when studying specific subclasses or known structures. On the other hand, a strategy focused on neutral losses can be employed when the goal is a more exhaustive search in a class of compounds. It is the careful choice and combination of these strategies that allows for a robust and differentiated exploration of complex mass spectrometry datasets.

5 Conclusion

In the present study, we demonstrate the potential of combining PCA, spectral similarity networking and MassQL as an effective approach for decoding mass fragmentation pathways of beauvericins. To achieve this objective, we employed these approaches in a combination of different collision energies from ESI-MS/MS experiments from a *F. oxysporum* extract to determine both the mass fragmentation pathways and the identification of known and novel hexadepsipeptides.

PCA and FBMN offered great insights into the correlation of the type of ions (protonated and sodiated adducts) and fragmentation patterns. Sodiated ions have a chelation effect on the sodium and the carbonyl from the peptide bond, further stabilizing the cycle and increasing the energy necessary for fragmentation. Based on this distinction, two sets of particular fragments were discovery for the identification of these hexadepsipeptides: ($[M + H]^+$) *m/z* 134, 244, 262, and 362 and ($[M + Na]^+$) *m/z* 266, 284 and 384. Currently, most methods for the screening of beauvericin are targeted to a few product ions. Hence, the identification of these seven diagnostic fragments could help in the search for this molecular class, decreasing analysis time and overestimation of compounds.

By using these fragments, MassQL accurately found other analogues of the same molecular class, identifying 18 beauvericins in this fungi extract, including 4 which were not found when analyzing FBMN alone. The superior potential of MassQL in detecting beavuericin analoges is due to the direct search of diagnostic ions, overcoming the inaccuracies in clustering and annotations of FBMN caused by the occurrence of multiple fragmentation events during MS analysis. Hence, the implementation of this tool enabled the automated search for beauvericins, streamlining the identification process and enhancing the efficiency and accuracy of analogue detection.

Lastly, FBMN analysis of sodiated beauvericins at 70 eV revealed subclasses with distinct amino acid residues, allowing distinction between beauvericins (beauvericin and beauvericin D) and two previously unknown structural isomers with an unusual methionine sulfoxide residue (MePhe-Hiv-MePhe-Hiv-MeMe(O)-Hiv and MePhe-Hiv-MePhe-Hiv-Me(O)-Hiv). Beauvericin D has been previously isolated from *Beauveria* species; however, it has never been reported for *F. oxysporum*. Similarly, these novel compounds had never been reported for fungi species, although previous studies have already shown the potential of bacteria to incorporate these residues into cyclic depsipeptides.

Ultimately, this approach revealed the correlation between adducts and the fragmentation patterns, the identification of beauvericin clusters, the characterization of known and unknown analogs, and the differentiation between structural isomers. The combination of different tools could shed light on conserved MS characteristics, facilitating the identification of metabolites in complex mixtures.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?accession=MSV000091616, MSV000091616.

Author contributions

Conceptualization: DS, AZ, and AP; Experiments: AZ and JV; Computational analysis: DS and AZ; Writing—original draft: DS, AZ, and AP; Writing—review and editing: DS, AZ, AP, and IC-G. Supervision and Funding Acquisition: IC-G. For experiments: JV performed fungi fermentation and extraction of the culture broths; AZ acquired the LC-MS/MS data. For computational analysis: DS processed the data using MZMine3, performed PCA, and submitted/analyzed the FBMN. AZ performed MassQL analysis. DS, AZ, and AP wrote the original draft of paper, which was revised by IC-G. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo [FAPESP grants number 2022/01073-0 and INCTBioNat 2014/50926-0] and by Conselho Nacional de Desenvolvimento Científico e Tecnológico–CNPq-INCT [grant number 465637/2014-0].

Acknowledgments

We would also like to thank Prof. Dr. Norberto Peporine Lopes for his theoretical contribution and Daniel Carretero Molina for helping to edit the figures.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1238475/full#supplementary-material

References

Böcker, S., and Dührkop, K. (2016). Fragmentation trees reloaded. *J. Cheminformatics* 8, 5. doi:10.1186/s13321-016-0116-8

Böcker, S., Letzel, M. C., Lipták, Z., and Pervukhin, A. (2009). Sirius: Decomposing isotope patterns for metabolite identification. *Bioinformatics* 25, 218–224. doi:10.1093/bioinformatics/btn603

Caesar, L. K., Joshua, J., KvalheimOlav, M., Cech, R. A., Cech, B., et al. (2018). Integration of biochemometrics and molecular networking to identify antimicrobials in Angelica keiskei. *Planta Med.* 84, 721–728. doi:10.1055/a-0590-5223

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., et al. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* 30, 918–920. doi:10.1038/nbt.2377

de Jonge, N. F., Louwen, J. J. R., Chekmeneva, E., Camuzeaux, S., Vermeir, F. J., Jansen, R. S., et al. (2023). MS2Query: Reliable and scalable MS2 mass spectra-based analogue search. *Nat. Commun.* 14, 1752. doi:10.1038/s41467-023-37446-4

De Souza, L. P., Alseekh, S., Brotman, Y., and Fernie, A. R. (2020). Network-based strategies in metabolomics data analysis and interpretation: From molecular networking to biological interpretation. *Expert Rev. Proteomics* 17, 243–255. doi:10.1080/14789450.2020.1766975

Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., et al. (2019). Sirius 4: A rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* 16, 299–302. doi:10.1038/s41592-019-0344-8

Dührkop, K., Nothias, L. F., Fleischauer, M., Reher, R., Ludwig, M., Hoffmann, M. A., et al. (2021). Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. *Nat. Biotechnol.* 39, 462–471. doi:10.1038/s41587-020-0740-8

Fox Ramos, A. E., Alcover, C., Evanno, L., Maciuk, A., Litaudon, M., Duplais, C., et al. (2017). Revisiting previously investigated plants: A molecular networking-based study of Geissospermum laeve. *J. Nat. Prod.* 80, 1007–1014. doi:10.1021/acs.jnatprod.6b01013

Fukuda, T., Arai, M., Tomoda, H., and Omura, S. (2004). New beauvericins, potentiators of antifungal miconazole activity, Produced by Beauveria sp. FKI-1366. II. Structure elucidation. *J. Antibiotics* 57, 117–124. doi:10.7164/antibiotics.57.117

Gunasekera, S. P., Ritson-Williams, R., and Paul, V. J. (2008). Carriebowmide, a new cyclodepsipeptide from the marine cyanobacterium lyngbya polychroa. *J. Nat. Prod.* 71, 2060–2063. doi:10.1021/np800453t

Gupta, S., Montllor, C., and Hwang, Y. S. (1995). Isolation of novel beauvericin analogues from the fungus Beauveria bassiana. *J. Nat. Prod.* 58, 733–738. doi:10.1021/np50119a012

Hamill, R. L., Higgens, C. E., Boaz, H. E., and Gorman, M. (1969). The structure op beauvericin, a new depsipeptide antibiotic toxic to. *Tetrahedron Lett.* 10, 4255–4258. doi:10.1016/S0040-4039(01)88668-8

Hohmann, L. J., Eng, J. K., Gemmill, A., Klimek, J., Vitek, O., Reid, G. E., et al. (2008). Quantification of the compositional information provided by immonium ions on a quadrupole-time-of-flight mass spectrometer. *Anal. Chem.* 80, 5596–5606. doi:10.1021/ac8006076

Jarmusch, A. K., Aron, A. T., Petras, D., Phelan, V. V., Bittremieux, W., Acharya, D. D., et al. (2022). A universal language for finding mass spectrometry data patterns. bioRxiv. doi:10.1101/2022.08.06.503000

Johnson, C. H., Ivanisevic, J., and Siuzdak, G. (2016). Metabolomics: Beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell. Biol.* 17, 451–459. doi:10.1038/nrm.2016.25

Klein-Júnior, L. C., Cretton, S., Allard, P. M., Genta-Jouve, G., Passos, C. S., Salton, J., et al. (2017). Targeted isolation of monoterpene indole alkaloids from Palicourea sessilis. *J. Nat. Prod.* 80, 3032–3037. doi:10.1021/acs.jnatprod.7b00681

Li, Y., He, N., Luo, M., Hong, B., and Xie, Y. (2020). Application of untargeted tandem mass spectrometry with molecular networking for detection of enniatins and beauvericins from complex samples. *J. Chromatogr. A* 1634, 461626. doi:10.1016/j.chroma.2020.461626

Lindon, J. C., Nicholson, J. K., Holmes, E., and Everett, J. R. (2000). Metabonomics: Metabolic processes studied by NMR spectroscopy of biofluids. *Concepts Magnetic Reson.* 12, 289–320. doi:10.1002/1099-0534(2000)12:5<289:AID-CMR3>3.0.CO;2-W

Liuzzi, V. C., Mirabelli, V., Cimmarusti, M. T., Haidukowski, M., Leslie, J. F., Logrieco, A. F., et al. (2017). Enniatin and beauvericin biosynthesis in Fusarium species: Production profiles and structural determinant prediction. *Toxins* 9, 45. doi:10.3390/toxins9020045

Mohimani, H., Gurevich, S., Shlemov, S., Mikheenko, S., Korobeynikov, A., Cao, L., et al. (2018). Dereplication of microbial metabolites through database search of mass spectra. *Nat. Commun.* 9, 4035. doi:10.1038/s41467-018-06082-8

Naman, C. B., Rattan, R., Nikoulina, S. E., Lee, J., Miller, B. W., Moss, N. A., et al. (2017). Integrating molecular networking and biological assays to target the isolation of a cytotoxic cyclic octapeptide, samoamide A, from an American Samoan marine cyanobacterium. *J. Nat. Prod.* 80, 625–633. doi:10.1021/acs.jnatprod.6b00907

Nothias, L. F., Nothias-Esposito, M., da Silva, R., Wang, M., Protsyuk, I., Zhang, Z., et al. (2018). Bioactivity-based molecular networking for the discovery of drug leads in natural product bioassay-guided fractionation. *J. Nat. Prod.* 81, 758–767. doi:10.1021/acs.jnatprod.7b00737

Nothias, L. F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., et al. (2020). Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* 17, 905–908. doi:10.1038/s41592-020-0933-6

Olivon, F., Allard, P. M., Koval, A., Righi, D., Genta-Jouve, G., Neyts, J., et al. (2017). Bioactive natural products prioritization using massive multi-informational molecular networks. *ACS Chem. Biol.* 12, 2644–2651. doi:10.1021/acschembio.7b00413

Pallerla, P., Bhumireddy, S. R., Lingampally, S. S., Ragi, N. C., and Sripadi, P. (2019). ESI-MS/MS analysis of protonated *N* -methyl amino acids and their immonium ions. *J. Mass Spectrom.* 54, 761–771. doi:10.1002/jms.4426

Pilo, A. L., and McLuckey, S. A. (2014). Oxidation of methionine residues in polypeptide ions via gas-phase ion/ion chemistry. *J. Am. Soc. Mass Spectrom.* 25, 1049–1057. doi:10.1007/s13361-014-0861-8

Pilon, A. C., Gu, H., Raftery, D., Bolzani, V. da S., Lopes, N. P., Castro-Gamboa, I., et al. (2019). Mass spectral similarity networking and gas-phase fragmentation reactions in the structural analysis of flavonoid glycoconjugates. *Anal. Chem.* 91, 10413–10423. doi:10.1021/acs.analchem.8b05479

Renaud, J. B., Kelman, M. J., McMullin, D. R., Yeung, K. K. C., and Sumarah, M. W. (2017). Application of C8 liquid chromatography-tandem mass spectrometry for the analysis of enniatins and bassianolides. *J. Chromatogr. A* 1508, 65–72. doi:10.1016/j.chroma.2017.05.070

Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., Dyrlund, T. S., et al. (2023). Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* 41, 447–449. doi:10.1038/s41587-023-01690-2

Schmid, R., Petras, D., Nothias, L. F., Wang, M., Aron, A. T., Jagels, A., et al. (2021). Ion identity molecular networking for mass spectrometry-based

metabolomics in the GNPS environment. *Nat. Commun.* 12, 3832. doi:10.1038/s41467-021-23953-9

Selegato, D. M., Freire, R. T., Tannüs, A., and Castro-Gamboa, I. (2016). New dereplication method applied to NMR-based metabolomics on different fusarium species isolated from Rhizosphere of Senna spectabilis. *J. Braz. Chem. Soc.* 27, 139. doi:10.5935/0103-5053.20160139

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–504. doi:10.1101/gr.1239303

Sy-Cordero, A. A., Pearce, C. J., and Oberlies, N. H. (2012). Revisiting the enniatins: A review of their isolation, biosynthesis, structure determination and biological activities. *J. Antibiotics* 65, 541–549. doi:10.1038/ja.2012.71

Tolosa, J., Rodríguez-Carrasco, Y., Ferrer, E., and Mañes, J. (2019). Identification and quantification of enniatins and beauvericin in animal feeds and their ingredients by LC-QTRAP/MS/MS. *Metabolites* 9, 33. doi:10.3390/metabo9020033

Urbaniak, M., Waśkiewicz, A., Koczyk, G., Błaszczyk, L., and Stępień, Ł. (2020). Divergence of beauvericin synthase gene among Fusarium and trichoderma species. *J. Fungi* 6, 288. doi:10.3390/jof6040288

Vuckovic, D. (2012). Current trends and challenges in sample preparation for global metabolomics using liquid chromatography–mass spectrometry. *Anal. Bioanal. Chem.* 403, 1523–1548. doi:10.1007/s00216-012-6039-y

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* 34, 828–837. doi:10.1038/nbt.3597

Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* 15, 473–484. doi:10.1038/nrd.2016.32

Xu, X., Zhao, S., Yu, Y., Chen, Z., Shen, H., and Zhou, L. (2016). Beauvericin K, a new antifungal beauvericin analogue from a marine-derived Fusarium sp. *Nat. Product. Commun.* 11, 1934578X1601101. doi:10.1177/1934578X1601101213

frontiers | Frontiers in Molecular Biosciences

TYPE Review
PUBLISHED 08 August 2023
DOI 10.3389/fmolb.2023.1215039

OPEN ACCESS

EDITED BY
Guillermo Moyna,
Universidad de la República, Uruguay

REVIEWED BY
Steven O'Reilly,
STipe Therapeutics, Denmark
Blaž Burja,
University of Zurich, Switzerland

*CORRESPONDENCE
Carolina Ramírez-Santana,
heily.ramirez@urosario.edu.co

RECEIVED 01 May 2023
ACCEPTED 27 July 2023
PUBLISHED 08 August 2023

CITATION
Morales-González V, Galeano-Sánchez D, Covaleda-Vargas JE, Rodriguez Y, Monsalve DM, Pardo-Rodriguez D, Cala MP, Acosta-Ampudia Y and Ramírez-Santana C (2023), Metabolic fingerprinting of systemic sclerosis: a systematic review.
Front. Mol. Biosci. 10:1215039.
doi: 10.3389/fmolb.2023.1215039

Metabolic fingerprinting of systemic sclerosis: a systematic review

Victoria Morales-González[1], Daniel Galeano-Sánchez[1], Jaime Enrique Covaleda-Vargas[1], Yhojan Rodriguez[1], Diana M. Monsalve[1], Daniel Pardo-Rodriguez[2], Mónica P. Cala[2], Yeny Acosta-Ampudia[1] and Carolina Ramírez-Santana[1]*

[1]Center for Autoimmune Diseases Research (CREA), School of Medicine and Health Sciences, Universidad Del Rosario, Bogotá, Colombia, [2]Metabolomics Core Facility—MetCore, Vicepresidency for Research, Universidad de Los Andes, Bogotá, Colombia

Introduction: Systemic sclerosis (SSc) is a chronic autoimmune disease, marked by an unpredictable course, high morbidity, and increased mortality risk that occurs especially in the diffuse and rapidly progressive forms of the disease, characterized by fibrosis of the skin and internal organs and endothelial dysfunction. Recent studies suggest that the identification of altered metabolic pathways may play a key role in understanding the pathophysiology of the disease. Therefore, metabolomics might be pivotal in a better understanding of these pathogenic mechanisms.

Methods: Through a systematic review of the literature following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Guidelines (PRISMA), searches were done in the PubMed, EMBASE, Web of Science, and Scopus databases from 2000 to September 2022. Three researchers independently reviewed the literature and extracted the data based on predefined inclusion and exclusion criteria.

Results: Of the screened studies, 26 fulfilled the inclusion criteria. A total of 151 metabolites were differentially distributed between SSc patients and healthy controls (HC). The main deregulated metabolites were those derived from amino acids, specifically homocysteine (Hcy), proline, alpha-N-phenylacetyl-L-glutamine, glutamine, asymmetric dimethylarginine (ADMA), citrulline and ornithine, kynurenine (Kyn), and tryptophan (Trp), as well as acylcarnitines associated with long-chain fatty acids and tricarboxylic acids such as citrate and succinate. Additionally, differences in metabolic profiling between SSc subtypes were identified. The diffuse cutaneous systemic sclerosis (dcSSc) subtype showed upregulated amino acid-related pathways involved in fibrosis, endothelial dysfunction, and gut dysbiosis. Lastly, potential biomarkers were evaluated for the diagnosis of SSc, the identification of the dcSSc subtype, pulmonary arterial hypertension, and interstitial lung disease. These potential biomarkers are within amino acids, nucleotides, carboxylic acids, and carbohydrate metabolism.

Discussion: The altered metabolite mechanisms identified in this study mostly point to perturbations in amino acid-related pathways, fatty acid beta-oxidation, and in the tricarboxylic acid cycle, possibly associated with inflammation, vascular damage, fibrosis, and gut dysbiosis. Further studies in targeted metabolomics are

required to evaluate potential biomarkers for diagnosis, prognosis, and treatment response.

KEYWORDS

systemic sclerosis, metabolomics, metabolic pathways, biomarkers, amino acids

1 Introduction

Systemic sclerosis is a chronic and rare autoimmune disease of the connective tissue, whose etiology remains unknown, and the pathogenesis is still partially understood, thus representing a clinical challenge and an unmet medical need (Denton and Khanna, 2017; Tsou et al., 2021). SSc is characterized by a pathogenic triad consisting of microvascular damage, innate and adaptive immune system abnormalities with autoantibody production and cell-mediated autoimmunity, and fibroblast dysfunction with excessive collagen deposition, which leads to progressive fibrosis of the skin and internal organs (Allanore et al., 2015; Pattanaik et al., 2015).

A striking characteristic of the disease is the variability within patients, with great heterogeneity in the clinical manifestations, the serological profiles, and the disease progression rate. SSc is characterized by high morbidity and mortality, higher than any other rheumatic disease (Elhai et al., 2017). Worldwide, the prevalence of SSc is approximately 17.6 cases per 100,000 population, and the incidence rate is 1.4 per 100,000 person-years; however, there is great variability among geographic populations (Bairkdar et al., 2021). Recent studies have reported a mortality rate of 1.39–5.1 times higher than the general population (Elhai et al., 2012; Rubio-Rivas et al., 2014; Hao et al., 2017; Kang et al., 2018). SSc has a disease-related mortality rate of approximately 55%, with the leading causes of death being pulmonary complications such as interstitial lung disease (ILD), followed by pulmonary arterial hypertension (PAH) (Tyndall et al., 2010). ILD has a 19%–52% prevalence in SSc patients (Perelas et al., 2020; Kuwana et al., 2022), and approximately 40% of these patients die within 10 years of diagnosis (Akter et al., 2014). The pathogenesis of SSc-ILD begins with a permanent injury to the alveolar epithelium, secondary to an activation of the immune system promoting profibrotic stimuli that induce fibroblast recruitment and differentiation to a myofibroblast phenotype and extracellular matrix (ECM) overproduction (Nihtyanova and Denton, 2020). Moreover, in SSc patients, PAH occurs with a prevalence of 5%–15% (Morrisroe et al., 2017; Naranjo and Hassoun, 2021). PAH is characterized by arterial remodeling, and an increased pulmonary vascular resistance secondary to abnormal vascular proliferation, disequilibrium in vasodilators, proliferative mediators, and thrombosis of the pulmonary vasculature resulting in right heart failure, which can eventually lead to death (Launay et al., 2017).

In relation to skin involvement, SSc can be subclassified into diffuse cutaneous systemic sclerosis (dcSSc) and limited cutaneous systemic sclerosis (lcSSc) (Young and Khanna, 2015). LcSSc is the most frequent subtype presentation (Coral-Alvarado et al., 2009), characterized by a gradual and early onset of Raynaud's phenomenon and skin fibrosis restricted to certain areas, such as the face and distal extremities, with minor systemic involvement (Herrick, 2018). In contrast, in the dcSSc subtype, Raynaud's phenomenon coexists with skin fibrosis extended proximally to knees, elbows, and the trunk with a more aggressive progression characterized by severe internal organ manifestations, mainly in the gastrointestinal tract, lungs, heart, and kidneys (Kowalska-Kępczyńska, 2022).

A better understanding of the pathophysiology of SSc is crucial to tackle the processes leading to disease progression and to discover effective therapies to improve the long-term survival of SSc patients. Recent studies suggest that metabolic perturbations may play an important role in SSc pathogenesis and are exhibited in different patients as a result of the disease heterogeneity and erratic course (Wishart, 2016; Cambiaghi et al., 2017; O'Reilly, 2022). Therefore, metabolomics may play an important role in understanding the pathophysiology of the disease.

Metabolic characterization represents a promising approach that can be applied for diagnosis, disease typing, and individual treatment of SSc, as well as biomarker discovery (Zhang et al., 2015a). Thus, this review aimed to identify altered metabolic pathways possibly responsible for the mechanisms associated with the appearance of SSc, in order to improve diagnosis, prognosis, and treatment. Additionally, it identifies metabolites that allow the segregation of patients with dcSSc and lcSSc.

2 Methods

2.1 Information sources and search strategy

A systematic review of the literature was conducted following the recommendations of the PRISMA guidelines (Page et al., 2021). Published studies related to the topic were retrieved after a literature search in four databases: PubMed, EMBASE, Web of Science, and Scopus, from January 2000 to September 2022. The references listed in the articles were manually searched. Only English and Spanish language articles were included. A search strategy combining MESH terms and free words were developed: ("Metabolomics," "Untargeted Metabolomics," "Metabolomic Fingerprinting," "Metabonomic," "Targeted Metabolomics," "Metabolic footprinting," "Metabolic profiling," "Metabolome," "Metabolic profile," "Lipidomics," "Lipidome" OR "Lipidomes") AND ("Scleroderma, Systemic," "Scleroderma, Diffuse," "Scleroderma, Limited," "CREST Syndrome," "Systemic Sclerosis," "Systemic Scleroderma" OR "Sclerosis, Systemic"). Supplementary Table S1 depicts the search strategy 1. This systematic review was not registered. Protocol was not written prior to the elaboration of the systematic review.

2.2 Eligibility criteria

Studies meeting the following criteria were included: 1) analytical observational studies (i.e., cross-sectional, case-control,

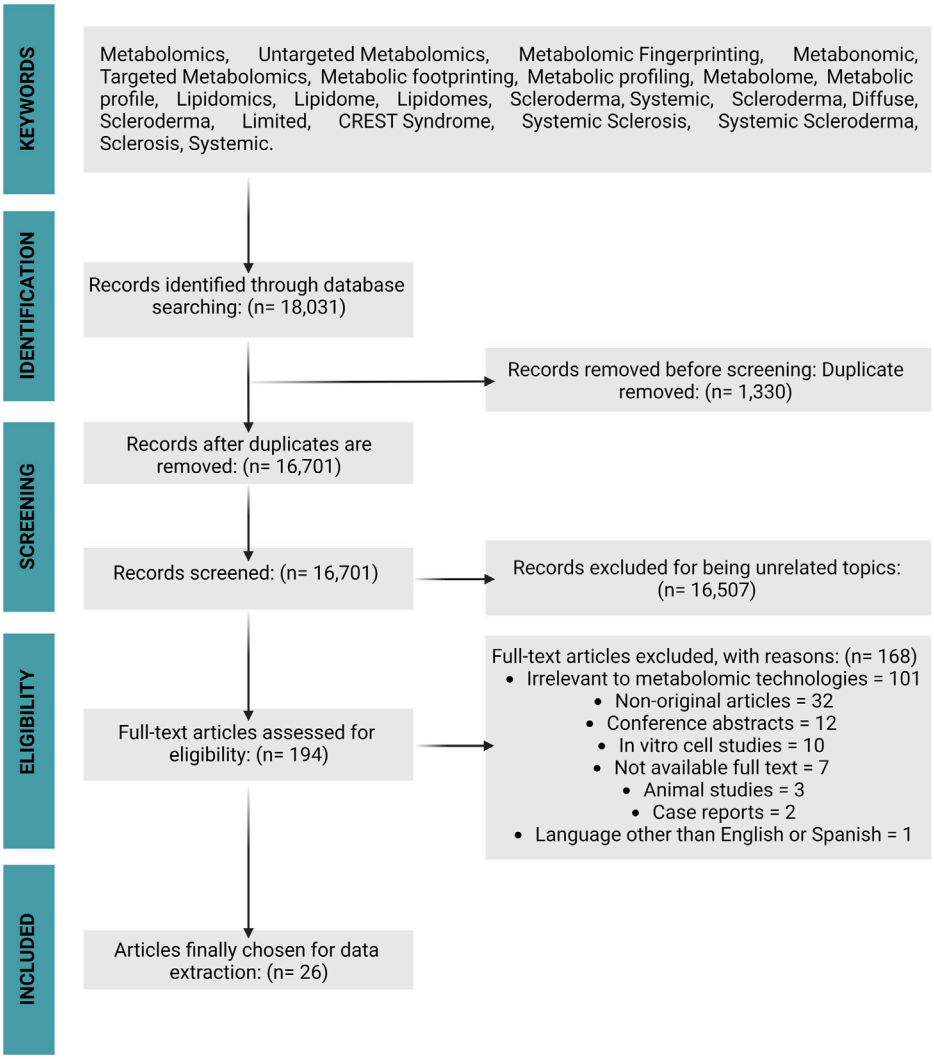


FIGURE 1
Preferred Reporting Items for Systematic Reviews and Meta-analyses flow chart.

and cohort studies) that evaluated altered metabolites in biological samples such as serum, plasma, urine, and exhaled breath by high-throughput techniques in patients diagnosed with SSc in comparison to HC; 2) studies published in English and/or Spanish and 3) studies implemented in adults. The exclusion criteria were as follows: 1) animal or *in vitro* cell studies; 2) non-original articles; 3) conference abstracts, guidelines, or editorials; 4) studies using irrelevant metabolomics techniques and 5) article data incomplete or missing.

2.3 Study selection

Three reviewers (VM, JC, DG), after removing duplicate articles, independently reviewed all the selected studies in the initial research in a two-step procedure assessing their eligibility. In the first phase, all identified titles and abstracts were evaluated to determine which records were possibly eligible for inclusion. Subsequently, the potentially relevant articles were selected and assessed again in the second phase. In this step, a full-text review was done to determine the eligible records according to the above criteria. Discrepancies in the final decision were resolved by consensus. The reasons for excluding studies were recorded. The primary outcome was to identify the differences in metabolic patterns between patients with SSc and HC. Secondary outcomes included comparing the differences in metabolic patterns across dcSSc and lcSSc subtypes and cardiopulmonary complications, as well as identifying potential metabolite biomarkers for SSc diagnosis and classification.

2.4 Data extraction and result synthesis

Data from each study were manually extracted and transferred into a Microsoft Excel form to include the following variables: (Denton and Khanna, 2017): publication information including first author, year of publication, and study geographic location; (Tsou et al., 2021); patients characteristics including age and sex;

(Allanore et al., 2015); sample size; (Pattanaik et al., 2015); sample type; (Elhai et al., 2017); methods used for metabolite identification and analysis, and (Bairkdar et al., 2021) differentially distributed metabolites across comparison groups. Key metabolites features were manually extracted based on statistical significance (p-values below a threshold of 0.05 or an area under the receiver-operator curve (AUC) greater than 0.70). These metabolites were then categorized according to the body fluid that was studied (plasma, serum, urine, and exhaled breath) and imported into the software MetaboAnalyst 5.0 for the generation of metabolic pathways enrichment analysis, which provides p values adjusted for multiple testing and uses the high-quality SMPDB metabolic pathways as the metabolite set library. There were no methods required for data conversion or the processing of missing summary statistics. Three reviewers (VM, JC, DG) independently extracted the information. Consensus was used to settle any inconsistencies or missing information. Tables were used to present the extracted metabolites.

2.5 Quality assessment

The quality of the eligible studies was evaluated using the QUADOMICS evaluation tool (Lumbreras et al., 2008). This scale represents an adaptation of the Quality Assessment of Diagnostic Accuracy Assessment (QUADAS) which assesses the quality of studies on omics-based research. This evaluation tool has 16 items, each of which can be answered with "yes," "no" or "unclear." The quality of the included articles was evaluated by three researchers independently, and the discrepancies were resolved by consensus after a comprehensive discussion. The PRISMA checklist for systematic reviews is presented in Supplementary Table S2.

3 Results

3.1 Study selection

A total of 18,031 records were retrieved from the initial database search, of which 1,330 duplicates were removed by electronic and manual double examination, obtaining a total of 16,701 articles. These were screened by titles and abstracts, excluding 16,507 for being unrelated to the topic of interest. The full text of the remaining 194 articles was fully assessed for eligibility, and finally, 26 articles fulfilled the inclusion criteria. Figure 1 displays the search results and the selection strategy.

3.2 Study characteristics

The characteristics of the selected studies are summarized in Table 1. In the selected studies, a total of 2004 individuals were enrolled, including 1,338 patients diagnosed with SSc, most of them within the lcSSc subtype and 666 HC. Most patients diagnosed with SSc included in the studies were women (87.9%). Patients had an average age of 56.3 years. Most of the selected articles were analytical cross-sectional studies (n = 23). Twenty-three of the twenty-six studies conducted metabolomics comparisons between SSc cases and HC, two studies compared metabolomic patterns between SSc patients with PAH and SSc patients without PAH (Thakkar et al., 2016; Deidda et al., 2017) and one study compared SSc cases with systemic lupus erythematosus (LES) (Bengtsson et al., 2016). Fourteen studies assessed the metabolome in plasma, eight in serum, two in urine, one in urine and plasma, and one study evaluated in exhaled breath. Most of these studies used high-performance liquid chromatography quadrupole time-of-flight mass spectrometry (HPLC-Q-TOF-MS), high-performance liquid chromatography with fluorescence detection (HPLC-FLD); and to a lesser extent high-performance liquid chromatography quadrupole-linear ion-trap hybrid mass spectrometry (HPLC-QTRAP-MS), high-performance liquid chromatography with on-line UV system. (HPLC-UV) and gas chromatography electron impact mass spectrometry (GC-EI-MS). Proton nuclear magnetic resonance spectroscopy (H NMR), and capillary gas chromatography with flame ionization detection (CGC-FID) were also described.

3.3 Quality assessment

The results of the methodologic quality assessment by the QUADOMICS tool are summarized in Supplementary Table S3. Because all the studies included in this evaluation were in phase I, the second and 14th QUADOMICS items were not applicable. Overall, the studies meet the majority of the QUADOMICS criterion, indicating that the quality of the included studies is good. All studies described the selection criteria (item 1) and the sample type (item 3), but none met item 12, indicating that the index test findings were interpreted with knowledge of the reference standard. Twenty of the 26 studies found fully comparable data between SSc patients and HC in terms of crucial characteristics including gender and age. Most studies did not avoid overfitting due to the lack of an independent validation set.

3.4 Metabolites and metabolic pathways associated with systemic sclerosis

In total, 151 altered metabolites were identified in the selected studies. Since the expression of the metabolites can be influenced by the sample used (Kaluarachchi et al., 2018; Lau et al., 2018) the altered metabolites were detailed according to each fluid.

3.4.1 Plasma

Fourteen studies assessed the metabolome in plasma samples of SSc patients (Caramaschi et al., 2007; Caramaschi et al., 2003; Tikly et al., 2006; Szamosi et al., 2009; McNearney et al., 2010; ichiro et al., 2014; Atteritano et al., 2016; Bellocchi et al., 2018; Ottria et al., 2020; Smolenska et al., 2020; Bögl et al., 2022; Geroldinger-Simić et al., 2021; Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Neumann Andersen et al., 2000). The greatest alterations found are grouped into two chemical classes: amino acids and lipids.

According to amino acids, four studies evaluated plasma Hcy levels in SSc patients vs. HC (Caramaschi et al., 2007; Caramaschi et al., 2003; Szamosi et al., 2009; ichiro et al., 2014). Consistently

TABLE 1 Characteristics of studies included in the systematic review.

Sample type	Sample size	Gender (M/F)	Age average (years)	Analytical technique	Altered metabolites in SSc vs. HC	Study
Plasma	SSc: 27	4/23	56.9	GC-MS[a]	*Increased:* Nitrate	Neumann Andersen et al. (2000)
	Controls: 27				*Decreased:* NR	
	SSc: 71	10/61	58.5	HPLC-FLD	*Increased:* Homocysteine	Caramaschi et al. (2003)
	Controls: 30				*Decreased:* NR	
	SSc: 15	NR	42.2	GC-MS[a]	*Increased:* Malondialdehyde	Tikly et al. (2006)
	Controls: 13				*Decreased:* NR	
	SSc: 60	4/56	54.6	HPLC-FLD	*Increased:* Homocysteine	Caramaschi et al. (2007)
	Controls: 30				*Decreased:* NR	
	SSc: 40	38/2	58.4	HPLC-MS[a]	*Increased:* NR	Atteritano et al. (2016)
	Controls: 40				*Decreased:* Vitamin D	
	SSc: 59	7/52	56.5	UHPLC-Q-TOF-MS	*Increased:* DL-2-aminooctanoic acid, Diacylglycerol 38:5, 1-(9Z-pentadecenoyl)-glycero-3-phosphate, phosphatidylcholine 36:4, 2,4-dinitrobenzenesulfonic acid, alpha-N-phenylacetyl-l-glutamine	Bellocchi et al. (2018)
	Controls: 28				*Decreased:* NR	
	SSc: 20	3/17	57	UHPLC-Orbitrap-MS	*Increased:* Lauric acid, myristic acid, arachidic acid, carnitine, isovaleryl-carnitine	Ottria et al. (2020)
	Controls: 7				*Decreased:* Octanoyl-carnitine, palmitoyl-carnitine	
	SSc: 42	7/35	59.9	HPLC-TQ-MS	*Increased:* Glutamine, proline, 1-methylhistidine, betaine, methylnicotinamide, asymmetric dimethylarginine	Smolenska et al. (2020)
	Controls: 27				*Decreased:* Tryptophan	
	SSc: 52	8/44	60	HPLC-IM-Q-TOF-MS	*Increased:* Phosphatidylcholine 34:1, 34:2, 34:3; sphingomyelin 33:1, 35:1, 35:2	Geroldinger-Simić et al. (2021)
	Controls: 48				*Decreased:* NR	
	SSc: 52	8/44	60	HPLC-IM-Q-TOF-MS	*Increased:* Kynurenine, dimethylarginine, citrulline, ornithine, phenylacetylglutamine, 1-methylhistidine, 3-methylhistidine	Bögl et al. (2022)
	Controls: 48				*Decreased:* Tryptophan, OH-tryptophan, alanine, lysophosphatidylcholine 22:4a, 22:4b, 20:2; sphingomyelin 34:1, 40:3	
	SSc: 59	7/52	56.5	HPLC-ESI-QTOF-MS	*Increased:* Alpha-N-phenyl acetyl-L-glutamine, butyrylcarnitine, valerylcarnitine, 2-4-dinitrobenzenesulfonic acid, oleic acid, 1-arachidonoylglycerol monoacylglycerol (20:4), monoacylglycerol (20:5)	Fernández-Ochoa ÁQuirantes-Piné et al. (2019)
	Controls: 28				*Decreased:* NR	

(Continued on following page)

10.3389/fmolb.2023.1215039

TABLE 1 (*Continued*) Characteristics of studies included in the systematic review.

Sample type	Sample size	Gender (M/F)	Age average (years)	Analytical technique	Altered metabolites in SSc vs. HC	Study
Serum	SSc: 10	0/10	47	HPLC-QTRAP-MS	*Increased:* Arachidonoyl-lysophosphatidic acid, sphingosine 1-phosphate	Tokumura et al. (2009)
	Controls: 13				*Decreased*: NR	
	SSc: 68	0/68	67.6	HPLC-TQ-MS	*Increased:* 17β-estradiol, estrone	Aida-Yasuoka et al. (2013)
	Controls: 35				*Decreased*: NR	
	SSc: 19	3/16	55	GC-TOF-MS	*Increased:* Aminomalonic acid, arachidonic acid, arginine, aspartic acid, beta-alanine, cholesterol, inositol-1-phosphate, lauric acid, oleamide, ornithine-1,5-lactam, picolinic Acid, pyroglutamic acid, ribose, succinic Acid, urea, uric acid	Bengtsson et al. (2016)
	Controls: 18				*Decreased*: Alanine, cysteine, lactic acid, malic acid, nonanoic acid, taurine, threonic acid	
	SSc: 37	8/29	58.7	H-NMRS GC-MS[a]	*Increased:* Glutamine, 3-OH-butyrate	Murgia et al. (2018)
	Controls: 20				*Decreased:* Citrate, aspartate, alanine, choline, glutamate, glutarate, glycerate, threonate	
	SSc: 97	16/81	59	HPLC-FLD	*Increased:* Kynurenine	Campochiaro et al. (2019)
	Controls: 10				*Decreased:* Tryptophan	
	SSc: 36	6/30	61.5	UHPLC-Q-TOF-MS	*Increased:* 1-methyladenosine	Meier et al. (2020)
	Controls: 12				*Decreased*: L-tryptophan, L-tyrosine	
	SSc: 30	6/24	46.3	UHPLC-Q-TOF-MS	*Increased:* Vitamin E, alpha-N-phenylacetyl-L-glutamine, L-glutamine, L-isoleucine, phenol, 2-oxoadipic acid, 1-palmitoyl-2-hydroxy-sn, glycero-3-phosphoethanolamine, chenodeoxycholate, indoxyl sulfate, D-quinovose	Sun et al. (2022)
	Controls: 30				*Decreased:* 3b-hydroxy-5-cholenoic acid, 1-stearoyl-glycerol, trans-dehydroandrosterone, 4-nonylphenol, norethindrone acetate, cis-9,10-epoxystearic acid, 16-hydroxypalmitic acid, 2-ethyl-2-hydroxybutyric acid, stearic acid, hexadecanedioic acid, 3 hydroxy caproic acid, androsterone sulfate, benzenebutanoic acid, pregnenolone sulfate, arachidonic acid, dodecanoic acid, palmitic acid, myristic acid, cholesterol 3-sulfate, caprylic acid, Cis-(6,9,12)-linolenic acid, alpha-ketocaproic acid, azelaic acid	

(Continued on following page)

10.3389/fmolb.2023.1215039

TABLE 1 (*Continued*) Characteristics of studies included in the systematic review.

Sample type	Sample size	Gender (M/F)	Age average (years)	Analytical technique	Altered metabolites in SSc vs. HC	Study
Urine	SSc: 59	7/52	56.5	HPLC-Q-TOF-MS	*Increased:* D-Sorbitol, N-cyclohexylformamide, Ser-Pro-Pro, dihydroxy-1H-indole glucuronide, 2-(2-phenylacetoxy)propinylglycine, alpha-N-phenylacetyl—L glutamine, pyroglutamic acid	Fernández-Ochoa ÁQuirantes-Piné et al. (2019)
	Controls: 28				*Decreased*: N-Methylnicotinamide, proline betaine, creatinine, vinylacetylglycine, N1-methyl-4pyridine-3-carboxamide, N1-methyl-2-pyridine-5-carboxamide, hydroxyprolyl-valine, L-beta-aspartyl-L-Leucine, Hypaphorine, 2-octenoyl-carnitine, decatrienoylcarnitine, 2-nonenoylcarnitine, 2,6-dimethylheptanoyl carnitine, 9-decenoylcarnitine, 9-hydroxydodecenoylcarnitine, undecenoyl carnitine	
	SSc: 11	0/11	51	GC-EI-MS	*Increased:* 15-F-2t-isoprostane	Cracowski et al. (2002)
	Controls: 11				*Decreased*: NR	
	SSc: 43	1/42	54.1	HPLC-UV	*Increased:* 8-isoprostaglandin-F2a	Volpe et al. (2006)
	Controls: 43				*Decreased*: NR	
Exhaled breath	SSc: 46	NR	54	CGC-FID	*Increased:* Ethane	Cope et al. (2006)
	Controls: 21				*Decreased*: Ethanol	

[a]Detector is not specified in the document.

Increased and decreased metabolites shown in the table correspond to altered metabolites in SSc, patients compared to HC. Abbreviations: SSc, Systemic sclerosis; GC-MS, Gas chromatography-mass spectrometry; HPLC-FLD, High-performance liquid chromatography with fluorescence detection; UHPLC-Q-TOF-MS, Ultra-high-performance liquid chromatography quadrupole time-of-flight mass spectrometry; UHPLC-Orbitrap-MS, Ultra-high-performance liquid chromatography coupled with ion trap mass spectrometry; HPLC-TQ-MS, High-performance liquid chromatography coupled to triple-stage quadrupole mass spectrometer; HPLC-IM-Q-TOF-MS, High-performance liquid chromatography coupled to ion mobility quadrupole time-of-flight mass spectrometry; HPLC-ESI-QTOF-MS, High-performance liquid chromatography coupled to electrospray ionization and quadrupole time-of-flight mass spectrometry; HPLC-QTRAP-MS, High-performance liquid chromatography quadrupole-linear ion trap hybrid mass spectrometry; GC-TOF-MS, Gas chromatography time-of-flight mass spectrometry; H NMRS, proton nuclear magnetic resonance spectrometry; HPLC-QTOF-MS, High-performance liquid chromatography quadrupole time-of-flight mass spectrometry; GC-EI-MS, gas chromatography electron impact mass spectrometry; HPLC-UV, High-performance liquid chromatography with on-line UV, system; CGC-FID, capillary gas chromatography flame ionization detection; NR, Not Reported.

10.3389/fmolb.2023.1215039

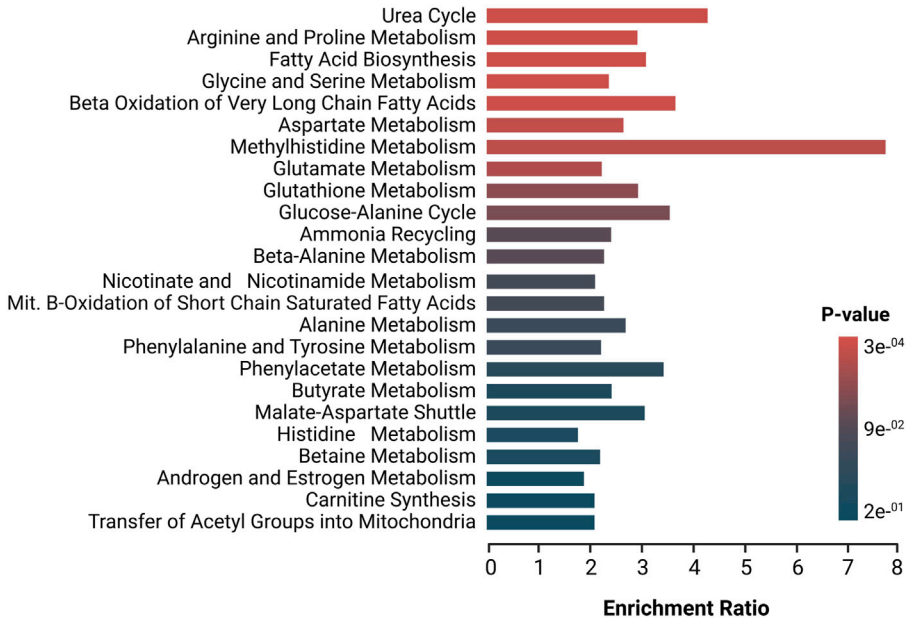


FIGURE 2
Enrichment analysis of altered pathways in SSc, The significance of pathway alteration is indicated according to the color scale. The bars in red and blue represent the biosynthetic pathways of greater and lesser impact, respectively.

increased levels of Hcy were found across the studies; however, two of them did not find significant differences in Hcy levels in patients with SSc compared to HC, yet, researchers did find significant differences in the concentrations in patients with vascular and thromboembolic manifestations (Szamosi et al., 2009; ichiro et al., 2014). One study looked at endogenous enkephalin levels in early SSc; nevertheless, they found no significant changes in enkephalin levels between SSc and HC patients, although they did find that low levels were associated with Raynaud syndrome, myositis, and telangiectasias (McNearney et al., 2010). ADMA, Kyn, 1-Methylhistidine, alpha-N-phenylacetyl-L-glutamine, glutamine, proline, citrulline, and ornithine were consistently increased across the studies (Bellocchi et al., 2018; Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Smolenska et al., 2020; Bögl et al., 2022). On the contrary, Trp and alanine were found with downward trends (Smolenska et al., 2020; Bögl et al., 2022).

Lipid content was represented by the classes carnitines, fatty acids (FA), glycerolipids, glycerophospholipids, sphingolipids, and steroids. Short-chain carnitines such as carnitine, butyrylcarnitine, and valerylcarnitine were increased across the studies (Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Ottria et al., 2020). On the contrary, acylcarnitines associated with long-chain fatty acids: octanoyl-carnitine, and palmitoyl-carnitine, were observed with downward trends in SSc patients when compared to HC (Ottria et al., 2020) Regarding FA, high levels of saturated (e.g., lauric acid, myristic acid, and arachidic acid) and unsaturated FA (e.g., oleic acid) were observed (Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Ottria et al., 2020) Similarly, trends in other lipid metabolites were consistently observed. For example, glycerolipids such as DG 38:5, MG 20:4, and MG 20:5 were consistently elevated (Bellocchi et al., 2018; Fernández-Ochoa ÁQuirantes-Piné et al., 2019), yet trends observed in glycerophospholipids, and sphingolipids were not uniform (Bellocchi et al., 2018; Geroldinger-Simić et al., 2021; Bögl et al., 2022). For example, metabolites such as 1-(9Z-pentadecenoyl)-glycero-3-phosphate, phosphatidylcholine 34:1, 34:2, 34:3, 36:4, and sphingomyelin 33:1, 35:1, 35:2 were found to be increased (Bellocchi et al., 2018; Geroldinger-Simić et al., 2021), while lysophosphatidylcholine 22:4, 22:4, 20:2, sphingomyelin 34:1, 40:3 were decreased (Bögl et al., 2022).

Lastly, it was found that vitamin D levels were decreased in patients with SSc, and this deficiency was linked to scleroderma and increased systolic pulmonary artery pressure (Atteritano et al., 2016).

3.4.2 Serum

Seven studies assessed the metabolome in serum SSc patients compared to HC (Tokumura et al., 2009; Aida-Yasuoka et al., 2013; Bengtsson et al., 2016; Murgia et al., 2018; Campochiaro et al., 2019; Meier et al., 2020; Sun et al., 2022). The most significant changes identified in serum are classified into amino acids, lipids, and tricarboxylic acids.

Trends of some altered amino acids in plasma are conserved in serum samples. For example, increased levels of Kyn, alpha-N-phenylacetyl-L-glutamine, glutamine, and ornithine were observed in both fluids (Bengtsson et al., 2016; Murgia et al., 2018; Campochiaro et al., 2019; Sun et al., 2022), meanwhile, Trp, glutamate, and alanine levels remained decreased in serum and plasma (Bengtsson et al., 2016; Murgia et al., 2018; Campochiaro et al., 2019; Meier et al., 2020). Discrepancies were found in levels of aspartic acid. Bengtsson et al. (2016) reported increased levels of aspartic acid in SSc patients when compared to LES patients; however, decreased levels of these metabolites were also observed (Murgia et al., 2018).

The altered lipid content in serum is diverse. In the case of FA and steroids, no clear trends were observed. Some FA, such as lauric acid, 3-OH-butyrate, cholesterol, and chenodeoxycholate, were found to be increased (Bengtsson et al., 2016; Murgia et al., 2018; Sun et al., 2022). On the contrary, FA such as nonanoic acid, azelaic acid, cis-9,10-epoxystearic acid, 16-hydroxypalmitic acid, 2-ethyl-2-hydroxybutyric acid, stearic acid, hexadecanedioic acid, hydroxycaproic acid, dodecanoic acid, palmitic acid, myristic acid and caprylic acid, cis-(6,9,12)-linolenic acid were downregulated (Bengtsson et al., 2016; Sun et al., 2022). Other lipid compounds, such as arachidonoyl-lysophosphatidic acid, were found to be increased in serum from SSc patients (Tokumura et al., 2009). Regarding steroids, 17β-estradiol, estrone, and vitamin E were found to be increased (Aida-Yasuoka et al., 2013; Sun et al., 2022), while androsterone sulfate, pregnenolone sulfate, and cholesterol 3-sulfate were decreased (Sun et al., 2022).

Metabolites indicating altered energy metabolism were identified; however, trends were not consistent. Increased levels of tricarboxylic acid cycle (TCA) metabolites such as succinate were observed (Bengtsson et al., 2016), while citrate and malate levels were found to be decreased (Bengtsson et al., 2016; Murgia et al., 2018). Increased levels of other metabolites such as purines, pyrimidines, and carbohydrates such as uric acid, 1-methyladenosine, picolinic acid, ribose, and D-quinovose were observed (Bengtsson et al., 2016; Meier et al., 2020; Sun et al., 2022) as well as low levels of lactate, choline, taurine, threonate and glycerate in SSc patients (Bengtsson et al., 2016; Murgia et al., 2018).

3.4.3 Urine

Three studies assessed the metabolome in urine (Cracowski et al., 2002; Volpe et al., 2006; Fernández-Ochoa ÁQuirantes-Piné et al., 2019). The most significant changes found in urine samples are chemically categorized into amino acids and carnitines. Urinary levels of alpha-N-phenylacetyl-L-glutamine were found to be increased, as they were in plasma and serum samples (Fernández-Ochoa ÁQuirantes-Piné et al., 2019). Additionally, pyroglutamic acid was upregulated in urine samples in concordance with serum samples (Fernández-Ochoa ÁQuirantes-Piné et al., 2019). Consistent with the results in plasma, urine samples reported decreased levels of acylcarnitines associated with long-chain fatty acids in SSc patients compared to HC (Fernández-Ochoa ÁQuirantes-Piné et al., 2019). Lastly, high levels of 8-isoprostaglandin-F2a were found in urine samples of SSc patients, which were related to more severe lung involvement and active patterns in nailfold video capillaroscopy (Volpe et al., 2006).

3.4.4 Exhaled breath

One study evaluated metabolites in exhaled breath of SSc patients (Cope et al., 2006). They found high breath ethane concentrations which were inversely associated with the diffusing capacity for carbon monoxide, and decreased levels of ethanol concentrations, compared to HC.

3.4.5 Metabolic pathways

Deregulated metabolites were imported to the MetaboAnalyst platform for the generation of metabolic pathway analyses and SMPDB metabolic pathways were used as a library of metabolite clusters. Figure 2 depicts the altered metabolic pathways in patients with SSc versus HC. Based on the hypergeometric *p*-value test, this software shows whether a metabolic pathway is more strongly represented in the list of compounds. The pathways in red represent the most significant deregulated pathways in SSc patients, while the pathways in blue represent the least significant deregulated pathways in these patients. Several of the amino acids that were found to be deregulated participate in the metabolism of the different pathways that were found to be significantly enriched. As for the urea cycle, several amino acids involved in this cycle, such as glutamic acid, alanine, aspartic acid, ornithine, arginine, urea, glutamine, and citrulline, were found to deregulate, as well as various amino acids involved in the arginine and proline metabolism (e.g., creatinine, glycine, glutamic acid, proline, aspartic acid, ornithine, succinic acid, urea, arginine, and citrulline) several who also participate in the glycine and serine metabolism such as betaine, glycine, alanine, Hcy, arginine, and glyceric acid making them part of the most significant enriched pathways. Similarly, in FA biosynthesis and beta-oxidation of very long-chain FA, several deregulated FA and carnitines were found to be involved in these two pathways (e.g., butyric acid, caprylic acid, myristic acid, dodecanoic acid, caproic acid, and L-carnitine).

3.5 Altered metabolism in lcSSc and dcSSc subtypes

Deregulated metabolites in dcSSc versus lcSSc subtypes and HC are described in Table 2. Overall, 7 of the 26 selected studies conducted metabolomics comparisons between dcSSc and lcSSc subtypes in order to discriminate them (Tokumura et al., 2009; Murgia et al., 2018; Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Smolenska et al., 2020; Geroldinger-Simić et al., 2021; Bögl et al., 2022; Sun et al., 2022). Of these, one study evaluated serum samples from SSc subtypes and compared them to HC, finding that levels of sphingosine 1-phosphate were significantly increased in dcSSc patients versus HC (Tokumura et al., 2009). In serum samples from dcSSc patients compared to lcSSc patients, levels of several amino acids, such as valine, glutamate, lysine, and betaine, such as valine, glutamate, lysine, and betaine, carbohydrates, including fructose, glycerol, and glycerate and carboxylic acids, such as acetate and glutarate were found to be significantly increased (Murgia et al., 2018; Sun et al., 2022). On the contrary, levels of glutamine, lactate, and glucose were significantly decreased (Murgia et al., 2018), as well as glycerophosphocholines (Sun et al., 2022). As for plasma samples, dcSSc patients also had higher concentrations of various amino acids and derivatives and phosphatidylcholine species (e.g., Kynurenine, citrulline, ornithine, N(G)-nitro-L-arginine methyl ester (L-NAME), beta-alanine, Phosphatidylcholine 32:0), and lower levels of sphingomyelins and glycerophosphoethanolamines (Smolenska et al., 2020; Geroldinger-Simić et al., 2021; Bögl et al., 2022). On the other hand, urinary metabolites detected in dcSSc patients showed increased levels of amino acids, including L-arogenate and indospicine and N(5-amino-2hydroxybenzoyl)glycine, and decreased levels of carnitines and decreased levels of carnitines in comparison to the lcSSc subtype (Fernández-Ochoa ÁQuirantes-Piné et al., 2019).

TABLE 2 Deregulated metabolites in dcSSc vs. HC and lcSSc subtypes.

Sample type	Sample size	Analytical technique	Altered metabolites in dcSSc vs. lcSSc	Study
Serum	dcSSc: 7	HPLC-QTRAP-MS	*Increased:* Sphingosine 1-phosphate	Tokumura et al. (2009)
	lcSSc: 3		*Decreased:* NR	
	Controls: 13			
	dcSSc: 14	H-NMRS GC- MS[a]	*Increased:* Valine, acetate, fructose, glutamate, glycerol, lysine, glycerate, glutarate	Murgia et al. (2018)
	lcSSc: 23		*Decreased:* Sorbitol, glucose, lactate, glutamine	
	dcSSc: 12	UHPLC-Q-TOF-MS	*Increased:* Trans-dehydroandrosterone, betaine, 1-stearoyl-2-oleoyl-sn-glycerol 3-phosphocholine	Sun et al. (2022)
	lcSSc: 18		*Decreased:* 1-palmitoyl-sn-glycero-3-phosphocholine	
Plasma	dcSSc: 21	HPLC-TQ-MS	*Increased:* Sarcosine, beta-alanine, methylnicotinamide, N(G)-nitro-L-arginine methyl ester (L-NAME)	Smolenska et al. (2020)
	lcSSc: 21		*Decreased:* NR	
	dcSSc: 11	HPLC-IM-Q-TOF-MS	*Increased:* Phosphatidylcholine 32:0	Geroldinger-Simić et al. (2021)
	lcSSc: 39		*Decreased:* Phosphatidylethanolamine 38:5, 38:6; sphingomyelin 32:2, 40:4, 30:1	
	dcSSc: 11	HPLC-IM-Q-TOF-MS	*Increased:* Kynurenine, citrulline, ornithine, phenylacetylglutamine	Bögl et al. (2022)
	lcSSc: 39		*Decreased:* Tryptophan, lysophosphatidylcholine 22:4	
Urine	dcSSc: 10	HPLC-ESI-QTOF-MS	*Increased:* L-arogenate, N (5-amino-2hydroxybenzoyl)glycine, indospicine	Fernández-Ochoa ÁQuirantes-Piné et al. (2019)
	lcSSc: 43		*Decreased:* 3-methylglutarylcarnitine, 5-hydroxyindoleacetic acid	

[a]Detector is not specified in the document.

Increased and decreased metabolites shown in the table correspond to altered metabolites in dcSSc, patients compared lcSSc, or HC. Abbreviations; dcSSc, diffuse cutaneous systemic sclerosis; lcSSc, limited cutaneous systemic sclerosis; HPLC-QTRAP-MS, High-performance liquid chromatography quadrupole-linear ion trap hybrid mass spectrometry; H NMRS, proton nuclear magnetic resonance spectrometry; GC-MS, Gas chromatography-mass spectrometry; UHLPC-Q-TOF-MS, Ultra-high-performance liquid chromatography quadrupole time-of-flight mass spectrometry; HPLC-TQ-MS, High-performance liquid chromatography coupled to triple-stage quadrupole mass spectrometer; HPLC-IM-Q-TOF-MS, High-performance liquid chromatography coupled to ion mobility quadrupole time-of-flight mass spectrometry; HPLC-ESI-QTOF-MS, high-performance liquid chromatography coupled to electrospray ionization and quadrupole time-of-flight mass spectrometry; NR, Not reported.

In order to further understand these alterations, deregulated metabolites in patients with dcSSc subtype were imported to MetaboAnalyst 5.0 platform for the generation of metabolic pathways analysis, identifying relevant pathways associated with the development of this subtype, using the high-quality SMPDB metabolic pathways as the metabolite set library (Figure 3). Pathways in red represent the most significant deregulated pathways in dcSSc patients. In contrast, the pathways presented in blue represent the least significant deregulated pathways in these patients. The biosynthetic pathways with the greatest impact were the aspartate metabolism, urea cycle, amino sugar metabolism, glycine and serine metabolism, and phenylacetate metabolism. These pathways are associated with alterations in amino acids (e.g., aspartate metabolism, urea cycle, glycine and serine metabolism, and phenylacetate metabolism) with several deregulated amino acids such as beta-alanine, glutamic acid, glutamine, citrulline, ornithine, betaine, and alpha-N-Phenylacetyl-L-glutamine. On the other hand, carbohydrates such as fructose, glyceric acid, and carboxylic acids (e.g., acetic acid) were also involved.

3.6 Metabolites associated with pulmonary complications in SSc patients

Of the selected studies, nine evaluated the alterations in the metabolomic profile of pulmonary complications in SSc patients, particularly PAH (Thakkar et al., 2016; Deidda et al., 2017) and ILD (Caramaschi et al., 2003; ichiro et al., 2014; Smolenska et al., 2020; Geroldinger-Simić et al., 2021; Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Sun et al., 2022; Meier et al., 2020). Table 3 describes the identified metabolites associated with pulmonary complications in SSc patients Thakkar et al. (2016) found ADMA levels to be significantly higher and L-arginine levels were significantly lower in SSc-PAH compared with Non-PAH patients. Additionally, studies found that SSc-PAH patients had higher amounts of carboxylic acids (e.g., lactate), and lipoproteins, and lower levels of amino acids, notably L-arginine, in comparison to SSc Non-PAH patients (Deidda et al., 2017). In terms of altered metabolites associated with ILD patients, increased levels of amino acids, such as Hcy, arginine, and valine, and fructosamines derived from branch-chain amino acids were discovered, while lower levels of

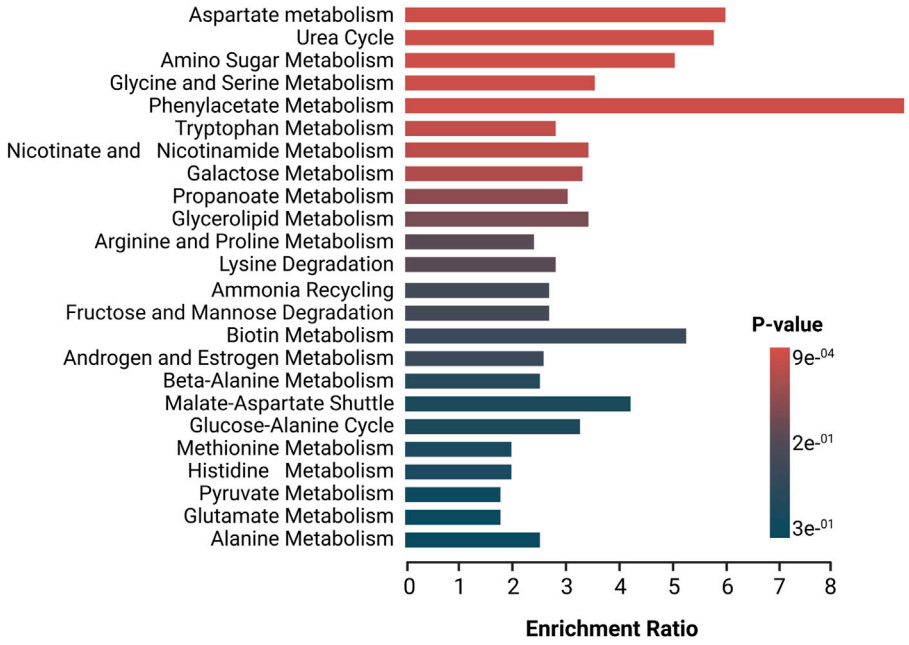


FIGURE 3
Enrichment analysis of altered pathways in dcSSc. The significance of pathway alteration is indicated according to the color scale. The bars in red and blue represent the biosynthetic pathways of greater and lesser impact, respectively.

glycerophosphoethanolamines (e.g., phosphatidylethanolamine 36:3, 38:5, 38:6) and steroids, including androsterone sulfate were found compared to non-ILD (Caramaschi et al., 2003; ichiro et al., 2014; Smolenska et al., 2020; Geroldinger-Simić et al., 2021; Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Sun et al., 2022). One study compared progressive ILD versus stable ILD patients finding increased levels of branched-chain amino acids (BCAAs) and one purine (xanthosine) and decreased levels of adenosine monophosphate in patients with progressive ILD (Meier et al., 2020). Lastly, one study compared urine samples between non-ILD patients and ILD patients, discovering that the latter group also had increased levels of amino acids (e.g., valyl valine, kynurenic acid, L-proline, proline-histidine) (Fernández-Ochoa ÁQuirantes-Piné et al., 2019).

3.7 Metabolites as potential biomarkers

Potential biomarkers identified in the review are described in Table 4. Of all selected articles, five studies assessed the diagnostic capability of biomarkers using AUC, reporting at least one biomarker with an AUC > 0.7 (Thakkar et al., 2016; Murgia et al., 2018; Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Meier et al., 2020; Sun et al., 2022). Of these, three studies evaluated the potential of metabolic biomarkers or panels to diagnose SSc (Murgia et al., 2018; Fernández-Ochoa ÁQuirantes-Piné et al., 2019; Meier et al., 2020), and one study examined diagnostic biomarkers for PAH (Thakkar et al., 2016) finding that serum ADMA levels ≥0.7 μM in PAH patients had a sensitivity of 86.7% and a specificity of 90.0%, two studies evaluated biomarkers for the classification of dcSSc subtype (Murgia et al., 2018; Sun et al., 2022), one study evaluated biomarkers for ILD (Sun et al., 2022), and one study evaluated biomarkers to distinguish progressive SSc-ILD from stable SSc-ILD (Meier et al., 2020). The latter validated their results using an enzymatic assay, obtaining similar results with significantly higher values detected in progressive SSc-ILD patients compared to stable SSc-ILD. These results were also found in another cohort of SSc- ILD patients (Meier et al., 2020).

3.8 Targeted metabolomics analysis in patients with systemic sclerosis

A total of 10 studies utilized targeted techniques to assess metabolite levels, primarily focusing on amino acids, in patients with SSc (Table 5). In plasma, SSc patients exhibited higher concentrations of metabolites such as Hcy, glutamine, proline, 1-methylhistidine, ADMA, betaine, malondialdehyde, and methylnicotinamide when compared to healthy controls (Caramaschi et al., 2003; Tikly et al., 2006; Caramaschi et al., 2007; Smolenska et al., 2020) Conversely, vitamin D and Trp showed lower concentrations in SSc patients, revealing a distinct trend compared to healthy individuals (Atteritano et al., 2016; Smolenska et al., 2020). Additionally, studies conducted on serum found elevated concentrations of arachidonoyl (20:4)-LPA in SSc patients compared to healthy individuals (Tokumura et al., 2009). Furthermore, there were increased levels of ADMA and symmetric dimethylarginine and decreased levels of arginine observed in SSc-PAH patients compared to SSc Non-PAH individuals (Thakkar et al., 2016). Lastly, urine and exhaled breath investigations demonstrated higher concentrations of 15-

TABLE 3 Deregulated metabolites associated with pulmonary complications in SSc patients.

Pulmonary complications	Sample type	Sample size	Analytical technique	Altered metabolites	Study
PAH	Serum	SSc- PAH: 15	HPLC-FLD	*Increased:* Asymmetric dimethylarginine, symmetric dimethylarginine	Thakkar et al. (2016)
		SSc Non-PAH: 30		*Decreased:* L-Arginine	
	Plasma	SSc- PAH: 8	H NMRS	*Increased:* Acetoacetate, Alanine, Lactate, VLDL, LDL	Deidda et al. (2017)
		SSc Non-PAH: 10		*Decreased:* γ-Aminobutyrate, arginine, betaine, choline, creatinine, glucose, glutamate, glycine, histidine, phenylalanine, tyrosine	
ILD	Plasma	SSc-ILD: 65	HPLC-FLD	*Increased:* Homocysteine	ichiro et al. (2014)
		SSc Non-ILD: 151		*Decreased:* NR	
		SSc-ILD: 62	HPLC-FLD	*Increased:* Homocysteine	Caramaschi et al. (2003)
		SSc Non-ILD: 9		*Decreased:* NR	
		SSc-ILD: 18	HPLC-Q-TOF-MS	*Increased:* N-(1-deoxy-1-fructosyl)-Valine, N-(1-deoxy-1-fructosyl)-leucine, N-(1-deoxy-1-fructosyl)-Isoleucine	Fernández-Ochoa ÁQuirantes-Piné et al. (2019)
		SSc Non-ILD: 41		*Decreased:* NR	
		SSc-ILD: 26	HPLC-TQ-MS	*Increased:* Valine, Arginine	Smolenska et al. (2020)
		SSc Non-ILD: 16		*Decreased:* NR	
		SSc-ILD: 14	HPLC-IM-Q-TOF-MS	*Increased:* NR	Geroldinger-Simić et al. (2021)
		SSc Non-ILD: 38		*Decreased:* Phosphatidylethanolamine 36:3, 38:5, 38:6	
	Serum	Stable SSc-ILD: 12	UHPLC-Q-TOF-MS	*Increased:* L-Leucine, L-Isoleucine, Xanthosine	Meier et al. (2020)
		Progressive SSc-ILD: 12		*Decreased:* Adenosine monophosphate	
		SSc-ILD: 19	UHPLC-Q-TOF-MS	*Increased:* L-Glutamine	Sun et al. (2022)
		SSc Non-ILD: 11		*Decreased:* Ile-Ala, Androsterone sulfate	
	Urine	SSc-ILD: 18	HPLC-Q-TOF-MS	*Increased:* Valyl valine, kynurenic acid, L-proline, proline-histidine, quinolinic acid, β-D-glucopyrapyranosil anthranilate	Fernández-Ochoa ÁQuirantes-Piné et al. (2019)
		SSc Non-ILD: 41		*Decreased:* NR	

Increased and decreased metabolites shown in the table correspond to altered metabolites in PAH, patients compared to non-PAH, patients and in ILD, patients compared to non-ILD, patients. Abbreviations: PAH, pulmonary arterial hypertension; ILD, interstitial lung disease; HPLC-SPE, High-performance liquid chromatography with solid phase extraction; H NMRS, proton nuclear magnetic resonance spectrometry; HPLC-FLD, High-performance liquid chromatography with fluorescence detection; HPLC-Q-TOF-MS, High-performance liquid chromatography quadrupole time-of-flight mass spectrometry; HPLC-MS, High-performance liquid chromatography-mass spectrometry; HPLC-IM-Q-TOF-MS, high-performance liquid chromatography quadrupole time-of-flight mass spectrometry; UHPLC-MS, ultra-high-performance liquid chromatography-mass spectrometry; UHPLC-Q-TOF-MS, ultra-high-performance liquid chromatography quadrupole time-of-flight mass spectrometry; NR, not reported.

F-2t-isoprostane, 8-isoprostaglandin-F2a, and ethane, while the ethanol concentration was reduced in SSc patients (Cracowski et al., 2002; Cope et al., 2006; Volpe et al., 2006). However, comparing metabolites across studies can be complex due to variations in cohorts, analytical techniques, and statistical analysis. Higher concentrations of Hy (11.1–11.8 μmol/L) were observed in the plasma of SSc patients compared to healthy individuals (3.5–6.9 μmol/L), with an approximate increase of 45.41% in Hcy found in SSc patients (Caramaschi et al., 2003; Caramaschi et al., 2007).

4 Discussion

In this study, we systematically reviewed 26 studies on the metabolomic profiling of SSc and summarized key findings on the dysregulation of major metabolic pathways in SSc, primarily amino acid-related pathways, lipid metabolism, and the TCA cycle. To our knowledge, this is the first systematic review of metabolomic analysis in SSc.

Altered amino acid metabolism was a common finding in analyzed samples of SSc patients, possibly associated with protein

TABLE 4 Potential metabolite biomarkers for diagnosis and classification of SSc.

	Sample type	Analytical technique	Metabolites	ROC curve	Validation	Study
SSc	Serum	H-NMRS	Aspartate	AUC: 0.81 (CL 95% 0.7–0.93)	NR	Murgia et al. (2018)
		GC-MS[a]	Alanine Citrate			
		UHPLC-Q-TOF-MS	L-tryptophan	AUC 0.884 (CI 95% 0.788–0.981)	NR	Meier et al. (2020)
			1-methyl-adenosine	AUC 0.822 (CI 95% 0.705–0.939)		
			L-tyrosine	AUC 0.812 (CI 95% 0.667–0.958)		
	Urine	HPLC-Q-TOF -MS	N1-methyl-4-pyridine-3-carboxamide	AUC: 0.818 (CI 95% 0.709–0.903	NR	Fernández-Ochoa ÁQuirantes-Piné et al. (2019)
			N1-methyl-2-pyridine-5-carboxamide	AUC 0.766 (CI 95% 0.674–0.858)		
			D-sorbitol	AUC: 0.802 (CI 95% 0.688–0.881)		
			2,6 Dimethyl-heptonoylcarnitine	AUC: 0.778 (CI 95% 0.67–0.86)		
	Plasma	HPLC-Q-TOF -MS	Alpha- N-phenylacetyl-L-glutamine	AUC. 0.766 (CI 95% 0.656–0.86)	NR	Fernández-Ochoa ÁQuirantes-Piné et al. (2019)
			1-arachidonoylglycerol monoacylglycerol (20:4) Monoacylglycerol (20:5)	AUC: 0.793 (CI95% 0.687–0.875)		
				AUC: 0.748 (CI 95% 0.64–0.858)		
dcSSc	Serum	UHPLC-Q-TOF-MS	1-Palmitoyl-sn-glycero-3-phosphocholine	AUC 0.650[b]	NR	Sun et al. (2022)
			Trans-dehydroandrosterone	AUC 0.720[b]		
			Betaine	AUC 0.771[b]		
			1-stearoyl-2-oleoyl-sn-glycerol 3-phosphocholine	AUC 0.725[b]		
		H-NMRS GC-MS[a]	Acetate	AUC: 0.84 (CI 95% 0.7–0.98)	NR	Murgia et al. (2018)
			Fructose			
			Glutamate			
			Glutamine			
			Glycerol			
			Glutarate			
PAH	Serum	HPLC-FLD	Asymmetric dimethylarginine	AUC 0.86 (CI 95% 0.7–1.0)	NR	Thakkar et al. (2016)
			Symmetric dimethylarginine	AUC 0.88 (CI 95% 0.74–1.0)		
Progressive ILD	Serum	UHPLC-Q-TOF-MS	L-Leucine	AUC 0.847 (CI 95% 0.695–1.00)	External validation Paris Cohort	Meier et al. (2020)
			L-Isoleucine	AUC 0.826 (CI 95% 0.656–0.997)	Progressive SSc-ILD (n = 7)	
			Adenosine monophosphate	AUC 0.785 (CI 95% 0.598–0.971)	Controls (n = 27)	
			Xanthosine	AUC 0.771 (CI 95% 0.551–0.9)		

(Continued on following page)

TABLE 4 (*Continued*) Potential metabolite biomarkers for diagnosis and classification of SSc.

	Sample type	Analytical technique	Metabolites	ROC curve	Validation	Study
ILD	Serum	UHPLC-Q-TOF-MS	Ile-Ala	AUC 0.807[b]	NR	Sun et al. (2022)
			L-Glutamine	AUC 0.756[b]		
			Androsterone sulfate	AUC 0.778[b]		

[a]Detector is not specified in the document.
[b]IC, 95% Not reported.
Abbreviations: SSc, Systemic sclerosis; dcSSc, diffuse cutaneous systemic sclerosis; PAH, pulmonary arterial hypertension; ILD, interstitial lung disease; H NMRS, proton nuclear magnetic resonance spectrometry; GC-MS, Gas chromatography-mass spectrometry; UHPLC-Q-TOF-MS, Ultra-high-performance liquid chromatography quadrupole time-of-flight mass spectrometry; HPLC-Q-TOF-MS, high-performance liquid chromatography quadrupole time-of-flight mass spectrometry; HPLC-FLD, High-performance liquid chromatography with fluorescence detection; ROC, receiver operating characteristic curve; AUC, area under the curve; NR, not reported.

synthesis and catabolic processes for energy production (Akram et al., 2011). Increased Hcy levels cause vascular injury by supporting oxidative stress through the production of reactive oxygen species, thus inhibiting antioxidant enzymes, and inducing low-density lipoprotein oxidation in arterial muscle cells (Zhang et al., 2018). Moreover, it induces endothelial dysfunction by inactivating anticoagulant substances (Škovierová et al., 2016). It is noteworthy that both Szamosi et al. (2009) and ichiro et al. (2014) found no difference in Hcy levels between SSc patients and HC; however, significantly higher Hcy concentrations were found in patients with vascular or thromboembolic events in comparison to SSc patients without these manifestations (Szamosi et al., 2009). Likewise, elevated Hcy levels were positively correlated with SSc-ILD (ichiro et al., 2014).

The upregulation of amino acid metabolites such as glutamine, ornithine, proline, and citrulline can lead to the augmentation of collagen synthesis with subsequent fibrosis of the skin and internal organs (Ung et al., 2021). Urea cycle intermediates, such as ornithine and citrulline are involved in proline synthesis (Albaugh et al., 2017). Proline a key component in the synthesis of collagen and the ECM (Karna et al., 2020), is increased in transforming growth factor beta (TGFβ) stimulated fibroblasts, increasing collagen formation and accounting for fibrosis (Schwörer et al., 2020). Glutamine promotes the novo synthesis of proline and sustains collagen synthesis in fibroblasts (Kay et al., 2021). Glutaminolysis is required for the formation of *a*-ketoglutarate, one of the main collagen I precursors, implying that glutamine metabolism is also important in the development of fibrosis (Ge et al., 2018). Hamanaka et al. (2019) found that the conversion of glutamine to glutamate is required for collagen protein production induced by TGFβ stimulated lung fibroblasts. Additionally, Bernard et al. (2018), tested the role of glutaminolysis in TGF-β1-dependent myofibroblast development and found that TGF-1-differentiated myofibroblasts were compared to controls, glutamate concentrations increased, but glutamine levels decreased, indicating accelerated glutaminolysis. This was linked to TGF- β1 induced mRNA and protein production of the glutaminase (GLS) isoform GLS1 that converts glutamine to glutamate. In this case, extracellular glutamine depletion inhibited TGF-β induced myofibroblast differentiation. Also, glutaminolysis is considered one of the main energy sources for effector T cells and facilitates Th17 proinflammatory phenotype (Cruzat et al., 2018). Several studies have found that the level of Th17 cells in SSc patients is increased compared to HC (Maddur et al., 2012). In SSc, Th17 cells release cytokines that can promote the proliferation and migration of dermal vascular smooth muscle cells inducing endothelial inflammation (Xing et al., 2013). Nevertheless, conflicting roles in Th17 function and fibrosis development have been discussed (Wei et al., 2022). Some studies claimed that Th17 cells induced type I collagen synthesis and secretion to promote fibrosis of murine SSc models (Wilson et al., 2010; Lei et al., 2016; Ramani and Biswas, 2019); however, other researchers propose Th17 cells decrease type I collagen production by dermal fibroblasts (Brembilla et al., 2013; Chizzolini et al., 2018). Additionally, it has been revealed that increased expression of endothelial CCR6, a surface marker of Th17 cell subsets, contributes to the development of SSc vasculopathy (Ikawa et al., 2021). These actions are associated with chronic inflammation and fibrosis, bolstering Th17 cell function in SSc patients.

Furthermore, activation of arginine methyltransferases by inflammation and oxidative stress leads to increased levels of ADMA, the principal endogenous inhibitor of nitric oxide synthase (NOS) (Zhang et al., 2015b). It also leads to an impairment of nitric oxide synthesis, contributing to the augmentation of vasoconstrictor episodes and pathological changes in the vascular system, generating endothelial dysfunction and vascular remodeling (Curtiss et al., 2019). The kynurenine pathway (KP) plays an important role in autoimmune disorders (Boros and Vécsei, 2019). In conditions characterized by inflammation, proinflammatory cytokines such as interferon γ, IL-6, and tumor necrosis factor induce Trp conversion to Kyn by the immune regulatory enzyme indoleamine-2,3-dioxygenase (IDO) (Zou, 2015). Upregulation of IDO enzyme can inhibit mTOR, a regulator of T cell differentiation, and therefore can inhibit effector T cells while promoting regulatory T cells (Treg) (Kurniawan et al., 2020; Sharabi and Tsokos, 2020); however, it has also been demonstrated *in vitro* that Kyn can stimulate mTORC1 activity (Qin et al., 2022). On the other hand, Kyn binds to the aryl hydrocarbon receptor in T cells and dendritic cells promoting the conversion of effector T cells into Treg and promoting IDO induction, therefore establishing a loop to maintain immunotolerance (Lionetto et al., 2021). Additionally, Trp deprivation via IDO mediates cell cycle arrest in the mid-G1 resulting in T cell death and suppression of antigen-specific T cell responses (Krupa and Kowalska, 2021). The Kyn/Trp ratio

TABLE 5 Relevant metabolites in SSc patients identified through targeted metabolomics studies.

Sample type	Sample size	Analytical technique	Quantified metabolite	Concentration patients vs. controls (mean ± SD)	Study
Plasma	SSc: 60	HPLC-FLD	Homocysteine	11.8 (10.3–14.5) vs. 6.5 (5.4–8.8) μmol/L[b]	Caramaschi et al. (2007)
	Controls: 30				
	SSc: 15	GC-MS[a]	Malondialdehyde	20.3 vs. 2.48 nmol/L	Tikly et al. (2006)
	Controls: 13				
	SSc: 71	HPLC-FLD	Homocysteine	11.1 vs. 6.9 μmol/L	Caramaschi et al. (2003)
	Controls: 30				
	SSc: 40	HPLC-MS[a]	Vitamin D	25.77 (±12.84) vs. 35.08 (±9.07) ng/mL	Atteritano et al. (2016)
	Controls: 40				
	SSc: 42 Controls: 27	HPLC-TQ-MS	Glutamine	689 (±122.3) vs. 618.4 (±165.3) μmol/L	Smolenska et al. (2020)
			Proline	178.8 (±55.2) vs. 152.5 (±47.3) μmol/L	
			1-Methylhistidine	5.7 (±3.9) vs. 4.1 (±1.5) μmol/L	
			Asymmetric dimethylarginine	0.344 (±0.112) vs. 0.289 (±0.10) μmol/L	
			Betaine	64.8 (±20.8) vs. 52.8 (±17.8) μmol/L	
			Tryptophan	32.5 (±9.6) vs. 40.8 (±12.3) μmol/L	
			Methylnicotinamide	0.312 (±0.166) vs. 0.232 (±0.106) μmol/L	
Serum	SSc: 10	HPLC-QTRAP-MS	Arachidonoyl (20:4)-LPA	2.54 (±0.15) vs. 1.15 (±0.37) nmol/mL	Tokumura et al. (2009)
	Controls: 13				
	SSc- PAH: 15	HPLC-FLD	Arginine	97.28 (±27.4) vs. 117.45 (±26.07) μmol[c]	Thakkar et al. (2016)
	SSc Non-PAH: 30		Asymmetric dimethylarginine Symmetric dimethylarginine	0.76 (±0.14) vs. 0.59 (±0.07) μmol[c] 0.76 (±0.26) vs. 0.46 (±0.07) μmol[c]	
Urine	SSc: 11	GC-EI-MS	15-F-2t-isoprostane	178 (±32) vs. 95 (±1) μmoles/mmole of creatinine	Cracowski et al. (2002)
	Controls: 11				
	SSc: 43	HPLC-UV	8-isoprostaglandin-F2a	341.7 vs. 147.6 pg/mg creatinine	Volpe et al. (2006)
	Controls: 43				
Exhaled breath	SSc: 46	GC-FID	Ethane	5.27 vs. 2.72 μmol ml^{-1} CO_2	Cope et al. (2006)
	Controls: 21		Ethanol	32.5 vs. 76.0 μmol ml^{-1} CO_2	

[a]Detector is not specified in the document.
[b]IQR, interquartile range.
[c]Comparison between SSc, patients with PAH vs. SSc, patients without PAH.
Abbreviations: SSc, Systemic sclerosis; SD, standard deviation; PAH, pulmonary arterial hypertension; GC-MS, Gas chromatography-mass spectrometry; HPLC-TQ-MS, High-performance liquid chromatography coupled to triple-stage quadrupole mass spectrometer; HPLC-QTRAP-MS, High-performance liquid chromatography quadrupole-linear ion trap hybrid mass spectrometry; HPLC-FLD, High-performance liquid chromatography with fluorescence detection; GC-EI-MS, gas chromatography electron impact mass spectrometry; HPLC-UV, High-performance liquid chromatography with on-line UV, system; CGC-FID, capillary gas chromatography flame ionization detection.

is a useful marker of IDO activation that reflects the state of immune activation in proinflammatory disorders. Therefore, an increased Kyn/Trp ratio in patients with SSc can be an indicator of increased inflammation and immune system activation.

Alterations in the TCA cycle, the central metabolic pathway for aerobic metabolic processes (Cavalcanti et al., 2014; Arnold and Finley, 2023), were also observed in analyzed samples of SSc patients compared to HC. Increased levels of succinate, a TCA cycle intermediate, were also found. In a study by Tannahill et al. (2013), bone-marrow-derived macrophages stimulated with lipopolysaccharides (LPS) showed that succinate is induced by LPS, impairing propyl hydroxylase (PHD) activity that leads to hypoxia-inducible factor 1-α (HIF-1 α) stabilization and activation, enhancing IL-1b production during inflammation. Therefore, LPS-induced succinate can serve as a signal to enhance IL-1b expression via HIF-1 α. Furthermore, the accumulation of succinate in lung tissue and myofibroblast can contribute to metabolic dysregulation in fibroblasts disrupting PHD activity and enhancing HIF-1 α, promoting the development of lung fibrosis (Wang et al., 2021a). Additionally, Henderson et al. (2020) performed *in vitro* studies in dermal fibroblasts derived from SSc patients and stimulated isolated normal healthy dermal fibroblasts (NHDFs) with TGF-β1 to activate fibrotic pathways and measured succinate levels. After stimulation, they found significantly higher levels of this metabolite in NHDFs, as

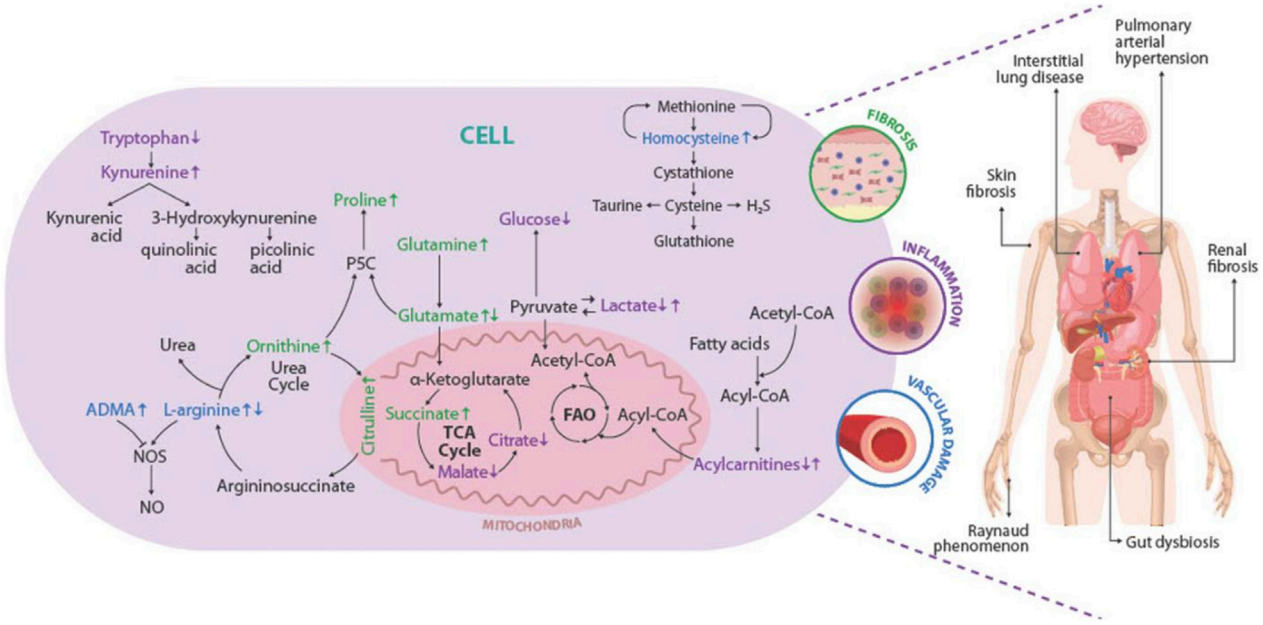


FIGURE 4
Overview of the deregulations found in metabolites and metabolic pathways and their association with the pathophysiology of the SSc. The main deregulated metabolites and metabolic pathways identified in SSc are represented inside the cell. The metabolites linked with fibrosis are shown in green, those related to the presentation of vascular injury in blue, and those associated with the development of inflammation in purple. Increased metabolite levels are shown by up arrows, whereas decreased metabolite levels are represented by down arrows. Representation of the pathogenic triad of SSc: fibrosis of the skin and internal organs (green), vascular damage (blue), and inflammation (purple). The main clinical manifestations and complications of SSc are represented in the human figure. P5C, sphingosine-1-phosphate receptor; TCA cycle: tricarboxylic acid cycle; FAO, fatty acid beta-oxidation; NO, Nitric oxide; NOS, Nitric oxide synthase; ADMA, Asymmetric dimethylarginine; H_2S, hydrogen sulfide.

well as elevated levels of succinate receptor GPR91 in SSc dermal fibroblasts suggesting that succinate released from macrophages can activate fibroblast to undergo fibrotic changes leading to enhanced ECM. Succinate activation of GPR91 has been shown to be important for fibroblast activation and ECM formation in murine intestinal fibrosis and non-alcoholic steatohepatitis (NASH)-associated fibrosis, as well as in fibrotic lung tissue from idiopathic pulmonary fibrosis patients and bleomycin-induced mice (Macias-Ceja et al., 2019; Liu et al., 2020).

On the contrary, citrate levels were significantly decreased. As a TCA cycle intermediate, citrate is crucial for energy production (Iacobazzi and Infantino, 2014). The high consumption of citrate to meet energy charges could explain its low levels, reflecting a reduction in energy availability and increased demand under inflammatory conditions. Various studies have reported decreased levels in samples of patients with autoimmune diseases underlining its importance in immune-mediated inflammatory pathologies (Ouyang et al., 2011; Alonso et al., 2016). Moreover, Yang et al. (2015) demonstrated that a decrease in citric acid accompanied by a decrease in glucose means an increase in energy consumption. Increased glycolysis plays a critical role in fibroblast differentiation and the progression of fibrosis (Zhu et al., 2019). It has been shown that TGF-β1 can cause a rewiring of cellular metabolism, including a shift toward glycolysis, uncoupling from mitochondrial oxidative phosphorylation, and increasing glutamine metabolism (Hewitson and Smith, 2021). In experimental models of SSc, the profibrotic M2 macrophages isolated from bleomycin-induced fibrotic mouse lungs showed increased glycolysis, suggesting its importance in assuring energy efficiency (Xie et al., 2017).

In the case of lipid metabolism, SSc patients showed alterations in carnitines, FA, glycerophospholipids, glycerolipids, sphingolipids, and steroids. Acyl-carnitines play an important role in cellular energy metabolism as a transporter of FA chains into the mitochondria, where long-chain FA are further oxidized (Fielding et al., 2018). Therefore, the downregulation of acyl-carnitines leads to perturbations in fatty acids oxidation (FAO), subsequently increasing FA metabolism (Beger et al., 2018). Furthermore, perturbations of FAO can shift T helper cell differentiation towards a proinflammatory Th17 phenotype (Slack et al., 2015). Additionally, *in vitro* studies have demonstrated that FA accumulation in non-adipose tissues, defined as lipotoxicity, due to FAO inhibition, promotes inflammation, oxidative stress, and fibrosis in renal tubular epithelial cells (Kang et al., 2015a). Moreover, in macrophages, increased FA metabolism can induce a switch to a profibrotic M2 phenotype, playing an important role in fibrosis (Nomura et al., 2016; Wang et al., 2017). In tissues, the healing process depends on whether the initial insult persists or not. If the insult persists, chronic activation of M2 can directly regulate the development and progression of fibrotic lung diseases through the production of chemokines, tissue inhibitor of metalloproteinases, and fibronectin, as well as the capability of M2 to differentiate into fibrocyte-like cells that express collagen, opposite to their primary anti-inflammatory activity through the release of TGF-β, IL10, and arginase, controlling wound healing and tissue regeneration (Braga et al., 2015; Kishore and Petrek, 2021). In

the kidney, M2 macrophages induce Th2-type immune responses, secrete large amounts of TGF-β and anti-inflammatory cytokines, transform into myofibroblasts in the injured kidney, inhibit immune responses, and promote wound healing and tissue fibrosis (Wang et al., 2021b).

Regarding sphingolipids, increased levels of sphingosine 1-phosphate (S1P) were observed. Deregulation of S1P in the pulmonary endothelium can lead to vasoconstrictive episodes and vascular remodeling increasing pulmonary vascular resistance (Gluschke et al., 2022). Moreover, S1P influences antigen uptake and presentation by dendritic cells (Arlt et al., 2014). Additionally, the S1P receptor, sphingosine-1-phosphate, modulates early fibrogenesis (Schmidt et al., 2017).

DcSSc and lcSSc subtypes are each characterized by different clinical manifestations, and disease progression (Herrick, 2018); however, the dcSSc subtype is associated with more severe and aggressive organ involvement (Smeets et al., 2020). Therefore, early identification of the disease subtype is imperative to achieve the effectiveness of therapeutic interventions. In this review, analyzed samples of dcSSc patients in comparison with lcSSc demonstrated significantly increased levels of amino acid-related pathways, involved in fibrosis, endothelial dysfunction, and gut dysbiosis. L-NAME, a NOS inhibitor was increased in the dcSSc subtype, suggesting a more severe endothelial dysfunction that could lead to vascular complications (Dooley et al., 2006). As for gut dysbiosis, phenylacetylglutamine (PAG), a gut microbiota-derived metabolite (Teufel et al., 2010), is consistently upregulated in SSc patients. A reduced number of commensal bacteria promotes an excess of substrate favoring PAG formation, suggesting that its deregulation could be an indicator of gastrointestinal involvement (Poesen et al., 2016).

In SSc patients, cardiopulmonary complications are the leading cause of death (Bruni et al., 2021). In this context, screening for PAH and ILD in SSc has emerged as an important consideration. Thakkar et al. (2016) displayed decreased levels of L-arginine, a common substrate of NOS (Cziráki et al., 2020), and increased levels of ADMA, suggesting an association with SSc-PAH. Decreased levels of amino acids with protective effects against endothelial dysfunction and anti-inflammatory effects by inhibition of proinflammatory cytokines, and oxidative stress reduction (Zhong et al., 2003; Hasegawa et al., 2012; Zhao et al., 2018), such as glycine, histidine, and betaine were also found in these patients. Furthermore, elevated levels of low-density lipoprotein in the lungs may lead to lipotoxicity, inducing inflammation and oxidative stress, which causes pulmonary vascular remodeling (Calvier et al., 2022). In SSc-ILD patients, metabolites such as Hcy, proline, glutamine, and BCAAs were elevated, which are involved in the amino acid pathways associated with fibrosis and inflammation. The upregulation of BCAAs can enhance proinflammatory phenotype by activating nuclear factor kappa B in immune cells and over-expression of IL-6 and tumor necrosis factor (Zhenyukh et al., 2017). On the other hand, phosphatidylethanolamine (PE) levels were decreased. Vazquez-de-Lara et al. (2018) demonstrated that PE could attenuate bleomycin-induced lung fibroblast, by decreasing the soluble collagen concentration in mice lungs.

To summarize, Figure 4 depicts a graphic overview of the main deregulated metabolites and metabolic pathways identified in SSc, as well as the role that these could play in the pathophysiology of the disease, leading to the appearance of clinical manifestations and complications associated with the disease.

Metabolomics, as a fast-developing technique in biomedical research, can be used to identify novel biomarkers (Kang et al., 2015b) and as a promising predictive or personalized medicine research technique (Zhou and Zhong, 2022). However, this systematic review has certain limitations, the most significant one is the challenge of comparing the metabolomic acquired across studies due to different limitation factors. Due to the scarcity of available quantitative data, the potential application of meta-analysis is limited, reducing the capacity to make more solid and generalizable conclusions. Variations in sample sources, sample preparation techniques, and metabolite detection methods may be to blame for the heterogeneity of results among studies. Another key constraint is the requirement for further research to validate metabolomics findings in multiple cohorts or independent populations. Validation of results is critical to ensuring the robustness and therapeutic usefulness of the proposed biomarkers. Likewise, most of the studies evaluated were cross-sectional, preventing us from determining a causal association in the metabolic changes associated with SSc. Despite these limitations, metabolomics remains a valuable tool, offering a unique opportunity to understand the metabolic basis of the disease and develop new diagnostic and treatment strategies.

In this review, potential biomarkers were described for the diagnosis of SSc, the identification of the dcSSc subtype, and the identification of primary pulmonary complications such as PAH, and ILD. These potential biomarkers were mainly within amino acids, nucleotides, carboxylic acids, and carbohydrate metabolism. More data are necessary concerning the specificity of biomarkers; as well as external validation studies in other and larger populations; however, we expect that metabolomics will provide more accurate and more validated biomarkers for the detection of SSc.

5 Conclusion

The data extracted from the 26 studies showed distinct metabolic profiles between SSc patients and HC and distinct profiles between SSc subtypes, generating new insights for non-invasive prognostic and early diagnostic biomarkers to improve individualized treatment and delay disease progression. Although the metabolic profile can still be affected by a series of other factors, the results obtained suggest the presence of a metabolic fingerprint of the disease. The disrupted metabolite mechanisms identified in this study, mainly, but not exclusively, involving amino acids and lipid metabolism, as well as TCA cycle dysregulation are associated with autoimmune inflammation, vascular damage, fibrosis, and gut dysbiosis, which might be relevant for the development of SSc. Nevertheless, further studies are required to evaluate the role of these alterations in the pathophysiology of the disease, as well as to assess whether these metabolomic networks have potential as treatment targets or as biomarkers not only for diagnosis but also for prognosis and treatment response.

Author contributions

CR-S, VM-G, YR, and MC contributed to the conceptualization and design of the study. Acquisition of data was performed by VM-G, JC-V, and DG-S. Methodology was designed by CR-S, VM-G, MC, and DP-R. Center Coordination: YA-A and CR-S. Writing and editing: VM-G, DP-R, MC, DM, YA-A, and CR-S. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Universidad Del Rosario (ABN011) Bogota, Colombia.

Acknowledgments

The authors would like to express their gratitude to CREA members for their contributions and fruitful discussions during the manuscript preparation process.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1215039/full#supplementary-material

References

Akter, T., Silver, R. M., and Bogatkevich, G. S. (2014). Recent advances in understanding the pathogenesis of scleroderma-interstitial lung disease. *Curr. Rheumatol. Rep.* 16 (4), 411. doi:10.1007/s11926-014-0411-1

Aida-Yasuoka, K., Peoples, C., Yasuoka, H., Hershberger, P., Thiel, K., Cauley, J. A., et al. (2013). Estradiol promotes the development of a fibrotic phenotype and is increased in the serum of patients with systemic sclerosis. *Arthritis Res. Ther.* 15 (1), R10. doi:10.1186/ar4140

Akram, M., Asif, M., Uzair, M., Naveed, A., Madni, M. A., Ali Shah, D. S., et al. (2011). Amino acids: A review article. *J. Med. Plants Res.* 5, 3997–4000.

Albaugh, V. L., Mukherjee, K., and Barbul, A. (2017). Proline precursors and collagen synthesis: Biochemical challenges of nutrient supplementation and wound healing. *J. Nutr.* 147, 2011–2017. doi:10.3945/jn.117.256404

Allanore, Y., Simms, R., Distler, O., Trojanowska, M., Pope, J., Denton, C. P., et al. (2015). Systemic sclerosis. *Nat. Rev. Dis. Prim.* 1 (1), 15002. doi:10.1038/nrdp.2015.2

Alonso, A., Julià, A., Vinaixa, M., Domènech, E., Fernández-Nebro, A., Cañete, J. D., et al. (2016). Urine metabolome profiling of immune-mediated inflammatory diseases. *BMC Med.* 14 (1), 133. doi:10.1186/s12916-016-0681-8

Arlt, O., Schwiebs, A., Japtok, L., Rüger, K., Katzy, E., Kleuser, B., et al. (2014). Sphingosine-1-Phosphate modulates dendritic cell function: Focus on non-migratory effects ***in vitro*** and ***in vivo*** cellular physiology and biochemistry. *Cell. Physiol. biochem.* 34 (1), 27–44. doi:10.1159/000362982

Arnold, P. K., and Finley, L. W. S. (2023). Regulation and function of the mammalian tricarboxylic acid cycle. *J. Biol. Chem.* 299 (2), 102838. doi:10.1016/j.jbc.2022.102838

Atteritano, M., Santoro, D., Corallo, G., Visalli, E., Buemi, M., Catalano, A., et al. (2016). Skin involvement and pulmonary hypertension are associated with vitamin D insufficiency in scleroderma. *Int. J. Mol. Sci.* 17 (12), 2103. doi:10.3390/ijms17122103

Bairkdar, M., Rossides, M., Westerlind, H., Hesselstrand, R., Arkema, E. V., and Holmqvist, M. (2021). Incidence and prevalence of systemic sclerosis globally: A comprehensive systematic review and meta-analysis. *Rheumatology* 60 (7), 3121–3133. doi:10.1093/rheumatology/keab190

Beger, R. D., Bhattacharyya, S., Gill, P. S., and James, L. P. (2018). "Acylcarnitines as translational biomarkers of mitochondrial dysfunction," in *Mitochondrial dysfunction caused by drugs and environmental toxicants* (Hoboken, NJ, USA: John Wiley & Sons, Inc.), 383–393.

Bellocchi, C., Fernández-Ochoa ÁMontanelli, G., Vigone, B., Santaniello, A., Milani, C., et al. (2018). Microbial and metabolic multi-omic correlations in systemic sclerosis patients. *Ann. N. Y. Acad. Sci.* 1421 (1), 97–109. doi:10.1111/nyas.13736

Bengtsson, A. A., Trygg, J., Wuttge, D. M., Sturfelt, G., Theander, E., Donten, M., et al. (2016). Metabolic profiling of systemic lupus erythematosus and comparison with primary sjögren's syndrome and systemic sclerosis. *PLoS One* 11 (7), e0159384. doi:10.1371/journal.pone.0159384

Bernard, K., Logsdon, N. J., Benavides, G. A., Sanders, Y., Zhang, J., Darley-Usmar, V. M., et al. (2018). Glutaminolysis is required for transforming growth factor-β1–induced myofibroblast differentiation and activation. *J. Biol. Chem.* 293 (4), 1218–1228. doi:10.1074/jbc.RA117.000444

Bögl, T., Mlynek, F., Himmelsbach, M., Sepp, N., Buchberger, W., and Geroldinger-Simić, M. (2022). Plasma metabolomic profiling reveals four possibly disrupted mechanisms in systemic sclerosis. *Biomedicines* 10 (3), 607. doi:10.3390/biomedicines10030607

Boros, F. A., and Vécsei, L. (2019). Immunomodulatory effects of genetic alterations affecting the kynurenine pathway. *Front. Immunol.* 10, 2570. doi:10.3389/fimmu.2019.02570

Braga, T. T., Agudelo, J. S. H., and Camara, N. O. S. (2015). Macrophages during the fibrotic process: M2 as friend and foe. *Front. Immunol.* 6, 602. doi:10.3389/fimmu.2015.00602

Brembilla, N., Montanari, E., Truchetet, M. E., Raschi, E., Meroni, P., and Chizzolini, C. (2013). Th17 cells favor inflammatory responses while inhibiting type I collagen deposition by dermal fibroblasts: Differential effects in healthy and systemic sclerosis fibroblasts. *Arthritis Res. Ther.* 15 (5), R151. doi:10.1186/ar4334

Bruni, C., Guignabert, C., Manetti, M., Cerinic, M. M., and Humbert, M. (2021). The multifaceted problem of pulmonary arterial hypertension in systemic sclerosis. *Lancet Rheumatol.* 3, e149. Available at: http://www.thelancet.com/article/S2665991320303568/fulltext.

Calvier, L., Herz, J., and Hansmann, G. (2022). Interplay of low-density lipoprotein receptors, LRPs, and lipoproteins in pulmonary hypertension. *JACC Basic Transl. Sci.* 7 (2), 164–180. doi:10.1016/j.jacbts.2021.09.011

Cambiaghi, A., Ferrario, M., and Masseroli, M. (2017). Analysis of metabolomic data: Tools, current strategies and future challenges for omics data integration. *Brief. Bioinform* 18, 498–510. doi:10.1093/bib/bbw031

Campochiaro, C., Lytton, S., Nihtyanova, S., Fuchs, D., Ong, V. H., and Denton, C. P. (2019). Elevated kynurenine levels in diffuse cutaneous and anti-RNA polymerase III positive systemic sclerosis. *Clin. Immunol.* 199, 18–24. doi:10.1016/j.clim.2018.12.009

Caramaschi, P., Martinelli, N., Biasi, D., Carletto, A., Faccini, G., Volpe, A., et al. (2003). Homocysteine plasma concentration is related to severity of lung impairment in scleroderma. *J. Rheumatol.* 30 (2), 298–304.

Caramaschi, P., Volpe, A., Canestrini, S., Bambara, L. M., Faccini, G., Carletto, A., et al. (2007). Correlation between homocysteine plasma levels and nailfold videocapillaroscopic patterns in systemic sclerosis. *Clin. Rheumatol.* 26 (6), 902–907. doi:10.1007/s10067-006-0425-9

Cavalcanti, J. H. F., Esteves-Ferreira, A. A., Quinhones, C. G. S., Pereira-Lima, I. A., Nunes-Nesi, A., Fernie, A. R., et al. (2014). Evolution and functional implications of the tricarboxylic acid cycle as revealed by phylogenetic analysis. *Genome Biol. Evol.* 6 (10), 2830–2848. doi:10.1093/gbe/evu221

Chizzolini, C., Dufour, A. M., and Brembilla, N. C. (2018). Is there a role for IL-17 in the pathogenesis of systemic sclerosis? *Immunol. Lett.* 195, 61–67. doi:10.1016/j.imlet.2017.09.007

Cope, K. A., Solga, S. F., Hummers, L. K., Wigley, F. M., Diehl, A. M., and Risby, T. H. (2006). Abnormal exhaled ethane concentrations in scleroderma. *Biomarkers* 11 (1), 70–84. doi:10.1080/13547500500515046

Coral-Alvarado, P., Pardo, A. L., Castaño-Rodriguez, N., Rojas-Villarraga, A., and Anaya, J. M. (2009). Systemic sclerosis: A world wide global analysis. *Clin. Rheumatol.* 28 (7), 757–765. doi:10.1007/s10067-009-1144-9

Cracowski, J. L., Carpentier, P. H., Imbert, B., Cachot, S., Stanke-Labesque, F., Bessard, J., et al. (2002). Increased urinary F2-isoprostanes in systemic sclerosis, but not in primary Raynaud's phenomenon: Effect of cold exposure. *Arthritis Rheum.* 46 (5), 1319–1323. doi:10.1002/art.10261

Cruzat, V., Macedo Rogero, M., Noel Keane, K., Curi, R., and Newsholme, P. (2018). Glutamine: Metabolism and immune function, supplementation and clinical translation. *Nutrients* 10 (11), 1564. doi:10.3390/nu10111564

Curtiss, P., Schwager, Z., Lo Sicco, K., and Franks, A. G. (2019). The clinical effects of l-arginine and asymmetric dimethylarginine: Implications for treatment in secondary Raynaud's phenomenon. *J. Eur. Acad. Dermatology Venereol.* 33 (3), 497–503. doi:10.1111/jdv.15180

Cziráki, A., Lenkey, Z., Sulyok, E., Szokodi, I., and Koller, A. (2020). L-Arginine-Nitric oxide-asymmetric dimethylarginine pathway and the coronary circulation: Translation of basic science results to clinical practice. *Front. Pharmacol.* 11, 569914. doi:10.3389/fphar.2020.569914

Deidda, M., Piras, C., Cadeddu Dessalvi, C., Locci, E., Barberini, L., Orofino, S., et al. (2017). Distinctive metabolomic fingerprint in scleroderma patients with pulmonary arterial hypertension. *Int. J. Cardiol.* 241, 401–406. doi:10.1016/j.ijcard.2017.04.024

Denton, C. P., and Khanna, D. (2017). Systemic sclerosis. *Lancet* 390 (10103), 1685–1699. doi:10.1016/S0140-6736(17)30933-9

Dooley, A., Gao, B., Bradley, N., Abraham, D. J., Black, C. M., Jacobs, M., et al. (2006). Abnormal nitric oxide metabolism in systemic sclerosis: Increased levels of nitrated proteins and asymmetric dimethylarginine. *Rheumatology* 45 (6), 676–684. doi:10.1093/rheumatology/kei276

Elhai, M., Meune, C., Avouac, J., Kahan, A., and Allanore, Y. (2012). Trends in mortality in patients with systemic sclerosis over 40 years: A systematic review and meta-analysis of cohort studies. *Rheumatology* 51 (6), 1017–1026. doi:10.1093/rheumatology/ker269

Elhai, M., Meune, C., Boubaya, M., Avouac, J., Hachulla, E., Balbir-Gurman, A., et al. (2017). Mapping and predicting mortality from systemic sclerosis. *Ann. Rheum. Dis.* 76 (11), 1897–1905. doi:10.1136/annrheumdis-2017-211448

Fernández-Ochoa ÁQuirantes-Piné, R., Borrás-Linares, I., Gemperline, D., Alarcón Riquelme, M. E., Beretta, L., et al. (2019). Urinary and plasma metabolite differences detected by HPLC-ESI-QTOF-MS in systemic sclerosis patients. *J. Pharm. Biomed. Anal.* 162, 82–90. doi:10.1016/j.jpba.2018.09.021

Fielding, R., Riede, L., Lugo, J., and Bellamine, A. (2018). l-Carnitine supplementation in recovery after exercise. *Nutrients* 10 (3), 349. doi:10.3390/nu10030349

Ge, J., Cui, H., Xie, N., Banerjee, S., Guo, S., Dubey, S., et al. (2018). Glutaminolysis promotes collagen translation and stability via α-Ketoglutarate-mediated mTOR activation and proline hydroxylation. *Am. J. Respir. Cell. Mol. Biol.* 58 (3), 378–390. doi:10.1165/rcmb.2017-0238OC

Geroldinger-Simić, M., Bögl, T., Himmelsbach, M., Sepp, N., and Buchberger, W. (2021). Changes in plasma phospholipid metabolism are associated with clinical manifestations of systemic sclerosis. *Diagnostics* 11 (11), 2116. doi:10.3390/diagnostics11112116

Gluschke, H., Siegert, E., Minich, W. B., Hackler, J., Riemekasten, G., Kuebler, W. M., et al. (2022). Autoimmunity to sphingosine-1-phosphate-receptors in systemic sclerosis and pulmonary arterial hypertension. *Front. Immunol.* 13, 935787. doi:10.3389/fimmu.2022.935787

Hamanaka, R. B., O'Leary, E. M., Witt, L. J., Tian, Y., Gökalp, G. A., Meliton, A. Y., et al. (2019). Glutamine metabolism is required for collagen protein synthesis in lung fibroblasts. *Am. J. Respir. Cell. Mol. Biol.* 61 (5), 597–606. doi:10.1165/rcmb.2019-0008OC

Hao, Y., Hudson, M., Baron, M., Carreira, P., Stevens, W., Rabusa, C., et al. (2017). Early mortality in a multinational systemic sclerosis inception cohort. *Arthritis & Rheumatology.* 69 (5), 1067–1077. doi:10.1002/art.40027

Hasegawa, S., Ichiyama, T., Sonaka, I., Ohsaki, A., Okada, S., Wakiguchi, H., et al. (2012). Cysteine, histidine and glycine exhibit anti-inflammatory effects in human coronary arterial endothelial cells. *Clin. Exp. Immunol.* 167 (2), 269–274. doi:10.1111/j.1365-2249.2011.04519.x

Henderson, J., Duffy, L., Stratton, R., Ford, D., and O'Reilly, S. (2020). Metabolic reprogramming of glycolysis and glutamine metabolism are key events in myofibroblast transition in systemic sclerosis pathogenesis. *J. Cell. Mol. Med.* 24 (23), 14026–14038. doi:10.1111/jcmm.16013

Herrick, A. L. (2018). Systemic sclerosis: Clinical features and management. *Medicine* 46 (2), 131–139. doi:10.1016/j.mpmed.2017.11.007

Hewitson, T. D., and Smith, E. R. (2021). A metabolic reprogramming of glycolysis and glutamine metabolism is a requisite for renal fibrogenesis—why and how? *Front. Physiol.* 12, 645857. doi:10.3389/fphys.2021.645857

Iacobazzi, V., and Infantino, V. (2014). Citrate – new functions for an old metabolite. *Biol. Chem.* 395 (4), 387–399. doi:10.1515/hsz-2013-0271

ichiro, M. S., Toki, S., Yamada, K., Uchiyama, A., and Ishikawa, O. (2014). Elevated plasma homocysteine level is possibly associated with skin sclerosis in a series of Japanese patients with systemic sclerosis. *J. Dermatol* 41, 986–991. doi:10.1111/1346-8138.12642

Ikawa, T., Miyagawa, T., Fukui, Y., Toyama, S., Omatsu, J., Awaji, K., et al. (2021). Endothelial CCR6 expression due to FLI1 deficiency contributes to vasculopathy associated with systemic sclerosis. *Arthritis Res. Ther.* 23 (1), 283. doi:10.1186/s13075-021-02667-9

Kaluarachchi, M., Boulangé, C. L., Karaman, I., Lindon, J. C., Ebbels, T. M. D., Elliott, P., et al. (2018). A comparison of human serum and plasma metabolites using untargeted 1H NMR spectroscopy and UPLC-MS. *Metabolomics* 14 (3), 32. doi:10.1007/s11306-018-1332-1

Kang, G. W., Jung, K. H., Lee, Y. S., Kim, H. J., Yoon, D. Y., Lee, S. H., et al. (2018). Incidence, prevalence, mortality and causes of death in systemic sclerosis in korea: A nationwide population-based study. *Br. J. Dermatology* 178 (1), e37–e39. doi:10.1111/bjd.15838

Kang, H. M., Ahn, S. H., Choi, P., Ko, Y. A., Han, S. H., Chinga, F., et al. (2015a). Defective fatty acid oxidation in renal tubular epithelial cells has a key role in kidney fibrosis development. *Nat. Med.* 21 (1), 37–46. doi:10.1038/nm.3762

Kang, J., Zhu, L., Lu, J., and Zhang, X. (2015b2015b). Application of metabolomics in autoimmune diseases: Insight into biomarkers and pathology. *J. Neuroimmunol.* 279, 25–32. doi:10.1016/j.jneuroim.2015.01.001

Karna, E., Szoka, L., Huynh, T. Y. L., and Palka, J. A. (2020). Proline-dependent regulation of collagen metabolism. *Cell. Mol. Life Sci.* 77 (10), 1911–1918. doi:10.1007/s00018-019-03363-3

Kay, E. J., Koulouras, G., and Zanivan, S. (2021). Regulation of extracellular matrix production in activated fibroblasts: Roles of amino acid metabolism in collagen synthesis. *Front. Oncol.* 11, 719922. doi:10.3389/fonc.2021.719922

Kishore, A., and Petrek, M. (2021). Roles of macrophage polarization and macrophage-derived miRNAs in pulmonary fibrosis. *Front. Immunol.* 12, 678457. doi:10.3389/fimmu.2021.678457

Kowalska-Kępczyńska, A. (2022). Systemic scleroderma—definition, clinical picture and laboratory diagnostics. *J. Clin. Med.* 11 (9), 2299. doi:10.3390/jcm11092299

Krupa, A., and Kowalska, I. (2021). The kynurenine pathway—new linkage between innate and adaptive immunity in autoimmune endocrinopathies. *Int. J. Mol. Sci.* 22 (18), 9879. doi:10.3390/ijms22189879

Kurniawan, H., Soriano-Baguet, L., and Brenner, D. (2020). Regulatory T cell metabolism at the intersection between autoimmune diseases and cancer. *Eur. J. Immunol.* 50 (11), 1626–1642. doi:10.1002/eji.201948470

Kuwana, M., Saito, A., Sakamoto, W., Raabe, C., and Saito, K. (2022). Incidence rate and prevalence of systemic sclerosis and systemic sclerosis-associated interstitial lung disease in Japan: Analysis using Japanese claims databases. *Adv. Ther.* 39 (5), 2222–2235. doi:10.1007/s12325-022-02078-5

Lau, C. H. E., Siskos, A. P., Maitre, L., Robinson, O., Athersuch, T. J., Want, E. J., et al. (2018). Determinants of the urinary and serum metabolome in children from six European populations. *BMC Med.* 16 (1), 202. doi:10.1186/s12916-018-1190-8

Launay, D., Sobanski, V., Hachulla, E., and Humbert, M. (2017). Pulmonary hypertension in systemic sclerosis: Different phenotypes. *Eur. Respir. Rev.* 26 (145), 170056. doi:10.1183/16000617.0056-2017

Lei, L., Zhao, C., Qin, F., He, Z. Y., Wang, X., and Zhong, X. N. (2016). Th17 cells and IL-17 promote the skin and lung inflammation and fibrosis process in a bleomycin-induced murine model of systemic sclerosis. *Clin. Exp. Rheumatol.* 34, 14–22.

Lionetto, L., Ulivieri, M., Capi, M., De Bernardini, D., Fazio, F., Petrucca, A., et al. (2021). Increased kynurenine-to-tryptophan ratio in the serum of patients infected with SARS-CoV2: An observational cohort study. *Biochimica Biophysica Acta (BBA) - Mol. Basis Dis.* 1867 (3), 166042. doi:10.1016/j.bbadis.2020.166042

Liu, X., Xie, L., Du, K., Liu, C., Zhang, N., Gu, C., et al. (2020). Succinate-GPR-91 receptor signalling is responsible for nonalcoholic steatohepatitis-associated fibrosis: Effects of DHA supplementation. *Liver Int.* 40 (4), 830–843. doi:10.1111/liv.14370

Lumbreras, B., Porta, M., Márquez, S., Pollán, M., Parker, L. A., and Hernández-Aguado, I. (2008). Quadomics: An adaptation of the Quality Assessment of Diagnostic Accuracy Assessment (QUADAS) for the evaluation of the methodological quality of studies on the diagnostic accuracy of '-omics'-based technologies. *Clin. Biochem.* 41 (16–17), 1316–1325. doi:10.1016/j.clinbiochem.2008.06.018

Macias-Ceja, D. C., Ortiz-Masiá, D., Salvador, P., Gisbert-Ferrándiz, L., Hernández, C., Hausmann, M., et al. (2019). Succinate receptor mediates intestinal inflammation and fibrosis. *Mucosal Immunol.* 12 (1), 178–187. doi:10.1038/s41385-018-0087-3

Maddur, M. S., Miossec, P., Kaveri, S. V., and Bayry, J. (2012). Th17 cells: Biology, pathogenesis of autoimmune and inflammatory diseases, and therapeutic strategies. *Am. J. Pathol.* 181 (1), 8–18. doi:10.1016/j.ajpath.2012.03.044

McNearney, T. A., Sluka, K. A., Ahn, C., Reveille, J. D., Fischbach, M., and Mayes, M. D. (2010). Plasma endogenous enkephalin levels in early systemic sclerosis: Clinical and laboratory associations. *Clin. Exp. Rheumatol.* 28 (2 Suppl. 58), S7–S11.

Meier, C., Freiburghaus, K., Bovet, C., Schniering, J., Allanore, Y., Distler, O., et al. (2020). Serum metabolites as biomarkers in systemic sclerosis-associated interstitial lung disease. *Sci. Rep.* 10 (1), 21912. doi:10.1038/s41598-020-78951-6

Morrisroe, K., Stevens, W., Huq, M., Prior, D., Sahhar, J., Ngian, G. S., et al. (2017). Survival and quality of life in incident systemic sclerosis-related pulmonary arterial hypertension. *Arthritis Res. Ther.* 19 (1), 122. doi:10.1186/s13075-017-1341-x

Murgia, F., Svegliati, S., Poddighe, S., Lussu, M., Manzin, A., Spadoni, T., et al. (2018). Metabolomic profile of systemic sclerosis patients. *Sci. Rep.* 8 (1), 7626. doi:10.1038/s41598-018-25992-7

Naranjo, M., and Hassoun, P. M. (2021). Systemic sclerosis-associated pulmonary hypertension: Spectrum and impact. *Diagn. (Basel)* 11 (5), 911. doi:10.3390/diagnostics11050911

Neumann Andersen, G., Caidahl, K., Kazzam, E., Petersson, A. S., Waldenström, A., Mincheva-Nilsson, L., et al. (2000). Correlation between increased nitric oxide production and markers of endothelial activation in systemic sclerosis: Findings with the soluble adhesion molecules E-selectin, intercellular adhesion molecule 1, and vascular cell adhesion molecule 1. *Arthritis Rheum.* 43 (5), 1085–1093. doi:10.1002/1529-0131(200005)43:5<1085:AID-ANR19>3.0.CO;2-7

Nihtyanova, S. I., and Denton, C. P. (2020). Pathogenesis of systemic sclerosis associated interstitial lung disease. *J. Scleroderma Relat. Disord.* 5 (2_Suppl. l), 6–16. doi:10.1177/2397198320903867

Nomura, M., Liu, J., RoviraII, Gonzalez-Hurtado, E., Lee, J., Wolfgang, M. J., et al. (2016). Fatty acid oxidation in macrophage polarization. *Nat. Immunol.* 17 (3), 216–217. doi:10.1038/ni.3366

O'Reilly, S. (2022). Metabolic perturbations in systemic sclerosis. *Curr. Opin. Rheumatol.* 34 (1), 91–94. doi:10.1097/BOR.0000000000000824

Ottria, A., Hoekstra, A. T., Zimmermann, M., van der Kroef, M., Vazirpanah, N., Cossu, M., et al. (2020). Fatty acid and carnitine metabolism are dysregulated in systemic sclerosis patients. *Front. Immunol.* 11, 822. doi:10.3389/fimmu.2020.00822

Ouyang, X., Dai, Y., Wen, J., and Wang, L. (2011). ^{1}H NMR-based metabolomic study of metabolic profiling for systemic lupus erythematosus. *Lupus* 20 (13), 1411–1420. doi:10.1177/0961203311418707

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 74, 790–799. doi:10.1016/j.rec.2021.07.010

Pattanaik, D., Brown, M., Postlethwaite, B. C., and Postlethwaite, A. E. (2015). Pathogenesis of systemic sclerosis. *Front. Immunol.* 6, 272. doi:10.3389/fimmu.2015.00272

Perelas, A., Silver, R. M., Arrossi, A. V., and Highland, K. B. (2020). Systemic sclerosis-associated interstitial lung disease. *Lancet Respir. Med.* 8 (3), 304–320. doi:10.1016/S2213-2600(19)30480-1

Poesen, R., Claes, K., Evenepoel, P., de Loor, H., Augustijns, P., Kuypers, D., et al. (2016). Microbiota-derived phenylacetylglutamine associates with overall mortality and cardiovascular disease in patients with CKD. *J. Am. Soc. Nephrol.* 27 (11), 3479–3487. doi:10.1681/ASN.2015121302

Qin, Y., Gao, C., and Luo, J. (2022). Metabolism characteristics of Th17 and regulatory T cells in autoimmune diseases. *Front. Immunol.* 13, 828191. doi:10.3389/fimmu.2022.828191

Ramani, K., and Biswas, P. S. (2019). Interleukin-17: Friend or foe in organ fibrosis. *Cytokine* 120, 282–288. doi:10.1016/j.cyto.2018.11.003

Rubio-Rivas, M., Royo, C., Simeón, C. P., Corbella, X., and Fonollosa, V. (2014). Mortality and survival in systemic sclerosis: Systematic review and meta-analysis. *Semin. Arthritis Rheum.* 44 (2), 208–219. doi:10.1016/j.semarthrit.2014.05.010

Schmidt, K. G., Herrero San Juan, M., Trautmann, S., Berninger, L., Schwiebs, A., Ottenlinger, F. M., et al. (2017). Sphingosine-1-Phosphate receptor 5 modulates early-stage processes during fibrogenesis in a mouse model of systemic sclerosis: A pilot study. *Front. Immunol.* 8, 1242. doi:10.3389/fimmu.2017.01242

Schwörer, S., Berisa, M., Violante, S., Qin, W., Zhu, J., Hendrickson, R. C., et al. (2020). Proline biosynthesis is a vent for TGFβ-induced mitochondrial redox stress. *EMBO J.* 39 (8), e103334. doi:10.15252/embj.2019103334

Sharabi, A., and Tsokos, G. C. (2020). T cell metabolism: New insights in systemic lupus erythematosus pathogenesis and therapy. *Nat. Rev. Rheumatol.* 16 (2), 100–112. doi:10.1038/s41584-019-0356-x

Škovierová, H., Vidomanová, E., Mahmood, S., Sopková, J., Drgová, A., Červeňová, T., et al. (2016). The molecular and cellular effect of homocysteine metabolism imbalance on human health. *Int. J. Mol. Sci.* 17 (10), 1733. doi:10.3390/ijms17101733

Slack, M., Wang, T., and Wang, R. (2015). T cell metabolic reprogramming and plasticity. *Mol. Immunol.* 68 (2), 507–512. doi:10.1016/j.molimm.2015.07.036

Smeets, R. L., Kersten, B. E., Joosten, I., Kaffa, C., Alkema, W., Koenen, H. J. P. M., et al. (2020). Diagnostic profiles for precision medicine in systemic sclerosis; stepping forward from single biomarkers towards pathophysiological panels. *Autoimmun. Rev.* 19 (5), 102515. doi:10.1016/j.autrev.2020.102515

Smolenska, Z., Zabielska-Kaczorowska, M., Wojteczek, A., Kutryb-Zajac, B., and Zdrojewski, Z. (2020). Metabolic pattern of systemic sclerosis: Association of changes in plasma concentrations of amino acid-related compounds with disease presentation. *Front. Mol. Biosci.* 7, 585161. doi:10.3389/fmolb.2020.585161

Sun, C., Zhu, H., Wang, Y., Han, Y., Zhang, D., Cao, X., et al. (2022). Serum metabolite differences detected by HILIC UHPLC-Q-TOF MS in systemic sclerosis. *Clin. Rheumatol.* 42, 125–134. doi:10.1007/s10067-022-06372-z

Szamosi, S., Csiki, Z., Szomják, E., Szolnoki, E., Szőke, G., Szekanecz, Z., et al. (2009). Plasma homocysteine levels, the prevalence of methylenetetrahydrofolate reductase gene C677T polymorphism and macrovascular disorders in systemic sclerosis: Risk factors for accelerated macrovascular damage? *Clin. Rev. Allergy Immunol.* 36 (2–3), 145–149. doi:10.1007/s12016-008-8105-y

Tannahill, G. M., Curtis, A. M., Adamik, J., Palsson-McDermott, E. M., McGettrick, A. F., Goel, G., et al. (2013). Succinate is an inflammatory signal that induces IL-1β through HIF-1α. *Nature* 496 (7444), 238–242. doi:10.1038/nature11986

Teufel, R., Mascaraque, V., Ismail, W., Voss, M., Perera, J., Eisenreich, W., et al. (2010). Bacterial phenylalanine and phenylacetate catabolic pathway revealed. *Proc. Natl. Acad. Sci.* 107 (32), 14390–14395. doi:10.1073/pnas.1005399107

Thakkar, V., Stevens, W., Prior, D., Rabusa, C., Sahhar, J., Walker, J. G., et al. (2016). The role of asymmetric dimethylarginine alone and in combination with N-terminal pro-B-type natriuretic peptide as a screening biomarker for systemic sclerosis-related pulmonary arterial hypertension: A case control study. *Clin. Exp. Rheumatol.* 1, 129–136. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27214686.

Tikly, M., Channa, K., Theodorou, P., and Gulumian, M. (2006). Lipid peroxidation and trace elements in systemic sclerosis. *Clin. Rheumatol.* 25 (3), 320–324. doi:10.1007/s10067-005-0013-4

Tokumura, A., Carbone, L. D., Yoshioka, Y., Morishige, J., Kikuchi, M., Postlethwaite, A., et al. (2009). Elevated serum levels of arachidonoyl-lysophosphatidic acid and sphingosine 1-phosphate in systemic sclerosis. *Int. J. Med. Sci.* 6, 168–176. doi:10.7150/ijms.6.168

Tsou, P. S., Varga, J., and O'Reilly, S. (2021). Advances in epigenetics in systemic sclerosis: Molecular mechanisms and therapeutic potential. *Nat. Rev. Rheumatol.* 17 (10), 596–607. doi:10.1038/s41584-021-00683-2

Tyndall, A. J., Bannert, B., Vonk, M., Airo, P., Cozzi, F., Carreira, P. E., et al. (2010). Causes and risk factors for death in systemic sclerosis: A study from the EULAR scleroderma trials and research (EUSTAR) database. *Ann. Rheum. Dis.* 69 (10), 1809–1815. doi:10.1136/ard.2009.114264

Ung, C. Y., Onoufriadis, A., Parsons, M., McGrath, J. A., and Shaw, T. J. (2021). Metabolic perturbations in fibrosis disease. *Int. J. Biochem. Cell. Biol.* 139, 106073. doi:10.1016/j.biocel.2021.106073

Vazquez-de-Lara, L., Tlatelpa-Romero, B., Romero, Y., Fernández-Tamayo, N., Vazquez-de-Lara, F., Justo-Janeiro, M. J., et al. (2018). Phosphatidylethanolamine induces an antifibrotic phenotype in normal human lung fibroblasts and ameliorates bleomycin-induced lung fibrosis in mice. *Int. J. Mol. Sci.* 19 (9), 2758. doi:10.3390/ijms19092758

Volpe, A., Biasi, D., Caramaschi, P., Mantovani, W., Bambara, L. M., Canestrini, S., et al. (2006). Levels of F2-isoprostanes in systemic sclerosis: Correlation with clinical features. *Rheumatology* 45 (3), 314–320. doi:10.1093/rheumatology/kei151

Wang, T., Liu, H., Lian, G., Zhang, S. Y., Wang, X., and Jiang, C. (2017). HIF1 -induced glycolysis metabolism is essential to the activation of inflammatory macrophages. *Mediat. Inflamm.* 2017, 9029327. doi:10.1155/2017/9029327

Wang, X., Chen, J., Xu, J., Xie, J., Harris, D. C. H., and Zheng, G. (2021b). The role of macrophages in kidney fibrosis. *Front. Physiol.* 12, 12. doi:10.3389/fphys.2021.705838

Wang, Z., Chen, L., Huang, Y., Luo, M., Wang, H., Jiang, Z., et al. (2021a). Pharmaceutical targeting of succinate dehydrogenase in fibroblasts controls bleomycin-induced lung fibrosis. *Redox Biol.* 46, 102082. doi:10.1016/j.redox.2021.102082

Wei, L., Abraham, D., and Ong, V. (2022). The yin and Yang of IL-17 in systemic sclerosis. *Front. Immunol.* 13, 885609. doi:10.3389/fimmu.2022.885609

Wilson, M. S., Madala, S. K., Ramalingam, T. R., Gochuico, B. R., Rosas, I. O., Cheever, A. W., et al. (2010). Bleomycin and IL-1beta-mediated pulmonary fibrosis is IL-17A dependent. *J. Exp. Med.* 207 (3), 535–552. doi:10.1084/jem.20092121

Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* 15, 473–484. doi:10.1038/nrd.2016.32

Xie, N., Cui, H., Ge, J., Banerjee, S., Guo, S., Dubey, S., et al. (2017). Metabolic characterization and RNA profiling reveal glycolytic dependence of profibrotic phenotype of alveolar macrophages in lung fibrosis. *Am. J. Physiology-Lung Cell. Mol. Physiology* 313 (5), L834-L844–44. doi:10.1152/ajplung.00235.2017

Xing, X., Yang, J., Yang, X., Wei, Y., Zhu, L., Gao, D., et al. (2013). IL-17A induces endothelial inflammation in systemic sclerosis via the ERK signaling pathway. *PLoS One* 8 (12), e85032. doi:10.1371/journal.pone.0085032

Yang, X. Y., Zheng, K D., Zheng, G., Zou, H., Wang, J. M., et al. (2015). Energy metabolism disorder as a contributing factor of rheumatoid arthritis: A comparative proteomic and metabolomic study. *PLoS One* 10 (7), e0132695. doi:10.1371/journal.pone.0132695

Young, A., and Khanna, D. (2015). Systemic sclerosis: Commonly asked questions by rheumatologists. *JCR J. Clin. Rheumatology* 21 (3), 149–155. doi:10.1097/RHU.0000000000000232

Zhang, A., Sun, H., Yan, G., Wang, P., and Wang, X. (2015a). Metabolomics for biomarker discovery: Moving to the clinic. *Biomed. Res. Int.* 2015, 354671. doi:10.1155/2015/354671

Zhang, L., Wan, Y. N., Zhao, J. H., Wang, Y. J., Wangxin, Y., Yan, J. W., et al. (2015b). The association between systemic sclerosis, arginine and asymmetric dimethylarginine. *Inflammation* 38 (1), 218–223. doi:10.1007/s10753-014-0025-9

Zhang, Y. J., Zhang, L., Huang, X. L., Duan, Y., Yang, L. J., and Wang, J. (2018). The association between homocysteine and systemic sclerosis: A review of the literature and meta-analysis. *Mod. Rheumatol.* 28 (4), 681–689. doi:10.1080/14397595.2017.1386844

Zhao, G., He, F., Wu, C., Li, P., Li, N., Deng, J., et al. (2018). Betaine in inflammation: Mechanistic aspects and applications. *Front. Immunol.* 9, 1070. doi:10.3389/fimmu.2018.01070

Zhenyukh, O., Civantos, E., Ruiz-Ortega, M., Sánchez, M. S., Vázquez, C., Peiró, C., et al. (2017). High concentration of branched-chain amino acids promotes oxidative stress, inflammation and migration of human peripheral blood mononuclear cells via mTORC1 activation. *Free Radic. Biol. Med.* 104, 165–177. doi:10.1016/j.freeradbiomed.2017.01.009

Zhong, Z., Wheeler, M. D., Li, X., Froh, M., Schemmer, P., Yin, M., et al. (2003). L-Glycine: A novel antiinflammatory, immunomodulatory, and cytoprotective agent. *Curr. Opin. Clin. Nutr. Metab. Care* 6 (2), 229–240. doi:10.1097/00075197-200303000-00013

Zhou, J., and Zhong, L. (2022). Applications of liquid chromatography-mass spectrometry based metabolomics in predictive and personalized medicine. *Front. Mol. Biosci.* 9, 1049016. doi:10.3389/fmolb.2022.1049016

Zhu, H., Chen, W., Liu, D., and Luo, H. (2019). The role of metabolism in the pathogenesis of systemic sclerosis. *Metabolism* 93, 44–51. doi:10.1016/j.metabol.2018.12.004

Zou, M. H. (2015). Tryptophan-kynurenine pathway is dysregulated in inflammation and immune activation. *Front. Biosci.* 20 (7), 1116–1143. doi:10.2741/4363

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 25 September 2023
DOI 10.3389/fmolb.2023.1206074

OPEN ACCESS

EDITED BY
Guillermo Moyna,
Universidad de la República, Uruguay

REVIEWED BY
Hugo Cerecetto,
Universidad de la República, Uruguay
Angel Marcelo Padilla,
University of Georgia, United States

*CORRESPONDENCE
Claudia Cuervo,
claudia.cuervo@javeriana.edu.co
Jorge Robles,
jrobles@javeriana.edu.co

RECEIVED 14 April 2023
ACCEPTED 13 September 2023
PUBLISHED 25 September 2023

CITATION
Pardo-Rodriguez D, Lasso P, Santamaría-Torres M, Cala MP, Puerta CJ, Méndez Arteaga JJ, Robles J and Cuervo C (2023), *Clethra fimbriata* hexanic extract triggers alteration in the energy metabolism in epimastigotes of *Trypanosoma cruzi*.
Front. Mol. Biosci. 10:1206074.
doi: 10.3389/fmolb.2023.1206074

Clethra fimbriata hexanic extract triggers alteration in the energy metabolism in epimastigotes of *Trypanosoma cruzi*

Daniel Pardo-Rodriguez[1,2,3,4], Paola Lasso[5], Mary Santamaría-Torres[4], Mónica P. Cala[4], Concepción J. Puerta[1], Jonh Jairo Méndez Arteaga[3], Jorge Robles[2]* and Claudia Cuervo[1]*

[1]Grupo de Enfermedades Infecciosas, Pontificia Universidad Javeriana, Bogotá, Colombia, [2]Grupo de Fitoquímica, Pontificia Universidad Javeriana, Bogotá, Colombia, [3]Grupo de Productos Naturales, Universidad del Tolima, Tolima, Colombia, [4]Metabolomics Core Facility—MetCore, Vice-Presidency for Research, Universidad de los Andes, Bogotá, Colombia, [5]Grupo de Inmunobiología y Biología Celular, Pontificia Universidad Javeriana, Bogotá, Colombia

Chagas disease (ChD), caused by *Trypanosoma cruzi*, is endemic in American countries and an estimated 8 million people worldwide are chronically infected. Currently, only two drugs are available for therapeutic use against *T. cruzi* and their use is controversial due to several disadvantages associated with side effects and low compliance with treatment. Therefore, there is a need to search for new tripanocidal agents. Natural products have been considered a potential innovative source of effective and selective agents for drug development to treat *T. cruzi* infection. Recently, our research group showed that hexanic extract from *Clethra fimbriata* (CFHEX) exhibits anti-parasitic activity against all stages of *T. cruzi* parasite, being apoptosis the main cell death mechanism in both epimastigotes and trypomastigotes stages. With the aim of deepening the understanding of the mechanisms of death induced by CFHEX, the metabolic alterations elicited after treatment using a multiplatform metabolomics analysis (RP/HILIC-LC-QTOF-MS and GC-QTOF-MS) were performed. A total of 154 altered compounds were found significant in the treated parasites corresponding to amino acids (Arginine, threonine, cysteine, methionine, glycine, valine, proline, isoleucine, alanine, leucine, glutamic acid, and serine), fatty acids (stearic acid), glycerophospholipids (phosphatidylcholine, phosphatidylethanolamine and phosphatidylserine), sulfur compounds (trypanothione) and carboxylic acids (pyruvate and phosphoenolpyruvate). The most affected metabolic pathways were mainly related to energy metabolism, which was found to be decrease during the evaluated treatment time. Further, exogenous compounds of the triterpene type (betulinic, ursolic and pomolic acid) previously described in *C. fimbriata* were found inside the treated parasites. Our findings suggest that triterpene-type compounds may contribute to the activity of CFHEX by altering essential processes in the parasite.

KEYWORDS

Chagas disease, hexanic extract of Clethra fimbriata, energy metabolism, multiplatform untargeted metabolomics, triterpenes, Trypanosoma cruzi

1 Introduction

The parasite *Trypanosoma cruzi* is the etiological agent of Chagas disease (ChD), a neglected tropical disease that affects more than 8 million people worldwide (WHO, 2021). Although, ChD is endemic to American countries, population movement has led to its spread to non-endemic countries (Klein et al., 2012), making it a global public health concern (Rassi et al., 2010; Tarleton, 2016; Pérez-Molina and Molina, 2017). In Colombia, it has been reported that approximately 436,000 people are infected and that about 11% of the population is at risk of acquiring the infection (Rassi et al., 2010; World Health Organization, 2015; Olivera et al., 2019).

There is no vaccine available for ChD, but there have been two drugs used since the 1970s: Nifurtimox (NFX) and Benznidazole (BNZ) (Nunes et al., 2013; Bustamante and Tarleton, 2014; Bern, 2015). However, the difficulty in conducting clinical trials during the chronic phase of the infection has made it challenging to determine the effectiveness of treatment during this phase (Urbina, 2010; Urbina, 2015). Additionally, the treatment with NFX and BNZ is associated with several issues, including high toxicity, side effects and a prolonged treatment time (Bern, 2015; Manne-Goehler et al., 2016). Furthermore, the presence of *T. cruzi* isolates with different degrees of susceptibility to these drugs has been reported (Castro et al., 2006; Mejía-Jaramillo et al., 2012). In addition to this, it has been found that *T. cruzi* has the ability to enter a dormant state known as non-replicating amastigote, which allows the parasite to resist the pharmacological stress induced by BNZ (Sánchez-Valdéz et al., 2018). Therefore, the development of safer and more efficient therapeutic alternatives for the treatment of ChD is essential.

Plants have been used for a long time in the treatment of multiple diseases and recently have gained renewed interest as a starting point to propose new natural products with specific bioactivities (Schmidt et al., 2012; Lopera Valle et al., 2013). In fact, an estimated 60% of currently available drugs are derived from natural products (Newman and Cragg, 2016; Newman and Cragg, 2020), suggesting the importance of natural sources in drug discovery. Colombia's distinct geographical location affords a varied range of ecosystems that support one of the world's highest diversity and dispersion of animals, fungi, and plants (Gori et al., 2022). As a result, the country is considered a promising source of chemical structures with specific biological activities.

Our research group recently evaluated the trypanosomicidal effects of the native Colombian plant *Clethra fimbriata*, finding that the ethanolic and hexanic extracts are effective against the different stages of *T. cruzi*. Further, the extracts induce the production of cytokines and cytotoxic molecules in $CD4^+$ and $CD8^+$ T cells from healthy donors, an effect that may be associated with the high content of pentacyclic triterpenes found in *C. fimbriata* (Castañeda et al., 2021; Pardo-Rodriguez et al., 2022). Based on these findings, this research focused on associating metabolic alterations and death mechanisms induced after treatment of epimastigotes with the CFHEX extract, using a multiplatform untargeted metabolomics approach.

2 Materials and methods

2.1 Plant material and extraction

Plant material was collected under the "Permit for wild species specimen collection of biological diversity for research with non-commercial purposes" (Permiso marco de recolección de especímenes de especies silvestres de la diversidad biológica para investigación con fines no comerciales) granted to the Pontificia Universidad Javeriana (Resolution 778 of 7 July 2017) issued by the "National Environmental Licensing Authority". *C. fimbriata* was collected in Majuy Hill, Via Cota, Cundinamarca, Colombia, and taxonomically identified by the Colombian National Herbarium (voucher specimen number COL 610805). *C. fimbriata* aerial parts were dried and crushed, followed by extraction by successive maceration (Five extractions) with 1:10 sample to solvent ratio, using hexane (CFHEX). Obtained extract was concentrated by evaporation *in vacuo*. Prior to biological tests, the extract was resuspended in ethanol.

2.2 Parasite maintenance

T. cruzi Y-strain epimastigotes (MHOM/BR/00/Y); a discrete typing unit (DTU TcII) (Pavia et al., 2012), were maintained in the exponential growth phase in Liver Infusion Tryptose (LIT) medium supplemented with 15% heat inactivated fetal bovine serum (FBSi) (Eurobio), 100 U/mL penicillin and 100 µg/mL streptomycin (Eurobio), at 26°C.

2.3 Sample preparation

Epimastigotes of *T. cruzi* were cultured in LIT medium supplemented with 15% FBSi at 26 °C. Once they reached an exponential growth phase, 1×10^8 parasites were transferred to fresh culture medium, and incubated with the IC_{90} (690 µg/mL) of the CFHEX for 36 h. As a negative control the parasites were incubated with fresh culture medium. After the incubation time, the parasites were washed three times with phosphate buffered saline (PBS) at 4 °C and immediately frozen in liquid nitrogen and kept at −80 °C until further processing. Each treatment was evaluated in six independent biological replicates.

2.3.1 Metabolite extraction

For the extraction of metabolites from the treated and control parasites, 500 µL of a solution of MeOH-water (4:1 v/v) were added to each of the samples. Then, two-3 mm tungsten carbide beads were added to each of the cryovials and vortexed at 3,200 rpm for 1 min. Subsequently, the samples were taken to the Tissue Lyser to perform 5 cycles of 30 Hz for 1 min. The samples were centrifuged at 15,700 *g*, at 4 °C for 20 min and filtered through 0.22 µm filters. Finally, 200 µL of the extracts were taken, which were used for subsequent analysis by reverse phase liquid chromatography coupled to mass spectrometry with a time-of-flight analyzer (RP-LC-QTOF/MS) in positive and negative polarity and hydrophilic interaction chromatography coupled to mass spectrometry with a time-of-flight analyzer (HILIC-LC-QTOF-MS) (Rojo et al., 2015).

2.3.2 Untargeted metabolomics by RP-LC-QTOF-MS

Samples were analyzed using an Agilent Technologies 1,260 liquid chromatography system coupled to a 6545 Q-TOF quadrupole time-of-flight mass analyzer with electrospray ionization. Five µL of the sample were injected onto a C_{18} column (InfinityLab Poroshell 120 EC-C18 (100 × 3.0 mm, 2.7 µm)) at 30 °C and compound gradient elution: 0.1% (v/v) of formic acid in Milli-Q® water (Phase A) and 0.1% (v/v) of formic acid in acetonitrile (Phase B) with a constant flow of 0.4 mL/min. The elution gradient started at 25% with respect to B and increased over 35 min to 95% B. Finally, the gradient decreased to 36% B over a period of 1 min and was maintained for a further 9 min until the system was rebalanced. Detection by mass spectrometry was performed in positive and negative ESI mode in full scan from 100 to 1,100 m/*z*. Throughout the analysis, two reference masses were used for mass correction: *m/z* 121.0509 $[C_5H_4N_4]^+$, *m/z* 922.0098 $[C_{18}H_{18}O_6N_3P_3F_{24}]^+$ in positive mode and *m/z* 112.9856 $[C_2O_2F_3(NH_4)]^-$, *m/z* 1,033.9881 [($C_{18}H_{18}O_6N_3P_3F_{24}$+trifluoroacetic acid)-H]$^-$ in negative mode.

2.3.3 Untargeted metabolomics by HILIC-LC-QTOF-MS

Two µL of the sample was injected onto a Kinetex HILIC 100 A column (150 × 3.0 mm, 2.6 µm) at 40 °C and a gradient elution composed of: 10 mM ammonium acetate in acetonitrile:water (50:50) (Phase A) and 10 mM ammonium acetate in acetonitrile: water (95:5) (Phase B) with a constant flow of 0.4 mL/min. The elution gradient started at 99% with respect to B and decreased for 15 min until reaching 50% B, where it was maintained for 1 min. Finally, the gradient increased to 99% B and was maintained for an additional 6 min until the system re-equilibrated. Detection by mass spectrometry was performed in negative ESI mode in full scan from 50 to 1,100 m/*z*. Throughout the analysis, two reference masses were used for mass correction: *m/z* 112.9856 $[C_2O_2F_3(NH_4)]^-$ and *m/z* 1033.9881 [$C_{18}H_{18}O_6N_3P_3F_{24}$+trifluoroacetic acid)-H]$^-$.

2.3.4 Untargeted metabolomics by GC-QTOF-MS

One hundred µL of the extracts were dried in a speedvac for 3 h at 35 °C. 10 µL of O-methoxyamine in pyridine (15 mg/mL) were added and vortexed at 3,200 rpm for 10 min. Subsequently, samples were kept in the dark for 16 h and 10 µL of *N,O*-Bistrifluoroacetamide with 1% trimethylsilyl chloride were added and incubated at 70 °C for 1 h. Finally, the samples were allowed to cool to room temperature for 30 min, 100 µL of methyl stearate in heptane as internal standard (10 mg/L) were added and the blend was vortexed for 10 min at 3,200 rpm. The derivatized samples were immediately analyzed according to the following methodology.

For data acquisition, an Agilent Technologies 7890B gas chromatograph coupled to an Agilent Technologies GC/Q-TOF 7250 time-of-flight mass selective detector, equipped with a split/splitless injection port (250 °C, ratio split 30) and an Agilent Technologies 7693A autosampler. The electron ionization (EI) source was operated at 70 eV. An Agilent Technologies J&W HP-5MS column (30 m, 0.25 mm, 0.25 µm) was used. The carrier gas flow was helium at a constant flow of 0.7 mL/min. The oven temperature was programmed from 60 °C (1 min) to 325 °C (10 min). The temperature of the transfer line to the detector, the source filament and the quadrupole were maintained at 280 °C, 230 °C and 150 °C, respectively. Detection by mass spectrometry was carried out between 50 and 600 m/*z* at a speed of 5 spectra/min.

2.3.5 Quality controls samples (QC)

Quality control (QC) samples were prepared by mixing equal volumes of the metabolic extract from each sample. Subsequently, the preparation and analysis of the QC samples were performed following the procedures described above in each of the analytical platforms. To determine the reproducibility of sample preparation and the stability of the analytical platform used, several QC elutions were performed until the analytical system equilibrated. Subsequently, the QC samples were analyzed every three randomly injected samples.

2.4 Data processing and analysis

The data obtained by LC-MS was processed using the Agilent MassHunter Profinder program for deconvolution, alignment and integration of the data using the recursive and molecular feature extraction algorithms. Treatment of the GC-MS data obtained consisted of deconvolution and identification of the metabolites using the Agilent MassHunter Unknowns Analysis program and the Fiehn and NIST libraries. Then the alignment of the retention times was carried out in the Agilent Mass Profiler Professional program, the results were exported to the Agilent MassHunter Quantitative program for data integration. Finally, the data obtained from GC-MS and LC-MS data processing were manually inspected. Then, the data was filtered by presence and by reproducibility, keeping only the metabolites present or absent in 100% of the samples belonging to the same group and with a coefficient of variation in the QC of less than 20%.

The identification of the molecular characteristics with statistically significant differences between the two groups (treated and untreated parasites) was carried out using univariate (UVA) and multivariate (MVA) statistical analysis. Regarding the UVA analysis, the *p-value* was determined by nonparametric tests (Mann-Whitney U test) using SIMCA-P + 16.0 (Umetrics). For the MVA analyses, an unsupervised principal component analysis (PCA) was performed to observe the unsupervised distribution of the analyzed samples. Subsequently, supervised orthogonal partial least squares discriminant analysis (OPLS-DA) models were performed to select the molecular features responsible for the separation between the groups. The MVA was performed using the SIMCA-P+16.0 software (Umetrics). The selected statistically significant molecular characteristics met at least one of the following requirements: 1) UVA: *p-value* < *0.05* and 2) MVA: Variance Important in Projection (VIP) > 1.

2.4.1 Annotation of statistically significant molecular features

The metabolites obtained by GC-MS analysis were identified using the Fiehn version 2013 libraries and MassHunter Personal Compound Database and Library Manager Software B.08.00. Whereas the significant characteristics obtained by LC-MS were putatively assigned in the CEU Mass Mediator annotator (http://ceumass.eps.uspceu.es) by matching the exact observed mass of each compound with the available *m/z* values. online at METLIN (http://metlin.scripps.

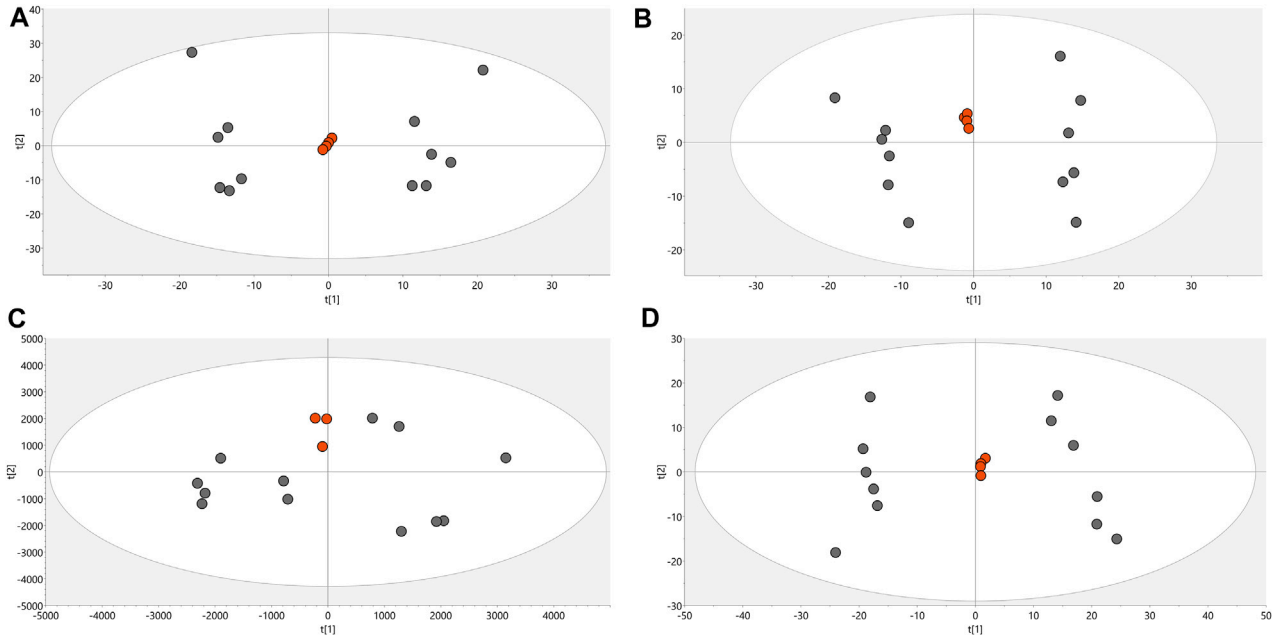


FIGURE 1
PCA score charts. **(A)**. LC/MS (+) R^2: 0.662, Q^2: 0.323. **(B)**. LC/MS (–) R^2: 0.841, Q^2: 0.654. **(C)**. GC/MS(+) R^2: 0.532, Q^2: 0.177. **(D)**. HILIC/MS(–) R^2: 0.813, Q^2: 0.66. Dots in orange color denote quality control, gray dots correspond to samples.

edu), KEGG (http://genome.jp/kegg), and LIPIDMAPS (http://lipidMAPS.org), using the following likely adducts: $[M + H]^+$, $[M + H\text{-}[H_2O]]^+$, $[M + Na]^+$, and $[M\text{-}H]^-$, $[M + \text{Formic acid-}H]^-$, $[M + Cl]^-$, $[M\text{-}H\text{-}[H_2O]]^-$ for positive and negative ionization modes, respectively. Furthermore, to confirm the identity of the metabolite, MS/MS analysis was performed.

2.4.2 Altered metabolite pathway mapping

The analysis of the affected metabolic pathways in the treated parasites was performed using the "Pathway Analysis" tool of the MetaboAnalyst 5.0 server (http://www.metaboanalyst.ca/). For which, the altered compounds were annotated and compared with the *Trypanosoma brucei* (KEGG) metabolome available on the same server.

3 Results

Conditions associated with exposure to CFHEX (IC_{90}: 690 µg/mL) in the epimastigote stage of *T. cruzi* were evaluated using untargeted metabolomics analysis. A multiplatform approach was used to detect the largest possible number of altered metabolites. The performance of the different analytical platforms was evaluated using unsupervised PCA models. The inspection of the clusters evidenced the clear grouping of the samples belonging to the quality control in the different analytical platforms used (Figure 1, orange dots). After verifying the performance of each analytical platform, the supervised orthogonal partial least squares regression method (OPLS-DA) was implemented to maximize the differences between the group made up of treated parasites and the group of untreated parasites (Figure 2) and to identify the molecular features with greater weight in the separation of the groups. The Pareto scaling method was used before the statistical analysis.

The OPLS-DA scoring plot presented in Figure 2 showed a clear separation between the groups: CFHEX-treated parasites (yellow dots) and untreated parasites (violet dots). Likewise, the variables R^2 and Q^2, which measure the goodness of fit and the predictive capacity of the model developed from the data matrix, respectively, show adequate settings ($R^2 > 0.402$) and good predictive capacity ($Q^2 > 0.961$) (Triba et al., 2015). Finally, to evaluate the reliability of the models, the variance cross-validation (CV-ANOVA) was performed, evidencing highly significant models in the four platforms analyzed (CV-ANOVA <0.05) (Eriksson et al., 2008). The individual differentiating metabolites were determined by a combination of MVA (VIP>1) and UVA ($p < 0.05$), obtaining a total of 154 altered compounds in the treated parasites, of which 25.16%, 30.97%, 14.84% and 12.26% were identified by LC/MS–, LC/MS+, GC/MS and HILIC/MS, respectively. In addition, the identification of some metabolites was achieved simultaneously by several platforms: 11.61% of compounds were found by $LC/MS^{-/+}$, 1.29% by LC/MS^+ and GC/MS, 1.94% by LC/MS^- and HILIC/MS, and 1.94% by $LC/MS^{-/+}$ and HILIC/MS.

The consolidated analysis of the metabolic modifications of the treated parasites showed 57.41% of the metabolites increased and 42.58% decreased. From these, 77.41% of the fluctuations were related to lipid chemical classes as follows: glycerophospholipids (65 compounds, 42.58%), fatty acids (36 compounds, 23.22%), sphingolipids (9 compounds, 5.81%), glycerolipids (6 compounds, 3.87%) and steroidal lipids (3 compounds, 1.93%). The remaining 22.58% of the changed metabolites belonged to chemical classes such as organic acids (21 compounds, 13.54%), nucleosides (5 compounds, 3.22%), oxygenated organic compounds (4 compounds, 2.58%), among others (5 compounds, 3.22%) (Figure 3).

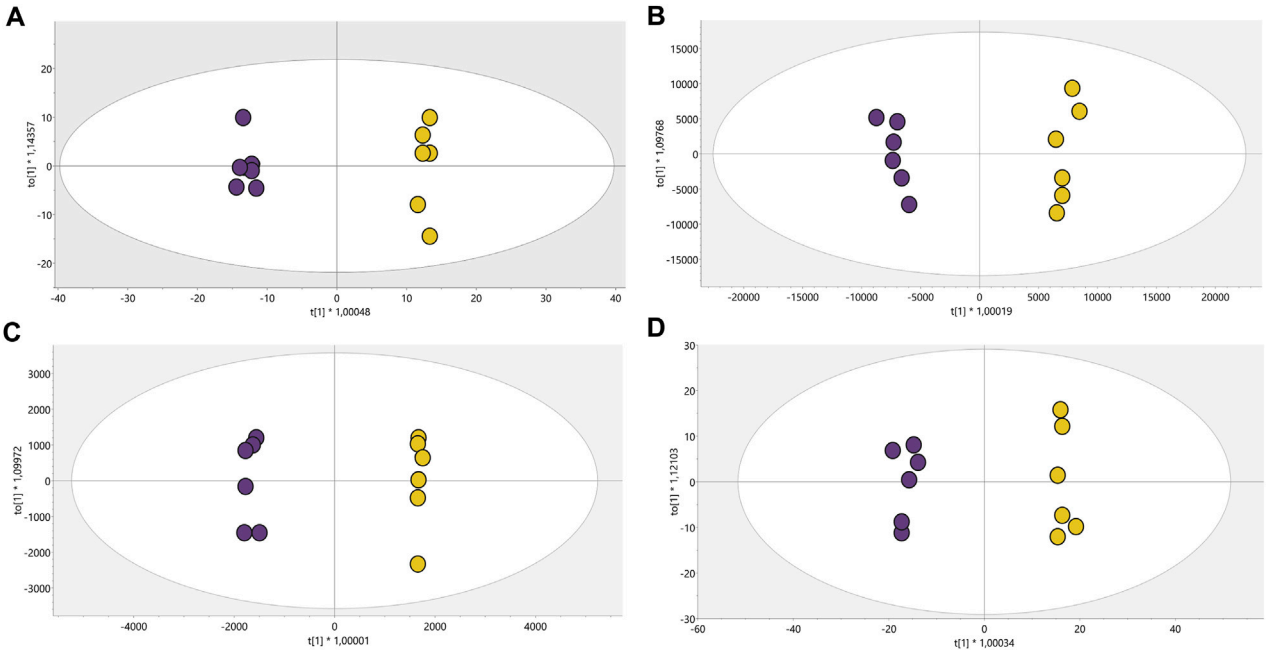


FIGURE 2
OPLS-DA models with Pareto scaling for metabolic analysis of treated parasites and control group. **(A)**. LC/MS (+): R^2: 0.402, Q^2: 0.961, CV ANOVA: $5.2e^{-5}$, **(B)**. LC/MS(–): R^2: 0.789, Q^2: 0.976, CV ANOVA: $1.0e^{-5}$, **(C)**. GC/MS: R^2: 0.715, Q^2: 0.978, CV ANOVA: $5.7e^{-4}$, **(D)**. HILIC/MS: R^2: 0.713, Q^2: 0.982, CV ANOVA: $3.2e^{-6}$. Dots in yellow color denote parasites treated with CFHEX and the violet dots the control group (untreated parasites).

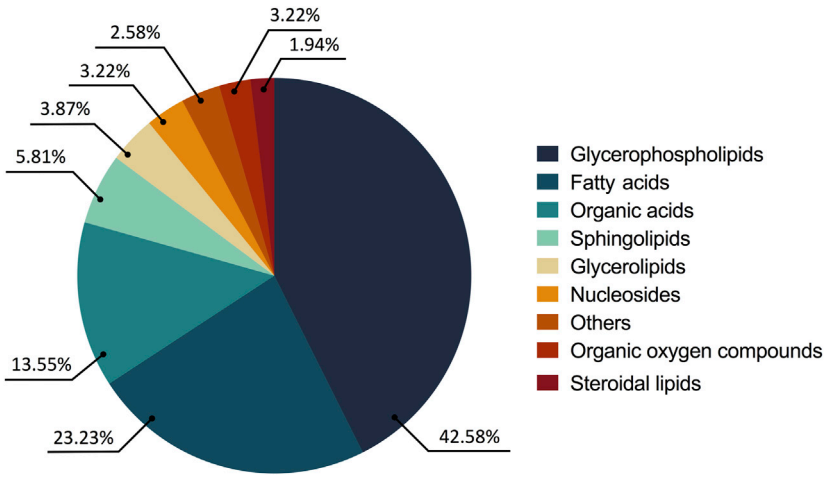


FIGURE 3
Altered metabolites in the treated parasites according to their chemical classes. The chemical classes are shown according to the color code.

Due to the number of modified metabolites found in parasites treated with CFHEX, it was decided to group them into two large groups. Supplementary Tables S1 and S2 summarize the lipidic and non-lipidic metabolites identified in the treated parasites and present information regarding retention times, coefficient of variation of the chromatographic signal in the QC group, statistical parameters for its selection, probable adducts, fold change (FC), and type of confirmation, among others. Besides, the set of altered metabolites between the two groups was analyzed using heat maps that allow the visualization of metabolite patterns changing between the groups. Thus, blue colors indicate decreased metabolite levels and red colors indicate increased metabolites in treated parasites (Figure 4; Figure 5). Besides, lipid metabolism experienced the greatest variation, with the glycerophospholipid, fatty acid (FA) and sphingolipid classes being the largest representatives (Supplementary Table S1). The group of glycerophospholipids was mainly constituted by lysophospholipids (LPL), of which lysophosphatidylcholines (LPC),

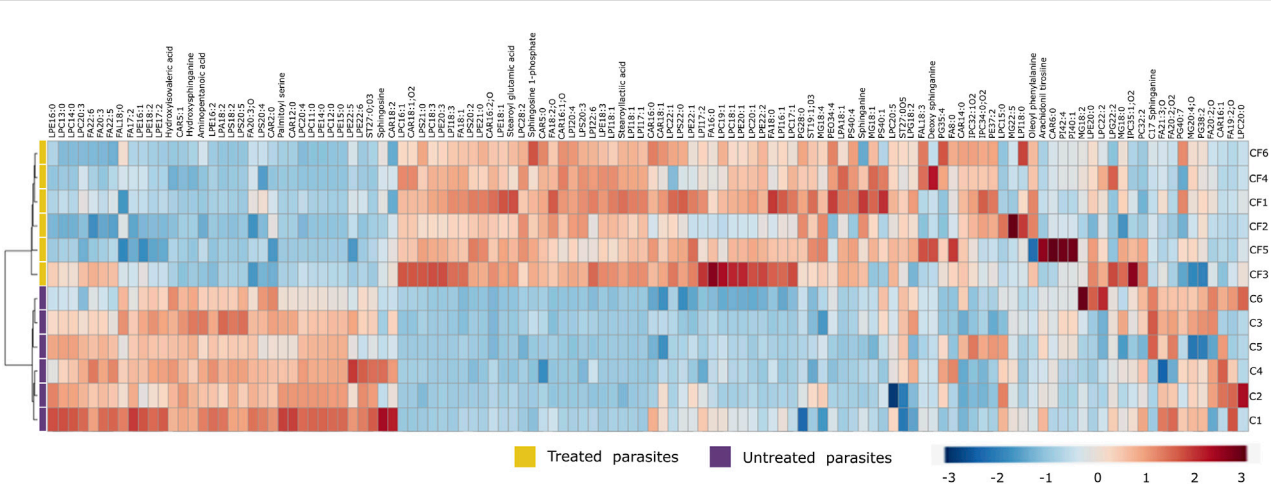


FIGURE 4
Heat map of lipid metabolites with statistically significant variation between *Trypanosoma cruzi* epimastigotes treated with CFHEX and the control group. The columns correspond to each altered metabolite identified, the rows correspond to the analyzed samples divided into the clades: CFHEX-treated parasites (yellow) and untreated parasites (violet). The level of variation is indicated on the right side on a color intensity scale representing relative abundance, where red colors denote metabolite increase and blue colors denote metabolite decrease.

lysophosphatidylethanolamines (LPE), lysophosphatidylinositols (LPI), lysophosphatidylglycerol (LPG) and lysophosphatidylserine (LPS), were observed to be increased (Figure 4, colored metabolites on the red color scale). The trends found in the FA subclass (fatty amides, and fatty esters (carnitines)), as well as in the sphingolipids subclass (phosphosphingolipids, sphingoid bases, steroidal lipids and glycerolipids) were also found to have an upward trend (Figure 4, metabolites on the red color scale).

On the other hand, non-lipid compounds were mainly grouped into the classes of nucleosides, acids, and oxygenated organic compounds (Supplementary Table S2). Particularly, in the group of nucleosides, metabolites such as uridine diphosphate (UDP), UDP-acetylgalactosamine, UDP-galactose, 5′-methylthioadenosine and adenosine were identified, all of which showed downward trends (Figure 5, metabolites in the blue color scale). Similar trends were also found in amino acids and peptides (proline, glycine, L-valine, L-isoleucine, serine, Pro-Pro, Pro-Phe, and Pro-Pro-Phe), carboxylic acids (pyruvate and phosphoenolpyruvate), oxygenated organic compounds (glycerate 3 phosphate, glyceric acid, and D-alose), and in other metabolites such as trypanothione, protoporphyrin IX, aminopentanoic acid, hydroxyisovaleric acid, hydroxybutyric acid, xanthine, hypoxanthine and L-carnitine (Figure 5, metabolites in blue color scale).

After an individual analysis of the total altered metabolites in the CFHEX-treated parasites, it was found that the greatest upward changes (FC > 3) occurred in the following compounds: lysophosphatidic acid (18:1); lysophosphatidylcholines (22:1, 18:3, 16:1); lysophosphatidylethanolamines (21:0, 20:0, 20:3, 18:1, 18:3, 17:1), phosphatidylethanolamine (37:2); lysophosphatidylglycerol (28:0); lysophosphatidylinositols (20:4, 18:1, 18:3, 17:1) phosphatidylinositol (40:1); lysophosphatidylserines (21:0, 20:2, 20:3), phosphatidylserine (40:4); dodecanoylcarnitine (12:0), tetradecanoylcarnitine (14:0), palmitoylcarnitine (16:0), hydroxyhexadecanoylcarnitine (16:1), hydroxypalmitoleoylcarnitine (16:1; O), oleoylcarnitine (18:1), linoleoylcarnitine (18:2), carboxyheptadecanoylcarnitine (18:1; O_2); hexadecanoyl sphinganine phosphomyo-inositol (34:0; O_2), hexadecanoyl eicosaphingenine phosphomyo-inositol (36:1; O_2); hydroxyeicosadienoic acid (20:2; O), Prostaglandin F2α (20:2; O_2); cholestane derivatives (27:0; O_5, 27:0; O_3), oleoyl glycerol, stearoyl lactic acid, stearoyl glutamic acid, palmitoyl serine, arachidonoyl tyrosine and ethyl glycine.

In contrast, the greatest downward changes (FC < 0.5) were observed in the lipids: lysophosphatidylcholines (22:2, 20:4, 18:2, 14:0, 13:0, 12:0, 11:0), lysophosphatidylethanolamines (22:5, 15:0, 14:0), lysophosphatidylglycerol (18:2, 16:2), phosphatidylglycerols (40:7, 36:4), lysophosphatydilsernines (20:5, 18:2), ethylacryloylcarnitine (5:1); tetradecanoyl sphingenine phosphomyoinositol (31:1; O_2) and in the non-lipid metabolites L-carnitine, hydroxybutyric acid, hydroxyisovaleric acid, trypanothione, UDP-acetylgalactosamine, proline, and the peptides Pro-Phe and Pro-Pro-Phe.

The summary of the main routes affected in the treated parasites is presented in Figure 6. The results of the pathway enrichment analysis with the topology analysis show that while the metabolic pathways with the greatest displacement in the Y-axis, denoted in dark colors, indicate the greatest changes in the metabolic pathway according to the number of metabolites participating in each affected pathway, the displacement on the X-axis, related to the size of the circle, indicates the impact on the metabolic pathway according to the importance of the altered metabolite, measured according to the number of connections that it presents on the pathway (Figure 6) (Xia et al., 2011). The metabolic pathways that were observed to be the most affected by CFHEX treatment (denoted in the red color scale and larger circles) were related to amino acid metabolism, which was found to be preferentially decreased. Thus, the biosynthesis of transfer RNA stands out due to the alteration of glycine, serine, valine, alanine, isoleucine, leucine, proline, and glutamate aminoacyl-tRNAs. Changes in the metabolism of glycine, serine, threonine, alanine, aspartate, glutamate, cysteine, and methionine were

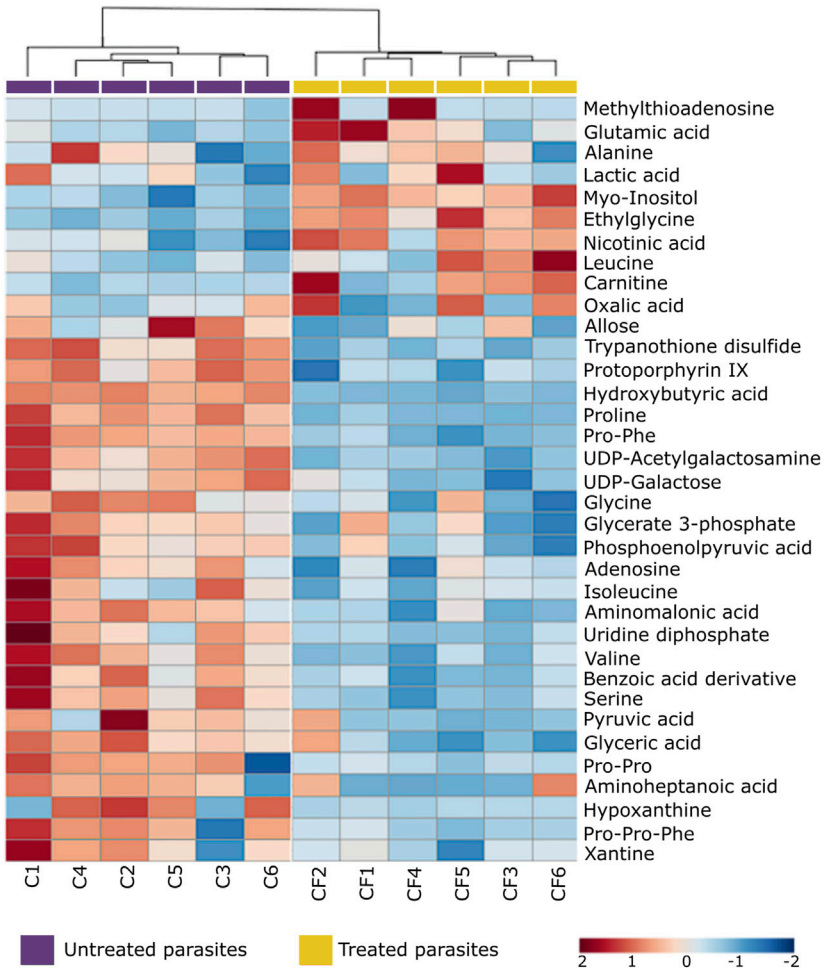


FIGURE 5
Heat map of non-lipid metabolites with statistically significant variation between *Trypanosoma cruzi* epimastigotes treated with CFHEX and the control group. The columns correspond to each altered metabolite identified, the rows correspond to the analyzed samples divided into the clades: CFHEX-treated parasites (yellow) and untreated parasites (violet). The level of variation is indicated on the right side on a color intensity scale representing relative abundance, where red colors denote metabolite increase and blue colors denote metabolite decrease.

also observed as well as on the arginine biosynthesis. Another altered pathway was glycolysis caused by the decrease of phosphoenolpyruvate, glycerate 3-phosphate, and pyruvate. Finally, the metabolism of sphingolipids and glycerophospholipids was also affected by the overproduction of sphingosine 1 phosphate, sphingosine, sphinganine, phosphatidic acid, phosphatidylcholines and phosphatidylethanolamines.

In addition to the metabolomic alterations described in the treated parasites, a series of compounds exogenous to *T. cruzi* were detected and found exclusively in the treated parasites and those would be constituents of the CFHEX extract (Pardo-Rodriguez et al., 2022). These compounds were: betulinic acid, ursolic acid, pomolic acid and their oxidized forms, betulonic acid, ursonic acid and pomonic acid. All these compounds are classified as pentacyclic triterpenes (compounds with a skeleton of 30 carbons that form five cycles) that are widely distributed in plant families and had been previously reported in *C. fimbriata* (Pardo-Rodriguez et al., 2022) (Figure 7).

4 Discussion

T. cruzi is a protozoan with a complex life cycle that involves vertebrate and invertebrate hosts and extracellular and intracellular stages (Rassi et al., 2010), exposing it, among other factors, to various sources of carbon as glucose and lipids in the mammalian host and amino acids, mainly proline in the insect vector (Cazzulo et al., 1985; Cazzulo, 1992; Michels et al., 2021). In particular, the epimastigote stage can use multiple carbon sources, however, in culture media it preferentially uses carbohydrates as a substrate for energy metabolism during the exponential phase and amino acids during the stationary phase (Cazzulo et al., 1985; Cazzulo, 1992; Silber et al., 2002; Barisón et al., 2017).

Glucose metabolism occurs in a similar way to other trypanosomatids, part of the enzymes of the glycolytic pathway in *T. cruzi* are compartmentalized within organelles of peroxisomal origin called glycosomes (Van Hellemond et al., 2005; Quiñones et al., 2020; Michels et al., 2021). Additionally, *T. cruzi* has a

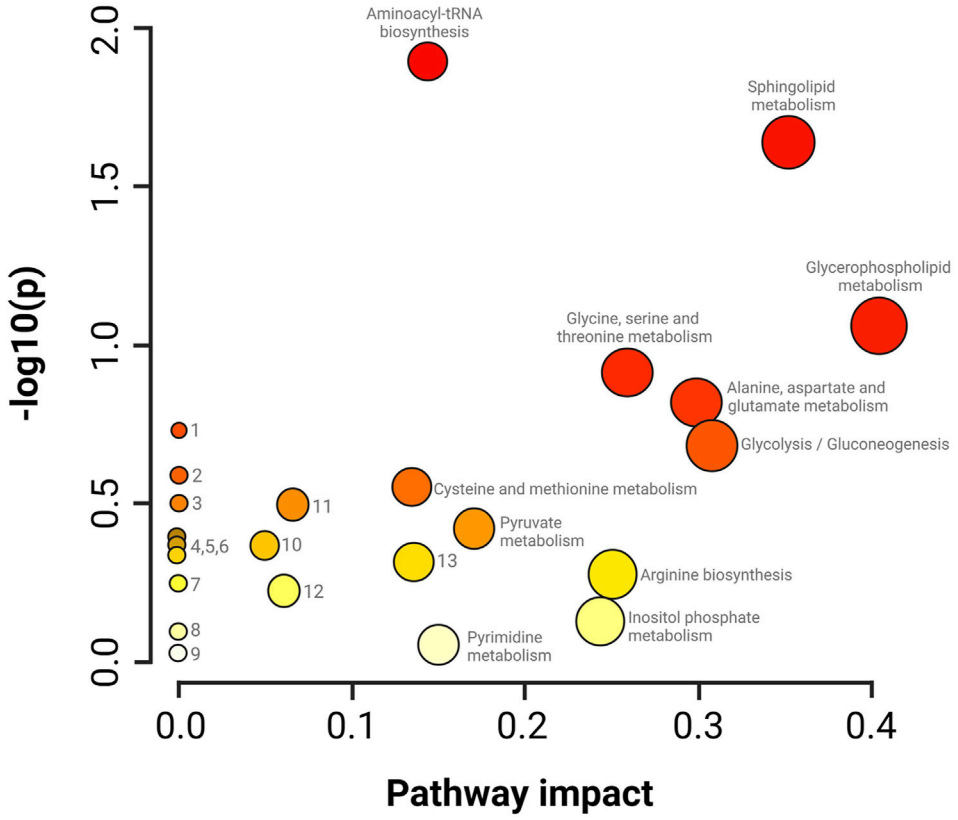


FIGURE 6
Impact analysis on metabolic pathways. enrichment (y-axis) and topology (x-axis) analysis. 1. Glyoxylate and dicarboxylate metabolism, 2. Taurine and hypotaurine metabolism, 3. Nitrogen metabolism, 4. Porphyrin metabolism, 5. Citrate cycle, 6. Valine, leucine and isoleucine degradation, 7. Nicotinate metabolism and nicotinamide, 8. Metabolism of lipoic acid, 9. Metabolism of aminosugars and sugar nucleotides, 10. Degradation of fatty acids, 11. Metabolism of arginine and proline, 12. Metabolism of purines, 13. Metabolism of glutathione. The darker colors indicate a greater number of changes in the metabolic pathway, while the size of the circle corresponds to the impact on the pathway according to the importance of the altered metabolite.

mitochondrial system that includes enzymes of the tricarboxylic acid cycle (TAC), as well as mitochondrial electron transport chain complexes (Stoppani et al., 1980; César Carranza et al., 2009), present at all stages of the parasite's life cycle. The glycolytic pathway is organized such that the first seven enzymes that catabolize glucose to glycerate 3-phosphate (G3P) are found within glycosomes, while the last three enzymes in the pathway reside in the cytosol (Bringaud et al., 2006; Michels et al., 2021), with the latter three enzymes being phosphoglycerate mutase, enolase and pyruvate kinase responsible for the transformation of G3P, phosphoenolpyruvate (PEP) and pyruvate (PYR), respectively.

G3P, PEP and PIR were found decreased after epimastigotes-CFHEX treatment, which may suggest energetic alterations in the treated parasites, as well as effects on metabolic pathways that use these substrates or derivatives of them. On the other hand, exposure of epimastigotes to CFHEX extract also reduced the levels of some nucleosides, which in normal conditions contribute to energy generation. The pathway that allows the formation of UDP-galactose is linked to the generation of glucose 6-phosphate and this last metabolite could be incorporated into the glycolytic pathway to generate ATP and feed other metabolic pathways such as pentose phosphate (Roper and Ferguson, 2003). Other biosynthetic pathways that participate as carbon and energy sources through the generation of intermediates for the TAC, such as the amino acids proline (Silber et al., 2002; Silber et al., 2005; Martins et al., 2009; Paes et al., 2013), valine, isoleucine, and serine (Hamptont, 1971; Sylvester and Krassner, 1976; Cazzulo, 1994; Silber et al., 2005; Manchola et al., 2016), were also found downward trends.

The effects observed on processes related to carbohydrate and amino acid-dependent energy production in epimastigotes are like those reported after nutritional stress induced by long periods of starvation. Souza et al. evaluated the role of FA oxidation after depriving *T. cruzi* epimastigotes cultures of glucose, finding that, in the absence of glucose, lipid droplets become the main sources of FA, which help the body survive nutritional stress by producing acetyl-CoA that fuels TAC, contributing to mitochondrial ATP production (Souza et al., 2021). In summary, epimastigotes use FA as carbon and energy source when glucose and amino acids are not available. In this context, the use of lipids as a source of carbon and energy in the treated parasites is suggested by the findings that the family of lipids conjugated to carnitine presented upward trends.

Interestingly, carnitines were the family of compounds that presented the greatest alteration in the study. Acyl-carnitines have the biological function of transporting FA into the mitochondria to be substrates for β-oxidation. To do this,

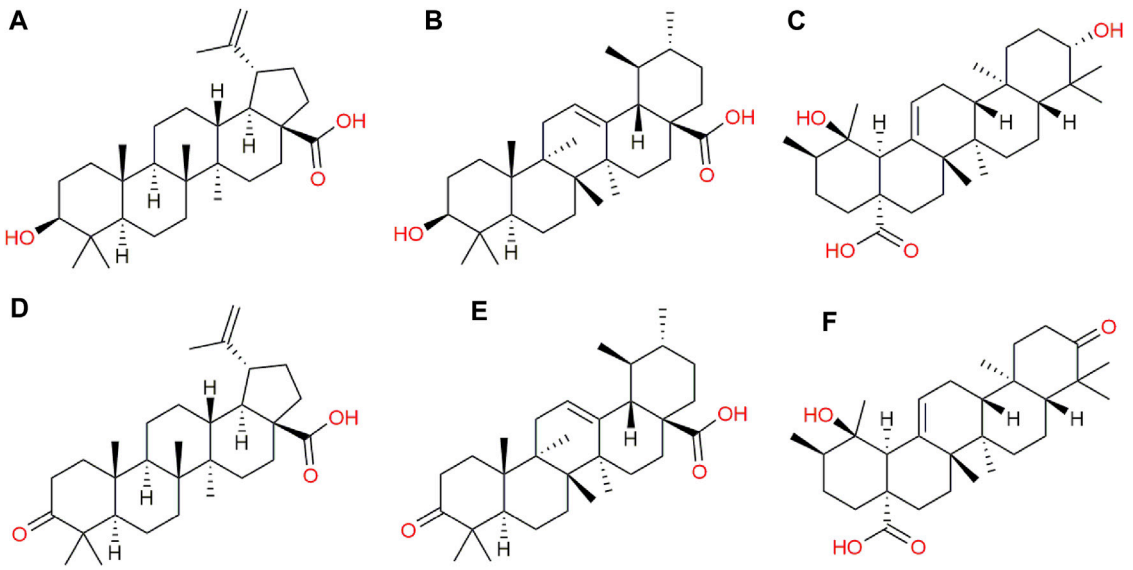


FIGURE 7
Exogenous metabolites found in epimastigotes treated with CFHEX extract. **(A)** Betulinic acid, **(B)** Ursolic acid, **(C)** Pomolic acid, **(D)** Betulonic acid, **(E)** Ursonic acid, **(F)** Pomonic acid.

carnitines form conjugates with the FA that will be oxidized, generating NADH, $FADH_2$ and acetyl CoA in each round (Longo et al., 2016). Both NADH and $FADH_2$ enter the electron transport chain to produce ATP (Carracedo et al., 2013). On the other hand, the acetyl CoA generated in the oxidation processes can feed the production of TAC intermediates or be part of new lipid synthesis (Ginger et al., 2000; Carracedo et al., 2013). However, it cannot be ruled out that the accumulation of carnitines is due to a malfunction of mitochondrial activity that prevents the catabolism of this type of metabolite.

FA can be acquired in three ways: exogenous FA that enter cells; FA that arise through *de novo* synthesis from acetyl-coA; and fatty acids that are released within the cell by hydrolysis of acylated proteins, phospholipids, and triglycerides (Koundouros and Poulogiannis, 2020). The lipid composition found in *T. cruzi* varies according to the stage analyzed; however, they are mainly represented by triacylglycerides, phosphatidylcholines, phosphatidylethanolamines and phosphatidylinositols (Cunha-E-Silva et al., 2002; Booth and Smith, 2020). These types of compounds incorporated the greatest global changes in the metabolism of the treated parasites, with a large part of these showing upward trends. Of these, the group of LPL had the largest number of representatives. LPL are metabolic intermediates normally generated through the active hydrolyzation of phospholipases. These enzymes cleave intracellular phospholipids from the cell membrane, generating a variety of products such as LPL, FA, diacylglycerols, phosphocholine, phosphoinositides and phosphatidic acid, among others (Belaunzarán et al., 2011). All these intermediate metabolites can contribute to the generation of ATP through oxidation to acetyl-coA and then be incorporated into TAC. However, the role of this mechanism in energy generation during nutritional stress events is still unknown. Other studies carried out in pathologies in which there is an increase in reactive oxygen species (ROS), such as diabetes and obesity, have associated the selective loss of a glycerophospholipid fatty acyl residue with the overproduction of ROS through lipid peroxidation. However, this mechanism is still under study (Reis and Spickett, 2012; Fuchs, 2014; Chen et al., 2018; Engel et al., 2021).

The use of glycerophospholipids as a substrate to produce intermediates such as LPL can affect the composition of biological membranes and, consequently, their selective permeability, leading to cell lysis. Importantly, the accumulation of this type of lipid in the treated parasites entails toxic effects, since high concentrations alter the structure of the membrane and cause cell lysis. Finally, some research suggests that FA oxidation may be a permanent source of reactive oxygen species, which can cause endoplasmic reticulum stress and changes in mitochondrial membrane potentials, causing apoptotic death (Bowes et al., 1993; Ly et al., 2017; Tan et al., 2020).

Although the observed alterations have been related to a possible mechanism of energy alteration, the increases or deficiencies of some metabolites suggest involvement in other biological processes. Changes in the content of amino acids such as proline, glutamate, serine, glycine, and leucine can affect protein synthesis, resistance to nutritional, osmotic, thermal, and oxidative stress, as well as invasion, replication, and metacyclogenesis processes, among others (Contreras et al., 1985; Pereira et al., 2002; Silber et al., 2005; Magdaleno et al., 2011; Paes et al., 2013; Marchese et al., 2018). On the other hand, alterations in nucleosides would directly affect the construction of macromolecules such as glycoconjugates, DNA and RNA, which would affect the invasion and processes dependent on nucleic acids, such as replication, transcription and protein synthesis (Landfear et al., 2004; MacRae et al., 2006). Finally, both energy imbalances and low trypanothione levels, observed in parasites treated with CFHEX, may suggest an increase in reactive oxygen species (ROS) as well as imbalances in redox potentials (Fairlamb et al., 1985; Bond

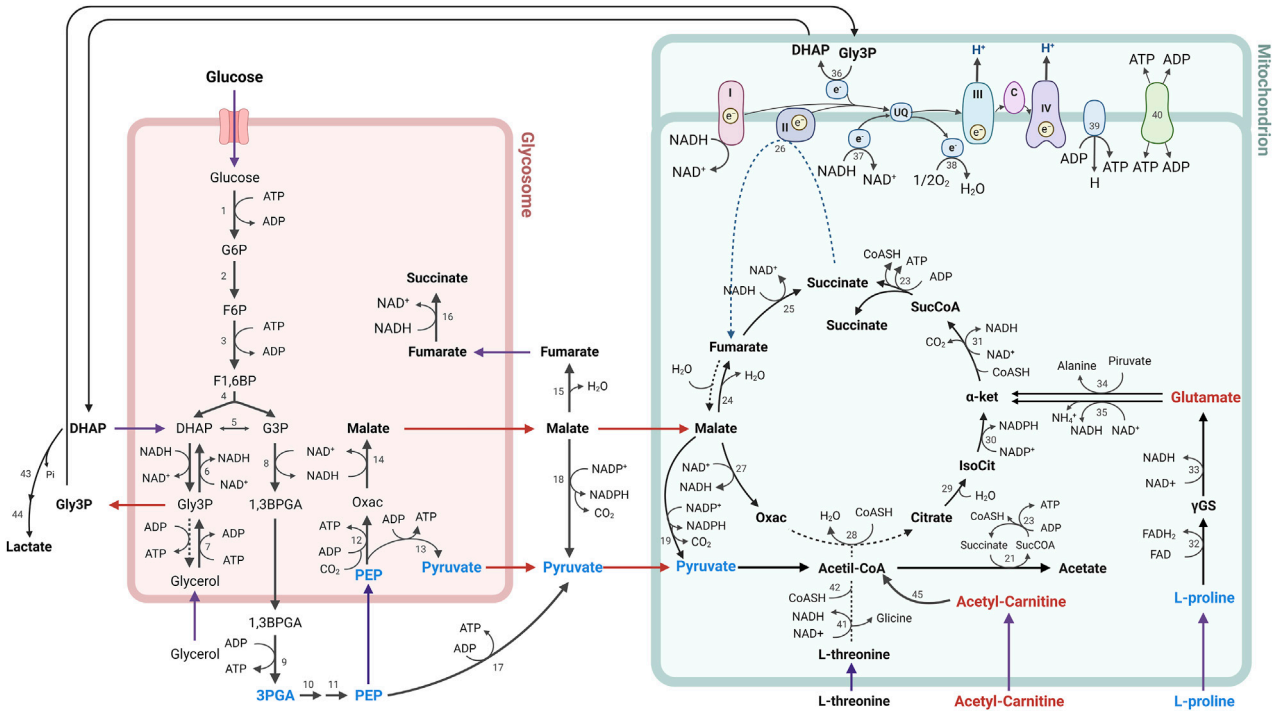


FIGURE 8
Metabolic processes for energy generation in *Trypanosoma cruzi*. The metabolites found in lower concentration are represented in blue, while the metabolites found in higher concentration in parasites treated with CFHEX compared to untreated parasites are represented in red. Numbers indicate participating enzymes: 1, hexokinase; 2, glucose-6-phosphate isomerase; 3, phosphofructokinase; 4, aldolase; 5, triose-phosphate isomerase; 6, glycerol-3-phosphate dehydrogenase; 7, glycerol kinase; 8, glyceraldehyde-3-phosphate dehydrogenase; 9, phosphoglycerate kinase B; 10, phosphoglycerate mutase; 11, enolase; 12, phosphoenolpyruvate carboxykinase; 13, pyruvate phosphate dikinase; 14, glycosomal malate dehydrogenase; 15, cytosolic fumarase; 16, glycosomal NADH-dependent fumarate reductase, 17, pyruvate kinase; 18, cytosolic malic enzyme; 19, mitochondrial malic enzyme; 20, pyruvate dehydrogenase complex; 21, acetate:succinate CoA-transferase; 22, acetyl-CoA thioesterase; 23, succinyl-CoA synthetase; 24, mitochondrial fumarase; 25, mitochondrial NADH-dependent fumarate reductase; 26, succinate dehydrogenase (respiratory chain complex II); 27, mitochondrial malate dehydrogenase; 28, citrate synthase; 29, aconitase; 30, isocitrate dehydrogenase; 31, α-ketoglutarate dehydrogenase; 32, L-proline dehydrogenase; 33, pyrroline-5-carboxylate dehydrogenase; 34, alanine aminotransferase; 35, glutamate dehydrogenase; 36, FAD-dependent mitochondrial glycerol-3-phosphate dehydrogenase; 37, rotenone-insensitive NADH dehydrogenase; 38, alternative oxidase; 39, FoF1-ATP synthase; 40, ADP/ATP carrier; 41, L-threonine dehydrogenase; 42, AKCT, 2-amino-3-ketobutyrate CoA-transferase; 43, methylglyoxal reductase; 44, lactaldehyde dehydrogenase; 45, Fatty acid β-oxidation; I, III, IV, respiratory chain complexes.

et al., 1999; Schönfeld et al., 2010; Menna-Barreto and De Castro, 2014). ROS can oxidize macromolecules such as lipids, nucleic acids, and proteins. Besides, it can cause endoplasmic reticulum stress and mitochondrial damage, alterations that can promote, among many other effects, programmed cell death (Jarzab and Stryjecka-Zimmer, 2008).

Previously, the analysis of the mechanisms of death induced by the CFHEX extract in epimastigotes and trypomastigotes of *T. cruzi* evidenced a trypanosomicidal mechanism associated with programmed cell death. The present investigation found a series of highly expressed proapoptotic compounds such as sphingosine, sphingosine 1 phosphate and ceramides (Cuvillier, 2002; Davaille et al., 2002; Koeller and Heise, 2011), which support the effects of programmed cell death previously evidenced (Pardo-Rodriguez et al., 2022). The trypanosomicidal effects can be related to the presence of exogenous terpenes found in the parasites and previously described in *C. fimbriata* (Pardo-Rodriguez et al., 2022). Ursolic acid induced a significant reduction in amastigotes in RAW macrophage cultures infected with trypomastigotes compared with the untreated cultures (Vanrell et al., 2020). Other studies conducted in models of acute infection in BALB/c albino male mice found that oral treatment with ursolic acid at a concentration of 20 mg/kg/day reduced parasitemia, measured at the parasitemic peak, after infection with trypomastigotes of strain Y by 60% (Da Silva Ferreira et al., 2013). On the other hand, betulinic acid inhibits cellular populations of all three stages of *T. cruzi* without causing toxicity in LLC-MK2 cells at the concentrations used. (200-1,600 μM). Furthermore, the treated parasites displayed alterations in cell membrane integrity, mitochondrial membrane potential, and reservosome inflammation, along with an increased production of reactive oxygen species (Sousa et al., 2017).

In summary, the metabolic alterations observed in parasites treated with CFHEX may reflect energetic alterations associated with glucose, hexoses and some amino acids which suggests that the energy demands of the parasite could be supplied from the β-oxidation of FA and the production of TAC intermediates (Figure 8). Although, with the results obtained here, it is not possible to specify the mechanism by which the alteration of energy metabolism is taking place, investigations carried out with betulinic, ursolic and oleanolic acids have shown a deprivation of glycolytic metabolism in cancer cells. In these studies, the glycolytic decline is associated with signaling

pathways such as mTOR, and AKT, which modulate the expression of glycolytic pathway enzymes such as hexokinase, phosphofructokinase, and pyruvate kinase. Additionally, it has been observed that apoptotic processes induced by oxidative stress are favored in cells treated with ursolic, betulinic and pomolic acids (Liu et al., 2014; Lewinska et al., 2017; Zheng et al., 2019; Wang et al., 2021). Other studies have found inhibitory effects in trypanosomatid species after exposure to betulinic and ursolic acids associated with alterations in the mitochondrial membrane potential and increases in oxidative stress, which in turn implies effects on oxidative phosphorylation (Yamamoto et al., 2015; Bossolani et al., 2017; Sousa et al., 2017; Albuquerque et al., 2020).

With this panorama, two possible hypotheses arise as an explanation for glucose deprivation or depletion. The triterpenes that constitute CFHEX (betulinic, ursolic and pomolic acid), which were found inside the treated parasites, can inhibit the uptake of glucose or its metabolization by modulating signaling pathways (Liu et al., 2014; Lewinska et al., 2017; Zheng et al., 2019; Wang et al., 2021). Likewise, and although it is an approach little studied, the possible inhibition exerted by triterpenes on glucose transporters and enzymes of the glycolytic pathway of *T. cruzi* cannot be ruled out. This last point becomes relevant when comparing the identities of the transporters and enzymes present in trypanosomatids with respect to those found in humans. In summary, the inhibition of glycolysis would induce the consumption of other carbon sources such as amino acids and lipids, which would serve as generators of TAC intermediates, that would directly feed mitochondrial oxidative phosphorylation, making the treated parasites dependent on this last mechanism to meet energy demands.

However, a second mechanism has been proposed for certain triterpenes, and it has been found that those compounds could modify the potentials of mitochondrial membranes (Ψm) (Hordyjewska et al., 2019), which, in turn, may impact electron transport and oxidative phosphorylation. These findings may suggest that the triterpenes found as major compounds in CFHEX have the ability to uncouple mitochondrial function and promote metabolic changes that induce a death similar to apoptosis in the treated parasites.

Collectively, our results showed that the hexanic extract of *C. fimbriata* induces alterations in energy metabolism in *T. cruzi* epimastigotes that are compatible with CFHEX induced apoptosis-like death. The finding of pentacyclic triterpenes in the metabolome of *C. fimbriata* treated parasites renders this extract as a novel source of triterpene compounds, which in the future may contribute to new alternatives for the control of *T. cruzi* infection. To our knowledge, this is the first investigation where the general metabolic context of *T. cruzi* correlated to the exposure of extracts rich in pentacyclic triterpenes. The metabolomic analyses performed on the epimastigote stage reflect the changes generated generated from the CFHEX treatment, however, they would only indicate the steady states of the metabolites at the selected treatment time (36 h). Therefore, the discussion was carried out by correlating the states of the altered metabolites with their biological functions. It is necessary for future studies to establish the metabolic pathway fluxes involved in the inhibition mechanisms induced in the *T. cruzi* stages by the CFHEX extract.

5 Conclusion

The effect of the hexanic extract of *C. fimbriata* on the metabolism of *T. cruzi* epimastigotes is related to multiple biosynthetic pathways, among which those associated with energy metabolism stand out, which in turn is connected to the induction of apoptosis-like observed in treated parasites. The constituent triterpenes of CFHEX such as betulinic and ursolic acids were found inside the treated parasites and may be related to the trypanosomicidal effects together with the metabolomic modifications.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

DP, MC, and MS designed the experiments; DP, MS, and PL executed the *in vitro* experiments, and acquired and interpreted the data; DP, CC, JR, and JM analyzed and interpreted the *in vitro* results; DP, MC, PL, JR, CP, and JM participated in the discussion of the results; DP, MC, PL, CC, CP, and JR drafted the manuscript; CC, JM and JR are the leaders of the project and designed the experiments, interpreted the results, and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Vicerrectoría de Investigación from the Pontificia Universidad Javeriana and belongs to the project "Evaluación del efecto tripanocida inducido por extractos de especies vegetales colombianas sobre diferentes unidades discretas de tipificación (UDTs) de *Trypanosoma cruzi*" (ID Proposal 00008279). DP was supported by Ph.D. scholarships from MinCiencias "Ministerio de Ciencia, Tecnología e Innovación, Colombia"; Convocatoria 755–2016, Formación de Capital Humano de Alto Nivel para el Departamento de Tolima. PL was funded by Pontificia Universidad Javeriana, Ministerio de Ciencia, Tecnología e Innovación, Ministerio de Educación Nacional, Ministerio de Industria, Comercio y Turismo and ICETEX, 2ª Convocatoria Ecosistema Científico–Colombia Científica 792-2017, Program "Generación de alternativas terapéuticas en cáncer a partir de plantas a través de procesos de investigación y desarrollo traslacional, articulados en sistemas de valor sostenibles ambiental y económicamente" (Contract no. FP44842-221-2018). We thank the Pontificia Universidad Javeriana for its collaboration with the publication (Proposal ID 10567).

Acknowledgments

The authors thank Pontificia Universidad Javeriana for its support for financing this work and its publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1206074/full#supplementary-material

References

Albuquerque, R. D. D. G., Oliveira, A. P., Ferreira, C., Passos, C. L. A., Fialho, E., Soares, D. C., et al. (2020). Anti-*Leishmania amazonensis* activity of the terpenoid fraction from *Eugenia pruniformis* leaves. *Acad Bras Cienc* 92, 20201181–e20201214. doi:10.1590/0001-3765202020201181

Barisón, M. J., Rapado, L. N., Merino, E. F., Pral, E. M. F., Mantilla, B. S., Marchese, L., et al. (2017). Metabolomic profiling reveals a finely tuned, starvationinduced metabolic switch in *Trypanosoma cruzi* epimastigotes. *J. Biol. Chem.* 292, 8964–8977. doi:10.1074/jbc.M117.778522

Belaunzarán, M. L., Lammel, E. M., and De Isola, E. L. D. (2011). Phospholipases A in trypanosomatids. *Enzyme Res.* 2011, 392082. doi:10.4061/2011/392082

Bern, C. (2015). Chagas' disease. *N. Engl. J. Med.* 373, 456–466. doi:10.1056/NEJMra1410150

Bond, C. S., Zhang, Y., Berriman, M., Cunningham, M. L., Fairlamb, A. H., and Hunter, W. N. (1999). Crystal structure of *Trypanosoma cruzi* trypanothione reductase in complex with trypanothione, and the structure-based discovery of new natural product inhibitors. *Structure* 7, 81–89. doi:10.1016/S0969-2126(99)80011-2

Booth, L. A., and Smith, T. K. (2020). Lipid metabolism in *Trypanosoma cruzi*: A review. *Mol. Biochem. Parasitol.* 240, 111324. doi:10.1016/j.molbiopara.2020.111324

Bossolani, G. D. P., Ueda-Nakamura, T., Silva, S. O., Dias Filho, B. P., Costa, T. O. G., Quintanilla, R. H. R., et al. (2017). Anti-trypanosoma activity and synergistic effects of natural and semi-synthetic triterpenes and predominant cell death through autophagy in amastigote forms. *J. Braz Chem. Soc.* 28, 2473–2489. doi:10.21577/0103-5053.20170103

Bowes, A. E., Samad, A. H., Jiang, P., Weaver, B., and Mellors, A. (1993). The acquisition of lysophosphatidylcholine by African trypanosomes. *J. Biol. Chem.* 268, 13885–13892. doi:10.1016/s0021-9258(19)85185-2

Bringaud, F., Rivière, L., and Coustou, V. (2006). Energy metabolism of trypanosomatids: adaptation to available carbon sources. *Mol. Biochem. Parasitol.* 149, 1–9. doi:10.1016/j.molbiopara.2006.03.017

Bustamante, J. M., and Tarleton, R. L. (2014). Potential new clinical therapies for Chagas disease. *Expert Rev. Clin. Pharmacol.* 7, 317–325. doi:10.1586/17512433.2014.909282

Carracedo, A., Cantley, L. C., and Pandolfi, P. P. (2013). Cancer metabolism: fatty acid oxidation in the limelight. *Nat. Rev. Cancer* 13, 227–232. doi:10.1038/nrc3483

Castañeda, S. J., Suta-Velásquez, M., Mateus, J., Pardo-Rodriguez, D., Puerta, C. J., Cuéllar, A., et al. (2021). Preliminary chemical characterization of ethanolic extracts from Colombian plants with promising anti - *Trypanosoma cruzi* activity. *Exp. Parasitol.* 223, 108079. doi:10.1016/j.exppara.2021.108079

Castro, J. A., deMecca, M. M., and Bartel, L. C. (2006). Toxic side effects of drugs used to treat Chagas' disease (American Trypanosomiasis). *Hum. Exp. Toxicol.* 25, 471–479. doi:10.1191/0960327106het653oa

Cazzulo, J. J. (1992). Energy metabolism in *Trypanosoma cruzi*. *Subcell. Biochem.* 18, 235–257. doi:10.1007/978-1-4899-1651-8_7

Cazzulo, J. J., Franke de Cazzulo, B. M., Engel, J. C., and Cannata, J. J. B. (1985). End products and enzyme levels of aerobic glucose fermentation in trypanosomatids. *Mol. Biochem. Parasitol.* 16, 329–343. doi:10.1016/0166-6851(85)90074-X

Cazzulo, J. J. (1994). Intermediate metabolism in *Trypanosoma cruzi*. *J. Bioenerg. Biomembr.* 26, 157–165. doi:10.1007/BF00763064

César Carranza, J., Kowaltowski, A. J., Mendonça, M. A. G., De Oliveira, T. C., Gadelha, F. R., and Zingales, B. (2009). Mitochondrial bioenergetics and redox state are unaltered in *Trypanosoma cruzi* isolates with compromised mitochondrial complex i subunit genes. *J. Bioenerg. Biomembr.* 41, 299–308. doi:10.1007/s10863-009-9228-4

Chen, Z., Zang, L., Wu, Y., Nakayama, H., Shimada, Y., Shrestha, R., et al. (2018). Lipidomic profiling on oxidized phospholipids in type 2 diabetes mellitus model zebrafish. *Anal. Sci.* 34, 1201–1208. doi:10.2116/analsci.18P281

Contreras, V. T., Salles, J. M., Thomas, N., Morel, C. M., and Goldenberg, S. (1985). *In vitro* differentiation of *Trypanosoma cruzi* under chemically defined conditions. *Mol. Biochem. Parasitol.* 16, 315–327. doi:10.1016/0166-6851(85)90073-8

Cunha-E-Silva, N. L., Atella, G. C., Porto-Carreiro, I. A., Morgado-Diaz, J. A., Pereira, M. G., and De Souza, W. (2002). Isolation and characterization of a reservosome fraction from *Trypanosoma cruzi*. *FEMS Microbiol. Lett.* 214, 7–12. doi:10.1111/j.1574-6968.2002.tb11317.x

Cuvillier, O. (2002). Sphingosine in apoptosis signaling. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* 1585, 153–162. doi:10.1016/S1388-1981(02)00336-0

Da Silva Ferreira, D., Esperandim, V. R., Toldo, M. P. A., Kuehn, C. C., Do Prado Júnior, J. C., Cunha, W. R., et al. (2013). *In vivo* activity of ursolic and oleanolic acids during the acute phase of *Trypanosoma cruzi* infection. *Exp. Parasitol.* 134, 455–459. doi:10.1016/j.exppara.2013.04.005

Davaille, J., Li, L., Mallat, A., and Lotersztajn, S. (2002). Sphingosine 1-phosphate triggers both apoptotic and survival signals for human hepatic myofibroblasts. *J. Biol. Chem.* 277, 37323–37330. doi:10.1074/jbc.M202798200

Engel, K. M., Schiller, J., Galuska, C. E., and Fuchs, B. (2021). Phospholipases and reactive oxygen species derived lipid biomarkers in healthy and diseased humans and animals – a Focus on lysophosphatidylcholine. *Front. Physiol.* 12, 732319–732328. doi:10.3389/fphys.2021.732319

Eriksson, L., Trygg, J., and Wold, S. (2008). CV-ANOVA for significance testing of PLS and OPLS® models. *J. Chemom.* 22, 594–600. doi:10.1002/cem.1187

Fairlamb, A., Blackburn, P., Ulrich, P., Chait, B., and Cerami, A. (1985). Trypanothione: A novel bis(glutathionyl)spermidine cofactor for glutathione reductase in trypanosomatids. *Science* 227, 1485–1487. doi:10.1126/science.3883489

Fuchs, B. (2014). Mass spectrometry and inflammation - MS methods to study oxidation and enzyme-induced changes of phospholipids. *Anal. Bioanal. Chem.* 406, 1291–1306. doi:10.1007/s00216-013-7534-5

Ginger, M. L., Prescott, M. C., Reynolds, D. G., Chance, M. L., and Goad, J. L. (2000). Utilization of leucine and acetate as carbon sources for sterol and fatty acid biosynthesis by Old and New World Leishmania species, *Endotrypanum monterogeii* and *Trypanosoma cruzi*. *Eur. J. Biochem.* 267, 2555–2566. doi:10.1046/j.1432-1327.2000.01261.x

Gori, B., Ulian, T., Bernal, H. Y., and Diazgranados, M. (2022). Understanding the diversity and biogeography of Colombian edible plants. *Sci. Rep.* 12 (1), 7835–7915. doi:10.1038/s41598-022-11600-2

Hamptont, J. R. (1971). *Serine metabolism in the culture form of* Trypanosoma cruzi: *distribution of 14c from serine into metabolic fractions*. United Kingdom: Pergnmon Press.

Hordyjewska, A., Ostapiuk, A., Horecka, A., and Kurzepa, J. (2019). Betulin and betulinic acid: triterpenoids derivatives with a powerful biological potential. *Phytochem. Rev.* 18 (3), 929–951. doi:10.1007/s11101-019-09623-1

Jarzab, A., and Stryjecka-Zimmer, M. (2008). Oxidative stress and apoptosis. *Ann. Univ. Mariae Curie-Sklodowska. Sect. B* 63, 67–71. doi:10.2478/v10079-008-0010-6

Klein, N., Hurwitz, I., and Durvasula, R. (2012). Globalization of chagas disease: A growing concern in nonendemic countries. *Epidemiol. Res. Int.* 2012, 1–13. doi:10.1155/2012/136793

Koeller, C. M. E., and Heise, N. (2011). The sphingolipid biosynthetic pathway is a potential target for chemotherapy against chagas disease. *Enzyme Res.* 2011, 648159. doi:10.4061/2011/648159

Koundouros, N., and Poulogiannis, G. (2020). Reprogramming of fatty acid metabolism in cancer. *Br. J. Cancer* 122, 4–22. doi:10.1038/s41416-019-0650-z

Landfear, S. M., Ullman, B., Carter, N. S., and Sanchez, M. A. (2004). Nucleoside and nucleobase transporters in parasitic protozoa. *Eukaryot. Cell* 3, 245–254. doi:10.1128/EC.3.2.245-254.2004

Lewinska, A., Adamczyk-Grochala, J., Kwasniewicz, E., Deregowska, A., and Wnuk, M. (2017). Ursolic acid-mediated changes in glycolytic pathway promote cytotoxic autophagy and apoptosis in phenotypically different breast cancer cells. *Apoptosis* 22, 800–815. doi:10.1007/s10495-017-1353-7

Liu, J., Wu, N., Ma, L., Liu, M., Liu, G., Zhang, Y., et al. (2014). Oleanolic acid suppresses aerobic glycolysis in cancer cells by switching pyruvate kinase type M isoforms. *PLoS One* 9, e91606–e91609. doi:10.1371/journal.pone.0091606

Longo, N., Frigeni, M., Pasquali, M., Biophys, B., and Author, A. (2016). Carnitine transport and fatty acid oxidation. *Biochim. Biophys. Acta* 1863, 2422–2435. doi:10.1016/j.bbamcr.2016.01.023

Lopera Valle, J. S., Rojas Jiménez, S., and Mejía Ochoa, M. (2013). Actividad tripanocida de plantas latinoamericanas, una futura alternativa terapéutica para la Enfermedad de Chagas. *Rev. Ibero-Latinoam. Parasitol.* 72, 22–30.

Ly, L. D., Xu, S., Choi, S.-K., Ha, C.-M., Thoudam, T., Cha, S.-K., et al. (2017). Oxidative stress and calcium dysregulation by palmitate in type 2 diabetes. *Exp. Mol. Med.* 49, e291. doi:10.1038/emm.2016.157

MacRae, J. I., Obado, S. O., Turnock, D. C., Roper, J. R., Kierans, M., Kelly, J. M., et al. (2006). The suppression of galactose metabolism in *Trypanosoma cruzi* epimastigotes causes changes in cell surface molecular architecture and cell morphology. *Mol. Biochem. Parasitol.* 147, 126–136. doi:10.1016/j.molbiopara.2006.02.011

Magdaleno, A., Suárez Mantilla, B., Rocha, S. C., Pral, E. M. F., and Silber, A. M. (2011). The involvement of glutamate metabolism in the resistance to thermal, nutritional, and oxidative stress in *Trypanosoma cruzi*. *Enzyme Res.* 2011, 486928. doi:10.4061/2011/486928

Manchola, N. C., Rapado, L. N., Barisón, M. J., and Silber, A. M. (2016). Biochemical Characterization of branched chain amino acids uptake in *Trypanosoma cruzi*. *J. Eukaryot. Microbiol.* 63, 299–308. doi:10.1111/jeu.12278

Manne-Goehler, J., Umeh, C. A., Montgomery, S. P., and Wirtz, V. J. (2016). Estimating the burden of chagas disease in the United States. *PLoS Negl. Trop. Dis.* 10, e0005033–e0005037. doi:10.1371/journal.pntd.0005033

Marchese, L., Nascimento, J. D. F., Damasceno, F. S., Bringaud, F., Michels, P. A. M., and Silber, A. M. (2018). The uptake and metabolism of amino acids, and their unique role in the biology of pathogenic trypanosomatids. *Pathogens* 7, 36–45. doi:10.3390/pathogens7020036

Martins, R. M., Covarrubias, C., Rojas, R. G., Silber, A. M., and Yoshida, N. (2009). Use of L-proline and ATP production by *Trypanosoma cruzi* metacyclic forms as requirements for host cell invasion. *Infect. Immun.* 77, 3023–3032. doi:10.1128/IAI.00138-09

Mejía-Jaramillo, A. M., Fernández, G. J., Montilla, M., Nicholls, R. S., and Triana-Chávez, O. (2012). Sensibilidad al benzonidazol de cepas de *Trypanosoma cruzi* sugiere la circulación de cepas naturalmente resistentes en Colombia. *Biomedica* 32, 196–205. doi:10.1590/S0120-41572012000300007

Menna-Barreto, R. F. S., and De Castro, S. L. (2014). The double-edged sword in pathogenic trypanosomatids: the pivotal role of mitochondria in oxidative stress and bioenergetics. *Biomed. Res. Int.* 2014, 614014. doi:10.1155/2014/614014

Michels, P. A. M., Villafraz, O., Pineda, E., Alencar, M. B., Cáceres, A. J., Silber, A. M., et al. (2021). Carbohydrate metabolism in trypanosomatids: new insights revealing novel complexity, diversity and species-unique features. *Exp. Parasitol.* 224, 108102. doi:10.1016/j.exppara.2021.108102

Newman, D. J., and Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* 79, 629–661. doi:10.1021/acs.jnatprod.5b01055

Newman, D. J., and Cragg, G. M. (2020). Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* 83, 770–803. doi:10.1021/acs.jnatprod.9b01285

Nunes, M. C. P., Dones, W., Morillo, C. A., Encina, J. J., Ribeiro, A. L., and Council on Chagas Disease of the Interamerican Society of Cardiology (2013). Chagas disease: an overview of clinical and epidemiological aspects. *J. Am. Coll. Cardiol.* 62, 767–776. doi:10.1016/j.jacc.2013.05.046

Olivera, M. J., Fory, J. A., Porras, J. F., and Buitrago, G. (2019). Prevalence of chagas disease in Colombia: A systematic review and meta-analysis. *PLoS One* 14, 0210156–e210218. doi:10.1371/journal.pone.0210156

Paes, L. S., Suárez Mantilla, B., Zimbres, F. M., Pral, E. M. F., Diogo de Melo, P., Tahara, E. B., et al. (2013). Proline dehydrogenase regulates redox state and respiratory metabolism in *Trypanosoma cruzi*. *PLoS One* 8, e69419. doi:10.1371/journal.pone.0069419

Pardo-Rodriguez, D., Lasso, P., Mateus, J., Mendez, J., Puerta, C. J., Cuéllar, A., et al. (2022). A terpenoid-rich extract from *Clethra fimbriata* exhibits anti-*Trypanosoma cruzi* activity and induces T cell cytokine production. *Heliyon* 8, e09182. doi:10.1016/j.heliyon.2022.e09182

Pavia, P. X., Thomas, M. C., López, M. C., and Puerta, C. J. (2012). Molecular characterization of the short interspersed repetitive element SIRE in the six discrete typing units (DTUs) of *Trypanosoma cruzi*. *Exp. Parasitol.* 132, 144–150. doi:10.1016/j.exppara.2012.06.007

Pereira, C. A., Alonso, G. D., Ivaldi, S., Silber, A., Alves, M. J. M., Bouvier, L. A., et al. (2002). Arginine metabolism in *Trypanosoma cruzi* is coupled to parasite stage and replication. *FEBS Lett.* 526, 111–114. doi:10.1016/S0014-5793(02)03157-5

Pérez-Molina, J. A., and Molina, I. (2017). Chagas disease. *Lancet* 6736, 82–94. doi:10.1016/S0140-6736(17)31612-4

Quiñones, W., Acosta, H., Gonçalves, C. S., Motta, M. C. M., Gualdrón-López, M., and Michels, P. A. M. (2020). Structure, properties, and function of glycosomes in *Trypanosoma cruzi*. *Front. Cell Infect. Microbiol.* 10, 25–11. doi:10.3389/fcimb.2020.00025

Rassi, A., Rassi, A., and Marin-Neto, J. A. (2010). Chagas disease. *Lancet* 375, 1388–1402. doi:10.1016/S0140-6736(10)60061-X

Reis, A., and Spickett, C. M. (2012). Chemistry of phospholipid oxidation. *Biochim. Biophys. Acta Biomembr.* 1818, 2374–2387. doi:10.1016/j.bbamem.2012.02.002

Rojo, D., Canuto, G. A. B., Castilho-martins, E. A., Marina, F., Barbas, C., López-Gonzálvez, Á., et al. (2015). A multiplatform metabolomic approach to the basis of antimonial action and resistance in *Leishmania infantum*. *PLoS One* 10, 0130675–e130720. doi:10.1371/journal.pone.0130675

Roper, J. R., and Ferguson, M. A. J. (2003). Cloning and characterisation of the UDP-glucose 4′-epimerase of *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* 132, 47–53. doi:10.1016/j.molbiopara.2003.07.002

Sánchez-Valdéz, F. J., Padilla, A., Wang, W., Orr, D., and Tarleton, R. L. (2018). Spontaneous dormancy protects *Trypanosoma cruzi* during extended drug exposure. *Elife* 7, e34039. doi:10.7554/eLife.34039

Schmidt, T. J., Khalid, S. A., Romanha, A. J., Alves, T. M. A., Biavatti, M. W., Brun, R., et al. (2012). The potential of secondary metabolites from plants as drugs or leads against protozoan neglected diseases - Part I. *Curr. Med. Chem.* 19, 2128–2175. doi:10.2174/092986712800229023

Schönfeld, P., Wieckowski, M. R., Lebiedzińska, M., and Wojtczak, L. (2010). Mitochondrial fatty acid oxidation and oxidative stress: lack of reverse electron transfer-associated production of reactive oxygen species. *Biochim. Biophys. Acta Bioenerg.* 1797, 929–938. doi:10.1016/j.bbabio.2010.01.010

Silber, A. M., Colli, W., Ulrich, H., Alves, M. J. M., and Pereira, C. A. (2005). Amino acid metabolic routes in *Trypanosoma cruzi*: possible therapeutic targets against chagas' disease. *Curr. Drug Targets Infect. Disord.* 5, 53–64. doi:10.2174/1568005053174636

Silber, A. M., Tonelli, R. R., Martinelli, M., Colli, W., and Alves, M. J. M. (2002). Active transport of L-proline in *Trypanosoma cruzi*. *J. Eukaryot. Microbiol.* 49, 441–446. doi:10.1111/j.1550-7408.2002.tb00225.x

Sousa, P. L., Souza, R. O. S., Tessarolo, L. D., de Menezes, R. R. P. P. B., Sampaio, T. L., Canuto, J. A., et al. (2017). Betulinic acid induces cell death by necrosis in *Trypanosoma cruzi*. *Acta Trop.* 174, 72–75. doi:10.1016/j.actatropica.2017.07.003

Souza, R. O. O., Damasceno, F. S., Marsiccobetre, S., Biran, M., Murata, G., Curi, R., et al. (2021). Fatty acid oxidation participates in resistance to nutrient-depleted environments in the insect stages of *Trypanosoma cruzi*. *PLoS Pathog.* 17, e1009495. doi:10.1371/journal.ppat.1009495

Stoppani, A. O. M., Docampo, R., De Boiso, J. F., and Frasch, A. C. C. (1980). Effect of inhibitors of electron transport and oxidative phosphorylation on *Trypanosoma cruzi* respiration and growth. *Mol. Biochem. Parasitol.* 2, 3–21. doi:10.1016/0166-6851(80)90044-4

Sylvester, D., and Krassner, S. M. (1976). Proline metabolism in *Trypanosoma cruzi* epimastigotes. *Comp. Biochem. Physiology – Part B* 55, 443–447. doi:10.1016/0305-0491(76)90318-7

Tan, S. T., Ramesh, T., Toh, X. R., and Nguyen, L. N. (2020). Emerging roles of lysophospholipids in health and disease. *Prog. Lipid Res.* 80, 101068. doi:10.1016/j.plipres.2020.101068

Tarleton, R. L. (2016). Chagas disease: A solvable problem, ignored. *Trends Mol. Med.* 22, 835–838. doi:10.1016/j.molmed.2016.07.008

Triba, M. N., B, L. L. M., Amathieu, R., Goossens, C., Bouchemal, N., Nahon, P., et al. (2015). PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Mol. Biosyst.* 11, 13–19. doi:10.1039/c4mb00414k

Urbina, J. A. (2015). Recent clinical trials for the etiological treatment of chronic chagas disease: advances, challenges and perspectives. *J. Eukaryot. Microbiol.* 62, 149–156. doi:10.1111/jeu.12184

Urbina, J. A. (2010). Specific chemotherapy of chagas disease: relevance, current limitations and new approaches. *Acta Trop.* 115, 55–68. doi:10.1016/j.actatropica.2009.10.023

Van Hellemond, J. J., Bakker, B. M., and Tielens, A. G. M. (2005). Energy metabolism and its compartmentation in *Trypanosoma brucei*. *Adv. Microb. Physiology* 50, 199–226. doi:10.1016/S0065-2911(05)50005-5

Vanrell, M. C., Martinez, S. J., Muñoz, L. I., Salassa, B. N., and Romano, P. S. (2020). Ursolic acid promotes clearance of *Trypanosoma cruzi* amastigotes in the host cell. in 1st International Electronic Conference on Biomolecules: Natural and Bio-Inspired Therapeutics for Human Diseases. 01 December 2020, Basel, Switzerland: MDPI, 1–12.

Wang, S., Chang, X., Zhang, J., Li, J., Wang, N., Yang, B., et al. (2021). Ursolic acid inhibits breast cancer metastasis by suppressing glycolytic metabolism via activating SP1/Caveolin-1 signaling. *Front. Oncol.* 11, 745584–745616. doi:10.3389/fonc.2021.745584

WHO (2021). *Chagas disease (also known as American trypanosomiasis) Fact sheet*. Available at: https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis) (Accessed April 06, 2023).

World Health Organization (2015). Weekly epidemiological record (WER). *WHO* 90, 33–40. doi:10.1016/j.actatropica.2012.04.013

Xia, J., Wishart, D. S., and Valencia, A. (2011). MetPA: A web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* 27, 2342–2344. doi:10.1093/bioinformatics/btq418

Yamamoto, E. S., Campos, B. L. S., Jesus, J. A., Laurenti, M. D., Ribeiro, S. P., Kallás, E. G., et al. (2015). The effect of ursolic acid on *Leishmania* (Leishmania) *amazonensis* is related to programed cell death and presents therapeutic potential in experimental cutaneous leishmaniasis. *PLoS One* 10, 0144946–e145019. doi:10.1371/journal.pone.0144946

Zheng, Y., Liu, P., Wang, N., Wang, S., Yang, B., Li, M., et al. (2019). Betulinic acid suppresses breast cancer metastasis by targeting GRP78-Mediated glycolysis and er stress apoptotic pathway. *Oxid. Med. Cell Longev.* 2019, 8781690. doi:10.1155/2019/8781690

Glossary

CAR	Carnitine
CER	Ceramide
CFHEX	C. fimbriata hexanic extract
GC-MS	Gas chromatography coupled to mass spectrometry
EtOH	Ethanol
FA	fatty acid
FADH2	Flavin adenine dinucleotide
FBS	Fetal Bovine Serum
FC	Fold change
PEP	Phosphoenolpyruvate
G3P	Glycerate 3 phosphate
HILIC	Hydrophilic Interaction Liquid Chromatography
IC90	90% Inhibitory Concentration
SL	Steroidal lipid
LC-MS	Liquid chromatography coupled to mass spectrometry
LIT	Liver Infusion Tryptose
LPC	Lysophosphatidylcholines
LPE	Lysophosphatidylethanolamines
LPG	Lysophosphatidylglycerol
LPI	Lysophosphatidylinositols
LPL	Lysophospholipids
LPS	Lysophosphatidylserine
MeOH	Methanol
MG	Monoglyceride
NADH	Nicotine Adenine Dinucleotide
OPLS-DA	Orthogonal Partial Least Squares with Discriminant Analysis
PCA	Principal Component Analysis
PIR	Pyruvate
PLS	Partial Least Squares
QC	Quality Control
TAC	Tricarboxylic Acid Cycle
ROS	Reactive oxygen species
TR	Retention Time
UDP	Uridine Diphosphate
VIP	Importance of the variable in the projection

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 02 October 2023
DOI 10.3389/fmolb.2023.1223863

OPEN ACCESS

EDITED BY
Pablo Hoijemberg,
Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

REVIEWED BY
Luca Narduzzi,
University of Granada, Spain
Paniz Jasbi,
Arizona State University, United States

*CORRESPONDENCE
Maria Clara Arrieta-Echeverri,
marriet1@eafit.edu.co
Laura Sierra-Zapata,
lsierra3@eafit.edu.co

RECEIVED 16 May 2023
ACCEPTED 11 September 2023
PUBLISHED 02 October 2023

CITATION
Arrieta-Echeverri MC, Fernandez GJ, Duarte-Riveros A, Correa-Álvarez J, Bardales JA, Villanueva-Mejía DF and Sierra-Zapata L (2023), Multi-omics characterization of the microbial populations and chemical space composition of a water kefir fermentation.
Front. Mol. Biosci. 10:1223863.
doi: 10.3389/fmolb.2023.1223863

Multi-omics characterization of the microbial populations and chemical space composition of a water kefir fermentation

Maria Clara Arrieta-Echeverri[1]*, Geysson Javier Fernandez[2], Adriana Duarte-Riveros[3], Javier Correa-Álvarez[1], Jorge Adalberto Bardales[3], Diego Fernando Villanueva-Mejía[1] and Laura Sierra-Zapata[1]*

[1]Research Group CIBIOP, School of Applied Sciences and Engineering, Universidad EAFIT, Medellín, Antioquia, Colombia, [2]Infectious Diseases Biology and Control Group (BCEI), Universidad de Antioquia UdeA, Medellín, Colombia, [3]Iluma Innovation Labs, Iluma Alliance, Durham, NC, United States

In recent years, the popularity of fermented foods has strongly increased based on their proven health benefits and the adoption of new trends among consumers. One of these health-promoting products is water kefir, which is a fermented sugary beverage based on kefir grains (symbiotic colonies of yeast, lactic acid and acetic acid bacteria). According to previous knowledge and the uniqueness of each water kefir fermentation, the following project aimed to explore the microbial and chemical composition of a water kefir fermentation and its microbial consortium, through the integration of culture-dependent methods, compositional metagenomics, and untargeted metabolomics. These methods were applied in two types of samples: fermentation grains (inoculum) and fermentation samples collected at different time points. A strains culture collection of ~90 strains was established by means of culture-dependent methods, mainly consisting of individuals of *Pichia membranifaciens*, *Acetobacter orientalis*, *Lentilactobacillus hilgardii*, *Lacticaseibacillus paracasei*, *Acetobacter pomorum*, *Lentilactobacillus buchneri*, *Pichia kudriavzevii*, *Acetobacter pasteurianus*, *Schleiferilactobacillus harbinensis*, and *Kazachstania exigua*, which can be further studied for their use in synthetic consortia formulation. In addition, metabarcoding of each fermentation time was done by 16S and ITS sequencing for bacteria and yeast, respectively. The results show strong population shifts of the microbial community during the fermentation time course, with an enrichment of microbial groups after 72 h of fermentation. Metataxonomics results revealed *Lactobacillus* and *Acetobacter* as the dominant genera for lactic acid and acetic acid bacteria, whereas, for yeast, *P. membranifaciens* was the dominant species. In addition, correlation and systematic analyses of microbial growth patterns and metabolite richness allowed the recognition of metabolic enrichment points between 72 and 96 h and correlation between microbial groups and metabolite abundance (e.g., Bile acid conjugates and *Acetobacter tropicalis*). Metabolomic analysis also evidenced the production of bioactive compounds in this fermented matrix, which have been associated with biological activities, including antimicrobial and antioxidant. Interestingly, the chemical family of Isoschaftosides (C-glycosyl flavonoids) was also found, representing an important finding since this compound, with hepatoprotective and anti-inflammatory activity, had not been previously

reported in this matrix. We conclude that the integration of microbial biodiversity, cultured species, and chemical data enables the identification of relevant microbial population patterns and the detection of specific points of enrichment during the fermentation process of a food matrix, which enables the future design of synthetic microbial consortia, which can be used as targeted probiotics for digestive and metabolic health.

KEYWORDS

microbial communities, fermentation dynamics, probiotics, digestive health, multiomics approach, fermented food analysis

1 Introduction

Microbial consortia are present in a wide range of environments including soils, biofilms, and food products, such as beer, kombucha, and dairy products (Padmaperuma et al., 2019). These associations play an important role in soil management and nutrient mobilization, and they have also been studied due to their potential in the industry and economic importance, as they are involved in the development of fermented foods, which confer nutritional properties to their consumers (Madigan et al., 2015; McCaughey et al., 2022). Fermented products have long been considered basic foods in many countries because the fermentation process is an old technique to produce, conserve, or transform the organoleptic properties of foods and beverages (Fiorda et al., 2017; Villarreal-Morales et al., 2018; Bengoa et al., 2019). Most of these matrices are spontaneous or can be fermented using a starter culture (culture-dependent ferments), so there are several variables in the process of their fermentation, including the source of microorganisms, their nutritional ingredients, and environmental conditions, resulting in thousands of different variations of these products (Villarreal-Morales et al., 2018; Dimidi et al., 2019; Sharma and Yaiphathoi, 2020; Safak et al., 2023).

In recent years, the popularity of these foods, including kombucha and kefir, has increased based on the potential health benefits that have been ascribed to them based on the metabiotics, which can be grouped into prebiotics, probiotics, and postbiotics produced by these products (Marco et al., 2020; Pihurov et al., 2023). These biological activities can be related to the production of biochemical reactions triggered by multiple microorganisms that result in the release of vitamins, amino acids, exopolysaccharides (EPS), and organic acids, among other bioactive compounds (Diez-Ozaeta and Astiazaran, 2022). Water kefir has been studied, mainly for its impact on the immune system and gastrointestinal health, being beneficial for preventing non-communicable diseases such as lactose malabsorption, diabetes, obesity, inflammation, and cardiovascular conditions through the modulation of gut microbiota (Fiorda et al., 2017; Dimidi et al., 2019; Calatayud et al., 2021; Araújo et al., 2023). The composition of water kefir is known to be a stable microbial community of lactic acid bacteria (LAB), acetic acid bacteria (AAB), and yeasts, as shown by both culture-dependent and culture-independent-based studies (Laureys and De Vuyst, 2014; Zanirati et al., 2015; Farag et al., 2020).

For many years, research projects on water kefir have relied on conventional culture-dependent methods consisting of the isolation and culturing of microbes prior to their identification according to either morphological, biochemical, or genetic characteristics (Jianzhong et al., 2009). These culture-dependent approaches aim to test different culture conditions (growth temperature, pH, carbon, and nitrogen source) that could help optimize culturing methods for target microbes in fermented food matrices such as water kefir (Wuyts et al., 2020). Furthermore, molecular culture-independent methodologies, specifically metabarcoding and shotgun sequencing, have proven to be a powerful tool to provide a more complete microbial diversity spectrum in food samples, especially for those microbial groups that are difficult to isolate by culture-dependent methods (Verce et al., 2019). Recent studies that have characterized the microbial populations of water kefir have found that the most common microorganisms are lactic acid bacteria of the *Lactobacillus*, *Leuconostoc,* and *Lactococcus* genus, acetic acid bacteria of the genus *Acetobacter* and *Gluconobacter,* and yeast such as *Saccharomyces cerevisiae* and *Zygotorulaspora florentina* (Gulitz et al., 2011; Fiorda et al., 2017; Chen et al., 2021; Spizzirri et al., 2023). Nevertheless, given the variability of substrates and water kefir grain origins, a diverse composition of species and strains can be found in each beverage (Moretti et al., 2022). For instance, the same report by Moretti et al. (2022) exemplifies the uniqueness of a WK beverage prepared in Antioquia, Colombia, which is called "*Arroz de indio*" or "*Indiecitos*", based on the fermentation of "*Aguapanela*" or dry sugar cane solution, sharing some similarities with but not identical to the WK system of this study.

Furthermore, the health benefits of water kefir are not only related to the presence of certain microorganisms but also to the compounds produced during fermentation. Several studies have focused on determining the concentration of specific compounds including simple and short chain sugars, exopolysaccharides (EPS), organic acids, amino acids, and volatile compounds (Chen et al., 2021; Patel et al., 2022). Metabolite screening was carried out through different analytical platforms to perform targeted and untargeted studies such as gas chromatography/mass spectrometry (GC-MS), chromatography/mass spectrometry (GC/MS), liquid chromatography/mass spectrometry (LC/MS), and capillary electrophoresis/mass spectrometry (CE/MS). However, LC/MS has proven to deliver accurate qualitative and quantitative capability and provides the convenience of simultaneous multi-component analysis (Azi et al., 2021). For the water kefir biological system, most existing studies have been performed using targeted metabolomic techniques, the two most commonly used techniques being high-pressure liquid chromatogram coupled with triple quadrupole mass spectrometry (HPLC-MSMS) and headspace/solid phase microextraction coupled with gas chromatogram and mass spectrometry (HS/SPME-GC–MS) (Hu et al., 2014; Patel et al., 2022). These methods have resulted in the identification of bioactive

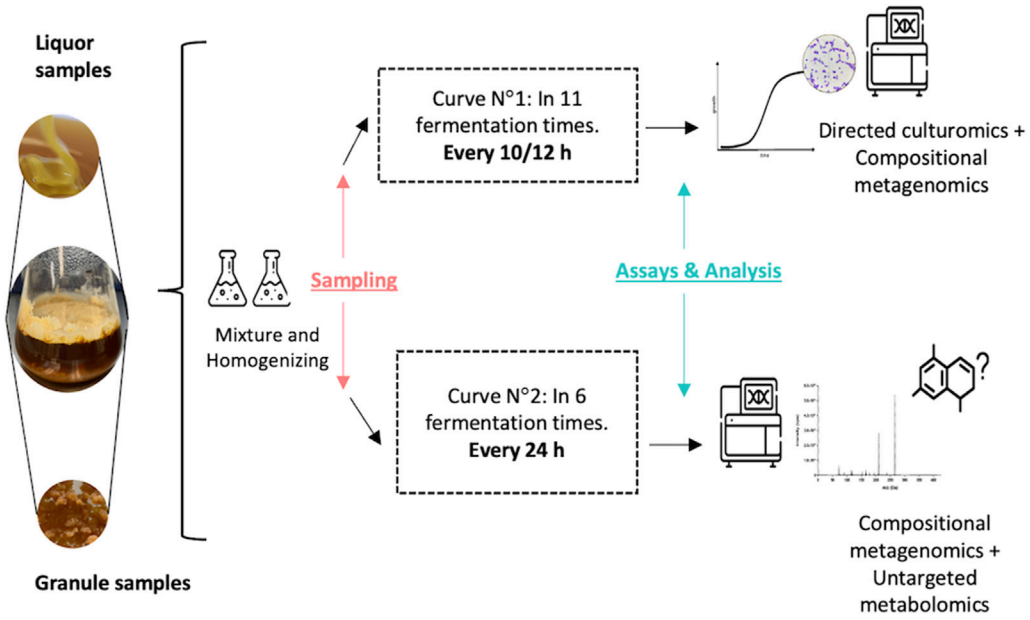


FIGURE 1
Experimental setup for microbial consortium sampling.

compounds such as organic acids, amino acids, flavonoids, and phenols, which have been reported as health-promoting agents (Azi et al., 2021).

Based on the statements above, evidence suggests that understanding complex biological systems, such as water kefir, requires the application of multiple methods that provide different perspectives (Weckx et al., 2019). Accordingly, the following project aimed to explore and characterize the microbial chemical space of a water kefir fermentation by integrating three different -omics approaches. Also of interest when developing the study was identifying different points of microbial enrichment during fermentation, which can serve as the knowledge basis for the future rational design of synthetic consortia, with a potential application as a health-beneficial food supplement. This can be reached using the knowledge provided by this study on the WK matrix and the strains collection isolated during its development. We consider that this approach not only deciphers the microbial chemical diversity of a locally produced (in Colombia) and modified fermented food but also enables its further use as a therapeutic ingredient, such as in probiotics and postbiotics.

2 Materials and methods

2.1 Water kefir fermentation and sampling

The water kefir fermentation was obtained from a private company in Colombia (Rionegro, Antioquia). As the origin of the sample remains unknown, the applied methods did not require the management of permits for the collection of biological material or access to genetic resources under the authorization of MinAmbiente (Colombia). Figure 1 shows the schematic diagram of the experimental setup. Two independent fermentation curves were performed using 12 g of kefir grains (which can be obtained for reproducibility assays upon request to the authors) in 100 mL of diluted molasses (14° Bx) on each sterile glass flask, then fermentation was set at room temperature (18°C) until the specific sampling time. Briefly, for the first curve, fermentation samples (grains and liquor) were collected every 10–12 h by duplicates, for 5 days, to perform microbial composition analysis and culturable-dependent population analysis. In total, 22 fermentation liquor samples (corresponding to data from 11 times by duplicates) and 2 inoculum (grains) samples were collected. The second growth curve was sampled every 24 h for 5 days (six different time points, in duplicates—14 samples), for the identification of non-culturable microorganisms and performing non-targeted metabolomic analysis (chemical space screening).

2.2 Culture-dependent methods

Three semi-selective and one non-selective culture media were chosen: MRS media, for lactic acid bacteria (62 g agar per L of distilled water liter (PanReac AppliChem, Darmstadt, Germany)); YM agar media for yeast (5 g of Peptone, 3 g Yeast extract, 3 g Malt extract, 10 g Dextrose, and 20 g Agar per L of distilled water, all substances from PanReac AppliChem, Darmstadt, Germany); WL, for acetic acid bacteria, (80,25 g per L of distilled water, Sigma-Aldrich, Saint Louis, United States); and MM, standardized culture media for total molasses degrading microbial biomass, which consist of diluted molasses to 14° Bx plus 15 g/L of bacteriological agar. Culture media were inoculated with 50 µL of serial dilutions of kefir grains homogenized with sterile 0.9%

saline solution and liquor at different fermentation times (10^{-2}, 10^{-3}, 10^{-4}) then incubated at room temperature (+/−25°C) for 72 h (3 days). Each dilution was cultured in duplicates. Afterward, colonies were counted in each dilution plate, differentiating by morphotypes, and registering CFU/mL for each culture media and fermentation time. Simultaneously, the isolation of representative morphotypes found in each culture media and fermentation time was registered. The purified morphotypes were Gram stained as a preliminary classification and then cryopreserved by duplicates at −80°C for downstream DNA extraction to constitute the first version of the strains collection derived from this microbial consortium.

2.3 Strains collection derived from representative morphotypes

Total DNA extraction from the resulting isolates was performed using the DNeasy Ultraclean Microbial Kit (QIAGEN, Hilden, Germany) for molecular identification through Sanger sequencing performed in the Macrogen Inc. sequencing facility (Seoul, South Korea). Primers used to perform the analysis were 27F (5′ AGAGTTTGATCMTGGCTCAG 3′) and 1492R (5′ TACGGYTACCTTGTTACGACTT 3′) for bacterial 16S rRNA hypervariable region (M. Y. Chen et al., 2022), and for yeast, the internal transcribed spacer ITS1 (5′ TCCGTAGGTGAACCTGCGG 3′) and ITS4 (5′ TCCTCCGCTTATTGATATGC 3′) (Taheur et al., 2017; Tan et al., 2022). Then, obtained data was uploaded to Geneious Prime ® 2021.2.1 software (https://www.geneious.com) to perform sequence trimming, alignment, and finding of consensus for subsequent BLAST classification (using 100% accuracy parameter). Finally, the collection of purified strains was stored at the Laboratory Center of Universidad EAFIT in Medellín, Colombia.

2.4 Total DNA extraction and library preparation

Genomic DNA extraction from 1.8 mL of sample (homogenized kefir grains and fermentation liquor) was performed using the DNeasy Ultraclean Microbial Kit (QIAGEN, Hilden, Germany) with the following modifications: treatment of the kefir grains with 0.9% saline solution prior to extraction and two additional steps to improve the cell lysis consisting of lysozyme incubation (37°C) for 15 min and water bath (70 °C) for 10 min + enzyme treatment with Proteinase K (QIAGEN, 2020). Extracted DNA for each sample was verified for its quality and integrity through electrophoresis gel and nanodrop quantification. Genomic samples were used to construct multi-amplicon libraries using the SWIFT AMPLICON® 16S + ITS PANEL protocol (Swift Biosciences, Ann Arbor, United States) with a primers pool covering all variable regions of the 16S rRNA gene (V1-V9) and the fungal ITS 1 and ITS 2 genes for the identification of bacteria and yeast (Swift Biosciences, 2018; Gao and Zhang, 2019). The sequencing process was performed using the Iseq 100 system with a 2 × 150 bp read length (Illumina, United States) at EAFIT University sequencing facility (AXOMICS).

2.5 Metabarcoding sequencing and data preprocessing

Sequence data processing was subjected to quality checks and analyzed using Qiime 2 workflow (Bolyen et al., 2019). For the bacterial analysis, obtained reads were divided by each region (V1–V9 of the 16S rRNA gene), and reads from the V4 region were selected to arrange the data into Amplicon Sequence Variants (ASVs) using DADA2, also correcting errors in sequences by removing singletons, chimeric sequences, and dereplicating data (Callahan et al., 2016). Data was then clustered into OTUs (Operational Taxonomic Units) with a 97% similarity. The taxonomic classifier for the classification was SILVA 138 SSU and GTDB databases for bacterial reads and UNITE for yeast identification (Beccati et al., 2017; Kõljalg et al., 2020; Parks et al., 2020). After quality checks and taxonomic categorization were done, alfa and beta diversity metrics were determined using the Phyloseq package and the integrated development environment for R, RStudio (version 1.4.1106) (RStudio Team, 2020).

2.6 Untargeted metabolomics analysis

As stated in Figure 1, samples were collected from a second curve at five different fermentation times (0, 72, 82, 92, and 120 h) and one media control (molasses) was also included. These samples were stored at −80°C until metabolite extraction was performed. Each fermentation time had four replicates resulting in a complete dataset of 24 samples to be analyzed. Liquor and control extracts were obtained by methanolic extraction using 50% MeOH. Briefly, 20 mL of each homogenized sample was mixed with 4X solvent in sterile glass flasks; afterward, the solution was sonicated at 30% amplitude for 20 min (pulse function 1 min on + 30 s off) and agitated in the dark at 200 rpm for 4 h. The resulting solution was centrifuged at maximum speed (4,500 rpm) for 15 min and filtered with a vacuum pump and Whatman filter papers (grade 2). Solvent evaporation was performed with a Rotavapor R-300 using the manufacturer's instructions (BÜCHI, Flawil, Switzerland). Extracts were vacuum evaporated with the Concentrator Plus system (Eppendorf; Hamburg, Germany). The final solid residue was Weighed and resuspended in 2 mL of 50% MeOH for downstream procedures.

2.7 Metabolomic data acquisition and pre-processing

The resulting 24 extracts (4 replicates per sample) were sent to the Metabolomic Core Facility (MetCore) at the Universidad de los Andes in Bogotá-Colombia. Samples were vortexed at 3,200 rpm and filtered on 0.22 µm filters. Then, 80 µL of each extract was taken for subsequent untargeted metabolomic analysis using an Agilent Technologies 1260 Liquid Chromatography system coupled to a 6545 Q-TOF quadrupole time-of-flight mass analyzer with electrospray ionization (RP-LC/MS-QTOF). For the reverse phase, the injection volume of the samples was 10 uL and the compound separation was done on a C18 column (InfinityLab Poroshell 100 × 3.0 mm, 2.7 µm) at 40°C. The mobile phases

TABLE 1 Parameters used for Classical Molecular Networking in GNPS platform.

Parameters	Values
Minimum pairs cosine	0,7
Minimum fragmented ions	6
Cluster size	2
Minimum matched peaks	6
Search analogs	Do search

used for elution were composed of 0.1% (v/v) formic acid in Milli-Q water (Phase A) and 0.1% (v/v) formic acid in acetonitrile (Phase B) pumped at 0.4 mL/min. The mass detection was performed in positive ESI mode on autoMS/MS from 50 to 2000 m/z. Tandem mass spectrometry data obtained was pretreated and converted to mzML format using MSConvert GUI (Adusumilli and Mallick, 2017). Throughout the analysis, two reference masses were used for mass correction: *m/z* 121.0509 ($C_5H_4N_4$), and *m/z* 922.0098 ($C_{18}H_{18}O_6N_3P_3F_{24}$) corresponding to protonated purine and protonated hexakis, respectively.

2.8 Molecular networking and metabolites identification

mzML converted files were uploaded to the GNPS server using an FTP client with the corresponding input parameters. Then, the files (ftp://massive.ucsd.edu/MSV000091955/) were used to perform Classical Molecular Networking) in the same platform (Wang et al., 2016), and the resulting networks were visualized on Cytoscape version 3.9.1 (Shannon et al., 2003). Different analyses were run by changing some of the network parameters, including min pairs cosine value, min of fragment ions, and number of matched peaks with final parameters stated in Table 1. After these analyses, the molecular network that best suited the objective (broad identification with rigor of a minimum of six peaks identified) was selected to perform manual curation and annotation of the chemical families. The resulting feature table with all nodes' information was downloaded in ".csv" format from Cytoscape version 3.9.1. Then, each subnetwork information was downloaded and treated as a cluster or chemical family. Each annotation or library identification within a subnetwork was revised for its level of identification (either gold or bronze and number of shared peaks) with reference to the spectra and peak list downloaded from public databases (MassBank of North America—MoNA (https://mona.fiehnlab.ucdavis.edu/) and CEU Mass Mediator (Gil-De-La-Fuente et al., 2019). In this way, when a single node or feature was annotated within a cluster (subnetwork), all neighbor nodes were grouped and named under the chemical family of this annotated compound, since the clustering algorithm groups by structural similarity according to the parameters set for the network (Table 1) (Schrimpe-Rutledge et al., 2016). Thus, each cluster was named when a single or multiple features in it were correctly annotated by library detection in GNPS, and the reference MS/MS of the compound was checked from public repositories using MassBank of North America—MoNA (https://mona.fiehnlab.ucdavis.edu/) and CEU Mass Mediator (Gil-De-La-Fuente et al., 2019). This procedure was specifically done for nodes denoting a weak annotation (six shared peaks and bronze standard). For the nodes that were not identified by GNPS libraries and with the aim of expanding the networks' annotation, a manual search of selected features was done using the precursor mass value as a query in chemical databases, namely, PubChem (Kim et al., 2021) ChemSpider (Pence and Williams, 2010), the Atlas of Natural Products (Santen et al., 2022). Each precursor mass on each node was also checked to unveil the presence of any of the three most common adducts, e.g [M + H+]. If a hit was found, the reference MS/MS was compared with the network feature, and if more than six peaks were found in common, an annotation hit was called. As an alternative way of propagating annotation of the network, if a cluster had more than one identified node, the chemical family that grouped all the annotations of the cluster was selected to name the cluster (e.g., haematommic acid and L-beta-3-phenyllactic acid were identified in the same cluster, both with a benzene ring in the structure and belonging to the more general category of phenolic acids). Finally, the modified network with the annotated notes was downloaded in a high-resolution image format (".png" with 600 DPI), with the annotated chemical families highlighted in circles and named after each search.

2.9 Statistical analysis using metaboanalyst 5.0

Simultaneously to the molecular networks, raw data obtained from RP-LC/MS-QTOF was converted to mzML format and uploaded to the Metaboanalyst 5.0 (https://www.metaboanalyst.ca/) (Pang et al., 2022). This platform was used to perform multivariate analyses on all samples from the 'LC-MS Spectra Processing' option. Principal components analysis (PCA) was performed using the default parameters in order to validate the significance of the sampling and to obtain insights about the treatments used in the experiment (fermentation times).

2.10 Statistical and integrative analyses

After the analyses were performed, statistical methods were applied to visualize the distribution patterns between the two types of samples (kefir grains and liquor) and the fermentation times using the integrated development environment for R, RStudio (version 1.4.1106) (Rstudio Team, 2020). Filtered data from the compositional metagenomic approach consist of ASVs that had more than 32 sequencing reads. Then, dominant species ASVs were selected to perform statistical validation through a non-parametric Mann—Whitney *U* test (p-value <0.05) using the function wilcox.test() in the R Stats package, looking for significant differences between the treatments (fermentation times and sample origin) (Supplementary Additional file S4) (Rstudio Team, 2020). In addition, the Pearson correlation test was conducted as a primary approximation to the association between observed ASV and annotated chemical families through the R function cor() from the R Corrplot package (Wei and Simko, 2021).

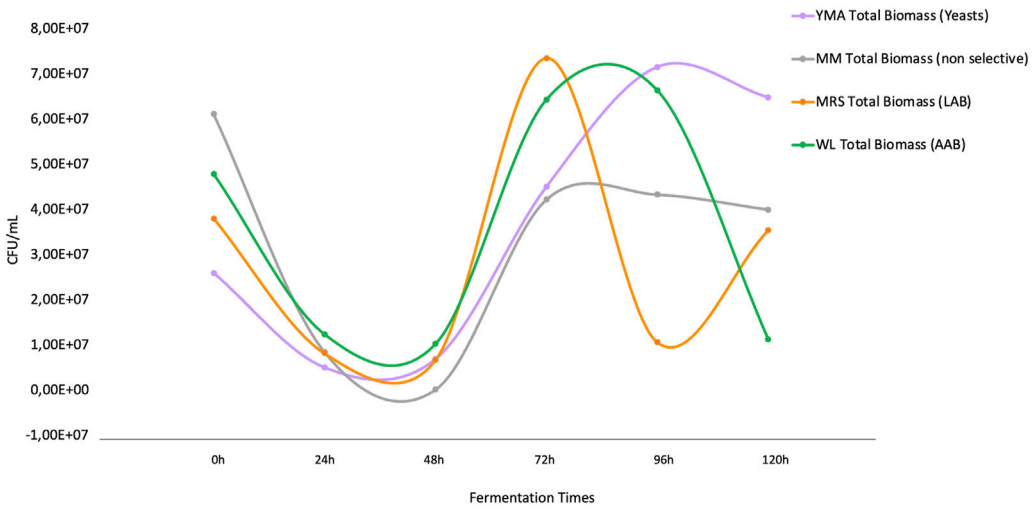


FIGURE 2
Population dynamics based on **culture-dependent methods** of the four microbial groups in the study. The orange line indicates the growth of lactic acid bacteria (MRS media), the green line indicates acetic-acid bacteria (WL media), the purple line indicates yeasts (YM media), and the gray line indicates total molasses degrading biomass (MM media).

3 Results

3.1 Culturable microbial communities in water kefir

The dynamics of representative microbial groups cultured from the fermentation are shown in Figure 2. It was observed that the groups under study had a relatively similar growth pattern during the fermentation process until 72 h post-inoculation of the grains was reached. Molasses culture media (MM) reflects an average of the culturable biomass, non-biased to specific nutritional compositions of the other media but supporting the growth of the entire community in the fermentation. The total biomass in the media was the lowest after 72 h (4.22E+07 CFU/mL), reflecting a stationary phase with a slight negative slope. Instead, for the other culture media that were selective for specific microbial groups (See Materials and Methods), high microbial growth was observed at 72 h for *Lactobacillus* sp. communities that were selectively grown in MRS media (Figure 2), which decreased rapidly in nearly 1 order of magnitude and retook a second stage of growth after 96 h. For acetic acid bacteria growing selectively in WL medium, the peak growth was observed before 72 h and lasted until after 96 h of fermentation, then rapidly decreasing and alternating with the regrowth of *Lactobacillus* sp. For yeasts growing in YM agar media, the peak growth is observed at the end of the fermentation, at exactly 96 h, after which, it starts to decline.

3.2 The strains collection includes the most representative groups reported for water kefir, exhibiting probiotic potential

A total of 95 samples were isolated, 73.7% represented by bacteria and 26.3% by yeasts. In addition, they were classified by the source of isolation resulting in 63.12% of the strains coming from fermentation liquor and 29.47% from kefir grains (Table 1, Supplementary Additional file S1). Among the microorganisms identified by Sanger sequencing, we could identify *Lentilactobacillus hilgardii, Lactobacillus buchneri, Lacticaseibacillus paracasei, Schleiferilactobacillus harbinensis, Acetobacter pasteurianus*, and *A. tropicalis* (Gulitz et al., 2011; Olivo et al., 2017; Lynch et al., 2021; Rodríguez et al., 2022) and yeasts such as *Pichia membranifaciens* and *Kazachstania exigua* (Nejati et al., 2020; Moretti et al., 2022). These microorganisms have been previously reported to have probiotic effects and, thus, provide a useful resource for the underlying interest of the project in providing a knowledge and resource basis for the future development of health-beneficial microbially derived supplements, such as probiotics and postbiotics.

3.3 Total microbial communities obtained through compositional metagenomics reflect similar patterns to what is observed in culturable population dynamics and suggest an enrichment of species at 72 h of fermentation

The successful sequencing of the prepared libraries derived from the grains inoculum (2) and fermentation liquor (22) resulted in a dataset of 24 samples with a combined size of approximately 6.23 Gbp and a quality score (Q30) of 94.48% (Supplementary Table S2, Supplementary Additional file S1). The obtained results from compositional metagenomics (V4 region of the 16S rRNA gene) support the hypothesis that the microbial consortium corresponds to a water kefir-derived product since the main groups belong to lactic acid bacteria, acetic acid bacteria, and yeasts (Fiorda et al., 2017; Çevik et al., 2019; Guzel-Seydim et al., 2021; Patel et al., 2022; Yerlikaya et al., 2022). According to the relative abundances of the six dominant species that can be observed in Figure 3, it was possible to observe that

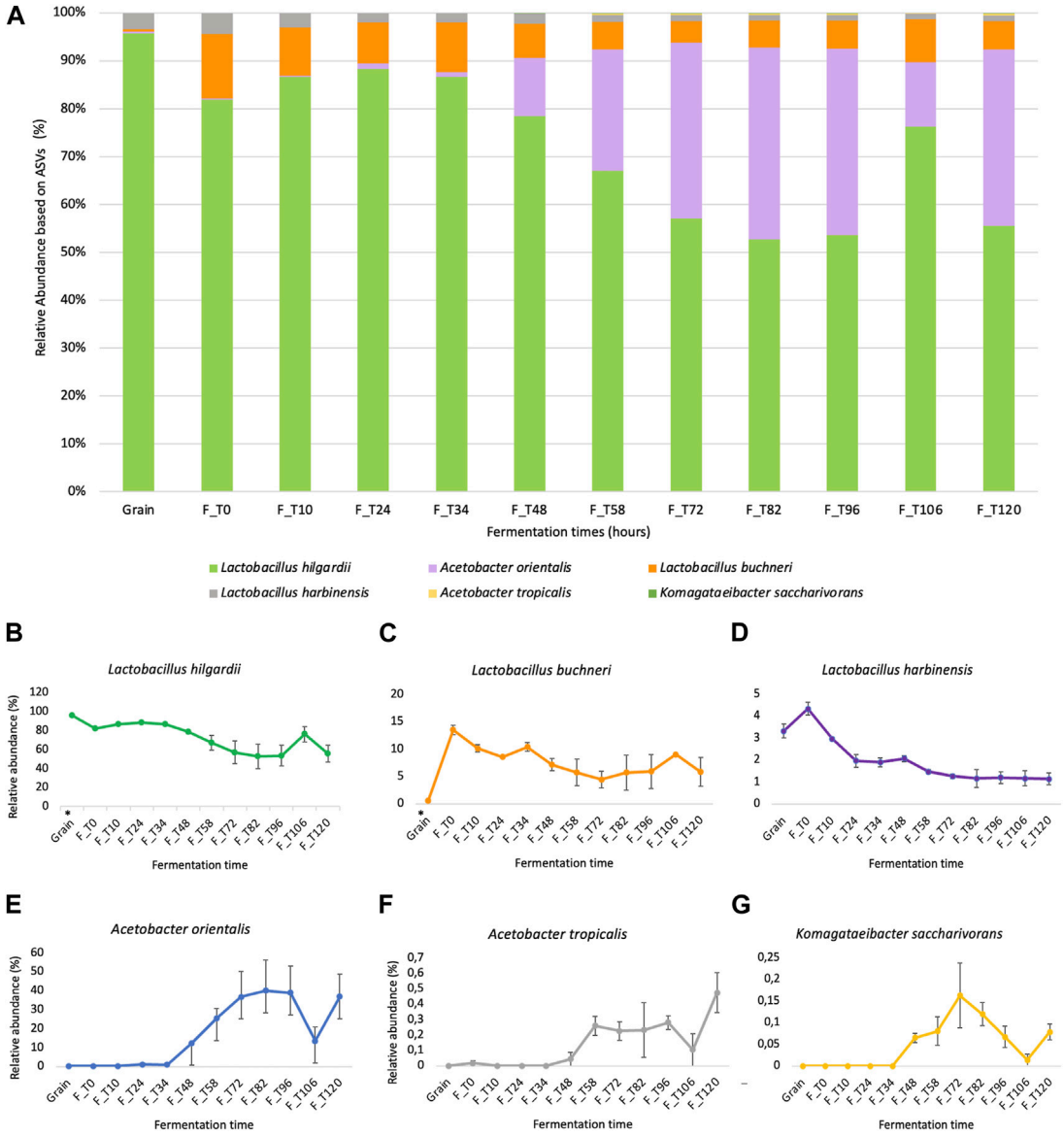


FIGURE 3
Relative abundance of dominant bacterial ASV identified through compositional metagenomics analysis. **(A)** Relative abundance of the six dominant bacterial ASVs on each fermentation time sample between 0 h and 120 h, including the grain inoculum. **(B–G)** Ind Individual kinetics showing the behavior of each identified ASV over time (*L. hilgardii*, *L. buchneri*, *L. harbinensis*, *A. orientalis*, *A. tropicalis*, and *K. saccharivorans*). The (*) in grain samples for *L. hilgardii* and *L. buchneri* indicates statistically significant differences in the abundance values for these groups according to sample origin (Mann-Whitney *U* test, *p*-value = 0.007246).

L. hilgardii was dominant in the grain sample, representing ~90% of it, and at early fermentation times until 34 h, starting to decrease after 48 h of the process to ~50% representation. For acetic acid bacteria instead, they increased from 1%–12%–~40% between 34 and 48 h of fermentation.

The kefir grains that were used as fermentation inoculum were mostly represented by lactic acid bacteria, which is consistent with previous studies in sugary beverages (Yerlikaya et al., 2022) (Supplementary Additional file S2). The results are partially in accordance with the observations in the population dynamics of culturable microbial groups (Figure 2); other species, mainly *Acetobacter orientalis* and *Lactobacillus buchneri*, unlike those in the grain (mainly *L. hilgardii*), increased significantly after 72 h, which can be explained from the perspective of the multispecies interactions and metabolite production taking place in the system. For the case of *L. hilgardii* and *L. buchneri,* statistically significant differences were found between the grains and fermentation liquor (*p*-value = 0.007246) through the Mann-Whitney *U* test. For the members of the family Acetobacteraceae, no significant test were performed since this group was not detected on the grain inoculum, but it is obvious that the population change is significant, passing from almost non detectable to representing ~40% of the microbial community in the matrix. Regarding the metabarcoding results for the ITS region (ITS1-ITS2), all obtained reads were classified as

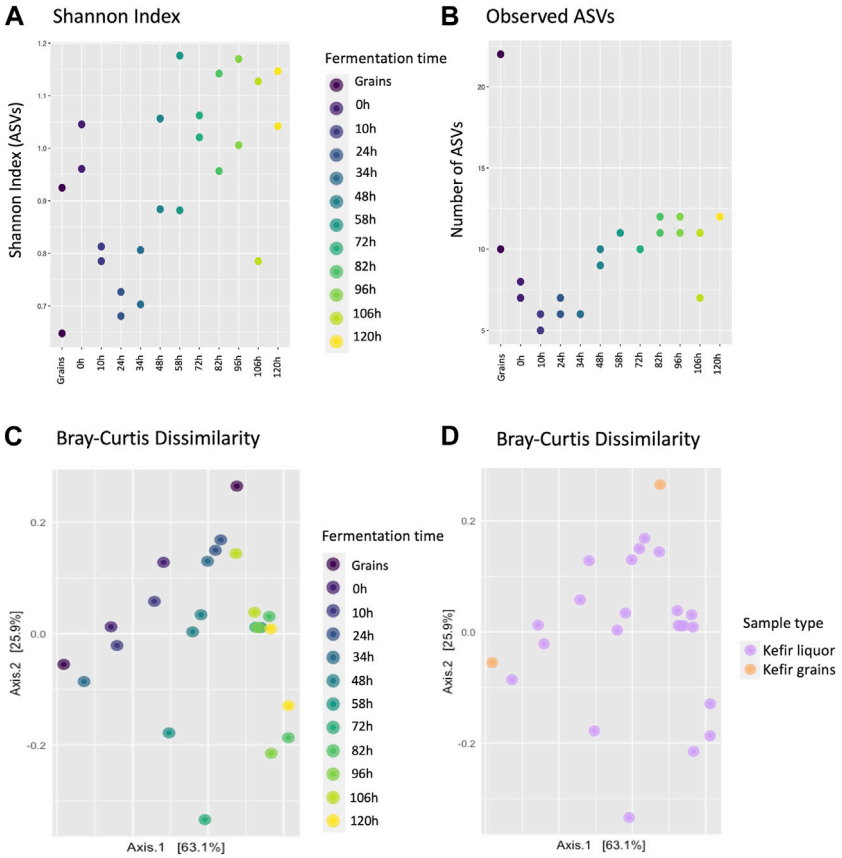


FIGURE 4
Estimation of diversity indices of the microbial communities during the fermentation process, based on the normalized ASVs counts/ per volume of sample for the microbial communities identified during the fermentation process. **(A)** Scatter plot showing the Shannon index (alpha diversity) estimated for each fermentation time, according to the identification of ASVs and grouped by fermentation sample. **(B)** Scatter plot of the number of identified ASVs on each sample classified by fermentation sample. **(C)** Scatterplot for the Bray-Curtis dissimilarity index (beta diversity) for classified ASVs on the fermentation liquor samples and grains inoculum. **(D)** Scatterplot for the Bray-Curtis dissimilarity index estimated by fermentation sample type.

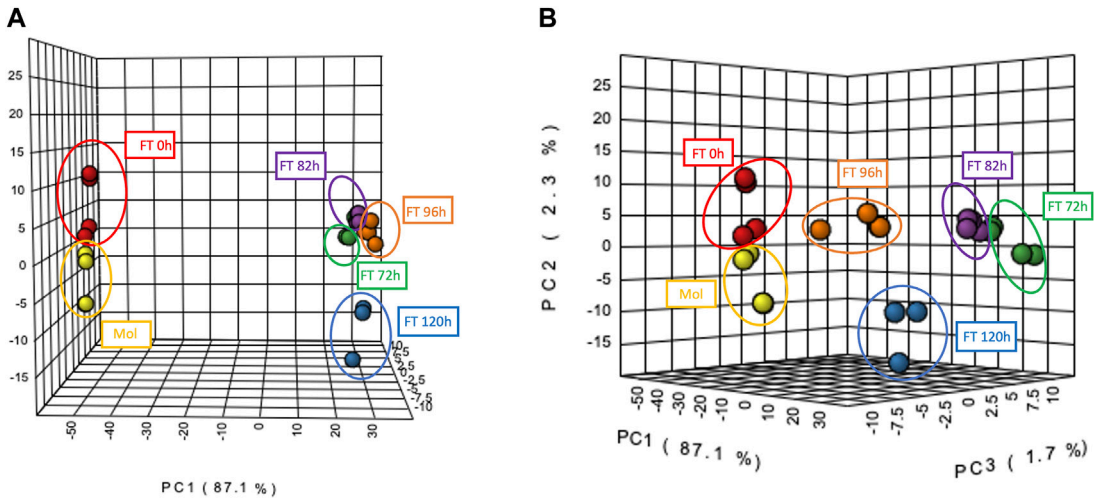


FIGURE 5
Principal Components Analysis (PCA) among molasses and fermentation samples generated by Metaboanalyst 5.0. **(A)** Red and Yellow clusters represent samples from molasses and fermentation time 0 h; green, purple, blue and orange clusters represent samples from the latest fermentation times from 72 to 120 h. **(B)** Distribution of the four replicates for each of the six fermentation samples.

P. membranifaciens, given its high abundance. Besides, there is evidence of coverage of the panel used for the metataxonomic libraries preparation favoring the 16S rRNA gene regions for bacterial identification (Figure 1, Supplementary Additional file S1).

3.4 Biodiversity of the water kefir fermentation decreases in the first stages of fermentation and restarts again after 48 h post-inoculum

The biodiversity of the grains inoculum and fermentation liquors was measured in terms of ecological indices using the Phyloseq package from the integrated development environment for R, RStudio (version 1.4.1106) (Supplementary Additional file S2). Alpha diversity was measured as the richness or dominance of species in the microbial communities of each sample. Shannon index values between 0.7 and 1.2 were obtained (Figure 4A). These diversity values indicated a gradual decrease in diversity between 10 and 34 h, which increased again after 48 h of fermentation.

In the same way, it was observed that the number of identified ASVs presents differences according to each fermentation time (Figure 4B). This result coincides with previous studies that show that the grain and liquor environment differ in microbial composition, with the grains typically having a higher microbial load (Moretti et al., 2022; Patel et al., 2022). Regarding the replicates, differences were observed between the values of the Shannon index and the number of ASVs for some of the replicates at each time, specifically in the grain inoculum and fermentation liquor at 48, 58, 82, 96, and 106 h.

In terms of b-diversity, the Bray-Curtis dissimilarity index was calculated and visualized by each fermentation time and the sample type. There was not a clear clustering based on the fermentation time; however, it was possible to observe a "vanishing" pattern from the top to the bottom of the graph (Figure 4C). According to the sample type, it was found that the two grain samples differed from each other, and for the fermentation liquor, there was no clustering pattern (Figure 4D).

3.5 Chemical space of the biological system composed by the water kefir community clusters in two groups according to multivariate statistical analysis

In order to analyze the system from a chemical point of view, a multivariate statistical analysis using Metaboanalyst was performed and allowed to group the samples by their chemical composition in two clusters based on the principal component 1 (PC1), suggesting that samples corresponding to molasses and early stages of fermentation exhibit a similar metabolite profile among them, while the samples from the latest times (72–120 h) of the fermentation share a common chemical profile (Figure 5A). This clustering was consistent with enrichment points identified by previous methods, indicating microbial interactions that could lead to the release of different amounts of metabolites. In addition, the four replicates of each treatment (fermentation time or control/molasses) were able to cluster together. This indicates that there are no significant differences within each replicate of a treatment between fermentation times (Figure 5B).

3.6 Untargeted metabolomics evidences a rich chemical biology derived from the water kefir fermentation, with different reported bioactivities

757 nodes and 260 compound annotations were obtained from the molecular network generated by GNPS bioinformatic platform (https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=2c80ff4b1e2a49fea7e844644198bbf3) with the selected criteria (Table 1). After manual curation and annotation, it was possible to identify 18 chemical families, such as phenolic acids, quinolines, flavonoids, monoterpenoids, organic acids such as lactic acid, and amino acids (Figure 6, Supplementary Additional file S3 Spectra_confirmation). It was possible to highlight the fermentation sample where each family was prevalent based on the number of spectra identified in each sample. At the beginning of the fermentation (Molasses and FT 0 h), the chemical families identified were glycerolipids and monoterpenoids. Families such as phenolic acids, flavonoids, and benzene products were found during the latest times (between 72 and 120 h).

3.7 Bacterial group presence in specific times of water kefir fermentation correlates with the enrichment of chemical families identified through metabolomics

The correlation was estimated between the six dominant ASVs identified down to the species level and the six dominant chemical families from the fermentation extracts (Figure 7). *Lactobacillus hilgardii*, *L. buchneri*, and *L. harbinensis* decrease in a direct correlation with glycerolipidss, a chemical family of compounds that plays important roles in cell signaling, membrane trafficking, and anchoring of membrane proteins (Henry et al., 2012). *Acetobacter orientalis* and *A. tropicalis* were positively correlated with four out of six chemical families. Furthermore, some metabolites of interest (e.g., phenolic acids) were plotted individually with observed dominant ASVs (Figures 7B–E), finding that *A. tropicalis* and bile acids conjugate correlate directly (Figure 7C), while *L. hilgardii* and glycerolipids have a direct correlation more pronounced at later stages of the fermentation, while *A. tropicalis* and the same chemical family have a negative correlation (Figures 7A, D, respectively), suggesting the sensibility of this species for these compounds or a capacity for degradation. Furthermore, phenolic acid concentration inversely correlates with populations of *L. harbinensis*.

4 Discussion

The use of different 'omics' techniques to study fermented matrices has significantly increased during the last decade, given the broader scope they offer to gain an understanding of the different layers involved in any biological system. Nevertheless, no reports of the integration of culture-dependent methods with compositional metagenomics have

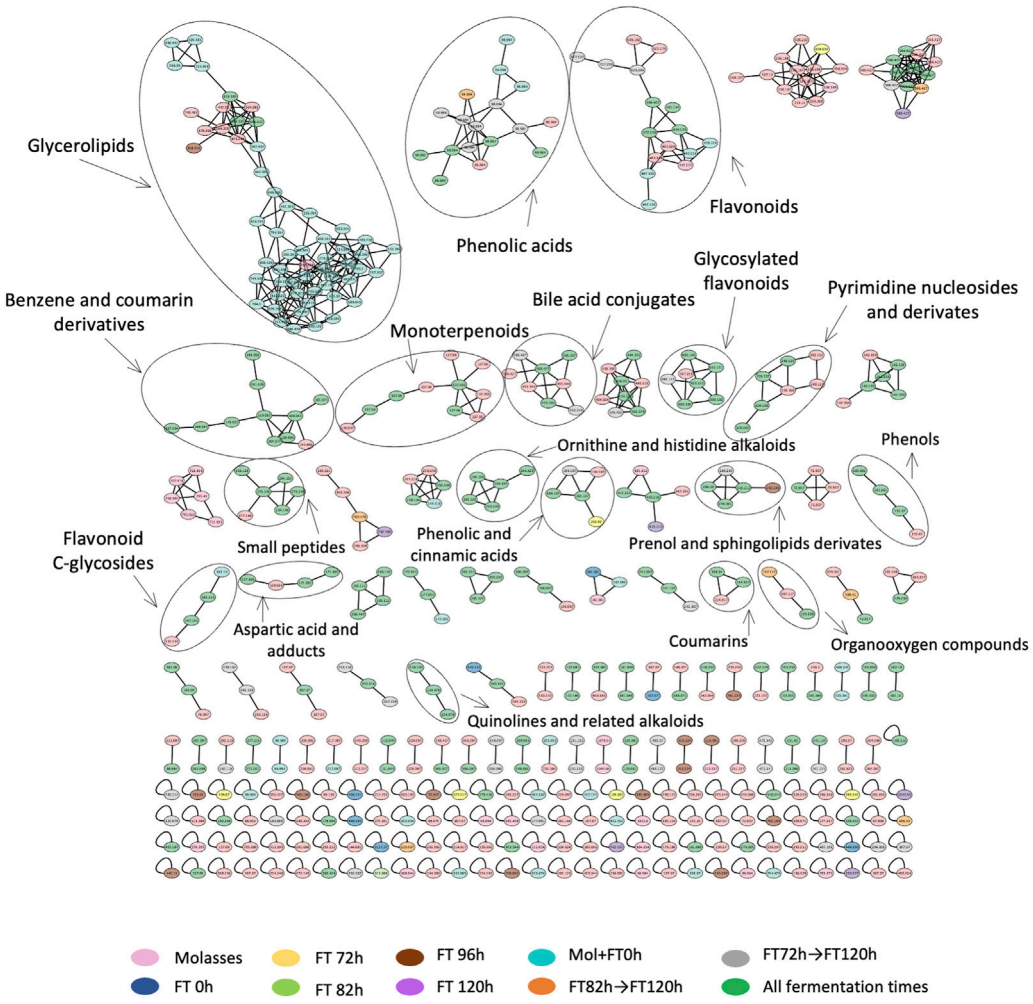


FIGURE 6
Curated molecular network with identified chemical families from the extracts processed by RP-LC/MS-QTOF. Nodes are colored by treatments: molasses control (pink), 0 h liquor (dark blue), 72 h liquor (yellow), 82 h liquor (green), 96 h liquor (brown), 120 h liquor (purple), molasses +0 h liquor (cyan), from 82 h to 120 h (orange), from 72 h to 120 h (gray), and presence in all groups (dark green).

been found prior to this research, making this the first of its kind. According to the findings of the study, the WK fermentation of the study, including its modifications according to local Colombian culture, still corresponds to a water kefir-derived product. The before is stated, since the main microbial groups found belong to lactic acid bacteria, acetic acid bacteria, and yeasts, as is expected from this kind of biological system (Fiorda et al., 2017; Çevik et al., 2019; Guzel-Seydim et al., 2021; Patel et al., 2022; Yerlikaya et al., 2022). The majority of reports on the composition of water kefir report *Lactobacillus* and *Acetobacter* as the representative genera in these fermentation processes (Gulitz et al., 2011; Fiorda et al., 2017; Verce et al., 2019). Some previous studies have also described dominant bacteria of the genera *Streptococcus*, *Leuconostoc,* and *Lactococcus* and yeast genera such as *Dekkera* and *Saccharomyces* (Laureys et al., 2022; Pihurov et al., 2023). Contrary to this last report, we found in our WK a hybrid matrix with a major presence of *Lactobacillus* sp. and *Acetobacter* sp but an absence of *Lactococcus*, *Leuconostoc, and Streptococcus* sp. on the bacterial composition. Regarding yeasts, we found a high abundance of *P. membranifaciens, Pichia kudriavzevii,* and *K. exigua.* This finding is interesting since some studies, which even change substrate conditions (Laureys and de Vuyst, 2016; Zannini et al., 2023), have not detected *P. membranifaciens* nor *K. exigua*, denoting the uniqueness of this WK fermentation and its chemical space. Specifically, this species, *K. exigua,* has been more commonly found in olive brine, wine, and other fermented matrices (Jood, I et al., 2017). There have also been some reports of its presence in superficial and subterraneous waters, sediments, and soils in crude extracting zones (SIB Colombia, 2023). According to the relative abundances of dominant species (Figure 3), *L. hilgardii* was observed to be dominant in the grain sample and during early fermentation times, starting to decrease after 48 h of the fermentation, which is probably related to the proliferation of other community members that compete for substrate and generate a novel chemical environment. Furthermore, this species has been associated with water kefir fermentations and is thought to be responsible for the growth of kefir grains due to the production of EPS, such as dextran, which is a sucrose derivate with great potential in the industry due to its relative stability and good solubility (Lebeaux et al., 2014; Laureys and de Vuyst, 2016; Lynch et al., 2021). Lactic acid bacteria are a commonly

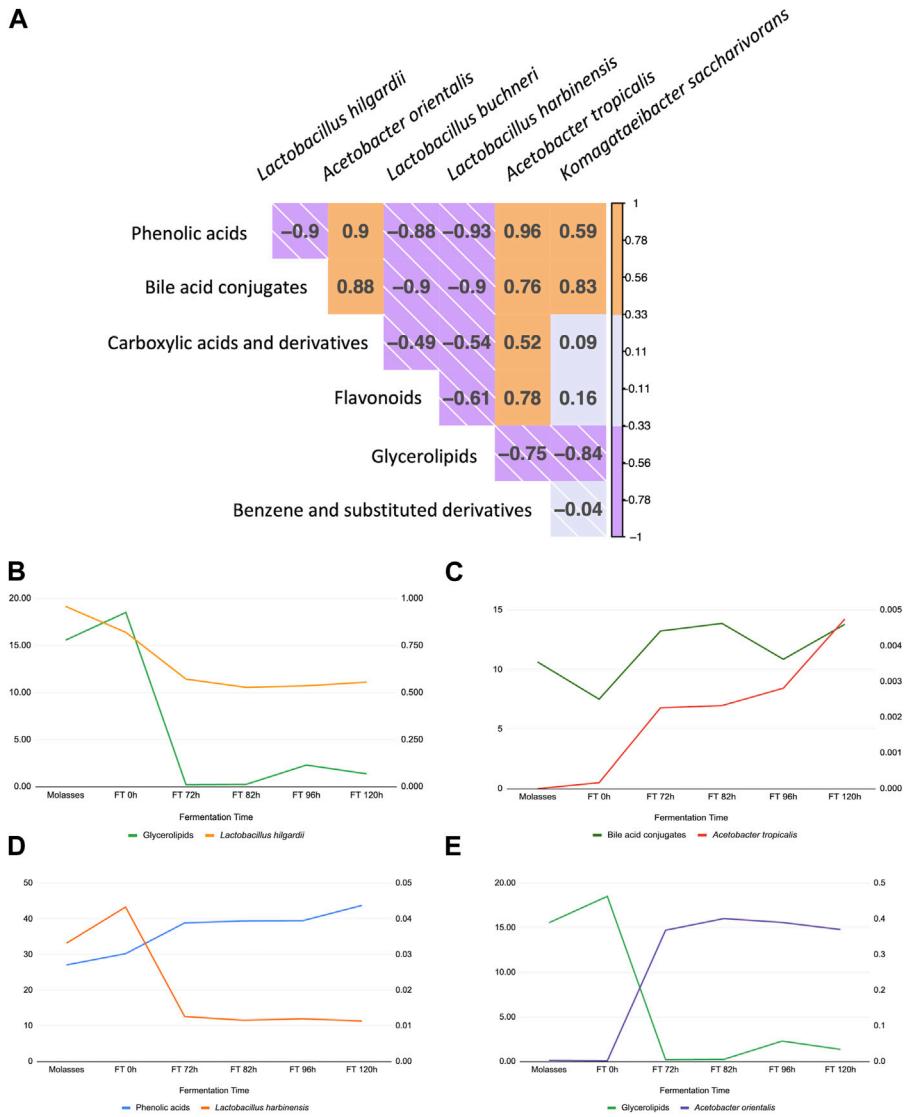


FIGURE 7
Correlation between compositional metagenomics and identified chemical families during the fermentation of the microbial consortia. **(A)** Correlation visualization between dominant ASVs, three *Lactobacillus,* and three *Acetobacter* taxa with dominant chemical families reported from GNPS molecular network. **(B–E)** Individual plots showing the change in each metabolite and taxa of interest over time (*L. hilgardii, L. harbinensis, A. orientalis,* and *A. tropicalis*), supporting the association obtained by Pearson's correlation.

known group due to their probiotic effects and the production of organic acids that preserve and improve the aromatic and bioactive qualities of water kefir and other fermented foods (Azi et al., 2021; Spizzirri et al., 2023). Even though we found a high abundance of various *Lactobacillus* species among the most abundant microbial groups, several other species of acetic acid bacteria and some yeasts increased after 72 h, suggesting that multispecies interactions and metabolite production contributed to their growth. These observations could be explained by the associations between different microorganisms, e.g., the acidification of the media by *Lactobacillus* species and the use of end-products as an energy source can improve the growth of acetic acid bacteria and yeasts; meanwhile, essential nutrients released by yeasts can support bacterial growth (Pendón et al., 2021). Regarding the fermentation substrate, molasses has been previously reported as a low-cost alternative for industrial fermentation of water kefir with a variety of potential biological activities, such as antioxidant capacity, due to the presence of sugar-derived compounds that induce the growth of microorganisms of interest as well as the production of metabolites that facilitate multispecies interactions (Deseo et al., 2020; Mordenti et al., 2021).

As we stated before, the relative number of reads for lactic acid bacteria decreased in time, and we observed that acetic acid bacteria species appeared in a relatively high abundance after 34 h, which could be due to their ability to grow in low concentrations of oxygen and, at the beginning of the fermentation, there is high oxygen availability and higher sugar levels in the media (Pendón et al., 2021) (Figures 3E, G). The genus *Acetobacter* has been reported in the microbial communities of kefir-based fermentations, but its role has not been fully elucidated, except for the fact that they can contribute to the flavor and aroma of the final fermented product and

metabolize sugar and alcohol through the pentose phosphate pathway to accumulate large amounts of diverse fermentative products including D-sorbitol, ascorbic acid, and the prebiotic levan, which are derived from the principal metabolite of this group, acetic acid (Azi et al., 2021; Patel et al., 2022; Yerlikaya et al., 2022). *Acetobacter orientalis* and *A. tropicalis* have been reported as the dominant species in previous studies. These species are known as fermentation stabilizers and contributors to the aroma of water kefir (Gulitz et al., 2011; Martínez-Torres et al., 2017; Guzel-Seydim et al., 2021). We corroborate these findings, with the two most prominent acid bacteria species being *A. orientalis* and *A. tropicalis* but correlating their appearance in the fermentation with phenolic acids of several classes and molecular weights, with monoglycerides, isoleucine derivatives, and flavones (Figures 6, 7). Also, an interesting proportion of the acetic acid bacteria *Komagataeibacter saccharivorans* was found in this system. This one is a producer of cellulose and has been mostly found in milk kefir but not in water kefir grains.

On the other side, regarding biodiversity measures, low diversity values for index at the start of the fermentation could be explained due to the dominant microorganisms found in the matrix that belong to a few genera, such as *Lactobacillus* and *Acetobacter*. These diversity values indicate a gradual decrease in diversity between 10 and 34 h, which increased again after 48 h of fermentation. This behavior was associated with previous studies, such as that of Patel et al. (2022) and their diversity analysis in kefir grains and liquid from the sequencing of the V3-V4 region of the 16S rRNA gene. Similarly, it was observed that the number of identified ASVs presented differences according to the fermentation sample. This result coincides with previous studies in which it has already been established that the grain and liquor environments differ in microbial composition, with the grains typically having a higher microbial load (Moretti et al., 2022; Patel et al., 2022). This result could be explained since the kefir grains are the main source of microorganisms, and biodiversity in the beverage can vary based on the substrate, culture conditions, and microbial interactions that occur during the fermentation (Fiorda et al., 2017; Pendón et al., 2021; Moretti et al., 2022). Apparent differences between replicates were also observed, which could be related to population dynamics and/or sample manipulation, considering that they differ more in the value of diversity and not in the number of taxa identified (e.g., grain replicates). In terms of b-diversity, the results suggest that at the beginning of fermentation, there are more different species, while at the end of the process, the samples share a greater number of species. Furthermore, no clustering was observed for fermentation liquor samples. This can be explained by the fact that the microbial consortium is a complex and dynamic biological system that can exhibit different characteristics under similar conditions.

The results of metabolite screening were consistent with previous targeted studies designed to look for organic acids, alcohol levels, produced EPS, and other compounds derived from the fermentation process of water kefir grains (Plessas et al., 2017; Azizi et al., 2021; Moretti et al., 2022; Esatbeyoglu et al., 2023). From the identified metabolites, it was observed that there could be a correlation between the microorganisms found and the biological potential of the fermented product. For example, inhibitory effects against pathogens have been associated with the production of protective organic acids by lactic acid bacteria and ethanol produced by yeasts (Yerlikaya, 2019). The formation of organic acids released by LAB species in water kefir is an important indicator of enhanced metabolic activity since they may be used as a substrate by other groups of microorganisms (Bulat and Ali, 2021; Satir, 2022). Previous studies have shown that ß-glycoside enzyme and phenolic acids (e.g., haematommic acid and lactic acid) produced during the microbial fermentation release glycosylated or bound flavonoids such as saponarin and puerarin and produce new polyphenols (Azi et al., 2021). Furthermore, water kefir has a high antioxidant potential due to the phenolic compounds and enzymes found as fermentation end-products (Cai et al., 2020; Yerlikaya et al., 2022). Most of the bioactive components are classified as polyphenolic compounds in nature, which include the identified cinnamic acid, coumarin, thymol, and myrcene (Satir, 2022). In the generated molecular network, it was found that this chemical family was dominant in the latest times of fermentation, suggesting its relationship with acetic acid bacteria metabolism and other multispecies interactions. In addition to their role against oxidative stress, some phenolics have been reported as antimicrobial agents, including benzoic acid, which was also annotated in the molecular network (Rodrigues et al., 2016; Azi et al., 2022). In this way, the increase in the total phenolic compounds is strongly correlated with an increase in antioxidant activity, which is also associated with an anti-inflammatory effect of water kefir fermentations (Rodrigues et al., 2016). From the identified metabolites, it was possible to annotate and validate by reference spectra a compound of interest that had not been reported for water kefir-based fermentations. Isoschaftoside is a C-glycosyl flavonoid originally extracted from root exudates of *Abrus cantoniensis* (Guan et al., 2022). Generally, C-glycosyl flavonoids are part of the diet and have been reported to have a wide range of pharmacological activities including blood-lipid-lowering, hypoglycemia, neuroprotective, antitumor, and antioxidant capacities (Tremmel et al., 2021; Guan et al., 2022). The C-glycosylation of flavonoids gives rise to more stable, biologically active metabolites with different spectral properties and increased solubility in polar media compared to O-glycosides (Brazier-Hicks et al., 2009; Vanegas et al., 2018; Khodzhaieva et al., 2021).

From the resulting molecular network, an estimate of the relative abundances for each chemical family was obtained based on the number of spectra found on each sample. The dominant chemical family among the fermentation was that of phenolic acids, with increasing values after 72 h, which suggests their association with microbial interactions. As mentioned above, phenolic acids are known as bioactive compounds with a wide range of applications and have been associated as end-products from lactic-acid bacteria metabolism, which could explain their relative high abundance since the beginning of the fermentation, where lactic-acid bacteria are more abundant and increase in the latest times, possibly related to the presence of yeast and acetic acid bacteria (Păcularu-Burada et al., 2022).

Finally, correlation primary analyses suggested that *Lactobacillus hilgardii*, *L. buchneri,* and *L. harbinensis* showed a positive correlation with glycerolipids, a chemical family of compounds that play important roles in cell signaling, membrane trafficking, and anchoring of membrane proteins (Henry et al., 2012), an association that could be explained by their decrease in

relative abundance over the fermentation course and the changes in the fermentation conditions (e.g., pH values, oxygen levels, multispecies interactions), which could also support its negative correlation with *Acetobacter*, representing microorganisms that are tolerant to these environmental conditions or that can degrade this family of compounds. *Acetobacter orientalis* and *A. tropicalis* were reported to be positively correlated with five out of six chemical families, including phenolic acids of several classes and molecular weights, with monoglycerides, isoleucine derivatives, and flavones, which support the hypothesis that the latest fermentation times represent key stages for bioactive compound detection and the production of interesting molecules from the therapeutic and functional nutrition point of view (Figure 7). These findings altogether, namely, the integration of metataxonomic data with the enrichment of chemical families, and the identification of representative strains isolated from this biological system (water kefir fermentation) represent a straightforward approach to unlocking the potential of fermented foods with proven health benefits, taking a step further toward the design of targeted formulations based in microorganisms and their metabolites (probiotics or live microbial biotherapeutics, postbiotics). Strategies like the one we propose in this research can translate the benefits of fermented foods beyond their prescription to be consumed daily into active pharmaceutical products or active ingredients since most of them do not naturally reach the required concentrations to be considered therapeutic or efficient, but the rational design of a product based on their properties and their microbial strains can be harnessed using bioprocesses and biotechnology to achieve this status.

5 Conclusion

Functional foods are gaining interest due to the increase in non-communicable diseases like diabetes, obesity, and cardio-metabolic conditions. The results presented in this work correspond to a primary study developed on a water kefir product, a type of fermented food locally produced in Colombia, with modifications to its traditional way of preparation. The results of culture-dependent and molecular methods showed consistent findings in microbial richness and metabolite production increasing between 72–96 h, with dominant microorganisms identified as *L. hilgardii* (LAB), *A. orientalis* (AAB), and *P. membranifaciens* (yeast). Untargeted metabolomics using molecular networking allowed the generation of hypotheses on which small molecules are being produced during fermentation. Phenolic acids, flavonoids, and monoterpenoids are of great interest given the reported health benefits represented in the most abundant chemical families, which are produced to a greater extent by bacteria, namely, by *Acetobacter* species, specifically *L. hilgardii* for beneficial lipids and *A. tropicalis* for phenolic compounds and bile acid conjugates. In addition, we annotated and subsequently confirmed by reference spectra the presence of Isoschaftosides, a group of compounds that can be promising for the development of products derived from this microbial consortium. This study is the first of its kind in a fermented water kefir matrix locally produced in Colombia and the first to report the enrichment of chemical families, such as Isoschaftosides and other flavonoids produced by fermentation of a WK-derived microbial consortium, This study, despite being a first approximation of the chemical space of a Colombian-based WK fermentation and its correlation to the microbial taxa involved, provides a solid basis for future studies elucidating the mechanism of action of these functional fermented foods by directed metabolomics or analytical chemistry or on the isolation of chemical compounds of nutritional and therapeutic interest from fermented food matrices. We conclude that this study contributes to the existing knowledge on the dynamics of kefir fermentation and highlights the unique biological potential that each version can exhibit, as well as providing specific knowledge that could be easily applied to the rational development of novel probiotic and postbiotic ingredients for functional nutrition.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

Conceptualization: MA-E, LS-Z, DV-M, JC-A, and AD-R; Methodology: LS-Z, GF, and MA-E; Statistical Analysis: MA-E, LS-Z, and GF; Investigation: MA-E, LS-Z, GF, JC-A, and DV-M; Resources: AD-R, JB, LS-Z, DV-M, and JC-A; Writing—original draft: MA-E, LS-Z, and GF; Supervision: LS-Z, GF, AD-R, and JB; Project administration: LS-Z, DV-M, and JC-A. All authors contributed to the article and approved the submitted version.

Funding

This project was carried out within the framework of the Tax Benefit for investment in Science and Technology call for projects of 2019 of the Ministry of Science and Technology of Colombia (Minciencias), granted to the collaboration between ILUMA Alliance companies and EAFIT University. The project funded all materials and technological services needed to complete the experiments and analysis, and the scientific personnel for the development of the experiments, specifically MA-E, as well as paying the tuition fees for her MSc. In Biosciences. EAFIT University funded the time that scientific advisors of the project (LS-Z, DV-M, and JC-A) dedicated to the project, and the company ILUMA Alliance funded the time that company researchers dedicated to the development of this investigation (AD-R and JB). The project included other lines of research and its broad title was “Application of advanced molecular biotechnology techniques to the development of functional ingredients for human and animal nutrition”. It kicked off in November 2019 and was concluded in January 2023.

Acknowledgments

The authors gratefully acknowledge MinCiencias (Colombian Ministry of Science, Technology, and Innovation) and allied

company, Iluma Alliance, for their financial support and provision of the biological system. They also acknowledge the support of EAFIT University, which provided the laboratory infrastructure and the scientific team that contributed to the different stages of this project.

Conflict of interest

Authors AD-R and JB were employed by the company Iluma Alliance.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1223863/full#supplementary-material

References

Adusumilli, R., and Mallick, P. (2017). Data conversion with ProteoWizard MsConvert. *Methods Mol. Biol.* 1550, 339–368. doi:10.1007/978-1-4939-6747-6_23

Araújo, C. S., Macedo, L. L., and Teixeira, L. J. Q. (2023). Evaluation of mid-infrared spectra associated with chemometrics for the determination of physicochemical properties during fermentation of a new strawberry-based beverage with water kefir grains. *J. Food Compos. Analysis* 123, 105490. doi:10.1016/j.jfca.2023.105490

Azi, F., Li, Z., Xu, P., and Dong, M. (2022). Transcriptomic analysis reveals the antibacterial mechanism of phenolic compounds from kefir fermented soy whey against Escherichia coli 0157:H7 and Listeria monocytogenes. *Int. J. Food Microbiol.* 383, 109953. doi:10.1016/j.ijfoodmicro.2022.109953

Azi, F., Tu, C., Meng, L., Li, Z., Cherinet, M. T., Ahmadullah, Z., et al. (2021). Metabolite dynamics and phytochemistry of a soy whey-based beverage bio-transformed by water kefir consortium. *Food Chem.* 342, 128225. doi:10.1016/j.foodchem.2020.128225

Azizi, N. F., Kumar, M. R., Ong Abdullah, J., Khalid, M., and Omar, A. R. (2021). Kefir and its biological activities. *Foods* 10 (6), 1210. doi:10.3390/foods10061210

Beccati, A., Jan, G., Quast, C., Yilmaz, P., and Glöckner, F. O. (2017). SILVA tree viewer: interactive web browsing of the SILVA phylogenetic guide trees. *BMC Bioinforma.* 18 (1), 1–4. doi:10.1186/s12859-017-1841-3

Bengoa, A. A., Iraporda, C., Garrote, G. L., and Abraham, A. G. (2019). Kefir micro-organisms: their role in grain assembly and health properties of fermented milk. *J. Appl. Microbiol.* 126 (3), 686–700. doi:10.1111/jam.14107

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37 (8), 852–857. doi:10.1038/s41587-019-0209-9

Brazier-Hicks, M., Evans, K. M., Gershater, M. C., Puschmann, H., Steel, P. G., and Edwards, R. (2009). The C-glycosylation of flavonoids in cereals. *J. Biol. Chem.* 284 (27), 17926–17934. doi:10.1074/jbc.M109.009258

Bulat, T., and Ali, T. (2021). Influences of oxidation-reduction potential on kefir: microbial counts, organic acids, volatile compounds and sensory properties. *LWT - Food Sci. Technol.* 144, 111195. doi:10.1016/j.lwt.2021.111195

Cai, Y., Sounderrajan, A., and Serventi, L. (2020). Water kefir: A review of its microbiological profile, antioxidant potential and sensory quality. *Acta Sci. Nutr. Health* 4 (6), 10–17. doi:10.31080/asnh.2020.04.0706

Calatayud, M., Börner, R. A., Ghyselinck, J., Verstrepen, L., De Medts, J., Van den Abbeele, P., et al. (2021). Water kefir and derived pasteurized beverages modulate gut microbiota, intestinal permeability, and cytokine production *in vitro*. *Nutrients* 13 (11), 3897. doi:10.3390/nu13113897

Callahan, B. J., Paul, J., Holmes, S. P., and Johnson, A. J. A. (2016). DADA2: high resolution sample inference from illumina amplicon data. *Nat. Methods* 13 (7), 581–583. doi:10.1038/nmeth.3869

Çevik, T., Özdemir, N., Taş, T. K., and Kök Taş, T. (2019). The effect of different sugars on water kefir grains. *Turkish J. Agric. - Food Sci. Technol.* 7, 40–45. doi:10.24925/turjaf.v7isp1.40-45.2687

Chen, M. Y., Wu, H. T., Chen, F. F., Wang, Y. T., Chou, D. L., Wang, G. H., et al. (2022). Characterization of Tibetan kefir grain-fermented milk whey and its suppression of melanin synthesis. *J. Biosci. Bioeng.* 133 (6), 547–554. doi:10.1016/j.jbiosc.2022.02.006

Chen, Z., Tian, E., Ye, T., Yang, X., Xue, Y., Shen, Y., et al. (2021). Effect of lactic acid bacteria and yeasts on the structure and fermentation properties of Tibetan kefir grains. *Int. Dairy J.* 114, 104943. doi:10.1016/j.idairyj.2020.104943

Deseo, M. A., Elkins, A., Rochfort, S., and Kitchen, B. (2020). Antioxidant activity and polyphenol composition of sugarcane molasses extract. *Food Chem.* 314, 126180. doi:10.1016/j.foodchem.2020.126180

Diez-Ozaeta, I., and Astiazaran, O. J. (2022). Fermented foods: an update on evidence-based health benefits and future perspectives. *Food Res. Int.* 156, 111133. doi:10.1016/j.foodres.2022.111133

Dimidi, E., Cox, S. R., Rossi, M., and Whelan, K. (2019). Fermented foods: definitions and characteristics, impact on the gut microbiota and effects on gastrointestinal health and disease. *Nutr.* 11, 1806. doi:10.3390/nu11081806

Esatbeyoglu, T., Fischer, A., Legler, A. D. S., Oner, M. E., Wolken, H. F., Köpsel, M., et al. (2023). Physical, chemical, and sensory properties of water kefir produced from Aronia melanocarpa juice and pomace. *Food Chem.* X, 100683. doi:10.1016/j.fochx.2023.100683

Farag, M. A., Jomaa, S. A., Hesham, R., and El-Seedi, A. H. R. (2020). The many faces of kefir fermented dairy products: quality characteristics, flavour chemistry, nutritional value, health benefits, and safety. *Nutrients* 12, 346. doi:10.3390/nu12020346

Fiorda, S., Assumpção, F., Porto de Souza Vandenberghe, L., Ricardo Soccol, C., Pagnoncelli, M. G. B., Vandenberghe, L. P. D. S., et al. (2017). Microbiological, biochemical, and functional aspects of sugary kefir fermentation - a review. *Food Microbiol.* 66, 86–95. doi:10.1016/j.fm.2017.04.004

Gao, W., and Zhang, L. (2019). Comparative analysis of the microbial community composition between Tibetan kefir grains and milks. *Food Res. Int.* 116, 137–144. doi:10.1016/j.foodres.2018.11.056

Gil-De-La-Fuente, A., Godzien, J., Saugar, S., Garcia-Carmona, R., Badran, H., Wishart, D. S., et al. (2019). CEU mass mediator 3.0: A metabolite annotation tool. *J. Proteome Res.* 18 (2), 797–802. doi:10.1021/acs.jproteome.8b00720

Guan, S., Sun, L., Wang, X., Huang, X., and Luo, T. (2022). Isoschaftoside inhibits lipopolysaccharide-induced inflammation in microglia through regulation of HIF-1 α -mediated metabolic reprogramming. *Evidence-Based Complementary Altern. Med.* 2022, 5227335. doi:10.1155/2022/5227335

Gulitz, A., Stadie, J., Wenning, M., Ehrmann, M. A., and Vogel, R. F. (2011). The microbial diversity of water kefir. *Int. J. Food Microbiol.* 151 (3), 284–288. doi:10.1016/j.ijfoodmicro.2011.09.016

Guzel-Seydim, S., Zeynep, B., Gökırmaklı, Ç., and Greene, A. K. (2021). A comparison of milk kefir and water kefir: physical, chemical, microbiological and functional properties. *Trends Food Sci. Technol.* 113, 42–53. doi:10.1016/j.tifs.2021.04.041

Henry, K., Kohlwein, S. D., and Carman, G. M. (2012). Metabolism and regulation of glycerolipids in the yeast *Saccharomyces cerevisiae*. *Genetics* 190 (2), 317–349. doi:10.1534/genetics.111.130286

Hu, J. B., Gunathilake, S., Chen, Y. C., and PawelUrban, L. (2014). On the dynamics of kefir volatome. *RSC Adv.* 4 (55), 28865–28870. doi:10.1039/c4ra02990a

Jianzhong, Z., Liu, X., Jiang, H., and Dong, M. (2009). Analysis of the microflora in Tibetan kefir grains using denaturing gradient gel electrophoresis. *Food Microbiol.* 26 (8), 770–775. doi:10.1016/j.fm.2009.04.009

Jood, I., Hoff, J. W., and Setati, M. E. (2017). Evaluating fermentation characteristics of *Kazachstania* spp. and their potential influence on wine quality. *World J. Microbiol. Biotechnol.* 33 (7), 129. doi:10.1007/s11274-017-2299-1

Khodzhaieva, R. S., Gladkov, E. S., Alexander, K., and Roshal, A. D. (2021). Progress and achievements in glycosylation of flavonoids. *Front. Chem.* 9, 637994. doi:10.3389/fchem.2021.637994

Kim, S., Chen, J., Cheng, T., Gindulyte, A., Jia, H., He, S., et al. (2021). PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res.* 49 (D1), D1388–D1395. doi:10.1093/nar/gkaa971

Kõljalg, U., Nilsson, H. R., Schigel, D., Tedersoo, L., Larsson, K. H., May, T. W., et al. (2020). The taxon hypothesis paradigm—on the unambiguous detection and communication of taxa. *Microorganisms* 8 (12), 1910–1924. doi:10.3390/microorganisms8121910

Laureys, D., and de Vuyst, L. (2014). Microbial species diversity, community dynamics, and metabolite kinetics of water kefir fermentation. *Appl. Environ. Microbiol.* 80 (8), 2564–2572. doi:10.1128/AEM.03978-13

Laureys, D., and de Vuyst, L. (2016). The water kefir grain inoculum determines the characteristics of the resulting water kefir fermentation process. *J. Appl. Microbiol.* 122 (3), 719–732. doi:10.1111/jam.13370

Laureys, D., Leroy, F., Vandamme, P., and De Vuyst, L. (2022). Backslopping time, rinsing of the grains during backslopping, and incubation temperature influence the water kefir fermentation process. *Front. Microbiol.* 13, 871550. doi:10.3389/fmicb.2022.871550

Lebeaux, D., Chauhan, A., Létoffé, S., Fischer, F., de Reuse, H., Beloin, C., et al. (2014). PH-mediated potentiation of aminoglycosides kills bacterial persisters and eradicates *in vivo* biofilms. *J. Infect. Dis.* 210 (9), 1357–1366. doi:10.1093/infdis/jiu286

Lynch, K. M., Wilkinson, L. D., and Arendt, E. K. (2021). An update on water kefir: microbiology, composition and production. *Int. J. Food Microbiol.* 345, 109128. doi:10.1016/j.ijfoodmicro.2021.109128

Madigan, M. T., Martinko, J. M., Bender, K. S., Buckley, D. H., and Stahl, D. A. (2015). *13va edición. Biología de Los microorganismos*. Madrid, España: Pearson.

Marco, M. L., Hill, C., Hutkins, R., Slavin, J., Tancredi, D. J., Merenstein, D., et al. (2020). Should there Be a recommended daily intake of microbes? *J. Nutr.* 150 (12), 3061–3067. doi:10.1093/jn/nxaa323

Martínez-Torres, A., Gutiérrez-Ambrocio, S., Heredia-del-Orbe, P., Villa-Tanaca, L., and Hernández-Rodríguez, C. (2017). Inferring the role of microorganisms in water kefir fermentations. *Int. J. Food Sci. Technol.* 52 (2), 559–571. doi:10.1111/ijfs.13312

McCaughey, C. S., Trebino, M. A., Yildiz, F. H., and Sanchez, L. M. (2022). Utilizing imaging mass spectrometry to analyze microbial biofilm chemical responses to exogenous compounds. *Methods Enzym.* 665, 281–304. doi:10.1016/bs.mie.2021.11.014

Mordenti, A. L., Giaretta, E., Campidonico, L., Parazza, P., and Formigoni, A. (2021). A review regarding the use of molasses in animal nutrition. *Animals* 11, 115. doi:10.3390/ani11010115

Moretti, A. F., Moure, M. C., Quiñoy, F., Esposito, F., Simonelli, N., Medrano, M., et al. (2022). Water kefir, a fermented beverage containing probiotic microorganisms: from ancient and artisanal manufacture to industrialized and regulated commercialization. *Future Foods* 5, 100123. doi:10.1016/j.fufo.2022.100123

Nejati, F., Junne, S., Kurreck, J., and Neubauer, P. (2020). Quantification of major bacteria and yeast species in kefir consortia by multiplex TaqMan QPCR. *Front. Microbiol.* 11 (1291), 1291. doi:10.3389/fmicb.2020.01291

Olivo, D., Galván, M., López-Rodríguez, G., Suárez-Diéguez, T., González-Unzaga, M., Anaya-Cisneros, L., et al. (2017). Actividad biológica y potencial terapéutico de Los probióticos y el kefiran del grano de Kéfir. Avaliable At: www.reibci.org.

Păcularu-Burada, B., Aida Vasile, M., Elena Bahrim, G., and Bahrim, G. E. (2022). Novel insights into different kefir grains usefulness as valuable multiple starter cultures to achieve bioactive gluten-free sourdoughs. *LWT* 165, 113670. doi:10.1016/j.lwt.2022.113670

Padmaperuma, G., Butler, T. O., Faqih, A. B., Almalki, W. J., and Vaidyanathan, S. (2019). Microbial consortia: concept and application in fruit crop management. *Fruit. Crops* 2019, 353–366. doi:10.1016/B978-0-12-818732-6.00025-3

Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N., et al. (2022). Using MetaboAnalyst 5.0 for LC-HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. *Nat. Protoc.* 17, 1735–1761. doi:10.1038/s41596-022-00710-w

Parks, D. H., Rinke, C., Mussig, A. J., and Hugenholtz, P. (2020). A complete domain-to-species taxonomy for bacteria and archaea. *Nat. Biotechnol.* 38 (9), 1079–1086. doi:10.1038/s41587-020-0501-8

Patel, S. H., Tan, J. P., Börner, R. A., Zhang, S. J., Priour, S., Lima, A., et al. (2022). A temporal view of the water kefir microbiota and flavour attributes. *Innovative Food Sci. Emerg. Technol.* 80, 103084. doi:10.1016/j.ifset.2022.103084

Pence, H. E., and Williams., A. (2010). Chemspider: an online chemical information resource. *J. Chem. Educ.* 87, 1123–1124. doi:10.1021/ed100697w

Pendón, M. D., Iraporda, C., Medrano, M., Garrote, G. L., and Abraham, A. G. (2021). Water kefir: factors affecting grain growth and health-promoting properties of the fermented beverage. *J. Appl. Microbiol.* 133 (1), 162–180. doi:10.1111/jam.15385

Pihurov, M., Păcularu-Burada, B., Cotârleț, M., Grigore-Gurgu, L., Borda, D., Stănciuc, N., et al. (2023). Kombucha and water kefir grains microbiomes' symbiotic contribution to postbiotics enhancement. *Foods* 12 (13), 2581. doi:10.3390/foods12132581

Plessas, S., Nouska, C., Mantzourani, I., Kourkoutas, Y., Alexopoulos, A., and Bezirtzoglou, E. (2017). Microbiological exploration of different types of kefir grains. *Fermentation* 3, 1. doi:10.3390/fermentation3010001

QIAGEN (2020). *DNeasy UltraClean microbial Kit handbook*. United States: QIAGEN. Available at: https://www.qiagen.com/us/resources/resourcedetail?id=a733f3a9-5a8b-4003-b28e-55bb050bad1e&lang=en.

Rodrigues, K. L., de Souza Ferreira, C., Rodrigues, M. R., Ferreira, C. D. S., Moraes, G. d. O. I., Coimbra, R. S., et al. (2016). A novel beer fermented by kefir enhances anti-inflammatory and anti-ulcerogenic activities found isolated in its constituents. *J. Funct. Foods* 21, 58–69. doi:10.1016/j.jff.2015.11.035

Rodríguez, M. A., Fernández, L. A., Díaz, M. L., Pérez, M., Corona, M., and Reynaldi, F. J. (2022). Microbiological and chemical characterization of water kefir: an innovative source of potential probiotics for bee nutrition. *Rev. Argent. Microbiol.* 55, 176–180. doi:10.1016/j.ram.2022.09.003

RStudio Team (2020). *RStudio*. Boston, MA: Integrated Development Environment for R. Avaliable At: http://www.rstudio.com/.

Şafak, H., Gün, İ., Tudor Kalit, M., and Kalit, S. (2023). Physico-chemical, microbiological and sensory properties of water kefir drinks produced from demineralized whey and dimrit and shiraz grape varieties. *Foods* 12 (9), 1851. doi:10.3390/foods12091851

Santen, J. A. V., Poynton, E. F., Iskakova, D., Mcmann, E., Alsup, T. A., Clark, T. N., et al. (2022). The natural products Atlas 2.0: A database of microbially-derived natural products. *Nucleic Acids Res.* 50 (D1), D1317–D1323. doi:10.1093/nar/gkab941

Satir, G. (2022). The effects of fermentation with water kefir grains on two varieties of tigernut (Cyperus esculentus L.) milk. *LWT - Food Sci. Technol.* 171, 114164. doi:10.1016/j.lwt.2022.114164

Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., and McLean, J. A. (2016). Untargeted metabolomics strategies—challenges and emerging directions. *J.* Am. Soc. Mass Spectrom. 27 (12), 1897–1905. doi:10.1007/s13361-016-1469-y

Shannon, P., Markiel, A., Owen, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi:10.1101/gr.1239303

Sharma, I., and Yaiphathoi, S. (2020). Role of microbial communities in traditionally fermented foods and beverages in North east India. *Recent Adv. Microb. Divers.* 2020, 445–470. doi:10.1016/b978-0-12-821265-3.00019-0

SIB Colombia (2023). Gbif. Avaliable At: https://www.gbif.org/es/dataset/a77265a0-2ae0-4c84-84dd-fadb11e152be#description (Accessed on August 21, 2023).

Spizzirri, U. G., Loizzo, M. R., Aiello, F., Prencipe, S. A., and Restuccia, D. (2023). Non-dairy kefir beverages: formulation, composition, and main features. *J. Food Compos. Analysis* 117, 105130. doi:10.1016/j.jfca.2023.105130

Swift Biosciences (2018). Swift amplicon 16S+ITS panel. Avaliable At: https://swiftbiosci.com/wp-content/uploads/2019/02/18-2328_Swift-Amplicon-16SITS-Panel_Protocol.pdf.

Taheur, F. B., Fedhila, K., Chaieb, K., Kouidhi, B., Bakhrouf, A., and Abrunhosa, L. (2017). Adsorption of aflatoxin B1, zearalenone and ochratoxin A by microorganisms isolated from kefir grains. *Int. J. Food Microbiol.* 251, 1–7. doi:10.1016/j.ijfoodmicro.2017.03.021

Tan, L. L., Tan, C. H., Ng, N. K. J., Tan, Y. H., Conway, P. L., and Loo, S. C. J. (2022). Potential probiotic strains from milk and water kefir grains in Singapore—use for defense against enteric bacterial pathogens. *Front. Microbiol.* 13, 857720. doi:10.3389/fmicb.2022.857720

Tremmel, M., Kiermaier, J., and Heilmann, J. (2021). *In vitro* metabolism of six C-glycosidic flavonoids from passiflora incarnata l. *Int. J. Mol. Sci.* 22 (12), 6566. doi:10.3390/ijms22126566

Vanegas, S., Garcia, K., Eichenberger, M., Fischer, D., Mortensen, U. H., and Naesby, M. (2018). Indirect and direct routes to C-glycosylated flavones in Saccharomyces cerevisiae. *Microb. Cell. Factories* 17 (107), 107. doi:10.1186/s12934-018-0952-5

Verce, M., de Vuyst, L., and Weckx, S. (2019). Shotgun metagenomics of a water kefir fermentation ecosystem reveals a novel oenococcus species. *Front. Microbiol.* 10, 479–516. doi:10.3389/fmicb.2019.00479

Villarreal-Morales, R., Sandra, L., Montañez-Saenz, J. C., Aguilar-González, C. N., and Rodriguez-Herrera, R. (2018). Metagenomics of traditional beverages. *Adv. Biotechnol. Food Industry* 14, 301–326. doi:10.1016/B978-0-12-811443-8.00011-6

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Yao, P., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* 34 (8), 828–837. doi:10.1038/nbt.3597

Weckx, S., Kerrebroeck, S. V., and de Vuyst, L. (2019). Omics approaches to understand sourdough fermentation processes. *Int. J. Food Microbiol.* 302, 90–102. doi:10.1016/j.ijfoodmicro.2018.05.029

Wei, T., and Simko, V. (2021). R package 'Corrplot': visualization of a correlation matrix (version 0.92). Avaliable At: https://github.com/taiyun/corrplot.

Wuyts, S., Fl van den Broek, M., and Lebeer, S. (2020). Applications of plant-based fermented foods and their microbes. *Curr. Opin. Biotechnol.* 61, 45–52. doi:10.1016/j.copbio.2019.09.023

Yerlikaya, O., Akan, E., and Kinik, Ö. (2022). The metagenomic composition of water kefir microbiota. *Int. J. Gastron. Food Sci.* 30, 100621. doi:10.1016/j.ijgfs.2022.100621

Yerlikaya, O. (2019). Probiotic potential and biochemical and technological properties of Lactococcus lactis Ssp. lactis strains isolated from raw milk and kefir grains. *J. Dairy Sci.* 102 (1), 124–134. doi:10.3168/jds.2018-14983

Zanirati, M., Ferreira, D., Abatemarco, M., Robert Nicoli, J., Nunes, Á. C., and Neumann, E. (2015). Selection of lactic acid bacteria from Brazilian kefir grains for potential use as starter or probiotic cultures. *Anaerobe* 32, 70–76. doi:10.1016/j.anaerobe.2014.12.007

Zannini, E., Lynch, K. M., Nyhan, L., Elke, K., O' Riordan, P., Luk, D., et al. (2023). Influence of substrate on the fermentation characteristics and culture-dependent microbial composition of water kefir. *Fermentation* 9, 28. doi:10.3390/fermentation9010028

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 10 November 2023
DOI 10.3389/fmolb.2023.1235160

OPEN ACCESS

EDITED BY
Guillermo Moyna,
Universidad de la República, Uruguay

REVIEWED BY
Jinping Gu,
Zhejiang University of Technology, China
Florence Mehl,
University of Zurich, Switzerland
Bonell Patino-Escobar,
University of California, San Francisco, United States

*CORRESPONDENCE
Mónica P. Cala,
mp.cala10@uniandes.edu.co
Susana Fiorentino,
susana.fiorentino@javeriana.edu.co

†These authors have contributed equally to this work and share first authorship

RECEIVED 05 June 2023
ACCEPTED 27 October 2023
PUBLISHED 10 November 2023

CITATION
Arévalo C, Rojas L, Santamaria M, Molina L, Arbeláez L, Sánchez P, Ballesteros-Ramírez R, Arevalo-Zambrano M, Quijano S, Cala MP and Fiorentino S (2023), Untargeted metabolomic and lipidomic analyses reveal lipid dysregulation in the plasma of acute leukemia patients.
Front. Mol. Biosci. 10:1235160.
doi: 10.3389/fmolb.2023.1235160

Untargeted metabolomic and lipidomic analyses reveal lipid dysregulation in the plasma of acute leukemia patients

Cindy Arévalo[1†], Laura Rojas[1†], Mary Santamaria[2], Luisana Molina[3], Lina Arbeláez[3], Paula Sánchez[3], Ricardo Ballesteros-Ramírez[1], Monica Arevalo-Zambrano[3], Sandra Quijano[1,3], Mónica P. Cala[2*] and Susana Fiorentino[1*]

[1]Grupo de Inmunobiología y Biología Celular, Facultad de Ciencias, Pontificia Universidad Javeriana, Bogotá, Colombia, [2]MetCore—Metabolomics Core Facility, Vice-Presidency for Research, Universidad de Los Andes, Bogotá, Colombia, [3]Hospital Universitario San Ignacio, Bogotá, Colombia

Acute leukemias (AL) are aggressive neoplasms with high mortality rates. Metabolomics and oxidative status have emerged as important tools to identify new biomarkers with clinical utility. To identify the metabolic differences between healthy individuals (HI) and patients with AL, a multiplatform untargeted metabolomic and lipidomic approach was conducted using liquid and gas chromatography coupled with quadrupole-time-of-flight mass spectrometry (LC-QTOF-MS or GC-QTOF-MS). Additionally, the total antioxidant capacity (TAC) was measured. A total of 20 peripheral blood plasma samples were obtained from patients with AL and 18 samples from HI. Our analysis revealed 135 differentially altered metabolites in the patients belonging to 12 chemical classes; likewise, the metabolic pathways of glycerolipids and sphingolipids were the most affected in the patients. A decrease in the TAC of the patients with respect to the HI was evident. This study conducted with a cohort of Colombian patients is consistent with observations from other research studies that suggest dysregulation of lipid compounds. Furthermore, metabolic differences between patients and HI appear to be independent of lifestyle, race, or geographic location, providing valuable information for future advancements in understanding the disease and developing more global therapies.

KEYWORDS

untargeted metabolomics, acute leukemia, lipids, antioxidant capacity, acute myeloid leukemia, acute lymphoid leukemia

Introduction

AL are a heterogeneous group of hematological malignancies that involve blocked hematopoietic progenitors and accumulate in the early phases of cell differentiation, leading to marrow failure, and can be classified into two large groups: acute myeloid leukemia (AML) and acute lymphoid leukemia (ALL) (B-ALL or T-ALL) (Arber et al., 2016). In the process of clonal evolution, leukemic cells accumulate mutations that can lead to aberrant metabolic programs that are necessary to meet bioenergetic and biosynthetic demands and maintain a redox balance for tumor survival and proliferation (DeBerardinis and Chandel, 2016; Romer-Seibert and Meyer, 2021). In the Colombian population, remission rates are low

compared to those in developed countries, so exploring the metabolic profiles in these cases is intriguing (Combariza et al., 2007; Ballesteros-Ramírez et al., 2020; Sossa et al., 2021). Increased glucose consumption by leukemic progenitor cells is beneficial for producing metabolic intermediates involved in other metabolic pathways and for ATP generation. On the other hand, leukemia-initiating cells (LICs) in AML are characterized by their dependence on oxidative phosphorylation and branched-chain amino acids, low glucose consumption, low production of reactive oxygen species (ROS), and high levels of glutathione (GSH) (Jones et al., 2018). While in B-ALL, LICs that carry alterations in the PAX5 (Paired Box 5) or IKZF (IKAROS Family Zinc Finger 1) genes are related to unlimited glucose consumption (Boag et al., 2006).

The evaluation of specialized tumor metabolism can be measured at the systemic level by evaluating the TAC and using metabolomic platforms, such as mass spectrometry (MS) and nuclear magnetic resonance (NMR), to identify unique metabolic fingerprints, which may allow the identification of biomarkers and therapeutic targets (Muthu and Nordström, 2019; Schmidt et al., 2021). Some authors have demonstrated its usefulness in blood plasma for diagnosis (Morad et al., 2022), response to treatment and follow-up of patients with AL (Naz et al., 2013; Grønningsæter et al., 2019; Kim et al., 2021), and to assess cellular response to antileukemic agents (Dhakshinamoorthy et al., 2015). In this sense, metabolic differences between HI and adult patients with AL have been reported using ^{1}H NMR (Musharraf et al., 2016; Yang et al., 2021). Other authors have managed to define groups of metabolites (mainly energy) associated with prognosis in patients with AML using gas chromatography-time-of-flight mass spectrometry (GC-TOF-MS) or liquid chromatography-mass spectrometry (LC-MS) (Chen et al., 2014; Dong et al., 2019). However, there are few publications describing differential metabolites between ALL and AML by LC-MS or ^{1}H NMR (Musharraf et al., 2017; Hao et al., 2022). In addition, few studies have focused on the oxidative stress profile in AL. Particularly, Naz et al. (2013) showed that patients in complete remission decreased their total antioxidant status (Naz et al., 2013); however, in pediatric patients with ALL, the differences in the levels of some antioxidants are not clear (Olaniyi et al., 2011).

Tumor metabolism is influenced by intrinsic factors, such as genetic alterations or lineage/tissue of origin, and extrinsic factors, such as access to nutrients and oxygen, interaction with cells in the microenvironment, and exposure to radiation or chemotherapy (Vander Heiden and DeBerardinis, 2017). In other words, population characteristics such as race, genetics, food culture, or healthy habits can also influence tumor metabolism and have great relevance to the risk of development or progression of the disease, the risk of recurrence, the risk of disease, and mortality for some types of cancer (Faulds and Dahlman-Wright, 2012; Islami et al., 2018; Peng et al., 2022).

Based on the aforesaid considerations, the current study examined the metabolomic profiles of cohorts comprising Colombian patients diagnosed with AL, using an untargeted metabolomic and lipidomic approach by LC-QTOF-MS and GC-QTOF-MS. Furthermore, the study established the total antioxidant capacities within these cohorts. To our knowledge, this is the first report to investigate the metabolic and lipid alterations associated with AL, specifically in the Colombian population. These findings hold significant potential for the development of future diagnostic and prognostic biomarkers for this population.

Materials and methods

Study participants

Between 2019 and 2020, twenty patients with AL who attended the Hospital Universitario San Ignacio (Bogotá D.C., Colombia) were linked to this study. They were patients older than 18 years, who were diagnosed for the first time and had not received previous therapy. The study was approved by the Ethics Committee of the Hospital Universitario San Ignacio and the Centro Javeriano de Oncología (Bogotá D.C., Colombia). Following the Declaration of Helsinki, written informed consent was obtained from all participants prior to clinical data collection and sample collection. The diagnosis was made according to the World Health Organization classification of tumors of hematopoietic and lymphoid tissues 2017 (Arber et al., 2016).

Sample collection

Peripheral blood samples were obtained from twenty patients with a *de novo* diagnosis of acute leukemia before starting chemotherapy treatment. Samples were collected with a minimum fast of 8 h in K_2EDTA tubes, and peripheral blood was centrifuged exactly 4 h after collection, at 3,500 rpm at 4°C for 10 min. The plasma obtained was aliquoted and stored at −80°C until processing. As a control group, 18 plasmas were collected from HI matched by age and sex under the same conditions. According to the NCCN Clinical Practice Guidelines in Oncology ® (Chang et al., 2021; Pollyea et al., 2021) and given the number of samples collected, the response to treatment at the end of induction was divided into two groups of patients: Complete remission (CR) [including CR with negative EMR (minimal residual disease)] and non-responders (NR) (including CR with a partial hematological response (CRp), CR with an incomplete hematological response (Cri) and CR with positive EMR or unknown), patients with premature death and NR).

Untargeted metabolomic and lipidomic analysis

Metabolomic analysis by LC-QTOF-MS and GC-QTOF-MS

For metabolomic analysis by reverse phase liquid chromatography coupled to mass spectrometry with a time-of-flight analyzer (RP-LC-QTOF-MS), samples were extracted using 40 µL of plasma mixed with 120 µL of cold methanol and ethanol (1:1, v/v) (−20°C) and samples were vortex-mixed for 5 min. Samples were incubated for 20 min at −20°C to precipitate proteins. Then, the samples were centrifuged for 10 min (16,000 g, 4°C). Metabolomic analysis was performed using the UHPLC system (Agilent 1260 Infinity LC System) coupled with the Q-TOF LC/MS system (Agilent Technologies, Waldronn, Germany) equipped with an electrospray ionization (ESI) source 2 µL of the extracted

sample was injected into the InfinityLab Poroshell 120 EC-C18 (100 mm × 3.0 mm, 2.7 μm) column at 30°C using 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in acetonitrile (B) as a mobile phase with a flow rate of 0.3 mL/min. Gradient elution started with 25% B and increased to 95% within 35 min. Then, the gradient returned to initial conditions at 35.1 min and held there for 8 min to allow column re-equilibrium. For constant mass correction, two reference masses were used and continuously infused into the system: *m/z* 121.0509 $(C_5H_4N_4 + H)^+$ and m/z 922.0098 $(C_{18}H_{18}O_6N_3P_3F_{24} + H)^+$ for positive ionization mode (ESI+) and *m/z* 112.9856 $(C_2O_2F_3 - NH_4)^-$ and *m/z* 1033.9881 $(C_{18}H_{18}O_6N_3P_3F_{24} + FA\text{-}H)^-$ for negative ionization mode (ESI -). The system was operated in full scan mode from 100 to 1,100 *m/z*; the capillary voltage was set to 3000, the drying gas flow rate was 12 L/min at 290°C, the gas nebulizer 52 psi, fragmentor voltage was 175 V and the skimmer 65 V and octopole radio frequency voltage (OCT RF Vpp) 750 V for both, positive and negative ionization modes. Data were collected in centroid mode at a scan rate of 1.02 spectrum per second.

For metabolomic analysis by gas chromatograph coupled to mass spectrometry with a time-of-flight analyzer (GC-QTOF-MS), samples were extracted using 140 μL of plasma mixed with 420 μL of cold methanol (−20°C) and vortex-mixed for 5 min. Samples were incubated for 20 min at −20°C to precipitate proteins. Then, the samples were centrifuged for 10 min (16,000 g, 4°C). An aliquot of 100 μL was transferred into glass inserts and evaporated to dryness in a speed vacuum concentrator (Thermo Scientific). The dry residue was dissolved in 10 μL of methoxyamine hydrochloride in pyridine (15 mg/mL) and vortex-mixed for 5 min. The samples were incubated for 16 h at room temperature in the dark. The silylation process was followed by adding 10 μL of bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS). After vortex-mixing (5 min) and incubation for 1 h at 70°C, the samples were diluted with 50 μL of internal standard (methyl stearate in heptane C18:0, 10 ppm). GC-QTOF-MS experiments were performed on an Agilent Technologies 7890B GC system coupled to 7250 QTOF mass spectrometer system (Agilent Technologies). Derivatized samples were injected (1 μL) with a split ratio of 30: 1 onto an HP-5MS capillary column (30 m × 0.25 mm; 0.25 μm) (Agilent Technologies) at a constant gas flow (helium) of 0.7 mL/min. The injector temperature was 250°C. The temperature gradient was kept at 60°C for 1 min and then programmed to 320°C at 10°C/min. Mass spectra were recorded at 70 eV in full scan mode with *m/z* values ranging from 50 to 600. The transfer line, filament source, and quadrupole temperature were fixed at 280°C, 230°C, and 150°C, respectively.

Lipidomic analysis by LC-QTOF-MS

For lipids extraction, 100 μL of plasma was extracted with 350 μL of cold methanol and 350 μL of MTBE and vortex mixed for 5 min. Then, the samples were centrifuged at 13,000 g for 10 min at room temperature. Lipidomic analysis was performed using the same RP-LC–QTOF–MS system employed for metabolomics analysis. A 1 μL of the extracted sample was injected into the InfinityLab Poroshell 120 EC-C8 (100 mm × 2.1 mm, 2.7 μm) column at 60°C using 5 mM ammonium formate in Mili-Q water) (A) and 5 mM ammonium formate in isopropanol: methanol (15:85) (B) as a mobile phase with a 0.4 mL/min flow rate. Gradient elution started with 75% B, then increased to 96% within 23 min, and kept there for 13 min, then increased to 100% and kept constant for 4 min. Then, the gradient returned to initial conditions at 42 min and held there for 11 min to allow column re-equilibrium. The same reference masses were used throughout the analysis as described in metabolomics analysis by LC-QTOF-MS for positive and negative ionization modes. The system was operated in full scan mode from 100 to 1,800 *m/z*; a capillary voltage was set to 3,000 V, the drying gas flow rate was 12 L/min at 290°C; and the gas nebulizer 45 psi, fragmentor voltage 175 V, the skimmer 65 V and octopole radio frequency voltage (OCT RF Vpp) 750 V. Data were collected in centroid mode at a scan rate of 1.02 spectrum per second.

Quality assurance (QA) and quality control (QC) procedures

Quality assurance and quality control procedures were implemented according to published guidelines to reduce unwanted variation (26) (Kirwan et al., 2022). Pure solvents and extraction blanks were evaluated at the beginning of each sequence to ensure the cleanliness of equipment and materials used in sample preparation. To equilibrate the chromatographic system, pooled samples (QC) were injected, which were prepared by mixing equal volumes of each plasma sample using the same procedure for both metabolomic (LC and GC) and lipidomic analysis. To monitor the system's stability, these QC samples were injected every ten samples. Additionally, biological samples were randomized within the sequence to reduce the possibility of bias (Sumner et al., 2007).

Data treatment

All raw LC-QTOF-MS datasets were processed using *Agilent MassHunter Profinder B.10.0* Software for deconvolution, alignment, and integration, using algorithms such as molecular feature extraction and recursive feature extraction; then, the raw data were inspected manually to remove background noise and unrelated ions. For GC-QTOF-MS, samples were normalized by internal standards prior to the statistical analysis. Finally, for all platforms, the data was filtered by presence and reproducibility, and the coefficient of variation (CV) in the QC lower than 20% to LC (or 30% to GC) was used for statistical analysis.

Statistical analysis

For both LC-MS and GC-MS data, the identification of the molecular characteristics with statistical differences between HI and AL patients were carried out using univariate and multivariate statistical analysis. First, the *p*-value was determined by Mann-Whitney U test (nonparametric tests) with a Benjamini-Hochberg False Discovery Rate *post hoc* correction (FDR) using MatLab (R2019b, Mathworks, Inc., Natick), while for the multivariate analysis, an unsupervised principal component analysis (PCA) and orthogonal partial least squares regression (OPLS-DA) was applied using SIMCA-P + 16.0 software. The statistically significant variables were selected based on *p*-value with FDR $p < 0.05$ and variance important in projection (VIP) > 1 with Jack-knife confident interval (JK).

Metabolites identification

The identification of metabolites was carried out based on a 4-level confidence system for high-resolution mass spectrometry analysis following the parameters (Schymanski et al., 2014). Metabolites by LC-MS were annotated using various online database (http://hmdb.ca), (http://genome.jp/keg), (https://massbank.eu/MassBank/), (http://lipidmaps.org) and (http://metlin.scripps.edu) utilized for this purpose CEU Mass Mediator tool (http://ceumass.eps.uspceu.es/). The metabolite's identity was confirmed by iterative MS/MS data with Agilent Lipid Annotator software, MS-DIAL 4.80 (http://prime.psc.riken.jp/compms/msdial/main.html), and CFM-ID 4.0 (https://cfmid.wishartlab.com/) for *in silico* mass spectral fragmentation. For GC-QTOF-MS chromatograms were deconvoluted and compared with Fiehn GC-MS Metabolomics RTL Library (Kind et al., 2009).

Pathway analysis

Metabolic pathway analysis was performed with the MetaboAnalyst 5.0 tool (http://www.MetaboAnalyst.ca/), integrating enrichment and topology pathway approaches. A list of identified significant metabolite compound names was loaded and processed using the "*Homo sapiens*" library. The KEGG pathway information was obtained in October 2019, and the specific pathway analysis parameters were the visualization method by scatter plot (testing significant features), enrichment method (hypergeometric test), topology analysis (relative-betweenness centrality), and selecting a pathway library by *H. sapiens*.

Determination of TAC in plasma

The antioxidant capacity was evaluated using the e-BQClab device (Bioquochem, Asturias, Spain) that measures the redox potential, which is expressed in micro coulombs (μC). The results in μC were transformed to Trolox Equivalent Antioxidant Capacity Units (TEAC). The e-BQClab device using electrochemistry can distinguish between fast and slow antioxidants: the Q1 value refers to the antioxidant capacity of the compounds with the highest free radical scavenging rate (examples, uric acid, GSH, vitamin E), while the Q2 value refers to the antioxidant capacity of the compounds with the lowest rate of free radical uptake (examples, polyphenols, resveratrol). The QT value is the sum of both. The measurement was performed in duplicate using 50 μL of all collected plasma samples.

Results

Characteristics of AL patients and HI

The mean age at diagnosis of the patients evaluated was 45.8 years (range 21–76), and 55% were female patients. We collected 9 patients with B-ALL, 9 patients with AML, 1 patient with acute promyelocytic leukemia, and 1 patient with mixed-phenotype acute leukemia (B-lymphoid and myeloid differentiation). The karyotype was normal in 40% (8) of the patients, abnormal in 50% (10), and there was no growth in 10% (2). According to the risk categories, 80% of the patients were classified as high-risk and the remaining 20% as intermediate-risk. The mean white blood cell count was 76,786 cells/μL (interval 800–403,000), hemoglobin was 9.0 g/dL (3.2–12), platelets were 68,130 cells/μL (interval 8,300–228,000), and the mean number of tumor cells over the total nucleated cells in bone marrow was 68.8% (interval 20–93.9). The first phase of chemotherapy (induction therapy) was based on the PETHEMA protocols for all patients, 7x3 or 5-Azacytidine for AML patients, and AIDA-PETHEMA for acute promyelocytic leukemia patients. Of the total number of patients at the end of induction, 10% (2) achieved CR, 70% (14) were NR, and 20% (4) could not be evaluated. In total, 18 samples were collected from HI; the group had a mean age of 31.9 years (range 19–61), and 55% were women. The clinical and demographic data of the patients are summarized in Table 1, along with the HI data. Clinical data related to treatment and response are detailed in Supplementary Table S1.

Alterations in lipid metabolism at the plasma level, differentiate HI from AL patients

Multiplatform metabolomic and lipidomic analyses of AL and HI plasma samples were conducted using different approaches aimed at detecting the largest possible number of metabolites. The performance of the different analytical platforms was evaluated by clustering the quality control (QC) samples using PCA models. In these models, a clear grouping of the QC samples belonging to each analytical platform was observed, indicating reliable, consistent performance and the conservation of biological variation across the platforms used (Supplementary Figure S1). Following the supervised OPLS-DA analyses, a discrimination between the HI group (green dots) and the AL patients (red dots) was observed for each platform, as depicted in Figure 1. This suggests distinct metabolomic profiles associated with the development of leukemia. The results indicated acceptable values ranging between 0.972 and 0.925 for R^2 and 0.670 and 0.852 for Q^2 in the cross-validation test in the metabolomic and lipidomic analyses on all analytical platforms used (Wheelock and Wheelock, 2013). On the other hand, volcano plots were generated to show the metabolites that were significant [$p < 0.05$, Log2(FC) < 1.3] by univariate analysis (Supplementary Figure S2).

A total of 328 metabolites differentially expressed between HI and AL were determined using a combination of MVA (VIP > 1 with JK), UVA ($p < 0.05$) applied on adjusted p-values, and fold change Log2(FC) > 1 and < 1). The detected metabolites during data processing across the different analytical techniques used are presented in Supplementary Table S2, and A typical metabolic fingerprint from each platform is presented in Supplementary Figures S3–S5. The compounds altered between AL patients and HI showed 61.28% (201 metabolites) downregulated metabolites and 38.72% (127 metabolites) upregulated metabolites. In respect of downregulated metabolites, we found glycerophospholipids (47.25%) to be the most representative, sphingolipids (23.37%), and a lower percentage (<7%) of sterol lipids, steroids, amino acids, bile acids, fatty acyls, organic acids, organooxygen compounds, and carnitines (Figure 2A). In the upregulated group, we observed that 47.25% of those metabolites were glycerophospholipids, 10.26% sphingolipids, 49.57% glycerolipids,

10.3389/fmolb.2023.1235160

TABLE 1 Summary description of the clinical characteristic of AL patients and date of healthy individuals.

Characteristic	*n* (%)
Sex (female)	11 (55)
Age (years), median (interval)	45.8 (21–76)
Immunophenotype	
B-ALL	9 (45)
AML	9 (45)
B/M AL	1 (5)
AML M3	1 (5)
Karyotype	
Normal	8 (40)
Abnormal	10 (50)
No growth	2 (10)
Risk	
Intermedium	4 (20)
High	16 (80)
Hematological parameters	
WBC count (μL), median (interval)	76,786 (800–403,000)
Hb (g/dL), median (interval)	9.0 (3.2–12)
Platelet count (μL), median (interval)	68,130 (8,300–228,000)
Tumor cells in bone marrow (%), median (interval)	68,8 (20–93,9)
Healthy Individuals	
Sex (female)	10 (55)
Age (years), median (interval)	31.9 (19–61)

B-ALL, B-acute lymphoid leukemia; AML, Acute myeloid leukemia; B/M AL, mixed-phenotype acute leukemia patient (B-lymphoid and myeloid differentiation); AML M3, Promyelocytic—acute myeloid leukemia; Hb, Hemoglobin; WBC, White blood cells.

4.27% fatty acyls, and less than 8% corresponded to sterol lipids, amino acids, bile acids, organic acids, imidazopyrimidines, and steroids (Figure 2B).

For greater reliability, the Log2(FC) was adjusted to ≥ 1.5 and ≤ 0.5, and 135 metabolites. The set of altered metabolites (FC > 1.5 or < 0.5) between the two groups was analyzed using heatmaps, which enable the visualization of patterns of metabolite changes among the groups. Therefore, blue colors indicate decreased metabolite levels, while red colors indicate increased metabolite levels in AL patients (Figure 3). The clustering analysis in the heatmap reveals a clear grouping of samples from AL patients (green) and HI patients (red), indicating similarity in the metabolomic profiles among individuals within each group.

The most significant variations between AL patients and HI are observed in the glycerophospholipids group, which includes metabolites such as lysophosphatidylcholines, phosphatidylcholines (PCs), phosphatidylserine, phosphatidylethanolamine, lysophosphatidylethanolamines, and some sphingolipids like sphingomyelin. These metabolites were predominantly found to be downregulated in AL patients (represented by blue colors in the heatmap). In contrast, the glycerolipids group, including triacylglycerols (TG), diacylglycerols (DG), some organic acids like pyruvic acid and hydroxyglutaric acid, and amino acids such as 4-acetamido-amino butanoic acid, glutamic acid, amino butanoic acid, and leucylproline, showed trends towards upregulation in AL patients (represented by red colors in Figure 3). These results suggest significant alterations in biochemical metabolites, particularly lipid compounds, in the plasma of AL patients compared to the control group.

Metabolic pathways associated with sphingolipid and glycerophospholipids are altered in AL patients

For a better understanding of the metabolic dysregulation between the two groups, differential metabolites were imported into MetaboAnalyst 5.0 to perform the Metabolomic Pathway Analysis. The *x*-axis represents the pathway impact value computed from pathway topological analysis, and the *y*-axis is the-log of the *p*-value obtained from pathway enrichment analysis. The pathways that were most significantly changed are characterized by both a high-log(*p*) *value* and a high impact value (top right region).

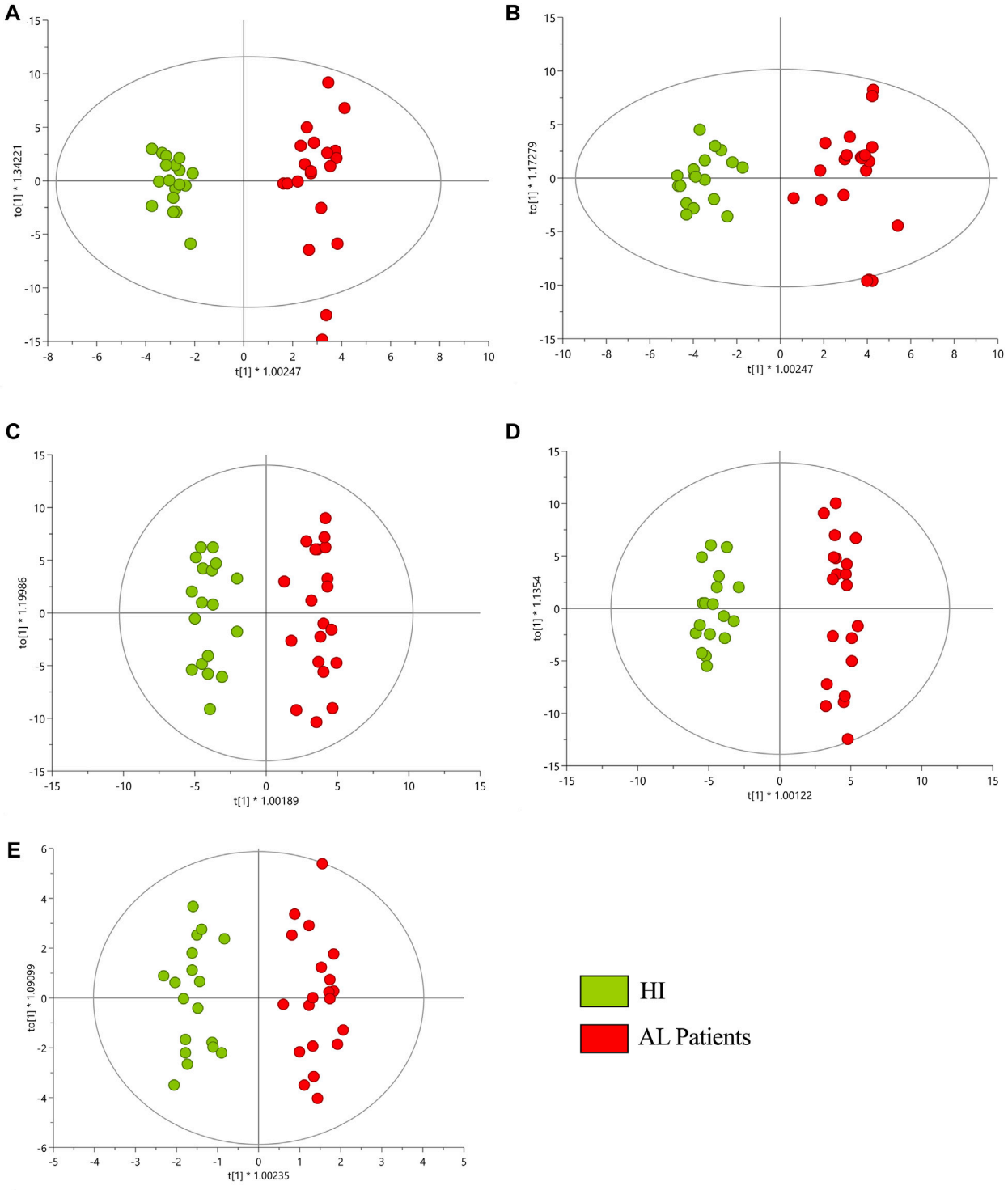


FIGURE 1
OPLS-DA score plot with Pareto scaling for a metabolic analysis of HI and AL patients. **(A)** GM-LC/MS (+) R^2: 0.958, Q^2: 0.852, $p_{CV\text{-}ANOVA}$: 1.948e^{-11}. **(B)** GM-LC/MS (–) R^2: 0.925, Q^2: 0.838, $p_{CV\text{-}ANOVA}$: 7.3713e^{-11}. **(C)** GL-LC/MS (+) R^2: 0.945, Q^2: 0.670, $p_{CV\text{-}ANOVA}$: 1.986e^{-4}. **(D)** GL-LC/MS (–) R^2: 0.972, Q^2: 0.789, $p_{CV\text{-}ANOVA}$: 6.991e^{-8}. **(E)** GC/MS R^2: 0.937, Q^2: 0.743, $p_{CV\text{-}ANOVA}$: 9.262e^{-7}. Red dots correspond to AL patients, and green dots are HI.

The node color of each pathway is determined by the *p*-value (red = lowest *p*-value and highest statistical significance), and the node radius (size) is based on the pathway impact factor, with the biggest indicating the highest impact (Mashabela et al., 2022). The metabolic pathways that were significantly altered in patients with AL compared with HI were sphingolipids, glycerophospholipids, alanine, aspartate, and glutamate metabolism (Figure 4). The metabolites identified within these altered pathways are a

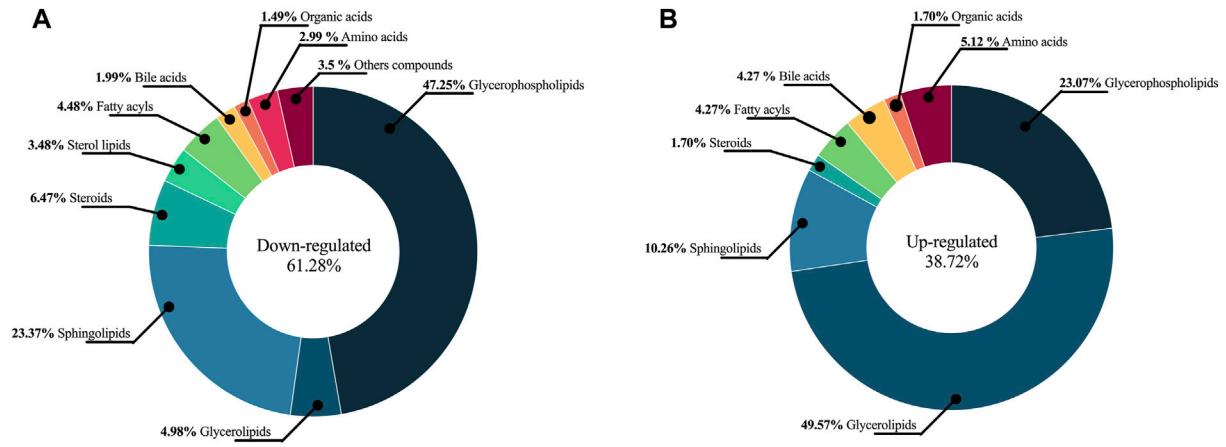


FIGURE 2
The chemical classes altered between AL patients and HI. The chemical classes and percentages are shown in a pie chart. **(A)** Downregulated metabolites **(B)** Upregulated metabolites.

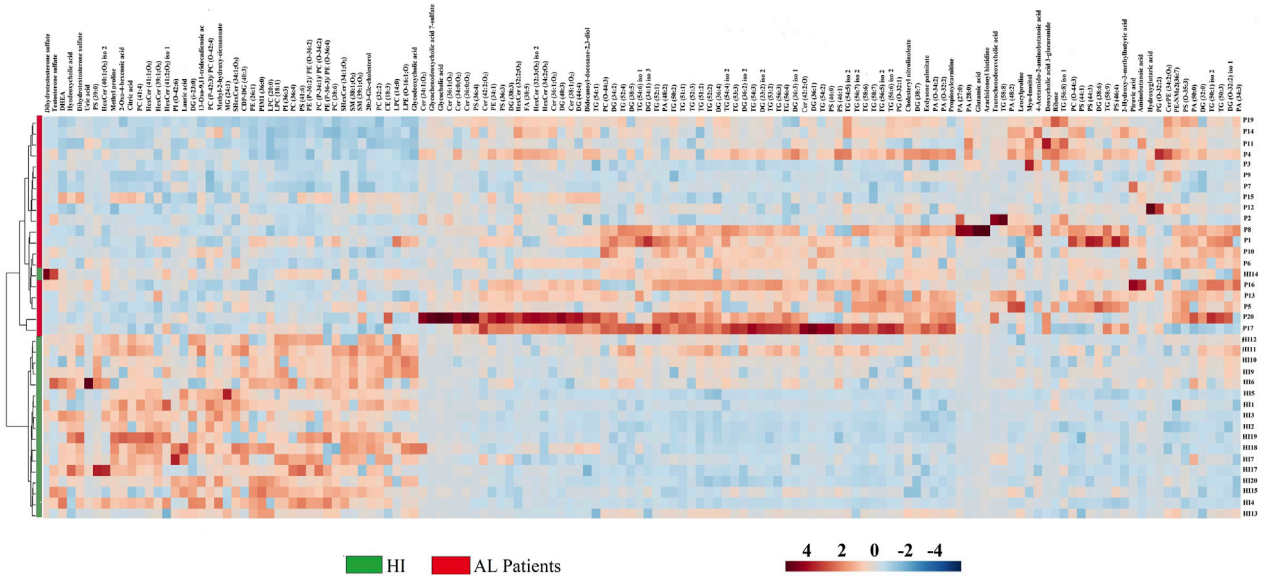


FIGURE 3
Heatmap of the metabolites and lipids with statistically significant variation. Fold Change ≥ 1.5 and ≤ 0.5, $p < 0.05$ between AL patients and HI.

significant increase in pyruvate and AA and a significant decrease in linoleic acid, phosphatidylcholine, phosphatidylethanolamine, and lysophosphatidylcholine were found in the patients. A summary graph of the main altered 328 metabolites and their participation in the different metabolic pathways associated with leukemia is shown in Figure 5.

The HI have higher TAC than the AL patients

This work shows a higher concentration of slow antioxidants (Q2) than fast antioxidants (Q1), both in HI and in patients. Also, significant differences in rapid antioxidants (Q1) between HI and patients with AL were observed. Regarding the QT value, the patients presented less TAC than the HI (Figure 6; Supplementary Table S3). It is possible that there is a relationship between the decrease in TAC and the lipid alterations present in the patients. The decrease in antioxidant systems would favor an increase in ROS, which could stimulate survival signals or oxidize macromolecules such as lipids, inducing cell death (Barrera, 2012).

Discussion

Thanks to the technical progress of metabolomics and the complexity of biological samples, the simultaneous use of several analytical platforms allows expanding the coverage of identification

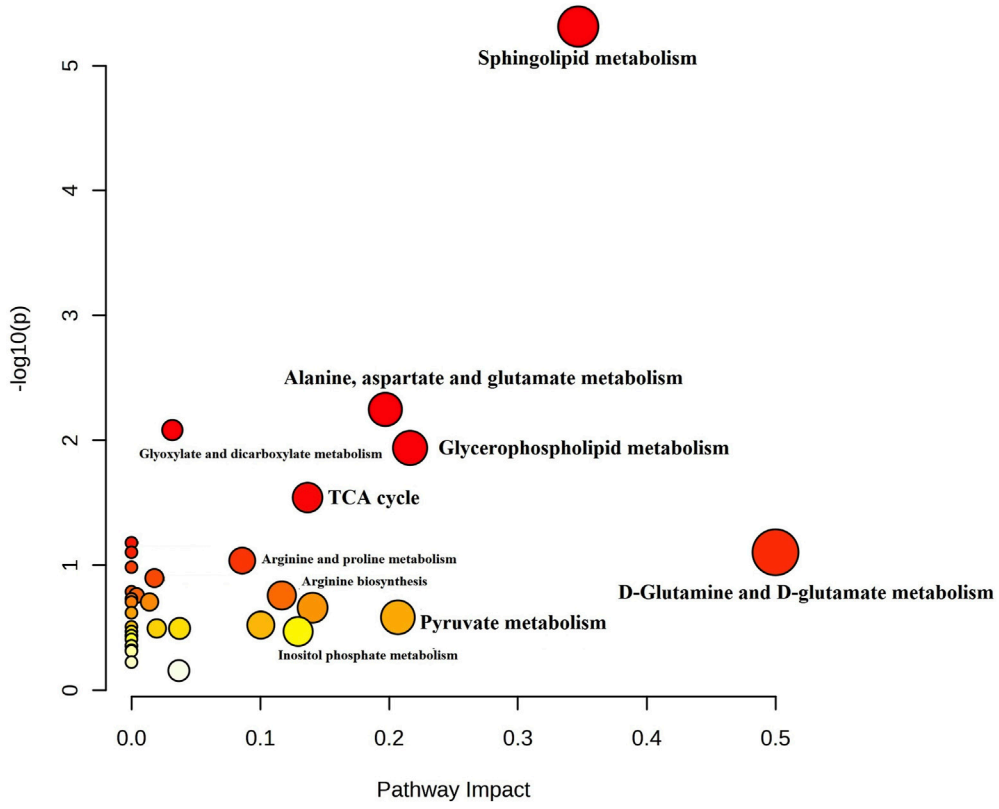


FIGURE 4
Summary of Metabolomic Pathway Analysis (MetPA) as generated by MetaboAnalyst 5.0 software package in AL patients. Using the KEGG database (All the matched pathways are displayed as circles). The color of each circle is based on *p*-values (darker colors indicate more significant changes of metabolites in the corresponding pathway). In contrast, the circle size corresponds to the pathway impact score. The most impacted pathways having high statistical significance scores are annotated (*p*-value < 0.05; pathway impact values ≥ 0.2.

and characterization of metabolites, making it possible to find new non-invasive biomarkers for the diagnosis and prognosis of AL, as well as delving into the characteristics and biological differences of the lineage of origin of the disease, lymphoid or myeloid (Bruno et al., 2018). In this study, metabolic differences were analyzed at the plasma level between 18 HI and 20 AL patients, using metabolomics by LC-QTOF-MS and GC-QTOF-MS and lipidomics by LC-QTOF-MS.

Among the results obtained, it was shown that the primary metabolic alterations in patients with AL are related to lipid metabolism, in agreement with previous reports in the literature (Musharraf et al., 2016). Lipids are essential components of malignant tumors, as they are necessary for the growth and spread of the tumor. Fatty acids, cholesterol, and phospholipids are the most important sources of energy production, function as signaling molecules, and participate in the biogenesis of cell membranes. They can be provided by the tumor microenvironment or by cancer cells themselves through the activation of *de novo* synthesis pathways. Importantly, especially cells of the immune system, cancer-associated fibroblasts, and cancer-associated adipocytes, can also undergo changes in lipid content, hindering or promoting tumor aggressiveness (Fu et al., 2020; Vasseur and Guillaumond, 2022). In AML, lipids have been used to identify genetic signatures related to prognosis, the immunological panorama, and characteristics of the tumor microenvironment (Ding et al., 2022) and, in turn, as markers to predict the risk of acute graft-and-host disease (aGvHD) from allogeneic hematopoietic stem cell transplantation (alloHSCT) (Liu et al., 2019).

Fatty acids are the main building blocks of several lipid species, they can be channeled into various metabolic pathways to synthesize complex lipid species, including glycerolipids such as DG and TG, glycerophospholipids such as phosphatidic acid, phosphatidylethanolamine, PS, phosphatidylglycerol and phosphatidylcholine, sphingolipids and cholesterol (CL) including cholesterol ester (Koundouros and Poulogiannis, 2020). We observed a decrease in glycerophospholipids and an increase in glycerolipids in AL patients. Since glycerophospholipids are the main constituents of cell membranes, it is possible that they are being rapidly consumed by proliferating cells at the expense of Increased glycerolipids that serve as central intermediates in glycerophospholipid synthesis or as lipid storage molecules (Pan et al., 2021).

Within the increase in glycerolipids, TG was the most relevant. Altered glycerophospholipid metabolism has previously been associated with disease progression in pediatric ALL patients (Yunnuo et al., 2014) and an increase in TG in conjunction with a decrease in CL has been reported in both AML and ALL (Nahid et al., 2013), which has been related to a poor response to treatment (Guzmán and Sandoval, 2004). Particularly, Pabst et al. (2017)

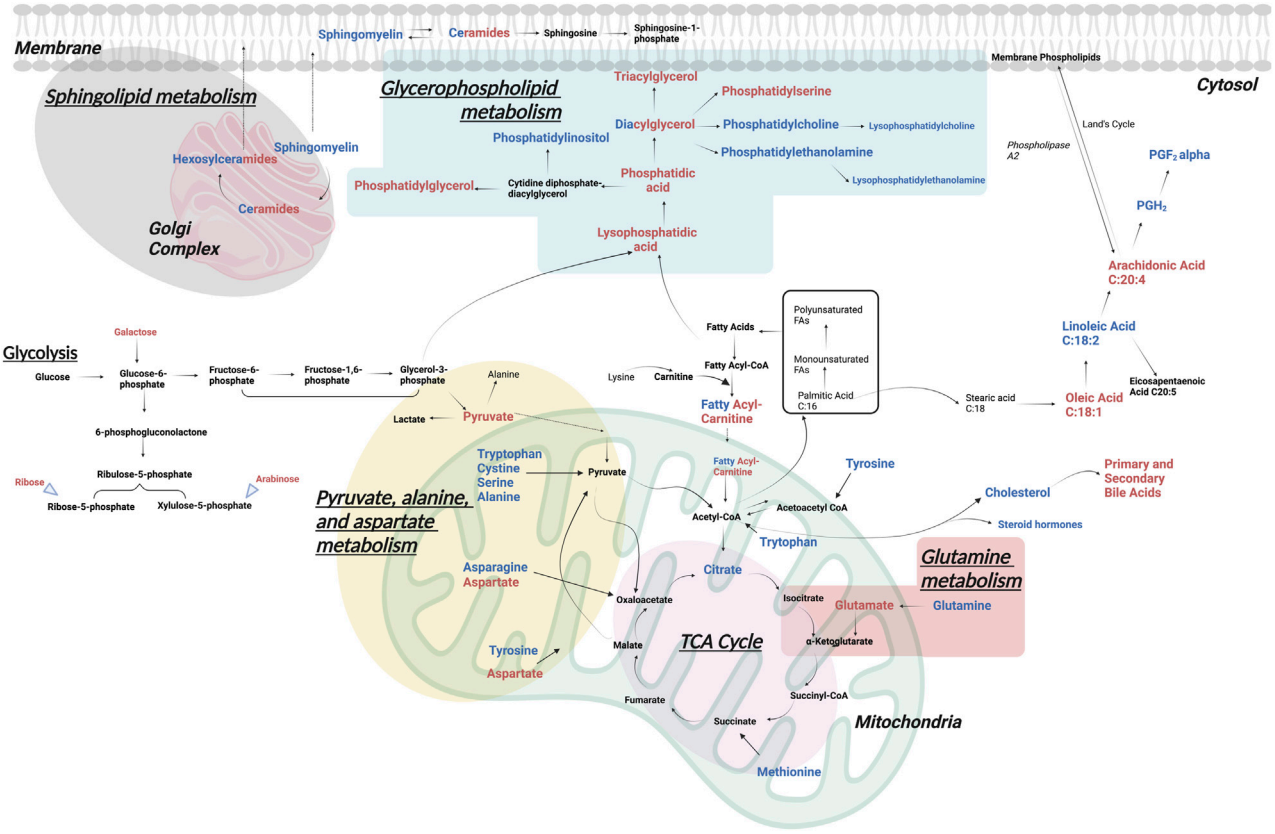


FIGURE 5
Significantly altered metabolomic pathways in AL. The main metabolic pathways altered in patients with AL vs. HI, identified by different metabolomic and lipidomic analytical platforms in peripheral blood plasma, are represented. In red upregulated and in blue downregulated metabolites.

analyzed 20 samples from individuals with AML and 20 HI by GC-MS and ultraperformance liquid chromatography-electrospray ionization-quadrupole time-of-flight mass spectrometry (UPLC-ESI-QTOFMS), obtaining similar results, a decrease in PCs, cholesterol ester, and CL in the patients, however, they found a reduction in TG in the patients, probably associated with consumption by proliferating AML cells. From the point of view of the evolution of the disease, it has been described in other works, that patients with myelodysplastic syndromes who progressed to AL had higher TG levels than those who did not evolve (Qiao et al., 2022) and that in patients with ALL, after a 5-year disease-free period, they developed dyslipidemia with increased plasma TG, increased LDL CL, and decreased HDL CL, which was associated with an increased risk of atherosclerotic disease (Morel et al., 2017). These findings indicate that metabolic alterations at the lipid level with increased TG play an essential role in leukemogenesis, maintenance, and tumor progression but are also associated with clinical complications in these patients, such as atherosclerosis.

It should be noted that LICs or their equivalents in other cancer stem cells tumor models as initiating cells of the leukemogenesis process have a profile of genetic alterations associated with high risk and a specific metabolic profile, which are relevant as mechanisms implicated in treatment resistance and disease relapse (Marchand and Pinho, 2021). In particular, the increase in the synthesis of lipids in the LICs favors their self-renewal capacity by increasing the production of NADPH, which is an essential cofactor in reducing oxidized GSH to reduced GSH and in the maintenance of low ROS levels (Liu et al., 2022). Interestingly, Ito et al. (2012) showed that the PML gene controls asymmetric and symmetric HSC division through PPARδ activity, a regulator of fatty acid synthesis. In prostate cancer, a high lipid diet may accelerate tumor cell proliferation by increasing levels of insulin-like growth factor 1, IL-1α, IL-1β, IL -6, or TNF-α (Xu et al., 2014) or through activation of signaling pathways such as MCP-1/CCR2 (monocyte chemoattractant protein-1/C-C Motif Chemokine Receptor 2) (Huang et al., 2012). In addition, a high lipid diet accelerates the development of AML in a murine knock-in model for MLL-AF9 through the activation of the FLT3 receptor (Fms Related Receptor Tyrosine Kinase 3) on the membrane of c-KIT + primitive hematopoietic stem cells, with subsequent activation of the JAK3-STAT3 (Janus kinase/signal transducer and activator of transcription) signaling pathway (Hermetet et al., 2020).

Of the lipids, the ones best characterized in oncogenic signaling are phosphoinositols (PI) and ceramides/sphingolipids (SF), which we found decreased in patients compared to HI. However, this is one of the few studies that reflect this alteration at plasma levels in patients with AL (Calderon-Rodríguez et al., 2019). PIs are precursors of phosphoinositide, such as PI(3,4,5)P3, which can promote tumorigenesis by activating the AKT/mTORC1/2 pathway, which is frequently altered in AL (Nepstad et al.,

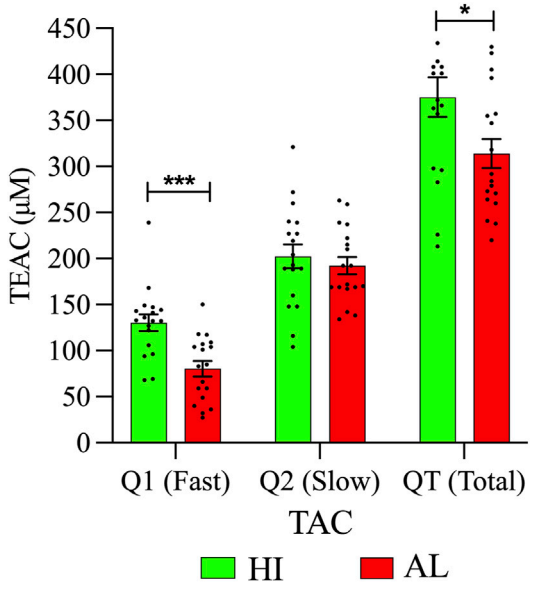


FIGURE 6
Plasma metabolic differences between AL patients and HI. TAC levels in the peripheral blood plasma of patients, and HI are expressed in Trolox Equivalent Antioxidant Capacity (TEAC). In all cases, data are represented as the mean ± SEM.

2020), while SF participates in regular signals of cell survival or apoptosis (Ogretmen, 2018). Relevantly, the activation of cell signaling pathways due to mutations in oncogenes such as Ras and FLT3 plays a vital role in the metabolic reprogramming of leukemic blasts (Wojcicki et al., 2020). In fact, recent *in vitro* studies have shown that distinct genetic changes in AML are associated with improved dynamics and metabolism of different types of lipids, such as ceramides. Additionally, it has been found that patients with abnormal karyotypes, particularly those who have recurrent AML genetic changes like the t(8; 21)(q22; q22.1);*RUNX1-RUNX1T1* translocation or inv(16)(p13.1q22) inversion, have higher levels of ceramide/sphingolipid production (Stefanko et al., 2017).

The other altered metabolites were amino acids. Leucylproline, 4-acetamido-amino butanoic acid, glutamate, amino butanoic acid, and methylmethylproline were found to be increased in the patients. Leucylproline is a dipeptide formed by leucine and proline residues. On the one hand, leucine is part of the branched-chain amino acids (BCAA), which have been shown to be essential for the proliferation of leukemic cells (independent of their lineage). Since it supports the synthesis of non-essential amino acids and the TCA cycle (Tabe et al., 2019). Most patients with AML and ALL have a high level of BCAA transporters (BCAT1), while serum BCAA levels are reduced, suggesting active absorption of BCAA (Kikushige et al., 2023) and the formation of secondary metabolites (dipeptides). We found a reduction in glutamine and an increase in glutamate, suggesting an active metabolism of glutamine. Particularly, glutamate has been described as an exquisite source for leukemic cells since it promotes a tumor phenotype by participating in signaling reactions; it is a source of nitrogen for DNA synthesis and other amino acids; it participates in redox reactions through GSH; and it is a source of biomass and energy as it is incorporated into the TCA cycle (Kreitz et al., 2019). Carnitines are a fundamental part of the synthesis of fatty acids and are vital mediators for tumor metabolic plasticity (Melone et al., 2018). Like us, Morad et al. (2022) demonstrated a reduction in plasma O-acetyl carnitine in patients with ALL and AML. However, the metabolism of carnitines must be studied in depth because some chemotherapeutic drugs interfere with the absorption, synthesis, and excretion of carnitine in non-tumor tissues, leading to secondary carnitine deficiency and therefore multi-organ toxicity, which can be reversed with carnitine treatment without affects effectiveness. Anticancer by affects effectiveness anticancer (Sayed-Ahmed, 2010).

Tumor cells develop a mechanism where they adjust to the high ROS by expressing elevated levels of antioxidant proteins to detoxify them while maintaining pro-tumorigenic signaling and resistance to apoptosis. At the systemic level, a reduction in the expression of antioxidant enzymes and antioxidant capacity has been reported in AL samples (Rasool et al., 2015; Chaudhary et al., 2023). The reduction in TAC could reflect the consumption of endogenous antioxidants due to the generation of free radicals by the leukemic process (mutations or enzyme alterations) and maintain the redox balance The lower concentration of fast antioxidants (Q1) is expected given their oxidative potential. The increase in free radicals can cause lipid peroxidation, where polyunsaturated fatty acids are more susceptible, such as arachidonic acid. Iron-dependent lipid peroxidation is an important driver of ferroptosis, and ferroptosis is critically involved in the pathogenesis of AL. Interestingly, circulating antioxidants related to dietary intake (vitamin C, carotenoids, vitamin A, and vitamin E) may impact tumor progression in some types of cancer and vary depending on the dietary culture (Abenavoli et al., 2019; Yin et al., 2022).

We recognize that our study consolidates a low number of patients, however, there are not a huge number of primary acute leukemia samples in our region, which is reflected by the number of patients included in recent papers in Colombia (Calderon-Rodríguez et al., 2019) and in Latin America (Aguirre-Guillén et al., 2017). Therefore, our study is pioneering in a different social and cultural context and our findings may inspire more research on the metabolism of malignant hemopathies. We will verify the data in a larger cohort of patients in the medium term. Likewise, this metabolic profile will help us to follow the evolution of acute leukemia patients recruited in a clinical trial in which a new medicant directed to metabolism regulation is tested. We also consider integrating other omics techniques, particularly transcriptomics with metabolomics, to strengthen longitudinal studies, offering the opportunity to design and apply personalized treatments and advance in the search for biomarkers predictive of clinical response.

Conclusion

This study reinforces previous observations of lipid abnormalities in patients with acute leukemia (AL), highlighting the significance of these metabolic dysregulations in the disease. Our findings indicate that glycerophospholipid metabolism, sphingolipid metabolism, and the metabolism of alanine, aspartate, glutamate, and glutamine are the primary deregulated pathways in AL patients. Additionally, we observed a lower total antioxidant capacity (TAC)

in AL patients, reflecting the consumption of antioxidants during the leukemogenic process. These metabolic findings contribute to a deeper understanding of the physiological characteristics of leukemia and provide valuable insights for targeted therapeutic interventions and personalized treatment strategies. Furthermore, to our knowledge, this research represents the first metabolomics investigation conducted on the Colombian population, underscoring the novelty of our results. Future validation studies are necessary to confirm these findings with a larger cohort and elucidate their clinical implications.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Hospital Universitario San Ignacio and the Centro Javeriano de Oncología (Bogotá D.C., Colombia) (FM-CIE-0202-19, 2 May 2019). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

Actively involved in the samples collection and clinical data abstraction, analyzed and interpreted data and wrote the manuscript, CA; analyzed and interpreted data and wrote the manuscript, LR; analyzed data, performed the statistical analysis, and wrote the manuscript, MS; involved in the samples collection and clinical data abstraction, LM, LA, PS, and MA; performed the statistical analysis and analyzed data, RB-R; conceptualized, supervised the research and wrote the manuscript, SQ; conceptualized and supervised the research, MC; conceptualized, supervised the research, wrote the manuscript and responsible of funding acquisition, SF. All authors contributed to the article and approved the submitted version.

Funding

Ministerio de Ciencia, Tecnología e Innovación (Colciencias), Ministerio de Educación Nacional, Ministerio de Industria, Comercio y Turismo e ICETEX (792–2017 [a]a Convocatoria Ecosistema Científico–- Colombia Científica para la Financiación de Proyectos de I + D + i), and Vicerrectoría de Investigaciones, Pontificia Universidad Javeriana, Bogotá D.C., Colombia (Contract No. FP44842-221-2018).

Acknowledgments

We acknowledge the financial resources. We extend a special thanks to all the patients and families, the Hospital Universitario San Ignacio, the Centro Javeriano de Oncología (Hematology, Nursing, and Clinical Laboratory Service), and the Grupo de Inmunobiología y Biología Celular of the Pontificia Universidad Javeriana, Bogotá D.C., Colombia.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1235160/full#supplementary-material

References

Abenavoli, L., Boccuto, L., Federico, A., Dallio, M., Loguercio, C., Di Renzo, L., et al. (2019). Diet and non-alcoholic fatty liver disease: the mediterranean way. *Int. J. Environ. Res. public health* 16 (17), 3011. doi:10.3390/ijerph16173011

Aguirre-Guillén, W. A., Angeles-Floriano, T., López-Martínez, B., Reyes-Morales, H., Zlotnik, A., and Valle-Rios, R. (2017). Omics techniques and biobanks to find new biomarkers for the early detection of acute lymphoblastic leukemia in middle-income countries: a perspective from Mexico. *Bol. medico del Hosp. Infant. Mex.* 74 (3), 227–232. doi:10.1016/j.bmhimx.2017.03.007

Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., et al. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 127 (20), 2391–2405. doi:10.1182/blood-2016-03-643544

Ballesteros-Ramírez, R., Quijano, S., Solano, J., Ordoñez-Reyes, C., Herrera, M. V., Murillo, R., et al. (2020). Influence of dose intensity in consolidation with HIDAC and other clinical and biological parameters in the survival of AML. *J. Cancer Epidemiol.* 2020, 8021095. doi:10.1155/2020/8021095

Barrera, G. (2012). Oxidative stress and lipid peroxidation products in cancer progression and therapy. *ISRN Oncol.* 2012, 137289. doi:10.5402/2012/137289

Boag, J. M., Beesley, A. H., Firth, M. J., Freitas, J. R., Ford, J., Hoffmann, K., et al. (2006). Altered glucose metabolism in childhood pre-B acute lymphoblastic leukaemia. *Leukemia* 20 (10), 1731–1737. doi:10.1038/sj.leu.2404365

Bruno, C., Patin, F., Bocca, C., Nadal-Desbarats, L., Bonnier, F., Reynier, P., et al. (2018). The combination of four analytical methods to explore skeletal muscle metabolomics: better coverage of metabolic pathways or a marketing argument? *J. Pharm. Biomed. Anal.* 148, 273–279. doi:10.1016/j.jpba.2017.10.013

Calderon-Rodríguez, S. I., Sanabria-Salas, M. C., and Umaña-Perez, A. (2019). A comparative proteomic study of plasma in Colombian childhood acute lymphoblastic leukemia. *PloS one* 14 (8), e0221509. doi:10.1371/journal.pone.0221509

Chang, J. H., Poppe, M. M., Hua, C. H., Marcus, K. J., and Esiashvili, N. (2021). Acute lymphoblastic leukemia. *Pediatr. Blood Cancer* 68 (Suppl. 2), e28371. doi:10.1002/pbc.28371

Chaudhary, P., Kumari, S., Dewan, P., Gomber, S., Ahmed, R. S., and Kotru, M. (2023). Chemotherapy-Induced oxidative stress in pediatric acute lymphoblastic leukemia. *Cureus* 15, e35968. doi:10.7759/cureus.35968

Chen, W. L., Wang, J. H., Zhao, A. H., Xu, X., Wang, Y. H., Chen, T. L., et al. (2014). A distinct glucose metabolism signature of acute myeloid leukemia with prognostic value. *Blood* 124 (10), 1645–1654. doi:10.1182/blood-2014-02-554204

Combariza, J. F., Casas, C. P., Rodríguez, M., Cardona, A. F., Ospina, É., and Grajales, M. (2007). Supervivencia en adultos con leucemia linfoide aguda *de novo* tratados con el esquema HyperCVAD en el Instituto Nacional de Cancerología (Colombia), entre enero de 2001 y junio de 2005. *Rev. Colomb. Cancerol.* 11.

DeBerardinis, R. J., and Chandel, N. S. (2016). Fundamentals of cancer metabolism. *Sci. Adv.* 2 (5), e1600200. doi:10.1126/sciadv.1600200

Dhakshinamoorthy, S., Dinh, N.-T., Skolnick, J., and Styczynski, M. P. (2015). Metabolomics identifies the intersection of phosphoethanolamine with menaquinone-triggered apoptosis in an *in vitro* model of leukemia. *Mol. Biosyst.* 11 (9), 2406–2416. doi:10.1039/c5mb00237k

Ding, L., Jiaming, L., Wei, Y., Wenbin, G., Wenping, S., Weng, Z., et al. (2022). A distinct lipid metabolism signature of acute myeloid leukemia with prognostic value. *Front. Oncol.* 12, 876981. doi:10.3389/fonc.2022.876981

Dong, W., Guangguo, T., Haibo, W., Peng, C., Jie, H., and Yanyu, W. (2019). Identification of novel serum biomarker for the detection of acute myeloid leukemia based on liquid chromatography-mass spectrometry. *J. Pharm. Biomed. analysis* 166, 357–363. doi:10.1016/j.jpba.2019.01.022

Faulds, M. H., and Dahlman-Wright, K. (2012). Metabolic diseases and cancer risk. *Curr. Opin. Oncol.* 24 (1), 58–61. doi:10.1097/CCO.0b013e32834e0582

Fu, Y., Zou, T., Shen, X., Nelson, P. J., Li, J., Wu, C., et al. (2020). Lipid metabolism in cancer progression and therapeutic strategies. *MedComm* 2 (1), 27–59. doi:10.1002/mco2.27

Grønningsæter, I. S., Fredly, H. K., Gjertsen, B. T., Hatfield, K. J., and Bruserud, Ø. (2019). Systemic metabolomic profiling of acute myeloid leukemia patients before and during disease-stabilizing treatment based on all-trans retinoic acid, valproic acid, and low-dose chemotherapy. *Cells* 8 (10), 1229. doi:10.3390/cells8101229

Guzmán, M., and Sandoval, M. (2004). Colesterol y triglicéridos como marcadores bioquímicos del estado de la enfermedad del paciente con leucemia linfocítica aguda. *An. Fac. Med.* 65 (4), 225–230. doi:10.15381/anales.v65i4.1374

Hao, X., Hui-Tao, Z., Hong-Wen, X., Chun-Lan, H., and Mei-Zhou, H. (2022). Serum metabolomics coupling with clinical laboratory indicators reveal taxonomic features of leukemia. *Front. Pharmacol.* 13, 794042. doi:10.3389/fphar.2022.794042

Hermetet, F., Mshaik, R., Simonet, J., Callier, P., Delva, L., and Quéré, R. (2020). High-fat diet intensifies MLL-AF9-induced acute myeloid leukemia through activation of the FLT3 signaling in mouse primitive hematopoietic cells. *Sci. Rep.* 10 (1), 16187. doi:10.1038/s41598-020-73020-4

Huang, M., Narita, S., Numakura, K., Tsuruta, H., Saito, M., Inoue, T., et al. (2012). A high-fat diet enhances proliferation of prostate cancer cells and activates MCP-1/CCR2 signaling. *Prostate* 72 (16), 1779–1788. doi:10.1002/pros.22531

Islami, F., Goding Sauer, A., Miller, K. D., Siegel, R. L., Fedewa, S. A., Jacobs, E. J., et al. (2018). Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA a cancer J. Clin.* 68 (1), 31–54. doi:10.3322/caac.21440

Ito, K., Carracedo, A., Weiss, D., Arai, F., Ala, U., Avigan, D. E., et al. (2012). A PML–PPAR-δ pathway for fatty acid oxidation regulates hematopoietic stem cell maintenance. *Nat. Med.* 18 (9), 1350–1358. doi:10.1038/nm.2882

Jones, C. L., Stevens, B. M., D'Alessandro, A., Reisz, J. A., Culp-Hill, R., Nemkov, T., et al. (2018). Inhibition of amino acid metabolism selectively targets human leukemia stem cells. *Cancer Cell* 34 (5), 724–740.e724. doi:10.1016/j.ccell.2018.10.005

Kikushige, Y., Miyamoto, T., Kochi, Y., Semba, Y., Ohishi, M., Irifune, H., et al. (2023). Human acute leukemia uses branched-chain amino acid catabolism to maintain stemness through regulating PRC2 function. *Blood Adv.* 7 (14), 3592–3603. doi:10.1182/bloodadvances.2022008242

Kim, H. K., Son, S. Y., Oh, J. S., Song, Y. N., Byun, J. M., Koh, Y., et al. (2021). Metabolic profiling during acute myeloid leukemia progression using paired clinical bone marrow serum samples. *Metabolites* 11 (9), 586. doi:10.3390/metabo11090586

Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S., et al. (2009). FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal. Chem.* 81 (24), 10038–10048. doi:10.1021/ac9019522

Kirwan, J. A., Gika, H., Beger, R. D., Bearden, D., Dunn, W. B., Goodacre, R., et al. (2022). Quality assurance and quality control reporting in untargeted metabolic phenotyping: mQACC recommendations for analytical quality management. *Metabolomics* 18 (9), 70. doi:10.1007/s11306-022-01926-3

Koundouros, N., and Poulogiannis, G. (2020). Reprogramming of fatty acid metabolism in cancer. *Br. J. Cancer* 122 (1), 4–22. doi:10.1038/s41416-019-0650-z

Kreitz, J., Schönfeld, C., Seibert, M., Stolp, V., Alshamleh, I., Oellerich, T., et al. (2019). Metabolic plasticity of acute myeloid leukemia. *Cells* 8 (8), 805. doi:10.3390/cells8080805

Liu, H., Zhang, Z., Song, L., Gao, J., and Liu, Y. (2022). Lipid metabolism of cancer stem cells. *Oncol. Lett.* 23 (4), 119. doi:10.3892/ol.2022.13239

Liu, Y., Huang, A., Chen, Q., Chen, X., Fei, Y., Zhao, X., et al. (2019). A distinct glycerophospholipid metabolism signature of acute graft versus host disease with predictive value. *JCI Insight* 4 (16), e129494. doi:10.1172/jci.insight.129494

Marchand, T., and Pinho, S. (2021). Leukemic stem cells: from leukemic niche biology to treatment opportunities. *Front. Immunol.* 12, 775128. doi:10.3389/fimmu.2021.775128

Mashabela, M. D., Piater, L. A., Steenkamp, P. A., Dubery, I. A., Tugizimana, F., and Mhlongo, M. I. (2022). Comparative metabolite profiling of wheat cultivars (*Triticum aestivum*) reveals signatory markers for resistance and susceptibility to stripe rust and aluminium (Al3+) toxicity. *Metabolites* 12 (2), 98. doi:10.3390/metabo12020098

Melone, M. A. B., Valentino, A., Margarucci, S., Galderisi, U., Giordano, A., and Peluso, G. (2018). The carnitine system and cancer metabolic plasticity. *Cell death Dis.* 9 (2), 228. doi:10.1038/s41419-018-0313-7

Morad, H. M., Abou-Elzahab, M. M., Aref, S., and El-Sokkary, A. M. A. (2022). Diagnostic value of ^{1}H NMR-based metabolomics in acute lymphoblastic leukemia, acute myeloid leukemia, and breast cancer. *ACS Omega* 7 (9), 8128–8140. doi:10.1021/acsomega.2c00083

Morel, S., Leahy, J., Fournier, M., Lamarche, B., Garofalo, C., Grimard, G., et al. (2017). Lipid and lipoprotein abnormalities in acute lymphoblastic leukemia survivors. *J. lipid Res.* 58 (5), 982–993. doi:10.1194/jlr.M072207

Musharraf, S. G., Siddiqui, A. J., Shamsi, T., Choudhary, M. I., and Rahman, A. U. (2016). Serum metabonomics of acute leukemia using nuclear magnetic resonance spectroscopy. *Sci. Rep.* 6 (February), 30693–30699. doi:10.1038/srep30693

Musharraf, S. G., Siddiqui, A. J., Shamsi, T., and Naz, A. (2017). SERUM metabolomics of acute lymphoblastic leukaemia and acute myeloid leukaemia for probing biomarker molecules. *Hematol. Oncol.* 35 (4), 769–777. doi:10.1002/hon.2313

Muthu, M., and Nordström, A. (2019). Current status and future prospects of clinically exploiting cancer-specific metabolism—why is tumor metabolism not more extensively translated into clinical targets and biomarkers? *Int. J. Mol. Sci.* 20 (6), 1385. doi:10.3390/ijms20061385

Nahid, E., Nasrin, D., Fariba, N., Mitra, Z., and Sakine, A. (2013). Serum lipid profile alterations in acute leukemia before and after chemotherapy. *Med. J Mashhad Univ. Med. Sci.* 60 (4), 9.

Naz, A., Shamsi, T. S., Sattar, A., and Mahboob, T. (2013). Oxidative stress and total antioxidant status in acute leukemia at diagnosis and post remission induction phase. *Pak. J. Pharm. Sci.* 26 (6), 1123–1130.

Nepstad, I., Hatfield, K. J., Grønningsæter, I. S., and Reikvam, H. (2020). The PI3K-Akt-mTOR signaling pathway in human acute myeloid leukemia (AML) cells. *Int. J. Mol. Sci.* 21 (8), 2907. doi:10.3390/ijms21082907

Ogretmen, B. (2018). Sphingolipid metabolism in cancer signalling and therapy. *Nat. Rev. Cancer* 18 (1), 33–50. doi:10.1038/nrc.2017.96

Olaniyi, J. A. A., Akinloye, A., Awosika, O., Rahamon, E., and Arinola, S. (2011). Antioxidant levels of acute leukaemia patients in Nigeria. *Sierra Leone J. Biomed. Res.* 3, 5.

Pabst, T., Kortz, L., Fiedler, G. M., Ceglarek, U., Idle, J. R., and Beyoğlu, D. (2017). The plasma lipidome in acute myeloid leukemia at diagnosis in relation to clinical disease features. *BBA Clin.* 7, 105–114. doi:10.1016/j.bbacli.2017.03.002

Pan, M., Qin, C., and Han, X. (2021). Lipid metabolism and lipidomics applications in cancer research. *Adv. Exp. Med. Biol.* 1316, 1–24. doi:10.1007/978-981-33-6785-2_1

Peng, G., Pakstis, A. J., Gandotra, N., Cowan, T. M., Zhao, H., Kidd, K. K., et al. (2022). Metabolic diversity in human populations and correlation with genetic and ancestral geographic distances. *Mol. Genet. metabolism* 137 (3), 292–300. doi:10.1016/j.ymgme.2022.10.002

Pollyea, D. A., Bixby, D., Perl, A., Bhatt, V. R., Altman, J. K., Appelbaum, F. R., et al. (2021). NCCN guidelines insights: acute myeloid leukemia, version 2.2021. *J. Natl. Compr. Canc Netw.* 19 (1), 16–27. doi:10.6004/jnccn.2021.0002

Qiao, W., Young, E., Feng, C., Liu, S., Jin, J., Noor, L., et al. (2022). Association between abnormal lipid profile and inflammation and progression of myelodysplastic syndrome to acute leukemia. *Exp. Hematol. Oncol.* 11 (58), 58. doi:10.1186/s40164-022-00309-7

Rasool, M., Farooq, S., Malik, A., Shaukat, A., Manan, A., Asif, M., et al. (2015). Assessment of circulating biochemical markers and antioxidative status in acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) patients. *Saudi J. Biol. Sci.* 22 (1), 106–111. doi:10.1016/j.sjbs.2014.09.002

Romer-Seibert, J. S., and Meyer, S. E. (2021). Genetic heterogeneity and clonal evolution in acute myeloid leukemia. *Curr. Opin. Hematol.* 28 (1), 64–70. doi:10.1097/moh.0000000000000626

Sayed-Ahmed, M. M. (2010). Role of carnitine in cancer chemotherapy-induced multiple organ toxicity. *Saudi Pharm. J. SPJ official Publ. Saudi Pharm. Soc.* 18 (4), 195–206. doi:10.1016/j.jsps.2010.07.008

Schmidt, D. R., Patel, R., Kirsch, D. G., Lewis, C. A., Vander Heiden, M. G., and Locasale, J. W. (2021). Metabolomics in cancer research and emerging applications in clinical oncology. *CA A Cancer J. Clin.* 71 (4), 333–358. doi:10.3322/caac.21670

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., et al. (2014). Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ. Sci. Technol.* 48 (4), 2097–2098. doi:10.1021/es5002105

Sossa, C. A., Salazar, V., Rosales, L., Idrobo, M., Reyes-Castellanos, H., Mantilla, J., et al. (2021). AML-425: acute myeloid leukemia: a multicenter experience in Colombia, on behalf of ACHO's renehoc investigators author links open overlay panel. *Clin. Lymphoma Myeloma Leuk.* 21, 2. doi:10.1016/S2152-2650(21)01735-3

Stefanko, A., Thiede, C., Ehninger, G., Simons, K., and Grzybek, M. (2017). Lipidomic approach for stratification of acute myeloid leukemia patients. *PLoS ONE* 12 (2), e0168781. doi:10.1371/journal.pone.0168781

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis chemical analysis working group (CAWG) metabolomics standards initiative (MSI). *Metabolomics* 3 (3), 211–221. doi:10.1007/s11306-007-0082-2

Tabe, Y., Lorenzi, P. L., and Konopleva, M. (2019). Amino acid metabolism in hematologic malignancies and the era of targeted therapy. *Blood* 134 (13), 1014–1023. doi:10.1182/blood.2019001034

Vander Heiden, M. G., and DeBerardinis, R. J. (2017). Understanding the intersections between metabolism and cancer biology. *Cell* 168 (4), 657–669. doi:10.1016/j.cell.2016.12.039

Vasseur, S., and Guillaumond, F. (2022). Lipids in cancer: a global view of the contribution of lipid pathways to metastatic formation and treatment resistance. *Oncogenesis* 11 (1), 46. doi:10.1038/s41389-022-00420-8

Wheelock, Å. M., and Wheelock, C. E. (2013). Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Mol. Biosyst.* 9 (11), 2589–2596. doi:10.1039/c3mb70194h

Wojcicki, A. V., Kasowski, M. M., Sakamoto, K. M., and Lacayo, N. (2020). Metabolomics in acute myeloid leukemia. *Mol. Genet. Metab.* 130 (4), 230–238. doi:10.1016/j.ymgme.2020.05.005

Xu, H., Hu, M.-B., Bai, P.-D., Zhu, W.-H., Ding, Q., and Jiang, H.-W. (2014). Will metformin postpone high-fat diet promotion of TRAMP mouse prostate cancer development and progression? *Int. Urology Nephrol.* 46 (12), 2327–2334. doi:10.1007/s11255-014-0823-x

Yang, F., Li, Q., Xiang, J., Zhang, H., Sun, H., Ruan, G., et al. (2021). NMR-based plasma metabolomics of adult B-cell acute lymphoblastic leukemia. *Mol. Omics* 17 (1), 153–159. doi:10.1039/d0mo00067a

Yin, L., Yan, H., Chen, K., Ji, Z., Zhang, X., Ji, G., et al. (2022). Diet-derived circulating antioxidants and risk of digestive system tumors: a mendelian randomization study. *Nutrients* 14 (16), 3274. doi:10.3390/nu14163274

Yunnuo, B., Haitao, Z., Xiaohan, S., Changhao, S., and Lihong, R. (2014). Biomarker identification and pathway analysis by serum metabolomics of childhood acute lymphoblastic leukemia. *Clin. chimica acta; Int. J. Clin. Chem.* 436, 207–216. doi:10.1016/j.cca.2014.05.022

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 15 November 2023
DOI 10.3389/fmolb.2023.1295216

OPEN ACCESS

EDITED BY
Xavier Domingo-Almenara,
Eurecat, Spain

REVIEWED BY
Elizabeth R. Lusczek,
University of Minnesota Twin Cities,
United States
Ana Paula Valente,
Federal University of Rio de Janeiro, Brazil

*CORRESPONDENCE
Nicolás Nin,
niconin@hotmail.com
Guillermo Moyna,
gmoyna@fq.edu.uy

†These authors have contributed equally to this work

RECEIVED 15 September 2023
ACCEPTED 30 October 2023
PUBLISHED 15 November 2023

CITATION
Hurtado JI, López-Radcenco A, Izquierdo-García JL, Rodríguez F, Moyna G, Greif G and Nin N (2023), A comparative NMR-based metabolomics study of lung parenchyma of severe COVID-19 patients.
Front. Mol. Biosci. 10:1295216.
doi: 10.3389/fmolb.2023.1295216

A comparative NMR-based metabolomics study of lung parenchyma of severe COVID-19 patients

Joaquín I. Hurtado[1†], Andrés López-Radcenco[2†], José Luis Izquierdo-García[3,4,5], Fernando Rodríguez[6], Guillermo Moyna[2*], Gonzalo Greif[1] and Nicolás Nin[6,7*]

[1]Laboratorio de Interacción Hospedero Patógeno, Unidad de Biología Molecular, Institut Pasteur de Montevideo, Montevideo, Uruguay, [2]Departamento de Química del Litoral, Universidad de la República, Paysandú, Uruguay, [3]Grupo de Resonancia Magnética Nuclear e Imagen en Biomedicina, Instituto Pluridisciplinar, Universidad Complutense de Madrid, Madrid, Spain, [4]Departamento de Química en Ciencias Farmacéuticas, Facultad de Farmacia, Universidad Complutense de Madrid, Madrid, Spain, [5]Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Madrid, Spain, [6]Centro de Referencia COVID 1, Hospital Español, Administración de Servicios de Salud del Estado, Montevideo, Uruguay, [7]Centro de Referencia COVID 2, Instituto Nacional de Ortopedia y Traumatología, Administración de Servicios de Salud del Estado, Montevideo, Uruguay

COVID-19 was the most significant infectious-agent-related cause of death in the 2020-2021 period. On average, over 60% of those admitted to ICU facilities with this disease died across the globe. In severe cases, COVID-19 leads to respiratory and systemic compromise, including pneumonia-like symptoms, acute respiratory distress syndrome, and multiorgan failure. While the upper respiratory tract and lungs are the principal sites of infection and injury, most studies on the metabolic signatures in COVID-19 patients have been carried out on serum and plasma samples. In this report we attempt to characterize the metabolome of lung parenchyma extracts from fatal COVID-19 cases and compare them with that from other respiratory diseases. Our findings indicate that the metabolomic profiles from fatal COVID-19 and non-COVID-19 cases are markedly different, with the former being the result of increased lactate and amino acid metabolism, altered energy pathways, oxidative stress, and inflammatory response. Overall, these findings provide additional insights into the pathophysiology of COVID-19 that could lead to the development of targeted therapies for the treatment of severe cases of the disease, and further highlight the potential of metabolomic approaches in COVID-19 research.

KEYWORDS

biomarkers, COVID-19, ICU patients, lung parenchyma, NMR-based metabolomics

1 Introduction

As experienced during the 2020–2023 COVID-19 pandemic, SARS-CoV-2 infections can result in a variety of respiratory conditions, including pneumonias-like symptoms, acute respiratory distress syndrome (ARDS), and multiorgan failure (Chavez et al., 2021). Potential risk factors for mortality among patients admitted to ICU included age, obesity, and comorbidities such as hypertension, diabetes, and cardiovascular disease (Ejaz et al., 2020; Ahlström et al., 2021; Booth et al., 2021). It was also observed that the clinical symptoms of COVID-19 could be influenced by viral load as well as by respiratory and gut microbiota

dysbiosis (Liu et al., 2020; Brosseau et al., 2021). While most of the patients diagnosed with COVID-19 attended the disease at home, 13%-14% needed hospitalization in moderate care facilities, and between 5%-6% were admitted to intensive care units (Verity et al., 2020; Gosangi et al., 2022). Hospital mortality was between 30%–60% in case series reported in the first wave, increasing significantly for patients admitted to the ICU who required mechanical ventilation (Abate et al., 2020; Bastos et al., 2021; Estenssoro et al., 2021; Kurtz et al., 2021; Ranzani et al., 2021; Dongelmans et al., 2022). Uruguay was no exception, and towards the end of 2020 the average number of new cases increased exponentially to over 400 cases per day (GUIAD-COVID-19, 2022). In addition, the most prevalent viral variant during the first wave was B.1.1.28 (now designated as P.6), and vaccines were not yet available (Moreno et al., 2020; Elizondo et al., 2021; Rego et al., 2021).

A number of studies have established that SARS-CoV-2 infections set off a chain of events that can lead to a cytokine storm, an immune system overreaction that may result in ARDS (Koçak Tufan et al., 2021), which is the most frequent complication of severe COVID-19 cases. However, there are still several aspects of the disease that remain unknown. In order to elucidate the pathophysiological effects of COVID-19 and improve clinical care through the selection of appropriate treatments, particularly for patients with severe manifestations of the disease, a thorough understanding of the metabolic alterations and early acute lung injury biomarkers are required.

Metabolomic profiling can complement the lack of knowledge regarding the molecular mechanisms underlying clinical manifestations and pathogenesis of COVID-19. Consequently, several studies have employed metabolomic approaches to better understand the metabolic pathways involved in COVID-19 pathogenesis (Ansone et al., 2021; Chen et al., 2022; Murali et al., 2023). Serum-based metabolomic studies in COVID-19 patients revealed altered glycolytic pathways as well as amino acid, lipid, and anaplerotic metabolism, suggesting an impact on energy pathways, inflammatory response, and oxidative stress, and confirming the systemic nature of the disease (Kimhofer et al., 2020; Lorente et al., 2021; Shi et al., 2021; Valdés et al., 2022). Additionally, metabolomic studies have been conducted in different biofluids, including sweat, saliva and used face masks, as well as exhaled breath, serum and plasma, to identify differential metabolites and metabolic changes associated with COVID-19 (Barberis et al., 2020; Barberis et al., 2021; Hasan et al., 2021). However, there are no studies focusing on changes in the metabolic profile in lung tissue, which is SARS-CoV-2 primary site of infection. In the present communication we use an NMR-based non-targeted metabolomics approach to characterize the metabolome of lung parenchyma from fatal COVID-19 cases and compare it with other fatal respiratory diseases. As discussed herein, we found statistically significant differences between metabolites related to energy metabolism and inflammatory processes, revealing a unique metabolic profile in the infected tissue.

2 Materials and methods

2.1 Sample acquisition and experimental design

The inclusion criteria comprised adults 18 years or older admitted to the ICU with respiratory sepsis and respiratory failure and which had received mechanical ventilation. Clinical information was obtained by retrospective chart review, and data of the Acute Physiology and Chronic Health disease Classification System II (APACHE-II) scores on admission, arterial oxygen pressure/inspired fraction of oxygen (PaO_2/FiO_2 or PAFI), the need of vasopressor support, renal or multiorgan failure, and the presence of comorbidities, such as diabetes, hypertension, or obesity, were collected. Fragments of lung tissue were collected during clinical autopsies performed on ICU patients deceased between November 2020 and February 2021 who had SARS-CoV-2 infection confirmed by RT-qPCR (n = 8). As stated above there was no vaccination strategy in place at the time, and therefore none of these patients had received immunization. In addition, lung fragments from non-COVID-19 deceased patients were collected between December 2016 and June 2018 at the same facility and with the same ethical safeguards. This group included microbiological and serological positive results for *Klebsiella pneumoniae*, *Leptospira interrogans*, and respiratory syncytial virus (n = 7). In all cases, tissue samples were obtained in the first 2 h post-mortem and stored at −80 °C until processed for NMR analysis.

2.2 NMR sample preparation and data acquisition

An adaptation of previously published methods was followed (Nakayasu et al., 2016). Briefly, lung tissue samples between 50 and 100 mg in wet weight were homogenized and extracted with 0.7 mL MeOH/H_2O (4:3) in a bullet blender (Next Advance, United States). Subsequently, chloroform was added to reach a final $CHCl_3$/MeOH/H_2O ratio of 8:4:3, vortexed for 5 min, and centrifuged for 5 min at 5,000 g. The aqueous phases were lyophilized and resuspended in a phosphate buffer prepared in D_2O (pH 7.4) (Dona et al., 2014).

Water-suppressed 1D-NOESY ^{1}H NMR spectra of aqueous tissue extracts were obtained at 25 °C on a Bruker AVANCE III 500 operating at a ^{1}H frequency of 500.13 MHz. A spectral width of 10 kHz, a data size of 32 K, and a total of 128 scans were employed to record each spectrum, using a relaxation delay of 4 s between scans. 1D-TOCSY and HSQC spectra were acquired and processed using parameters provided with the spectrometer.

2.3 NMR data processing

NMR data were processed and analyzed with MNova (version 14.0, MestreLab Research, S.L., Santiago de Compostela, Spain). Free induction decays were zero-filled to 64 K points and apodized with a 0.3 Hz exponential window function prior to Fourier transformation. All spectra were manually phase- and baseline-corrected, and referenced to the anomeric proton signal of α-glucose (5.22 ppm). Spectra were manually aligned, and the data was normalized to the total spectral area after excluding the residual water resonance region and regions without signals. No binning was employed to construct the data matrices used for the multivariate statistical analyses.

2.4 Metabolite identification and estimation of relative concentrations

Metabolites were identified by comparison of ^{1}H NMR data against spectral repositories, including the Biological Magnetic Resonance Bank (BMRB) (Hoch et al., 2023), the Human Metabolome Database (HMDB) (Wishart et al., 2022), and Chenomx (version 9, Chenomx, Inc., Edmonton, Canada). When required, metabolite identification was confirmed with data from 1D-TOCSY and HSQC spectra.

Given the characteristics of lung parenchyma and the difficulties of obtaining precise dry weights in biologically-hazardous samples, variations in metabolite levels were estimated using relative concentrations. This figure was computed as the ratio between the area from individual metabolite ^{1}H NMR signals and the total area of the spectrum.

2.5 Statistical analysis

Multivariate statistical analyses, including principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA), were carried out with the PLS_Toolbox package (version 8.5, Eigenvector Research Inc., Manson, WA, United States) implemented for MATLAB (revision 2014a, The MathWorks Inc., Natick, MA, United States). For all models, the data was mean-centered and scaled using a Pareto factor (Van Den Berg et al., 2006). Analysis of the data was first performed with PCA, which reduces data dimensionality and facilitates the identification of clusters or trends (Wold et al., 1987; Trygg and Wold, 2002; Trygg et al., 2006). The PCA scores plot was also employed to identify strong outliers outside the 95% significance region of Hotelling's T2 ellipse. Cross-validation of OPLS-DA models was achieved using the random subset method, which involved 20 iterations over data split into 5 equally-sized parts. Receiver operating characteristic (ROC) curves were plotted, and areas under the curves were calculated to ensure the goodness of fit of the resulting models (Ekelund, 2012; Simundic, 2012). Permutation tests with 100 iterations were also performed to determine the degree of over-fitting and further validate the discriminant analyses (Ni et al., 2008). When needed, statistical total correlation spectroscopy (STOCSY) analyses were performed with an in-house MATLAB script based on the algorithm described elsewhere (Cloarec et al., 2005).

Pairwise *t*-test comparisons were carried out between continuous demographic variables as well as between the relative concentrations of all identified metabolites in COVID-19 and non-COVID-19 samples using GraphPad Prism (version 7.0, GraphPad Software, Inc., San Diego, CA, United States).

2.6 Metabolic pathways analyses

Metabolic pathway analysis was performed using the Pathway Analysis module of Metaboanalyst v.5.0 (Xia et al., 2011; Chong et al., 2019), which combines results from robust pathway enrichment analysis with pathway topology analysis to identify the most relevant pathways involved in the conditions under study (Aittokallio and Schwikowski, 2006; Kankainen et al., 2011). The selected pathway enrichment analysis method was GlobalAncova (Hummel et al., 2008), the node importance measure for topological analysis was out-degree centrality, and KEGG metabolic pathways were used as the backend knowledgebase.

TABLE 1 Demographic and clinical characteristics of the study population upon admission in ICU. Variations in continuous variabbles with *p*-values <0.05 are indicated with bold numbers.

Parameter	COVID-19	Non-COVID-19	*p*-value
Cohort size (*n*)	8	7	-
Mean age	68.6 ± 8.2	57.3 ± 17.1	0.992
Female	3 (37%)	3 (50%)	-
COPD[a]	4 (50%)	2 (33%)	-
Diabetes	3 (38%)	0 (0%)	-
Hypertension	7 (88%)	1 (17%)	-
Obesity	3 (38%)	0 (0%)	-
Renal failure	5 (63%)	2 (33%)	-
APACHE-II score	20.6 ± 8.4	19.2 ± 10.2	>0.999
PAFI[b] on day 1	115 ± 31	230 ± 162	**0.001**
Vasopressor support	8 (100%)	5 (83%)	-
Days of mechanical ventilation	16.4 ± 5.3	9.7 ± 7.1	>0.999
Length of ICU stay	17.6 ± 4.9	11.2 ± 8.3	>0.999

[a]Chronic obstructive pulmonary disease.
[b]Arterial oxygen pressure/inspired fraction of oxygen (PaO2/FiO2 or PAFI).

3 Results

3.1 Clinical characteristics of study patients

All patients in this study had been diagnosed with pneumonia, presented respiratory sepsis, and exhibited high APACHE-II scores upon admission to the ICU (Table 1). They all required mechanical ventilation, and more than 80% were on vasopressor support. The average ICU stay was 17.6 ± 4.9 days for COVID-19, and 11.2 ± 8.3 for non-COVID-19 patients. When compared to non-COVID patients, those with COVID-19 had a higher percentage of comorbidities on admission (diabetes, hypertension, chronic obstructive pulmonary disease, and obesity) and a lower PAFI score.

3.2 Metabolomic analysis

We initially compared ^{1}H NMR profiles from lung tissue extracts from COVID-19 autopsies against those from non-COVID-19 autopsies (Figure 1). As shown in Figure 2A, a PCA derived from the ^{1}H NMR data showed good discrimination between groups despite the low number of samples. Indeed, inspection of the loading plot from an OPLS-DA model obtained

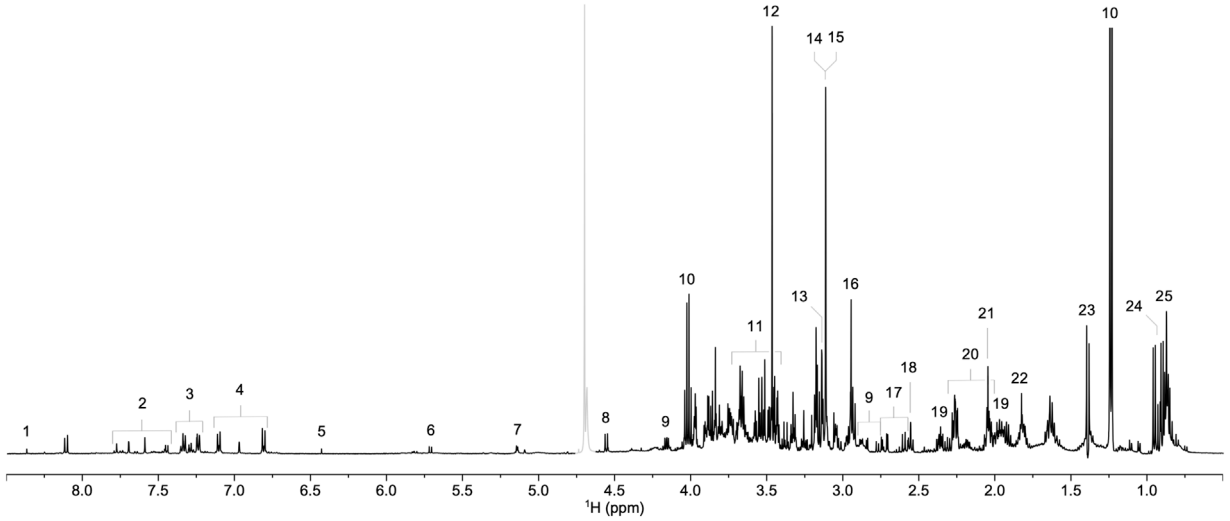


FIGURE 1
Representative ^{1}H NMR spectrum of a lung parenchyma extract sample. Signals corresponding to formate (1), tryptophan (2), phenylalanine (3), tyrosine (4), fumarate (5), uracil (6), α-glucose (7), β-glucose (8), asparagine (9), lactate (10), glycerol-3-phosphate (11), glycine (12), betaine (13), choline (14), phosphocholine (15), creatine (16), citrate (17), pyruvate (18), glutamine (19), glutamate (20), methionine (21), acetate (22), alanine (23), valine (24), and isoleucine (25) are annotated. The grayed-out region corresponds to the residual HDO signal.

with the same data identified an important number of discriminating ^{1}H signals (Figures 2B, C). Dereplication using a combination of STOCSY analyses, classical 1D and 2D NMR experiments, and comparison to data from various ^{1}H spectral repositories allowed us to identify 21 metabolites (Figure 2C), 11 of which had significant differences in levels among the two cohorts (Table 2). The relative concentrations of the amino acids valine, alanine, methionine, glycine, tryptophane, phenylalanine, tyrosine, and asparagine were significantly increased in samples from COVID-19 patients. On the other hand, choline and glycerol-3-phosphate levels, as well as that of the metabolic intermediate succinate, were significantly lower among these samples.

3.3 Pathway analysis results

Metabolic pathway analysis was performed to identify the most relevant pathways involved in COVID-19 lung autopsy (Figure 3). This pathway analysis identified alterations in amino acids biosynthesis and degradation, anaplerotic alanine-aspartate-glutamate metabolism, glycine-serine-threonine metabolism, synthesis and degradation of ketone bodies and glycerophospholipid metabolism.

4 Discussion

One of the most salient aspects from the results presented above is the general increase in the levels of essential amino acids, generally recognized as sepsis biomarkers (Mierzchala-Pasierb et al., 2020; Ahn et al., 2021), in patients with COVID-19. Indeed, branched chain amino acids (BCAAs), including isoleucine and valine (Table 2), are involved in stress, energy, and muscle metabolism (Neinast et al., 2019). BCAAs have different metabolic routes, with valine going solely to carbohydrates (glucogenic), leucine solely to fats (ketogenic), and isoleucine being both a glucogenic and a ketogenic amino acid. These metabolites can also regulate immune responses and influence viral infection (Atila et al., 2021). Hence, the maintenance of metabolic homeostasis is essential for the body's normal physiological functioning, and disruptions in metabolic homeostasis could potentially facilitate virus infection. Our results in lung autopsies of COVID-19 patients show a significant enrichment in valine (Table 2). This is also evidenced by the metabolic pathway analysis, which revealed that valine, leucine, and isoleucine degradation and, to a lesser extent synthesis pathways, are significantly affected (Figure 3). High levels of BCAAs are associated with metabolic encephalopathy, often linked with respiratory suppression, epileptic seizures, and brain damage due to lack of oxygen (Ozturk et al., 2022). These results contrast those from a previous study conducted in serum, where the metabolic profiles of patients with ARDS due to COVID-19 and H1N1 were compared (Lorente et al., 2021). This report by Lorente and coworkers is particularly noteworthy, as it presents a footprint analysis in patients with the same severity of ARDS. On the other hand, most existing metabolomic studies contrast SARS-CoV-2 infected patients with healthy controls and cannot discern between metabolic dysregulations caused by the virus or the development of ARDS. These authors found that amino acid metabolism was decreased in COVID-19 patients, and the concentration of BCAAs, including isoleucine and valine, were also lower when compared with influenza A patients. Although different biofluids are commonly used for biomarker discovery, it is necessary to consider lung tissue metabolome as a complementary input. Indeed, it is not uncommon to find that certain metabolites are decreased in serum but increased in the tissue (Bernatchez and McCall, 2020).

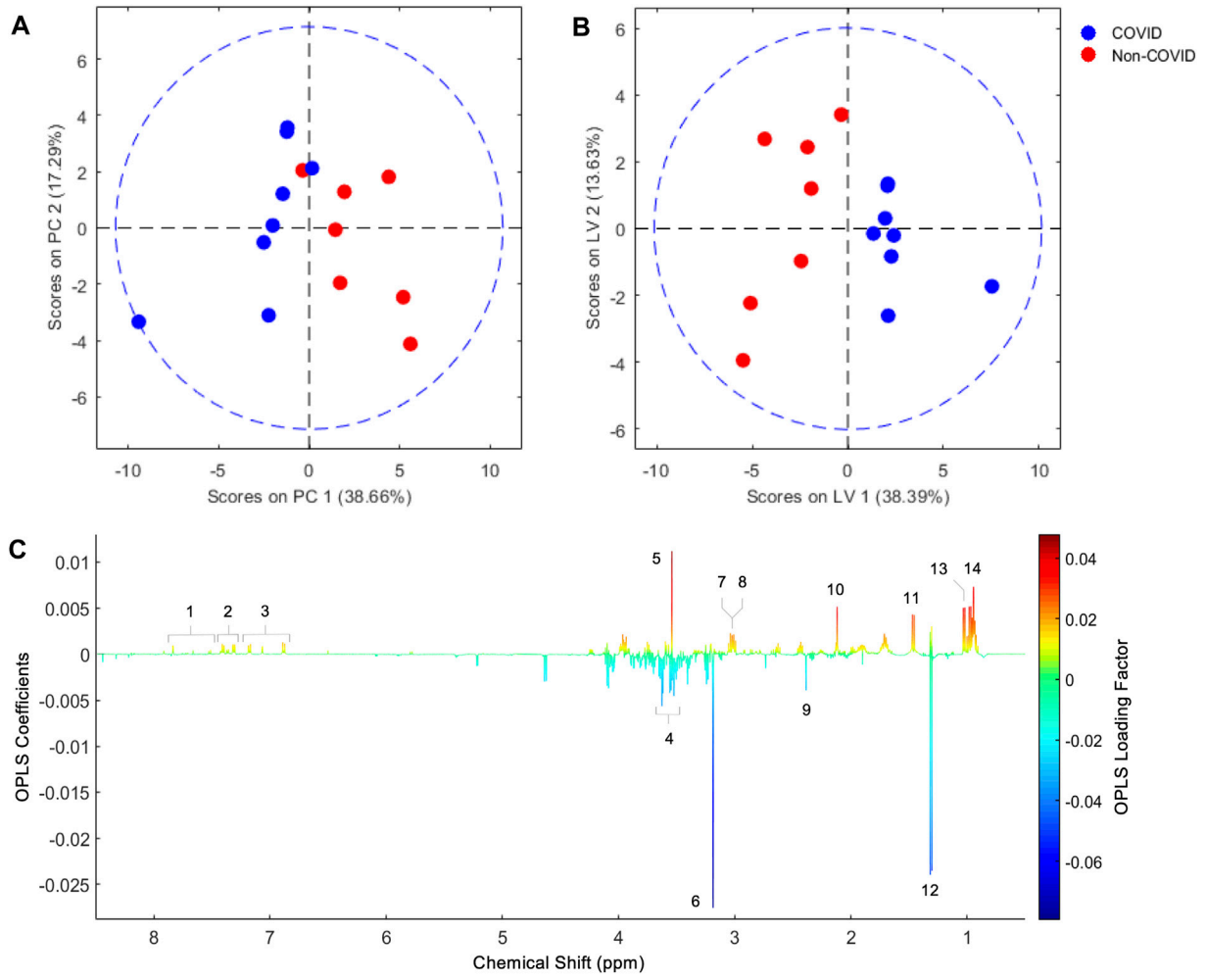


FIGURE 2
PCA score plot obtained from lung parenchyma extract ^{1}H NMR data **(A)**, and score and loading factor plots obtained from the OPLS-DA comparing COVID-19 and non-COVID samples (**B** and **C**, respectively). Metabolites that differentiate the COVID-19 from the non-COVID-19 cohorts are annotated in the loading factor plot, including tryptophan (1), phenylalanine (2), tyrosine (3), glycerol-3-phosphate (4), glycine (5), choline (6), creatine (7), asparagine (8), succinate (9), methionine (10), alanine (11), lactate (12), valine (13), and isoleucine (14). The OLPS-DA model had R^2Y and Q^2Y coefficients of 0.75 and 0.32, respectively, and its ROC curve had an AUC value of 0.98 (Supplementary Figures S1, S2).

Other metabolites found to be significantly more abundant in patients with COVID-19 were tyrosine, phenylalanine, and tryptophan. Absorption of the latter metabolite is mediated by angiotensin converting enzyme 2 (ACE2), the primary receptor of SARS-CoV-2, and has been recognized as a marker of inflammation in severe COVID-19 cases (Takeshita and Yamamoto, 2022). Similarly, elevated plasma or serum levels of tyrosine are observed in a variety of ailments, including hyperphenylalaninemia, sepsis, severe burns, transient tyrosinemia and hyperphenylalaninemia of the newborn, phlebotomus fever, viral hepatitis, or hepatic encephalopathy (Rosen et al., 1977; Watanabe et al., 1979; Rudnick and Ebach, 2004; Ansone et al., 2021). High levels of this non-essential amino acid synthetized from phenylalanine have also been detected in septic patients (Freund et al., 1978). Also, increased phenylalanine serum concentrations have been associated with immunological activation and an increased risk of cardiovascular events in sepsis and other viral infections (Ansone et al., 2021). This could be explained due to muscle tissue catabolism leading to amino acid release, which, together with the body's differential metabolic capacity for different amino acids, results in their accumulation. Indeed, despite muscle tissue is easily able to oxidize BCAAs to support its own energy requirements, aromatic amino acids as well as sulfur-containing amino acids such as taurine, cysteine, and methionine are not as easily metabolized, and may account for the increase in the levels of tyrosine seen during sepsis (Freund et al., 1978). It has also been reported that as disease severity progresses, there is a significant increase in phenylalanine serum concentrations (Martínez-Gómez et al., 2022). Taken together with our results, these findings support the idea that these aromatic amino acids could be used as biomarkers of COVID-19 severity.

Additionally, succinate was found significantly depleted in COVID-19 patients. This metabolite plays a key role in hypoxia,

TABLE 2 Metabolite relative concentrations in COVID-19 and non-COVID-19 patients. Variations with *p*-values <0.05 are indicated with bold numbers.

Metabolite	COVID-19	Non-COVID-19	Fold change[a]	*p*-value
Alanine	1.842 ± 0.205	1.306 ± 0.304	−1.41	**0.001**
Asparagine	0.145 ± 0.047	0.070 ± 0.030	−2.07	**0.002**
β-Hydroxybutyrate	0.246 ± 0.116	0.205 ± 0.063	−1.20	0.217
Betaine	0.395 ± 0.330	0.330 ± 0.211	−1.20	0.327
Choline	3.262 ± 0.808	5.341 ± 1.662	1.64	**0.008**
Creatine	0.655 ± 0.236	0.523 ± 0.214	−1.25	0.139
Glucose	0.269 ± 0.090	0.679 ± 0.561	2.52	0.081
Glutamate	2.309 ± 0.426	2.465 ± 0.488	1.07	0.263
Glycine	1.645 ± 0.229	1.202 ± 0.207	−1.37	**0.001**
Glycerol-3-phosphate	0.027 ± 0.004	0.036 ± 0.006	1.33	**0.002**
Histidine	0.129 ± 0.123	0.040 ± 0.014	−3.25	0.052
Isoleucine	1.508 ± 2.527	0.400 ± 0.132	−3.77	0.128
Lactate	15.958 ± 4.933	17.677 ± 3.053	1.11	0.214
Methionine	0.212 ± 0.116	0.109 ± 0.038	−1.94	**0.021**
Phenylalanine	0.564 ± 0.254	0.272 ± 0.067	−2.07	**0.007**
Phosphocholine	1.050 ± 0.342	1.010 ± 0.250	−1.04	0.400
Succinate	0.006 ± 0.002	0.013 ± 0.004	2.17	**0.002**
Tyrosine	0.415 ± 0.134	0.180 ± 0.052	−2.31	**0.001**
Tryptophan	0.054 ± 0.026	0.028 ± 0.003	−1.92	**0.011**
Uracil	0.074 ± 0.018	0.059 ± 0.022	−1.25	0.102
Valine	1.480 ± 0.677	0.723 ± 0.246	−2.05	**0.008**

[a]Fold changes were computed according to the guidelines of Vinaixa and coworkers (Vinaixa et al., 2012).

where it acts inhibiting the prolyl hydroxylase domain-containing enzymes (PHD) (Yang et al., 2012). Under normal oxygenation, PHD constantly degrades the hypoxia-inducible transcription factor (HIF). This O_2-sensitive factor mediates the response to hypoxia through the expression of genes that regulate cellular energy production, biosynthesis, cell growth, and redox homeostasis (Yang et al., 2014). In our cohort of severe COVID-19 patients lower initial PAFI scores were observed, indicating decreased blood oxygenation (Yang et al., 2012). While increased succinate levels would be expected in this scenario, it is known that mechanical ventilation periods like the ones experienced by our patients lead to succinate downregulation (Mussap and Fanos, 2021). As previously reported, these results indicate that despite high sensitivity, changes in succinate levels are not suitable indicators of disease severity or patient prognosis (Mussap and Fanos, 2021).

Choline levels were also found to be significantly lower in COVID-19 samples. This has also been reported in serum from severe COVID-19 patients, where an increase in the consumption of this trimethylamine caused by activation of macrophage innate immune receptors was linked to extracellular cytokine secretion (Sanchez-Lopez et al., 2019). The presence of pro-inflammatory components in bronchoalveolar lavage fluid is elevated even in severe COVID-19 patients treated with glucocorticoids, suggesting that slowing down the cytokine storm is a critical strategy for disease control (Barberis et al., 2020).

Similarly, we found a significant drop in glycerol-3-phosphate levels among COVID-19 samples. This phosphorylated polyol is tightly related to phospholipid metabolism, which is now known to be deregulated in COVID-19 patients based on serum metabolomic analyses (Shen et al., 2020; Shi et al., 2021). More importantly, it has been reported that the decrease in the levels of this species are directly related to severity in COVID-19 patients (Wu et al., 2020). Although the reduction in glycerol-3-phosphate concentration at the site of SARS-CoV-2 infection warrants further investigation, our results corroborate that this metabolite could be considered as a biomarker of severe manifestations of the disease.

Finally, lactate was the most widely expressed metabolite across both cohorts with no statistically significant differences between them. This finding is consistent with the known fact that

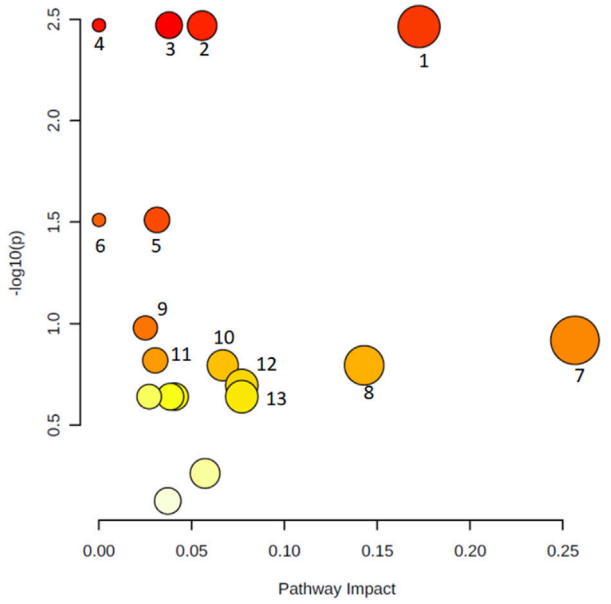


FIGURE 3
Metabolic pathway analysis of the set of metabolites identified in lung autopsies of COVID-19 patients and other fatal pneumonias. Y-axis represents the statistical *p*-values from enrichment analysis, and the X-axis represents the pathway impact value calculated from pathway topology analysis. The node colors represent the *p*-values (lowest: light yellow; highest: dark red) and the node radius indicate the pathway impact values. Dysregulated metabolic pathways include aminoacyl-tRNA biosynthesis (1), pantothenate and CoA biosynthesis (2), valine, leucine, and isoleucine degradation (3), valine, leucine, and isoleucine biosynthesis (4), alanine, aspartate, and glutamate metabolism (5), selenocompound metabolism (6), glycine, serine, and threonine metabolism (7), synthesis and degradation of ketone bodies (8), arginine and proline metabolism (9), butanoate metabolism (10), Cysteine and methionine metabolism (11), Glycerophospholipid metabolism (12), and glyoxylate and dicarboxylate metabolism (13).

high plasma lactate concentration is a marker of poor prognosis and an indicative of metabolic acidosis in critically ill patients, and was expected to be higher in both groups (Martha et al., 2021).

In conclusion, distinct metabolic signatures associated with energy metabolism and inflammatory pathways differentiate COVID-19 from fatal pneumonias caused by other respiratory infections. In particular, we found a significant increase in the levels of branched-chain, aromatic, and sulfur-containing amino acids in lung tissue from fatal COVID-19 cases. Many of these have been recognized as sepsis and inflammatory markers and are associated with lung injury, a condition that commonly leads to severe refractory hypoxemia and is one of the main causes of mortality in COVID-19 patients (Dhont et al., 2020; Donina, 2022; Ribeiro et al., 2022).

To our knowledge, this is the first comparative metabolomic study employing lung tissue samples from COVID-19 patients. In spite of the heterogeneity and wide range of symptoms observed, our findings provide additional insights into the pathogenesis of COVID-19 and have helped identify potential biomarkers for disease severity and treatment efficacy. Notwithstanding, the nature of the samples led to small cohorts affected differently by comorbidities. Some of these, such as diabetes, could have a sizable impact on the metabolic pathways identified as altered in our analyses (Felig et al., 1977). Therefore, the preliminary results reported in this work should be further corroborated in larger scale studies.

Data availability statement

The dataset presented in this study can be found at Mendeley Data with doi: 10.17632/bg45mx8rxd.1 (https://data.mendeley.com/datasets/bg45mx8rxd).

Ethics statement

The study was approved by the Research Bioethics and Ethics Committee from the Hospital Español, a state-managed public hospital which served as the principal COVID-19 reference in Uruguay during the pandemic. An informed consent form was signed by direct family members before autopsy.

Author contributions

JH: Data curation, Formal Analysis, Investigation, Methodology, Writing–original draft. AL-R: Data curation, Formal Analysis, Investigation, Methodology, Writing–review and editing. JLI-G: Conceptualization, Formal Analysis, Methodology, Writing–review and editing. FR: Investigation, Writing–review and editing. GM: Conceptualization, Data curation, Methodology, Writing–review and editing. GG: Funding acquisition, Project administration, Writing–review and editing. NN: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Writing–original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the Agencia Nacional de Investigación e Innovación (ANII, award FSS_X_2019_1_155219), the Fondo para la Convergencia Estructural del MERCOSUR (FOCEM), and the Ministerio de Ciencia e Innovación (MCIN, award PID 2019-10656RJ-I00). Additional financial support was received from the Programa de Desarrollo de las Ciencias Básicas (PEDECIBA).

Acknowledgments

We wish to thank Prof. Gonzalo Moratorio for his recommendations regarding the preparation of lung parenchyma extracts from COVID-19 patients suitable for NMR experiments.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1295216/full#supplementary-material

References

Abate, S. M., Ali, S. A., Mantfardo, B., and Basu, B. (2020). Rate of intensive care unit admission and outcomes among patients with coronavirus: a systematic review and Meta-analysis. *PLoS One* 15, e0235653. doi:10.1371/journal.pone.0235653

Ahlström, B., Frithiof, R., Hultström, M., Larsson, I. M., Strandberg, G., and Lipcsey, M. (2021). The Swedish COVID-19 intensive care cohort: risk factors of ICU admission and ICU mortality. *Acta Anaesthesiol. Scand.* 65 (4), 525–533. doi:10.1111/aas.13781

Ahn, S., Lee, S. H., Chung, K. Y., Ku, N. S., Hyun, Y., Chun, S., et al. (2021). Development and validation of a novel sepsis biomarker based on amino acid profiling. *Clin. Nutr.* 40 (6), 3668–3676. doi:10.1016/j.clnu.2021.05.008

Aittokallio, T., and Schwikowski, B. (2006). Graph-based methods for analysing networks in cell biology. *Briefings Bioinforma.* 7 (3), 243–255. doi:10.1093/bib/bbl022

Ansone, L., Briviba, M., Silamikelis, I., Terentjeva, A., Perkons, I., Birzniece, L., et al. (2021). Amino acid metabolism is significantly altered at the time of admission in hospital for severe COVID-19 patients: findings from longitudinal targeted metabolomics analysis. *Microbiol. Spectr.* 9 (3), e0033821. doi:10.1128/spectrum.00338-21

Atila, A., Alay, H., Yaman, M. E., Akman, T. C., Cadirci, E., Bayrak, B., et al. (2021). The serum amino acid profile in COVID-19. *Amino acids* 53 (10), 1569–1588. doi:10.1007/s00726-021-03081-w

Barberis, E., Amede, E., Khoso, S., Castello, L., Sainaghi, P. P., Bellan, M., et al. (2021). Metabolomics diagnosis of COVID-19 from exhaled breath condensate. *Metabolites* 11 (12), 847. doi:10.3390/metabo11120847

Barberis, E., Timo, S., Amede, E., Vanella, V. V., Puricelli, C., Cappellano, G., et al. (2020). Large-scale plasma analysis revealed new mechanisms and molecules associated with the host response to SARS-CoV-2. *Int. J. Mol. Sci.* 21 (22), 8623. doi:10.3390/ijms21228623

Bastos, L. S., Ranzani, O. T., Souza, T. M. L., Hamacher, S., and Bozza, F. A. (2021). COVID-19 hospital admissions: Brazil's first and second waves compared', *Lancet Respir. Med.* 9, e82-e83. doi:10.1016/S2213-2600(21)00287-3

Bernatchez, J. A., and McCall, L. (2020). Insights gained into respiratory infection pathogenesis using lung tissue metabolomics. *PLOS Pathog.* 16 (7), e1008662. doi:10.1371/journal.ppat.1008662

Booth, A., Reed, A. B., Ponzo, S., Yassaee, A., Aral, M., Plans, D., et al. (2021). Population risk factors for severe disease and mortality in COVID-19: a global systematic review and meta-analysis. *PLOS ONE* 16 (3), e0247461. doi:10.1371/journal.pone.0247461

Brosseau, L. M., Escandón-Vargas, K., Ulrich, A. K., Rasmussen, A. L., Roy, C. J., Bix, G. J., et al. (2021). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) dose, infection, and disease outcomes for coronavirus disease 2019 (COVID-19): a review. *Clin. Infect. Dis.* 75 (1), e1195–e1201. doi:10.1093/cid/ciab903

Chavez, S., Long, B., Koyfman, A., and Liang, S. Y. (2021). Coronavirus Disease (COVID-19): a primer for emergency physicians. *Am. J. Emerg. Med.* 44, 220–229. doi:10.1016/j.ajem.2020.03.036

Chen, W., Yao, M., Chen, M., Ou, Z., Yang, Q., He, Y., et al. (2022). Using an untargeted metabolomics approach to analyze serum metabolites in COVID-19 patients with nucleic acid turning negative. *Front. Pharmacol.* 13, 964037. doi:10.3389/fphar.2022.964037

Chong, J., Wishart, D. S., and Xia, J. (2019). Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis', *Curr. Protoc. Bioinforma.* 68, e86. doi:10.1002/cpbi.86

Cloarec, O., Dumas, M., Craig, A. W., Barton, R., Trygg, J., Hudson, J., et al. (2005). Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic ^{1}H NMR data sets. *Anal. Chem.* 77 (5), 1282–1289. doi:10.1021/ac048630x

Dhont, S., Derom, E., Van Braeckel, E., Depuydt, P., and Lambrecht, B. N. (2020). The pathophysiology of 'happy' hypoxemia in COVID-19. *Respir. Res.* 21 (1), 198. doi:10.1186/s12931-020-01462-5

Dona, A. C., Jiménez, B., Schäfer, H., Humpfer, E., Spraul, M., Lewis, M. R., et al. (2014). Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal. Chem.* 86 (19), 9887–9894. doi:10.1021/ac5025039

Dongelmans, D. A., Termorshuizen, F., Brinkman, S., Bakhshi-Raiez, F., Arbous, M. S., De Lange, D. W., et al. (2022). Characteristics and outcome of COVID-19 patients admitted to the ICU: a nationwide cohort study on the comparison between the first and the consecutive upsurges of the second wave of the COVID-19 pandemic in The Netherlands. *Ann. Intensive Care* 12 (1), 5. doi:10.1186/s13613-021-00978-3

Donina, Z. A. (2022). Causes of hypoxemia in COVID-19. *J. Evol. Biochem. Phys.* 58, 73–80. doi:10.1134/S0022093022010070

Ejaz, H., Alsrhani, A., Zafar, A., Javed, H., Junaid, K., Abdalla, A. E., et al. (2020). COVID-19 and comorbidities: deleterious impact on infected patients. *J. Infect. Public Health* 13, 1833–1839. doi:10.1016/j.jiph.2020.07.014

Ekelund, S. (2012). ROC curves—what are they and how are they used? Point of care. *J. Near-patient Test. Technol.* 11 (1), 16–21. doi:10.1097/poc.0b013e318246a642

Elizondo, V., Harkins, G. W., Mabvakure, B., Smidt, S., Zappile, P., Marier, C., et al. (2021). SARS-CoV-2 genomic characterization and clinical manifestation of the COVID-19 outbreak in Uruguay. *Emerg. Microbes Infect.* 10 (1), 51–65. doi:10.1080/22221751.2020.1863747

Estenssoro, E., Loudet, C., Rios, F., Edul, V. S. K., Plotnikow, G., Andrian, M., et al. (2021). Clinical characteristics and outcomes of invasively ventilated patients with COVID-19 in Argentina (SATICOVID): a prospective, multicentre cohort study. *Lancet Respir. Med.* 9 (9), 989–998. doi:10.1016/s2213-2600(21)00229-0

Felig, P., Wahren, L., Sherwin, R., and Palaiologos, G. (1977). Amino acid and protein metabolism in diabetes mellitus. *Arch. Intern Med.* 137 (4), 507–513. doi:10.1001/archinte.1977.03630160069014

Freund, H. R., Ryan, J. A., Jr, and Fischer, J. E. (1978). Amino acid derangements in patients with sepsis: treatment with branched chain amino acid rich infusions. *Ann. Surg.* 188 (3), 423–430. doi:10.1097/00000658-197809000-00017

Gosangi, B., Rubinowitz, A. N., Irugu, D., Gange, C., Bader, A., and Cortopassi, I. (2022). COVID-19 ARDS: a review of imaging features and overview of mechanical ventilation and its complications. *Emerg. Radiol.* 29 (1), 23–34. doi:10.1007/s10140-021-01976-5

GUIAD-COVID-19 (2022). *Estadisticasuy*. Available online: https://guiad-covid.github.io/estadisticasuy.html (accessed on July 26, 2021).

Hasan, M. K., Suleiman, M. M., and Perez-Lopez, A. (2021). Metabolomics in the diagnosis and prognosis of COVID-19. *Front. Genet.* 12, 721556. doi:10.3389/fgene.2021.721556

Hoch, J. C., Baskaran, K., Burr, H., Chin, J., Eghbalnia, H. R., Fujiwara, T., et al. (2023). Biological magnetic resonance data bank. *Nucleic Acids Res.* 51, D368–D376. doi:10.1093/nar/gkac1050

Hummel, M., Meister, R., and Mansmann, U. (2008). GlobalANCOVA: exploration and assessment of gene group effects. *Bioinformatics* 24 (1), 78–85. doi:10.1093/bioinformatics/btm531

Kankainen, M., Gopalacharyulu, P., Holm, L., and Orešič, M. (2011). MPEA-metabolite pathway enrichment analysis. *Bioinformatics* 27 (13), 1878–1879. doi:10.1093/bioinformatics/btr278

Kimhofer, T., Lodge, S., Whiley, L., Gray, N. S., Loo, R. L., Lawler, N. G., et al. (2020). Integrative modeling of quantitative plasma lipoprotein, metabolic, and amino acid data reveals a multiorgan pathological signature of SARS-CoV-2 infection. *J. Proteome Res.* 19 (11), 4442–4454. doi:10.1021/acs.jproteome.0c00519

Koçak Tufan, Z., Kayaaslan, B., and Mer, M. (2021). COVID-19 and sepsis. *Turkish J. Med. Sci.* 51, 3301–3311. doi:10.3906/sag-2108-239

Kurtz, P., Bastos, L. S., Dantas, L. F., Zampieri, F. G., Soares, M., Hamacher, S., et al. (2021). Evolving changes in mortality of 13,301 critically ill adult patients with COVID-

19 over 8 months. *Intensive Care Med.* 47 (5), 538–548. doi:10.1007/s00134-021-06388-0

Liu, Y., Yan, L., Wan, L., Xiang, T., Le, A., Liu, J., et al. (2020). Viral dynamics in mild and severe cases of COVID-19. *Lancet Infect. Dis.* 20 (6), 656–657. doi:10.1016/s1473-3099(20)30232-2

Lorente, J. A., Nin, N., Villa, P., Vasco, D., Miguel-Coello, A. B., Rodriguez, I. R., et al. (2021). Metabolomic diferences between COVID-19 and H1N1 influenza induced ARDS. *Crit. Care* 25 (1), 390. doi:10.1186/s13054-021-03810-3

Martha, J. W., Wibowo, A., and Pranata, R. (2021). Prognostic value of elevated lactate dehydrogenase in patients with COVID-19: a systematic review and meta-analysis. *Postgrad. Med. J.* 98, 422–427. doi:10.1136/postgradmedj-2020-139542

Martínez-Gómez, L. E., Ibarra-González, I., Fernández-Lainez, C., Tusie, T., Moreno-Macías, H., Martinez-Armenta, C., et al. (2022). Metabolic reprogramming in SARS-CoV-2 infection impacts the outcome of COVID-19 patients. *Front. Immunol.* 13, 936106. doi:10.3389/fimmu.2022.936106

Mierzchala-Pasierb, M., Lipinska-Gediga, M., Fleszar, M. G., Lesnik, P., Placzkowska, S., Serek, P., et al. (2020). Altered profiles of serum amino acids in patients with sepsis and septic shock - preliminary findings. *Archives Biochem. biophysics* 691, 108508. doi:10.1016/j.abb.2020.108508

Moreno, P., Moratorio, G., Iraola, G., Fajardo, A., Aldunate, F., Pereira-Gómez, M., et al. (2020). *An effective COVID-19 response in South America: the Uruguayan Conundrum*. medRxiv. doi:10.1101/2020.07.24.20161802

Murali, R., Wanjari, U. R., Mukherjee, A. G., Gopalakrishnan, A. V., Kannampuzha, S., Namachivayam, A., et al. (2023). Crosstalk between COVID-19 infection and kidney diseases: a review on the metabolomic approaches. *Vaccines* 11 (2), 489. doi:10.3390/vaccines11020489

Mussap, M., and Fanos, V. (2021). Could metabolomics drive the fate of COVID-19 pandemic? A narrative review on lights and shadows. *Clin. Chem. Laboratory Med. (CCLM)* 59 (12), 1891–1905. doi:10.1515/cclm-2021-0414

Nakayasu, E. S., Nicora, C. D., Sims, A. C., Burnum-Johnson, K. E., Kim, Y. H., Kyle, J. E., et al. (2016). MPLEx: a robust and universal protocol for single-sample integrative proteomic, metabolomic, and lipidomic analyses. *mSystems* 1 (3), e00043. doi:10.1128/msystems.00043-16

Neinast, M. D., Jang, C., Hui, S., Murashige, D., Chu, Q., Morscher, R. J., et al. (2019). Quantitative analysis of the whole-body metabolic fate of branched-chain amino acids. *Cell Metab.* 29 (2), 417–429. doi:10.1016/j.cmet.2018.10.013

Ni, N., Nandi, S., Kreyssig, A., Goldman, A. I., Mun, E. D., Bud'ko, S. L., et al. (2008). First-order structural phase transition inCaFe2As2. *Phys. Rev. B* 78 (1), 014523. doi:10.1103/physrevb.78.014523

Ozturk, A., Bayraktar, N., Bayraktar, M., Ibrahim, B., Bozok, T., and Resat, C. M. (2022). Evaluation of amino acid profile in serum of patients with Covid-19 for providing a new treatment strategy. *J. Med. Biochem.* 41 (4), 526–533. doi:10.5937/jomb0-37514

Ranzani, O. T., Bastos, L. S., Gelli, J. G. M., Marchesi, J. F., Baião, F. A., Hamacher, S., et al. (2021). Characterisation of the first 250 000 hospital admissions for COVID-19 in Brazil: a retrospective analysis of nationwide data. *Lancet Respir. Med.* 9 (4), 407–418. doi:10.1016/s2213-2600(20)30560-9

Rego, N., Salazar, C., Paz, M. A., Costábile, A., Fajardo, A., Ferrés, I., et al. (2021). Emergence and spread of a B.1.1.28-Derived P.6 lineage with Q675H and Q677H spike mutations in Uruguay. *Viruses* 13 (9), 1801. doi:10.3390/v13091801

Ribeiro, A., Mendonça, M., Sabina Sousa, C., Trigueiro Barbosa, M., and Morais-Almeida, M. (2022). Prevalence, presentation and outcomes of silent hypoxemia in COVID-19. Clinical medicine insights: circulatory, respiratory and pulmonary medicine. *Clin. Med. Insights. Circ. Respir. Pulm. Med.* 16, 11795484221082761. doi:10.1177/11795484221082761

Rosen, H. S., Yoshimura, N. N., Hodgman, J. M., and Fischer, J. E. (1977). Plasma amino acid patterns in hepatic encephalopathy of differing etiology. *Gastroenterology* 72 (3), 483–487. doi:10.1016/s0016-5085(77)80261-8

Rudnick, D. A., and Ebach, D. R. (2004). Tyrosinemia. *Encycl. Gastroenterology* 2004, 538–541. doi:10.1016/b0-12-386860-2/00772-3

Sanchez-Lopez, E., Zhong, Z., Stubelius, A., Sweeney, S. R., Booshehri, L. M., Antonucci, L., et al. (2019). Choline uptake and metabolism modulate macrophage IL-1B and IL-18 production. *Cell Metab.* 29 (6), 1350–1362. doi:10.1016/j.cmet.2019.03.011

Shen, B., Yi, X., Sun, Y., Bi, X., Du, J., Zhang, C., et al. (2020). Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell* 182 (1), 59–72. doi:10.1016/j.cell.2020.05.032

Shi, D., Yan, R., Lv, L., Jiang, H., Lu, Y., Sheng, J., et al. (2021). The serum metabolome of COVID-19 patients is distinctive and predictive. *Metabolism-clinical Exp.* 118, 154739. doi:10.1016/j.metabol.2021.154739

Simundic, A. (2012). Diagnostic accuracy—Part 1: basic concepts sensitivity and specificity, ROC analysis, STARD statement. *J. Near-patient Test. Technol.* 11 (1), 6–8. doi:10.1097/poc.0b013e318246a5d6

Takeshita, H., and Yamamoto, K. (2022). Tryptophan metabolism and COVID-19-induced skeletal muscle damage: is ACE2 a key regulator? *Front. Nutr.* 9, 868845. doi:10.3389/fnut.2022.868845

Trygg, J., Holmes, E., and Lundstedt, T. (2006). Chemometrics in metabonomics. *J. Proteome Res.* 6 (2), 469–479. doi:10.1021/pr060594q

Trygg, J., and Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). *J. Chemom.* 16 (3), 119–128. doi:10.1002/cem.695

Valdés, A., Moreno, L., Rello, S. R., Orduña, A., Bernardo, D., and Cifuentes, A. (2022). Metabolomics study of COVID-19 patients in four different clinical stages. *Sci. Rep.* 12 (1), 1650. doi:10.1038/s41598-022-05667-0

Van Den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., and Van Der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* 7 (1), 142. doi:10.1186/1471-2164-7-142

Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., et al. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* 20 (6), 669–677. doi:10.1016/s1473-3099(20)30243-7

Vinaixa, M., Samino, S., Saez, I., Duran, J., Guinovart, J. J., and Yanes, O. (2012). A guideline to univariate statistical analysis for LC/MS-Based untargeted metabolomics-derived data. *Metabolites* 2, 775–795. doi:10.3390/metabo2040775

Watanabe, A., Takesue, A., Higashi, T., and Nagashima, N. (1979). Serum amino acids in hepatic encephalopathy-effects of branched chain amino acid infusion on serum aminogram. *Acta hepato-gastroenterologica* 26 (5), 346–357.

Wishart, D. S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., et al. (2022). HMDB 5.0: the human metabolome Database for 2022. *Nucleic acids Res.* 50 (D1), D622–D631. doi:10.1093/nar/gkab1062

Wold, S., Esbensen, K. H., and Geladi, P. (1987). Principal component analysis. *Chemom. Intelligent Laboratory Syst.* 2 (1–3), 37–52. doi:10.1016/0169-7439(87)80084-9

Wu, D., Shu, T., Yang, X., Song, J. Y., Zhang, M., Yao, C., et al. (2020). Plasma metabolomic and lipidomic alterations associated with COVID-19. *Natl. Sci. Rev.* 7 (7), 1157–1168. doi:10.1093/nsr/nwaa086

Xia, J., Wishart, D. S., and Valencia, A. (2011). MetPA: a web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* 27 (13), 2342–2344. doi:10.1093/bioinformatics/btq418

Yang, M., Soga, T., Pollard, P. J., and Adam, J. (2012). The emerging role of fumarate as an oncometabolite. *Front. Oncol.* 2, 85. doi:10.3389/fonc.2012.00085

Yang, M., Su, H., Soga, T., Kranc, K. R., and Pollard, P. J. (2014). Prolyl hydroxylase domain enzymes: important regulators of cancer metabolism. *Hypoxia* 127, 127–142. doi:10.2147/hp.s47968

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 13 May 2024
DOI 10.3389/fmolb.2024.1250413

OPEN ACCESS

EDITED BY
Pablo Hoijemberg,
Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

REVIEWED BY
Paniz Jasbi,
Arizona State University, United States
Stefano Cacciatore,
International Centre for Genetic Engineering and Biotechnology (ICGEB), South Africa

*CORRESPONDENCE
Sara Londoño-Osorio,
slondonoo1@eafit.edu.co
Laura Sierra-Zapata,
lsierra3@eafit.edu.co

RECEIVED 30 June 2023
ACCEPTED 20 February 2024
PUBLISHED 13 May 2024

CITATION
Londoño-Osorio S, Leon-Carreño L, Cala MP and Sierra-Zapata L (2024), The gut metabolome in a cohort of pregnant and lactating women from Antioquia-Colombia. *Front. Mol. Biosci.* 11:1250413. doi: 10.3389/fmolb.2024.1250413

The gut metabolome in a cohort of pregnant and lactating women from Antioquia-Colombia

Sara Londoño-Osorio[1]*, Lizeth Leon-Carreño[2], Mónica P. Cala[2] and Laura Sierra-Zapata[1]*

[1]CIBIOP Research Group, School of Applied Sciences and Engineering, Universidad EAFIT, Medellín, Colombia, [2]MetCore–Metabolomics Core Facility, Vice-Presidency for Research, Universidad de Los Andes, Bogotá, Colombia

Nutrition during the perinatal period is an essential component of health and one that can severely impact the correct development of a human being and its overall condition, in all the subsequent stages of life. The availability of several compounds, mainly macronutrients and micronutrients, plays a key role in the balanced nutrition of both mother and baby and is a process with direct relation to the gut microbiome. Thus, we hereby refer to the set of small molecules derived from gut microbiome metabolism as the gut metabolome. These continuous processes occurring in the gut of a gestating or lactating mother related to microbial communities and nutrients, can be revealed by metabolomics. In this study, we explore for the first time the gut metabolome of pregnant and lactating women, from our region of Antioquia-Colombia, applying untargeted metabolomics by LC-QTOF-MS, and molecular networking. Regarding the gut metabolome composition of the cohort, we found, key metabolites that can be used as biomarkers of microbiome function, overall metabolic health, dietary intake, pharmacology, and lifestyle. In our cohort, pregnant women evidenced a significantly higher abundance of prostaglandins, alkaloids, corticosteroids, organosilicons, and natural toxins, while in lactating women, lipids stand out. Our results suggest that unveiling the metabolic phenotype of the gut microbiome of an individual, by untargeted metabolomics, allows a broad visualization of the chemical space present in this important niche and enables the recognition of influential indicators of the host's health status and habits, especially of women during this significant perinatal period. This study constitutes the first evidence of the use of untargeted LC-QTOF-MS coupled with molecular networking analysis, of the gut microbiome in a Colombian cohort and establishes a methodology for finding relative abundances of key metabolites, with potential use in nutritional and physiological state assessments, for future personalized health and nutrition practices.

KEYWORDS

gut microbiome, pregnancy, lactation, perinatal nutrition, untargeted metabolomics, molecular networking, gut metabolome

Introduction

All nutrients come from the diet, and diet is one of the most important aspects impacting and modulating health and the gut microbiota. This 'microbial' organ within our guts, and the set of genes it contains, called the microbiome (El Hage et al., 2017) have been extensively studied over the last decade. Several of these studies, now published in prestigious journals, have uncovered that dysbiosis, or an imbalance of the intestinal microbial communities (microbiota) and the decrease in ecological diversity within the gut, are related to gastrointestinal, metabolic, and autoimmune diseases, mental disorders, and even some types of cancer (Derrien and Veiga, 2017; Deng et al., 2021; Zhao et al., 2021; Christovich and Luo, 2022; Horn et al., 2022). Since the gut microbiome has a crucial role in the absorption and metabolism of nutrients, both macro and micro, aiming for a balanced microbial community in the gut, helps maintain the host homeostasis, and builds the intestinal barrier (DAS & Nair, 2019). The presence or absence of specific microbial genera or species has been associated with multiple diseases, most of them, non-communicable ones, such as inflammatory bowel disease, diabetes, obesity, some types of cancer, Parkinson's, and Alzheimer's, among others (Novakovic et al., 2020; Zhang et al., 2020; Bardenhorst et al., 2023). Regarding micronutrient absorption, Hadadi and collaborators (2021) addressed the importance of the gut microbiome for maintaining the balance of the host vitamins and minerals. They also address the micronutrient-microbiome axis as a bidirectional entity, and according to other studies, several micronutrient deficiencies could be positively or negatively associated with the gut microbiota (Hadadi et al., 2021). Another study carried out by Maynard and Weinkove has revealed that certain host microbes, such as *C. elegans* and *E.coli,* play a role in the effective supplementation of micronutrients by the secretion of siderophores (iron and B12), or the uptake and conversion into more readily absorbable derivatives or micronutrients, such is the case of folic acid (Maynard and Weinkove, 2020). Moreover, Bielik and Kolisek (2021) reported the positive effect of probiotics on mineral absorption, stating they are promising due to their ability to modulate the composition and metabolism of the gut microbiota (Bielik and Kolisek, 2021).

On the other hand, the perinatal period is marked by hormonal, immunological, and—especially during the late stages of healthy pregnancies without complications—by inflammatory changes that alter the function and bacterial composition of the mother's gut (Mandal et al., 2016). Estrogen and progesterone also impact this composition through their effect on bacterial metabolism and the increase in abundance of pathogenic bacteria (Edwards et al., 2017). It is also known that the gut microbiota contributes to the regulation of glucose metabolism in pregnancy (Brantsæter et al., 2011). For example, the abundance of the genus *Collinsella* sp. Is positively correlated with circulating insulin, and low dietary fiber intake was associated with a gut microbiota favoring lactate fermentation, while high fiber intake promotes short-chain fatty acid-producing bacteria (Fu et al., 2022). Related to this, low dietary fiber may enable the overgrowth of *Collinsella* sp. and alter the overall fermentation pattern in gut microbiota (Gomez-Arango et al., 2018). This suggests that dietary choices during pregnancy can modify the nutritional ecology of the gut microbiota. Besides, in a study conducted on pregnant women, it was shown that there are significant differences in the relative abundance of several genera in women on a vegetarian diet, specifically a reduction in *Collinsella* sp., *Holdemania* sp., and an increase in the relative abundances of *Roseburia* sp. and Lachnospiraceae sp. (Barrett et al., 2018). The most recent research on gut microbiome during the perinatal period in mice shows that the characteristic microbiota of the third trimester of pregnancy, increases weight gain, insulin resistance, and a greater inflammatory response when transferred to germ-free mice (Koren et al., 2012). Studies in other populations different from the American and European ones, such as those from Latin America, the Caribbean, Asia and African, or from women and children's cohorts, are urgently needed as well as their underlying data (Magne et al., 2016), in order to properly acknowledge the gut microbiome in world-population scale, and be able to develop solutions to improve the health status of the groups belonging to these communities, in need of tools for this purpose.

In the quest for the characterization of generalizable traits of the gut microbiome, metabolomics has appeared as one of the most useful techniques to study it, being defined as a comprehensive analysis of all metabolites in a biological system with their proper identification and quantification (Fiehn, 2002), and is recognized as a powerful top-down systems biology approach, for understanding the genetics-environment-health paradigm and identifying clinically relevant biomarkers (Moco et al., 2013). Metabolomics studies within the gut, which we name here the gut metabolome, have been increasing in the last years due to the strong relationship found between some gut microbiome metabolic pathways and diseases, especially non-communicable ones, and due to the involvement of the gut microbiota in several biochemical functions directly associated with perturbations that can lead to the development of diseases (De Preter et al., 2015). Moreover, the identification and relative quantification of metabolites in these environments can point out lifestyle and dietary habits, and nutrient balance in the gut, which in turn, allows the highlighting of specific disease predispositions (Vernocchi et al., 2012), such as a mineral or vitamin deficiency of (Lai et al., 2022; Wan et al., 2022), an excess of an inflammatory molecule (Zhang et al., 2021), among others. Metabolomics is a technique that can be performed over different biological matrices such as cells, tissues, stool samples, and biofluids such as plasma, saliva, urine, and blood. The sample selection will always depend on the research or clinical question, but biofluids are typically used to identify biomarkers, whereas tissues and cells are used to investigate mechanisms associated with the pathophysiological process (Chetwynd et al., 2017). Regarding human stool samples, which reflect the gut metabolome, most of the published research has focused on characterizing its complex bacterial composition using next-generation microbial DNA sequencing and sophisticated metagenomic techniques. However, a growing number of microbiome researchers are recognizing that considerable information could be gained by using a more integrative approach that also includes comprehensive fecal metabolite analysis (Karu et al., 2018; Haffner et al., 2022).

One of the techniques widely used to study the gut (fecal) metabolome, is liquid chromatography coupled to mass spectrometry (LC-MS), which does not usually include derivatization steps. The technique can be performed in a

targeted or untargeted mode, depending on the experimental design, and multiple approaches can be taken to analyze the raw data, thus allowing the recognition of multiple chemical families and the greater elucidation of the chemical space, phenotype, and nutrients composition of the gut. In this research, we used classical molecular networking and untargeted metabolomics to make a pilot and first approach toward the characterization of the chemical space of the gut microbiota (gut metabolome) of women from Antioquia, Colombia. These women conform to a pilot cohort (n = 23) of pregnant, 7) lactating 9), and reproductive-age women 7) acting as controls. By using LC-QTOF-MS/MS metabolomic techniques and data analysis, we aimed at the identification and quantification of several compounds of nutritional importance for the baby's appropriate development, which are supplied by the mother during these fundamental stages of pregnancy and lactation. As stated above, macronutrients, micronutrients, and derived metabolites play a key role in the balanced nutrition of both mother and baby, and both are intrinsically related to the gut microbiome. Thus, with this pilot study, we wish to contribute to the maternal nutritional body of knowledge in our area of the world since to date, there are no published studies that explore the chemical diversity of the Colombian female population during the mentioned stages, despite these being key interventional periods for nutrition. It is our wish that the knowledge derived from this pilot study and its validation in larger cohorts can help avoid future developmental complexities in an individual during later stages of their lives also avoiding future health complications. Thus, the relevance of this kind of pilot study and as mentioned earlier, the further validation of its preliminary results in larger cohorts is evident, to broaden our knowledge of the gut microbiome chemical space and phenotype in the populations in Colombia, Latin America, and the Caribbean.

Materials and methods

Study cohort and sample collection

A group of twenty-three women volunteers, from Antioquia, Colombia between 23 and 35 years old were enrolled in the study, between August 2020 and May 2021. Nine of them were lactating, seven were pregnant and seven were control group (non-pregnant or lactating women) of reproductive age (Figure 1). Both pregnant and lactating women were enrolled since they complied with a healthy pregnancy/lactation stage, without complications. Average values of the different variables measured for each group of the cohort are detailed in Table 1, as well as detailed information for each volunteer (age, height, weight, pregnancy or *postpartum* week, lipid profile) which was saved as correlated metadata for the study. As inclusion criteria, the selected cohort must declare non-consumption of antibiotics in the past 6 months before the sample collection. Two different samples were taken from each volunteer, a blood sample was collected in collaboration with Abad Laboratory, to measure the lipid profile (low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides, and total cholesterol), and a stool sample was provided. This last sample was processed at Universidad EAFIT, within the next 24 h of collection, in an anaerobic chamber (Vinyl Anaerobic Chamber Type B from Coy Laboratory Products). Briefly, 200 mg of it was homogenized in 1 mL pH 7.2 buffer solution (0.05% K2HPO4, 0.05% KH2PO4, 0.05% MgSO4 x 7H2O, 0.0005% FeSO4 x 7H2O, 0.005% (NH4) 2SO4, 0.1% cysteine, 0.001% resazurin, and 20% glycerol) (Hayashi et al., 2002) and stored at −80°C for any subsequent use.

Sample treatment

Frozen Stool samples were weighted, lyophilized at −80°C for 72 h, and weighed again to determine the removed water content percentage. Then, nitrogen gas was injected for 10 min into each sample to guarantee an inert environment. For extraction, 60 mg of each lyophilized sample was mixed with 300 µL of MeOH and vortex-mixed for 5 min. Subsequently, samples were taken to an ultrasound for 30 min and vortex-mixed again for 5 min. Finally, samples were centrifuged at 180,00x g, 4°C for 15 min and 100 µL of the extract was used for the analysis by LC-QTOF-MS (Cheng et al., 2020).

Metabolomic analysis

Data acquisition for untargeted metabolomics and molecular networking using RP-LC/MS and HILIC-LC/MS

Metabolomics data from fecal samples were acquired using an Agilent Technologies 1,260 Liquid Chromatography system coupled to a 6545 Q-TOF quadrupole time-of-flight mass analyzer with electrospray ionization. For the reversed-phase, 2 µL of the sample was injected into a C18 column (InfinityLab Poroshell 120-EC 100 × 2.1 mm, 1.9 µm) at 40°C. The mobile phases used for elution were composed of 0.1% (v/v) formic acid in Milli-Q water (Phase A) and 0.1% (v/v) formic acid in acetonitrile (Phase B) pumped at 0.4 mL/min with a gradient starting at 5% B, increased at 96% B in 15 min and kept there 1 min and then, at 16.1 min, going back to the initial conditions until 20 min. Detection by mass spectrometry was performed in positive ESI mode in full scan and autoMS/MS from 50 to 1,100 m/*z* and 20eV. Throughout the analysis, two reference masses were used for mass correction: *m/z* 121.0509 $[C_5H_4N_4 + H]^+$, and *m/z* 922.0098 $[C_{18}H_{18}O_6N_3P_3F_{24} + H]^+$, corresponding to protonated purine and protonated hexakis, respectively.

For hydrophilic interaction chromatography, 5 µL of the sample was injected into a HILIC-Z (InfinityLab Poroshell 100 × 2.1 mm, 1.9 µm) column, which was thermostated at 30°C. The elution gradient was composed of 10% (200 mM ammonium formate pH 3): 90% H_2O (Phase A) and 10% (200 mM ammonium formate pH 3): 90% ACN (Phase B) with a constant flow of 0.5 mL/min. The chromatography gradient started at 100% of phase B and decreased to 70% B in 10 min. The starting condition was returned by minute 11 and kept there for 5 min for re-equilibration time. Data were collected in negative mode operated in full scan and MS/MS mode at 20 eV from 50 to 1,100 m/*z*.

Data processing and analysis for untargeted metabolomic analysis approach

The full scan raw data from RP-LC/MS and HILIC-LC/MS was processed using Agilent MassHunter Profinder Software B.08.00.

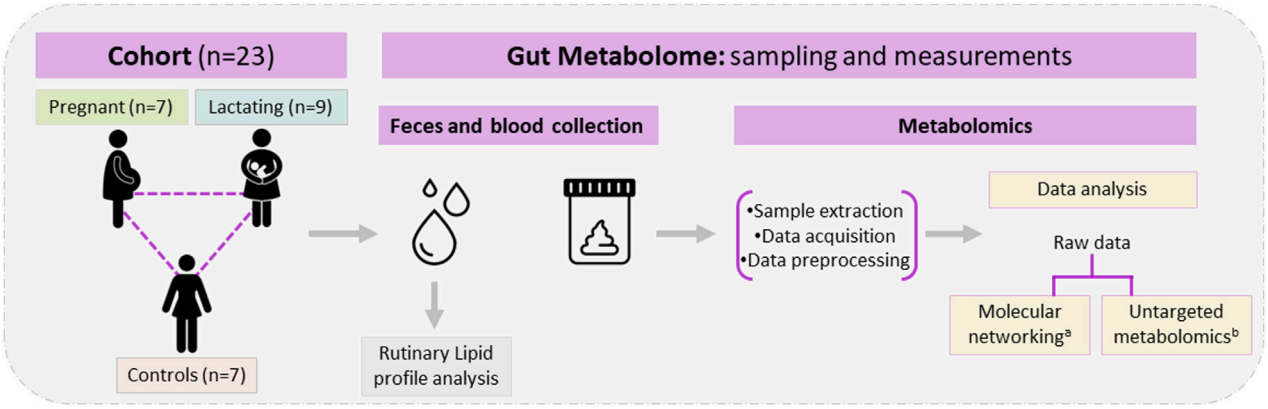


FIGURE 1
Study design, cohort composition, samples used, and metabolomic analysis. [a]**Classical molecular network was performed using GNPS and Cytoscape.** [b]**Untargeted metabolomics used MetaboAnalyst 5.0, MATLAB R2021b, and CEU mass mediator, the libraries Kegg, HMDB, METLIN, lipid maps, and MS-DIAL 4,80 software for metabolite identification.**

TABLE 1 Characteristics of the study cohort. Values of total cholesterol, LDL, HDL, and triglycerides are presented in (mg/dL) units.

	Lactating	Pregnant	Control
Number of volunteers	9	7	7
Age (years)	32.9 ± 2	28.6 ± 4.4	30.4 ± 10.7
BMI	22.5 ± 2.1	25.7 ± 67	21.4 ± 1.2
Total Cholesterol	215.67 ± 55.1	243 ± 50.4*	175.6 ± 32.8*
LDL	129.59 ± 51.8	127.9 ± 39.4	101.5 ± 29.2
HDL	62.3 ± 6.8*	78.9 ± 15.6*	60.5 ± 12.1*
Triglycerides	100.8 ± 49.8*	181 ± 45.1*	68.2 ± 17.3*
Gestational week	0	27.29 ± 3.9	0
***Postartum* week**	13.3 ± 15.9	0	0

Super index *: Triglycerides and HDL (p-value = 0.01 and 0.03 respectively) were statistically different in pregnant women compared to lactating women. Triglycerides had a (p-value = 0.01) for lactating women vs. the control group. Triglycerides, HDL, and total.cholesterol (p-value = 0.002, 0.03, and 0.04 respectively) were statistically significant between the pregnant and control group.

The software uses the Molecular Feature Extraction (MFE) technique and Recursive Feature Extraction algorithms for noise reduction, feature deconvolution, and alignment. The data matrices from each platform were filtered by presence and reproducibility, keeping only the metabolites detected in at least 80% of all stool samples and using a threshold of 20% based on the coefficient of variation (CV) of metabolite levels in the quality controls (QCs). Differences among the groups were explored using both multivariate (MVA) and univariate (UVA) statistical analyses. For MVA, a partial least-squares discriminant analysis PLS-DA model was used for sample classification and to detect differences between the groups using MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/MetaboAnalyst/ModuleView.xhtml). Metabolites with variable importance in projection (VIP) ≥ 1 and a jackknifing confidence interval that did not include zero were considered statistically significant from the PLS-DA models. The univariate analysis employed in this study used the Mann-Whitney U test in MATLAB R2021b to evaluate the significant differences between each metabolite (p-value < 0.05) in the following comparisons: Lactating vs Control, Pregnant vs Lactating, and Pregnant vs Control.

Metabolite identification

To annotate statistically significant metabolites, the CEU Mass Mediator tool (http://ceumass.eps.uspceu.es/) was used, which matches metabolites with libraries, in addition to analyzing their correspondence with the mass spectral library and the generated molecular formula. The databases Kegg (http://genome.jp/keg), HMDB (http://hmdb.ca), METLIN (http://metlin.scripps.edu), and Lipid MAPS (http://lipidmaps.org), as well as the software MS-DIAL 4.80 (http://prime.psc.riken.jp/compms/msdial/main.html) and Agilent MassHunter qualitative analysis software, were also utilized for this purpose. The identification level assigned to each compound was according to the Metabolomics Standards

Initiative (MSI) by Fiehn (Sumner et al., 2007) where level 1 corresponds to the metabolites identified by reference standard, level 2 to those that have MS/MS spectrum match and molecular formula, level 3 with unequivocal molecular formula, and level 4 only with *m/z* database match.

Data processing and analysis for molecular networking approach

For classical molecular networking, raw data (.d files) obtained from the data acquisition with C18 and HILIC columns, were converted into (.mzXML) format using MSconverGUI (Holman et al., 2014). Once the data were confirmed to be reproducible and a separation between groups was observed, the datasets were uploaded to GNPS web platform GNPS–Analyze, Connect, and Network with your Mass Spectrometry Data (ucsd.edu) (Wang et al., 2016) under de massive code MSV000088880 MassIVE Dataset Summary (ucsd.edu) for C18 data, and MSV000089161 MassIVE Dataset Summary (ucsd.edu) for HILIC data. Two classical molecular networks were built to visualize the features present in the samples' chemical space, and clustered by chemical families. In a second layer of information, each feature was classified by color, as being part of either one cohort group, two of them, or being a shared feature across the three groups in the study. Each group (lactating, pregnant, control) had seven volunteers meaning seven different datasets that act as replicates of the chemical space of the said physiological state; we included a fourth group which consisted of a mix of pure standards of dietary choline derivatives as a control for this specific micronutrient, highly important during pregnancy and lactation. Several of choline's biochemical route derivatives in the gut microbiome were included, these being acetylcholine, betaine, phosphatidylcholine, choline chloride, and trimethylamine. The network parameters set in the GNPS platform were (Min pairs cosine: 0.75, Min fragmented ions: 0.6, Min matched peaks: 6, Cluster size: 2, Analog search: do search). Then, the generated molecular networks were exported to Cytoscape (Ideker, 2003), following manual annotation and curation of the clusters.

Network curation and annotation

This procedure was followed as proposed by Sierra-Zapata et al. (2020). The total features table was exported from Cytoscape as (.csv) file to analyze the abundances of each feature based on the spectral count and the identification provided by the platform (GNPS) for each feature. A query was used to extract the nodes information of each sub-network (Supplementary Tables S1, S2) and based on the library hit found for a feature through GNPS, we assigned a chemical family name to each sub-network, by looking at the metabolite's functionality in PubChem. When non-conclusive, a search in ChemSpider and the Human metabolome database was done as well. In the cases where a unique node from the sub-network was annotated, the entire sub-network was labeled by the same chemical family, and when different nodes were identified, the family name was given following the functionality that grouped all of them. This is done in accordance with the algorithm of GNPS, where a single node's annotation, can be propagated to its neighboring nodes connected by edges, given structural similarity clustering (Wang et al., 2016). In Cytoscope, the nodes were colored according to their presence in each group of the cohort: light blue for the lactating group, dark green for the pregnant group, orange for the control group, red for standard metabolites, and purple for the group of metabolites present in both lactating and pregnant women. This network was exported in (.pdf) format with the precursor mass available as a label on the nodes, and the chemical family was then added as a circle grouping the cluster of nodes. Given the family name and its abundance (in numbers of spectra) among the treatments, the relative abundance for each chemical family was calculated in each group of the cohort to see any statistical difference (Supplementary Material S1, S2). Also, a PCA was performed into MetaboAnalyst using the raw data to visualize any clustering of the chemical space of the cohort's groups.

Results

Untargeted metabolomics analysis by RP-LC/MS and HILIC-LC/MS

The cohort of volunteers and their characteristics, from where the data were obtained, can be revised in Figure 1 and Supplementary Datasheet S1. First, untargeted metabolomics analysis of the data acquired from the volunteers' samples, according to the methods described above, was performed. A multivariate analysis (MVA) was made using PLS-DA (Figure 2) to compare data from Lactating vs. Control, Pregnant vs. Control, and Lactating vs. Pregnant treatments. The PLS-DA model shows values of R^2 ranging from 0.95 to 0.99 and Q^2 from 0.21 to 0.40 indicating a clear separation between the comparison of features acquired by RP-LC/MS and HILIC-LC/MS and thus can be considered a good feature selector model. Then, a univariate analysis (UVA) was conducted to identify the differential metabolites between the proposed comparisons. A total of 200 differential molecular features were identified in both platforms through UVA and MVA analyses, considering those with a $p\text{-}value < 0.05$ or VIP>1. Among them, 85 metabolites were statistically different when comparing lactating and control groups, 67 metabolites when comparing pregnant and control groups, and 48 metabolites between pregnant and lactating groups. Supplementary Table S3 shows the metabolites that were detected as up or downregulated among the groups of the study, including the significance metrics provided by the MVA and UVA (VIP and *p*-value).

For pregnant women compared to lactating ones, it has been found that piperine, benzenoids, hydroxypregnene, glycerophosphoserines, glycerophosphates, deoxyinosine, prostaglandins, biotin, and steroids (Figure 3A, Supplementary Table S3) Pregnant vs. Lactating) were upregulated. Specifically, hydroxypregnene, deoxyinosine, prostaglandins, and steroids were detected as diminished in lactating women vs controls, thus being differentially detected in the guts of pregnant, lactating, and women of reproductive age.

Then, when comparing pregnant women with non-pregnant nor lactating controls (Figure 3B) increased levels of alkaloids, bile acids, carbohydrates, corticosteroid hormones, some fatty acids, glycerophosphocholines, glycerophosphoserines, glycerophosphates, sphingolipids, and sterols were found. Meanwhile,

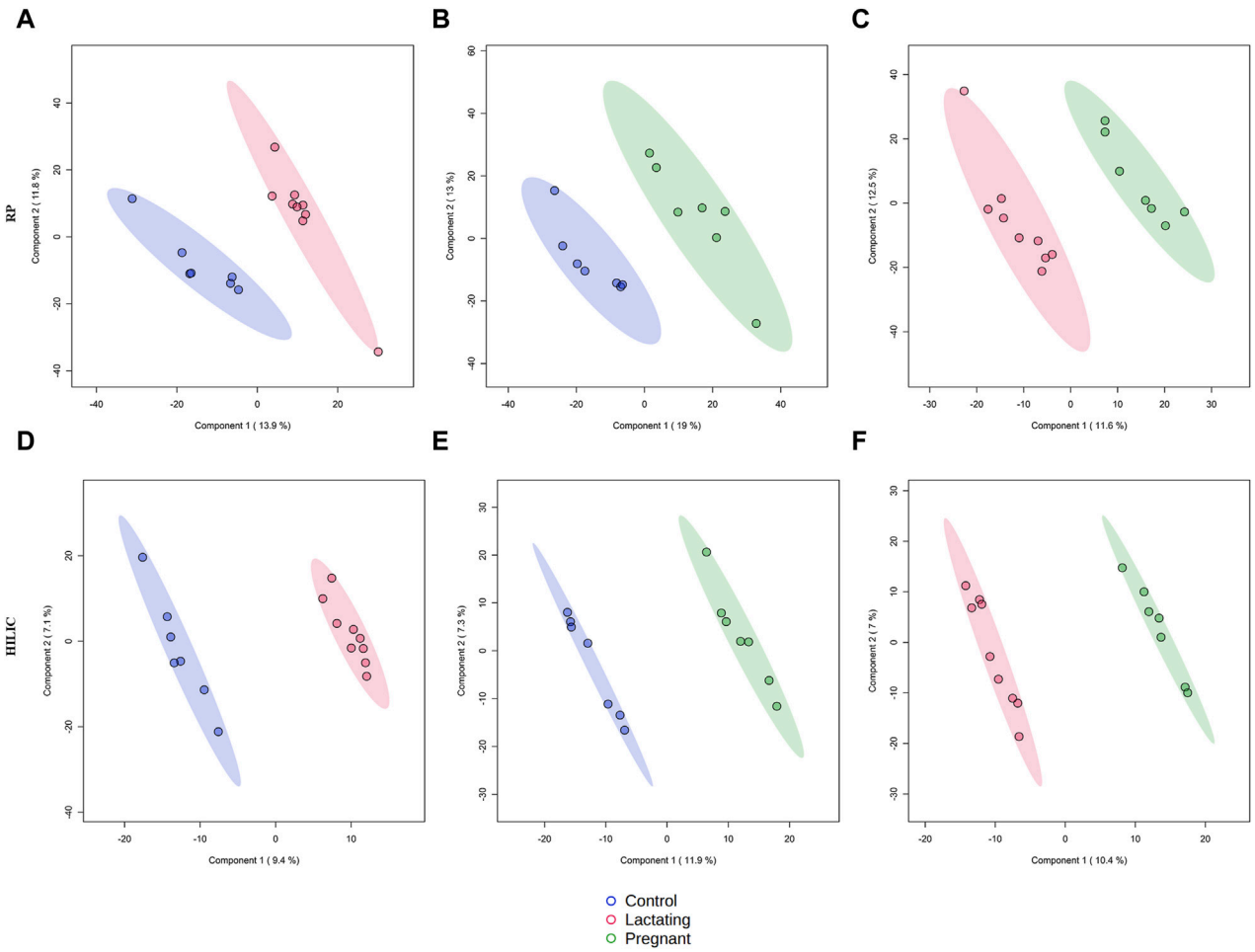


FIGURE 2
Supervised PLS-DA models for metabolomics by reverse-phase (RP) in positive mode and Hydrophilic interaction chromatography (HILIC) in negative mode. **(A)** R^2:0.95404, Q^2:0.21976; **(B)** R^2: 0.956, Q^2: 0.29733; **(C)** R^2: 0.9692, Q^2: 0.12499; **(D)** R^2: 0.99479, Q^2: 0.37436; **(E)** R^2: 0.99663, Q^2: 0.27773; **(F)** R^2: 0.99622, Q^2: 0.40789.

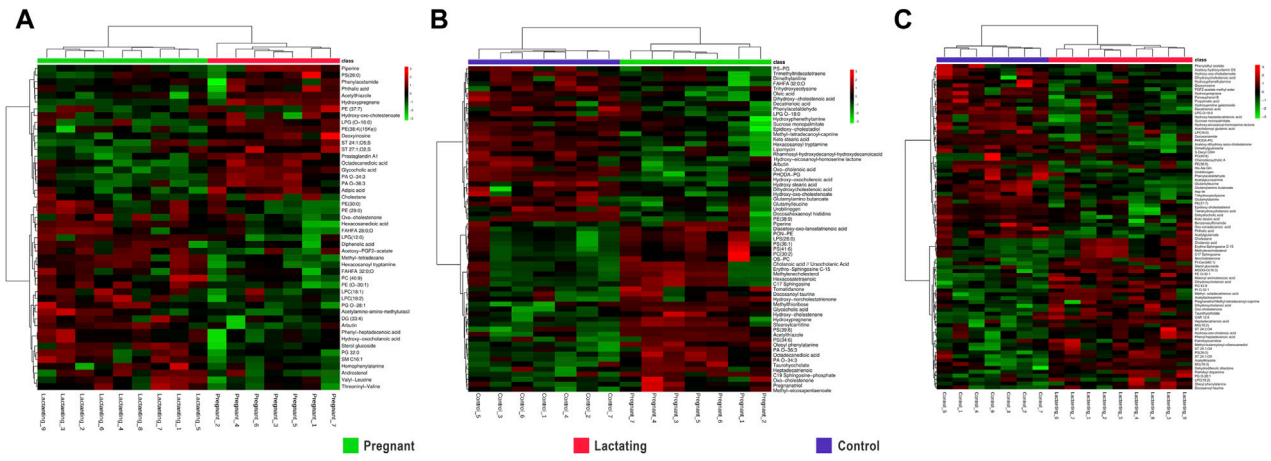


FIGURE 3
Hierarchical Clustering Heatmaps of metabolome data **(A)**. Pregnant vs. Lactating **(B)**. Pregnant vs. Control **(C)**. Lactating vs. control.

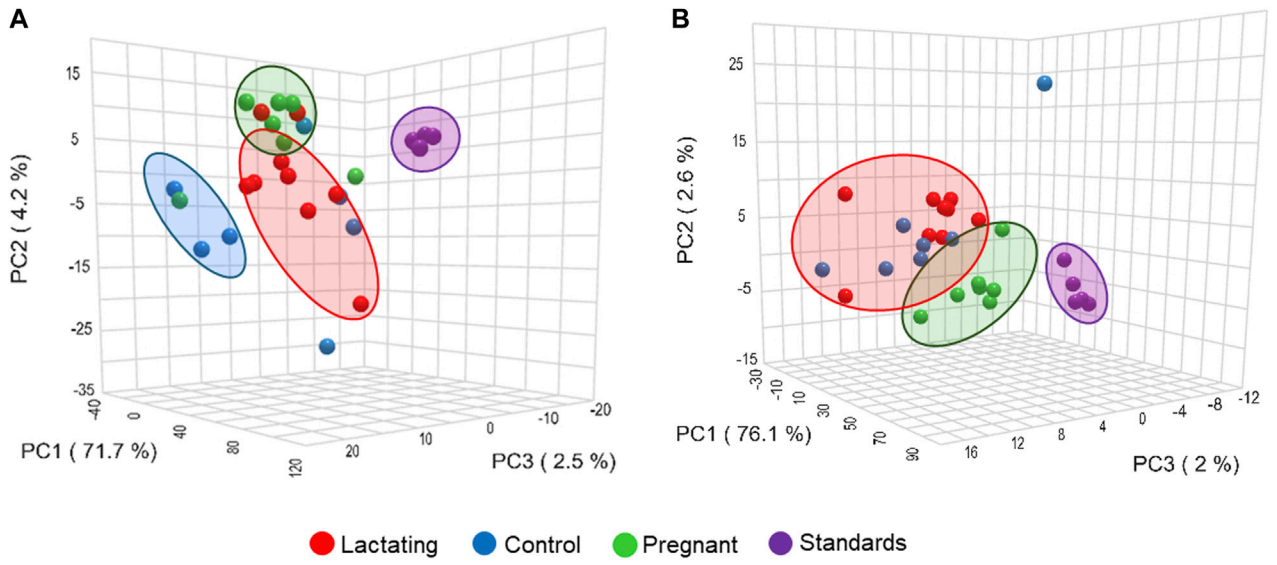


FIGURE 4
Principal component analysis (PCA) from LC-MS/MS raw data, **(A)** using C18 column and **(B)** HILIC column.

glycerophosphoglycerols, steroids, and 53% of the total fatty acids found showed a decrease in the pregnant group. Specifically, glycerophosphocholines and glycerophosphoserines, corticosteroid hormones, bile acids, fatty acids, carbohydrates, and sterols are increased both in lactating and pregnant women gut metabolome when compared to control women in the cohort, as found by this methodology of untargeted metabolomics.

Importantly, for lactating women compared to non-pregnant nor lactating women of reproductive age, it has been found that amines, phthalic acid, urobilinogen, acetylglucosamine, corticosteroid hormones such as hydroxypregnene, fatty amides, glycerophosphoglycerols, glycerophosphoethanolamines, glycerophosphoinositol, prostaglandins, peptides and proteins, polyketides, steroids, and vitamin D were mostly downregulated in a range of 0.1 to 0.7 fold change (Figure 3C; Lactating vs. Control). On the other hand, in the lactating group, there were also notable upregulations compared to controls observed in various chemical families, such as bile acids, carnitines, ceramides, glycerolipids, glycerophosphocholines, glycerophosphoserines, and palmitoyl dopamine. Furthermore, amino acids and derivates showed a 40% increase, as did carbohydrates (50%), benzoic acids (67%), fatty acids (63%), corticosteroid hormones (67%), and sterols (80%).

Chemical space defined by molecular networking

Before analyzing the data by molecular networking, a principal component analysis was performed by MetaboAnalyst (Xia et al., 2009), evidencing that for the HILIC platform, samples from the pregnant group clustered together and correlated (Figure 4B), separating themselves from the other cluster of control and lactating groups of volunteers. However, in the analysis by RP-LC grouping is not as evident as in the data obtained by HILIC, although a distinction is still observed between the volunteers in each group (control vs lactating vs pregnant ones, Figure 4A). When running the classical molecular network at the GNPS platform, we obtained 382 annotated metabolites (nodes or features) out of 1,583 (24% of the chemical space identified), for the C18 column, and 118 out of 465 (~25% of the chemical space identified) for the HILIC column. A chemical family was assigned as the name to each sub-network that had at least one annotated metabolite, getting a total of 32 chemical families for the C18 column (Figure 5), and the relative abundance compared across groups of the study of the most biologically significant of them (20) is shown in (Figure 7A).

Among the chemical families identified, we observed the following in concordance with the untargeted metabolomics approach (results presented in the section above): glucosylceramides, sphingolipids, bilirubin metabolism, phosphocholine and derivatives, indole acetic acid and related hormones, glucosyIceramides, diacylglycerol, primary bile acids, amino acid (tryptophan), carnitine derivatives, omega-6, omega-3 fatty acids and derivatives. Of these, carnitine derivatives, ceramides, lysophospholipids, phosphocholine derivatives, secondary bile acids, and tryptophan were in higher abundance in lactating women. Nevertheless, we also observed a larger identification of phytonutrients or plant-derived metabolites in the gut metabolome of the cohort such as monoterpenoids, terpenoids, and aromatic compounds from plants, vitamin E related compounds, coumarin-related compounds, caffeic acid and from rosemary plant; as well as pharmaceutical molecules (glucocorticoids and allergic conditions treatment, human aspirin metabolism, cardiac glycoside, anticoagulant warfarin) when using this molecular networking untargeted metabolomics approach. We also detected compounds with a broad classification as organosilicons and other natural toxins, in the gut metabolome of the cohort, in a significantly higher abundance in control women (of reproductive age).

For the HILIC column, 11 chemical families were annotated (Figure 6) and the relative abundances of all of them are shown in

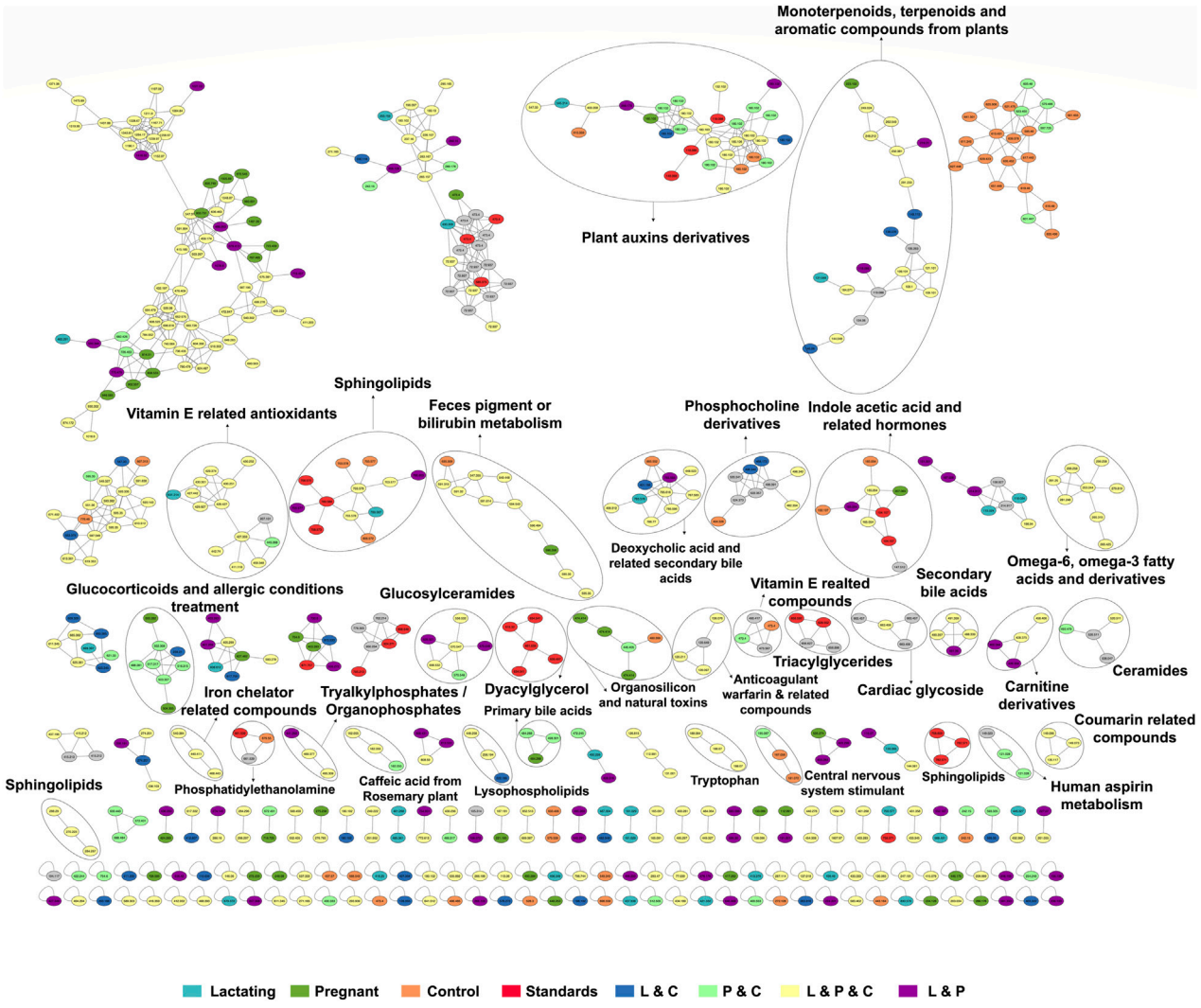


FIGURE 5
The molecular network created in GNPS for LC-MS/MS data acquired using the C18 column. Each node represents a single metabolite precursor mass (feature) and the color indicates the physiological state where each metabolite was found. Lactating (light blue), pregnant (green), control (orange), lactating and pregnant (purple), lactating and control (dark blue), pregnant and control (light green), lactating, pregnant and control (yellow), and the standards (red). Grey nodes denote the confluence of that feature in all groups within the study.

(Figure 7B). The HILIC column, as a method able to detect polar compounds, allowed us to identify the following chemical families: betaine, phytonutrients as lignans and neolignans from plants, gIycoside and lipids derivatives from plant food sources, plant polyphenols, raffinose trisaccharides, glucosinolates and lipopolysaccharides which are proven prebiotic compounds (Zhang et al., 2022), glutamic acid derivatives, benzene derivatives, sphingolipids.

Discussion

The gut microbiome of humans is estimated to comprise around 45 million non-redundant genes (Sender et al., 2016; Tierney et al., 2019). When compared to the human genome and its approximately 20.000 genes, the microbiota exceeds this capacity more than 1,000 times, evidencing its profound potential to influence the biochemical environment of the host (Lee-Sarwar et al., 2020). More surprisingly, the Human Microbiome Project discovery of metabolic pathway abundances in the gut is relatively consistent across populations, while taxonomic composition varies between individuals. This denotes that a core set of conserved pathways is associated with microbial genes, but their abundance varies depending on the taxonomic composition of this environment (Huttenhower et al., 2012). Besides, microbial functions are closely reflected by the composition of the metabolome, or better said, the collection of small molecules present in a sample. Although the human body houses many discrete microbiomes and metabolomes, the gut is taxonomically the most diverse and largest site (Thursby and Juge, 2017). Gathering the ideas exposed before, gut metabolome studies are now being considered the next Frontier to unveil the gut microbiome and are becoming prevalent in studies concerning this so-called organ, since they reflect the phenotype of the individual and thus, provide a

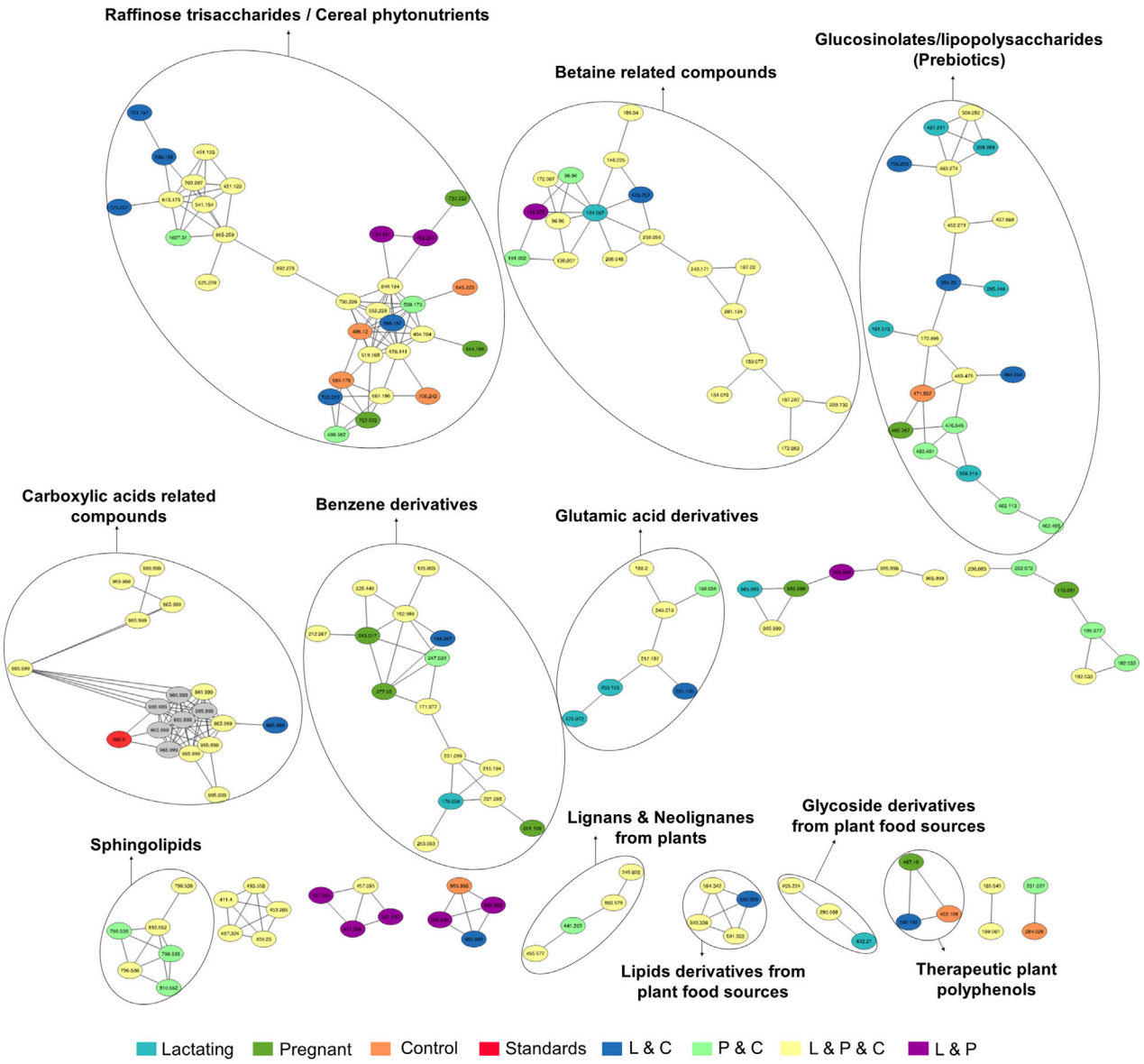


FIGURE 6
Molecular network created in GNPS for LC-MS/MS data acquired using HILIC column. Each node represents a single metabolite mass and the color indicates the physiological state where each metabolite was found. Lactating (light blue), pregnant (green), control (orange), lactating and pregnant (purple), lactating and control (dark blue), pregnant and control (light green), lactating, pregnant and control (yellow), and the standards (red). Grey nodes denote the confluence of that feature in all groups within the study.

more accurate perspective of the biochemical and metabolic processes taking place in this environment.

In this respect, our study offers various conclusions, some of them new and others reinforcing previous evidence from the scientific community on the gut microbiome and its associated metabolome. First, the gut metabolome or chemical space associated with the gut is a rich environment where important biomarkers of health can be detected (Figures 3–7). This is important in the way that the same metabolites are not always detected in serum metabolomics (Dhakan et al., 2019; Wen et al., 2020). Second, although our cohort is small (n = 23), the gut metabolome of pregnant, lactating, and women of reproductive age from our region of Antioquia (Colombia), evidences structural differences between groups both in its composition and relative abundance, denoting a plausible different core composition of microbial and host metabolism (Figures 2–6). In the case of microbial metabolism, these differences can be attributed to the differential taxonomic communities associated with each group of women according to their physiological state, as stated before (Koren et al., 2012). Specifically in this study, our findings suggest that, for the physiological stages of pregnancy and lactation, metabolites related to fats mobilization and membrane formation such as glycerophosphocholines, glycerophosphoserines, and fatty acids; hormones (corticosteroids); bile acids; carbohydrates and sterols are increased, being significantly overexpressed both in lactating

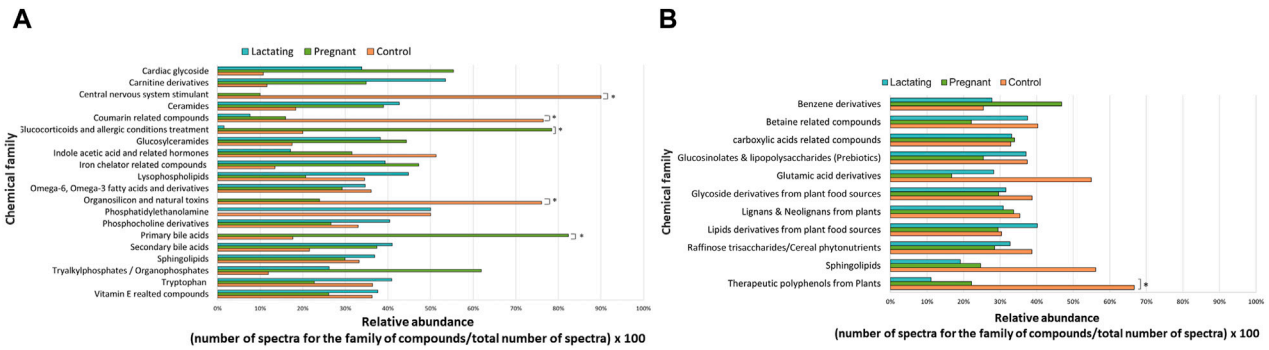


FIGURE 7
Relative abundance of the chemical families among the groups, detected by C18 **(A)** and HILIC **(B)** column. Asterisks represent those chemical families that are significantly different between groups (fold change ratio >1.8).

and pregnant women's gut metabolome (Figures 3–7). This result is expected since, specifically for lipids metabolism, previously published findings reporting multiple physiological changes that occur in healthy, gestating women, which contribute to the alterations in lipid profiles, mainly to support the developing fetus to whom cholesterol and essential fatty acids are essential for normal development (Wild and Feingold, 2000). Also, larger doses of foods rich in healthy fats are needed to meet the metabolic demand, especially for the nutrient choline, which is highly available in fats (Zeisel, 2013). Our findings, also correlate with a transformation of the gut microbiota into a proinflammatory immune state as pregnancy progresses (Koren et al., 2012; Trevisanuto et al., 2013), since prostaglandins, a main biomarker of both the promotion and resolution of inflammation (Ricciotti and Fitzgerald, 2011), are increased in this last group (Figure 3A). Also, corticosteroids such as hydroxypregnene are decreased in the lactating stage but increased in pregnancy, which could be an indication of prescription of corticosteroids to treat symptoms of autoimmune conditions or of inflammation, as well as being one of the most important antenatal therapies available to improve newborn outcomes before anticipated preterm birth (El-Sayed et al., 2017). Also, interestingly, glucocorticoid compounds with anti-inflammatory and immunosuppressive effects are commonly used to treat inflammatory bowel disease, asthma, allergies, and rheumatic diseases and are upregulated among the pregnant group as well, denoting a normal behavior in pregnancy where maternal glucocorticoids critically rise reaching up to a 20-fold increase of mid-pregnancy concentrations (Solano and Arck, 2020). Also, in lactating women, we found an increased ratio of palmitoyl dopamine, which is an endogenous, long-chain, linear fatty acid dopamide with entourage effects in the endocannabinoid system (Matsumoto et al., 2016). This is of high interest since its biological significance in lactation is understudied and it is then an interesting metabolite to further analyze as a biomarker.

Continuing with deeper insights into the unique metabolic traits of each group of the cohort, it can be observed in Figures 3A,C, that sphingolipids and ceramides are upregulated among the lactating group. These metabolites are involved in the regulation of insulin resistance during the perinatal period (Rico et al., 2017). It is also abundant in human breast milk and has a positive impact on cognitive functions and brain development of the infant (Dei Cas et al., 2020). In addition, prostaglandins are known to affect uterine contractility and cervical ripening and are important in the initiation of labor (Wood et al., 2021). These we found as being upregulated in pregnant women, which denotes the correct reflection of the gut chemical environment with the state of the individual. Also, these findings are in accordance with what Liang and collaborators found in 2022. They found nine metabolites differentially expressed in stool samples from pregnant women in the third trimester and full term. These included levels of lipids and lipid-like molecules, such as long-chain fatty acids and 21-hydroxysteroids, being upregulated in pregnant women compared to full-term, whereas the levels of amino acids and dipeptides showed a downregulation. On the other hand, 20-hydroxyarachidonic acid and palmitic acid were enriched at the time of full-term pregnancy. Other metabolites like cyclohexylsulfamate, 3,3-dimethylacrylic acid, hydroxyisocaproic acid, and phenylalanylphenylalanine (Phe-Phe) were also identified in fecal samples from Chinese pregnant women (Liang et al., 2022).

In summary, for the chemical space composition of the gut or as called by us in this research, the gut metabolome of our cohort, we mainly observe metabolites that are either produced by the gut microbiome bacterial metabolism or modified by it. Examples of these compounds are bile acids, bilirubin (van Best et al., 2020; Garcia et al., 2022), tryptophan (Stoll et al., 2016; Gao et al., 2020), hormones (Jiang et al., 2021; Marć et al., 2022), glycerophosphocholines. Thus, we can suggest that the gut metabolome can be seen as a reflection of an appropriate gut microbiome profile, understating appropriate as the balance of the communities according to what has been reported for a healthy state in a certain condition. Also, we prove that fecal samples, which contain small and large molecules from the gut microbiome, can indeed reflect the net result of nutrient ingestion, digestion, and absorption by both gut bacteria and the gastrointestinal tract (Ulaszewska et al., 2019.)

Another interesting finding in this study, regarding a more general behavior of the population, is the reinforcement that gut metabolomics reflects diet, drug consumption, and pharmacokinetics, even if the person does not declare it or the initial data collected in the enrolling questionnaires, such as the

one used in this study. For example, metabolites such as coumarin, omega-3 and omega-6 fatty acids, and vitamin E were detected, which are associated mostly with a plant-based diet (Pistollato et al., 2015; Sebastiani et al., 2019); drugs such as antihistamines and anticoagulants were found in the volunteers that declared its consumption. Central nervous system stimulants like caffeine were found significantly higher in the control group, and it is coherent with behavior during pregnancy and lactation, a time when women avoid high doses of this metabolite.

These findings are in accordance with the ones by (Pires et al., 2019), where the authors found significant metabolic changes in the chemical ecology of the gut environment between populations of individuals living in the Amazon, and those from an urban, industrialized setting, which was mainly attributed to dietary differences as well as diverse patterns of environmental exposure. Furthermore, organosilicons and other toxins coming from the heating of food, plastics, and agrochemicals, which can be harmful as they accumulate over time only when they have small particle sizes, are significantly abundant among the group of pregnant women from our region (Antioquia) which is not an encouraging finding from the public health perspective. Considering these compounds, specifically, those with a low silicon particle size, can overcome biological membranes and skin barriers, being possibly transferred to the baby (Dixon and Williamson, 2016), and can be endocrine disruptors. Organosilicon compounds are widely encountered in commercial products such as sealants, adhesives, coatings, medical products, and cosmetics (Mojsiewicz-Pienkowska et al., 2016).

Additionally, other, natural, compounds were found at toxic levels suggesting the ability of the methodology followed in the study to detect abnormal levels of naturally present molecules. Such is the case of volunteer 14, part of the pregnant women group, who showed significantly higher levels of primary bile acids which could be related to cholestasis and cause irreversible toxicity to the fetus (Mazzella et al., 2001).

We would like to highlight as well, that studies in other populations different from the American and European ones, such as Latin peoples, Asian people, African people, or women and children, are urgently needed as well as obataining the underlying data to properly acknowledge the gut microbiome and its associated chemical space on a world-populationscale. This will allow an appropriate and significant characterization of the gut microbiome of other countries and regions, as well as of different conditions such as healthy pregnancies and lactation in women. Furthermore, gut microbial communities change with age and sex; with one study showing a strong positive association between age and alpha diversity in young adults (less than 40 years old), and women were found to have more diversity than men (De la Cuesta-Zuluaga et al., 2019), thus groups like women, which have been previously excluded from study cohort in microbiome studies, should be included. Finally, there is a need to convert findings like the ones in this study into affordable and accessible strategies to measure gut health in every population. This reinforces the need for more studies on the gut metabolome in larger, and the use of ordinated (e.g., PCAs), clustering approaches, supervised models, or the employment of unsupervised models like NMF which have the added advantage that pre-calculated signatures of bacterial assemblages can be reapplied to even a single metagenome, removing the need for large cohort sizes capturing microbiome variation (Frioux et al., 2023).

Overall, this exploratory study serves as a starting point to describe the gut metabolome of healthy pregnant and lactating women from Antioquia, Colombia, a special population regarding the requirements of these physiological states and the profound impact that maternity can have on child development in terms of healthy growth, but also in its adequate cognitive development, as well as its regional nature. The two approaches to analyzing metabolomics data were complementary in the study, we could say that molecular networking serves as the starting point to have a broad panoramic view of the metabolites present in the chemical space. These can be later found in a more stringent and quantitative way by the untargeted metabolomic analysis.

Conclusion

Gut metabolomics studies can shed light on the phenotype differences of a population with a specific condition, such as healthy pregnancies and lactation, from others. In this study, within a cohort of women from Antioquia, Colombia, we found that lactating women can be differentiated from other pregnant and reproductive-age non-pregnant nor lactating women by a gut metabolomic profile enriched in carnitine derivatives, glycerophosphocholines, bile acids, ceramides, glycerolipids, and glycerophosphoserines. Pregnant and lactating women, when compared to reproductive age controls, are enriched in glycerophosphocholines and glycerophosphoserines, corticosteroid hormones, bile acids, fatty acids, carbohydrates, and sterols. These metabolites can be further studied in a larger population, to scale their occurrence, and plausibly develop preventive biomarkers for healthy pregnancies. Metabolites such as toxins, xenobiotics, and environmental contaminants, which can be missed by other techniques, and are ubiquitous harmful foreign chemicals present in the environment, were detected in this study in fertile age, non-pregnant nor lactating women denoting a presence in the diet and lifestyle of women that can become pregnant in the future, posing a risk to the infant's health. The metabolite Palmitoyl domamine was found as upregulated in lactating women, being reported for the first time in a gut metabolomics study, and in this specific population. Fibers and phytonutrients such as lignans and neolignans, glycosides, and lipid-derivatives from plant food sources, plant polyphenols, raffinose trisaccharides, glucosinolates, and lipopolysaccharides which are proven prebiotic compounds, were also found by molecular networking analysis in this cohort, denoting the capacity of this method to detect dietary compounds. Although our cohort is still limited for scaling these conclusions to a population level, this research sets an initial basis in our country and region, Latin America, for future population level measurements of a normal gut metabolome composition during the important perinatal period, which can provide valuable information to enhance public health nutrition strategies in middle-income countries.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The studies involving humans were approved by Institutional Ethics in Research Committee of Universidad EAFIT, formalized by the board of directors–Acts 457 of 01 October of 2014 and 474 of 26 April of 2017, in consonance with the established by resolutions 008,430 of October 4 of 1993 from the Ministry of Social Protection and 2,378 of 2008 from the Ministry of Health, among other. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization: LS-Z, SL-O, MC; Experiments and assays development in the lab: SL-O and LL-C. Methodology, Statistical Analysis: all authors contributed equally. Writing and original draft: SL-O, LL-C and LS-Z; Supervision: LS-Z and MC; Project administration: LS-Z, MC. All authors contributed to the article and approved the submitted version.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project was carried out within the framework of an internal grant from Universidad EAFIT from the 2019 call for projects, it also received funding from Minciencias (Colombia's Science, technology, and Innovation Ministry) through approved proposal in young researchers by funding SL-O's master in biosciences tuition fees. EAFIT University funded the time scientific advisor LS-Z dedicated to the project and Universidad de los Andes through the METCORE facility center funded the time that MC and LL-C devoted to the research.

Acknowledgments

The authors gratefully acknowledge support from Minciencias (Colombian Ministry of Science, Technology, and Innovation) for the financial support, and Daniel Pardo Rodriguez for his comments on the manuscript and revision. We also are deeply grateful to all the volunteers of the study for allowing the advancement of science through their provision of samples and their will to be part of the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2024.1250413/full#supplementary-material

References

Bardenhorst, S. K., Cereda, E., Severgnini, M., Barichella, M., Pezzoli, G., Keshavarzian, A., et al. (2023). Gut microbiota dysbiosis in Parkinson disease: a systematic review and pooled analysis. *Eur. J. Neurology* 30 (11), 3581–3594. doi:10.1111/ene.15671

Barrett, H. L., Gomez-Arango, L. F., Wilkinson, S. A., McIntyre, H. D., Callaway, L. K., Morrison, M., et al. (2018). A vegetarian diet is a major determinant of gut microbiota composition in early pregnancy. *Nutrients* 10 (7), 890. doi:10.3390/nu10070890

Bielik, V., and Kolisek, M. (2021). Bioaccessibility and bioavailability of minerals in relation to a healthy gut microbiome. *Int. J. Mol. Sci.* 22 (13), 6803. doi:10.3390/ijms22136803

Brantsæter, A. L., Myhre, R., Haugen, M., Myking, S., Sengpiel, V., Magnus, P., et al. (2011). Intake of probiotic food and risk of preeclampsia in primiparous women: the Norwegian Mother and Child Cohort Study. *Am. J. Epidemiol.* 174 (7), 807–815. doi:10.1093/aje/kwr168

Cheng, K., Brunius, C., Fristedt, R., and Landberg, R. (2020). An LC-QToF MS based method for untargeted metabolomics of human fecal samples. *Metabolomics* 16 (4), 46–48. doi:10.1007/s11306-020-01669-z

Chetwynd, A. J., Dunn, W. B., and Rodriguez-Blanco, G. (2017). "Collection and preparation of clinical samples for metabolomics," in *Metabolomics: from fundamentals to clinical applications. Advances in experimental medicine and biology*. Editor A. Sussulini (Cham: Springer). doi:10.1007/978-3-319-47656-8_2

Christovich, A., and Luo, X. M. (2022). Gut microbiota, leaky gut, and autoimmune diseases. *Front. Immunol.* 13 (June), 946248–946257. doi:10.3389/fimmu.2022.946248

Das, B., and Nair, G. B. (2019). Homeostasis and dysbiosis of the gut microbiome in health and disease. *J. Biosci.* 44 (5), 117–118. doi:10.1007/s12038-019-9926-y

Dei Cas, M., Paroni, R., Signorelli, P., Mirarchi, A., Cerquiglini, L., Troiani, S., et al. (2020). Human breast milk as source of sphingolipids for newborns: comparison with infant formulas and commercial cow's milk. *J. Transl. Med.* 18 (1), 481–513. doi:10.1186/s12967-020-02641-0

De la Cuesta-Zuluaga, J., Kelley, S. T., Chen, Y., Escobar, J. S., Mueller, N. T., Ley, R. E., et al. (2019). Age and sex-dependent patterns of gut microbial diversity in human adults. *BioRxiv* 4 (4), 1–12. doi:10.1128/mSystems.00261-19

Deng, Y. L., Tang, D. R., Hou, P. F., Shen, W. T., Li, H. L., Wang, T., et al. (2021). Dysbiosis of gut microbiota in patients with esophageal cancer. *Microb. Pathog.* 150 (2020), 104709. doi:10.1016/j.micpath.2020.104709

De Preter, V., Machiels, K., Joossens, M., Arijs, I., Matthys, C., Vermeire, S., et al. (2015). Faecal metabolite profiling identifies medium-chain fatty acids as discriminating compounds in IBD. *Gut* 64 (3), 447–458. doi:10.1136/gutjnl-2013-306423

Derrien, M., and Veiga, P. (2017). Rethinking diet to aid human–microbe symbiosis. *Trends Microbiol.* 25 (Issue 2), 100–112. doi:10.1016/j.tim.2016.09.011

Dhakan, D. B., Maji, A., Sharma, A. K., Saxena, R., Pulikkan, J., Grace, T., et al. (2019). The unique composition of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered using multi-omics approaches. *GigaScience* 8 (3), giz004–20. doi:10.1093/gigascience/giz004

Dixon, P. H., and Williamson, C. (2016). The pathophysiology of intrahepatic cholestasis of pregnancy. *Clin. Res. Hepatology Gastroenterology* 40 (2), 141–153. doi:10.1016/j.clinre.2015.12.008

Edwards, S. M., Cunningham, S. A., Dunlop, A. L., and Corwin, E. J. (2017). The maternal gut microbiome during pregnancy. *MCN Am. J. Maternal/Child Nurs.* 42 (6), 310–317. doi:10.1097/NMC.0000000000000372

El Hage, R., Hernandez-Sanabria, E., and Van de Wiele, T. (2017). Emerging trends in "smart probiotics": functional consideration for the development of novel health and industrial applications. *Front. Media S.A* 8, 1889. doi:10.3389/fmicb.2017.01889

El-Sayed, Y. Y., Borders, A. E. B., and Gyamfi-Bannerman, C. (2017). Committee opinion No. 713: antenatal corticosteroid therapy for fetal maturation. *Obstetrics Gynecol.* 130 (Issue 2), E102–e109. doi:10.1097/AOG.0000000000002237

Fiehn, O. (2002). Metabolomics - the link between genotypes and phenotypes. *Plant Mol. Biol.* 48 (1–2), 155–171. doi:10.1023/A:1013713905833

Frioux, C., Ansorge, R., Waszak, S. M., Hildebrand, F., Nedjad, C. G., Fritscher, J., et al. (2023). Enterosignatures define common bacterial guilds in the human gut microbiome. *Cell Host Microbe.* 31, 1111–1125.e6. doi:10.1016/j.chom.2023.05.024

Fu, J., Zheng, Y., Gao, Y., and Xu, W. (2022). Dietary fiber intake and gut microbiota in human health. *Microorganisms* 10 (12), 2507–2518. doi:10.3390/microorganisms10122507

Gao, K., Mu, C. L., Farzi, A., and Zhu, W. Y. (2020). Tryptophan metabolism: a link between the gut microbiota and brain. *Adv. Nutr.* 11 (3), 709–723. doi:10.1093/advances/nmz127

Garcia, C. J., Kosek, V., Beltrán, D., Tomás-Barberán, F. A., and Hajslova, J. (2022). Production of new microbially conjugated bile acids by human gut microbiota. *Biomolecules* 12 (5), 687. doi:10.3390/biom12050687

Gomez-Arango, L. F., Barrett, H. L., Wilkinson, S. A., Callaway, L. K., McIntyre, H. D., Morrison, M., et al. (2018). Low dietary fiber intake increases Collinsella abundance in the gut microbiota of overweight and obese pregnant women. *Gut Microbes* 9 (3), 189–201. doi:10.1080/19490976.2017.1406584

Hadadi, N., Berweiler, V., Wang, H., and Trajkovski, M. (2021). Intestinal microbiota as a route for micronutrient bioavailability. *Curr. Opin. Endocr. Metabolic Res.* 20, 100285. doi:10.1016/j.coemr.2021.100285

Holman, J. D., Tabb, D. L., and Mallick, P. (2014). Employing proteowizard to convert raw mass spectrometry data. *Curr Protoc Bioinformatics*), 1–9. doi:10.1002/0471250953.bi1324s46

Haffner, J. J., Katemauswa, M., Kagone, T. S., Hossain, E., Jacobson, D., Flores, K., et al. (2022). Untargeted fecal metabolomic analyses across an industrialization gradient reveal shared metabolites and impact of industrialization on fecal microbiome-metabolome interactions. *MSystems* 7 (6), e0071022. doi:10.1128/msystems.00710-22

Horn, J., Mayer, D. E., Chen, S., and Mayer, E. A. (2022). Role of diet and its effects on the gut microbiome in the pathophysiology of mental disorders. *Transl. Psychiatry* 12 (1), 164. doi:10.1038/s41398-022-01922-0

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., et al. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486 (7402), 207–214. doi:10.1038/nature11234

Hayashi, H., Sakamoto, M., and Benno, Y. (2002). Phylogenetic analysis of the human gut microbiota using 16s rdna clone libraries and strictly anaerobic culture-based methods. *Microbiology and Immunology* 46, 535–548. doi:10.1111/j.1348-0421.2002.tb02731.x

Ideker, T., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Jiang, L., Fei, H., Tong, J., Zhou, J., Zhu, J., Jin, X., et al. (2021). Hormone replacement therapy reverses gut microbiome and serum metabolome alterations in premature ovarian insufficiency. *Front. Endocrinol.* 12 (December), 794496–794513. doi:10.3389/fendo.2021.794496

Karu, N., Deng, L., Slae, M., Guo, A. C., Sajed, T., Huynh, H., et al. (2018). A review on human fecal metabolomics: methods, applications and the human fecal metabolome database. *Anal. Chim. Acta* 1030, 1–24. doi:10.1016/j.aca.2018.05.031

Koren, O., Goodrich, J. K., Cullender, T. C., Spor, A., Laitinen, K., Bäckhed, H. K., et al. (2012). Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell* 150 (3), 470–480. doi:10.1016/j.cell.2012.07.008

Lai, Y., Masatoshi, H., Ma, Y., Guo, Y., and Zhang, B. (2022). Role of vitamin K in intestinal health. *Front. Immunol.* 12 (January), 791565–791619. doi:10.3389/fimmu.2021.791565

Lee-Sarwar, K. A., Lasky-Su, J., Kelly, R. S., Litonjua, A. A., and Weiss, S. T. (2020). Metabolome–microbiome crosstalk and human disease. *Metabolites* 10 (5), 181–210. doi:10.3390/metabo10050181

Liang, X., Wang, R., Luo, H., Liao, Y., Chen, X., Xiao, X., et al. (2022). The interplay between the gut microbiota and metabolism during the third trimester of pregnancy. *Front. Microbiol.* 13 (December), 1059227–1059315. doi:10.3389/fmicb.2022.1059227

Magne, F., O'Ryan, M., Vidal, R., and Farfan, M. (2016). The human gut microbiome of Latin America populations: a landscape to be discovered. *Curr. Opin. Infect. Dis.* 29 (5), 528–537. doi:10.1097/QCO.0000000000000300

Mandal, S., Godfrey, K. M., McDonald, D., Treuren, W. V., Bjørnholt, J. V., Midtvedt, T., et al. (2016). Fat and vitamin intakes during pregnancy have stronger relations with a proinflammatory maternal microbiota than does carbohydrate intake. *Microbiome* 4, 55–11. doi:10.1186/s40168-016-0200-3

Marć, M. A., Jastrząb, R., and Mytych, J. (2022). Does the gut microbial metabolome really matter? The connection between GUT metabolome and neurological disorders. *Nutrients* 14 (19), 3967. doi:10.3390/nu14193967

Matsumoto, Y., Ito, A., Uesugi, M., and Kittaka, A. (2016). Efficient N-acyldopamine synthesis. *Chem. Pharm. Bull.* 64 (7), 935–940. doi:10.1248/cpb.c16-00162

Maynard, C., and Weinkove, D. (2020). Bacteria increase host micronutrient availability: mechanisms revealed by studies in *C. elegans*. *Genes Nutr.* 15 (1), 4–11. doi:10.1186/s12263-020-00662-4

Mazzella, G., Nicola, R., Francesco, A., Patrizia, S., Luciano, B., Anna, M., et al. (2001). Ursodeoxycholic acid administration in patients with cholestasis of pregnancy: effects on primary bile acids in babies and mothers. *Hepatology* 33 (3), 504–508. doi:10.1053/jhep.2001.22647

Moco, S., Collino, S., Rezzi, S., and Martin, F. P. J. (2013). Metabolomics perspectives in pediatric research. *Pediatr. Res.* 73 (4–2), 570–576. doi:10.1038/pr.2013.1

Mojsiewicz-Pienkowska, K., Jamrógiewicz, M., Szymkowska, K., and Krenczkowska, D. (2016). Direct human contact with siloxanes (silicones) - safety or risk part 1. Characteristics of siloxanes (silicones). *Front. Pharmacol.* 7 (MAY), 132–138. doi:10.3389/fphar.2016.00132

Novakovic, M., Rout, A., Kingsley, T., Kirchoff, R., Singh, A., Verma, V., et al. (2020). Role of gut microbiota in cardiovascular diseases. *World J. Cardiol.* 12 (4), 110–122. doi:10.4330/wjc.v12.i4.110

Pires, E. S., Hardoim, C. C. P., Miranda, K. R., Secco, D. A., Lobo, L. A., de Carvalho, D. P., et al. (2019). The gut microbiome and metabolome of two riparian communities in the Amazon. *Front. Microbiol.* 10 (September), 2003–2013. doi:10.3389/fmicb.2019.02003

Pistollato, F., Cano, S. S., Elio, I., Vergara, M. M., Giampieri, F., and Battino, M. (2015). Plant-based and plant-rich diet patterns during gestation: beneficial effects and possible shortcomings. *Adv. Nutr.* 6 (5), 581–591. doi:10.3945/an.115.009126

Ricciotti, E., and Fitzgerald, G. A. (2011). Prostaglandins and inflammation. *Arteriosclerosis, Thrombosis, Vasc. Biol.* 31 (5), 986–1000. doi:10.1161/ATVBAHA.110.207449

Rico, J. E., Samii, S. S., Mathews, A. T., Lovett, J., Haughey, N. J., and McFadden, J. W. (2017). Temporal changes in sphingolipids and systemic insulin sensitivity during the transition from gestation to lactation. *PLoS ONE* 12 (5), 1–20. doi:10.1371/journal.pone.0176787

Sebastiani, G., Barbero, A. H., Borrás-Novel, C., Casanova, M. A., Aldecoa-Bilbao, V., Andreu-Fernández, V., et al. (2019). The effects of vegetarian and vegan diet during pregnancy on the health of mothers and offspring. *Nutrients* 11 (3), 1–29. doi:10.3390/nu11030557

Sender, R., Fuchs, S., and Milo, R. (2016). Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* 14 (8), 10025333–e1002614. doi:10.1371/journal.pbio.1002533

Sierra-Zapata, L., Álvarez, J. C., and Romero-Tabarez, M. (2020). Inducible antibacterial activity in the bacillales by triphenyl tetrazolium chloride. *Sci Rep 10, 5563 (2020)*. doi:10.1038/s41598-020-62236-z

Solano, M. E., and Arck, P. C. (2020). Steroids, pregnancy and fetal development. *Front. Immunol.* 10 (January), 3017–3113. doi:10.3389/fimmu.2019.03017

Stoll, M. L., Kumar, R., Lefkowitz, E. J., Cron, R. Q., Morrow, C. D., and Barnes, S. (2016). Fecal metabolomics in pediatric spondyloarthritis implicate decreased metabolic diversity and altered tryptophan metabolism as pathogenic factors. *Genes Immun.* 17 (7), 400–405. doi:10.1038/gene.2016.38

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis: chemical analysis working group (CAWG) metabolomics standards initiative (MSI). *Metabolomics* 3 (3), 211–221. doi:10.1007/s11306-007-0082-2

Thursby, E., and Juge, N. (2017). Introduction to the human gut microbiota. *Biochem. J.* 474 (11), 1823–1836. doi:10.1042/BCJ20160510

Tierney, B. T., Yang, Z., Luber, J. M., Beaudin, M., Wibowo, M. C., Baek, C., et al. (2019). The landscape of genetic content in the gut and oral human microbiome. *Cell Host Microbe* 26 (2), 283–295. doi:10.1016/j.chom.2019.07.008

Trevisanuto, D., Peruzzetto, C., Cavallin, F., Vedovato, S., Cosmi, E., Visentin, S., et al. (2013). Fetal placental inflammation is associated with poor neonatal growth of preterm infants: a case-control study. *J. Maternal-Fetal Neonatal Med.* 26 (15), 1484–1490. doi:10.3109/14767058.2013.789849

Ulaszewska, M. M., Weinert, C. H., Trimigno, A., Portmann, R., Andres Lacueva, C., Badertscher, R., et al. (2019). Nutrimetabolomics: an integrative action for metabolomic analyses in human nutritional studies. *Mol. Nutr. Food Res.* J63 (1), e1800384. doi:10.1002/mnfr.201800384

Van Best, N., Rolle-Kampczyk, U., Schaap, F. G., Basic, M., Olde Damink, S. W. M., Bleich, A., et al. (2020). Bile acids drive the newborn's gut microbiota maturation. *Nat. Commun.* 11 (1), 3692. doi:10.1038/s41467-020-17183-8

Vernocchi, P., Vannini, L., Gottardi, D., Del Chierico, F., Serrazanetti, D. I., Ndagijimana, M., et al. (2012). Integration of datasets from different analytical techniques to assess the impact of nutrition on human metabolome. *Front. Cell. Infect. Microbiol.* 2 (December), 156. doi:10.3389/fcimb.2012.00156

Wan, Z., Zheng, J., Zhu, Z., Sang, L., Zhu, J., Luo, S., et al. (2022). Intermediate role of gut microbiota in vitamin B nutrition and its influences on human health. *Front. Nutr.* 9 (December), 1031502–1031521. doi:10.3389/fnut.2022.1031502

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* 34 (8), 828–837. doi:10.1038/nbt.3597

Wen, K., Tao, L., Tao, Z., Meng, Y., Zhou, S., Chen, J., et al. (2020). Fecal and serum metabolomic signatures and microbial community profiling of postmenopausal osteoporosis mice model. *Front. Cell. Infect. Microbiol.* 10 (November), 535310–535311. doi:10.3389/fcimb.2020.535310

Wild, R., and Feingold, K. R. (2000). "Effect of pregnancy on lipid metabolism and lipoprotein levels," in *Endotext* (South Dartmouth (MA): MDText.com, Inc).

Wood, E. M., Hornaday, K. K., and Slater, D. M. (2021). Prostaglandins in biofluids in pregnancy and labour: a systematic review. *PLoS ONE* 16 (11 November), e0260115–e0260122. doi:10.1371/journal.pone.0260115

Xia, J., Psychogios, N., Young, N., and Wishart, D. S. (2009). MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* 37 (Suppl. 2), 652–660. doi:10.1093/nar/gkp356

Zeisel, S. H. (2013). Nutrition in pregnancy: the argument for including a source of choline. *Int. J. Women's Health* 5 (1), 193–199. doi:10.2147/IJWH.S36610

Zhang, M., Zhao, D., Zhou, G., and Li, C. (2020). Dietary pattern, gut microbiota, and alzheimer's disease. *J. Agric. Food Chem.* 68 (46), 12800–12809. doi:10.1021/acs.jafc.9b08309

Zhang, Y., Jiang, C., Huang, S., Sun, J., Song, X., Nishanbaev, S. Z., et al. (2022). Effects of polyphenols and glucosinolates in broccoli extract on human gut microorganisms based on simulation *in vitro*. *ACS Omega* 7 (49), 45096–45106. doi:10.1021/acsomega.2c05523

Zhang, Y., Wang, Y., Ke, B., and Du, J. (2021). TMAO: how gut microbiota contributes to heart failure. *Transl. Res.* 228, 109–125. doi:10.1016/j.trsl.2020.08.007

Zhao, Y., Liu, Y., Li, S., Peng, Z., Liu, X., Chen, J., et al. (2021). Role of lung and gut microbiota on lung cancer pathogenesis. *J. Cancer Res. Clin. Oncol.* 147 (8), 2177–2186. doi:10.1007/s00432-021-03644-0

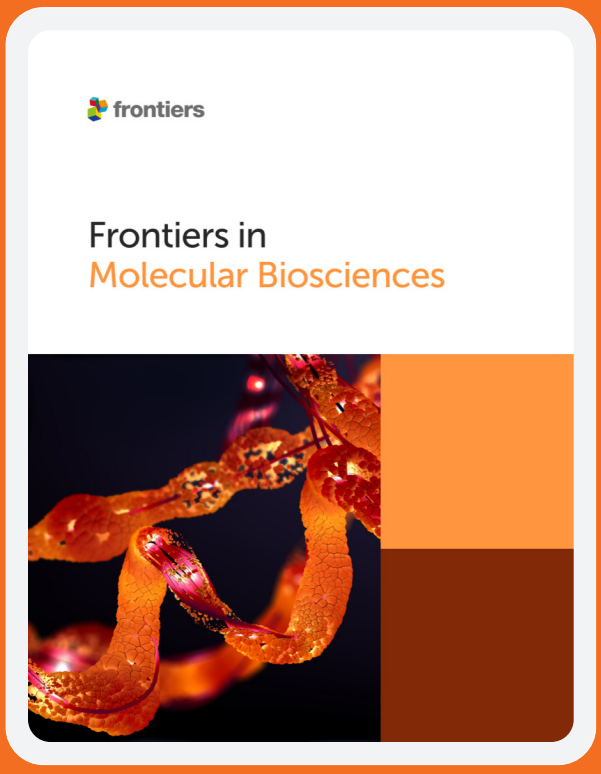
Frontiers in
Molecular Biosciences

Explores biological processes in living organisms on a molecular scale

Focuses on the molecular mechanisms underpinning and regulating biological processes in organisms across all branches of life.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact