METABOLOMICS PERSPECTIVES FOR CLINICAL MEDICINE

EDITED BY: Michal Jan Markuszewski, Philip Britz-McKibbin, Serge Rudaz and Danuta Dudzik

PUBLISHED IN: Frontiers in Molecular Biosciences

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-343-8
DOI 10.3389/978-2-88974-343-8

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

METABOLOMICS PERSPECTIVES FOR CLINICAL MEDICINE

Topic Editors:
Michal Jan Markuszewski, Medical University of Gdansk, Poland
Philip Britz-McKibbin, McMaster University, Canada
Serge Rudaz, Université de Genève, Switzerland
Danuta Dudzik, Medical University of Gdańsk, Poland

Citation: Markuszewski, M. J., Britz-McKibbin, P., Rudaz, S., Dudzik, D., eds. (2022). Metabolomics Perspectives for Clinical Medicine. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-343-8

Table of Contents

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 15 October 2020
doi: 10.3389/fmolb.2020.585161

Metabolic Pattern of Systemic Sclerosis: Association of Changes in Plasma Concentrations of Amino Acid-Related Compounds With Disease Presentation

Zaneta Smolenska[1], Magdalena Zabielska-Kaczorowska[2,3], Anna Wojteczek[1], Barbara Kutryb-Zajac[3] and Zbigniew Zdrojewski[1]*

[1] Department of Internal Medicine, Connective Tissue Diseases and Geriatrics, Medical University of Gdańsk, Gdańsk, Poland, [2] Department of Human Physiology, Medical University of Gdańsk, Gdańsk, Poland, [3] Department of Biochemistry, Medical University of Gdańsk, Gdańsk, Poland

OPEN ACCESS

Edited by:
Serge Rudaz,
Université de Genève, Switzerland

Reviewed by:
Roccaldo Sardella,
University of Perugia, Italy
Ann Van Schepdael,
KU Leuven, Belgium

***Correspondence:**
Zaneta Smolenska
zaneta.smolenska@gumed.edu.pl

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

Received: *19 July 2020*
Accepted: *22 September 2020*
Published: *15 October 2020*

Citation:
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

Objective: Amino acids (AA) and their derivatives play an integral role in the synthesis of structural and regulatory elements in human organisms and therefore pathologies such as systemic sclerosis that may alter the blood pattern of these compounds. This study aimed to evaluate changes in plasma concentrations of amino acid-related metabolites in systemic sclerosis in a search for potential biomarkers and mechanisms of the disease.

Methods: Plasma samples from 42 patients diagnosed with systemic sclerosis (SSc) according to the 2013 American College of Rheumatology and European League Against Rheumatism ACR/EULAR classification criteria were compared to 27 matched healthy controls. Liquid chromatography/mass spectrometry was applied for the analysis of 36 amino acid-related metabolites.

Results: The analysis of plasma AA metabolite patterns revealed the number of changes including an increase (20%) in concentrations of NO synthase (NOS) inhibitor asymmetric dimethylarginine (ADMA) in SSc vs. healthy subjects. Furthermore, SSc patients had higher glutamine, proline, betaine, 1-methylhistidine, and methylnicotinamide levels, while the concentration of tryptophan was lower. The specific metabolic pattern was identified for several aspects of disease presentation. Most interestingly NOS inhibitor L-NAME was elevated in patients with diffuse systemic sclerosis or telangiectasia.

Conclusions: These results provide further evidence for the involvement of endothelium-dependent pathways in the mechanisms and presentation of SSc. Endothelial dysfunction biomarkers may be useful in the assessment of presentation and prognosis in SSc.

Keywords: systemic sclerosis, amino acids, metabolome, endothelium, inflammation

INTRODUCTION

Systemic sclerosis (SSc) is an autoimmune connective tissue disease characterized by fibrosis of the tissues, preceded by microvascular alterations and immune dysfunction (Zanatta et al., 2019). The proliferation of fibroblasts with the production of excessive extracellular matrix leads to the thickening of the skin and damage to internal organs. It is a highly heterogeneous disease with a wide variety of clinical presentations and is connected with increased morbidity and mortality (Barsotti et al., 2019). SSc can occur in two main cutaneous subsets: diffuse systemic sclerosis (dcSSc) and limited systemic sclerosis (lcSSc) (Young and Khanna, 2015). In the lcSSc subset, skin fibrosis is mainly restricted to distal extremities and the face, with slow accumulation of organ involvement, while Raynaud's phenomenon typically occurs even years before any skin and visceral changes. By contrast, the dcSSc subset presents a more aggressive progression characterized by severe internal organ manifestations with skin thickening extended proximally to elbows and knees as well as the trunk. In the dcSSc subtype, anti-topoisomerase I autoantibodies (anti-Scl-70) are more frequent, while in the lcSSc subtype, anti-centromere autoantibodies are dominant features (Kranenburg et al., 2016).

In 2013, EULAR published new criteria for SSc, which can help early diagnose this illness (Van Den Hoogen et al., 2013). These include main clinical features as extending of skin involvement, finger changes as sclerodactyly, puffy fingers, digital ulcerations and pitting scars, telangiectasia, the presence of lungs complications (pulmonary hypertension and/or interstitial disease) but other manifestations as from GI tract, heart, arthritis with tendon inflammation, or calcinosis in the skin also occur commonly in SSc. Many studies have shown that dcSSc characterized by rapid progression of skin thickness has been associated with earlier onset of heart, lung involvement, and also with increased disease severity and mortality rates due to this organ involvement (Kumánovics et al., 2017). Predicting which organs are likely to be involved, and better understanding who is at the risk for potentially devastating complications is critical for appropriate patient counseling.

Metabolomic profiling including the analysis of plasma amino acid balance is a new, promising biochemical approach that can be applied for SSc screening, diagnosis, disease typing, and treatment monitoring, as well as may be promising for SSc biomarker discovery (Sandlers, 2020). Amino acids play a very integral role as metabolic intermediates and in the building block of proteins, while their plasma concentration highly depends on the organs' functions and pathological conditions that alter human body metabolism (Bröer and Bröer, 2017).

This study aimed to evaluate the changes in plasma concentrations of amino acids and related metabolites in SSc depending on the clinical characteristics of patients and organ involvement regarding potential mechanisms of the disease. The additional goal was to investigate whether individual amino acids and their derivatives can be potential SSc biomarkers.

MATERIALS AND METHODS

Patients

This study included 42 adult patients (35 women, 7 men, aged 59.9 ± 12.4; min–max: 28–78 years) diagnosed with SSc referred to the Outpatient Rheumatology Clinic of the Department of Internal Diseases, Connective Tissue Diseases and Geriatrics of the University Hospital in Gdańsk, Poland. The control group represented age and sex-matched individuals (n = 27) with no report of any pathology. All the investigation involving human subjects was performed in compliance with standards of the Declaration of Helsinki (1975/83). The study has been approved by the Bioethics Committee at the Medical University of Gdańsk, Poland. The diagnosis of SSc was established according to the 2013 American College of Rheumatology and European League Against Rheumatism ACR/EULAR classification criteria. The SSc patients were categorized into the limited (lcSSc, n = 21) or the diffuse cutaneous (dcSSc, n = 21) subsets. The modified Rodnan skin score was used for grading the cutaneous extension of sclerosis. Organ involvement was evaluated using double-contrast esophagography, high-resolution computed tomography (HRCT) scans, pulmonary function tests (PFTs), two-dimensional echocardiogram, X-ray, and ultrasonography of the musculoskeletal system. All patients were subjected to basic routine laboratory and serological tests. Further clinical features are summarized in **Table 1**.

The Analysis of Plasma Amino Acids and Nicotinamide Metabolites

The concentrations of plasma amino acids and related compounds were determined using liquid chromatography/mass spectrometry (LC/MS) according to our described procedure (Olkowicz et al., 2017) that expanded earlier use of the reversed-phase ion-paring approach for analysis of amino acids (Chaimbault et al., 1999). Briefly, an aliquot of plasma (0.05 mL) was spiked with internal standards and deproteinized with 0.1 mL of acetonitrile followed by maintenance on ice for 15 min. The tubes were then centrifuged at 4°C, 12,000 × g for 5 min. The supernatant was collected and freeze-dried. Samples were then dissolved in 0.1 mL of water and analyzed with the use of ion-pair high-performance liquid chromatography with mass detection. Chromatographic separation was performed using 2.5 μm Synergy Hydro-RP 50 × 2.0 mm column. The mobile phase was delivered at 0.2 mL/min in a gradient from 0 to 60% acetonitrile in 12 min. The mass detector (TSQ Vantage, Thermo, United States) with heated electrospray (HESI-2) ion source was operating in positive MS2 mode for the detection of amino acids. The electrospray cone voltage was set at 4.5 kV and a heated capillary temperature was 275°C. Sheath gas flow was set for 35 arbitrary units. Post column make-up flow of methanol with 0.05% formic acid at 0.2 mL/min was used to improve ionization efficiency. The identity of individual amino acids and their derivatives was confirmed by the similarity of molecular weights, fragmentation pattern, and chromatographic retention time. The procedure allowed interference-free quantitation of all analyzed compounds including challenging molecule pairs such as SDMA

TABLE 1 | Demographic and clinical presentation of systemic sclerosis patients (n = 42).

Gender	85% F/15% M
Age (y)	59.9 ± 12.4 (28–78)
C-reactive protein (mg/L)	4.9 ± 4.3 (0.2–45.9)
Erythrocyte sedimentation rate (mm/h)	15.3 ± 12.3 (1.6–66.0)
Calcinosis	14%
Joint involvement	73%
Lung involvement	61%
Dysphagia	28%
Telangiectasia	65%
Scleroderma	49%
Upper GI tract involvement	45%
Heart involvement	17%

and ADMA that despite similar molecular weights produced different fragments in MS^2 mode that were used for quantitation. Calibration was based on curve prepared in spiked plasma matrix as we have described in detail earlier (Olkowicz et al., 2017).

Statistical Analyses

The significance of differences in the level of analyzed compounds between patients and controls or their association with pathology were evaluated by Student t-test or by Mann-Whitney's U-test when variables distribution was not symmetrical (as confirmed by Shapiro-Wilks' W test) or the groups of patients were much different from the sample size. All results are presented as mean ± S.D. with the minima and maxima observed. $p < 0.05$ was considered as a significant difference.

RESULTS

The comparison of amino acid (AA)-related metabolites concentration between systemic sclerosis patients and healthy control group is shown in **Table 2**. The patients have higher concentration of glutamine (12% increase over control = IOC), proline (17% IOC), 1-methylhistidine (39% IOC), betaine (23% IOC), methylnicotinamide (MNA, 35% IOC) and asymmetric dimethylarginine (ADMA, 19% IOC). The level of tryptophan was lower in the SSc group than in controls by 20%.

We have demonstrated many alterations in plasma concentrations of amino acid-related compounds that were dependent on specific disease presentations within the SSc patient group (**Table 3**). Patients with dcSSc revealed higher concentrations of sarcosine, β-alanine, MNA and N(G)-nitro-L-arginine Methyl Ester (L-NAME) than the lcSSc group. Disease presentation with calcinosis was associated with a significant elevation of glutamate, sarcosine, proline, tyrosine, 3-methylhistidine, and ornithine concentrations. Patients with joint pain have a lower level of plasma glutamine but higher levels of ornithine and 1-methylhistidine. Lung involvement was associated with higher concentrations of valine and arginine. Patients with telangiectasia have a higher concentration of glutamate, lysine, and L-NAME. The presence of extensive skin changes (scleroderma) was associated with lower concentrations of a broad range of AA such as asparagine, sarcosine, proline, histidine, ornithine, citrulline, and phenylalanine.

DISCUSSION

In this study, we investigated the plasma amino acid (AA) profile in patients with SSc to detect potential biomarkers and to relate particular changes in the AA pattern with clinical characteristics of patients and organ involvement in the context of potential pathological mechanisms. We indicated a group of AA related compounds that have prevalent fluctuations in this pathology. The most significant changes between SSc patients and healthy controls were observed in plasma 1-methylhistidine (39% increase over control, IOC) > MNA (35% IOC) > betaine (23% IOC) > tryptophan (20% decrease over control, DOC) > ADMA (19% IOC) > proline (17% IOC) > glutamine (12% IOC). Moreover, we have demonstrated significant alterations in the plasma AA levels within the SSc patient group that deserve attention as potential molecules that may predict the specific course of the disease and are worth further studies as clinical biomarkers. The disturbed balance of the AA metabolites that we observed is known to be associated with vascular endothelial dysfunction, inflammation, or methylation disturbances.

Cardiovascular dysfunction is of fundamental importance in SSc pathogenesis from the early onset of the disease through the late clinical complications (Silva et al., 2015; Smoleńska et al., 2019). In this study, we revealed an increased plasma concentration of the endogenous NO synthase (NOS) inhibitor – ADMA, that was increased in the entire SSc patients' group in comparison to healthy controls. The impaired nitric oxide production leads to the augmentation of vasoconstrictor episodes and pathological changes in the vascular system such as inflammatory and thrombotic signaling or vascular remodeling (Flavahan, 2015). Moreover patients with SSc expressed a decreased level of plasma tryptophan. This AA could be utilized by endothelial cells and via the kynurenine pathway may direct them to ROS-dependent apoptosis (Duran and San Martín, 2014). Interestingly, patients with specific clinical presentations such as dcSSc and telangiectasia revealed an elevated level of other NOS inhibitor - L-NAME, suggesting increased severity of the endothelial dysfunction that could mark active disease pattern and more extended vascular injury. Therefore, our research highlights the importance of multi-parameter profiles of endothelium analysis in patients with SSc that will help to track the severity of endothelial dysfunction and contribute to the monitoring of endothelium protection therapy.

The pathogenesis of SSc involves the Th1-related early inflammatory phase, which is then followed by a switch to Th2, leading to irreversible fibrosis. However, in some subgroups of SSc patients, there is a prolonged active inflammatory response. Therefore, specific inflammatory markers were proposed as important candidates for the diagnosis and differentiation of the form of the disease (Ross et al., 2018). Overall, patients with inflammatory SSc seems to represent a subgroup

TABLE 2 | The comparison of amino acid metabolite concentrations in systemic sclerosis patients with control healthy subjects.

	Control (*n* = 27)				Systemic sclerosis (*n* = 42)			
	Mean	**S.D.**	**Min**	**Max**	**Mean**	**S.D.**	**Min**	**Max**
Alanine	223.5	119.6	65.0	528.3	265.5	134.6	83.4	682.6
Asparagine	36.7	13.0	9.1	60.9	36.5	13.7	15.8	83.4
Aspartate	9.7	4.3	4.6	23.5	7.8	4.9	1.1	20.2
Arginine	67.8	22.1	17.8	113.6	76.7	26.8	38.2	167.7
Cystine	16.5	8.3	2.8	35.6	18.6	9.5	0.9	41.3
Glutamate	86.7	52.1	4.8	225.2	98.0	42.7	36.1	205.3
Glutamine	618.4	165.3	231.0	893.1	689.0*	122.3	457.2	971.1
Glycine	219.3	107.4	54.3	541.3	216.4	85.1	82.5	464.7
Histidine	88.6	21.6	31.7	120.5	93.0	27.7	43.6	166.0
Isoleucine	96.9	41.8	27.7	213.9	102.9	33.5	41.0	198.9
Leucine	129.5	36.3	28.3	190.6	131.3	31.9	81.2	233.7
Lysine	146.1	43.4	32.6	218.9	139.6	30.3	74.6	212.5
Methionine	33.3	12.4	4.7	55.5	29.8	9.0	13.5	52.7
Phenylalanine	69.0	24.2	20.1	140.9	74.8	30.1	42.0	207.5
Proline	152.5	47.3	23.7	247.1	178.8*	55.2	93.7	342.6
Serine	94.8	39.3	22.2	200.7	93.1	24.8	32.2	152.7
Threonine	141.8	52.6	34.5	238.5	141.2	41.2	77.1	282.6
Tryptophan	40.8	12.3	13.3	61.8	32.5**	9.6	11.5	66.6
Tyrosine	73.1	24.5	15.3	130.2	70.8	24.9	32.1	166.8
Valine	242.7	80.5	55.0	470.6	240.5	49.8	147.6	389.9
1-methylhistidine	4.1	1.5	0.9	7.0	5.7*	3.9	1.6	26.4
3-methylhistidine	5.1	5.6	0.5	21.7	6.4	7.5	0.0	39.6
α-aminobutyrate	26.4	9.6	6.6	51.0	25.5	9.1	9.9	52.8
β-aminobutyrate	5.1	1.3	3.0	8.3	5.6	1.3	3.5	8.3
Betaine	52.8	17.8	19.2	82.8	64.8*	20.8	29.4	117.0
β-alanine	6.1	3.7	0.5	13.1	7.1	4.3	1.3	17.6
Citrulline	39.5	15.1	9.7	69.1	40.0	13.5	17.7	81.0
Hydroxyproline	5.9	4.8	1.8	23.3	6.7	3.6	2.1	20.7
Methylnicotinamide	0.232	0.106	0.065	0.550	0.312*	0.166	0.112	0.959
Ornithine	45.6	15.9	8.5	68.6	52.2	16.2	25.8	100.8
Sarcosine	7.5	1.9	3.2	11.0	7.8	2.7	3.4	17.0
Taurine	66.7	29.6	22.8	134.4	60.0	24.0	34.7	156.6
ADMA	0.289	0.100	0.096	0.512	0.344*	0.112	0.101	0.576
L-NAME	0.051	0.015	0.026	0.081	0.056	0.018	0.030	0.102

*Values are in μmol/L. *p < 0.05; **p < 0.01.*

with higher morbidity and mortality. Accordingly, patients with inflammatory SSc showed a faster decline of forced vital capacity (FVC) over time and received more frequently immunosuppressive treatment (e.g., cyclophosphamide, CYC) (Mitev et al., 2019). It has been demonstrated that these patients showed persistent long-term CRP elevations and even treatment with CYC did not affect CRP levels. Presumably, macrophages, less affected by CYC treatment, maintain inflammation in this subgroup. In our patient group, systemic inflammation may be responsible for the increased concentration of MNA in dcSSc, while elevated sarcosine and β-alanine may relate to more severe pathological processes and muscle injury. Moreover, the consistent increase of betaine, 1-methylhistidine, and MNA (by 20–40%) in SSc patients in comparison to healthy controls implies the activation of methylated derivatives formation. One challenge is that these compounds are produced by completely different methylation mechanisms: methylation of histidine residues in skeletal muscle proteins in the case of 1-methyhistidine, methylation of free nicotinamide by liver-specific enzyme in the case of MNA, and methylation of phospholipid elements in the case of betaine (Houweling et al., 2012; Obeid, 2013; Pissios, 2017). One common factor that is needed in all these processes is S-adenosylmethionine. However, it needs to be established whether the production or turnover of this molecule is indeed stimulated in patients with SSc.

The disease presentation with calcinosis was associated with a significant elevation of glutamate, sarcosine, proline, tyrosine, 3-methylhistidine, and ornithine concentration. It is difficult to speculate on the mechanism but the broad range and unidirectional trend for amino acid changes may indicate its reduced utilization in general that may be related to impaired mobility. Patients with joint pain have a lower level of plasma

TABLE 3 | The concentrations of amino acid metabolites in plasma of systemic sclerosis patients that were significantly different ($p < 0.05$) in patients with specific pattern of disease.

	dcSSc (*n* = 21)				lcSSc (*n* = 21)			
	Mean	**S.D.**	**Min**	**Max**	**Mean**	**S.D.**	**Min**	**Max**
β -alanine	6.05	3.76	1.32	13.3	9.27	4.7	2.83	17.62
Methylnicotinamide	0.31	0.17	0.15	0.95	0.35	0.12	0.22	0.70
Sarcosine	7.35	2.91	3.81	17.03	8.98	2.36	4.22	14.35
L-NAME	0.052	0.012	0.034	0.074	0.065	0.020	0.035	0.102
	Calcinosis present (*n* = 6)				**No calcinosis (*n* = 36)**			
Glutamate	146.6	57.5	36.1	205.2	98.3	37.2	41.7	201.5
Proline	246.8	35.9	214.8	315.0	178.5	50.1	102.4	342.6
Tyrosine	99.0	36.1	54.7	148.4	68.9	22.4	42.6	166.7
3-methylhistidine	14.5	13.4	2.3	39.6	5.2	5.7	0.05	22.7
Ornithine	69.3	15.0	44.1	87.7	52.8	15.5	32.6	100.8
Sarcosine	10.1	2.1	8.6	14.3	7.9	2.7	3.8	17.0
	Joint pain (*n* = 28)				**No joint pain (*n* = 14)**			
Glutamine	690.6	134.9	457.2	971.1	716.3	121.8	540.8	906.3
Ornithine	58.2	17.1	32.6	100.8	47.5	12.0	33.6	74.5
1-methylhistidine	6.9	1.56	26.3	4.9	4.2	2.5	6.5	1.1
	Lung involvement (*n* = 26)				**No lung involvement (*n* = 16)**			
Arginine	81.3	26.4	43.4	167.7	63.7	22.5	38.1	123.9
Valine	249.0	48.6	147.6	346.7	222.5	34.5	178.4	286.6
	Telangiectasia (*n* = 24)				**No telangiectasia (*n* = 18)**			
Glutamate	112.7	44.2	36.1	205.2	89.0	39.4	45.2	201.5
Lysine	148.0	30.1	88.3	212.5	130.8	29.9	74.5	203.7
L-NAME	0.07	0.02	0.03	0.10	0.05	0.13	0.03	0.08
	Scleroderma (*n* = 20)				**No scleroderma (*n* = 22)**			
Asparagine	34.9	15.2	17.8	83.3	42.0	11.3	20.0	73.0
Proline	167.1	38.7	102.5	235.6	205.0	62.7	100.0	342.7
Histidine	86.6	26.5	49.0	166.0	104.3	26.57	48.9	146.8
Phenylalanine	68.0	24.3	42.0	144.3	85.4	39.6	44.1	207.5
Citrulline	35.7	10.3	17.6	60.7	45.2	16.0	23.3	81.0
Ornithine	49.6	14.4	32.6	77.1	60.8	16.1	44.1	100.8
Sarcosine	7.2	2.6	4.0	13.2	9.2	2.3	5.6	17.0

Values are in μ mol/L.

glutamine but higher of ornithine and 1-methylhistidine. That again may relate to impaired mobility or in the case of glutamine to impaired immune system function (Cruzat et al., 2018).

The presence of extensive skin changes (scleroderma) was associated with lower concentrations of asparagine, sarcosine, proline, histidine, ornithine, citrulline, and phenylalanine. Such a broad range of changes may indicate accelerated protein synthesis associated with this specific disease presentation. In turn, proline deficiency, which is an essential amino acid during periods of increased body stress, in patients with scleroderma can be an indicator of poor prognosis (Liang et al., 2013). Extensive skin involvement corresponds to more severe internal organ manifestation and increased impairment (Wu, 2009). The differences in intestinal absorption or diet may also contribute. Histidine and phenylalanine are nutritionally essential amino acids and their lower levels suggest that there are disturbances in absorption from the digestive tract in scleroderma patients. Even 90% of cases with SSc involve the digestive tract suggesting the need for appropriate nutritional care, treatment, and dedicated monitoring (Miller et al., 2018).

One compound that differentiated several groups was sarcosine. This molecule is an intermediate of glycine metabolism, being a specific substrate of sarcosine dehydrogenase that is a mitochondrial enzyme, which converts sarcosine to glycine. Sarcosine is formed from betaine via dimethylglycine by dimethylglycine dehydrogenase (Augustin et al., 2016). Concentration of sarcosine was lower in dcSSc patients and in those with scleroderma while higher concentration

was observed in calcinosis presentation. No differences in betaine concentration indicate that its conversion to glycine is responsible for the changes in sarcosine concentration. While the exact mechanism of differences in sarcosine concentration is difficult to propose at present, this molecule is worth further investigation as a biomarker of specific disease complications.

CONCLUSION

This study confirmed that the pathobiology of systemic sclerosis interferes with the plasma amino acid profile and particular amino acid metabolites can be selected as potential biomarkers and predictive factors that differentiate individual subtypes of systemic sclerosis.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/supplementary material.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee at the Medical University of Gdańsk. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ZS conceived the study, conducted data analysis, and wrote the manuscript. MZ-K conducted biochemical analysis. AW conducted clinical data analysis. BK-Z wrote the manuscript. ZZ corrected the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Science Centre of Poland (2019/35/D/NZ3/03512 and 2016/23/B/NZ4/03877).

REFERENCES

Augustin, P., Hromic, A., Pavkov-Keller, T., Karl, G., and Peter, M. (2016). Structure and biochemical properties of recombinant human dimethylglycine dehydrogenase and comparison to the disease-related H109R variant. *FEBS J.* 283, 3587–3603. doi: 10.1111/febs.13828

Barsotti, S., Martina, O., Veronica, C., Marco, D. B., Gemma, L., and Alessandra, D. R. (2019). One year in review 2019: systemic sclerosis. *Clin. Exp. Rheumatol.* 37 (Suppl. 119), S3–S14.

Bröer, S., and Bröer, A. (2017). Amino acid homeostasis and signalling in mammalian cells and organisms. *Biochem. J.* 474, 1935–1963. doi: 10.1042/BCJ20160822

Chaimbault, P., Petritis, K., Elfakir, C., and Dreux, M. (1999). Determination of 20 underivatized proteinic amino acids by ion-pairing chromatography and pneumatically assisted electrospray mass spectrometry. *J. Chromatogr. A* 855, 191–202. doi: 10.1016/s0021-9673(99)00685-8

Cruzat, V., Macedo Rogero, M., Noel Keane, K., Rui, C., and Philip, N. (2018). Glutamine: metabolism and immune function, supplementation and clinical translation. *Nutrients* 10:1564. doi: 10.3390/nu10111564

Duran, C., and San Martín, A. (2014). Do endothelial cells eat tryptophan to die? *Circ. Res.* 114, 406–408. doi: 10.1161/CIRCRESAHA.113.303150

Flavahan, N. A. (2015). A vascular mechanistic approach to understanding Raynaud phenomenon. *Nat. Rev. Rheumatol.* 11, 146–158. doi: 10.1038/nrrheum.2014.195

Houweling, M., van der Drift, S. G. A., Jorritsma, R., and Tielens, A. G. M. (2012). Technical note: quantification of plasma 1- and 3-methylhistidine in dairy cows by high-performance liquid chromatography-tandem mass spectrometry. *J. Dairy Sci.* 95, 3125–3130. doi: 10.3168/jds.2011-4769

Kranenburg, P., van den Hombergh, W. M. T., Knaapen-Hans, H. K. A., van den Hoogen, F. H. J., Jaap, F., and Madelon, C. V. (2016). Survival and organ involvement in patients with limited cutaneous systemic sclerosis and antitopoisomerase-I antibodies: Determined by skin subtype or auto-antibody subtype? A long-term follow-up study. *Rheumatology* 55, 2001–2008. doi: 10.1093/rheumatology/kew298

Kumánovics, G., Péntek, M., Bae, S., Daniela, O. P., Dinesh, K., and Daniel, E. F. (2017). Assessment of skin involvement in systemic sclerosis. *Rheumatology* 56, 53–66.

Liang, X., Zhang, L., Natarajan, S. K., and Becker, D. F. (2013). Proline mechanisms of stress survival. *Antioxid. Redox. Signal.* 19, 998–1011. doi: 10.1089/ars.2012.5074

Miller, J. B., Gandhi, N., Clarke, J., and McMahan, Z. (2018). Gastrointestinal involvement in systemic sclerosis: an update. *J. Clin. Rheumatol.* 24, 328–337. doi: 10.1097/rhu.0000000000000626

Mitev, A., Christ, L., Feldmann, D., Moritz, B., Kim, M., and Anna-Maria, K. (2019). Inflammatory stays inflammatory: a subgroup of systemic sclerosis characterized by high morbidity and inflammatory resistance to cyclophosphamide. *Arthritis Res. Ther.* 21:262. doi: 10.1186/s13075-019-2057-x

Obeid, R. (2013). The metabolic burden of methyl donor deficiency with focus on the betaine homocysteine methyltransferase pathway. *Nutrients* 5, 3481–3495. doi: 10.3390/nu5093481

Olkowicz, M., Debski, J., Jablonska, P., Michal, D., and Ryszard, T. S. (2017). Application of a new procedure for liquid chromatography/mass spectrometry profiling of plasma amino acid-related metabolites and untargeted shotgun proteomics to identify mechanisms and biomarkers of calcific aortic stenosis. *J. Chromatogr. A* 1517, 66–78. doi: 10.1016/j.chroma.2017.08.024

Pissios, P. (2017). Nicotinamide N-Methyltransferase: more than a Vitamin B3 clearance enzyme. *Trends Endocrinol. Metab.* 28, 340–353. doi: 10.1016/j.tem.2017.02.004

Ross, L., Stevens, W., Rabusa, C., Wilson, M., Ferdowsi, N., and Walker, J. (2018). The role of inflammatory markers in assessment of disease activity in systemic sclerosis. *Clin. Exp. Rheumatol.* 36(Suppl. 113), 126–134.

Sandlers, Y. (2020). *Amino Acids Profiling for the Diagnosis of Metabolic Disorders. In: Biochemical Testing - Clinical Correlation and Diagnosis*. London: IntechOpen.

Silva, I., Teixeira, A., Oliveira, J., Isabel, A., Rui, A., and Carlos, V. (2015). Predictive value of vascular disease biomarkers for digital ulcers in systemic sclerosis patients. *Clin. Exp. Rheumatol.* 33(4 Suppl. 91), S127–S130.

Smoleńska, Ż, Puzio, E., Dorniak, K., Marta, G. B., and Zbigniew, Z. (2019). Heart in systemic sclerosis – pathogenesis and diagnosis based on new imaging methods. *Forum Reumatol.* 5, 181–189. doi: 10.5603/FR.2019.0020

Van Den Hoogen, F., Khanna, D., Fransen, J., Philip, J. C., Christopher, P. D., and Oliver, D. (2013). 2013 classification criteria for systemic sclerosis: an american college of rheumatology/European league against rheumatism

collaborative initiative. *Arthritis Rheum.* 65, 2737–2747. doi: 10.1002/art.38098

Wu, G. (2009). Amino acids: metabolism, functions, and nutrition. *Amino Acids* 37, 1–17. doi: 10.1007/s00726-009-0269-0

Young, A., and Khanna, D. (2015). Systemic sclerosis: commonly asked questions by rheumatologists. *J. Clin. Rheumatol.* 21, 149–155. doi: 10.1097/RHU.0000000000000232

Zanatta, E., Claudia, C., D'Amico, G., d'Humières, T., Carlo, D. L., and Francesco, T. (2019). Inflammation and coronary microvascular dysfunction in autoimmune rheumatic diseases. *Int. J. Mol. Sci.* 20:5563. doi: 10.3390/ijms20225563

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 14 December 2020
doi: 10.3389/fmolb.2020.609806

Distinctive Metabolomics Patterns Associated With Insulin Resistance and Type 2 Diabetes Mellitus

Xinyun Gu[1], Mohammed Al Dubayee[2], Awad Alshahrani[2], Afshan Masood[3], Hicham Benabdelkamel[3], Mahmoud Zahra[4], Liang Li[1], Anas M. Abdel Rahman[4,5,6] and Ahmad Aljada[4]**

[1] Department of Chemistry, University of Alberta, Edmonton, AB, Canada, [2] Department of Medicine, College of Medicine, King Saud bin Abdulaziz University for Health Sciences, King Abdullah International Medical Research Center, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia, [3] Obesity Research Center, College of Medicine, King Saud University, Riyadh, Saudi Arabia, [4] Department of Biochemistry and Molecular Medicine, College of Medicine, Alfaisal University, Riyadh, Saudi Arabia, [5] Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia, [6] Department of Chemistry, Memorial University of Newfoundland, St. John's, NL, Canada

Obesity is associated with an increased risk of insulin resistance (IR) and type 2 diabetes mellitus (T2DM) which is a multi-factorial disease associated with a dysregulated metabolism and can be prevented in pre-diabetic individuals with impaired glucose tolerance. A metabolomic approach emphasizing metabolic pathways is critical to our understanding of this heterogeneous disease. This study aimed to characterize the serum metabolomic fingerprint and multi-metabolite signatures associated with IR and T2DM. Here, we have used untargeted high-performance chemical isotope labeling (CIL) liquid chromatography-mass spectrometry (LC-MS) to identify candidate biomarkers of IR and T2DM in sera from 30 adults of normal weight, 26 obese adults, and 16 adults newly diagnosed with T2DM. Among the 3633 peak pairs detected, 62% were either identified or matched. A group of 78 metabolites were up-regulated and 111 metabolites were down-regulated comparing obese to lean group while 459 metabolites were up-regulated and 166 metabolites were down-regulated comparing T2DM to obese groups. Several metabolites were identified as IR potential biomarkers, including amino acids (Asn, Gln, and His), methionine (Met) sulfoxide, 2-methyl-3-hydroxy-5-formylpyridine-4-carboxylate, serotonin, L-2-amino-3-oxobutanoic acid, and 4,6-dihydroxyquinoline. T2DM was associated with dysregulation of 42 metabolites, including amino acids, amino acids metabolites, and dipeptides. In conclusion, these pilot data have identified IR and T2DM metabolomics panels as potential novel biomarkers of IR and identified metabolites associated with T2DM, with possible diagnostic and therapeutic applications. Further studies to confirm these associations in prospective cohorts are warranted.

Keywords: type 2 diabetes mellitus, insulin resistance, obesity, untargeted metabolomics profiling, clinical metabolic panel, chemical isotope labeling liquid chromatography

Abbreviations: AAA, aromatic amino acids; ACN, acetonitrile; Asn, asparagine; BCAA, branched-chain amino acid; BMI, body mass index; CIL, chemical isotope labeling; DnsCl, dansyl chloride; FA, formic acid; Gln, glutamine; HG, hyperglycemia; His, histidine; HOMA-IR, homeostatic model assessment; INSR, insulin receptor; IR, insulin resistance; IRS-1, insulin receptor substrate-1; LC-MS, liquid chromatography-mass spectrometry; LDL, low-density lipoprotein; LHD, lower high-density lipoprotein; NEFAs, non-esterified fatty acids; PLS-DA, partial least squares discriminant analysis; T2DM, type 2 diabetes mellitus; Trig, triglycerides.

OPEN ACCESS

Edited by:
Philip Britz-McKibbin, McMaster University, Canada

Reviewed by:
Michal Jan Markuszewski, Medical University of Gdańsk, Poland
Hai-long Piao, Chinese Academy of Sciences, China

****Correspondence:***
Anas M. Abdel Rahman
aabdelrahman46@kfshrc.edu.sa
Ahmad Aljada
aaljada@alfaisal.edu

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

Received: *24 September 2020*
Accepted: *23 November 2020*
Published: *14 December 2020*

Citation:
Gu X, Al Dubayee M, Alshahrani A, Masood A, Benabdelkamel H, Zahra M, Li L, Abdel Rahman AM and Aljada A (2020) Distinctive Metabolomics Patterns Associated With Insulin Resistance and Type 2 Diabetes Mellitus. Front. Mol. Biosci. 7:609806. doi: 10.3389/fmolb.2020.609806

INTRODUCTION

Obesity is considered the most crucial factor in the development of several metabolic diseases such as T2DM. The prevalence of obesity in the Middle East is increasing, where about 70% of males above 20 years old are overweight compared to females (74%) in Saudi Arabia (Ng et al., 2014). Obesity and T2DM are conditions that are interlinked biochemically, metabolically, and at multiple levels making it difficult to discern the differences in their pathologies. Although closely interconnected, not all individuals with obesity develop diabetes and remain metabolically healthy, while a majority of patients diagnosed with T2DM are obese (predominant central obesity). The strong bidirectional relationship existing between obesity and T2DM is causally linked by IR (Al-Goblan et al., 2014) as a result of an increased chronic low-grade inflammation and oxidative stress, which are characteristics of both states. IR is characterized by a decreased tissue responsiveness to circulating insulin levels leading to defects in uptake and oxidation of glucose, a decrease in glycogen synthesis, decreased ability to suppress lipid oxidation, and the existence of a pro-oxidant state. The presence of IR far precedes the onset and presentation of the clinical symptoms of T2DM due to HG, delaying its prediction, diagnosis, and management by several years (Sas et al., 2015).

Insulin resistance, commonly observed in patients with obesity, affects multiple organs, including the adipose tissue, muscle, and liver, and attenuates insulin signaling pathways. In obese individuals, adipose tissue releases bigger amounts of NEFAs that promote triglyceride accumulation, resulting in worsening IR and β-cell dysfunction (Scheen, 2003). Shortly after an acute increase in plasma NEFA levels in humans, IR starts to develop. On the other hand, when the level of plasma NEFA decreases, as in antilipolytic agent used cases, peripheral insulin uptake improves. It has also been proposed the connection of NEFA and fatty acids delivery and intracellular metabolism to the levels of intracellular content of fatty acid metabolites such as diacylglycerol (DAG), which activates a serine (Ser)/threonine kinase cascade leading to Ser/threonine phosphorylation of IRS-1 and INSR substrate-2 (IRS-2), and a reduced ability of these molecules to activate PI3K (Snel et al., 2012). Subsequently, events downstream of INSR signaling are diminished due to the lipotoxicity giving rise to IR in obese individuals. Initiation of IR forms the first phase in the pathogenesis of T2DM followed sequentially by elevations in plasma glucose levels (that stimulate β-cells to secrete higher amounts of insulin), oxidative stress (that accelerates β-cell insufficiency) exacerbating the existing HG, ultimately leading to apoptosis (β-cell death) and development of overt T2DM (Cernea and Dobreanu, 2013).

Metabolomics has recently become a powerful method to measure subtle biochemical changes in several diseases (Adamski, 2016). Several metabolites associated with IR and obesity have been identified in T2DM including BCAAs, AAA, mannose, fructose, α-hydroxybutyrate, and phospholipids (Wang et al., 2011; Ferrannini et al., 2013; Yu et al., 2016). The onset of T2DM is relatively long, and symptoms of T2DM can occur at a very late stage without acute metabolic disturbances, making it difficult for early diagnosis (Sas et al., 2015). The current clinical practice for T2DM diagnoses, such as fasting glucose and glucose tolerance tests, lacks efficient early diagnose of T2DM. Identification of sensitive *in vivo* biomarkers that could reflect the early onset of T2DM would be crucial for the identification of high-risk asymptomatic diabetic individuals for better prevention. To identify metabolomics patterns for IR and T2DM individuals, a high-performance CIL LC-MS was utilized in this study. Our goal is to identify potential metabolic biomarkers for IR and T2DM, other than HOMA-IR and HG. CIL is used to modify the chemical and physical properties of metabolites for much-improved separation and enhanced detection sensitivity, thereby increasing the number of detectable metabolites (Jacob et al., 2019a; Dahabiyeh et al., 2020). Using differential isotope labeling also provides more accurate and precise quantification of metabolite concentration differences in comparative samples (i.e., relative quantification) (Jacob et al., 2019a; Dahabiyeh et al., 2020).

MATERIALS AND METHODS

Subjects

All subjects were recruited from a primary healthcare hospital located at King Abdulaziz Medical City in Riyadh, Saudi Arabia. All subjects underwent a medical check-up at the Department of Medicine and were screened for medical history. The anthropometric measurements included weight, height, waist, and hip circumferences. The study participants comprised of three groups: 30 adults of normal weight, 26 obese adults, and 16 adults newly diagnosed with T2DM. Exclusion criteria included (1) patients with coronary event or procedure (myocardial infarction, unstable angina, coronary artery bypass surgery, or coronary angioplasty) in the previous 3 months; (2) patient on steroids; (3) hepatic disease (transaminase $>$ 3 times normal); (4) renal impairment (serum creatinine $>$ 1.5 mg/dL); (5) history of drug or alcohol abuse; (6) participation in any other concurrent clinical trials; (7) any other life-threatening diseases; and (8) use of an investigational agent within 30 days of study. Institutional review board (IRB) approval was obtained from both King Abdulaziz Medical City Ethics Committee (Protocol # RC12/105), and King Faisal Specialist Hospital and Research Center (KFSHRC) (RAC# 2170 013), and all study participants signed a written informed consent form. All volunteers were properly instructed to fast for 12 h before the day appointed for vein puncture.

Chemicals and Reagents

Liquid chromatography-mass spectrometry grade reagents, including water, ACN, methanol, and FA, were purchased from Fisher Scientific (Ottawa, ON). ^{13}C-DnsCl was available from Nova Medical Testing, Inc. (Edmonton, Canada) with the procedures published previously (Zhao et al., 2019).

Metabolomic Profiling Workflow

Supplementary Figure S1 shows the schematic of the overall metabolomics analysis workflow. Each sample was derivatized by ^{12}C-DnsCl, while a pooled sample generated by mixing of

aliquots of all individual samples was labeled by ^{13}C-DnsCl. The ^{13}C-labeled pool sample served as an internal standard for all ^{12}C-labeled individual samples. The sample amount of each sample was normalized using the LC-UV method (Wu and Li, 2012). The ^{12}C-labeled individual sample was mixed with the same mole amount of ^{13}C-labeled pool. The mixture was injected onto LC-MS. All the labeled metabolites were detected as peak pairs on mass spectra. The peak area ratios were used for quantitative metabolomics analysis; the same ^{13}C-labeled pool was spiked into all ^{12}C-labeled individual samples, and thus the peak ratio values of a labeled metabolite in different samples reflected the concentration differences of this metabolite in these samples. In other words, every ^{12}C-labeled metabolite from an individual sample had its corresponding ^{13}C-labeled metabolite in the pooled sample as a reference, resulting in high accuracy for relative quantification (Jacob et al., 2019a; Dahabiyeh et al., 2020).

Serum Samples and Dansylation Labeling

Serum samples including lean control (n = 30), obese (n = 26), and T2DM (n = 16) were collected and stored at −80°C. A 15 μL sample was used and the metabolites were extracted by protein precipitation with 45 μL of methanol. After 2 h incubation at −20°C, 45 μL of supernatant was dried and then mixed with 25 μL of water, 12.5 μL of ACN, 12.5 μL of sodium carbonate/sodium bicarbonate buffer, and 25 μL of ^{12}C-DnsCl or ^{13}C-DnsCl (18 mg/mL in ACN). The mixture was incubated at 40°C for 45 min and 5 μL of 250 mM NaOH were added and incubated for 10 min at 40°C. Twenty-five μL of 425 mM FA in 1:1 ACN/H_2O was added to consume excess NaOH.

LC-UV

Before LC-MS injections, sample normalization was performed to minimize variations in the total sample amount of individual samples when comparing samples. A step-gradient LC-UV method measured the total concentration of dansyl labeled metabolites (Wu and Li, 2012). In brief, 5 μL of the labeled sample was injected into a Phenomena's Kinetes C18 column (2.1 mm × 5 cm, 1.7 μm particle size, 100 Å pore size) connected to a Waters ACQUITY UPLC system (Waters, Milford, MA, United States). Mobile phase A was 0.1% (v/v) FA in 5% (v/v) ACN, and mobile phase B was 0.1% (v/v) FA in ACN. The 6.5 min LC gradient including: t = 0 min, 0% B; t = 1 min, 0% B; t = 1.1 min, 95% B; t = 2.6 min, 95% B; t = 3.1 min, 0% B, and the flow rate was 0.45 mL/min. PDA detector was operated at 338 nm. The area under the peak representing the total concentration of dansyl-labeled metabolites was integrated using Waters Empower (V6.00).

LC-MS

Each sample was labeled by ^{12}C-DnsCl and mixed in equal mole amount with a ^{13}C-labeled pool sample based on the quantification results from LC-UV analysis. The samples were analyzed by a Dionex Ultimate 3000 UHPLC System (Thermo Scientific, Sunnyvale, CA, United States) connected to Maxis II quadrupole time-of-flight (Q-TOF) mass spectrometer (Bruker, Billerica, MA, United States). The analytes were separated using a reversed-phase Eclipse Plus C18 column (2.1 mm × 10 cm, 1.8 μm particle size, 95 Å pore size) (Agilent Inc., Santa Clara, CA, United States). Mobile phase A was 0.1% (v/v) FA in 5% (v/v) ACN, and solvent B was 0.1% (v/v) FA in ACN. The LC gradient was: t = 0 min, 20% B; t = 3.5 min, 35% B; t = 18 min, 65% B; t = 21 min, 99% B; t = 34 min, 99% B, and flow rate of 0.18 mL/min. MS conditions were as follows: polarity, positive; dry temperature, 230°C; dry gas, 8 L/min; capillary voltage, 4500 V; nebulizer, 1.0 bar; endplate offset, 500 V; spectra rate, 1.0 Hz.

The quality control (QC) sample was prepared by mixing the ^{12}C- and ^{13}C-labeled pooled samples in equal mole. A QC injection was performed every 15 LC-MS sample runs. In total, there were 14 QC samples injected and analyzed. Peak pairs with ratio values having >±25% RSD in the QC samples were filtered out.

Data Analysis

The MS spectra of the detected analytes were converted into.cvs files using Bruker Daltonics Data Analysis 4.3 software. The raw data generated from multiple LC-MS runs were processed by peak picking, peak pairing, and peak-pair filtering to remove redundant peaks (IsoMS; Zhou et al., 2014). IsoMS files from each injection were aligned together based on the peak's accurate mass and retention time to generate the aligned file. The missing peak pair information in aligned files was re-extracted from raw data by Zerofill software (Huan and Li, 2015). The final metabolite-intensity data file was used for statistical analysis after normalization and/or scaling. The PLS-DA was performed by MetaboAnalyst[1].

Metabolite identification was carried out using the three-tier metabolite identification approach (Zhao et al., 2019). In tier 1, peak pairs were searched against a labeled metabolite library (CIL Library) based on accurate mass and retention time. The CIL Library (i.e., dansyl amines and phenols) contains 711 experimental entries, including metabolites and dipeptides (Huan et al., 2015). In tier 2, linked identity library (LI Library) was used for identification of the remaining peak pairs. LI Library includes metabolic-pathway-related metabolites (more than 7000 entries extracted from the KEGG database), providing high-confidence putative identification results based on accurate mass and predicted retention time matches. In tier 3, the remaining peak pairs were searched, based on accurate mass match, against the MyCompoundID (MCID) library composed of 8021 known human endogenous metabolites (zero-reaction library) and their predicted metabolic products from one metabolic reaction (375,809 compounds) (one-reaction library) and two metabolic reactions (10,583,901 compounds) (two-reaction library) (Li et al., 2013). The identified features were processed for building the IR and T2DM models using Multiple Professional Profiler (MPP) Software (Agilent Inc., Santa Clara, CA, United States) to construct the Venn diagrams, clustering heat maps, and the profiles.

[1] www.metaboanalyst.ca

RESULTS

Demographic Data of Study Participants

Demographic data of study participants are summarized in **Table 1**. The study groups are lean, obese, and newly diagnosed obese with T2DM. Both obese and T2DM groups were significantly older than the normal weight group. Few subjects from groups 2 and 3 were on statins or other cholesterol-lowering agents, angiotensin-converting enzyme inhibitors (ACE-I) or other anti-hypertensives, non-steroidal anti-inflammatory drugs, or anti-oxidants and they were on stable doses for the last 2 months of their participation in the study. T2DM group had significantly higher LDL and Trig and HDL when compared to the healthy disease-free groups.

Metabolomics Results

Based on the unique characteristics of the peak pair of the CIL LC-MS method, 3633 peak pairs were detected in the participants' samples (the full data can be found in https://www.ebi.ac.uk/metabolights/MTBLS2098) (**Supplementary Table S1**). The IsoMS software filtered out redundant peak pairs such as those from adduct ions, dimers, multimers, etc., to retain only one peak pair ($[M + H]^+$) for each metabolite. Thus, the number of peak pairs detected reflects the number of detected metabolites. From the detected peak pairs, 216 metabolites were positively identified using both retention time and accurate mass searching against the labeled metabolite library (CIL Library). One hundred thirty-five peak pairs were putatively identified based on accurate mass and predicted retention time matches by searching against the LI Library. Six hundred and eleven and 1296 metabolites were putatively matched with the zero-reaction and one-reaction library via accurate mass only by searching against the MCID library, respectively. Thus, 62.2% of the 3633 peak pairs detected were either identified or matched, which shows the significant coverage of the submetabolome using the dansylation labeling LC-MS method for the serum samples analyzed in this study.

TABLE 1 | Demographic data of study participants.

	Lean (n = 30)	Obese (n = 26)	T2DM (n = 16)
Age (years)	25.7 ± 5.77	36.2 ± 12.12†	49.4 ± 12.12†‡
Geniler (FM)	(11/19)	(17/9)	(4/12)
BMI (kg/m^2)	23.0 ± 1.48	38.8+8.59†	32.7 ± 7.61†
Glucose (mmol/L)	5.1 ± 0.5	5.4 ± 0.67	10.2 ± 4.99†‡
HbA1c	–	5.3 ± 1.73	8.7 ± 2.82†
LDL(mmol/L)	2.5 ± 0.94	3.0 ± 0.81	3.6 ± 0.72†
HDL (mmol/L)	1.35 ± 0.16	1.18 ± 0.27†	1.01 ± 0.21†
Trig (mmol/L)	0.83 ± 0.33	1.26 ± 0.61	2.01 ± 1.04†
Insulin (μU/mL)	4.59 ± 2.04	9.98 ± 6.17†	7.44 ± 8.53
HOMA-IR	1.03 ± 0.56	2.30 ± 1.44†	2.80 ± 2.55

Results are presented as Mean ± SD. †$p < 0.05$ vs normal weight subjects. ‡$P < 0.05$ vs obese.

Statistical Analysis Between Study Groups

Multivariate statistical analysis was performed to analyze the serum metabolome dataset. PLS-DA was first performed to reveal the distinct separation between the groups visually. The metabolome dataset was analyzed to see the separation between lean and obese groups, as shown in **Supplementary Figure S2A**, where the clusters of two groups were separated with $Q^2 = 0.737$ and $R^2 = 0.972$. Univariate analysis to further analyze the metabolome changes using volcano plots was performed on the metabolome set. In the volcano plot, the x-axis is the fold change (FC) of the obese group over the lean control group, and the y-axis is the p-value from the t-test for comparing the two groups., The q-value (false discovery rate) less than 0.05, and FC > 1.5 (or FC < 0.67) were used to determine metabolites with significance, that has been calculated using R Script. Herein, the cutoff p-value equals q-value, which is 0.05. The FC criterion chosen was based on the technical accuracy and reproducibility, i.e., for dansylation LC-MS, the errors and RSD values are less than ± 25%. Thus, we conservatively used ± 50% change as the criterion. In **Supplementary Figure S2B**, a total of 189 metabolites were dysregulated. Among them, 78 metabolites were up-regulated (FC > 1.5) and 111 metabolites were down-regulated (FC < 0.67) comparing obese to lean group. By searching against our dansyl standard library using these 189 metabolites, 30 of them were positively identified.

A clear separation was observed with $Q^2 = 0.885$ and $R^2 = 0.985$ from the PLS-DA score plot of T2DM and obese groups in **Supplementary Figure S2C**. The clear separation illustrates that obese and T2DM groups experienced some significant metabolome alterations. From the volcano plot showed in **Supplementary Figure S2D**, 459 metabolites were up-regulated (FC > 1.5), and 166 metabolites were down-regulated (FC < 0.67) comparing T2DM to obese groups. The cut-off p-value here is 0.038 (when q-value = 0.05). Sixty-seven metabolites out of 625 were positively identified using the Dnsyl library. PLS-DA analysis was also performed to the lean vs T2DM. Two clusters were well separated on the PLS-DA score plot with $Q^2 = 0.809$ and $R^2 = 0.977$ as shown in **Supplementary Figure S2E**. From the volcano plot showed in **Supplementary Figure S2F**, 189 metabolites were up-regulated (FC > 1.5), and 117 metabolites were down-regulated (FC < 0.67) comparing lean to T2DM groups. The cut-off p-value here is 0.068 (when q-value = 0.05). Three hundred and five metabolites were common between the three groups and were used in the downstream analysis for building the IR and T2DM models.

Metabolic Profile for Study Confounder

Metabolomics expression in human serum is highly sensitive to specific physiological changes such as age, BMI, LDL-cholesterol (LDL-C), etc. (Jacob et al., 2019b). In this study, BMI, age, and LDL-C are the main confounders that were considered in the downstream data analysis. The values of these confounders for the study participants were integrated into the metabolomics dataset. Pearson similarity test (R = 0.95–1) reveals 27 metabolites depend on BMI (**Figure 1A**), and 56 metabolites on age

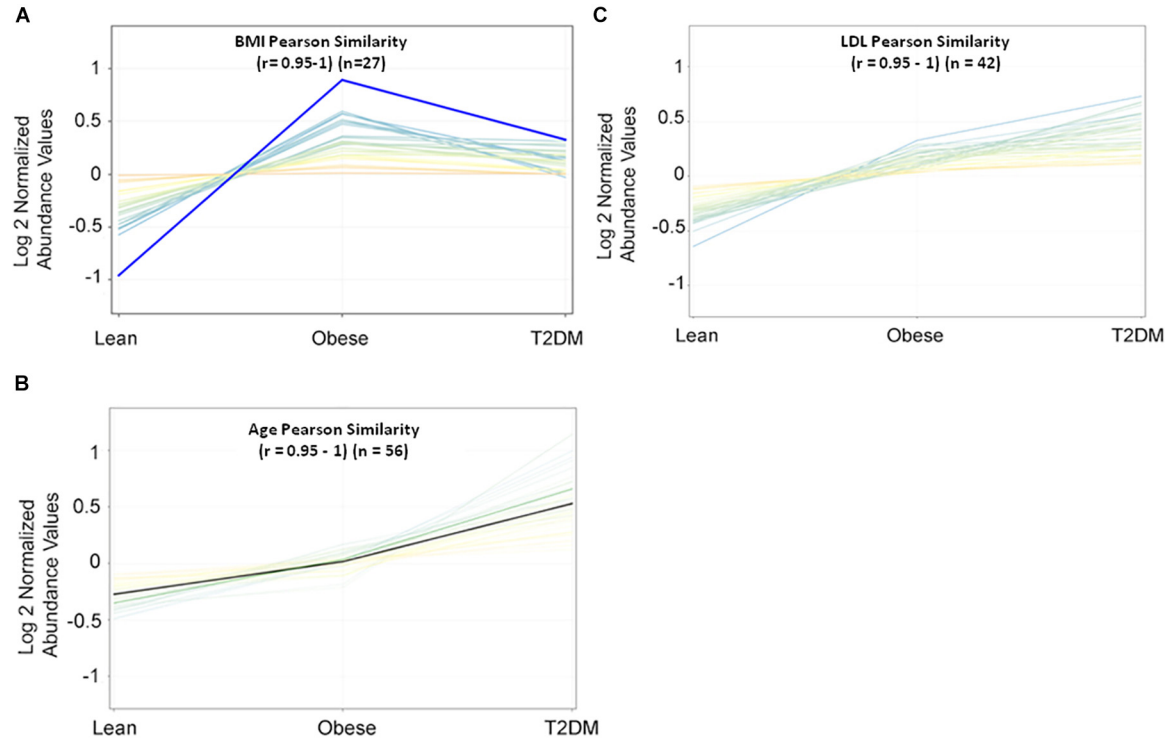

FIGURE 1 | Pearson similarity tests (r = 0.95–1) were filtered out with 27, 56, and 42 metabolites as **(A)** BMI-dependent; **(B)** age-dependent, and **(C)** LDL-C-dependent, respectively.

(**Figure 1B**), while 42 depend on LDL-C (**Figure 1C**). These confounders-related metabolites were excluded from IR and T2DM metabolic profiles. HOMA-IR and glucose-dependent metabolites were also determined using the same similarity approach as the other diabetic confounders and used to compare them with the final metabolic pattern of IR and T2DM.

IR Metabolic Pattern

Insulin resistance metabolic pattern was built using a model where dysregulated metabolites in obesity remained unchanged in T2DM compared to the lean group. After one-way ANOVA and Tukey honest significant difference (HSD) analysis, significantly different metabolites for each pair of groups were demonstrated in a Venn diagram by applying IR metabolic model on overall detected features, 351 identified and unidentified metabolites fell within this IR pattern (**Supplementary Figure S3A**), only 66 feature were up-regulated and 100 down-regulated in both obese and T2DM compared to lean as shown in **Supplementary Figures S3B–D**. The identified metabolites between the study groups (n = 305) were further analyzed to extract the IR metabolic pattern (**Figure 2A**). IR metabolic group were 43 metabolites that are statistically significant between both the lean vs obese and lean vs T2DM groups, and insignificant between obese vs T2DM. **Figure 2B** shows the breakdown of the 43 metabolites based on FC analysis (FC $>$ 1.5 or $<$0.67), where 18 metabolites (G18) were up-regulated in both obese and T2DM compared to lean group (**Figure 2C**), while nine metabolites (G9) were down-regulated in both obese and T2DM compared to lean (**Figure 2D**). Among the identified IR metabolic panel, the up- and down-regulated metabolites (G18, and G9, respectively) were further analyzed to exclude BMI-, age, and LDL-C-related metabolites (**Supplementary Figures S4A,B**). Only nine metabolites were down-regulated in obesity and T2DM and are independent of these three confounders (**Supplementary Figure S4B**).

T2DM Metabolomics Pattern

In this study, a metabolomics pattern for T2DM has been determined by extracting the metabolites that significantly unchanged between the lean and obese groups and significantly dysregulated in T2DM compared to both lean and obese groups. As shown in **Supplementary Figure S5A**, out of 3633 detected features, 605 metabolites were dysregulated in T2DM compared to both lean and obese groups based on one-way ANOVA Tukey HSD cutoff (FDRp $<$ 0.05). A total number of 529 significantly changed metabolites were filtered out based on FC analysis (FC $>$ 1.5 or $<$ 0.67) (**Supplementary Figure S5B**). One hundred eighty metabolites were down-regulated (G180) (**Supplementary Figure S5**C), and 349 (G349) were up-regulated (**Supplementary Figure S5D**) in T2DM compared to both lean and obese groups. Applying the same analysis on the identified molecules (n = 305), 62 metabolites were dysregulated in T2DM compared to both lean and obese groups (FDRp $<$ 0.05), as shown in **Figure 3A**. Fifty-six metabolites out of 60 were significantly dysregulated

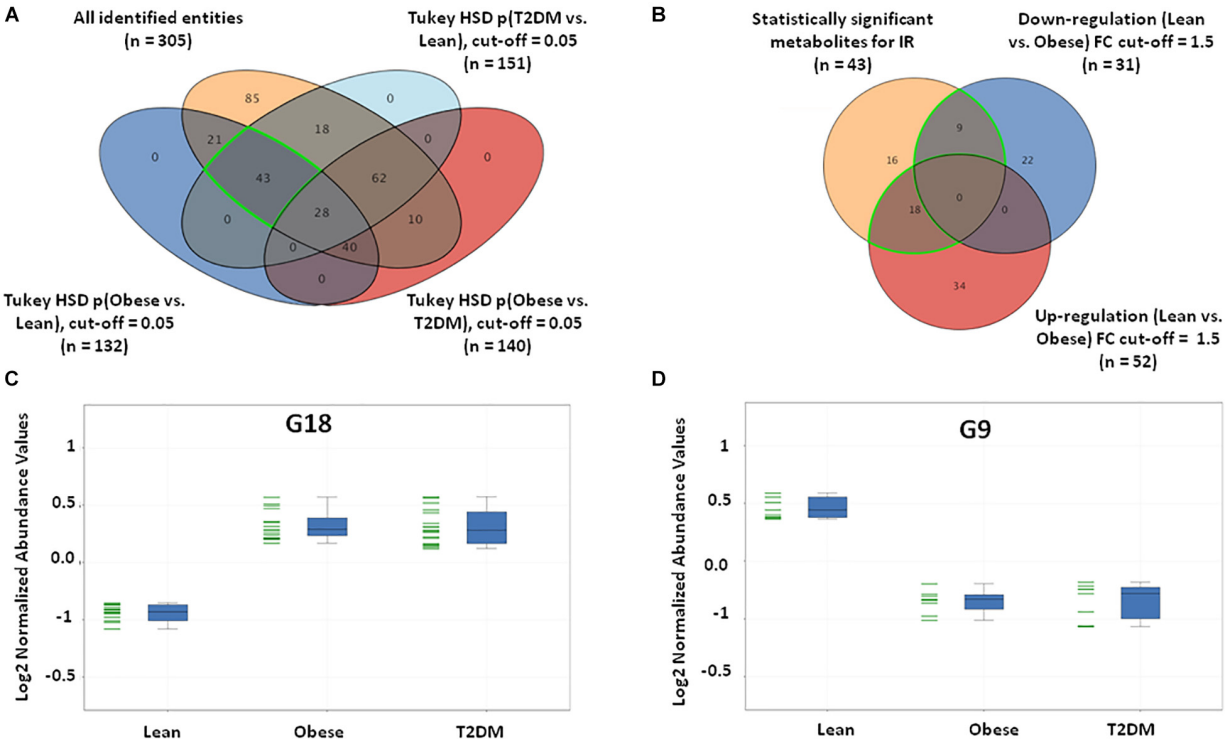

FIGURE 2 | IR metabolic model based on the positively identified features. **(A)** A Venn diagram showing that out of 3633 detected features, only 305 (8%) metabolites were positively identified. 43 (G43) significant metabolites were considered a metabolic pattern for IR, where they were statistically significant between lean vs obese and lean vs T2D, and insignificant between obese vs T2D (FDR-corrected p-value < 0.05). **(B)** A Venn diagram showing the number of the up- and down-regulated metabolites after fold change analysis (cutoff l.5) on the significant metabolites (G43) (nine and 18 metabolites, respectively). **(C)** The expression profile of G18 and **(D)** G9 metabolites.

based on FC analysis (FC $>$ 1.5 or $<$ 0.67) (**Figure 3B**), where 31 metabolites were up-regulated (**Figure 3C**), and 23 down-regulated (**Figure 3D**) in T2DM compared to other groups. After applying the confounder filters (BMI, age, and LDL-C) on these 54 metabolites (**Supplementary Figures S6A,B**), T2DM metabolic profile, 19, and 23 metabolites were up-regulated and down-regulated as BMI, age, and LDL-C-independent metabolites, respectively, as summarized in **Figure 4**.

Biomarkers Evaluation for IR and T2DM Metabolic Patterns

The nine metabolites of IR metabolic pattern are independent of glucose, HOMA-IR, and insulin as shown in **Supplementary Figures S7A–C**. On the other hand, the 42 metabolites of T2DM metabolic pattern were found to be HOMA-IR- and insulin-independent. However, only 11 metabolites were found to be glucose-dependent and up-regulated in T2DM compared to other groups (**Supplementary Figure S7A**). Heat maps that show cluster analysis of the entire average expression of each metabolite for IR specific panels (n = 9) (**Figure 5A**) and T2DM specific (n = 42) (**Figure 5B**) were generated after excluding all confounder related metabolites. The glucose-dependent metabolites are highlighted in these heat maps, which were created by Entities Hierarchical clustering for the average normalized data, and the similarity-based on Pearson. Also, two metabolites were gender-dependent and are up-regulated in males (prolyl-leucine and prolyl-isoleucine). The nine metabolites that represent IR metabolic pattern were analyzed as potential biomarkers, where the area under the curve (AUC) of the receiver operating characteristic (ROC) analysis was found 0.77 for the top changed five metabolites; serotonin, 2-methyl-3-hydroxy-5-formylpyridine-4-carboxylate, Asn, His, and methionine (Met) sulfoxide, when the comparison was done between the lean and obese groups (**Figures 6A,B**). Another comparison for the ability of this panel to predict IR in T2DM patients was performed for the same set of metabolites and found five metabolites (serotonin, 2-methyl-3-hydroxy-5-formylpyridine-4-carboxylate, His, Met sulfoxide, and 4, 6-dihydroxyquinoline) to have AUC 0.79 (**Figures 6C,D**).

The T2DM metabolic panel with 42 metabolites was evaluated for being used as potential biomarkers using the AUC of the ROC analysis. Pipecolate, cytidine, homogentistic acid, cystenyl-glycine, phosphoethanolamin, 7-caraboxy-7-carbaguanidine, glutaminyl-leucine, 3,4-dihydroxymandelate, hydroquinone, and alanyl-Ser were found to be the highest to predict hyperglycemic diabetic patients from lean and obese with AUC 0.958 and 0.975, respectively.

The top-scoring IPA metabolomic networks "cell-to-cell signaling and interaction, molecular transport, small molecule

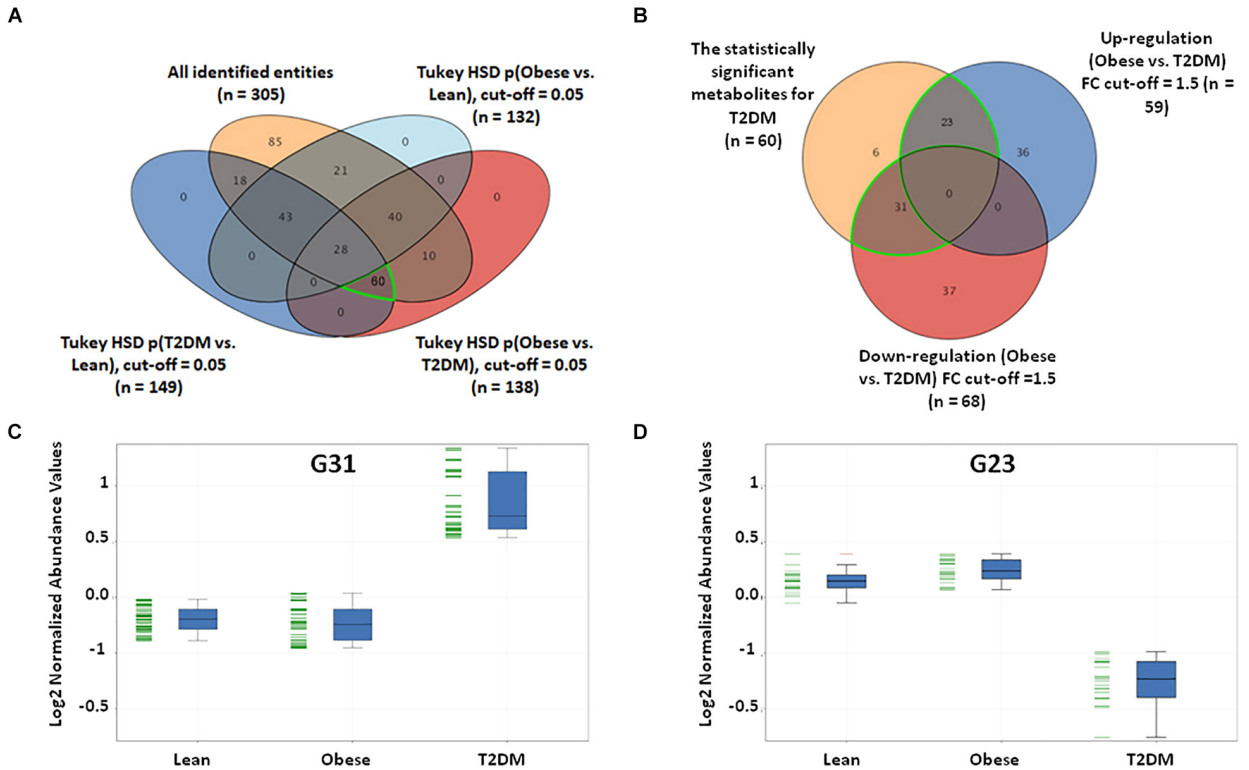

FIGURE 3 | T2DM metabolic model based on the identified features only. **(A)** Among the positively identified metabolites, 62 features were considered T2DM metabolic profile, where they are statistically significant between lean and T2DM, and obese and T2DM, and insignificant between lean and obese (FDR-corrected *p*-value < 0.051). **(B)** Fold change analysis (cutoff 13) separates these 62 significant metabolites into up- and down-regulated in T2DM compared to both lean and obese groups (31 and 23 metabolites, respectively). **(C)** Representative profile of the up-regulated (G31) and **(D)** down-regulated (G23) metabolites.

biochemistry" (**Figure 8A**), and "cellular compromise, lipid metabolism, small molecule biochemistry" (**Figure 8B**), are depicted for IR and T2DM metabolic patterns, respectively.

DISCUSSION

Insulin resistance is observed when higher than normal insulin concentrations are needed to achieve normal metabolic responses or when normal insulin concentrations fail to achieve a normal metabolic response (Kahn, 1978; Campbell et al., 1988). IR can be identified earlier than insulin secretion failure and is not always associated with the development of diabetes when islet cell secretion can keep up with normal insulin demand. Many methods and indices are available for the estimation of IR. At present, the most reliable reference methods available for estimating IR are hyperinsulinemic-euglycemic clamp and intravenous glucose tolerance test. The glucose clamp approach has several limitations such as time-consuming, labor-intensive, expensive, and requires an experienced operator to manage the technical difficulties. Other simple methods, from which indices can be derived, include homeostasis model assessment (HOMA-IR), quantitative insulin sensitivity check index (QUICKI), and Matsuda index developed by Matsuda and DeFronzo (1999). These indices are used in epidemiological and clinical studies to predict diabetes development in a non-diabetic population. HOMA-IR is a model of the relationship between insulin and glucose dynamics that predicts fasting steady-state insulin and glucose concentrations for a wide range of possible combinations of IR and β-cell function. HOMA-IR values inversely connected to insulin sensitivity (Gutch et al., 2015). Nevertheless, HOMA-IR has limitations in subjects with a lower BMI, a lower β-cell function, and high fasting glucose levels such as lean T2DM with insulin secretory defects (Kang et al., 2005). Recently, Quantose IR Test has been introduced commercially. Quantose IR Test is a fasting blood test that measures a panel of biomarkers comprised of a small organic acid [alpha-hydroxybutyric acid (AHB), two lipids (oleic acid and linoleoylglycerophosphocholine (LGPC)], and insulin. In our present study, the top five metabolites; serotonin, 2-methyl-3-hydroxy-5-formylpyridine-4-carboxylate, Asn, His, and Met sulfoxide were found to have a significantly high discriminatory capacity in identifying IR. Aside from these, we also identified three amino acids (Gln, Asn, and His) and 4,6 dihydroxyquinoline [product of tryptophan (Trp) metabolism] to be associated with the IR metabolic profile. A high-fat diet in animal models reduced the levels of Gln and Asn (gluconeogenic amino acids) and 4,6-dihydroxyquinoline (Liu et al., 2017). Higher levels of Gln and His (suppressor of hepatic gluconeogenesis) are known to be significantly associated with a lower risk for incident T2DM (Kimura et al., 2013; Chen et al.,

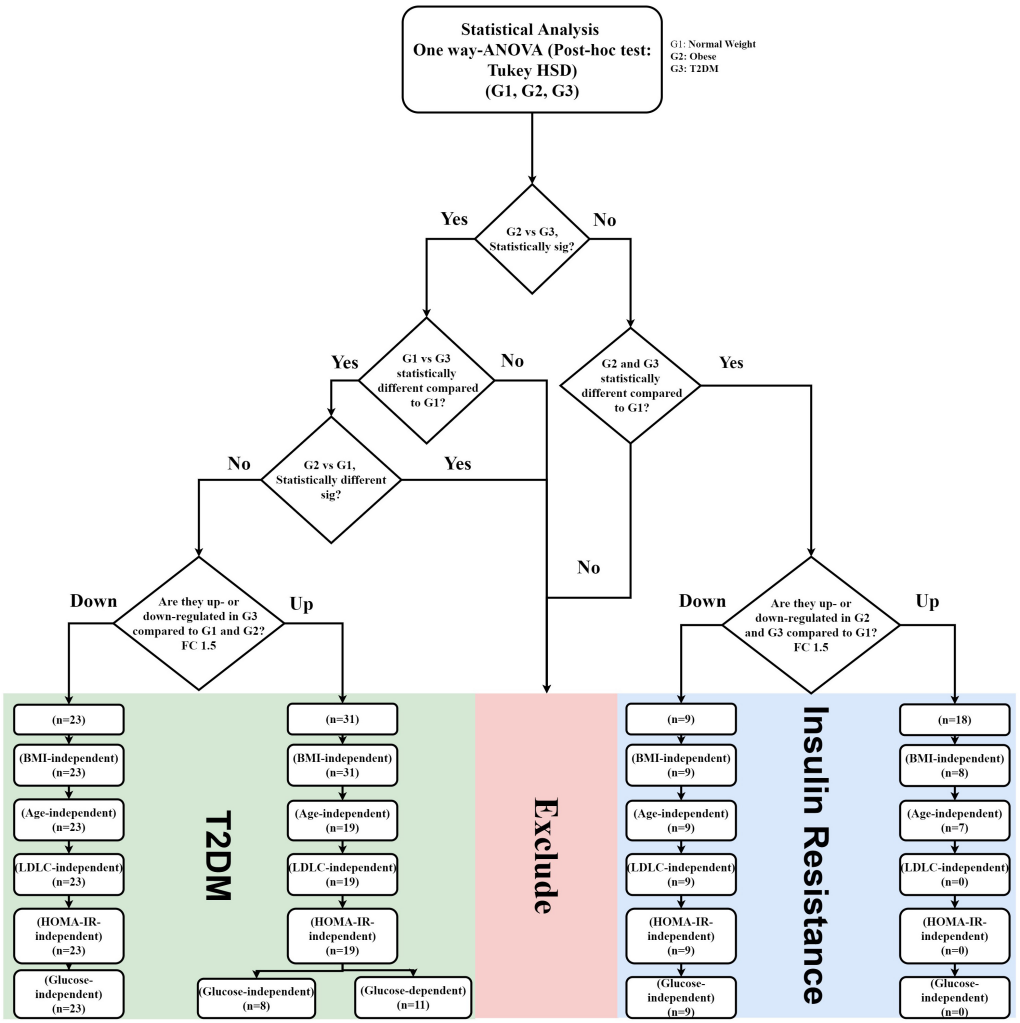

FIGURE 4 | IR and T2DM metabolic panel data mining scheme. IR metabolic model was generated based on the fold change difference (FC cutoff 1.5) and statistically significant features between lean vs obese and lean vs T2D, and insignificance between obese vs T2D (FDR-corrected p-value < 0.05; $n = 27$ metabolites; 18 up-regulated, and nine down-regulated) T2DM metabolic panels were built based on fold change (FC cutoff 1.5), and statistical significance between lean and T2DM, and obese and T2DM, and insignificance between lean and obese (FDR-corrected p-value < 0.05) ($n = 54$ metabolites: 31 were up-regulated, and 23 down-regulated). These panels underwent several filtration stages to exclude BMI, age, and LDL-C effects. Eventually, the remaining metabolites in each panel were correlated with HOMA-IR and glucose levels.

2019) while lower levels of circulating Asn are associated with increased BMI, IR, and HG (Banerji, 2015; Ottosson et al., 2018). Consistent with these findings, we found a decrease in the levels of Asn in obese T2DM compared to the obese and the lean groups, while the levels of Gln and His were lowered more in the obese than the obese T2DM patients.

Several studies have shown a correlation between certain amino acids and the development of diabetes years later. The mechanism by which elevations in plasma of certain amino acids links to the development of T2DM is currently unclear (Yamada et al., 2015; Chen et al., 2019; Vangipurapu et al., 2019). Glutamate (Glu) was the most strongly associated metabolite with T2DM, followed by increased levels of BCAA (Ottosson et al., 2018). Glu and Asn were both associated with a composite endpoint of developing T2DM or coronary artery disease (CAD; Ottosson et al., 2018). On the other hand, high Gln concentrations were associated with a decreased risk of incident T2DM (Chen et al., 2019) and decreased blood glucose in adolescents with T1DM after exercise while IR was unaltered during the euglycemic clamp (Torres-Santiago et al., 2017). Gln supplementation also reduced waist circumference in overweight and obese humans and improved insulin sensitivity in DIO Wistar rats (Abboud et al., 2019). Similarly, His oral supplementation improved IR (DiNicolantonio et al., 2018). His metabolism in T2DM may affect insulin sensitivity. His metabolism by the gut microbiota, in some T2DM patients, increases imidazole propionate levels which can decrease insulin sensitivity (Koh et al., 2018). Reduction of Gln, Asn, and His in obese and T2DM in this study is consistent with the previous reports of their role in increasing insulin sensitivity.

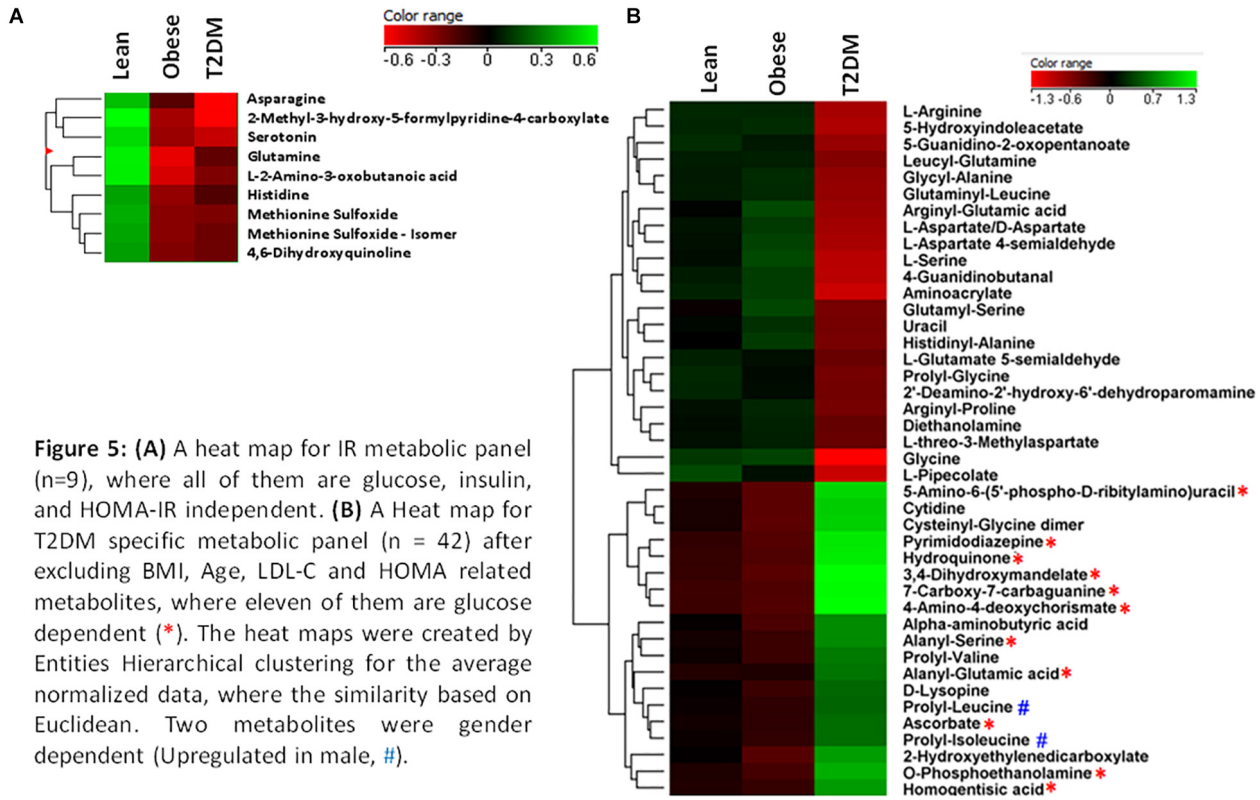

Figure 5: (A) A heat map for IR metabolic panel (n=9), where all of them are glucose, insulin, and HOMA-IR independent. **(B)** A Heat map for T2DM specific metabolic panel (n = 42) after excluding BMI, Age, LDL-C and HOMA related metabolites, where eleven of them are glucose dependent (*). The heat maps were created by Entities Hierarchical clustering for the average normalized data, where the similarity based on Euclidean. Two metabolites were gender dependent (Upregulated in male, #).

FIGURE 5 | (A) A heat map for IR metabolic panel (*n* = 9), where all of them are glucose, insulin, and HOMA-IR independent. **(B)** A heat map for T2DM specific metabolic panel (*n* = 42) after excluding BMI, age, LDL-C, and HOMA related metabolites, where 11 of them are glucose dependent (*). The heat maps were created by entities hierarchical clustering for the average normalized data, where the similarity based on Euclidean. Two metabolites were gender dependent (up-regulated in male, #).

Metabolites detected in the serum were decreased in obese and T2DM groups compared to lean subjects. Prominent decreases were also observed for metabolites from amino acids including 4,6-dihydroxyquinoline, Met sulfoxide, and L-2-amino-3-oxobutanoic acid. It is interesting to note that 4,6-dihydroxyquinoline was reported to be inhibited in high-fat diet-fed rats compared to normal diet controls (Liu et al., 2017). On the other hand, Met is one of the most susceptible to reactive oxygen species (ROS), resulting in both S and R diasteroisomeric forms (oxidation) of Met sulfoxide. Two Met residues in serum albumin (Met-111 and Met-147) are highly oxidized to Met sulfoxide in patients with diabetes (Suzuki et al., 2016) and the higher Met sulfoxide content in apoA-I from diabetic patients is consistent with lipid peroxidation products levels in plasma (Brock et al., 2008). Oxidative damage, mainly Met sulfoxide residues apolipoprotein B100 of LDL, was also increased in T2DM (Rabbani et al., 2010). The data on Met sulfoxide in this study contradict these studies and require further examination to elucidate the reduction of Met sulfoxide and its relation to IR.

Other serum metabolites that were reduced in IR include 2-methyl-3-hydroxy-5-formylpyridine-4-carboxylate, and serotonin or 5-hydroxytryptamine (5-HT). Reduction in 2-methyl-3-hydroxy-5-formylpyridine-4-carboxylate, an intermediate metabolite in vitamin B6 metabolism, observed in this study is consistent with the fact that vitamin B6 has been reported to help regulate blood glucose levels and insulin release (Liu et al., 2016). Moreover, low B6 levels have been associated with diabetic complications, such as neuropathy and retinopathy (Ellis et al., 1991; Nix et al., 2015), and to help in reducing diabetes complications (Kannan and Jain, 2004). Serotonin improves insulin sensitivity through serotonylation of Rab4, which likely represents the converging point between insulin and serotonin signaling cascades (Al-Zoairy et al., 2017). Bioinformatic and network pathway analysis carried out using IPA identified dysregulation of insulin as the central node in the pathway related to IR (**Figure 8A**). The second node with the highest connectivity in the network was serotonin, which showed that the highly interconnected regulation of serotonin with insulin altering the insulin signaling pathway. Moreover, 5-hydroxyindoleacetic acid (5HIAA), a breakdown product of serotonin, is down-regulated in the urine of diabetic patients. Serotonin plays a key role in controlling insulin secretion and its absence could lead to diabetes (Robinson, 2009). Elevation of the brain serotonin level may be regarded as an effective approach to treat T2DM and its complications (Derkach et al., 2015).

Branched-chain amino acids, including essential amino acids, play key roles in the energy homeostasis regulation, nutrition metabolism, gut health, immunity, and diseases. A positive

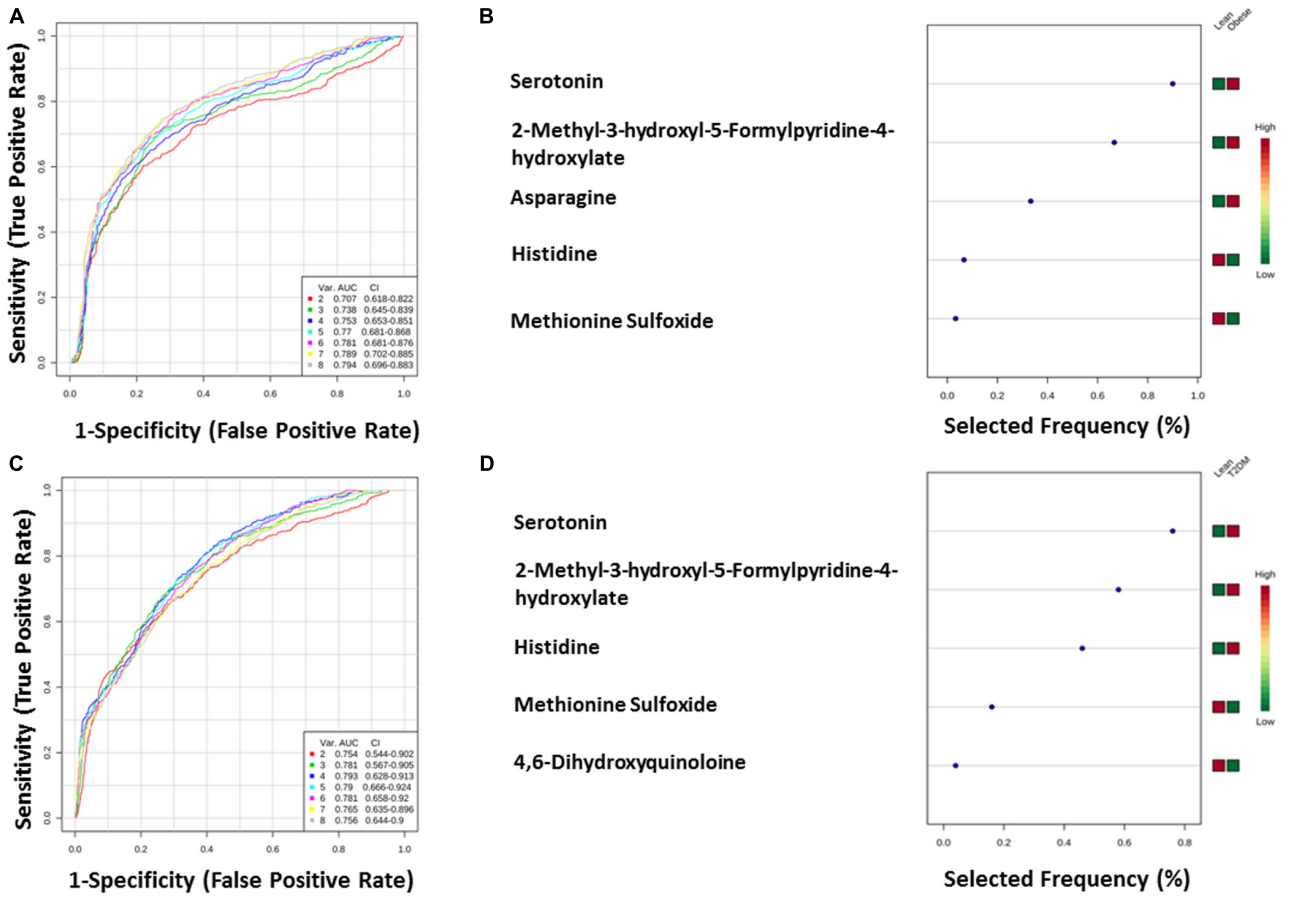

FIGURE 6 | (A,B) Lean vs obese ROC analysis for IR top five metabolites. **(C,D)** Lean vs T2DM.

association between increased circulating BCAAs with higher T2DM risk (Lotta et al., 2016) and IR in obese or diabetic patients (Zhao et al., 2016) has been reported. Similarly, AAA are strongly associated with the development of T2DM and IR (Yang et al., 2018). In a recent prospective study, aimed at identifying novel metabolic biomarkers predictive of future diabetes in 11,896 young adults from four Finnish cohorts, the strongest biomarkers of diabetes risk were BCAA and AAA (Ahola-Olli et al., 2019). In another targeted metabolomics platform, BCAA, AAA [phenylalanine (Phe) and tyrosine (Tyr)], Glu/Gln, Met, and C3 and C5 acylcarnitines were found to be strongly associated with IR (Newgard et al., 2009; Chen et al., 2016). A dramatic drop in BCAA and C3 and C5 acylcarnitines was observed in obese cases with T2DM following gastric bypass or gastric sleeve (Laferrere et al., 2011; Magkos et al., 2013). Furthermore, Leu, Ile, Val, Phe, and Tyr levels in plasma were also found to be associated with future development of T2DM (Wang et al., 2011; Chen et al., 2016). However, results seem to be controversial in different races, diets, and distinct tissues (Zhao et al., 2016). Recently, Lone et al. reported an association between five essential [Ile, Leu, lysine (Lys), Phe, and Val] and five non-essential [alanine (Ala), Glu, Gln, glycine (Gly), and Tyr] amino acids and the prevalence of T2DM (Lu et al., 2019). Association with the incidence of T2DM and four essential (Ile, Leu, Trp, and Val) and two non-essential (Gln and Tyr) amino acids was also reported while the accumulation of Gln and Gly was associated with T2DM lower risk (Lu et al., 2019). Moreover, abnormal circulating amino acid profiles in obesity, T2DM, and metabolic syndrome as measured by UPLC-TQ-MS demonstrated a decline in serum Gly and an increase in Val, Ile, Glu, and proline (Pro) in obesity, metabolic syndrome, and T2DM (Okekunle et al., 2017). In our study, arginine (Arg), Ser, Asp, and Gly were inhibited in T2DM whereas Asn, Gln, and His serum levels were lower in IR. The metabolomics profile of T2DM in our study showed the involvement of AAA through the presence of their intermediates. Metabolites of Phe, namely, homogentistic acid and 4-Amino-4-deoxychorismate, and of Tyr; 3,4 hydroxymandelic acid and hydroquinone, were identified. All these variations stem from using different methodologies and instrumentations and the fact that T2DM is a disease caused by a complex interchange between genetic, epigenetic, and environmental factors (diet and activity level), and that diabetes affects many major organs, including the heart, blood vessels, nerves, eyes, and kidneys. Additionally, genetic factors can make some people more vulnerable to diabetes. Thus, it

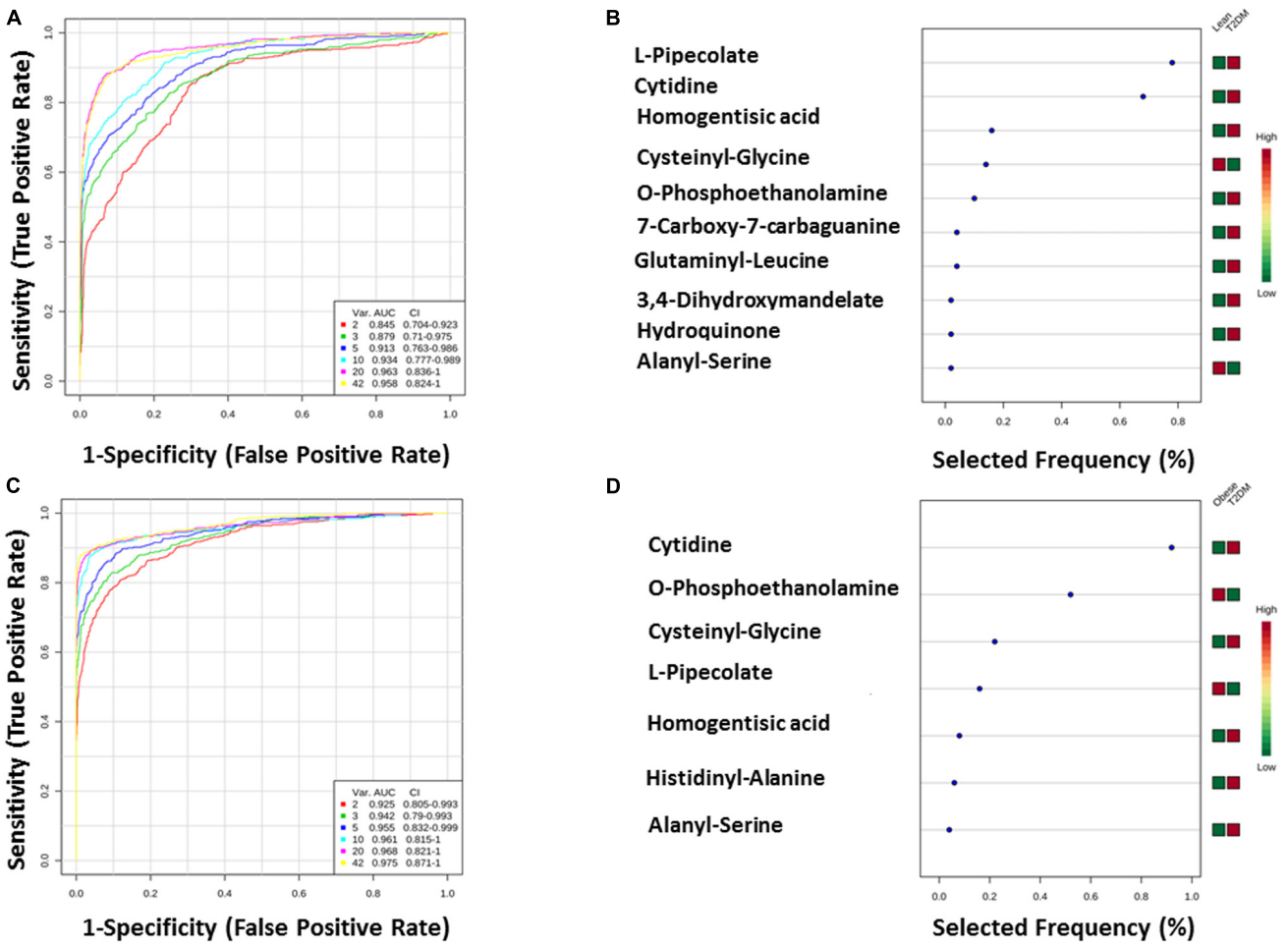

FIGURE 7 | (A,B) Lean vs T2DM ROC analysis for T2DM top 10 metabolites. **(C,D)** Obese vs T2DM.

is hard to identify distinct metabolic patterns that could serve as metabolic biomarkers for IR and T2DM. Moreover, in our study, we excluded age, BMI, and LDL-C in the analysis. This resulted in a reduction of the number of metabolites correlating with IR and T2DM. Age was considered as a confounder factor since both obese and T2DM groups were significantly older than the normal weight group. This represents a limitation for the study. Other confounders could impact IR and T2DM metabolome panels.

An interesting metabolite that was inhibited in T2DM in our study is pipecolate or 2 aminoadipic acid (2-AAA) which is an intermediate of the Lys degradation pathway. Previous studies have shown that circulating pipecolate levels were strongly associated with obesity and metabolic syndrome and had the ability to predict the risk of future T2DM, HG, increasing insulin secretion in early IR, and had a lesser role in the setting of advanced IR or T2DM (Wang et al., 2013; Libert et al., 2018). In turn, pipecolate reported to enhance insulin secretion in cell-based, islet, and animal model systems (Wang et al., 2013), and contribute to a compensatory mechanism by up-regulating insulin secretion to maintain glucose homeostasis in early IR. It has been found to independently act on β-cells of the pancreas to regulate the release of insulin at glucose-dependant concentrations. It augments the release of insulin at 2.5 mmol/L whereas higher levels of glucose (>11.1 mmol/L) inhibit this augmentation. This was also seen in our metabolic profile where levels of pipecolic acid were reduced in T2DM compared to the obese while those in the obese were higher than their lean counterparts. In another study, Wang et al. showed that treatment of diet-induced obesity with pipecolic acid significantly reduced body weight, fat accumulation, and lowered fasting glucose. Pipecolate regulating glycolipid metabolism is independent of diet and exercise, implying that improving its level can be a mean to treat diabetes (Xu et al., 2019). Our observations of a decrease in the levels of pipecolate in established T2DM cases are consistent with the previous findings and a possible explanation could be that pipecolate contribution to maintaining glucose homeostasis is overcome in established diabetes. In this scenario, an early measurement of pipecolate could serve as a novel potential metabolic marker of hyperglycemia that would predict a predisposition to T2DM and would be used in diabetes risk assessment.

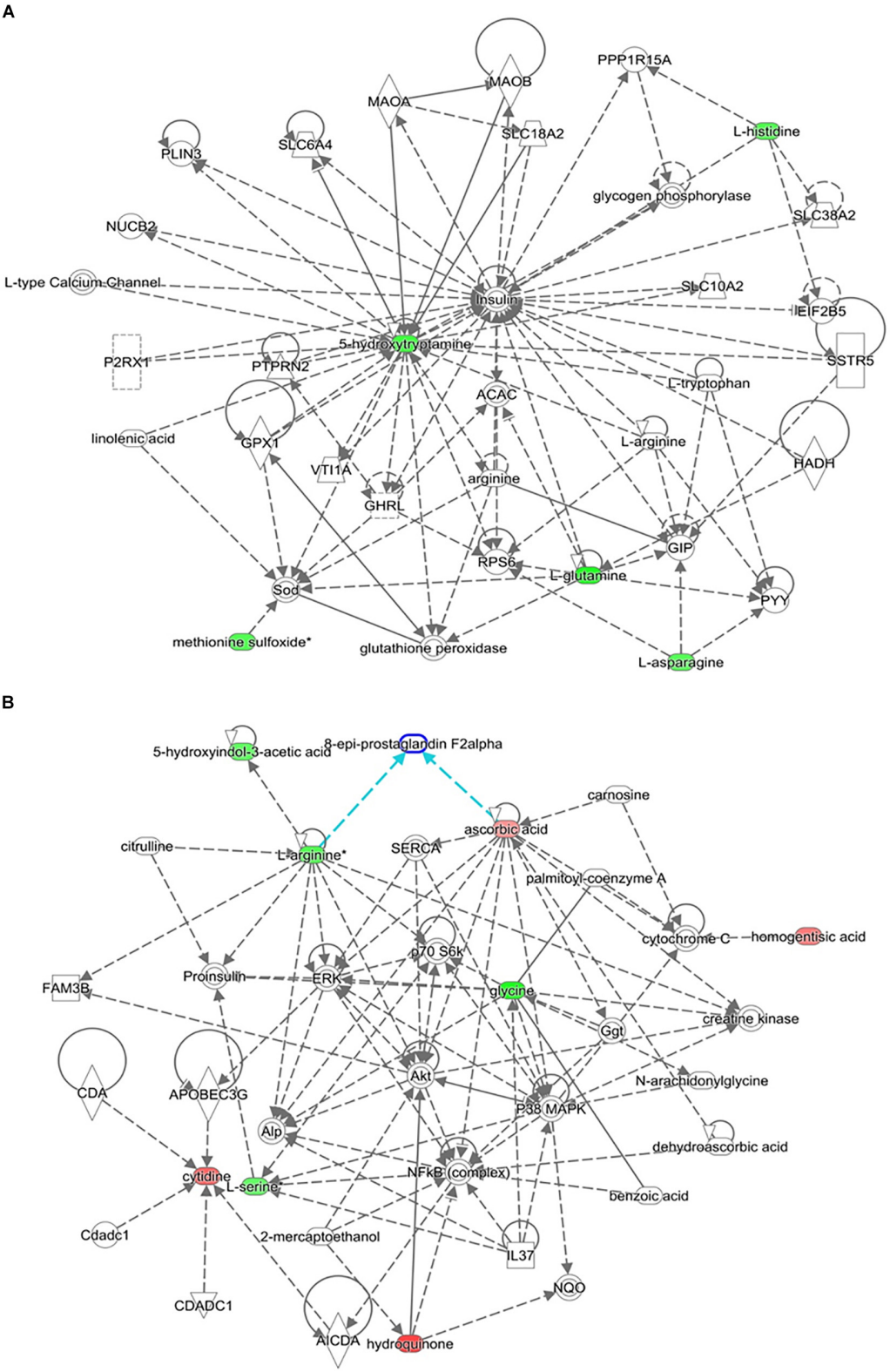

FIGURE 8 | Depicted top scoring I PA metabolomic networks. **(A)** "Cell-to-cell signaling and interaction, molecular transport, small molecule biochemistry" for IR metabolomic panel and **(B)** "cellular compromise, lip id metabolism, small molecule biochemistry" for T2DM metabolomic panel. The dotted lines indicate indirect and the straight lines indicate direct relationships. Nodes colored red represent up-regulation and green represent down-regulation. The interaction networks were generated through the use of IPA (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/).

Few dipeptides are known to be associated with physiological or cell-signaling effects such as ophthalmic acid in cystic fibrosis (DiBattista et al., 2019). However, most are simply short-lived intermediates on their way to degradation pathways following further proteolysis. His-Ala is a dipeptide resulting from incomplete breakdown of protein digestion or protein catabolism. In this study, His-Ala is inhibited in T2DM. Interestingly, His-Ala has been patented for reducing uric acid (Patent# JP2004359663A). High levels of uric acid in the blood are associated with increased risk of developing diabetes (Xiong et al., 2019). On the other hand, T2DM is associated with high serum uric acid levels and thus levels of His-Ala could be inhibited by uric acids. Peptidases play a pivotal role in the production, degradation, and regulation of peptides *in vivo* (Tiruppathi et al., 1990; Rosenblum and Kozarich, 2003). Prolyl peptidases are characterized by a biochemical preference for cleaving Pro-containing peptides. Prolyl peptidases family includes prolyl endopeptidase, prolyl endopeptidase-like, dipeptidyl peptidase 4 (DPP4), DPP7, DPP8, DPP9, and fibroblast activation protein (Lone et al., 2010). DPP4 (also known as CD26) selectively cleaves dipeptides from peptides and proteins containing Pro or Ala in the N-terminal penultimate position (Gorrell et al., 2001; Kirby et al., 2009). This proteolysis can alter activities of target substrates, including the functional activity of bioactive peptides or facilitated degradation of macromolecules by other peptidases. DPP4 plays a major role in glucose and insulin metabolism (Rohrborn et al., 2015). DPP4 does cleave Ala containing dipeptides and thus lowering levels of His-Ala could result from higher levels of DPP4 in diabetes. Similarly, six serum dipeptides had concentrations lower in T2DM than obese and control groups. These include Glu-Ser, Gly-Ala, Prol-Gly, Leu-Gln, Arg-Pro, and Arg-Glu dipeptides. Lower serum levels of these peptides could result from increased peptidases and could represent a target for T2DM treatment. On the other hand, other serum dipeptides concentrations are increased in T2DM. These include dipeptides Cys-Gly, Ala-Ser, Ala-Glu, and BCAA-Pro dipeptides, Pro-Val, Pro-Leu, Pro-Ile. Among these dipeptides, only Ala-Glu and Ala-Ser are glucose-dependent metabolites. Further investigations are needed to clarify the role of these dipeptides in diabetes.

We identified phosphoethanolamine (PE) and diethanolamine in T2DM metabolomics profile. PE and diethanolamine are substrates for the synthesis of phospholipids, phosphatidylcholine (PC), and phosphatidylethanolamine (PtdE). PE is the crucial metabolite determining the rate-limiting step of the reaction, to produce CDP–ethanolamine via the cytidine dependant CDP-ethanolamine metabolic pathways, which, together with DAG, generates PtdE (Steenbergen et al., 2005). Phospholipids are major components of all cellular membranes with PtdE being the major phospholipid of the mitochondrion. The CDP-ethanolamine pathway is an important regulator of hepatic lipid homeostasis (Leonardi et al., 2009). The involvement of PtdE and more importantly the PtdE/PC ratio points to alterations in phospholipid pathway that regulates muscle IR, insulin sensitivity, and to the presence of endoplasmic reticulum stress and mitochondrial dysfunction (Fu et al., 2011; Meikle et al., 2013). Interestingly, we also identified two metabolites participating in the folate metabolism, namely, 7-craboxy-7-carbaguanidine and pyrimidodiazepine. Folate is a cofactor known to regulate major metabolic pathways including the phospholipid pathways. Decreased levels of folate in an animal study showed optimal folate levels determine the synthesis of PC via the methylation of PtdE, giving further credence to our findings of dysregulation of phospholipids with T2DM (Zhao et al., 2018). Besides, folate deficiency is known to predispose to obesity, lipid disorders, and T2DM, and an increase in the metabolites involved in its synthesis may account for a compensatory mechanism (Li et al., 2017).

Another interesting finding in our metabolomic profiling of T2DM is the dysregulation of the pyrimidine metabolic pathway; cytidine, 5-amino-6-(5'-phospho-D-ribitylamino) uracil, uracil, and aminoacrylate. Cytidine is an important molecule required for the synthesis of di/triphosphates that act as a fundamental source of energy for cellular reactions and are involved in signaling pathways. Aside from being intracellular energy molecules, nucleotides play an important role as extracellular signaling molecules and have been described for adenosine (Burnstock and Novak, 2013) and uridine (Yamamoto et al., 2010) in terms of glucose regulation, IR, and diabetes. We found a decrease in the levels of uracil in our study. Previous studies have shown changes in the levels of nitrogen compounds, such as nucleotides, nucleosides, and their metabolites, vary considerably, depending on the degree of IR in obese subjects (Yang et al., 2018). By the same inference, this role can be extended to cytidine which is being reported in this study. Network pathway analysis relating to HG centered around dysregulation of signaling pathways related to ERK, p38 MAPK, and Akt (**Figure 8B**). Ser and threonine kinases are known regulators of cellular functions that include glucose metabolism, glycogen synthesis, protein synthesis, cell proliferation, cell hypertrophy, and cell death. These signaling pathways also regulate proinsulin, another node identified in the network map. Interactions of Akt and p38 MAPK are important mediators of insulin action via modulating INSR substrate and GLUT 4 activity. HG alters the Akt and members of the MAP kinase family signaling proteins there contributing to type 2 diabetes (Rane et al., 2001). Interpretation of the metabolome database thus contributes to the development of a comprehensive and accessible dataset of detected metabolites in obese and obese T2DM plasma samples for the discovery of disease associations and diagnoses in future research.

The current study is the lack of an independent external validation cohort. We plan to expand this work in the future by recruiting more subjects from multiple centers. Another limitation of the current study is the coverage of small molecules. While CIL LC-MS offers a high-coverage analysis of a chemical-group-based submetabolome, the current work only profiled the amine/phenol submetabolome. Other submetabolomes (e.g., acids, carbonyls, and hydroxyls) need to be examined in the future. In addition, we did not examine the lipidome. The results reported in this study demonstrated significant changes

in comparative groups, suggesting that global metabolome and lipidome analysis is warranted in future studies where larger cohorts of samples will be collected in order to increase both coverage and statistical power.

CONCLUSION

Metabolomics is an emerging approach for studying metabolic changes connected to disease development and progression. Metabolite-profiling techniques improvement is providing the increased extent of coverage of the human metabolome and advances have led to the application to defining predictive biomarkers and pathways for diseases including T2DM. Different metabolomic profiles have been reported in obesity and T2DM. This study identified nine metabolomics profile for IR and 42 for T2DM after the exclusion of confounders (age, BMI, and LDL-C). Identification and characterization of the metabolomic pattern in obese subjects might aid in identifying subjects at a high risk of developing metabolic diseases such as T2DM, thus allowing early treatment intervention. Future studies are required to establish causal relationships between metabolic biomarkers identified in this study for IR and T2DM and to examine their predictive values for developing T2DM.

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 human participants were reviewed and approved by King Abdulaziz Medical City Ethics Committee (Protocol # RC12/105), and King Faisal Specialist Hospital and Research Center (KFSHRC) (RAC# 2170 013). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ASA and AAR designed the projected, performed the data analysis, supervised the clinical and experimental activity, drafted the manuscript, and approved the final version. XG and LL generated the metabolomics data and performed part of the metabolome data analysis. AM and HB interpreted the biochemical pathways and helped in the final draft. MA, AwA, and MZ recruited the patients and collected the clinical data. All the authors approved the final version.

FUNDING

This project was funded by the Department of Genetics, Research Center, at King Faisal Specialist Hospital and Research Center (KFSHRC), and Alfaisal University-Office of Research and Graduate Studies (ORG).

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Executive Director of the Research Center, KFSHRC, to Dr. Brian Meyer, Chairman of the Department of Genetics for the funding supports, and to all the patients who contributed to this project. The work performed at the Li-node of the Metabolomics Innovation Centre (TMIC) of Canada was partially funded by Genome Canada and Canada Foundation for Innovation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.609806/full#supplementary-material

REFERENCES

Abboud, K. Y., Reis, S. K., Martelli, M. E., Zordao, O. P., Tannihao, F., De Souza, A. Z. Z., et al. (2019). Oral Glutamine Supplementation Reduces Obesity, Pro-Inflammatory Markers, and Improves Insulin Sensitivity in DIO Wistar Rats and Reduces Waist Circumference in Overweight and Obese Humans. *Nutrients* 11:536. doi: 10.3390/nu11030536

Adamski, J. (2016). Key elements of metabolomics in the study of biomarkers of diabetes. *Diabetologia* 59, 2497–2502. doi: 10.1007/s00125-016-4044-y

Ahola-Olli, A. V., Mustelin, L., Kalimeri, M., Kettunen, J., Jokelainen, J., Auvinen, J., et al. (2019). Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts. *Diabetologia* 62, 2298–2309. doi: 10.1007/s00125-019-05001-w

Al-Goblan, A. S., Al-Alfi, M. A., and Khan, M. Z. (2014). Mechanism linking diabetes mellitus and obesity. *Diabetes Metab. Syndr. Obes.* 7, 587–591. doi: 10.2147/dmso.s67400

Al-Zoairy, R., Pedrini, M. T., Khan, M. I., Engl, J., Tschoner, A., Ebenbichler, C., et al. (2017). Serotonin improves glucose metabolism by Serotonylation of the small GTPase Rab4 in L6 skeletal muscle cells. *Diabetol. Metab. Syndr.* 9:1.

Banerji, J. (2015). Asparaginase treatment side-effects may be due to genes with homopolymeric Asn codons (Review-Hypothesis). *Int. J. Mol. Med.* 36, 607–626. doi: 10.3892/ijmm.2015.2285

Brock, J. W., Jenkins, A. J., Lyons, T. J., Klein, R. L., Yim, E., Lopes-Virella, M., et al. (2008). Increased methionine sulfoxide content of apoA-I in type 1 diabetes. *J. Lipid. Res.* 49, 847–855. doi: 10.1194/jlr.m800015-jlr200

Burnstock, G., and Novak, I. (2013). Purinergic signalling and diabetes. *Purinergic Signal* 9, 307–324. doi: 10.1007/s11302-013-9359-2

Campbell, P. J., Mandarino, L. J., and Gerich, J. E. (1988). Quantification of the relative impairment in actions of insulin on hepatic glucose production and peripheral glucose uptake in non-insulin-dependent diabetes mellitus. *Metabolism* 37, 15–21. doi: 10.1016/0026-0495(88)90023-6

Cernea, S., and Dobreanu, M. (2013). Diabetes and beta cell function: from mechanisms to evaluation and clinical implications. *Biochem. Med.* 23, 266–280. doi: 10.11613/bm.2013.033

Chen, S., Akter, S., Kuwahara, K., Matsushita, Y., Nakagawa, T., Konishi, M., et al. (2019). Serum amino acid profiles and risk of type 2 diabetes among Japanese adults in the Hitachi Health Study. *Sci. Rep.* 9:7010.

Chen, T., Ni, Y., Ma, X., Bao, Y., Liu, J., Huang, F., et al. (2016). Branched-chain and aromatic amino acid profiles and diabetes risk in Chinese populations. *Sci. Rep.* 6:20594.

Dahabiyeh, L. A., Malkawi, A. K., Wang, X., Colak, D., Mujamammi, A. H., Sabi, E. M., et al. (2020). Dexamethasone-Induced Perturbations in Tissue Metabolomics Revealed by Chemical Isotope Labeling LC-MS analysis. *Metabolites* 10:42. doi: 10.3390/metabo10020042

Derkach, K. V., Bondareva, V. M., Chistyakova, O. V., Berstein, L. M., and Shpakov, A. O. (2015). The Effect of Long-Term Intranasal Serotonin Treatment on Metabolic Parameters and Hormonal Signaling in Rats with High-Fat Diet/Low-Dose Streptozotocin-Induced Type 2 Diabetes. *Int. J. Endocrinol.* 2015:245459.

DiBattista, A., Mcintosh, N., Lamoureux, M., Al-Dirbashi, O. Y., Chakraborty, P., and Britz-Mckibbin, P. (2019). Metabolic Signatures of Cystic Fibrosis Identified in Dried Blood Spots For Newborn Screening Without Carrier Identification. *J. Proteome. Res.* 18, 841–854.

DiNicolantonio, J. J., Mccarty, M. F., and Jh, O. K. (2018). Role of dietary histidine in the prevention of obesity and metabolic syndrome. *Open Heart* 5:e000676. doi: 10.1136/openhrt-2017-000676

Ellis, J. M., Folkers, K., Minadeo, M., Vanbuskirk, R., Xia, L. J., and Tamagawa, H. (1991). A deficiency of vitamin B6 is a plausible molecular basis of the retinopathy of patients with diabetes mellitus. *Biochem. Biophys. Res. Commun.* 179, 615–619. doi: 10.1016/0006-291x(91)91416-a

Ferrannini, E., Natali, A., Camastra, S., Nannipieri, M., Mari, A., Adam, K. P., et al. (2013). Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. *Diabetes* 62, 1730–1737. doi: 10.2337/db12-0707

Fu, S., Yang, L., Li, P., Hofmann, O., Dicker, L., Hide, W., et al. (2011). Aberrant lipid metabolism disrupts calcium homeostasis causing liver endoplasmic reticulum stress in obesity. *Nature* 473, 528–531. doi: 10.1038/nature09968

Gorrell, M. D., Gysbers, V., and Mccaughan, G. W. (2001). CD26: a multifunctional integral membrane and secreted protein of activated lymphocytes. *Scand J. Immunol.* 54, 249–264. doi: 10.1046/j.1365-3083.2001.00984.x

Gutch, M., Kumar, S., Razi, S. M., Gupta, K. K., and Gupta, A. (2015). Assessment of insulin sensitivity/resistance. *Indian J. Endocrinol. Metab.* 19, 160–164.

Huan, T., and Li, L. (2015). Counting missing values in a metabolite-intensity data set for measuring the analytical performance of a metabolomics platform. *Anal. Chem.* 87, 1306–1313. doi: 10.1021/ac5039994

Huan, T., Wu, Y., Tang, C., Lin, G., and Li, L. (2015). DnsID in MyCompoundID for rapid identification of dansylated amine- and phenol-containing metabolites in LC-MS-based metabolomics. *Anal. Chem.* 87, 9838–9845. doi: 10.1021/acs.analchem.5b02282

Jacob, M., Gu, X., Luo, X., Al-Mousa, H., Arnaout, R., Al-Saud, B., et al. (2019a). Metabolomics Distinguishes DOCK8 Deficiency from Atopic Dermatitis: Towards a Biomarker Discovery. *Metabolites* 9, 274. doi: 10.3390/metabo9110274

Jacob, M., Lopata, A. L., Dasouki, M., and Abdel Rahman, A. M. (2019b). Metabolomics toward personalized medicine. *Mass Spectrom Rev.* 38, 221–238.

Kahn, C. R. (1978). Insulin resistance, insulin insensitivity, and insulin unresponsiveness: a necessary distinction. *Metabolism* 27, 1893–1902. doi: 10.1016/s0026-0495(78)80007-9

Kang, E. S., Yun, Y. S., Park, S. W., Kim, H. J., Ahn, C. W., Song, Y. D., et al. (2005). Limitation of the validity of the homeostasis model assessment as an index of insulin resistance in Korea. *Metabolism* 54, 206–211. doi: 10.1016/j.metabol.2004.08.014

Kannan, K., and Jain, S. K. (2004). Effect of vitamin B6 on oxygen radicals, mitochondrial membrane potential, and lipid peroxidation in H2O2-treated U937 monocytes. *Free Radic Biol. Med.* 36, 423–428. doi: 10.1016/j.freeradbiomed.2003.09.012

Kimura, K., Nakamura, Y., Inaba, Y., Matsumoto, M., Kido, Y., Asahara, S., et al. (2013). Histidine augments the suppression of hepatic glucose production by central insulin action. *Diabetes* 62, 2266–2277. doi: 10.2337/db12-1701

Kirby, M., Yu, D. M., O'connor, S., and Gorrell, M. D. (2009). Inhibitor selectivity in the clinical application of dipeptidyl peptidase-4 inhibition. *Clin. Sci.* 118, 31–41. doi: 10.1042/cs20090047

Koh, A., Molinaro, A., Stahlman, M., Khan, M. T., Schmidt, C., Manneras-Holm, L., et al. (2018). Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell* 175, 947–961. doi: 10.1016/j.cell.2018.09.055

Laferrere, B., Reilly, D., Arias, S., Swerdlow, N., Gorroochurn, P., Bawa, B., et al. (2011). Differential metabolic impact of gastric bypass surgery versus dietary intervention in obese diabetic subjects despite identical weight loss. *Sci. Transl. Med.* 3, 80–82.

Leonardi, R., Frank, M. W., Jackson, P. D., Rock, C. O., and Jackowski, S. (2009). Elimination of the CDP-ethanolamine pathway disrupts hepatic lipid homeostasis. *J. Biol. Chem.* 284, 27077–27089. doi: 10.1074/jbc.m109.031336

Li, J., Goh, C. E., Demmer, R. T., Whitcomb, B. W., Du, P., and Liu, Z. (2017). Association between Serum Folate and Insulin Resistance among U.S. Nondiabetic Adults. *Sci. Rep.* 7:9187.

Li, L., Li, R., Zhou, J., Zuniga, A., Stanislaus, A. E., Wu, Y., et al. (2013). MyCompoundID: using an evidence-based metabolome library for metabolite identification. *Anal. Chem.* 85, 3401–3408. doi: 10.1021/ac400099b

Libert, D. M., Nowacki, A. S., and Natowicz, M. R. (2018). Metabolomic analysis of obesity, metabolic syndrome, and type 2 diabetes: amino acid and acylcarnitine levels change along a spectrum of metabolic wellness. *PeerJ.* 6:e5410. doi: 10.7717/peerj.5410

Liu, R., Li, H., Fan, W., Jin, Q., Chao, T., Wu, Y., et al. (2017). Leucine Supplementation Differently Modulates Branched-Chain Amino Acid Catabolism, Mitochondrial Function and Metabolic Profiles at the Different Stage of Insulin Resistance in Rats on High-Fat Diet. *Nutrients* 9:565. doi: 10.3390/nu9060565

Liu, Z., Li, P., Zhao, Z. H., Zhang, Y., Ma, Z. M., and Wang, S. X. (2016). Vitamin B6 Prevents Endothelial Dysfunction, Insulin Resistance, and Hepatic Lipid Accumulation in Apoe (-/-) Mice Fed with High-Fat Diet. *J. Diabetes Res.* 2016:1748065.

Lone, A. M., Nolte, W. M., Tinoco, A. D., and Saghatelian, A. (2010). Peptidomics of the prolyl peptidases. *AAPS J.* 12, 483–491. doi: 10.1208/s12248-010-9208-y

Lotta, L. A., Scott, R. A., Sharp, S. J., Burgess, S., Luan, J., Tillin, T., et al. (2016). Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. *PLoS Med.* 13:e1002179. doi: 10.1371/journal.pmed.1002179

Lu, Y., Wang, Y., Liang, X., Zou, L., Ong, C. N., Yuan, J. M., et al. (2019). Serum Amino Acids in Association with Prevalent and Incident Type 2 Diabetes in A Chinese Population. *Metabolites* 9:14. doi: 10.3390/metabo9010014

Magkos, F., Bradley, D., Schweitzer, G. G., Finck, B. N., Eagon, J. C., Ilkayeva, O., et al. (2013). Effect of Roux-en-Y gastric bypass and laparoscopic adjustable gastric banding on branched-chain amino acid metabolism. *Diabetes* 62, 2757–2761. doi: 10.2337/db13-0185

Matsuda, M., and DeFronzo, R. A. (1999). Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. *Diabetes Care* 22, 1462–1470. doi: 10.2337/diacare.22.9.1462

Meikle, P. J., Wong, G., Barlow, C. K., Weir, J. M., Greeve, M. A., Macintosh, G. L., et al. (2013). Plasma lipid profiling shows similar associations with prediabetes and type 2 diabetes. *PLoS One* 8:e74341. doi: 10.1371/journal.pone.0074341

Newgard, C. B., An, J., Bain, J. R., Muehlbauer, M. J., Stevens, R. D., Lien, L. F., et al. (2009). A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* 9, 311–326. doi: 10.1016/j.cmet.2009.02.002

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., et al. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 384, 766–781.

Nix, W. A., Zirwes, R., Bangert, V., Kaiser, R. P., Schilling, M., Hostalek, U., et al. (2015). Vitamin B status in patients with type 2 diabetes mellitus with and without incipient nephropathy. *Diabetes Res. Clin. Pract.* 107, 157–165. doi: 10.1016/j.diabres.2014.09.058

Okekunle, A. P., Li, Y., Liu, L., Du, S., Wu, X., Chen, Y., et al. (2017). Abnormal circulating amino acid profiles in multiple metabolic disorders. *Diabetes Res. Clin. Pract.* 132, 45–58. doi: 10.1016/j.diabres.2017.07.023

Ottosson, F., Smith, E., Melander, O., and Fernandez, C. (2018). Altered Asparagine and Glutamate Homeostasis Precede Coronary Artery Disease and Type 2 Diabetes. *J. Clin. Endocrinol. Metab.* 103, 3060–3069. doi: 10.1210/jc.2018-00546

Rabbani, N., Chittari, M. V., Bodmer, C. W., Zehnder, D., Ceriello, A., and Thornalley, P. J. (2010). Increased glycation and oxidative damage to apolipoprotein B100 of LDL cholesterol in patients with type 2 diabetes and effect of metformin. *Diabetes* 59, 1038–1045. doi: 10.2337/db09-1455

Rane, M. J., Coxon, P. Y., Powell, D. W., Webster, R., Klein, J. B., Pierce, W., et al. (2001). p38 Kinase-dependent MAPKAPK-2 activation functions as 3-phosphoinositide-dependent kinase-2 for Akt in human neutrophils. *J. Biol. Chem.* 276, 3517–3523. doi: 10.1074/jbc.m005953200

Robinson, R. (2009). Serotonin's role in the pancreas revealed at last. *PLoS Biol.* 7:e1000227. doi: 10.1371/journal.pbio.1000227

Rohrborn, D., Wronkowitz, N., and Eckel, J. (2015). DPP4 in Diabetes. *Front. Immunol.* 6:386. doi: 10.3389/fimmu.2015.00386

Rosenblum, J. S., and Kozarich, J. W. (2003). Prolyl peptidases: a serine protease subfamily with high potential for drug discovery. *Curr. Opin. Chem. Biol.* 7, 496–504. doi: 10.1016/s1367-5931(03)00084-x

Sas, K. M., Karnovsky, A., Michailidis, G., and Pennathur, S. (2015). Metabolomics and diabetes: analytical and computational approaches. *Diabetes* 64, 718–732. doi: 10.2337/db14-0509

Scheen, A. J. (2003). Pathophysiology of type 2 diabetes. *Acta Clin. Belg.* 58, 335–341.

Snel, M., Jonker, J. T., Schoones, J., Lamb, H., De Roos, A., Pijl, H., et al. (2012). Ectopic fat and insulin resistance: pathophysiology and effect of diet and lifestyle interventions. *Int. J. Endocrinol.* 2012:983814.

Steenbergen, R., Nanowski, T. S., Beigneux, A., Kulinski, A., Young, S. G., and Vance, J. E. (2005). Disruption of the phosphatidylserine decarboxylase gene in mice causes embryonic lethality and mitochondrial defects. *J. Biol. Chem.* 280, 40032–40040. doi: 10.1074/jbc.m506510200

Suzuki, S., Kodera, Y., Saito, T., Fujimoto, K., Momozono, A., Hayashi, A., et al. (2016). Methionine sulfoxides in serum proteins as potential clinical biomarkers of oxidative stress. *Sci. Rep.* 6:38299.

Tiruppathi, C., Miyamoto, Y., Ganapathy, V., Roesel, R. A., Whitford, G. M., and Leibach, F. H. (1990). Hydrolysis and transport of proline-containing peptides in renal brush-border membrane vesicles from dipeptidyl peptidase IV-positive and dipeptidyl peptidase IV-negative rat strains. *J. Biol. Chem.* 265, 1476–1483.

Torres-Santiago, L., Mauras, N., Hossain, J., Weltman, A. L., and Darmaun, D. (2017). Does oral glutamine improve insulin sensitivity in adolescents with type 1 diabetes? *Nutrition* 34, 1–6. doi: 10.1016/j.nut.2016.09.003

Vangipurapu, J., Stancakova, A., Smith, U., Kuusisto, J., and Laakso, M. (2019). Nine Amino Acids Are Associated With Decreased Insulin Secretion and Elevated Glucose Levels in a 7.4-Year Follow-up Study of 5,181 Finnish Men. *Diabetes* 68, 1353–1358. doi: 10.2337/db18-1076

Wang, T. J., Larson, M. G., Vasan, R. S., Cheng, S., Rhee, E. P., Mccabe, E., et al. (2011). Metabolite profiles and the risk of developing diabetes. *Nat. Med.* 17, 448–453.

Wang, T. J., Ngo, D., Psychogios, N., Dejam, A., Larson, M. G., Vasan, R. S., et al. (2013). 2-Aminoadipic acid is a biomarker for diabetes risk. *J. Clin. Invest.* 123, 4309–4317.

Wu, Y., and Li, L. (2012). Determination of total concentration of chemically labeled metabolites as a means of metabolome sample normalization and sample loading optimization in mass spectrometry-based metabolomics. *Anal. Chem.* 84, 10723–10731.

Xiong, Q., Liu, J., and Xu, Y. (2019). Effects of Uric Acid on Diabetes Mellitus and Its Chronic Complications. *Int. J. Endocrinol.* 2019:9691345.

Xu, W. Y., Shen, Y., Zhu, H., Gao, J., Zhang, C., Tang, L., et al. (2019). 2-Aminoadipic acid protects against obesity and diabetes. *J. Endocrinol.* 243, 111–123.

Yamada, C., Kondo, M., Kishimoto, N., Shibata, T., Nagai, Y., Imanishi, T., et al. (2015). Association between insulin resistance and plasma amino acid profile in non-diabetic Japanese subjects. *J. Diabetes Investig.* 6, 408–415.

Yamamoto, T., Inokuchi, T., Ka, T., Yamamoto, A., Takahashi, S., Tsutsumi, Z., et al. (2010). Relationship between plasma uridine and insulin resistance in patients with non-insulin-dependent diabetes mellitus. *Nucleosides Nucleotides Nucleic Acids* 29, 504–508.

Yang, Q., Vijayakumar, A., and Kahn, B. B. (2018). Metabolites as regulators of insulin sensitivity and metabolism. *Nat. Rev. Mol. Cell Biol.* 19, 654–672.

Yu, D., Moore, S. C., Matthews, C. E., Xiang, Y. B., Zhang, X., Gao, Y. T., et al. (2016). Plasma metabolomic profiles in association with type 2 diabetes risk and prevalence in Chinese adults. *Metabolomics* 12:3.

Zhao, M., Yuan, M. M., Yuan, L., Huang, L. L., Liao, J. H., Yu, X. L., et al. (2018). Chronic folate deficiency induces glucose and lipid metabolism disorders and subsequent cognitive dysfunction in mice. *PLoS One* 13:e0202910. doi: 10.1371/journal.pone.0202910

Zhao, S., Li, H., Han, W., Chan, W., and Li, L. (2019). Metabolomic Coverage of Chemical-Group-Submetabolome Analysis: Group Classification and Four-Channel Chemical Isotope Labeling LC-MS. *Anal. Chem.* 91, 12108–12115.

Zhao, X., Han, Q., Liu, Y., Sun, C., Gang, X., and Wang, G. (2016). The Relationship between Branched-Chain Amino Acid Related Metabolomic Signature and Insulin Resistance: A Systematic Review. *J. Diabetes Res.* 2016:2794591.

Zhou, R., Tseng, C. L., Huan, T., and Li, L. (2014). IsoMS: automated processing of LC-MS data generated by a chemical isotope labeling metabolomics platform. *Anal. Chem.* 86, 4675–4679.

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 12 January 2021
doi: 10.3389/fmolb.2020.593976

The Role of Serum Metabolomics in Distinguishing Chronic Rhinosinusitis With Nasal Polyp Phenotypes

Shaobing Xie[1,2†], *Hua Zhang*[1,2†], *Yongzhen Liu*[1,2], *Kelei Gao*[1,2], *Junyi Zhang*[1,2], *Ruohao Fan*[1,2], *Shumin Xie*[1,2], *Zhihai Xie*[1,2], *Fengjun Wang*[1,2*] *and Weihong Jiang*[1,2*]

[1] *Department of Otolaryngology Head and Neck Surgery, Xiangya Hospital of Central South University, Changsha, China,*
[2] *Hunan Province Key Laboratory of Otolaryngology Critical Diseases, Changsha, China*

Background: Chronic rhinosinusitis with nasal polyps (CRSwNP) is a heterogeneous disease characterized by different clinical features and treatment responsiveness. This study aimed to compare the serum metabolomics profiles between eosinophilic CRSwNP (eCRSwNP) and non-eosinophilic CRSwNP (neCRSwNP) and healthy controls (HC) and explore objective biomarkers for distinguishing eCRSwNP before surgery.

Methods: Serum samples were collected from 33 neCRSwNP patients, 37 eCRSwNP patients, and 29 HC. Serum metabolomics profiles were investigated by ultra-high-performance liquid chromatography–mass spectrometry.

Results: The analysis results revealed that neCRSwNP, eCRSwNP, and HC exhibited distinctive metabolite signatures. In addition, eCRSwNP could be distinguished from neCRSwNP referring to their serum metabolic profiles, and the top ten different metabolites were citrulline, choline, linoleic acid, adenosine, glycocholic acid, L-serine, triethanolamine, 4-guanidinobutyric acid, methylmalonic acid, and L-methionine, which were related to several most important pathways including arginine and proline metabolism; glycine, serine, and threonine metabolism; linoleic acid metabolism; and purine metabolism. Among these distinctive metabolites, citrulline, linoleic acid, adenosine, and 4-guanidinobutyric acid showed good predictabilities, and the serum levels of citrulline, linoleic acid, and adenosine were significantly correlated with tissue eosinophil (T-EOS) percentage and T-EOS count.

Conclusion: eCRSwNP patients exhibited discriminative serum metabolic signatures in comparison with neCRSwNP patients and HC. These results suggested that metabolomics profiles contributed to understanding the pathophysiological mechanisms of CRSwNP and distinguishing its phenotypes

Keywords: chronic rhinosinusitis with nasal polyps, eosinophil, metabolomics, metabolites, biomarker

OPEN ACCESS

Edited by:
Philip Britz-McKibbin, McMaster University, Canada

Reviewed by:
Koel Chaudhury, Indian Institute of Technology Kharagpur, India
Michal Jan Markuszewski, Medical University of Gdansk, Poland

***Correspondence:**
Fengjun Wang
187308584@qq.com
Weihong Jiang
jiangwh68@126.com

[†] *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: *08 September 2020*
Accepted: *03 December 2020*
Published: *12 January 2021*

Citation:
Xie S, Zhang H, Liu Y, Gao K, Zhang J, Fan R, Xie S, Xie Z, Wang F and Jiang W (2021) The Role of Serum Metabolomics in Distinguishing Chronic Rhinosinusitis With Nasal Polyp Phenotypes. Front. Mol. Biosci. 7:593976. doi: 10.3389/fmolb.2020.593976

INTRODUCTION

Chronic rhinosinusitis (CRS) is a common inflammatory disease characterized by inonasal mucosa paranasal sinuses with nasal blockage, rhinorrhea, post-nasal discharge, and olfactory dysfunction (Yao et al., 2017). Previous studies reported that CRS affected ∼5.5–28% of the general population worldwide, and the prevalence still continues to increase

(Chitsuthipakorn et al., 2018; Grayson et al., 2019; Li et al., 2019; Yao et al., 2019). Based on the presence or absence of nasal polyp, CRS is grouped into chronic rhinosinusitis with nasal polyps (CRSwNP) and chronic rhinosinusitis without nasal polyps (CRSsNP) (Bayar Muluk et al., 2019; Qing et al., 2019; Hoy, 2020). Due to disease heterogeneity, CRSwNP is further classified into eosinophilic CRSwNP (eCRSwNP) and non-eosinophilic CRSwNP (neCRSwNP), and these two phenotypes have obviously different disease characteristics, treatments, and prognosis (Ho et al., 2018; Fujieda et al., 2019; Yao et al., 2020). In comparison with neCRSwNP, eCRSwNP exhibits more serious disease symptoms, a higher rate of comorbid asthma, poorer treatment response, and a higher risk of recidivism (Ho et al., 2018; Hoy, 2020). Thus, a pre-operative examination that could discriminate eCRSwNP from neCRSwNP was pivotal to develop personalized treatments and follow-up. However, endotyping of CRSwNP is extremely challenging for rhinologists because of a lack of objective approaches. Therefore, it is urgently needed to develop an objective indicator or biomarker to distinguish CRSwNP phenotypes before surgery that can improve the prognosis and long-term management strategies.

Metabolomics is a burgeoning omics technology which provides opportunities to establish a powerful exploratory tool for monitoring disease status and help to expound the pathogenesis of diseases (Kelly et al., 2017; Turi et al., 2018; Spertini, 2020). Recent studies utilized metabolomics analysis to evaluate the metabolic signature in airway inflammatory diseases, such as asthma (Reisdorph and Wechsler, 2013), allergic rhinitis (Ma et al., 2020), pneumonia (Ning et al., 2018), and chronic obstructive pulmonary disease (Adamko et al., 2015) and identified several biomarkers and major metabolic pathways which might improve the understanding of these disorders and develop novel therapy target. However, no previous study has employed metabolomics to analyze metabolites and metabolic pathway changes in the serum of CRSwNP patients and explore objective biomarkers to distinguish eCRSwNP before surgery.

Therefore, we aimed to evaluate the serum metabolic signatures of CRSwNP and explore the association between the metabolite differences and CRSwNP phenotypes. In the present study, ultra-high-performance liquid chromatography–mass spectrometry (UHPLC-MS) was utilized to investigate serum metabolic profiles in eCRSwNP, comparing with neCRSwNP and healthy control (HC). Linear regression analysis was conducted to assess the correlation between different metabolites and tissue eosinophil (T-EOS) percentage and T-EOS count in CRSwNP patients.

MATERIALS AND METHODS

Participants and Settings

We recruited 70 consecutive patients with CRSwNP from June 2018 to October 2018 in our tertiary clinic. CRSwNP was diagnosed referring to the guidelines of the European Position Paper on Rhinosinusitis and Nasal Polyps 2012 (Fokkens et al., 2012). Exclusion criteria are as follows: (1) other nasal or sinus diseases, such as fungal sinusitis, allergic rhinitis, cystic fibrosis, aspirin-exacerbated respiratory disease, and tumor; (2) treatment including antibiotics, oral or systemic corticosteroids, immunotherapy, or anti-allergic drugs 4 weeks before the surgery; (3) inflammatory, septic diseases or autoimmune diseases; (4) age <18 years or >75 years old; (5) severe heart, kidney, or other organ dysfunction; and (6) pregnant condition. All CRSwNP patients received routine preoperative examination, including blood tests, nasal endoscopy, computed tomography (CT) or magnetic resonance imaging (MRI), chest X-rays, and electrocardiography. All participants scored their nasal symptoms by utilizing the widely accepted visual analog scale (VAS) as previously described (Zhu et al., 2020). Preoperative CT score was recorded using the Lund–Mackay staging system (Lund and Mackay, 1993). A total of 29 age- and gender-matched healthy volunteers with no evidence of rhinitis or rhinosinusitis, diabetes mellitus, or inflammatory or autoimmune conditions were enrolled as healthy controls (HC).

Diagnosis of eCRSwNP and neCRSwNP

During the surgery, nasal polys were obtained from all patients with CRSwNP, then immersed in 10% formalin and embedded with paraffin wax. The embedded tissues were sectioned at 5-μm thickness and were stained with hematoxylin and eosin (H&E) for the visualization of eosinophils. The numbers of eosinophils, lymphocytes, neutrophils, and plasma cells were counted in 10 randomly selected high-power fields by two observers who were blinded to the clinical data. eCRSwNP was diagnosed when the tissue eosinophils (T-EOS) percentage was higher than 10% of total inflammatory cells, otherwise defined as neCRSwNP (Hu et al., 2012; Zhong et al., 2020).

Serum Sample Collection and Preparation

Fasting peripheral whole blood from CRSwNP patients and HC were collected with vacuum blood collection tubes in the morning. The blood samples were centrifuged at 1,200 g for 10 min at 4°C within 1 h of venipuncture. The serum samples were collected and stored at −80°C. Serum samples were thawed on ice and vortexed thoroughly. The serum samples were mixed with 300 μL methanol and vortexed for 30 s and incubated at −40°C for 1 h and centrifuged at 12,000 g for 10 min at 4°C. 100 μL of supernatant was transferred to a fresh tube vial for UHPLC-MS analysis (Dunn et al., 2011; Naz et al., 2014). The quality control (QC) sample was utilized as previous described to assess the stability and reliability of the analytical system (Liu et al., 2016).

UHPLC-MS Analysis

The untargeted metabolomic analysis was performed by utilizing a 1,290 Infinity series UHPLC System (Waters Corporation, Milford, MA, USA). The mobile phase was composed of 25 mmol/L ammonium acetate in water was applied as phase A, and 25 mmol/L ammonia in acetonitrile was used as phase B. The analysis procedure was processed as previously described (Zhao et al., 2019). The Triple TOF 6600 mass spectrometry (AB Sciex, Boston, MA, USA) was used to obtain spectra data, and the acquisition software (Analyst TF 1.7, AB Sciex, Framingham, MA, USA) continuously evaluated the full-scan survey MS data. In each cycle, the most intensive 12 precursor ions (intensity >

100) were chosen for MS/MS at collision energy (CE) of 30 eV. The cycle time was 0.56 s. Electrospray ionization (ESI) source conditions were set as previous study described (Liu et al., 2016; Zhao et al., 2019).

Data Processing and Analysis

MS raw data (.wiff) files were converted to the mzXML format by Proteo Wizard and processed by R package XCMS V3.2. The process includes peak deconvolution, alignment, and integration. Peak extraction and alignment were performed by Proteo Wizard and analyzed by R package as previous study described (Kuhl et al., 2012; Zhao et al., 2019). Metabolites identification refers to the In-house MS2 database. The processed data was exported to SIMCA (Version 14.1, Umetrics, Umea, Sweden) for multivariate analysis. Orthogonal partial least squares-discriminant analysis (OPLS-DA) was conducted to identify the major latent metabolites in the data matrix (Yang et al., 2020). The quality of the models was validated by R2Ycum (goodness of fit) and Q2cum (goodness of consistency). Meanwhile, the 200 permutations of cross-test were conducted to reduce the risk of overfitting and possibilities of false-positive findings. Metabolites contributing were calculated based on the variable importance for project (VIP) values (VIP $>$ 1.0) and *P*-values ($P < 0.05$) (Wang et al., 2018). A volcano plot was presented to project the metabolic regulations of the remarkable shifts in metabolites. The receiver operating characteristic (ROC) analysis was applied to the serum data to assess the performance of potential biomarker, and the area under the curve (AUC) was calculated to evaluate the sensitivity and specificity. To identify associated metabolic pathways, the pathway analysis was conducted using MetaboAnalyst 3.0.

Statistical Analysis

Continuous variables are described as mean ± standard deviation (SD). When the variables distributed normally, one-way analysis of variance (ANOVA) or Student's *t*-test was used, otherwise Kruskal–Wallis H test or Mann–Whitney *U*-test was performed. Discontinuous variables were described as number (percentage) and compared using Chi-square test. To evaluate the correlation between different metabolites and T-EOS percentage and T-EOS count in CRSwNP patients, Spearman's correlation analyses were performed. Significant difference was accepted when *P*-value $<$ 0.05. All statistical analyses were conducted on SPSS statistics software version 23.0 (IBM, Chicago, IL, USA).

RESULTS

Baseline Characteristics of All Subjects

Demographic and clinical characteristics of all subjects are listed in the **Table 1**. Among 70 CRSwNP patients, 33 (47.14%) patients were identified as eCRSwNP, and the other 37 (52.86%) patients were defined as neCRSwNP. Compared to the HC and neCRSwNP groups, the eCRSwNP group showed higher levels of blood eosinophil (B-EOS) count and B-EOS percentage (all $P < 0.001$). However, no statistical difference was observed in age, gender, rate of smoking, drinking, and BMI among three groups, and VAS score and Lund–Mackay score between the neCRSwNP and eCRSwNP groups. Typical histological findings of neCRSwNP and eCRSwNP are exhibited in **Figures 1A,B**. The T-EOS count and percentage in the eCRSwNP patients were significantly higher than those in the neCRSwNP (all $P < 0.001$, **Figures 1C,D**).

Metabolomic Signatures of neCRSwNP vs. HC

The OPLS-DA model exhibited a clear and distinctive clustering between neCRSwNP and HC (**Figure 2A**), R2X (cum), R2Y (cum), and Q2 were 0.236, 0.724, and 0.202, respectively. The OPLS-DA model was then assessed by permutation analysis, and all permuted R2s were below or around 0.6 and all permuted Q2s were below 0, which means that all R2s and Q2s are lower than the original on the right (**Figure 2B**). Thus, this suggests that the model fittings were valid and predictive. The potential differential metabolites were selected referring to the contribution of VIP (VIP $>$ 1 and $P < 0.05$). Finally, a total of 20 metabolites including 11 upregulated and nine downregulated for distinguishing neCRSwNP from HC were detected by UHPLC-MS analysis and they are shown in **Figure 2C**. In addition, metabolic pathway analysis results showed that cysteine and methionine metabolism and purine metabolism were the major involved metabolic pathways (**Figure 2D**).

Metabolomic Signatures of eCRSwNP vs. HC

Figure 3A shows that eCRSwNP patients were distinguished from HC based on serum metabolic profiles. In the OPLS-DA model, R2X (cum), R2Y (cum), and Q2 were 0.254, 0.695, and 0.391, respectively, and the model was assessed by permutation analysis, and analysis results suggested that the model fittings were valid and predictive (**Figure 3B**). Compared to HC, 49 metabolites were expressed at significantly different concentrations in the eCRSwNP group including 39 upregulated and 10 downregulated (**Figure 3C**). The most affected pathways including arginine and proline metabolism and linoleic acid metabolism are displayed in the **Figure 3D**.

Metabolomic Signatures of eCRSwNP vs. neCRSwNP

Figure 4A exhibits that serum metabolomic profiles of eCRSwNP patients had significantly different serum metabolomics profiles in comparison with neCRSwNP patients. In the OPLS-DA model, R2X (cum), R2Y (cum), and Q2 were 0.215, 0.509, and 0.244, respectively. The permutation analysis results demonstrated that the discriminating models were reliable (**Figure 4B**). In the eCRSwNP group, 24 metabolites were observed at different levels including 11 upregulated and 13 downregulated in comparison with the neCRSwNP group (**Figure 4C**). The most important pathways including arginine and proline metabolism; glycine, serine, and threonine metabolism; purine metabolism; and linoleic acid metabolism are displayed in the **Figure 4D**. Results of top 10 potential discriminant metabolites are displayed in the **Table 2**, and their relative serum concentrations between two groups are comparatively shown in **Figure 5**.

TABLE 1 | Clinical characteristics of subjects.

Variables	HC ($n = 29$)	neCRSwNP ($n = 33$)	eCRSwNP ($n = 37$)	*P*-value
Age (years), mean ± SD	28.5 ± 8.5	32.8 ± 12.2	35.1 ± 14.8	0.103
Gender (male/female), *n*	14/15	18/15	19/18	0.885
Smoking (yes/no), *n*	10/19	17/16	19/18	0.306
Drinking (yes/no), *n*	7/22	8/25	12/25	0.673
BMI (kg/m^2), mean ± SD	22.2 ± 1.8	22.7 ± 1.8	22.3 ± 1.5	0.570
B-EOS counts (10^6/L), mean ± SD	81.2 ± 24.0	174.7 ± 84.4	405.8 ± 159.3	< 0.001
B-EOS percentage, %	1.2 ± 0.8	2.2 ± 1.0	4.5 ± 1.2	< 0.001
VAS score, mean ± SD	–	5.5 ± 1.5	5.9 ± 1.7	0.300
Lund-Mackay score, mean ± SD	–	18.8 ± 3.9	18.2 ± 3.8	0.499

eCRSwNP, eosinophilic chronic rhinosinusitis with nasal polyps; neCRSwNP, non-eosinophilic chronic rhinosinusitis with nasal polyps; HC, healthy control; SD, standard deviation; BMI, body mass index; B-EOS, blood eosinophil; VAS, visual analog scale.

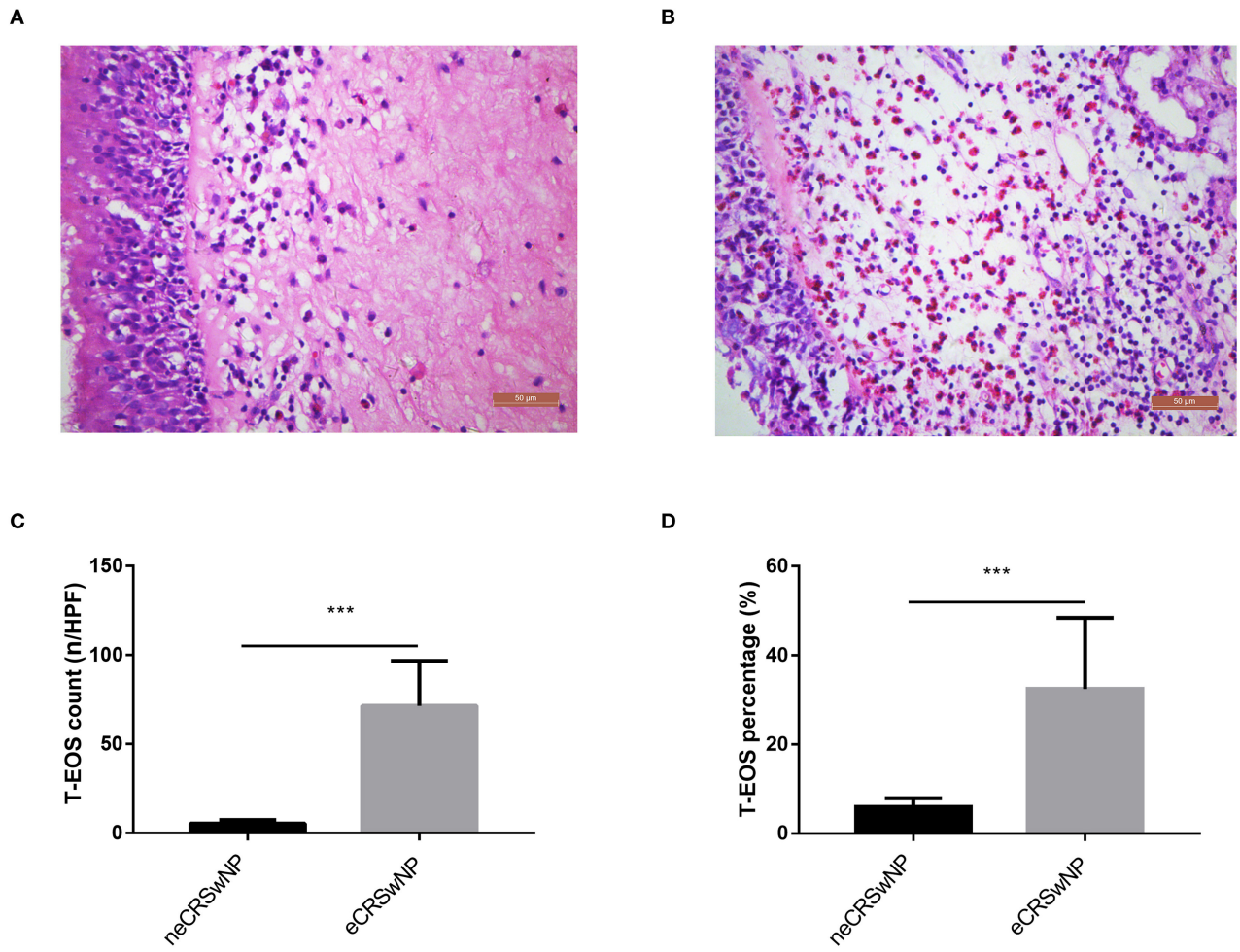

FIGURE 1 | Representative H&E staining of neCRSwNP and eCRSwNP. **(A)** neCRSwNP. **(B)** neCRSwNP. **(C,D)** Comparison of T-EOS percentage and T-EOS count per HPF between neCRSwNP and eCRSwNP. Mann–Whitney U test was utilized. eCRSwNP, eosinophilic chronic rhinosinusitis with nasal polyps; neCRSwNP, non-eosinophilic chronic rhinosinusitis with nasal polyps; T-EOS, tissue eosinophil; H&E, hematoxylin and eosin; HPF, high-power field. ***$P < 0.001$.

The ROC curves of these distinctive metabolites are depicted in **Supplementary Figure 1**, and analysis results are shown in **Supplementary Table 1**. Citrulline, linoleic acid, adenosine, and 4-guanidinobutyric acid exhibited good accuracy for distinguishing eCRSwNP (AUC > 0.7), and they were included in Spearman's correlation analysis to explore their association

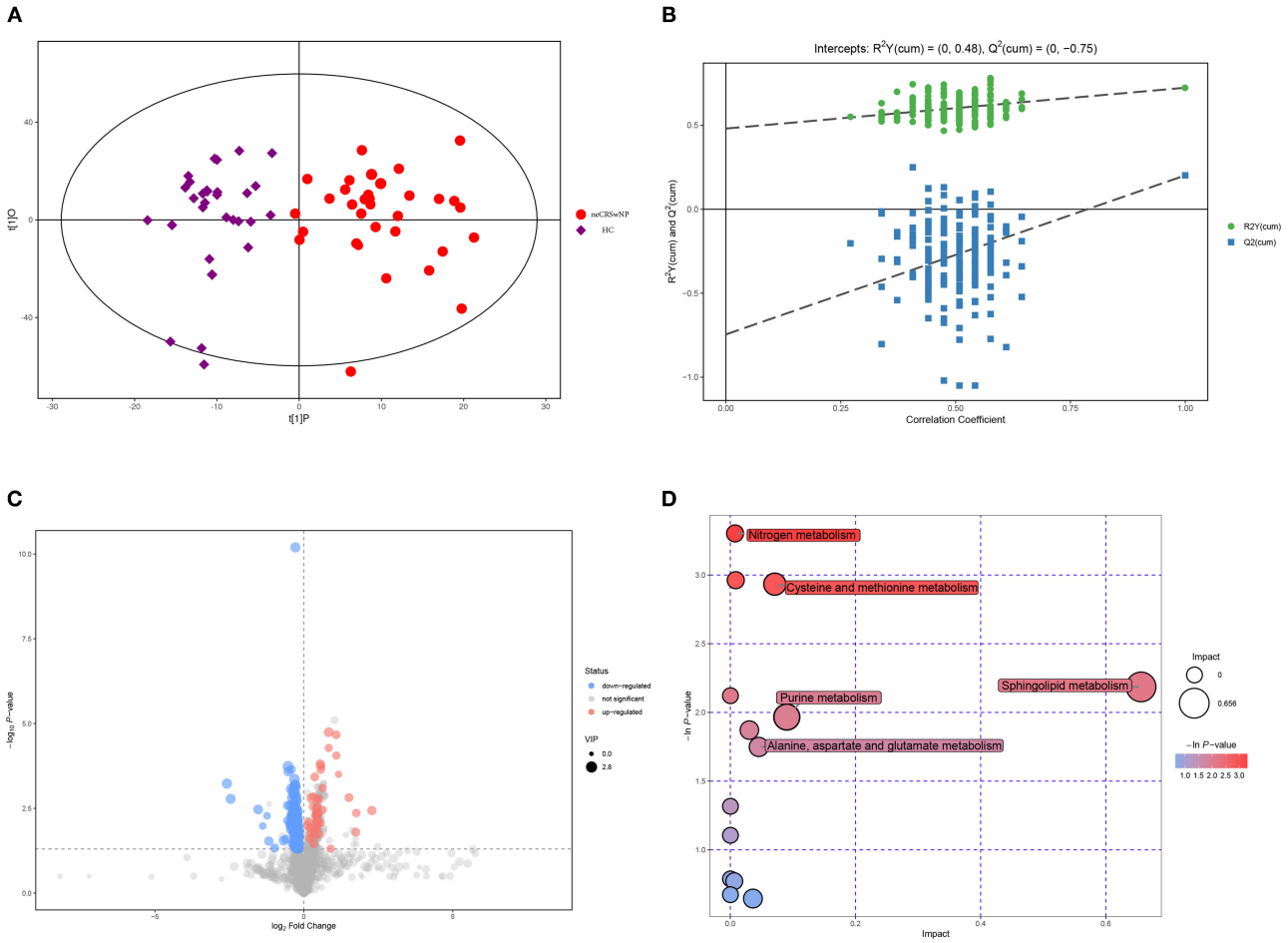

FIGURE 2 | Metabolomic analysis of serum samples of neCRSwNP and HC. **(A,B)** OPLS-DA model and permutation test of the OPLS-DA model. **(C)** Volcano plot. **(D)** Metabolic pathway bubble chart. neCRSwNP, non-eosinophilic chronic rhinosinusitis with nasal polyps; HC, healthy control; OPLS-DA, orthogonal partial least square-discriminate analysis.

with the severity of eosinophils infiltration in the nasal polys. The serum levels of citrulline and adenosine were positively correlated with T-EOS percentage and T-EOS count, while linoleic acid levels were negatively correlated with T-EOS percentage and T-EOS count (**Supplementary Figure 2**).

DISCUSSION

CRSwNP is a complex disease with persistent inflammation in the nasal and sinonasal mucosa, and its physiopathologic mechanisms are poorly clarified (Yamada et al., 2019). Considering the heterogeneity, CRSwNP is divided into eCRSwNP and neCRSwNP, and these two phenotypes have distinctive clinical and pathologic features, drug sensitivity, prognosis, and recurrence rate (Sivrice et al., 2020). Thus, discriminating eCRSwNP from neCRSwNP through a simple and reliable method before surgery is important to promote the precision medicine and improve the management strategies. Up to now, tissue pathological evaluation with H&E staining is the golden standard to diagnose eCRSwNP, which is invasive and relatively subjective and inapplicable to patients who prefer non-surgical treatment (Brescia et al., 2020). Therefore, it is urgently needed to develop an easy, minimally invasive, objective, and feasible method or biomarker to identify subtypes of CRSwNP before treatment. Our study is the first one to describe an innovative application of metabolomics analysis in exploring metabolic signatures to distinguish CRSwNP phenotypes. Our analysis results showed that eCRSwNP exhibited discriminative serum metabolites and metabolic pathway in comparison with neCRSwNP and HC. These results suggested that serum metabolomics was useful for developing objective biomarkers for distinguishing eCRSwNP, and the metabolites and metabolic pathway highlighted in the present study will help us to improve the understanding of underlying pathogenesis of eCRSwNP and explore new therapeutic targets.

We firstly reported that the arginine and proline metabolism pathway was disturbed in eCRSwNP patients. Previous studies demonstrated that arginine metabolism was pivotal in the nitric oxide (NO) synthesis and associated with cellular metabolism, inflammation, and immune response (Xu et al.,

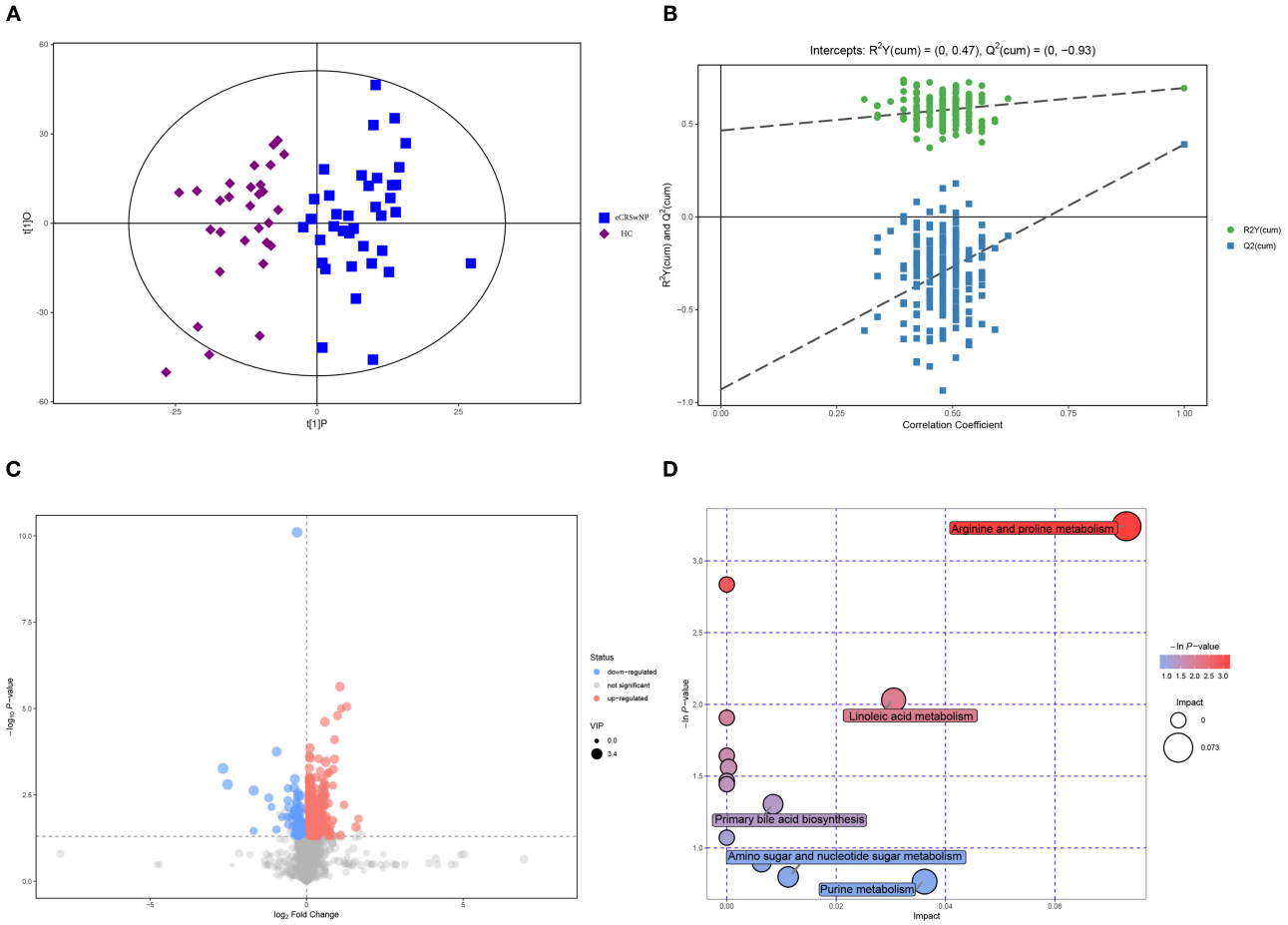

FIGURE 3 | Metabolomic analysis of serum samples of eCRSwNP and HC. **(A,B)** OPLS-DA model and permutation test of the OPLS-DA model. **(C)** Volcano plot. **(D)** Metabolic pathway bubble chart. eCRSwNP, eosinophilic chronic rhinosinusitis with nasal polyps; HC, healthy control; OPLS-DA, orthogonal partial least square-discriminate analysis.

2016). In a recent study, Xu et al. (2017) found that arginine metabolism was regulated in the asthma patients, and the elevated level of arginine promoted the production of fraction of exhaled nitric oxide and then aggravated asthma symptoms. In another study, the researchers reported that the concentrations of ornithine, citrulline, creatine, creatinine, and sarcosine were increased in the serum of asthma patients, and they suggested that arginine metabolism was the most crucial in the development of asthma (Quan-Jun et al., 2017). Liang et al. (2019) observed that arginine metabolism was significantly changed in the serum of commuters who exposure to automobile exhaust, and the arginine metabolism dysfunction increased oxidative stress and inflammation response and aggravated air pollution toxicity. In the present study, we also found that the serum concentrations of citrulline and 4-guanidinobutyric acid were significantly elevated in the eCRSwNP group, and the AUCs for discriminating eCRSwNP were 0.791 and 0.809, respectively, and the serum levels of citrulline were positively correlated with T-EOS percentage and T-EOS count. Citrulline and 4-guanidinobutyric acid were the downstream products of arginine metabolism, and they were proved to participate in regulating T cell proliferation and differentiation and promote inflammatory response in several diseases including asthma and allergic rhinitis (Xu et al., 2017). In addition, citrulline is a key molecule in the citrulline–arginine–NO cycle, and it has been demonstrated to maintain the high NO production and promote the cellular metabolism and the inflammation response (King et al., 2004; Xu et al., 2016). Thus, we suggested that arginine metabolism associated with the development of eCRSwNP, and citrulline could distinguish eCRSwNP and associate with the severity of eosinophils infiltration. Further studies are needed to discover the underlying mechanism.

We firstly found that linoleic acid metabolism was disturbed strongly in the serum of eCRSwNP, and the serum levels of linoleic acid were decreased in eCRSwNP patients compared to neCRSwNP patients and HC and were negatively correlated with T-EOS percentage and T-EOS count, which meant linoleic acid might be a promising biomarker for discriminating eCRSwNP and a novel therapeutic target. Recently, increasing evidence

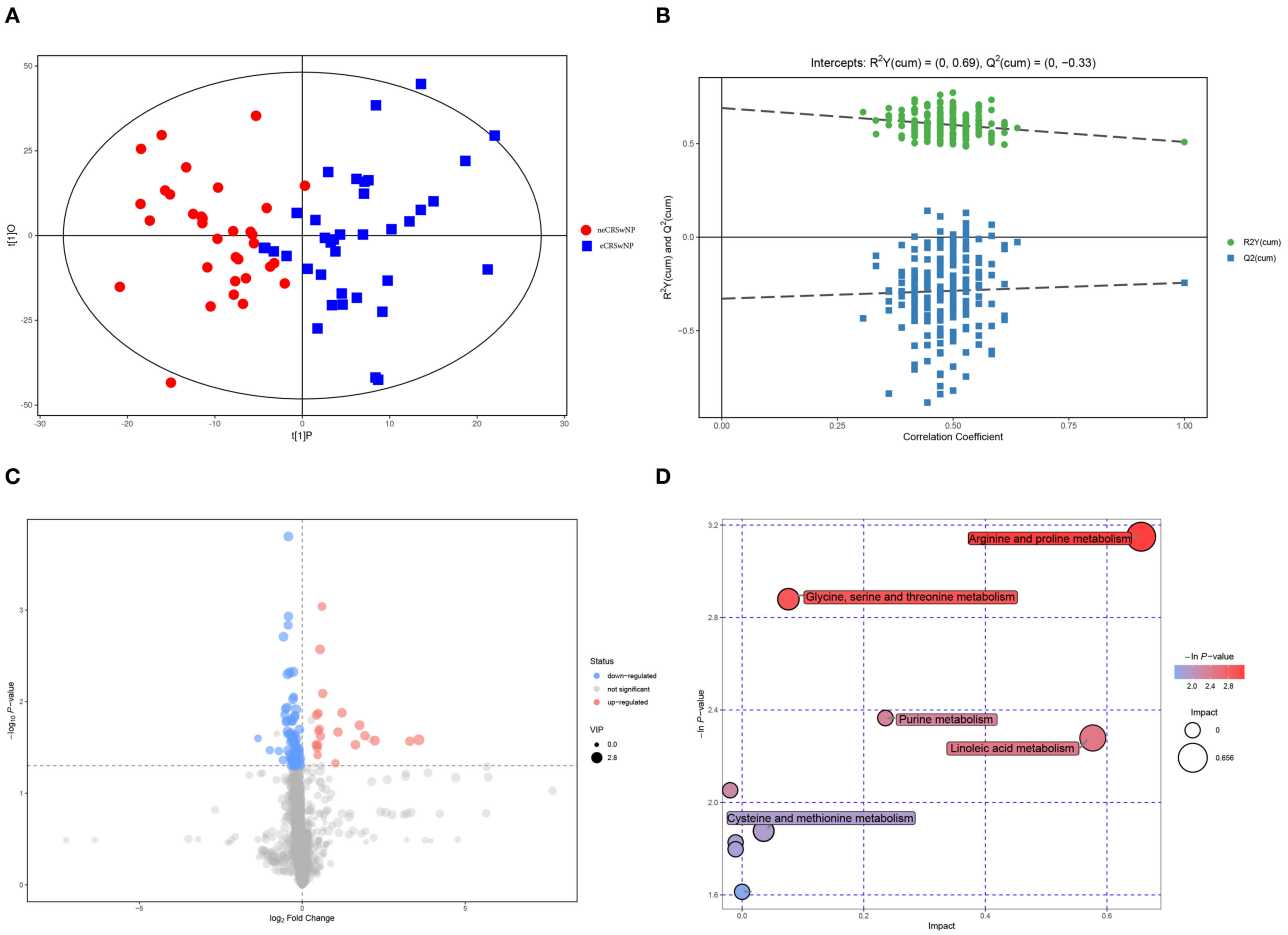

FIGURE 4 | Metabolomic analysis of serum samples of neCRSwNP and neCRSwNP. **(A,B)** OPLS-DA model and permutation test of the OPLS-DA model. **(C)** Volcano plot. **(D)** Metabolic pathway bubble chart. neCRSwNP, non-eosinophilic chronic rhinosinusitis with nasal polyps; eCRSwNP, eosinophilic chronic rhinosinusitis with nasal polyps; OPLS-DA, orthogonal partial least square-discriminate analysis.

demonstrated that fatty acid metabolism played emerging roles in regulating immune responses in allergic and inflammatory diseases (Arita, 2016; Ishihara et al., 2019). A previous study reported that oleic acid could reduce the production of inflammatory cells and eosinophils in bronchial alveolar lavage fluid, and IgE in serum of mouse models, then suppressed the occurrence and development of asthma (Lee et al., 2019). Previous publications showed that polyunsaturated fatty acids could affect the functions of T cells via inhibiting its proliferation and activation, and also could suppress the activation and secretion of mast cells (Yu and Björkstén, 1998; Wang and Kulka, 2015; Arita, 2016; Radzikowska et al., 2019). Linoleic acid, a common polyunsaturated fatty acid, has been proved to be crucial in activating both autophagy and antioxidation in a synergistic feedback loop and greatly aids in the prevention and treatment of multiple inflammatory disease (Wang and Kulka, 2015; Lee et al., 2019). Therefore, we have reasons to believe that linoleic acid may play a pivotal role in the eCRSwNP and can serve as an objective indication for distinguishing CRSwNP phenotypes.

Another interesting finding was that the serum concentrations of adenosine were most indicative of distinguishing eCRSwNP and reflecting the severity of eosinophil infiltration in the nasal polys tissue. Adenosine, an endogenous purine nucleoside, can be accumulated during different physiologic and pharmacologic processes, such as hypoxia, trauma, and inflammation, and several studies suggested that it played an important role in modulating mast cell, monocytes, and T cell functions (Gomez et al., 2013; Yuryeva et al., 2015). A recent study reported that adenosine was produced in high concentrations in the serum of chronic obstructive pulmonary diseases, and the serum levels of adenosine significantly correlated with disease severity (Singh Patidar et al., 2018). Mao et al. (2020) found that the levels of adenosine were significantly increased in the plasma of chronic spontaneous urticaria and associated with disease activity, and they also observed that plasma adenosine was a promising biomarker for predicting treatment outcomes. Vass et al. (2006) demonstrated that the adenosine concentrations were elevated in the exhaled breath of allergic rhinitis and positively correlated with NO concentrations. Collectively, these

TABLE 2 | Top 10 metabolites discriminating eCRSwNP from neCRSwNP.

Metabolites	VIP	*P*	FC	AUC	Pathways
Citrulline	2.73	< 0.001	4.11	0.791	Arginine and proline metabolism
Glycine	2.47	0.008	0.36	0.544	Glycine, serine and threonine metabolism
Linoleic acid	2.13	< 0.001	0.41	0.823	Linoleic acid metabolism
Adenosine	2.08	< 0.001	2.48	0.902	Purine metabolism
Glycocholic acid	1.94	0.005	2.06	0.627	Primary bile acid biosynthesis
L-Serine	1.90	0.024	0.52	0.615	Glycine, serine and threonine metabolism
Triethanolamine	1.84	< 0.001	1.93	0.524	Glycerophospholipid metabolism
4-Guanidinobutyric acid	1.80	0.002	2.37	0.809	Arginine and proline metabolism
Methylmalonic acid	1.79	0.036	1.77	0.672	Pyrimidine metabolism
L-methionine	1.72	0.010	0.62	0.690	Cysteine and methionine metabolism

eCRSwNP, eosinophilic chronic rhinosinusitis with nasal polyps; neCRSwNP, non-eosinophilic chronic rhinosinusitis with nasal polyps; VIP, variable importance for project; FC, fold change; AUC, area under the curve.

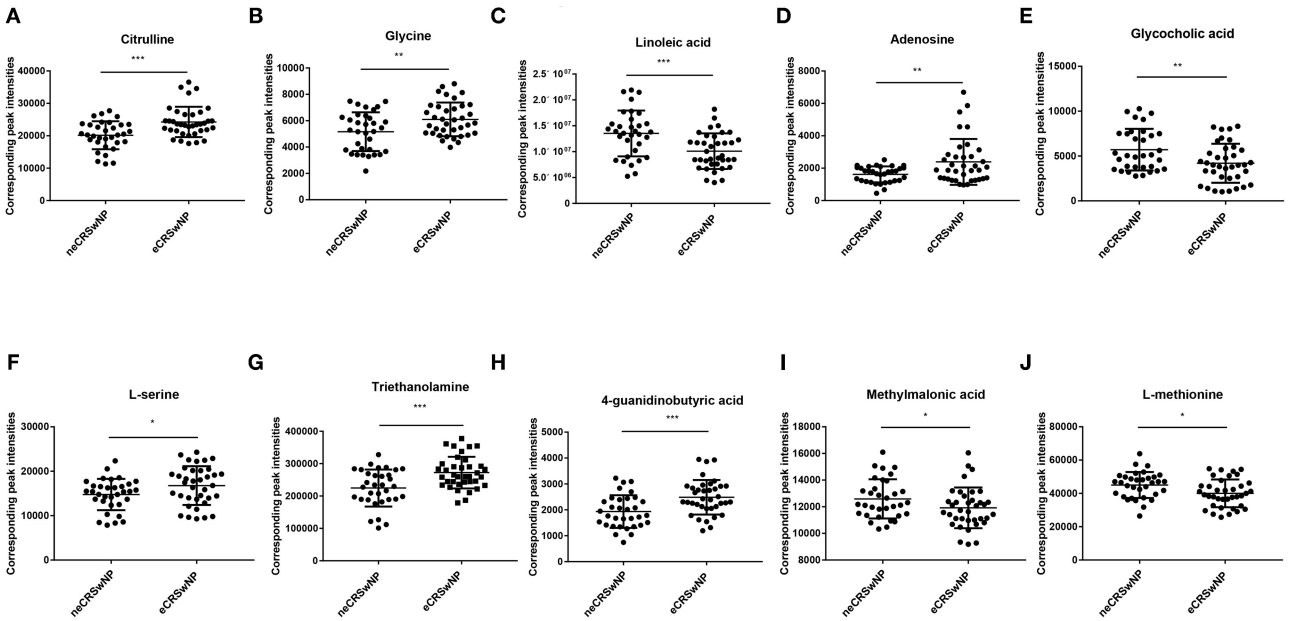

FIGURE 5 | Top 10 most discriminant metabolites in their relative levels in neCRSwNP group and neCRSwNP group. neCRSwNP, non-eosinophilic chronic rhinosinusitis with nasal polyps; eCRSwNP, eosinophilic chronic rhinosinusitis with nasal polyps. Mann–Whitney *U*-test was used for the statistical analysis. $^{*}P < 0.05$, $^{**}P < 0.001$, $^{***}P < 0.001$.

studies provide a reasonable explanation for our observation of elevated adenosine in eCRSwNP and its value in distinguishing CRSwNP phenotypes.

Of note, abnormal glycine, serine, and threonine metabolism was also found in the present study, and the serum levels of glycine and L-serine were decreased in the eCRSwNP patients. Accordingly, serine is one of the crucial amino acids in the synthesis of human proteins, and L-serine, another isoform of serine, was proven to be pivotal in suppressing the production of reactive oxygen species and reducing oxidative stress in several inflammatory diseases (Rodriguez et al., 2019). A previous study reported that L-serine provided components for nerve function and exerted anti-inflammatory properties, it could relieve chronic pain in low-back and knee pain patients (Sasahara et al., 2020). Glycine has been previously demonstrated to be critical in controlling the levels of oxygen species, and it exhibits anti-inflammatory and immunomodulatory effects in several disorders (Yang et al., 2020). Alonso et al. (2016) utilized nuclear magnetic resonance to analyze the urine metabolites of several inflammatory diseases and found that the serum levels of glycine were significantly decreased. Therefore, we ultimately believed that glycine and L-serine were associated with eCRSwNP, and they might serve as novel metabolic biomarkers for discriminating CRSwNP phenotypes.

Our study has several limitations. First, the sample size is relatively small, and a validation cohort study is needed to confirm the conclusions. Second, all included patients are from single centers with the same ethnicity and region, which might

limit their generalization. Third, there is a wide variation in diagnostic criteria of eCRSwNP among previous reports, and no clear criteria currently exists regarding the cutoff value, this may limit the applicability of our findings. Lastly, because CRSwNP is a nasal and sinus disease mainly characterized by local inflammation and metabolic changes, the degrees of systemic metabolic disturbance and metabolic pathway dysfunction are relatively low, and these may partly influence the accuracy and predictability of OPLS-DA and permutation models. Future studies with larger sample sizes and unified diagnostic criteria are needed to validate and strengthen our present conclusion.

In conclusion, we have demonstrated that serum metabolomics could be utilized to distinguish CRSwNP phenotypes and establish metabolic signatures which might reflect the severity of eosinophil infiltration. These results suggested that metabolomic profiles contributed to understanding the pathophysiological mechanisms of eCRSwNP.

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 the ethical committee of Xiangya Hospital of Central South University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ShaoX and HZ wrote the manuscript. HZ and YL collected the sample. KG and JZ performed the data analysis. RF and ShuX provided statistical support. ZX, FW, and WJ designed the research study. All authors reviewed the manuscript and approved the final version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 81770985, 81873695, and 81800917) and Natural Science Foundation of Hunan Province (Nos. 2020JJ4910, 2018JJ2662, and 2018JJ2632).

ACKNOWLEDGMENTS

We are grateful for the help of Shanghai BIOTREE Biological Technology Co., Ltd. (Shanghai, China) in the metabolite detection and data analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.593976/full#supplementary-material

Supplementary Figure 1 | The ROC analysis results of top 10 most discriminant metabolites for distinguishing CRSwNP phenotypes. ROC, receiver operating characteristics; CRSwNP, chronic rhinosinusitis with nasal polyps; AUC, area under the curve.

Supplementary Figure 2 | Relationship between eosinophilic infiltration and metabolites with good predictabilities. **(A–D)** Correlation between T-EOS percentage and citrulline, linoleic acid, adenosine and 4-guanidinobutyric acid. **(E–H)** Correlation between T-EOS percentage and citrulline, linoleic acid, adenosine and 4-guanidinobutyric acid. T-EOS, tissue eosinophil. Spearman correlation analysis was used.

Supplementary Table 1 | ROC analysis results of top ten metabolites for discriminating eCRSwNP from neCRSwNP.

REFERENCES

Adamko, D. J., Nair, P., Mayers, I., Tsuyuki, R. T., Regush, S., and Rowe, B. H. (2015). Metabolomic profiling of asthma and chronic obstructive pulmonary disease: a pilot study differentiating diseases. *J. Allergy Clin. Immunol.* 136, 571-80.e3. doi: 10.1016/j.jaci.2015.05.022

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.* 14:133. doi: 10.1186/s12916-016-0681-8

Arita, M. (2016). Eosinophil polyunsaturated fatty acid metabolism and its potential control of inflammation and allergy. *Allergol. Int.* 65, S2–S5. doi: 10.1016/j.alit.2016.05.010

Bayar Muluk, N., Cingi, C., Scadding, G. K., and Scadding, G. (2019). Chronic rhinosinusitis-could phenotyping or endotyping aid therapy? *Am. J. Rhinol. Allergy.* 33, 83–93. doi: 10.1177/1945892418807590

Brescia, G., Alessandrini, L., Giacomelli, L., Parrino, D., Zanotti, C., Tealdo, G., et al. (2020). A classification of chronic rhinosinusitis with nasal polyps based on structured histopathology. *Histopathology* 76, 296–307. doi: 10.1111/his.13969

Chitsuthipakorn, W., Seresirikachorn, K., Sommer, D. D., McHugh, T., and Snidvongs, K. (2018). Endotypes of chronic rhinosinusitis across ancestry and geographic regions. *Curr. Allergy Asthma Rep.* 18:46. doi: 10.1007/s11882-018-0800-z

Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., et al. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 6, 1060–1083. doi: 10.1038/nprot.2011.335

Fokkens, W. J., Lund, V. J., Mullol, J., Bachert, C., Alobid, I., Baroody, F., et al. (2012). European position paper on rhinosinusitis and nasal polyps 2012. *Rhinology* 23:3. doi: 10.4193/Rhino50E2

Fujieda, S., Imoto, Y., Kato, Y., Ninomiya, T., Tokunaga, T., Tsutsumiuchi, T., et al. (2019). Eosinophilic chronic rhinosinusitis. *Allergol. Int.* 68, 403–412. doi: 10.1016/j.alit.2019.07.002

Gomez, G., Nardone, V., Lotfi-Emran, S., Zhao, W., and Schwartz, L. B. (2013). Intracellular adenosine inhibits IgE-dependent degranulation of human skin mast cells. *J. Clin. Immunol.* 33, 1349–1359. doi: 10.1007/s10875-013-9950-x

Grayson, J. W., Cavada, M., and Harvey, R. J. (2019). Clinically relevant phenotypes in chronic rhinosinusitis. *J. Otolaryngol. Head Neck Surg.* 48:23. doi: 10.1186/s40463-019-0355-6

Ho, J., Hamizan, A. W., Alvarado, R., Rimmer, J., Sewell, W. A., and Harvey, R. J. (2018). Systemic predictors of eosinophilic chronic

rhinosinusitis. *Am. J. Rhinol. Allergy*. 32, 252–257. doi: 10.1177/1945892418779451

Hoy, S. M. (2020). Dupilumab: a review in chronic rhinosinusitis with nasal polyps. *Drugs* 80, 711–717. doi: 10.1007/s40265-020-01298-9

Hu, Y., Cao, P. P., Liang, G. T., Cui, Y. H., and Liu, Z. (2012). Diagnostic significance of blood eosinophil count in eosinophilic chronic rhinosinusitis with nasal polyps in Chinese adults. *Laryngoscope*. 122, 498–503. doi: 10.1002/lary.22507

Ishihara, T., Yoshida, M., and Arita, M. (2019). Omega-3 fatty acid-derived mediators that control inflammation and tissue homeostasis. *Int. Immunol*. 31, 559–567. doi: 10.1093/intimm/dxz001

Kelly, R. S., Dahlin, A., McGeachie, M. J., Qiu, W., Sordillo, J., Wan, E. S., et al. (2017). Asthma metabolomics and the potential for integrative omics in research and the clinic. *Chest*. 151, 262–277. doi: 10.1016/j.chest.2016.10.008

King, N. E., Rothenberg, M. E., and Zimmermann, N. (2004). Arginine in asthma and lung inflammation. *J. Nutr*. 134, 2830S−2836S. doi: 10.1093/jn/134.10.2830S

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R., and Neumann, S. (2012). CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal. Chem*. 84, 283–289. doi: 10.1021/ac202450g

Lee, S.-Y., Bae, C.-S., Seo, N.-S., Na, C.-S., Yoo, H. Y., Oh, D.-S., et al. (2019). Camellia japonica oil suppressed asthma occurrence via GATA-3 and IL-4 pathway and its effective and major component is oleic acid. *Phytomedicine* 57, 84–94. doi: 10.1016/j.phymed.2018.12.004

Li, X., Li, C., Zhu, G., Yuan, W., and Xiao, Z. A. (2019). TGF-β1 induces epithelial-mesenchymal transition of chronic sinusitis with nasal polyps through MicroRNA-21. *Int. Arch. Allergy Immunol*. 179, 304–319. doi: 10.1159/000497829

Liang, D., Ladva, C. N., Golan, R., Yu, T., Walker, D. I., Sarnat, S. E., et al. (2019). Perturbations of the arginine metabolome following exposures to traffic-related air pollution in a panel of commuters with and without asthma. *Environ. Int*. 127, 503–513. doi: 10.1016/j.envint.2019.04.003

Liu, S., Liang, Y. Z., and Liu, H. T. (2016). Chemometrics applied to quality control and metabolomics for traditional Chinese medicines. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci*. 1015–1016, 82–91. doi: 10.1016/j.jchromb.2016.02.011

Lund, V. J., and Mackay, I. S. (1993). Staging in rhinosinusitus. *Rhinology*. 31, 183–184.

Ma, G. C., Wang, T. S., Wang, J., Ma, Z. J., and Pu, S. B. (2020). Serum metabolomics study of patients with allergic rhinitis. *Biomed. Chromatogr*. 34:e4739. doi: 10.1002/bmc.4739

Mao, M., Liu, H., Yan, S., Yuan, Y., Liu, R., Wu, Y., et al. (2020). Plasma adenosine is linked to disease activity and response to treatment in patients with chronic spontaneous urticaria. *Allergy*. doi: 10.1111/all.14502. [Epub ahead of print].

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

Ning, P., Zheng, Y., Luo, Q., Liu, X., Kang, Y., Zhang, Y., et al. (2018). Metabolic profiles in community-acquired pneumonia: developing assessment tools for disease severity. *Crit. Care*. 22:130. doi: 10.1186/s13054-018-2049-2

Qing, X., Zhang, Y., Peng, Y., He, G., Liu, A., and Liu, H. (2019). Mir-142-3p regulates inflammatory response by contributing to increased TNF-α in chronic rhinosinusitis with nasal polyposis. *Ear Nose Throat J*. 100, NP50–NP56. doi: 10.1177/0145561319847972

Quan-Jun, Y., Jian-Ping, Z., Jian-Hua, Z., Yong-Long, H., Bo, X., Jing-Xian, Z., et al. (2017). Distinct metabolic profile of inhaled budesonide and salbutamol in asthmatic children during acute exacerbation. *Basic Clin. Pharmacol. Toxicol*. 120, 303–311. doi: 10.1111/bcpt.12686

Radzikowska, U., Rinaldi, A. O., Çelebi Sözener, Z., Karaguzel, D., Wojcik, M., Cypryk, K., et al. (2019). The influence of dietary fatty acids on immune responses. *Nutrients*. 11:2990. doi: 10.3390/nu11122990

Reisdorph, N., and Wechsler, M. E. (2013). Utilizing metabolomics to distinguish asthma phenotypes: strategies and clinical implications. *Allergy*. 68, 959–962. doi: 10.1111/all.12238

Rodriguez, A. E., Ducker, G. S., Billingham, L. K., Martinez, C. A., Mainolfi, N., Suri, V., et al. (2019). Serine metabolism supports macrophage IL-1β production. *Cell Metab*. 29, 1003–1011.e4. doi: 10.1016/j.cmet.2019.01.014

Sasahara, I., Yamamoto, A., Takeshita, M., Suga, Y., Suzuki, K., Nishikata, N., et al. (2020). l-serine and EPA relieve chronic low-back and knee pain in adults: a randomized, double-blind, placebo-controlled trial. *J Nutr*. 150, 2278–2286. doi: 10.1093/jn/nxaa156

Singh Patidar, B., Meena, A., Kumar, M., Menon, B., Rohil, V., and Kumar Bansal, S. (2018). Adenosine metabolism in COPD: a study on adenosine levels, 5'-nucleotidase, adenosine deaminase and its isoenzymes activity in serum, lymphocytes and erythrocytes. *COPD* 15, 559–571. doi: 10.1080/15412555.2018.1537365

Sivrice, M. E., Okur, E., Yasan, H., Tüz, M., Kumbul, Y., and Akin, V. (2020). Can the systemic immune inflammation index preoperatively predict nasal polyp subtypes? *Eur Arch Otorhinolaryngol*. 277, 3045–3050. doi: 10.1007/s00405-020-06174-6

Spertini, F. (2020). Metabolomics and allergy: opening pandora's box. *J. Allergy Clin. Immunol*. 145, 782–784. doi: 10.1016/j.jaci.2020.01.012

Turi, K. N., Romick-Rosendale, L., Ryckman, K. K., and Hartert, T. V. (2018). A review of metabolomics approaches and their application in identifying causal pathways of childhood asthma. *J. Allergy Clin. Immunol*. 141, 1191–1201. doi: 10.1016/j.jaci.2017.04.021

Vass, G., Huszár, E., Augusztinovicz, M., Baktai, G., Barát, E., Herjavecz, I., et al. (2006). The effect of allergic rhinitis on adenosine concentration in exhaled breath condensate. *Clin. Exp. Allergy* 36, 742–747. doi: 10.1111/j.1365-2222.2006.02496.x

Wang, W., Zhao, L., He, Z., Wu, N., Li, Q., Qiu, X., et al. (2018). Metabolomics-based evidence of the hypoglycemic effect of Ge-Gen-Jiao-Tai-Wan in type 2 diabetic rats via UHPLC-QTOF/MS analysis. *J. Ethnopharmacol*. 219, 299–318. doi: 10.1016/j.jep.2018.03.026

Wang, X., and Kulka, M. (2015). n-3 polyunsaturated fatty acids and mast cell activation. *J. Leukoc. Biol*. 97, 859–871. doi: 10.1189/jlb.2RU0814-388R

Xu, W., Comhair, S. A. A., Janocha, A. J., Lara, A., Mavrakis, L. A., Bennett, C. D., et al. (2017). Arginine metabolic endotypes related to asthma severity. *PLoS ONE* 12:e0183066. doi: 10.1371/journal.pone.0183066

Xu, W., Ghosh, S., Comhair, S. A., Asosingh, K., Janocha, A. J., Mavrakis, D. A., et al. (2016). Increased mitochondrial arginine metabolism supports bioenergetics in asthma. *J. Clin. Invest*. 126, 2465–2481. doi: 10.1172/JCI82925

Yamada, T., Miyabe, Y., Ueki, S., Fujieda, S., Tokunaga, T., Sakashita, M., et al. (2019). Eotaxin-3 as a plasma biomarker for mucosal eosinophil infiltration in chronic rhinosinusitis. *Front. Immunol*. 10:74. doi: 10.3389/fimmu.2019.00074

Yang, Y., Wu, Z., Li, S., Yang, M., Xiao, X., Lian, C., et al. (2020). Targeted blood metabolomic study on retinopathy of prematurity. *Invest. Ophthalmol. Visual Sci*. 61:12. doi: 10.1167/iovs.61.2.12

Yao, Y., Xie, S., and Wang, F. (2019). Identification of key genes and pathways in chronic rhinosinusitis with nasal polyps using bioinformatics analysis. *Am. J. Otolaryngol*. 40, 191–196. doi: 10.1016/j.amjoto.2018.12.002

Yao, Y., Xie, S., Yang, C., Zhang, J., Wu, X., and Sun, H. (2017). Biomarkers in the evaluation and management of chronic rhinosinusitis with nasal polyposis. *Eur. Arch. Otorhinolaryngol*. 274, 3559–3566. doi: 10.1007/s00405-017-4547-2

Yao, Y., Yang, C., Yi, X., Xie, S., and Sun, H. (2020). Comparative analysis of inflammatory signature profiles in eosinophilic and noneosinophilic chronic rhinosinusitis with nasal polyposis. *Biosci. Rep*. 40:BSR20193101. doi: 10.1042/BSR20193101

Yu, G., and Björkstén, B. (1998). Polyunsaturated fatty acids in school children in relation to allergy and serum IgE levels. *Pediatr. Allergy Immunol*. 9, 133–138. doi: 10.1111/j.1399-3038.1998.tb00359.x

Yuryeva, K., Saltykova, I., Ogorodova, L., Kirillova, N., Kulikov, E., Korotkaya, E., et al. (2015). Expression of adenosine receptors in monocytes from patients with bronchial asthma. *Biochem. Biophys. Res. Commun*. 464, 1314–1320. doi: 10.1016/j.bbrc.2015.07.141

Zhao, H., Cheng, N., Wang, Q., Zhou, W., Liu, C., Liu, X., et al. (2019). Effects of honey-extracted polyphenols on serum antioxidant capacity and metabolic phenotype in rats. *Food Funct*. 10, 2347–2358. doi: 10.1039/C8FO02138D

Zhong, B., Yuan, T., Du, J., Tan, K., Yang, Q., Liu, F., et al. (2020). The role of preoperative blood eosinophil counts in distinguishing chronic rhinosinusitis with nasal polyps phenotypes. *Int. Forum Allergy Rhinol*. doi: 10.1002/alr.22636. [Epub ahead of print].

Zhu, Z., Wang, W., Zhang, X., Wang, X., Zha, Y., Chen, Y., et al. (2020). Nasal fluid cytology and cytokine profiles of eosinophilic and non-eosinophilic chronic rhinosinusitis with nasal polyps. *Rhinology*. 58, 314–322. doi: 10.4193/Rhin19.275

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 14 April 2021
doi: 10.3389/fmolb.2021.632950

Comprehensive Metabolomics Identified the Prominent Role of Glycerophospholipid Metabolism in Coronary Artery Disease Progression

*Hui Chen[1,2,3†], Zixian Wang[1,2,3,4†], Min Qin[1,2,3], Bin Zhang[3,5], Lu Lin[2], Qilin Ma[6], Chen Liu[7], Xiaoping Chen[8], Hanping Li[3], Weihua Lai[2] and Shilong Zhong[1,2,3,4]**

[1]Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology, Guangzhou, China, [2]Department of Pharmacy, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [3]Guangdong Provincial Key Laboratory of Coronary Heart Disease Prevention, Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [4]School of Biology and Biological Engineering, South China University of Technology, Guangzhou, China, [5]Department of Cardiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [6]Department of Cardiology, Xiangya Hospital, Central South University, Changsha, China, [7]Department of Cardiology, the First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China, [8]Department of Clinical Pharmacology, Xiangya Hospital, Central South University, Changsha, China

OPEN ACCESS

Edited by:
Danuta Dudzik, Medical University of Gdańsk, Poland

Reviewed by:
Donghai Lin, Xiamen University, China
Fumio Matsuda, Osaka University, Japan

***Correspondence:**
Shilong Zhong
zhongsl@hotmail.com

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

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

Received: *24 November 2020*
Accepted: *09 February 2021*
Published: *14 April 2021*

Citation:
Chen H, Wang Z, Qin M, Zhang B, Lin L, Ma Q, Liu C, Chen X, Li H, Lai W and Zhong S (2021) Comprehensive Metabolomics Identified the Prominent Role of Glycerophospholipid Metabolism in Coronary Artery Disease Progression. Front. Mol. Biosci. 8:632950. doi: 10.3389/fmolb.2021.632950

Background: Coronary stenosis severity determines ischemic symptoms and adverse outcomes. The metabolomic analysis of human fluids can provide an insight into the pathogenesis of complex disease. Thus, this study aims to investigate the metabolomic and lipidomic biomarkers of coronary artery disease (CAD) severity and to develop diagnostic models for distinguishing individuals at an increased risk of atherosclerotic burden and plaque instability.

Methods: Widely targeted metabolomic and lipidomic analyses of plasma in 1,435 CAD patients from three independent centers were performed. These patients were classified as stable coronary artery disease (SCAD), unstable angina (UA), and myocardial infarction (MI). Associations between CAD stages and metabolic conditions were assessed by multivariable-adjusted logistic regression. Furthermore, the least absolute shrinkage and selection operator logistic-based classifiers were used to identify biomarkers and to develop prediagnostic models for discriminating the diverse CAD stages.

Results: On the basis of weighted correlation network analysis, 10 co-clustering metabolite modules significantly ($p < 0.05$) changed at different CAD stages and showed apparent correlation with CAD severity indicators. Moreover, cross-comparisons within CAD patients characterized that a total of 72 and 88 metabolites/lipid species significantly associated with UA (vs. SCAD) and MI (vs. UA), respectively. The disturbed pathways included glycerophospholipid metabolism, and cysteine and methionine metabolism. Furthermore, models incorporating metabolic and lipidomic profiles with traditional risk factors were constructed. The combined model that incorporated 11 metabolites/lipid species and four traditional risk factors represented better discrimination of UA and MI (C-statistic = 0.823, 95% CI, 0.783–0.863) compared with the model involving risk factors alone (C-statistic = 0.758, 95% CI, 0.712–0.810). The

combined model was successfully used in discriminating UA and MI patients ($p < 0.001$) in a three-center validation cohort.

Conclusion: Differences in metabolic profiles of diverse CAD subtypes provided a new approach for the risk stratification of unstable plaque and the pathogenesis decipherment of CAD progression.

Keywords: coronary artery disease, metabolome, lipidome, severity, glycerophospholipid metabolism, diagnostic marker

INTRODUCTION

Coronary artery disease (CAD) refers to underlying coronary artery atherosclerotic lesions that cause vascular lumen stenosis or occlusion and insufficient blood supply and result in myocardial ischemia, hypoxia, or necrosis (Malakar et al., 2019). Despite the advances in medical treatment, percutaneous coronary intervention, and surgical therapies, atherosclerotic CAD persists as a major clinical problem leading to a significant proportion of mortality of aging populations (Zhou et al., 2019; Virani et al., 2020). CAD can be stratified into stable coronary artery disease (SCAD), unstable angina (UA), and myocardial infarction (MI) according to the clinical symptoms, the extent of arterial blockage, and the condition of myocardial damage (Shao et al., 2020). Atherosclerotic plaque accumulation and development become chronic, complicated, and dynamic over time. The detailed mechanisms of plaque formation and development are poorly known. Thus, novel biomarkers for patients with risks of plaque instability and rupture need to be identified to delay onset and improve treatment.

Emerging metabolomics is a powerful tool to systematically investigate the functional small molecule in biological fluid samples. An abnormal metabolome can reportedly characterize CAD, further providing clues for physiological and pathological explorations (Wishart, 2016; Tzoulaki et al., 2019). Elevated plasma trimethylamine N-oxide levels can predict a future risk of major adverse cardiac events (MACE) and an increased prevalence of cardiovascular disease (CVD) (Dannenberg et al., 2020; Gencer et al., 2020). Short-chain fatty acids and primary and secondary bile acids affect CVD progression (Fan et al., 2016; Tang et al., 2019). Previous studies highlighted the key role of lipid species in the formation and subsequent disruption of atherosclerotic plaques, including ceramides, sphingomyelin, phosphatidylcholines, and cholesterol esters (Wang D. D. et al., 2017; Poss et al., 2020). Altered lipid metabolism correlated with inflammation and oxidative stress, such as the oxidation of phospholipids and cholesterol in LDL and played an important part in the formation of lipid-laden foam cells within the intima to the necrotic lipid core of unstable plaque (Meikle et al., 2011; Lu et al., 2017; Zhong et al., 2019). However, the relationship between plasma metabolic profiling and detailed characterization and quantification of atherosclerosis burden at different CAD stages needs to be systematically elucidated.

The goals of the present study were to comprehensively investigate the plasma metabolomic and lipidomic signatures associated with increased CAD severity and to evaluate the significantly differential metabolites and lipid species for their use in discriminating the subgroups of CAD, thereby providing an enhanced understanding of disease progression. Herein, we performed a widely targeted metabolomic and lipidomic evaluation in plasma of patients with SCAD, UA, and MI and identified specific features of metabolite profiles that are associated with increase in CAD severity and can be used to differentiate these three subgroups. The subsequent pathway analysis revealed that glycerophospholipid metabolism was the most significantly altered metabolic pathway. Disease diagnostic classifiers for discriminating between different CAD subgroups were constructed and validated based on novel metabolic markers and traditional risk factors.

MATERIALS AND METHODS

Study Population

An overview of the workflow is depicted in **Figure 1**. This study was a two-stage study that included a total of 1,435 Chinese subjects with CAD. In the discovery cohort (N = 942), we evaluated the association of plasma metabolome and lipidome with CAD using consecutively enrolled samples with clinical and demographic information obtained from Guangdong Provincial People's Hospital (Cai et al., 2018) in 2010–2014. In the verification cohort (N = 493), we enrolled multi-center patients with CAD from three centers (including Guangdong Provincial People's Hospital, Xiangya Hospital of Central South University, and the First Affiliated Hospital of Sun Yat-sen University) from 2017 to 2018.

All subjects were 18–80 years and met the diagnostic criteria of CAD. They were further stratified into three subgroups (SCAD, UA, and MI) on the basis of a detailed diagnosis performed by cardiologists, their symptoms, ischemic changes in electrocardiogram, laboratory measurements, and coronary angiographic results. The specific diagnostic criteria of CAD subtypes are summarized under **Supplemental Materials: Supplementary Methods**. The exclusion criteria were as follows: 1) severe renal dysfunction, serum creatinine >3.0 mg/dl, renal transplantation, or dialysis; 2) liver dysfunction, alanine aminotransferase >135 U/L, or cirrhosis; 3) during pregnancy or lactation; 4) malignant tumors or hemodialysis; 5) autoimmune disorders; and 6) unavailable information. Demographic information, medication history, and biochemical measurements were collected according to

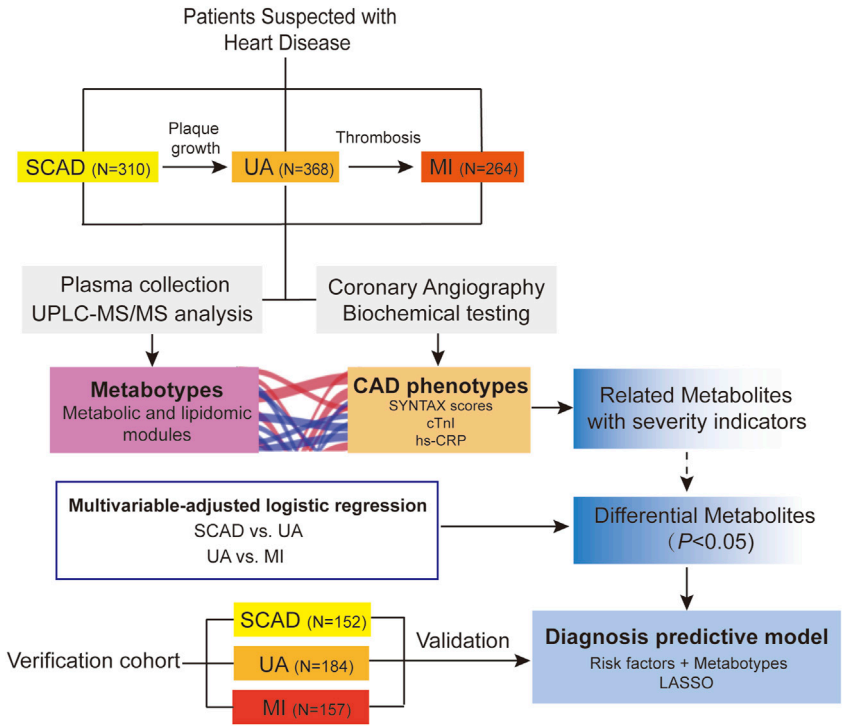

FIGURE 1 | Overview of workflow chart for data generation and analysis. CAD, coronary artery disease; cTnI, cardiac troponin I; hs-CRP, high-sensitivity C-reactive protein; LASSO, least absolute shrinkage and selection operator; MI, myocardial infarction; SCAD, stable coronary disease; UA, unstable angina; SYNTAX scores, Synergy between percutaneous coronary intervention with TAXUS and Cardiac Surgery scores; UPLC-MS/MS, ultra-performance liquid chromatography-tandem mass spectrometry.

standard procedures and obtained from the hospital electronic case system.

Ethics Statement

The study fully complies with the guidance of the Helsinki Declaration. The Medical Ethical Review Committee of Guangdong Provincial People's Hospital granted ethics approval (GDREC2010137 and GDREC2017071H). Written informed consent was obtained from all subjects.

Plasma Sample Collection

Each eligible subject fasted for at least 8 h to minimize the influence of nutrition on metabolite levels. The subjects' venous blood samples were collected in ethylene diamine tetraacetic acid (EDTA) vacutainer tubes in the morning (between 9 AM and 12 PM) after overnight fasting and cooled in a freezer (4°C) immediately. Plasma was separated by centrifugation (2095 *g*, 10 min, 4°C) within 2 h and refrigerated at −80°C until analysis.

Severity Evaluation of Coronary Artery Disease Via Angiographic Analysis

Coronary angiography (CAG) was performed to define the extent and severity of CAD in patients with suspected symptoms whose clinical characteristics and results of noninvasive testing indicated a high likelihood of CAD and who are amenable to, and candidates for coronary revascularization (Fihn et al., 2014). CAG was performed using the standard technique and images of coronary angiograms were obtained from Syngo Dynamics cardiovascular imaging software (Siemens Medical Solutions, United States, Inc, Malvern, Pennsylvania). The complexity and burden of atherosclerotic CAD were evaluated using an angiographic scoring system (SYNTAX scores) (Thuijs et al., 2019; Takahashi et al., 2020) and diagnosed by two professional cardiologists blinded to the clinical outcome (details are presented in the **Supplemental Materials: Supplementary Methods**).

Widely Targeted Metabolomic Analysis and Data Preprocessing

The hydrophilic and hydrophobic compounds were extracted from each plasma sample and detected via ultra-performance liquid chromatography and electrospray ionization-tandem mass spectrometry (UPLC-ESI-MS/MS) system in the positive and negative ionization modes in Metware Biotechnology (Wuhan, China). Details for the sample preparation and UPLC-MS/MS experiment parameters are provided in the **Supplemental Materials: Supplementary Methods**.

In total, 202 metabolites (including nucleosides, hormones, carbohydrates, organic acids and derivatives, and amino acids and derivatives) and 667 lipid species (including ceramides, cholesteryl esters, diacylglycerol, lysophosphatidic acid,

lysophosphatidylcholine, lysophosphatidylethanolamine, lysophosphatidylserine, monoglyceride, phosphatidic acid, phosphatidylcholines, phosphatidylglycerol, phosphatidylserine, phosphatidylethanolamine, and triacylglycerol) were identified and quantified.

Quality control (QC) samples were utilized for the normalization of the data. A QC sample was created via pooling aliquots from all samples and was injected every 10 samples throughout the run to assess the instrument's stability. Highly stable QC data showed that the run had great repeatability and reliability (**Supplementary Figure S1**).

For metabolomic and lipidomic analyses, raw signals with more than half of the missing rate in the QC samples (those with zero ion intensity) were removed. Missing metabolomic data were imputed by replacing the missing value with a minimum value of the metabolite quantified. To adjust signal drift, we applied the Quality Control–based Robust LOESS (LOcally Estimated Scatterplot Smoothing) Signal Correction (QC–RLSC) algorithm for analytical batch effect correction (Luan et al., 2018), which is an effective way to normalize the metabolic features to the QC samples within an analytical block. The dataset of discovery cohort after batch effect correction is available in Supplemental Materials: **Supplementary Table S3**. The dataset was then scaled by pareto scaling with procedures of mean centering and scaling to the square root of standard deviation (van den Berg et al., 2006). Then, the matrix was exported for further analysis.

Clustering of Metabolites Using Weighted Correlation Network Analysis.

A metabolic network was constructed by the weighted correlation network analysis (WCNA), which used metabolites' pairwise correlations to identify modules of highly correlated metabolites (Pei et al., 2017). An unsigned weighted metabolite co-expression network was constructed. Considering the scale-free topology fit index and mean connectivity, the soft-thresholding power $\beta = 4$ and min module size = 5 were chosen for the analysis. Spearman correlation between metabolite modules and clinical parameters was calculated using R. The Benjamini–Hochberg method was used to control the false discovery rate (FDR). Hub metabolites indicated a high degree of connectivity in biological interaction networks and ,thus, they were considered biologically important. Clusters of co-abundant plasma metabolites were identified using the "WGCNA" package in R.

Statistical Analysis

Among the baseline characteristics of the study population, continuous variables were described using medians (interquartile ranges) and were compared using Mann–Whitney U tests (non-normal distribution). Categorical variables were presented as counts (percentages) and were compared with Chi-squared tests. Statistical significance was determined as $p < 0.05$.

The linear regression analysis, adjusted for traditional Framingham risk factors, including age, sex, hypertension, diabetes mellitus, smoking, low-density lipoprotein cholesterol (LDLC), high-density lipoprotein cholesterol (HDLC), and triglycerides (TG) (Senthong et al., 2016), was applied to examine the associations of metabolomic and lipidomic profiles with SYNTAX score, SYNTAX score II, hs-CRP, and cardiac troponin I (cTnI) levels.

To assess the association of individual metabolomic and lipidomic signatures against the different stages of CAD, we performed adjusted logistic regression of metabolomic and lipidomic profiles against SCAD vs. UA and UA vs. MI to estimate the odds ratios (ORs) and 95% confidence intervals (CIs). To avoid potential confounders, traditional risk factors, including age, sex, hypertension, diabetes mellitus, smoking, LDLC, HDLC, and TG, were used as covariates for adjustment. Subjects with missing covariates were omitted. Statistical significance was determined as a p-value of <0.05. Open database sources, including the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (http://www.genome.jp/kegg/) and the MetaboAnalyst (https://www.metaboanalyst.ca) (version 4.0), were used to identify the highly enriched metabolic pathways based on the significantly differential levels of metabolites and lipid species.

In the development of diagnostic models to classify CAD subgroups, firstly, the following were added to develop the traditional risk factor-based model in a stepwise regression (forward and backward) with the aim to minimize the Akaike information criterion (AIC): age, sex, hypertension, diabetes mellitus, smoking, LDLC, HDLC, and TG (traditional risk factors); APOA and Lp(a) (risk lipid traits); and left ventricular ejection fraction (LVEF, heart function indicator). This procedure was performed within 10 iterations of a 5-fold cross-validation framework ("MASS", "caret" packages). Subsequently, metabolites and lipid species that were nominally significantly ($p < 0.05$) associated with UA (vs. SCAD) and MI (vs. UA) in the adjusted logistic regression analysis were included into least absolute shrinkage and selection operator (LASSO) penalized models ("glmnet" package) to further reduce the number of markers and select the most powerful predictive features. In the LASSO selection analysis, the optimal value for the tuning parameter λ was determined via 5-fold cross-validation (200 iterations). We adopted the largest value of lambda, such that the error was within one standard error of the minimum, known as "1-se" lambda. The relative contribution of features to classification assignment (UA vs. SCAD and MI vs. UA) was determined by the occurrence frequency in our multivariate model training. The feature was selected with an occurrence frequency of more than 100 times.

To evaluate the predictability of the models, a binary logistic regression model was then fitted using the chosen biomarkers as the covariates; this model was generated as follows: combined diagnostic score (probability) = 1/1 + exp [-(intercept + coefficient1 (biomarker1) + coefficient2 (biomarker2) . . . + coefficient n (biomarker n))]. The area under the curve (AUC, equivalently

TABLE 1 | Baseline characteristics of discovery cohort.

	SCAD (*N* = 310)	UA (*N* = 368)	MI (*N* = 264)	*p* Value	
				SCAD vs. UA	UA vs. MI
Age, years	63.2 (56.9, 70.4)	65.0 (57.5, 72.5)	61.6 (52.9, 68.9)	0.066	<0.001
Male	243 (78.4)	286 (77.7)	228 (86.4)	0.91	0.0081
SBP, mmHg	132 (120, 145)	130 (120, 143)	124 (110, 135)	0.37	<0.001
BMI, kg/m^2	24 (22, 27)	24 (22, 26)	24 (21, 26)	0.2	0.16
Current smokers	79 (25.6)	98 (27.1)	97 (37.0)	0.74	0.01
Comorbidities					
Hypertension	200 (64.7)	235 (63.9)	129 (48.9)	0.88	<0.001
Hyperlipidemia	39 (12.6)	41 (11.1)	24 (9.1)	0.64	0.48
Arrhythmia	30 (9.7)	38 (10.3)	11 (4.2)	0.89	<0.001
Diabetes mellitus	83 (26.7)	100 (27.2)	75 (28.4)	0.99	0.80
Laboratory data					
ALT, U/L	22.0 (17.9, 29.0)	23.5 (18.0, 33.0)	28.0 (19.0, 39.0)	0.049	<0.001
AST, U/L	23.0 (19.0, 27.0)	24.0 (20.0, 29.0)	26.0 (21.0, 36.0)	0.064	<0.001
GLUC, mmol/L	5.6 (5.0, 7.1)	5.8 (5.0, 7.3)	6.0 (5.1, 7.9)	0.21	0.13
eGFR, ml/min/1.73 m^2	90.1 (76.0, 106.5)	87.5 (70.8, 103.6)	87.7 (71.2, 103.0)	0.08	0.57
CK, U/L	89.5 (64.0, 122.0)	86.0 (63.0, 116.0)	86.0 (62.1, 132.8)	0.41	0.60
CKMB, U/L	5.9 (4.3, 8.2)	6.6 (4.7, 9.2)	6.9 (5.0, 9.2)	0.04	0.34
TC, mmol/L	4.1 (3.5, 5.0)	4.2 (3.5, 4.8)	4.0 (3.5, 4.7)	0.47	0.18
TG, mmol/L	1.3 (1.0, 1.9)	1.4 (1.0, 1.9)	1.3 (1.0 1.8)	0.29	0.51
LDLC, mmol/L	2.4 (1.9, 3.1)	2.5 (2.0, 3.1)	2.4 (1.9, 3.0)	0.73	0.52
HDLC, mmol/L	0.97 (0.84, 1.12)	0.94 (0.81, 1.12)	0.85 (0.72, 0.99)	0.1	<0.001
APOA, g/L	1.07 (0.90, 1.23)	1.01 (0.89, 1.21)	0.93 (0.80, 1.09)	0.11	<0.001
Lp(a), mg/dL	151.0 (76.0, 400.9)	169.4 (80.9, 357.3)	238.1 (118.9, 457.8)	0.98	0.0053
CREA, μmol/L	80.7 (69.0, 93.0)	81.4 (71.0, 97.0)	85.0 (73.5, 100.0)	0.21	0.026
BNP, pg/mL	114.4 (41.2, 278.0)	168.7 (59.9, 549.0)	670.1 (280.3, 1749.0)	0.0025	<0.001
hs-CRP, mg/L	2.3 (0.7, 4.5)	2.1 (1.0, 6.4)	6.4 (2.2,15.2)	0.14	<0.001
cTnI, μg/mL	0.01 (0.005, 0.04)	0.02 (0.008, 0.02)	0.3 (0.04, 1.9)	0.025	<0.001
Medication					
β-blockers	277 (89.6)	319 (86.9)	239 (90.5)	0.33	0.20
ACEI or ARB	191 (61.8)	225 (61.3)	182 (68.9)	0.91	0.058
CCBs	95 (30.7)	113 (30.8)	49 (18.6)	1	<0.001
PPIs	152 (49.2)	176 (48.0)	127 (48.1)	0.81	1
Cardiac function					
SYNTAX score	13.0 (8.0, 23.0)	14.0 (9.0, 22.0)	19.0 (10.0, 27.1)	0.76	<0.001
SYNTAX score II	26.0 (22.0, 32.0)	27.0 (21.0, 34.0)	28.0 (22.0, 34.0)	0.09	0.57
Counts of Long-lesion				0.53	0.0017
1	74 (23.9)	87 (23.6)	97 (36.7)		
2	25 (8.1)	30 (8.2)	23 (10.2)		
3	5 (1.6)	2 (0.5)	3 (1.1)		
4	-	1 (0.3)	-		
No. of SV				0.032	0.067
1	75 (26.8)	118 (35.0)	61 (25.3)		
2	110 (39.3)	109 (32.3)	81 (33.6)		
3	77 (27.5)	99 (29.4)	90 (37.3)		
LVEF	65.0 (61.0, 69.0)	65.0 (60.0, 69.0)	54.0 (45.0, 63.0)	0.31	<0.001
LVMI	112.5 (97.9, 132.8)	112.7 (95.6, 135.8)	125.6 (103.7, 150.2)	0.73	0.0015

Data are shown as median (interquartile range) or n (%). p values were calculated using Mann–Whitney U test for non-normally distributed continuous variables and the Chi-squared test for categorical variables. ACEI, angiotensin converting enzyme inhibitor; ALT, alanine aminotransferase; APOA, apolipoprotein A; ARB, angiotensin receptor Blocker; AST, aspartate aminotransferase; BMI, body mass index; BNP, B-type natriuretic peptide; CCB, calcium channel blocker; CK, creatine kinase; CKMB, MB isoenzyme of creatine kinase; CREA, Creatinine; cTnI, cardiac troponin I; eGFR, estimated glomerular filtration rate; GLUC, glucose; HDLC, high-density lipoprotein cholesterol; hs-CRP, high-sensitivity C-reactive protein; LDLC, low-density lipoprotein cholesterol; Lp(a), lipoprotein(a); LVEF, left ventricular ejection fraction; LVMI, left ventricular mass index; MI, myocardial infarction; No of SV, No. of stenosed vessels; PPI, proton pump inhibitor; SBP, systolic blood pressure; SCAD, stable coronary artery disease; SYNTAX, Synergy between PCI with TAXUS and Cardiac Surgery; TC, total cholesterol; TG, triacylglycerol; UA, unstable angina.

known as C-statistic) of the receiver operating characteristic (ROC) was applied to calculate the proportions of concordant pairs among all pairs of observation with 1.0 indicating perfect prediction accuracy. Moreover, the continuous net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were calculated in assessing the models. The 95% confidence intervals (CIs) were estimated for each parameter. The difference of combined diagnostic scores between CAD subgroups in the validation cohort was examined by the Wilcoxon rank sum test.

All the above analyses were conducted on the R platform (version 3.6.1, http://www.R-project.org/).

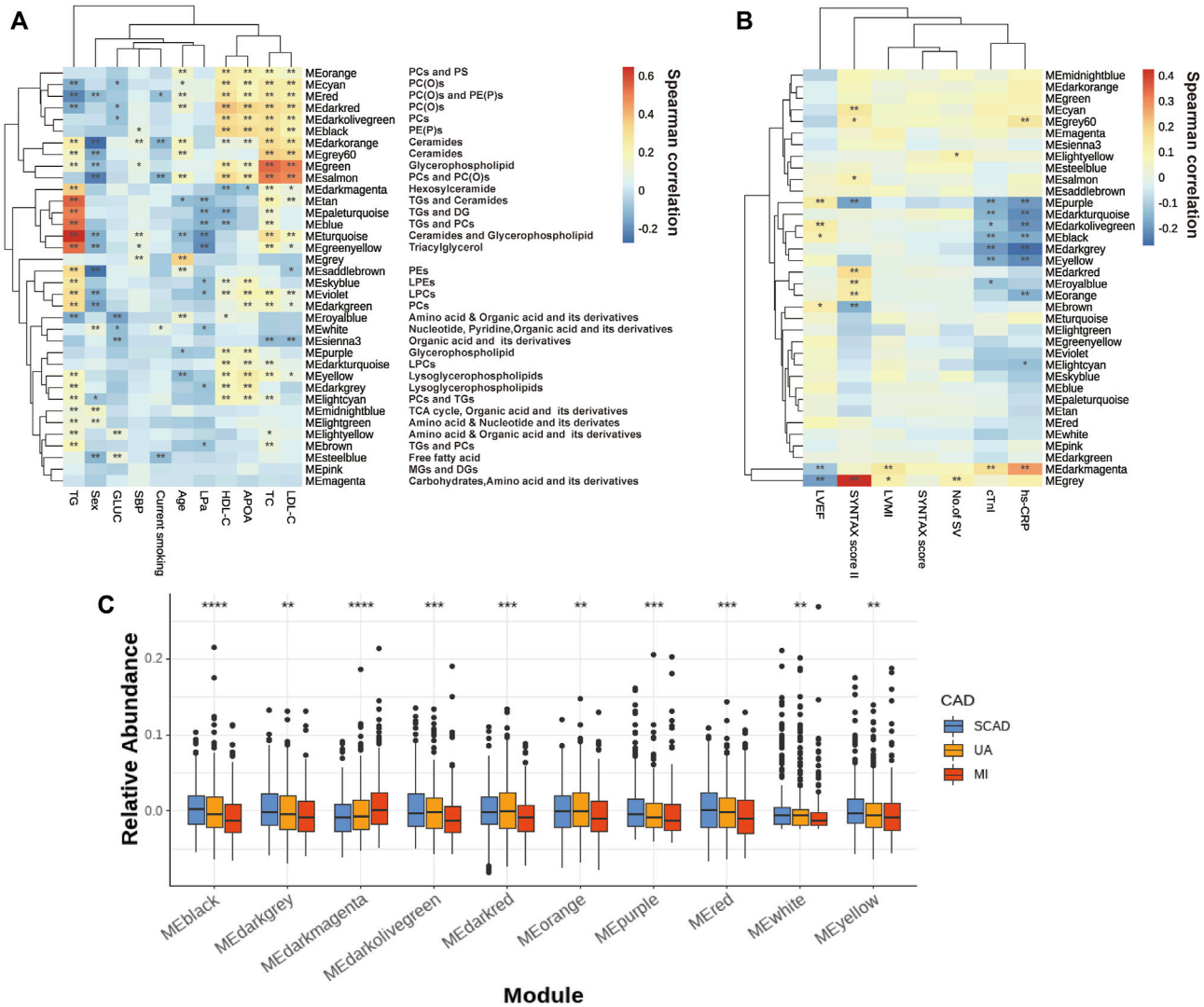

FIGURE 2 | Assessment of the metabolite modules associated with the CAD progression. **(A)** Heatmap of metabolite modules and major CAD risk factors. **(B)** Heatmap of metabolite modules and major CAD phenotypes. **(C)** Boxplot of 10 significantly altered metabolite modules between CAD subgroups compared by the Wilcoxon rank sum test. In **(A,B)**, the colors varying from blue to orange indicate negative to positive correlations, and *FDR< 0.05, **FDR <0.01 by the spearman correlation. In **(C)**, boxes represent the inter-quartile ranges, lines inside the boxes denote medians and the asterisk represents *p* values < 0.05 by the Wilcoxon rank sum test. APOA, apolipoprotein A; GLUC, glucose; HDLC, high-density lipoprotein cholesterol; LDLC, low-density lipoprotein cholesterol; LP(a), lipoprotein(a); LVEF, left ventricular ejection fraction; LVMI, left ventricular mass index; No. of SV, No. of stenosed vessels; SBP, systolic blood pressure; TC, total cholesterol; TG, triacylglycerol; other abbreviations as in **Figure 1**.

RESULTS

Characteristics of the Study Population

A total of 1,435 CAD patients were included from three independent centers in China (**Figure 1**). The discovery cohort included 942 participants enrolled at Guangdong Provincial People's Hospital, which were further classified into the following groups on the basis of the guidelines for diagnosis: SCAD ($N = 310$), UA ($N = 368$), and MI ($N = 264$). The baseline characteristics and laboratory data of each group are shown in **Table 1**. With disease shifting, the disturbance in lipid metabolism occurred with decreasing HDL-C and APOA but increasing Lp(a). Inflammatory state increased, as significant differences in hs-CRP levels were found between UA vs. MI ($p < 0.001$). The systemic atherosclerotic burden of CAD was determined using SYNTAX score system, and the median scores of each group were as follows: SCAD, 13.0 (8.0, 23.0); UA, 14.0 (9.0, 22.0); and MI, 19.0 (10.0, 27.1). The SYNTAX score showed a significant difference between SCAD vs. MI ($p < 0.001$) and UA vs. MI ($p < 0.001$). The MI group exhibited a higher proportion of three-stenosed vessels (37.56%), a larger left ventricular mass index (LVMI), and a lower LVEF. The median levels of cTnI, an indicator of myocardial infarction, were 0.01 (0.005, 0.04), 0.02 (0.008, 0.02), and 0.3 (0.04, 1.9) μg/ml in the SCAD, UA, and MI

groups, respectively. Significant differences in cTnI levels were found with SCAD vs. UA ($p = 0.025$) and UA vs. MI ($p < 0.001$).

The validation cohort from three centers included 493 participants. Their baseline characteristics are summarized in **Supplementary Table S11**.

Identification of Modules Associated With Multiple Clinical Traits

In the WCNA, 756 of the metabolites and lipid species in the discovery set were parsed into 35 co-abundance modules, whereas the gray module comprised unassigned metabolites and lipids due to weak correlation with others. However, each metabolite and lipid were further analyzed individually.

The correlations of thirty-five eigenmetabolites of the modules and external traits are shown in **Figures 2A,B** (Detailed annotation and information are listed in **Supplementary Tables S2–S5**). We identified 16 of 35 modules (45.7%) that were significantly associated with major CAD phenotypes (either SYNTAX scores, number of stenosed vessels, LVEF, LVMI, or cTnI levels). Moreover, by abundance cross-comparison, 10 of these 35 modules (28.6%) showed significant differences with $p < 0.05$ between CAD stages (**Figure 2C**). Notably, the eigenmetabolites in darkmagenta module (hexocylceramides) were positively correlated with hs-CRP levels (Rho = 0.28, FDR <0.001), cTnI (Rho = 0.17, FDR = 0.0034), LVMI (Rho = 0.14, FDR = 0.0030), and TG (Rho = 0.35, FDR <0.001) but negatively correlated with LVEF (Rho = −0.12, FDR = 0.0077) and HDL-C (Rho = −0.12, FDR = 0.0095) (**Figure 2A**). Moreover, the dark gray module (lysoglycerophospholipids) was negatively correlated with hs-CRP (Rho = −0.27, FDR <0.001) and cTnI (Rho = −0.19, FDR <0.001) but positively correlated with APOA (Rho = 0.23, FDR <0.001), HDLC (Rho = 0.18, FDR <0.001), and TG (Rho = 0.17, FDR <0.001, **Figure 2A**).

Several modules also showed strong correlations with the conventional lipid traits (**Supplementary Table S2**). Except for the modules with triglyceride inside, sphingolipids such as Cer and glycerophospholipids such as PEs and LPCs showed a close correlation with TC, LDLC, HDLC, and APOA. For example, the black module contains PE(P)s and dark red module contains PC(O), both of which showed a decreasing tendency with disease development (**Figure 2C**); these were positively correlated with HDLC (Rho = 0.334, FDR = 1.20E-23; Rho = 0.371, FDR= 2.06E-29) and APOA (Rho = 0.318, FDR= 4.83E-18; Rho = 0.339, FDR= 1.28E-20). Moreover, green (glycerophospholipids) and gray60 (ceramides), were positively correlated with TC (Rho = 0.541, FDR = 3.84E-68; Rho = 0.360, FDR = 1.00E-27) and LDLC (Rho = 0.506, FDR = 1.93E-58; Rho = 0.336, FDR = 6.43E-24).

Correlations Between Plasma Metabolite Levels and Severity Indicators

The linear regression analysis of metabolites and lipid species to the SYNTAX scores (atherosclerotic burden), cTnI (myocardial necrosis), and hs-CRP (inflammatory state) was conducted by adjusting for traditional risk factors, including age, sex, hypertension, diabetes mellitus, smoking, LDLC, HDLC, and TG. Numerous metabolites and lipid species showed strong association for one or more severity indicators (**Supplementary Tables S6–S9**). Three lipid species, namely, the hexosylceramide HexCer(d18:1/22:0) and the alkylphosphatidylcholine PC(O-32:0) and PC(O-42:3), were consistently significantly ($p < 0.05$) correlated with four severity indicators (**Supplementary Figure S2A**, **Supplementary Tables S6–S9**). We also found that high HexCer (d18:1/22:0) exhibited a high proportion of three-stenosed vessels (stenosed defined as >50%) with a univariate estimate of 0.11 ± 0.039, $p = 0.0057$, and an adjusted estimate (for the traditional risk factors above) of 0.082 ± 0.040, $p = 0.042$ (**Supplementary Figure S2B**).

Changes in the Plasma Metabolomic Features Between Different Coronary Artery Disease Subgroups

We focused on SCAD vs. UA for transition from coronary stability to instability and UA vs. MI for cardiac events. The logistic regression analysis of the metabolic and lipidomic profiles against UA (vs. SCAD) adjusting for traditional risk factors, identified 72 metabolites/lipid species that were significantly ($p < 0.05$) associated with UA (**Figure 3A**). The regression analysis against MI (vs. UA) conducted by adjusting for traditional risk factors identified 88 metabolites/lipid species that were significantly ($p < 0.05$) associated with MI (**Figure 3B**). The enrichment pathway analysis of significantly differential metabolites and lipid species for SCAD vs. UA and UA vs. MI are presented in **Supplementary Figure S3** and **Supplementary Table S10**. For SCAD vs. UA, the metabolism pathway significantly changed in glycerophospholipid metabolism ($p = 7.22E-05$, FDR = 6.06E-03) and valine, leucine, and isoleucine biosynthesis ($p = 1.25E-03$, FDR = 5.26E-02). Furthermore, the pathway analysis revealed that cysteine and methionine metabolism ($p = 4.88E-03$, FDR = 0.263) and glycerophospholipid metabolism ($p = 6.26E-03$, FDR = 0.263) were the main perturbed pathways for UA vs. MI. The glycerophospholipid metabolism was the most significantly altered pathway among all paired comparisons.

Generating Optimal Diagnosis Models for Subgroup Identification and Prediction

We focused on UA vs. MI for the prediction of cardiac events. In the first model (the traditional model), 11 conventional CAD risk factors, including age, sex, hypertension, diabetes mellitus, smoking, TG, LDLC, HDLC, APOA, Lp(a), and LVEF, were considered in the stepwise variable selection modeling. Finally five variables, namely, age, hypertension, TG, HDLC, and LVEF were retained in the model with minimal AIC and had an AUC value of 0.758 (**Figure 4B**; **Table 2**). In the second model (metabolic model), 88 metabolites/lipid species that were significantly ($p < 0.05$) associated with MI (vs. UA) were considered as input variables. The model was obtained by the LASSO logistic analysis (5-fold cross validation, 200 repeats, **Table 3**) and consisted of 16 metabolic biomarkers that performed similarly as the traditional model with continuous NRI of −0.0763 (95% CI, −0.271–0.118, $p = 0.443$) and IDI of −0.0178 (95% CI, −0.0669–0.0313, $p = 0.478$; **Tables 2, 3**). The third model (combined model) incorporated the 11 most predictive metabolic biomarkers to four conventional risk

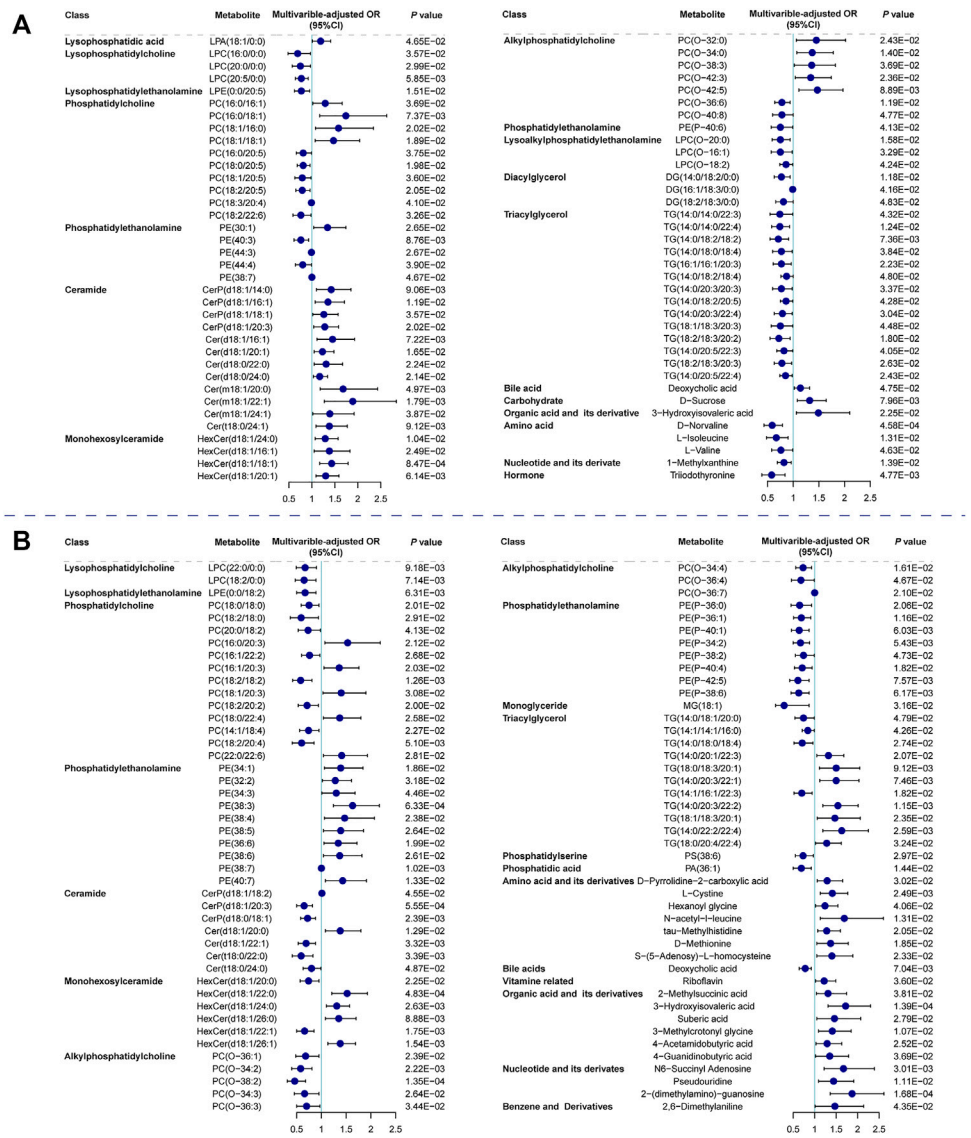

FIGURE 3 | Relationship between metabolic features against UA (vs. SCAD) and MI (vs. UA). Forest plot of odds ratios and 95% confidence intervals for logistic regression of individual metabolites/lipid species against **(A)** SCAD vs. UA ($p < 0.05$) **(B)** UA vs. MI ($p < 0.05$), adjusting for age, sex, hypertension, diabetes mellitus, smoking, LDLC, HDLC and TG. Cer, ceramide; CI, confidence interval; DG, diacylglycerol; HexCer, hexosylceramide; LPA, lysophosphatidic acid; LPC, lysophosphatidylcholine; LPC(O), lysoalkylphosphatidylethanolamine; LPE, lysophosphatidylethanolamine; MG, monoglyceride; OR, odds ratio; PA, phosphatidic acid; PC, phosphatidylcholine; PC(O), alkylphosphatidylcholine; PE, phosphatidylethanolamine; PE(P), phosphatidylethanolamine; PS, phosphatidylserine; other abbreviations as in **Figures 1**, **2**.

factors from LASSO selection (**Table 3**). It yielded better discrimination for the prediction of MI than the traditional model with an increased AUC from 0.758 to 0.823 (**Figure 4B**), a continuous NRI of 0.751 (95% CI, 0.571–0.932, $p < 0.0001$), and an IDI of 0.105 (95% CI, 0.072–0.137, $p < 0.0001$; **Table 2**).

However, the discriminating performance between SCAD and UA was not as satisfactory. The characteristics at baseline of the discovery cohort did not show many differences, and the traditional model based on AIC selection only included LVEF as the predictor with a poor AUC of 0.526. Nevertheless, the utilization of metabolic and lipidomic biomarkers provided another approach for discrimination. On the basis of the 72 variables ($p < 0.05$) selected from adjusted logistic regression, LASSO logistic analyses were further applied to identify the most predictive biomarkers. The optimized model consisting of 17 features exhibited a considerable performance with an AUC of 0.688 (**Tables 2, 4**). The ROC curve of SCAD vs. UA is plotted in **Figure 4A**. The diagnostic efficiency of the metabolic model showed a small improvement compared with that of the traditional model with an AUC from 0.562 to 0.688, a continuous NRI of 0.474 (95% CI, 0.289–0.659, $p < 0.0001$),

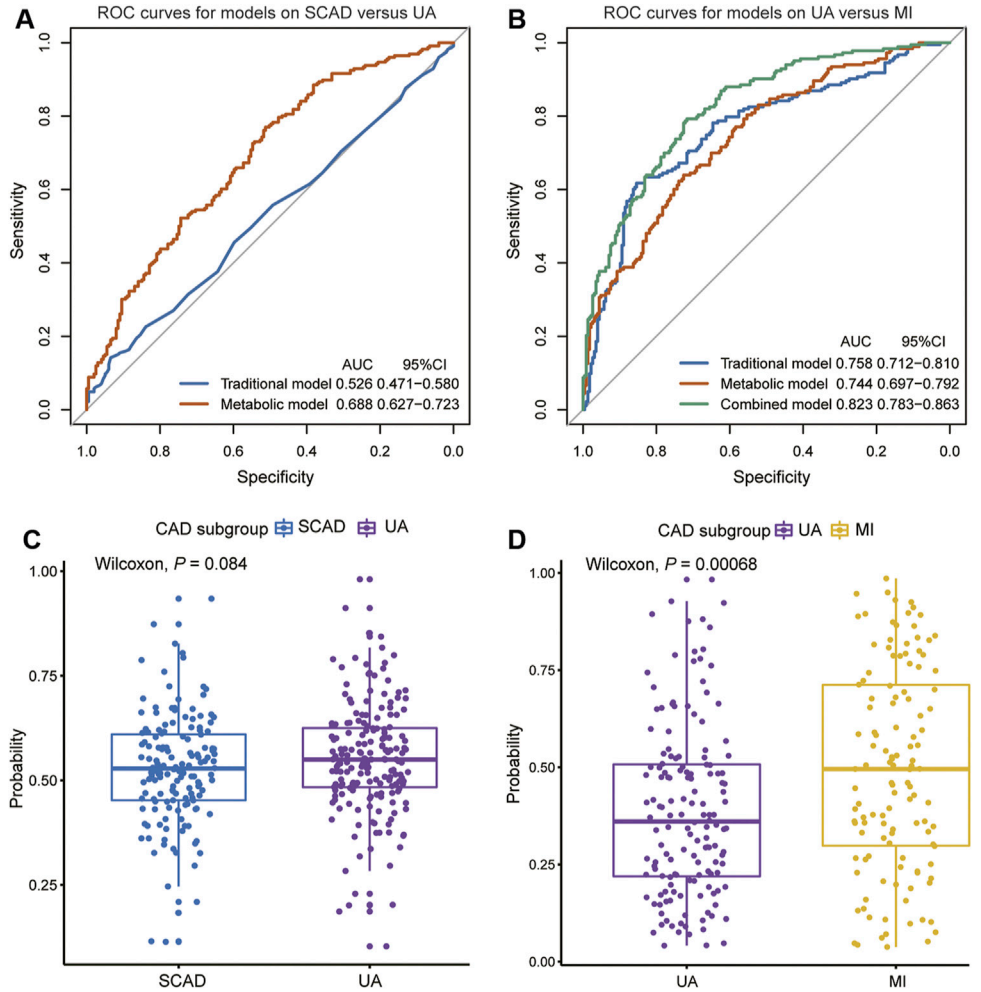

FIGURE 4 | Diagnostic performances in discovery cohort are shown via ROC curves between **(A)** SCAD vs. UA **(B)** UA vs. MI. The combined diagnosis score in validation cohort were compared between **(C)** SCAD and UA patients **(D)** UA and MI patients. AUC, area under curve; CI, confidence interval; ROC, receiver operating characteristic; other abbreviations as in **Figure 1**.

TABLE 2 | Model performance measures (95% CIs) for discrimination of CAD subtypes in the discovery cohort.

	Prediction of UA (vs. SCAD)				
Feature	**AUC**	**IDI**	***p* Value**	**Continuous NRI**	***p* Value**
Traditional model[a]	0.526 (0.471–0.580)				
Metabolic model[b]	0.688 (0.627–0.723)	0.105 (0.0749–0.135)	<0.0001	0.474 (0.289–0.659)	<0.0001
	Prediction of MI (vs. UA)				
Feature	**AUC**	**IDI**	***p* Value**	**Continuous NRI**	***p* Value**
Traditional model[c]	0.758 (0.712–0.810)				
Metabolic model[d]	0.744 (0.697–0.792)	−0.0178 (−0.0669–0.0313)	0.443	−0.0763 (−0.271–0.118)	0.478
Combined model[e]	0.823 (0.783–0.863)	0.105 (0.072–0.137)	<0.0001	0.751 (0.571–0.932)	<0.0001

[a]*Traditional model for UA (vs. SCAD) based on LVEF.*

[b]*Metabolic model for UA vs. SCAD based on: D-Norvaline, LPC(20:5/0:0), HexCer(d18:1/18:1), Cer(m18:1/22:1), 3,3′,5-Triiodo-L-thyronine, LPC(20:0/0:0), D-Sucrose, TG (18:2/18:3/20:2), 3-Hydroxy-3-methyl butyric acid, L-Isoleucine, Deoxycholic acid, 1-Methylxanthine, LPC (16:0/0:0), PC(18:3/20:4), PE (40:3), PC(O-42:5), and TG (14:0/20:3/20:3).*

[c]*Traditional model for MI (vs. UA) based on: age, hypertension, TG, HDLC and LVEF.*

[d]*Metabolic model for MI (vs. UA) based on: TG (14:0/20:3/22:2), 3-Hydroxy-3-methyl butyric acid, HexCer(d18:1/22:0), HexCer(d18:1/22:1), PA (36:1), PC(O-38:2), D-Methionine, Deoxycholic acid, PC(18:2/18:2), HexCer(d18:1/26:1), PC(O-34:2), PC(18:2/20:2), L-Cystine, TG (14:1/16:1/22:3), 3-Methylcrotonyl glycine, and MG (18:1).*

[e]*Combined model for MI (vs. UA) based on: age, LVEF, HexCer(d18:1/22:1), 3-Hydroxy-3-methyl butyric acid, CerP (d18:1/20:3), Cer(d18:1/22:1), PC(18:2/18:2), PC(16:0/20:3), HDLC, Hypertension, PC(18:2/20:4), PC(O-38:2), Deoxycholic acid, L-Cystine, and D-Methionine.*

*AUC, area under the curve; CAD, coronary artery disease; CI, confidence interval; IDI, integrated discrimination improvement; NRI, net reclassification improvement; other abbreviations as in **Tables 1, 4**.*

TABLE 3 | Feature inclusion frequency using LASSO based feature selection for MI (vs. UA).

	Metabolites-only model			Metabolites and risk factor model		
	Variable	**Frequency**	**Coefficient**	**Variable**	**Frequency**	**Coefficient**
1	TG (14:0/20:3/22:2)	200	0.16	Age	200	–0.02
2	3-Hydroxy-3-methyl butyric acid	200	0.13	LVEF	200	–0.05
3	HexCer(d18:1/22:0)	200	0.08	HexCer(d18:1/22:1)	200	–0.13
4	HexCer(d18:1/22:1)	200	–0.16	3-Hydroxy-3-methyl butyric acid	200	0.19
5	PA (36:1)	200	–0.17	CerP (d18:1/20:3)	199	–0.08
6	PC(O-38:2)	200	–0.22	Cer(d18:1/22:1)	185	–0.05
7	D-Methionine	199	0.12	PC (18:2/18:2)	174	–0.02
8	Deoxycholic acid	199	–0.06	PC (16:0/20:3)	166	0.15
9	PC (18:2/18:2)	199	–0.12	HDLC	149	–0.04
10	HexCer (d18:1/26:1)	182	0.06	Hypertension	126	–0.04
11	PC (O-34:2)	182	–0.03	PC (18:2/20:4)	126	–0.04
12	PC (18:2/20:2)	182	–0.07	PC(O-38:2)	126	–0.02
13	L-Cystine	137	0.02	Deoxycholic acid	126	–0.01
14	TG (14:1/16:1/22:3)	137	–0.02	L-Cystine	101	0.01
15	3-Methylcrotonyl glycine	104	0.02	D-Methionine	101	0.02
16	MG (18:1)	104	–0.05			

LASSO based feature selection was performed within a 5-fold cross-validation framework (200 iterations). Variables selected with frequency >100 times and their average coefficient (Coefficient) were indicated. MI, myocardial infarction; PA, phosphatidic acid; MG, monoglyceride, other abbreviations as in ***Table 4****.*

TABLE 4 | Feature inclusion frequency using LASSO based feature selection for UA (vs. SCAD).

	Metabolites-only model			Metabolites and risk factor model		
	Variable	**Frequency**	**Coefficient**	**Variable**	**Frequency**	**Coefficient**
1	D-Norvaline	195	–0.26	PE (40:3)	70	–0.03
2	LPC (20:5/0:0)	194	–0.08	Cer(m18:1/22:1)	70	0.06
3	HexCer(d18:1/18:1)	194	0.09	Cer(t18:0/24:1)	70	0.01
4	Cer(m18:1/22:1)	191	0.26	HexCer(d18:1/18:1)	70	0.02
5	3,3′,5-Triiodo-L-thyronine	191	–0.13	D-Norvaline	70	–0.05
6	LPC (20:0/0:0)	184	–0.08	LPC (20:5/0:0)	68	–0.01
7	D-Sucrose	184	0.05	CerP (d18:1/20:3)	58	0.01
8	TG (18:2/18:3/20:2)	177	–0.04	PC (16:0/16:1)	47	0.01
9	3-Hydroxy-3-methyl butyric acid	177	0.08	PC (18:3/20:4)	47	0.00
10	L-isoleucine	177	–0.07	LPA (18:1/0:0)	37	0.00
11	Deoxycholic acid	167	0.02	Deoxycholic acid	37	0.00
12	1-Methylxanthine	167	–0.03	1-Methylxanthine	37	0.00
13	LPC (16:0/0:0)	167	–0.05	LPC (16:0/0:0)	37	–0.01
14	PC (18:3/20:4)	153	0.00	D-Sucrose	24	0.00
15	PE (40:3)	153	–0.01	LVEF	8	0.00
16	PC(O-42:5)	153	0.05	L-isoleucine	3	0.00
17	TG (14:0/20:3/20:3)	153	–0.05	HDLC	2	0.00
18				LDLC	2	0.00
19				PE (38:7)	2	0.00
20				PC(O-42:5)	2	0.00

LASSO based feature selection was performed within a 5-fold cross-validation framework (200 iterations). Variables selected with frequency >100 times and their average coefficient (Coefficient) were indicated. Cer, ceramide; HDLC, high-density lipoprotein cholesterol; HexCer, hexosylceramide; LASSO, least absolute shrinkage and selection operator; LDLC, low-density lipoprotein cholesterol; LPA, lysophosphatidic acid; LPC, lysophosphatidylcholine; LVEF, left ventricular ejection fraction; PC, phosphatidylcholine; PC(O), alkylphosphatidylcholine; PE, phosphatidylethanolamine; SCAD, stable coronary artery disease; TG, triacylglycerol; UA, unstable angina.

and a IDI of 0.105 (95% CI, 0.0749–0.135, $p < 0.0001$; **Table 2**).

We subsequently assessed the optimal model for ability to differentiate among CAD subgroups in the validation cohort (**Supplementary Table S11**). The validation cohort was also divided into the following groups: SCAD ($N = 152$); UA ($N = 184$); and MI ($N = 157$). We used the established optimal LASSO models to further demonstrate the potential ability of subgroup discrimination. Consistently, the combined diagnostic score could help differentiate UA vs. MI patients ($p < 0.001$, **Figure 4D**). Similarly, the performance on SCAD and UA patients was not satisfactory ($p = 0.084$, **Figure 4C**).

DISCUSSION

In this study, we demonstrated that the plasma metabolomic and lipidomic signatures changed dynamically with CAD

progression, implying that CAD may involve a universal metabolomic and lipidomic disturbance. A total of 72 and 88 metabolites/lipid species have been identified to be significantly associated with UA (vs. SCAD) and MI (vs. UA), respectively. Moreover, the pathway analysis of these potential biomarkers indicated that glycerophospholipid metabolism exhibited the most significantly altered metabolic pathway in all paired comparisons. Lastly, the newly developed combined diagnostic models improved stratification performance of CAD subtypes compared with the traditional risk model, offering further evidence of dysbiotic metabolome and lipidome and highlighting its potential to distinguish various stages of CAD.

Specifically, the co-clustering modules within lipid classes including phosphatidylcholine (PC), lysophosphatidylcholine (LPC), lysophosphatidylethanolamine (LPE), phosphatidylethanolamine (PE(P)), and alkylphosphatidylcholine (PC(O)) tended to decrease with plaque instability and were inversely correlated with CAD severity and myocardial markers. Moreover, modules containing five PCs were positively correlated with HDLC (as seen in darkolivegreen module, Rho = 0.314, FDR = 5.77E-21) and primarily decreased in the MI group. Different PC species showed diverse effects on CAD progression. We observed that PCs with longer and more unsaturated acyl chain had an inverse association with UA (vs. SCAD). PCs are the most abundant membrane lipids in mammals (van Meer et al., 2008) and are the key structural molecules in the surface monolayer of HDL particles (Kontush et al., 2013). Shorter and highly saturated acyl chains of PC molecules confer less fluidity of the lipid monolayer, thereby directly affecting HDL's ability to accept cholesterol from peripheral tissues and phospholipid hydroperoxides from low-density lipoproteins (Kontush et al., 2013; Toledo et al., 2017).

PC in lipoproteins or from cell membrane can be further hydrolyzed on the sn-2 position fatty acid to generate LPC and free fatty acid by the phospholipase A2 enzyme (Norris et al., 2014). Although the catalysis of phospholipase A2 was expected to generate LPC to promote inflammation and atherosclerosis development (Huang et al., 2020; Schmitz and Ruebsaamen, 2010), most LPC species exhibited a negative association with UA (vs. SCAD) and MI (vs. UA). As shown in **Figure 3**, LPC(16:0/0:0), LPC(20:0/0:0), and LPC(20:5/0:0) were decreased in UA patients compared with SCAD patients. LPC(22:0/0:0) and LPC(18:2/0:0) were decreased in MI patients (vs. UA), which is consistent with previously reported results (Fan et al., 2016; Lu et al., 2017). LPC is reportedly an inducer of endothelial dysfunction and a regulator of vascular tone (Zhang et al., 2009; Paapstel et al., 2018). Lower levels of LPC in the circulation may result from the increase in the catabolism of these species or to their more efficient removal from blood circulation into the tissues, either in the form of modified lipoprotein or directly from albumin (Meikle et al., 2011).

One of the prominent features observed was that a number of PE(P) species with polyunsaturated fatty acids displayed a significant inverse association with MI compared with UA patients (**Figure 3B**). Alkylphospholipids [alkylphosphatidylcholine, PC(O) and alkylphosphatidylethanolamine, PE(O)] and alkenylphospholipids [primarily presented as phosphatidylcholine, PC(P), and phosphatidylethanolamine, PE(P) species, equivalently known as plasmalogens] have been proposed to protect against atherosclerosis due to their antioxidant characteristics and a high proportion of polyunsaturated fatty acids and alkyl/alkenyl linked fatty acids. They are more susceptible to oxidation under heightened oxidative stress (Lessig and Fuchs, 2009; Ford, 2010). In addition, plasmalogens are essential for intracellular cholesterol transport (Munn et al., 2003) and HDLC-mediated cholesterol efflux (Maeba et al., 2018). Recently, the inclusion of plasmalogens into reconstituted HDL improved the lipoprotein anti-apoptotic activity on endothelial cells (Sutter et al., 2015). Therefore, low plasmalogens levels in plasma may reflect the high oxidative stress and the action of reactive oxygen species on these lipids.

However, the module containing ceramides was elevated with disease shifting and was positively associated with CAD severity, myocardial markers, and inflammatory state. Notably, hexosylceramide species played an important role in the development of CAD. Specific hexosylceramide species [e.g., HexCer(22:0/0:0)] were related to the enhanced coronary atherosclerosis burden. Both mono- and dihexosylceramide have a direct association with the risk of future cardiovascular events in patients with type 2 diabetes, which is a potential atherogenesis-contributing factor (Alshehry et al., 2016).

Some plasma ceramide levels were observed to be up-regulated with the disease shifting direction as SCAD, UA, and MI and positively correlated with atherosclerosis burden quantified by the SYNTAX score and SYNTAX score II and the evidence of subclinical myonecrosis quantified by cTnI. This result corroborates the finding that elevated plasma ceramide levels are independent biomarkers of MACE (Laaksonen et al., 2016). Cer (d18:1/20:1) was significantly elevated in UA (vs. SCAD) and we previously reported that Cer (d18:1/20:1) was negatively related with LVEF and could serve as an independent predictor of MACE and all-cause mortality (Qin et al., 2020). As the metabolites of sphingolipid, ceramides are considered lipotoxic inducers of disturbed glucose homeostasis and insulin resistance and causative agents in the pathophysiology of atherosclerosis (Chaurasia and Summers, 2015; Laaksonen et al., 2016). Studies in rodent models revealed that the inhibition of ceramide synthesis prevents ischemic cardiomyopathy-related heart failure post hypoxia or MI while simultaneously diminishing ventricular remodeling and lowering cell death rates and changing the abundance of proinflammatory detrimental neutrophils (Park and Goldberg, 2012; Hadas et al., 2020). The underlying functions of ceramides involve the promotion of lipoprotein transport into the arterial wall, platelet activation, and endothelial dysfunction via uncoupling of NO signaling pathways (Chaurasia and Summers, 2015; Meikle and Summers, 2017).

In addition to the use of certain lipid species of sphingolipids and glycerophospholipids as predictors of CAD progression, the downregulation of deoxycholic acid in MI, which plays key roles in bile acid and cholesterol metabolism, indicated that the

metabolism of cholesterol and phospholipids might be inhibited (Sayin et al., 2013). Moreover, low triiodothyronine was inversely associated with UA occurrence, which indicated a close link between thyroid function and atherosclerosis process. Since triiodothyronine is the most biologically active thyroid hormone, it plays a vital role in regulating heart rate, contractile force, and peripheral arterial resistance (Jabbar et al., 2017). A meta-analysis of 56 studies showed that a reduced serum triiodothyronine level was further associated with the increased risk of all-cause and cardiogenic death, and was an independent predictor of MACE (Wang B. et al., 2017). Lastly, a number of amino acids and their derivatives were altered with CAD shifting. Elevated levels of plasma cystine (the disulphide form of cysteine) were positively associated with MI (vs. UA) and werepositively correlated with SYNTAX score II and hs-CRP, which is indicated to link with a higher oxidative stress and endothelial dysfunction (Oliveira and Laurindo, 2018). A high level of methionine served as a strong predictor for MI (vs. UA) selected by LASSO. A previous study has shown that methionine promotes atherosclerotic plaques independent of homocysteine levels in the rodent model (Selhub and Troen, 2016).

Our study had some limitations that needed be considered. First, due to the upgrading of analytical platform and technical issues with the mass spectrometry, the metabolites and lipid species detected were not in accordance, thereby resulting in the lack of four independent predictors for UA (vs. SCAD) model and one for MI (vs. UA) model which affected the model estimation in the verification cohort. Second, site-to-site and observer-to-observer variations in the evaluation of coronary stenosis may exist, leading to diagnostic bias. Third, the improvement in AUC for a model is often very minor, yet the category-free NRI may overstate the incremental value of a biomarker. Last, our study population tended to consist of middle-aged to elderly Chinese patients. Thus, other ethnicities within Asia and other races, such as Caucasians and Africans, should be included in future studies.

CONCLUSION

Multiple plasma metabolites and lipid species differed between CAD subgroups, and the alterations were correlated with CAD severity. The metabolites involved in glycerophospholipid metabolism appeared to be a predominant alteration in CAD progression. A small number of these biomarkers significantly improved the diagnostic value for differentiating patients between CAD types. These findings may help to predict disease progression and clinical outcome and indicate the potential for novel intervention strategies to attenuate disease 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 reviewedand approved by the Medical Ethical Review Committee of Guangdong Provincial People's Hospital (GDREC2010137 and GDREC2017071H). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SZ was the principal investigator of this study and designed the study. HC and ZW performed the data analysis and drafted the manuscript. MQ and SZ assisted in statistical analysis and critically revised the manuscript. BZ, LL, QM, CL, XC, HL, and WL were responsible for patient recruitment and clinical data collection. All authors reviewed and approved the final manuscript.

FUNDING

This work was funded by grants from the National Natural Science Foundation of China (No. 81872934 (Principle investigator: SZ), 81673514 (Principle investigator: SZ)), the Key-Area Research and Development Program of Guangdong Province, China (No. 2019B020229003 (Principle investigator: SZ)), Science and Technology Planning Project of Guangdong Province, China (No. 2017B030314041 (Principle investigator: Jiyan Chen)). The funders were not involved in designing the study; collecting, analyzing, or interpreting the data; or in writing or submitting the manuscript for publication.

ACKNOWLEDGMENTS

We are grateful to the Metware Biotechnology Co., Ltd. for their contribution and assistance in metabolomic analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.632950/full#supplementary-material.

REFERENCES

Alshehry, Z. H., Mundra, P. A., Barlow, C. K., Mellett, N. A., Wong, G., McConville, M. J., et al. (2016). Plasma lipidomic profiles improve on traditional risk factors for the prediction of cardiovascular events in type 2 diabetes mellitus. *Circulation* 134 (21), 1637–1650. doi:10.1161/CIRCULATIONAHA.116.023233

Cai, L., Bai, X., Lei, H., Wu, H., Liu, Y., Zhu, Q., et al. (2018). High plasma exposure of statins associated with increased risk of contrast-induced acute kidney injury in Chinese patients with coronary artery disease. *Front. Pharmacol.* 9, 427. doi:10.3389/fphar.2018.00427

Chaurasia, B., and Summers, S. A. (2015). Ceramides - lipotoxic inducers of metabolic disorders. *Trends Endocrinol. Metab.* 26 (10), 538–550. doi:10.1016/j.tem.2015.07.006

Dannenberg, L., Zikeli, D., Benkhoff, M., Ahlbrecht, S., Kelm, M., Levkau, B., et al. (2020). Targeting the human microbiome and its metabolite TMAO in cardiovascular prevention and therapy. *Pharmacol. Ther.* 213, 107584. doi:10.1016/j.pharmthera.2020.107584

Fan, Y., Li, Y., Chen, Y., Zhao, Y. J., Liu, L. W., Li, J., et al. (2016). Comprehensive metabolomic characterization of coronary artery diseases. *J. Am. Coll. Cardiol.* 68 (12), 1281–1293. doi:10.1016/j.jacc.2016.06.044

Fihn, S. D., Blankenship, J. C., Alexander, K. P., Bittl, J. A., Byrne, J. G., Fletcher, B. J., et al. (2014). ACC/AHA/AATS/PCNA/SCAI/STS focused update of the guideline for the diagnosis and management of patients with stable ischemic heart disease. *J. Am. Coll. Cardiol.* 64 (18), 1929–1949. doi:10.1016/j.jacc.2014.07.017

Ford, D. A. (2010). Lipid oxidation by hypochlorous acid: chlorinated lipids in atherosclerosis and myocardial ischemia. *Clin. Lipidol.* 5 (6), 835–852. doi:10.2217/clp.10.68

Gencer, B., Li, X. S., Gurmu, Y., Bonaca, M. P., Morrow, D. A., Cohen, M., et al. (2020). Gut microbiota-dependent trimethylamine N-oxide and cardiovascular outcomes in patients with prior myocardial infarction: a nested case control study from the PEGASUS-TIMI 54 trial. *J. Am. Heart Assoc.* 9 (10), e015331. doi:10.1161/JAHA.119.015331

Hadas, Y., Vincek, A. S., Youssef, E., Żak, M. M., Chepurko, E., Sultana, N., et al. (2020). Altering sphingolipid metabolism attenuates cell death and inflammatory response after myocardial infarction. *Circulation* 141 (11), 916–930. doi:10.1161/CIRCULATIONAHA.119.041882

Huang, F., Wang, K., and Shen, J. (2020). Lipoprotein-associated phospholipase A2: the story continues. *Med. Res. Rev.* 40 (1), 79–134. doi:10.1002/med.21597

Jabbar, A., Pingitore, A., Pearce, S. H. S., Zaman, A., Iervasi, G., and Razvi, S. (2017). Thyroid hormones and cardiovascular disease. *Nat. Rev. Cardiol.* 14 (1), 39–55. doi:10.1038/nrcardio.2016.174

Kontush, A., Lhomme, M., and Chapman, M. J. (2013). Unraveling the complexities of the HDL lipidome. *J. Lipid Res.* 54 (11), 2950–2963. doi:10.1194/jlr.R036095

Laaksonen, R., Ekroos, K., Sysi-Aho, M., Hilvo, M., Vihervaara, T., Kauhanen, D., et al. (2016). Plasma ceramides predict cardiovascular death in patients with stable coronary artery disease and acute coronary syndromes beyond LDL-cholesterol. *Eur. Heart J.* 37 (25), 1967–1976. doi:10.1093/eurheartj/ehw148

Lessig, J., and Fuchs, B. (2009). Plasmalogens in biological systems: their role in oxidative processes in biological membranes, their contribution to pathological processes and aging and plasmalogen analysis. *Curr. Med. Chem.* 16 (16), 2021–2041. doi:10.2174/092986709788682164

Lu, J., Chen, B., Chen, T., Guo, S., Xue, X., Chen, Q., et al. (2017). Comprehensive metabolomics identified lipid peroxidation as a prominent feature in human plasma of patients with coronary heart diseases. *Redox Biol.* 12, 899–907. doi:10.1016/j.redox.2017.04.032

Luan, H., Ji, F., Chen, Y., and Cai, Z. (2018). statTarget: a streamlined tool for signal drift correction and interpretations of quantitative mass spectrometry-based omics data. *Anal. Chim. Acta* 1036, 66–72. doi:10.1016/j.aca.2018.08.002

Maeba, R., Kojima, K. I., Nagura, M., Komori, A., Nishimukai, M., Okazaki, T., et al. (2018). Association of cholesterol efflux capacity with plasmalogen levels of high-density lipoprotein: a cross-sectional study in chronic kidney disease patients. *Atherosclerosis* 270, 102–109. doi:10.1016/j.atherosclerosis.2018.01.037

Malakar, A. K., Choudhury, D., Halder, B., Paul, P., Uddin, A., and Chakraborty, S. (2019). A review on coronary artery disease, its risk factors, and therapeutics. *J. Cell. Physiol.* 234 (10), 16812–16823. doi:10.1002/jcp.28350

Meikle, P. J., and Summers, S. A. (2017). Sphingolipids and phospholipids in insulin resistance and related metabolic disorders. *Nat. Rev. Endocrinol.* 13 (2), 79–91. doi:10.1038/nrendo.2016.169

Meikle, P. J., Wong, G., Tsorotes, D., Barlow, C. K., Weir, J. M., Christopher, M. J., et al. (2011). Plasma lipidomic analysis of stable and unstable coronary artery disease. *Arterioscler Thromb. Vasc. Biol.* 31 (11), 2723–2732. doi:10.1161/ATVBAHA.111.234096

Munn, N. J., Arnio, E., Liu, D., Zoeller, R. A., and Liscum, L. (2003). Deficiency in ethanolamine plasmalogen leads to altered cholesterol transport. *J. Lipid Res.* 44 (1), 182–192. doi:10.1194/jlr.m200363-jlr200

Norris, P. C., Gosselin, D., Reichart, D., Glass, C. K., and Dennis, E. A. (2014). Phospholipase A2 regulates eicosanoid class switching during inflammasome activation. *Proc. Natl. Acad. Sci. U.S.A.* 111 (35), 12746–12751. doi:10.1073/pnas.1404372111

Oliveira, P. V. S., and Laurindo, F. R. M. (2018). Implications of plasma thiol redox in disease. *Clin. Sci. (Lond)* 132 (12), 1257–1280. doi:10.1042/CS20180157

Paapstel, K., Kals, J., Eha, J., Tootsi, K., Ottas, A., Piir, A., et al. (2018). Inverse relations of serum phosphatidylcholines and lysophosphatidylcholines with vascular damage and heart rate in patients with atherosclerosis. *Nutr. Metab. Cardiovasc. Dis.* 28 (1), 44–52. doi:10.1016/j.numecd.2017.07.011

Park, T. S., and Goldberg, I. J. (2012). Sphingolipids, lipotoxic cardiomyopathy, and cardiac failure. *Heart Fail. Clin.* 8 (4), 633–641. doi:10.1016/j.hfc.2012.06.003

Pei, G., Chen, L., and Zhang, W. (2017). WGCNA application to proteomic and metabolomic data analysis. *Methods Enzymol.* 585, 135–158. doi:10.1016/bs.mie.2016.09.016

Poss, A. M., Maschek, J. A., Cox, J. E., Hauner, B. J., Hopkins, P. N., Hunt, S. C., et al. (2020) Machine learning reveals serum sphingolipids as cholesterol-independent biomarkers of coronary artery disease. *J. Clin. Invest.* 130 (3), 1363–1376. doi:10.1172/JCI131838

Qin, M., Zhu, Q., Lai, W., Ma, Q., Liu, C., Chen, X., et al. (2020). Insights into the prognosis of lipidomic dysregulation for death risk in patients with coronary artery disease. *Clin. Transl. Med.* 10 (5), e189. doi:10.1002/ctm2.189

Sayin, S. I., Wahlström, A., Felin, J., Jäntti, S., Marschall, H. U., Bamberg, K., et al. (2013). Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cell Metab.* 17 (2), 225–235. doi:10.1016/j.cmet.2013.01.003

Schmitz, G., and Ruebsaamen, K. (2010). Metabolism and atherogenic disease association of lysophosphatidylcholine. *Atherosclerosis* 208 (1), 10–18. doi:10.1016/j.atherosclerosis.2009.05.029

Selhub, J., and Troen, A. M. (2016). Sulfur amino acids and atherosclerosis: a role for excess dietary methionine. *Ann. N.Y. Acad. Sci.* 1363, 18–25. doi:10.1111/nyas.12962

Senthong, V., Li, X. S., Hudec, T., Coughlin, J., Wu, Y., Levison, B., et al. (2016). Plasma trimethylamine N -oxide, a gut microbe-generated phosphatidylcholine metabolite, is associated with atherosclerotic burden. *J. Am. Coll. Cardiol.* 67 (22), 2620–2628. doi:10.1016/j.jacc.2016.03.546

Shao, C., Wang, J., Tian, J., and Tang, Y. D. (2020). Coronary artery disease: from mechanism to clinical practice. *Adv. Exp. Med. Biol.* 1177, 1–36. doi:10.1007/978-981-15-2517-9_1

Sutter, I., Velagapudi, S., Othman, A., Riwanto, M., Manz, J., Rohrer, L., et al. (2015). Plasmalogens of high-density lipoproteins (HDL) are associated with coronary artery disease and anti-apoptotic activity of HDL. *Atherosclerosis* 241 (2), 539–546. doi:10.1016/j.atherosclerosis.2015.05.037

Takahashi, K., Serruys, P. W., Fuster, V., Farkouh, M. E., Spertus, J. A., Cohen, D. J., et al. (2020). Redevelopment and validation of the SYNTAX score II to individualise decision making between percutaneous and surgical revascularisation in patients with complex coronary artery disease: secondary analysis of the multicentre randomised controlled SYNTAXES trial with external cohort validation. *Lancet* 396 (10260), 1399–1412. doi:10.1016/S0140-6736(20)32114-0

Tang, W. H. W., Bäckhed, F., Landmesser, U., and Hazen, S. L. (2019). Intestinal microbiota in cardiovascular health and disease. *J. Am. Coll. Cardiol.* 73 (16), 2089–2105. doi:10.1016/j.jacc.2019.03.024

Thuijs, D., Kappetein, A. P., Serruys, P. W., Mohr, F. W., Morice, M. C., Mack, M. J., et al. (2019). Percutaneous coronary intervention versus coronary artery bypass grafting in patients with three-vessel or left main coronary artery disease: 10-year follow-up of the multicentre randomised controlled SYNTAX trial. *Lancet.* 394 (10206), 1325–1334. doi:10.1016/S0140-6736(19)31997-X

Toledo, E., Wang, D. D., Ruiz-Canela, M., Clish, C. B., Razquin, C., Zheng, Y., et al. (2017). Plasma lipidomic profiles and cardiovascular events in a randomized intervention trial with the Mediterranean diet. *Am. J. Clin. Nutr.* 106 (4), 973–983. doi:10.3945/ajcn.116.151159

Tzoulaki, I., Castagné, R., Boulangé, C. L., Karaman, I., Chekmeneva, E., Evangelou, E., et al. (2019). Serum metabolic signatures of coronary and

carotid atherosclerosis and subsequent cardiovascular disease. *Eur. Heart J.* 40 (34), 2883–2896. doi:10.1093/eurheartj/ehz235

van den Berg, R. A., Hoefsloot, H. C., 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, 142. doi:10.1186/1471-2164-7-142

van Meer, G., Voelker, D. R., and Feigenson, G. W. (2008). Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cel Biol* 9 (2), 112–124. doi:10.1038/nrm2330

Virani, S. S., Alonso, A., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., et al. (2020). Heart disease and stroke statistics-2020 update: a report from the American heart association. *Circulation* 141 (9), e139–e596. doi:10.1161/cir.0000000000000757

Wang, B., Liu, S., Li, L., Yao, Q., Song, R., Shao, X., et al. (2017). Non-thyroidal illness syndrome in patients with cardiovascular diseases: a systematic review and meta-analysis. *Int. J. Cardiol.* 226, 1–10. doi:10.1016/j.ijcard.2016.10.039

Wang, D. D., Toledo, E., Hruby, A., Rosner, B. A., Willett, W. C., Sun, Q., et al. (2017). Plasma ceramides, mediterranean diet, and incident cardiovascular disease in the PREDIMED trial (Prevención con Dieta Mediterránea). *Circulation* 135 (21), 2028–2040. doi:10.1161/CIRCULATIONAHA.116.024261

Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* 15 (7), 473–484. doi:10.1038/nrd.2016.32

Zhang, R., Bai, N., So, J., Laher, I., MacLeod, K. M., and Rodrigues, B. (2009). The ischemic metabolite lysophosphatidylcholine increases rat coronary arterial tone by endothelium-dependent mechanisms. *J. Mol. Cell Cardiol.* 47 (1), 112–120. doi:10.1016/j.yjmcc.2009.03.026

Zhong, S., Li, L., Shen, X., Li, Q., Xu, W., Wang, X., et al. (2019). An update on lipid oxidation and inflammation in cardiovascular diseases. *Free Radic. Biol. Med.* 144, 266–278. doi:10.1016/j.freeradbiomed.2019.03.036

Zhou, M., Wang, H., Zeng, X., Yin, P., Zhu, J., Chen, W., et al. (2019). Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 394 (10204), 1145–1158. doi:10.1016/S0140-6736(19)30427-1

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 14 May 2021
doi: 10.3389/fmolb.2021.682600

Metabolomics Reveals Differences in Aqueous Humor Composition in Patients With and Without Pseudoexfoliation Syndrome

Diana Anna Dmuchowska[1]†, Karolina Pietrowska[2†], Pawel Krasnicki[1], Tomasz Kowalczyk[2], Magdalena Misiura[3], Emil Tomasz Grochowski[1], Zofia Mariak[1], Adam Kretowski[2,4] and Michal Ciborowski[2]**

[1]Department of Ophthalmology, Medical University of Bialystok, Bialystok, Poland, [2]Metabolomics Laboratory, Clinical Research Center, Medical University of Bialystok, Bialystok, Poland, [3]Department of Pharmaceutical Analysis, Medical University of Bialystok, Bialystok, Poland, [4]Department of Endocrinology, Diabetology and Internal Medicine, Medical University of Bialystok, Bialystok, Poland

OPEN ACCESS

Edited by:
Michal Jan Markuszewski,
Medical University of Gdansk, Poland

Reviewed by:
Ursula Schlötzer-Schrehardt,
University of Erlangen Nuremberg, Germany
Chen Yang,
Chinese Academy of Sciences (CAS), China

***Correspondence:**
Michal Ciborowski
michal.ciborowski@umb.edu.pl
Diana Anna Dmuchowska
diana.dmuchowska@umb.edu.pl

†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: 18 March 2021
Accepted: 04 May 2021
Published: 14 May 2021

Citation:
Dmuchowska DA, Pietrowska K, Krasnicki P, Kowalczyk T, Misiura M, Grochowski ET, Mariak Z, Kretowski A and Ciborowski M (2021) Metabolomics Reveals Differences in Aqueous Humor Composition in Patients With and Without Pseudoexfoliation Syndrome. Front. Mol. Biosci. 8:682600. doi: 10.3389/fmolb.2021.682600

Pseudoexfoliation syndrome (XFS) is stress- or inflammation-induced elastosis accompanied by excessive production of microfibrils and their deposition in the anterior segment of the eye. Approximately 60–70 million people are affected by XFS worldwide. It is a component of a systemic disorder, considered a major risk factor for accelerated cataract formation, cataract surgery complications and development of glaucoma, which untreated or inadequately treated may lead to blindness. Moreover, XFS has been associated with cardiovascular and cerebrovascular morbidity, dementia, sensorineural hearing loss and pelvic organ prolapse. The pathogenesis of XFS has not been fully elucidated yet. Aqueous humor (AH) is a transparent fluid filling the anterior and posterior chambers of the eye. Determination of AH metabolites that are characteristic for XFS may provide valuable information about the molecular background of this ocular disorder. The aim of this study was to compare the composition of AH in XFS and non-XFS patients undergoing cataract surgery. The AH samples from 34 patients (15 with XFS and 19 without) were analyzed using liquid chromatography coupled to a Quadrupole Time-of-Flight mass spectrometer (LC-QTOF-MS). The obtained metabolic fingerprints were analyzed using multivariate statistics. Eleven statistically significant metabolites were identified. Compared with the non-XFS group, the AH of patients with XFS contained significantly lower levels of amino acids and their derivatives, for example, arginine (–31%, VIP = 2.38) and homo-arginine (–19%, VIP = 1.38). Also, a decrease in the levels of two acylcarnitines, hydroxybutyrylcarnitine (–29%, VIP = 1.24) and decatrienoylcarnitine (–46%, VIP = 1.89), was observed. However, the level of indoleacetaldehyde in XFS patients was significantly higher (+96%, VIP = 2.64). Other significant metabolites were two well-recognized antioxidants, ascorbic acid (–33%, VIP = 2.11) and hydroxyanthranilic acid (–33%, VIP = 2.25), as well as S-adenosylmethionine, a compound with anti-inflammatory properties (–29%, VIP = 1.93). Metabolic pathway analysis demonstrated that the identified metabolites belonged to eight metabolic pathways, with cysteine and methionine metabolism as well as arginine and proline metabolism being the most

frequently represented. XFS can be associated with enhanced oxidative stress and inflammation, as well as with the disturbances of cellular respiration and mitochondrial energy production. Implementation of non-targeted metabolomics provided a better insight into the still not fully understood pathogenesis of XFS.

Keywords: ophthalmology, pseudoexfoliation syndrome, aqueous humor, metabolomics, mass spectrometry

INTRODUCTION

Pseudoexfoliation is a stress- and inflammation-induced deposition of extracellular fibrillary protein material visible in the anterior segment of the eye. The complex is highly cross-linked and glycosylated. Among others, it contains elastic microfibrillar components, as well as noncollagenous components, such as laminin, nidogen, and fibronectin. The condition is caused by excessive production and reduced degradation of these constituents (Kivelä, 2018; Zenkel, 2018). Aqueous humor (AH) is a transparent fluid filling the anterior and posterior chamber of the eye. AH maintains intraocular pressure and provides nutrients for avascular ocular tissues. It is a mixture of electrolytes, organic solutes, growth factors, cytokines and proteins (Pietrowska et al., 2018a) Pseudoexfoliation syndrome (XFS) can be found on a routine ophthalmic examination. It presents as whitish deposits on the anterior capsule of the crystalline lens and pupil margin. It can also be found throughout the anterior segment of the eye including trabecular meshwork, Schlemm canal, zonules, and ciliary body (Ritch and Schlötzer-Schrehardt, 2001). About 60–70 million people are affected by XFS worldwide. The prevalence of XFS varies across populations, with more frequent occurrence in people of Scandinavian descent. XFS is an age-related progressive condition (Aboobakar et al., 2017). It is considered a major risk factor for accelerated cataract formation, lens subluxation, cataract surgery complications, and the development of glaucoma, which untreated or inadequately treated may lead to blindness (Vazquez-Ferreiro et al., 2016; Vazquez-Ferreiro et al., 2017). Moreover, XFS is a component of a systemic disorder, as similar deposits were found in other (non-ocular) organs, e.g. skin, heart, lungs, liver and kidneys. XFS has been associated with cardiovascular and cerebrovascular morbidity, dementia, sensorineural hearing loss and pelvic organ prolapse (Aboobakar et al., 2017; Aviv et al., 2017; Scharfenberg et al., 2019). It is unclear whether XFS is an ocular condition with systemic implications or systemic disease with ocular manifestation (Aviv et al., 2017). The pathogenesis of XFS is multifactorial and has not been fully elucidated yet. Systemic, environmental and genetic factors have been implicated (Vazquez and Lee, 2014).

Metabolomics provides information about the current biochemical status of a given biological material. Metabolomics analysis of the AH enabled to detect hundreds of molecules and dysregulated metabolic pathways. As metabolomics reflects the phenotype more accurately than many other omics technologies, the results can be more easily translatable to clinical practice (Kell et al., 2005). This is the first study on the metabolomics of AH in XFS. To the best of our knowledge, the metabolomics of AH was the subject of only one published study involving patients with pseudoexfoliation glaucoma (Myer et al., 2020). Moreover, some authors analyzed the proteomics of AH in pseudoexfoliation glaucoma (Kliuchnikova et al., 2016; Sharma et al., 2018; Botling Taube et al., 2019) and plasma metabolomics in XFS (Leruez et al., 2018). The aim of this study was to compare the composition of AH in XFS and non-XFS patients undergoing cataract surgery. Identification of metabolites distinctive for XFS might provide valuable information about the molecular background of this ocular disorder. There is a need for biomarkers to accurately assess the risk of glaucoma, its progression rate and response to treatment in patients with XFS. Identification of such biomarkers would allow for tailored follow-up and treatment of this condition. This is a vitally important objective, given that nearly half of XFS patients will eventually develop pseudoexfoliation glaucoma that if inadequately treated, is a potentially blinding disease (Ritch, 2008).

This study identified novel metabolites related to XFS, as well as metabolic pathways that are disturbed during the course of this condition. These findings contribute to a better understanding of XFS pathophysiology and may help to identify potential novel therapeutic targets.

MATERIALS AND METHODS

Study Participants and Sample Collection

The study included AH samples from 34 patients undergoing cataract surgery. The patients were divided into two groups: with XFS and without (controls). The presence of XFS was assessed on slit-lamp examination. The groups were sex-, age- and BMI-matched. The XFS group included 15 patients (10 women, mean age ±SD = 80.5 ± 5.7 years, mean BMI ±SD = 26.8 ± 3.2 kg/m^2), and the control group was comprised of 19 patients (11 women, mean age ±SD = 80.1 ± 4.1 years, mean BMI ±SD = 26.4 ± 4.2 kg/m^2). There were no major differences in systemic comorbidities or medications used (**Supplementary Tables S1 and S2**). The presence of concomitant ocular disorders and/or diabetes mellitus was an exclusion criterion from the study.

Before the cataract extraction, the anterior chamber of the eye was punctured using a 30 G needle; approximately 50–100 µl of AH was aspirated, transferred to Eppendorf tubes (Eppendorf, Hamburg, Germany), frozen and stored at −80°C until the analysis.

The protocol of the study was reviewed and approved by the Medical Ethics Committee of the Medical University of Bialystok (decisions no. R-I-002/154/2014 and R-I-002/140/2018) and conformed with the provisions of the Declaration of Helsinki.

The patients provided their written informed consent to participate in this study.

Chemicals

Purified water was obtained using the Milli-Q Integral 3 system (Millipore SAS, Molsheim, France). Zomepirac sodium salt (used as an internal standard, IS), L-serine, L-arginine hydrochloride monohydrate, ascorbic acid, L-homoarginine hydrochloride, LC-MS grade methanol acetonitrile, formic acid and LC grade ethanol were purchased from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). Pure p. a. ammonium solution (25%) was purchased from Avantor Performance Materials (Gliwice, Poland). The API-TOF reference mass solution kit (G1969–850001) and tuning solutions, ESI-L low concentration tuning mix (G1969–85000) and ESI-TOF Biopolymer Analysis reference masses (G1969–850003) were purchased from Agilent Technologies (Santa Clara, California, United States).

Sample Treatment

AH samples were treated as described elsewhere (Pietrowska et al., 2017; Pietrowska et al., 2018b). Briefly, protein precipitation and metabolite extraction were performed by 1-min vortex-mixing of the equal volumes of the AH sample and freeze cold (−20°C) methanol/ethanol (1:1) mixture containing 1 ppm of Zomepirac. Following the extraction, the samples were stored on ice for 10 min and centrifuged at 21,000 × *g* for 20 min at 4°C. The supernatant was filtered through a 0.22 μm nylon filter. Blank extraction followed the same protocol but using the freeze cold (−20°C) methanol/ethanol (1:1) mixture solely.

Aqueous Humor Metabolic Fingerprinting

The extracted samples were analyzed using an LC-MS system consisting of 1290 Infinity UHPLC (Agilent, Santa Clara, California, United States) with a degasser, two binary pumps and a thermostated autosampler coupled to a 6550 Q-TOF-MS detector (Agilent, Santa Clara, California, United States). The analyses were carried out in a positive (+) and negative (−) ion mode. The samples were analyzed using two different types of chromatography, hydrophilic interaction liquid chromatography (HILIC) for polar compounds and reversed-phase liquid chromatography (RP) for less polar and non-polar compounds. The samples were analyzed against a quality control (QC) sample prepared by mixing several AH samples. The mixture was prepared using some of the samples included in this project and additional samples, as the volume obtained by mixing spare samples from this project was not sufficient. The QC sample was prepared according to the same protocol as the other samples; it was injected ten times at the beginning of the sequence to equilibrate the LC column and later injected again at intervals (every 3–4 samples) to control the stability of the LC-MS system. The samples were analyzed using our standard AH fingerprinting methods (Pietrowska et al., 2017; Pietrowska et al., 2018b). Detailed LC-MS parameters are listed in the Supplementary Material.

Liquid Chromatography-Mass Spectrometry Data Treatment

The raw data collected by the analytical instrumentation were cleaned of background noise and unrelated ions with the Molecular Feature Extraction (MFE) tool of the Mass Hunter Qualitative Analysis Software B.06.00 (Agilent, Santa Clara, California, United States). The MFE algorithm uses the accuracy of the mass measurements to group ions related by charge-state envelope, isotopic distribution and/or the presence of adducts and dimers. The MFE then creates a list of all possible compounds described by mass, retention time (RT) and abundance. The limit for the background noise for data extraction by the MFE was individually selected for each type of chromatography and each ion mode. The values of 1500, 800, 1200, and 1000 were used for HILIC (+), HILIC (−), RP (+), and RP (−), respectively. The following adduct settings: +H, +Na, +K for positive ion mode and −H, +HCOO, +Cl for negative ion mode were applied to identify the co-eluting adducts of the same feature. Dehydration neutral losses were also allowed. Additionally, +NH4 was included in the list of possible adducts for data recorded in HILIC ESI + mode. Only metabolic features with a quality score ≥80% were accepted to preserve a good quality of the data. The sample alignment, as well as data cleaning and filtering, were performed using Mass Profiler Professional 12.6.1 (Agilent, Santa Clara, California, United States). The parameters applied for the alignment were 1% for RT and 15 ppm for the mass variation. During the first step of the data treatment, the signals present in the blank sample were separated from the signals present in the biological samples by the use of the Venn diagram. Before the statistical analysis, a quality assurance (QA) protocol was implemented to keep solely the repetitively measured metabolic features. As the QC sample was prepared by mixing additional AH samples, not only metabolic features with CV <30% but also those absent in QC samples were accepted. Additionally, the features were filtered to keep only those present in at least 80% of the samples in at least one of the studied groups. Missing values were replaced as described by Armitage et al. (2015).

Statistical Analysis

Principal component analysis (PCA) was used to check the quality of the data (clustering of the QC samples) and to detect potential outliers. A multivariate statistical analysis based on the orthogonal partial least squares discriminant analysis (OPLS-DA) models was carried out to identify the metabolites that discriminated the XFS group from the controls. The validity of the models was evaluated based on the results of a permutation test and *p*-value provided by cross-validated analysis of variance (CV-ANOVA). Additionally, to assess the predictive accuracy of the OPLS-DA models, for each model receiver operating characteristic (ROC) analysis was performed. The contribution of each metabolite to the observed sample discrimination was assessed based on the volcano plots obtained by plotting variable importance in the projection (VIP) against loading values scaled as correlation coefficient values [*p*(corr)] generated based on the obtained OPLS-DA models. Variables with VIP >1.0 and absolute *p*(corr) >0.4 were

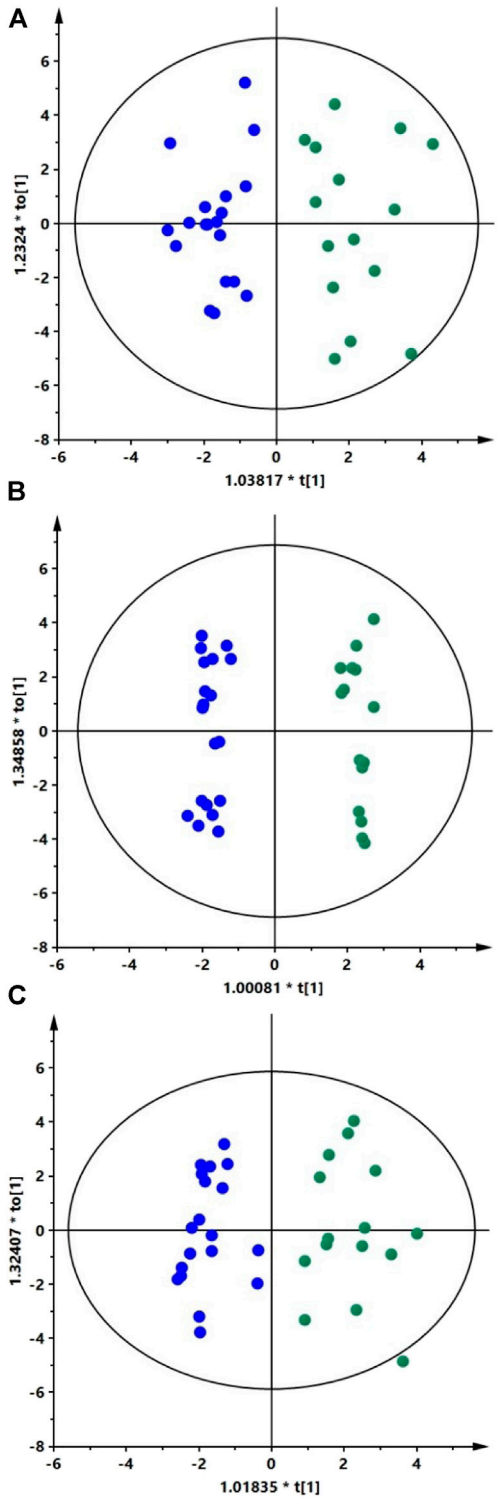

FIGURE 1 | Classification of XFS patients and controls based on AH metabolic fingerprinting data.

considered significant. Multivariate calculations and plots were obtained with SIMCA−P + 13.0.3.0 (Umetrics, Umeå, Sweden) or with SIMCA 17 (Sartorius Stedim Data Analytics AB, Gottingen, Germany). Additionally, for each metabolite *p*-value was calculated using t-test or Mann-Whitney nonparametric U test, depending on the normality of the data distribution (assessed by the Shapiro-Wilk test). Obtained *p*-values were corrected by Benjamini-Hochberg false discovery rate (FDR).

Metabolite Identification

The metabolites were identified based on the MS/MS fragmentation, as described previously (Pietrowska et al., 2017). Accurate masses of features were searched against the METLIN, KEGG, LIPIDMAPS, and HMDB databases, accessed simultaneously by CEU Mass Mediator (an on-line search engine, http://ceumass.eps.uspceu.es/mediator/). Putative identities were then confirmed by matching the experimental MS/MS spectra with the MS/MS spectra from the databases or with the fragmentation spectra and retention times obtained for the metabolite's standard. The experiments were repeated under identical chromatographic conditions as the primary analysis. Ions were targeted for collision-induced dissociation (CID) fragmentation on the fly based on the previously determined accurate mass and retention time. The identity of carnitines was confirmed based on the already described fragmentation pattern (Piszcz et al., 2016).

Metabolic Pathway Analysis

The pathway analysis was performed with MetaboAnalyst 4.0 (http://www.metaboanalyst.ca/). This on-line tool analyses the impact of particular compounds on biochemical pathways specifically for metabolomics studies (Chong et al., 2018).

RESULTS

We analyzed AH samples from cataract patients with XFS (n = 15) and without (n = 19). The metabolites in extracted AH samples were separated by two types of liquid chromatography (RP and HILIC) and then detected with a QTOF mass analyzer. Four datasets were obtained with 296, 120, 211, and 150 metabolic features for HILIC (+), HILIC (−), RP (+) and RP (−), respectively. PCA models were obtained to verify the quality of the obtained data. Clustering of the QC samples (**Supplementary Figure S1**) indicated the proper quality of the data. OPLS-DA models were used to identify statistically significant metabolites differentiating XFS patients from the controls. The models were obtained for each dataset (**Figure 1**). The study groups could not be differentiated based on the negative ion mode data obtained using RP chromatography.

The OPLS-DA scatter plots show a clear distinction between patients with XFS (green dots) and the controls (blue dots). Panels A, B, and C show models obtained for the HILIC (+) data (R2 = 0.635, Q2 = 0.345), HILIC (−) data (R2 = 0.559, Q2 = 0.471) and RP (+) data (R2 = 0.624, Q2 = 0.417), respectively. The results of permutation tests and CV-ANOVA showed that models were statistically valid. For each model obtained Q2 intercept values and *p*-values were as follow: HILIC (+) data (Q2 = −0.342, p = 0.01), HILIC (−) data (Q2 = −0.616,

TABLE 1 | Metabolites differentiating significantly AH of patients with XFS from AH of non-XFS controls.

Name	Change (%)	*p*(corr)	VIP	*p*-value	Corrected *p*-value	Monoisotopic mass [Da]
L-serine*	−30.76	−0.44	1.01	0.02	0.07	105.0426
3-hydroxy anthranilic acid	−33.10	−0.50	2.25	0.05	0.3	153.0426
	−38.22	−0.40	2.12	0.002	0.2	153.0426
Indoleacetaldehyde	+96.36	0.64	2.64	0.006	0.2	159.0684
2-hydroxycinnamic acid/m-coumaric acid (co-elution)	−25.24	−0.51	1.23	0.01	0.2	164.0473
L-arginine*	−30.68	−0.59	2.38	0.04	0.3	174.1117
Ascorbic acid*	−32.81	−0.50	2.11	0.02	0.3	176.0321
Homo-L-arginine*	−18.72	−0.41	1.38	0.01	0.2	188.1273
Ergothioneine	−66.42	−0.41	1.81	0.05	0.3	229.0884
	−64.76	−0.44	2.48	0.03	0.3	229.0885
Hydroxybutyrylcarnitine	−28.86	−0.46	1.24	0.02	0.2	247.142
Decatrienoylcarnitine	−45.83	−0.45	1.89	0.002	0.09	309.194
S-adenosyl-L-methioninate	−29.15	−0.59	1.93	0.0004	0.04	398.1372

The direction of change indicates increased (+) or decreased (−) abundance of a metabolite in the AH of patients with XFS in comparison to its abundance in the AH of the controls. The VIP values were calculated based on OPLS-DA models built separately for each method and ion mode.
**The identity of these metabolites was confirmed by the LC-MS/MS analysis of the standard.*

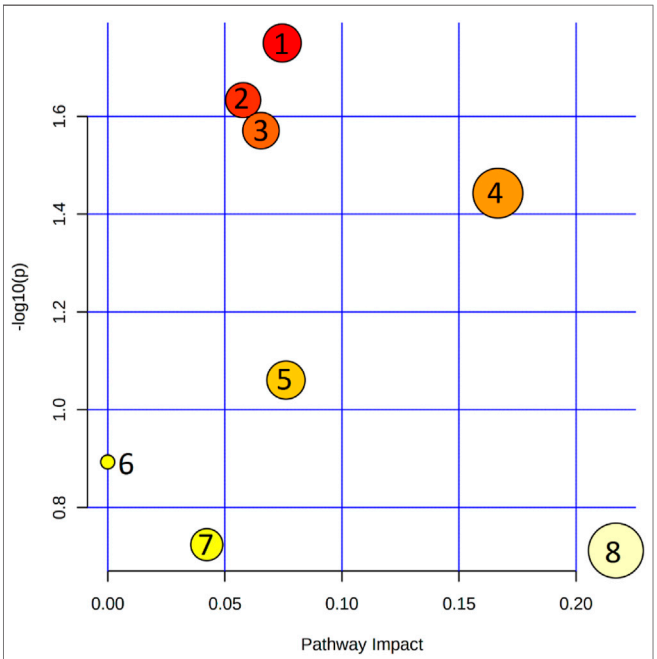

FIGURE 2 | Pathway analysis of metabolites that were shown to differ significantly between the metabolic profiles of aqueous humor in patients with XFS and without. 1. Cysteine and methionine metabolism; 2. Arginine and proline metabolism; 3. Tryptophan metabolism; 4. Aminoacyl-tRNA biosynthesis; 5. Arginine biosynthesis; 6. Sphingolipid metabolism; 7. Glyoxylate and dicarboxylate metabolism; 8. Glycine, serine and threonine metabolism.

p = 0.0008), and RP (+) data (Q2 = −0.345, p = 0.008). AUC values obtained for all three models based on ROC analysis were equal to 1.0, indicating perfect predictive accuracy.

Eleven metabolites (**Table 1**; **Supplementary Table S3**) were shown to discriminate significantly patients with XFS and the controls. Compared with the controls, AH of patients with XFS contained significantly lower levels of amino acids, organic acids and acylcarnitines. Indoleacetaldehyde was the only metabolite, the level of which was significantly higher in XFS patients than in the controls. Biochemical pathway analysis involving the significant metabolites mentioned above identified several altered pathways in the AH of XFS patients, especially the metabolism of amino acids and aminoacyl-tRNA biosynthesis (**Figure 2**). Among other alterations, the disturbances of tryptophan metabolism were found in XFS patients, with a resultant decrease in 3-hydroxyanthranilic acid level and an increase in indoleacetaldehyde level.

DISCUSSION

The aim of this study was to identify potential differences in the metabolic composition of AH obtained from patients with and without XFS. We focused on XFS, rather than on its consequence, pseudoexfoliation glaucoma, to characterize an initial pathological stage of the latter condition. Our findings imply that XFS is associated with enhanced oxidative stress and inflammation, as well as with the disruption of cellular respiration and mitochondrial energy production. Gut microbiota might also play a role in the pathogenesis of this condition. These observations are consistent with the results of previous studies (Koliakos et al., 2008; Erdurmuş et al., 2011; Zenkel et al., 2011; Aboobakar et al., 2017; Borrás, 2018; Papadopoulou et al., 2018; Schlötzer-Schrehardt, 2018; Sharma et al., 2018; Botling Taube et al., 2019; Ghaffari Sharaf et al., 2020).

To the best of our knowledge, only one metabolomics study of AH from patients with pseudoexfoliation glaucoma, not XFS, has been conducted thus far (Myer et al., 2020). Similarly to Myer et al. (2020), we have found decreased levels of L-arginine. Interestingly, according to Leruez et al. neither arginine nor tryptophan proved to be significant in the plasma of XFS (Leruez et al., 2018), which implies that the metabolism of these amino acids in XFS is affected locally, in the anterior chamber of the eye, rather than systemically. L-arginine is utilized by NO synthases for the synthesis of NO (Stuehr, 2004). Another substrate for NO synthase is Homo-L-arginine,

which in this study was also found decreased in AH of patients with XFS (Tsikas and Wu, 2015). Thus, the reduction of arginine and homoarginine levels in patients with XFS points to a dysfunction of the NO pathway, which plays a significant role in regulating AH outflow balance in the eye (Kotikoski et al., 2002; Chang et al., 2015) and seems to be involved in the pathogenesis of glaucoma (Polak et al., 2007). Lower levels of NO in AH of XFS patients with glaucoma have already been reported by Kotikoski et al. (2002). L-arginine has also antioxidant and anti-inflammatory properties (Jobgen et al., 2006; Thomas et al., 2008; Tosun et al., 2012; Pimentel et al., 2013; Wang et al., 2014; Erdinest et al., 2015; Park et al., 2020).

As an important one-carbon donor to the folate cycle, serine contributes to nucleotide synthesis, methylation reactions and the antioxidant defense (Bleich et al., 2004). Folate deficiency is a risk factor for XFS (Aboobakar et al., 2017). A decreased level of serine might be associated with a reduced concentration of S-adenosyl-methionine. S-adenosyl-methionine has anti-inflammatory activity and is an intermediate in the homocysteine biosynthesis pathway. Its decreased level in patients with XFS might reflect the increased activity of the methionine cycle and excessive production of homocysteine. Although homocysteine was not identified as a significant metabolite in the present study, it has been well-documented in both plasma and AH of patients with XFS. It is unclear whether hyperhomocysteinemia is a cause or consequence of XFS (Bleich et al., 2004; Aboobakar et al., 2017; Leruez et al., 2018). According to Rebecca et al. hyperhomocysteinemia in XFS causes structural changes promoting elastin aggregation (Rebecca et al., 2019). Serine metabolism also intersects with the methionine cycle (Yang and Vousden, 2016), in which methionine synthase remethylates homocysteine in a vitamin B12-dependent reaction (Yang and Vousden, 2016; Koc and Kaya, 2020). It is suggested that the intake of vitamin B6 and vitamin B12 may reduce the risk of pseudoexfoliation glaucoma (Kang et al., 2014). This observation might support the theory about the increased activity of the methionine pathway and a deficit of serine and S-adenosyl-methionine in patients with XFS.

In line with Koliakos et al. (2003) we have found decreased level of ascorbic acid, a potent antioxidant. It implies that XFS is associated with enhanced oxidative stress within the eye. Furthermore, vitamin C supplementation is believed to protect against the XFS progression (Koliakos et al., 2003). Due to the reduced concentration of ascorbic acid in AH, patients with XFS might be protected less against UV radiation. This hypothesis is supported by the observation that prolonged time spent outdoors is an independent risk factor for the development of XFS (Kang et al., 2012; Pasquale et al., 2014). The role of vitamin C in cataract prevention was described by Weikel et al. (2014). Aside from the reduced levels of ascorbic acid, we also found decreased levels of 3-hydroxyanthranilic acid in the AH of XFS patients. The latter is one of the tryptophan-derived compounds that act as physical filters for bands from the UVA spectrum (Wilson et al., 2016). UV radiation is implicated in the photodamage to the human eye. Decreased levels of tryptophan-derived constituents impair protection from UV light, lead to enhanced oxidative damage and accumulation of modified proteins implicated in nuclear cataract formation (Tweeddale et al., 2016). Moreover, Laganovska et al. suggested that XFS might be associated with disturbances in the kynurenine pathway, the primary route for tryptophan catabolism (Laganovska et al., 2003). This might lead to the enhancement of oxidative stress in the eye, a key factor in the pathogenesis of ocular diseases (Kruk et al., 2015). Chronic oxidative stress may disrupt the balance between matrix metalloproteinases and their tissue inhibitors, which leads to the accumulation of extracellular matrix fibrils constituting pseudoexfoliation material (Zenkel et al., 2011; Schlötzer-Schrehardt, 2012). Enhanced oxidative stress in the eye was also reported by other authors, along with mitochondrial dysfunction (Koliakos et al., 2008; Erdurmuş et al., 2011; Zenkel et al., 2011; Borrás, 2018; Papadopoulou et al., 2018; Schlötzer-Schrehardt, 2018; Sharma et al., 2018; Botling Taube et al., 2019; Ghaffari Sharaf et al., 2020). The latter might play a role in XFS progression. Non-functioning mitochondria were observed *in vitro* models of XFS; the mitochondrial dysfunction might negatively affect the efficiency of respiration, leading to a decrease in ATP levels and enhanced synthesis of reactive oxygen species (Want et al., 2016). Acylcarnitines are involved in the mitochondrial metabolism of lipids and fatty acids and are essential for the proper function of the eye (Pescosolido et al., 2008; Pescosolido et al., 2009). Carnitines may protect against selenite-induced cataract, and their loss may be a marker of disease development (Ritch, 2007). As postulated recently, the reduced level of carnitines in XFS might reflect mitochondrial dysfunction (Want et al., 2016; Ghaffari Sharaf et al., 2020). Our findings seem to support this theory, as the levels of hydroxybutyrylcarnitine and decatrienoylcarnitine in XFS patients were significantly lower than in the controls. The deficit of carnitines in the AH might reflect enhanced β-oxidation and elevated levels of mitochondrial acetyl-CoA, conditions that eventually lead to the development of cellular oxidative stress (Schönfeld et al., 2010). In summary, we have demonstrated that the AH of XFS patients shows a decreased antioxidant content and increased oxidative stress factors.

Some of the metabolites identified in this study (e.g. indoleacetaldehyde, 2-hydroxycinnamic acid/m-coumaric acid, ergothioneine) might be related to the human microbiome, especially the intestinal microflora (Filannino et al., 2014; Cumming et al., 2018; Gao et al., 2018). Several authors suggested a link between gut microbiota and eye diseases (Lin, 2018; Shivaji, 2019; Petrillo et al., 2020). AH from patients with XFS contained significantly higher levels of indoleacetaldehyde, a precursor in the bacterial synthesis of indoleacetic acid from tryptophan (Gao et al., 2018; Roager and Licht, 2018). An elevated level of indoleacetic acid in mammalian cells may cause many disorders, such as disruption of apoptosis, protein degradation and cell cycle progression arrest (Holland et al., 2012; Zhao et al., 2015). While the function of ergothioneine in mammalian cells is still not fully understood, recent evidence suggests that this metabolite is a powerful antioxidant; cells deficient in ergothioneine were shown to be more prone to oxidative stress, with resultant increased damage to mitochondrial DNA,

protein oxidation and peroxidation of lipids (Paul and Snyder, 2010; Cheah and Halliwell, 2012). Hence, the deficit of ergothioneine in the AH of patients with XFS seems to support the notion that this condition is associated with enhanced oxidative stress in the eye.

This study had several potential limitations. The sample size was relatively small and from a limited geographical area. Hence, it is unclear whether the results could be generalized to other ethnic groups, especially considering the uneven distribution of XFS globally. Moreover, the study focused solely on XFS rather than on pseudoexfoliation glaucoma. Finally, not all metabolites could be analyzed with the LC-MS approach.

An interesting direction of future research would be a comparative analysis of the AH and plasma metabolomics in the same patients, especially given that XFS is a systemic condition associated with increased vascular permeability of the blood-aqueous barrier (Scharfenberg et al., 2019). From a clinical perspective, the results of this study might justify further research on the anti-oxidative and anti-inflammatory treatment of XFS.

CONCLUSION

The results of this study suggest that the pathogenesis of XFS may involve enhanced oxidative stress and inflammation, as well as the dysfunction of mitochondria and gut microbiota. The knowledge of metabolites and metabolic pathways involved in XFS pathogenesis might facilitate the development of novel prevention and treatment strategies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion 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 The Medical Ethics Committee of the Medical University of Bialystok (decisions no. R-I-002/154/2014 and R-I-002/140/2018). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conceptualization: DD, KP, PK, ZM, and MC; methodology and software: DD, KP, PK, TK, MM, and MC; formal analysis: KP, TK, MM, and EG; investigation: DD, KP, PK, TK, EG, and MC; resources: ZM, AK, and MC; data curation: KP, MM, and MC; writing-original draft preparation: DD, KP, TK, and EG; writing-review and editing: ZM, AK, and MC; visualization: KP, MM, and MC; supervision: ZM, AK, and MC. All authors have read and approved the final manuscript.

FUNDING

This work was supported by the Medical University of Bialystok, Poland (Grant No. N/ST/ZB/16/001/1157). This study was conducted with the use of equipment purchased by the Medical University of Bialystok as a part of the RPOWP 2007–2013 Funding, Priority I, Axis 1.1, Contract no. UDA-RPPD.01.01.00-20-001/15-00 dated 26.06.2015.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.682600/full#supplementary-material

REFERENCES

Aboobakar, I. F., Johnson, W. M., Stamer, W. D., Hauser, M. A., and Allingham, R. R. (2017). Major Review: Exfoliation Syndrome; Advances in Disease Genetics, Molecular Biology, and Epidemiology. *Exp. Eye Res.* 154, 88–103. doi:10.1016/j.exer.2016.11.011

Armitage, E. G., Godzien, J., Alonso-Herranz, V., López-Gonzálvez, Á., and Barbas, C. (2015). Missing Value Imputation Strategies for Metabolomics Data. *ELECTROPHORESIS* 36 (24), 3050–3060. doi:10.1002/elps.201500352

Aviv, U., Ben Ner, D., Sharif, N., Gur, Z., and Achiron, A. (2017). Pseudoexfoliation: An Ocular Finding with Possible Systemic Implications. *Isr. Med. Assoc. J.* 19 (1), 49– 54.

Bleich, S., Roedl, J., Von Ahsen, N., Schlötzer-Schrehardt, U., Reulbach, U., Beck, G., et al. (2004). Elevated Homocysteine Levels in Aqueous Humor of Patients with Pseudoexfoliation Glaucoma. *Am. J. Ophthalmol.* 138 (1), 162–164. doi:10.1016/j.ajo.2004.02.027

Borrás, T. (2018). Growth Factors, Oxidative Damage, and Inflammation in Exfoliation Syndrome. *J. Glaucoma* 27 (Suppl. 1), S54–S60. doi:10.1097/IJG.0000000000000904

Botling Taube, A., Konzer, A., Alm, A., and Bergquist, J. (2019). Proteomic Analysis of the Aqueous Humour in Eyes with Pseudoexfoliation Syndrome. *Br. J. Ophthalmol.* 103 (8), 1190–1194. doi:10.1136/bjophthalmol-2017-310416

Chang, J. Y. H., Stamer, W. D., Bertrand, J., Read, A. T., Marando, C. M., Ethier, C. R., et al. (2015). Role of Nitric Oxide in Murine Conventional Outflow Physiology. *Am. J. Physiology-Cell Physiol.* 309 (4), C205–C214. doi:10.1152/ajpcell.00347.2014

Cheah, I. K., and Halliwell, B. (2012). Ergothioneine; Antioxidant Potential, Physiological Function and Role in Disease. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1822 (5), 784–793. doi:10.1016/j.bbadis.2011.09.017

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 (W1), W486–W494. doi:10.1093/nar/gky310

Cumming, B. M., Chinta, K. C., Reddy, V. P., and Steyn, A. J. C. (2018). Role of Ergothioneine in Microbial Physiology and Pathogenesis. *Antioxid. Redox Signaling* 28 (6), 431–444. doi:10.1089/ars.2017.7300

Erdinest, N., Shohat, N., Moallem, E., Yahalom, C., Mechoulam, H., Anteby, I., et al. (2015). Nitric Oxide Secretion in Human Conjunctival Fibroblasts Is Inhibited by Alpha Linolenic Acid. *J. Inflamm.* 12, 59. doi:10.1186/s12950-015-0104-1

Erdurmuş, M., Yağcı, R., Atış, Ö., Karadağ, R., Akbaş, A., and Hepşen, I. F. (2011). Antioxidant Status and Oxidative Stress in Primary Open Angle Glaucoma and Pseudoexfoliative Glaucoma. *Curr. Eye Res.* 36 (8), 713–718. doi:10.3109/02713683.2011.584370

Filannino, P., Gobbetti, M., De Angelis, M., and Di Cagno, R. (2014). Hydroxycinnamic Acids Used as External Acceptors of Electrons: an

Energetic Advantage for Strictly Heterofermentative Lactic Acid Bacteria. *Appl. Environ. Microbiol.* 80 (24), 7574–7582. doi:10.1128/AEM.02413-14

Gao, J., Xu, K., Liu, H., Liu, G., Bai, M., Peng, C., et al. (2018). Impact of the Gut Microbiota on Intestinal Immunity Mediated by Tryptophan Metabolism. *Front. Cel. Infect. Microbiol.* 8, 13. doi:10.3389/fcimb.2018.00013

Ghaffari Sharaf, M., Damji, K. F., and Unsworth, L. D. (2020). Recent Advances in Risk Factors Associated with Ocular Exfoliation Syndrome. *Acta Ophthalmol.* 98 (2), 113–120. doi:10.1111/aos.14298

Holland, A. J., Fachinetti, D., Han, J. S., and Cleveland, D. W. (2012). Inducible, Reversible System for the Rapid and Complete Degradation of Proteins in Mammalian Cells. *Proc. Natl. Acad. Sci.* 109 (49), E3350–E3357. doi:10.1073/pnas.1216880109

Jobgen, W. S., Fried, S. K., Fu, W. J., Meininger, C. J., and Wu, G. (2006). Regulatory Role for the Arginine-Nitric Oxide Pathway in Metabolism of Energy Substrates. *J. Nutr. Biochem.* 17 (9), 571–588. doi:10.1016/j.jnutbio.2005.12.001

Kang, J. H., Loomis, S. J., Wiggs, J. L., Willett, W. C., and Pasquale, L. R. (2014). A Prospective Study of Folate, Vitamin B6, and Vitamin B12Intake in Relation to Exfoliation Glaucoma or Suspected Exfoliation Glaucoma. *JAMA Ophthalmol.* 132 (5), 549–559. doi:10.1001/jamaophthalmol.2014.100

Kang, J. H., Loomis, S., Wiggs, J. L., Stein, J. D., and Pasquale, L. R. (2012). Demographic and Geographic Features of Exfoliation Glaucoma in 2 United States-based Prospective Cohorts. *Ophthalmology* 119 (1), 27–35. doi:10.1016/j.ophtha.2011.06.018

Kell, D. B., Brown, M., Davey, H. M., Dunn, W. B., Spasic, I., and Oliver, S. G. (2005). Metabolic Footprinting and Systems Biology: the Medium Is the Message. *Nat. Rev. Microbiol.* 3 (7), 557–565. doi:10.1038/nrmicro1177

Kivelä, T. T. (2018). Histopathology of Exfoliation Syndrome. *J. Glaucoma* 27 (Suppl. 1), S38–S43. doi:10.1097/IJG.0000000000000947

Kliuchnikova, A. A., Samokhina, N. I., Ilina, I. Y., Karpov, D. S., Pyatnitskiy, M. A., Kuznetsova, K. G., et al. (2016). Human Aqueous Humor Proteome in Cataract, Glaucoma, and Pseudoexfoliation Syndrome. *Proteomics* 16 (13), 1938–1946. doi:10.1002/pmic.201500423

Koc, H., and Kaya, F. (2020). Relationship between Homocysteine Levels, Anterior Chamber Depth, and Pseudoexfoliation Glaucoma in Patients with Pseudoexfoliation. *Int. Ophthalmol.* 40 (7), 1731–1737. doi:10.1007/s10792-020-01341-4

Koliakos, G. G., Befani, C. D., Mikropoulos, D., Ziakas, N. G., and Konstas, A. G. P. (2008). Prooxidant-antioxidant Balance, Peroxide and Catalase Activity in the Aqueous Humour and Serum of Patients with Exfoliation Syndrome or Exfoliative Glaucoma. *Graefes Arch. Clin. Exp. Ophthalmol.* 246 (10), 1477–1483. doi:10.1007/s00417-008-0871-y

Koliakos, G. G., Konstas, A. G., Schlötzer-Schrehardt, U., Hollo, G., Katsimbris, I. E., Georgiadis, N., et al. (2003). 8-Isoprostaglandin F2a and Ascorbic Acid Concentration in the Aqueous Humour of Patients with Exfoliation Syndrome. *Br. J. Ophthalmol.* 87 (3), 353–356. doi:10.1136/bjo.87.3.353

Kotikoski, H., Moilanen, E., Vapaatalo, H., and Aine, E. (2002). Biochemical Markers of the L-Arginine-Nitric Oxide Pathway in the Aqueous Humour in Glaucoma Patients. *Acta Ophthalmol. Scand.* 80 (2), 191–195. doi:10.1034/j.1600-0420.2002.800214.x

Kruk, J., Kubasik-Kladna, K., and Y. Aboul-Enein, H. (2015). The Role Oxidative Stress in the Pathogenesis of Eye Diseases: Currnt Status and a Dual Role of Physical Activity. *Mrmc* 16 (3), 241–257. doi:10.2174/1389557516666151120114605

Laganovska, G., Martinsons, A., Pitrans, B., Widner, B., and Fuchs, D. (2003). Kynurenine and Neopterin in the Aqueous Humor of the Anterior Chamber of the Eye and in Serum of Cataract Patients. *Adv. Exp. Med. Biol.* 527, 367–374. doi:10.1007/978-1-4615-0135-0_42

Leruez, S., Bresson, T., Chao de la Barca, J. M., Marill, A., de Saint Martin, G., Buisset, A., et al. (2018). A Plasma Metabolomic Signature of the Exfoliation Syndrome Involves Amino Acids, Acylcarnitines, and Polyamines. *Invest. Ophthalmol. Vis. Sci.* 59 (2), 1025–1032. doi:10.1167/iovs.17-23055

Lin, P. (2018). The Role of the Intestinal Microbiome in Ocular Inflammatory Disease. *Curr. Opin. Ophthalmol.* 29 (3), 261–266. doi:10.1097/ICU.0000000000000465

Myer, C., Abdelrahman, L., Banerjee, S., Khattri, R. B., Merritt, M. E., Junk, A. K., et al. (2020). Aqueous Humor Metabolite Profile of Pseudoexfoliation Glaucoma Is Distinctive. *Mol. Omics* 16, 425–435. doi:10.1039/c9mo00192a

Papadopoulou, G., Zisimopoulos, D., Kalaitzopoulou, E., Makri, O. E., Tsapardoni, F. N., Georgakopoulos, C. D., et al. (2018). Age-related Aqueous Humor (AH) and Lens Epithelial Cell/capsule Protein Carbonylation and AH Protein Concentration in Cataract Patients Who Have Pseudoexfoliative Diseases. *Mol. Vis.* 24, 890–901.

Park, J. W., Piknova, B., Jenkins, A., Hellinga, D., Parver, L. M., and Schechter, A. N. (2020). Potential Roles of Nitrate and Nitrite in Nitric Oxide Metabolism in the Eye. *Sci. Rep.* 10 (1), 13166. doi:10.1038/s41598-020-69272-9

Pasquale, L. R., Jiwani, A. Z., Zehavi-Dorin, T., Majd, A., Rhee, D. J., Chen, T., et al. (2014). Solar Exposure and Residential Geographic History in Relation to Exfoliation Syndrome in the United States and Israel. *JAMA Ophthalmol.* 132 (12), 1439–1445. doi:10.1001/jamaophthalmol.2014.3326

Paul, B. D., and Snyder, S. H. (2010). The Unusual Amino Acid L-Ergothioneine Is a Physiologic Cytoprotectant. *Cell Death Differ* 17 (7), 1134–1140. doi:10.1038/cdd.2009.163

Pescosolido, N., Imperatrice, B., and Karavitis, P. (2008). The Aging Eye and the Role of L-Carnitine and its Derivatives. *Drugs R. D* 9 (Suppl. 1), 3–14. doi:10.2165/0126839-200809001-00002

Pescosolido, N., Imperatrice, B., Koverech, A., and Messano, M. (2009). L-carnitine and Short Chain Ester in Tears from Patients with Dry Eye. *Optom. Vis. Sci.* 86 (2), E132–E138. doi:10.1097/OPX.0b013e318194e767

Petrillo, F., Pignataro, D., Lavano, M. A., Santella, B., Folliero, V., Zannella, C., et al. (2020). Current Evidence on the Ocular Surface Microbiota and Related Diseases. *Microorganisms* 8 (7), 1033. doi:10.3390/microorganisms8071033

Pietrowska, K., Dmuchowska, D. A., Krasnicki, P., Mariak, Z., Kretowski, A., and Ciborowski, M. (2018a). Analysis of Pharmaceuticals and Small Molecules in Aqueous Humor. *J. Pharm. Biomed. Anal.* 159, 23–36. doi:10.1016/j.jpba.2018.06.049

Pietrowska, K., Dmuchowska, D. A., Krasnicki, P., Bujalska, A., Samczuk, P., Parfieniuk, E., et al. (2018b). An Exploratory LC-MS-based Metabolomics Study Reveals Differences in Aqueous Humor Composition between Diabetic and Non-diabetic Patients with Cataract. *Electrophoresis* 39, 1233–1240. doi:10.1002/elps.201700411

Pietrowska, K., Dmuchowska, D. A., Samczuk, P., Kowalczyk, T., Krasnicki, P., Wojnar, M., et al. (2017). LC-MS-Based Metabolic Fingerprinting of Aqueous Humor. *J. Anal. Methods Chem.* 2017, 1–13. doi:10.1155/2017/6745932

Pimentel, A. M. L., Pereira, N. R., Costa, C. A., Mann, G. E., Cordeiro, V. S. C., de Moura, R. S., et al. (2013). L-arginine-nitric Oxide Pathway and Oxidative Stress in Plasma and Platelets of Patients with Pre-eclampsia. *Hypertens. Res.* 36 (9), 783–788. doi:10.1038/hr.2013.34

Piszcz, J., Armitage, E. G., Ferrarini, A., Rupérez, F. J., Kulczynska, A., Bolkun, L., et al. (2016). To Treat or Not to Treat: Metabolomics Reveals Biomarkers for Treatment Indication in Chronic Lymphocytic Leukaemia Patients. *Oncotarget* 7, 22324–22338. doi:10.18632/oncotarget.8078

Polak, K., Luksch, A., Berisha, F., Fuchsjaeger-Mayrl, G., Dallinger, S., and Schmetterer, L. (2007). Altered Nitric Oxide System in Patients with Open-Angle Glaucoma. *Arch. Ophthalmol.* 125 (4), 494–498. doi:10.1001/archopht.125.4.494

Rebecca, M., Gayathri, R., Bhuvanasundar, R., Sripriya, K., Shantha, B., and Angayarkanni, N. (2019). Elastin Modulation and Modification by Homocysteine: a Key Factor in the Pathogenesis of Pseudoexfoliation Syndrome? *Br. J. Ophthalmol.* 103 (7), 985–992. doi:10.1136/bjophthalmol-2018-312088

Ritch, R. (2007). Natural Compounds: Evidence for a Protective Role in Eye Disease. *J. Can. D'ophtalmol* 42 (3), 425–438. doi:10.3129/can.j.ophthalmol.i07-044

Ritch, R., and Schlötzer-Schrehardt, U. (2001). Exfoliation Syndrome. *Surv. Ophthalmol.* 45 (4), 265–315. doi:10.1016/s0039-6257(00)00196-x

Ritch, R. (2008). The Management of Exfoliative Glaucoma. *Prog. Brain Res.* 173, 211–224. doi:10.1016/S0079-6123(08)01115-1

Roager, H. M., and Licht, T. R. (2018). Microbial Tryptophan Catabolites in Health and Disease. *Nat. Commun.* 9 (1), 3294. doi:10.1038/s41467-018-05470-4

Scharfenberg, E., Rauscher, F. G., Meier, P., and Hasenclever, D. (2019). Pseudoexfoliation Syndrome: Analysis of Systemic Comorbidities of 325 PEX-Positive Patients Compared with 911 PEX-Negative Patients. *Graefes Arch. Clin. Exp. Ophthalmol.* 257 (11), 2471–2480. doi:10.1007/s00417-019-04438-4

Schlötzer-Schrehardt, U. (2018). Molecular Biology of Exfoliation Syndrome. *J. Glaucoma* 27 (Suppl. 1), S32–S37. doi:10.1097/IJG.0000000000000903

Schlötzer-Schrehardt, U. (2012). Neue Pathogenetische Erkenntnisse Zum Pseudoexfoliations-Syndrom/Glaukom. *Ophthalmologe* 109 (10), 944–951. doi:10.1007/s00347-012-2531-1

Schönfeld, P., Więckowski, 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 (Bba) - Bioenerg.* 1797 (6-7), 929–938. doi:10.1016/j.bbabio.2010.01.010

Sharma, S., Chataway, T., Klebe, S., Griggs, K., Martin, S., Chegeni, N., et al. (2018). Novel Protein Constituents of Pathological Ocular Pseudoexfoliation Syndrome Deposits Identified with Mass Spectrometry. *Mol. Vis.* 24, 801–817.

Shivaji, S. (2019). Connect between Gut Microbiome and Diseases of the Human Eye. *J. Biosci.* 44 (5). doi:10.1007/s12038-019-9931-1

Stuehr, D. J. (2004). Enzymes of the L-Arginine to Nitric Oxide Pathway. *J. Nutr.* 134 (10 Suppl. l), 2748S–2751S. doi:10.1093/jn/134.10.2748S

Thomas, D. D., Ridnour, L. A., Isenberg, J. S., Flores-Santana, W., Switzer, C. H., Donzelli, S., et al. (2008). The Chemical Biology of Nitric Oxide: Implications in Cellular Signaling. *Free Radic. Biol. Med.* 45 (1), 18–31. doi:10.1016/j.freeradbiomed.2008.03.020

Tosun, M., Erdurmus, M., Bugdayci, G., Celebi, S., and Alcelik, A. (2012). Aqueous Humour and Serum Concentration of Asymmetric Dimethyl Arginine in Pseudoexfoliation Syndrome. *Br. J. Ophthalmol.* 96 (8), 1137–1140. doi:10.1136/bjophthalmol-2012-301901

Tsikas, D., and Wu, G. (2015). Homoarginine, Arginine, and Relatives: Analysis, Metabolism, Transport, Physiology, and Pathology. *Amino Acids* 47 (9), 1697–1702. doi:10.1007/s00726-015-2055-5

Tweeddale, H. J., Hawkins, C. L., Janmie, J. F., Truscott, R. J. W., and Davies, M. J. (2016). Cross-linking of Lens Crystallin Proteins Induced by Tryptophan Metabolites and Metal Ions: Implications for Cataract Development. *Free Radic. Res.* 50 (10), 1116–1130. doi:10.1080/10715762.2016.1210802

Vazquez, L. E., and Lee, R. K. (2014). Genomic and Proteomic Pathophysiology of Pseudoexfoliation Glaucoma. *Int. Ophthalmol. Clin.* 54 (4), 1–13. doi:10.1097/IIO.0000000000000047

Vazquez-Ferreiro, P., Carrera-Hueso, F. J., Fikri-Benbrahim, N., Barreiro-Rodriguez, L., Diaz-Rey, M., and Ramón Barrios, M. A. (2017). Intraocular Lens Dislocation in Pseudoexfoliation: a Systematic Review and Meta-Analysis. *Acta Ophthalmol.* 95 (3), e164–e169. doi:10.1111/aos.13234

Vazquez-Ferreiro, P., Carrera-Hueso, F. J., Jornet, J. E. P., Fikri-Benbrahim, N., Diaz-Rey, M., and Sanjuan-Cerveró, R. (2016). Intraoperative Complications of Phacoemulsification in Pseudoexfoliation: Metaanalysis. *J. Cataract Refract Surg.* 42 (11), 1666–1675. doi:10.1016/j.jcrs.2016.09.010

Wang, Y.-T., Piyankarage, S. C., Williams, D. L., and Thatcher, G. R. J. (2014). Proteomic Profiling of Nitrosative Stress: Protein S-Oxidation Accompanies S-Nitrosylation. *ACS Chem. Biol.* 9 (3), 821–830. doi:10.1021/cb400547u

Want, A., Gillespie, S. R., Wang, Z., Gordon, R., Iomini, C., Ritch, R., et al. (2016). Autophagy and Mitochondrial Dysfunction in Tenon Fibroblasts from Exfoliation Glaucoma Patients. *PLoS One* 11 (7), e0157404. doi:10.1371/journal.pone.0157404

Weikel, K. A., Garber, C., Baburins, A., and Taylor, A. (2014). Nutritional Modulation of Cataract. *Nutr. Rev.* 72 (1), 30–47. doi:10.1111/nure.12077

Wilson, K., Auer, M., Binnie, M., Zheng, X., Pham, N. T., Iredale, J. P., et al. (2016). Overexpression of Human Kynurenine-3-Monooxygenase Protects against 3-Hydroxykynurenine-Mediated Apoptosis through Bidirectional Nonlinear Feedback. *Cell Death Dis* 7, e2197. doi:10.1038/cddis.2016.87

Yang, M., and Vousden, K. H. (2016). Serine and One-Carbon Metabolism in Cancer. *Nat. Rev. Cancer* 16 (10), 650–662. doi:10.1038/nrc.2016.81

Zenkel, M. (2018). Extracellular Matrix Regulation and Dysregulation in Exfoliation Syndrome. *J. Glaucoma* 27 (Suppl. 1), S24–S28. doi:10.1097/IJG.0000000000000902

Zenkel, M., Krysta, A., Pasutto, F., Juenemann, A., Kruse, F. E., and Schlötzer-Schrehardt, U. (2011). Regulation of Lysyl Oxidase-like 1 (LOXL1) and Elastin-Related Genes by Pathogenic Factors Associated with Pseudoexfoliation Syndrome. *Invest. Ophthalmol. Vis. Sci.* 52 (11), 8488–8495. doi:10.1167/iovs.11-8361

Zhao, L., Liu, P., Guo, G., and Wang, L. (2015). Combination of Cytokinin and Auxin Induces Apoptosis, Cell Cycle Progression Arrest and Blockage of the Akt Pathway in HeLa Cells. *Mol. Med. Rep.* 12 (1), 719–727. doi:10.3892/mmr.2015.3420

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 14 May 2021
doi: 10.3389/fmolb.2021.662792

Understanding Systemic and Local Inflammation Induced by Nasal Polyposis: Role of the Allergic Phenotype

María Isabel Delgado-Dolset[1], David Obeso[1,2], Javier Sánchez-Solares[1], Leticia Mera-Berriatua[1], Paloma Fernández[1], Coral Barbas[2], Miguel Fresnillo[3], Tomás Chivato[1,4], Domingo Barber[1], María M. Escribese[1,5]† and Alma Villaseñor[1]*†*

[1]Department of Basic Medical Sciences, Faculty of Medicine, Institute of Applied Molecular Medicine (IMMA), San Pablo CEU Universities, Madrid, Spain, [2]Centre for Metabolomics and Bioanalysis (CEMBIO), Department of Chemistry and Biochemistry, Faculty of Pharmacy, San Pablo CEU Universities, Madrid, Spain, [3]Otorhinolaringology Service, HM Montepríncipe Hospital, Madrid, Spain, [4]Department of Clinic Medical Sciences, Faculty of Medicine, San Pablo CEU Universities, Madrid, Spain, [5]Department of Basic Medical Sciences, Faculty of Medicine, San Pablo CEU Universities, Madrid, Spain

OPEN ACCESS

Edited by:
Serge Rudaz,
Université de Genève, Switzerland

Reviewed by:
Monica Cala,
University of Los Andes, Colombia
Gino Marioni,
University of Padua, Italy

***Correspondence:**
Alma Villaseñor
alma.villasenor@ceu.es
María M. Escribese
mariamarta.escribesealonso@ceu.es

†These authors share last authorship

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

Received: 01 February 2021
Accepted: 27 April 2021
Published: 14 May 2021

Citation:
Delgado-Dolset MI, Obeso D, Sánchez-Solares J, Mera-Berriatua L, Fernández P, Barbas C, Fresnillo M, Chivato T, Barber D, Escribese MM and Villaseñor A (2021) Understanding Systemic and Local Inflammation Induced by Nasal Polyposis: Role of the Allergic Phenotype. Front. Mol. Biosci. 8:662792. doi: 10.3389/fmolb.2021.662792

Chronic rhinosinusitis with nasal polyps (CRSwNP) is characterized by persistent symptoms associated to the development of nasal polyps. To this day, the molecular mechanisms involved are still not well defined. However, it has been suggested that a sustained inflammation as allergy is involved in its onset. In this exploratory study, the aim was to investigate the effect of the allergic status in the development of CRSwNP. To achieve this, we recruited 22 patients with CRSwNP and classified them in non-allergic and allergic using ImmunoCAP ISAC molecular diagnosis. Plasma samples were analyzed using liquid chromatography coupled to mass spectrometry (LC-MS). Subsequently, significant metabolites from plasma that were commercially available were then analyzed by targeted analysis in some nasal polyps. Additionally, nasal polyp and nasal mucosa samples were examined for eosinophils, neutrophils, $CD3^+$ and $CD11c^+$ cells, as well as collagen deposition and goblet cell hyperplasia. We found that 9 out of the 22 patients were sensitized to some aeroallergens (named as allergic CRSwNP). The other 13 patients had no sensitizations (non-allergic CRSwNP). Regarding metabolomics, bilirubin, cortisol, lysophosphatidylcholines (LPCs) 16:0, 18:0 and 20:4 and lysophosphatidylinositol (LPI) 20:4, which are usually related to a sustained allergic inflammation, were unexpectedly increased in plasma of non-allergic CRSwNP compared to allergic CRSwNP. LPC 16:0, LPC 18:0 and LPI 20:4 followed the same trend in nasal polyp as they did in plasma. Comparison of nasal polyps with nasal mucosa showed a significant increase in eosinophils ($p < 0.001$) and neutrophils ($p < 0.01$) in allergic CRSwNP. There were more eosinophils in polyps of non-allergic CRSwNP than in their nasal mucosa ($p < 0.01$). Polyps from non-allergic CRSwNP had less eosinophils than the polyps of allergic CRSwNP ($p < 0.05$) and reduced amounts of collagen compared to their nasal mucosa ($p < 0.001$). Our data suggests that there is a systemic inflammatory response associated to CRSwNP in the absence of allergy, which could be accountable for the nasal polyp development. Allergic CRSwNP presented a higher number of eosinophils in nasal polyps,

suggesting that eosinophilia might be connected to the development of nasal polyps in this phenotype.

Keywords: metabolomics, nasal polyps, allergy, targeted analysis, untargeted analysis

INTRODUCTION

Chronic rhinosinusitis with nasal polyps (CRSwNP) is a disease characterized by persistent inflammatory symptoms in nasal and paranasal mucosa that result in the development of a nasal polyp (Fokkens et al., 2012). This is an outgrowth of tissue that arises into the nasal cavity (Schleimer, 2017). The prevalence of the disease is estimated between 1–4% of the population (Fokkens et al., 2012; Chen et al., 2020). Symptoms include nasal blockage and itching, rhinorrhea, sneezing, facial pain, headache and smell impairment or loss (Georgy and Peters, 2012; Chen et al., 2020). Treatment starts with intranasal topic corticosteroids in the milder cases, followed by surgical extirpation or biological drugs, such as omalizumab, in the most severe ones (Lund, 1995; Georgy and Peters, 2012).

CRSwNP presents several endotypes, which are subtypes of diseases defined either by having the same molecular mechanism or because they respond to the same treatment (Anderson, 2008). In this sense, the analysis of cells and molecules involved in inflammation and the underlying pathophysiological mechanism is essential in CRSwNP (Brescia et al., 2020).

There are several comorbidities associated to CRSwNP appearance and recurrence, including asthma, allergic rhinitis, cystic fibrosis or aspirin-exacerbated respiratory disease (AERD) (Lund, 1995; Wynn and Har-El, 2004; Pearlman et al., 2010). In addition, nasal polyps relapse in up to 60–70% of the patients (Wynn and Har-El, 2004).

There are significant differences in histological features between nasal mucosa and nasal polyps, such as the development of oedema or fibrosis, goblet cell hyperplasia and/or squamous metaplasia; the thickening of the basal membrane; and the infiltration of inflammatory cells, including lymphocytes, eosinophils and neutrophils (Ferreira Couto et al., 2008).

However, the molecular mechanisms involved in the development of nasal polyps are still not well defined (Hopkins, 2019). Due to the inflammatory features described in several cohorts of CRSwNP (Brescia et al., 2018), most of the hypothesis agree that long-term maintained inflammation plays a key role in this process (Tos et al., 2010; Chojnowska et al., 2013; Brescia et al., 2021) and, therefore, inflammatory diseases such as allergy and/or asthma might be associated with its onset.

The role of allergy in CRSwNP development has been extensively discussed. In their 2014 review, Wilson *et al* (Wilson et al., 2014) examined the existing evidence both for and against an association of these two diseases. Although the role of allergy in CRSwNP was inconclusive, the review showed a higher rate of positive skin prick test among CRSwNP patients; and greater improvement in patients with negative skin prick tests compared to those with positive ones. These differences suggest that allergic inflammation might play a role in nasal polyposis.

Metabolomics is one of the most promising tools in the identification of biomarkers. It allows the detection of dynamic changes and alterations in the metabolism that point to a given pathological state (Villaseñor et al., 2017). The metabolome is closely linked to the phenotype and can be an extremely useful tool for diagnosing diseases and evaluating the effect of treatments. From a practical point of view, it uses very sensitive and specific techniques, such as liquid chromatography coupled to mass spectrometry (LC-MS), which allows the simultaneous detection of a great variety of metabolites in a biological sample (Crestani et al., 2020). Once the study subjects are well characterized, these results should be validated in a new and larger cohort of patients in subsequent studies. Compared to other omics, such as transcriptomics or genomics, the validation of metabolites found after exploratory studies is carried out through the development of analytical methodologies. This process is usually laborious depending on the number of metabolites and their physicochemical properties, and is conditioned by the availability of commercial standards.

Previous reports show that allergic inflammation has both systemic and local effects. As for the systemic role of allergy, we have previously demonstrated that severe allergic phenotypes induce significant changes in the plasma metabolome (Obeso et al., 2018). Similarly, allergic inflammation also induces local remodeling in oral mucosa (Aceves and Ackerman, 2009; Samitas et al., 2018; Rosace et al., 2019). In fact, this remodeling might result in the formation of the nasal polyp (Tos et al., 2010). However, to our knowledge, no metabolomic analysis has been performed for CRSwNP.

Here, our aim is to determine the role of allergic inflammation in patients with CRSwNP and how it affects both systemic and local features. Thus, we performed: 1) a systemic analysis, using a non-targeted metabolomics approach to achieve an overall picture of the metabolic profile in CRSwNP patients with and without allergy, and 2) a local analysis using both a targeted metabolomic analysis of the nasal polyps to evaluate if the systemic potential metabolites found are associated specifically to the nasal polyp, and a histological analysis to better characterize the local polyp environment. Our results support the idea that the nasal polyp has a specific inflammatory environment characterized by immune cells infiltrates, epithelial damage and the presence of inflammatory-related metabolites.

MATERIALS AND METHODS

Patients and Sample Collection

Twenty-two adult patients diagnosed with CRSwNP within an age range of 48 ± 8 years that arrived for the first time to the Otorhinolaryngology Service of the Hospital Madrid Monteprincipe (Spain) were included in this study. Patients

underwent a basic allergy history questionnaire. The inclusion criteria were: patients older than 18 years of age, with nasal polyps that needed to undergo a surgery to remove following their clinician's diagnosis. Patients received the same pharmacological pre-surgery treatment and their corticosteroid medication was removed two weeks prior the procedure. The exclusion criteria were: patients with concomitant inflammatory diseases such as autoimmune diseases or cancer. The ethical committee approved the study protocol and all subjects were informed and provided written consent prior to any procedure. Data were anonymized.

During endoscopic surgical procedures to remove the polyp, 5 mm biopsies of nasal polyp and nasal mucosa were obtained and either kept in RNAlater at −80°C or fixed in 4% paraformaldehyde (PFA) for 24–48 h. PFA-fixed samples were dehydrated and included in paraffin using Leica TP 1020 tissue processor. From these, 3 μm sections were obtained and used for histological and immunohistochemical analysis. Additionally, from the twenty-two patients, a blood sample from nineteen was obtained. About 20 ml of heparinized blood were collected to obtain plasma using a Ficoll-Paque (GE Healthcare™) density gradient centrifugation. Supernatants were stored at −80°C until their use.

ImmunoCAP ISAC

To determine the sensitization profile of the patients, ImmunoCAP ISAC® (Phadia, Uppsala, Sweden) with chips for 112 allergens assays were performed to detect specific IgE (sIgE) as described in the manufacturer protocol. Values above 0.3 ISU-E were considered positive.

Untargeted Plasma Metabolomic Analysis

Plasma samples were measured using a Liquid Chromatography coupled to Mass Spectrometry (LC-QTOF-MS Agilent series 6520). We followed previously described methodologies developed in our group (Ciborowski et al., 2010). Principal component analysis and heatmaps with hierarchical clustering were obtained using MetaboAnalyst v 5.0 webpage (https://www.metaboanalyst.ca).

Full descriptions are available in the Supplementary Information. Metabolite annotation was carried out using the online advanced CEU Mass Mediator tool (Godzien et al., 2015; Dudzik et al., 2018; Gil-de-la-Fuente et al., 2019). Statistical analysis was performed using non-parametric Mann Whitney *U* test, with statistical significance set at 95% ($p < 0.05$) with a Benjamini-Hochberg (also known as False Discovery Rate, FDR). Annotation was confirmed through LC-MS/MS experiments using 20 eV for fragmentation. Data were uploaded to Metabolomics Workbench webpage (ID study: ST001733 and ST001734).

Target Method for Nasal Polyps

Sample Preparation

Three nasal polyp samples of each group collected in RNAlater were used to measure the metabolites from plasma that were commercially available. RNAlater solvent was removed by washing the tissue 3 times with PBS 1X. Then, the nasal polyp was frozen in liquid nitrogen for 30 s. The frozen sample was put in cryoPREPTMCP02 (Covaris, MA, United States) plastic bags and submerged again for 30 s in liquid nitrogen. Once the plastic bag was inside the automated cryoPREPTMCP02, two consecutive impact forces of levels 2 and 4 out of 6 were applied. The resultant powder was gathered and weighted. Then, 100 μL of cold methanol:ethanol (1:1, v/v) and 0.5 μL of internal standard (LPC 18:1-d7; 0.01 mM) were added per each 10 mg of tissue for metabolite extraction and protein precipitation. Samples were then vortex-mixed and homogenized using Tissue-Lyser LT homogenizer (Qiagen, Germany) for 5 min at 50 Hz, 3 times. Supernatant containing the metabolites was separated from the pellet by centrifugation (2,000 rcf for 10 min at 4°C). Then, an aliquot of 70 μL was transferred to an LC vial and diluted with 490 μL of mobile phase (5% water: 95% acetonitrile; both with 7.5 mM ammonium acetate and 0.1% acetic acid). All samples were randomized before metabolite extraction and for the corresponding analytical run.

Analysis of Nasal Polyp by Liquid Chromatography Coupled to Triple Quadrupole Mass Spectrometry (LC-QqQ-MS)

Samples were measured using Dynamic Molecular Reaction Monitoring (dMRM) in LC-QqQ-MS for the analysis of metabolites. We used HPLC system (1260 Infinity, Agilent Technologies) coupled to an ESI(AJS)-QqQ-MS (6470 Agilent Technologies). The metabolites were separated on a Kinetex hydrophilic interaction liquid chromatography (HILIC) silica column (150 mm × 2.1 mm, particle size 100Å, Phenomenex, United States) maintained at 25°C. The mobile phases consisted of: A) water, and B) acetonitrile, both with 7.5 mM ammonium acetate and 0.1% acetic acid, with a final pH of 4.0 in the aqueous phase. The flow rate was 0.5 ml/min with an injection volume of 5 μL. Gradient started with 5% of A for 2 min, then increased up to 50% until 12 min, and back to initial conditions until 22 min. The MS conditions were: 5500 V of capillary voltage in positive ESI mode, a nebulizer gas flow rate of 11.0 L/min; a source temperature of 250°C; and a source pressure of 60 psi. The sample tray temperature was maintained at 4°C. Each transition was optimized adjusting the fragmentor and collision energy voltages, which can be seen in **Table 1**.

Data Acquirement, Pre-processing and Treatment

Data were acquired in MassHunter Workstation B.05.00 (Agilent Technologies), and re-processed using MassHunter QQQ Quantitative Analysis B.08.00 (Agilent) where peak areas were integrated. Concentration of metabolites were calculated using external calibration curves with the standard addition method. Once the concentrations were obtained, all data treatment and statistical comparisons was performed using Excel 2016, MATLAB R 2018b and GraphPad Prism v8.1.2.

Immune Cell Quantification

Tissue Staining

We quantified eosinophils, neutrophils, $CD3^+$ and $CD11c^+$ cells, and checked for collagen deposition and goblet cell hyperplasia,

TABLE 1 | Optimized MS transitions and parameters for the targeted analysis. LPC: Lysophosphatidylcholine, LPI: Lysophosphatidylinositol, RT: Retention time, V: Volt, eV: Electronvolt.

Compound name	Precursor ion (*m/z*)	Product ion (*m/z*)	RT (min)	Fragmentor (V)	Collision energy (eV)	Polarity
Bilirubin	585.3	299.1	0.91	131	25	+
LPC 16:0	496.3	183.9	9.95	100	28	+
LPC 18:0	524.4	183.8	9.80	100	28	+
LPI 20:4	621.3	361.3	7.83	84	17	+

in nasal polyps and nasal mucosa samples. For eosinophil quantification, we adapted a Luna staining protocol. Samples were stained with a working solution prepared with 0.9 parts of homemade Weigert's Iron Hematoxylin, and 0.1 parts of commercial Briebrich-Scarlett solution (Sigma Aldrich, ref. HT151) for 5 min. Slides were then differentiated in 1% acid alcohol and 0.25% lithium carbonate solutions, and preparations were mounted.

Regarding goblet cell hyperplasia, we optimized a Periodic Acid-Schiff (PAS) staining. Briefly, samples were kept in a 0.5% periodic acid solution, then stained with Schiff's reactive (Merck, ref. 109033) and washed. Nuclei were stained with a 1:4 Harris Hematoxylin solution and differentiated with a 1% acid alcohol (1% of HCl in 70% ethanol) solution and a saturated lithium carbonate (Sigma-Aldrich, ref. 62470) solution. Preparations were mounted with DPX medium.

Regarding neutrophil, $CD3^+$ cells and $CD11c^+$ cells quantification, we performed immunohistochemical staining with anti-human neutrophil elastase (ab68672, ABCAM), anti-CD3 (MCA 1477, AbD Serotec) and anti-CD11c (NCL-L-CD11c-563, Novocastra) using the automatized system Leica Bond Max (Leica Biosystems), as previously described (Sanchez-Solares et al., 2019). For negative controls, the antibody was substituted with antibody diluent Bond™ (Leica Biosystems) for incubation.

Masson Trichrome staining (Sigma-Aldrich, ref HT15-1KT) was performed following the manufacturer's instructions.

Image Analysis

All slides were scanned with Leica SCN400 scanner (Leica Biosystems). Images were obtained for each staining using the Leica Scan Viewer software. Luna and neutrophil elastase-positive cells were counted in the whole sample, while $CD3^+$ cells and $CD11c^+$ cells were counted in five representative areas on the sample. Areas were measured using ImageJ v1.51j8 by at least two independent observers. Results are presented as number of cells per area. For goblet cell hyperplasia, PAS^+ stained areas in the epithelium were measured, while for collagen deposition, green areas of Masson staining were measured in the whole sample using Image-Pro Plus v4.5.0.29 for Windows (Media Cybernetics) software.

Statistical Analysis

GraphPad Prism v8.1.2 for Windows (GraphPad Software) was used for statistical analysis. We checked data distribution and then used *t*-test or Mann-Whitney *U* test to determine significant differences between the means accordingly. Statistical significance was set at 95% ($p < 0.05$).

RESULTS

Patient Classification

Patients with CRSwNP were classified according to their allergic sensitization phenotype by ImmunoCAP ISAC (**Figure 1**). As observed in the figure, 13 patients (59.1%) had undetectable sIgE levels and were classified as non-allergic CRSwNP. On the other hand, 9 (40.9%) had significant levels of sIgE. They were sensitized to either one or more perennial allergen sources (*i.e.* cat, dog, mites or *Alternaria*) and/or to a group of seasonal allergen sources. Thus, we consider them to be perennial allergic patients and classified them as allergic CRSwNP patients.

Untargeted Plasma Metabolomic Analysis

The plasma metabolic profile of each patient with CRSwNP with or without allergic sensitization was acquired. After data treatment, 535 features for LC-MS positive and 429 for LC-MS negative ionization modes in each plasma sample were obtained. These features passed the different quality filters (blank subtraction, presence in quality control samples (QCs; >50%) and patients (>75%), and coefficient of variation in QCs (<30%)).

To assess the correct performance of the LC-MS equipment and the quality of the acquired data, samples were projected on a PCA model (**Figure 2A**). Clustering of the QC injections in the non-supervized plot indicated the high quality of the data, while dispersion of the samples showed the biological variability of the patients. The groups were compared using a discriminant PLS-DA analysis, obtaining no model. Thus, a Mann Whitney-*U* test was used for the selection of significant features, finding a total of 13 and 26 features from LC-MS in positive and negative ionization modes, respectively. For these significant features, we assessed that they had a coefficient of variation (CV) on QCs lower than the percentage of change between the non-allergic CRSwNP compared to the allergic CRSwNP group. Additionally, we checked that metabolites had a % of change higher than 20% or that the feature matched the complementary polarity. These features were represented using a heatmap with hierarchical clustering for each ionization mode (**Figures 2B,C**). As observed, the significant features from the positive ionization mode were able to define a specific metabolic signature for each group, accurately clustering all the samples. On the other hand,

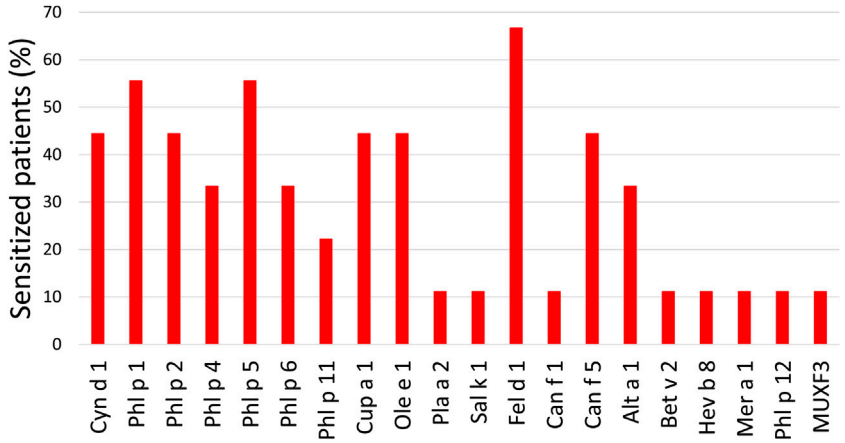

FIGURE 1 | Percentage of allergic patients sensitized to each allergen. Cyn d: *Cynodon dactylon* (Bermuda grass); Phl p: *Phleum pratense* (timothy); Cup a: *Cupressus arizonica* (cypress); Ole e: *Olea europaea* (olive); Pla a: *Platanus acerifolia* (London plane tree); Sal k: *Salsola kali* (saltwort); Fel d: *Felis domesticus* (cat); Can f: *Canis familiaris* (dog); Alt a: *Alternaria alternata* (*Alternaria* plant rot fungus); Bet v: *Betula verrucose* (European white birch); Hev b: *Hevea brasiliensis* (latex); Mer a: *Mercurialis annua* (annual mercury); MUXF3: Bromelain.

the significant features from the negative ionization mode were able to correctly cluster around 85% of the samples (16 out of 19).

We carried out an annotation for all the significant features. After tentative annotation and confirmation by MS/MS fragmentation experiments, we were able to annotate 14 features. Taking into account both polarities, eight metabolites had a unique annotation–comprizing sn-1 and sn-2 for lysophospholipids– (**Table 2**). We found increased levels of lysophosphatidylcholines (LPC 16:0, LPC 18:0 and LPC 20:4), a lysophosphatidylinositol (LPI 20:4), cortisol and bilirubin in the non-allergic CRSwNP compared to the allergic CRSwNP group.

Overall, specific systemic metabolic changes were defined for non-allergic and allergic CRSwNP patients.

Targeted Metabolomic Analysis of Nasal Polyps

The significant metabolites found in plasma that had available commercial standards were used to test their presence in the nasal polyp. Therefore, nasal polyp samples from six CRSwNP patients (three non-allergic and three allergic) were analyzed to confirm the results found in plasma. We applied univariate statistical analysis between the two groups, obtaining no statistically significant changes in any metabolite ($p > 0.05$, Mann Whitney *U* test). However, to study the metabolite patterns, these were represented in a heatmap using the average per group (**Figure 2D**). We observe that, despite the low number of samples, the LPC 16:0, LPC 18:0 and LPI 20:4 followed the same increasing trend in the non-allergic CRSwNP patients as we described in plasma samples. Finally, these metabolites in plasma and in the nasal polyps from untargeted and targeted analyses were represented using trajectories in **Figure 3**.

Histological Features

To better characterize the specific local features of nasal polyps and mucosa associated to allergy, we analyzed the infiltration of eosinophils, neutrophils, $CD3^+$ and $CD11c^+$; along with the hyperplasia of goblet cells and the abundance of collagen fibers in all of them.

Eosinophil quantification (**Figure 4A**) revealed that the number of these cells in nasal polyps was higher than in the nasal mucosa samples (9.7 ± 2.3 cells/area *vs* 0.71 ± 0.41 cells/area, $p < 0.001$). Importantly, the highest number of eosinophils was detected in nasal polyps from allergic CRSwNP patients (**Figure 4B**). In fact, polyps of non-allergic CRSwNP patients had significantly less eosinophils than those of allergic CRSwNP patients (16.6 ± 3.6 cells/area *vs* 4.5 ± 1.4 cells/area, $p < 0.05$).

We also found more neutrophils (**Figure 5A**) in nasal polyps than in the nasal mucosa (0.60 ± 0.10 cells/area *vs* 0.31 ± 0.10 cells/area, $p < 0.01$). Nasal polyps of allergic CRSwNP patients had the highest number of neutrophils (**Figure 5B**), being significantly different to their own nasal mucosa (0.70 ± 0.17 cells/area *vs* 0.16 ± 0.060 cells/area, $p < 0.01$). However, there were no significant differences in the number of neutrophils in nasal polyps between non-allergic and allergic CRSwNP patients ($p > 0.41$).

No statistical differences were observed between the nasal polyps of non-allergic CRSwNP and those of allergic CRSwNP regarding collagen deposition (**Figure 6**, $p > 0.66$), goblet cell hyperplasia (**Figure 7**, $p < 0.87$), infiltration of $CD3^+$ (**Figure 8**, $p < 0.22$) and $CD11c^+$ cells (**Figure 9**, $p > 0.41$).

However, more collagen deposition was observed in the nasal mucosa compared to the nasal polyps in the non-allergic CRSwNP ($p < 0.001$). Moreover, the same trend was found for the allergic CRSwNP patients, although the difference was not significant ($p > 0.059$) (**Figure 6B**).

There also seemed to be a higher area stained with PAS in the epithelium of nasal polyps compared to the nasal mucosa for both allergic and non-allergic CRSwNP groups (**Figure 7**). Although the difference was not significant ($p > 0.29$ for non-allergic CRSwNP and $p > 0.41$ for allergic CRSwNP patients), this fact suggests the presence of goblet cell hyperplasia.

Finally, for the infiltration of $CD3^+$ and $CD11c^+$ cells (**Figures 8**,**9**, respectively), no statistical differences were obtained between

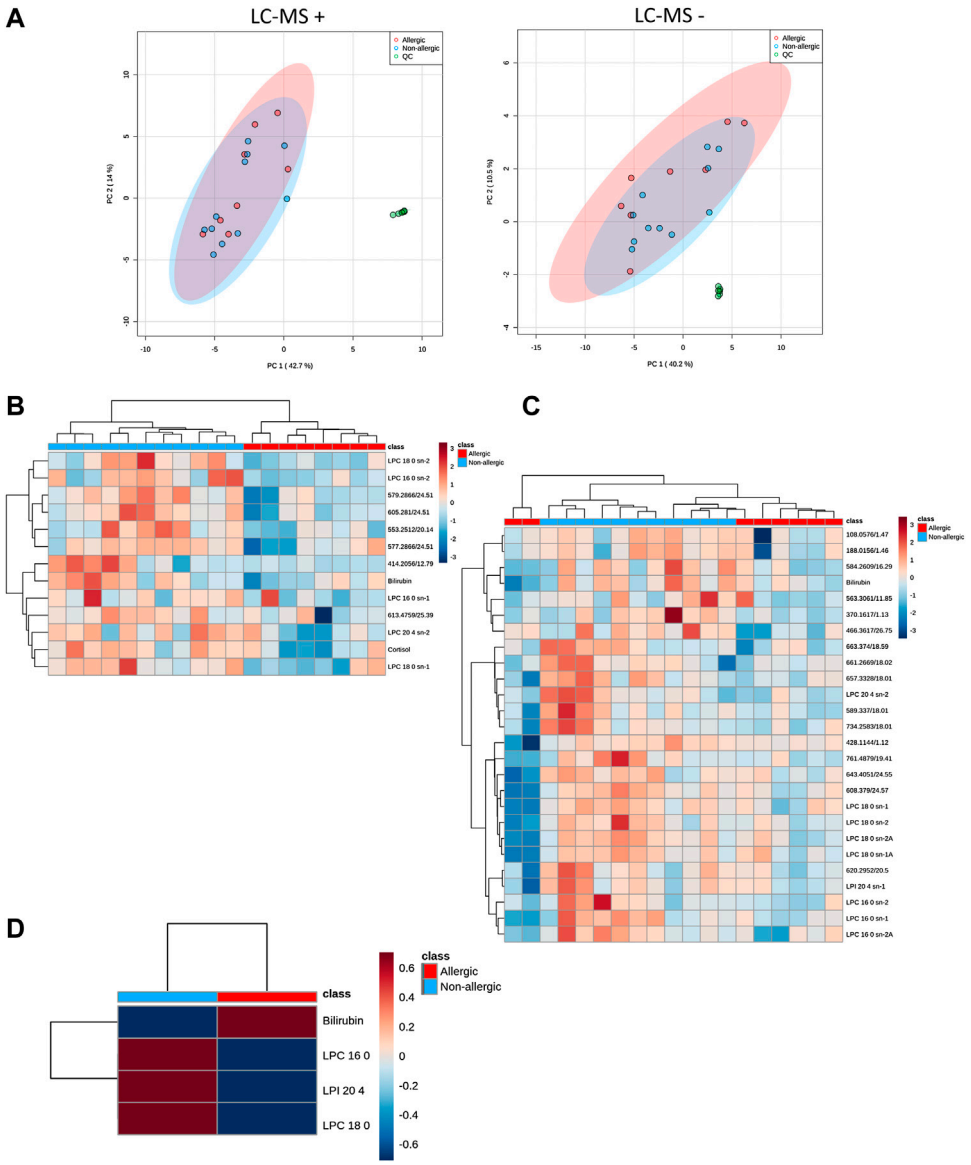

FIGURE 2 | (A). Unsupervized PCA for QC injections (●; green dot), allergic patients with CRSwNP (●; red dot) and non-allergic patients with CRSwNP (●; blue dot) models showing the quality of the data for LC-MS. Confidence regions for each group were set at 95% and are depicted with the corresponding color group. Data were centered scaled and Log transformed. **(B,C).** Significant signals were depicted on a heatmap using hierarchical clustering of the samples (represented in columns) and metabolites (in rows) for ESI + and ESI – , respectively. Red cells represent higher levels of the specific metabolite in that sample, whereas blue cells represent lower levels. Samples and metabolites are clustered according to their similarity. Mann-Whitney *U* test with a Benjamini-Hochberg correction was used to detect statistical significance ($p < 0.05$). Unknown features are represented by "Mass@Retention Time." **(C).** Heatmaps of significant plasma metabolomics signals using hierarchical clustering analysis of the experimental samples LC-MS positive polarity. Mann Whitney *U* test was used to detect statistical significance ($p < 0.05$). **(D).** Target metabolites analyzed by LC-QqQ-MS in nasal polyps were represented in a heatmap using hierarchical clustering of the samples (represented in columns) and metabolites (in rows) showing group averages.

nasal mucosa and nasal polyp for both allergic and non-allergic CRSwNP groups.

In summary, we found that nasal polyps present a higher immune cell infiltration (eosinophils and neutrophils) than nasal mucosa for both allergic and non-allergic CRSwNP patients. Moreover, allergic CRSwNP patients were characterized by eosinophilia in their nasal polyps compared to non-allergic CRSwNP patients. In the case of non-allergic CRSwNP patients, more collagen deposition in their nasal mucosa than in their nasal polyp was observed.

DISCUSSION

Nasal polyps are growths of inflamed nasal tissue that have been well-known for a long time. However, the molecular mechanisms

TABLE 2 | Pairwise comparisons showing the significant identified metabolites.

						Non-allergic CRSwNP *vs* allergic CRSwNP					
N°	**Technique**	**Compound**	**Adduct**	***m/z* (Da)**	**Mass (Da)**	**RT (min)**	**Error (ppm)**	**%CV on QCs**	**FC in non-allergic**	**% Change in non-allergic**	***p*-Value**	***p*-BH**
1	LC – MS –	Bilirubin	$[M-H]^-$	583.2549	584.2621	32.94	2	16.9	2.03	103.28	0.041	0.049
	LC – MS +		$[M + H]^+$	585.2702	584.2630	32.90	1	11.4	1.96	95.56	0.033	0.039
2	LC – MS +	Cortisol	$[M + H]^+$	363.2171	362.2099	3.66	1	12.4	2.20	119.55	0.026	0.034
3	LC – MS –	LPC 16:0 sn-1	$[M + FA]^-$	540.3301	495.3324	19.42	1	8.8	1.23	23.49	0.026	0.038
	LC – MS +		$[M + H]+$	496.3402	495.3330	19.36	1	13.5	1.27	27.17	0.036	0.042
4	LC – MS –	LPC 16:0 sn-2	$[M + FA]^-$	540.3305	495.3324	20.20	1	6.2	1.17	16.94	0.041	0.049
	LC – MS +		$[M + H]+$	496.3403	495.3331	20.14	1	10.5	1.25	25.45	0.018	0.032
5	LC – MS –	LPC 18:0 sn-2	$[M + FA]^-$	568.3623	523.3642	23.74	1	9.1	1.33	32.80	0.007	0.029
	LC – MS +		$[M + H]+$	524.3713	523.3641	23.66	0	12.6	1.36	36.19	0.005	0.025
6	LC – MS –	LPC 18:0 sn-1	$[M + FA]^-$	568.3618	523.3637	24.58	1	6.2	1.23	22.55	0.041	0.049
	LC – MS +		$[M + H]+$	524.3719	523.3647	24.52	2	11.1	1.30	30.01	0.010	0.031
7	LC – MS –	LPC 20:4 sn-2	$[M + Cl]^-$	578.3016	543.3324	18.59	1	10.8	1.42	42.11	0.026	0.034
	LC – MS +		$[M + H]+$	544.3401	543.3329	18.49	1	23.7	1.49	49.03	0.054	0.071
8	LC – MS –	LPI 20:4 sn-1	$[M-H]^-$	619.2880	620.2952	19.59	1	6.1	1.38	38.01	0.020	0.034

LC-MS: Liquid Chromatography coupled to Mass Spectrometry; RT: retention time; ppm: part per million; FC: fold change, was calculated as average of area in Non-allergic/average of area in Allergic; % change was calculated as (FC–1) × 100; LPC: Lysophosphatidylcholine; LPI: Lysophosphatidylinositol; FA: formic acid; BH: Benjamini-Hochberg (False Discovery Rate).

involved in the development of nasal polyps remain unclear. Additionally, even though allergy and CRSwNP have been traditionally linked, whether there is a connection between allergy and the development of nasal polyps or not is yet to be described. Here, we perform an original experimental design, aiming to elucidate systemic metabolic differences in CRSwNP patients with and without allergy.

From our patient cohort, 40% (9 out of 22) of the patients were allergic, close to the levels in average population which is around 30% (Bauchau and Durham, 2004; Bousquet et al., 2008).

A metabolic fingerprint in plasma of patients with CRSwNP was obtained. Non-allergic CRSwNP patients displayed an increase of LPCs (LPC 16:0, LPC 18:0, LPC 20:4) together with LPI 20:4, compared to allergic CRSwNP. These LPCs have been previously associated with systemic inflammation (Chiurchiù et al., 2018; Obeso et al., 2018). Moreover, they have been described in asthma (Ried et al., 2013; Comhair et al., 2015; Villaseñor et al., 2017) as arachidonic acid (AA) precursors. The free fatty acids from the LPCs are released after the action of lipase A2 to synthesize inflammatory mediators, that participate as precursors in the AA pathway (Balgoma et al., 2010; Arita, 2016; Bennett and Gilroy, 2016). Interestingly, LPI 20:4, which is one of the more abundant LPIs in plasma, has been related to a potent pro-inflammatory signaling in intestinal bowel disease and colorectal cancer in animal models (Grill et al., 2019) and in type 2 diabetes (Lu et al., 2016). Overall, the LPI 20:4 metabolite seems to be involved in the inflammatory response.

On the other hand, there is growing evidence that bilirubin exerts potent anti-inflammatory effects. Bilirubin is able to suppress inflammatory responses by preventing the migration of leukocytes into target tissues through the disruption of vascular cell adhesion molecule-1 (VCAM-1)-dependent cell signaling. In a previous study, bilirubin was shown to alleviate colitis (Vogel and Zucker, 2016). Additionally, nanoparticles containing bilirubin were used for the treatment of allergic lung inflammation disease in a mouse model obtaining amelioration of the disease (Kim et al., 2017). Therefore, bilirubin has been demonstrated that has anti-oxidative, anti-inflammatory and immunosuppressive functions in various diseases such as inflammatory bowel disease, cardiovascular disease, autoimmune disorders, cancer and type 2 diabetes mellitus (Peng et al., 2017). Mildly elevation of this metabolite is associated with better prognosis; thus, the decrease we observe in allergic CRSwNP patients might be related with a worse outcome. As bilirubin can be measured in a laboratory standard test, this metabolite can be compared in future studies in these patients.

Finally, cortisol has been described as a hormone which acts suppressing early inflammatory responses; however, when cortisol is not enough to control inflammation, it prepares immune cells for major subsequent inflammatory episodes (Yeager et al., 2018). Interestingly, circadian rhythms of salivary cortisol were found to be disrupted in patients with extensive nasal polyposis compared to controls (Fidan et al., 2013). The authors suggest that a therapy with cortisol-based drugs might be useful in the treatment of CRSwNP (Fidan et al., 2013). In this line, intrapolyp steroid injections appear to

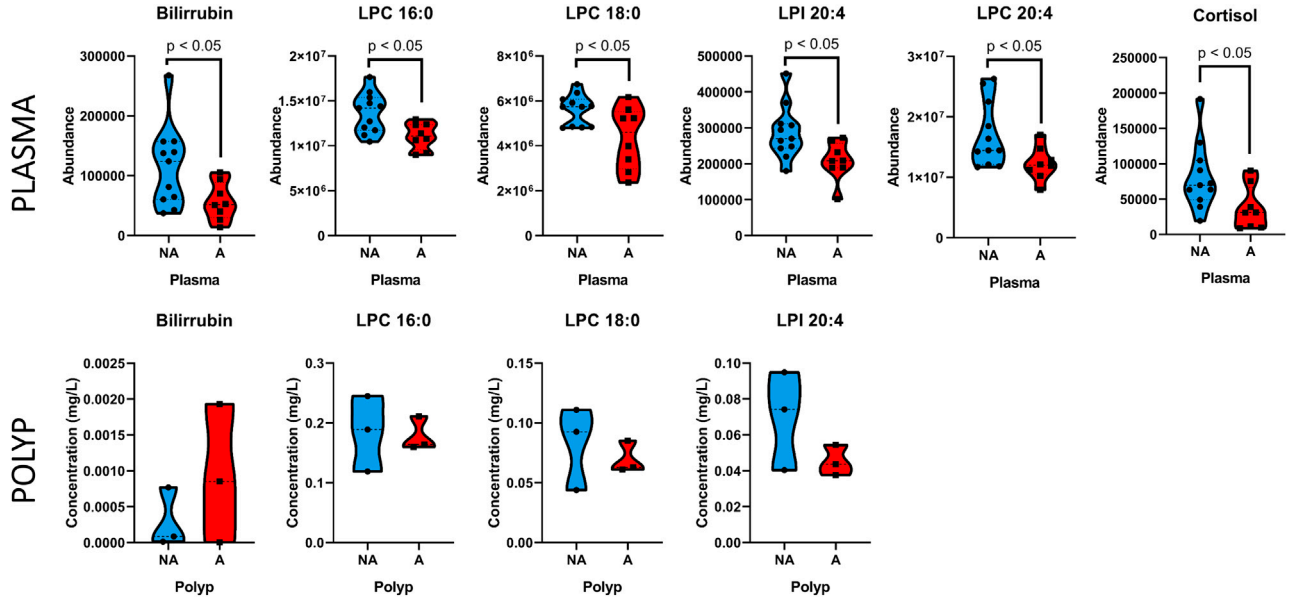

FIGURE 3 | Trajectories of representative metabolites in non-allergic CRSwNP (blue), and allergic CRSwNP (red) patients in plasma and nasal polyp samples. The median is indicated by the discontinuous line within the violin plot, and each sample's value is shown individually.

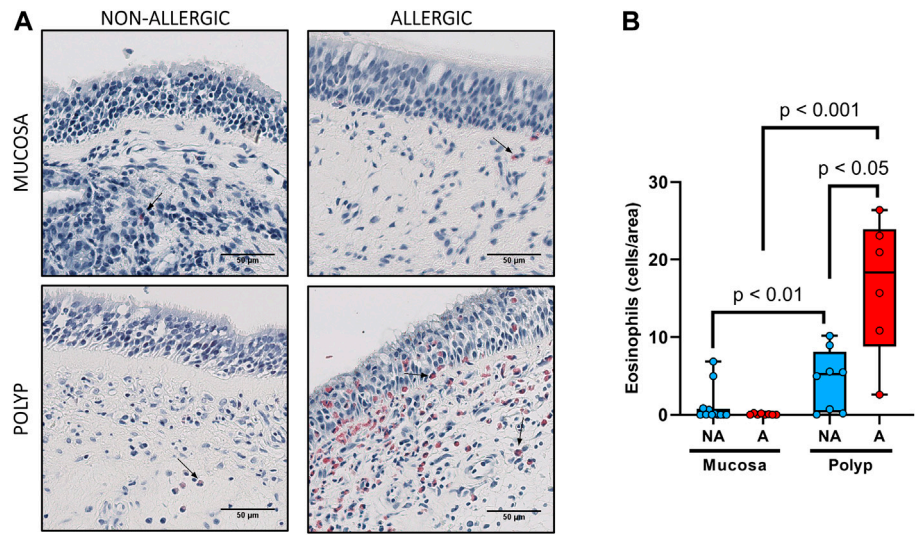

FIGURE 4 | Analysis of eosinophil infiltrates in nasal polyps and nasal mucosa. **(A).** Representative images of Luna staining in nasal mucosa (n = 20) and nasal polyps (n = 14) in all experimental groups (eosinophils: black arrow →, scale bar = 50 µm). **(B).** Comparison of the number of eosinophils per area in nasal polyps and nasal mucosa samples between non-allergic CRSwNP (NA) and allergic CRSwNP (A) patients. Results are presented as mean ± SEM. Mann-Whitney U test was used to state significant differences.

be effective and safe for the treatment of nasal polyps (Kırıs et al., 2016). Additionally, it has been shown that polyp-derived epithelial cells produce cortisol, which may be involved in the anti-inflammatory response established when patients receive treatment with glucocorticoid for nasal polyps (Kook et al., 2015). The elevation of this metabolite in the non-allergic CRSwNP patients might point to a specific pathway to cope with inflammation. Additionally, this suggests that, in allergic CRSwNP patients, the cortisol acts by recruiting immune cells instead of solving the inflammation by itself.

Our results demonstrate that allergy produces metabolic changes in plasma in patients with CRSwNP. The increase of LPC 16:0, LPC 18:0, LPC 20:4, LPI 20:4, cortisol and

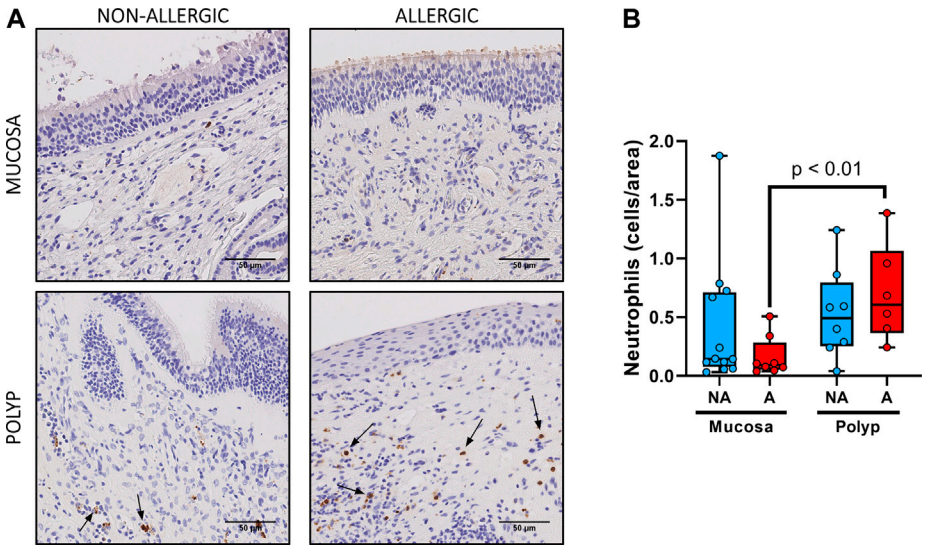

FIGURE 5 | Analysis of neutrophil infiltrates in nasal polyps and nasal mucosa. **(A).** Representative images of anti-neutrophil elastase immunostaining in nasal mucosa (n = 20) and nasal polyps (n = 14) in all experimental groups (neutrophils: black arrow →, scale bar = 50 μm). **(B).** Comparison of the number of eosinophils per area in nasal polyps and nasal mucosa samples between non-allergic CRSwNP (NA) and allergic CRSwNP (A) patients. Results are presented as mean ± SEM. Mann-Whitney U test was used to state significant differences.

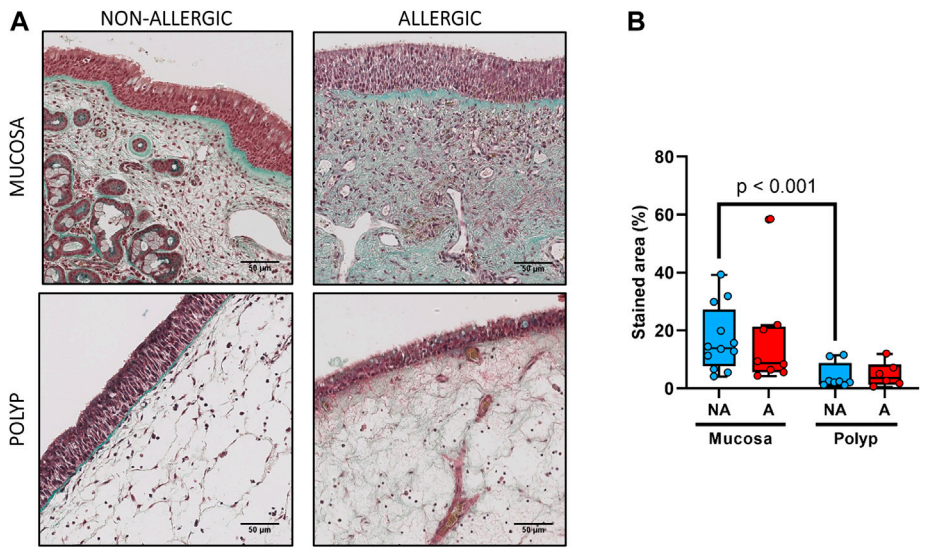

FIGURE 6 | Analysis of collagen fibers deposition in nasal polyps and nasal mucosa. **(A).** Representative images of Masson-Trichrome staining in nasal mucosa (n = 20) and nasal polyps (n = 14) in all experimental groups (collagen fibers: green, cytoplasm: red, erythrocytes: gold, scale bar = 50 μm). **(B).** Comparison of the green-stained area (collagen fibers) in nasal polyps and nasal mucosa samples between non-allergic CRSwNP (NA) and allergic CRSwNP (A) patients. Results are presented as mean ± SEM. Mann-Whitney U test was used to state significant differences.

bilirubin metabolites in the plasma of non-allergic CRSwNP patients points toward a systemic inflammatory response in the absence of allergy. These metabolites might participate in the development of a characteristic tissue environment, responsible for the differences observed at the histological and inflammatory infiltration levels. Therefore, we hypothesized that they might be also altered in the nasal polyp in the same way.

Consequently, we implemented a novel target methodology to analyze these significant metabolites from plasma that were commercially available in the nasal polyp, obtaining good analytical parameters. We observed that the trends of the LPC

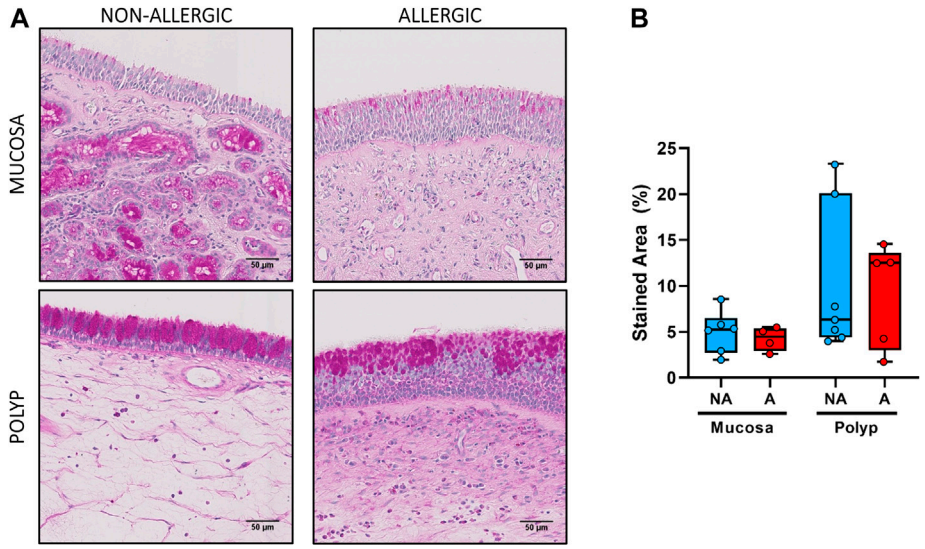

FIGURE 7 | Analysis of goblet cell hyperplasia in nasal polyps and nasal mucosa. **(A).** Representative images of PAS staining in nasal mucosa (n = 10) and nasal polyps (n = 12) in all experimental groups (mucopolysaccharides: pink, nuclei: purple, scale bar = 50 µm). **B.** Comparison of the PAS-positive (pink) stained area in the epithelium of nasal polyps and nasal mucosa samples between non-allergic CRSwNP (NA) and allergic CRSwNP (A) patients. Results are presented as mean ± SEM. Mann-Whitney U test was used to state significant differences. Only samples with sufficient epithelial integrity were included in this analysis.

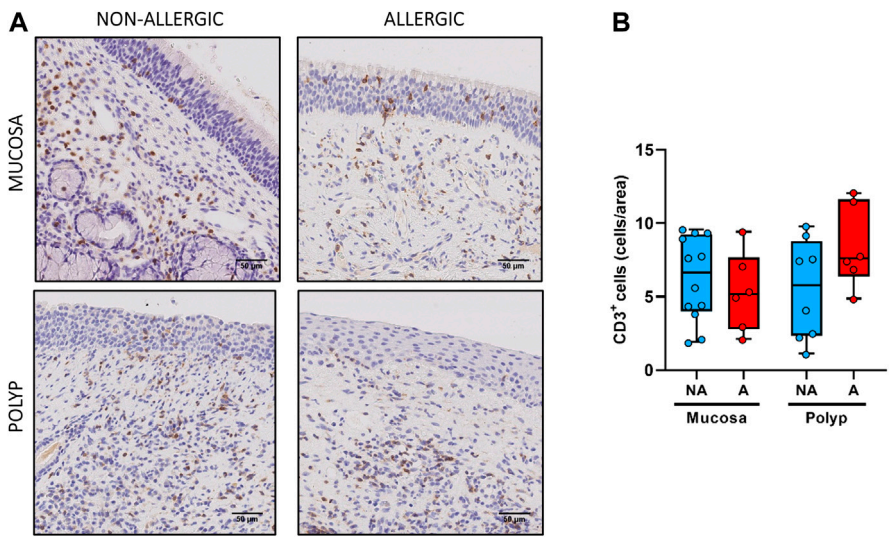

FIGURE 8 | Analysis of $CD3^+$ cells infiltrates in nasal polyps and nasal mucosa. **(A).** Representative images of anti-CD3 immunostaining in nasal mucosa (n = 18) and nasal polyps (n = 14) in all experimental groups ($CD3^+$ cells: black arrow →, scale bar = 50 µm). **(B).** Comparison of the number of $CD3^+$ cells per area in nasal polyps and nasal mucosa samples between non-allergic CRSwNP (NA) and allergic CRSwNP (A) patients. Results are presented as mean ± SEM. Mann-Whitney U test was used to state significant differences.

16:0, LPC 18:0 and LPI 20:4 metabolites are similar in plasma and nasal polyps samples. However, further studies are needed, including a higher number of tissue samples to validate these results.

Together, the metabolomic results demonstrate that allergy induces specific metabolic changes in CRSwNP patients. These are LPC 16:0, LPC 18:0, LPC 20:4, LPI 20:4, cortisol and bilirubin. Most of these metabolites were measured in nasal polyp, a biological sample that has not been previously analyzed by metabolomics. We showed that LPC 16:0, LPC 18:0 and LPI 20:4 follow the same trend in the nasal polyp than in plasma. This novel methodological procedure could be used in future studies to

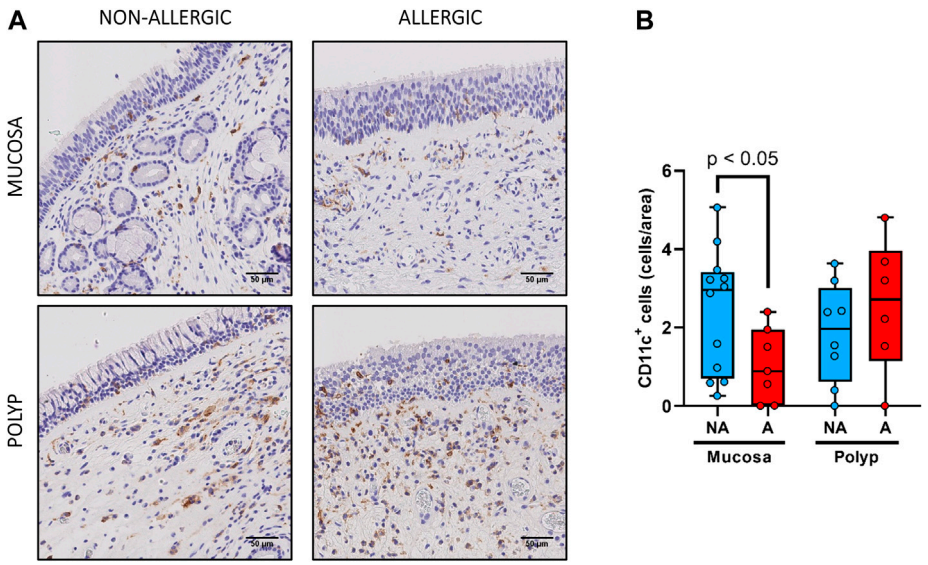

FIGURE 9 | Analysis of CD11c$^+$ cells infiltrates in nasal polyps and nasal mucosa. **(A).** Representative images of anti-CD11c immunostaining in nasal mucosa (n = 19) and nasal polyps (n = 14) in all experimental groups (CD11c$^+$ cells: black arrow →, scale bar = 50 μm). **(B).** Comparison of the number of CD11c$^+$ cells per area in nasal polyps and nasal mucosa samples between non-allergic CRSwNP (NA) and allergic CRSwNP (A) patients. Results are presented as mean ± SEM. Mann-Whitney U test was used to state significant differences.

better understand the local metabolomic environment of nasal polyps.

Histological characteristics in the nasal polyp and their nasal mucosa of these patients was also analyzed. Tissue eosinophils have been described both in nasal polyps and in allergic pathologies. In fact, treatment with anti-IL5 (mepolizumab) has shown a reduction in both nasal polyp and blood eosinophils, and a significant improvement in CRSwNP patients, resulting in better prognosis (Bachert et al., 2017). Eosinophils have been also measured locally and systemically. Therefore, a high number of tissue eosinophils and/or their proteins in nasal polyps have been related to more severe and maintained symptoms (Sun et al., 2017) and recurrence (Tos et al., 2010; Van Zele et al., 2014; Yang et al., 2018). High numbers of eosinophils in plasma correlate with higher allergy incidence and worse symptomatology as well (Erbek et al., 2007). Thus, it seems clear that eosinophils might play a role in the development and prognosis of allergic CRSwNP.

On the other hand, neutrophil quantification showed a similar distribution pattern between non-allergic CRSwNP and allergic CRSwNP patients, which was independent of the presence or absence of eosinophils. Although there are studies describing lower numbers of neutrophils in eosinophilic CRSwNP compared to the non-eosinophilic polyps (Schleimer, 2017); other authors (Pothoven et al., 2017; Kong and Kim, 2018) have recently reported similar results, linking the role of neutrophils in this disease to their expression of oncostatin M (OSM), a cytokine that has been found elevated in CRSwNP and that induces barrier dysfunction.

Additionally, a lower collagen deposition in nasal polyps compared to nasal mucosa for both allergic and non-allergic CRSwNP patients, suggest that the polyps of the study are edematous, rather than a fibrous (Brescia et al., 2021).

In brief, immune cell infiltration analysis revealed differential features between nasal polyp and nasal mucosa and suggest that a maintained allergy would enhances the inflammatory response mediated by eosinophils in nasal polyps but, surprisingly, not in nasal mucosa for allergic CRSwNP patients.

Separately, according to Brescia *et al* (Brescia et al., 2021), CRS appears to be a very heterogeneous inflammatory condition, with various emerging endotypes. Molecular and cellular screening together with clinical phenotyping in CRSwNP could be useful to define endotypes, and thus clarify the inflammatory mechanisms and allow the establishment of a more precise treatment. However, because of the aim of this exploratory study was to use allergy as the main classifying criterion, we are not able to describe our data in terms of CRS endotypes. This limitation should be addressed in the future including a more complete immunological, histopathological and clinical phenotyping approach, what will lead to an improvement in the endotyping capability, as it has been reported (Brescia et al., 2017; Kim and Cho, 2017; Bachert et al., 2018). In this sense, metabolomics could offer help in both 1) the endotyping of patients with CRSwNP, along with cellular and other molecular analysis; and 2) the understanding of the molecular mechanisms that underlay the endotypes.

Our findings suggest that the two phenotypes (non-allergic CRSwNP and allergic CRSwNP), although sharing histological features such as cellular infiltration, have distinctive metabolomic fingerprints and eosinophilia, which point toward different mechanisms of formation.

Previous studies from our group have followed this metabolomics approach in other respiratory allergy models and have found alterations in the AA pathway, as we report in this manuscript (Obeso et al., 2018; Barker-Tejeda et al., 2020). Although, surprisingly, this route was downregulated in allergic CRSwNP patients, an increased recruitment of eosinophils in the nasal polyp was observed. Therefore, we hypothesized that the development of the phenotype CRSwNP without allergy requires a great underlying uncontrolled systemic inflammation, which is different in the allergic phenotype. Thus, eosinophils seem to be accountant for the inflammation leading to CRSwNP development in the allergic phenotype, while non-allergic CRSwNP phenotype would be characterized by higher levels of AA precursors and other inflammatory mediators needed to develop the nasal polyp. The reason of the increase in these biological pathways is, however, yet to be defined.

This is an exploratory study, where the design is innovative and aimed to understand the effect of respiratory allergy in the development of CRSwNP; however, it has some limitations. Although the sample size was small, the samples (nasal polyps, nasal mucosa and plasma) were extensively characterized by metabolomics and histology. Moreover, despite the metabolic alterations that were observed between these two groups of patients not being able to generate a discriminant model (i.e. PLS-DA) capable of predicting new samples, the findings are promising in this field and would shed light on the mechanism by which patients without allergy develop CRSwNP. Therefore, in further studies, the validation of these results in a bigger cohort is needed.

Overall, CRSwNP is a pathology of high level of clinic-pathological complexity where the collection of biopsies is not an easy task. Although the role of histological study of biopsies is a complementary approach for the endotyping of nasal polyps, the inclusion of a metabolomic analysis has allowed us to identify biological processes associated with the allergic or non-allergic phenotypes and, therefore, could be helpful in the design of novel, less invasive treatments for these patients.

CONCLUSION

Our results demonstrate that patients with CRSwNP with and without allergy display systemic metabolic changes. Surprisingly, these metabolites (LPC 16:0, LPC 18:0, LPC 20:4, LPI 20:4, cortisol and bilirubin), which are associated with inflammation, appear to be increased in the absence of allergy, suggesting that non-allergic CRSwNP patients develop the nasal polyp after a sustained systemic inflammation. We have developed a new method for the analysis of nasal polyps using targeted metabolomics. With this, we spotted the same trends of LPC 16:0, LPC 18:0 and LPI 20:4 that were observed in plasma. The increased numbers of eosinophils in the nasal polyp of allergic CRSwNP patients hints that nasal polyps might develop by a local immune effector cell recruitment that ends in tissue remodeling. Finally, this is an exploratory study, where the significant metabolites were obtained by a semi-quantitative comparison between the groups. Thus, interpretation of these results should be made taking into account these limitations. Further validation studies in new cohort of samples using targeted quantitative methods must be carried out.

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://www.metabolomicsworkbench.org/, ST001733;https://www.metabolomicsworkbench.org/, ST001734.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethical committee of Hospital Madrid Monteprincipe (Spain). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TC was the PI and together with DB, ME and AV, which designed and supervised the research. MID-D, JS-S, LM-B and PF performed the immunohistochemical analysis. DO and AV performed the metabolomic analysis. MID-D, DO, CB, DB and AV contributed to the metabolomics interpretation. MF included all study patients. CB and ME supervised the metabolomic and immunohistochemical analysis, respectively. All authors contributed to the writing of the manuscript and have given approval to the final version of the manuscript.

FUNDING

This work was supported by Instituto de Salud Carlos III (project numbers PI19/00044 and PI18/01467) co-funded by European Regional Development Fund "Investing in your future" for the thematic network and co-operative research centers ARADyAL RD16/0006/0015. This work was also supported by the grant from Ministerio de Ciencia, Innovacion y Universidades RTI 2018–095166-B-I00. MID-D, JS-S, LM-B were supported by FPI-CEU predoctoral fellowships and A.V. is funded by a postdoctoral research fellowship from ARADyAL. Data have been presented at EAACI Congress 2018 (Munich, Germany) and EAACI Congress 2019 (Lisbon, Portugal).

ACKNOWLEDGMENTS

We would like to thank Virginia Garcia and Javier Moratinos of the Histology Core Facility at the Institute of Applied Molecular Medicine, Faculty of Medicine, San Pablo CEU University, Madrid, Spain, for their advice and expertize in this study, as well as Tomás Clive Barker Tejeda for his asserted comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.662792/full#supplementary-material

REFERENCES

Aceves, S. S., and Ackerman, S. J. (2009). Relationships between Eosinophilic Inflammation, Tissue Remodeling, and Fibrosis in Eosinophilic Esophagitis. *Immunol. Allergy Clin. North America* 29 (1), 197–211. doi:10.1016/j.iac.2008.10.003

Anderson, G. P. (2008). Endotyping Asthma: New Insights into Key Pathogenic Mechanisms in a Complex, Heterogeneous Disease. *The Lancet* 372 (9643), 1107–1119. doi:10.1016/s0140-6736(08)61452-x

Arita, M. (2016). Eosinophil Polyunsaturated Fatty Acid Metabolism and its Potential Control of Inflammation and Allergy. *Allergol. Int.* 65 (Suppl. l), S2–S5. doi:10.1016/j.alit.2016.05.010

Bachert, C., Sousa, A. R., Lund, V. J., Scadding, G. K., Gevaert, P., Nasser, S., et al. (2017). Reduced Need for Surgery in Severe Nasal Polyposis with Mepolizumab: Randomized Trial. *J. Allergy Clin. Immunol.* 140 (4), 1024–1031. doi:10.1016/j.jaci.2017.05.044

Bachert, C., Zhang, N., Hellings, P. W., and Bousquet, J. (2018). Endotype-driven Care Pathways in Patients with Chronic Rhinosinusitis. *J. Allergy Clin. Immunol.* 141 (5), 1543–1551. doi:10.1016/j.jaci.2018.03.004

Balgoma, D., Astudillo, A. M., Pérez-Chacón, G., Montero, O., Balboa, M. A., and Balsinde, J. (2010). Markers of Monocyte Activation Revealed by Lipidomic Profiling of Arachidonic Acid-Containing Phospholipids. *J.I.* 184 (7), 3857–3865. doi:10.4049/jimmunol.0902883

Barker-Tejeda, T. C., Bazire, R., Obeso, D., Mera-Berriatua, L., Rosace, D., Vazquez-Cortes, S., et al. (2020). Exploring Novel Systemic Biomarker Approaches in Grass-Pollen Sublingual Immunotherapy Using Omics. *Allergy*. 76 (4), 1199-1212. doi:10.1111/all.14565

Bauchau, V., and Durham, S. R. (2004). Prevalence and Rate of Diagnosis of Allergic Rhinitis in Europe. *Eur. Respir. J.* 24 (5), 758–764. doi:10.1183/09031936.04.00013904

Bennett, M., and Gilroy, D. W. (2016). Lipid Mediators in Inflammation. *Microbiol. Spectr.* 4 (6), 4. doi:10.1128/microbiolspec.MCHD-0035-2016

Bousquet, J., Khaltaev, N., Cruz, A. A., Denburg, J., Fokkens, W. J., Togias, A., et al. (2008). Allergic Rhinitis and its Impact on Asthma (ARIA) 2008 Update (In Collaboration with the World Health Organization, GA(2)LEN and AllerGen). *Allergy* 63 Suppl 86, 8–160. doi:10.1111/j.1398-9995.2007.01620.x

Brescia, G., Alessandrini, L., Giacomelli, L., Parrino, D., Zanotti, C., Tealdo, G., et al. (2020). A Classification of Chronic Rhinosinusitis with Nasal Polyps Based on Structured Histopathology. *Histopathology* 76 (2), 296–307. doi:10.1111/his.13969

Brescia, G., Alessandrini, L., and Marioni, G. (2021). Structured Histopathology for Endotyping and Planning Rational Treatment in Chronic Rhinosinusitis. *Am. J. Otolaryngol.* 42 (1), 102795. doi:10.1016/j.amjoto.2020.102795

Brescia, G., Barion, U., Zanotti, C., Giacomelli, L., Martini, A., and Marioni, G. (2017). The Prognostic Role of Serum Eosinophil and Basophil Levels in Sinonasal Polyposis. *Int. Forum Allergy Rhinol.* 7 (3), 261–267. doi:10.1002/alr.21885

Brescia, G., Zanotti, C., Parrino, D., Barion, U., and Marioni, G. (2018). Nasal Polyposis Pathophysiology: Endotype and Phenotype Open Issues. *Am. J. Otolaryngol.* 39 (4), 441–444. doi:10.1016/j.amjoto.2018.03.020

Chen, S., Zhou, A., Emmanuel, B., Thomas, K., and Guiang, H. (2020). Systematic Literature Review of the Epidemiology and Clinical Burden of Chronic Rhinosinusitis with Nasal Polyposis. *Curr. Med. Res. Opin.* 36 (11), 1897–1911. doi:10.1080/03007995.2020.1815682

Chiurchiù, V., Leuti, A., MaccarroneBioactive Lipids, M., and Inflammation, Chronic. (2018). Managing the Fire within. *Front. Immunol.* 9, 38. doi:10.3389/fimmu.2018.00038

Chojnowska, S., Kępka, A., Waszkiewicz, N., Zp, K., Duchnowska, E., Ościłowicz, K., et al. (2013). Etiopathogenesis of Nasal Polyps. *Prog. Heal Sci.* 3 (2), 151–159.

Ciborowski, M., Javier Rupérez, F., Martínez-Alcázar, M. P., Angulo, S., Radziwon, P., Olszanski, R., et al. (2010). Metabolomic Approach with LC–MS Reveals Significant Effect of Pressure on Diver's Plasma. *J. Proteome Res.* 9 (8), 4131–4137. doi:10.1021/pr100331j

Comhair, S. A. A., McDunn, J., Bennett, C., Fettig, J., Erzurum, S. C., and Kalhan, S. C. (2015). Metabolomic Endotype of Asthma. *J.I.* 195 (2), 643–650. doi:10.4049/jimmunol.1500736

Crestani, E., Harb, H., Charbonnier, L.-M., Leirer, J., Motsinger-Reif, A., Rachid, R., et al. (2020). Untargeted Metabolomic Profiling Identifies Disease-specific Signatures in Food Allergy and Asthma. *J. Allergy Clin. Immunol.* 145 (3), 897–906. doi:10.1016/j.jaci.2019.10.014

Dudzik, D., Barbas-Bernardos, C., García, A., and Barbas, C. (2018). Quality Assurance Procedures for Mass Spectrometry Untargeted Metabolomics. A reviewJournal of Pharmaceutical and Biomedical Analysis.a review, *J. Pharm. Biomed. Anal.* 147, 149–173. doi:10.1016/j.jpba.2017.07.044

Erbek, S. S., Erbek, S., Topal, O., and Cakmak, O. (2007). The Role of Allergy in the Severity of Nasal Polyposis. *Am. J. Rhinology* 21 (6), 686–690. doi:10.2500/ajr.2007.21.3062

Ferreira Couto, L. G., Fernades, A. M., Brandão, D. F., de Santi Neto, D., Pereira Valera, F. C., and Anselmo-Lima, W. T. (2008). Histological Aspects of Rhinosinusal Polyps. *Braz. J. Otorhinolaryngol.* 74 (2), 207–212. doi:10.1016/s1808-8694(15)31090-9

Fidan, V., Alp, H. H., Kalkandelen, S., and Cingi, C. (2013). Melatonin and Cortisol Rhythm in Patients with Extensive Nasal Polyposis. *Am. J. Otolaryngol.* 34 (1), 61–64. doi:10.1016/j.amjoto.2012.09.001

Fokkens, W. J., Lund, V. J., Mullol, J., Bachert, C., Alobid, I., Baroody, F., et al. (2012). EPOS 2012: European Position Paper on Rhinosinusitis and Nasal Polyps 2012. A Summary for Otorhinolaryngologists. *Rhin* 50 (1), 1–12. doi:10.4193/rhino50e2

Georgy, M. S., and Peters, A. T. (2012). Chapter 7: Nasal Polyps. *Allergy Asthma Proc.* 33 (Suppl. 1), S22–S23. doi:10.2500/aap.2012.33.3537

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

Godzien, J., Alonso-Herranz, V., Barbas, C., and Armitage, E. G. (2015). Controlling the Quality of Metabolomics Data: New Strategies to Get the Best Out of the QC Sample. *Metabolomics* 11 (3), 518–528. doi:10.1007/s11306-014-0712-4

Grill, M., Högenauer, C., Blesl, A., Haybaeck, J., Golob-Schwarzl, N., Ferreirós, N., et al. (2019). Members of the Endocannabinoid System Are Distinctly Regulated in Inflammatory Bowel Disease and Colorectal Cancer. *Sci. Rep.* 9 (1), 2358. doi:10.1038/s41598-019-38865-4

Hopkins, C. (2019). "*Chronic Rhinosinusitis with Nasal Polyps*,"Editor C. G. Solomon, N. Engl., and J. Med 381 (1), 55–63. doi:10.1056/nejmcp1800215

Kim, D. E., Lee, Y., Kim, M., Lee, S., Jon, S., and Lee, S.-H. (2017). Bilirubin Nanoparticles Ameliorate Allergic Lung Inflammation in a Mouse Model of Asthma. *Biomaterials* 140, 37–44. doi:10.1016/j.biomaterials.2017.06.014

Kim, D. W., and Cho, S. H. (2017). Emerging Endotypes of Chronic Rhinosinusitis and its Application to Precision Medicine. *Allergy Asthma Immunol. Res.* 9 (4), 299. doi:10.4168/aair.2017.9.4.299

Kong, I. G., and Kim, D. W. (2018). Pathogenesis of Recalcitrant Chronic Rhinosinusitis: The Emerging Role of Innate Immune Cells. *Immune Netw.* 18 (2), 1–12. doi:10.4110/in.2018.18.e6

Kook, J. H., Kim, H. J., Kim, K. W., Park, S. J., Kim, T. H., Lim, S. H., et al. (2015). The Expression of 11β-Hydroxysteroid Dehydrogenase Type 1 and 2 in Nasal Polyp-Derived Epithelial Cells and its Possible Contribution to Glucocorticoid Activation in Nasal Polyp. *Am. J. Rhinol,allergy* 29 (4), 246–250. doi:10.2500/ajra.2015.29.4185

Kırıs, M., Muderris, T., Yalçıner, G., Bercin, S., Sevil, E., and Gul, F. (2016). Intrapolyp Steroid Injection for Nasal Polyposis: Randomized Trial of Safety and Efficacy. *Laryngoscope* 126 (8), 1730–1735. doi:10.1002/lary.25945

Lu, Y., Wang, Y., Ong, C.-N., Subramaniam, T., Choi, H. W., Yuan, J.-M., et al. (2016). Metabolic Signatures and Risk of Type 2 Diabetes in a Chinese Population: an Untargeted Metabolomics Study Using Both LC-MS and GC-MS. *Diabetologia* 59 (11), 2349–2359. doi:10.1007/s00125-016-4069-2

Lund, V. J. (1995). Fortnightly Review: Diagnosis and Treatment of Nasal Polyps. *BMJ* 311 (7017), 1411–1414. doi:10.1136/bmj.311.7017.1411

Obeso, D., Mera-Berriatua, L., Rodríguez-Coira, J., Rosace, D., Fernández, P., Martín-Antoniano, I. A., et al. (2018). Multi-omics Analysis Points to Altered Platelet Functions in Severe Food-Associated Respiratory Allergy. *Allergy* 73 (11), 2137–2149. doi:10.1111/all.13563

Pearlman, A. N., Chandra, R. K., Conley, D. B., and Kern, R. C. (2010). "Epidemiology of Nasal Polyps," in *Nasal Polyposis*. Editors T. M. Önerci and B. J. Ferguson (Heidelberg, Berlin: Springer), 9–15. doi:10.1007/978-3-642-11412-0_2

Peng, Y.-F., Goyal, H., and Xu, G.-D. (2017). Serum Bilirubin Has an Important Role in Multiple Clinical Applications. *J. Lab. Precis. Med.* 2, 82. doi:10.21037/jlpm.2017.09.08

Pothoven, K. L., Norton, J. E., Suh, L. A., Carter, R. G., Harris, K. E., Biyasheva, A., et al. (2017). Neutrophils Are a Major Source of the Epithelial Barrier Disrupting Cytokine Oncostatin M in Patients with Mucosal Airways Disease. *J. Allergy Clin. Immunol.* 139 (6), 1966–1978. doi:10.1016/j.jaci.2016.10.039

Ried, J. S., Baurecht, H., Stückler, F., Krumsiek, J., Gieger, C., Heinrich, J., et al. (2013). "Integrative Genetic and Metabolite Profiling Analysis Suggests Altered Phosphatidylcholine Metabolism in Asthma,". Editor H-U. Simon and Allergy, 68 (5), 629–636. doi:10.1111/all.12110

Rosace, D., Gomez-Casado, C., Fernandez, P., Perez-Gordo, M., Dominguezdel, M. d. C. C., Vega, A., et al. (2019). Profilin-mediated Food-Induced Allergic Reactions Are Associated with Oral Epithelial Remodeling. *J. Allergy Clin. Immunol.* 143 (2), 681–690. doi:10.1016/j.jaci.2018.03.013

Samitas, K., Carter, A., Kariyawasam, H. H., and Xanthou, G. (2018). Upper and Lower Airway Remodelling Mechanisms in Asthma, Allergic Rhinitis and Chronic Rhinosinusitis: The One Airway Concept Revisited. *Allergy* 73 (5), 993–1002. doi:10.1111/all.13373

Sanchez-Solares, J., Delgado-Dolset, M. I., Mera-Berriatua, L., Hormias-Martin, G., Cumplido, J. A., Saiz, V., et al. (2019). Respiratory Allergies with No Associated Food Allergy Disrupt Oral Mucosa Integrity. *Allergy*. 74 (11), 2261-2265. doi:10.1111/all.13860

Schleimer, R. P. (2017). Immunopathogenesis of Chronic Rhinosinusitis and Nasal Polyposis. *Annu. Rev. Pathol. Mech. Dis.* 12 (1), 331–357. doi:10.1146/annurev-pathol-052016-100401

Sun, C., Ouyang, H., and Luo, R. (2017). Distinct Characteristics of Nasal Polyps with and without Eosinophilia. *Braz. J. Otorhinolaryngol.* 83 (1), 66–72. doi:10.1016/j.bjorl.2016.01.012

Tos, M., Larsen, P. L., Larsen, K., and Caye-Thomasen, P. (2010). "Pathogenesis and Pathophysiology of Nasal Polyps," in *Nasal Polyposis*. Editors T. M. Önerci and B. J. Ferguson (Heidelberg, Berlin: Springer), 53–63. doi:10.1007/978-3-642-11412-0_7

Van Zele, T., Holtappels, G., Gevaert, P., and Bachert, C. (2014). Differences in Initial Immunoprofiles between Recurrent and Nonrecurrent Chronic Rhinosinusitis with Nasal Polyps. *Am. J. Rhinol. allergy* 28 (3), 192–198. doi:10.2500/ajra.2014.28.4033

Villaseñor, A., Rosace, D., Obeso, D., Pérez-Gordo, M., Chivato, T., Barbas, C., et al. (2017). Allergic Asthma: an Overview of Metabolomic Strategies Leading to the Identification of Biomarkers in the Field. *Clin. Exp. Allergy* 47 (4), 442–456. doi:10.1111/cea.12902

Vogel, M. E., and Zucker, S. D. (2016). Bilirubin Acts as an Endogenous Regulator of Inflammation by Disrupting Adhesion Molecule-Mediated Leukocyte Migration. *Inflamm. Cel Signal* 3 (1), e1178. doi:10.14800/ics.1178

Wilson, K. F., McMains, K. C., and Orlandi, R. R. (2014). The Association between Allergy and Chronic Rhinosinusitis with and without Nasal Polyps: an Evidence-Based Review with Recommendations. *Int. Forum Allergy Rhinology* 4 (2), 93–103. doi:10.1002/alr.21258

Wynn, R., and Har-El, G. (2004). Recurrence Rates after Endoscopic Sinus Surgery for Massive Sinus Polyposis. *The Laryngoscope* 114 (5), 811–813. doi:10.1097/00005537-200405000-00004

Yang, H. H., Fang, H., You, Q. J., Han, L., Yang, Z. F., Yu, L. L., et al. (2018). [Predictive Study on Recurrence of Chronic Sinusitis with Nasal Polyps by Tissue Eosinophils and Sinus CT]. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* 53 (11), 842–846. doi:10.3760/cma.j.issn.1673-0860.2018.11.009

Yeager, M. P., Guyre, C. A., Sites, B. D., Collins, J. E., Pioli, P. A., and Guyre, P. M. (2018). The Stress Hormone Cortisol Enhances Interferon-υ-Mediated Proinflammatory Responses of Human Immune Cells. *Anesth. Analgesia* 127 (2), 556–563. doi:10.1213/ane.0000000000003481

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 14 May 2021
doi: 10.3389/fmolb.2021.682559

Gaining Insights Into Metabolic Networks Using Chemometrics and Bioinformatics: Chronic Kidney Disease as a Clinical Model

Julien Boccard[1,2,], Domitille Schvartz[3], Santiago Codesido[1,2], Mohamed Hanafi[4], Yoric Gagnebin[1,2], Belén Ponte[5], Fabien Jourdan[6] and Serge Rudaz[1,2]*

[1] School of Pharmaceutical Sciences, University of Geneva, Geneva, Switzerland, [2] Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, Geneva, Switzerland, [3] Translational Biomarker Group, Department of Internal Medicine Specialties, University of Geneva, Geneva, Switzerland, [4] Unité Statistique, Sensométrie et Chimiométrie, Nantes, France, [5] Service of Nephrology and Hypertension, Department of Medicine, Geneva University Hospitals (HUG), Geneva, Switzerland, [6] Toxalim, Research Centre in Food Toxicology, Université de Toulouse, INRAE, ENVT, INP-Purpan, UPS, Toulouse, France

OPEN ACCESS

Edited by:
Zheng-Jiang Zhu, Shanghai Institute of Organic Chemistry (CAS), China

Reviewed by:
Tao Huan, University of British Columbia, Canada
Li Chen, Fudan University, China

***Correspondence:**
Julien Boccard
julien.boccard@unige.ch

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

Received: *18 March 2021*
Accepted: *19 April 2021*
Published: *14 May 2021*

Citation:
Boccard J, Schvartz D, Codesido S, Hanafi M, Gagnebin Y, Ponte B, Jourdan F and Rudaz S (2021) Gaining Insights Into Metabolic Networks Using Chemometrics and Bioinformatics: Chronic Kidney Disease as a Clinical Model. Front. Mol. Biosci. 8:682559. doi: 10.3389/fmolb.2021.682559

Because of its ability to generate biological hypotheses, metabolomics offers an innovative and promising approach in many fields, including clinical research. However, collecting specimens in this setting can be difficult to standardize, especially when groups of patients with different degrees of disease severity are considered. In addition, despite major technological advances, it remains challenging to measure all the compounds defining the metabolic network of a biological system. In this context, the characterization of samples based on several analytical setups is now recognized as an efficient strategy to improve the coverage of metabolic complexity. For this purpose, chemometrics proposes efficient methods to reduce the dimensionality of these complex datasets spread over several matrices, allowing the integration of different sources or structures of metabolic information. Bioinformatics databases and query tools designed to describe and explore metabolic network models offer extremely useful solutions for the contextualization of potential biomarker subsets, enabling mechanistic hypotheses to be considered rather than simple associations. In this study, network principal component analysis was used to investigate samples collected from three cohorts of patients including multiple stages of chronic kidney disease. Metabolic profiles were measured using a combination of four analytical setups involving different separation modes in liquid chromatography coupled to high resolution mass spectrometry. Based on the chemometric model, specific patterns of metabolites, such as N-acetyl amino acids, could be associated with the different subgroups of patients. Further investigation of the metabolic signatures carried out using genome-scale network modeling confirmed both tryptophan metabolism and nucleotide interconversion as relevant pathways potentially associated with disease severity. Metabolic modules composed of chemically adjacent or close compounds of biological relevance were further investigated using carbon transfer reaction paths. Overall, the proposed integrative data analysis strategy allowed deeper insights into the

metabolic routes associated with different groups of patients to be gained. Because of their complementary role in the knowledge discovery process, the association of chemometrics and bioinformatics in a common workflow is therefore shown as an efficient methodology to gain meaningful insights in a clinical context.

Keywords: metabolomics, chemometrics, bioinformatics, integrative data analysis, chronic kidney disease, metabolic networks

INTRODUCTION

While efforts are still being made to improve both technological and computational aspects, metabolomics is now recognized as an essential approach to assess biochemical phenotypes in many application fields, including clinical research. Mass spectrometry (MS) has established itself as a major analytical detection technique by offering high sensitivity and substantial throughput (Zhang et al., 2020). Metabolomic experiments often generate large amounts of high-dimensional and complex biochemical data involving multiple signals measured from thousands of low molecular weight compounds. Dedicated strategies need thus to be applied to extract meaningful biological knowledge from the collected MS data (Boccard et al., 2010). Despite the considerable developments made to improve the different steps of the workflow (Pezzatti et al., 2020), assessing the metabolic diversity of a complex sample still constitutes a challenging analytical endeavor. The difficulty is mainly due to the large chemical space and concentration ranges covered by metabolites characterizing biological systems (Frainay et al., 2018). The integration of data collected from different sample preparation protocols, separation principles, ionization modes or analytical platforms has been recognized as an efficient strategy to improve the metabolome coverage of complex samples, thus potentially offering better understanding of the underlying biological mechanisms associated with a given phenotypic pattern (Richards et al., 2010). Dedicated data mining tools accounting adequately for metabolomic signals spread over multiple data tables are therefore needed, and chemometrics offers potent solutions for data integration based on dimensionality reduction approaches. More specifically, multivariate models able to handle multiple blocks of variables (multiblock) associated with different groups of observations (multigroup) are now established as effective methods for data integration in omics disciplines (Boccard and Rudaz, 2014).

An additional complexity frequently encountered in clinical research is due to a certain degree of heterogeneity in sample collection among different groups of individuals. Classical case–control studies usually involve a group of healthy volunteers used as a reference and compared to one or various pathological situations. On one hand, it could be difficult to collect measurements obtained by a highly invasive technique for the control group. On the other hand, a longitudinal follow-up of critical patients involving repeated measures at different time points may be required in standard protocols. This temporal follow-up offers rich information, but these types of longitudinal setups generate multiway data, i.e., three-dimensional tensors in this case (individuals × variables × time) making the global analysis of all available data more challenging.

In this context, Network Principal Component Analysis (NetPCA) has been recently proposed as a novel and generic approach to handle any type of structure composed of several data matrices (Codesido et al., 2020). Relationships between groups of observations or blocks of variables are translated into a network structure, where the nodes are standard two-dimensional data matrices. Two types of edges link the nodes to connect data tables characterized by the same observations or variables. This formalism translates any links between data matrices into optimization constraints to find principal directions of covariations by using the same model parameters (i.e., coefficients or loadings), that are then extracted using a set of linear models.

Handling data spread over multiple matrices constitutes a crucial step toward insightful data integration, but biological information is rarely obtained directly from statistical models, e.g., based on variables coefficients, or selected subsets of annotated metabolites. It is now well-recognized that a contextualization of metabolomic results is mandatory to go beyond simple associations and provide reliable mechanistic hypotheses (Kell, 2004). Converting subsets of up- or down-modulated biomarkers into biological processes and functions plays thus another critical role to go toward a functional description of the molecular events under study (Booth et al., 2013).

A first approach to this aim is to derive a biological meaning from a subset of relevant metabolites by retrieving metadata associated with each compound, such as chemical classes or known metabolic pathways. This can be done using a controlled vocabulary, i.e., ontologies allowing functional annotations, and bioinformatic tools that are designed to query this information stored in dedicated databases, e.g., KEGG (Kanehisa and Goto, 2000) and HMDB (Wishart et al., 2018). Biological processes and/or molecular functions can then be ranked using a statistical test (e.g., Fisher exact test or hypergeometric test) according to their probability to be represented more frequently than it would occur by random chance (Kankainen et al., 2011). The rationale behind this strategy is that the metabolites belonging to a metabolic pathway involved in the manifestation of a specific metabolic signature are expected to be modulated simultaneously. This is a rather strong hypothesis that may not be fulfilled in a real biological context and this type of analysis has therefore some limitations (Marco-Ramell et al., 2018).

Alternatively, describing metabolic networks as graphs constitutes a very efficient methodology to model biochemical reactions defining the metabolism. A metabolite-centric

representation can be gained using a compound graph defining metabolites as nodes, which are connected if they are substrate and product of the same reaction (Guimera and Amaral, 2005). Metabolic pathways can then be defined as subgraphs involving a series of metabolites belonging to the same metabolic process. Such a strategy allows to go much further in the biological interpretation of altered metabolic patterns, compared to standard metabolic pathways over-representation analysis. As enzymes drive most of these reactions, they can be related to specific proteins, which are synthesized from their corresponding genes. These biological links between genes and metabolites constitute valuable information that can be used to infer metabolic networks from genomes, following a systems biology approach. Major improvements of the sequencing technology and databases have indeed enabled the reconstruction of genome-scale metabolic networks of several model organisms (Yilmaz and Walhout, 2017). It is, however, to be noted that network curation and validation remain mandatory to correct for false positive or missing reactions, and guarantee an adequate stoichiometric balance. Various metrics can then be applied to highlight specific topological features of interest, such as path lengths, degree centrality and clustering coefficient, which allow hubs and modules of topological importance to be highlighted (Lacroix et al., 2008). Moreover, subnetworks focusing on specific sets of reactions related to modulated metabolites can be easily extracted and visualized (Frainay and Jourdan, 2017). Such a strategy is useful to reduce the complexity of large networks and to gain a better mechanistic understanding of the phenomenon under study, through the possibility to investigate a reduced list of biochemical reactions characterizing particular metabolic phenotypes. Resources for pathway mapping and/or network analysis include KEGG (Kanehisa and Goto, 2000), MetaCyc (Caspi et al., 2016), Recon (Thiele et al., 2013; Noronha et al., 2017), and MetExplore (Cottret et al., 2010, 2018).

This work presents the integrative analysis of data collected from several metabolomic studies designed to investigate the effects of chronic kidney disease (CKD). The selected analytical strategy was based on multiple chromatographic separation setups, including reversed-phase liquid chromatography (RPLC), hydrophilic interaction chromatography with amide (aHILIC), and polymeric zwitterionic stationary phases (ZICpHILIC) coupled to high-resolution mass spectrometry (HRMS). The heterogeneous and complex data structure generated was successfully handled using chemometrics for multiblock and multiset data integration, and state-of-the-art bioinformatic resources were implemented for an in-depth study of metabolic events.

MATERIALS AND METHODS

Chronic Kidney Disease Dataset

Plasma samples collected during four prospective observational monocentric CKD studies performed at a tertiary hospital (Geneva University Hospitals, Geneva, Switzerland) were considered. CKD severity stages were defined according to the glomerular filtration rate (GFR) criterion KDIGO guidelines (Levey et al., 2020). A first cohort of patients was recruited to assess the impact of different stages of CKD severity (3b–5) on plasma metabolites (Gagnebin et al., 2019). A second and third one explored the benefits offered by kidney transplantation, as well as the potential impact of kidney donation on healthy living donors (LKD) (Gagnebin et al., 2020b). Finally, the last one was designed to investigate plasmatic metabolites patterns of patients with end-stage renal disease (ESRD) undergoing regular hemodialysis (HD) (Gagnebin et al., 2020a). The latter were under chronic HD for at least 3 months with three dialysis per week. In addition, a control group of healthy volunteers was also included.

Plasma from all cohorts were collected for an integrative analysis of these different renal conditions. All samples were collected in the morning after an overnight fast and several times points were considered for individuals undergoing HD or transplantation, as well as LKD. Samples were directly thawed, aliquoted and stored at −80°C. More details about the inclusion criteria, GFR, HD characteristics and ethical concerns can be found in the specific reference of each study.

Sample Preparation and Analysis

Sample Preparation

Pipetting and liquid handling was carried out using a Tecan Freedom Evo-2 (Tecan, Switzerland) to ensure sample preparation repeatability. First, solvent protein precipitation was performed using cold methanol spiked with isotopically labeled standards [(d_5-indole)-L-tryptophan, 1,2-$^{13}C_2$-taurine and 2,2,4-d_3-DL-glutamic acid)], all from Cambridge Isotope Laboratories Inc. (Andover, United States) at 1.25 μg mL-1. A volume of 960 μL cold methanol containing the standards was added to 240 μL of thawed plasma. Samples were then vortexed for 20 s, mixed at 1,200 rpm for 30 min at 4°C and centrifuged at 15,000 × g for 20 min at 4°C. Supernatants were divided into 300 μL aliquots and dispatched on 96-well plates. Each extract was then dried for 8 h in a Thermo Fisher Scientific Savant 210A SpeedVac (Thermo Electron LED GmbH, Langenselbold, Germany) and stored at −80°C. Before analysis, samples were reconstituted in 60 μL of H_2O/MeCN (95:5, v/v) for RPLC and in 60 μL of H_2O:MeCN (25:75, v/v) for aHILIC and ZICpHILIC. Quality control (QC) and diluted QC (dQC) samples were prepared to assess and correct analytical variations using the same sample preparation procedures as the individual samples.

LC-MS Analysis

Liquid chromatography was carried out on a Waters H-Class Acquity UPLC system (Waters Corporation, Milford, MA, United States) using different separation modes. RPLC analysis was performed using a Phenomenex Kinetex C18 column (150 × 2.1 mm, 1.7 μm), and a SecurityGuard ULTRA pre-column. A gradient of mobile phase A (0.1% FA in water) and mobile phase B (0.1% FA in MeCN) was applied as follows: 2% B for 1 min, increased to 100% B over 14 min, held for 3 min, and then returned to 2% B to re-equilibrate the column for 7 min (total run time of 25 min) at a flow rate of 300 μL min-1 and a column temperature of 30°C. Separation using aHILIC was achieved on a Waters Acquity BEH Amide column

(150 × 2.1 mm, 1.7 μm), and a VanGuard^{TM} pre-column. A gradient of mobile phase A (H_2O:MeCN; 5:95 v/v) and mobile phase B (10 mM ammonium formate in H_2O:MeCN, 70:30 v/v adjusted at pH 6.50 in the aqueous part) was applied as follows: 0% B for 2 min, increased to 70% B over 18 min, held for 3 min, and then returned to 0% B to re-equilibrate the column for 7 min (total run time of 31 min) at a flow rate of 500 μL min-1 and a column temperature of 40°C.

ZICpHILIC analysis was carried out on a Merck SeQuant ZicpHILIC column (150 × 2.1 mm, 5 μm) and the appropriate guard kit was applied. A gradient of mobile phase A (MeCN) and mobile phase B (2.8 mM ammonium formate adjusted to pH 9.00) was applied as follows: 5% B for 1 min, increased to 51% B over 9 min, held for 3 min at 51% B and then returned to 5% B in 0.1 min before re-equilibrating the column for 6.9 min (total run time of 20 min) at a flow rate of 300 μL min-1 and a column temperature of 40°C. For all separation modes, a volume of 5 μL of the 393 samples was analyzed in eight batches using constrained randomization (Jonsson et al., 2015). QC samples were injected for system conditioning (15 first injections in each batch), while QCs and dQCs were analyzed regularly during the sequence (every 6 samples).

The UPLC system was coupled to a maXis 3G Q-TOF high-resolution MS from Bruker (Bruker Daltonik GmbH, Bremen, Germany) with an electrospray ionization source working in positive (ESI+) mode for RPLC and aHILIC, or negative (ESI−) mode for RPLC and ZICpHILIC. The instrument was operated using the following parameters: capillary voltage of –4.7 kV for ESI+ and 2.8 kV for ESI-, nebulizing gas pressure of 2.0 bar, drying gas temperature of 225°C for RPLC and 200°C for aHILIC and ZICpHILIC and a flow rate of 8.0 L min^{-1}. Data acquisition between 50 and 1,000 m/z was performed in profile mode at a rate of 2 Hz. In-run automatic calibration was achieved using formate adducts in the 90–1,247 m/z range and the quadratic plus high-precision calibration algorithm (Bruker Daltonics). The detailed analytical protocol can be found elsewhere (Gagnebin et al., 2019).

Data Processing and Analysis

Raw data processing was performed using Progenesis QI 2.3 (Non-linear Dynamics, Waters, Newcastle upon Tyne, United Kingdom). QCs and dQCs were used to monitor and control data acquisition quality, remove unreliable signals, and correct for within-batch drifts and between-batch effects. A filtering procedure was carried out to remove unreliable signals by applying a threshold of 50% for the dQC/QC ratio relative standard deviation (RSD) and a dQC/QC ratio between 0.2 and 0.8. LOESS regression was used for intra- and inter-batch normalization based on QC samples. Finally, the position of the QCs was assessed separately for each data block using Principal Component Analysis (data not shown).

Metabolite annotation was achieved using an in-house database containing experimental data from more than 900 authentic standard compounds measured in different chromatographic conditions. Briefly, level 1 annotation was achieved by matching *m/z* values, retention times, and isotopic patterns, while MS/MS spectra and collisional cross-section values were considered for confirmatory purpose. Cytoscape 3.8.2 was used to generate circular layout graphs. NetPCA was computed after unit variance scaling using the NetPCA Python package (Codesido et al., 2020).

Clustering was computed under the MATLAB 9.5 environment (The MathWorks, Natick, United States). MetExplore (Cottret et al., 2010) and MetExploreViz (Chazalviel et al., 2018) were used with the Recon3D human metabolic reconstruction (Brunk et al., 2018) for mapping identified metabolites, pathway over-representation analysis, network visualization and evaluation of carbon transfer reaction paths. A more detailed description of the data processing and analysis workflow, as well as the dataset, are available as **Supplementary Material**.

RESULTS

Metabolomic Dataset Structure

In total, 393 blood samples were collected from the pre-defined cohorts of patients grouped into different categories according to their renal status: 56 healthy control volunteers (CTRL), 69 CKD patients at intermediate stage (ICKD), 35 patients with ESRD undergoing HD (HD), 42 ESRD undergoing kidney graft (KG) and 24 healthy LKD (DV). Repeated measures were carried out for three groups: (i) HD: before (preHD) and after (postHD) blood dialysis on the mid-week session, (ii) KG: before (preKG), 1 week (postKG1), and 1 month (postKG2) after the graft, (iii) DV: before (preDV), 1 week (postDV1) and 1 year (postDV2) after the transplantation procedure.

All samples went through our standard analytical workflow already presented elsewhere (Pezzatti et al., 2020). Briefly, generic sample preparation was first carried out to cover a large chemical diversity of compounds and each sample was analyzed using four LC-HRMS setups, i.e., RPLC+, RPLC−, aHILIC+, and ZICpHILIC-. Raw data processing involving baseline correction, peak picking, adduct deconvolution and retention time alignment was carried out with Progenesis QI, while subsequent filtering and normalization based on QC samples (Broadhurst et al., 2018; Pezzatti et al., 2020) was performed using in house scripts. Metabolite annotation was then performed by matching reference values measured with standards in the same analytical conditions, and a scoring strategy was implemented to remove compounds annotated in more than one LC-MS setup. For that purpose, a peak quality score based on intensity, shape, and retention time was used to select the best analytical information (Pezzatti et al., 2019). In the present work, annotation was restricted to metabolites matching entries from our in-house database. This strategy led to a set of 218 compounds annotated at level 1 for each sample. These identified metabolites were spread over four blocks of variables corresponding to the different LC-MS setups as follows: 88 in RPLC+, 38 in RPLC−, 27 in aHILIC+, and 65 in ZICpHILIC-. A schematic diagram of the dataset structure composed of 40 data matrices involving the five different groups of individuals, the four blocks of variables,

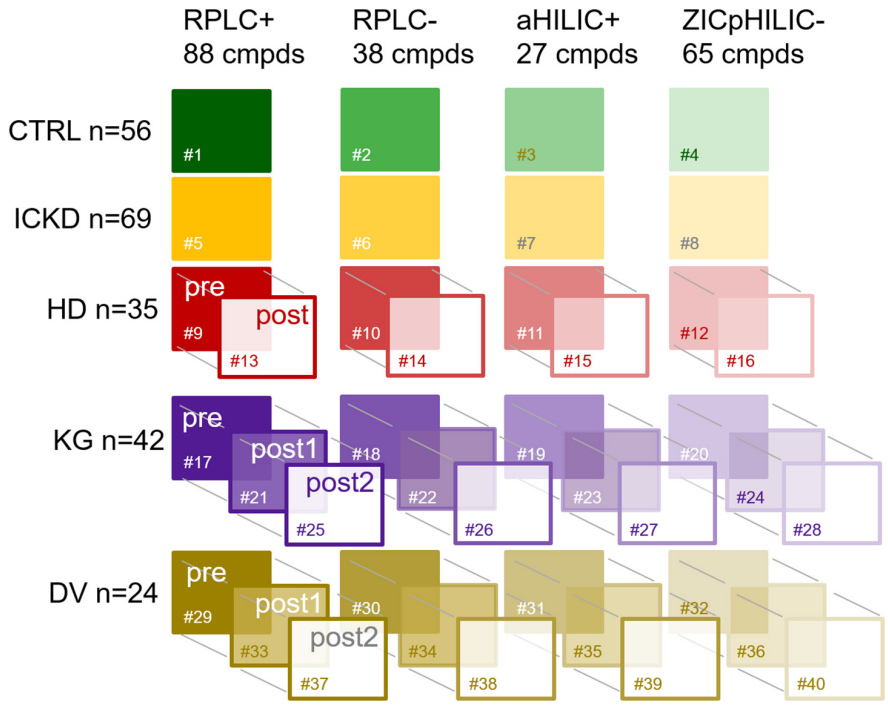

FIGURE 1 | Schematic diagram of the dataset structure composed of 40 data matrices involving the five different groups of individuals, the four blocks of variables, and repeated measurements. CTRL, control group; ICKD, intermediate chronic kidney disease; HD, hemodialysis; KG, kidney graft; DV, living kidney donors. Sharps indicate the numbering of the tables.

and repeated measurements for HD, KT, and KD is provided in **Figure 1**.

Network Principal Component Analysis Modeling

This type of complex data structure involving multiple groups of observations and blocks of variables, together with repeated measures for certain individuals, is not straightforward to handle efficiently, i.e., without breaking connections between data matrices linking their rows and/or columns. NetPCA was performed after unit variance scaling using connection links between data matrices to define the constraints of the model. For that purpose, a network incidence matrix was used to define the topology of the data structure. By these means, matchings between the same sets of observations (rows), variables (columns), or both (rows and columns) could be explicitly included in the optimization process, thus leading to global components accounting for these links. Circular layout graphs corresponding to the connections between (A) groups of observations and (B) blocks of variables are presented in **Figure 2**. Using this representation, blocks of data associated with the same individuals appear clearly (data matrices #1–4 for CTRL and #5–8 for ICKD), while tensorial substructures generated by repeated measures during the longitudinal follow-ups form larger clusters (data matrices #9–16 for HD, #17–28 for KG, and #29–40 for DV). Moreover, the four subsets of metabolites measured using the different LC-MS setups were also clearly visible.

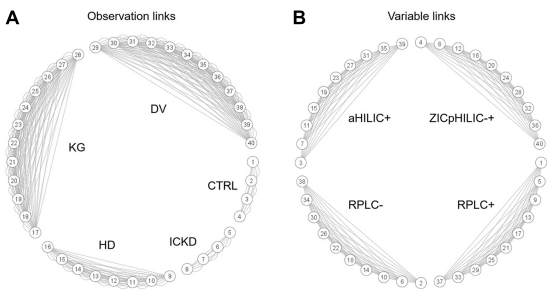

FIGURE 2 | Circular layout graphs describing network connections between **(A)** groups of samples and **(B)** blocks of variables. The numbering of the data matrices corresponds to that given in **Figure 1**.

A two-component NetPCA model was considered to investigate the main metabolic variations in the dataset. From the distribution of the groups on the score plot (**Figure 3**), it appeared that the first component, summarizing 27.0% of the total variability, strongly followed the severity of CKD. The CTRL and preDV groups of healthy individuals were located on the left (negative scores), ICKD transitional situation in the middle, while preHD and preKG patients characterized by ESRD requiring specific treatment were distributed on the right (positive scores). This important part of explained variance is consistent with prior knowledge regarding the massive kidney dysfunction associated with CKD status characterized by

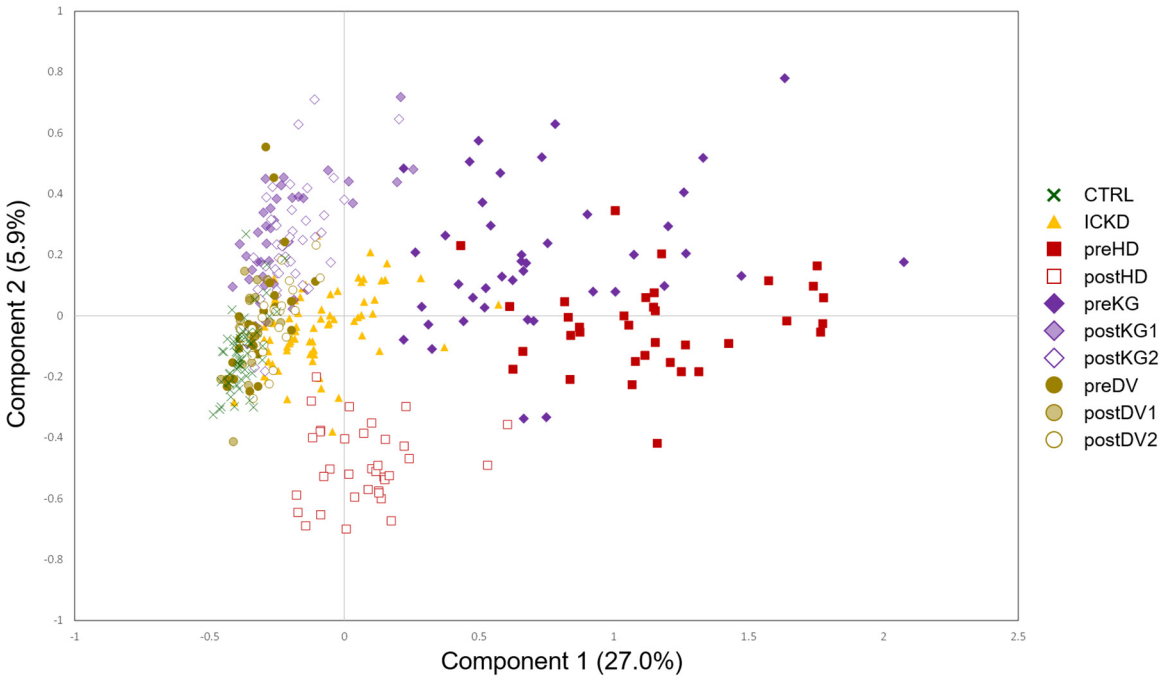

FIGURE 3 | NetPCA score plot of the two first global components. CTRL, dark green crosses; ICKD, filled orange triangles; preHD, filled red squares; postHD, empty red squares; preKG, filled violet diamonds; postKG1, light violet diamonds; postKG2, empty violet diamonds; preDV, filled brown circles; postDV1, light brown circles; postDV2, empty brown circles.

markedly decreased glomerular filtration. These characteristics are indeed shared by the preKG and preHD groups with the most severe kidney damage. It should also be noted that postHD samples are associated with a return to a situation similar to ICKD along this first component, compared to the preHD group. A similar observation can be made by comparing the samples of the KG group before (preKG) and after (postKG1 and postKG2) transplantation. As patients seem to recover (at least partially) a metabolic profile similar to less severe stages of renal disease, this suggests marked beneficial effects offered by hemodialysis and graft on many metabolite levels. In addition, kidney donation does not appear to affect the metabolomic profile of LKD over the long term, as postDV1 and postDV2 groups of samples were located close to healthy individuals on the left, with only slightly higher metabolites levels possibly resulting from reduced filtration capacity with a single kidney.

Finally, samples from the postHD group, i.e., metabolic profiles measured after dialytic therapy, are characterized by negative scores on the second component on which they appear separated from the other sample groups, revealing thus another pattern of metabolic profiles that was not observed in the case of endogenous renal alteration. It is to be noted that the second component summarized a trend driven only by a subset of samples with 5.9% of the total variability, thus more specific to the impact of hemodialysis. The NetPCA score plot of the two first global components is depicted in **Figure 3**.

Relative block influences were computed for both components, thus offering an objective way to evaluate the relative contributions of each data matrix to the global decomposition. Prior observations were confirmed, as the first component was mainly associated with variations from data blocks related to the preHD and preKG groups, while the second component summarized the trend associated with postHD samples. Interestingly, balanced contributions from the four LC-MS setups could be highlighted, underlining their overall agreement in terms of biochemical information. Relative blocks influences are summarized in **Table 1**.

Metabolite-Centric Analysis

The contributions of the variables to the components (or loadings) are helpful to investigate the trends associated with meaningful observations groupings and interpret multivariate models. Because NetPCA accounts explicitly for multigroup structures, these influences of the variables can be computed by taking only specific subsets of observations into account, i.e., subsets of metabolites associated with a specific renal status in this case. This provides also an objective basis to highlight specific differences between two groups, e.g., CTRL and postHD, by assessing potentially dissimilar patterns of compounds summarizing the variability of the two metabolic phenotypes. Based on this information, it is also possible to adopt another way of interpreting the model through a variable-centric approach. In this case, the contribution of each metabolite to the explained variability can be summarized using the different components of the model. A cumulative contribution for each component, according to its relative part of variability explained can be expressed as a percentage of the total variance of the metabolite. Ranking the variables according to their percentage

TABLE 1 | Relative block influences for the two first NetPCA components.

	NetPCA Component 1				NetPCA Component 2			
Group	**RPLC+**	**RPLC−**	**aHILIC+**	**ZICpHILIC−**	**RPLC+**	**RPLC−**	**aHILIC+**	**ZICpHILIC−**
CTRL	2.6%	2.6%	2.4%	2.3%	1.4%	1.5%	1.0%	1.6%
ICKD	0.3%	0.3%	0.3%	0.1%	0.1%	0.1%	0.1%	0.3%
preHD	13.3%	13.7%	13.1%	15.1%	0.1%	0.1%	0.0%	0.0%
postHD	0.0%	0.1%	0.1%	0.1%	11.2%	12.0%	14.0%	15.7%
preKG	4.6%	6.4%	2.2%	7.6%	2.3%	1.7%	0.7%	3.3%
postKG1	0.8%	0.8%	0.7%	0.4%	3.8%	4.1%	4.2%	4.0%
postKG2	0.5%	0.5%	0.4%	0.4%	3.8%	3.6%	3.5%	4.7%
preDV	0.7%	0.8%	0.7%	0.6%	0.1%	0.1%	0.1%	0.0%
postDV1	0.9%	0.9%	0.9%	0.8%	0.2%	0.2%	0.2%	0.1%

of variance explained for a given component of interest allows a straightforward extraction of the biological variability of interest. Explicitly accounting for the data structure provides thus a better understanding of the role of each variable in the decomposition. The percentage of explained variance for the 218 metabolites is summarized in **Supplementary Table 1**. A threshold of 20% was applied in order to underline relevant metabolites accounting for at least this proportion of their total variability on the different axes. This led to a large subset of 106 compounds markedly modulated according to disease severity on the first component, while a smaller pattern composed of 14 metabolites was associated with the second axis. From these results, it can be observed that a massive increase of the abundance of a large number of metabolites is associated with first component. This general accumulation of blood metabolites is in line with prior knowledge of the pathology (Zhao, 2013).

Levels of known uremic retention solutes, such as creatinine, adipate, allantoin, or hippuric acid were observed as positively correlated with the first component, i.e., levels that were increased according to disease severity (Boelaert et al., 2013). Moreover, numerous N-acetyl amino acids were also in the subset of metabolites with markedly augmented abundances, including N-acetylmethionine, N-acetyltryptophan, N-acetylphenylalanine, N-acetylleucine, N-acetylproline, N-acetyllysine, N-asparagine. These may be the hallmark of N-acetylation as an altered detoxification mechanism in CKD (Sekula et al., 2016). Other noticeably increased metabolites were kynurenic acid, formylmethionine, anthranilic acid, 5'-methylthioadenosine, myo-inositol and indoxyl sulfate. Tryptophan and guanidinoacetic acid (a precursor of creatine, creatinine and urea) were characterized by a decrease level when compared to healthy volunteers with fully efficient kidney function. Overall, these results were in agreement with previous studies (Hocher and Adamski, 2017; Kalim and Rhee, 2017).

Additionally, a subset of metabolites was highlighted as characteristic of the dialyzed patients (postHD), on the second component, involving various amino acids, such as lysine, arginine, methionine, proline, threonine, cysteine, valine and norvaline. Interpretation remains, however, challenging, as this observation could result from lower overall metabolite concentration and reduced ion suppression that may increase signal (Gagnebin et al., 2017). Moreover, HD patients were under chronic hemodialysis for at least 3 months with standard three dialysis per week, and this effect could be due to a cyclic equilibria of metabolite levels, as discussed with the medical team. An evaluation of the net balance would require quantitative plasmatic measurements as well as proper assessment of clearance.

Bioinformatics

In order to go beyond simple lists of potential biomarkers and propose biological hypotheses, further analysis of metabolic pathways was carried out using MetExplore (Cottret et al., 2010). The latter uses biosources, i.e., curated metabolic networks obtained from genome-scale reconstructions, to offer deeper insights into metabolic modulations and a better understanding of potential mechanisms leading to a specific metabolic phenotype. The Recon3D (Brunk et al., 2018) biosource of *Homo sapiens* derived from 2,990 genes was used as the most accurate reconstruction of the human metabolic network to date. Recon3D topology includes the cellular localization of metabolites (e.g., mitochondria, cytoplasm, etc.) but this description of biological compartments is not relevant for biological matrices such as plasma. A simplified version of the network was obtained by removing this information, and a single node was considered for each metabolite in the case of multiple occurrences in different compartments. This network was composed of 109 pathways involving 4,095 metabolites, 5,389 reactions and 3,099 enzymatic complexes. A subset of 134 metabolites were successfully mapped on Recon3D using MetExplore *Metabolite Identifier Matcher* module, corresponding to 61% of the pool of compounds identified in the samples.

Over-Representation Analysis

Potential biomarker metabolites related to the first components were investigated using over-representation analysis. The latter aims to highlight pathways the more likely associated with a subset of metabolites by assessing whether they contain significantly more differentially expressed compounds than expected by chance. Because hemodialysis cannot be considered as a biological process, differences between preHD and postHD conditions summarized by the second component should indeed not be associated with specific biological pathways. Therefore,

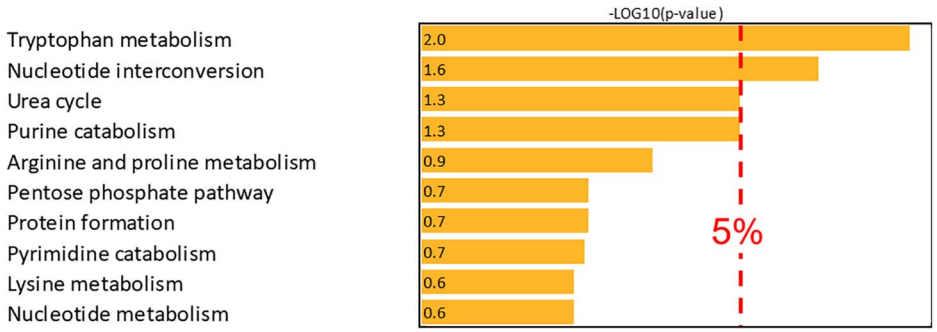

FIGURE 4 | Over-representation analysis (–log10 of the right-tailed Fisher test *p*-value corrected using the Benjamini-Hochberg False Discovery Rate procedure). The –log10 value of the 5% significance threshold is 1.301.

it was not relevant to further investigate potentially impacted metabolic pathways.

Metabolite mapping allowed 55 compounds associated with the first NetPCA component to be localized on the Recon3D network and over-representation analysis was therefore performed based on this subset. By these means, two pathways were reported as over-represented, namely *tryptophan metabolism* and *nucleotide interconversion*. The latter were thus objectively confirmed as metabolic pathways associated with CKD with 7 and 6 metabolites present in the subset of potential biomarkers, respectively. Moreover, *urea cycle* and *purine catabolism* were close to the significance threshold. A summary of the results of the over-representation analysis is proposed in **Figure 4**.

Network Analysis

Investigating metabolic pathways characterizing specific phenotypes provide valuable information about biological processes but are, however, limited to offer a global overview of the metabolism and its potential alterations. As many metabolites are involved in multiple interconnected pathways, the specificity of a metabolic signature is often difficult to guarantee. This is particularly crucial in a case such as CKD, because relevant phenotypic features of the disease may be spread over several pathways. As a consequence, it could be challenging to gain an overall understanding of the molecular events based only on pathway analysis.

Metabolic networks connecting all the pathways as a single object constitute therefore a very interesting alternative strategy. Subnetwork extraction was carried out based on significantly over-represented metabolic pathways associated with the first NetPCA component, namely *tryptophan metabolism* and *nucleotide interconversion*. By these means, it was then possible to highlight hubs, modules and bridges from the topology of the graph. This was done by investigating paths between pairs of metabolites, as a sequence of edges (i.e., reactions) connecting the starting node to the ending node. Due to the high number of edges composing a metabolic network, a very large amount of paths is possible, but by far not all of them are biologically relevant. Following the parsimony principle, the shortest path between two compounds might seem to be a suitable answer to make a choice among this multiplicity of alternatives. However, in many situations this solution is biochemically not optimal. Therefore, path search algorithms have been developed by incorporating biochemical rules to find the most relevant metabolic routes. Investigating lightest paths constitutes an efficient strategy, and it was performed based on the evaluation of the minimal squared degree sum of the nodes in the path, thus avoiding to give too much emphasis to uninformative ubiquitous compounds (Croes et al., 2006).

As a result, the two pathways highlighted as the most relevant using over-representation analysis, namely *tryptophan metabolism* and *nucleotide interconversion* could be efficiently displayed and linked in the subnetwork related to the first NetPCA component. Interestingly, the central role played by nucleotides and their derivatives could be highlighted, as metabolic hubs involved in a large number of reactions. Moreover, the reaction paths between tryptophan, kynurenic acid and indole-3-acetate are in line with prior studies reporting the link between renal dysfunction and the enzymatic activity of indoleamine 2,3-dioxygenase. The latter is responsible for tryptophan catabolism, the initial molecular event of the kynurenin pathway (Schefold et al., 2009). A decreased tryptophan plasmatic level is associated with CKD, while increased abundances of downstream products such as kynurenic acid have been reported (Kalim and Rhee, 2017). Related to these findings, a protein-bound uremic toxin from gut microbial origin, namely indoxyl sulfate was also highlighted as a metabolite of tryptophan associated with kidney failure (Niwa et al., 1994). Alterations in nucleotide interconversion may involve disturbed purine and pyrimidine metabolism. These have already been linked to both an increase in the prevalence and progression of the disease (Sekula et al., 2016; Shen et al., 2016). The degradation of purine derivatives generates hypoxanthine, which is further converted to xanthine and finally to uric acid, which is in accordance with known associations with CKD and uremic solutes. The subnetwork associated with these two pathways and the subset of altered metabolites related to the first NetPCA component is presented in **Figure 5**.

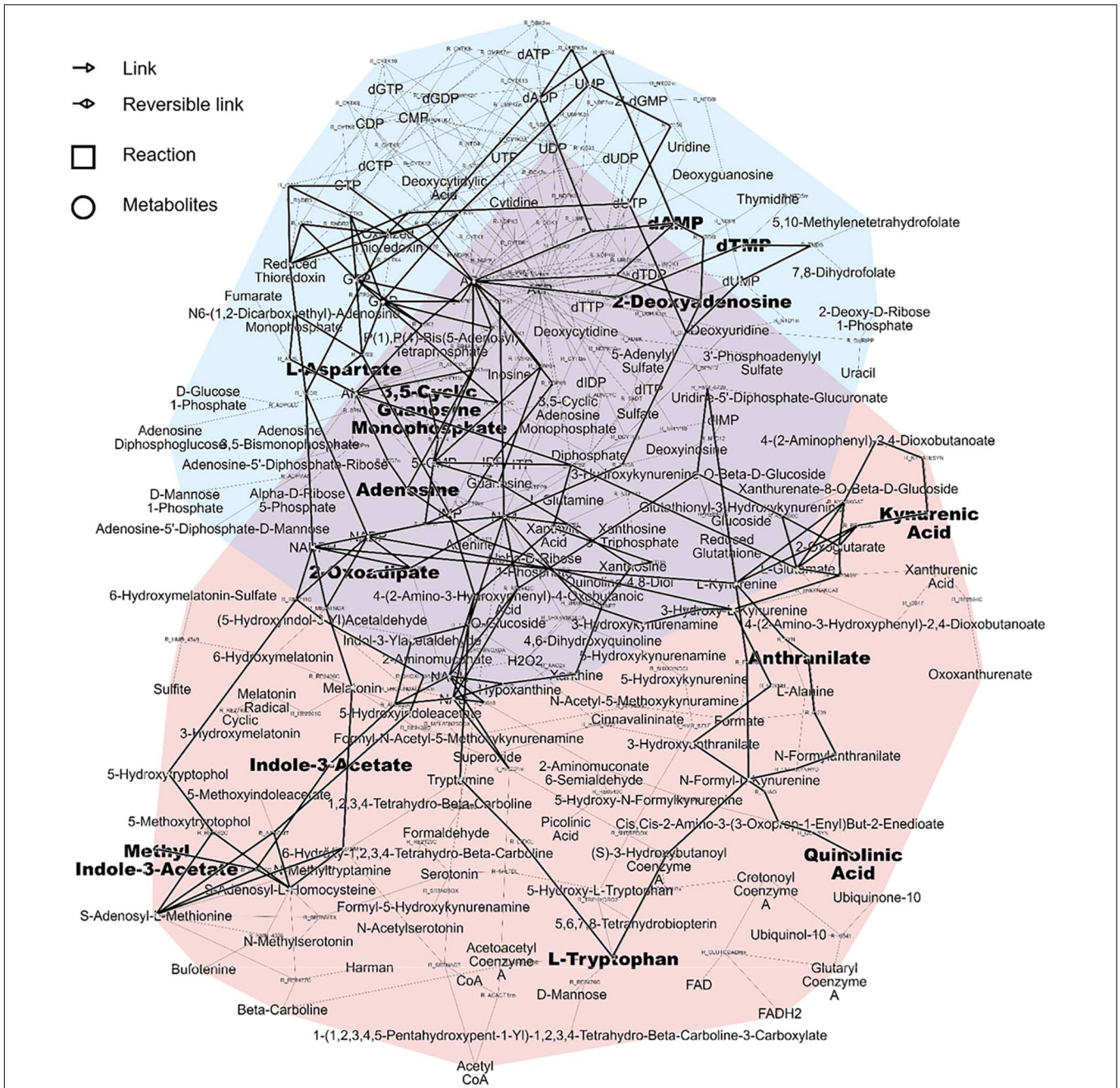

FIGURE 5 | Pink pathway: Tryptophan metabolism. Blue pathway: Nucleotide interconversion (purine and pyrimidine metabolism). Metabolites related to the first NetPCA component and the corresponding subnetwork are shown in bold.

Carbon Transfer Reaction Paths

Investigating the importance of individual nodes in a metabolic network is often challenging because these parameters are correlated with high degrees, therefore giving preference to highly connected pathways in terms of global topological properties. A node having many connections may therefore lack specificity in terms of metabolic information and not be a key determinant for a given process. Ubiquitous compounds playing an auxiliary role in metabolic reactions, e.g., H_2O, CO_2, or NAD, often constitute shortcuts in the graph and this also includes compounds associated with the *nucleotide interconversion* pathway, i.e., AMP, ADP, and ATP, as well as their deoxy counterparts. As it makes the investigation of local connectivity and bridges between modules challenging, a list of such side compounds was excluded from further topological investigations. Based on this, atom mapping was used for each reaction to evaluate the transfer between substrate and product atoms (Rahman et al., 2016), and edges not supporting any

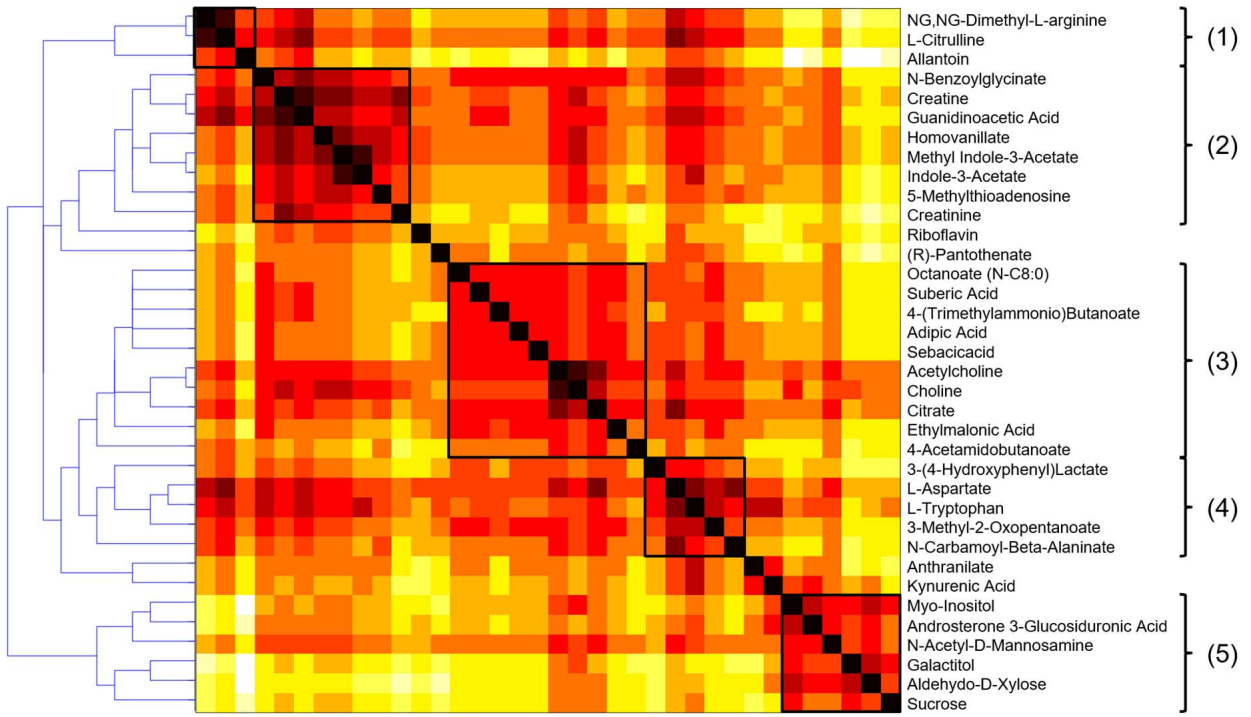

FIGURE 6 | HCA dendrogram and heatmap of the metabolic fingerprint distance matrix, with the five clusters highlighted as relevant biological modules. The darker, the shorter reaction path.

carbon atom transfer were removed from the compound graph. Conversely, those meeting this criterion were further considered as relevant for topological analysis (Frainay and Jourdan, 2017). Metabolic modules were then evaluated based on their reaction path in this graph using a distance matrix computed to summarize the different reaction paths between the compounds of the selected metabolic subset. Hierarchical clustering was then carried out on the distance matrix using complete-linkage to highlight potential biologically meaningful groupings. This agglomeration strategy helps to find compact clusters and avoid the chaining effect that would make the detection of biological modules more difficult. The HCA dendrogram and heatmap are presented in **Figure 6**.

A visual inspection of the dendrogram revealed five main clusters: (1) a small cluster composed of 3 amino acids of the urea cycle (N.N-dimethyl-L-arginine, L-citrulline and allantoin), that was connected with (2) a group of 8 metabolites associated with creatine metabolism (creatine, creatinine, guanidoacetic acid, hippuric acid, homovanillic acid, methyl indole-3-acetic acid, indole-3-acetic acid, 5-methylthioadenosine). It is well-known that creatinine blood levels reflect glomerular filtration efficiency and the urea to creatinine ratio can be used to characterize kidney function impairment (Duarte and Preuss, 1993). It constitutes therefore a clear positive control biomarker, reinforcing the biological validity of the findings.

A third cluster (3) composed of 10 metabolites included small carboxylic acids that could be potentially linked to disorders of fatty acid oxidation (adipic acid, sebacic acid, suberic acid, caprylic acid, deoxycarnitine, acetylcholine, choline, citric acid, ethylmalonic acid, 4-acetamidobutanoic acid) (Kang et al., 2015). Moreover, carnitine is necessary to transfer fatty acids for their oxidation in the mitochondria, and it also plays an important role in acetylcholine metabolism (Hoppel, 2003). The fourth (4) subset of 5 metabolites involving L-tryptophan, 3-(4-hydroxyphenyl)lactic acid, N-carbamoyl-beta-alanine, L-aspartate, and 3-methyl-2-oxovaleric acid, could be linked to aromatic amino acids metabolism and related compounds. Finally, (5) a cluster of sugar derivatives formed the last group (myo-inositol, N-acetyl-D-mannosamine, galactitol, D-xylose, and sucrose). Sugars are associated with the pathogenesis of diabetic nephropathy, but the underlying mechanisms involved are still unclear. Detrimental effects of advanced glycation end products constitute an interesting hypothesis (Dronavalli et al., 2008), but it remains, however, to be explored to better understand the progression of CKD. Box-plots of metabolites with representative alteration patterns from altered pathways are presented in **Figure 7**.

DISCUSSION

The association of chemometrics and bioinformatics in a common workflow was shown to be an effective approach for the integrative analysis of samples collected from several groups of patients suffering from multiple stages of CKD and/or undergoing different treatments. Despite heterogeneous

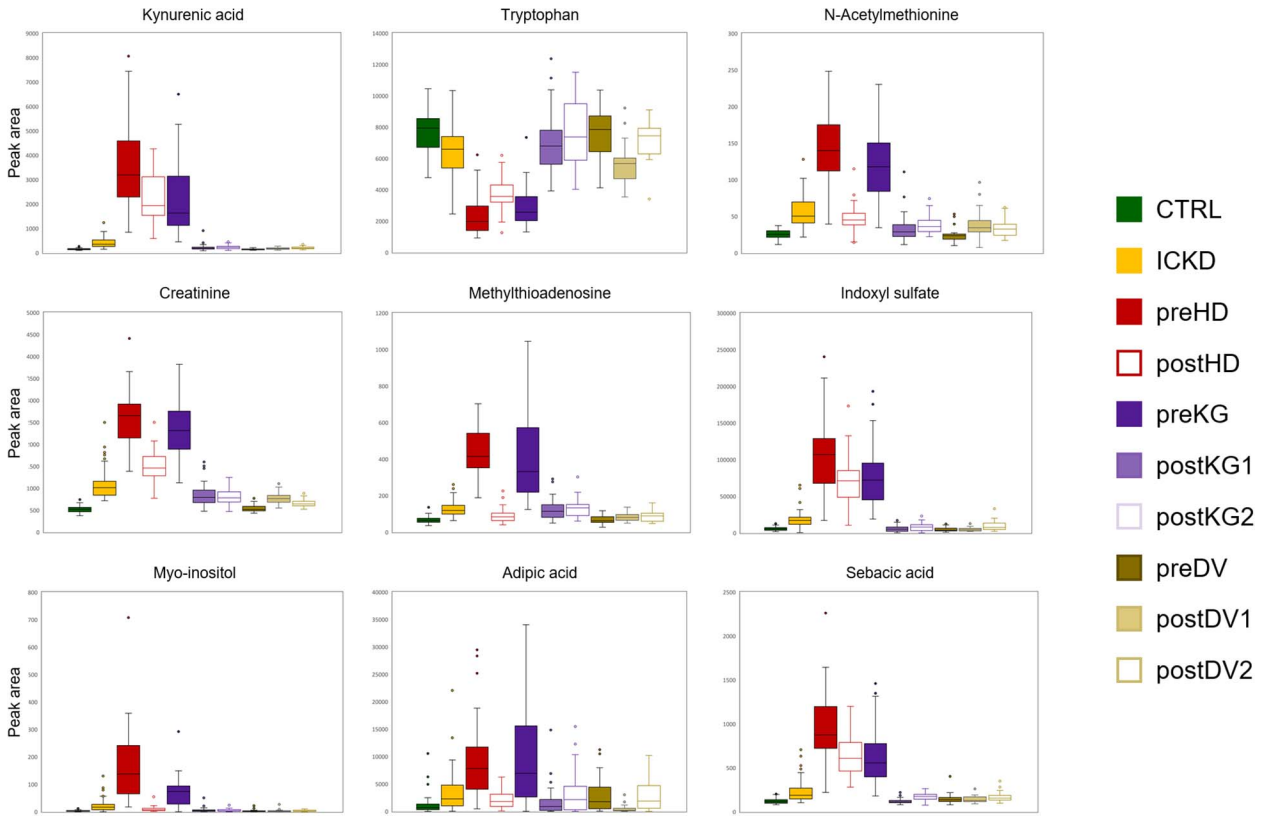

FIGURE 7 | Box-plots showing nine metabolites with representative alteration patterns from altered pathways.

sample collection, NetPCA allowed all groups of individuals to be included in a global model for an overall evaluation of their metabolic phenotypes. This integrative strategy underlined similar subsets of compounds describing the beneficial metabolic effects provided by hemodialysis and kidney graft, but also, for the first time, to compare the different alteration patterns on a common scale. The latter could be related to pathological modifications due to CKD and showed that kidney donors were only moderately affected by the decline in kidney function following organ donation. Notably, N-acetylation detoxification was reported to be altered in CKD, while tryptophan metabolism and nucleotide interconversion were highlighted using over-representation analysis. Further investigation using network reconstruction allowed deeper insights into their link to be gained. Tryptophan metabolism was already reported as an altered metabolic route in CKD, with decreased tryptophan plasmatic concentration and increased levels of downstream products such as kynurenic acid. Nucleotide interconversion can be considered as a generic biological process covering both purine and pyrimidine metabolism, both known to be affected during CKD. Further information about meaningful metabolic modules was finally obtained using hierarchical clustering based on reliable reaction paths. By these means, additional hypotheses involving creatine metabolism and urea cycle, carnitine and disorders of fatty acid oxidation, as well as aromatic amino acids metabolism and sugar derivatives could be drawn.

Although the coverage of some parts of metabolic networks or chemical families still needs to be improved and/or refined (e.g., lipids), the current databases already allow a fine assessment of the interconnectivity of the different metabolic pathways and their topology. This allows to put into perspective the differences observed between samples characterizing specific clinical or experimental situations, and to go further in the biological interpretation of the regulatory networks governing phenomena of interest. In addition, the annotation of metabolic networks is being actively carried out thus offering continuous improvements to grasp the complexity of the metabolism. With this aim in sight, the association of chemometrics and bioinformatics in a common workflow will certainly play a central role in the future of metabolomics.

DATA AVAILABILITY STATEMENT

The 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 the local ethics committee (Commission cantonale

d'éthique de la recherche (CCER), Geneva, Switzerland) and performed according to the Declaration of Helsinki. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

BP coordinated CKD clinical studies and sample collection. YG carried out sample preparation, LC-MS experiments. YG, SC, and JB managed data processing. SC, JB, and MH participated in chemometrics data analysis. JB, DS, and FJ participated in bioinformatics data analysis. BP, FJ, and SR managed the project. All authors participated in manuscript writing and correction.

FUNDING

This work was partly supported by the research fund of the Department of Internal Medicine of the University Hospital and the Faculty of Medicine of Geneva; this fund received an unrestricted grant from AstraZeneca Switzerland. BP was partly funded by the Marie Heim-Vötglin Grant of the Swiss National Science Foundation (SNSF nos. PMPDP3_171352 and PMPDP3_186203). This work was supported by the French Ministry of Research and National Research Agency as part of the French MetaboHUB, the National Metabolomics and Fluxomics Infrastructure (Grant ANRINBS-0010).

ACKNOWLEDGMENTS

We thank Chantal Martinez, study nurse, for her help in collecting the clinical data and Géraldine Poulain for her help in the sample and biobank processing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.682559/full#supplementary-material

REFERENCES

Boccard, J., and Rudaz, S. (2014). Harnessing the complexity of metabolomic data with chemometrics. *J. Chemometrics* 28, 1–9. doi: 10.1002/cem.2567

Boccard, J., Veuthey, J. L., and Rudaz, S. (2010). Knowledge discovery in metabolomics: an overview of MS data handling. *J. Separat. Sci.* 33, 290–304. doi: 10.1002/jssc.200900609

Boelaert, J., t'Kindt, R., Schepers, E., Jorge, L., Glorieux, G., Neirynck, N., et al. (2013). State-of-the-art non-targeted metabolomics in the study of chronic kidney disease. *Metabolomics* 10, 425–442. doi: 10.1007/s11306-013-0592-z

Booth, S. C., Weljie, A. M., and Turner, R. J. (2013). Computational tools for the secondary analysis of metabolomics experiments. *Comput. Struct. Biotechnol. J.* 4:e201301003. doi: 10.5936/csbj.201301003

Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J. I, Wilson, D., Lewis, M. R., et al. (2018). Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* 14:72.

Brunk, E., Sahoo, S., Zielinski, D. C., Altunkaya, A., Drager, A., Mih, N., et al. (2018). Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.* 36:272. doi: 10.1038/nbt.4072

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C. A. I, Keseler, M., et al. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 44, D471–D480.

Chazalviel, M., Frainay, C., Poupin, N., Vinson, F., Merlet, B., Gloaguen, Y., et al. (2018). MetExploreViz: web component for interactive metabolic network visualization. *Bioinformatics* 34, 312–313. doi: 10.1093/bioinformatics/btx588

Codesido, S., Hanafi, M., Gagnebin, Y., González-Ruiz, V., Rudaz, S., and Boccard, J. (2020). Network principal component analysis: a versatile tool for the investigation of multigroup and multiblock datasets. *Bioinformatics* doi: 10.1093/bioinformatics/btaa954 [Epub ahead of print].

Cottret, L., Frainay, C., Chazalviel, M., Cabanettes, F., Gloaguen, Y., Camenen, E., et al. (2018). MetExplore: collaborative edition and exploration of metabolic networks. *Nucleic Acids Res.* 46, W495–W502.

Cottret, L., Wildridge, D., Vinson, F., Barrett, M. P., Charles, H., Sagot, M. F., et al. (2010). MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Res.* 38, W132–W137.

Croes, D., Couche, F., Wodak, S. J., and van Helden, J. (2006). Inferring meaningful pathways in weighted metabolic networks. *J. Mol. Biol.* 356, 222–236. doi: 10.1016/j.jmb.2005.09.079

Dronavalli, S., Duka, I., and Bakris, G. L. (2008). The pathogenesis of diabetic nephropathy. *Nat. Clin. Pract. Endocrinol. Metab.* 4, 444–452.

Duarte, C. G., and Preuss, H. G. (1993). Assessment of renal-function – glomerular and tubular. *Clin. Lab. Med.* 13, 33–52. doi: 10.1016/s0272-2712(18)30459-1

Frainay, C., and Jourdan, F. (2017). Computational methods to identify metabolic sub-networks based on metabolomic profiles. *Brief. Bioinform.* 18, 43–56. doi: 10.1093/bib/bbv115

Frainay, C., Schymanski, E. L., Neumann, S., Merlet, B., Salek, R. M., Jourdan, F., et al. (2018). Mind the gap: mapping mass spectral databases in genome-scale metabolic networks reveals poorly covered areas. *Metabolites* 8:51. doi: 10.3390/metabo8030051

Gagnebin, Y., Jaques, D. A., Rudaz, S., de Seigneux, S., Boccard, J., and Ponte, B. (2020a). Exploring blood alterations in chronic kidney disease and haemodialysis using metabolomics. *Sci. Rep.* 10:19502.

Gagnebin, Y., Pezzatti, J., Lescuyer, P., Boccard, J., Ponte, B., and Rudaz, S. (2019). Toward a better understanding of chronic kidney disease with complementary chromatographic methods hyphenated with mass spectrometry for improved polar metabolome coverage. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 1116, 9–18. doi: 10.1016/j.jchromb.2019.03.031

Gagnebin, Y., Pezzatti, J., Lescuyer, P., Boccard, J., Ponte, B., and Rudaz, S. (2020b). Combining the advantages of multilevel and orthogonal partial least squares data analysis for longitudinal metabolomics: application to kidney transplantation. *Anal. Chim. Acta* 1099, 26–38. doi: 10.1016/j.aca.2019.11.050

Gagnebin, Y., Tonoli, D., Lescuyer, P., Ponte, B., de Seigneux, S., Martin, P. Y., et al. (2017). Metabolomic analysis of urine samples by UHPLC-QTOF-MS: impact of normalization strategies. *Anal. Chim. Acta* 955, 27–35. doi: 10.1016/j.aca.2016.12.029

Guimera, R., and Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature* 433, 895–900. doi: 10.1038/nature03288

Hocher, B., and Adamski, J. (2017). Metabolomics for clinical use and research in chronic kidney disease. *Nature Rev. Nephrol.* 13, 269–284. doi: 10.1038/nrneph.2017.30

Hoppel, C. (2003). The role of carnitine in normal and altered fatty acid metabolism. *Am. J. Kidney Dis.* 41, S4–S12.

Jonsson, P., Wuolikainen, A., Thysell, E., Chorell, E., Stattin, P., Wikstrom, P., et al. (2015). Constrained randomization and multivariate effect projections improve information extraction and biomarker pattern discovery in metabolomics

studies involving dependent samples. *Metabolomics* 11, 1667–1678. doi: 10.1007/s11306-015-0818-3

Kalim, S., and Rhee, E. P. (2017). An overview of renal metabolomics. *Kidney Int.* 91, 61–69. doi: 10.1016/j.kint.2016.08.021

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30.

Kang, H. M., Ahn, S. H., Choi, P., Ko, Y. A., Han, S. H., Chinga, F., et al. (2015). Defective fatty acid oxidation in renal tubular epithelial cells has a key role in kidney fibrosis development. *Nat. Med.* 21, 37–46. doi: 10.1038/nm.3762

Kankainen, M., Gopalacharyulu, P., Holm, L., and Oresic, M. (2011). MPEA-metabolite pathway enrichment analysis. *Bioinformatics* 27, 1878–1879. doi: 10.1093/bioinformatics/btr278

Kell, D. B. (2004). Metabolomics and systems biology: making sense of the soup. *Curr. Opin. Microbiol.* 7, 296–307. doi: 10.1016/j.mib.2004.04.012

Lacroix, V., Cottret, L., Thebault, P., and Sagot, M. F. (2008). An introduction to metabolic networks and their structural analysis. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 5, 594–617. doi: 10.1109/tcbb.2008.79

Levey, A. S., Eckardt, K. U., Dorman, N. M., Christiansen, S. L., Hoorn, E. J., Ingelfinger, J. R., et al. (2020). Nomenclature for kidney function and disease: report of a Kidney Disease: improving global outcomes (KDIGO) consensus conference. *Kidney Int.* 97, 1117–1129.

Marco-Ramell, A., Palau-Rodriguez, M., Alay, A., Tulipani, S., Urpi-Sarda, M., Sanchez-Pla, A., et al. (2018). Evaluation and comparison of bioinformatic tools for the enrichment analysis of metabolomics data. *BMC Bioinformatics* 19:1. doi: 10.1186/s12859-017-2006-0

Niwa, T., Ise, M., and Miyazaki, T. (1994). Progression of glomerular sclerosis in experimental uremic rats by administration of indole, a precursor of indoxyl sulfate. *Am. J. Nephrol.* 14, 207–212. doi: 10.1159/000168716

Noronha, A., Danielsdottir, A. D., Gawron, P., Johannsson, F., Jonsdottir, S., Jarlsson, S., et al. (2017). ReconMap: an interactive visualization of human metabolism. *Bioinformatics* 33, 605–607.

Pezzatti, J., Boccard, J., Codesido, S., Gagnebin, Y., Joshi, A., Picard, D., et al. (2020). Implementation of liquid chromatography-high resolution mass spectrometry methods for untargeted metabolomic analyses of biological samples: a tutorial. *Anal. Chim. Acta* 1105, 28–44. doi: 10.1016/j.aca.2019.12.062

Pezzatti, J., Gonzalez-Ruiz, V., Codesido, S., Gagnebin, Y., Joshi, A., Guillarme, D., et al. (2019). A scoring approach for multi-platform acquisition in metabolomics. *J. Chromatogr. A* 1592, 47–54. doi: 10.1016/j.chroma.2019.01.023

Rahman, S. A., Torrance, G., Baldacci, L., Cuesta, S. M., Fenninger, F., Gopal, N., et al. (2016). Reaction Decoder Tool (RDT): extracting features from chemical reactions. *Bioinformatics* 32, 2065–2066. doi: 10.1093/bioinformatics/btw096

Richards, S. E., Dumas, M. E., Fonville, J. M., Ebbels, T. M. D., Holmes, E., and Nicholson, J. K. (2010). Intra- and inter-omic fusion of metabolic profiling data in a systems biology framework. *Chemomet. Intell. Lab. Syst.* 104, 121–131. doi: 10.1016/j.chemolab.2010.07.006

Schefold, J. C., Zeden, J. P., Fotopoulou, C., von Haehling, S., Pschowski, R., Hasper, D., et al. (2009). Increased indoleamine 2,3-dioxygenase (IDO) activity and elevated serum levels of tryptophan catabolites in patients with chronic kidney disease: a possible link between chronic inflammation and uraemic symptoms. *Nephrol. Dial. Transplant.* 24, 1901–1908. doi: 10.1093/ndt/gfn739

Sekula, P., Goek, O. N., Quaye, L., Barrios, C., Levey, A. S., Romisch-Marg, W., et al. (2016). A metabolome-wide association study of kidney function and disease in the general population. *J. Am. Soc. Nephrol.* 27, 1175–1188.

Shen, K. Y., Johnson, D. W., and Gobe, G. C. (2016). The role of cGMP and its signaling pathways in kidney disease. *Am. J. Physiol. Renal Physiol.* 311, F671–F681.

Thiele, I., Swainston, N., Fleming, R. M. T., Hoppe, A., Sahoo, S., Aurich, M. K., et al. (2013). A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* 31:419.

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46, D608–D617.

Yilmaz, L. S., and Walhout, A. J. M. (2017). Metabolic network modeling with model organisms. *Curr. Opin. Chem. Biol.* 36, 32–39. doi: 10.1016/j.cbpa.2016.12.025

Zhang, X. W., Li, Q. H., Xu, Z. D., and Dou, J. J. (2020). Mass spectrometry-based metabolomics in health and medical science: a systematic review. *RSC Adv.* 10, 3092–3104. doi: 10.1039/c9ra08985c

Zhao, Y. Y. (2013). Metabolomics in chronic kidney disease. *Clin. Chim. Acta* 422, 59–69. doi: 10.1016/j.cca.2013.03.033

Conflict of Interest: The authors declare that this study received indirect partial funding from AstraZeneca. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 27 May 2021
doi: 10.3389/fmolb.2021.681400

Blood Microsampling to Monitor Metabolic Profiles During Physical Exercise

*Cindy Nix, Maryam Hemmati, Gaël Cobraiville, Anne-Catherine Servais and Marianne Fillet**

Laboratory for the Analysis of Medicines, Department of Pharmacy, CIRM, University of Liège, Liège, Belgium

Monitoring approaches and technical improvements are key factors to improve a sportsman's health, training, and recovery after an injury. In this study, a targeted metabolomics approach using microsampling with hemaPEN® was developed to measure changes in blood concentrations of nine amino acids and four organic acids before, during, and after exercise. The aim of this research project was to investigate if a reliable monitoring of metabolite levels during sports activity can be achieved by collecting one drop of whole blood at different time points. A hemaPEN device is an easy-to-use and noninvasive microsampling technique designed to collect four accurate and precise blood volumes simultaneously (10.96 µl). Twenty healthy volunteers between 19 and 30 years of age were included in this study. Physical activity consisted in running as fast as possible 1,600 m after 400 m warm-up. One drop of blood was collected at five time points: before exercise, after 800-m running, after 1,600 m, and 30 min and 60 min after finishing the exercise. The influence of physical activity on metabolite levels was evaluated using two ultrahigh-performance liquid chromatography coupled to tandem mass spectrometry (UHPLC-MS/MS) methods. Analytical performance criteria such as metabolite stability, method precision, trueness, and accuracy were found to be satisfactory. Expected significant metabolic changes were identified for lactic acid, main TCA cycle intermediates, and some amino acids (e.g., creatinine, choline, and taurine). This preliminary study performed on a small cohort demonstrated a high interest of using microsampling for fluxomics analysis, not only to collect quickly and easily biological samples during sports events but also because it is much easier to store and to process the samples than classical plasma/serum samples obtained by venipuncture. The present results open new avenue for fluxomics analysis in the context of health care.

Keywords: amino acids, organic acids, running, hemaPEN®, UHPLC-MS/MS, targeted metabolomics

OPEN ACCESS

Edited by:
Serge Rudaz,
Université de Genève, Switzerland

Reviewed by:
Francesco Botre,
University of Lausanne, Switzerland
Raphael Faiss,
University of Lausanne, Switzerland

****Correspondence:***
Marianne Fillet
marianne.fillet@uliege.be

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

Received: *16 March 2021*
Accepted: *30 April 2021*
Published: *27 May 2021*

Citation:
Nix C, Hemmati M, Cobraiville G, Servais A-C and Fillet M (2021) Blood Microsampling to Monitor Metabolic Profiles During Physical Exercise. Front. Mol. Biosci. 8:681400. doi: 10.3389/fmolb.2021.681400

Abbreviations: EMA, European Medicines Agency; ESI, electrospray ionization; MS/MS, tandem mass spectrometry; MRM, multiple reaction monitoring; NMT, neuromuscular training; QC, quality control; RSD, relative standard deviation; SD, standard deviation; TCA, tricarboxylic acid; UHPLC, ultrahigh-performance liquid chromatography; VAMS, volumetric absorptive microsampling; WADA, World Anti-doping Agency.

INTRODUCTION

Exercise is one key factor to sustain good health (Kelly et al., 2020). Indeed, it is widely proven that physical activity reduces the risk of obesity (Mielke et al., 2019), cardiovascular diseases (Porter et al., 2019), diabetes, hypertension (Aune et al., 2015), and depression (Matta et al., 2021). The World Health Organization (WHO) recommends at least 150 min of moderate exercise or 75 min of intense exercise per week for adults. Muscle strengthening activities and reduced sedentary time also constitute two critical points to keep fit. Despite all the recognized benefits of sports, worldwide around 1 in 4 adults is not sufficiently physically active (WHO, 2020). In order to promote physical exercise, various awareness campaigns and programs have recently been set up, such as the Project Smart which promotes physical activity for children through games (Julien et al., 2021) or the global action plan of the WHO to encourage sport activities (WHO, 2018). In order to avoid any risk of injury, it is crucial that sports practice is supervised and adapted to each person. For athletes, this consideration is particularly crucial since an injury can have a big impact on their health and their career as well as financial repercussions. Currently, the number of competitions in which athletes participate is increasing. Therefore, athletes intensify their training to be always competitive. Nevertheless, health professionals agree that overloading training and competitions can have serious consequences for the health of athletes (Soligard et al., 2016). To avoid the risk of injury, several strategies are available.

In those contexts of amateur and professional sports, fluxomics could be an interesting approach to implement personalized athlete monitoring to reduce the risk of injury, adapt the training, and speed up recovery after injury if any (Al-Khelaifi et al., 2019). Indeed, the measurement of metabolite levels in individuals at different time frames provides information concerning the evolution of their physiopathological state. For example, several studies have shown that exercise induced metabolic changes. Lewis *et al.* demonstrated that the concentrations of niacinamide, glucose-6-phosphate, pantothenate, and succinate increased in plasma after an exercise (Lewis et al., 2010). More recently, Stander *et al.* have shown that the serum concentrations of carbohydrates, fatty acids, TCA cycle intermediates, and ketones were increased after a marathon, while the levels of amino acids were reduced (Stander et al., 2018). Prado et al. even proposed the term "sportomics" to qualify the use of "-omics" sciences to better understand the metabolic changes induced by a physical activity (Prado et al., 2017). They performed untargeted metabolomics in urine before and after a soccer match. Different categories of metabolites were found to be interesting, among which are organic acids (Prado et al., 2017).

Although several studies showing metabolomic changes during or after physical activity have already been published (Sakaguchi et al., 2019), none allowed athletes to collect blood by themselves directly at the training site. In this study, we aimed to make the proof of concept that well-known biomarkers such as lactic acid or creatinine can easily be monitored during physical activity using a noninvasive and easy-to-handle microsampling device. Blood collection was achieved using a microsampling device named hemaPEN® (Trajan Scientific and Medical, VIC, Australia). This pen-shaped device allows the accurate and precise collection of samples from a single drop of blood and, depending on the studied compounds, is influenced in a limited way or not influenced by hematocrit compared to classical DBS (Deprez et al., 2019; Protti et al., 2020). Blood is absorbed by four capillaries and transferred to four paper discs. Each capillary collects 2.74 µl of blood. This device can be used everywhere by everyone without specific training (Trajan-HemaPEN brochure, 2020) (cf. **Figure 1**). In this study, we investigated whether it is possible to monitor metabolism changes in a reliable manner from 2.74 µl of blood not only before and after the exercise but also during the effort.

MATERIALS AND METHODS

Chemicals and Reagents

L-asparagine (≥98%), choline chloride (≥99%), creatine (≥99.5%), creatinine (≥98%), L-leucine (≥98%), L-methionine (≥98%), L-proline (≥99.5%), taurine (≥99%), L-valine (≥98%), 2-hydroxybutanoic acid, L-lactic acid (≥98%), malic acid (≥99%), and 2-oxo-glutaric acid (≥99%) were purchased from Sigma Aldrich (St. Louis, MO, United States). L-lactic acid-d_3 and choline-d_9 were obtained from Toronto Research Chemicals (Toronto, ON, Canada). L-asparagine $^{13}C_4$ (≥99%) and L-leucine-d_3 (≥99%) were obtained from Euriso-top (Saint-Aubin, France). Acetonitrile, formic acid, trifluoroacetic acid, isopropanol, methanol, and water (MS-grade) were bought from Biosolve (Valkenswaard, The Netherlands). hemaPEN® devices were obtained from Trajan Scientific and Medical (VIC, Australia) and Ostro® 96-well sample preparation plates from Waters Corporation (Dublin, Ireland). MiniCollect® safety lancets were purchased from Greiner Bio-One (Vilvoorde, Belgium).

Healthy Volunteers

The protocol of this study was reviewed and approved by the Ethical Committee of the University of Liège (Liège, Belgium), Nr Eudra CT: B7072020000041; ref: 2020/182. The participants provided their written informed consent to participate in the study. Twelve women and eight men between 19 and 30 years of age with a body mass index (BMI) between 18.5 and 25 were included in this study. The participants did not suffer from any chronic respiratory, inflammatory, cardiovascular, or metabolic diseases and were not smokers. The participants did not have any fever, and they have not been infected with COVID-19. They exercised at least 1 h per week (4.2 h/week in average). To minimize the impact of nutrition on metabolic changes, they had standard meals (breakfast and lunch) on the day of running and no excess food or drink the day before the exercise. The parameters of all the participants are described in **Supplementary Table S1**.

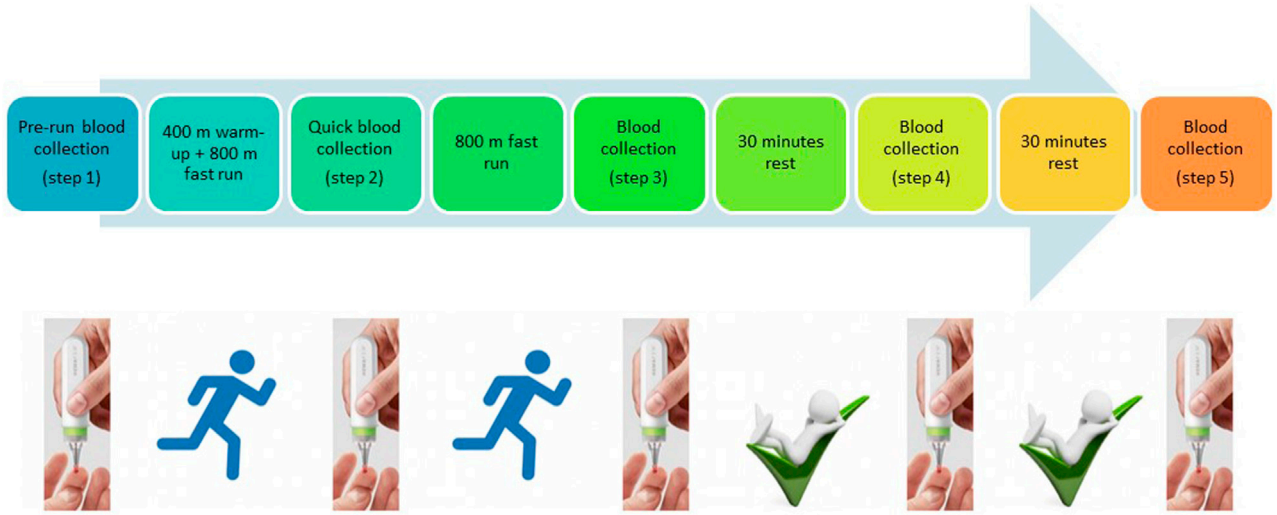

FIGURE 1 | Design of the study.

Physical Exercise

The physical exercise consisted of running 400 m as a warm-up that was followed by running 1,600 m at high speed on athletic tracks.

Sample Collection

Blood samples were collected at five different time points. At every time point, four replicate samples were collected. The first sample was collected at rest, shortly before running (Step 1), the second sample after running of 400-m warm-up and 800-m running (Step 2), the third sample after the completion of 1,600-m running (Step 3), and the fourth and fifth samples at 30 min (Step 4) and 60 min (Step 5) after finishing the exercise, respectively. Blood was collected with a hemaPEN® device *via* a finger prick with a safety lancet (**Figure 1**).

Sample Storage

After collection, the hemaPEN® devices were kept at ambient temperature for 2 h to allow the samples to dry. The devices were then stored at −20°C before analysis.

Stock Solutions of Amino Acids and Organic Acids

Aqueous stock solutions of nine amino acids and four organic acids were prepared at specific concentrations: 200 mM choline chloride, 100 mM L-methionine, 150 mM L-asparagine, 150 mM L-leucine, 100 mM creatine, 200 mM taurine, 500 mM L-proline, 300 mM L-valine, 500 mM creatinine, 800 mM L-lactic acid, 150 mM 2-hydroxybutanoic acid, 120 mM malic acid, and 80 mM 2-oxoglutaric acid. Stocks solutions were stored at −80°C. Aqueous stock solutions of 10 mM L-lactic acid-d_3, choline-d_9, L-asparagine $^{13}C_4$, and L-leucine-d_3 were prepared to be used as internal standards. These solutions were stored at −80°C.

Calibration Standards and Quality Control

Calibration solutions were prepared by diluting aqueous stock solutions in human blood to reach the targeted concentrations. Separate calibration curves were prepared for the analytes in the range of their endogenous concentrations. Calibration ranges were as follows: 15–150 µM for 2-hydroxybutanoic acid, 640–6400 µM for lactic acid, 6–60 µM for malic acid, 2.4–24 µM for 2-oxoglutaric acid, 14–140 µM for asparagine, 5.5–55 µM for choline, 92.5–925 µM for creatine, 27–270 µM for creatinine, 28–280 µM for leucine, 5.5–55 µM for methionine, 55–550 µM for proline, 41.5–415 µM for taurine, and 56–560 µM for valine.

Five calibration levels were used to construct calibration curves. Concentration at the upper limit of the calibration curves (C5) was considered as the 100% concentration. This C5 solution was diluted to obtain the other levels: 10%=C1 (lowest level of the calibrations curves), 25%=C2, 50%=C3, and 75%=C4. The calibration standards were prepared with hemaPEN® devices and extracted according to the sample preparation procedure, and each calibration standard was injected six times. A quality control (QC) sample was prepared at C3 and injected after every twenty injections to check the system performance during the batch analysis.

Sample Preparation

Using Trajan's specially designed tool, the hemaPEN® devices were opened and the cartridge containing the four replicate samples removed. One paper disc was used for this study, and the other three were kept in the cartridge in a sealed box at −20°C for possible future analyses.

The extraction solvent was a mixture of acetonitrile and water (60:40, v/v) containing the internal standards in the concentrations as follows: 80 µM for L-lactic acid-d3, 1.25 µM for choline-d9, 5 µM for L-asparagine 13C4, and 7.5 µM for L-leucine-d3. The hemaPEN® paper disks were placed in an Ostro™ 96-well plate (Waters Corporation, Dublin, Ireland)

and were incubated with 200 µl extraction solvent for 5 min without agitation followed by an agitation step of 5 min using a ThermoMixer C (Eppendorf, Aarshot, Belgium) at 20°C and 850 rpm. The samples were then collected in a 96-well plate (Agilent Technologies, Waldbronn, Germany) by passing them through the Ostro™-plate using a vacuum manifold.

Before the analysis of organic acids by UHPLC-MS/MS, a volume of 100 µl of the extracted samples was evaporated during 60 min at 40°C in a CentriVAP Concentrator (LabConco, Kansas-City, MO, United States). The dried residues were reconstituted in 50 µl water and analyzed by UHPLC-MS/MS. The sample preparation procedure is described in **Supplementary Figure S1**.

UHPLC-MS/MS Analysis

The samples containing the amino acids and organic acids were analyzed with an ultrahigh-performance liquid chromatography–tandem mass spectrometry (UHPLC-MS/MS) method that was published previously by our group (Kok et al., 2019). LC-MS/MS analyzes were conducted on an Agilent® 1290 Infinity system coupled to an Agilent® 6495 triple quadrupole mass spectrometer.

Sample Randomization

Samples were analyzed in a randomized way to ensure that the results obtained are not influenced by the order of analysis.

Stability Evaluation

The stability of the analytes in the autosampler was assessed. To evaluate the stability in the autosampler, three samples at concentration level C1 (first level of the calibration curves) and three samples at concentration level C5 (upper limit of the calibration curve) were analyzed immediately after preparation and after 24-h storage at 4°C in the autosampler. The responses obtained after 24 h were compared to the responses obtained immediately after sample preparation.

Long-term stability was assessed using one of the four replicate samples collected with the hemaPEN® device. One paper sample disc collected after 1,600 m of running was analyzed after 5 months of storage and compared to the freshly analyzed samples. This study was possible because four replicates per sample are collected simultaneously with one hemaPEN® device. Stability is considered satisfactory if the responses do not vary more than 15% compared to fresh samples (Smith, 2012).

Performance Criteria of Analytical Methods

Response Function, Trueness, Precision, and Accuracy

For all the studied compounds, the most appropriate regression model as well as the trueness, the precision, and the accuracy of the methods were obtained performing the prevalidation of the methods. Calibration curves with five levels were prepared independently on three different days. The most suitable regression models were chosen according to the obtained accuracy profiles. The acceptance limits and the maximum risk of having future measurements falling out of these acceptance limits were set at ±20%. All accuracy profiles are presented in **Supplementary Figure S2**. Trueness of the analytical method was evaluated using the relative bias (%) at five concentration levels. Within-run precision was assessed by the repeatability (RSD%) and between-run precision by the intermediate precision (RSD%).

Matrix Effect

The matrix effect was assessed according to the method described by Matuszewski et al. (2003) at two concentration levels (midrange of the calibration curve and the upper limit of the calibration curve). For amino acids, neat standard solutions containing the appropriate concentration of all the amino acids were prepared in acetonitrile/water (6:4). For organic acids, a neat standard solution containing the appropriate concentration of all the organic acids was prepared in water. Post-extraction spiked matrices were prepared by extracting blank blood. For amino acids, extracted blank blood was directly spiked with a mixture containing all the studied amino acids at the appropriate concentration. For organic acids, 100 µl of extracted blank blood was evaporated during 60 min at 40°C in a CentriVAP Concentrator (LabConco, Kansas City, MO, United States) and reconstituted with 50 µl of water containing the appropriate concentration of all the organic acids studied. The matrix effect was calculated by dividing the peak areas obtained with the post-extraction spiked matrices by the peak areas obtained with the neat standard solutions. Neat standard solutions and post-extraction spiked matrices were prepared in triplicates for each concentration level.

Carryover

Carryover was assessed according to the guideline on bioanalytical method validation from the European Medicines Agency (EMA) (Smith, 2012). A solution at the highest concentration level was injected three times. After each injection, a blank was injected. For all compounds, the peak area obtained in blanks should not exceed 20% of the peak area obtained at the lowest level of the calibration curves. For the internal standards, the peak area obtained should not exceed 5%. Carryover was calculated according to the formula below:

$$\textit{Carryover}\ (\%) = \frac{\textit{Peak area obtained in a blank after an injection of a solution at the highest concentration level}}{\textit{Peak area obtained at the lowest level of the calibration curve.}}$$

Data Analysis

Data acquisition was performed with MassHunter Data Acquisition software (B.08.02, Agilent Technologies, Waldbronn, Germany). Data analysis was performed with Quantitative Analysis software (version 10.1, Agilent Technologies, Waldbronn, Germany). Peak areas corresponding to the endogenous concentrations of studied compounds in blood used to prepare calibration curves were subtracted from the values obtained for the calibration solutions (Vishwanathan et al., 2000). The concentration obtained at each step of the study was divided by the concentration obtained at step 1 (at rest before exercise) to obtain a change of magnitude factors. The means of the change of magnitude factors obtained at

each step were compared with a one-way ANOVA test for data with a normal distribution and a Friedman test for data not satisfying the normality test. These statistical analyzes were performed using GraphPad prism 6 and as well as the construction of the different graphs. Results of prevalidation were computed with e-noval 4.1 software (Pharmalex, Mont-Saint-Guibert, Belgium).

RESULTS

In this study, two previously developed UHPLC-MS/MS methods were used to quantify nine amino acids and four organic acids in whole blood after microsampling. Five samples were collected by each participant before, during, and after a physical activity consisting in 1,600 m of intensive running. This study focused on 13 metabolites because the influence of physical activity on their blood concentration is well-described in the literature (Stander et al., 2018; Schranner et al., 2020). Indeed, the aim of our study was not to discover new biomarkers but to investigate the potential of microsampling using hemaPEN® devices for fluxomics analysis.

Analytical Performance of the Methods

For all the studied compounds, the most adequate response functions to quantify the analytes over the entire range of concentrations were determined by applying different regression models and selecting the most suitable accuracy profiles. The selected models for which the tolerance limits are included within the acceptance limits (±20%) as well as the accuracy profiles are presented in **Supplementary Figure S2**.

In term of trueness, relative bias was below ±8.63% for all the organic acids and below ±8.24% for all the amino acids at the five investigated concentration levels. For within- and between-run precision, repeatability (RSD%) and intermediate precision (RSD %) were below 13.9% for all the compounds at the five concentration levels. The results obtained for trueness, repeatability, and intermediate precision are presented in **Supplementary Table S2**.

Stability

For all the studied compounds, the responses obtained directly after preparation and after 24-h storage in the autosampler or after 5 months storage at −20°C do not vary by more than 15%, except for 2-oxoglutaric acid, methionine, creatine, and taurine at long-term storage. For 2-oxoglutaric acid and methionine, concentrations obtained after 5-month storage were 35.1 and 56.5% lower than the nominal concentrations, respectively. Concentrations of creatine and taurine obtained after 5-month storage at −20°C were 17.6 and 15.7% higher than the nominal concentrations, respectively. Results concerning the stability in the autosampler and the long-term stability are presented in **Supplementary Tables S3, S4**, respectively.

Matrix Effect and Carryover

The matrix effect represents the influence of the matrix on the ionization of the compounds of interest in the ESI source (Matuszewski et al., 2003). Indeed, human blood contains a large number of compounds. These compounds can interfere with the ionization of the quantified analytes. A matrix effect lower than 100% means that the matrix reduces ionization efficiency, while a matrix effect higher than 100% means that the signal is enhanced in the matrix. The obtained results are presented in **Supplementary Table S5**.

The matrix effect was lower than 100% (in average 76.6%) for all the compounds, except for asparagine (117%). The ionization of creatine does not seem to be influenced by the matrix (99.9%). For creatinine, proline, valine, and methionine, the matrix has a more important impact on the ionization process. The results obtained at the different concentration levels were comparable (**Supplementary Table S5**).

Concerning the carryover, the peak areas obtained in the blank samples did not exceed 2.07% of the peak area obtained at the lowest level of the calibration curve, for all the compounds. Percentages of carryover are presented in **Supplementary Figure S3**.

QC Samples

The relative standard deviation (RSD) of the concentrations obtained for the QC samples was below 10.8% for all the amino acids and below 10.0% for all the organic acids. In clinical laboratories, the Levey–Jennings charts are commonly used to report values obtained for QC samples. An example of the Levey–Jennings chart obtained with the QC samples injected along all the analyses of runners' samples are given in **Supplementary Figure S4**. This type of graph is usually interpreted, thanks to the Westgard rules. According to these rules, no QC sample should exceed ±3 SD; four consecutive QC samples should not exceed +1 SD or −1 SD; two consecutive QC samples should not exceed ±2 SD, and ten consecutive QC samples should not fall on one side of the mean (Dasgupta and Wahed, 2014). In this case, QC samples met all these criteria.

Monitoring of Metabolites Levels Before, During, and After Exercise

In this study, the blood concentrations of three amino acids, namely, creatinine, choline, and taurine, increased significantly during the exercise (**Figure 2**). Results are expressed in change of magnitude factors compared to the blood concentration before running. The blood concentration of creatinine increased by a factor 1.22 during the exercise, and no statistically significant decrease was observed during the 60 min following the physical activity. The blood concentration of choline and taurine increased by factors 1.21 and 1.31 during running, respectively. A decrease was already observed 30 min after the end of the physical activity. The means of the change of magnitude factors observed at each step were compared with a one-way ANOVA test for creatinine and taurine and with a Friedman test for choline.

Regarding the organic acids, the 1,600-m running induced an increase of the blood concentration of lactic acid, malic acid, and

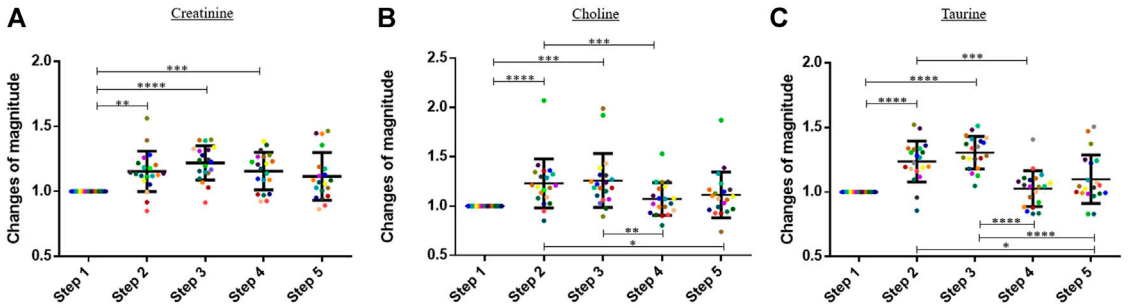

FIGURE 2 | Graphs representing the "changes of magnitude" obtained during the five steps of the study. Step 1: before physical exercise, step 2: after 400-m warm-up and 800-m running, step 3: after the completion of 1,600-m running, step 4: 30 min after finishing the exercise, and step 5: 60 min after finishing the exercise. Each color represents one participant. Black lines represent the mean ± SD of the change of magnitude factor obtained at each step ($n = 20$). **(A)** Creatinine, **(B)** choline, **(C)** taurine: amino acids for which an increase in blood concentration was observed during the exercise (*, p-value < 0.05; **, p-value < 0.01; ***, p-value < 0.001; ****, and p-value < 0.0001).

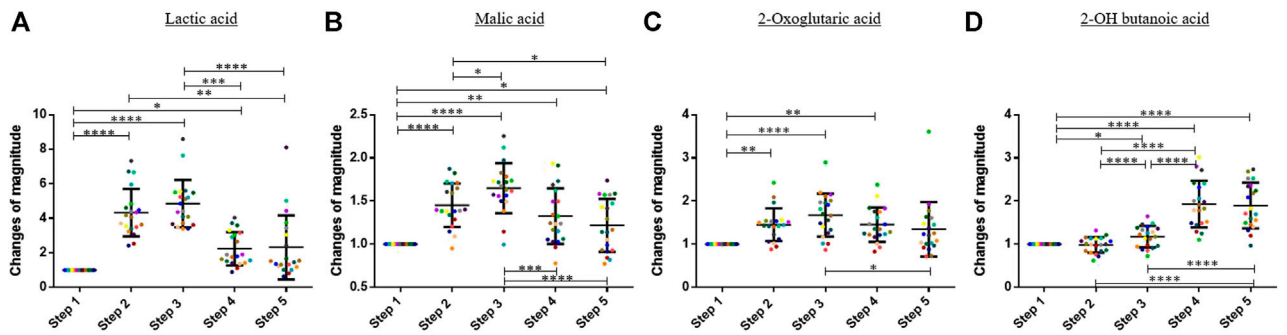

FIGURE 3 | Graphs representing the "changes of magnitude" obtained during the five steps of the study. Step 1: before physical exercise, step 2: after 400-m warm-up and 800-m running, step 3: after the completion of 1,600-m running, step 4: 30 min after finishing the exercise, and step 5: 60 min after finishing the exercise. Each color represents one participant. Black lines represent the mean ± SD of the change of magnitude factor obtained at each step ($n = 20$). **(A)** Lactic acid, **(B)** malic acid, **(C)** oxoglutaric acid, **(D)** 2-OH butanoic acid: organic acids for which an increase in blood concentration was observed during and/or after the exercise (*, p-value < 0.05; **, p-value < 0.01; ***, p-value < 0.001; ****, and p-value < 0.0001).

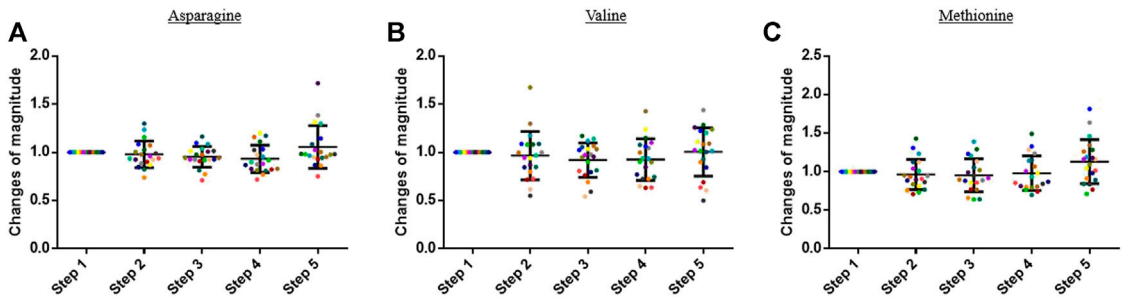

FIGURE 4 | Graphs representing the "changes of magnitude" obtained during the five steps of the study. Step 1: before physical exercise, step 2: after 400-m warm-up and 800-m running, step 3: after the completion of 1,600-m running, step 4: 30 min after finishing the exercise, and step 5: 60 min after finishing the exercise. Each color represents one participant. Black lines represent the mean ± SD of the change of magnitude factor obtained at each step ($n = 20$). **(A)** Asparagine, **(B)** valine, and **(C)** methionine: examples of compound for which no change in blood concentration was observed during the study.

2-oxoglutaric acid. The blood concentration of these compounds decreased during the 60 min after finishing the exercise (**Figure 3**). On average, the blood concentration of lactic acid, malic acid, and 2-oxoglutaric acid increased by factors 4.86, 1.65, and 1.67 between step 1 and step 3, respectively. A Friedman test was performed to compare the means of the change of magnitude factors for lactic acid, malic acid, and 2-oxoglutaric acid.

For 2-OH butanoic acid, the highest blood concentration was observed 30 min after finishing the exercise (*see* **Figure 3**). A slight decrease was observed at step 5, but this decrease was not

statistically significant. The results obtained at the different steps were compared with a one-way ANOVA for this compound.

The concentration of the other analyzed compounds did not significantly vary in this study. Some examples are presented in **Figure 4**.

DISCUSSION

In this study, we have demonstrated the potential of microsampling to follow up easily several metabolites. The hemaPEN® device offers many advantages compared to classical blood sampling. Indeed, this type of blood collection can be done by everyone and everywhere. The volume of blood collected is low, and thus, multiple collections is not an issue. Moreover, this type of sampling is less invasive than a venipuncture. The integrity of the sample is guaranteed because the paper discs on which the blood is absorbed are not accessible once the sample has been collected. They are protected from light and from humidity, thanks to a fully integrated desiccant. Compared to other microsampling devices that are available, for example, in the VAMS device from Neoteryx or the new Microsampling Wing™ from Shimadzu, four replicates of the same sample are collected with one drop of blood. This constitutes a key advantage because one paper disc can be analyzed, while the three others can be stored for future analyzes or used as backup if an issue is encountered during the first analysis. In this study, we demonstrated the potential of the hemaPEN® for the follow-up of athletes, but many other applications could be developed in the near future such as drug therapeutic monitoring or anti-doping tests. The World Anti-Doping Agency (WADA) recommends long-term sample storage and re-analysis programs for a more effective detection of doping (Word Anti-Doping Agency, 2020). In the near future, WADA also wants to develop guidelines for the collection, transport, analysis, and storage of DBS samples to implement this type of sampling in routine analyses. In this context, the hemaPEN® device could be interesting for anti-doping tests (World Anti-Doping Agency, 2020).

In this study, the most appropriate regression models were chosen, thanks to the accuracy profiles with acceptance limits below ±20%. Both UHPLC-MS/MS methods showed acceptable trueness and precision. The stability of the studied analytes in the autosampler conformed to EMA guidelines. The long-term stability at −20°C was satisfactory, except for 2-oxoglutaric acid, methionine, creatine, and taurine. Depending on the metabolites, stability might be affected during storage. Nevertheless, since relative quantification is performed in this kind of monitoring, the impact of the degradation should not be an issue, provided that all the compared samples are prepared simultaneously.

Importantly, we observed that the Levey–Jennings charts obtained for the QC samples during the analysis of the runners' samples met all the criteria described in the Westgard rules. All the samples were therefore analyzed in appropriate conditions.

Since the cohort was small (20 individuals), statistical analysis was performed using the one-way ANOVA or Friedman test. Since the reliability of the analytical process involving self-microsampling was demonstrated in this study, it would be interesting to conduct additional studies involving larger cohorts to investigate the influence of different parameters such as BMI, sex, age, ethnicity, intensity of the efforts, and genetic predisposition, on metabolic changes. Such studies could include a broader panel of targeted metabolites or could be done by untargeted metabolomics. In those cases, dedicated multivariate analysis should be employed (i.e., PCA analysis).

Concerning the change in magnitude factors observed in our study, low variation in blood concentrations of amino acids was observed. Nevertheless, an increase in the blood concentration of taurine and creatinine was observed. Cuisinier *et al.* and Medelli *et al.* already described an increased concentration of taurine in plasma after a marathon (Cuisinier et al., 2001) and after a cycling competition (Medelli et al., 2003). Several other studies mentioned an increase in serum creatinine levels after physical activity (Hodgson et al., 2017). This increase in creatinine can be explained by an increased muscle breakdown (Hodgson et al., 2017; Omassoli et al., 2019). Pechlivanis *et al.* observed a decrease in leucine and valine serum concentrations after sprint running (Pechlivanis et al., 2013). After a marathon, Stander *et al.* observed a reduced concentration of methionine, valine, and leucine (Stander et al., 2018). Nevertheless, for the amino acids, conclusions are not consistent across all the studies (Schranner et al., 2020). The duration and the intensity of exercise or the fitness of the participants could explain these discrepancies between studies. A decrease in blood concentrations of amino acids can be explained by the fact that the body activates the catabolism of these molecules to find energy when carbohydrate stores are depleted. In our study, exercise duration was short. This is probably why the body may not have used all of its carbohydrate stores. Therefore, the concentrations of the amino acids were not decreased. Future studies with a longer exercise duration should be conducted to confirm this hypothesis.

Concerning the organic acids, lactate represents one key compound. Indeed, several articles already described an increase in lactate during exercise (Pechlivanis et al., 2013; Gáspari et al., 2015). This accumulation in lactate reflects the shift from an aerobic to an anaerobic energy production system. Monitoring this shift constitutes an essential parameter to adapt athletes' training in endurance sports (Etxegarai et al., 2019). In our study, the blood concentration of lactic acid was found to increase significantly during exercise, showing the potential of our method to follow up athletes. Thirty minutes after the end of the physical exercise, the concentration of lactic acid decreased to reach a concentration near the nominal one. The blood concentration of two other organic acids, that is, malic acid and 2-oxoglutaric acid, intermediates in the TCA cycle, also increased during physical effort. This finding corresponds to the conclusions previously obtained in other studies. Indeed, Stander *et al.* found elevated serum concentrations of these molecules after a marathon (Stander et al., 2018). Recently Schranner *et al.* published a review bringing together 27 studies and 57 experiments. They concluded that the blood concentrations of intermediates of the TCA cycle such as malate or 2-oxoglutaric acid were upregulated by exercise as well as the concentration of lactate. This increase in lactate was particularly noted with high-intensity exercise. The concentration of these three organic acids decreased during the 60 min after physical activity, as observed in our study (Schranner et al., 2020). Schranner's review also reported an increase in the concentration of ketone bodies during exercise. In our study, 2-hydroxybutanoic acid, a ketone body, was found to increase during and after exercise. Therefore, the

conclusions obtained in this study for the organic acids are in agreement with the literature.

Finally, this study demonstrates that microsampling with hemaPEN® followed by targeted metabolomics has a great potential for the follow-up of metabolic profiles. This study used two UHPLC-MS/MS methods to quantify 13 metabolites using only 2.74 µl of whole blood and offers the main advantage that sampling can be performed directly at the training site without special skill. In the future, targeted and untargeted metabolomics approaches on broader cohorts could be investigated to expand the panel of applications.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

This study involving human participants was reviewed and approved by the Ethical Committee of the University of Liège (Liège, Belgium) Nr Eudra CT: B7072020000041; ref: 2020/182. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CN: experiments and redaction. MH: experiments. GC: technical support. A-CS: data processing and manuscript handling. MF: design of the study and article handling.

FUNDING

This research was conducted by the financial support provided by the National Fund for Scientific Research (FNRS) (CN: ID 40001038 and MH: ID 31239052), the European funds of regional development (FEDER), and Walloon Region of Belgium as part of the operational program "Walloon-2020. EU".

ACKNOWLEDGMENTS

Authors thank all the participants of this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.681400/full#supplementary-material

REFERENCES

Al-Khelaifi, F., Abraham, D., Diboun, I., and Elrayess, M. A. (2019). "Proteomics and metabolomics research in exercise and sport," in *Sports, Exercise, and Nutritional Genomics: Current Status and Future Directions* (Amsterdam: Elsevier), 539–566. doi:10.1016/B978-0-12-816193-7.00023-3

Aune, D., Norat, T., Leitzmann, M., Tonstad, S., and Vatten, L. J. (2015). Physical Activity and the Risk of Type 2 Diabetes: A Systematic Review and Dose-Response Meta-Analysis. *Eur. J. Epidemiol.* 30, 529–542. doi:10.1007/s10654-015-0056-z

Cuisinier, C., Ward, R. J., Francaux, M., Sturbois, X., and De Witte, P. (2001). Changes in Plasma and Urinary Taurine and Amino Acids in Runners Immediately and 24 H after a Marathon. *Amino Acids*. 20, 13–23. doi:10.1007/s007260170062

Dasgupta, A., and Wahed, A. (2014). "Laboratory Statistics and Quality Control," in *Clin. Chem. Immunol. Lab. Qual. Control.* (Amsterdam: . Elsevier), 47–66. doi:10.1016/b978-0-12-407821-5.00004-8

Deprez, S., Paniagua-González, L., Velghe, S., and Stove, C. P. (2019). Evaluation of the Performance and Hematocrit Independence of the HemaPEN as a Volumetric Dried Blood Spot Collection Device. *Anal. Chem.* 91, 14467–14475. doi:10.1021/acs.analchem.9b03179

Etxegarai, U., Portillo, E., Irazusta, J., Koefoed, L., and Kasabov, N. (2021). A Heuristic Approach for Lactate Threshold Estimation for Training Decision-Making: An Accessible and Easy to Use Solution for Recreational Runners. *Eur. J. Oper. Res.* 291, 427–437. doi:10.1016/j.ejor.2019.08.023

Gáspari, A. F., Berton, R., Lixandrão, M. E., Perlotti Piunti, R., Chacon-Mikahil, M. P. T., and Bertuzzi, R. (2015). The Blood Lactate Concentration Responses in a Real Indoor Sport Climbing Competition. *Sci. Sports*. 30, 228–231. doi:10.1016/j.scispo.2015.05.002

Hodgson, L., Walter, E., Venn, R., Galloway, R., Pitsiladis, Y., Sardat, F., et al. (2017). Acute Kidney Injury Associated with Endurance Events-Is it a Cause for Concern? A Systematic Review. *BMJ Open Sport Exerc. Med.* 3, e000093. doi:10.1136/bmjsem-2015-000093

Julien, C., Castelli, D., Bray, D., Lee, S., Burson, S., and Jung, Y. (2021). Project SMART: A Cooperative Educational Game to Increase Physical Activity in Elementary Schools. *Smart Health*. 19, 100163. doi:10.1016/j.smhl.2020.100163

Kelly, R. S., Kelly, M. P., and Kelly, P. (2020). Metabolomics, Physical Activity, Exercise and Health: A Review of the Current Evidence. *Biochim Biophys Acta Mol Basis Dis.* 1866, 165936. doi:10.1016/j.bbadis.2020.165936

Kok, M. G. M., Nix, C., Nys, G., and Fillet, M. (2019). Targeted Metabolomics of Whole Blood Using Volumetric Absorptive Microsampling. *Talanta*. 197, 49–58. doi:10.1016/j.talanta.2019.01.014

Lewis, G. D., Farrell, L., Wood, M. J., Martinovic, M., Rowe, G. C., Souza, A., et al. (2010). Exercise Metabolomics. *Sci. Transl Med.* 2, 33ra37. doi:10.1126/scitranslmed.3001006.Metabolic

Matta, P. N., Baul, T. D., Loubeau, K., Sikov, J., Plasencia, N., Sun, Y., et al. (2021). Low Sports Participation Is Associated with Withdrawn and Depressed Symptoms in Urban, School-Age Children. *J. Affective Disord.* 280, 24–29. doi:10.1016/j.jad.2020.11.076

Matuszewski, B. K., Constanzer, M. L., and Chavez-Eng, C. M. (2003). Strategies for the Assessment of Matrix Effect in Quantitative Bioanalytical Methods Based on HPLC–MS/MS. *Anal. Chem.* 75, 3019–3030. doi:10.1021/ac020361s

Medelli, J., Lounana, J., and Hill, D. (2003). Variation in Plasma Amino Acid Concentrations during a Cycling Competition. *J. Sports Med. Phys. Fitness* 43, 236–242.

Mielke, G., Bailey, T., Burton, N., and Brown, W. (2019). Associations between Participation in Recreational Sports with Hypertension, Diabetes and Obesity in a Cohort of Australian Adults. *J. Sci. Med. Sport*. 22, S56. doi:10.1016/j.jsams.2019.08.249

Omassoli, J., Hill, N. E., Woods, D. R., Delves, S. K., Fallowfield, J. L., Brett, S. J., et al. (2019). Variation in Renal Responses to Exercise in the Heat with Progressive Acclimatisation. *J. Sci. Med. Sport* 22, 1004–1009. doi:10.1016/j.jsams.2019.04.010

Pechlivanis, A., Kostidis, S., Saraslanidis, P., Petridou, A., Tsalis, G., Veselkov, K., et al. (2013). 1H NMR Study on the Short- and Long-Term Impact of Two Training Programs of Sprint Running on the Metabolic Fingerprint of Human Serum. *J. Proteome Res.* 12, 470–480. doi:10.1021/pr300846x

Porter, A. K., Schilsky, S., Evenson, K. R., Florido, R., Palta, P., Holliday, K. M., et al. (2019). The Association of Sport and Exercise Activities with Cardiovascular Disease Risk: The Atherosclerosis Risk in Communities (ARIC) Study. *J. Phys. Act. Heal.* 16, 698–705. doi:10.1123/jpah.2018-0671

Prado, E., Souza, G. H. M. F., Pegurier, M., Vieira, C., Lima-Neto, A. B. M., Assis, M., et al. (2017). Non-targeted Sportomics Analyses by Mass Spectrometry to Understand Exercise-Induced Metabolic Stress in Soccer Players. *Int. J. Mass Spectrom.* 418, 1–5. doi:10.1016/j.ijms.2017.02.002

Protti, M., Marasca, C., Cirrincione, M., Cavalli, A., Mandrioli, R., and Mercolini, L. (2020). Assessment of Capillary Volumetric Blood Microsampling for the Analysis of Central Nervous System Drugs and Metabolites. *Analyst*. 145, 5744–5753. doi:10.1039/d0an01039a

Sakaguchi, C., Nieman, D., Signini, E., Abreu, R., and Catai, A. (2019). Metabolomics-based Studies Assessing Exercise-Induced Alterations of the Human Metabolome: A Systematic Review. *Metabolites*. 9, 164. doi:10.3390/metabo9080164

Schranner, D., Kastenmüller, G., Schönfelder, M., Römisch-Margl, W., and Wackerhage, H. (2020). Metabolite Concentration Changes in Humans after a Bout of Exercise: a Systematic Review of Exercise Metabolomics Studies. *Sports Med Open.* 6. doi:10.1186/s40798-020-0238-4

Smith, G. (2012). European Medicines Agency Guideline on Bioanalytical Method Validation: What More Is There to Say?. *Bioanalysis*. 4, 865–868. doi:10.4155/bio.12.44

Soligard, T., Schwellnus, M., Alonso, J.-M., Bahr, R., Clarsen, B., Dijkstra, H. P., et al. (2016). How Much Is Too Much? (Part 1) International Olympic Committee Consensus Statement on Load in Sport and Risk of Injury. *Br. J. Sports Med.* 50, 1030–1041. doi:10.1136/bjsports-2016-096581

Stander, Z., Luies, L., Mienie, L. J., Keane, K. M., Howatson, G., Clifford, T., et al. (2018). The Altered Human Serum Metabolome Induced by a Marathon. *Metabolomics*. 14, 1–11. doi:10.1007/s11306-018-1447-4

Trajan-HemaPEN brochure (2020). Available at: https://cdn.shopify.com/s/files/1/0767/9441/files/BR-0536-G_RevC.pdf?v=1593616400 (Accessed January 7, 2021).

Vishwanathan, K., Tackett, R. L., Stewart, J. T., and Bartlett, M. G. (2000). Determination of Arginine and Methylated Arginines in Human Plasma by Liquid Chromatography-Tandem Mass Spectrometry. *J. Chromatogr. B: Biomed. Sci. Appl.* 748, 157–166. doi:10.1016/s0378-4347(00)00399-6

WHO (2018). *Global Action Plan on Physical Activity 2018-2030: More Active People for a Healthier World*. Available at: https://www.who.int/publications/i/item/9789241514187 (Accessed January 4, 2021).

WHO (2020). *Physical activity*. Available at: https://www.who.int/news-room/fact-sheets/detail/physical-activity (Accessed January 4, 2021).

World Anti-Doping Agency (2020). *WADA Welcomes Enhanced Long-Term Sample Storage and Re-analysis Program*. Available at: https://www.wada-ama.org/en/media/news/2020-12/wada-welcomes-enhanced-long-term-sample-storage-and-re-analysis-program (Accessed January 11, 2021).

Conflict of Interest: 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.

The handling editor declared a past collaboration with one of the authors (MF).

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 15 June 2021
doi: 10.3389/fmolb.2021.680159

Urine NMR Metabolomics Profile of Preterm Infants With Necrotizing Enterocolitis Over the First Two Months of Life: A Pilot Longitudinal Case-Control Study

Jean-Charles Picaud[1], *Anna De Magistris*[2], *Michele Mussap*[3]*, *Sara Corbu*[3], *Angelica Dessì*[3], *Antonio Noto*[4], *Vassilios Fanos*[3] *and Flaminia Cesare Marincola*[5]

[1]Neonatology Unit, Croix-Rousse University Hospital, Hospices Civils de Lyon, Lyon, France, [2]Pediatrics and Neonatology Division of, Azienda USL Romagna, Santa Maria Delle Croci Hospital, Ravenna, Italy, [3]Department of Surgical Sciences, University of Cagliari, Cagliari, Italy, [4]Department of Biomedical Sciences, University of Cagliari, Cagliari, Italy, [5]Department of Chemical and Geological Sciences, Cittadella Universitaria di Monserrato, University of Cagliari, Cagliari, Italy

OPEN ACCESS

Edited by:
Danuta Dudzik,
Medical University of Gdańsk, Poland

Reviewed by:
Simon Eaton,
University College London, United Kingdom
Sanjay Keshav Patole,
KEM Hospital for Women, United States

***Correspondence:**
Michele Mussap
mumike153@gmail.com

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

Received: *13 March 2021*
Accepted: *20 May 2021*
Published: *15 June 2021*

Citation:
Picaud J-C, De Magistris A, Mussap M, Corbu S, Dessì A, Noto A, Fanos V and Cesare Marincola F (2021) Urine NMR Metabolomics Profile of Preterm Infants With Necrotizing Enterocolitis Over the First Two Months of Life: A Pilot Longitudinal Case-Control Study. Front. Mol. Biosci. 8:680159. doi: 10.3389/fmolb.2021.680159

Objective: To investigate changes in the urine metabolome of very low birth weight preterm newborns with necrotizing enterocolitis (NEC) and feed intolerance, we conducted a longitudinal study over the first 2 months of life. The metabolome of NEC newborns was compared with two control groups that did not develop NEC: the first one included preterm babies with feed intolerance, while the second one preterm babies with good feed tolerance.

Methods: Newborns developing NEC within the 3 weeks of life were identified as early onset NEC, while the remaining as late onset NEC. Case-control matching was done according to the gestational age (±1 week), birth weight (± 200 g), and postnatal age. A total of 96 urine samples were collected and analyzed. In newborns with NEC, samples were collected before, during and after the diagnosis over the first 2 months of life, while in controls samples were collected as close as possible to the postnatal age of newborns with NEC. Proton nuclear magnetic resonance (^{1}H NMR) spectroscopy was used for metabolomic analysis. Data were analyzed by univariate and multivariate statistical analysis.

Results: In all the preterm newborns, urine levels of betaine, glycine, succinate, and citrate positively correlated with postnatal age. Suberate and lactate correlated with postnatal age in preterms with NEC and in controls with food intolerance, while *N,N*-dimethylglycine (*N,N*-DMG) correlated only in controls with good digestive tolerance. Preterm controls with feed intolerance showed a progressive significant decrease of *N*-methylnicotinamide and carnitine. Lactate, betaine, myo-inositol, urea, creatinine, and *N,N*-dimethylglycine discriminated late-onset NEC from controls with good feed tolerance.

Conclusion: Our findings are discussed in terms of contributions from nutritional and clinical managements of patients and gut microbiota.

Keywords: metabolomics, proton nuclear magnetic resonance spectroscopy, necrotizing enterocolitis, prematurity, urine

INTRODUCTION

Cellular metabolism plays a crucial role both in health and disease, mirroring interactions between the host genome, environment, and microbiome. Environmental and lifestyle factors as well as traumatic events or diseases, have the potential to alter the individual metabolic phenotype both directly, by inducing perturbations in various metabolic pathways, and indirectly, by promoting epigenetic changes, which in turn lead to changes in gene expression, transcripts, and ultimately in the metabolic profile of a given cell, tissue, or biological fluid (Eicher et al., 2020). However, the organism's rapid response to any exogenous or endogenous factor altering the cellular and tissue homeostasis (e.g., asphyxia, sepsis, gut dysbiosis) can be unveiled and monitored overt time NEITHER by genomics nor by transcriptomics and proteomics. These "omics" offer specific but tardive information on any biological change; conversely, metabolomics, the science that identifies and quantifies endogenous and exogenous metabolites, represents an "instant omics" capable to provide information on the current status of a living system (Ashrafian et al., 2020). Indeed, changes in the individual metabolic profile occur much earlier than any clinically detectable sign or symptom, and thus, metabolomics is strategic in the context of the precision medicine approach (Karczewski and Snyder, 2018). One of the most important applications of metabolomic studies is the early identification of critically ill newborns at risk of adverse clinical outcomes during their stay in the neonatal intensive care unit (NICU) (Fanos et al., 2018; Bardanzellu et al., 2020; Locci et al., 2020). Necrotizing enterocolitis (NEC) is a life-threatening disease affecting almost exclusively preterm newborns, consisting of abnormal intestinal colonization followed by an immune-inflammatory response leading to the loss of intestinal barrier function and possible perforation of the intestine (Neu and Walker, 2011; Meister et al., 2020; Neu, 2020). The pooled estimated NEC incidence in very low birth weight (VLBW) newborns is approximately 7% (Alsaied et al., 2020), while it is less common in late premature and in full term newborns. NEC's pathogenesis is primarily marked by an abnormal inflammatory response and necrosis of the gut mucosa along the whole gastrointestinal tract (Niño et al., 2016). Further risk factors include sepsis, enteral formula feeding, prolonged antibiotic exposure, and gut dysbiosis (Raba et al., 2021). Recently, it was demonstrated that NEC is associated with elevated blood levels of CCR9 + CD4 + T cells as well as CCR9 + interleukin-17 (IL-17) producing Treg (previously called regulatory T cells); the histological NEC severity is positively and negatively correlated with their gut and blood concentration, respectively (Ma et al., 2019). As a consequence, the therapeutic modulation of lymphocyte balance may open new perspectives for improving NEC severity and outcome (Nguyen and Sangild., 2019). Recommendations on feeding practices, such as breastfeeding, the proper management of feeding intolerance, the application of feeding guidelines, and the implementation of probiotics with diet, can prevent NEC onset in critically ill newborns admitted in NICU (Bi et al., 2019). Several studies, recently revised in an elegant review (Agakidou et al., 2020), have investigated the metabolic profile of blood, plasma, serum, urine, stools, and intestinal epithelial cells in preterm neonates with NEC, opening new horizons on the molecular mechanisms associated with the disease and searching candidate biomarkers for the early diagnosis and prognosis of NEC. This pilot study aimed to explore the presence of the urinary metabolic signature in VLBW preterm newborns by using a proton nuclear magnetic resonance spectroscopy (^{1}H NMR)-based metabolomic approach. We investigated the dynamic changes of the urine metabolome in infants with NEC over the first 2 months of life by collecting samples at different time points, namely before, during, and after the diagnosis. Since all babies with NEC were also affected by feeding intolerance (FI), each of them was matched with two preterm newborns without NEC, the first one with FI and the second one with good digestive tolerance.

MATERIALS AND METHODS

Patients

This case-control study was conducted in the NICU, Hospital de la Croix Rousse (HCR), Hospices Civils de Lyon, Lyon, France. The study was approved by the local Ethics Committee (Comité de Protection des Personnes Sud-Est IV, Lyon) and performed following the approved guidelines. Infant parents signed informed consent forms before participation. We considered eligible for the study VLBW preterm babies recruited prospectively. Eighteen VLBW preterm infants were included: 6 with NEC and feeding intolerance (group 1, NEC); 6 with feeding intolerance without any sign of NEC (group 2, FI); and 6 with good digestive tolerance without NEC (group 3, GDT). We considered feeding intolerance the inability of the baby to ingest and digest enteral nutrition. This condition became clinically evident with the appearance of (a) gastric residues (more than 50% of the ingested food after 2–3 consecutive meals); (b) biliary or hemorrhagic color of residues; (c) abdominal distension with discomfort on palpation and gaseous dilation of the loops of the small intestine (Lucchini et al., 2011). Good feeding tolerance was defined as the ability of the preterm infant to safely ingesting and digesting the prescribed enteral feeding without complications associated with aspiration, infection, and gastrointestinal dysfunction (Shulman et al., 1998). In group 1, three babies developed the disease within the first 3 weeks of life (hereafter called *early-onset* NEC), while the remaining three developed NEC 6–8 weeks after birth (hereafter called *late-onset* NEC). Each baby belonging to group 1 NEC was matched with two babies, the first one belonging to group 2 FI and the second one to control group 3 GDT. In order to reduce uninformative variations that could interfere with the identification of relevant information encoded in the experimental spectral dataset, controls were selected according to matched gestational age (± 1 week), birth weight (± 200 g), and postnatal age at the time of urine sampling (± 7 days). NEC was defined as the presence of clinical evidence fulfilling modified Bell's stage criteria (Bell et al., 1978; Juhl et al., 2019) and was confirmed by radiological pneumatosis intestinalis. All NEC cases were Bell stage II. Neonates with major congenital abnormalities (including those of the

gastrointestinal tract) were excluded as controls from the study. Total parenteral nutrition was used for all infants up to 2–3 weeks of life (Darmaun et al., 2018). As soon as tolerated, enteral feeding was gradually introduced using a milk bank or expressed breast milk provided by their mother. According to the European Milk Bank Association (EMBA) Working Group recommendations (Arslanoglu et al., 2019), milk was fortified in both cases. Fortema® (Bledina, Villefranche-sur-Saône, France) for protein and carbohydrate intakes, Liquigen® (Nutricia, Saint-Ouen, France) for lipid intakes and the multicomponent fortifier Nutriprem® (Bledina, Villefranche-sur-Saône, France) were used. In babies manifesting feeding intolerance, enteral nutrition was interrupted and then gradually reintroduced when clinical conditions went back to normal. During the study period, no change was made in the enteral parenteral nutritional policy.

Sample Collection and Preparation

Urine samples were collected over approximately two months after birth. In babies with NEC (group 1), urine samples were collected at various intervals: before, during, and after the disease's onset. In babies belonging to groups 2 FI and 3 GDT, samples were collected at the day of life as close as possible to those collected in babies with NEC. Samples (volume 1–2 ml) were collected using a cotton ball inserted into the disposable diaper; the urine was aspired by a syringe and then transferred to a sterile 2 ml vial and immediately frozen at −80°C until their shipping to the metabolab of the University of Cagliari. Before analysis, 800 µL of thawed urine were transferred into a 1.5 ml centrifuge microtube, and then 8 µL of sodium azide (10% w/w) were added to avoiding any possible bacterial growth. The sample was then centrifuged at 12,000 *g* for 10 min at + 4°C. To stabilize the pH of urine samples, 630 µL of supernatant were mixed with 70 µL of phosphate buffer solution [1.5M KH_2PO_4, 1% sodium 3-trimethylsilyl-propionate-2,2,3,3-d4 (TSP, 98 atom % D), pH 7.4]. Finally, 650 µL were placed into a 5 mm wide NMR tube.

Proton Nuclear Magnetic Resonance Spectroscopy Analysis

The analysis was conducted at 300K by using a Varian UNITY INOVA 500 spectrometer (Agilent Technologies, CA, United States) operating at 499.839 MHz. A standard 1-D pulse sequence NOESY was used with water suppression. For each urine spectrum, a total of 128 scans were collected in 64k data points over a spectral width of 6,000 Hz using a relaxation delay of 2 s, an acquisition time of 1.5 s, and mixing of 0.1 s. Before Fourier transformation, the free induction decay was multiplied with 0.3 Hz exponential line broadening spectra. All spectra were phased, and baseline corrected using MestReNova (Version 8.1, Mestrelab Research SL, Santiago de Compostela, Spain). The chemical shift scale was set by assigning a value of δ = 0.00 ppm to the internal standard TSP signal. After correction for misalignments in chemical shift, primarily due to pH-dependent signals and deleting the regions containing the water and TSP signals, the NMR spectra were binned into 0.0025 ppm buckets over a chemical shift range of 0.5–9.5 ppm. Bins were normalized to the sum of total spectral area to compensate for the overall concentration differences and used as a dataset (97 x 4,423) for multivariate analysis. The assignment of the metabolites in the ^{1}H NMR spectra was performed according to literature data (Diaz et al., 2016; Scalabre et al., 2017), the Human Metabolome database, available at http://www.hmdb.ca (Wishart et al., 2018), and Chenomx NMR suite 8.1 software (evaluation version, Chenomx, Edmonton, Canada). All the samples were analyzed simultaneously.

Data Processing and Statistical Analysis

For multivariate statistical analysis, data were Pareto scaled (Misra, 2020). Multivariate statistical analysis of the NMR dataset consisted of principal component analysis (PCA), orthogonal projection to latent structures (OPLS) regression, and orthogonal projection to latent structures discriminant analysis (OPLS-DA) supported by the SIMCA software (version 16.0, Umetrics, Umeå, Sweden). OPLS was used in the case of a continuous Y-matrix (i.e., multiple time points) and OPLS-DA for identifying discriminant metabolites in a pairwise comparison between case and control groups. The quality of OPLS and OPLS-DA models were evaluated through the following parameters: the cumulative values of total Y explained variance, i.e., goodness of fit (R^2Y), and the Y predictable variation, i.e., the goodness of predictability (Q^2). The latter was extracted by the default method of 7-fold internal cross-validation of SIMCA. Additionally, the models were tested for overfitting using permutation testing (n = 400). The models' significance was further assessed by an ANOVA based on the cross-validated predictive residuals (CV-ANOVA) with a p-value ≤ 0.05 (Eriksson et al., 2008). The models were considered valid if the permutation test and the CV-ANOVA test were significant. The variables with the most significant contributions to OPLS and OPLS-DA models were identified by exploring the correlation coefficients line plots by following two criteria: absolute p and $p(corr)$ values were set to be greater than 0.05 and 0.5, respectively. p-value represents each variable's importance and $p(corr)$ its reliability (Cloarec et al., 2005). The univariate statistical analysis was performed by the GraphPad Prism Statistics software package, version 8.1.2 (GraphPad Prism Software Inc., San Diego, CA, United States), to measure the Pearson's correlation coefficient between metabolites and postnatal time and compare the variation in the abundance of discriminant metabolites between groups. The magnitude of variation was evaluated by calculating the effect size (ES) adjusted for small sample number (Berben et al., 2012). Effect sizes were classified small between 0.2 and 0.5, medium between 0.5 and 0.8, and large when greater than 0.8. The non-parametric Mann-Whitney U test was used for the univariate statistical approach; a p-value ≤ 0.05 was considered statistically significant.

RESULTS

The main characteristics of the study population are reported in **Table 1**. Group 1 NEC does not significantly differ from group 2 FI and 3 GDT for gestational age, birth weight, mean Apgar scores at 5 min ($p > 0.05$). The male:female ratio was 1:1 for each group. A total of 97 urine samples were collected

TABLE 1 | Characteristics of study population.

Variable, description	NEC (*n* = 6)	Controls (*n* = 12)	
		FI-PT (*n* = 6)	GDT-PT (*n* = 6)
Gestational age (weeks, mean ± SD)	27.1 ± 1.6	27.2 ± 1.3	27.7 ± 1.6
Male/Female, n	3/3	3/3	3/3
Birth weight (g, means ± SD)	1,016 ± 104	920 ± 104	950 ± 65
Cesarean section delivery, n	2	5	3
IUGR, n	1	4	1
Apgar score: ≤5 at 5 min, n	1	2	2
Early-onset of NEC (< 25 days), n	3	/	/
Late-onset of NEC (> 40 days), n	3	/	/
Antibiotics, n	6	6	6

Abbreviation: FI-PT, preterm with feed intolerance; GDT-PT, preterm with good digestive tolerance; IUGR, intrauterine growth restriction; NEC, necrotizing enterocolitis; SD, standard deviation.

during the first 2 months of life: 38 in group 1 NEC (18 *early-onset* NEC, 20 *late-onset* NEC), 27 in group 2 FI, and 32 in control group 3 GDT. A preliminary PCA analysis was performed for searching any inherent separation among samples and the presence of outliers. The total amount of variance explained by the first two principal components (PCs) was 40%. The PC1 *vs.* PC2 scores plot clearly indicates the absence of any cluster (**Figure 1A**). Conversely, the scores plot unveils a similar, unidirectional temporal trend for groups 2 FI and 3 GDT, scores shifting from the left to the right side of the PC1 axis as age increases (**Figure 1B**). The temporal shift of the urine metabolome of the *early-onset* NEC subgroup follows the same trajectory along the PC1 axis as those exhibited by groups 2 and 3 (**Figures 2A–C**). On the other hand, the time course of the *late-onset* NEC subgroup is similar to that of the other groups, but only until the diagnosis of the disease; soon after, the trajectory inverts the direction, approaching the metabolic profile of the first days of life (**Figures 2D–F**). The PCA loadings plot revealed that this behavior was mainly related to the urinary gluconate concentration (**Supplementary Figure S1**). Indeed, in the urine sample of the *late-onset* NEC subgroup, rigorously collected after the diagnosis, the content of gluconate was much higher than that found in the urine of the same babies collected just before the diagnosis; moreover, gluconate concentration was comparable with that observed in the urine collected during the first day of life in all the neonates. Gluconate is a nutrient degraded by gluconokinase to generate 6-phosphogluconate, playing a crucial physiological role (Ramachandran et al., 2006; Riganti et al., 2012); however, during the total parenteral nutrition (TPN) at birth and after the clinical diagnosis of NEC, the intravenous calcium administration was a source of exogenous gluconate in babies with NEC. The subgroup of *early-onset* NEC babies developed the disease during TPN, while in the *late-onset* subgroup, NEC was diagnosed during enteral nutrition (EN). In detail, the latter received two cycles of TPN, namely at birth and after the onset of NEC, alternated by an EN cycle. Based on these findings, the analysis of the temporal trajectories of *late-onset* NEC scores in the PCA model showed that the inversion of the trend observed after the diagnosis was associated with the introduction of the second cycle of TPN therapy, leading to the reasonable conclusion that the nutritional intervention may be the main source of this dynamic modification. Since the urine NMR spectra of all infants under TPN showed very intense signals due to exogenous gluconate, we removed these signals before statistical analysis to avoid any TPN contribution to the spectral profile. A more detailed insight into the time dependence of urine metabolome of group 1 NEC and groups 2 and 3 was undertaken by OPLS regression by using the ^{1}H NMR urine spectral data as an independent variable and the postnatal sampling days as Y-variable. An OPLS model was built separately for each group. Model performances are summarised in **Table 2**, while **Figure 3** depicts the corresponding scores and loading plots. The models built for group 2 FI and 3 GDT demonstrated good modeling and predictive abilities, while a lower but acceptable predictivity characterized the model for group 1 NEC. The robustness of the models was validated by the permutation test (n = 400) and CV-ANOVA. The corresponding loadings plots (**Figures 3D–F**) allow identifying the most significant metabolic signature associated with postnatal age. Metabolites positively correlating with postnatal age were: betaine, glycine, citrate, and succinate in all the groups; *N,N*-dimethylglycine (DMG) in control group 3 GDT; suberate and lactate in group 1 NEC and group 2 FI; creatinine in control group 3 GDT only. Besides, unassigned resonances at δ = 3.95 and δ = 3.74 were inversely correlated with postnatal age in all the groups, while carnitine and *N*-methylnicotinammide (*N*-MNA) only in group 2 FI. These time-dependent changes were also investigated by the univariate statistical analysis; in particular, the Pearson's correlation coefficient (r) was computed. Results confirmed the statistical significance of the findings mentioned above- ($p < 0.05$) (**Supplementary Table S1**). To identify the metabolic signature(s) associated with NEC, the two subgroups of *early*-and *late-onset* NEC were compared with the corresponding matched groups 2 and 3 by using the OPLS-DA approach. We included only spectra of urine collected either immediately before or at the diagnosis time, based on the assumption that they may provide more information on the presence of metabolic perturbations

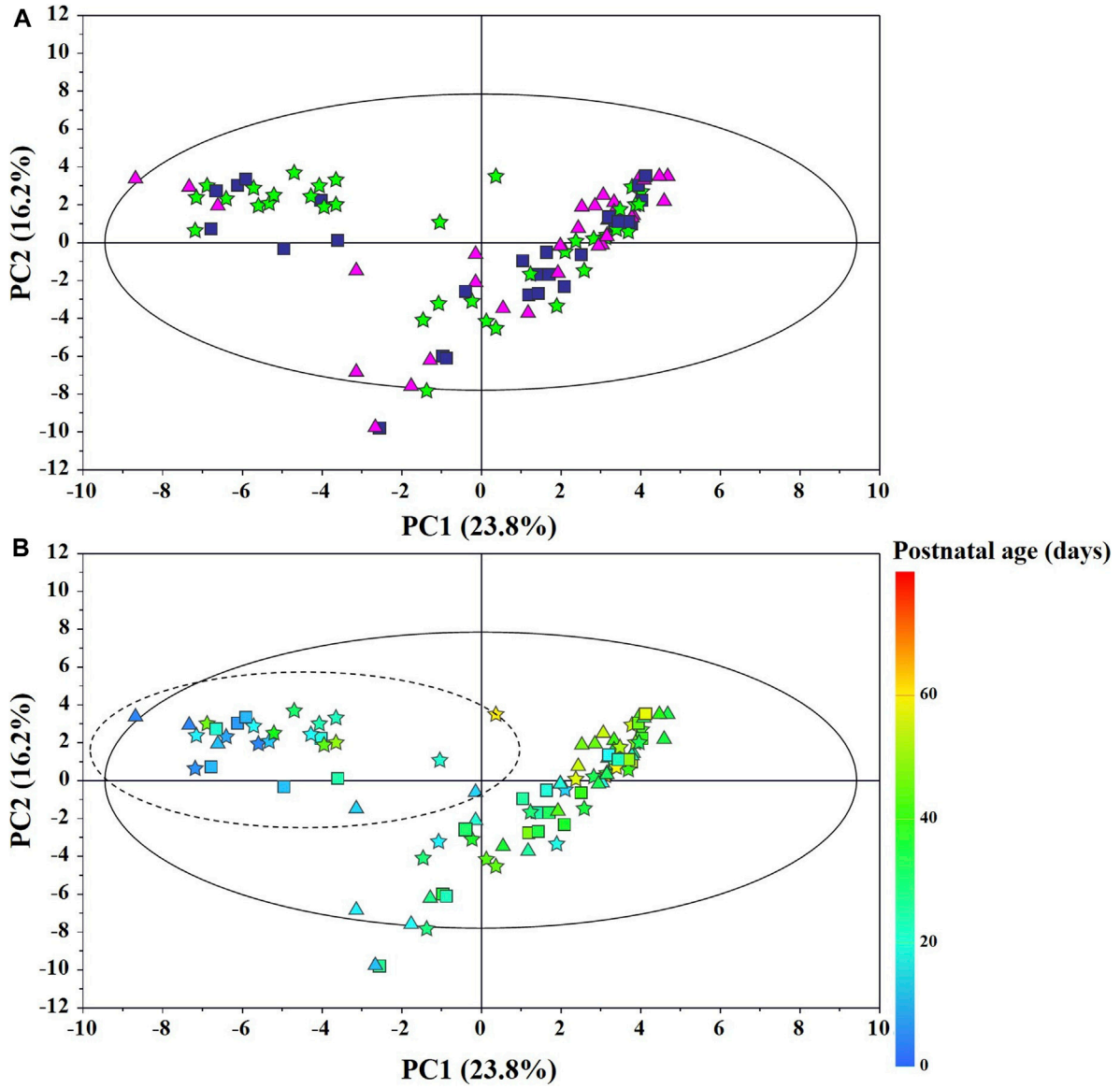

FIGURE 1 | PCA scores plot from the model built with the ^{1}H-NMR spectra of infant urine samples: **(A)** ★, group 1 NEC; ■, group 2 FI (feed intolerance without NEC); ▲, control group 3 GDT (good digestive tolerance without NEC); **(B)** Scores are coloured according to the postnatal age. Samples in the dotted circle were collected during total parenteral nutrition (TPN).

associated with the disease. In **Table 3**, we reported the quality parameters of the pairwise OPLS-DA models. No model comparing the *early-onset* NEC subgroup with groups 2 and 3 was found significant; conversely, the model comparing the *late-onset* NEC subgroup with control group 3 GDT showed a significant group separation (p = 0.02) as reported in **Figure 4A**. The analysis of the OPLS-DA loadings plot (**Figure 4B**) showed the leading metabolites responsible for sample discrimination, providing an assessment of the main statistically significant differences between the two groups: lactate was more abundant in the *late-onset* NEC subgroup, while *N,N*-DMG, betaine, creatinine, myo-inositol, and urea were more abundant in group 3. These findings were further supported by the univariate statistical analysis for assessing significant differences in the relative content of these metabolites between the two groups (**Figure 5**).

DISCUSSION

Despite the impressive body size of literature on clinical metabolomics-based studies in human disease, the number of metabolomics-based studies on NEC is small, and the identification of reliable candidate omics-based biomarkers for the prediction and the early diagnosis of NEC is still far from being definitive. After the exclusion of metabolomics-based

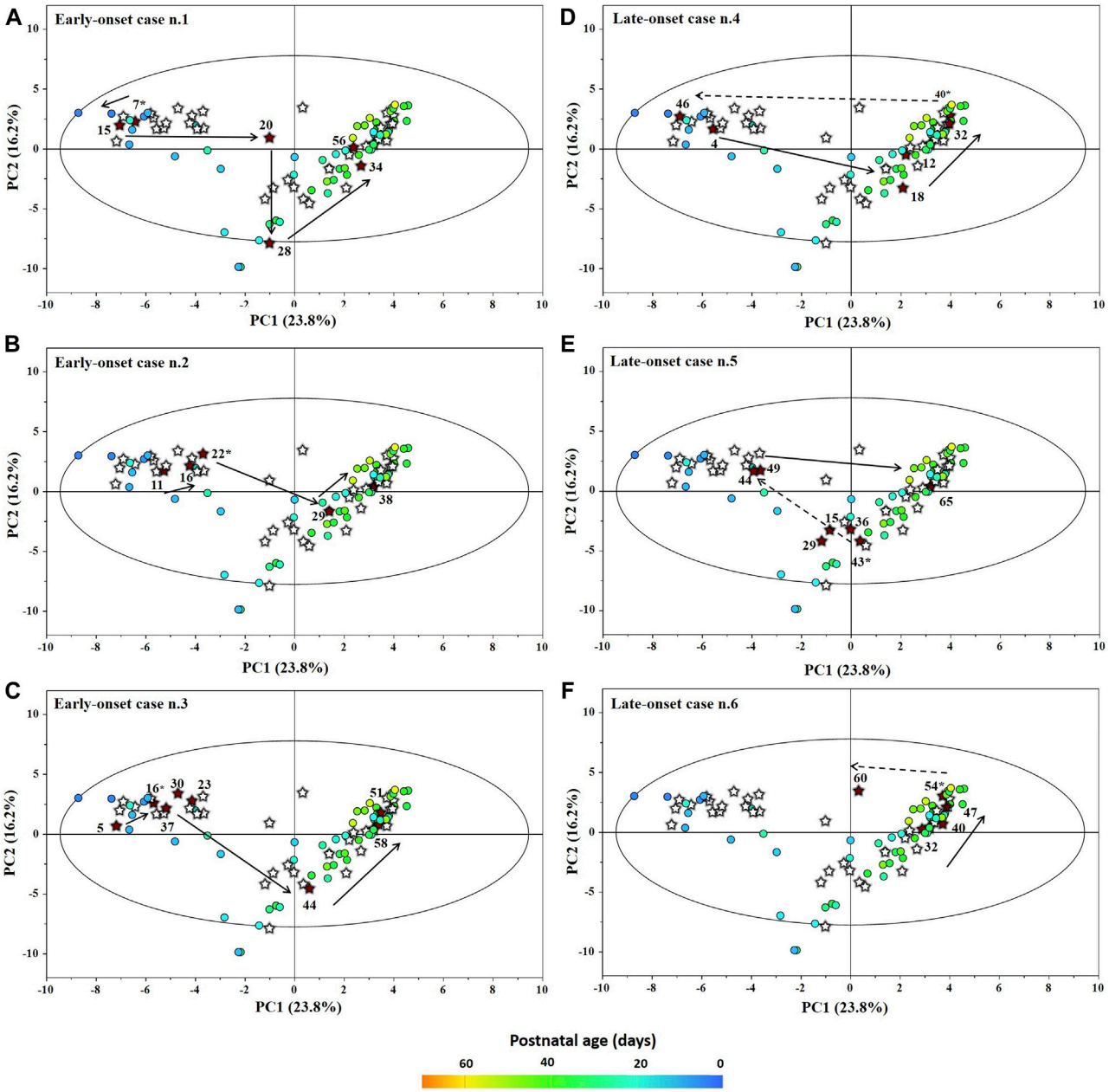

FIGURE 2 | Temporal trajectories (obtained from the model built with the ^{1}H-NMR spectra) of each individual early-onset **(A–C)** and late-onset **(D–F)** NEC baby in the PC1 vs PC2 scores plot. FI and GDT (o); NEC (★). Scores are colored according to the postnatal age. Numbers denotes the post-natal age of NEC at the time of sampling and the asterisk marks the day of the disease onset.

studies enrolling babies with sepsis (both early-onset and late-onset sepsis), approximately ten studies have utilized metabolomics alone or combined with other omics in preterm newborns with NEC; five were based only on metabolomics in various biological fluids, including serum (Wilcock et al., 2016; Wang et al., 2019), stools (Rusconi et al., 2018), urine (Thomaidou et al., 2019), and dry blood spots (Sinclair et al., 2020). A study combined metabolomics with proteomics in serum samples (Stewart et al., 2016a), and four studies integrated metagenomics with metabolomics in urine (Morrow et al., 2013) and stools (Stewart et al., 2016b; Wandro et al., 2018; Brehin et al., 2020). The heterogeneity of patient cohorts, patients and samples size, samples type, analytical methods, length of patient monitoring, diagnostic criteria of NEC, nutrition, and the presence of potentially confounding factors such as comorbidities (sepsis, bronchopulmonary dysplasia) hampers an adequate comparison between our results and those previously published. Overall, the time-dependent shift of scores observed

TABLE 2 | Statistical parameters of the OPLS models derived from the ^{1}H-NMR spectra of urine samples from group 1 NEC, group 2 FI, and group 3 GDT cases.[a]

Data set	R^2Y	Q^2Y	Permutation test[b]		*p*-value[c]
			R^2Y intercept	Q^2Y intercept	
Group 1 NEC	0.743	0.398	0.509	−0.463	$2.01 \cdot 10^{-3}$
Group 2 FI	0.813	0.567	0.531	−0.545	$7.26 \cdot 10^{-3}$
Group 3 GDT	0.880	0.638	0.461	−0.430	$1.05 \cdot 10^{-5}$

[a]*Gluconate signals were removed from the dataset.*
[b]*n = 400.*
[c]p-value *obtained from cross validation ANOVA (CV-ANOVA). NEC, necrotizing enterocolitis; FI, food intolerance; GDT, good digestive tolerance.*

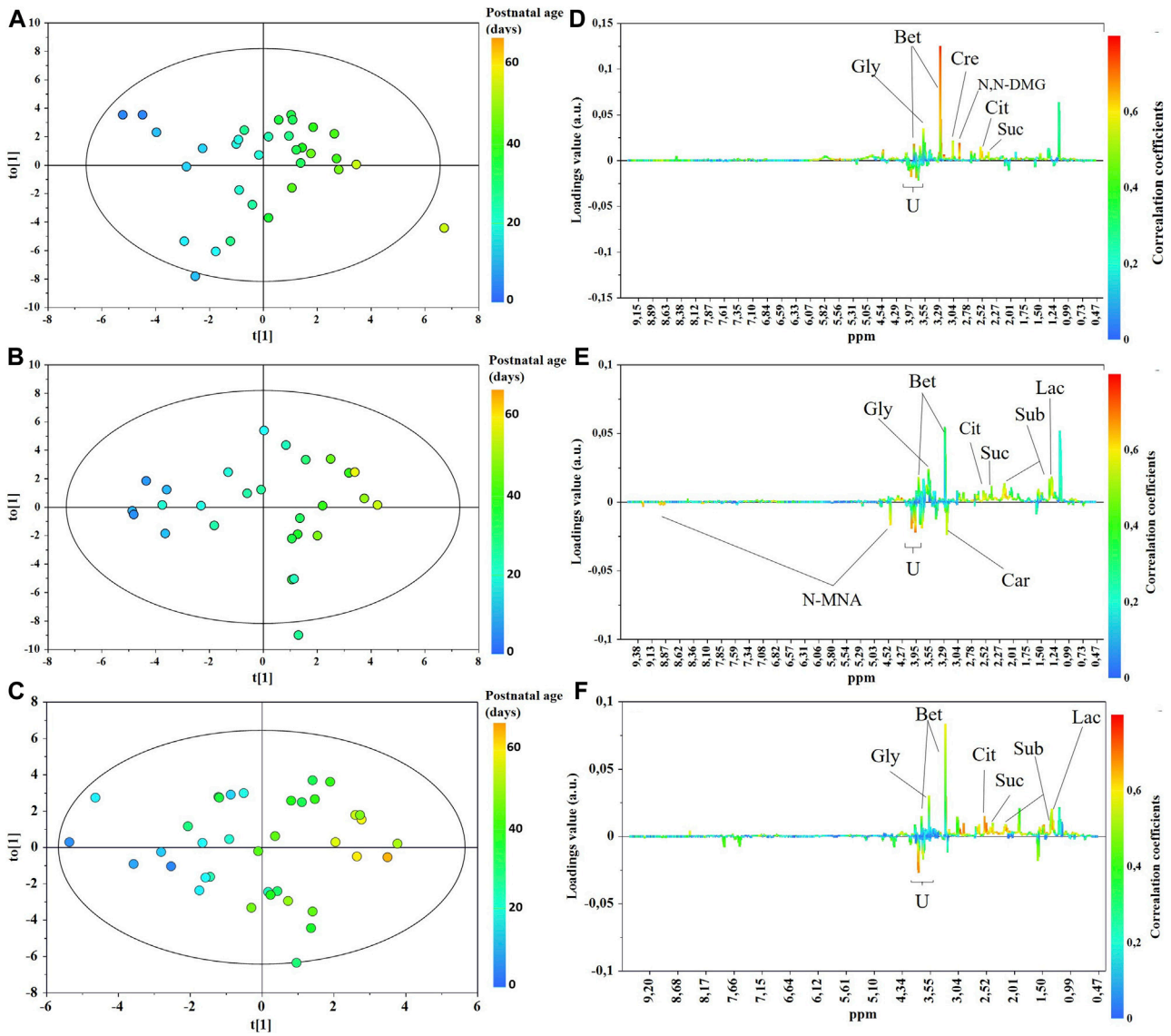

FIGURE 3 | OPLS scores **(A–C)** and loadings line **(D–F)** plots of the ^{1}H-NMR urine spectra from control group 3 GDT **(top)**, group 2 FI **(middle)**, and group 1 NEC **(bottom)**. The scores are coloured according to the postnatal age. Abbreviations: Bet, betaine; Car, carnitine; Cre, creatinine; Cit, citrate; *N,N*-DMG, *N,N*-dimethylglycine; *N*-MNA, *N*-methylnicotinamide; Gly, glycine; Lac, lactate; Sub, suberate; Suc, succinate.

TABLE 3 | Statistical parameters for the OPLS-DA models built for the pairwise comparison between cases and controls.[a]

Pairwise comparison				Permutation test		
	R^2X	R^2Y	Q^2Y	R^2Y intercept	Q^2Y intercept	*p*-value
Early-onset NEC *vs* FI	0.325	0.906	0.263	0.860	0.114	1
Early-onset NEC *vs* GDT	0.395	0.811	0.023	0.740	−0.034	0.99
Late-onset NEC *vs* FI	0.434	0.844	0.461	0.624	−0.554	0.07
Late-onset NEC *vs* GDT	0.273	0.805	0.624	0.627	−0.606	0.02

[a]*The models were considered valid only if the permutation test and* p-value *obtained from the cross validation ANOVA (CV-ANOVA) test (p < 0.05) were satisfied at the same time. NEC, necrotizing enterocolitis; FI, food intolerance; GDT, good digestive tolerance.*

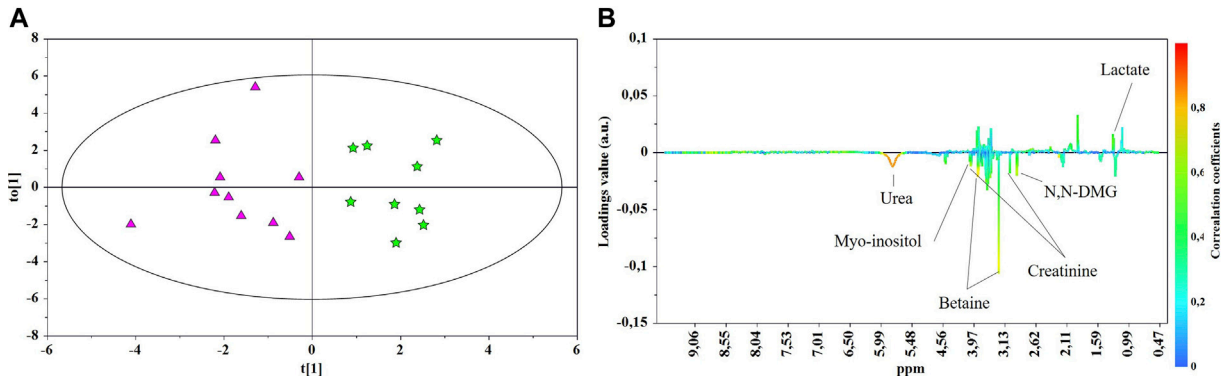

FIGURE 4 | OPLS-DA scores **(A)** and correlation loading **(B)** plots for the pair-wise comparison between the *late-onset* NEC subgroup (★) and control group 3 GDT (▲). ^{1}H NMR spectra of urine samples collected just prior to and at the days of NEC diagnosis were analyzed.

in all the babies enrolled in our study confirms the dynamic postnatal metabolic change due to the influence of age, height, and weight progression, with the concomitant development and maturation of organs and tissues, as previously found elsewhere (Cesare Marincola et al., 2016; Scalabre et al., 2017).

Betaine, a trimethylated glycine derived either from diet or the oxidation of choline, increased over time in all the three groups of infants, while *N,N*-dimethylglycine (*N,N*-DMG), derived from the loss of a methyl group from betaine, increased only in control group 3, GDT. Betaine and *N,N*-DMG are methyl donors in several metabolic pathways, including the homocysteine and the DNA methylation. Betaine is involved in osmotic homeostasis, protecting cells from dehydration and kidneys from injuries; betaine is also an anti-oxidant and is involved in neurodevelopment and immune functions. In the urine of healthy adults, betaine is almost completely absent; conversely, in healthy newborns' urine, high betaine levels are usually present, reflecting dietary choline disposal (Davies et al., 1988). Breast milk is an important dietary source of choline which is essential during infant growth and development. Interestingly, urine betaine and *N,N*-DMG levels were significantly reduced in the subgroup of babies with *late-onset* NEC (20 samples) compared with control group 3 GDT (32 samples), confirming previous findings that associated the reduction of urine betaine with NEC (Thomaidou et al., 2019). Data from the literature evidence that urine betaine was found decreased over the first 48-h of life in various groups of full-term infants with impaired growth and this trend was followed by the increase of betaine at the end of the first week of life (Marincola et al., 2015). Conversely, high urine levels of betaine were observed in full-term infants with congenital cytomegalovirus infection (Fanos et al., 2013) and hypoxic-ischemic encephalopathy (Locci et al., 2018). Thus, the decrease of urine betaine in babies with *late-onset* NEC may be associated with prematurity and kidney dysfunction rather than with sepsis and inflammation. In our preterms, the influence of infant growth and maturation over time is reflected by the positive correlation between postnatal age and glycine, succinate, and citrate. However, in the group of babies with *early-* and *late-onset* NEC (38 samples), urinary succinate and citrate were significantly reduced (fold change −0.161 and −0.163, respectively) compared to control group 3 GDT. On the other hand, they were closely comparable between group 2 FI (27 samples) and group 3 GDT (32 samples). Conversely, urine glycine abundance was almost equal between group 1 NEC and control group 3 GDT and significantly increased in group 2 FI (fold change 0.161). The reduction in succinate and citrate in group 1 NEC may be related to the impairment of the tricarboxylic acid (TCA or Krebs) cycle in babies with NEC, leading to decreased carbohydrates, amino acids, and lipids availability. Indeed, as newborns gain weight during the early postnatal age, the increase of urine succinate and citrate may reflect the high metabolic turnover due to the increasing energy demand (Moltu et al., 2014; Scalabre et al., 2017).

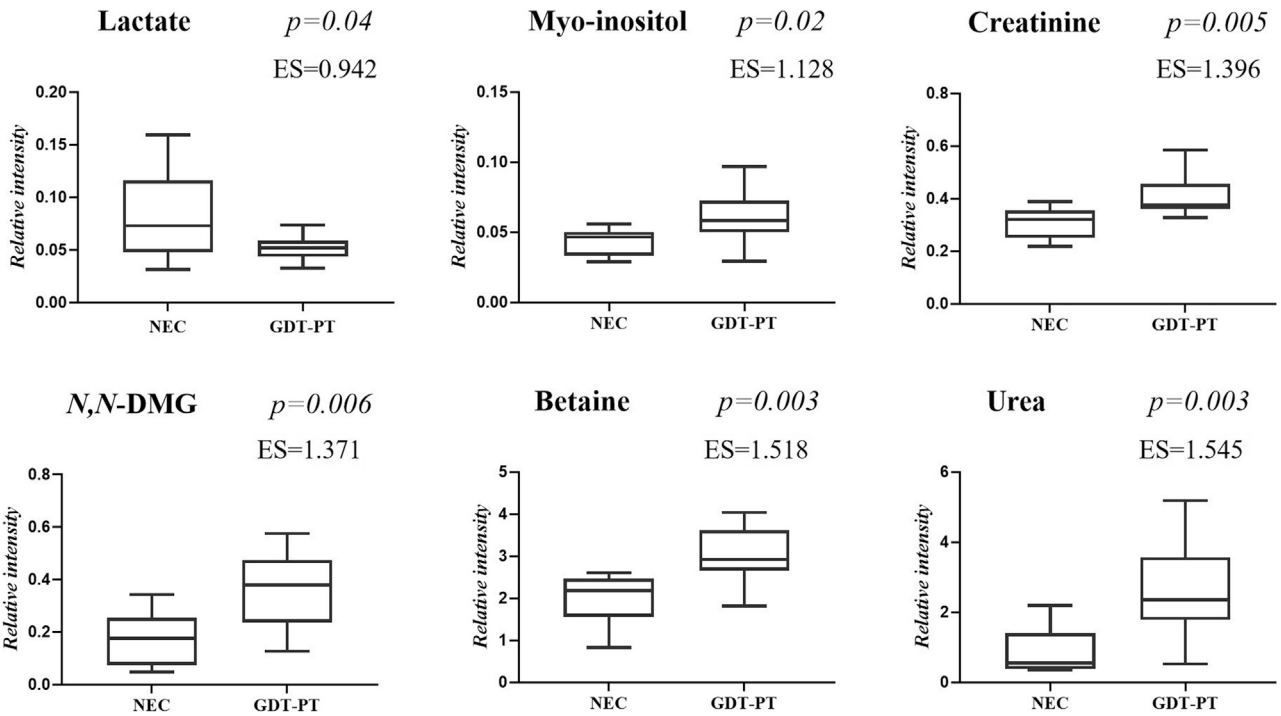

FIGURE 5 | Selected metabolites discriminating control group 3 GDT and *late-onset* NEC subgroup the day before the onset of the disease and at the disease onset. Values in the box plots are shown as the normalized peak areas of the metabolites.

A positive correlation between urine creatinine and postnatal age was observed in control group 3 GDT but not in group 1 NEC; a weak correlation was also found in group 2 FI. Creatinine abundance was significantly lower in group 1 NEC compared with group 3 GDT (fold change −0.74) and with group 2 FI (fold change −0.68), confirming similar results (fold change −0.35) previously published elsewhere (Thomaidou et al., 2019). Based on our results and data from the literature, we could argue that in preterm babies with good digestive tolerance (group 3 GDT), the positive correlation between creatinine and postnatal age is due at least to two factors: the progressive maturation of the kidney leading to the increase in glomerular filtration rate (GFR), even though slower than in full-term infants (Gubhaju et al., 2014), and the progressive increase of muscle mass. The latter is closely related to the rate of protein synthesis, which depends on feeding (Davis and Fiorotto, 2009). In groups 1 NEC and 2 FI, the replicated interruptions overtime of the enteral feeding, in response to clinical symptoms of food digestive intolerance, slows down the synthesis of proteins, with a negative consequence on the growth of organs and tissues, such as the kidney and muscle mass, and ultimately with the decrease of urine creatinine excretion. In the urine samples of the subgroup of babies with *late-onset* NEC (20 samples), we observed high lactate levels compared with those of the control group 3 GDT, especially close to the onset of the disease. A possible explanation may be the impaired TCA production of energy associated with NEC. In babies with *early-onset* NEC and in those with food intolerance (group 2 FI), the predominance of hyperlactatemia over the first 48–72 h of life in preterm infants may reduce differences between groups (Junior et al., 2021); later, lactate levels may better discriminate critically ill preterm infants with NEC or other acute diseases, from preterm infants with non-severe acute disease. Increased levels of lactate may also originate from different sources. Lactate produced by human metabolism is primarily the levorotatory isomer L-lactate; conversely, D-lactate is prevalently produced by bacterial fermentation of undigested carbohydrates in the gastrointestinal tract, and only a small fraction of this isomer is endogenously formed from methylglyoxal through the glyoxalase system (Adeva-Andany et al., 2014). High levels of D-lactate have been found in the urine and plasma of preterm babies with NEC (Garcia et al., 1984; Lei et al., 2016); this finding suggests that D-lactate may be considered an index of increased enteric bacterial activity (Grishin et al., 2013). Although ^{1}H NMR is unable to distinguish the lactate enantiomers, we cannot rule out that the high abundance of lactate in the urine of infants with *late-onset* NEC may derive at least in part from the accumulation of D-lactate. The positive correlation between lactate and postnatal age in group 2 FI and in the subgroup *late-onset* NEC, together with the concomitant decrease in *N*-methylnicotinamide, seem to confirm an imbalance of the host–microbial metabolism in these infants. Indeed, *N*-methylnicotinamide has been utilized as an index of the suppression of the gut microbiome in an experimental study in an animal model on NEC (Jiang et al., 2017).

Myo-inositol, an inositol stereoisomer mediating cell signal transduction in response to a variety of hormones,

neurotransmitters, and growth factors, was decreased in the urine of babies with *late-onset* (fold change −0.20), but not *early-onset* NEC, and in the urine of babies with FI (group 2, fold change −0.17), compared with control group 3 GDT. Factors such as kidney impairment and perturbances in the metabolism of glucose and lung surfactant influence the urinary level of myo-inositol in preterm infants; however, myo-inositol is a natural constituent of breast milk and is commonly added to formula milk (Brion et al., 2021). Therefore, the nutritional intake strongly influences the myo-inositol concentration. It is reasonable to argue that the dietary restriction, applied after multiple episodes of feeding intolerance prior to the onset of the NEC and their management during the disease, contributes to decreasing urine myo-inositol in *late-onset* NEC and FI groups. Our result confirms similar results previously reported elsewhere (Thomaidou et al., 2019); it is also supported by the simultaneous decrease of urea, a nutritional index reflecting the protein intake, in babies with NEC and FI.

Our study is affected by several limitations. First, the small number of preterm infants limits the strength of results; however, this is a pilot study. Second, the lack of any taxonomic characterization of the gut microbiota hampers to elucidate the significance of metabolic alterations originating from dysbiosis and abnormal gut microbiota fermentation. Third, this single-center study hampers the recruitment of a large number of patients and the comparison of the effects of the therapeutic management on the urine metabolome between different centers. Fourth, this study adopted a single analytical platform. Combining the highly quantitative and reproducible nature of ^{1}H NMR spectroscopy with the high sensitivity and specificity of MS may improve the panels of detectable metabolites, and potentially the reliability and accuracy of statistical models. A further limitation is the lack of ANOVA for repeated measurements (RM ANOVA); however, this limitation does not hamper the identification of candidate biomarkers for NEC, derivable by the OPLS-DA model. The strengths of this study are the analysis of longitudinal data and the classification of infants with NEC in *early-onset* (sample size = 18) and *late-onset* NEC (sample size = 20). In a previous study, babies with NEC were divided based on the gut microbiota composition (Morrow et al., 2013); unfortunately, a specific bacterial fingerprint associated with NEC was never identified unambiguously. Thus, even that study reported no definitive metabolic data. Overall, previous metabolomics-based studies on preterm infants with NEC are often inconclusive, even when metabolomics was combined with proteomics or microbiomics; (Stewart et al., 2016a; Wilcock et al., 2016; Brehin et al., 2020). Our study confirms that the urine metabolome of infants with NEC is significantly different from that of preterms infants with food intolerance but without NEC and from that of preterm infants with good digestive tolerance. However, the identification of robust candidate biomarkers of NEC requires the system biology approach based, at least, on metabolomics and microbiomics for defining an early accurate diagnosis of the disease and predicting the risk of an adverse clinical outcome much earlier than the clinical onset of the disease.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion 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 the Ethics Committee (Comité de Protection des Personnes Sud-Est IV) at the Croix-Rousse University Hospital, Hospices Civils de Lyon. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

J-CP and VF conceived the study and defined the methodology, ADM was responsible of the medical care of infants, collected and stored biological samples, SC conducted ^{1}H NMR experiments and statistical analysis, FCM supervised the ^{1}H NMR experiments and validated data, AD and ADM supervised the database and selected participants, J-CP, VF, AD, AN, FCM participated to data interpretation, FCM and SC wrote the original draft and prepared figures and tables, MM supervised the manuscript, supplementary material, tables, and figures and wrote the final version. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.680159/full#supplementary-material

REFERENCES

Adeva-Andany, M., López-Ojén, M., Funcasta-Calderón, R., Ameneiros-Rodríguez, E., Donapetry-García, C., Vila-Altesor, M., et al. (2014). Comprehensive Review on Lactate Metabolism in Human Health. *Mitochondrion* 17, 76–100. doi:10.1016/j.mito.2014.05.007

Agakidou, E., Agakidis, C., Gika, H., and Sarafidis, K. (2020). Emerging Biomarkers for Prediction and Early Diagnosis of Necrotizing Enterocolitis in the Era of Metabolomics and Proteomics. *Front. Pediatr.* 8, 602255. doi:10.3389/fped.2020.602255

Alsaied, A., Islam, N., and Thalib, L. (2020). Global Incidence of Necrotizing Enterocolitis: a Systematic Review and Meta-Analysis. *BMC Pediatr.* 20 (1), 344. doi:10.1186/s12887-020-02231-5

Arslanoglu, S., Boquien, C.-Y., King, C., Lamireau, D., Tonetto, P., Barnett, D., et al. (2019). Fortification of Human Milk for Preterm Infants: Update and Recommendations of the European Milk Bank Association (EMBA) Working Group on Human Milk Fortification. *Front. Pediatr.* 7, 76. doi:10.3389/fped.2019.00076

Ashrafian, H., Sounderajah, V., Glen, R., Ebbels, T., Blaise, B. J., Kalra, D., et al. (2020). Metabolomics - the Stethoscope for the 21st century. *Med. Princ Pract.* doi:10.1159/000513545 Epub ahead of print

Bardanzellu, F., Piras, C., Atzei, A., Neroni, P., and Fanos, V. (2020). Early Urinary Metabolomics in Patent Ductus Arteriosus Anticipates the Fate: Preliminary Data. *Front. Pediatr.* 8, 613749. doi:10.3389/fped.2020.613749

Bell, M. J., Ternberg, J. L., Feigin, R. D., Keating, J. P., Marshall, R., Barton, L., et al. (1978). Neonatal Necrotizing Enterocolitis. *Ann. Surg.* 187 (1), 1–7. doi:10.1097/00000658-197801000-00001

Berben, L., Sereika, S. M., and Engberg, S. (2012). Effect Size Estimation: Methods and Examples. *Int. J. Nurs. Stud.* 49 (8), 1039–1047. doi:10.1016/j.ijnurstu.2012.01.015

Bi, L.-w., Yan, B.-l., Yang, Q.-y., Li, M.-m., and Cui, H.-l. (2019). Probiotic Strategies to Prevent Necrotizing Enterocolitis in Preterm Infants: a Meta-Analysis. *Pediatr. Surg. Int.* 35 (10), 1143–1162. doi:10.1007/s00383-019-04547-5

Brehin, C., Dubois, D., Dicky, O., Breinig, S., Oswald, E., and Serino, M. (2020). Evolution of Gut Microbiome and Metabolome in Suspected Necrotizing Enterocolitis: A Case-Control Study. *Jcm* 9 (7), 2278. doi:10.3390/jcm90722789

Brion, L. P., Phelps, D. L., Phelps, D. L., Ward, R. M., Nolen, T. L., Hallman, N. M. K., et al. (2021). Blood Myo-Inositol Concentrations in Preterm and Term Infants. *J. Perinatol.* 41 (2), 247–254. doi:10.1038/s41372-020-00799-5

Cesare Marincola, F., Corbu, S., Lussu, M., Noto, A., Dessì, A., Longo, S., et al. (2016). Impact of Early Postnatal Nutrition on the NMR Urinary Metabolic Profile of Infant. *J. Proteome Res.* 15 (10), 3712–3723. doi:10.1021/acs.jproteome.6b00537

Cloarec, O., Dumas, M. E., Trygg, J., Craig, A., Barton, R. H., Lindon, J. C., et al. (2005). Evaluation of the Orthogonal Projection on Latent Structure Model Limitations Caused by Chemical Shift Variability and Improved Visualization of Biomarker Changes in1H NMR Spectroscopic Metabonomic Studies. *Anal. Chem.* 77 (2), 517–526. doi:10.1021/ac048803i

Darmaun, D., Lapillonne, A., Simeoni, U., Picaud, J.-C., Rozé, J.-C., Saliba, E., et al. (2018). Parenteral Nutrition for Preterm Infants: Issues and Strategy. *Arch. de Pédiatrie* 25 (4), 286–294. doi:10.1016/j.arcped.2018.02.005

Davies, S. E. C., Chalmers, R. A., Randall, E. W., and Iles, R. A. (1988). Betaine Metabolism in Human Neonates and Developing Rats. *Clinica Chim. Acta* 178 (3), 241–249. doi:10.1016/0009-8981(88)90232-x

Davis, T. A., and Fiorotto, M. L. (2009). Regulation of Muscle Growth in Neonates. *Curr. Opin. Clin. Nutr. Metab. Care* 12 (1), 78–85. doi:10.1097/MCO.0b013e32831cef9f

Diaz, S. O., Pinto, J., Barros, A. S., Morais, E., Duarte, D., Negrão, F., et al. (2016). Newborn Urinary Metabolic Signatures of Prematurity and Other Disorders: A Case Control Study. *J. Proteome Res.* 15 (1), 311–325. doi:10.1021/acs.jproteome.5b00977

Eicher, T., Kinnebrew, G., Patt, A., Spencer, K., Ying, K., Ma, Q., et al. (2020). Metabolomics and Multi-Omics Integration: A Survey of Computational Methods and Resources. *Metabolites* 10, 202. doi:10.3390/metabo10050202

Eriksson, L., Trygg, J., and Wold, S. (2008). CV-ANOVA for Significance Testing of PLS and OPLS Models. *J. Chemometrics* 22, 594–600. doi:10.1002/cem.1187

Fanos, V., Locci, E., Noto, A., Lazzarotto, T., Manzoni, P., Atzori, L., et al. (2013). Urinary Metabolomics in Newborns Infected by Human Cytomegalovirus: a Preliminary Investigation. *Early Hum. Dev.* 89 (Suppl. 1), S58–S61. doi:10.1016/S0378-3782(13)70017-3

Fanos, V., Pintus, R., and Dessì, A. (2018). Clinical Metabolomics in Neonatology: From Metabolites to Diseases. *Neonatology* 113 (4), 406–413. doi:10.1159/000487620

Garcia, J., Smith, F. R., and Cucinell, S. A. (1984). Urinary D-Lactate Excretion in Infants with Necrotizing Enterocolitis. *J. Pediatr.* 104 (2), 268–270. doi:10.1016/s0022-3476(84)81010-0

Grishin, A., Papillon, S., Bell, B., Wang, J., and Ford, H. R. (2013). The Role of the Intestinal Microbiota in the Pathogenesis of Necrotizing Enterocolitis. *Semin. Pediatr. Surg.* 22 (2), 69–75. doi:10.1053/j.sempedsurg.2013.01.002

Gubhaju, L., Sutherland, M. R., Horne, R. S. C., Medhurst, A., Kent, A. L., Ramsden, A., et al. (2014). Assessment of Renal Functional Maturation and Injury in Preterm Neonates during the First Month of Life. *Am. J. Physiology-Renal Physiol.* 307 (2), F149–F158. doi:10.1152/ajprenal.00439.2013

Jiang, P., Trimigno, A., Stanstrup, J., Khakimov, B., Viereck, N., Engelsen, S. B., et al. (2017). Antibiotic Treatment Preventing Necrotising Enterocolitis Alters Urinary and Plasma Metabolomes in Preterm Pigs. *J. Proteome Res.* 16 (10), 3547–3557. doi:10.1021/acs.jproteome.7b00263

Juhl, S. M., Hansen, M. L., Gormsen, M., Skov, T., and Greisen, G. (2019). Staging of Necrotising Enterocolitis by Bell's Criteria Is Supported by a Statistical Pattern Analysis of Clinical and Radiological Variables. *Acta Paediatr.* 108 (5), 842–848. doi:10.1111/apa.14469

Junior, L. K. O., Carmona, F., Aragon, D. C., and Gonçalves-Ferri, W. A. (2021). Evaluation of Urine Output, Lactate Levels and Lactate Clearance in the Transitional Period in Very Low Birth Weight Preterm Infants. *Eur. J. Pediatr.* 180 (1), 91–97. doi:10.1007/s00431-020-03717-1

Karczewski, K. J., and Snyder, M. P. (2018). Integrative Omics for Health and Disease. *Nat. Rev. Genet.* 19 (5), 299–310. doi:10.1038/nrg.2018.4

Lei, G., Zhang, J., Wang, X., and Chen, M. (2016). Plasma D-Lactate Levels in Necrotizing Enterocolitis in Premature Infants. *Iran J. Pediatr.* In Press (2), e4403. doi:10.5812/ijp.4403

Locci, E., Bazzano, G., Demontis, R., Chighine, A., Fanos, V., and d'Aloja, E. (2020). Exploring Perinatal Asphyxia by Metabolomics. *Metabolites* 10 (4), 141. doi:10.3390/metabo10040141

Locci, E., Noto, A., Puddu, M., Pomero, G., Demontis, R., Dalmazzo, C., et al. (2018). A Longitudinal 1H-NMR Metabolomics Analysis of Urine from Newborns with Hypoxic-Ischemic Encephalopathy Undergoing Hypothermia Therapy. Clinical and Medical Legal Insights. *PLoS One* 13 (4), e0194267. doi:10.1371/journal.pone.0194267)

Lucchini, R., Bizzarri, B., Giampietro, S., and De Curtis, M. (2011). Feeding Intolerance in Preterm Infants. How to Understand the Warning Signs. *J. Maternal-Fetal Neonatal Med.* 24 (Suppl. 1), 72–74. doi:10.3109/14767058.2011.607663

Ma, F., Li, S., Gao, X., Zhou, J., Zhu, X., Wang, D., et al. (2019). Interleukin-6-mediated CCR9+ Interleukin-17-Producing Regulatory T Cells Polarization Increases the Severity of Necrotizing Enterocolitis. *EBioMedicine* 44, 71–85. doi:10.1016/j.ebiom.2019.05.042

Marincola, F. C., Dessì, A., Pattumelli, M. G., Corbu, S., Ossicini, C., Ciccarelli, S., et al. (2015). 1H NMR-Based Urine Metabolic Profile of IUGR, LGA, and AGA Newborns in the First Week of lifeH NMR-Based Urine Metabolic Profile of IUGR, LGA, and AGA Newborns in the First Week of Life. *Clinica Chim. Acta* 451 (Pt A), 28–34. doi:10.1016/j.cca.2015.08.008

Meister, A. L., Doheny, K. K., and Travagli, R. A. (2020). Necrotizing Enterocolitis: It's Not All in the Gut. *Exp. Biol. Med. (Maywood)* 245 (2), 85–95. doi:10.1177/1535370219891971

Misra, B. B. (2020). Data Normalization Strategies in Metabolomics: Current Challenges, Approaches, and Tools. *Eur. J. Mass. Spectrom. (Chichester)* 26, 165–174. doi:10.1177/1469066720918446

Moltu, S., Sachse, D., Blakstad, E., Strømmen, K., Nakstad, B., Almaas, A., et al. (2014). Urinary Metabolite Profiles in Premature Infants Show Early Postnatal Metabolic Adaptation and Maturation. *Nutrients* 6 (5), 1913–1930. doi:10.3390/nu6051913

Morrow, A. L., Lagomarcino, A. J., Schibler, K. R., Taft, D. H., Yu, Z., Wang, B., et al. (2013). Early Microbial and Metabolomic Signatures Predict Later Onset of Necrotizing Enterocolitis in Preterm Infants. *Microbiome* 1 (1), 13. doi:10.1186/2049-2618-1-13

Neu, J. (2020). Necrotizing Enterocolitis: The Future. *Neonatology* 117 (2), 240–244. doi:10.1159/000506866)

Neu, J., and Walker, W. A. (2011). Necrotizing Enterocolitis. *N. Engl. J. Med.* 364 (3), 255–264. doi:10.1056/NEJMra1005408

Nguyen, D. N., and Sangild, P. T. (2019). Pathogenesis and Biomarkers for Necrotizing Enterocolitis: Getting Any Closer?. *EBioMedicine* 45, 13–14. doi:10.1016/j.ebiom.2019.06.029

Niño, D. F., Sodhi, C. P., and Hackam, D. J. (2016). Necrotizing Enterocolitis: New Insights into Pathogenesis and Mechanisms. *Nat. Rev. Gastroenterol. Hepatol.* 13 (10), 590–600. doi:10.1038/nrgastro.2016.119

Raba, A. A., O'Sullivan, A., and Miletin, J. (2021). Pathogenesis of Necrotising Enterocolitis: The Impact of the Altered Gut Microbiota and Antibiotic

Exposure in Preterm Infants. *Acta Paediatr.* 110 (2), 433–440. doi:10.1111/apa.15559

Ramachandran, S., Fontanille, P., Pandey, A., and Larroche, C. (2006). Gluconic Acid: Properties, Applications and Microbial Production. *Food Technol. Biotechnol.* 44, 185–195.

Riganti, C., Gazzano, E., Polimeni, M., Aldieri, E., and Ghigo, D. (2012). The Pentose Phosphate Pathway: an Antioxidant Defense and a Crossroad in Tumor Cell Fate. *Free Radic. Biol. Med.* 53 (3), 421–436. doi:10.1016/j.freeradbiomed.2012.05.006

Rusconi, B., Jiang, X., Sidhu, R., Ory, D. S., Warner, B. B., and Tarr, P. I. (2018). Gut Sphingolipid Composition as a Prelude to Necrotizing Enterocolitis. *Sci. Rep.* 8 (1), 10984. doi:10.1038/s41598-018-28862-4

Scalabre, A., Jobard, E., Demède, D., Gaillard, S., Pontoizeau, C., Mouriquand, P., et al. (2017). Evolution of Newborns' Urinary Metabolomic Profiles According to Age and Growth. *J. Proteome Res.* 16 (10), 3732–3740. doi:10.1021/acs.jproteome.7b00421

Shulman, R. J., Schanler, R. J., Lau, C., Heitkemper, M., Ou, C.-N., and Smith, E. O. B. (1998). Early Feeding, Feeding Tolerance, and Lactase Activity in Preterm Infants. *J. Pediatr.* 133 (5), 645–649. doi:10.1016/s0022-3476(98)70105-2

Sinclair, T. J., Ye, C., Chen, Y., Zhang, D., Li, T., Ling, X. B., et al. (2020). Progressive Metabolic Dysfunction and Nutritional Variability Precedes Necrotizing Enterocolitis. *Nutrients* 12 (5), 1275. doi:10.3390/nu12051275

Stewart, C. J., Embleton, N. D., Marrs, E. C. L., Smith, D. P., Nelson, A., Abdulkadir, B., et al. (2016b). Temporal Bacterial and Metabolic Development of the Preterm Gut Reveals Specific Signatures in Health and Disease. *Microbiome* 4 (1), 67. doi:10.1186/s40168-016-0216-8

Stewart, C. J., Nelson, A., Treumann, A., Skeath, T., Cummings, S. P., Embleton, N. D., et al. (2016a). Metabolomic and Proteomic Analysis of Serum from Preterm Infants with Necrotising Entercolitis and Late-Onset Sepsis. *Pediatr. Res.* 79 (3), 425–431. doi:10.1038/pr.2015.235

Thomaidou, A., Chatziioannou, A. C., Deda, O., Benaki, D., Gika, H., Mikros, E., et al. (2019). A Pilot Case-Control Study of Urine Metabolomics in Preterm Neonates with Necrotizing Enterocolitis. *J. Chromatogr. B* 1117, 10–21. doi:10.1016/j.jchromb.2019.04.019

Wandro, S., Osborne, S., Enriquez, C., Bixby, C., Arrieta, A., and Whiteson, K. (2018). The Microbiome and Metabolome of Preterm Infant Stool Are Personalized and Not Driven by Health Outcomes, Including Necrotizing Enterocolitis and Late-Onset Sepsis. *mSphere* 3 (3), e00104–18. doi:10.1128/mSphere.00104-18

Wang, F., Li, W., Wang, G., Yu, M., Zhong, J., Xu, C., et al. (2019). Gas Chromatography-Mass Spectrometry Based Serum Metabolic Analysis for Premature Infants and the Relationship with Necrotizing Enterocolitis: a Cross-Sectional Study. *Ital. J. Pediatr.* 45 (1), 54. doi:10.1186/s13052-019-0646-6

Wilcock, A., Begley, P., Stevens, A., Whatmore, A., and Victor, S. (2016). The Metabolomics of Necrotising Enterocolitis in Preterm Babies: an Exploratory Study. *J. Maternal-Fetal Neonatal Med.* 29 (5), 758–762. doi:10.3109/14767058.2015.1017462

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 (D1), D608–D617. doi:10.1093/nar/gkx1089

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 29 July 2021
doi: 10.3389/fmolb.2021.665661

Comprehensive Metabolic Signature of Renal Dysplasia in Children. A Multiplatform Metabolomics Concept

*Szymon Macioszek[1†], Renata Wawrzyniak[1†], Anna Kranz[2], Marta Kordalewska[1], Wiktoria Struck-Lewicka[1], Danuta Dudzik[1], Margot Biesemans[1], Michał Maternik[2], Aleksandra M. Żurowska[3] and Michał J. Markuszewski[1]**

[1]Department of Biopharmaceutics and Pharmacodynamics, Medical University of Gdańsk, Gdańsk, Poland, [2]Department of Pediatrics, Nephrology and Hypertension, Medical University of Gdańsk, Gdańsk, Poland, [3]Centre for Rare Diseases, Medical University of Gdańsk, Gdańsk, Poland

OPEN ACCESS

Edited by:
Manuel Portero-Otin,
Universitat de Lleida, Spain

Reviewed by:
Mariona Jové,
Universitat de Lleida, Spain
Roccaldo Sardella,
University of Perugia, Italy
Ann Van Schepdael,
KU Leuven, Belgium
Hai-long Piao,
Dalian Institute of Chemical Physics, (CAS), China

***Correspondence:**
Michał J. Markuszewski
michal.markuszewski@gumed.edu.pl

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

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

Received: 08 February 2021
Accepted: 19 July 2021
Published: 29 July 2021

Citation:
Macioszek S, Wawrzyniak R, Kranz A, Kordalewska M, Struck-Lewicka W, Dudzik D, Biesemans M, Maternik M, Żurowska AM and Markuszewski MJ (2021) Comprehensive Metabolic Signature of Renal Dysplasia in Children. A Multiplatform Metabolomics Concept. Front. Mol. Biosci. 8:665661. doi: 10.3389/fmolb.2021.665661

Renal dysplasia is a severe congenital abnormality of the kidney parenchyma, which is an important cause of end-stage renal failure in childhood and early adulthood. The diagnosis of renal dysplasia relies on prenatal or postnatal ultrasounds as children show no specific clinical symptoms before chronic kidney disease develops. Prompt diagnosis is important in terms of early introduction of nephroprotection therapy and improved long-term prognosis. Metabolomics was applied to study children with renal dysplasia to provide insight into the changes in biochemical pathways underlying its pathology and in search of early indicators for facilitated diagnosis. The studied cohort consisted of 72 children, 39 with dysplastic kidneys and 33 healthy controls. All subjects underwent comprehensive urine metabolic profiling with the use of gas chromatography and liquid chromatography coupled to mass spectrometry, with two complementary separation modes of the latter. Univariate and multivariate statistical calculations identified a total of nineteen metabolites, differentiating the compared cohorts, independent of their estimated glomerular filtration rate. Seven acylcarnitines, xanthine, and glutamine were downregulated in the urine of renal dysplasia patients. Conversely, renal dysplasia was associated with higher urinary levels of dimethylguanosine, threonic acid or glyceric acid. This is the first metabolomic study of subjects with renal dysplasia. The authors define a characteristic urine metabolic signature in children with dysplastic kidneys, irrespective of renal function, linking the condition with altered fatty acid oxidation, amino acid and purine metabolisms.

Keywords: renal dysplasia, metabolomics, pediatric nephrology, multiplatform approaches, LC-MS, GC-MS

INTRODUCTION

Renal dysplasia, though classified as a rare disease (birth prevalence: 1/2,300) is one of the major causes of chronic kidney disease in childhood. This congenital abnormality of the kidneys is due to early abnormal kidney development which results in malformation of the normal histologic structure of the kidney with the characteristic presence of embryological tissue in the form of undifferentiated and metaplastic tissues. Due to an accompanying reduction in the number of overall nephrons, renal dysplasia may lead to chronic kidney disease (CKD) and with time progress to end-stage renal disease (ESRD). Renal dysplasia is one of the most frequent underlying pathologies in children

requiring renal replacement therapy (13.5%) (www.orpha.net). Its treatment focuses mainly on slowing down the progression of CKD.

Renal dysplasia is usually symptomless before complications of CKD develop. It is diagnosed through prenatal or postnatal radiological screening. Ultrasonography reveals a normal-sized or small kidney with increased echogenicity and either absent or poor corticomedullary differentiation, frequently accompanied by the presence of small cysts. The extent of dysplastic changes is extremely variable and when mild is difficult to visualize. Early diagnosis is hampered by the lack of available biomarkers of abnormal kidney differentiation in the initial period of stable renal function. Prompt diagnosis of renal dysplasia is important in terms of management and long term prognosis due to the risk of future end-stage renal disease.

Metabolomics is an advanced tool providing insight into molecular processes occurring in a living organism and enabling the observation of disturbances in metabolic pathways resulting from changes in both genome and proteome or from environmental factors. In untargeted metabolomics, a whole set of metabolites present in a studied biological matrix is analysed and subsequently evaluated. Due to the wide range of compounds with various physicochemical properties, usually several complementary analytical techniques are used to cover the whole metabolome. Initially, metabolomics was used in search of disease biomarkers that are challenging for traditional diagnosis (López-López et al., 2018; Parfieniuk et al., 2018) and increasingly for elucidating molecular mechanisms of various disorders such as cancers (Armitage and Barbas, 2014; Kaushik and DeBerardinis, 2018), cardiovascular diseases (Rhee and Gerszten, 2012; Ussher et al., 2016) but also renal diseases (Hocher and Adamski, 2017; Kalim and Rhee, 2017). Detailed knowledge about the molecular basis of a disease may be a starting point for the proposition of new therapeutic targets.

The kidney diseases that are most frequently studied by metabolomics include CKD, diabetic nephropathy, renal cell carcinoma, and acute kidney injury (Kalim and Rhee, 2017). Kidney function, assessed by the estimated glomerular filtration rate (eGFR), has been related to about one third of the detected metabolites in both general and CKD populations (Benito et al., 2018). Studies including the pediatric population are scarce since only one study performed on children was found. Atzori et al. (2010) collected a group of children with various nephrouropathies (renal dysplasia, vesico-ureteral reflux, urinary tract infection, acute kidney injury, and others), and compared their ^{1}H NMR-based metabolic profile with healthy children. However, only five children with renal dysplasia were employed, and for statistical analysis dysplasia samples were combined with other pathologies.

The aim of the performed study was to search for urine metabolites which may discriminate children with dysplastic kidneys from those with normally developed ones, taking into account the confounding presence of metabolites characteristic for decreased glomerular filtration rate associated with CKD, which is a characteristic hallmark for this disorder.

MATERIALS AND METHODS

Study Design and Population

72 children were enrolled in the study; 39 subjects with renal dysplasia [mean age 5.68 years (range 0.08–17.40)] and 33 healthy controls [mean age 7.28 years (range 0.09–17.69)]. The majority of children in both cohorts were below 5 years of age. Males were predominant in both the study (66.6%) and control groups (60.6%). Renal dysplasia was diagnosed by ultrasonography. Dysplastic changes were present in both kidneys (bilateral renal dysplasia) in the majority of subjects (61.5%). The renal function was assessed based on the eGFR, calculated with the new Schwartz formula. CKD with a decreased renal function (eGFR <60 ml/min/1.73 m^2) was present in 41% of the studied cohort with dysplastic kidneys. The clinical features of the studied pediatric cohorts are presented in **Table 1**. The study was approved by the local Bioethics Committee of the Medical University of Gdańsk (NKBBN/499/2016, NKBBN/493/2018).

Chemicals and Reagents

The LC-MS grade methanol and acetonitrile were purchased from Fisher Scientific (Loughborough, United Kingdom). The mobile phase additive formic acid was from Chem-Lab (Zedelgem, Belgium). Ammonium formate, pentadecanoic acid, pyridine, urease, methoxyamine hydrochloride, *N*,*O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS) and alkane standard mixture for GC were from Sigma-Aldrich (United States, Switzerland, Germany). The Milli-Q PLUS system (Millipore, Austria) was used to obtain ultrapure water for sample dilution and urease solution preparation.

Sample Collection and Preparation

First morning urine samples were collected to minimalize the effects of diet but also of circadian rhythm or physical activity. The samples were collected in 1.5 ml Eppendorf tubes and placed immediately at −80°C. The samples were stored frozen until the day of analyses, when they were thawed at room temperature.

Untargeted GC-MS and LC-MS Metabolomic Analysis

The urine samples were analyzed along with the quality control (QC) and blank samples, using two complementary analytical platforms namely, gas chromatography coupled to triple quadrupole mass spectrometry (GC-QQQ/MS) and liquid chromatography coupled to time-of-flight mass spectrometry (LC-TOF-MS). Additionally, in terms of the LC technique, two complementary separation modes were used: reversed-phase (RP) and hydrophilic interaction chromatography (HILIC), both in positive (RP+, HILIC+) and negative (RP−, HILIC−) ionization modes. In RP, lipids and other nonpolar metabolites can be separated while HILIC is suitable for the separation of polar compounds such as amino acids, nucleosides, sugars, organic acids, or amines. This yielded five analytical batches for each urine sample. The detailed protocols of

TABLE 1 | Clinical features of the studied cohorts of 72 children (39 with renal dysplasia and 33 healthy controls).

	Cohort of subjects with renal dysplasia	Cohort of healthy controls	*p*-value
Total number enrolled	39	33	—
Females (%)	13/39 (33.3)	13/33 (39.4)	0.540
Males (%)	26/39 (66.6)	20/33 (60.6)	0.540
Mean age ± SD	5.68 ± 5.84	7.28 ± 5.62	0.135
Range age in years	(0.08–17.40)	(0.09–17.69)	—
<5 years age (%)	23/39 (58.9)	15/33 (45.5)	0.215
>5 years age (%)	16/39 (41.1)	18/33 (54.5)	0.215
Mean BMI ± SD	17.05 ± 5.12	16.48 ± 2.65	0.560
Range BMI	(7.07–19.46)	(8.89–25)	—
eGFR >60 ml/min/1.73 m^2 (%)	21/37 (56.8)	—	—
eGFR <60 ml/min/1.73 m^2 (%)	16/37 (43.2)	—	—
Bilateral dysplasia (%)	25/39 (64.1)	—	—
Unilateral dysplasia (%)	14/39 (35.9)	—	—

sample preparation and analysis for both LC-MS and GC-MS are provided in the Supplementary Material.

Agilent 1,200 HPLC system coupled to a 6,224 TOF/MS system (Agilent Technologies, Germany) was used to determine the urine metabolic fingerprints. Reversed-phase separation was achieved using a 2.1 mm × 100 mm, 1.8 μm, Zorbax Extend-C18 column (Agilent Technologies, United States), with mobile phase consisting of 0.1% formic acid in water and 0.1% formic acid in acetonitrile. In HILIC mode, 10 mM ammonium formate water solution and acetonitrile were used to enable separation of polar compounds in a Poroshell 120 HILIC 4.6 x 50 mm, 2.7 μm column (Agilent Technologies, United States).

Complementary GC-MS analysis was conducted on a GCMS-TQ8030 system (Shimadzu, Japan). The chromatographic separation was performed in a Zebron ZB-5MS column (30 m × 0.25 mm, 0.25 μm) with helium as a carrier gas. The scan mode from 50 m/*z* to 600 m/*z* was applied.

Data Processing and Analysis

The software used for raw data processing included: MassHunter Qualitative Analysis version B.06.00 (Agilent Technologies, Germany), MassHunter DA Reprocessor version B.08.00 (Agilent Technologies, Germany), Mass Profiler Professional (MPP, Agilent Technologies, Germany) and Automated Mass Spectra Detection and Identification System (AMDIS, National Institute of Standards and Technology, United States). Only peaks with intensity higher than 5,000 counts and present in at least 80% of samples in a group of children with dysplasia or healthy controls were retained for further data processing. Raw data were normalized with the use of probabilistic quotient normalization (PQN) to correct for differences in urine dilution between the patients.

For statistical analysis, Matlab 2014a (Mathworks, Natrick, MA, United States), SIMCA 16 (Sartorius Stedim Biotech, Sweden) and Metaboanalyst 4.0 (https://www.metaboanalyst.ca) were employed. First, principal component analysis (PCA) was applied for each dataset to verify whether the QC samples were measured identically, regardless of their position in the analytical run. To examine the differences between the renal dysplasia group and the healthy controls, *t*-test or Mann-Whitney *U* test with multiple testing correction was applied, depending on data distribution and equality of variances. The variables with corrected p value ≤ 0.05 were considered as significantly differentiating the compared groups. The same methodology was used to compare renal dysplasia patients with normal and decreased eGFR to separate the influence of impaired renal function from metabolic changes due to abnormal kidney structure.

A supervised multivariate statistical method - partial least squares discriminant analysis (PLS-DA), was used to analyse the predictive power of the metabolites to identify patients with dysplasia, considering relationships between all metabolites, in contrast to univariate methods. For PLS-DA models built in SIMCA 16, CV-ANOVA values were calculated to assess their reliability. Based on the PLS-DA models, VIP (variable importance in projection) and SR (selectivity ratio) values were calculated to select compounds that are potentially related to the differentiation between the groups. A VIP coefficient higher than one indicates the variables' relevance for the differentiation between the compared groups. SR, a further tool for ranking variables importance in regression models, was used for the selection of metabolites that have a different abundance in renal dysplasia patients and healthy controls. For correlation analysis of metabolite abundances with the eGFR values of the patients, Spearman's rank correlation was calculated.

Metabolite Identification

In the LC-MS analysis, identification of the analytical signals was the last step of the workflow, following the statistical comparisons. The metabolites were annotated using the measured accurate mass and isotopic distribution pattern, while their identity was confirmed after a fragmentation pattern analyses. For the confirmation of metabolite structures, MS/MS analysis on HPLC-Q-TOF/MS System 6550A (Agilent Technologies, Germany) was implemented. Therefore, the metabolite identification was provided at level 2 according to Metabolomics Standards Initiative. For GC-MS data, annotated metabolites were selected based on their retention indices, calculated from the retention times of the alkane mixture. Further identification of the signals was possible by comparing the metabolite fragments in mass spectra libraries, such as NIST 11 and an in-house library of urinary metabolites.

RESULTS

Untargeted GC-MS and LC-MS Metabolomic Analysis

Urine metabolic fingerprints from 72 children (39 with renal dysplasia and 33 healthy controls) were measured by means of two complementary analytical platforms, LC-TOF-MS (RP and HILIC in both positive and negative ionization modes) and GC-QQQ/MS. The yielded datasets consisted of 252 (RP+), 54 (RP−), 383 (HILIC+), 32 (HILIC−) and 66 features (GC-MS) after alignment and filtration. **Figure 1** illustrates the metabolic fingerprints obtained from the urine of a child with dysplastic kidneys by the two complementary platforms and for LC–MS in the four different separation modes.

PCA models built on PQN-normalized and log-transformed data demonstrated the clustering of the QC samples on the PCA score plots verifying the stability of the analytical system and the method reproducibility. The wide spread of studied urine samples from children with dysplastic kidneys and healthy controls confirmed the validity of the applied experimental procedures and the negligibility of the analytical variability in comparison to the obtained biological variability. The distribution of the samples from both studied cohorts, in comparison to the QC samples, is presented in **Figure 2**.

Comparison of Urine Metabolomic Profiles Derived From Children With Renal Dysplasia and Healthy Controls

The total number of statistically significant variables in the comparison of healthy and disease groups was 44 for RP-LC-MS(+), 10 for RP-LC-MS(−), 52 for HILIC-LC-MS(+), 10 for HILIC-LC-MS(−), and 7 for GC-MS. All of the features detected by LC-MS were subjected to metabolite identification with the use of available databases, and their identity was confirmed by fragmentation patterns analyses with the use of LC-QTOF/MS. Finally, 28 significant metabolites from univariate statistical analysis were successfully identified. Multivariate PLS-DA analysis distinguished 10 relevant metabolites with a VIP value higher than 1, and eight metabolites with an SR value higher than 0.5 that significantly differentiated children with dysplastic kidneys from healthy controls. **Figure 3** illustrates a significant separation of urine samples between children with dysplastic kidneys and healthy controls by multivariate PLS-DA analysis. Significant metabolites obtained from all analytical techniques and ionization modes are compiled in **Table 2**.

Metabolic Changes Potentially Associated With Renal Function due to Reduced eGFR in Patients With Dysplastic Kidneys

In this study, the urinary metabolomic signature of children with renal dysplasia in comparison to the healthy pediatric controls, was evaluated using complementary analytical platforms and advanced statistics. The observed urinary metabolite changes derived mainly from the purine, lipid and amino acid metabolism as well as from glycolysis, the TCA cycle and the urea cycle. Among the 28 metabolites which were significantly different in renal dysplasia subjects in comparison to the healthy controls, nine were found to differentiate subjects with normal and reduced eGFR (**Table 3**). The highest correlation with eGFR values was calculated for metabolites shown in **Table 4**. Due to the presence of CKD (eGFR <60 ml/min/m2) in a significant proportion of the studied cohort, further statistical comparisons were conducted in patients with renal dysplasia based on eGFR criteria. The final set of 19 metabolites which significantly differed subjects with renal dysplasia independently of eGFR from healthy controls is listed in **Table 5**.

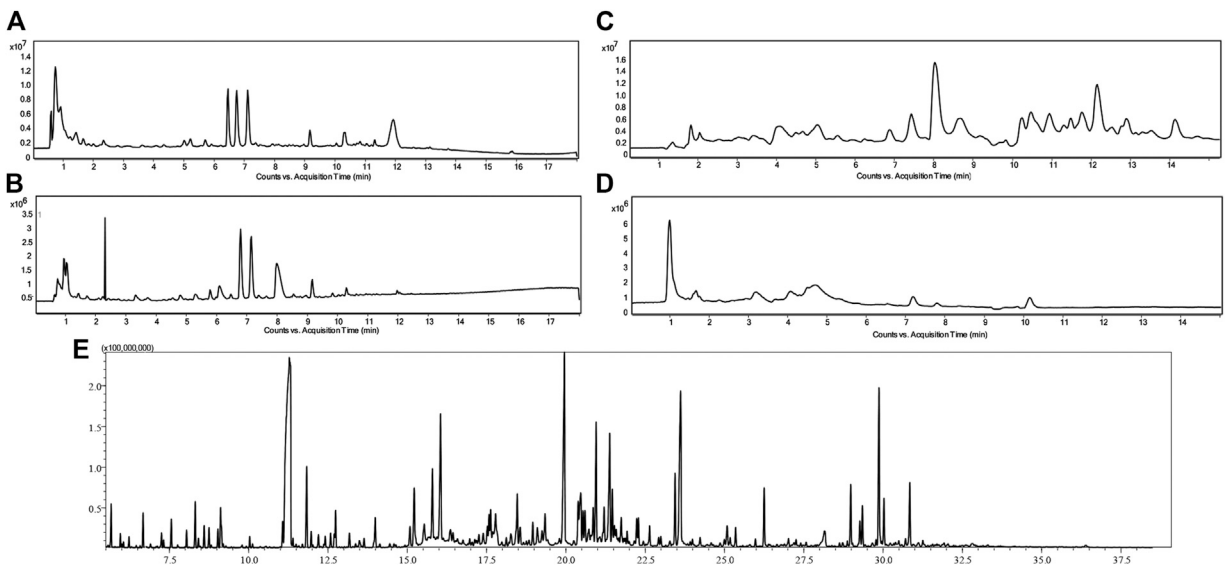

FIGURE 1 | Chromatograms of urine metabolic fingerprints measured by means of RP LC-TOF-MS in positive **(A)** and negative **(B)** ionization modes, HILIC LC-TOF-MS in both positive **(C)** and negative **(D)** ionization modes and GC-QQQ/MS **(E)** in a child with dysplastic kidneys.

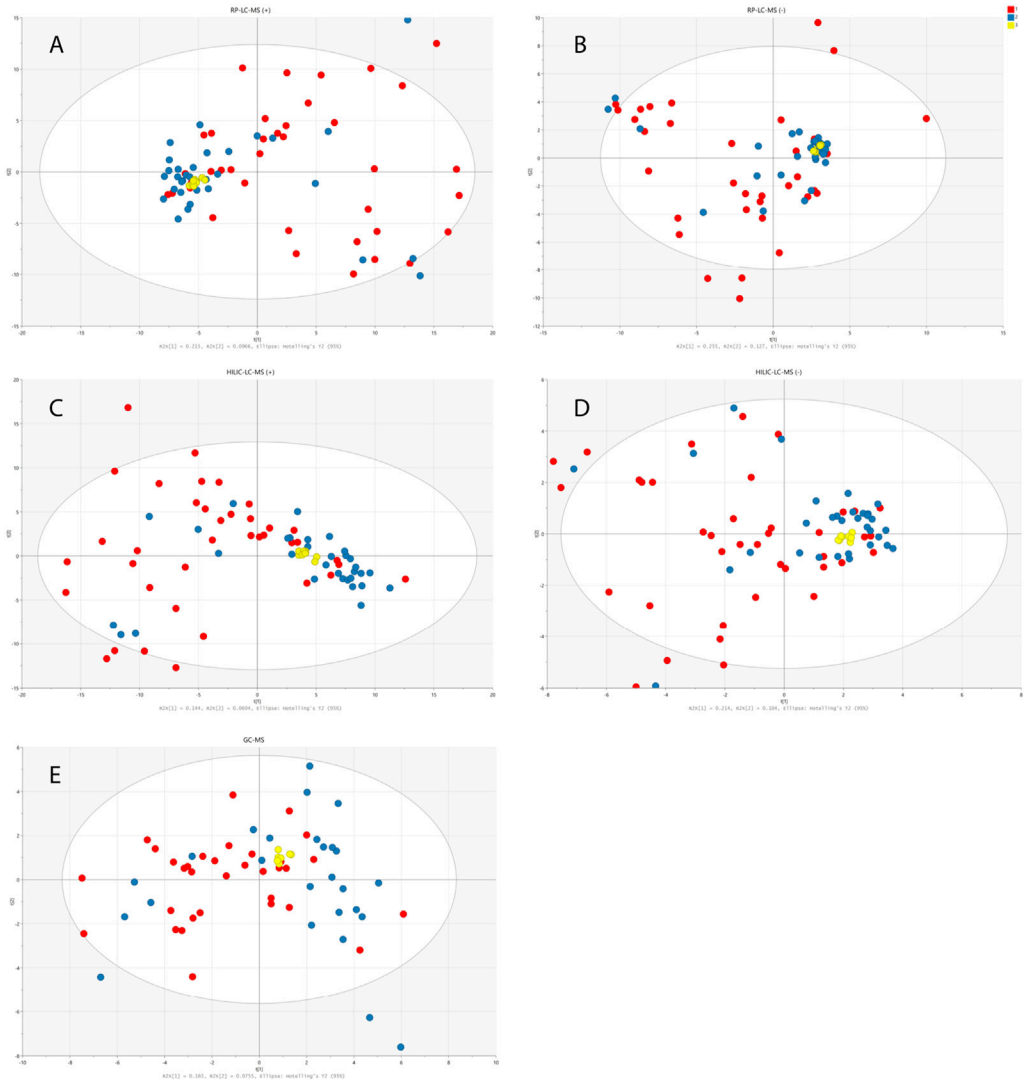

FIGURE 2 | PCA models built on data obtained from RP-LC-TOF-MS analysis in positive **(A)** and negative **(B)** ionization modes, HILIC-LC-TOF-MS analysis in positive **(C)** and negative **(D)** ionization modes and from GC-MS analysis **(E)**. Red circles correspond to samples from disease group, blue ones represent healthy controls and yellow ones QC samples.

Furthermore, no significant differences were observed between male and female patients, according to the applied uni- and multivariate statistical techniques. The age of the subjects did not influence the differences in findings between the cohorts.

DISCUSSION

Renal dysplasia constitutes a complex and multifaceted disorder characterized by abnormal renal cell differentiation, which leads to the presence of primitive tubules, interstitial fibrosis, renal cysts and cartilage in the renal parenchyma (Phua and Ho, 2016). The most common etiologies of renal dysplasia include both intrinsic defects in the renal parenchyma's differentiation and functional or structural obstruction of the lower urinary tract (Woolf et al., 2004). Recently, several genetic mutations, mainly associated with Six2, Wnt, Bmp7, and Hnf1β, and copy number variations have been identified in patients with renal dysplasia (Weber et al., 2006; Braun et al., 2016; Verbitsky et al., 2019). Nevertheless, the pathophysiology and underlying molecular mechanisms of renal dysplasia still remain poorly explored and understood in spite of it being one of the most common causes of renal failure in neonates and a leading cause of CKD in childhood.

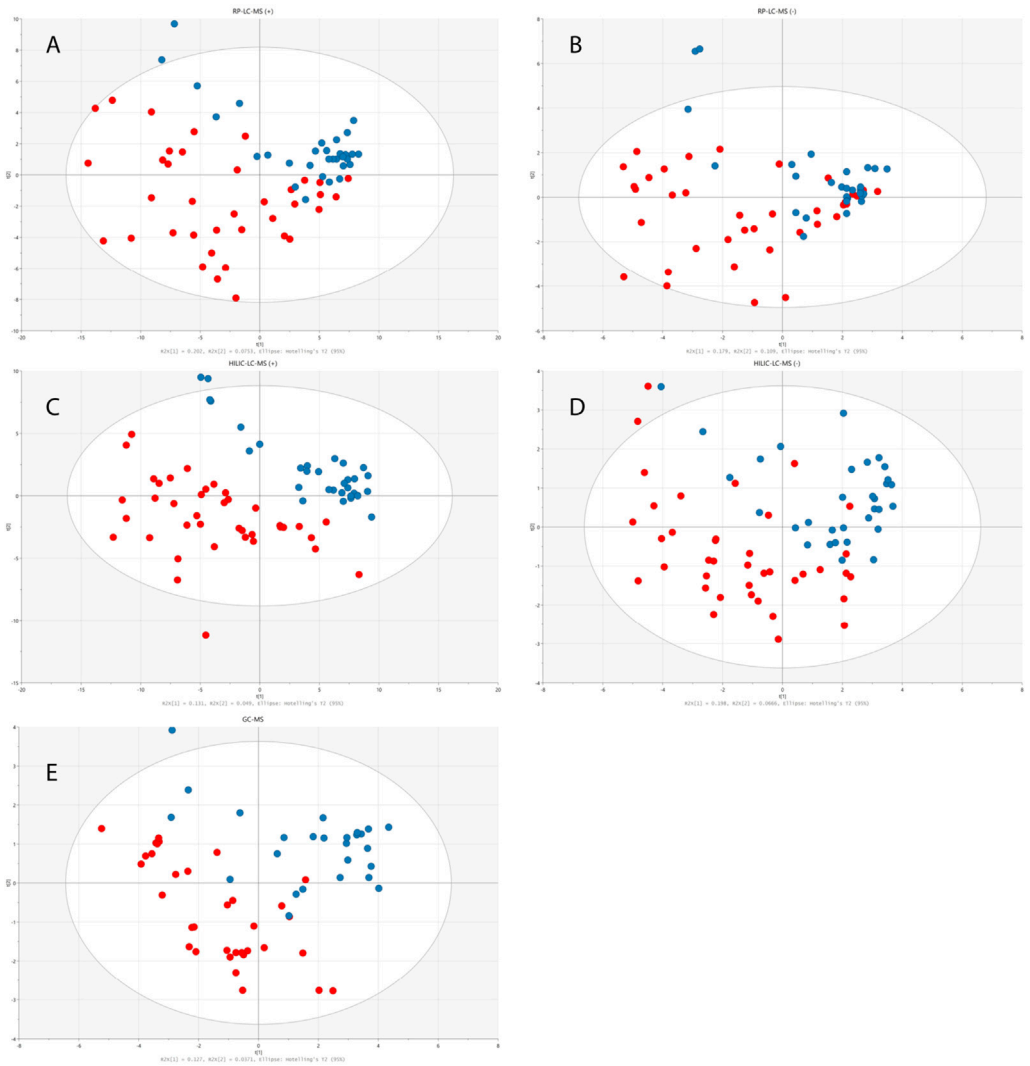

FIGURE 3 | PLS-DA models built on data obtained from RP-LC-TOF-MS analysis in positive **(A)** (R^2 = 0.25, Q^2 = 0.282, CV-ANOVA $p < 0.001$) and negative **(B)** (R^2 = 0.453, Q^2 = 0.236, CV-ANOVA p = 0.013) ionization modes, HILIC-LC-TOF-MS analysis in positive **(C)** (R^2 = 0.227, Q^2 = 0.505, CV-ANOVA $p < 0.001$) and negative **(D)** (R^2 = 0.264, Q^2 = 0.271, CV-ANOVA $p < 0.001$) ionization modes and from GC-MS analysis **(E)** (R^2 = 0.164, Q^2 = 0.174, CV-ANOVA p = 0.017). Red circles correspond to samples from disease group, while blue ones represent healthy controls.

To the authors' best knowledge, no publicly available studies have used metabolomics to investigate renal dysplasia, especially in children. Benito et al. (2018) have evaluated the metabolic indicators of CKD in a cohort which included a pediatric population. The results of this study show increased levels of sphingosine-1-phosphate, n-butyrylcarnitine, and cis-4-decenoylcarnitine in plasma from patients with CKD and decreased level of bilirubin. The authors have stressed that kidney function (estimated by the eGFR) is related to about one third of the detected metabolites in both the general and CKD populations (Benito et al., 2018).

In this study, the observed urinary metabolite changes derived mainly from purine, lipid and amino acid metabolism as well as from glycolysis, the TCA cycle and the urea cycle. Altered levels of acetylasparagine, trimethylamine-N-Oxide, betaine, dimethylarginine, hippuric acid, uric acid, and hypoxanthine were potentially more characteristic of impaired kidney function measured by decreased eGFR. Most of these metabolites have been previously described in terms of CKD pathophysiology. The main biochemical pathways associated with renal dysplasia and/or a decreased eGFR are graphically displayed in **Figure 4**.

Metabolic Changes Related to CKD Pathophysiology

In a recent study based on plasma untargeted metabolomics, the increase in some N-acetyl amino acids was observed in all stages of CKD (*N*-acetylmethionine, *N*-acetylserine, *N*-acetyltryptophan, *N*-acetylglycine, *N*-acetylphenylalanine,

TABLE 2 | Statistically significant metabolites differentiating patients with renal dysplasia from healthy controls.

Compound	Analytical technique	Retention time (min)	Fragmentation pattern	Statistical analysis	Change in dysplasia	Fold change	Biochemical pathway
Methylguanosine	RP (+)	3.1	166.0843	UV	Decrease	–35%	Purine metabolism
6-Keto-decanoylcarnitine	RP (+)	10.1	60.0806, 85.0283, 271.1543	UV	Decrease	–44%	Fatty acid metabolism
Dodecanedioylcarnitine	RP (+)	10.9	60.0806, 85.0284	UV	Decrease	–36%	Fatty acid metabolism
	HILIC(+)	10.5	297.1695	UV		–37%	
Hydroxyisovaleroylcarnitine	RP (+)	1.6	60.0807, 85.0284, 185.0809	UV	Decrease	–31%	Fatty acid metabolism
Hydroxydecanoylcarnitine	RP (+)	10.3	60.0806, 85.0283, 255.1593	UV	Decrease	–55%	Fatty acid metabolism
Citric acid	RP (–)	1.0	111.0088, 85.0295, 87.0088	VIP, SR	Increase	+31%	Citric acid cycle (TCA)
Hippuric acid	RP (–)	6.8	134.0615, 77.0399, 56.0144	VIP	Decrease	–10%	Phenylalanine metabolism
Furoic acid	RP (–)	1.0	67.0190	SR	Increase	+16%	Microbial metabolism
Dimethylguanosine	RP (–)	5.2	178.0736,220.0839, 192.0892	SR	Increase	+18%	Degradation product of tRNA
Betaine	HILIC(+)	11.5	58.0650, 59.0728	UV, VIP, SR	Increase	+80%	Glycine and serine metabolism, methionine metabolism
Nonanoylcarnitine	HILIC(+)	10.2	60.0805, 85.0283, 243.1591	UV	Decrease	–46%	Fatty acidmetabolism
Tiglylcarnitine	HILIC(+)	11.4	60.0807,85.0284	SR	Decrease	–54%	Fatty acid metabolism
Butyrylcarnitine	HILIC(+)	11.7	60.0805, 85.0282, 173.0808	UV	Decrease	–55%	Fatty acidmetabolism
Trimethylamine N-oxide	HILIC(+)	11.9	58.0649, 59.0728	VIP	Decrease	–34%	Microbial metabolism in diverse environments
Carnitine	HILIC(+)	12.7	57.0334, 60.0807, 85.0284, 103.0391	VIP	Increase	+24%	Fatty acid metabolism
Dimethylarginine	HILIC(+)	13.2	70.0650, 88.0868, 116.0706, 158.1288	VIP	Decrease	–42%	L-arginine derivative
Xanthine	HILIC(–)	2.8	80.9652, 108.0206	UV	Decrease	–43%	Purine metabolism
Uric acid	HILIC(–)	4.6	124.0143, 96.0195, 69.0086	VIP, SR	Increase	+228%	Purine metabolism
Indoxyl sulfate	HILIC(–)	1.0	79.9571,80.9557, 132.0452	UV	Decrease	–18%	Tryptophan metabolite
p-Cresol sulfate	HILIC(–)	1.0	107.0501, 79.9572	VIP, SR	Decrease	–40%	Microbial metabolism
Glutamine	HILIC(–)	7.2	127.0513,128.0353, 109.0407, 101.0720	UV	Decrease	–35%	D-Glutamine and D-glutamate metabolism
Hexadecanoic acid	GC	23.0	313.0,117.0,132.0	UV, SR	Increase	+38%	Fatty acid biosynthesis
Threonic acid	GC	15.9	292.0, 205.0, 220.0	UV, VIP	Increase	+20%	Ascorbate and aldarate metabolism
Glyceric acid	GC	12.3	292.0, 189.0, 133.0	UV	Increase	+234%	Glycerolipid metabolism, glycine and serine metabolism
Arabitol	GC	18.5	307.0,217.0,103.0	UV, VIP	Increase	+51%	Pentose and glucuronate interconversions
Lactose	GC	30.1	361.0,204.0,319.0	UV	Increase	+15%	Lactose synthesis, galactose metabolism
Aconitic acid	GC	18.9	375.0,229.0,285.0	UV	Decrease	–30%	Citric acid cycle (TCA)
Lactic acid	GC	7.3	117.0,191.0	UV	Increase	+270%	Glycolysis/ Gluconeogenesis

UV- univariate statistical analysis, VIP- variable importance in projection, SR-selectivity ratio.

N-acetylleucine, *N*-acetylproline, *N*-acetyllysine, *N*-acetylasparagine, and *N*-acetylaspartic acid) (Gagnebin et al., 2019). *N*-acetylserine and *N*-acetyllysine were indicated as risk factors of end-stage renal disease for type 1 diabetes and CKD (Niewczas et al., 2017). Sekula et al. (2016) observed that *N*-acetylation might be a crucial detoxification mechanism in CKD. Moreover, *N*-acetylalanine was also observed to be correlated with GFR (Sekula et al., 2016). In the presented study, an increased urinary level of *N*-acetylasparagine was observed in patients with renal dysplasia and was correlated with a decreased eGFR. The above alterations of *N*-acetylated compounds in plasma and urine may indicate a link between *N*-acetylation and renal dysfunction. Trimethylamine-N-oxide (TMAO), a small amine plasma molecule, originates mainly from the intestinal microbiota's metabolism. Gut microbiota produce trimethylamine (TMA) from food products containing TMA

TABLE 3 | Statistically significant metabolites differentiating patients with normal and reduced estimated glomerular filtration rate (eGFR).

Compound	Analytical technique	Statistical analysis	Biochemical pathway
Methylguanosine	RP (+)	UV	Purine metabolism
Citric acid	RP (–)	VIP	Citric acid cycle (TCA)
Hippuric acid	RP (–)	VIP	Phenylalanine metabolism
Betaine	HILIC(+)	VIP, SR	Glycine and serine metabolism, methionine metabolism
Trimethylamine N-oxide	HILIC(+)	VIP	Microbial metabolism in diverse environments
Carnitine	HILIC(+)	VIP	Fatty acid metabolism
Dimethylarginine	HILIC(+)	UV, VIP	L-arginine derivative
Uric acid	HILIC(–)	VIP	Purine metabolism
p-Cresol sulfate	HILIC(–)	VIP, SR	Microbial metabolism

or TMAO and dietary precursors such as choline, phosphatidylcholine, betaine, and carnitine (Al-Waiz et al., 1992; Zhang et al., 1999; Bain et al., 2005). TMA is subsequently absorbed through the intestinal barrier into the bloodstream, then N-oxidized by the hepatic enzyme flavin-containing monooxygenase isoform 3 (FMO3) and excreted as TMAO with urine (de la Huerga et al., 1951; Bell et al., 1991). Thus, TMAO levels could be a result of various production processes, including dietary precursor intake, endogenous TMA production from gut microbiota, TMA and TMAO intestinal absorption, as well as FMO3 enzymatic activity and its renal excretion (Pelletier et al., 2019). Some previous studies have already reported TMAO accumulation in CKD patients (Gagnebin et al., 2019; Pelletier et al., 2019). In this study, a decrease in the urinary level of TMAO was observed in renal dysplasia patients with altered eGFR. A recent study assessed the plasma TMA, TMAO, choline, betaine, and carnitine concentrations in the consecutive stages of CKD (Pelletier et al., 2019) using the measured glomerular filtration rate (mGFR) and the renal clearance. TMAO, choline, and carnitine were inversely correlated with the mGFR in CKD patients.

The elimination of circulating betaine in humans is mainly due to its metabolism rather than renal excretion (Schwahn et al., 2003). However, some previous studies reported that renal excretion of betaine is elevated in patients with kidney injury (Lever et al., 1994; Dellow et al., 1999). In this study the increased urinary level of betaine was observed in renal dysplasia patients with accompanying decreased eGFR, which rather supports the latter opinion. Similar findings were reported by Missailidis at al. (Missailidis et al., 2016), who observed a decrease in the plasma betaine level that was associated with a declined renal function, with the lowest levels observed in stage 5 CKD patients. The results may vary due to different disease stages in these studies. Hippuric acid constitutes the glycine conjugate of benzoic acid, which originates from the phenylalanine metabolism (Zhao et al., 2013). It is primarily eliminated from the blood by the kidneys, through active tubular secretion via organic anion transporters (Deguchi et al., 2005). Additionally, hippuric acid represents one of the well-known protein-bound uremic toxins. In this study, the decreased urinary level of this metabolite was observed as significant in the statistical comparisons between the renal dysplasia patients and the control group, as well as among the renal dysplasia patients with differences in eGFR values. An earlier study also showed the increased tissue level of hippuric acid in CKD rats in comparison to the control group, probably due to the reduced renal clearance of these metabolites. Nevertheless, hippuric acid in humans is also an excretory product of environmental-toxic exposures, dietary protein degradation, and resynthesis by intestinal microbial metabolism of quinic acid through the shikimate pathway (Pero, 2010).

In the presented study, alterations in the purine metabolism were observed in both comparisons, namely in renal dysplasia patients as compared to the control group and in renal dysplasia patients regarding the differences in the eGFR. The metabolism of uric acid is a complex process that includes hepatic production and renal as well as gut excretion (Maiuolo et al., 2016). Uric acid constitutes the end product of both exogenous and endogenous purine metabolisms (Chaudhary et al., 2013). The endogenous production takes place mainly in the liver, intestines, muscles, kidneys, and the vascular endothelium (Chaudhary et al., 2013). Approximately two-thirds of the uric acid load are eliminated by the kidneys, while the remaining one-third is excreted by the gastrointestinal system. Almost all uric acid is filtered from the glomeruli and the amount of its excretion is regulated by post-glomerular reabsorption and secretion (Maiuolo et al., 2016). Reabsorption of uric acid occurs at the S1 segment of the proximal tubule and approximately 10% of the filtered uric acid appears in the urine (Chaudhary et al., 2013). Therefore, hyperuricemia is considered as a crucial risk factor for the renal

TABLE 4 | Metabolites with the highest correlation with eGFR value (Spearman's rank correlation).

Compound	r	Analytical technique
3-Methylglutarylcarnitine	–0.58	RP+
Methyluric acid	–0.44	RP+
	–0.40	RP-
	–0.60	HILIC-
Methylguanosine	0.77	RP+
	0.63	HILIC+
N-Acetylasparagine	–0.68	HILIC+
Guanidinosuccinic acid	–0.48	HILIC+
Dimethylarginine	0.71	HILIC+
Hypoxanthine	0.72	HILIC+
Xanthine	0.47	HILIC-
Ethanolamine	0.63	GC
Uracil	0.58	GC
D-Allose	0.38	GC
Phenylalanine	0.38	GC

TABLE 5 | Metabolic signature of renal dysplasia, unrelated to eGFR value.

Compound	Change in dysplasia	Biochemical pathway
6-Keto-decanoylcarnitine	Decrease	Fatty acid metabolism
Dodecanedioylcarnitine	Decrease	Fatty acid metabolism
Hydroxyisovaleroylcarnitine	Decrease	Fatty acid metabolism
Hydroxydecanoylcarnitine	Decrease	Fatty acid metabolism
Furoic acid	Increase	Microbial metabolism
Dimethylguanosine	Increase	Degradation product of tRNA
Nonanoylcarnitine	Decrease	Fatty acidmetabolism
Tiglylcarnitine	Decrease	Fatty acid metabolism
Butyrylcarnitine	Decrease	Fatty acidmetabolism
Xanthine	Decrease	Purine metabolism
Indoxyl sulfate	Decrease	Tryptophan metabolite
Glutamine	Decrease	D-Glutamine and D-glutamate metabolism
Hexadecanoic acid	Increase	Fatty acid biosynthesis
Threonic acid	Increase	Ascorbate and aldarate metabolism
Glyceric acid	Increase	Glycerolipid metabolism, glycine and serine metabolism
Arabitol	Increase	Pentose and glucuronate interconversions
Lactose	Increase	Lactose synthesis, galactose metabolism
Aconitic acid	Decrease	Citric acid cycle (TCA)
Lactic acid	Increase	Glycolysis/Gluconeogenesis

dysfunction, hypertension, hyperlipidemia, diabetes, and obesity (Maiuolo et al., 2016). Hyperuricemia may be a consequence of the increased production or impaired renal excretion, as well as of a combination of both processes (Su et al., 2014). In this study the increased urinary levels of uric acid and hypoxanthine were observed in renal dysplasia patients as compared to the control group, also in terms of eGFR differences. There are many previous studies that indicated the blood hyperuricemia and disturbed purine nucleotide metabolism as potential contributory risk factors in the development and progression of CKD (Johnson et al., 2013; Mazumder et al., 2018; Zoccali and Mallamaci, 2018; Oh et al., 2019). Thus, the disturbances of the purine metabolism observed in this study, may rather indicate kidney dysfunction than renal dysplasia.

Metabolomic Signature Potentially Characteristic for Renal Dysplasia

The metabolic changes, observed in this study as statistically significant only between renal dysplasia patients and control

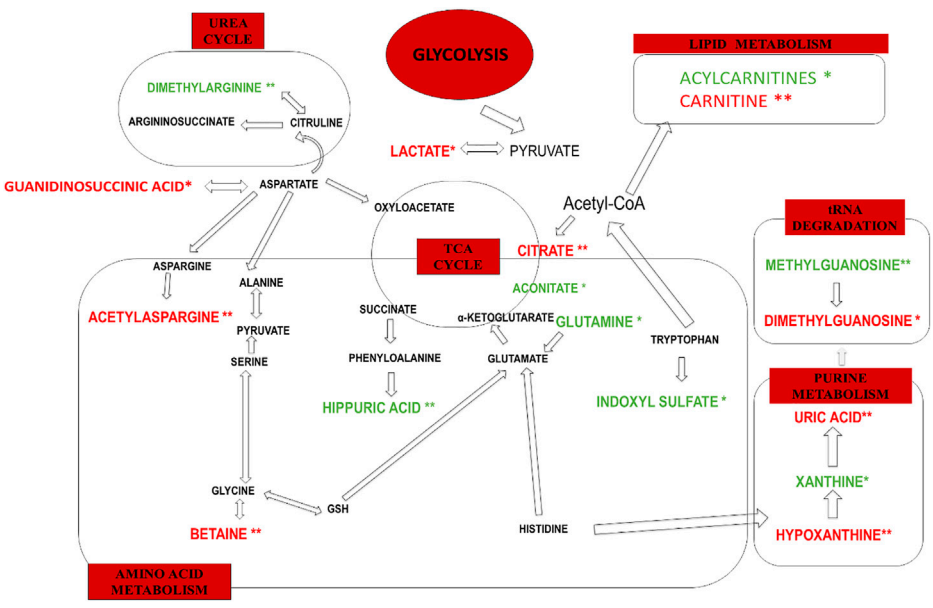

FIGURE 4 | The main biochemical pathways altered only in renal dysplasia (*) and renal dysplasia accompanied by decreased eGFR (**). Decreased levels of urinary metabolites are marked in green. Increased urinary metabolite levels are marked in red.

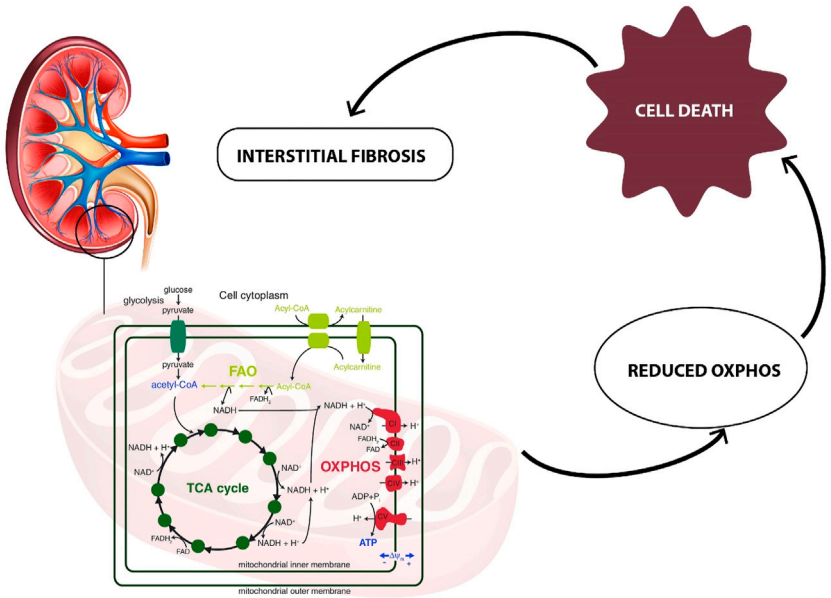

FIGURE 5 | Relationship between metabolic reprogramming of mitochondrial metabolism (FAO, OXPHOS) and kidney interstitial fibrosis.

group, include decreased urinary levels of acylcarnitines, indoxyl sulfate, xanthine, aconitate, glutamine as well as increased urinary levels of lactate, dimethylguanosine, and guanidinosuccinic acid.

Acylcarnitines, esters of L-carnitine, and fatty acids play a crucial role in the cellular metabolism (Li et al., 2019). The main function of acylcarnitines constitutes long-chain fatty acids (LCFAs) metabolism, as they transport activated LCFAs into the mitochondria for subsequent β-oxidation to provide energy for various cell processes (Tarasenko et al., 2018). Acylcarnitines are also involved in glycolysis, TCA cycle, branched-chain amino acid metabolism, fatty acid peroxidation, and ketone bodies production. Therefore, they are key factors regulating the balance of the intracellular sugar and lipid metabolism (Qu et al., 2016). In this study, the decreased urinary levels of a few acylcarnitines (6-keto-decanoylcarnitine, dodecanedioylcarnitine, hydroxyisovaleroylcarnitine, hydroxydecanoylcarnitine, nonanoylcarnitine, butyrylcarnitine) were observed as statistically significant in renal dysplasia patients as compared to the control group. These alterations can be explained by blood accumulation of acylcarnitines, potentially associated with mitochondrial dysfunction. The matrix of mitochondria constitutes a central place for metabolic pathways such as the TCA cycle and oxidative phosphorylation (OXPHOS) (Podrini et al., 2020). The reduced forms of nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide ($FADH_2$) derived from the glycolysis pathway. The fatty acid oxidation (FAO) and TCA cycle are energy-rich molecules containing a pair of electrons with high transfer potential. These electrons are used to reduce molecular oxygen to water and large amount of free energy is released, which subsequently can be used for adenosine triphosphate (ATP) generation. OXPHOS constitutes a process involved in ATP production as a result of the electron transfer from NADH or $FADH_2$ to O_2.

Several renal diseases, including tubular disorders, chronic tubulointerstitial nephritis, cystic renal disease, and glomerular diseases were reported as mitochondrial cytopathies affecting the OXPHOS activity (Ueda et al., 2004; Au et al., 2007; Emma et al., 2011). One of the recent reports, indicated mitochondrial damage as a key feature of renal inflammation and fibrosis (Chung et al., 2019). In this study, the human kidney tissue and kidney tissue samples collected from animal models with fibrosis were analyzed. The significant mitochondrial defect, including the loss of the mitochondrial transcription factor A (TFAM) in kidney tubular cells, resulting in a reduced OXPHOS, was observed. Additionally, the kidney histological analysis was performed and significant epithelial atrophy, dilated tubules, and interstitial fibrosis were indicated (Chung et al., 2019). Interstitial fibrosis is frequently present in dysplastic kidneys and increases with the progression of CKD. Since FAO is the main energy source for renal proximal tubular epithelial cells, the reduced FAO process would impact the lipid metabolism (Zhou and Liu, 2016). It could lead to disruption of balance between fatty acid synthesis and consumption, as well as dysregulation of intracellular lipid accumulation. Inhibition of FAO in tubular epithelial cells *in vitro* results in ATP depletion, apoptosis, cell dedifferentiation, and intracellular lipid deposition (Kang et al., 2015). The potential relationship between metabolic reprogramming of mitochondrial metabolism (FAO, OXPHOS) and kidney fibrosis is presented in **Figure 5**.

The metabolic alterations related to the glutamine, aconitate, and lactate levels were also observed to be statistically significant in the comparison between renal dysplasia patients and the control group. All these alterations indicate the involvement of the aerobic glycolysis (a Warburg-like effect), glutamine anaplerosis and the dysregulation of fatty acid biosynthesis. These metabolic changes are also connected with the above

described reduced FAO and OXPHOS activities. Recently, the same metabolic reprogramming related to mitochondrial dysfunction has been observed in polycystic kidney disease (Podrini et al., 2020). Additionally, Zhao et al. (2016) reported these metabolic alterations in renal fibrosis in rats, which suggests that the cells utilize mainly the glucose and lipid metabolism to maintain energy homeostasis during renal fibrotic process.

In this report, the decrease in urinary level of indoxyl sulfate was observed in renal dysplasia patients as compared to control group. The gut microbiota convert the dietary tryptophan to indole, which is absorbed by the intestine and subsequently metabolized to indoxyl sulfate in the liver (Pezzatti et al., 2019). Indoxyl sulfate represents another established uremic toxin (Pezzatti et al., 2019). The above-mentioned metabolic alteration observed in this study, may be related to renal dysfunction and altered kidney elimination of indoxyl sulfate. However, in the recent ^{1}H-NMR-based kidney and urine untargeted metabolomics of renal interstitial fibrosis rats, the altered level of indoxyl sulfate was observed (Zhao et al., 2016). Zhao et al. (2013) reported the increased level of indoxyl sulfate in the kidney of a rat model of early renal injury, using UPLC Q-TOF/HSMS untargeted metabolomics approach. Moreover, many of previous studies, marked this metabolite as indicator of renal function, vascular disease and mortality in CKD patients (Gika et al., 2012; Gagnebin et al., 2019; Pezzatti et al., 2019). By contrast, in the longitudinal metabolomics studies, plasma levels of indoxyl sulfate showed no association with incident CKD (Rhee et al., 2013) or with CKD progression (Niewczas et al., 2014). The observed results of the above-mentioned studies may be associated with various biological models used, different range of patients' age or the stage of the disease. However, indoxyl sulfate seems to be a metabolic indicator of early kidney dysfunction or renal histopathological changes.

Xanthine represents a significant metabolic byproduct of guanine triphosphate (GTP) or guanine metabolism and is derived from the purine metabolism pathway (Giulia Battelli et al., 2016). The elevated plasma or urinary xanthine levels may result from the inhibition of xanthine oxidase or from the blockage of the metabolism of xanthine to uric acid (Marro et al., 1997). In this study, the decline in urinary level of xanthine was observed in renal dysplasia patients as compared to the control group. The decrease of this metabolite in the kidneys of rats with renal fibrosis was also recently reported (Zhou and Liu, 2016). This alteration may underline the reduced purine metabolism and bioenergy production. The same trend in xanthine levels in the kidneys was observed in an animal model of early renal injury (Zhao et al., 2013).

Dimethylguanosine, a modified nucleotide indirectly associated with purine metabolism, constitutes a degradation product of transfer RNA and is mainly excreted by the kidneys (Tsalik et al., 2015). In this study, the elevation in urinary level of dimethylguanosine was observed in the renal dysplasia group as compared to the healthy subjects. Previously, consistently higher serum levels of this metabolite were reported in polycystic kidney disease (Grams et al., 2017). Probably, the early decline in kidney blood flow in polycystic kidney disease compared with GFR differentially affects the secretion of small molecules by the proximal tubule. Dimethylguanosine was also indicated to have reduced urinary excretion in patients with kidney failure related to proximal tubule function (Niwa et al., 1998). Additionally, the elevated plasma level of dimethylguanosine was observed in patients with acute renal injury (Tsalik et al., 2015).

Guanidinosuccinic acid (GSA), a derivative of L-arginine, is a precursor of nitric oxide (NO) which tends to accumulate in uremic plasma (Zerra and Josephson, 2019). GSA constitutes also an example of a well-known uremic toxin. GSA impairs the secondary wave of ADP-induced platelet aggregation as well as the release and synthesis of thromboxane A2 in platelets in advanced renal disease (Zerra and Josephson, 2019). In this study, the statistically significant elevation in urinary level of GSA was observed in the renal dysplasia subjects as compared to the control group. Previously, the increased serum level of GSA was reported in end-stage renal failure patients (de Deyn et al., 2003). Additionally, the accumulation of guanidino compounds has been associated with neurological, cardiovascular, hematological, and immunological complications of renal failure (Ringoir et al., 1988). The increased levels of GSA were also reported in the plasma and kidneys of a rat model of polycystic kidney disease (Torremans et al., 2006).

In summary, the urine metabolic changes discovered in children with dysplastic kidneys seem to be characteristic and point towards the presence of altered fatty acid oxidation, amino acid, and purine metabolisms in this parenchymal disorder. However, there are some limitations of this research. The studied renal dysplasia group included mostly young children on different diets and with a significant proportion of subjects with a decreased GFR. Furthermore, while decreased renal function has been reported to influence the metabolic findings in renal diseases, the influence of age and diet requires more extensive investigation. It is also difficult to unequivocally interpret the recognized metabolic signature that is present in children with renal dysplasia as specific for this disorder. It may also represent early abnormalities of the initial stages of CKD which are poorly described. Another option may be that the obtained metabolic signature is due to the presence of renal cysts or ongoing processes of kidney fibrosis.

Integration of the obtained metabolic data with further proteomic, genomic or transcriptomics research may unravel the still poorly understood mechanisms of progression of renal dysplasia in the future. Additionally, it may facilitate earlier recognition of renal dysplasia and enable the introduction of novel therapies for nephroprotective management of children with this congenital abnormality.

CONCLUSION

The novel application of a comprehensive metabolomic analysis enabled the recognition of a characteristic urinary metabolic profile for renal dysplasia, allowing the evaluation of different metabolic pathways involved in this disorder. Metabolites associated with the decreased eGFR were excluded to eliminate

the influence of decreased kidney function which has been recognized as an important confounding factor. The main biochemical pathways that have been found to be altered in dysplastic kidneys include the glycolysis pathway, the lipid, purine and amino acid metabolism, and the TCA and urea cycles. We suggest that the decreased levels of acylcarnitines in the urine of the renal dysplasia subjects are caused by their accumulation in the blood, due to mitochondrial dysfunction. In consequence, oxidative phosphorylation and fatty acid oxidation may be disturbed, leading to ATP depletion, apoptosis, cell dedifferentiation, and intracellular lipid deposition. A further validation of the reported results is necessary and should be performed in larger populations of children with renal dysplasia, notably in those with a normal renal function.

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 Independent Bioethics Committee for Scientific Research at Medical University of Gdańsk. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

SM and RW performed metabolomics analysis, data treatment, statistical analysis, and writing of the manuscript. AK participated in the collection of biological samples and clinical data, in the analysis of data and writing of the manuscript. MK contributed in analytical measurements and data analysis. WS-L, DD, and MB participated in data processing and treatment as well as revision of the manuscript. MM participated in the collection and analysis of data and revision of the manuscript. AZ conceived the original idea and participated in the design, analysis, and writing and revision of the manuscript. MJM contributed in the study design, writing, and revision of the manuscript.

FUNDING

This work was supported by the National Science Centre in Poland (Grant no 2015/19/N/NZ7/03397 and 2012/07/E/NZ7/0441). Publication of the article was supported by the project POWR.03.02.00-00-I026/17-00 co-financed by the European Union through the European Social Fund under the Operational Programme Knowledge Education Development 2014–2020.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.665661/full#supplementary-material

REFERENCES

Al-Waiz, M., Mikov, M., Mitchell, S. C., and Smith, R. L. (1992). The Exogenous Origin of Trimethylamine in the Mouse. *Metabolism.* 41 (2), 135–136. doi:10.1016/0026-0495(92)90140-6

Armitage, E. G., and Barbas, C. (2014). Metabolomics in Cancer Biomarker Discovery: Current Trends and Future Perspectives. *J. Pharm. Biomed. Anal.* 87, 1–11. doi:10.1016/j.jpba.2013.08.041

Atzori, L., Antonucci, R., Barberini, L., Locci, E., Cesare Marincola, F., Scano, P., et al. (2010). 1H NMR-Based Metabolic Profiling of Urine From Children with Nephrouropathies. *Front. Biosci.* E2 (2), 725–732. doi:10.2741/e132

Au, K. M., Lau, S. C., Mak, Y. F., Lai, W. M., Chow, T. C., Chen, M. L., et al. (2007). Mitochondrial DNA Deletion in a Girl With Fanconi's Syndrome. *Pediatr. Nephrol.* 22 (1), 136–140. doi:10.1007/s00467-006-0288-y

Bain, M., Fornasini, G., and Evans, A. (2005). Trimethylamine: Metabolic, Pharmacokinetic and Safety Aspects. *Curr. Drug Metab.* 6 (3), 227–240. doi:10.2174/1389200054021807

Bell, J. D., Lee, J. A., Lee, H. A., Sadler, P. J., Wilkie, D. R., and Woodham, R. H. (1991). Nuclear Magnetic Resonance Studies of Blood Plasma and Urine From Subjects With Chronic Renal Failure: Identification of Trimethylamine-N-Oxide. *BBA - Mol. Basis Dis.* 1096 (2), 101–107. doi:10.1016/0925-4439(91)90046-C

Benito, S., Sánchez-Ortega, A., Unceta, N., Andrade, F., Aldámiz-Echevarria, L., Goicolea, M. A., et al. (2018). Untargeted Metabolomics for Plasma Biomarker Discovery for Early Chronic Kidney Disease Diagnosis in Pediatric Patients Using LC-QTOF-MS. *Analyst.* 143 (18), 4448–4458. doi:10.1039/c8an00864g

Braun, D. A., Schueler, M., Halbritter, J., Gee, H. Y., Porath, J. D., Lawson, J. A., et al. (2016). Whole Exome Sequencing Identifies Causative Mutations in the Majority of Consanguineous or Familial Cases With Childhood-Onset Increased Renal Echogenicity. *Kidney Int.* 89 (2), 468–475. doi:10.1038/ki.2015.317

Chaudhary, K., Malhotra, K., Sowers, J., and Aroor, A. (2013). Uric Acid - Key Ingredient in the Recipe for Cardiorenal Metabolic Syndrome. *Cardiorenal Med.* 3, 208–220. doi:10.1159/000355405

Chung, K. W., Dhillon, P., Huang, S., Sheng, X., Shrestha, R., Qiu, C., et al. (2019). Mitochondrial Damage and Activation of the STING Pathway Lead to Renal Inflammation and Fibrosis. *Cel Metab.* 30 (4), 784–799. doi:10.1016/j.cmet.2019.08.003

De Deyn, P. P., Vanholder, R., and D'Hooge, R. (2003). Nitric Oxide in Uremia: Effects of Several Potentially Toxic Guanidino Compounds. *Kidney Int.* 63, S25–S28. doi:10.1046/j.1523-1755.63.s84.9.x

De La Huerga, J., Popper, H., and Steigmann, F. (1951). Urinary Excretion of Choline and Trimethylamines After Intravenous Administration of Choline in Liver Diseases. *J. Lab. Clin. Med.* 38 (6), 904–910. doi:10.5555/uri:pii:0022214351900765

Deguchi, T., Takemoto, M., Uehara, N., Lindup, W. E., Suenaga, A., and Otagiri, M. (2005). Renal Clearance of Endogenous Hippurate Correlates With Expression Levels of Renal Organic Anion Transporters in Uremic Rats. *J. Pharmacol. Exp. Ther.* 314 (2), 932–938. doi:10.1124/jpet.105.085613

Dellow, W. J., Chambers, S. T., Lever, M., Lunt, H., and Robson, R. A. (1999). Elevated Glycine Betaine Excretion in Diabetes Mellitus Patients Is Associated With Proximal Tubular Dysfunction and Hyperglycemia. *Diabetes Res. Clin. Pract.* 43 (2), 91–99. doi:10.1016/S0168-8227(98)00115-6

Emma, F., Montini, G., Salviati, L., and Dionisi-Vici, C. (2011). Renal Mitochondrial Cytopathies. *Int. J. Nephrol.* 2011, 1–10. doi:10.4061/2011/609213

Gagnebin, Y., Pezzatti, J., Lescuyer, P., Boccard, J., Ponte, B., and Rudaz, S. (2019). Toward a Better Understanding of Chronic Kidney Disease with Complementary Chromatographic Methods Hyphenated with Mass

Spectrometry for Improved Polar Metabolome Coverage. *J. Chromatogr. B.* 1116, 9–18. doi:10.1016/j.jchromb.2019.03.031

Gika, H. G., Theodoridis, G. A., Earll, M., and Wilson, I. D. (2012). A QC Approach to the Determination of Day-To-Day Reproducibility and Robustness of LC-MS Methods for Global Metabolite Profiling in Metabonomics/Metabolomics. *Bioanalysis.* 4 (18), 2239–2247. doi:10.4155/bio.12.212

Giulia Battelli, M., Polito, L., Bortolotti, M., and Bolognesi, A. (2016). Xanthine Oxidoreductase in Drug Metabolism: Beyond a Role as a Detoxifying Enzyme. *Curr. Med. Chem.* 23 (35), 4027–4036. doi:10.2174/0929867323666160725091915

Grams, M. E., Tin, A., Rebholz, C. M., Shafi, T., Köttgen, A., Perrone, R. D., et al. (2017). Metabolomic Alterations Associated With Cause of CKD. *Clin. J. Am. Soc. Nephrol.* 12 (11), 1787–1794. doi:10.2215/CJN.02560317

Hocher, B., and Adamski, J. (2017). Metabolomics for Clinical Use and Research in Chronic Kidney Disease. *Nat. Rev. Nephrol.* 13 (March), 269. doi:10.1038/nrneph.2017.30

Johnson, R. J., Nakagawa, T., Jalal, D., Sánchez-Lozada, L. G., Kang, D.-H., and Ritz, E. (2013). Uric Acid and Chronic Kidney Disease: Which Is Chasing Which? *Nephrol. Dial. Transplant.* 28, 2221–2228. doi:10.1093/ndt/gft029

Kalim, S., and Rhee, E. P. (2017). An Overview of Renal Metabolomics. *Kidney Int.* 91, 61–69. doi:10.1016/j.kint.2016.08.021

Kang, H. M., Ahn, S. H., Choi, P., Ko, Y.-A., Han, S. H., Chinga, F., et al. (2015). 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

Kaushik, A. K., and DeBerardinis, R. J. (2018). Applications of Metabolomics to Study Cancer Metabolism. *Biochim. Biophys. Acta (Bba) - Rev. Cancer.* 1870, 2–14. doi:10.1016/j.bbcan.2018.04.009

Lever, M., Sizeland, P. C., Bason, L. M., Hayman, C. M., Robson, R. A., and Chambers, S. T. (1994). Abnormal Glycine Betaine Content of the Blood and Urine of Diabetic and Renal Patients. *Clinica Chim. Acta* 230 (1), 69–79. doi:10.1016/0009-8981(94)90090-6

Li, S., Gao, D., and Jiang, Y. (2019). Function, Detection and Alteration of Acylcarnitine Metabolism in Hepatocellular Carcinoma. *Metabolites.* 9, 36. doi:10.3390/metabo9020036

López-López, Á., López-Gonzálvez, Á., Barker-Tejeda, T. C., and Barbas, C. (2018). A Review of Validated Biomarkers Obtained Through Metabolomics. *Expert Rev. Mol. Diagn.* 18, 557–575. doi:10.1080/14737159.2018.1481391

Maiuolo, J., Oppedisano, F., Gratteri, S., Muscoli, C., and Mollace, V. (2016). Regulation of Uric Acid Metabolism and Excretion. *Int. J. Cardiol.* 213, 8–14. doi:10.1016/j.ijcard.2015.08.109

Marro, P. J., Baumgart, S., Delivoria-Papadopoulos, M., Zirin, S., Corcoran, L., McGaurn, S. P., et al. (1997). Purine Metabolism and Inhibition of Xanthine Oxidase in Severely Hypoxic Neonates Going onto Extracorporeal Membrane Oxygenation1. *Pediatr. Res.* 41 (4 I), 513–520. doi:10.1203/00006450-199704000-00010

Mazumder, M. K., Phukan, B. C., Bhattacharjee, A., and Borah, A. (2018). Disturbed Purine Nucleotide Metabolism in Chronic Kidney Disease Is a Risk Factor for Cognitive Impairment. *Med. Hypotheses.* 111, 36–39. doi:10.1016/j.mehy.2017.12.016

Missailidis, C., Hällqvist, J., Qureshi, A. R., Barany, P., Heimbürger, O., Lindholm, B., et al. (2016). Serum Trimethylamine-N-Oxide Is Strongly Related to Renal Function and Predicts Outcome in Chronic Kidney Disease. *Plos One* 11 (1), e0141738. doi:10.1371/journal.pone.0141738

Niewczas, M. A., Mathew, A. V., Croall, S., Byun, J., Major, M., Sabisetti, V. S., et al. (2017). Circulating Modified Metabolites and a Risk of ESRD in Patients With Type 1 Diabetes and Chronic Kidney Disease. *Dia Care.* 40 (3), 383–390. doi:10.2337/dc16-0173

Niewczas, M. A., Sirich, T. L., Mathew, A. V., Skupien, J., Mohney, R. P., Warram, J. H., et al. (2014). Uremic Solutes and Risk of End-Stage Renal Disease in Type 2 Diabetes: Metabolomic Study. *Kidney Int.* 85 (5), 1214–1224. doi:10.1038/ki.2013.497

Niwa, T., Takeda, N., and Yoshizumi, H. (1998). RNA Metabolism in Uremic Patients: Accumulation of Modified Ribonucleosides in Uremic Serum. *Kidney Int.* 53, 1801–1806. doi:10.1046/j.1523-1755.1998.00944.x

Oh, T. R., Choi, H. S., Kim, C. S., Bae, E. H., Ma, S. K., Sung, S.-A., et al. (2019). Hyperuricemia Has Increased the Risk of Progression of Chronic Kidney Disease: Propensity Score Matching Analysis From the KNOW-CKD Study. *Sci. Rep.* 9 (1), 6681. doi:10.1038/s41598-019-43241-3

Orphanet (2020). Renal Dysplasia. Available at: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=EN&Expert=93108 (Accessed July 5, 2021).

Parfieniuk, E., Zbucka-Kretowska, M., Ciborowski, M., Kretowski, A., and Barbas, C. (2018). Untargeted Metabolomics: An Overview of its Usefulness and Future Potential in Prenatal Diagnosis. *Expert Rev. Proteomics.* 15, 809–816. doi:10.1080/14789450.2018.1526678

Pelletier, C. C., Croyal, M., Ene, L., Aguesse, A., Billon-Crossouard, S., Krempf, M., et al. (2019). Elevation of Trimethylamine-N-Oxide in Chronic Kidney Disease: Contribution of Decreased Glomerular Filtration Rate. *Toxins.* 11 (11), 635. doi:10.3390/toxins11110635

Pero, R. (2010). Health Consequences of Catabolic Synthesis of Hippuric Acid in Humans. *Curr. Clin. Pharmacol.* 5 (1), 67–73. doi:10.2174/157488410790410588

Pezzatti, J., González-Ruiz, V., Codesido, S., Gagnebin, Y., Joshi, A., Guillarme, D., et al. (2019). A Scoring Approach for Multi-Platform Acquisition in Metabolomics. *J. Chromatogr. A.* 1592, 47–54. doi:10.1016/j.chroma.2019.01.023

Phua, Y. L., and Ho, J. (2016). Renal Dysplasia in the Neonate. *Curr. Opin. Pediatr.* 28, 209–215. doi:10.1097/MOP.0000000000000324

Podrini, C., Cassina, L., and Boletta, A. (2020). Metabolic Reprogramming and the Role of Mitochondria in Polycystic Kidney Disease. *Cell Signal.* 67, 109495. doi:10.1016/j.cellsig.2019.109495

Qu, Q., Zeng, F., Liu, X., Wang, Q. J., and Deng, F. (2016). Fatty Acid Oxidation and Carnitine Palmitoyltransferase I: Emerging Therapeutic Targets in Cancer. *Cell Death Dis.* 7, e2226. doi:10.1038/cddis.2016.132

Rhee, E. P., Clish, C. B., Ghorbani, A., Larson, M. G., Elmariah, S., McCabe, E., et al. (2013). A Combined Epidemiologic and Metabolomic Approach Improves CKD Prediction. *J. Am. Soc. Nephrol.* 24 (8), 1330–1338. doi:10.1681/ASN.2012101006

Rhee, E. P., and Gerszten, R. E. (2012). Metabolomics and Cardiovascular Biomarker Discovery. *Clin. Chem.* 58, 139–147. doi:10.1373/clinchem.2011.169573

Ringoir, S., Schoots, A., and Vanholder, R. (1988). Uremic Toxins. *Kidney Int. Suppl.* 24 (Suppl. 24), S4–S9.

Schwahn, B. C., Hafner, D., Hohlfeld, T., Balkenhol, N., Laryea, M. D., and Wendel, U. (2003). Pharmacokinetics of Oral Betaine in Healthy Subjects and Patients With Homocystinuria. *Br. J. Clin. Pharmacol.* 55 (1), 6–13. doi:10.1046/j.1365-2125.2003.01717.x

Sekula, P., Goek, O.-N., Quaye, L., Barrios, C., Levey, A. S., Römisch-Margl, W., et al. (2016). A Metabolome-Wide Association Study of Kidney Function and Disease in the General Population. *J. Am. Soc. Nephrol.* 27 (4), 1175–1188. doi:10.1681/ASN.2014111099

Su, J., Wei, Y., Liu, M., Liu, T., Li, J., Ji, Y., et al. (2014). Anti-Hyperuricemic and Nephroprotective Effects of Rhizoma Dioscoreae Septemlobae Extracts and its Main Component Dioscin via Regulation of MOAT1, MURAT1 and MOCT2 in Hypertensive Mice. *Arch. Pharm. Res.* 37 (10), 1336–1344. doi:10.1007/s12272-014-0413-6

Tarasenko, T. N., Cusmano-Ozog, K., and McGuire, P. J. (2018). Tissue Acylcarnitine Status in a Mouse Model of Mitochondrial β-oxidation Deficiency during Metabolic Decompensation Due to Influenza Virus Infection. *Mol. Genet. Metab.* 125 (1–2), 144–152. doi:10.1016/j.ymgme.2018.06.012

Torremans, A., Marescau, B., Kränzlin, B., Gretz, N., Billiouw, J.-M., Vanholder, R., et al. (2006). Biochemical Validation of a Rat Model for Polycystic Kidney Disease: Comparison of Guanidino Compound Profile with the Human Condition. *Kidney Int.* 69 (11), 2003–2012. doi:10.1038/sj.ki.5000443

Tsalik, E. L., Willig, L. K., Rice, B. J., Van Velkinburgh, J. C., Mohney, R. P., McDunn, J. E., et al. (2015). Renal Systems Biology of Patients With Systemic Inflammatory Response Syndrome. *Kidney Int.* 88 (4), 804–814. doi:10.1038/ki.2015.150

Ueda, Y., Ando, A., Nagata, T., Yanagida, H., Yagi, K., Sugimoto, K., et al. (2004). A Boy With Mitochondrial Disease: Asymptomatic Proteinuria Without Neuromyopathy. *Pediatr. Nephrol.* 19 (1), 107–110. doi:10.1007/s00467-003-1318-7

Ussher, J. R., Elmariah, S., Gerszten, R. E., and Dyck, J. R. B. (2016). The Emerging Role of Metabolomics in the Diagnosis and Prognosis of Cardiovascular Disease. *J. Am. Coll. Cardiol.* 68, 2850–2870. doi:10.1016/j.jacc.2016.09.972

Verbitsky, M., Westland, R., Perez, A., Kiryluk, K., Liu, Q., Krithivasan, P., et al. (2019). The Copy Number Variation Landscape of Congenital Anomalies of the Kidney and Urinary Tract. *Nat. Genet.* 51 (1), 117–127. doi:10.1038/s41588-018-0281-y

Weber, S., Moriniere, V., Knüppel, T., Charbit, M., Dusek, J., Ghiggeri, G. M., et al. (2006). Prevalence of Mutations in Renal Developmental Genes in Children With Renal Hypodysplasia: Results of the ESCAPE Study. *J. Am. Soc. Nephrol.* 17 (10), 2864–2870. doi:10.1681/ASN.2006030277

Woolf, A. S., Price, K. L., Scambler, P. J., and Winyard, P. J. (2004). Evolving Concepts in Human Renal Dysplasia. *J. Am. Soc. Nephrol.* 15 (4), 998–1007. doi:10.1097/01.ASN.0000113778.06598.6F

Zerra, P. E., and Josephson, C. D. (2019). *Transfusion Medicine and Hemostasis*. Third Edition. Amsterdam, Netherlands: Elsevier Science. doi:10.1016/B978-0-12-813726-0.00047-7

Zhang, A. Q., Mitchell, S. C., and Smith, R. L. (1999). Dietary Precursors of Trimethylamine in Man: A Pilot Study. *Food Chem. Toxicol.* 37 (5), 515. doi:10.1016/S0278-6915(99)00028-9

Zhao, L., Dong, M., Liao, S., Du, Y., Zhou, Q., Zheng, H., et al. (2016). Identification of Key Metabolic Changes in Renal Interstitial Fibrosis Rats Using Metabonomics and Pharmacology. *Sci. Rep.* 6, 27194. doi:10.1038/srep27194

Zhao, Y.-Y., Lei, P., Chen, D.-Q., Feng, Y.-L., and Bai, X. (2013). Renal Metabolic Profiling of Early Renal Injury and Renoprotective Effects of Poria Cocos Epidermis Using UPLC Q-TOF/HSMS/MSE. *J. Pharm. Biomed. Anal.* 81-82, 202–209. doi:10.1016/j.jpba.2013.03.028

Zhou, D., and Liu, Y. (2016). Understanding the Mechanisms of Kidney Fibrosis. *Nat. Rev. Nephrol.* 12 (2), 68–70. doi:10.1038/nrneph.2015.215

Zoccali, C., and Mallamaci, F. (2018). Uric Acid in Chronic Kidney Disease: The Quest for Causality Continues. *Nephrol. Dial. Transplant.* 33, 193–195. doi:10.1093/ndt/gfx341

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 30 July 2021
doi: 10.3389/fmolb.2021.681955

Free Triiodothyronine Connected With Metabolic Changes in Patients With Coronary Artery Disease by Interacting With Other Functional Indicators

Xiao-xue Tian[1,2,3†], Shu-fen Zheng[1,2,3†], Ju-e Liu[2,3†], Yuan-yuan Wu[2,3], Lu Lin[2,3], Hong-mei Chen[2,3], Li-wen Li[2,3], Min Qin[2,3], Zi-xian Wang[2,3], Qian Zhu[2,3], Wei-hua Lai[2] and Shilong Zhong[1,2,3,4]**

[1]School of Pharmaceutical Sciences, Southern Medical University, Guangzhou, China, [2]Department of Pharmacy, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [3]Guangdong Provincial Key Laboratory of Coronary Heart Disease Prevention, Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [4]Department of Cardiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China

This study aims to evaluate the association between free triiodothyronine (FT3) and outcomes of coronary artery disease (CAD) patients, as well as to assess the predictive power of FT3 and related functional markers from the perspective of potential mechanism. A total of 5104 CAD patients with an average follow-up of three years were enrolled into our study. Multivariate Cox regression was used to evaluate the associations between FT3, FT4 (free thyroxin), FT3/FT4 and death, MACE. We developed and validated an age, biomarker, and clinical history (ABC) model based on FT3 indicators to predict the prognosis of patients with CAD. In the multivariable Cox proportional hazards model, FT3 and FT3/FT4 were independent predictors of mortality (Adjusted HR = 0.624, 95% CI = 0.486–0.801; adjusted HR = 0.011, 95% CI = 0.002–0.07, respectively). Meanwhile, emerging markers pre-brain natriuretic peptide, fibrinogen, and albumin levels are significantly associated with low FT3 ($p < 0.001$). The new risk death score based on biomarkers can be used to well predict the outcomes of CAD patients (C index of 0.764, 95% CI = 0.731–0.797). Overall, our findings suggest that low levels of FT3 and FT3/FT4 are independent predictors of death and MACE risk in CAD patients. Besides, the prognostic model based on FT3 provides a useful tool for the death risk stratification of CAD patients.

Keywords: free triiodothyronine, coronary artery disease, pre-brain natriuretic peptide, fibrinogen, metabolism, risk model, prognostic markers

Abbreviations: FT3, free triiodothyronine; FT4, free thyroxine; ALB, albumin; FIB, fibrinogen; proBNP, pro-B-type natriuretic peptide; HR, hazard ratio; 95%CI, 95% confidence interval.

OPEN ACCESS

Edited by:
Michal Jan Markuszewski, Medical University of Gdansk, Poland

Reviewed by:
Gaurav Sharma, University of Texas Southwestern Medical Center, United States
Naqiong Wu, Chinese Academy of Medical Sciences, China

****Correspondence:***
Shilong Zhong
zhongsl@hotmail.com
Wei-hua Lai
laiweihuax@163.com

[†]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: *17 March 2021*
Accepted: *12 July 2021*
Published: *30 July 2021*

Citation:
Tian X, Zheng S, Liu J, Wu Y, Lin L, Chen H, Li L, Qin M, Wang Z, Zhu Q, Lai W and Zhong S (2021) Free Triiodothyronine Connected With Metabolic Changes in Patients With Coronary Artery Disease by Interacting With Other Functional Indicators. Front. Mol. Biosci. 8:681955. doi: 10.3389/fmolb.2021.681955

INTRODUCTION

Coronary artery disease (CAD), also known as coronary atherosclerotic heart disease (CHD), is an inflammatory atherosclerotic disease that manifests as stable angina, unstable angina, myocardial infarction, and sudden cardiac death (álvarez-álvarez et al., 2017.; Pothineni et al., 2017; Musunuru and Kathiresan, 2019). Despite the prominent enrichment of treatments, patients with CAD have poor prognoses and high mortality (Li et al., 2015; Elmariah et al., 2018). The prognosis of CAD is affected by many factors, and the prognostic value of traditional risk factors for CAD is limited. A complex relationship exists between thyroid hormone (THs) levels and outcomes of CAD. THs play a central role in many cellular processes, including differentiation, growth, metabolism, and physiology (Boelaert; Gorman, 2006). Changes in TH concentrations in plasma, especially low triiodothyronine (T3) levels, represent hormonal imbalance, which is usual among patients suffering from an acute coronary event (Iervasi, 2008; Xu, 2012; Hao, 2014).

T3, the biologically active form of THs, is derived mainly from the peripheral transformation of precursor thyroxine (T4). Variations in the concentrations of T3 and T4 in plasma may exert a wide range of functions in several mechanisms, including heart dysfunction (Iervasi et al., 2003; Yazıcı et al., 2017). The levels of free triiodothyronine (FT3) and free thyroxin (FT4) are important indicators of the metabolic status of THs and have gained increasing attention as markers of many acute diseases (Iervasi et al., 2003; Zhang et al., 2016). The value of FT3 as an independent risk factor for the prognosis of patients with CAD remains controversial (Coceani et al., 2010; Wang et al., 2013; Lamprou et al., 2017; Wang et al., 2017). In addition, guidelines recommended for routine screening of low T3 syndrome in patients with CAD have not been pushed (Levine et al., 2016). Low T3 syndrome, known as nonthyroidal illness syndrome and euthyroid sick syndrome, is characterized by low serum levels of total T3 and FT3 with normal levels of thyroid stimulating hormone (TSH) and FT4; this condition is deemed as a strong prognostic determinant of chronic and systolic heart failure (HF) (Iervasi et al., 2003; Cooper and Biondi, 2012; Zhang et al., 2012; Jabbar et al., 2015; Yazıcı et al., 2017).

To the best of our knowledge, precise identification and discriminatory risk evaluation are important prerequisites to targeted treatment and prevention in high risk of all-cause death and major adverse cardiovascular (MACE) for CAD. Identifying novel and overlooked biomarkers, which not only guide the diagnosis and prognosis of patients with CAD but also detect new molecular mechanisms to elucidate the pathological progress of CAD, is important due to the limited predictive power of few predictors based on genetic factors and traditional clinical risk biomarkers available for risk stratification in patients with CAD. Hyperthyroidism is linked to an increased risk of thrombus (Kim et al., 2017). Data from epidemiological studies indicate that patients with low thyroid hormone levels are at higher risk of heart failure, and the prognosis of heart failure is also worse (Vale et al., 2019). A study suggested that a rise in thyroxine level is associated with the increase of FVIII, FIX, VWF and fibrinogen levels (Debeij et al., 2010). It has been reported that there is a positive correlation between FT3 level and cardiac ejection fraction, and a significant negative correlation with NT-proBNP. Although the imbalance of FT3 and FT4 levels is the most common manifestation of thyroid dysfunctions in patients with CAD, previous data on the prevalence of low FT3 and high FT4 levels in CAD patients are insufficient, and their prognostic roles are unclear. Thus, we evaluated the association of FT3 with other functional parameters and outcomes in 5104 CAD patients and assessed the predictive power of FT3 combined with other markers. The levels of thyroid hormone and related biomarkers are predictive of death in CAD, likely expressing different and synergistic pathogenic pathways, and their combined assay significantly improves the ability of risk stratification.

METHODS

Study Populations

During the prospective study period, 5,104 Chinese CAD participants from Guangdong Provincial People's Hospital, who underwent coronary angiography, were included. Patients were consecutively enrolled between January 2010 and December 2013 and followed up for all-cause death and MACE up to five years. The exclusion criteria for patients in single-center cohort study included the following: 1) age <18 years or >80 years, 2) renal insufficiency (defined as serum creatinine concentration > two times the upper limit of normal [230 μmol/L], history of renal transplantation or dialysis), 3) hepatic insufficiency (defined as serum transaminase concentration > two times the upper limit of normal [80 U/L], or a diagnosis of cirrhosis), 4) being pregnant or lactating, 5) advanced cancer or hemodialysis, 6) history of thyroid problems and use of antithyroid drugs or thyroid hormone medication, and 7) incomplete information about cardiovascular events during follow-up.

The primary endpoint of interest was all-cause death, followed by MACE. MACE is the occurrence of cardiac death, nonfatal myocardial infarctions, coronary revascularization, and cerebral infarction. All participants were followed up prospectively for the study endpoints by inpatient and outpatient hospital visits and telephone contacts with the patients or their families. At each follow-up assessment (every 6 months), the participants were questioned about new adverse cardiovascular events. From August 2010 to August 2018, all patients were followed up for the primary endpoint (all-cause death) and secondary endpoint (MACE), with a mean follow-up time of 3 years. Baseline information, including demographics, medical history, biochemical measurements, and medication, were obtained from the hospital information database. This study was approved by the Medical Ethical Review Committee of Guangdong Provincial People's Hospital and conducted according to the Declaration of Helsinki. Informed consent was obtained from all individual participants included in the study.

Study Design

Eligible participants were evaluated for demographic factors (e.g., age and gender), clinical characteristics (e.g., The Synergy

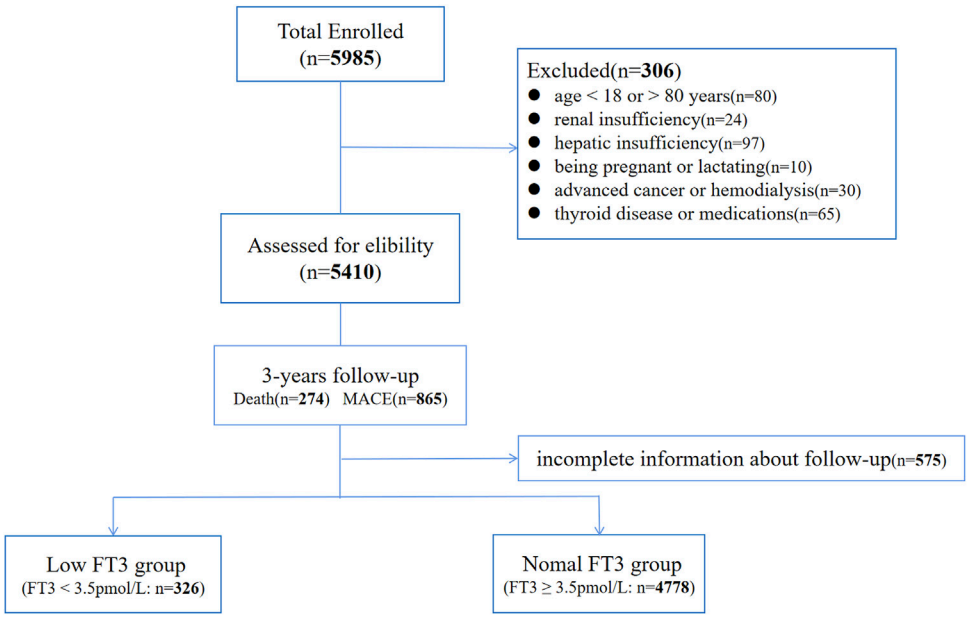

FIGURE 1 | Study design (Flowchart of the study population).

between PCI with TAXUS and Cardiac Surgery score, left ventricular ejection fraction [LVEF]), CAD risk factors (e.g., diabetes mellitus, smoking, hypertension, body mass index, and dyslipidemia), and current medication use. Baseline characteristics, including demographics, medical history, biochemical measurements, and medication, were obtained from the hospital information database. Follow-up data on mortality (time and cause of death) were used in the analysis as the primary outcome of interest. In some cases, these outcome data were unavailable. The serum concentrations of T3, FT3, T4, FT4, and of TSH have been determined and recorded. Normal range was as follows: T3, from 70 ng/dl to 170 ng/dl; FT3, from 3.28 to 6.47 pmol/L; T4, from 4.5 μg/dl to 12.0 μg/dl; FT4, from 7.64 to 16.03 pmol/L; TSH, from 0.49 to 4.91 μIU/ml, pro-brain-type natriuretic peptide (proBNP), from 0 to 125 pg/ml.

Statistical Analysis

Continuous variables were expressed as mean ± SD, and categorical variables are described using number and percentage. Student's t-test was used to compare continuously normal variables, the Wilcoxon rank-sum test was used for continuously abnormal variables, and the χ^2-test was used for categorical variables. Cox proportional hazards analysis was performed to investigate the associations of clinical parameters, levels of FT3, FT4 or FT3/FT4 ratio with outcomes. In the multivariable analysis. We implemented three models to evaluate the impact of potential confounders. Model 1 included age, sex. Model 2 further adjusted for a history of diabetes, heart failure, aspartate aminotransferase, apolipoprotein a, creatine, creatine kinase MB, high-density lipoprotein cholesterol, glucose, lipoprotein (a), medication history of calcium channel blockers and proton pump inhibitors. Model 3 incorporated the components of model 2, heart rate, systolic blood pressure and diastolic blood pressure. Variables with $p < 0.05$ in the univariate Cox regression analysis were included in the multivariate Cox proportional hazard regression models by an improved backward stepwise procedure to select covariates based on the Akaike information criterion (AIC). The Spearman rank correlation coefficient (Spearman ρ) was used to determine the associations among FT3, FT4, FT3/FT4, and other risk factors. In addition, the Harrell–Lee C-index was used to quantify the additional predictive value of the new model over the traditional model containing common clinical variables. The area under the curve (AUC) of time-dependent receiver-operating characteristic (ROC) analysis was used to quantify the predictive performance. In addition, each selected biomarker was assigned a corresponding score based on its value on the nomogram. Thus, the total score of death and MACE risk was calculated by nomogram, and the cutoff points of the total score detected by x-tile analysis could be used to divide CAD patients into different categories.

The final model was represented as a nomogram. Calibration curves (2000 bootstrap resamples) were generated to verify the accuracy of the nomogram. Decision curve analyses were performed to assess the clinical utility of the nomogram. Finally, we used X-tile software (Version 3.6.2, calculated by the "rms package" of R software) to calculate the optimal cut-off points for linear prediction and established a prognostic risk stratification. Kaplan-Meier method was used to estimate the survival probabilities and obtain survival curves, which are used to illustrate the cumulative incidence of clinical endpoints based on the cut-off values, and the log-rank (Mantel–Cox) test was adopted to compare survival curves. A two-tailed $p < 0.05$ was

TABLE 1 | Baseline characteristics.

Characteristics	Value N (%) or mean ± SD			
	Total (N = 5,104)	Low FT3 (N = 326)	Normal FT3(N = 4,778)	*p* Value
Demographic data				
Age	64.17 ± 10.61	69.79 ± 9.35	63.79 ± 10.57	<0.0001
Sex (male)	3,886 (76.3%)	224 (69.14%)	3,632 (76.84%)	0.0017
BMI, kg/m2	24.68 ± 4.99	23.84 ± 4.59	24.73 ± 5.02	0.0005
Smoke	977 (21.72%)	56 (22.13%)	912 (21.69%)	0.8695
Comorbidities				
Arrhythmia	481 (9.43%)	42 (12.88%)	435 (9.19%)	0.0272
Diabetes	1,425 (27.92%)	110 (33.74%)	1,300 (27.46%)	0.052
Heart failure	346 (6.78%)	46 (14.11%)	295 (6.23%)	<0.0001
Hypertension	3,053 (59.82%)	213 (65.34%)	2,811 (59.35%)	0.0356
Hyperlipidemia	623 (12.21%)	32 (9.82%)	589 (12.44%)	0.173
Baseline biochemical measurements				
ALT, U/L	31.31 ± 32.69	41.09 ± 66.79	30.67 ± 28.89	0.1013
AST, U/L	35.64 ± 55.71	59.37 ± 106.21	34.07 ± 50.29	<0.0001
CK, U/L	173.38 ± 435.84	296.58 ± 620.98	165.63 ± 420.73	0.0497
eGFR, ml/min/1.73 m^2	100.37 ± 365.33	66.62 ± 29.07	102.94 ± 379.86	<0.0001
CKMB, U/L	10.52 ± 18.4	13 ± 24.26	10.37 ± 18	0.5443
CHOL, mmol/L	4.48 ± 1.21	4.51 ± 1.51	4.48 ± 1.19	0.5021
LDLC, mmol/L	2.76 ± 0.97	2.73 ± 1.14	2.77 ± 0.96	0.0769
HDLC, mmol/L	0.98 ± 0.24	0.96 ± 0.27	0.99 ± 0.24	0.0566
TRIG, mmol/L	1.68 ± 1.31	1.61 ± 1.52	1.69 ± 1.29	0.0007
GLUC, mmol/L	6.79 ± 2.99	7.43 ± 3.22	6.74 ± 2.97	<0.0001
Lpa, mg/L	290.4 ± 325.83	353.84 ± 377.45	286.53 ± 321.31	0.0004
APOA, g/L	1.09 ± 0.26	1 ± 0.26	1.09 ± 0.26	<0.0001
TSH,mu/L	1.93 ± 3.47	3.72 ± 10.68	1.82 ± 2.22	0.5547
FT4,pmol/L	11.72 ± 2.86	11.85 ± 3.29	11.62 ± 2.22	0.0856
Medication				
β-blockers	3,746 (76.39%)	258 (79.63%)	3,457 (76.16%)	0.155
ACEIs	2,612 (53.32%)	195 (60.19%)	2,394 (52.8%)	0.0103
CCBs	1,458 (29.76%)	111 (34.26%)	1,334 (29.42%)	0.0669
PPIs	2,356 (48.09%)	206 (63.58%)	2,136 (47.11%)	<0.0001

SD = standard deviation; BMI = body mass index; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CK = creatine kinase; eGFR = estimated glomerular filtration rate; CKMB = creatine kinase MB; CHOL = cholesterol; LDLC = low-density lipoprotein cholesterol; HDLC = high-density lipoprotein cholesterol; TRIG = triglyceride; GLUC = glucose; Lpa = lipoprotein (a); APOA = apolipoprotein a; ACEIs = angiotensin converting enzyme inhibitors; CCBs = calcium channel blockers; PPIs = proton pump inhibitors.

statistically significant. All tests were performed using SAS (version 9.4) and R (version 3.6.2, http://www.R-project.org/).

RESULTS

Baseline Characteristics

A total of 5104 CAD patients were investigated, including 3,886 males (76.3%) and 1,218 females (23.7%), with an average age of 64.2 years at baseline, and a flowchart of the study population is shown in **Figure 1**. **Table 1** provides the baseline characteristics of the participants. During an average of 3 years of follow-up, 274 (5.61%) died and 865 (16.9%) had MACE events. The mean ± standard deviations of FT3, FT4, and FT3/FT4 were 4.4 ± 0.93 pmol/L, 11.73 ± 2.86 pmol/L, and 0.39 ± 0.1 pmol/L, respectively. FT3 < 3.5 pmol/L and FT4 > 17.8 pmol/L are divided into low FT3 group ($N = 326$) and high FT4 group ($N = 105$). In addition, patients with FT3/FT4 <0.33 were classified as the low FT3/FT4 group (N = 1,194) in accordance with the lower quartile lower limit. The proportion of patients with HF was higher in the low FT3 group (14.11 vs 6.23%, $p < 0.0001$), and no significant difference was found in the proportion of patients with smoking history. The common clinical indicators in the low FT3 group were higher than those in the normal FT3 group. Myocardial damage indexes of proBNP, hydroxybutyrate dehydrogenase (HBDH), cardiac troponin (cTnl), high-sensitivity C-reactive protein (hs-CRP), and high-sensitivity cardiac troponin (hs-cTnl) in the low FT3 group were higher than those in the normal FT3 group (all $p < 0.001$). However, the level of albumin (ALB), triglyceride (TRIG), total protein (TP), hemoglobin (HGB), and lymphocyte in the low FT3 group was lower than that in the normal FT3 group (**Supplementary Table S1**).

Associations Between Thyroid Hormones and Other Functional Parameters

A significant correlation exists between FT3/FT4 and ALB ($r = 0.381$, $p < 0.001$) and followed by proBNP ($r = -0.285$, $p < 0.001$). The results showed that ALB ($r = 0.348$), TP ($r = 0.268$), HCT ($r = 0.279$), and HGB ($r = 0.276$) were positively correlated with FT3 concentration ($p < 0.001$). Meanwhile, proBNP ($r = -0.194$), HBDH ($r = -0.156$), and DBIL ($r = -0.161$), and myocardial

TABLE 2 | Cox proportional hazards analysis for Death.

Characteristics	Univariate analysis		Multivariate analysis	
	HR (95%CI)	*p* value	HR (95%CI)	*p* value
Demographic data				
Age	1.055 (1.041–1.069)	3.33E-15	1.035 (1.02–1.051)	5.34E-06
Sex	1.098 (0.823–1.465)	0.5248		
Smoke	1.786 (1.209–2.639)	0.0036		
BMI	0.914 (0.864–0.966)	0.0015		
Comorbidities				
Arrhythmia	2.11 (1.555–2.862)	1.63E-06		
Diabetes	1.378 (1.155–1.643)	0.0004		
Heart failure	4.048 (3.071–5.334)	3.11E-23	2.484 (1.808–3.412)	1.94E-08
Hypertension	1.461 (1.133–1.882)	0.0034		
Hyperlipidemia	0.612 (0.388–0.965)	0.0345		
Medication				
β-blockers	0.965 (0.707–1.318)	0.8248		
ACEIs	1.088 (0.85–1.393)	0.5019		
CCBs	1.552 (1.212–1.985)	0.0005		
PPIs	1.291 (1.013–1.646)	0.0392		
Biochemical measurements				
eGFR	0.979 (0.974–0.984)	4.34E–15		
ALT	1.001 (0.999–1.004)	0.2539		
AST	1.002 (1–1.003)	0.0061		
APOA	0.366 (0.218–0.612)	0.0001	0.328 (0.18–0.597)	0.00026
CHOL	0.875 (0.786–0.973)	0.014		
CK	1 (1–1)	0.0385		
CKMB	1.004 (1–1.009)	0.0527		
GLUC	1.067 (1.039–1.096)	2.35E–06	1.068 (1.03–1.108)	0.0003
HDLC	0.484 (0.294–0.797)	0.0043		
LDLC	0.864 (0.758–0.985)	0.0283		
LPa	1.001 (1–1.001)	0.0004	1 (1–1)	0.018
TRIG	0.914 (0.813–1.028)	0.1343		
FT3	0.411 (0.336–0.503)	7.69E–18	0.577 (0.454–0.734)	7.17E-06
FT4	1.051 (1.032–1.069)	3.49E–08		
FT3/FT4	4.12E-04 (9.05E-05-0.00188)	7.45E–24	0.01 (0.001–0.064)	7.37E-07
TSH	0.998 (0.966–1.032)	0.9268		

injury indicators Troponin I, high-sensitivity troponin T were negatively correlated with FT3 concentration ($p < 0.001$) (**Table 2**), Similar patterns of correlation strength were also found in FT3/FT4 ratio. In addition, statistically significant associations were also found between coagulation markers and thyroid hormone levels.

Clinical Outcomes

After an average of three years of follow-up, 116 (35.58%) of the 326 patients had MACE events and 63 (20.93%) died in the low FT3 group. **Supplementary Tables S2–S3** provides univariable Cox modeling results for clinical indices and biomarkers in relation to mortality and MACE. As shown, age conferred increased risk for death (HR = 1.055, 95% CI = 1.041–1.069, and p = 3.33E-15). Biomarker data demonstrated a strong association between elevated FIB and risk for death (HR = 1.33, 95% CI = 1.23–1.43, and p = 8.06E-13). Low ALB levels may prompt increased death MACE risks (HR = 0.859, 95% CI = 0.837–0.881, and p = 3.98E-32; HR = 0.933, 95% CI = 0.918–0.947, and p = 6.98E-18, respectively). In the multivariable Cox regression analyses, FT3, FT4 and FT3/FT4 were all independent predictors of mortality after controlling for age and other clinical markers ($p < 0.05$) (**Table 3**). Notably, multivariate regression yielded similar results regarding relation between FT3, FT3/FT4 and MCACE risk, while after adjustment by clinical variables, no significant correlation was observed between FT4 and MACE risk (Adjusted HR = 1.008, 95% CI = 0.985–1.032). The cumulative survival curves of MACE and death in patients with different FT3, FT4, and FT3/FT4 levels are shown in **Figure 2**. Results demonstrate consistently the prognositic difference in different FT3, FT4 or FT3/FT4 ratio groups.

Development and Validation of the Predictive Model Based on FT3 and FT4

After univariate Cox regression analysis of 15 indicators significantly related to FT3 and FT4 (**Supplementary Table S3**), the indicators with $p < 0.05$ were screened out as the candidate variables for subsequent model constructions. A model including all candidate predictors were fitted. Then, the importance of each predictor variable in the model was measured by partial chi-square statistics minus degrees of freedom. Finally, the top three relevant candidate indicators were obtained. In the fitted predictive death model, proBNP (X^2-df = 139.77), ALB (X^2-df = 120.33), and HGB (X^2-df = 88.67) were significantly

TABLE 3 | Cox proportional hazards analysis for MACE.

Characteristics	Univariate analysis		Multivariate analysis	
	HR (95%CI)	*p* value	HR (95%CI)	*p* value
Demographic data				
Age	1.014 (1.007–1.021)	0.0002	1.008 (1–1.016)	0.032
Sex	1.218 (1.03–1.441)	0.0212	1.281 (1.039–1.581)	0.02
Smoke	1.092 (0.881–1.353)	0.4218		
BMI	0.982 (0.96–1.004)	0.1023		
Comorbidities				
Arrhythmia	1.296 (1.056–1.589)	0.0129		
Diabetes	1.345 (1.212–1.492)	2.54E-08	1.44 (1.218–1.704)	1.94E-05
Heart failure	1.934 (1.584–2.36)	8.60E-11	1.795 (1.463–2.201)	1.90E-08
Hypertension	1.274 (1.108–1.465)	0.0007		
Hyperlipidemia	1.051 (0.855–1.293)	0.637		
Medication				
β-blockers	0.907 (0.768–1.072)	0.2519		
ACEIs	1.025 (0.893–1.176)	0.73		
CCBs	1.432 (1.245–1.647)	5.27E-07	1.366 (1.152–1.62)	0.0003
PPIs	1.207 (1.053–1.384)	0.0068		
Biochemical measurements				
eGFR	0.998 (0.997–1)	0.0487		
ALT	1.001 (1–1.003)	0.1452		
AST	1.001 (1.001–1.002)	5.15E-05		
APOA	0.532 (0.404–0.699)	6.09E-06		
CHOL	1.032 (0.979–1.089)	0.2405		
CK	1 (1–1)	6.00E-04		
CKMB	1.004 (1.002–1.007)	4.00E-04		
GLUC	1.047 (1.028–1.066)	1.01E-06		
HDLC	0.514 (0.388–0.68)	3.17E-06		
LDLC	1.055 (0.987–1.128)	0.1144		
LPa	1 (1–1.001)	5.30E-06	1 (1–1)	0.0003
TRIG	0.991 (0.941–1.042)	0.7163		
FT3	0.79 (0.71–0.879)	1.42E-05	0.831 (0.735–0.939)	0.003
FT4	1.027 (1.011–1.044)	0.0009		
FT3/FT4	0.104 (0.0465–0.233)	3.54E-08	0.312 (0.121–0.8)	0.015
TSH	0.985 (0.959–1.011)	0.254		

correlated with death. However, proBNP (X^2-df = 61.56) and ALB (X^2-df = 60.97) were strongly independent predictors of MACE, and FIB (X^2-df = 37.56) had greater prognostic value than other candidate variables involved in the outcome of MACE (**Figure 3**).

Our aim is to develop and validate a new predictive model based on FT3, including age, biomarker, and clinical history (ABC). HF, which was significantly correlated with endpoint events after multivariate Cox correction based on AIC stepwise regression, was selected as the clinical variable. The final predictive model established by age and candidate markers was represented via a nomogram (**Figure 4**).

On the one hand, the final MACE predictive models including age, HF, FT3, FIB, and proBNP showed a C index of 0.621 (95% CI = 0.596–0.646), which was significantly higher than the traditional model consisting of common clinical variables (C index of 0.586, 95% CI = 0.554–0.618) (**Table 4**). On the other hand, the death predictive models of the single AUC of ALB, proBNP, and FT3/FT4 obtained 0.666, 0.708, and 0.634, respectively. The MACE predictive models of the single AUC of FIB, proBNP, and FT3 obtained 0.597, 0.594, and 0.55, respectively (**Figure 5**).

The final model was represented by a nomogram, which showed that low FT3/FT4 values; low ALB levels, high FIB levels, and high proBNP levels were associated with low survival rates. The calibration curves of the nomogram for the predicted five-year mortality and MACE risk demonstrated consistency with the prediction and observation of the primary cohort (**Figure 6**). Decision curve analysis showed that the use of nomogram to predict mortality adds more benefit than either the treat-all-patients scheme or the treat-none scheme (**Figure 7**). Finally, the prognostic model for death was divided into three groups according to the total score, as follows: low risk (<118), medium risk (118–124) and high risk (≥124). Similarly, the scores of low-, medium-, and high-risk of MACE groups were <113, 113–122, and ≥122, respectively. The cumulative survival curve of risk stratification is shown in **Figure 8**.

DISCUSSION

Our study demonstrated that the levels of FT3 and FT3/FT4 were independent predictors of death and MACE risk in patients with CAD on multivariate analyses including conventional CAD risk factors. We also proved that thyroid hormone indicators were significantly related to the emerging biomarkers proBNP, FIB,

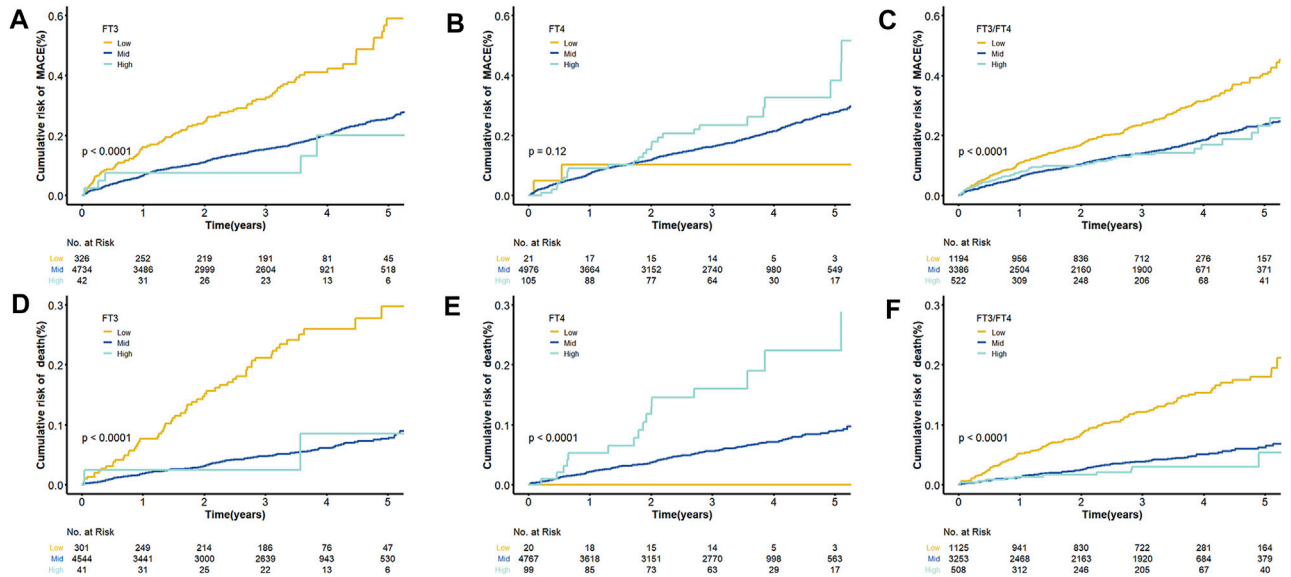

FIGURE 2 | Cumulative Kaplan–Meier Curves for FT3, FT4 and the ratio of FT3 and FT4. **(A)** Cumulative Kaplan–Meier Curves for MACE according to FT3 quartiles; **(B)** Cumulative Kaplan–Meier Curves for MACE according to FT4 quartiles; **(C)** Cumulative Kaplan–Meier Curves for MACE according to FT3/FT4 quartiles; **(D)** Cumulative Kaplan–Meier Curves for death according to FT3 quartiles; **(E)** Cumulative Kaplan–Meier Curves for death according to FT4 quartiles; **(F)** Cumulative Kaplan–Meier Curves for death according to FT3/FT4 quartiles.

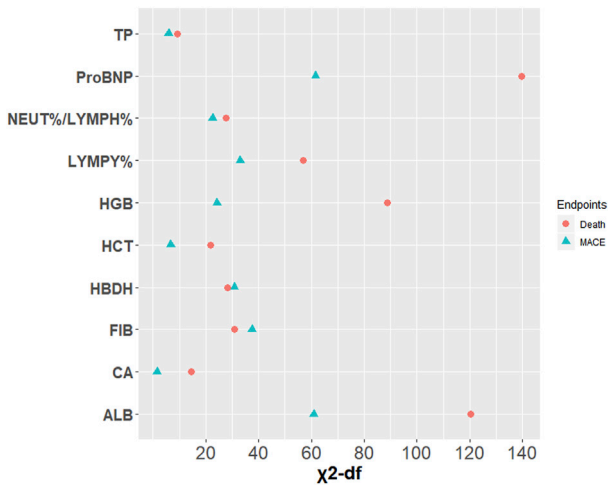

FIGURE 3 | Relative importance of each variable in the full model. Measured by partial Wald χ^2 minus the predictor degrees of freedom.

and ALB, which suggested that thyroid hormone levels may have an impact on adverse outcomes by the key biological pathways, such as heart function and injury, liver function, and coagulation. Finally, according to the predictive characteristics of FT3 and other cardiovascular biomarkers, we developed and validated a new, well-calibrated, biomarker-based risk score to assess the death risk of patients with CAD.

In our study, FT3 and FT3/FT4 are powerful prognostic markers. In recent years, the consequences on the cardiovascular system of milder forms of thyroid dysfunction have been increasingly recognized (Gomberg-Maitland and Frishman, 1998). The further clinical and experimental evidences suggest that a low FT3 level is a strong predictor of a poor prognosis in patients with chronic cardiovascular diseases. However, the relationship between low FT3 and adverse prognosis of CAD patients had a controversy in the last few years (Iervasi et al., 2003; Boelaert and Franklyn, 2005; Coceani et al., 2009; Iervasi, 2009; Ichiki, 2014). In this study, we found that more deaths and MACE occurred in the low FT3 group. Previous studies showed that FT3 promoted the progression of atherosclerosis probably due to the inhibition of lipoprotein enzyme activity and the decreasing clearance rate of total cholesterol (Chapidze et al., 2006; Asvold et al., 2007; Maria Moreno et al., 2019). Although FT3 is associated with lipid levels and linked with the increased mortality of CAD, the mechanism is worthy of further study due to the result of the interaction of many factors (Danese et al., 2000; Duntas, 2002; Erem, 2006).

On the basis of the relationship between FT3 and the prognosis of CAD, we also analyzed the correlation between FT3 and other clinical indicators of CAD patients to further understand whether or not the function was influenced by low FT3 with other common risk factors of CAD. Low T3 syndrome is associated with many traditional risk factors of CAD, such as proBNP, FIB, and ALB (Knezl et al., 2008; Chuang et al., 2014; Brozaitiene et al., 2016). Liang *et al.* reported that BNP was increased six-fold, and its promoter activity increased three to fivefold following T3 treatment (Liang et al., 2003). Afandi *et al.* reported that the decrease in thyroid hormone binding proteins is often a consequence of the acute phase response by impaired synthesis, rapid breakdown, and movement out of the plasma

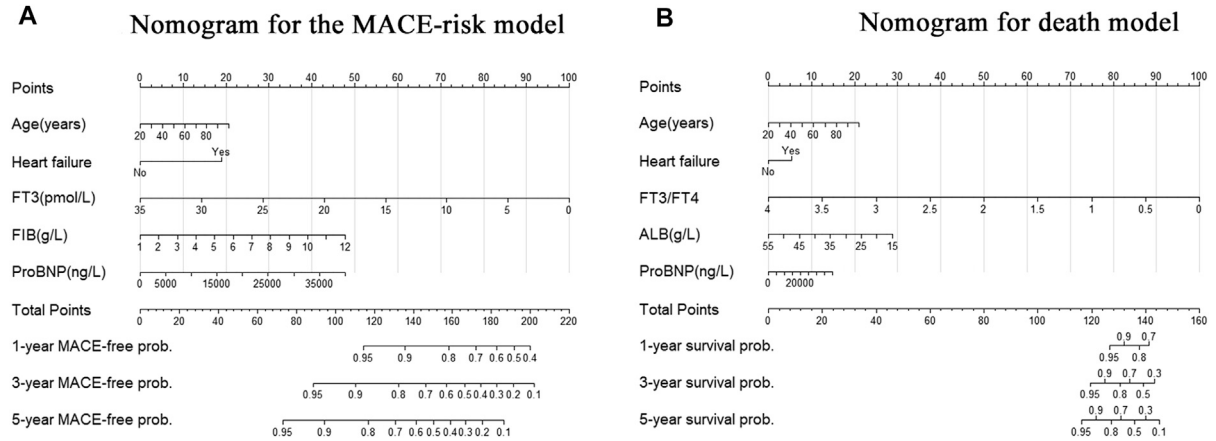

FIGURE 4 | Nomogram for the new biomarker-based risk score. **(A)** Nomogram for the final biomarker-based ABC-MACE risk score; **(B)** Nomogram for the final biomarker-based ABC-death risk score. Note: The scores of each variable were added to obtain the total score, and a vertical line was drawn on the total score to obtain the corresponding probability of outcomes.

TABLE 4 | C indices of models.

Model	C Index (95%CI)	
	Death	MACE
Model 1	0.693 (0.633–0.752)	0.586 (0.554–0.618)
Model 2	0.751 (0.715–0.787)	0.621 (0.596–0.646)
Model 3	0.764 (0.731–0.797)	0.614 (0.590–0.638)
Model 4	0.758 (0.724–0.793)	0.619 (0.596–0.643)

Model 1 of Death: Traditional model (age + sex + HyperT + DM + CHOL + HDLC + BMI + smoking + TRIG)
Model 2 of Death:age + HF + FT3/FT4+HGB + ProBNP
Model 3 of Death:age + HF + FT3/FT4+ALB + ProBNP
Model 4 of Death:age + HF + FT3/FT4+ALB + HGB
Model 1 of MACE: Traditional model (age + sex + HyperT + DM + CHOL + HDLC + BMI + smoking + TRIG)
Model 2 of MACE:age + HF + FT3+FIB + ProBNP
Model 3 of MACE:age + HF + FT3+ALB + ProBNP
Model 4 of MACE:age + HF + FT3+ALB + FIB

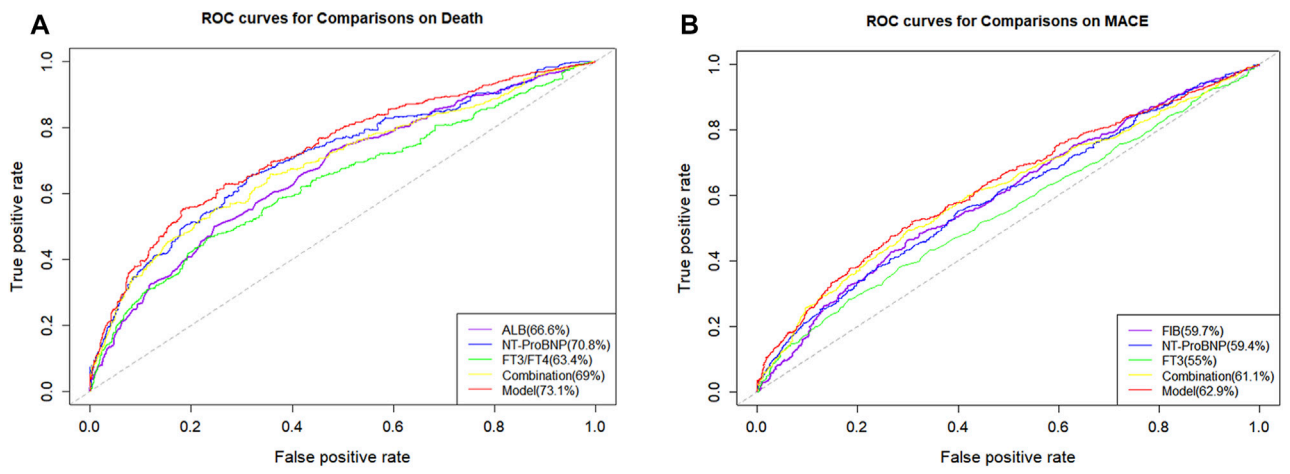

FIGURE 5 | Receiver operating characteristics curves (ROC) of different parameters and models. **(A)** ROC curves of different parameters and models for predicting all-cause mortality; **(B)** ROC curves of different parameters and models for predicting MACE.

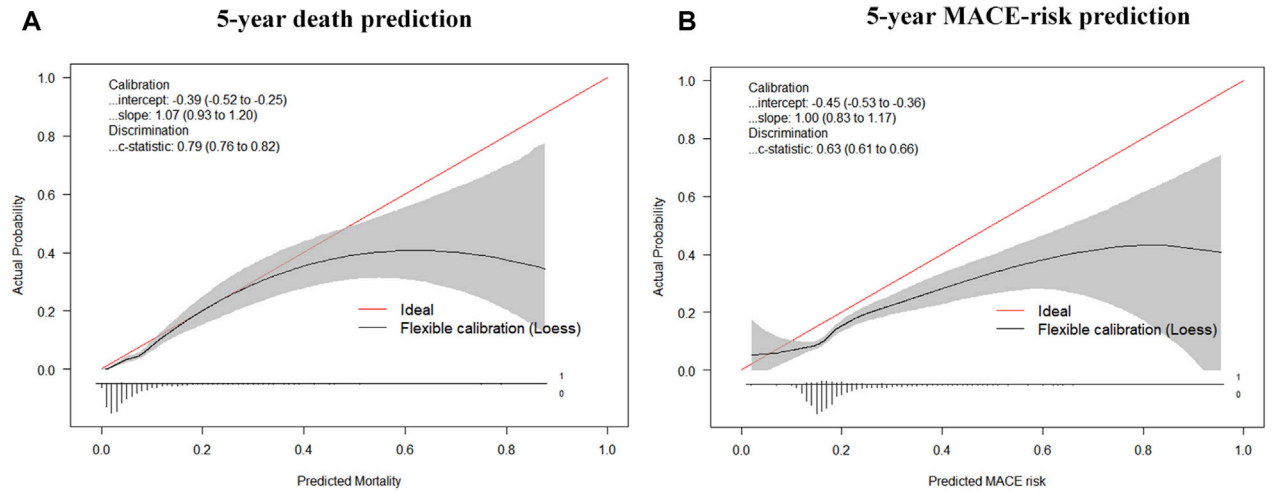

FIGURE 6 | Calibration curves for the nomogram model. **(A)** Calibration curves of the prognostic nomogram for 5-years overall survival; **(B)** Calibration curves of the nomogram predicting risk of MACE.

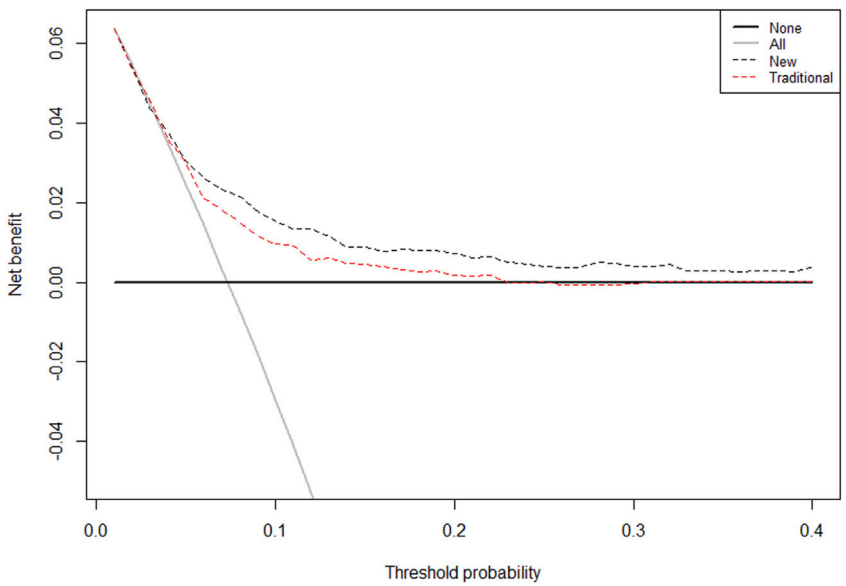

FIGURE 7 | The Decision Curves Analysis (DCA) curve of the prognostic nomogram for predicting probability of overall survival at 5 years. Note: The traditional group-composed by parameters of age, sex, hypertension, diabetes, CHOL, HDLC, BMI, smoking, TRIG; Novel group: age, HF, FT3/FT4, ALB, and proBNP. The horizontal axis represents the threshold value, which is the reference probability of whether a patient receives treatment, and the vertical axis represents the net benefit rate after the advantages minus the disadvantages. Under the same threshold probability, the larger net benefit implies that patients can obtain the maximum benefit using the diagnosis of this model. The closer the curve in the DCA graph is to the top, the higher the value of the model diagnosis will be.

space (Afandi et al., 2000). However, there are no detailed data on thyroid parameters in patients with acute liver failure (ALF) so far. One study showed that more than 50% of ALF patients exhibit abnormal thyroid parameters, and their prognosis was worse than that of normal thyroid patients (Anastasiou et al., 2015). In addition, a study suggested that hyperthyroidism was associated with an increased risk of thrombosis, and patients with hyperthyroidism show higher levels of fibrinogen (Kim et al., 2017). Another study confirmed that VWF and fibrinogen mediate up to 10% of the association between FT4 and cardiovascular disease (Bano et al., 2019). In our study, the level of coagulation parameters we observed changes significantly with the increase of FT4 level, and seems to be less affected by FT3 level. It can be concluded from the existing literature that patients with hypothyroidism are at increased risk of bleeding complications due to impaired coagulation and

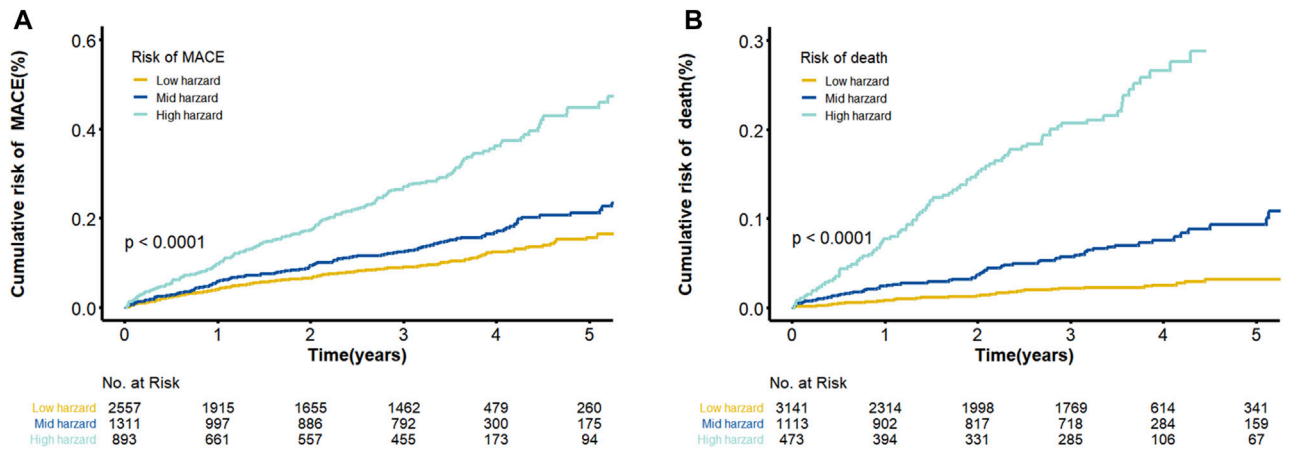

FIGURE 8 | Cumulative Kaplan–Meier curves of the prediction model. **(A)** Cumulative risk of MACE by predicted 5-years ABC-MACE risk group; **(B)** Cumulative probability of mortality by predicted 5-years ABC-death risk group. Note: Survival curves stratified by the score calculated by the nomogram scoring system [low risk, moderate risk and high risk].

fibrinolysis (Elbers et al., 2018). Notably, thyroid function with the pathophysiological mechanism of the heart and the prognosis of acute myocardial infarction is still being explored. A study demonstrated a significant relationship between the suppression of thyroid axis function, increased inflammation markers and increased NT-pro-BNP levels in CAD patients undergoing rehabilitation after ACS. They have also demonstrated that measures of FT4 and FT3/FT4 ratio together with NT-pro-BNP can be important prognostic markers of negative long-term outcomes (i.e., mortality) (Brozaitiene et al., 2016). A study aimed at explaining the relationship between T3 and cardiac function found that in the multiple regression analysis, FT3 levels were positively correlated with cardiac ejection fraction, and significantly negatively correlated with NT-proBNP(She et al., 2018). This is consistent with the correlation trend results of FT3 and proBNP obtained in this study. Overall, this study provides further insights into the relationship between thyroid hormones and liver function, thrombosis, and heart function. Whether the effect of thyroid hormones on cardiovascular events and cardiovascular death is mediated by coagulation factors or regulated by heart and liver functions requires further research in the future.

Identifying new biomarkers to effectively improve the prediction for prognosis and guide the individualized treatment of patients with CAD is critical (Malakar et al., 2019). As a quantitative tool for assessing risk and benefit, clinical predictive models provide risk estimates for the presence of the disease (diagnosis) or an event in the future course of disease (prognosis) for individual patients (Steyerberg and Vergouwe, 2014). We developed and validated a well-calibrated new risk score based on biomarkers to assess the risk of death in CAD patients. The novel ABC death score includes age, clinical history of HF, and biomarkers of proBNP, FT3/FT4, and ALB. The developed model including multiple and novel risk factors is useful for estimating prognosis in patients with CHD to inform treatment decisions or for use as a risk stratification tool in future research.

Although this novel ABC–CHD model has few variables, it provides reliable CV death prediction compared with models composed of traditional clinical factors. In our model, the most important prognostic variable is proBNP, which is a well-known marker of stress and dysfunction of cardiomyocytes (Beatty et al., 2015). Recent evidence suggests that proBNP also plays a role in metabolic pathways, including lipolysis and regulation of blood glucose levels, which are important in the pathophysiology of CAD (Zois et al., 2014). A study suggested that proBNP is independent of other prognostic markers, including systolic and diastolic dysfunction, left ventricular mass index, inducible ischemia, exercise capacity, C-reactive protein (CRP), cTnT, and New York Heart Association classification, which can predict cardiovascular morbidity and mortality (Bibbins-Domingo et al., 2007). Increasing concentration of proBNP may also be a signal of vascular dysfunction, in which the natriuretic peptides produce changes in vascular smooth muscle proliferation or contractility, in part via cyclic guanosine monophosphate cascades or nitric oxide synthesis (Ahluwalia et al., 2004).

In summary, these findings suggest that a simple test of biomarkers (i.e., proBNP and ALB) may stratify the risk of prognosis in patients with CAD and provide treatment strategies aimed at reducing future cardiovascular disease morbidity and mortality. The novel model can bring more incremental risk prediction (C index difference: 0.071) than the traditional predictive models of death. Furthermore, considering the powerful performance of the novel ABC-CHD model with a small number of variables provides a tool that may be easier to use in the clinical environment.

This study has several limitations. First, this study was based on a single center. Larger studies enrolling a more diverse population are needed to verify these findings and the external validity of prognostic models. Second, because measurements of

FT3, FT4, and TSH were performed only once during initial hospitalization, we were unable to account for possible variations in thyroid function over time. Third, more patients with low FT3 have a history of HF compared with normal patients. Thus, the outcomes of patients may be biased. Further mechanism studies about low FT3 on the risk of death should be considered in the interaction with proBNP, FIB, and ALB. This complex relation merits further well-designed investigations.

CONCLUSION

This study demonstrated that low levels of FT3 and FT3/FT4 are independent predictors of death and MACE risk in CAD patients. The associations of thyroid hormone with other functional parameters (i.e. heart function and injury, liver function, and coagulation markers) warrant further study. A novel risk score for the prediction of death including age, HF, and three biomarkers was successfully developed and validated. It can be widely used to complement clinical assessment and guide management based on death and MACE risk stratification in patients with CAD because it is based on a small number of readily available biomarkers and clinical factors.

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 author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethical Review Committee of Guangdong Provincial People's Hospital (No. GDREC2010137 and GDREC2017071H) and conducted according to the Declaration of Helsinki. The patients/participants provided their written informed consent (No. 20100910, 20170211) to participate in this study.

AUTHOR CONTRIBUTIONS

X-XT, S-FZ, and J-EL wrote manuscript. X-XT, S-FZ, and Y-YW designed research. X-XT, S-FZ, MQ, QZ, Z-XW, and J-EL performed research. S-FZ, X-XT, MQ, QZ, and J-EL analyzed data.

FUNDING

This study was funded by the National Nature Science Foundation of China (No. 81872934 (Principle investigators: Shilong Zhong), 81673514 (Principle investigators: Shilong Zhong); http://www.nsfc.gov.cn/), the National key research and development program (No. 2017YFC0909301 (Principle investigators: Shilong Zhong); https://service.most.gov. cn/index/), the Key research and development program of Guangdong Province, China (2019B020229003 (Principle investigators: Shilong Zhong); http://pro.gdstc.gd.gov.cn/egrantweb/), and the Guangdong Provincial People's Hospital Clinical Research Fund (Y012018085). No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.681955/full#supplementary-material.

REFERENCES

Afandi, B., Schussler, G. C., Arafeh, A.-H., Boutros, A., and Yap, M. G. (2000). Selective Consumption of Thyroxine-Binding Blobulin during Cardiac Bypass Surgery. *Metabolism-clinical Exp.* 49 (2), 270–274. doi:10.1016/s0026-0495(00)91611-1

Ahluwalia, A., MacAllister, R. J., and Hobbs, A. J. (2004). Vascular Actions of Natriuretic Peptides. Cyclic GMP-dependent and -independent Mechanisms. *Basic Res. Cardiol.* 99 (2), 83–89. doi:10.1007/s00395-004-0459-6

álvarez-álvarez, M. M., Zanetti, D., Carreras-Torres, R., Moral, P., and Athanasiadis, G. (2017). A Survey of Sub-saharan Gene Flow into the Mediterranean at Risk Loci for Coronary Artery Disease. *Eur. J. Hum. Genet. Ejhg* 25 (4), 472–476. doi:10.1038/ejhg.2016.200

Anastasiou, O., Sydor, S., Sowa, J. P., Manka, P., and Canbay, A. (2015). Higher Thyroid-Stimulating Hormone, Triiodothyronine and Thyroxine Values Are Associated with Better Outcome in Acute Liver Failure. *PLoS ONE* 10 (7), e0132189. doi:10.1371/journal.pone.0132189

Asvold, B. O., Vatten, L. J., Nilsen, T. I. L., and Bjoro, T. (2007). The Association between TSH within the Reference Range and Serum Lipid Concentrations in a Population-Based Study. The HUNT Study. *Eur. J. Endocrinol.* 156 (2), 181–186. doi:10.1530/eje.1.02333

Bano, A., Chaker, L., de Maat, M., Atiq, F., Kavousi, M., Franco, O. H., et al. (2019). Thyroid Function and Cardiovascular Disease: The Mediating Role of Coagulation Factors. *J. Clin. Endocrinol. Metab.* 104 (8), 3203–3212. doi:10.1210/jc.2019-00072

Beatty, A. L., Ku, I. A., Bibbins-Domingo, K., Christenson, R. H., DeFilippi, C. R., Ganz, P., et al. (2015). Traditional Risk Factors versus Biomarkers for Prediction of Secondary Events in Patients with Stable Coronary Heart Disease: From the Heart and Soul Study. *J. Am. Heart Assoc.* 4 (7). doi:10.1161/jaha.114.001646

Bibbins-Domingo, K., Gupta, R., Na, B., Wu, A. H., Schiller, N. B., and Whooley, M. A. (2007). N-terminal Fragment of the Prohormone Brain-type Natriuretic Peptide (NT-proBNP), Cardiovascular Events, and Mortality in Patients with Stable Coronary Heart Disease. *Jama* 297 (2), 169–176. doi:10.1001/jama.297.2.169

Boelaert, K., and Franklyn, J. A. (2005). Boelaert Thyroid Hormone in Health and Disease. *J. Endocrinol.* 187 (1), 1–15. doi:10.1677/joe.1.06131

Brozaitiene, J., Mickuviene, N., Podlipskyte, A., Burkauskas, J., and Bunevicius, R. (2016). Relationship and Prognostic Importance of Thyroid Hormone and N-Terminal Pro-B-type Natriuretic Peptide for Patients after Acute Coronary Syndromes: a Longitudinal Observational Study. *Bmc Cardiovasc. Disord.* 16 (1), 45. doi:10.1186/s12872-016-0226-2

Chapidze, G., Dolidze, N., Williams, M., Sharadze, E., and Latsabidze, N. (2006). Peculiarities of Lipid Profile Parameters in Cardiac Patients with Hypo- and Hyperthyroidism. *Georgian Med. News* 4 (133), 44–46.

Chuang, C. P., Jong, Y. S., Wu, C. Y., and Lo, H. M. (2014). Impact of Triiodothyronine and N-Terminal Pro-B-type Natriuretic Peptide on the Long-Term Survival of Critically Ill Patients with Acute Heart Failure. *Am. J. Cardiol.* 113 (5), 845–850. doi:10.1016/j.amjcard.2013.11.039

Coceani, M., Iervasi, G., Pingitore, A., Carpeggiani, C., and L'Abbate, A. (2009). Thyroid Hormone and Coronary Artery Disease: from Clinical Correlations to Prognostic Implications. *Clin. Cardiol.* 32 (7), 380–385. doi:10.1002/clc.20574

Coceani, M., Iervasi, G., Pingitore, A., Carpeggiani, C., and L'Abbate, A. (2010). Thyroid Hormone and Coronary Artery Disease: From Clinical Correlations to Prognostic Implications. *Clin. Cardiol.* 32 (7), 380–385. doi:10.1002/clc.20574

Cooper, D. S., and Biondi, B. (2012). Subclinical Thyroid Disease. *Lancet* 379 (9821), 1142–1154. doi:10.1016/s0140-6736(11)60276-6

Danese, M. D., Ladenson, Paul, W., Meinert, C. L., and Powe, N. R. (2000). Clinical Review 115: Effect of Thyroxine Therapy on Serum Lipoproteins in Patients with Mild Thyroid Failure: A Quantitative Review of the Literature. *J. Clin. Endocrinol. Metab.* 85 (9), 2993–3001. doi:10.1210/jcem.85.9.6841

Debeij, J., Cannegieter, S. C., Zaane, B. V., Smit, J. W. A., Corssmit, E. P. M., Rosendaal, F. R., et al. (2010). The Effect of Changes in Thyroxine and Thyroid-Stimulating Hormone Levels on the Coagulation System. *J. Thromb. Haemost. Jth* 8 (12), 2823–2826. doi:10.1111/j.1538-7836.2010.04054.x

Duntas, L. H. (2002). Thyroid Disease and Lipids. *Thyroid* 12 (4), 287–293. doi:10.1089/10507250252949405

Elbers, L. P. B., Fliers, E., and Cannegieter, S. C. (2018). The Influence of Thyroid Function on the Coagulation System and its Clinical Consequences. *J. Thromb. Haemost* 16 (4), 634–645. doi:10.1111/jth.13970

Elmariah, S., Farrell, L. A., Furman, D., Lindman, B. R., Shi, X., Morningstar, J. E., et al. (2018). Association of Acylcarnitines with Left Ventricular Remodeling in Patients with Severe Aortic Stenosis Undergoing Transcatheter Aortic Valve Replacement. *Jama Cardiol.* 3 (3), 242–246. doi:10.1001/jamacardio.2017.4873

Erem, C. (2006). Blood Coagulation, Fibrinolytic Activity and Lipid Profile in Subclinical Thyroid Disease: Subclinical Hyperthyroidism Increases Plasma Factor X Activity. *Clin. Endocrinol.* 64 (3), 323–329. doi:10.1111/j.1365-2265.2006.02464.x

Gomberg-Maitland, M., and Frishman, W. H. (1998). Frishman Thyroid Hormone and Cardiovascular Disease. *Am. Heart J.* 135–18796. doi:10.1016/s0002-8703(98)70081-x

Gorman, C. A.. (2006). Euthyroid Sick Syndrome: An Overview. *Thyroid* 7 (1), 125–132. doi:10.1089/thy.1997.7.125

Hao, P.. (2014). The Effect of Low FT3 Levels on Coronary Artery Calcification and MACE in Outpatients with Suspected Coronary Artery Disease. *Coron. Artery Dis.* 25 (5), 427–432. doi:10.1097/MCA.0000000000000095

Ichiki, T. (2014). Thyroid Hormone and Atherosclerosis. *Vasc. Pharmacol.* 52 (3-4), 151–156. doi:10.1016/j.vph.2009.09.004

Iervasi, G. (2008). Acute Effects of Triiodothyronine (T3) Replacement Therapy in Patients with Chronic Heart Failure and Low-T3 Syndrome: A Randomized, Placebo-Controlled Study. *J. Clin. Endocrinol. Metab.* 93 (4), 1351. doi:10.1210/jc.2007-2210

Iervasi, G., Pingitore, A., Landi, P., Raciti, M., Ripoli, A., Scarlattini, M., et al. (2003). Low-T3 Syndrome: a strong Prognostic Predictor of Death in Patients with Heart Disease. *Circulation* 107 (5), 708–713. doi:10.1161/01.cir.0000048124.64204.3f

Iervasi, G. (2009). Thyroid Hormones and the Cardiovascular System: Pathophysiology and Interventions. *Biomed. Pharmacother.* 63 (10), 0–753. doi:10.1016/j.biopha.2009.08.003

Jabbar, A., Ingoe, L., Pearce, S., Zaman, A., and Razvi, S. (2015). Thyroxine in Acute Myocardial Infarction (ThyrAMI) - Levothyroxine in Subclinical Hypothyroidism post-acute Myocardial Infarction: Study Protocol for a Randomised Controlled Trial. *Trials* 16 (1), 115. doi:10.1186/s13063-015-0621-5

Kim, N., Gu, J. Y., Yoo, H. J., Han, S. E., Kim, Y. I., Nam-Goong, I. S., et al. (2017). Contact System Activation and High Thrombin Generation in Hyperthyroidism. *Eur. J. Endocrinol.* 176, 583–589. doi:10.1530/EJE-16-0835

Knezl, V., Soukup, T., Okruhlicová, L., Slezák, J., and Tribulová, N. (2008). Thyroid Hormones Modulate Occurrence and Termination of Ventricular Fibrillation by Both Long-Term and Acute Actions. *Physiol. Res.* 57 (Suppl. 22), S91–S96. doi:10.33549/physiolres.931557

Lamprou, V., Varvarousis, D., Polytarchou, K., Varvarousi, G., and Xanthos, T. (2017). The Role of Thyroid Hormones in Acute Coronary Syndromes: Prognostic Value of Alterations in Thyroid Hormones. *Clin. Cardiol.* 40 (8), 528–533. doi:10.1002/clc.22689

Levine, G. N., Bates, E. R., Bittl, J. A., Brindis, R. G., Fihn, S. D., Fleisher, L. A., et al. (2016). ACC/AHA Guideline Focused Update on Duration of Dual Antiplatelet Therapy in Patients with Coronary Artery Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J. Am. Coll. Cardiol.* 68 (10), 1082–1115. doi:10.1016/j.jacc.2016.03.513

Li, J., Li, X., Wang, Q., Hu, S., Wang, Y., Masoudi, F. A., et al. (2015). ST-segment Elevation Myocardial Infarction in China from 2001 to 2011 (The China PEACE-Retrospective Acute Myocardial Infarction Study): a Retrospective Analysis of Hospital Data. *Lancet.* 385 (9966), 441–451. doi:10.1016/S0140-6736(14)60921-1

Liang, F., Webb, P., Marimuthu, A., Zhang, S., and Gardner, D. G. (2003). Triiodothyronine Increases Brain Natriuretic Peptide (BNP) Gene Transcription and Amplifies Endothelin-dependent BNP Gene Transcription and Hypertrophy in Neonatal Rat Ventricular Myocytes. *J. Biol. Chem.* 278 (17), 15073–15083. doi:10.1074/jbc.m207593200

Malakar, A. K., Choudhury, D., Halder, B., Paul, P., Uddin, A., and Chakraborty, S. (2019). A Review on Coronary Artery Disease, its Risk Factors, and Therapeutics. *J. Cell Physiol.* 234 (10), 16812–16823. doi:10.1002/jcp.28350

Maria Moreno, P. d. L., Lombardi, A., Elena Silvestri, and Lanni, A. (2019). Fernando Goglia Metabolic Effects of Thyroid Hormone Derivatives. *Thyroid* 18 (2), 239–253. doi:10.1089/thy.2007.0248

Musunuru, K., and Kathiresan, S. (2019). Genetics of Common, Complex Coronary Artery Disease. *Cell* 177 (1), 132–145. doi:10.1016/j.cell.2019.02.015

Pothineni, N. V. K., Swathi, S., Kevin, K., Shirazi, L. F., Francesco, R., Shah, P. K., et al. (2017). Infections, Atherosclerosis, and Coronary Heart Disease. *Eur. Heart J.* 43, 43. doi:10.1093/eurheartj/ehx362

She, J., Feng, J., Deng, Y., Sun, L., Wu, Y., Guo, M., et al. (2018). Correlation of Triiodothyronine Level with In-Hospital Cardiac Function and Long-Term Prognosis in Patients with Acute Myocardial Infarction. *Dis. markers* 2018, 1–8. doi:10.1155/2018/5236267

Steyerberg, E. W., and Vergouwe, Y. (2014). Towards Better Clinical Prediction Models: Seven Steps for Development and an ABCD for Validation. *Eur. Heart J.* 35 (29), 1925–1931. doi:10.1093/eurheartj/ehu207

Vale, C., Neves, J. S., Von Hafe, M., Borges-Canha, M., and Leite-Moreira, A. (2019). The Role of Thyroid Hormones in Heart Failure. *Cardiovasc. Drugs Ther.* 33, 179–188. doi:10.1007/s10557-019-06870-4

Wang, B., Liu, S., Li, L., Yao, Q., Song, R., Shao, X., et al. (2017). Non-thyroidal Illness Syndrome in Patients with Cardiovascular Diseases: a Systematic Review and Meta-Analysis. *Int. J. Cardiol.* 226, 1–10. doi:10.1016/j.ijcard.2016.10.039

Wang, J., Ma, X., Qu, S., Li, Y., and Xu, J. (2013). High Prevalence of Subclinical Thyroid Dysfunction and the Relationship between Thyrotropin Levels and Cardiovascular Risk Factors in Residents of the Coastal Area of China. *Exp. Clin. Cardiol.* 18 (1), 16–20. doi:10.1161/CIRCOUTCOMES.111.000039

Xu, Y. (2012). A Low fT3 Level as a Prognostic Marker in Patients with Acute Myocardial Infarctions. *Jpn. J. Med.* 51 (21), 3009–3015. doi:10.2169/internalmedicine.51.7902

Yazıcı, S., Kırış, T., Ceylan, U. S., Terzi, S., Erdem, A., Atasoy, I., et al. (2017). Yeşilçimen Relation of Low T3 to One-Year Mortality in Non–ST-elevation Acute Coronary Syndrome Patients. *J. Clin. Lab. Anal.* 31, e22036. doi:10.1002/jcla.22036

Zhang, B., Peng, W., Wang, C., Li, W., and Xu, Y. (2012). A Low fT3 Level as a Prognostic Marker in Patients with Acute Myocardial Infarctions. *Jpn. J. Med.* 51 (21), 3009–3015. doi:10.2169/internalmedicine.51.7902

Zhang, M., Sara, J. D., Yasushi, M., Hossein, G., Bell, M. R., Rajiv, G., et al. (2016). Clinical Outcomes of Patients with Hypothyroidism Undergoing Percutaneous Coronary Intervention. *Eur. Heart J.* 37 (26), 2055–2065. doi:10.1093/eurheartj/ehv737

Zois, N. E., Bartels, E. D., Hunter, I., Kousholt, B. S., Olsen, L. H., and Goetze, J. P. (2014). Natriuretic Peptides in Cardiometabolic Regulation and Disease. *Nat. Rev. Cardiol.* 11 (7), 403–412. doi:10.1038/nrcardio.2014.64

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 03 August 2021
doi: 10.3389/fmolb.2021.676349

A Cross-Platform Metabolomics Comparison Identifies Serum Metabolite Signatures of Liver Fibrosis Progression in Chronic Hepatitis C Patients

*Meera Shanmuganathan[1], Mohammad Omair Sarfaraz[2], Zachary Kroezen[1], Holly Philbrick[1], Richel Poon[1], Andrew Don-Wauchope[2], Marco Puglia[3], David Wishart[4] and Philip Britz-McKibbin[1]**

[1]Department of Chemistry and Chemical Biology, McMaster University, Hamilton, ON, Canada, [2]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada, [3]Department of Medicine, Division of Gastroenterology, McMaster University, Hamilton, ON, Canada, [4]Departments of Biological Sciences and Computing Science, University of Alberta, Edmonton, AB, Canada

OPEN ACCESS

Edited by:
Wolfram Weckwerth, University of Vienna, Austria

Reviewed by:
Fumio Matsuda, Osaka University, Japan
Sofia Moco, Vrije Universiteit Amsterdam, Netherlands

***Correspondence:**
Philip Britz-McKibbin
britz@mcmaster.ca

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

Received: *05 March 2021*
Accepted: *19 July 2021*
Published: *03 August 2021*

Citation:
Shanmuganathan M, Sarfaraz MO, Kroezen Z, Philbrick H, Poon R, Don-Wauchope A, Puglia M, Wishart D and Britz-McKibbin P (2021) A Cross-Platform Metabolomics Comparison Identifies Serum Metabolite Signatures of Liver Fibrosis Progression in Chronic Hepatitis C Patients. Front. Mol. Biosci. 8:676349. doi: 10.3389/fmolb.2021.676349

Metabolomics offers new insights into disease mechanisms that is enhanced when adopting orthogonal instrumental platforms to expand metabolome coverage, while also reducing false discoveries by independent replication. Herein, we report the first inter-method comparison when using multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) and nuclear magnetic resonance (NMR) spectroscopy for characterizing the serum metabolome of patients with liver fibrosis in chronic hepatitis C virus (HCV) infection ($n = 20$) and non-HCV controls ($n = 14$). In this study, 60 and 30 serum metabolites were detected frequently (>75%) with good technical precision (median CV < 10%) from serum filtrate samples ($n = 34$) when using standardized protocols for MSI-CE-MS and NMR, respectively. Also, 20 serum metabolite concentrations were consistently measured by both methods over a 500-fold concentration range with an overall mean bias of 9.5% ($n = 660$). Multivariate and univariate statistical analyses independently confirmed that serum choline and histidine were consistently elevated ($p < 0.05$) in HCV patients with late-stage (F2-F4) as compared to early-stage (F0-F1) liver fibrosis. Overall, the ratio of serum choline to uric acid provided optimal differentiation of liver disease severity ($AUC = 0.848$, $p = 0.00766$) using a receiver operating characteristic curve, which was positively correlated with liver stiffness measurements by ultrasound imaging ($r = 0.606$, $p = 0.0047$). Moreover, serum 5-oxo-proline concentrations were higher in HCV patients as compared to non-HCV controls ($F = 4.29$, $p = 0.0240$) after adjustment

Abbreviations: m/z:RMT, accurate mass:relative migration time; BGE, background electrolyte; Cl-Tyr, 4-chlorotyrosine; CV, coefficient of variance; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; F-Phe, 3-flurorophenylalanine; HCV, hepatitis C virus; IS, internal standard; LC, liquid chromatography; MS, mass spectrometry; MAGMET, magnetic resonance for metabolomics; METAVIR, meta-analysis of histological data in viral hepatitis; MSI-CE-MS, multisegment injection-capillary electrophoresis-mass spectrometry; NMR, nuclear magnetic resonance; NMS, naphthalene monosulfonic acid; PCA, principal component analysis; PLS-DA, partial least squares-discriminant analysis; ROC, receiver operating characteristic curve.

for covariates (age, sex, BMI), indicative of elevated oxidative stress from glutathione depletion with the onset and progression of liver fibrosis. Both instrumental techniques enable rapid yet reliable quantification of serum metabolites in large-scale metabolomic studies with good overlap for biomarker replication. Advantages of MSI-CE-MS include greater metabolome coverage, lower operating costs, and smaller sample volume requirements, whereas NMR offers a robust platform supported by automated spectral and data processing software.

Keywords: metabolomics (OMICS), capillary electrophoresis-mass spectrometry, nuclear magnetic resonance, serum, biomarkers, liver fibrosis, hepatitis C virus infection

INTRODUCTION

High-field nuclear magnetic resonance (NMR) spectroscopy and high-resolution mass spectrometry (MS) are two core instrumental platforms used in discovery-based metabolomics research. Selection of one or more analytical method(s) is dependent on several factors, including infrastructure/operating costs, sample volume/workup requirements, sample throughput, as well as selectivity and sensitivity that impact overall metabolome coverage. In general, NMR offers excellent reproducibility and quantitative performance together with unambiguous metabolite identification, and thus is well suited for longitudinal metabolomic studies, metabolic flux analysis, and non-invasive analysis of tissue specimens (Emwas et al., 2019). However, concentration sensitivity and resolution are limited without isotopic enrichment when using fast spectral acquisition protocols in one-dimensional proton (^{1}H)-NMR. These constraints result in quantification of typically a few dozen polar serum metabolites depending on magnet field strength, probe design, and spectral processing strategies (Wishart, 2019; Casadei-Gardini, et al., 2020). NMR metabolite coverage can be further expanded to include various lipoprotein, cholesterol and fatty acid species when ultrafiltration is avoided (Würtz et al., 2017). In contrast, MS-based metabolomic methods are more accessible when using bench-top instrumentation with greater sensitivity and resolution especially when coupled to one or more high efficiency separation techniques (Kuehnbaum and Britz-McKibbin, 2013). For instance, liquid chromatography (LC)-MS using separation mechanisms (e.g., reversed-phase, hydrophilic interaction) provide exceptional metabolome coverage (Rhee et al., 2019) especially when using chemical isotope labeling methods (Han and Li, 2018). Yet, separations in LC-MS are generally constrained by lower throughput and complicated data pre-processing when performing non-targeted metabolite profiling (i.e., time alignment, peak picking), where a major fraction of molecular features comprise compounds with unknown chemical structures (da Silva et al., 2015). Alternatively, multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) offers a multiplexed separation platform for metabolomics (Kuehnbaum et al., 2013) with higher sample throughput, improved quality control, and lower sample volume requirements (Nori de Macedo et al., 2017) Furthermore, customized serial injection configurations accelerate biomarker discovery using novel data workflows to encode mass spectral information temporally within a separation (DiBattista et al., 2017) while providing robust inter-batch adjustment in large-scale metabolomic studies (Shanmuganathan et al., 2021). Although there have been several cross-platform metabolomic analysis involving NMR and LC-MS (Nevedomskaya et al., 2011; Psychogios et al., 2011; Bruno et al., 2018; Bhatia et al., 2019), to the best of our knowledge no study to date has explored the potential benefits of combining CE-MS with NMR methodologies in metabolomics (Marshall and Powers, 2017).

Chronic hepatitis C virus (HCV) infection can lead to progressive liver disease with a high risk for death from cirrhosis and hepatocellular carcinoma if not treated early with pangenotypic direct-acting antiviral regimens (Ghany and Morgan. 2020). Most individuals (~85%) infected with HCV develop chronic infections, which contribute to a high burden of liver-related disease complications and spiraling healthcare costs (Myers et al., 2014). Similar to other forms of chronic liver disease, HCV infections are accompanied by liver fibrosis, a scarring process characterized by thickening of liver tissue and accumulation of extracellular matrix proteins with eventual loss of liver function (Khatun and Ray, 2019). Optimal patient care and treatment decisions are dependent on staging of disease progression, which has relied on a liver biopsy to assess the severity of fibrosis, such as the widely used Meta-analysis of Histological Data in Viral Hepatitis (METAVIR) scoring system (Goodman, 2007). However, liver biopsy is an expensive and invasive procedure with risks for patient bleeding, including other complications. It is also prone to both sampling and inter-subject variability depending on quality of biopsy, and thus has been increasingly supplanted by less invasive methods for liver fibrosis assessment, such as ultrasound imaging and blood-based liver protein panels. As a result, there is growing interest in metabolomics to identify novel biomarkers of hepatic fibrosis that offer greater specificity, sensitivity, and reproducibility, and accessibility in a clinical setting (Chang and Yang, 2019). This is urgently needed to augment diagnostic applications for chronic liver disease differentiation (Soga et al., 2011), monitoring treatment responses to therapy (Meoni et al., 2019), and risk assessment of advanced stages of liver disease (Diren and Idle, 2020), including patients co-infected with HCV/HIV (Naggie et al., 2020).

In this work, metabolomic analyses were performed on serum filtrate samples collected from HCV patients at different stages of liver fibrosis, as well as non-HCV controls when using MSI-CE-MS

TABLE 1 | Study characteristics of treatment-naïve hepatitis C virus (HCV) infected patients (n = 20) at different stages of liver fibrosis, and healthy non-HCV infected participants (n = 14).

Criteria	Non-HCV (n = 14)	HCV early-stage (n = 9)	HCV late-stage (n = 11)	p-value[a]
Age (years)	44 ± 15	57 ± 10	63 ± 11	**0.0318**; 0.216
Sex (male)	10 (77%)	8 (89%)	9 (82%)	—
BMI (kg/m^2)	24.4 ± 1.3	26.0 ± 5.8	26.5 ± 5.0	0.334; 0.843
FibroScan test (kPa)[b]	4.81 ± 0.55	5.39 ± 1.1	11.2 ± 3.6	0.111; **1.73 × 10^{-4}**
FibroTest score[c]	0.22 ± 0.17	0.63 ± 0.23	0.69 ± 0.23	**1.61 × 10^{-4}**; 0.592
γ-Glutamyltransferase (U/L)	29 ± 26	132 ± 125	76 ± 77	**8.69 × 10^{-3}**; 0.233
Total bilirubin (μM)	11.0 ± 6.3	14.8 ± 7.5	11.8 ± 6.8	0.216; 0.369
Alpha-2-macroglobulin (g/L)	1.82 ± 0.56	3.4 ± 1.2	4.4 ± 1.2	**4.08 × 10^{-4}**; 0.0747
Hapatoglobin (g/L)	1.03 ± 0.39	1.09 ± 0.51	1.40 ± 0.75	0.780; 0.337
Apolipoprotein A1 (g/L)	1.36 ± 0.20	1.44 ± 0.11	1.40 ± 0.30	0.311; 0.720
Alanine aminotransferase (U/L)	24.1 ± 8.4	82 ± 55	52 ± 36	**1.16 × 10^{-3}**; 0.151
Fibrosis grade/METAVIR score[d]				
F0	14	5	—	—
F1	—	4	—	—
F2	—	—	5	—
F3	—	—	2	—
F4	—	—	4	—
HCV genotype				
1A/1B	—	9/0	9/2	—

[a]*Student's t-test to assess statistical significance (**p < 0.05**) when comparing healthy non-HCV controls with early-stage HCV, as well as early-stage HCV with late-stage HCV patients, respectively.*

[b]*FibroScan test results to assess liver fibrosis based on transient elastography using ultrasound imaging.*

[c]*FibroTest uses an algorithm that combines 5 standard serum protein biomarkers, including γ-glutamyltransferase, total bilirubin, alpha-2-macroglobulin, haptoglobin, and apolipoprotein A1.*

[d]*METAVIR score to assess the extent of inflammation and fibrosis by histopathological examination of a liver biopsy.*

and ^{1}H-NMR. Standardized protocols were used for sample preparation, data acquisition, and data pre-processing allowing for an inter-method comparison of serum metabolites consistently measured by both techniques. Importantly, this cross-platform metabolomics study allowed for independent replication of serum biomarker candidates associated with liver fibrosis progression from chronic HCV infection, which complement tissue histopathology, serum liver protein panels, and ultrasound imaging techniques.

RESULTS

Clinical Characteristics of Study Participants

Demographic and clinical characteristics of all study participants, including healthy non-HCV participants (n = 14), and patients with chronic HCV infection at different stages of liver fibrosis (n = 20) are summarized in **Table 1**. In this pilot study, most recruited subjects were older, overweight male adults, with non-HCV participants generally being younger (p = 0.0318). HCV infected patients were treatment naïve at the time of recruitment with most having a HCV genotype A, as confirmed by positive serum anti-HCV antibodies and HCV RNA test results. Staging of liver fibrosis using the METAVIR scoring system was performed by tissue histopathology which confirmed no fibrosis (F0) in non-HCV controls, whereas minimal scarring/inflammation (early-stage, F0-F1, n = 9) or more advanced stages of fibrosis (late-stage, F2-F4, n = 11) in HCV patient sub-groups with four individuals having cirrhosis (F4). Similarly, FibroScan test results using transient elastography measurements revealed no differences in liver stiffness between non-HCV and early-stage liver fibrosis HCV patients (FibroScan <7.0 kPa, $p < 0.05$), in contrast to the late-stage fibrosis HCV patients (FibroScan >8.0 kPa, $p = 1.73 \times 10^{-4}$). As expected, a panel of blood liver proteins and FibroTest scores were elevated in HCV patients as compared to non-HCV controls. However, there were no significant differences in circulating liver protein levels between HCV sub-groups at different stages of liver fibrosis severity ($p > 0.05$).

Serum Metabolome Characterization by Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry

Serum samples were prepared by ultrafiltration to remove protein after dilution with recovery/internal standards prior to MSI-CE-MS analysis, which were analyzed under two configurations for cationic (pH 1.8, positive ion mode) and anionic (pH 8.5, negative ion mode) metabolites with full-scan data acquisition. Sample throughput in MSI-CE-MS is enhanced when using a serial sample injection format comprising 13 serum filtrates analyzed within a single run. In this case, duplicate analysis of each serum filtrate diluted in a distinctive pattern (1:2, 1:1, 2:1) together with a pooled QC sample were acquired in random sequence as shown for alanine and lactic acid in **Figure 1A**. Data pre-processing in MSI-CE-MS combined both targeted analysis of known serum metabolites, as well as a nontargeted screening strategy to authenticate unknown metabolites from a pooled serum sample as described elsewhere (Shanmuganathan et al., 2021). All serum metabolites were annotated based on their

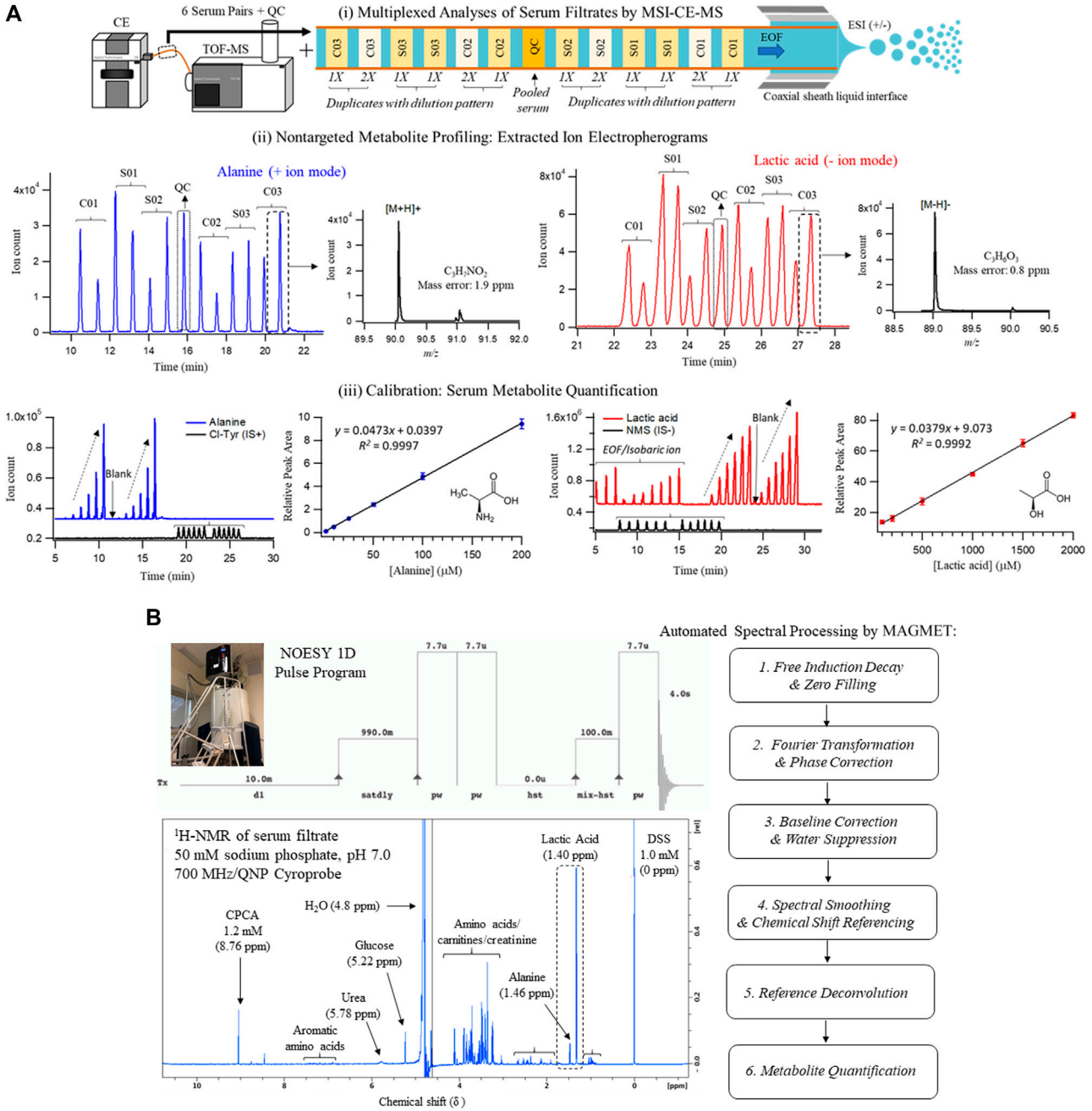

FIGURE 1 | Cross-platform method comparison for characterization of the serum metabolome from HCV patients with liver fibrosis and non-HCV controls when using **(A)** MSI-CE-MS and **(B)** ^{1}H-NMR. In both cases, diluted serum filtrate samples were analyzed using standardized protocols, which included 60 and 30 polar/ionic metabolites measured in most samples (>75%) with adequate precision (CV < 35%) by MSI-CE-MS and NMR, respectively. Each sample was analyzed by NMR with signal averaging (120 scans, 12 min/sample) followed by automated spectral processing by MAGMET using a targeted metabolite library, whereas six pairs of serum filtrates together with a pooled QC were analyzed in each MSI-CE-MS run (8 min/sample) using both targeted and nontargeted approaches with metabolite quantification using external calibration curves with internal standards. Serum metabolite concentrations measured consistently by both platforms are illustrated in this case for alanine and lactic acid.

characteristic accurate mass: relative migration time (*m/z*:RMT) under positive (p) or negative (n) ion mode detection after rejecting spurious signals, background ions and dataset redundancy (Saoi et al., 2020). Overall, 55 serum metabolites in this study were identified with high confidence (level 1) after spiking with authentic standards (*i.e.,* co-migration) and having low mass error (<5 ppm). This also allowed for their quantification using an external calibration curve with ion responses normalized to an internal standard (20 μM, 4-chlorotyrosine, Cl-Tyr or naphthalene monosulfonic acid, NMS) over a 100-fold dynamic range with good linearity (R^2 > 0.990). Otherwise, unknown metabolites (5 compounds, level 4) were annotated based on their most likely

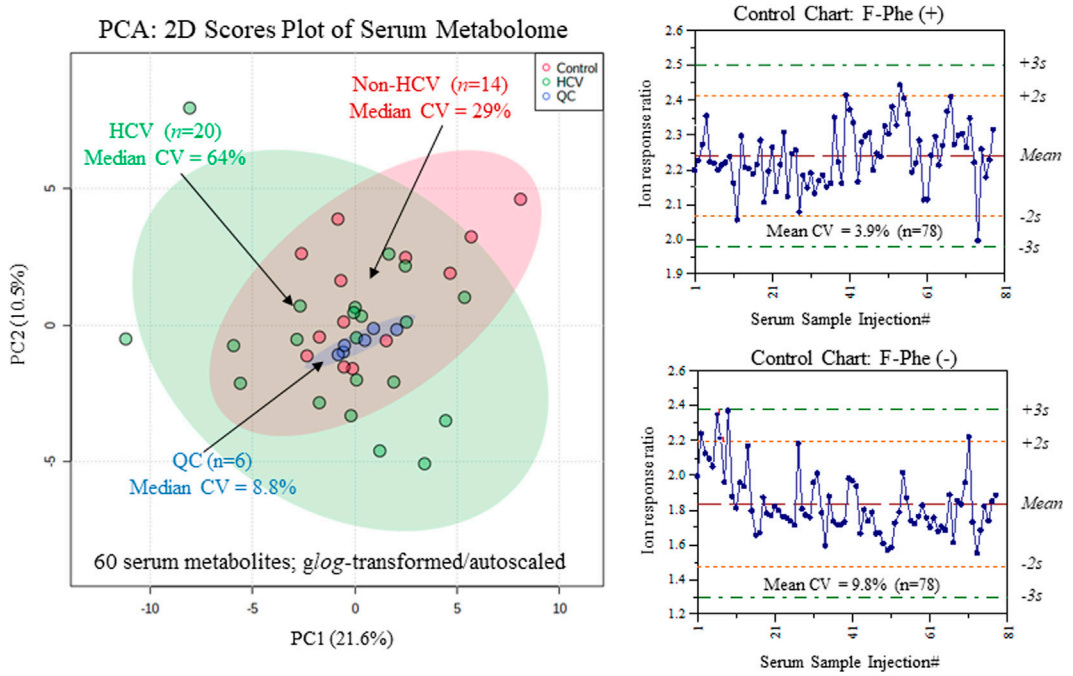

FIGURE 2 | Overview of the serum metabolome and data quality when using MSI-CE-MS as depicted in 2D scores plot using PCA, which compares the technical precision (median CV = 8.8%, n = 6) from repeat QC samples relative to the larger biological variance in chronic HCV liver fibrosis patients (median CV = 64%, n = 20) and non-HCV controls (median CV = 29%, n = 14). Control charts for the recovery standard (3-fluorophenylalanine, F-Phe) confirms acceptable intermediate precision (mean CV = 6.8%, n = 156) when analyzing all serum samples in this study by MSI-CE-MS in positive and negative ion modes.

molecular formula. Overall, 60 polar/ionic metabolites were consistently analyzed (median CV < 35%) in most serum samples (>75%) from non-HCV controls and HCV patients (n = 34) by MSI-CE-MS. **Figure 2** depicts a 2D scores plot from principal component analysis (PCA) confirming acceptable technical precision achieved for 60 serum metabolites from QC samples (median CV = 8.8%, n = 6) relative to the larger biological variance in non-HCV (median CV = 29%, n = 14) and HCV patients (median CV = 64%, n = 20). Also, control charts for the recovery standard (20 µM, 3-fluorophenylalanine, F-Phe) added to all serum samples and analyzed under both MSI-CE-MS configurations demonstrated good intermediate precision (mean CV = 6.8%, n = 156) with few data (~5%) exceeding warning limits (±2s).

Targeted Metabolite Profiling of Serum Filtrates by Nuclear Magnetic Resonance

A standardized approach was also used to prepare serum samples (280 µl) after ultrafiltration to remove protein followed by a 1.25-fold dissolution in a buffer system (70 µl, 250 mM phosphate, pH 7.0 with 54% *vol* D_2O) containing a chemical shift reference that also served as internal standard (1.0 mM 2,2-dimethyl-2-silapentane-5-sulfonate, DSS-d6). ^{1}H-NMR spectra (700 MHz) in 5 mm diameter tubes were acquired using a NOESY pulse sequence with water suppression after a manual shimming protocol, which generated an average line width for DSS-d6 of (0.95 ± 0.14) Hz (n = 40). Each ^{1}H-NMR spectrum required 12 min for 128 scans as shown for a pooled serum filtrate sample from the study cohort in **Figure 1B**. Most proton resonances for serum metabolites have signals clustered within distinct chemical shift windows (δ ~ 3.2–4.2 ppm; 1.8–2.7 ppm; 0.8–1.2 ppm) as compared to prominent peaks for lactic acid (methyl proton, δ = 1.40 ppm, doublet), alanine (methyl proton, δ = 1.46 ppm doublet), and D-glucose (α-anomeric proton, δ = 5.22 ppm, doublet). Raw FID NMR data was uploaded to a user-friendly webserver, Magnetic Resonance for Metabolomics (MAGMET) using an automated workflow for data pre-processing and spectral deconvolution from a library of 47 serum metabolites (Ravanbakhsh et al., 2015). After spectral processing, serum metabolite concentrations were calculated using a reference standard with known concentration (DSS-d5, 1,000 µM). Technical precision was assessed from the intermittent analysis of an external QC comprised of four amino acid standards (mean CV = 2.5%, n = 3), as well as internal QC of pooled serum from cohort (median CV = 9.0%, n = 3) which had higher variance (CV > 35%) for certain serum metabolites prone to spectral interferences (*e.g.,* arginine, methionine, leucine, hydroxyvaleric acid). Overall, 30 polar/ionic metabolites were reliably quantified by ^{1}H-NMR in most serum samples in this study.

Serum Metabolite Quantification: Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry vs Nuclear Magnetic Resonance

An inter-method comparison was next performed for serum metabolites measured with high detection frequency (>75% of

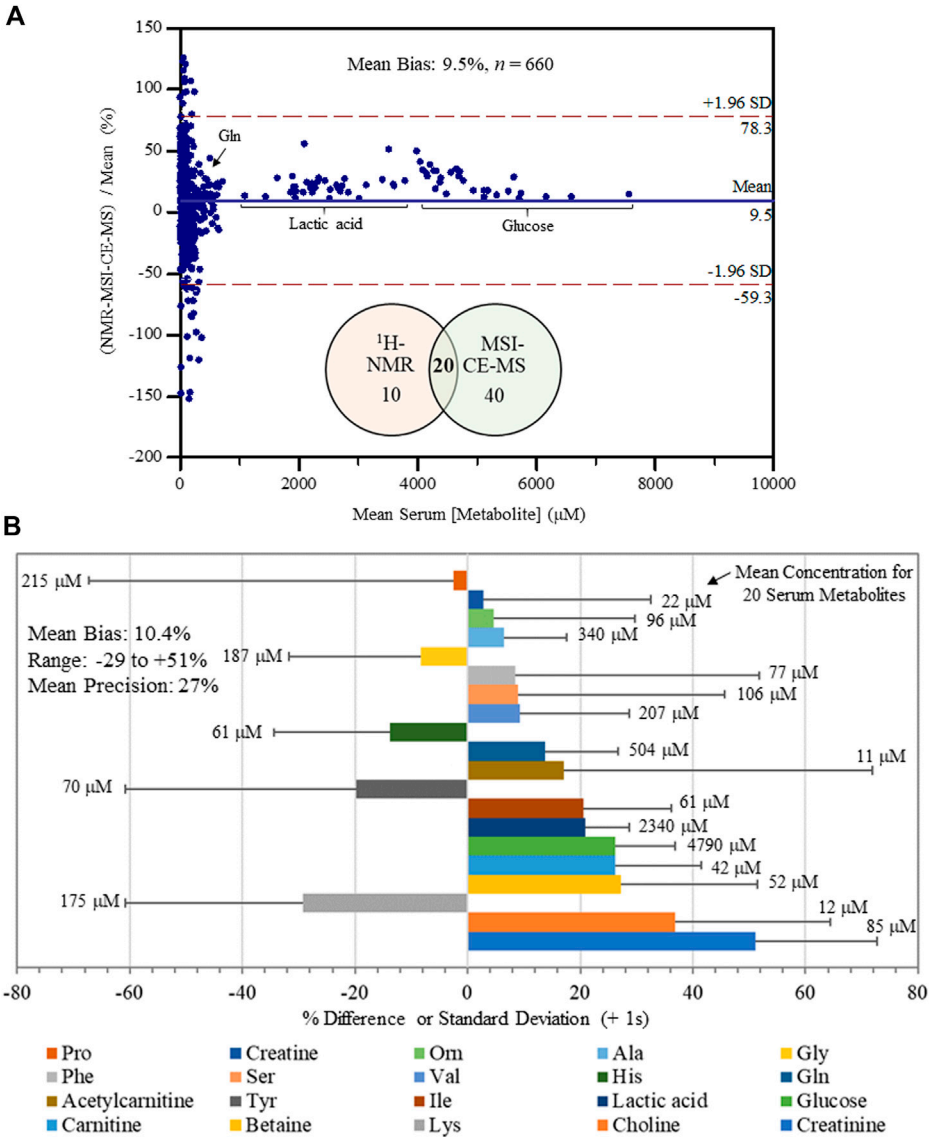

FIGURE 3 | Inter-method comparison for serum metabolite quantification by ^{1}H-NMR and MSI-CE-MS platforms in non-HCV and HCV patients (n = 34). **(A)** A Bland-Altman %difference plot confirms a normal distribution for 20 serum metabolite concentrations measured independently in 34 serum filtrate samples having a mean bias of 9.5% (n = 660) with few metabolites (~5%) exceeding mutual agreement limits (± 2s). **(B)** A summary of the mean concentration (μM) for each serum metabolite consistently measured by both instrumental platforms, as well as a bar graph depicting the mean bias (% difference) together with an error bar indicating the mean precision (+1 s) in reported bias across all samples (n = 34). Serum metabolites analyzed in this study were ranked ordered (top to bottom) from lowest to highest bias having an overall bias and precision of 10.4 and 27%, respectively.

all samples analyzed) and adequate technical precision (CV < 35%) by both MSI-CE-MS and NMR platforms. Concentration detection limits for ^{1}H-NMR under the acquisition conditions used were about 5 μM, but higher detection thresholds were evident for certain serum metabolites prone to chemical shift spectral overlap. In MSI-CE-MS, concentration sensitivity is metabolite dependent given the disparity in solute ionization efficiency (Chalcraft et al., 2009) with detection limits (S/N ~ 3) ranging from 0.2 to 0.5 μM when using a conventional coaxial sheath liquid interface with small volumes (~5 nL) introduced on-capillary. **Figure 3A** depicts a Bland-Altman %difference plot highlighting a normal data distribution with an overall mean bias of 9.5% (n = 660) based on 20 serum metabolite concentrations measured in 34 serum samples by MSI-CE-MS and NMR with few missing data (20 or 2.9% in total). There was a significant overlap in metabolome coverage (~67%) between both platforms that comprised primarily micromolar levels of polar/ionic metabolites from serum filtrates. Over a 500-fold dynamic range in serum metabolite concentrations was assessed in non-HCV controls and HCV patients (n = 34) ranging from D-glucose

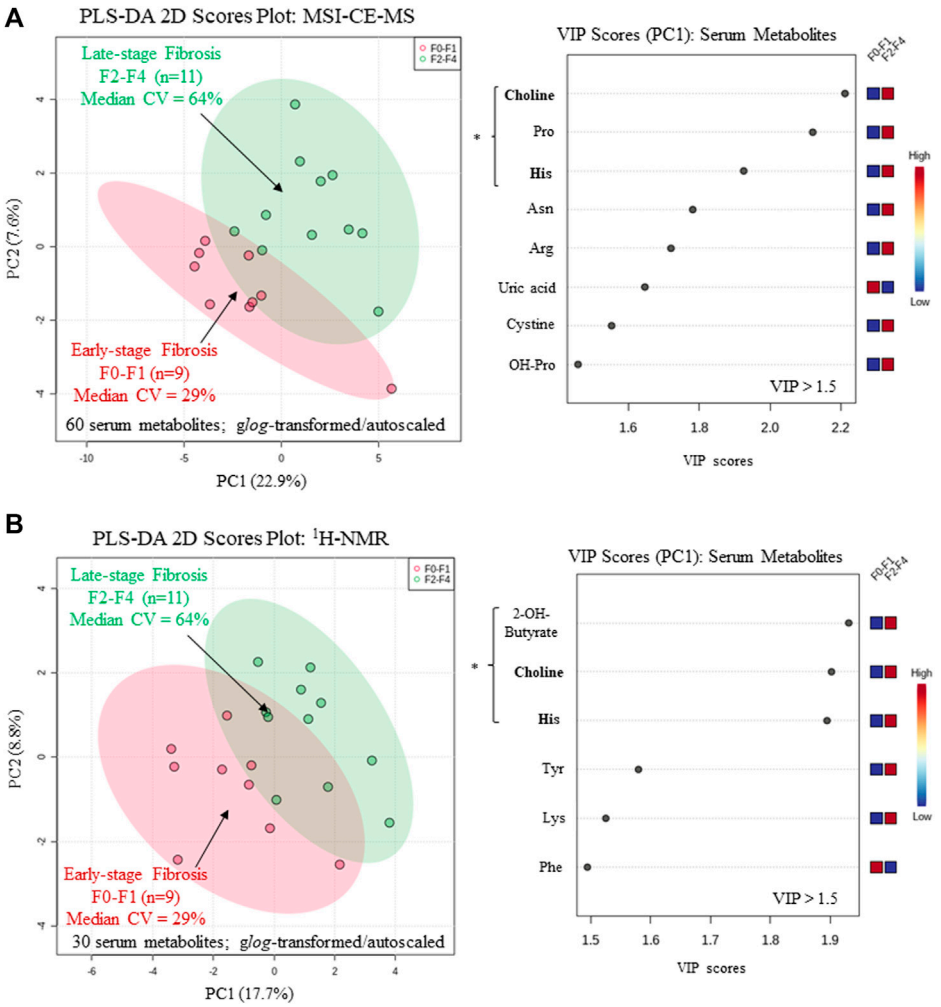

FIGURE 4 | Supervised multivariate data analysis of serum metabolome characterized by **(A)** MSI-CE-MS and **(B)** 1H-NMR when using a partial least-squares-discriminate analysis (PLS-DA) model for differentiation of chronic HCV patients with early (F0-F1) from late-stage (F2-F4) liver fibrosis as determined by histopathology. A variable importance in projection (VIP) scores (VIP >1.5) were used to rank order important serum metabolites associated with liver fibrosis progression highlighting consistent outcomes for choline and histidine (bolded) in both platforms. All serum metabolites were generalized *log*-transformed and autoscaled, where * indicates statistically significant serum metabolites between HCV sub-groups ($p < 0.05$).

(mean concentration of 4.8 mM) to *O*-acetyl-*L*-carnitine (mean concentration of 11 μM). Bias was evident among a sub-set of samples/serum metabolites (*e.g.,* proline, tyrosine, phenylalanine), yet only a small fraction of total data (~5.0%) exceeded mutual agreement limits (±2s). **Figure 3B** illustrates the bias distribution for each of the 20 serum metabolites (mean bias of 10.4% ranging from -29 to +51%) that are depicted as solid bars. Also, the mean precision in measured bias between the two platforms is 27%, which is shown as an error bar (+1s) for each serum metabolite. Overall, serum alanine, glycine, ornithine, valine, histidine, glutamine, isoleucine, lactic acid, glucose, carnitine and betaine had the most consistent measurements across both methods in terms of acceptable bias and variability (<25%) when using only a single internal standard for data normalization. An excel file in the **Supplementary Material** provides a complete list of serum metabolites and their responses/concentrations measured in non-HCV and HCV patients using MSI-CE-MS and NMR, including calibration curves and figures of merit acquired for 20 serum metabolites used in this inter-method comparison.

Serum Metabolites Differentiating Liver Fibrosis Progression in Hepatitis C

A major focus of this pilot study was to identify putative serum biomarkers that differentiate liver fibrosis in treatment naïve HCV patients. Complementary multivariate and univariate statistical analyses were performed on serum metabolome data to identify putative serum biomarkers that may enable less invasive assessment of liver fibrosis. **Figure 4** compares two partial least squares-discriminant analysis (PLS-DA) models

TABLE 2 | Cross-platform comparison of serum metabolites that differentiate HCV patients with late (F2-F4, $n = 11$) to early-stage (F0-F1, $n = 9$) liver fibrosis.

Serum Metabolite/ID	Mean FC	*p*-value	Effect size[a]
MSI-CE-MS	—	—	—
Choline	1.43	0.0312	1.03
Proline	1.41	0.0401	0.99
Histidine	1.19	0.0653	0.90
^{1}H-NMR	—	—	—
2-Hydroxybutyric acid	2.24	0.0307	1.04
Choline	1.32	0.0447	0.94
Histidine	1.21	0.0456	0.97

[a]*A two-tailed student's t-test with equal variance of log-transformed serum metabolite concentrations was used to differentiate liver fibrosis progression, where effect size is defined as eta^2.*

from MSI-CE-MS and ^{1}H-NMR data to rank order significant serum metabolites (VIP >1.5) that differentiate late-stage fibrosis (F2-F4, $n = 11$) from early-stage (F0-F1, $n = 9$) fibrosis in well-matched HCV patients based on their METAVIR scores (**Table 1**). Overall, serum choline and histidine were among the top ranked metabolites consistently elevated in late-stage as compared to early-stage fibrosis HCV patients by both MSI-CE-MS and NMR. **Table 2** confirms that both MSI-CE-MS (choline, proline, histidine) and NMR (2-hydroxybutyric acid, choline, histidine) identified four serum metabolites elevated (mean fold-change, FC > 1.2, $p < 0.05$, effect size >0.90) with more advanced stages of liver fibrosis when using a two-tailed student's t-test with equal variance. Hydroxybutyric acid isomers were not fully resolved by MSI-CE-MS in this study preventing their accurate quantification, whereas serum proline was not found to be different between HCV sub-groups using NMR, and thus not independently replicated. Several other serum metabolites also had higher serum concentrations with increasing liver fibrosis (*e.g.,* asparagine, arginine, tyrosine, hydroxyproline) in contrast to uric acid and phenylalanine (**Figure 4**); however, these trends were not statistically significant ($p > 0.05$).

Figure 5A depicts a receiver operating characteristic (ROC) curve for the ratio of serum choline to uric acid based on MSI-CE-MS data, which provided optimal discrimination between HCV liver fibrosis patient sub-groups (AUC = 0.848, $p = 0.00766$) not feasible by serum liver protein panels or FibroTest scores (**Table 1**). Moreover, **Figure 5B** confirms that there was a moderate positive

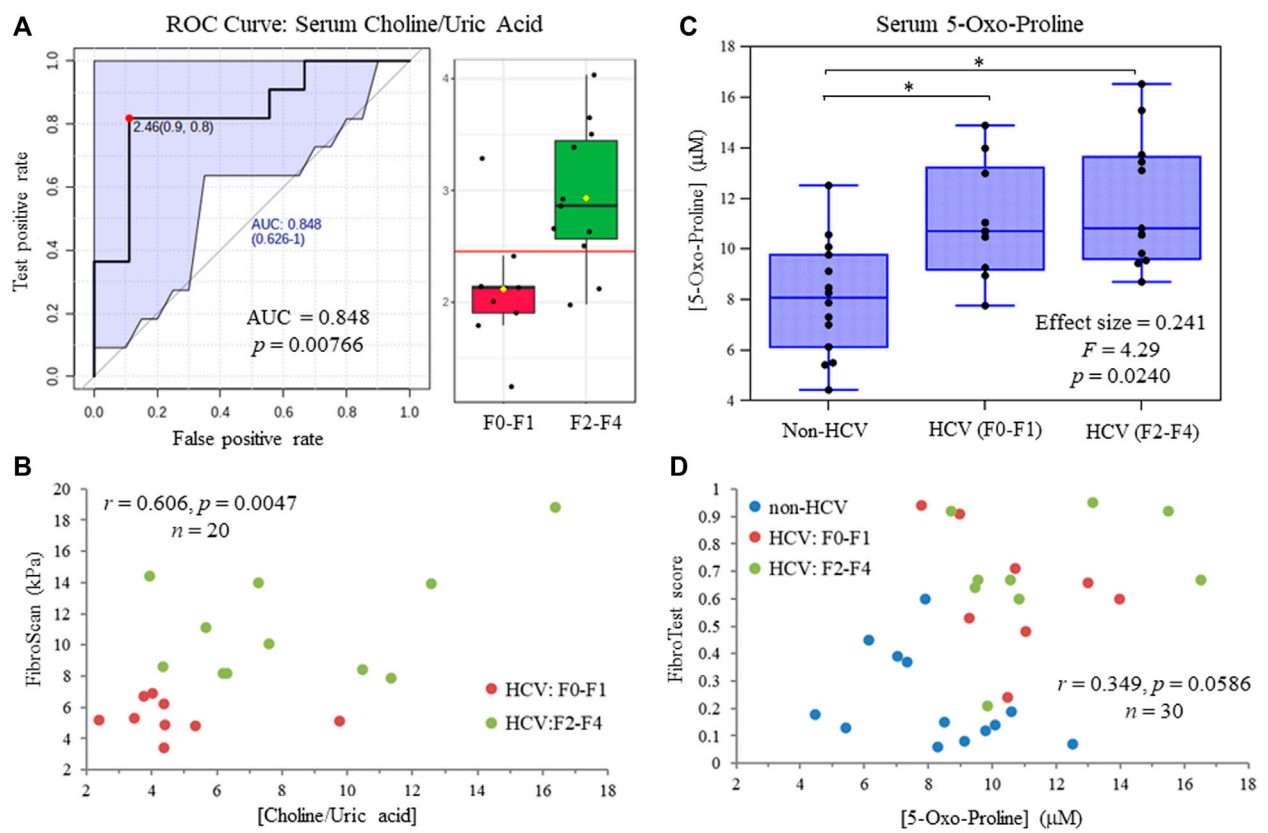

FIGURE 5 | (A) Receiver operating characteristic (ROC) curve highlighting that serum choline:uric acid ratio optimally differentiates late (F2-F4) from early (F0-F1) stage liver fibrosis in treatment naïve HCV patients with an area under the curve, AUC = 0.848. **(B)** A moderate positive correlation is shown between the serum choline to uric acid ratio and independent FibroScan results from ultrasound imaging of HCV patients at different stages of liver fibrosis. **(C)** ANCOVA (between-subject effects) with covariate adjustments confirms that serum 5-oxo-proline concentrations are elevated in HCV patients with early and late-stage liver fibrosis as compared to non-HCV controls (*$p < 0.05$). **(D)** A weak positive correlation is depicted for serum 5-oxo-proline concentrations and FibroTest scores from a blood liver protein panel that differentiate non-HCV controls from HCV patients with chronic hepatic inflammation.

correlation (r = 0.606, p = 0.0047, n = 20) between the serum choline to uric acid ratio and liver stiffness measurements from FibroScan test results (kPa) in HCV patients. Also, an ANCOVA (between-subject effects) with adjustment for covariates (age, sex and BMI) revealed that serum 5-oxo-proline concentrations measured only by MSI-CE-MS were consistently elevated (mean FC = 1.44, F = 4.29, p = 0.0240) in HCV patients with early and late-stage fibrosis as compared to non-HCV controls as shown in **Figure 5C**. In contrast, liver stiffness measurements based on FibroScan test results were not able to differentiate non-HCV controls from early-stage fibrosis (F0-F1) HCV patients (**Table 1**). **Figure 5D** highlights that there was a weak correlation between serum 5-oxo-proline concentrations and FibroTest score (r = 0.349, p = 0.0586, n = 30), which is derived from an age/sex-adjusted algorithm of five serum liver proteins.

DISCUSSION

Cross-Platform Serum Metabolomics Comparison

CE-MS based methods have not been widely used within the metabolomics community due to longstanding concerns regarding migration time variability and long-term robustness. However, these technical obstacles can be overcome with implementation of standardized protocols in large-scale metabolomic studies (Harada et al., 2018; Shanmuganathan et al., 2021). These protocols have been recently implemented in international ring trials to demonstrate cross-laboratory comparability (Drouin et al., 2020). Also, CE-MS inter-method comparisons increasingly demonstrate reliable quantification of metabolite concentrations relative to validated assays used in a clinical setting (DiBattista et al., 2017; Wild et al., 2019; Azab et al., 2020). Yet, there have been few cross-platform metabolomic studies involving CE-MS in conjunction with other widely used instrumental methods, such as LC-MS and GC-MS (Büscher et al., 2009; Naz et al., 2013; Rojo et al., 2015). To the best our knowledge, this work represents the first direct comparison between CE-MS and NMR for characterization of the serum metabolome following ultrafiltration.

There was considerable metabolome overlap between MSI-CE-MS and ^{1}H-NMR with 20 metabolites consistently measured in serum filtrate samples, including amino acids, carnitines, organic acids, and glucose (**Figure 3**). MAGMET uses automated spectral processing, phasing, water-removal, baseline correction, chemical shift referencing, peak picking, curve fitting and spectral deconvolution via a biofluid-specific reference library for targeted quantitative profiling of 47 serum metabolites (Ravanbakhsh et al., 2015); however, only 30 serum metabolites were reliably measured in most non-HCV controls and HCV patients. This was likely due to sub-optimal shimming contributing to higher-than-average spectral line widths (~0.95 Hz) and lower signal-to-noise for detecting several lower abundance metabolites (*e.g.,* hypoxanthine, acetoacetate) and organic solvents (*e.g.,* methanol, acetone) within the spectral library. In contrast, MSI-CE-MS takes advantage of both a targeted and nontargeted metabolomics data workflow with serum filtrates analyzed under two separation/ionization conditions for cationic and anionic metabolites (**Figure 1**); this process also allows for the discovery of unknown metabolites lacking authentic standards or reference data available in the Human Metabolome Database (Wishart et al., 2018). In this study, 60 serum metabolites (including five unknowns) were measured in most serum filtrates by MSI-CE-MS with good technical precision (CV < 10%) comparable to NMR (**Figure 2**). The greater serum metabolome coverage for MSI-CE-MS is mainly attributed to its improved resolution and lower detection limits as compared to 1D ^{1}H-NMR. Both methods had good mutual agreement for 20 metabolites measured in 34 serum samples from non-HCV and HCV patients with a mean bias of 9.5% (n = 660) and few outlier data (**Figure 3**). Better quantitative performance in MSI-CE-MS may be realized when using matching stable-isotope internal standards with multiple reaction monitoring (Saoi et al., 2020) to correct for potential ion suppression effects unlike discovery-based metabolomics using a single internal standard with full-scan data acquisition. Furthermore, concentration sensitivity with deeper metabolome coverage can be further enhanced when using sheathless or low-flow CE-MS interface designs (Zhang et al., 2017).

Similarly, NMR quantification using automated MAGMET processing was prone to spectral interferences and potential bias, which can be minimized with lower detection limits when using higher field magnets. Overall, multiplexed electrophoretic separations based on MSI-CE-MS (**Figure 1**) enable faster data acquisition than NMR (<8 min/sample [both ion modes] vs 12 min/sample) with a two-fold greater metabolome coverage and analogous reproducibility. Additionally, MSI-CE-MS requires far less sample volume than NMR (<5 µl [if required] serum vs 450 µl serum) that is optimal for analysis of volume-restricted biospecimens (Nori de Macedo et al., 2017), and single-cell analyses (Duncan et al., 2019). Expanded metabolome coverage can be further realized when using non-aqueous buffers systems in MSI-CE-MS for a diverse range of water-insoluble yet ionic lipids and fatty acids from serum ether extracts (Azab et al., 2020).

Serum Metabolites Signatures of Liver Fibrosis in Chronic Hepatitis C Virus Patients

Chronic HCV viral infections are one of several causes of liver disease, including fibrosis, cirrhosis, and hepatocellular carcinoma. Although liver biopsies remain the gold standard for diagnosis and staging severity, there is increasing use of ultrasound-based transient elastography for monitoring liver fibrosis progression as it correlates well with the METAVIR scoring system. However, this non-invasive method for assessing liver stiffness suffers higher failure rates for obese patients with ascites, and improved diagnostic confidence is achieved for mild disease states when combined with serum protein biomarker panels (Wilder and Patel, 2014). In our study, recruited HCV patients at different stages of liver fibrosis were well-matched (age, sex, BMI, HCV genotype, treatment naïve), with FibroScan test results differentiating moderate to severe fibrosis (F2-F3) or cirrhosis (F4) from no fibrosis (F0) or mild fibrosis (F1) patient sub-groups (**Table 1**).

There was no difference in liver stiffness measurements when comparing early-stage liver fibrosis (F0-F1) HCV patients to a younger/healthy non-HCV control unlike the FibroTest score primarily from elevated γ-glutamyl transferase and alpha-2-macroglobulin levels in circulation. Despite aberrant metabolic signatures of fibrosis reported in various liver diseases (Chang and Yang, 2019), few metabolomic studies have identified serum biomarkers to differentiate liver fibrosis progression caused by HCV infection (Sarfaraz et al., 2016).

In this work, we identified four serum metabolites that differentiate fibrosis progression in HCV patients, including choline, histidine, proline and 2-hydroxybutyric acid. Interestingly, choline and histidine from serum filtrates were elevated in severe fibrosis/cirrhosis (F2-F4) as compared to early-stage fibrosis (F0-F2) HCV patients as independently confirmed by both MSI-CE-MS and NMR methods (**Table 2**). Choline is an essential dietary nutrient required for the biosynthesis of phospholipids and donor in methylation reactions, which plays a critical role in maintaining liver function (Mehedint and Zeisel, 2013). Although choline deficient diets can contribute to fatty liver and hepatic fibrosis, previous NMR metabolomic studies have shown that elevated serum choline in HCV patients differentiates hepatocellular carcinoma from HCV controls without liver cancer (Wei et al., 2012). Yet, contradictory findings were reported by NMR with lower levels of serum choline in hepatocellular carcinoma as compared to patients with liver cirrhosis (Gao et al., 2009). In our case, chronic HCV patients with hepatocellular carcinoma were excluded from study recruitment. However, more advanced stages of liver fibrosis/cirrhosis may progress to cancer without effective anti-viral medications. Additionally, serum histidine was previously demonstrated by NMR to be elevated in HCV patients with increasing liver fibrosis and necroinflammation (Safaraz et al., 2016). Histidine is an essential amino acid that also functions as an important antioxidant and metal chelator, which has been shown to attenuate thioacetamide-induced liver fibrosis in rats (El-Batch et al., 2011). Independent replication by two orthogonal instrumental methods further supports our findings in a well-matched patient cohort despite a modest study power involving primarily older male participants. Overall, the serum choline to uric acid ratio was found to improve discrimination between different stages of liver fibrosis (**Figure 5A**) as compared to choline alone (AUC = 0.803, $p = 0.0317$). In our study, lower circulating levels of uric acid were associated with more advanced fibrosis in largely overweight male HCV patients (**Figure 4A**), although hyperuricemia has been associated with liver damage in patients with non-alcoholic fatty liver disease (Afzali et al., 2010; Zhou et al., 2016). Importantly, there was a positive correlation ($r = 0.606$, $p = 0.0047$) in the serum choline to uric acid ratio with FibroScan test results, which links aberrant metabolism in circulation to liver disease phenotype/pathology in HCV patients (**Figure 5B**). Further validation in a larger patient cohort is needed to reproduce our findings and better demonstrate its clinical utility when used in conjunction with ultrasound imaging techniques.

Serum 5-oxo-proline as measured by MSI-CE-MS was found to be elevated in both early and late-stage fibrosis HCV patients as compared to non-HCV controls (**Figure 5C**). Oxo-proline (or pyroglutamic acid) is an important yet infrequently measured amino acid intermediate within the glutathione cycle (Bachhawat and Yadav, 2018), which accumulates in circulation due to acquired 5-oxoprolinemia from hepatic oxidative insult and glutathione depletion (Gamarra et al., 2019). However, most studies to date have focused on a high anion gap metabolic acidosis from 5-oxoprolinemia following sepsis or acetaminophen toxicity (Liss et al., 2013) rather than liver fibrosis/inflammation from chronic HCV infection or non-alcoholic steatohepatitis (Saoi et al., 2020). Serum cystine concentrations were also elevated among HCV patients with liver fibrosis relative to non-HCV controls in this study indicative of elevated oxidative stress in liver diseases (Cichoz-Lach and Michalak, 2014). In fact, serum 5-oxo-proline and cystine were highly co-linear ($r = 0.619$, $p = 9.38 \times 10^{-5}$, $n = 34$), yet they were not measured by NMR preventing their independent replication. FibroScan test results did not differentiate early-stage fibrosis HCV patients (F0-F1) from non-HCV controls without fibrosis unlike specific serum liver proteins, or FibroTest scores (**Table 1**), which were weakly correlated with oxo-proline concentrations (**Figure 5D**). Elevated circulating concentrations of oxo-proline reflecting oxidative stress and impaired glutathione-dependent redox homeostasis offers a plausible biochemical mechanism associated with the onset and progression of liver fibrosis in HCV patients.

CONCLUDING REMARKS

We conducted the first cross-platform serum metabolomics study to compare the performance of MSI-CE-MS and NMR methods using standardized protocols, which was also applied to identify putative biomarkers of liver fibrosis from chronic HCV infection. Both techniques offered similar reproducibility with good mutual agreement and few outliers when quantifying 20 serum metabolites using a single internal standard. A targeted NMR metabolomics approach was facilitated by use of an automated spectral processing and deconvolution software together with a serum-specific metabolite library; however, sub-optimal shimming contributed to line width broadening and lower sensitivity with potential bias for certain serum metabolites prone to spectral interference. On the other hand, multiplexed separations by MSI-CE-MS offer faster data acquisition speeds, much lower sample volume requirements and greater metabolome coverage. Four serum metabolites were elevated in HCV patients with more advanced liver fibrosis severity, with choline and histidine being replicated independently by both instrumental platforms. Overall, the choline to uric acid ratio was found to optimally differentiate between late (F2-F4) and early (F0-F1) stages of liver fibrosis, which was also correlated well with liver stiffness measurements by ultra-sound imaging. Similar to serum liver protein panels, oxo-proline concentrations were higher in HCV patients with liver fibrosis as compared to non-HCV controls reflecting elevated oxidative stress and glutathione depletion from chronic inflammation. Further

validation of serum biomarker candidates in this pilot study is warranted in a larger cohort of HCV patients while evaluating their clinical utility as compared to FibroTest scores and FibroScan test results. Serum biomarkers of hepatic fibrosis offer a less invasive procedure to liver biopsies when monitoring disease progression and treatment interventions for HCV patients to prevent end-stage liver failure.

MATERIALS AND METHODS

Chemical and Reagents

All metabolite standards and buffers were purchased from Sigma-Aldrich (St. Louis, MO, United States). All LC-MS grade solvents, including acetonitrile, isopropanol, methanol, and water were obtained from Caledon Laboratories Ltd (Georgetown, ON, Canada). Calibrant solutions for serum metabolites were prepared by serial dilution of stock solutions (50 mM) in LC-MS grade water and stored refrigerated (4°C). A NMR Metabolomics Analysis Kit with access codes to MAGMET software were supplied by The Metabolomics Innovation Centre (Edmonton, AB, Canada). The kit includes Amicon filters with a 3 kDa molecular weight cut-off (MWCO), microcentrifuge tubes, NMR buffer (250 mM potassium phosphate, pH 7.0, 5 mM 2,2-dimethyl-2-silapentane-5 sulfonate, DSS-d_6, 5.84 mM 2-chloropyrimidine-5-carboxylic acid, CPCA, and D_2O 54% *vol* in H_2O) and a QC standard mixture (1.25 mM glycine, 1.25 mM alanine, 1.25 mM threonine and 1.25 mM aspartic acid).

Study Population and Sample Collection

The study approval was obtained from the McMaster University Health Research Ethics Board (REB Project #0932) and all study participants (20 patients with chronic hepatitis C and 14 participants as healthy controls) provided signed informed consent for study enrolment according to the Declaration of Helsinki. Relevant clinical and demographic information was also collected, including sex, age, body mass index (BMI), HCV genotype, cardiovascular risk factors (*e.g.,* hypertension, diabetes, and dyslipidemia), medication history (*e.g.,* lipid lowering therapies, oral hypoglycemics, and insulin), and habitual alcohol intake. An attending physician recruited patients if they meet the inclusion criteria while a research nurse, independent of the attending physician, obtained consent. Peripheral blood samples were collected during clinic visits. Following blood clotting (45 min at 25°C), the isolated serum samples were immediately transferred to cryovials and stored at −80°C. Routine clinical tests were also collected using Li-heparin, K-EDTA, and plain collection vials for total bilirubin, alanine aminotransferase, γ-glutamyltransferase, alpha-2-macroglogulin, haptoglobin and apolipoprotein A1. The study included patients chronically infected with HCV from the 2F Digestive Diseases Clinic at McMaster University (Hamilton, ON). Study inclusion criteria included: 1) adult patients (≥18 years of age) and 2) treatment naïve, chronic HCV patients (genotype 1, positive anti-HCV antibodies and HCV RNA in serum). Exclusion criteria included: 1) chronic hepatitis A, B, D and E, 2) conditions that may alter the accuracy of serum biomarkers of fibrosis: extra-hepatic biliary obstruction; immunosuppression (*e.g.,* due to HIV, medications); pregnancy; and systemic inflammatory conditions (*e.g.,* sepsis, inflammatory bowel disease), 3) excessive alcohol consumption defined as ≥ 40 g/d for men and ≥20 g/d for women, 4) antiviral therapy for HCV within the previous 6 months, 5) patients with NAFLD as determined by echogenic liver on ultrasound, and 6) patients with hepatocellular carcinoma. Percutaneous liver biopsy was performed under local anesthesia with an ultrasound guidance via the right costal approach (McGill, 2001). Although liver biopsies are considered the "gold standard" for staging fibrosis, there has been a decline in its use (Myers et al., 2014). FibroScan test results were used to non-invasively assess liver fibrosis progression in non-HCV and HCV patients at the Liver Clinic at McMaster University, and liver fibrosis was graded using the METAVIR scoring system. Also, the FibroTest score was calculated from a panel of five serum liver protein measurements after adjustment for age and sex.

Serum Filtrate Preparation Prior to Metabolomic Analyses

Prior to NMR and MSI-CE-MS analysis, frozen serum samples were slowly thawed on ice and a pooled quality control (QC) sample was prepared by taking 10 µl aliquots of serum from each participant in the study. All serum samples were processed according to the protocol provided by the TMIC NMR metabolomics kit (http://magmet.ca/spectra_collection). The serum pre-treatment protocol was as follows: 3 kDa MWCO filters (Amicon Ultra 0.5 ml Centrifugal Filter Unit, Millipore Sigma Inc.) were first rinsed with 500 µl of water and centrifuged for 15 min at 14,000 *g* to remove additives from the manufacturing process. The rinsing process was repeated five times after which the filters were air dried prior to serum processing. An aliquot of 450 µl of serum sample was then added to the pre-rinsed filter tube and centrifuged for 25 min at 14,000 *g*. The serum filtrate containing free circulating polar/ionic metabolites (*i.e.,* non-protein bound fraction) was then separately aliquoted for independent MSI-CE-MS and ^{1}H-NMR analysis. In this case, removal of serum protein by ultrafiltration reduced spectral interferences for metabolite quantification in NMR, as well as deleterious capillary surface adsorption and ion source contamination in MSI-CE-MS. An aliquot of 50 µl was used for MSI-CE-MS analysis while 280 µl was required for NMR analysis. The serum filtrate (280 µl) for NMR analysis was then diluted 1.25-fold with 70 µl of the NMR kit buffer and the solution was vortexed for 30 s and transferred to a 5 mm NMR tube. The serum filtrate (50 µl) for MSI-CE-MS analysis was diluted 4-fold in deionized water containing several internal/recovery standards, including 4-fluoro-*L*-phenylalanine (F-Phe, 20 µM), 3-chloro-*L*-tyrosine (Cl-Tyr, 20 µM), 2-naphthalenesulfonic acid (NMS, 20 µM), and $^{13}C_6$-*D*-glucose (^{13}C-glucose, 2 mM). The solution was vortexed for 30 s and a 20 µl aliquot was transferred into a polypropylene vial prior to MSI-CE-MS analysis.

Nuclear Magnetic Resonance Data Acquisition

Data was obtained on a Bruker Avance III 700 MHz NMR spectrometer (Bruker Biospin, Rheinstetten, Germany) equipped with a 5 mm QNP cryoprobe, operating at 700.17 MHz for ^{1}H and controlled by TopSpin software (v.3.5 for Linux OS). Data was collected at room temperature, using a noesypr 1d pulse program with water suppression (Ravanbakhsh et al., 2015). The acquisition and mixing time were set to 4 s and 100 ms, respectively. Spectra was acquired with eight steady state scans with a field width of ≤80 Hz and O1P and spectral width were set to 4.69 and 12 ppm, respectively. Each sample was shimmed using a manual shimming protocol which included Z6 shimming along the *Z-X-Y* axes before the Z-X-Y-XZ-YZ axes to maintain a peak linewidth for DSS-d5 (<1 Hz) required for automated MAGMET spectral processing. Each sample required about 12 min to complete 128 scans.

Nuclear Magnetic Resonance Data Spectral Profiling and Annotation

The raw NMR data was uploaded in a zip format (FID file) to the MAGMET webserver (http://magmet.ca/users/login) that is based on the automated ^{1}H-NMR spectral processing of serum filtrate samples as described previously (Ravanbakhsh et al., 2015). Briefly, a standardized workflow is used to process NMR spectral files, including Fourier transform/phase correction, baseline correction, water suppression, spectral smoothing, chemical shift referencing, followed by spectral deconvolution and metabolite quantification from a library of 47 serum metabolites. The following parameters were selected: the biofluid was set to serum, spectrometer frequency was set to 700 MHz, chemical shift (CS) reference and CS concentration were set to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS-d6) and 1,000 μM, respectively. The internal DSS-d6 standard was used to calculate the concentration of all metabolites detected in the serum filtrates by comparing the peak area of individual metabolites in the spectra with the known concentration of DSS-d6. Also, CPCA was used for optimal automated phase correction based on its stable chemical shift at 8.76 ppm. A list of identified metabolites with their absolute concentration was outputted as a table format from the webserver. Serum metabolites detected in more than 75% of serum samples analyzed in this study were included in the data matrix for statistical analysis and any missing values and/or non-detects were replaced with half of lowest detected value.

Serum Metabolomics by Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry

An Agilent 6230 time-of-flight mass spectrometer (TOF-MS) with a coaxial sheath liquid Jetstream electrospray ion source with heated nitrogen gas was equipped to an Agilent G7100A capillary electrophoresis (CE) unit and used for the analysis of polar/ionic metabolites under aqueous buffer conditions. An Agilent 1260 Infinity Isocratic pump, and a 1260 Infinity degasser were used to deliver the sheath liquid at a rate of 10 μl/min. Separations were performed on an uncoated open tubular fused-silica capillary with an internal diameter of 50 μm and outer diameter of 360 μm (Polymicro Technologies Inc., AZ, United States) with a total capillary length of 135 cm. About 7 mm of the polyimide coating was removed from both ends of the capillary using a capillary window maker (MicroSolv, Leland, NC, United States) to reduce sample carry-over and prevent polymer swelling and/or degradation of the outer polyimide coating. Purine (10 μl) and hexakis (2,2,3,3-tetrafluoropropoxy) phosphazine (HP-921, 10 μl) were added into 200 ml of sheath liquid (0.1% *vol* formic acid in 60:40 MeOH:H_2O, and 50:50 MeOH: H_2O for positive and negative ion mode, respectively) to allow for real-time mass correction while also monitoring for potential matrix-induced ion suppression effects during separations since constant mass signals were detected at *m/z* 121.0509 and 922.0098 for purine and HP-921, respectively. The instrument was operated in 2 GHz extended dynamic range under positive and negative ion modes that spanned a mass range of *m/z* 50–1700. The data acquisition rate was set to 500 ms/spectrum and both profile and centroid data was stored in a "*.d" file format. The electrospray ionization conditions were set to, 2000 V for the Vcap and nozzle voltage during separation while turned off during injection, nebulizer was turned off during injection but was set to 10 psi during separation, while the drying gas was delivered at 8 L/min at 300°C with a sheath gas flow of 3.5 L/min at 195°C. In addition, the MS voltage settings for the fragmentor, skimmer and Oct1 RF were set to 120, 65, and 750 V, respectively. Instrument control and data acquisition were performed using Agilent MassHunter Workstation LC/MS Data Acquisition Software (B.06.01). New capillaries were conditioned by flushing at high pressure (900 mbar) with methanol for 30 min, 1.0 M NaOH for 30 min, de-ionized water for 30 min, and background electrolyte (BGE) for 30 min. At the start of each day, the CE electrode and MS interface was wiped daily with isopropanol:water (50:50) to avoid salt build-up followed by mass calibration of TOF-MS instrument as preventative maintenance measures.

All serum filtrate samples were analyzed by MSI-CE-MS under two configurations prior to a 10 min capillary flush with BGE namely an acidic BGE under positive ion mode for cationic/zwitterionic metabolites (1 M formic acid with 15% *vol* acetonitrile, pH 1.8), and an alkaline BGE under negative ion mode for acidic metabolites (50 mM ammonium bicarbonate, pH 8.5 adjusted with ammonium hydroxide). Serial sample injections in MSI-CE-MS were performed by alternating a hydrodynamic injection for each serum filtrate (100 mbar for 5 s) followed by an electrokinetic injection of BGE (30 kV for 75 s) to initiate electrophoretic separation at the capillary inlet. This interrupted separation process was repeated for a total of 13 serum samples that were introduced in a randomized order within a single run by MSI-CE-MS to ensure no effective loss in separation performance (Saoi et al., 2019). An applied voltage of 30 kV at 25°C was used for all runs in both positive and negative modes while a pressure gradient of 2 mbar/min was applied during separation (total time of 45 min) to allow for faster elution of slower migrating metabolites. BGE and sheath liquid were degassed before use by

sonication for 10 min. Data normalization for peak integration used 20 μM 4-chlorotyrosine (Cl-Tyr) and naphthalene monosulfonic acid (NMS) as internal standards for positive and negative ion mode, respectively with the exception of glucose (total hexose) that used $^{13}C_6$-glucose as it co-migrates with the electroosmotic flow (Shanmuganathan et al., 2021). A recovery standard, 3-fluorophenylalanine (F-Phe) was added to all serum samples prior to ultrafiltration to monitor for technical precision using control charts. Each sample effectively required 4 min to analyze by MSI-CE-MS in each ion mode while including a pooled quality control (QC) sample in every run. Metabolite identification for serum metabolites was confirmed (*e.g.,* co-migration, accurate mass) by spiking authentic standards in pooled serum filtrates. Serum metabolite quantification by MSI-CE-MS was achieved using a six-point calibration curve (in duplicate) over a 100-fold dynamic range with good linearity ($R^2 \sim 0.998$) after least-squares linear regression with detection (S/N ~ 3) and quantification (S/N ~ 10) limits ranging from 0.2 to 0.5 μM to 1 and 2 μM, respectively as summarized in the excel file of the **Supplementary Material**.

Statistical Analysis

Raw MSI-CE-MS data (*.d format) was processed using Mass-Hunter Workstation Qualitative Analysis software (version B.06.00, Agilent Technologies, 2012). A comprehensive study of all detectable molecular features from the raw data was performed using Mass-Hunter Molecular Feature Extractor, Molecular Formula Generator tools, and an in-house compound database. Molecular features were extracted using a symmetric 10 ppm mass window and all ions were annotated using their accurate mass (*m/z*), relative migration time (RMT) normalized to an internal standard (Cl-Tyr, NMS, or ^{13}C-glucose), and ionization mode of detection (p: positive, n: negative). RMTs are reported since they are an important parameter used to exclude redundant adducts and/or fragment ion peaks, which exhibit identical RMTs as the parent compound. Peak smoothing was performed using a quadratic/cubic Savitzky-Golay function (7 points) prior to peak integration. Peak areas and migration times for all molecular features and internal standards were transferred to an Excel worksheet (Microsoft Office) and relative peak areas (RPA) for each unique molecular feature was saved as csv file. Molecular features detected in more than 75% of all serum samples analyzed with a coefficient of variance (CV < 35%) for QC samples were included in the final data matrix for further statistical analysis. Any non-detects were replaced by a value that was half the detection limit, where the limit of detection was set to the smallest value in the data set. Multivariate data analysis such as principal component analysis (PCA), partial least squares-discriminant analysis (PLS-DA) and receiver operating characteristic (ROC) curves were performed using the online webserver, MetaboAnalyst 5.0 (Chong et al., 2018), where data sets were (generalized) *log*-transformed (glog) and auto-scaled (PCA, PLS-DA) unless otherwise stated. Univariate statistical analysis, including student's t-test and ANCOVA (between-subjects with adjustments for age, sex and BMI), and data normality test (Shapiro-Wilk test, $\alpha = 0.05$) were performed using the Statistical Package for Social Science (IBM SPSS Statistical for Windows, Version 20.0. NY, United States). The inter-method comparison between MSI-CE-MS and ^{1}H-NMR serum metabolite concentrations was performed using Bland-Altman %difference plots with MedCalc statistical software (MedCal® Version 12.5, Ostend, Belgium). All electropherograms, mass spectra and graphs were displayed using Igor Pro (Wavemetrics Inc, OR, United States) or Microsoft Excel (Redmond, WA, United States).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article are available in a public repository, MetaboLights (https://www.ebi.ac.uk/metabolights/).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by McMaster University Health Research Ethics Board (REB Project #0932). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MOS, AD-W, and MP conceived the study, recruited patients and acquired liver fibrosis measurements. MS, HP, and RP performed the experiments, MS and ZK analyzed the data, and MS and PB-M performed statistical analysis and wrote the manuscript draft. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by funding from the Natural Sciences and Engineering Research Council of Canada, the Canada Foundation for Innovation (PB-M, DW), Genome Canada (PB-M, DW), as well as the Department of Pathology and Molecular Medicine Residents Research Grant Fund (MOS, AD-W).

ACKNOWLEDGMENTS

The kind advice of Aalim Weljie in the study design (MOS), as well as the generous technical support of Hilary Jenkins and Bob Berno (PB-M) at the NMR facility in the Department of Chemistry and Chemical Biology at McMaster University are also acknowledged. The authors wish to thank all individuals who consented to participate in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.676349/full#supplementary-material

REFERENCES

Afzali, A., Weiss, N. S., Boyko, E. J., and Ioannou, G. N. (2010). Association between Serum Uric Acid Level and Chronic Liver Disease in the United States. *Hepatology* 52, 578–589. doi:10.1002/hep.23717

Azab, S. M., Zamzam, A., Syed, M. H., Abdin, R., Qadura, M., and Britz-McKibbin, P. (2020). Serum Metabolic Signatures of Chronic Limb-Threatening Ischemia in Patients with Peripheral Artery Disease. *J. Clin. Med.* 9, 1877. doi:10.3390/jcm9061877

Bachhawat, A. K., and Yadav, S. (2018). The Glutathione Cycle: Glutathione Metabolism beyond the γ-glutamyl Cycle. *IUBMB Life* 70, 585–592. doi:10.1002/iub.1756

Bhatia, A., Sarma, S. J., Lei, Z., and Sumner, L. W. (2019). UHPLC-QTOF-MS/MS-SPE-NMR: A Solution to the Metabolomics Grand challenge of Higher-Throughput, Confident Metabolite Identifications. *Methods Mol. Biol.* 2037, 113–133. doi:10.1007/978-1-4939-9690-2_7

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

Büscher, J. M., Czernik, D., Ewald, J. C., Sauer, U., and Zamboni, N. (2009). Cross-platform Comparison of Methods for Quantitative Metabolomics of Primary Metabolism. *Anal. Chem.* 81, 2135–2143. doi:10.1021/ac8022857

Casadei-Gardini, A., Del Coco, L., Marisi, G., Conti, F., Rovesti, G., Ulivi, P., et al. (2020). 1H-NMR Based Serum Metabolomics Highlights Different Specific Biomarkers between Early and Advanced Hepatocellular Carcinoma Stages. *Cancers* 12, 241. doi:10.3390/cancers12010241

Chalcraft, K. R., Lee, R., Mills, C., and Britz-McKibbin, P. (2009). Virtual Quantification of Metabolites by Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry: Predicting Ionization Efficiency without Chemical Standards. *Anal. Chem.* 81, 2506–2515. doi:10.1021/ac802272u

Chang, M.-L., and Yang, S.-S. (2019). Metabolic Signature of Hepatic Fibrosis: From Individual Pathways to Systems Biology. *Cells* 8, 1423. doi:10.3390/cells8111423

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

Cichoz-Lach, H., and Michalak, A. (2014). Oxidative Stress as a Crucial Factor in Liver Diseases. *World J. Gastroenterol.* 20, 8082–8091. doi:10.3748/wjg.v20.i25.8082

da Silva, R. R., Dorrestein, P. C., and Quinn, R. A. (2015). Illuminating the Dark Matter in Metabolomics. *Proc. Natl. Acad. Sci. USA* 112, 12549–12550. doi:10.1073/pnas.1516878112

DiBattista, A., McIntosh, N., Lamoureux, M., Al-Dirbashi, O. Y., Chakraborty, P., and Britz-McKibbin, P. (2017). Temporal Signal Pattern Recognition in Mass Spectrometry: A Method for Rapid Identification and Accurate Quantification of Biomarkers for Inborn Errors of Metabolism with Quality Assurance. *Anal. Chem.* 89, 8112–8121. doi:10.1021/acs.analchem.7b01727

Diren, B., and Idle, J. R. (2020). Metabolomic and Lipidomic Biomarkers for Premalignant Liver Disease Diagnosis and Therapy. *Metabolites* 10, 50. doi:10.3390/metabo10020050

Drouin, N., van Mever, M., Zhang, W., Tobolkina, E., Ferre, S., Servais, A.-C., et al. (2020). Capillary Electrophoresis-Mass Spectrometry at Trial by Metabo-Ring: Effective Electrophoretic Mobility for Reproducible and Robust Compound Annotation. *Anal. Chem.* 92, 14103–14112. doi:10.1021/acs.analchem.0c03129

Duncan, K. D., Fyrestam, J., and Lanekoff, I. (2019). Advances in Mass Spectrometry Based Single-Cell Metabolomics. *Analyst* 144, 782–793. doi:10.1039/c8an01581c

El-Batch, M., Ibrahim, W., and Said, S. (2011). Effect of Histidine on Autotaxin Activity in Experimentally Induced Liver Fibrosis. *J. Biochem. Mol. Toxicol.* 25, 143–150. doi:10.1002/jbt.20370

Emwas, A.-H., Roy, R., McKay, R. T., Tenori, L., Saccenti, E., Gowda, G. A. N., et al. (2019). NMR Spectroscopy for Metabolomics Research. *Metabolites* 9, 123. doi:10.3390/metabo9070123

Gamarra, Y., Santiago, F. C., Molina-López, J., Castaño, J., Herrera-Quintana, L., Domínguez, Á., et al. (2019). Pyroglutamic Acidosis by Glutathione Regeneration Blockage in Critical Patients with Septic Shock. *Crit. Care* 23, 162. doi:10.1186/s13054-019-2450-5

Gao, H., Lu, Q., Liu, X., Cong, H., Zhao, L., Wang, H., et al. (2009). Application of1H NMR-Based Metabonomics in the Study of Metabolic Profiling of Human Hepatocellular Carcinoma and Liver Cirrhosis. *Cancer Sci.* 100, 782–785. doi:10.1111/j.1349-7006.2009.01086.x

Ghany, M. G., and Morgan, T. R. (2020). AASLD-IDSA Hepatitis C Guidance Panel.Hepatitis C Guidance 2019 Update: American Association for the Study of Liver Diseases-Infectious Diseases Society of America Recommendations for Testing, Managing, and Treating Hepatitis C Virus Infection. *Hepatology* 71, 686–721. doi:10.1002/hep.31060

Goodman, Z. D. (2007). Grading and Staging Systems for Inflammation and Fibrosis in Chronic Liver Diseases. *J. Hepatol.* 47, 598–607. doi:10.1016/j.jhep.2007.07.006

Han, W., and Li, L. (2018). Chemical Isotope Labeling LC-MS for Human Blood Metabolome Analysis. *Methods Mol. Biol.* 1730, 213–225. doi:10.1007/978-1-4939-7592-1_14

Harada, S., Hirayama, A., Chan, Q., Kurihara, A., Fukai, K., Iida, M., et al. (2018). Reliability of Plasma Polar Metabolite Concentrations in a Large-Scale Cohort Study Using Capillary Electrophoresis-Mass Spectrometry. *PLoS One* 13, e0191230. doi:10.1371/journal.pone.0191230

Khatun, M., and Ray, R. B. (2019). Mechanisms Underlying Hepatitis C Virus-Associated Hepatic Fibrosis. *Cells* 8, 1249. doi:10.3390/cells8101249

Kuehnbaum, N. L., and Britz-McKibbin, P. (2013). New Advances in Separation Science for Metabolomics: Resolving Chemical Diversity in a Post-Genomic Era. *Chem. Rev.* 113, 2437–2468. doi:10.1021/cr300484s

Kuehnbaum, N. L., Kormendi, A., and Britz-McKibbin, P. (2013). Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: a High-Throughput Platform for Metabolomics with High Data Fidelity. *Anal. Chem.* 85, 10664–10669. doi:10.1021/ac403171u

Liss, D. B., Paden, M. S., Schwarz, E. S., and Mullins, M. E. (2013). What Is the Clinical Significance of 5-oxoproline (Pyroglutamic Acid) in High Anion gap Metabolic Acidosis Following Paracetamol (Acetaminophen) Exposure?. *Clin. Toxicol.* 51, 817–827. doi:10.3109/15563650.2013.844822

Marshall, D. D., and Powers, R. (2017). Beyond the Paradigm: Combining Mass Spectrometry and Nuclear Magnetic Resonance for Metabolomics. *Prog. Nucl. Magn. Reson. Spectrosc.* 100, 1–16. doi:10.1016/j.pnmrs.2017.01.001

McGill, D. B. (2001). Liver Biopsy: when, How, by Whom, and where?. *Curr. Gastroenterol. Rep.* 3, 19–23. doi:10.1007/s11894-001-0036-1

Mehedint, M. G., and Zeisel, S. H. (2013). Choline's Role in Maintaining Liver Function. *Curr. Opin. Clin. Nutr. Metab. Care* 16, 339–345. doi:10.1097/MCO.0b013e3283600d46

Meoni, G., Lorini, S., Monti, M., Madia, F., Corti, G., Luchinat, C., et al. (2019). The Metabolic Fingerprints of HCV and HBV Infections Studied by Nuclear Magnetic Resonance Spectroscopy. *Sci. Rep.* 9, 4128. doi:10.1038/s41598-019-40028-4

Myers, R. P., Krajden, M., Bilodeau, M., Kaita, K., Marotta, P., Peltekian, K., et al. (2014). Burden of Disease and Cost of Chronic Hepatitis C Virus Infection in Canada. *Can. J. Gastroenterol. Hepatol.* 28, 243–250. doi:10.1155/2014/317623

Naggie, S., Lusk, S., Thompson, J. W., Mock, M., Moylan, C., Lucas, J. E., et al. (2020). Metabolomic Signature as a Predictor of Liver Disease Events in Patients with HIV/HCV Coinfection. *J. Infect. Dis.* 222, 2012–2020. doi:10.1093/infdis/jiaa316

Naz, S., García, A., and Barbas, C. (2013). Multiplatform Analytical Methodology for Metabolic Fingerprinting of Lung Tissue. *Anal. Chem.* 85, 10941–10948. doi:10.1021/ac402411n

Nevedomskaya, E., Mayboroda, O. A., and Deelder, A. M. (2011). Cross-platform Analysis of Longitudinal Data in Metabolomics. *Mol. Biosyst.* 7, 3214–3222. doi:10.1039/c1mb05280b

Nori de Macedo, A., Mathiaparanam, S., Brick, L., Keenan, K., Gonska, T., Pedder, L., et al. (2017). The Sweat Metabolome of Screen-Positive Cystic Fibrosis Infants: Revealing Mechanisms beyond Impaired Chloride Transport. *ACS Cent. Sci.* 3, 904–913. doi:10.1021/acscentsci.7b00299

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., et al. (2011). The Human Serum Metabolome. *PLoS One* 6, e16957. doi:10.1371/journal.pone.0016957

Ravanbakhsh, S., Liu, P., Bjordahl, T. C., Mandal, R., Grant, J. R., Wilson, M., et al. (2015). Accurate, Fully-Automated NMR Spectral Profiling for Metabolomics. *PLoS One* 10, e0124219. doi:10.1371/journal.pone.0124219

Rhee, E. P., Waikar, S. S., Rebholz, C. M., Zheng, Z., Perichon, R., Clish, C. B., et al. (2019). Variability of Two Metabolomic Platforms in CKD. *Clin. J. Am. Soc. Nephrol.* 14, 40–48. doi:10.2215/CJN.07070618

Rojo, D., Canuto, G. A. B., Castilho-Martins, E. A., Tavares, M. F. M., 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, e0130675. doi:10.1371/journal.pone.0130675

Saoi, M., Li, A., McGlory, C., Stokes, T., von Allmen, M. T., Phillips, S. M., et al. (2019). Metabolic Perturbations from Step Reduction in Older Persons at Risk for Sarcopenia: Plasma Biomarkers of Abrupt Changes in Physical Activity. *Metabolites* 9, 134. doi:10.3390/metabo9070134

Saoi, M., Sasaki, K., Sagawa, H., Abe, K., Kogiso, T., Tokushige, K., et al. (2020). High Throughput Screening of Serum γ-Glutamyl Dipeptides for Risk Assessment of Nonalcoholic Steatohepatitis with Impaired Glutathione Salvage Pathway. *J. Proteome Res.* 19, 2689–2699. doi:10.1021/acs.jproteome.9b00405

Sarfaraz, M. O., Myers, R. P., Coffin, C. S., Gao, Z. H., Shaheen, A. A. M., Crotty, P. M., et al. (2016). A Quantitative Metabolomics Profiling Approach for the Noninvasive Assessment of Liver Histology in Patients with Chronic Hepatitis C. *Clin. Translational Med.* 5, 33. doi:10.1186/s40169-016-0109-2

Shanmuganathan, S., Kroezen, Z., Gill, B., Azab, S., de Souza, R. J., Teo, K. K., et al. (2021). The Maternal Serum Metabolome by Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A Standardized Data Workflow for Large-Scale Epidemiological Studies. *Nat. Prot.* 16, 1966–1994. doi:10.1038/s41596-020-00475-0

Soga, T., Sugimoto, M., Honma, M., Mori, M., Igarashi, K., Kashikura, K., et al. (2011). Serum Metabolomics Reveals γ-glutamyl Dipeptides as Biomarkers for Discrimination Among Different Forms of Liver Disease. *J. Hepatol.* 55, 896–905. doi:10.1016/j.jhep.2011.01.031

Wei, S., Suryani, Y., Gowda, G. A. N., Skill, N., Maluccio, M., and Raftery, D. (2012). Differentiating Hepatocellular Carcinoma from Hepatitis C Using Metabolite Profiling. *Metabolites* 2, 701–716. doi:10.3390/metabo2040701

Wild, J., Shanmuganathan, M., Hayashi, M., Potter, M., and Britz-McKibbin, P. (2019). Metabolomics for Improved Treatment Monitoring of Phenylketonuria: Urinary Biomarkers for Non-invasive Assessment of Dietary Adherence and Nutritional Deficiencies. *Analyst* 144, 6595–6608. doi:10.1039/c9an01642b

Wilder, J., and Patel, K. (2014). The Clinical Utility of FibroScan(®) as a Noninvasive Diagnostic Test for Liver Disease. *Med. Devices (Auckl)* 7, 107–114. doi:10.2147/MDER.S46943

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–D617. doi:10.1093/nar/gkx1089

Wishart, D. S. (2019). NMR Metabolomics: A Look Ahead. *J. Magn. Reson.* 306, 155–161. doi:10.1016/j.jmr.2019.07.013

Würtz, P., Kangas, A. J., Soininen, P., Lawlor, D. A., Davey Smith, G., and Ala-Korpela, M. (2017). Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -omic Technologies. *Am. J. Epidemiol.* 186, 1084–1096. doi:10.1093/aje/kwx016

Zhang, W., Hankemeier, T., and Ramautar, R. (2017). Next-generation Capillary Electrophoresis-Mass Spectrometry Approaches in Metabolomics. *Curr. Opin. Biotechnol.* 43, 1–7. doi:10.1016/j.copbio.2016.07.002

Zhou, Y., Wei, F., and Fan, Y. (2016). High Serum Uric Acid and Risk of Nonalcoholic Fatty Liver Disease: A Systematic Review and Meta-Analysis. *Clin. Biochem.* 49, 636–642. doi:10.1016/j.clinbiochem.2015.12.010

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.

frontiers in Molecular Biosciences

REVIEW
published: 20 September 2021
doi: 10.3389/fmolb.2021.698337

Nuclear Magnetic Resonance Spectroscopy in Clinical Metabolomics and Personalized Medicine: Current Challenges and Perspectives

Marine P. M. Letertre[1], *Patrick Giraudeau*[1] *and Pascal de Tullio*[2]*

[1]*Université de Nantes, CNRS, CEISAM UMR 6230, Nantes, France,* [2]*Metabolomics Group, Center for Interdisciplinary Research of Medicine (CIRM), Department of Pharmacy, Université de Liège, Liège, Belgique*

OPEN ACCESS

Edited by:
Serge Rudaz,
Université de Genève, Switzerland

Reviewed by:
Young Hae Choi,
Leiden University, Netherlands
Benedicte Elena-Herrmann,
INSERM U1209 Institut pour l'Avancée des Biosciences (IAB), France
Claudio Luchinat,
University of Florence, Italy

***Correspondence:**
Pascal de Tullio
P.detullio@uliege.be

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

Received: *21 April 2021*
Accepted: *30 August 2021*
Published: *20 September 2021*

Citation:
Letertre MPM, Giraudeau P and de Tullio P (2021) Nuclear Magnetic Resonance Spectroscopy in Clinical Metabolomics and Personalized Medicine: Current Challenges and Perspectives. Front. Mol. Biosci. 8:698337. doi: 10.3389/fmolb.2021.698337

Personalized medicine is probably the most promising area being developed in modern medicine. This approach attempts to optimize the therapies and the patient care based on the individual patient characteristics. Its success highly depends on the way the characterization of the disease and its evolution, the patient's classification, its follow-up and the treatment could be optimized. Thus, personalized medicine must combine innovative tools to measure, integrate and model data. Towards this goal, clinical metabolomics appears as ideally suited to obtain relevant information. Indeed, the metabolomics signature brings crucial insight to stratify patients according to their responses to a pathology and/or a treatment, to provide prognostic and diagnostic biomarkers, and to improve therapeutic outcomes. However, the translation of metabolomics from laboratory studies to clinical practice remains a subsequent challenge. Nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry (MS) are the two key platforms for the measurement of the metabolome. NMR has several advantages and features that are essential in clinical metabolomics. Indeed, NMR spectroscopy is inherently very robust, reproducible, unbiased, quantitative, informative at the structural molecular level, requires little sample preparation and reduced data processing. NMR is also well adapted to the measurement of large cohorts, to multi-sites and to longitudinal studies. This review focus on the potential of NMR in the context of clinical metabolomics and personalized medicine. Starting with the current status of NMR-based metabolomics at the clinical level and highlighting its strengths, weaknesses and challenges, this article also explores how, far from the initial "opposition" or "competition", NMR and MS have been integrated and have demonstrated a great complementarity, in terms of sample classification and biomarker identification. Finally, a perspective discussion provides insight into the current methodological developments that could significantly raise NMR as a more resolutive, sensitive and accessible tool for clinical applications and point-of-care diagnosis. Thanks to these advances, NMR has a strong potential to join the other analytical tools currently used in clinical settings.

Keywords: nuclear magnetic resonance, clinical metabolomics, personalized medicine, spectroscopy, biomarkers

CLINICAL METABOLOMICS AND PERSONALIZED MEDICINE

Amongst "omics" approaches, metabolomics is generally presented as the last that appeared in terms of occurrence and development, but also as the final biological and biochemical stones in the complex networks of organisms. Indeed, this approach aims at identify and quantify organic low molecular weight molecules (<1,500 Da) belonging to different classes of metabolites (Nicholson and Lindon, 2008; Patti et al., 2012). These metabolites form the metabolome, which is composed of endogenous but also exogenous biochemicals coming from environment, life-style, food, medicines, microbiome and which could be involved in catabolic and anabolic reactions (**Figure 1**). Metabolomics is clearly correlated with the functionality of the organism, while the other "omics" such as genomics, transcriptomics and proteomics, are closest to its capabilities (**Figure 1**). While the applications of metabolomics are numerous and varied in areas such as food and natural products quality controls (Lee et al., 2017; Li et al., 2021), environmental studies (Viant, 2009; Bedia et al., 2018) or agriculture (Kumar et al., 2017), the most highlighted and probably the most promising application fields of this methodology are clinical metabolomics and personalized medicine (Wishart, 2016; Li B. et al., 2017; Kohler et al., 2017; Nielsen, 2017; Tolstikov et al., 2017; Trivedi et al., 2017; Jacob et al., 2019; Pang et al., 2019). Clinical metabolomics is a general terminology that deals with all the applications of this approach that involve human subjects. It includes fundamental studies of diseases (Nielsen, 2017), searches for new biomarkers discovery (Kohler et al., 2017) as well as for new therapeutic targets and drug development processes (Powers, 2014; Cuperlovic-Culf and Culf, 2016; Frédérich et al., 2016), epidemiology (Moayyeri et al., 2013; Chan et al., 2017; Yu et al., 2019) and, recently, appears as an interesting tool in the development of personalized or precision medicine (Wishart, 2016; Li B. et al., 2017; Trivedi et al., 2017; Jacob et al., 2019). Indeed, medical care is continuously evolving toward a more patient-centered approach.

Personalized medicine is probably the most important paradigm change in medical care that occurred during the last few years and is clearly the future of modern medicine (Di Sanzo et al., 2017; Elemento, 2020). This approach attempts to optimize the therapies and the patient care based on the individual patient characteristics (*i.e.*, genetic dispositions, phenotype, life-style, environmental parameters ...) and is expected to improve treatment efficacy and the quality of life of the patients. The keystone of this approach is linked to the ability of the clinicians 1) to precisely characterize the disease, 2) to stratify the patients (*i.e.*, according to genotype but also to phenotype), 3) to select the right treatments adapted to the disease and the patient conditions and 4) to follow the pathology and the treatment outcomes and to predict their evolutions. Thus, the classical clinical tools currently used must be improve so personalized medicine can reach such ambitious objectives. Modern and innovative approaches are instrumental to propose more robust and trustworthy preventive/prognostic solutions, in order to measure, integrate and model informative data that could help clinicians to select the best option for the patient care. In this way, biomarker discovery and measurement for longitudinal patient follow-up appear as the cornerstones of personalized medicine (Kohler et al., 2017)

Given all these demands and needs, it is clear that clinical metabolomics should play an important role in changing the way patient care is approached. Patients' metabolic profiles are very dynamic and can be influenced by external or internal stimuli, lifestyle and clinical changes and can be used to monitor and explore cellular or tissue homeostasis as well as physiological and pathological conditions (Wishart, 2016; Tolstikov et al., 2017). A robust, quantitative and reproducible study of their metabolome is then essential to accurately define their phenotype (Jacob et al., 2019). Moreover, human pathologies are often complex, with multiple molecular pathogenesis and heterogeneous clinical pictures between patients and are not only driven by genetics

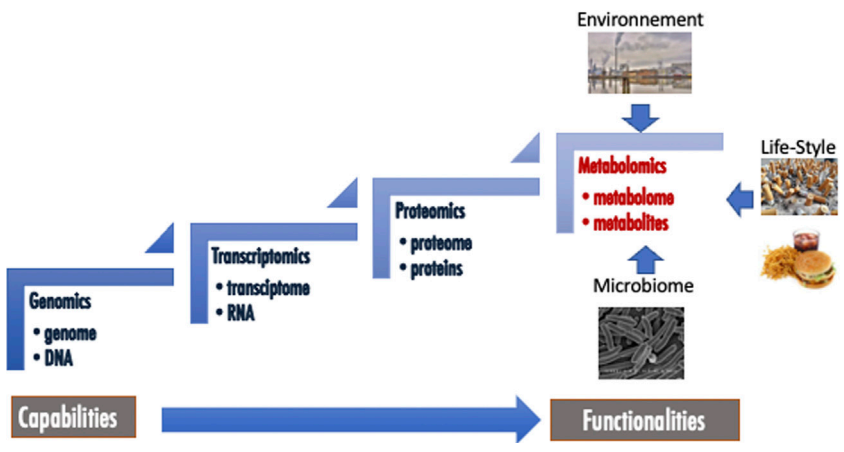

FIGURE 1 | Metabolomics in the field of other omics.

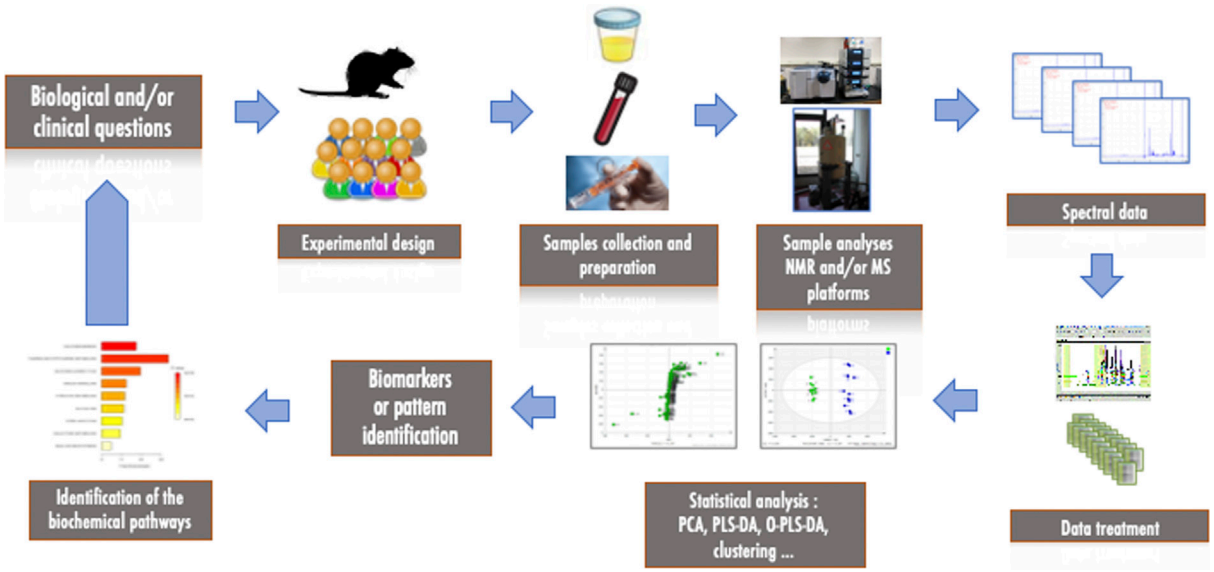

FIGURE 2 | Workflow of a metabolomics study.

but can also be strongly influenced by the function or dysfunction of different metabolic and biochemical networks. These networks are highly controlled by several internal but also external patient parameters, and therefore, the identification and integration of these parameters is essential for a better understanding of the mechanisms that led to the causality and the development of a pathology. Since it allows measuring the occurrence and variations of metabolites in organs, tissues and biofluids to be reported in a spatial and temporal manner, clinical metabolomics is to be an essential tool and will play a major role for the search for biomarkers, the identification of biochemical pathways involved in a pathology, the study of the environment and lifestyle influences and the treatment follow-up. However, there are still many obstacles and challenges to overcome in order to bring this approach from the laboratory to the clinical practice (Pinu et al., 2019).

The study of the metabolome and the monitoring of metabolites can be considered through two approaches: non-targeted approach and targeted approach. Non-targeted (or untargeted) metabolomics could be defined as the comprehensive and extensive measurement of a larger number of metabolites without a selection based on chemical class or biological activity. This approach is most commonly used for without *a priori* exploratory studies of pathologies and for the discovery of new specific biomarkers. Targeted or biology-driven metabolomics is the analysis of selected, chemically similar or groups of biochemically annotated metabolites such as known clinical biomarkers. It deals with quantitation or semi-quantitation of a set of known metabolites. It currently requires prior knowledge of the chemical or spectral properties of the metabolites of interest. This approach is used to study particular pathways, chemical families or biological activities and is mandatory for validating the metabolites identified by an untargeted strategy. Targeted metabolomics is particularly suited to the longitudinal studies and monitoring of patients and treatments that are essential in personalized medicine. Some classes of metabolites have led to the development of specific metabolomics fields. Lipids, which are considered an essential and crucial class of compounds, led to an "omics" approach named lipidomics, while sugars are studied in glycomics. Going one step further, fluxomics, which is the analysis of metabolic fluxes relying on labeled metabolic precursors, represents a very interesting approach for the in-depth study of the intracellular metabolism as well as the biochemical and metabolic pathways of an organism.

The classical workflow of a metabolomics study consists of several steps, as shown in **Figure 2**. The first step starts with the biological and/or clinical questions and leads to the experimental design, the choice of models and samples to be collected and analyzed (biofluids, biopsies, cells). The second important step is the measurement and analysis of the collected samples using high-throughput technological platforms. After the measurement and pre-processing of the data, statistical analyses are necessary to extract the most relevant information to interpret the results biologically and to identify metabolites or patterns that could be considered as biomarkers of the pathology of interest. Depending on the structure of the data, this usually requires reducing the size of the data via multivariate statistical analysis or applying classical univariate approaches. As these analyses deal with variance, all experimental and analytical variabilities must be minimized in order to reduce noise, avoid confounding factors and maximize response. Finally, the features that have been identified can be correlated to biochemical pathways and interpreted in the light of the original question and/or hypothesis. Metabolomics is thus a highly collaborative field that requires interaction between biologists, analytical chemists and statisticians.

The accurate and complete measurement of the metabolome is not an easy task at the analytical level (Kohler et al., 2017). Indeed,

the great diversity of metabolites both at the physico-chemical level, the broad range of concentrations (up to mmol/L at best for the most concentrated metabolites) as well as the associated dynamic range detection issues represent probably the most important and specific challenges for the classically used analytical methods. Moreover, the analysis of complex biological matrices is further hampered by the presence of proteins, high ionic strength and the sample heterogeneity. Therefore, pre-treatment of the samples is often necessary to reduce these problems and to adapt the samples to the analytical method. However, it is obvious that the more complex and time-consuming this treatment is, the higher the risk of altering the sample (in terms of metabolites composition) and the higher the risk of introducing experimental variability.

Although the first work identified as relating to metabolomics involved GC-MS in the 70th (Zlatkis and Liebich, 1971), Nuclear Magnetic Resonance (NMR) quickly appeared to be a powerful analytical technique in this field. This is mainly due to the fact that NMR is a highly robust, reproducible and non-destructive method that can be straightforwardly adapted to the analysis of complex media (Emwas et al., 2019; Wishart, 2019; Sahoo et al., 2020). Largely in minority before the 2000s, the use of mass spectrometry (MS), coupled with chromatographic techniques (Gas or liquid chromatographies- GC or LC) for metabolome analysis has also been developed (Gowda and Djukovic, 2014; Beale et al., 2018; Cui et al., 2018; Rampler et al., 2021). Indeed, the need to better understand and characterize the metabolome, as well as the advent of concepts such as biological systems and metabolic networks and the use of metabolomics in the discovery of disease-specific biomarkers, have made it necessary to increase the number of metabolites identified, especially those present in lower concentrations. In this context, mass spectrometry coupled with chromatographic techniques naturally appeared to be the most suitable analytical platform, thanks to its favorable limit of detection. This is mainly due to the rapid technological progress of this technique, the accessible cost of routine instrumentation as well as its impressive sensitivity compared to NMR. This have progressively reversed the situation to the point where at present the use of MS is applied in the majority of metabolomics studies. This technological evolution inevitably raises the question of the future of NMR in metabolomics and more particularly in clinical metabolomics. On the one hand, if one looks at the basic opposition of the two techniques by comparing their sensitivity and the number of metabolites detected, it's a done deal. On the other hand, if we look in more details at the situation, at the real needs in clinical metabolomics, and if we humbly remember that no single analytical technique can answer 100% of the questions, nor cover 100% of the needs, it is possible to consider these two approaches as perfectly complementary and equally essential, each with its advantages, covering the deficiencies of the other (this complementarity will be discussed in detail later in the article). The right tool(s) for the right biological question(s) must be the rule. Besides NMR and LC-MS (and to a lesser extent GC-MS), which are the two most widespread platforms, other approaches have also been or are being explored, such as vibrational spectroscopy (*i.e.*, FT-IR and Raman) (Du et al., 2020; Sherman et al., 2020) and capillary electrophoresis coupled or not with a MS detector (Maier and Schmitt-Kopplin, 2016; Sasaki et al., 2019).

There are many studies comparing the advantages and disadvantages of NMR and MS in the field of metabolomics and it is not our intention here to add one more (Frédérich et al., 2016; Kohler et al., 2016; Emwas et al., 2019; Wishart, 2019). Instead, we want to focus on what NMR can bring to clinical metabolomics and personalized medicine, how it can address the challenges of these fields and how its use provides them with a new opportunity and an added value (Markley et al., 2017; Emwas et al., 2019; Takis et al., 2019; Giraudeau, 2020). Thus, the three following chapters respectively highlight the current position of NMR in clinical metabolomics, the complementarity of NMR with MS and the recent and future developments of NMR in the same field.

NUCLEAR MAGNETIC RESONANCE IN CLINICAL METABOLOMICS AND PERSONALIZED MEDICINE

To fully understand the role that NMR can play in clinical metabolomics and personalized medicine, it is important to keep in mind the weaknesses and strengths of this technique in these particular applications. With this in mind, we will examine how this analytical approach has been used advantageously and successfully in a wide range of research and projects and how its weaknesses can be improved.

Initial Limitations of Nuclear Magnetic Resonance Approach

When discussing NMR and comparing it to other analytical techniques, particularly MS, its lack of sensitivity and resolution are often highlighted. In fact, despite significant technical improvements in recent years, the limit of detectable and quantifiable concentrations for hydrophobic metabolites in NMR is often in the micromolar range, - a few tens of micromolar at best on the typical metabolomics platforms. Moreover, the absence of separative techniques preceding the NMR analysis often leads to the overlapping of certain signals, which sometimes drastically reduces the resolution, especially in 1D-NMR. While multi-dimensional NMR techniques greatly improve the resolution (see *Recent and Future Developments in Nuclear Magnetic Resonance-Based Metabolomics*), sensitivity remains the main weak point of NMR. Knowing that many metabolites in biofluids have concentrations often close to or below the detection limit of NMR, it is obvious that this technique can only visualize a small part of the metabolome. However, this limitation should be counterbalanced, on the one hand, because the part of the metabolome that can be visualized and quantified by NMR is often of crucial importance, and on the other hand because reliable metabolomics analysis can be multiple and should not be boiled down to the detection of a maximum number of metabolites.

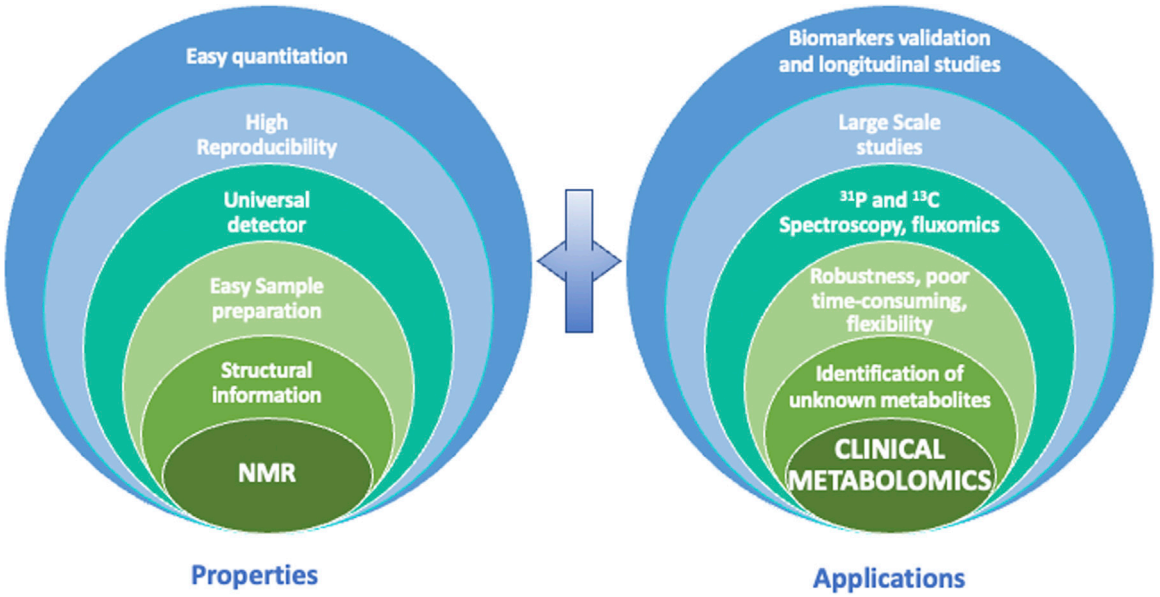

FIGURE 3 | Properties of NMR that allow specific advantages for clinical metabolomics.

Advantages and Specificities of the Nuclear Magnetic Resonance Approach and Its Applications

Although the limitations of NMR have been briefly stated, we must keep in mind some of its interesting characteristics to understand why NMR is an analytical technique of choice in the field of clinical metabolomics (**Figure 3**). First, NMR is highly reproducible, and it has intrinsic quantitative properties. Second, NMR is non-selective for analytes whose concentration is above the limit of detection, allowing almost universal detection for all organic molecules, depending of course of the sample preparation applied. Third, it provides crucial structural information owing to the high informative character of chemical shift and J-coupling information contained in NMR spectra. Fourth, NMR is non-destructive, which makes it possible to recover precious samples, and most importantly allows multiple 1D and 2D experiments on a single sample. As shown in **Figure 3**, all of these properties, which will be detailed above, have a positive impact on clinical metabolomics and enable many valuable NMR-based metabolomics studies.

Robustness and Reproducibility

The robustness and high reproducibility of NMR relies on the spectroscopic and physical measurement of samples. This capability is very important in the context of statistical data processing to minimize experimental variability and thus increase the sensitivity of the approach. Furthermore, with an adapted standardization of the acquisition parameters, NMR could also potentially allow the comparison and the integration of datasets from different instruments and/or performed at several sites. This potential will greatly facilitate biomarker validation. Indeed, the comparison of different datasets is one of the challenges that clinical metabolomics have to address in order to improve the quality and the robustness of the results and reach the standard required to enter into clinical practices. Moreover, coupled with automated samples preparator and changers, NMR, and more especially ^{1}H NMR, allows high throughput measurements of samples. Thus, this approach is currently the unique analytical platform that is adapted to large scale epidemiology but also to longitudinal studies as described in many recent publications (Jobard et al., 2017; Locci et al., 2018; Sliz et al., 2018; Welsh et al., 2018; Vignoli et al., 2019b; Debik et al., 2019; Deelen et al., 2019). For example, Vignoli et al. analyzed by NMR-based metabolomics the serum samples of 978 patients collected after an acute myocardial infarction and that were clinically followed during 2 years. The aim of this study was to explore if metabolomics profiles of patients could be correlated with a higher death risk and could enhance the existing prognostic risk models of death. Authors demonstrated on both training and validation sets that metabolomics data were relevant to identify high risk patients and that combination of these data with existing scoring methods was able to improve risk classification (Vignoli et al., 2019b). In a same epidemiologic approach, Deelen & al. measure with a standardized high-throughput NMR-based procedure, the metabolome signature of more than 40,000 individuals selected in several European cohorts. The aim to this study was to identify metabolites that could predict long term mortality. Using a stepwise procedure, they identified 14 circulating biomarkers that are independently associated with all-cause mortality. These markers could improve the existing score based on conventional risk factors and could potentially be used to help clinicians to define individual strategy for at risk patients (Deelen

et al., 2019). In another recent publication, urine NMR metabolomics has been identified as an interesting pipeline for large-scale epidemiology studies. This study demonstrated that it was possible to quantify 43 metabolites and to assigned more than 100 metabolites using a semi-automated methodology in a 1,004 individuals' cohort. Intra-assays measurement of metabolites concentrations highlighted that with a coefficient of variation (CV%) less than 5%, urine NMR could provide highly robust and accurate results. However, the authors also reported that, as expected, the intra-individuals' variations in the metabolites over 30 days, as well as inter-individuals' variations are very high (respectively CV > 20% and >40% for most of the metabolites). They conclude that high throughput urine NMR-based metabolomics could be an interesting and new base for epidemiologic and genetic applications (Tynkkynen et al., 2019). At the longitudinal level, Jobard et al. were able to follow-up by NMR-based metabolomics two types of treatments of HER-2 positive breast cancer patients (79 individuals) during 13 weeks (6 time-points). With this approach, the authors identified which treatment led to the most relevant impact on the patient's metabolism and highlighted that this effect is still observable several weeks after the end of the therapy. This work demonstrated that metabolomics could be used to predict clinical response or toxicity and tailored the treatment to patients (Jobard et al., 2017).

Detection and Quantitation

NMR is often presented as a universal detector. Indeed, any organic molecule with carbon, phosphorus, nitrogen or protons present in a solution will give a specific NMR signal or signals. Of course, for reasons of sensitivity and chemical and physical characteristics, compounds with one or more protons are the most easily detectable. Thus, within the detection limits of the system, NMR can visualize all the molecules present in a sample, without matrix or ionization effects that could affect the signal of certain compounds, and depending on the sample preparation applied. This property makes it possible to quickly visualize all the samples in a cohort and to easily identify outliers and possible subgroups. This is crucial for a better understanding of the structure of cohorts or groups. Moreover, this "universal" detection can be correlated with one of the most important properties of NMR, namely that it could be intrinsically quantitative. Indeed, not only signal intensity is directly proportional to the concentration of a molecule (taking into account the number of nuclei in the molecule), but if qNMR conditions are ensured, such as full relaxation, sufficient signal-to-noise and proper reference signal, the coefficient of proportionality is the same for all peaks, making it possible to quantify multiple analytes with a single internal or external reference (Holzgrabe et al., 2005). Hence, under controlled spectral conditions, NMR is one of the few techniques that allows quantification without the need for reference compounds or calibration curves. It is probably this feature that makes NMR unique and a tool of choice for clinical metabolomics. It is indeed clear that in the context of personalized medicine, the longitudinal aspect of the metabolomics studies is an essential point especially for patient follow-up and treatment evaluation. Comparison of metabolic profiles over time is not possible without a solid baseline and robust values. Moreover, no biomarkers could be useful without quantitation (Wishart, 2016). For multi-omics integration, as well as for translational purposes and clinical applications, absolute quantification appears as a keystone and a requirement of metabolomics studies (Wishart, 2016; Pinu et al., 2019). Depending on the biofluid or tissue examined, several protocols, recommendations and commercial solutions (such as Bruker IVDr NMR platform) have been proposed that allow quantification of 50–150 metabolites in one experiment through 1D NMR spectroscopy and in a range of concentrations from µM to mM and with a huge reproducibility (Nagana Gowda et al., 2015a; Emwas et al., 2016; Wallmeier et al., 2017; Jiménez et al., 2018; Amiel et al., 2019). This number may appear relatively small compared to the hundreds of metabolites that can be identified in MS, but it should be borne in mind that we are talking about quantification and not just identification. According to the importance of the field, commercial softwares, algorithms and workflows have been recently developed to facilitate automated or high throughput NMR quantification (*i.e.*, Batman, Bayesil, Speaq 2.0, SasMeQ, AQuA, SigMa, ChenomX) (Hao et al., 2012; Ravanbakhsh et al., 2015; Jung et al., 2016; Verhoeven et al., 2017; Beirnaert et al., 2018; Röhnisch et al., 2018; Khakimov et al., 2020). Some recent publications have demonstrated the interest of quantification in clinical metabolomics studies. For example, in a targeted approach, 27 metabolites (mainly amino acids) serum concentrations have been measured in a cohort of 157 smoker's patients with and without chronic obstructive pulmonary disease (COPD) and have highlighted that reduced amino acid concentrations could be associated with an increase incidence of respiratory exacerbation (Labaki et al., 2019). In a longitudinal study of ischemia reperfusion injury in adult cardiac surgery, NMR-based metabolomics on serum was used to follow-up patients up to 20 h post-operatively. 57 metabolites were quantified to get a longitudinal dataset that allow the exploration of the time-dependent alterations related to surgical trauma (Maltesen et al., 2020). NMR was also used to quantify short chain fatty acid (SCFA) in patients stools and demonstrated good correlations between high levels of SCFA, hypertension and on non-dipping blood pressure profile. This study highlighted the capability of NMR to easily quantify metabolites in stools and to also understand the impact of microbiota on human disease (Huart et al., 2019; Huart et al., 2021a).

Metabolite Identification

The development of two-dimensional approaches (COSY, TOCSY, HSQC, *etc*), coupled with methodological advances, should make it possible to increase the number of metabolites that can be detected and quantified (Féraud et al., 2020; Martineau et al., 2020). These developments are further discussed in *Recent and Future Developments in Nuclear Magnetic Resonance-Based Metabolomics*. Interestingly, NMR could also be used to guide MS quantification demonstrating

TABLE 1 | The main free access NMR databases.

Databases	NMR data	Number of metabolites	Automation search	Quant[a]	Others
HMDB	1D and 2D spectra, experimental and raw data, chemical shifts list	>100,000 (not all with measured NMR spectra)	1D and 2D search according to chemical shifts	No	Multiple information about the metabolites (concentrations in different fluids, physical and biological properties, enzymes and transporters, metabolic pathways . . .) Very complete
BML-NMR	1D ^{1}H and 2D J-resolved spectra, experimental and raw data, different spectral conditions	208	Possible with J-resolved data	Possible with J-resolved data	Many experimental conditions are proposed (experiment type, water suppression, buffer, excitation angle and relaxation delay)
BMRB	1D (proton and carbon) and 2D spectra, experimental and raw data, chemical shifts list	Repository for data Huge number of metabolites	1D and 2D search according to chemical shifts, to mass, to solvent and field strength	No	Many 1D and 2D spectral types, 3D structure. Not limited to metabolites
Metabolight	Mainly 1D proton spectra, spectral experimental conditions and raw data	Repository for data Huge number of metabolites	According to name	No	Database containing many metabolomics datasets

[a]*Quantification.*

the good complementarity between these two analytical platforms (see *Combining Nuclear Magnetic Resonance to Mass Spectrometry in Clinical Metabolomics*). Moreover, NMR is known to be the technique that provides the most structural information to characterize organic molecules. This can be extremely important for the identification of known or still unknown or undescribed metabolites (Dona et al., 2016; Wang et al., 2019). Accurate identification of metabolites is obviously essential at different levels for the accuracy and the relevance of the all the metabolomics studies. During the last decade, the development of metabolites databases and automated comparison tools increased the possibility to assign metabolites by comparing the NMR data between samples and reference spectra. We can cite HMDB (https://hmdb.ca), which is probably the most complete in terms of metabolites, BNL-NMR database (http://www.bml-nmr.org), BMRB (https://bmrb.io/metabolomics/), Metabolight (https://www.ebi.ac.uk/metabolights/) and some commercial software and platforms that allow the identification and quantification of metabolites (*i.e.*, Bruker IVDr platform and ChenomX software©). Many of these databases are interconnected to increase their potency and include more or less complex identification and search automated systems. **Table 1** describes the main characteristics of the open access and free databases. Obviously, these spectral databases as well as the comparison, identification and quantification tools still need to be improved by adding new metabolites spectra and data and by the development of more powerful algorithms for the automation of metabolite identification and quantification.

Multi-Nuclei Detection

Even if ^{1}H is the most frequently detected atom in NMR-based metabolomics, ^{31}P and especially ^{13}C can also be used in metabolomics. Even if the sensitivity of ^{31}P NMR is less than that of proton, it remains extremely interesting to explore. Indeed, phosphorus is an essential element in the biochemistry of the organisms. Phosphorylation or dephosphorylation of enzymes and proteins via kinases or phosphotransferases plays a key role in many processes, while certain phosphorylated metabolites and enzymes (NAD, NADH, NADP, UTP, CTP, ATP, ADP, AMP, *etc.*) are the mainstays of the energy machinery of cells. Moreover, phosphometabolites would represent more than 30% of the metabolites identified (Mazurek et al., 1997). Phosphorus NMR is still under-exploited at the moment, but because of its specificity, it represents a unique tool that is very interesting to develop for metabolomic applications (Bhinderwala et al., 2020). At the organic level, carbon is undoubtedly the most interesting element to examine, especially in NMR because it is present in all the molecules of interest and its chemical shift range is much wider than that of the proton, which considerably increases its resolution and facilitate the identification of the biomarkers. Unfortunately, ^{13}C NMR suffers from a low sensitivity, owing to the low natural abundance and gyromagnetic ratio of ^{13}C nuclei. Moreover, quantification is not as straightforward as with ^{1}H NMR. Therefore, direct carbon measurement is hardly ever applied in the field of clinical metabolomics. Still, ^{13}C spectroscopic information can be obtained with enhanced sensitivity via two-dimensional heteronuclear correlation experiments with inverse detection such as HSQC and HMBC. We have also to mention that ^{13}C and sometime ^{15}N observations are also very important in fluxomics which aims to quantify fluxes of metabolic reactions and is extremely important in *in vivo* and *in vitro* fundamental studies of biological systems (Crown and Antoniewicz, 2013; Niedenführ et al., 2015; Millard et al., 2017; Giraudeau, 2020). For example, by using labeled substrate (*i.e.*, ^{13}C labelled glucose or ^{15}N glutamine), this approach allows to follow the evolution of selected biochemical pathways by observing the labeled metabolites that are formed over time. NMR is particularly well suited for monitoring and quantifying the precursors and products of these biochemical pathways. It also allows to easily map the location of stable isotopes and to determine the incorporation points of markers in metabolites (Massou et al., 2007; Lane and Fan, 2017).

Sample Preparation

The nature of biological samples analyzed in clinical metabolomics is often complex and sometimes not directly compatible with analytical techniques, especially NMR and MS. This leads to the need to adapt pre-analytical protocols, for example by precipitating proteins, which considerably increases the complexity of the analyses by introducing a significant risk of variability. However, in NMR, it has been possible to develop spectral techniques that limit the pre-analytical steps and thus the manipulation of the samples (*e.g.*, CPMG pulse sequence to suppress protein signals, pre-saturation pulse sequences to suppress the water signal). At this level, NMR is therefore less time consuming and less likely to introduce undesirable experimental variability (Beckonert et al., 2007; Emwas et al., 2016; Vignoli et al., 2019a; Giskeødegård et al., 2019; Snytnikova et al., 2019). The absence of chromatographic techniques gives more flexibility to this approach and allows the rapid analysis of classical biofluids (blood, urine, saliva, cerebrospinal fluid) but also of various samples types (*e.g.*, biopsies, cells, feces, bronchoalveolar lavage fluid) (Ciaramelli et al., 2017; Kim et al., 2018; Romano et al., 2018; Albrecht et al., 2020; Duarte et al., 2020). Due to the nature of the analytical platforms used, most metabolomics experiments require the use of liquid samples. For solid samples such as cells or biopsies, this necessarily involves the insertion of a lysis step in the sample preparation process (Beckonert et al., 2007; Matheus et al., 2014; Kostidis et al., 2017; Mili et al., 2020). This step is sometimes difficult to implement and can lead to a lack of reproducibility and a loss of time. Direct observation of solid or semi-solid samples would limit these drawbacks. The use of high-resolution magic angle spinning (HRMAS) NMR spectroscopy enables the measurement of metabolites in intact tissue or cells and the detection of few dozens of compounds (Gogiashvili et al., 2019; Ruhland et al., 2019; Tilgner et al., 2019). Even if the resolution is lower than in classical high-resolution liquid NMR, this unique application can be particularly interesting at the clinical level, especially as a rapid diagnostic tool (10–15 min) for the analysis of biopsies. Moreover, recent progress in the miniaturization of such approach make it possible to limit its invasive character, hence promising application perspectives in clinics (Lucas-Torres et al., 2021).

Nuclear Magnetic Resonance-Based Lipidomics

Lipids are a large group of biomolecules, classified, due to their molecular weight, among the metabolites. Different subgroups are often distinguished including fatty acids, glycerolipids, phospholipids, sterols and more specifically, ceramides, sphingolipids, acyl-carnitines, lipoproteins. They play a key role in many biological processes since they can act as energy reservoir, signal molecules, protein traffickers and of course main constituents of plasma membranes (Gross and Han, 2011). Many diseases (such as cancer, diabetes, cardiovascular diseases) and pathological conditions are often accompanied by lipid dysregulation (Lydic and Goo, 2018; Guo et al., 2020). Initially included in metabolomics, the importance of this field, linked to the physicochemical specificities of these compounds, to their very huge number and to their essential biological role quickly led to the appearance of a distinct approach called lipidomics (Dennis, 2009). The analysis of lipids can face several critical issues: 1) the complexity of the samples, the huge number of compounds and their broad diversity of concentrations, 2) their nature and the high number of isomers and isobaric lipids and 3) their physicochemical properties. Mass spectrometry, generally coupled with liquid or gas separative techniques, is currently the analytical technique of choice for lipidome analysis, especially since the development of devices providing additional separation via ion mobility (Paglia et al., 2015; Jurowski et al., 2017; Leaptrot et al., 2019). Due to its lack of sensitivity and resolution, the use of NMR for lipidomics has been limited for a long time to fundamental studies such as 1) determining the structure of lipids of biological interest, 2) studying the structure and the composition of plasma membranes using ^{13}C-labeled precursors and ^{31}P NMR, 3) monitoring the impact of pathological conditions on lipid metabolism. More recently, NMR demonstrated its ability to be useful in classical lipidomic analyses and to be a very interesting and complementary tool to MS (Li J. et al., 2017; Gil et al., 2019). For example, ^{31}P spectroscopy could be chosen to monitor selectively and quantitatively phospholipid classes. Moreover, even if it does not allow, like MS, to finely separate lipids, proton NMR can nevertheless be used to carry out quantitative studies of lipids belonging to the different major classes. Several studies have thus demonstrated the interest of this approach in clinical lipidomics (Ouldamer et al., 2016; Curtarello et al., 2019; Bruzzone et al., 2020; Huart et al., 2021b) and specific workflows and tools (i.e., Lipspin) have been developed for semi-automated profiling (Barrilero et al., 2018; Amiel et al., 2019; Johnson et al., 2021). Recent developments in two-dimensional NMR also open new perspectives to increase the resolution and the identification of lipids (Marchand et al., 2018; Wang et al., 2020). Another specific aspect related to NMR is its ability to profile lipoproteins in blood (Soininen et al., 2015; Kostara et al., 2017; Jiménez et al., 2018). Lipoproteins are supramolecular lipid transporters classified by density ranging from Very Low-Density Lipoproteins (VLDL) to chylomicrons. These particles and more specifically their profile of distribution between the different subclasses is particularly important to measure in different pathological status such as cardiovascular diseases, metabolic syndrome neuropathologies and degenerative diseases (Catapano et al., 2016; Lambert et al., 2020). As described in different papers of the literature, ^{1}H NMR is probably the most adapted methodology to obtain fine quantitative profiles of lipoproteins (Jiménez et al., 2018).

Nuclear Magnetic Resonance Metabolomics and Personalized Medicine

Personalized medicine is a new paradigm in patient care and thus is still under development. It will be based on a better characterization of the patient, his physiological state and his response to treatment. Because it allows the discovery of biomarkers and the stratification of patients, metabolomics, as

well as pharmacometabolomics which evaluates their response to treatment, are and will certainly be part, with genomics, transcriptomics and proteomics of the key tools in this new way of approaching pathologies (Wishart, 2016; Jacob et al., 2019; Beger et al., 2020; Ashrafian et al., 2021). Its characteristics of reproducibility, robustness, speed and its ability to quantify metabolites, certainly position NMR as an analytical technique of choice in metabolomics-based personalized medicine, especially because longitudinal aspect is essential (Everett, 2017; Jacob et al., 2019). As mentioned previously, clinical metabolomics and personalized medicine approach are clearly connected. Therefore, most of the recent studies in clinical metabolomics are clearly oriented towards this personalization. Oncology is certainly the field of application where personalized medicine using "omics" is the most advanced (Yu and Snyder, 2016). Indeed, a fine classification of patients, a better evaluation of the efficacy of treatments and a clear vision of prognoses are essential for effective patient management. Thus, several recent studies and papers have demonstrated that NMR-based metabolomics can be effectively applied to precision oncology (Palmnas and Vogel, 2013; Hu et al., 2021; Vignoli et al., 2021).

Fingerprinting Approach and Application in Clinical Biology

Among all the possible approaches developed in metabolomics, fingerprinting is probably the one that has been applied first. It refers to the non-targeted and without a priori metabolomics studies that led to the identification of specific spectral or chemical patterns that could be related to a pathological or a particular status without identification of the metabolites (Kosmides et al., 2013). This approach is obviously not incompatible with the identification and quantification of biomarkers, but it is focused on a more holistic view of the metabolome and its possible transformation over time or under the effect of a pathology. The possible diagnostic application of fingerprinting is immediately obvious. However, it is equally obvious that such an application inevitably implies reproducibility, robustness, standardization of analytical methods and capability for high throughput analyses as it is the case in clinical biology. These are precisely among the strengths of NMR. As previously mentioned, NMR is indeed particularly well adapted to the study of large cohorts and thus to fingerprinting (Amathieu et al., 2014; Rzeznik et al., 2017; Takis et al., 2019). This approach faces several challenges and requires obviously the development of specific workflows and methodologies, especially for the multivariate analyses of the raw data (Zacharias et al., 2018; Markley et al., 2019). Besides the diagnostic models that metabolomic fingerprints can generate, it could also be very useful in a preventive framework, essential in the personalized approach to treatment. Indeed, regular observation of the metabolomic profile of patients would undoubtedly allow early identification of deviations that could be linked to the onset of certain pathologies (Takis et al., 2019).

Another important question concerning NMR-based metabolomics is its interest and its capacity to become one of the tools used in clinical practice and its interest compared to the techniques used until now in clinical biology. The numerous examples found in the literature demonstrate the strong potential of metabolomics in clinics, but the transition from laboratory to clinical practice is still a major challenge (Pinu et al., 2019). It can only be filled by the transition of metabolomics to the standards of quality, robustness and reproducibility required in clinical biology and NMR certainly has an important role to play (Ashrafian et al., 2021). It is clear that one of the first clinical applications of metabolomics remains the discovery of new biomarkers but that this approach, because of its holistic aspect, can bring much more to the understanding of pathologies, the prediction of their evolution, the stratification of patients and the evaluation and adaptation of treatments. Far from competing with existing tools, clinical metabolomics, once it masters and standardizes its analysis and data processing protocols, will undoubtedly provide essential information for improving patient care. NMR, thanks to its analytical qualities, its robustness and its ease of automation is undoubtedly a technological platform that will find its place among other instruments capable of providing clinicians with the data necessary for diagnosis and monitoring of patients.

For the sake of completeness, we should also highlight *in vivo* Magnetic Resonance Spectroscopy (MRS) which consists in localized NMR spectroscopic acquisitions performed within Magnetic Resonance Imaging (MRI) systems. It is the only technique that allows *in vivo* investigation of the human metabolome. (Rhodes, 2017; van de Weijer and Schrauwen-Hinderling, 2019). In this review, we have focused on classical high-resolution NMR spectroscopy without detailing *in vivo* or imaging applications, which are vast areas of clinical interest and would deserve a dedicated review. We will describe in the following sections how NMR limitations have been or will be challenged, what are the methodological and technological evolutions that will allow NMR to evolve in the near future and to remain a powerful analytical platform in metabolomics. We will also examine how this technique can be advantageously combined with other analytical approaches and why this complementarity may represent a solution for improving our knowledge and exploration of the metabolome.

COMBINING NUCLEAR MAGNETIC RESONANCE TO MASS SPECTROMETRY IN CLINICAL METABOLOMICS

The previous section demonstrated the strong potential that NMR spectroscopy has within the field of clinical metabolomics. However, it is well-known that NMR-based metabolomics has some drawbacks, namely its lack of sensitivity and the non-negligible signal overlap in routine 1D ^{1}H experiments of complex biological samples. This limits its application within several fields, including personalized medicine. Indeed, signal overlap makes the difficult task of metabolite identification and the subsequent biomarker discovery even more difficult. As such and as explained

previously, mass spectrometry based metabolomics became more popular than NMR spectroscopy in a vast majority of metabolomics applications (Letertre et al., 2021). However, MS techniques come with their set of drawbacks as well: lack of robustness and repeatability, and the difficulty to identify the biomarkers corresponding to the numerous features detected in MS spectra. These drawbacks are not to be ignored in clinical research, as the discovery of biomarkers of a given pathology, or biomarkers of an exposition to a therapeutic treatment, request robust and repeatable methods for intra and inter-laboratory comparison. To overcome the respective drawbacks of each techniques and to combine their strengths, the complementarity between both NMR spectroscopy and MS-based metabolomics techniques has been discussed several times in the past 15 years (Pan and Raftery, 2007; Marshall and Powers, 2017; Letertre et al., 2021). In the following section, different examples of studies combining both NMR and MS-based metabolomics applied in clinical settings are presented, and their advantages and drawbacks are discussed.

Nuclear Magnetic Resonance Hardware Hyphenation to Mass Spectrometry Hardware

Combining NMR with liquid chromatography (LC-NMR) and further with MS (LC-NMR-MS) through hardware hyphenation has been long done, especially in natural products analysis and the different ways of doing it have been nicely described recently (Gebretsadik et al., 2019). This approach has also been found useful for drug metabolism (Shockcor, 2002) and pharmaceutical research (Lindon et al., 2000; Lindon et al., 2002). For instance, the combination HPLC-NMR with an ion-trap MS made the identification of paracetamol metabolites and endogenous compounds in human urine possible (Shockcor et al., 1996). By successfully detecting phenylacetylglutamine, which was not possible by using only ^{1}H NMR, this triple-hyphenated system overcame the NMR signal overlap issue. One the other hand, the NMR part of the system was an essential tool to determine which isomers of the paracetamol-glucuronide conjugate was present in the sample (Shockcor et al., 1996). A similar investigation applied this system to characterize ibuprofen metabolism in human urine (Clayton et al., 1998). LC-NMR-MS was also used in parallel to ^{19}F NMR spectroscopy to investigate the metabolism of fluorinated novel drug candidates (Dear et al., 1998) or drug intermediates (Scarfe et al., 1999; Scarfe et al., 1998) within urine samples of animal models and without requesting specific radiolabeling. However, the community has lost interest in LC-NMR-MS since the past decade, most certainly due to the technical difficulties encountered by combining techniques coming with orthogonal analytical requirements (Silva Elipe, 2003).

Combining Nuclear Magnetic Resonance and Mass Spectrometry Datasets

To Increase Metabolic Coverage

Rather than hyphenating their respective hardware, combining the datasets of NMR and MS-based metabolomics workflows has had more success. The most obvious reason to use both NMR and MS-based metabolomics is to increase the metabolic coverage, thus increasing the chance of identifying new biomarkers. Indeed, it is well emphasized within the metabolomics community that no tools whatsoever offer a full coverage of the metabolic landscape. Several studies nicely supported this assessment by using a Venn diagram, which shows how the metabolite identification overlays between the different platforms used. One of the most famous example is a study of Human Serum Metabolome (Psychogios et al., 2011). By using five analytical platforms (NMR spectroscopy, LC-ESI-MS/MS, GC-MS, DFI-MS and TLC-GC-FID), the authors were able to identify 3,764 compounds, from which only 200 were commonly detected by at least two platforms. Furthermore, this effort was completed by quantitative data for some of the detected metabolites, and although some of the results differed between platforms, good agreement overall were found (Psychogios et al., 2011). The combination of GC-MS, LC-MS and NMR was also applied to explore a NIST standard reference material for human plasma and its application in clinical laboratories. A total of 353 metabolites were identified, and whilst GC-MS was the analytical technique showing the most of unique identification (65), and that LC-MS and NMR identifications were found to overlap, NMR still allowed to detect small sugars which were not directly accessible by LC-MS (Simón-Manso et al., 2013). Similar to the Human Serum Metabolome, a study focused on the mouse skeletal muscle metabolome by combining NMR, FIA-MS, GC-MS and LC-HRMS, highlighted 132 discriminant metabolites, from which only 17 were detected by more than one analytical platform (Bruno et al., 2018). Importantly, the analytical approach proposed in this article was aimed to be easily adapted for human clinical trials. In a final example, the effect of therapeutic treatment on human gastric cancer cells was assessed by metabolomics and lipidomics by using three analytical platforms (NMR spectroscopy, GC-MS and LC-MS). Once again, out of the 111 compounds detected, only 21 were commonly highlighted by at least two analytical techniques (*Goulitquer et al., 2018*). This proves the importance of using multiple platforms if the aim of a given study is to capture the metabolome as broadly as possible, or to carefully chose the appropriate platform if only a specific subpart of the metabolome is of interest, as the different requirements in term of sample preparation as well as the very essence of the analytical platform selected will give access only to a limited part of the metabolome.

To Correlate Variables Detected by Nuclear Magnetic Resonance and Mass Spectrometry Techniques

Another way to show the limited overlap that can be observed between NMR and MS datasets is to correlate their respective signals, as it was done in a study focusing on colorectal cancer and polyps serum samples that were analyzed by NMR spectroscopy and targeted LC-MS/MS (Deng et al., 2016). It is clear that in **Figure 4B**, the correlation taking into consideration only the variables that were commonly detected by both NMR and MS techniques represented only a small subset of all the features detected either by NMR or MS (**Figure 4A**). However, correlating the intrinsic covariance of signals detected by each of the

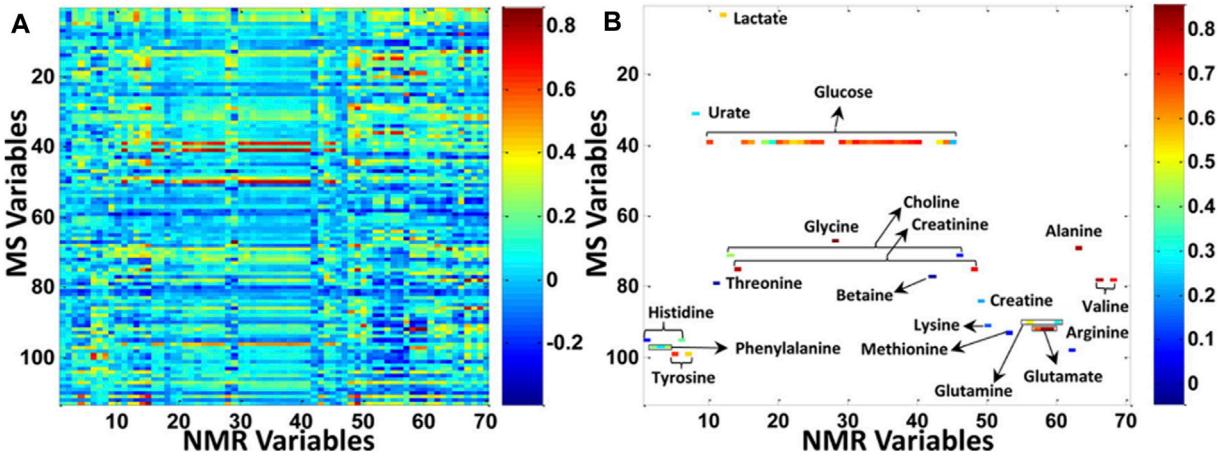

FIGURE 4 | Increasing the metabolic coverage of serum samples from patients suffering from colorectal cancer and polyps, by using both NMR and MS analytical platforms. **(A)** Correlation between all NMR and MS variables. **(B)** Correlation between the subset of metabolites (labeled in the figure) that can be detected by both NMR and MS. The *X* axis provides an index of all NMR variables in the data matrix, and the *Y* axis provides an index of all MS variables in the data matrix. Reproduced with permission from (Deng et al., 2016).

analytical techniques can also serve as a tool to confirm metabolite annotation made by one analytical technique, or to acquire deeper knowledge on biomolecular reactions and thus to enhance biomarker discovery. The first tool based on this methodology was the so-called SHY (statistical heterospectroscopy), based on a Pearson correlation method. Crockford *et al.* showed positive or negative correlations between NMR with LC-MS signals measured within urine samples of rats treated with hydrazine as a proof-of-concept (Crockford et al., 2006). This approach was further applied to human urine samples, for instance to highlight biological processes of inborn errors of metabolism by correlating NMR and DESI-MS signals (Pan et al., 2007) or to successfully investigate the xenometabolome of a random subset of an epidemiological study (Crockford et al., 2008). In this last study, new drug metabolites were discovered thanks to the correlation between NMR and MS signals but also to the use of MS tools to investigate ion fragmentations, such as MS^E and different collision energies (Crockford et al., 2008).

To Improve Statistical Models Through Multi-Block Data Integration

The second advantage of combining NMR and MS-based metabolomics datasets is to produce more robust and trustworthy multivariate statistical models. To do so, multi-block data fusion, or data integration, is gaining in popularity within the metabolomics community (Doeswijk et al., 2011; Boccard and Rudaz, 2014). Three levels are available to apply data integration, often referred as low-, mid- and high-levels. The difference between these levels have been clearly explained previously (Boccard and Rudaz, 2014). Briefly, low-level data fusion consists in taking the matrices obtained by each of the analytical technique as they are, without performing multivariate statistics on the individual blocks beforehand. Mid-level consists in reducing the individual matrices before integration (*e.g.*, by selecting the most discriminant variables), and high-level data fusion integrates only the global outputs of the individual statistical models. Once the fusion matrix has been produced, chemometrics can be applied, in a very similar way as for the individual matrix, by using unsupervised and supervised statistical analysis (Boccard and Rudaz, 2014).

In a first example which aimed at determining metabolic differences between serum samples from breast cancer patients and healthy controls, a mid-level data fusion approach was used to enhance the discriminative performance of unsupervised analyses and limit the misclassification of the supervised analyses performed on the individuals NMR and the direct analysis in real time (DART-MS) models (Gu et al., 2011). In that end, another supervised analysis was performed by setting up the Y variable to the first component of the unsupervised NMR model, which performed better than the MS model, and the X matric was set as the DART-MS dataset, containing more variables. The resulting model performed better in term of both discriminative and misclassification performances. However this is not always the case, as shown in another example where the discriminative power of supervised models based on the combined NMR and GC-MS datasets did not outperform the supervised models of the individual datasets (Teul et al., 2009). Still, correlations between the discriminant features of the multi-block model, which were detected by both the NMR and GC-MS datasets, offered a better understanding of the metabolic alterations lying in the plasma samples of patients suffering from stable carotid atherosclerosis compared to controls (Teul et al., 2009). This proves the benefits of combining those two techniques to either help increasing the metabolic coverage, improving multivariate analysis performance or have a deepest understanding of the biological processes.

This is also well exemplified in a study which investigated the metabolic profiles of human dopaminergic neuroblastoma cells treated with different neurotoxins and analyzed by both NMR

and high-throughput direct-ionization electrospray ionization MS (DI-ESI-MS) (Marshall et al., 2015). Firstly, the sample preparation was optimized to propose a dual analysis of the same sample which thus prevented extra sample handling. Secondly, the discriminant power of the multi-block principal component analysis (PCA) model which integrated the NMR dataset with the DI-ESI-MS dataset through a low-level approach, was clearly higher than the single NMR or DI-ESI-MS-based PCA models. Finally, the metabolite identification of the discriminant features were facilitated by accurate mass measurements and fragmentation patterns obtained by tandem MS experiments (Marshall et al., 2015). From top to bottom, this demonstrates how it is possible to successfully combined NMR and MS dataset to explore the effect of a specific treatment relevant to Parkinson's disease, and further clinical applications will most certainly be observed in the coming decade. However, careful attention should be paid to single block formatting during multi-block integration, especially through a low-level approach. Indeed, the aim is not to give too much weight to some variables or to one of the block considered for the integration (Boccard and Rudaz, 2014). In the study discussed previously combining NMR with DI-ESI-MS, each block were scaled to unit variance and by the square root of its variable count (Marshall et al., 2015).

As highlighted by the above examples, data integration has a lot of potential in clinical metabolomics, as the NMR and MS datasets can be fused with metadata describing life-style factors from large human cohorts. This was done in a study which aimed at observing metabolic alterations in human plasma samples from patient suffering from three different chronic diseases (acute coronary syndrome, breast and colon cancers) (Acar et al., 2017). The samples were analyzed by NMR and LC-MS (positive and negative ionization) and the integration of these three blocks with the metadata (**Figure 5A**), by using multiple kernel learning, provided a model which outperformed the individual models when it came to acute coronary syndrome (**Figure 5B**). However, the fusion of the different datasets did not improve the performance of the individual NMR model for the breast cancer samples, and of none of the individual models for the colon cancer samples (**Figure 5C**). In fact, the integration of the metadata can be useful, as it can help picking up novel confounding factors, as it was shown for metabolites linked to coffee consumption and smoking habits (Acar et al., 2017), but these very same confounding factors can also influence the selection of discriminative variables. This is a problem often encountered in metabolomics and even though several methods have been proposed to optimize the variable selection step, such as the one based on sparse multi-block PLSR for biomarker discovery (Karaman et al., 2015) or backward variable elimination from PLS-DA models combined with Monte Carlo Cross-Validation (Deng et al., 2016), this issue remains a current limitation of data fusion.

Using Nuclear Magnetic Resonance and Mass Spectrometry Strengths to Help Metabolite Identification

Increasing metabolic coverage and sensitivity also means more biomarkers to identify, which is obviously of major interest to understand their roles in specific diseases. The complementarity of the information that can be gathered by both NMR and MS-based techniques represents also an advantage when it comes to identifying biomarkers, especially when high resolution MS (HRMS) is used to acquire accurate mass measurements from parent compounds and their fragments in addition to the structural information obtained by 1D and 2D NMR spectroscopy. Detailed approaches to combine both have actually been proposed, such as SUMMIT MS/NMR (Bingol et al., 2015) or NMR/MS translator (Bingol and Brüschweiler, 2015a). The first one relies on HRMS measurements of a complex sample, from which putative molecular formulas and scaffolds can be proposed and NMR spectra predicted. Those predicted spectra are then compared to experimental HSQC NMR spectra, which have been deconvoluted for each of the sample metabolites (Bingol and Brüschweiler, 2017). To put it simply, NMR/MS Translator could be seen as the reverse of SUMMIT MS/NMR, as it starts with 1D and 2D NMR acquisition, so putative annotations can be made by comparing the experimental NMR spectra to databases. For the best hits, the MS spectra are predicted and then compared to experimental MS spectra (Bingol and Brüschweiler, 2015a). This last approach allowed the authors to identify new human urine metabolites which had never been reported previously. It has also been suggested by the same group that the SUMMIT MS/NMR approach could be applied on analytes which remained unidentified even following the application of the NMR/MS translator approach (Bingol and Brüschweiler, 2017). Upon the fact that some of the steps of these two approaches need to be automated in order to make the entire process more rapid, it could promote the identification of biomarkers in clinical research (Bingol and Brüschweiler, 2015b), but this has not really been widely applied yet.

Combining Nuclear Magnetic Resonance and Mass Spectrometry Techniques in a Quantitative Approach

Once the strengths of NMR and MS have been joined to increase the metabolic coverage, to provide more powerful statistical models and to identify new metabolites, new biomarkers of interest can be highlighted. However, approaches combining NMR and MS datasets often considers relative concentrations. Needless to say, that this is not satisfactory for clinical applications and on the contrary, absolute concentrations are needed for intra- and inter-laboratory comparison, as well as to compare data obtained with different analytical strategies. Recent methods have been proposed toward that goal. One, called « NMR-guided-MS quantitation », which consist in acquiring the absolute concentrations of analytes present in a randomly selected reference sample by NMR, which are then used as concentrations of reference for the rest of the samples analyzed by LC-MS/MS (Nagana Gowda et al., 2018). This method was applied to quantify 30 human serum metabolites in eight samples, and showed excellent correlations between the concentrations obtained by NMR and the ones obtained by NMR-guided MS, and good agreement between the NMR-guided MS approach and stable-isotope-labelled internal

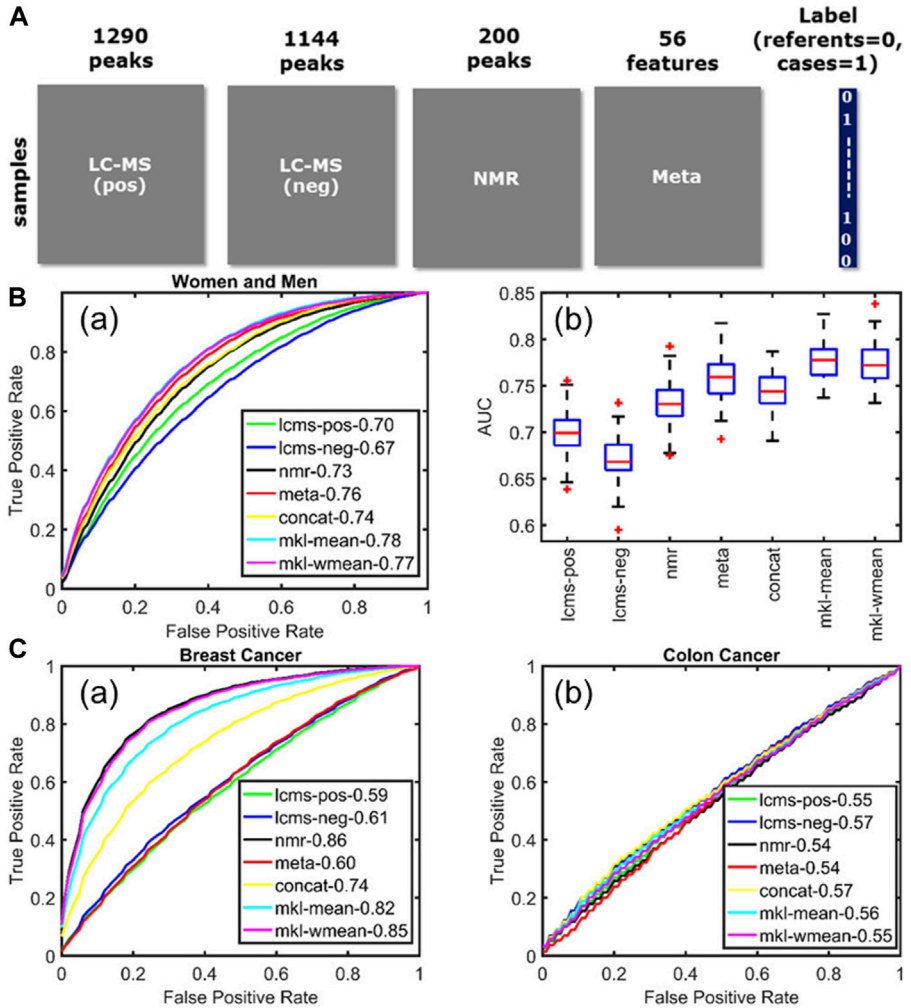

FIGURE 5 | Multiblock data fusion applied to study the metabolic alterations in human plasma samples from patient suffering from chronic diseases. **(A)** Data sets used in this study: metabolomics measurements (LC–MS and NMR), the metadata set containing life-style information, and the label information corresponding to each sample. **(B)** Acute coronary syndrome. **(a)** Average ROC curves showing the forecasting performance of individual data sets as well as fusion methods for women and men. **(b)** Boxplots summarize the performance of different approaches across 100 training/test sets. **(C)** Average ROC curves illustrating the forecasting performance of **(a)** breast cancer and **(b)** colon cancer. Figure reproduced with permission from (Acar et al., 2017).

standards (SIL IS) measurements by MS. However once considering each of the metabolites individually, even though most of them showed good correlations between NMR and NMR-guided MS (*e.g.*, R^2 = 0.989 for proline), some demonstrated very poor correlations (*e.g.*, R^2 = 0.207 for pyroglutamic acid) (Nagana Gowda et al., 2018). This needs to be seriously consider when it comes to clinical biomarker discovery as whatever explanation lying behind those poor correlations (*e.g.*, glutamine cyclization (Purwaha et al., 2014; Nagana Gowda et al., 2015a; Nagana Gowda et al., 2015b), multiple or poor signals, ion suppression . . .), it proves that it is wrong to assume that MS can provide stable measurements of all metabolites. To identify those unstable metabolites, or as an alternative to labour-intensive calibration curves, the NMR-guided MS can be of interest. Build on this approach, another one has been proposed, based on the derivatization of the reference sample with SIL IS and of the rest of the samples with unlabelled IS (Fei et al., 2019). This new approach, called the qNMR-MS, offers the possibility to reduce matrix effect but presents the drawback of adding additional sample handling, potentially limiting when large number of samples are considered.

Nuclear Magnetic Resonance and Mass Spectrometry Techniques as the Keystones of Fluxomics

A branch of metabolomics which combines both analytical platforms which gathered a lot of interest in clinical research is Stable Isotope-Resolved Metabolomics (SIRM), also referred as fluxomics analysis. SIRM offers the possibility to quantitatively apprehend metabolic pathways and fluxes by measuring

isotopomers, by NMR, and isotopologues, by MS, following labelling of a precursor molecule with stable isotope tracers. Most importantly, one of the advantages of fluxomics is that it can be done either *in vitro* or *in situ*. SIRM has thus the potential to lift the veil on the metabolic mechanism of numerous diseases, especially cancer (Lane et al., 2011; Lane et al., 2016; Lane et al., 2019). More active glycolysis and Krebs cycle, as well as an activated pyruvate carboxylation were for instance found to promote tumour development in lung tissues (Fan et al., 2009). The deep gain in knowledge on how diseases actually work can clearly improve personalized treatment (Fan et al., 2012). SIRM is probably the metabolomics branch presenting the finest achievements in combining NMR and MS analytical technologies to unravel disease understanding, but also the most complex one. However tremendous efforts have been done to promote rapid development of new computational tools to help SIRM and other metabolomics branches to be implemented in the long term within clinical laboratories. These computational tools will also certainly help the integration of SIRM findings, or metabolomics in general, with other kinds of datasets, such as genomics, transcriptomics, proteomics or clinical metadata to acquire a more in-depth knowledge of biomolecular mechanism of a pathology.

Combining Nuclear Magnetic Resonance and Mass Spectrometry Techniques for Personalized Medicine: Where Do We Stand?

Some studies apply both NMR and MS-based analytical strategies to clinical research to combine the respective biomarkers of interest, and include them in a common metabolic pathway analysis, as it was done to study the primary membranous glomerylonephritis and the subsequent nephrotic syndrome that it can cause in adults (Taherkhani et al., 2019). Others use NMR as a primary tool for open-profiling metabolomics, and then use subsequent LC-MS/MS to confirm the results obtained by NMR or to quantitatively target a subset of metabolites. This approach was used to analyze 244 human serum samples from the ECLISPE study and to identify biomarkers of chronic obstructive pulmonary disease (Ubhi et al., 2012), or to analyze 32 neonate urine samples and identify biomarkers related to late-onset sepsis (Sarafidis et al., 2017). Through an example of large epidemiological study performed on 4,680 urinary samples from the INTERMAP study, ^{1}H NMR was also used for metabolic phenotyping before applying GC-MS and LC-MS/MS to analyze the urinary amino acids (Chan et al., 2017). Surprisingly when it comes to large epidemiological cohorts, in the 47 studies reported in the COMETS initiative (Yu et al., 2019), relatively few are applying both NMR and MS-based metabolic profiling approaches. From those, it is worse mentioning the AIRWAVE study (Elliott et al., 2014), the MAC study (Chow et al., 2017), the MESA study (Bild et al., 2002) or the TwinsUK study (Moayyeri et al., 2013). The same observation can been done when it comes to metabolomics biomarkers from acute respiratory distress syndrome, chronic obstructive pulmonary disease and asthma (Bowler et al., 2017). In another review focusing on preeclampsia, 16 metabolomics studies were based on MS-data and 12 by NMR, but none were employing both (Kelly et al., 2017). However and as nicely pointed out in this review, combining both could provide more robust and accurate preeclampsia metabolic profile, as the metabolic coverage accessible with each methods in the studies reviewed were not always comparable, also because of the targeted approach often applied in MS which focus only on a subset of the metabolome (Kelly et al., 2017). The integration of NMR and MS metabolomics with other OMICS analytical platforms gathered a lot of interest in the last 3 years in the field of biomedical sciences (Manzoni et al., 2018), personalized medicine (Jacob et al., 2019), environmental health (Yao et al., 2019), microbiome research (Zimmermann et al., 2019) or toxicology (González-Ruiz et al., 2019). Integration of different OMICS platforms have even been of interest for personalized medicine in human space flight (Schmidt and Goodwin, 2013). Furthermore, optimized extraction protocol to analyzed on the same sample the metabolites, the proteins and the lipids have been proposed (Coman et al., 2016), so upon further computational development, the perspectives of integrating NMR-based metabolomics with MS-based metabolomics or other OMICS will certainly promote its application within clinical settings.

RECENT AND FUTURE DEVELOPMENTS IN NUCLEAR MAGNETIC RESONANCE-BASED METABOLOMICS

The first section of this review highlighted the major role that NMR plays as an analytical tool in clinical metabolomics. The second part described how this role can be further strengthened by combining NMR with other analytical techniques, especially MS-based metabolomics. Nevertheless, there are still major challenges posed to the NMR spectroscopists in order to further improve the potential of NMR spectroscopy within clinical metabolomics Indeed, NMR has well-known limitations. As mentioned in the previous section, the main limitation is a reduced sensitivity compared to other analytical methods and particularly MS. The sensitivity of NMR at high magnetic field (>500 MHz) is in the micromolar range. This is sufficient to detect major metabolites in biofluids or extracts, but relatively large sample amounts are often required, and the detection of less concentrated, specialized metabolites can be a challenge. A second limitation arises from resolution issues, since it can be difficult to separate overlapping metabolite signals in crowded spectral regions of the ^{1}H spectrum. Finally, a third reason why NMR is less used than MS in the clinical world is the relatively high purchase cost of NMR instruments (>1M€ for a 600 MHz spectrometer) and the associated consumption of cryofluids.

NMR spectroscopy would certainly be much more widely used in the clinical world if the above-mentioned limitations were circumvented. While these challenges are not new, several recent (<10 years) methodological advances in the NMR community have laid the foundations for a more sensitive, better resolved and

more accessible NMR spectroscopy (Giraudeau, 2020). This part focuses on such recent advances, which, in addition to the high robustness of NMR spectroscopy, have the potential to provoke a significant paradigm shift regarding the role of NMR for biomedical applications. Some of them have already proved their usefulness in the field while other rather offer mid-term perspectives, but in our view, all are of interest to the fields of clinical studies and personalized medicine.

Improving the Sensitivity

The sensitivity of NMR directly results from the level of nuclear polarization, generally determined by a Boltzmann law at thermal equilibrium. This results in relatively weak nuclear polarization levels, for instance, only 0.000008 for ^{1}H at 300 K in a 14 T magnetic field -the typical NMR metabolomics configuration. The most direct -but technologically challenging-approach consists in increasing the static B_0 field, since the sensitivity increases with $B_0^{3/2}$. NMR metabolomics experiments are typically performed between 500 and 800 MHz, but magnets up to 1.2 GHz are now commercially available, providing impressive results on biofluids (Banci et al., 2019). However, such very high field magnets only provide a modest sensitivity gain (a factor 2.8 between 600 MHz and 1.2 GHz) while their cost is at least ten times higher. On the hardware side, more promising perspectives probably arise from the development of more sensitive NMR probes. Cryogenically cooled probes can provide a signal to noise ratio (SNR) improvement by a factor 3 to 4, however they are expensive and show limited efficiency for samples with high salinity (Kovacs, 2005). On the one hand, higher field and cryoprobes are well suited to improve the limit of detection for a given sample volume, but on the other hand, numerous small volume probes have been developed to analyze mass-limited samples without compromising on sensitivity. These include microprobes that can accommodate sample volumes of a few tens of µL (Clendinen et al., 2014), but also recent microfluidic-based probes that can detect metabolites at sub-millimolar concentrations in sample volumes of *ca.* 2 µL (Finch et al., 2016). The incorporation of such microfluidic devices in NMR experiments also makes it possible to perform flow experiments, opening great avenues for time-resolved metabolomics. Patra *et al.* recently applied this approach to non-invasive metabolomic monitoring of microfluidic cultures with as few as 1,250 individual cells (Patra et al., 2021).

In addition to such magnet and probe advances that will certainly enhance the performance of clinical NMR metabolomics, great promises arise from hyperpolarization methods, which have been the focus of many exciting developments in the NMR community in the last 2 decades. Indeed, these approaches can enhance the sensitivity of NMR spectroscopy by up to four orders of magnitude by enhancing the nuclear polarization to values close to unity. The two most popular methods for hyperpolarization are para-hydrogen induced polarization (Duckett and Mewis, 2013) and dynamic nuclear polarization (Plainchont et al., 2018). Both have been discovered many decades ago, but only recent developments have made them applicable to the analysis of complex samples with metabolomics relevance.

The first approach, para-hydrogen induced polarization, is based on the transfer of hyperpolarization from H_2 in the para state to the nuclear spins of analytes (Duckett and Mewis, 2013). While the initial approach involved a chemical hydrogenation reaction, it was made more versatile and general by the development of the SABRE technique (signal amplification by reversible exchange) which involves the addition of a metal-based complex to reversibly transfer the hyperpolarization to the analytes (Lloyd et al., 2012). This method is very attractive for practical applications since it is simple and relatively cheap. However, it has a certain degree of selectivity since the SABRE catalyst mainly binds to compounds containing electron-donating heteroatoms such as nitrogen. SABRE-based hyperpolarization has already been successfully applied to quantify metabolites in natural extracts (Hermkens et al., 2016). Although it has not yet been applied to a metabolomics study, Tessari and co-workers were recently able to detect numerous metabolites at nanomolar concentrations in solid phase extracts of urine, which forms a promising perspective for metabolomics (Sellies et al., 2019).

Parallel developments in the NMR world have been focusing on another hyperpolarization technique, dissolution dynamic nuclear polarization (d-DNP, **Figure 6A**) proposed in 2003 by Ardenkjaer-Larsen and co-workers (Ardenkjaer-Larsen et al., 2003). This approach consists in mixing the sample with small amounts of free radicals, freezing it in a glass-forming solution at liquid Helium temperature and in a static magnetic field, then irradiating it by microwaves at the Larmor frequency of the unpaired electrons. Under such conditions, the very high polarization of the electrons is transferred to nuclei, leading to polarizations close to unity within a few minutes. The frozen sample is then rapidly transferred to a nearby NMR spectrometer, where classical spectra can be obtained with sensitivity enhancements by up to four orders of magnitude. The d-DNP approach is technically demanding but very general, since all metabolite signals can be enhanced in a non-selective fashion. A more fundamental limitation arises from the decrease of hyperpolarization during sample transfer, which occurs as a function of nuclear longitudinal relaxation times (T_1). As a consequence, most applications of d-DNP have been focusing on ^{13}C nuclei, since their T_1 can reach several tens of seconds, especially for quaternary carbons. In the MRI community, d-DNP has rapidly had a great impact on metabolic imaging, with the first injection of hyperpolarized pyruvate to humans in 2013 (Nelson et al., 2013), and not less than 25 undergoing clinical trials reported in 2019 (Ardenkjaer-Larsen, 2019). In NMR spectroscopy, d-DNP has also been widely used to investigate metabolic processes in real-time, for instance to get insight into enzymatic kinetics (Wilson et al., 2010). In this context, the application of d-DNP to extracts or biofluids opens promising perspectives to enhance the sensitivity of NMR metabolomics, and first steps towards this goal have been reported recently. In 2015, Dumez *et al.* showed that d-DNP could be applied to enhance the ^{13}C NMR signals in plant and cancer cell extracts at natural abundance (Dumez et al., 2015), and in 2016, the very good analytical repeatability (<4%) of the method was demonstrated (Bornet et al., 2016). Lerche *et al.*

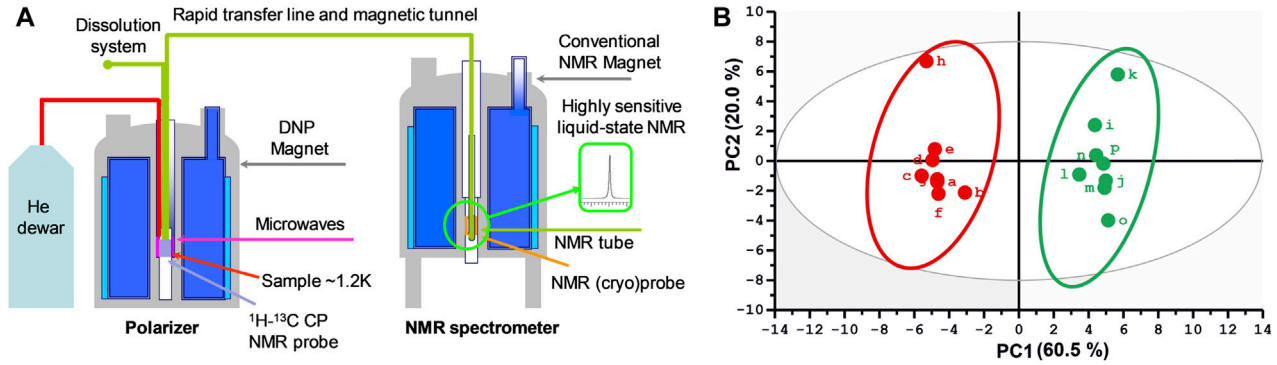

FIGURE 6 | Potential of dissolution dynamic nuclear polarization (d-DNP) for highly sensitive ^{13}C NMR metabolomics at natural abundance. **(A)** d-DNP experimental scheme, where the sample is first hyperpolarized at liquid He temperature in a glass-forming solvent by microwave irradiation of free radicals, then rapidly transferred through a magnetic tunnel to a classical NMR spectrometer where conventional detection occurs. **(B)** First application of this experimental scheme in a natural abundance ^{13}C metabolomics study, reported in 2020 by Dey *et al.* Scores plot of the principal component analysis (PCA) obtained from 40 spectral buckets from hyperpolarized ^{13}C spectra of 16 tomato fruit extracts at different ripening stages (green vs red). Integrals were normalized to Na-TSP-d_4 as an internal reference and to the weight of the sample used for extraction. Mean centering and unit variance scaling were used in PCA. **(B)** Reproduced from (Dey et al., 2020).

also reported a complementary approach relying on the incubation of the targeted biological material with a ^{13}C-labeled substrate (Lerche et al., 2017). More recently, in 2020, Dey *et al.* demonstrated, on the example of plant extracts, the first hyperpolarized metabolomics study at natural ^{13}C abundance (**Figure 6B**) (Dey et al., 2020). While these recent methods have not yet been applied to clinical metabolomics, they could pave the way towards the detection of biomarkers that were not accessible by NMR so far. In particular, ongoing technological developments to accelerate the sample transfer (Bowen and Hilty, 2010) -thus making d-DNP compatible with ^{1}H detection- and to increase the lifetime of hyperpolarized samples (Ji et al., 2017) could help spreading this promising approach in the metabolomics community.

Improving the Resolution

Typical samples of metabolomics relevance are extremely complex, since they contain a great diversity of metabolites. Although NMR is not as sensitive as MS, resulting spectra can be extremely complex and characterized by strong and numerous peak overlaps that can alter the high analytical performance of NMR. In a classical untargeted metabolomics workflow, several NMR signals pertaining to different metabolites can be observed within a single bucket, making it difficult to identify relevant biomarkers. When quantitative data are being sought, the accurate determination of peak areas is hampered by such overlaps, leading to errors in concentration determination. Signal processing methods can help deconvoluting individual metabolite contributions, but such methods often rely on databases which are specific of a given matrix prepared under specific conditions (Hao et al., 2012; Lacy et al., 2014; Ravanbakhsh et al., 2015). One can also rely on the detection of heteronuclei, such as ^{13}C, that offer a much broader chemical shift dispersion compared to ^{1}H (Clendinen et al., 2014). However, due to a limited sensitivity, the routine application of ^{13}C NMR metabolomics will require highly sensitive detection methods such as those reported in the previous paragraphs. Fortunately, in addition to these approaches, many innovative NMR methods -based on pulse sequence developments-have been developed to simplify the analysis of complex mixtures, that were successfully transferred to metabolomics in recent studies, and there is not much doubt that at least some of them will become part of the daily clinical metabolomics workflow in a near future.

The most widely used strategy to better separate overlapping signals in NMR of complex mixtures is to rely on multi-dimensional methods such as 2D NMR. Indeed, in 2D NMR spectra, peaks are spread along two orthogonal dimensions (typically ^{1}H–^{1}H or ^{1}H-^{13}C), hence reducing peak overlap while providing crucial information on atomic connectivity. 2D NMR has been used for decades to elucidate the molecular structure of chemical compounds, including metabolites. In metabolomics studies, popular experiments such as J-resolved spectroscopy, COSY (correlation spectroscopy), TOCSY (total correlation spectroscopy) or HSQC (heteronuclear single-quantum correlation) are generally applied to a subset of samples from a given study, or to a purified fraction of a biological matrix, to elucidate the structure of biomarkers (Mahrous and Farag, 2015). However, the systematic use of 2D NMR peaks volumes as a raw data in metabolomics workflow is still far from routine. Nevertheless, this would provide a great way to better extract individual metabolite variations, since 2D peaks are much less prone to overlap than their 1D counterparts. Indeed, early studies have shown the potential of using 2D NMR in metabolomics. For instance, Van *et al.* showed that 2D TOCSY spectra of mice urine samples allowed to better characterize concentration changes in low-concentrated metabolites compared to classical 1D NMR (Van et al., 2008). Since then, 2D spectra have been used in a number of metabolomics studies that highlighted the relevance of using 2D NMR data in such context (Robinette et al., 2011; Féraud et al., 2015; Puig-Castellví et al., 2018). An

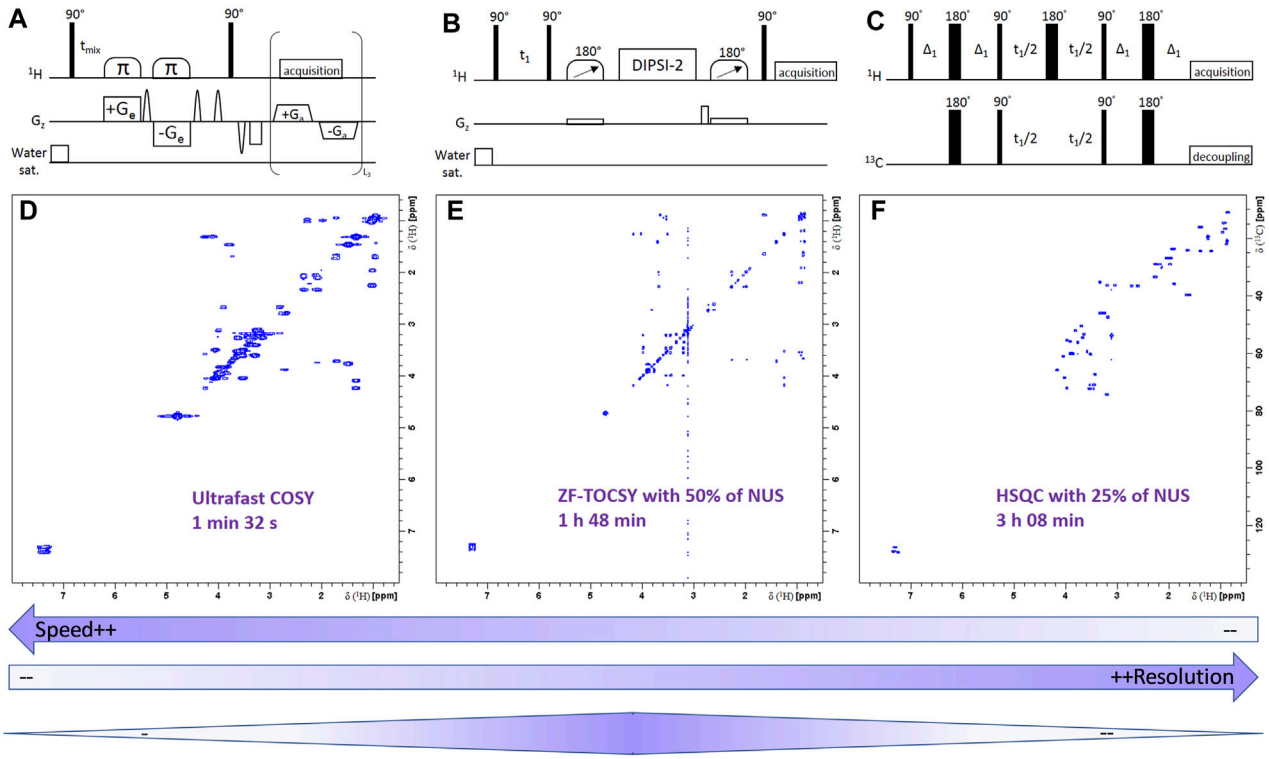

FIGURE 7 | Typical fast 2D NMR pulse sequences for high-throughput metabolomics and their corresponding spectrum obtained on a model mixture of metabolites **(A,D)** Ultrafast ^{1}H–^{1}H correlation spectroscopy (UF COS)Y **(B,E)** Z-filter ^{1}H–^{1}H total correlation spectroscopy (ZF-TOCSY) with 50% of non-uniform sampling in the indirect dimension **(C,F)** ^{1}H–^{13}C heteronuclear single-quantum correlation (HSQC) with 25% of non-uniform sampling in the indirect dimension. Pulse sequences offer complementary performances in terms of speed, resolution and sensitivity. Adapted with permission from (Martineau et al., 2020).

additional benefit of 2D NMR is that it requires a less advanced level of pre-processing, since data are already well separated at the acquisition stage (Féraud et al., 2019).

However, a major obstacle which has limited the systematic use of 2D NMR in metabolomics has been the long experiment time needed to record 2D spectra with sufficient sensitivity and resolution. Indeed, typical 2D NMR experiments may last between a few tens of minutes until many hours since they rely on the repetition of numerous 1D experiments with a slight incrementation of a specific delay in the pulse sequence. Such durations are not compatible with the high-throughput analysis of large sample cohorts. They may also not be compatible with time stability issues of some biological samples. Fortunately, several methodological developments have been carried out to accelerate the acquisition of 2D NMR spectra (Rouger et al., 2016). These include spectral aliasing (Njock et al., 2010), fast repetition techniques (Schanda, 2009), non-uniform sampling (NUS) (Mobli and Hoch, 2014) or ultrafast (UF) spectroscopy (Giraudeau and Frydman, 2014). The reader is referred to the aforementioned reviews for detailed explanations on the corresponding methodologies, but in summary, these methods can accelerate the acquisition of 2D NMR spectra while preserving a good sensitivity and resolution performance, leading to reasonable acquisition times. A particularly interesting feature of fast 2D NMR is that it offers many different pulse sequences with complementary features that make it possible to choose, for a given matrix and application, the best compromise between experiment time, sensitivity and resolution (**Figure 7**) (Martineau et al., 2020). Some of these methods have been successfully applied in metabolomics workflows. For instance, Marchand *et al.* showed that UF COSY and NUS TOCSY applied to pig lipid serum lipid extracts offered an improved detection of biomarkers characterizing the administration of a growth promoted (Marchand et al., 2018). Feraud *et al.* showed that NUS COSY spectra recorded in less than 10 min provided a fast and efficient approach for the profiling of human urine samples (Féraud et al., 2020). It is also worth highlighting that most 2D NMR approaches only provide information on relative metabolite concentration variations between samples. However, when associated with appropriate analytical procedures, fast 2D NMR can yield accurate absolute quantitative data, which can be an interesting alternative to 1D NMR when targeted quantification of metabolites (Marchand et al., 2017). This strategy has been applied, for instance, to the quantification of metabolites in cancer cell extracts (Martineau et al., 2011; Le Guennec et al., 2012).

Another promising perspective on the pulse sequence development side arises from pure-shift NMR methods. The term "pure-shift" refers to an ensemble of methodologies that

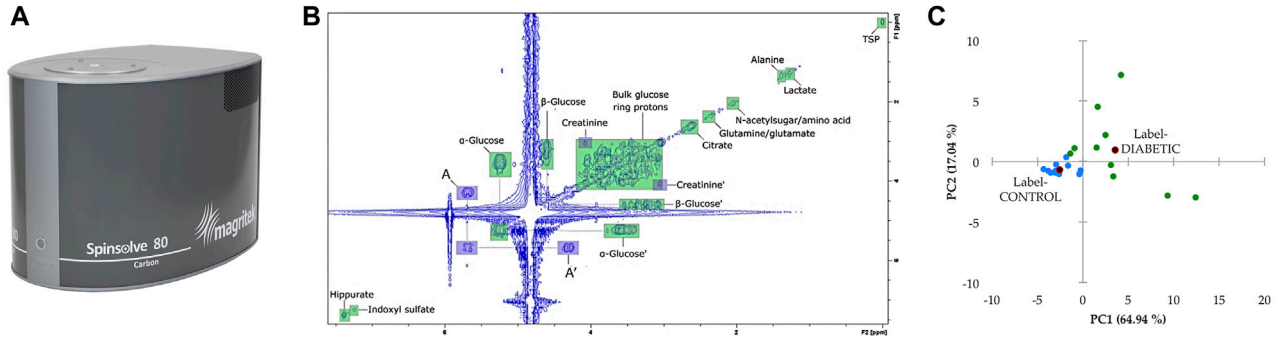

FIGURE 8 | Potential of benchtop NMR spectroscopy to make NMR metabolomics more accessible in a clinical context. **(A)** Typical benchtop NMR spectrometer **(B)** 2D ^{1}H–^{1}H COSY NMR of type 2 diabetes urinary profile acquired at 60 MHz using a benchtop NMR spectrometer. Creatinine blue squares represent the long-range connectivity cross-peak for this metabolite. Blue squares labelled A represent unassigned, unusual doublet resonances arising from 'mirroring' spectral signal located at δ = 5.13–5.29 ppm (in this case, reflecting the α-glucose cross-peak). **(C)** Principal component analysis (PCA) scores plot of PC2 (17.04% of total variance) versus PC1 (64.94% of total variance) for a preliminary investigation of distinctions between healthy control and type 2 diabetic cohorts, and also potential sample outliers. Colour codings: blue, urine samples collected from healthy controls; green, those from type 2 diabetes participants. The black points represent scores plot centroids for the two groups explored. PCA was performed using XLSTAT2014 software, and the dataset was TSP-normalised, generalised logarithmically (glog)-transformed and Pareto-scaled prior to analysis. **(A)** Courtesy of Magritek GmbH. **(B)** Reproduced from (Leenders et al., 2020) under Creative Commons Attribution 4.0 International License. **(C)** Reproduced from (Percival et al., 2018) under Creative Commons Attribution 4.0 International License.

aim at transforming all NMR multiplets into singlets (Castañar, 2017). In the case of ^{1}H NMR, pure-shift NMR provides a great way of reducing peak overlap in the case of complex mixtures, while retaining the simplicity of 1D spectra. These methodologies suffer from a strong sensitivity penalty but could be attractive for metabolomics workflows. Recent studies demonstrated the potential of pure-shift NMR in plant metabolomics (Lopez et al., 2019), and application to samples of clinical relevance may occur in the coming years.

Finally, an alternative to these pulse sequence approaches to simplify NMR spectra of complex mixtures is to rely on selective methods that reduce the number of observable analytes. While this strategy may seem counter-intuitive in metabolomics, such methods can be of interest when focusing on a limited set of targeted metabolites which can be of interest as markers of a given pathology. Most promising strategies rely on "chemosensing" methods such as the addition of charged nanoparticles that selectively suppress NMR signals of metabolites whose charge is opposite to those of the nanoparticles (Zhang et al., 2016), or the coating of nanoparticles with ligands that selectively bind to some classes of metabolites (Salvia et al., 2015).

Improving the Accessibility

A major challenge for a widespread clinical application of NMR spectroscopy lies in the limited accessibility to NMR instruments arising from their high cost, heaviness and high level of technicity -including the regular handling of cryogenic fluids. Very exciting perspectives arise along this direction from the recent development of compact NMR spectrometers, which have been made commercially available for a few years and have already known a great success in chemistry labs and industries (Singh and Blümich, 2016). Such spectrometers rely on permanent magnets that do not require any specific maintenance and which provide a medium magnetic field (1–2 T) yielding a ^{1}H resonance frequency between 40 and 100 MHz (**Figure 8A**) (Kuster et al., 2011). Benchtop NMR spectrometers are transportable (<100 kg), low-cost (<100,000 €) and most commercial models can easily be used in both static and flow configurations, which explains their success for chemical applications. Of course, they also have a reduced performance compared to high-field NMR spectrometer, with a lower sensitivity and a limited ability to separate overlapping peaks (owing to the small frequency range in Hz, while multiplet structures such as J-couplings are invariant to the magnetic field).

In this context, it goes without saying that benchtop NMR will not replace high-field NMR in the detection and structure elucidation of low-concentrated biomarkers. But it could play a significant role as an affordable and high-throughput metabolic profiling tool at the point-of-care, especially when sample amounts are not limited, *e.g.* urine samples. Along this line, impressive results have been achieved by Wilson and co-workers, who demonstrated that a 60 MHz commercial benchtop spectrometer could detect and even quantify major metabolites in urine within a few minutes, reaching limits of detection of *ca.* 25 μM (Percival et al., 2018). They also showed nicely resolved 2D COSY spectra (**Figure 8B**) (Leenders et al., 2020), and they eventually reported an efficient group separation between type 2 diabetes patients and healthy controls (**Figure 8C**). In another recent study, Izquierdo-Garcia *et al.* showed that a tuberculosis biomarker in urine -previously determined by high-field NMR- could also be detected by benchtop NMR (Izquierdo-Garcia et al., 2020).

While these results remain preliminary and will need to be validated at a larger scale, they highlight how NMR spectroscopy could soon make its way towards the patient's bed and help as a routine tool for rapid and accurate sample classification in a clinical context. In addition, recent developments have shown how high-resolution pulse sequences could be implemented on benchtop spectrometers (Gouilleux et al., 2020). These include solvent-suppression pulse sequences such as those used in routine

high-field NMR metabolomics, as well as methods which have been described above to improve the resolution such as fast 2D NMR or pure-shift approaches. Recent results highlighted the potential of such advanced benchtop NMR methods for sample classification (Gouilleux et al., 2018), and one can expect that clinical metabolomics will benefit from such advances in the coming years.

CONCLUSION

Thanks to intrinsic properties such as high reproducibility, the possibility to quantify, a high degree of structural information, its "universal" detection capacity for all organic molecules as well as its adaptability in the analysis of biological samples, NMR very early appeared as a platform of choice in clinicals metabolomics. The numerous publications, works and results based on NMR attest this fact and keep contributing to the development of this approach. The recent progresses in mass spectrometry coupled with liquid or gas chromatography, a more sensitive and higher resolution technique, have progressively led NMR to play a "second" role in metabolomic studies, raising the question of its future in the field. However, the applications of metabolomics in clinical research and personalized medicine have brought new needs and challenges for metabolomics, such as the analysis of large cohorts, the stratification and the longitudinal follow-up of patients and the identification and quantification of biomarkers. To face such multiple requirements and needs, NMR has real assets and opportunities. Indeed, the many recent instrumental and methodological developments aiming at improving both sensitivity and resolution, as well as the demonstration of its excellent complementarity with mass spectrometry, highlight the leading role of NMR spectroscopy in clinical metabolomics and in personalized medicine. The translation from laboratory studies to clinical practice is another challenge that metabolomics is facing and, in this respect, we are confident that NMR will be one of the key analytical platforms that can provide valuable and innovative solutions and opportunities with a view to a personalized approach to medicine. Considering the numerous promising perspectives mentioned in this review, there is no doubt that the recent and future developments will rekindle the flame of NMR spectroscopy in clinical metabolomics for the next decades.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

Authors from CEISAM acknowledge support of the European Research Council under the European Union's Horizon 2020 Research and Innovation Program (ERC Grant Agreement No. 814747/SUMMIT) the French National Infrastructure for Metabolomics and Fluxomics MetaboHUB-ANR-11-INBS-0010 (www.metabohub.fr) and the Corsaire metabolomics core facility (Biogenouest). PdT is Research Director of the FNRS.

REFERENCES

Acar, E., Gürdeniz, G., Savorani, F., Hansen, L., Olsen, A., Tjønneland, A., et al. (2017). Forecasting Chronic Diseases Using Data Fusion. *J. Proteome Res.* 16, 2435–2444. doi:10.1021/acs.jproteome.7b00039

Albrecht, B., Voronina, E., Schipke, C., Peters, O., Parr, M. K., Díaz-Hernández, M. D., et al. (2020). Pursuing Experimental Reproducibility: An Efficient Protocol for the Preparation of Cerebrospinal Fluid Samples for NMR-Based Metabolomics and Analysis of Sample Degradation. *Metabolites* 10, 251. doi:10.3390/metabo10060251

Amathieu, R., Triba, M. N., Nahon, P., Bouchemal, N., Kamoun, W., Haouache, H., et al. (2014). Serum 1H-NMR Metabolomic Fingerprints of Acute-On-Chronic Liver Failure in Intensive Care Unit Patients with Alcoholic Cirrhosis. *PLOS ONE* 9, e89230. doi:10.1371/journal.pone.0089230

Amiel, A., Tremblay-Franco, M., Gautier, R., Ducheix, S., Montagner, A., Polizzi, A., et al. (2019). Proton NMR Enables the Absolute Quantification of Aqueous Metabolites and Lipid Classes in Unique Mouse Liver Samples. *Metabolites* 10, 9. doi:10.3390/metabo10010009

Ardenkjaer-Larsen, J. H., Fridlund, B., Gram, A., Hansson, G., Hansson, L., Lerche, M. H., et al. (2003). Increase in Signal-To-Noise Ratio of > 10,000 Times in Liquid-State NMR. *Proc. Natl. Acad. Sci.* 100, 10158–10163. doi:10.1073/pnas.1733835100

Ardenkjaer-Larsen, J. H. (2019). Hyperpolarized MR - What's up Doc? *J. Magn. Reson.* 306, 124–127. doi:10.1016/j.jmr.2019.07.017

Ashrafian, H., Sounderajah, V., Glen, R., Ebbels, T., Blaise, B. J., Kalra, D., et al. (2021). Metabolomics: The Stethoscope for the Twenty-First Century. *Med. Princ Pract.* 30, 301–310. doi:10.1159/000513545

Banci, L., Barbieri, L., Calderone, V., Cantini, F., Cerofolini, L., Ciofi-Baffoni, S., et al. (2019). Biomolecular NMR at 1.2 GHz. *ArXiv191007462 Phys*. Available at: http://arxiv.org/abs/1910.07462 (Accessed December 30, 2020).

Barrilero, R., Gil, M., Amigó, N., Dias, C. B., Wood, L. G., Garg, M. L., et al. (2018). LipSpin: A New Bioinformatics Tool for Quantitative 1H NMR Lipid Profiling. *Anal. Chem.* 90, 2031–2040. doi:10.1021/acs.analchem.7b04148

Beale, D. J., Pinu, F. R., Kouremenos, K. A., Poojary, M. M., Narayana, V. K., Boughton, B. A., et al. (2018). Review of Recent Developments in GC-MS Approaches to Metabolomics-Based Research. *Metabolomics* 14, 152. doi:10.1007/s11306-018-1449-2

Beckonert, O., Keun, H. C., Ebbels, T. M. D., Bundy, J., Holmes, E., Lindon, J. C., et al. (2007). Metabolic Profiling, Metabolomic and Metabonomic Procedures for NMR Spectroscopy of Urine, Plasma, Serum and Tissue Extracts. *Nat. Protoc.* 2, 2692–2703. doi:10.1038/nprot.2007.376

Bedia, C., Cardoso, P., Dalmau, N., Gómez-Canela, C., Gorrochategui, E., et al. (2018). "Chapter Nineteen Applications of Metabolomics Analysis in Environmental Research," in *Comprehensive Analytical Chemistry Data Analysis for Omic Sciences: Methods and Applications*. Editors J. Jaumot, C. Bedia, and R. Tauler (Amsterdam: Elsevier), 533–582. doi:10.1016/bs.coac.2018.07.006

Beger, R. D., Schmidt, M. A., and Kaddurah-Daouk, R. (2020). Current Concepts in Pharmacometabolomics, Biomarker Discovery, and Precision Medicine. *Metabolites* 10, 129. doi:10.3390/metabo10040129

Beirnaert, C., Meysman, P., Vu, T. N., Hermans, N., Apers, S., Pieters, L., et al. (2018). Speaq 2.0: A Complete Workflow for High-Throughput 1D NMR Spectra Processing and Quantification. *PLoS Comput. Biol.* 14, e1006018. doi:10.1371/journal.pcbi.1006018

Bhinderwala, F., Evans, P., Jones, K., Laws, B. R., Smith, T. G., Morton, M., et al. (2020). Phosphorus NMR and its Application to Metabolomics. *Anal. Chem.* 92, 9536–9545. doi:10.1021/acs.analchem.0c00591

Bild, D. E., Bluemke, D. A., Burke, G. L., Detrano, R., Diez Roux, A. V., Folsom, A. R., et al. (2002). Multi-Ethnic Study of Atherosclerosis: Objectives and Design. *Am. J. Epidemiol.* 156, 871–881. doi:10.1093/aje/kwf113

Bingol, K., and Brüschweiler, R. (2015a). NMR/MS Translator for the Enhanced Simultaneous Analysis of Metabolomics Mixtures by NMR Spectroscopy and Mass Spectrometry: Application to Human Urine. *J. Proteome Res.* 14, 2642–2648. doi:10.1021/acs.jproteome.5b00184

Bingol, K., and Brüschweiler, R. (2015b). Two Elephants in the Room. *Curr. Opin. Clin. Nutr. Metab. Care* 18, 471–477. doi:10.1097/MCO.0000000000000206

Bingol, K., and Brüschweiler, R. (2017). Knowns and Unknowns in Metabolomics Identified by Multidimensional NMR and Hybrid MS/NMR Methods. *Curr. Opin. Biotechnol.* 43, 17–24. doi:10.1016/j.copbio.2016.07.006

Bingol, K., Bruschweiler-Li, L., Yu, C., Somogyi, A., Zhang, F., and Brüschweiler, R. (2015). Metabolomics Beyond Spectroscopic Databases: A Combined MS/NMR Strategy for the Rapid Identification of New Metabolites in Complex Mixtures. *Anal. Chem.* 87, 3864–3870. doi:10.1021/ac504633z

Boccard, J., and Rudaz, S. (2014). Harnessing the Complexity of Metabolomic Data with Chemometrics. *J. Chemom.* 28, 1–9. doi:10.1002/cem.2567

Bornet, A., Maucourt, M., Deborde, C., Jacob, D., Milani, J., Vuichoud, B., et al. (2016). Highly Repeatable Dissolution Dynamic Nuclear Polarization for Heteronuclear NMR Metabolomics. *Anal. Chem.* 88, 6179–6183. doi:10.1021/acs.analchem.6b01094

Bowen, S., and Hilty, C. (2010). Rapid Sample Injection for Hyperpolarized NMR Spectroscopy. *Phys. Chem. Chem. Phys.* 12, 5766–5770. doi:10.1039/c002316g

Bowler, R. P., Wendt, C. H., Fessler, M. B., Foster, M. W., Kelly, R. S., Lasky-Su, J., et al. (2017). New Strategies and Challenges in Lung Proteomics and Metabolomics. An Official American Thoracic Society Workshop Report. *Ann. ATS* 14, 1721–1743. doi:10.1513/AnnalsATS.201710-770WS

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

Bruzzone, C., Bizkarguenaga, M., Gil-Redondo, R., Diercks, T., Arana, E., García de Vicuña, A., et al. (2020). SARS-CoV-2 Infection Dysregulates the Metabolomic and Lipidomic Profiles of Serum. *iScience* 23, 101645. doi:10.1016/j.isci.2020.101645

Castañar, L. (2017). Pure Shift 1 H NMR: what Is Next? *Magn. Reson. Chem.* 55, 47–53. doi:10.1002/mrc.4545

Catapano, A. L., Graham, I., De Backer, G., Wiklund, O., Chapman, M. J., Drexel, H., et al. (2016). 2016 ESC/EAS Guidelines for the Management of Dyslipidaemias. *Atherosclerosis* 253, 281–344. doi:10.1016/j.atherosclerosis.2016.08.018

Chan, Q., Loo, R. L., Ebbels, T. M. D., Van Horn, L., Daviglus, M. L., Stamler, J., et al. (2017). Metabolic Phenotyping for Discovery of Urinary Biomarkers of Diet, Xenobiotics and Blood Pressure in the INTERMAP Study: An Overview. *Hypertens. Res.* 40, 336–345. doi:10.1038/hr.2016.164

Chow, W.-H., Chrisman, M., Daniel, C. R., Ye, Y., Gomez, H., Dong, Q., et al. (2017). Cohort Profile: The Mexican American Mano a Mano Cohort. *Int. J. Epidemiol.* 46, e3. doi:10.1093/ije/dyv016

Ciaramelli, C., Fumagalli, M., Viglio, S., Bardoni, A. M., Piloni, D., Meloni, F., et al. (2017). 1H NMR to Evaluate the Metabolome of Bronchoalveolar Lavage Fluid (BALf) in Bronchiolitis Obliterans Syndrome (BOS): Toward the Development of a New Approach for Biomarker Identification. *J. Proteome Res.* 16, 1669–1682. doi:10.1021/acs.jproteome.6b01038

Clayton, E., Taylor, S., Wright, B., and Wilson, I. D. (1998). The Application of High Performance Liquid Chromatography, Coupled to Nuclear Magnetic Resonance Spectroscopy and Mass Spectrometry (HPLC-NMR-MS), to the Characterisation of Ibuprofen Metabolites from Human Urine. *Chromatographia* 47, 264–270. doi:10.1007/BF02466530

Clendinen, C. S., Lee-McMullen, B., Williams, C. M., Stupp, G. S., Vandenborne, K., Hahn, D. A., et al. (2014). 13C NMR Metabolomics: Applications at Natural Abundance. *Anal. Chem.* 86, 9242–9250. doi:10.1021/ac502346h

Coman, C., Solari, F. A., Hentschel, A., Sickmann, A., Zahedi, R. P., and Ahrends, R. (2016). Simultaneous Metabolite, Protein, Lipid Extraction (SIMPLEX): A Combinatorial Multimolecular Omics Approach for Systems Biology. *Mol. Cell Proteom.* 15, 1435–1466. doi:10.1074/mcp.M115.053702

Crockford, D. J., Holmes, E., Lindon, J. C., Plumb, R. S., Zirah, S., Bruce, S. J., et al. (2006). Statistical Heterospectroscopy, an Approach to the Integrated Analysis of NMR and UPLC-MS Data Sets: Application in Metabonomic Toxicology Studies. *Anal. Chem.* 78, 363–371. doi:10.1021/ac051444m

Crockford, D. J., Maher, A. D., Ahmadi, K. R., Barrett, A., Plumb, R. S., Wilson, I. D., et al. (2008). 1H NMR and UPLC-MSE Statistical Heterospectroscopy: Characterization of Drug Metabolites (Xenometabolome) in Epidemiological Studies. *Anal. Chem.* 80, 6835–6844. doi:10.1021/ac801075m

Crown, S. B., and Antoniewicz, M. R. (2013). Publishing 13C Metabolic Flux Analysis Studies: A Review and Future Perspectives. *Metab. Eng.* 20, 42–48. doi:10.1016/j.ymben.2013.08.005

Cui, L., Lu, H., and Lee, Y. H. (2018). Challenges and Emergent Solutions for LC-MS/MS Based Untargeted Metabolomics in Diseases. *Mass. Spec. Rev.* 37, 772–792. doi:10.1002/mas.21562

Cuperlovic-Culf, M., and Culf, A. S. (2016). Applied Metabolomics in Drug Discovery. *Expert Opin. Drug Discov.* 11, 759–770. doi:10.1080/17460441.2016.1195365

Curtarello, M., Tognon, M., Venturoli, C., Silic-Benussi, M., Grassi, A., Verza, M., et al. (2019). Rewiring of Lipid Metabolism and Storage in Ovarian Cancer Cells after Anti-VEGF Therapy. *Cells* 8, 1601. doi:10.3390/cells8121601

Dear, G. J., Ayrton, J., Plumb, R., Sweatman, B. C., Ismail, I. M., Fraser, I. J., et al. (1998). A Rapid and Efficient Approach to Metabolite Identification Using Nuclear Magnetic Resonance Spectroscopy, Liquid Chromatography/mass Spectrometry and Liquid Chromatography/nuclear Magnetic Resonance Spectroscopy/sequential Mass Spectrometry. *Rapid Commun. Mass. Spectrom.* 12, 2023–2030.

Debik, J., Euceda, L. R., Lundgren, S., Gythfeldt, H. v. d. L., Garred, O., Borgen, E., et al. (2019). Assessing Treatment Response and Prognosis by Serum and Tissue Metabolomics in Breast Cancer Patients. *J. Proteome Res.* 18, 3649–3660. doi:10.1021/acs.jproteome.9b00316

Deelen, J., Kettunen, J., Fischer, K., van der Spek, A., Trompet, S., Kastenmüller, G., et al. (2019). A Metabolic Profile of All-Cause Mortality Risk Identified in an Observational Study of 44,168 Individuals. *Nat. Commun.* 10, 3346. doi:10.1038/s41467-019-11311-9

Deng, L., Gu, H., Zhu, J., Nagana Gowda, G. A., Djukovic, D., Chiorean, E. G., et al. (2016). Combining NMR and LC/MS Using Backward Variable Elimination: Metabolomics Analysis of Colorectal Cancer, Polyps, and Healthy Controls. *Anal. Chem.* 88, 7975–7983. doi:10.1021/acs.analchem.6b00885

Dennis, E. A. (2009). Lipidomics Joins the Omics Evolution. *Proc. Natl. Acad. Sci.* 106, 2089–2090. doi:10.1073/pnas.0812636106

Dey, A., Charrier, B., Martineau, E., Deborde, C., Gandriau, E., Moing, A., et al. (2020). Hyperpolarized NMR Metabolomics at Natural 13C Abundance. *Anal. Chem.* 92, 14867–14871. doi:10.1021/acs.analchem.0c03510

Di Sanzo, M., Cipolloni, L., Borro, M., La Russa, R., Santurro, A., Scopetti, M., et al. (2017). Clinical Applications of Personalized Medicine: A New Paradigm and Challenge. *Curr. Pharm. Biotechnol.* 18, 194–203. doi:10.2174/1389201018666170224105600

Doeswijk, T. G., Smilde, A. K., Hageman, J. A., Westerhuis, J. A., and van Eeuwijk, F. A. (2011). On the Increase of Predictive Performance with High-Level Data Fusion. *Analytica Chim. Acta* 705, 41–47. doi:10.1016/j.aca.2011.03.025

Dona, A. C., Kyriakides, M., Scott, F., Shephard, E. A., Varshavi, D., Veselkov, K., et al. (2016). A Guide to the Identification of Metabolites in NMR-Based Metabonomics/metabolomics Experiments. *Comput. Struct. Biotechnol. J.* 14, 135–153. doi:10.1016/j.csbj.2016.02.005

Du, J., Su, Y., Qian, C., Yuan, D., Miao, K., Lee, D., et al. (2020). Raman-guided Subcellular Pharmaco-Metabolomics for Metastatic Melanoma Cells. *Nat. Commun.* 11, 4830. doi:10.1038/s41467-020-18376-x

Duarte, D., Castro, B., Pereira, J. L., Marques, J. F., Costa, A. L., and Gil, A. M. (2020). Evaluation of Saliva Stability for NMR Metabolomics: Collection and Handling Protocols. *Metabolites* 10, 515. doi:10.3390/metabo10120515

Duckett, S. B., and Mewis, R. E. (2013). "Improving NMR and MRI Sensitivity with Parahydrogen," in *Hyperpolarization Methods in NMR Spectroscopy*. Editor L. T. Kuhn (Berlin, Heidelberg: Springer Berlin Heidelberg), 75–103.

Dumez, J.-N., Milani, J., Vuichoud, B., Bornet, A., Lalande-Martin, J., Tea, I., et al. (2015). Hyperpolarized NMR of Plant and Cancer Cell Extracts at Natural Abundance. *Analyst* 140, 5860–5863. doi:10.1039/c5an01203a

Elemento, O. (2020). The Future of Precision Medicine: towards a More Predictive Personalized Medicine. *Emerg. Top. Life Sci.* 4, 175–177. doi:10.1042/ETLS20190197

Elliott, P., Vergnaud, A.-C., Singh, D., Neasham, D., Spear, J., and Heard, A. (2014). The Airwave Health Monitoring Study of Police Officers and Staff in Great Britain: Rationale, Design and Methods. *Environ. Res.* 134, 280–285. doi:10.1016/j.envres.2014.07.025

Emwas, A.-H., Roy, R., McKay, R. T., Ryan, D., Brennan, L., Tenori, L., et al. (2016). Recommendations and Standardization of Biomarker Quantification Using NMR-Based Metabolomics with Particular Focus on Urinary Analysis. *J. Proteome Res.* 15, 360–373. doi:10.1021/acs.jproteome.5b00885

Emwas, A.-H., Roy, R., McKay, R. T., Tenori, L., Saccenti, E., Gowda, G. A. N., et al. (2019). NMR Spectroscopy for Metabolomics Research. *Metabolites* 9, 123. doi:10.3390/metabo9070123

Everett, J. R. (2017). NMR-based Pharmacometabonomics: A New Paradigm for Personalised or Precision Medicine. *Prog. Nucl. Magn. Reson. Spectrosc.* 102-103, 1–14. doi:10.1016/j.pnmrs.2017.04.003

Fan, T. W., Lane, A. N., Higashi, R. M., Farag, M. A., Gao, H., Bousamra, M., et al. (2009). Altered Regulation of Metabolic Pathways in Human Lung Cancer Discerned by 13C Stable Isotope-Resolved Metabolomics (SIRM). *Mol. Cancer* 8, 41. doi:10.1186/1476-4598-8-41

Fan, T. W.-M., Lorkiewicz, P. K., Sellers, K., Moseley, H. N. B., Higashi, R. M., and Lane, A. N. (2012). Stable Isotope-Resolved Metabolomics and Applications for Drug Development. *Pharmacol. Ther.* 133, 366–391. doi:10.1016/j.pharmthera.2011.12.007

Féraud, B., Govaerts, B., Verleysen, M., and de Tullio, P. (2015). Statistical Treatment of 2D NMR COSY Spectra in Metabolomics: Data Preparation, Clustering-Based Evaluation of the Metabolomic Informative Content and Comparison with 1H-NMR. *Metabolomics* 11, 1756–1768. doi:10.1007/s11306-015-0830-7

Féraud, B., Leenders, J., Martineau, E., Giraudeau, P., Govaerts, B., and de Tullio, P. (2019). Two Data Pre-processing Workflows to Facilitate the Discovery of Biomarkers by 2D NMR Metabolomics. *Metabolomics* 15, 63. doi:10.1007/s11306-019-1524-3

Féraud, B., Martineau, E., Leenders, J., Govaerts, B., de Tullio, P., and Giraudeau, P. (2020). Combining Rapid 2D NMR Experiments with Novel Pre-processing Workflows and MIC Quality Measures for Metabolomics. *Metabolomics* 16, 42. doi:10.1007/s11306-020-01662-6

Fei, Q., Wang, D., Jasbi, P., Zhang, P., Nagana Gowda, G. A., Raftery, D., et al. (2019). Combining NMR and MS with Chemical Derivatization for Absolute Quantification with Reduced Matrix Effects. *Anal. Chem.* 91, 4055–4062. doi:10.1021/acs.analchem.8b05611

Finch, G., Yilmaz, A., and Utz, M. (2016). An Optimised Detector for *In-Situ* High-Resolution NMR in Microfluidic Devices. *J. Magn. Reson.* 262, 73–80. doi:10.1016/j.jmr.2015.11.011

Frédérich, M., Pirotte, B., Fillet, M., and de Tullio, P. (2016). Metabolomics as a Challenging Approach for Medicinal Chemistry and Personalized Medicine. *J. Med. Chem.* 59, 8649–8666. doi:10.1021/acs.jmedchem.5b01335

Gebretsadik, T., Linert, W., Thomas, M., Berhanu, T., and Frew, R. (2019). LC-NMR for Natural Products Analysis: A Journey from an Academic Curiosity to a Robust Analytical Tool. *Sci* 1, 31. doi:10.3390/sci1010031

Gil, M., Samino, S., Barrilero, R., and Correig, X. (2019). "Lipid Profiling Using 1H NMR Spectroscopy," in *NMR-Based Metabolomics: Methods and Protocols Methods in Molecular Biology*. Editors G. A. N. Gowda and D. Raftery (New York, NY: Springer), 35–47. doi:10.1007/978-1-4939-9690-2_3

Giraudeau, P., and Frydman, L. (2014). Ultrafast 2D NMR: An Emerging Tool in Analytical Spectroscopy. *Annu. Rev. Anal. Chem.* 7, 129–161. doi:10.1146/annurev-anchem-071213-020208

Giraudeau, P. (2020). NMR-based Metabolomics and Fluxomics: Developments and Future Prospects. *Analyst* 145, 2457–2472. doi:10.1039/D0AN00142B

Giskeødegård, G. F., Andreassen, T., Bertilsson, H., Tessem, M.-B., and Bathen, T. F. (2019). The Effect of Sampling Procedures and Day-To-Day Variations in Metabolomics Studies of Biofluids. *Analytica Chim. Acta* 1081, 93–102. doi:10.1016/j.aca.2019.07.026

Gogiashvili, M., Nowacki, J., Hergenröder, R., Hengstler, J. G., Lambert, J., and Edlund, K. (2019). HR-MAS NMR Based Quantitative Metabolomics in Breast Cancer. *Metabolites* 9, 19. doi:10.3390/metabo9020019

González-Ruiz, V., Schvartz, D., Sandström, J., Pezzatti, J., Jeanneret, F., Tonoli, D., et al. (2019). An Integrative Multi-Omics Workflow to Address Multifactorial Toxicology Experiments. *Metabolites* 9, 79. doi:10.3390/metabo9040079

Gouilleux, B., Marchand, J., Charrier, B., Remaud, G. S., and Giraudeau, P. (2018). High-throughput Authentication of Edible Oils with Benchtop Ultrafast 2D NMR. *Food Chem.* 244, 153–158. doi:10.1016/j.foodchem.2017.10.016

Gouilleux, B., Farjon, J., and Giraudeau, P. (2020). Gradient-based Pulse Sequences for Benchtop NMR Spectroscopy. *J. Magn. Reson.* 319, 106810. doi:10.1016/j.jmr.2020.106810

Goulitquer, S., Croyal, M., Lalande, J., Royer, A.-L., Guitton, Y., Arzur, D., et al. (2018). Consequences of Blunting the Mevalonate Pathway in Cancer Identified by a Pluri-Omics Approach. *Cell Death Dis.* 9, 1–12. doi:10.1038/s41419-018-0761-0

Gowda, G. A. N., and Djukovic, D. (2014). Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. *Methods Mol. Biol.* 1198, 3–12. doi:10.1007/978-1-4939-1258-2_1

Gross, R. W., and Han, X. (2011). Lipidomics at the Interface of Structure and Function in Systems Biology. *Chem. Biol.* 18, 284–291. doi:10.1016/j.chembiol.2011.01.014

Gu, H., Pan, Z., Xi, B., Asiago, V., Musselman, B., and Raftery, D. (2011). Principal Component Directed Partial Least Squares Analysis for Combining Nuclear Magnetic Resonance and Mass Spectrometry Data in Metabolomics: Application to the Detection of Breast Cancer. *Analytica Chim. Acta* 686, 57–63. doi:10.1016/j.aca.2010.11.040

Guo, R., Chen, Y., Borgard, H., Jijiwa, M., Nasu, M., He, M., et al. (2020). The Function and Mechanism of Lipid Molecules and Their Roles in the Diagnosis and Prognosis of Breast Cancer. *Molecules* 25, 4864. doi:10.3390/molecules25204864

Hao, J., Astle, W., De Iorio, M., and Ebbels, T. M. D. (2012). BATMAN--an R Package for the Automated Quantification of Metabolites from Nuclear Magnetic Resonance Spectra Using a Bayesian Model. *Bioinformatics* 28, 2088–2090. doi:10.1093/bioinformatics/bts308

Hermkens, N. K. J., Eshuis, N., van Weerdenburg, B. J. A., Feiters, M. C., Rutjes, F. P. J. T., Wijmenga, S. S., et al. (2016). NMR-Based Chemosensing via P-H2 Hyperpolarization: Application to Natural Extracts. *Anal. Chem.* 88, 3406–3412. doi:10.1021/acs.analchem.6b00184

Holzgrabe, U., Deubner, R., Schollmayer, C., and Waibel, B. (2005). Quantitative NMR Spectroscopy-Applications in Drug Analysis. *J. Pharm. Biomed. Anal.* 38, 806–812. doi:10.1016/j.jpba.2005.01.050

Hu, R., Li, T., Yang, Y., Tian, Y., and Zhang, L. (2021). "NMR-Based Metabolomics in Cancer Research," in *Cancer Metabolomics: Methods And Applications Advances in Experimental Medicine and Biology*. Editor S. Hu (Cham: Springer International Publishing), 201–218. doi:10.1007/978-3-030-51652-9_14

Huart, J., Leenders, J., Taminiau, B., Descy, J., Saint-Remy, A., Daube, G., et al. (2019). Gut Microbiota and Fecal Levels of Short-Chain Fatty Acids Differ upon 24-Hour Blood Pressure Levels in Men. *Hypertension* 74, 1005–1013. doi:10.1161/HYPERTENSIONAHA.118.12588

Huart, J., Cirillo, A., Saint-Remy, A., Krzesinski, J.-M., de Tullio, P., and Jouret, F. (2021a). The Faecal Abundance of Short Chain Fatty Acids Is Increased in Men with a Non-dipping Blood Pressure Profile. *Acta Cardiol.* 76, 1–4. doi:10.1080/00015385.2021.1901020

Huart, J., Cirillo, A., Taminiau, B., Descy, J., Saint-Remy, A., Daube, G., et al. (2021b). Human Stool Metabolome Differs upon 24 H Blood Pressure Levels and Blood Pressure Dipping Status: A Prospective Longitudinal Study. *Metabolites* 11, 282. doi:10.3390/metabo11050282

Izquierdo-Garcia, J. L., Comella-del-Barrio, P., Campos-Olivas, R., Villar-Hernández, R., Prat-Aymerich, C., De Souza-Galvão, M. L., et al. (2020). Discovery and Validation of an NMR-Based Metabolomic Profile in Urine as TB Biomarker. *Sci. Rep.* 10, 22317. doi:10.1038/s41598-020-78999-4

Jacob, M., Lopata, A. L., Dasouki, M., and Abdel Rahman, A. M. (2019). Metabolomics toward Personalized Medicine. *Mass. Spec. Rev.* 38, 221–238. doi:10.1002/mas.21548

Ji, X., Bornet, A., Vuichoud, B., Milani, J., Gajan, D., Rossini, A. J., et al. (2017). Transportable Hyperpolarized Metabolites. *Nat. Commun.* 8, 13975. doi:10.1038/ncomms13975

Jiménez, B., Holmes, E., Heude, C., Tolson, R. F., Harvey, N., Lodge, S. L., et al. (2018). Quantitative Lipoprotein Subclass and Low Molecular Weight Metabolite Analysis in Human Serum and Plasma by 1H NMR Spectroscopy in a Multilaboratory Trial. *Anal. Chem.* 90, 11962–11971. doi:10.1021/acs.analchem.8b02412

Jobard, E., Trédan, O., Bachelot, T., Vigneron, A. M., Aït-Oukhatar, C. M., Arnedos, M., et al. (2017). Longitudinal Serum Metabolomics Evaluation of Trastuzumab and Everolimus Combination as Pre-operative Treatment for

HER-2 Positive Breast Cancer Patients. *Oncotarget* 8, 83570–83584. doi:10.18632/oncotarget.18784

Johnson, H., Puppa, M., Merwe, M., and Tipirneni-Sajja, A. (2021). CRAFT for NMR Lipidomics: Targeting Lipid Metabolism in Leucine-supplemented Tumor-bearing Mice. *Magn. Reson. Chem.* 59, 138–146. doi:10.1002/mrc.5092

Jung, Y.-S., Hyeon, J.-S., and Hwang, G.-S. (2016). Software-assisted Serum Metabolite Quantification Using NMR. *Analytica Chim. Acta* 934, 194–202. doi:10.1016/j.aca.2016.04.054

Jurowski, K., Kochan, K., Walczak, J., Barańska, M., Piekoszewski, W., and Buszewski, B. (2017). Analytical Techniques in Lipidomics: State of the Art. *Crit. Rev. Anal. Chem.* 47, 418–437. doi:10.1080/10408347.2017.1310613

Karaman, İ., Nørskov, N. P., Yde, C. C., Hedemann, M. S., Bach Knudsen, K. E., and Kohler, A. (2015). Sparse Multi-Block PLSR for Biomarker Discovery when Integrating Data from LC-MS and NMR Metabolomics. *Metabolomics* 11, 367–379. doi:10.1007/s11306-014-0698-y

Kelly, R. S., Giorgio, R. T., Chawes, B. L., Palacios, N. I., Gray, K. J., Mirzakhani, H., et al. (2017). Applications of Metabolomics in the Study and Management of Preeclampsia: a Review of the Literature. *Metabolomics* 13, 86. doi:10.1007/s11306-017-1225-8

Khakimov, B., Mobaraki, N., Trimigno, A., Aru, V., and Engelsen, S. B. (2020). Signature Mapping (SigMa): An Efficient Approach for Processing Complex Human Urine 1H NMR Metabolomics Data. *Analytica Chim. Acta* 1108, 142–151. doi:10.1016/j.aca.2020.02.025

Kim, H. K., Kostidis, S., and Choi, Y. H. (2018). NMR Analysis of Fecal Samples. *Methods Mol. Biol.* 1730, 317–328. doi:10.1007/978-1-4939-7592-1_24

Kohler, I., Verhoeven, A., Derks, R. J., and Giera, M. (2016). Analytical Pitfalls and Challenges in Clinical Metabolomics. *Bioanalysis* 8, 1509–1532. doi:10.4155/bio-2016-0090

Kohler, I., Hankemeier, T., van der Graaf, P. H., Knibbe, C. A. J., and van Hasselt, J. G. C. (2017). Integrating Clinical Metabolomics-Based Biomarker Discovery and Clinical Pharmacology to Enable Precision Medicine. *Eur. J. Pharm. Sci.* 109, S15–S21. doi:10.1016/j.ejps.2017.05.018

Kosmides, A. K., Kamisoglu, K., Calvano, S. E., Corbett, S. A., and Androulakis, I. P. (2013). Metabolomic Fingerprinting: Challenges and Opportunities. *Crit. Rev. Biomed. Eng.* 41, 205–221. doi:10.1615/critrevbiomedeng.2013007736

Kostara, C. E., Tsimihodimos, V., Elisaf, M. S., and Bairaktari, E. T. (2017). NMR-Based Lipid Profiling of High Density Lipoprotein Particles in Healthy Subjects with Low, Normal, and Elevated HDL-Cholesterol. *J. Proteome Res.* 16, 1605–1616. doi:10.1021/acs.jproteome.6b00975

Kostidis, S., Addie, R. D., Morreau, H., Mayboroda, O. A., and Giera, M. (2017). Quantitative NMR Analysis of Intra- and Extracellular Metabolism of Mammalian Cells: A Tutorial. *Analytica Chim. Acta* 980, 1–24. doi:10.1016/j.aca.2017.05.011

Kovacs, H., Moskau, D., and Spraul, M. (2005). Cryogenically Cooled Probes-A Leap in NMR Technology. *Prog. Nucl. Magn. Reson. Spectrosc.* 46, 131–155. doi:10.1016/j.pnmrs.2005.03.001

Kumar, R., Bohra, A., Pandey, A. K., Pandey, M. K., and Kumar, A. (2017). Metabolomics for Plant Improvement: Status and Prospects. *Front. Plant Sci.* 8, 1302. doi:10.3389/fpls.2017.01302

Küster, S. K., Danieli, E., Blümich, B., and Casanova, F. (2011). High-resolution NMR Spectroscopy under the Fume Hood. *Phys. Chem. Chem. Phys.* 13, 13172–13176. doi:10.1039/c1cp21180c

Labaki, W. W., Gu, T., Murray, S., Curtis, J. L., Yeomans, L., Bowler, R. P., et al. (2019). Serum Amino Acid Concentrations and Clinical Outcomes in Smokers: SPIROMICS Metabolomics Study. *Sci. Rep.* 9, 11367. doi:10.1038/s41598-019-47761-w

Lacy, P., McKay, R. T., Finkel, M., Karnovsky, A., Woehler, S., Lewis, M. J., et al. (2014). Signal Intensities Derived from Different NMR Probes and Parameters Contribute to Variations in Quantification of Metabolites. *PLoS ONE* 9, e85732. doi:10.1371/journal.pone.0085732

Lambert, V., Hansen, S., Schoumacher, M., Lecomte, J., Leenders, J., Hubert, P., et al. (2020). Pyruvate Dehydrogenase Kinase/lactate axis: a Therapeutic Target for Neovascular Age-Related Macular Degeneration Identified by Metabolomics. *J. Mol. Med.* 98, 1737–1751. doi:10.1007/s00109-020-01994-9

Lane, A. N., and Fan, T. W.-M. (2017). NMR-based Stable Isotope Resolved Metabolomics in Systems Biochemistry. *Arch. Biochem. Biophys.* 628, 123–131. doi:10.1016/j.abb.2017.02.009

Lane, A. N., Fan, T. W.-M., Bousamra, M., Higashi, R. M., Yan, J., and Miller, D. M. (2011). Stable Isotope-Resolved Metabolomics (SIRM) in Cancer Research with Clinical Application to NonSmall Cell Lung Cancer. *OMICS: A J. Integr. Biol.* 15, 173–182. doi:10.1089/omi.2010.0088

Lane, A. N., Higashi, R. M., and Fan, T. W.-M. (2016). Preclinical Models for Interrogating Drug Action in Human Cancers Using Stable Isotope Resolved Metabolomics (SIRM). *Metabolomics* 12, 118. doi:10.1007/s11306-016-1065-y

Lane, A. N., Higashi, R. M., and Fan, T. W.-M. (2019). NMR and MS-based Stable Isotope-Resolved Metabolomics and Applications in Cancer Metabolism. *TrAC Trends Anal. Chem.* 120, 115322. doi:10.1016/j.trac.2018.11.020

Le Guennec, A., Tea, I., Antheaume, I., Martineau, E., Charrier, B., Pathan, M., et al. (2012). Fast Determination of Absolute Metabolite Concentrations by Spatially Encoded 2D NMR: Application to Breast Cancer Cell Extracts. *Anal. Chem.* 84, 10831–10837. doi:10.1021/ac3033504

Leaptrot, K. L., May, J. C., Dodds, J. N., and McLean, J. A. (2019). Ion Mobility Conformational Lipid Atlas for High Confidence Lipidomics. *Nat. Commun.* 10, 985. doi:10.1038/s41467-019-08897-5

Lee, K.-M., Jeon, J.-Y., Lee, B.-J., Lee, H., and Choi, H.-K. (2017). Application of Metabolomics to Quality Control of Natural Product Derived Medicines. *Biomol. Ther.* 25, 559–568. doi:10.4062/biomolther.2016.249

Leenders, J., Grootveld, M., Percival, B., Gibson, M., Casanova, F., and Wilson, P. B. (2020). Benchtop Low-Frequency 60 MHz NMR Analysis of Urine: A Comparative Metabolomics Investigation. *Metabolites* 10, 155. doi:10.3390/metabo10040155

Lerche, M. H., Yigit, D., Frahm, A. B., Ardenkjær-Larsen, J. H., Malinowski, R. M., and Jensen, P. R. (2017). Stable Isotope-Resolved Analysis with Quantitative Dissolution Dynamic Nuclear Polarization. *Anal. Chem.* 90, 674–678. doi:10.1021/acs.analchem.7b02779

Letertre, M. P. M., Dervilly, G., and Giraudeau, P. (2021). Combined Nuclear Magnetic Resonance Spectroscopy and Mass Spectrometry Approaches for Metabolomics. *Anal. Chem.* 93, 500–518. doi:10.1021/acs.analchem.0c04371

Li, B., He, X., Jia, W., and Li, H. (2017a). Novel Applications of Metabolomics in Personalized Medicine: A Mini-Review. *Molecules* 22, 1173. doi:10.3390/molecules22071173

Li, J., Vosegaard, T., and Guo, Z. (2017b). Applications of Nuclear Magnetic Resonance in Lipid Analyses: An Emerging Powerful Tool for Lipidomics Studies. *Prog. Lipid Res.* 68, 37–56. doi:10.1016/j.plipres.2017.09.003

Li, S., Tian, Y., Jiang, P., Lin, Y., Liu, X., and Yang, H. (2021). Recent Advances in the Application of Metabolomics for Food Safety Control and Food Quality Analyses. *Crit. Rev. Food Sci. Nutr.* 61, 1448–1469. doi:10.1080/10408398.2020.1761287

Lindon, J. C., Nicholson, J. K., and Wilson, I. D. (2000). Directly Coupled HPLC-NMR and HPLC-NMR-MS in Pharmaceutical Research and Development. *J. Chromatogr. B: Biomed. Sci. Appl.* 748, 233–258. doi:10.1016/s0378-4347(00)00320-0

Lindon, J. C., Nicholson, J. K., and Wilson, I. D. (2002). "Biomedical and Pharmaceutical Applications of HPLC-NMR and HPLC-NMR-MS," in *On-Line LC-NMR and Related Techniques* (New-York, NY: John Wiley & Sons), 45–87. doi:10.1002/0470854820.ch3

Lloyd, L. S., Adams, R. W., Bernstein, M., Coombes, S., Duckett, S. B., Green, G. G. R., et al. (2012). Utilization of SABRE-Derived Hyperpolarization to Detect Low-Concentration Analytes via 1D and 2D NMR Methods. *J. Am. Chem. Soc.* 134, 12904–12907. doi:10.1021/ja3051052

Locci, E., Noto, A., Puddu, M., Pomero, G., Demontis, R., Dalmazzo, C., et al. (2018). A Longitudinal 1H-NMR Metabolomics Analysis of Urine from Newborns with Hypoxic-Ischemic Encephalopathy Undergoing Hypothermia Therapy. Clinical and Medical Legal Insights. *PLOS ONE* 13, e0194267. doi:10.1371/journal.pone.0194267

Lopez, J. M., Cabrera, R., and Maruenda, H. (2019). Ultra-Clean Pure Shift 1H-NMR Applied to Metabolomics Profiling. *Sci. Rep.* 9, 6900. doi:10.1038/s41598-019-43374-5

Lucas-Torres, C., Roumes, H., Bouchaud, V., Bouzier-Sore, A. K., and Wong, A. (2021). Metabolic NMR Mapping with Microgram Tissue Biopsy. *NMR Biomed.* 34, e4477. doi:10.1002/nbm.4477

Lydic, T. A., and Goo, Y. H. (2018). Lipidomics Unveils the Complexity of the Lipidome in Metabolic Diseases. *Clin. Transl. Med.* 7, 4. doi:10.1186/s40169-018-0182-9

Mahrous, E. A., and Farag, M. A. (2015). Two Dimensional NMR Spectroscopic Approaches for Exploring Plant Metabolome: A Review. *J. Adv. Res.* 6, 3–15. doi:10.1016/j.jare.2014.10.003

Maier, T. V., and Schmitt-Kopplin, P. (2016). Capillary Electrophoresis in Metabolomics. *Methods Mol. Biol. Clifton NJ* 1483, 437–470. doi:10.1007/978-1-4939-6403-1_21

Maltesen, R. G., Wimmer, R., and Rasmussen, B. S. (2020). A Longitudinal Serum NMR-Based Metabolomics Dataset of Ischemia-Reperfusion Injury in Adult Cardiac Surgery. *Sci. Data* 7, 198. doi:10.1038/s41597-020-0545-0

Manzoni, C., Kia, D. A., Vandrovcova, J., Hardy, J., Wood, N. W., Lewis, P. A., et al. (2018). Genome, Transcriptome and Proteome: the Rise of Omics Data and Their Integration in Biomedical Sciences. *Brief. Bioinform.* 19, 286–302. doi:10.1093/bib/bbw114

Marchand, J., Martineau, E., Guitton, Y., Dervilly-Pinel, G., and Giraudeau, P. (2017). Multidimensional NMR Approaches towards Highly Resolved, Sensitive and High-Throughput Quantitative Metabolomics. *Curr. Opin. Biotechnol.* 43, 49–55. doi:10.1016/j.copbio.2016.08.004

Marchand, J., Martineau, E., Guitton, Y., Le Bizec, B., Dervilly-Pinel, G., and Giraudeau, P. (2018). A Multidimensional 1H NMR Lipidomics Workflow to Address Chemical Food Safety Issues. *Metabolomics* 14, 60. doi:10.1007/s11306-018-1360-x

Markley, J. L., Brüschweiler, R., Edison, A. S., Eghbalnia, H. R., Powers, R., Raftery, D., et al. (2017). The Future of NMR-Based Metabolomics. *Curr. Opin. Biotechnol.* 43, 34–40. doi:10.1016/j.copbio.2016.08.001

Markley, J. L., Dashti, H., Wedell, J. R., Westler, W. M., and Eghbalnia, H. R. (2019). "Tools for Enhanced NMR-Based Metabolomics Analysis," in *NMR-Based Metabolomics: Methods And Protocols Methods in Molecular Biology*. Editors G. A. N. Gowda and D. Raftery (New York, NY: Springer), 413–427. doi:10.1007/978-1-4939-9690-2_23

Marshall, D. D., and Powers, R. (2017). Beyond the Paradigm: Combining Mass Spectrometry and Nuclear Magnetic Resonance for Metabolomics. *Prog. Nucl. Magn. Reson. Spectrosc.* 100, 1–16. doi:10.1016/j.pnmrs.2017.01.001

Marshall, D. D., Lei, S., Worley, B., Huang, Y., Garcia-Garcia, A., Franco, R., et al. (2015). Combining DI-ESI-MS and NMR Datasets for Metabolic Profiling. *Metabolomics* 11, 391–402. doi:10.1007/s11306-014-0704-4

Martineau, E., Giraudeau, P., Tea, I., and Akoka, S. (2011). Fast and Precise Quantitative Analysis of Metabolic Mixtures by 2D 1H INADEQUATE NMR. *J. Pharm. Biomed. Anal.* 54, 252–257. doi:10.1016/j.jpba.2010.07.046

Martineau, E., Dumez, J.-N., and Giraudeau, P. (2020). Fast Quantitative 2D NMR for Metabolomics and Lipidomics: A Tutorial. *Magn. Reson. Chem.* 58, 390–403. doi:10.1002/mrc.4899

Massou, S., Nicolas, C., Letisse, F., and Portais, J.-C. (2007). NMR-based Fluxomics: Quantitative 2D NMR Methods for Isotopomers Analysis. *Phytochemistry* 68, 2330–2340. doi:10.1016/j.phytochem.2007.03.011

Matheus, N., Hansen, S., Rozet, E., Peixoto, P., Maquoi, E., Lambert, V., et al. (2014). An Easy, Convenient Cell and Tissue Extraction Protocol for Nuclear Magnetic Resonance Metabolomics. *Phytochem. Anal.* 25, 342–349. doi:10.1002/pca.2498

Mazurek, S., Boschek, C. B., and Eigenbrodt, E. (1997). The Role of Phosphometabolites in Cell Proliferation, Energy Metabolism, and Tumor Therapy. *J. Bioenerg. Biomembr.* 29, 315–330. doi:10.1023/a:1022490512705

Mili, M., Panthu, B., Madec, A.-M., Berger, M.-A., Rautureau, G. J. P., and Elena-Herrmann, B. (2020). Fast and Ergonomic Extraction of Adherent Mammalian Cells for NMR-Based Metabolomics Studies. *Anal. Bioanal. Chem.* 412, 5453–5463. doi:10.1007/s00216-020-02764-9

Millard, P., Cahoreau, E., Heuillet, M., Portais, J.-C., and Lippens, G. (2017). 15N-NMR-Based Approach for Amino Acids-Based 13C-Metabolic Flux Analysis of Metabolism. *Anal. Chem.* 89, 2101–2106. doi:10.1021/acs.analchem.6b04767

Moayyeri, A., Hammond, C. J., Valdes, A. M., and Spector, T. D. (2013). Cohort Profile: TwinsUK and Healthy Ageing Twin Study. *Int. J. Epidemiol.* 42, 76–85. doi:10.1093/ije/dyr207

Mobli, M., and Hoch, J. C. (2014). Nonuniform Sampling and Non-Fourier Signal Processing Methods in Multidimensional NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 83, 21–41. doi:10.1016/j.pnmrs.2014.09.002

Nagana Gowda, G. A., Gowda, Y. N., and Raftery, D. (2015a). Expanding the Limits of Human Blood Metabolite Quantitation Using NMR Spectroscopy. *Anal. Chem.* 87, 706–715. doi:10.1021/ac503651e

Nagana Gowda, G. A., Gowda, Y. N., and Raftery, D. (2015b). Massive Glutamine Cyclization to Pyroglutamic Acid in Human Serum Discovered Using NMR Spectroscopy. *Anal. Chem.* 87, 3800–3805. doi:10.1021/ac504435b

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

Nelson, S. J., Kurhanewicz, J., Vigneron, D. B., Larson, P. E. Z., Harzstark, A. L., Ferrone, M., et al. (2013). Metabolic Imaging of Patients with Prostate Cancer Using Hyperpolarized [1-13C]Pyruvate. *Sci. Transl. Med.* 5, 198ra108. doi:10.1126/scitranslmed.3006070

Nicholson, J. K., and Lindon, J. C. (2008). Metabonomics. *Nature* 455, 1054–1056. doi:10.1038/4551054a

Niedenführ, S., Wiechert, W., and Nöh, K. (2015). How to Measure Metabolic Fluxes: a Taxonomic Guide for 13 C Fluxomics. *Curr. Opin. Biotechnol.* 34, 82–90. doi:10.1016/j.copbio.2014.12.003

Nielsen, J. (2017). Systems Biology of Metabolism: A Driver for Developing Personalized and Precision Medicine. *Cel Metab.* 25, 572–579. doi:10.1016/j.cmet.2017.02.002

Njock, G. B. B., Pegnyemb, D. E., Bartholomeusz, T. A., Christen, P., Vitorge, B., Nuzillard, J.-M., et al. (2010). Spectral Aliasing: A Super Zoom for 2D-NMR Spectra. Principles and Applications. *Chimia* 64, 235–240. doi:10.2533/chimia.2010.235

Ouldamer, L., Nadal-Desbarats, L., Chevalier, S., Body, G., Goupille, C., and Bougnoux, P. (2016). NMR-Based Lipidomic Approach to Evaluate Controlled Dietary Intake of Lipids in Adipose Tissue of a Rat Mammary Tumor Model. *J. Proteome Res.* 15, 868–878. doi:10.1021/acs.jproteome.5b00788

Paglia, G., Kliman, M., Claude, E., Geromanos, S., and Astarita, G. (2015). Applications of Ion-Mobility Mass Spectrometry for Lipid Analysis. *Anal. Bioanal. Chem.* 407, 4995–5007. doi:10.1007/s00216-015-8664-8

Palmnas, M., and Vogel, H. (2013). The Future of NMR Metabolomics in Cancer Therapy: Towards Personalizing Treatment and Developing Targeted Drugs? *Metabolites* 3, 373–396. doi:10.3390/metabo3020373

Pan, Z., and Raftery, D. (2007). Comparing and Combining NMR Spectroscopy and Mass Spectrometry in Metabolomics. *Anal. Bioanal. Chem.* 387, 525–527. doi:10.1007/s00216-006-0687-8

Pan, Z., Gu, H., Talaty, N., Chen, H., Shanaiah, N., Hainline, B. E., et al. (2007). Principal Component Analysis of Urine Metabolites Detected by NMR and DESI-MS in Patients with Inborn Errors of Metabolism. *Anal. Bioanal. Chem.* 387, 539–549. doi:10.1007/s00216-006-0546-7

Pang, H., Jia, W., and Hu, Z. (2019). Emerging Applications of Metabolomics in Clinical Pharmacology. *Clin. Pharmacol. Ther.* 106, 544–556. doi:10.1002/cpt.1538

Patra, B., Sharma, M., Hale, W., and Utz, M. (2021). Time-Resolved Non-invasive Metabolomic Monitoring of a Single Cancer Spheroid by Microfluidic NMR. *Sci. Rep.* 11, 53. doi:10.1038/s41598-020-79693-1

Patti, G. J., Yanes, O., and Siuzdak, G. (2012). Metabolomics: the Apogee of the Omics Trilogy. *Nat. Rev. Mol. Cel Biol.* 13, 263–269. doi:10.1038/nrm3314

Percival, B. C., Grootveld, M., Gibson, M., Osman, Y., Molinari, M., Jafari, F., et al. (2018). Low-Field, Benchtop NMR Spectroscopy as a Potential Tool for Point-of-Care Diagnostics of Metabolic Conditions: Validation, Protocols and Computational Models. *High-Throughput* 8, 2. doi:10.3390/ht8010002

Pinu, F. R., Goldansaz, S. A., and Jaine, J. (2019). Translational Metabolomics: Current Challenges and Future Opportunities. *Metabolites* 9, 108. doi:10.3390/metabo9060108

Plainchont, B., Berruyer, P., Dumez, J.-N., Jannin, S., and Giraudeau, P. (2018). Dynamic Nuclear Polarization Opens New Perspectives for NMR Spectroscopy in Analytical Chemistry. *Anal. Chem.* 90, 3639–3650. doi:10.1021/acs.analchem.7b05236

Powers, R. (2014). The Current State of Drug Discovery and a Potential Role for NMR Metabolomics. *J. Med. Chem.* 57, 5860–5870. doi:10.1021/jm401803b

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., et al. (2011). The Human Serum Metabolome. *PLOS ONE* 6, e16957. doi:10.1371/journal.pone.0016957

Puig-Castellví, F., Pérez, Y., Piña, B., Tauler, R., and Alfonso, I. (2018). Comparative Analysis of 1H NMR and 1H-13C HSQC NMR Metabolomics to Understand the Effects of Medium Composition in Yeast Growth. *Anal. Chem.* 90, 12422–12430. doi:10.1021/acs.analchem.8b01196

Purwaha, P., Silva, L. P., Hawke, D. H., Weinstein, J. N., and Lorenzi, P. L. (2014). An Artifact in LC-MS/MS Measurement of Glutamine and Glutamic Acid: In-Source Cyclization to Pyroglutamic Acid. *Anal. Chem.* 86, 5633–5637. doi:10.1021/ac501451v

Rampler, E., Abiead, Y. E., Schoeny, H., Rusz, M., Hildebrand, F., Fitz, V., et al. (2021). Recurrent Topics in Mass Spectrometry-Based Metabolomics and Lipidomics-Standardization, Coverage, and Throughput. *Anal. Chem.* 93, 519–545. doi:10.1021/acs.analchem.0c04698

Ravanbakhsh, S., Liu, P., Bjorndahl, T., Mandal, R., Grant, J. R., Wilson, M., et al. (2015). Accurate, Fully-Automated NMR Spectral Profiling for Metabolomics. *PLoS ONE* 10, e0124219. doi:10.1371/journal.pone.0124219

Rhodes, C. J. (2017). Magnetic Resonance Spectroscopy. *Sci. Prog.* 100, 241–292. doi:10.3184/003685017X14993478654307

Röhnisch, H. E., Eriksson, J., Müllner, E., Agback, P., Sandström, C., and Moazzami, A. A. (2018). AQuA: An Automated Quantification Algorithm for High-Throughput NMR-Based Metabolomics and its Application in Human Plasma. *Anal. Chem.* 90, 2095–2102. doi:10.1021/acs.analchem.7b04324

Robinette, S. L., Ajredini, R., Rasheed, H., Zeinomar, A., Dossey, F. C., Dossey, A. T., et al. (2011). Hierarchical Alignment and Full Resolution Pattern Recognition of 2D NMR Spectra: Application to Nematode Chemical Ecology. *Anal. Chem.* 83, 1649–1657. doi:10.1021/ac102724x

Romano, F., Meoni, G., Manavella, V., Baima, G., Tenori, L., Cacciatore, S., et al. (2018). Analysis of Salivary Phenotypes of Generalized Aggressive and Chronic Periodontitis through Nuclear Magnetic Resonance-Based Metabolomics. *J. Periodontol.* 89, 1452–1460. doi:10.1002/JPER.18-0097

Rouger, L., Gouilleux, B., and Giraudeau, P. (2016). "Fast N-Dimensional Data Acquisition Methods," in *Encyclopedia of Spectroscopy and Spectrometry*. Editors J. C. Lindon, G. E. Tranter, and D. Koppenaal. 3rd edition (Amsterdam: Elsevier), 588–596.

Ruhland, E., Bund, C., Outilaft, H., Piotto, M., and Namer, I. J. (2019). A Metabolic Database for Biomedical Studies of Biopsy Specimens by High-resolution Magic Angle Spinning Nuclear MR: a Qualitative and Quantitative Tool. *Magn. Reson. Med.* 82, 62–83. doi:10.1002/mrm.27696

Rzeznik, M., Triba, M. N., Levy, P., Jungo, S., Botosoa, E., Duchemann, B., et al. (2017). Identification of a Discriminative Metabolomic Fingerprint of Potential Clinical Relevance in Saliva of Patients with Periodontitis Using 1H Nuclear Magnetic Resonance (NMR) Spectroscopy. *PLOS ONE* 12, e0182767. doi:10.1371/journal.pone.0182767

Sahoo, N. K., Tejaswini, G., Madhusmita, S., and Muralikrishna, K. S. (2020). An Overview on NMR Spectroscopy Based Metabolomics. *Int. J. Pharm. Sci. Dev. Res.* 6, 016–020. doi:10.17352/ijpsdr.000029

Salvia, M.-V., Ramadori, F., Springhetti, S., Diez-Castellnou, M., Perrone, B., Rastrelli, F., et al. (2015). Nanoparticle-Assisted NMR Detection of Organic Anions: From Chemosensing to Chromatography. *J. Am. Chem. Soc.* 137, 886–892. doi:10.1021/ja511205e

Sarafidis, K., Chatziioannou, A. C., Thomaidou, A., Gika, H., Mikros, E., Benaki, D., et al. (2017). Urine Metabolomics in Neonates with Late-Onset Sepsis in a Case-Control Study. *Sci. Rep.* 7, 45506. doi:10.1038/srep45506

Sasaki, K., Sagawa, H., Suzuki, M., Yamamoto, H., Tomita, M., Soga, T., et al. (2019). Metabolomics Platform with Capillary Electrophoresis Coupled with High-Resolution Mass Spectrometry for Plasma Analysis. *Anal. Chem.* 91, 1295–1301. doi:10.1021/acs.analchem.8b02994

Scarfe, G. B., Wright, B., Clayton, E., Taylor, S., Wilson, I. D., Lindon, J. C., et al. (1998). 19F-NMR and Directly Coupled HPLC-NMR-MS Investigations into the Metabolism of 2-bromo-4- Trifluoromethylaniline in Rat: a Urinary Excretion Balance Study without the Use of Radiolabelling. *Xenobiotica* 28, 373–388. doi:10.1080/004982598239489

Scarfe, G. B., Wright, B., Clayton, E., Taylor, S., Wilson, I. D., Lindon, J. C., et al. (1999). Quantitative Studies on the Urinary Metabolic Fate of 2-Chloro-4-Trifluoromethylaniline in the Rat Using 19F-NMR Spectroscopy and Directly Coupled HPLCNMR-MS. *Xenobiotica* 29, 77–91. doi:10.1080/004982599238821

Schanda, P. (2009). Fast-pulsing Longitudinal Relaxation Optimized Techniques: Enriching the Toolbox of Fast Biomolecular NMR Spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* 55, 238–265. doi:10.1016/j.pnmrs.2009.05.002

Schmidt, M. A., and Goodwin, T. J. (2013). Personalized Medicine in Human Space Flight: Using Omics Based Analyses to Develop Individualized Countermeasures that Enhance Astronaut Safety and Performance. *Metabolomics* 9, 1134–1156. doi:10.1007/s11306-013-0556-3

Sellies, L., Reile, I., Aspers, R. L. E. G., Feiters, M. C., Rutjes, F. P. J. T., and Tessari, M. (2019). Parahydrogen Induced Hyperpolarization Provides a Tool for NMR Metabolomics at Nanomolar Concentrations. *Chem. Commun.* 55, 7235–7238. doi:10.1039/C9CC02186H

Sherman, L. M., Petrov, A. P., Karger, L. F. P., Tetrick, M. G., Dovichi, N. J., and Camden, J. P. (2020). A Surface-Enhanced Raman Spectroscopy Database of 63 Metabolites. *Talanta* 210, 120645. doi:10.1016/j.talanta.2019.120645

Shockcor, J. P., Unger, S. E., Wilson, I. D., Foxall, P. J. D., Nicholson, J. K., and Lindon, J. C. (1996). Combined HPLC, NMR Spectroscopy, and Ion-Trap Mass Spectrometry with Application to the Detection and Characterization of Xenobiotic and Endogenous Metabolites in Human Urine. *Anal. Chem.* 68, 4431–4435. doi:10.1021/ac9606463

Shockcor, J. P. (2002). "Application of On-Line LC-NMR and Related Techniques to Drug Metabolism Studies," in *On-Line LC-NMR and Related Techniques* (John Wiley & Sons), 89–108. doi:10.1002/0470854820.ch4

Silva Elipe, M. V. (2003). Advantages and Disadvantages of Nuclear Magnetic Resonance Spectroscopy as a Hyphenated Technique. *Analytica Chim. Acta* 497, 1–25. doi:10.1016/j.aca.2003.08.048

Simón-Manso, Y., Lowenthal, M. S., Kilpatrick, L. E., Sampson, M. L., Telu, K. H., Rudnick, P. A., et al. (2013). Metabolite Profiling of a NIST Standard Reference Material for Human Plasma (SRM 1950): GC-MS, LC-MS, NMR, and Clinical Laboratory Analyses, Libraries, and Web-Based Resources. *Anal. Chem.* 85, 11725–11731. doi:10.1021/ac402503m

Singh, K., and Blümich, B. (2016). NMR Spectroscopy with Compact Instruments. *TrAC Trends Anal. Chem.* 83, 12–26. doi:10.1016/j.trac.2016.02.014

Sliz, E., Kettunen, J., Holmes, M. V., Williams, C. O., Boachie, C., Wang, Q., et al. (2018). Metabolomic Consequences of Genetic Inhibition of PCSK9 Compared with Statin Treatment. *Circulation* 138, 2499–2512. doi:10.1161/CIRCULATIONAHA.118.034942

Snytnikova, O. A., Khlichkina, A. A., Sagdeev, R. Z., and Tsentalovich, Y. P. (2019). Evaluation of Sample Preparation Protocols for Quantitative NMR-Based Metabolomics. *Metabolomics* 15, 84. doi:10.1007/s11306-019-1545-y

Soininen, P., Kangas, A. J., Würtz, P., Suna, T., and Ala-Korpela, M. (2015). Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. *Circ. Cardiovasc. Genet.* 8, 192–206. doi:10.1161/CIRCGENETICS.114.000216

Taherkhani, A., Nafar, M., Arefi-Oskouie, A., Broumandnia, N., Parvin, M., Mahmoudieh, L., et al. (2019). Metabolomic Analysis of Membranous Glomerulonephritis: Identification of a Diagnostic Panel and Pathogenic Pathways. *Arch. Med. Res.* 50, 159–169. doi:10.1016/j.arcmed.2019.08.004

Takis, P. G., Ghini, V., Tenori, L., Turano, P., and Luchinat, C. (2019). Uniqueness of the NMR Approach to Metabolomics. *TrAC Trends Anal. Chem.* 120, 115300. doi:10.1016/j.trac.2018.10.036

Teul, J., Rupérez, F. J., Garcia, A., Vaysse, J., Balayssac, S., Gilard, V., et al. (2009). Improving Metabolite Knowledge in Stable Atherosclerosis Patients by Association and Correlation of GC-MS and 1H NMR Fingerprints. *J. Proteome Res.* 8, 5580–5589. doi:10.1021/pr900668v

Tilgner, M., Vater, T. S., Habbel, P., and Cheng, L. L. (2019). High-Resolution Magic Angle Spinning (HRMAS) NMR Methods in Metabolomics. *Methods Mol. Biol.* 2037, 49–67. doi:10.1007/978-1-4939-9690-2_4

Tolstikov, V., Akmaev, V. R., Sarangarajan, R., Narain, N. R., and Kiebish, M. A. (2017). Clinical Metabolomics: a Pivotal Tool for Companion Diagnostic Development and Precision Medicine. *Expert Rev. Mol. Diagn.* 17, 411–413. doi:10.1080/14737159.2017.1308827

Trivedi, D. K., Hollywood, K. A., and Goodacre, R. (2017). Metabolomics for the Masses: The Future of Metabolomics in a Personalized World. *New Horiz. Transl. Med.* 3, 294–305. doi:10.1016/j.nhtm.2017.06.001

Tynkkynen, T., Wang, Q., Ekholm, J., Anufrieva, O., Ohukainen, P., Vepsäläinen, J., et al. (2019). Proof of Concept for Quantitative Urine NMR Metabolomics Pipeline for Large-Scale Epidemiology and Genetics. *Int. J. Epidemiol.* 48, 978–993. doi:10.1093/ije/dyy287

Ubhi, B. K., Riley, J. H., Shaw, P. A., Lomas, D. A., Tal-Singer, R., MacNee, W., et al. (2012). Metabolic Profiling Detects Biomarkers of Protein Degradation in COPD Patients. *Eur. Respir. J.* 40, 345–355. doi:10.1183/09031936.00112411

van de Weijer, T., and Schrauwen-Hinderling, V. B. (2019). Application of Magnetic Resonance Spectroscopy in Metabolic Research. *Biochim. Biophys. Acta (BBA) - Mol. Basis Dis.* 1865, 741–748. doi:10.1016/j.bbadis.2018.09.013

Van, Q. N., Issaq, H. J., Jiang, Q., Li, Q., Muschik, G. M., Waybright, T. J., et al. (2008). Comparison of 1D and 2D NMR Spectroscopy for Metabolic Profiling. *J. Proteome Res.* 7, 630–639. doi:10.1021/pr700594s

Verhoeven, A., Slagboom, E., Wuhrer, M., Giera, M., and Mayboroda, O. A. (2017). Automated Quantification of Metabolites in Blood-Derived Samples by NMR. *Analytica Chim. Acta* 976, 52–62. doi:10.1016/j.aca.2017.04.013

Viant, M. R. (2009). Applications of Metabolomics to the Environmental Sciences. *Metabolomics* 5, 1–2. doi:10.1007/s11306-009-0157-3

Vignoli, A., Ghini, V., Meoni, G., Licari, C., Takis, P. G., Tenori, L., et al. (2019a). High-Throughput Metabolomics by 1D NMR. *Angew. Chem. Int. Ed.* 58, 968–994. doi:10.1002/anie.201804736

Vignoli, A., Tenori, L., Giusti, B., Takis, P. G., Valente, S., Carrabba, N., et al. (2019b). NMR-based Metabolomics Identifies Patients at High Risk of Death within Two Years after Acute Myocardial Infarction in the AMI-Florence II Cohort. *BMC Med.* 17, 3. doi:10.1186/s12916-018-1240-2

Vignoli, A., Risi, E., McCartney, A., Migliaccio, I., Moretti, E., Malorni, L., et al. (2021). Precision Oncology via NMR-Based Metabolomics: A Review on Breast Cancer. *Int. J. Mol. Sci.* 22, 4687. doi:10.3390/ijms22094687

Wallmeier, J., Samol, C., Ellmann, L., Zacharias, H. U., Vogl, F. C., Garcia, M., et al. (2017). Quantification of Metabolites by NMR Spectroscopy in the Presence of Protein. *J. Proteome Res.* 16, 1784–1796. doi:10.1021/acs.jproteome.7b00057

Wang, C., Zhang, B., Timári, I., Somogyi, Á., Li, D.-W., Adcox, H. E., et al. (2019). Accurate and Efficient Determination of Unknown Metabolites in Metabolomics by NMR-Based Molecular Motif Identification. *Anal. Chem.* 91, 15686–15693. doi:10.1021/acs.analchem.9b03849

Wang, C., Timári, I., Zhang, B., Li, D.-W., Leggett, A., Amer, A. O., et al. (2020). COLMAR Lipids Web Server and Ultrahigh-Resolution Methods for Two-Dimensional Nuclear Magnetic Resonance- and Mass Spectrometry-Based Lipidomics. *J. Proteome Res.* 19, 1674–1683. doi:10.1021/acs.jproteome.9b00845

Welsh, P., Rankin, N., Li, Q., Mark, P. B., Würtz, P., Ala-Korpela, M., et al. (2018). Circulating Amino Acids and the Risk of Macrovascular, Microvascular and Mortality Outcomes in Individuals with Type 2 Diabetes: Results from the ADVANCE Trial. *Diabetologia* 61, 1581–1591. doi:10.1007/s00125-018-4619-x

Wilson, D. M., Keshari, K. R., Larson, P. E. Z., Chen, A. P., Hu, S., Criekinge, M. V., et al. (2010). Multi-compound Polarization by DNP Allows Simultaneous Assessment of Multiple Enzymatic Activities *In Vivo*. *J. Magn. Reson.* 205, 141–147. doi:10.1016/j.jmr.2010.04.012

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

Wishart, D. S. (2019). NMR Metabolomics: A Look Ahead. *J. Magn. Reson.* 306, 155–161. doi:10.1016/j.jmr.2019.07.013

Yao, X., Cao, D., Wang, F., Zhang, W., Ma, C., and Song, M. (2019). An Overview of Omics Approaches to Characterize the Effect of Perfluoroalkyl Substances in Environmental Health. *TrAC Trends Anal. Chem.* 121, 115367. doi:10.1016/j.trac.2018.12.021

Yu, K.-H., and Snyder, M. (2016). Omics Profiling in Precision Oncology. *Mol. Cell Proteomics* 15, 2525–2536. doi:10.1074/mcp.O116.059253

Yu, B., Zanetti, K. A., Temprosa, M., Albanes, D., Appel, N., Barrera, C. B., et al. (2019). The Consortium of Metabolomics Studies (COMETS): Metabolomics in 47 Prospective Cohort Studies. *Am. J. Epidemiol.* 188, 991–1012. doi:10.1093/aje/kwz028

Zacharias, H., Altenbuchinger, M., and Gronwald, W. (2018). Statistical Analysis of NMR Metabolic Fingerprints: Established Methods and Recent Advances. *Metabolites* 8, 47. doi:10.3390/metabo8030047

Zhang, B., Xie, M., Bruschweiler-Li, L., and Brüschweiler, R. (2016). Nanoparticle-Assisted Removal of Protein in Human Serum for Metabolomics Studies. *Anal. Chem.* 88, 1003–1007. doi:10.1021/acs.analchem.5b03889

Zimmermann, M., Zimmermann-Kogadeeva, M., Wegmann, R., and Goodman, A. L. (2019). Mapping Human Microbiome Drug Metabolism by Gut Bacteria and Their Genes. *Nature* 570, 462–467. doi:10.1038/s41586-019-1291-3

Zlatkis, A., and Liebich, H. M. (1971). Profile of Volatile Metabolites in Human Urine. *Clin. Chem.* 17, 592–594. doi:10.1093/clinchem/17.7.592

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.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 02 November 2021
doi: 10.3389/fmolb.2021.708052

NMR-Based Metabolomics Analysis Predicts Response to Neoadjuvant Chemotherapy for Triple-Negative Breast Cancer

*Xiangming He[1,2], Jinping Gu[3], Dehong Zou[2], Hongjian Yang[2], Yongfang Zhang[2], Yuqing Ding[2] and Lisong Teng[1]**

[1]The First Affiliated Hospital, Zhejiang University School of Medicine (FAHZU), Hangzhou, China, [2]Chinese Academy of Sciences, Institute of Basic Medicine and Cancer (IBMC), The Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, China, [3]Key Laboratory for Green Pharmaceutical Technologies and Related Equipment of Ministry of Education, College of Pharmaceutical Sciences, Zhejiang University of Technology, Hangzhou, China

Triple-negative breast cancer (TNBC) is the most fatal type of breast cancer (BC). Due to the lack of relevant targeted drug therapy, in addition to surgery, chemotherapy is still the most common treatment option for TNBC. TNBC is heterogeneous, and different patients have an unusual sensitivity to chemotherapy. Only part of the patients will benefit from chemotherapy, so neoadjuvant chemotherapy (NAC) is controversial in the treatment of TNBC. Here, we performed an NMR spectroscopy–based metabolomics study to analyze the relationship between the patients' metabolic phenotypes and chemotherapy sensitivity in the serum samples. Metabolic phenotypes from patients with pathological partial response, pathological complete response, and pathological stable disease (pPR, pCR, and pSD) could be distinguished. Furthermore, we conducted metabolic pathway analysis based on identified significant metabolites and revealed significantly disturbed metabolic pathways closely associated with three groups of TNBC patients. We evaluated the discriminative ability of metabolites related to significantly disturbed metabolic pathways by using the multi-receiver–operating characteristic (ROC) curve analysis. Three significantly disturbed metabolic pathways of glycine, serine, and threonine metabolism, valine, leucine, and isoleucine biosynthesis, and alanine, aspartate, and glutamate metabolism could be used as potential predictive models to distinguish three types of TNBC patients. These results indicate that a metabolic phenotype could be used to predict whether a patient is suitable for NAC. Metabolomics research could provide data in support of metabolic phenotypes for personalized treatment of TNBC.

Keywords: metabolomics, triple-negative breast cancer, metabolic pathway, ROC curve, NMR spectroscopy

OPEN ACCESS

Edited by:
Danuta Dudzik,
Medical University of Gdańsk, Poland

Reviewed by:
Alma Villaseñor,
CEU San Pablo University, Spain
Sofia Moco,
Vrije Universiteit Amsterdam, Netherlands

****Correspondence:***
Lisong Teng
lsteng@zju.edu.cn

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

Received: *11 May 2021*
Accepted: *10 September 2021*
Published: *02 November 2021*

Citation:
He X, Gu J, Zou D, Yang H, Zhang Y, Ding Y and Teng L (2021) NMR-Based Metabolomics Analysis Predicts Response to Neoadjuvant Chemotherapy for Triple-Negative Breast Cancer. Front. Mol. Biosci. 8:708052. doi: 10.3389/fmolb.2021.708052

INTRODUCTION

According to the American Cancer Society estimates, in the female patients, breast cancer (BC) was the tumor with the highest incidence (about 30%) among the new invasive cancer cases in the US in 2020; in addition, BC had the second highest mortality rate, accounting for 15% among the new cancer death cases (Siegel et al., 2020). Approximately 10–20% of all invasive BC cases were triple-negative breast cancer (TNBC) (Kumar and Aggarwal, 2016). Due to the lack of estrogen receptor

(ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) expression, TNBC lacked standardized treatment strategies (Yin et al., 2020). Chemotherapy is still the most common treatment option for TNBC (Masuda et al., 2017). Neoadjuvant chemotherapy (NAC) is controversial in the treatment of TNBC. A part of the TNBC patients were sensitive to chemotherapy drugs, and about 30–40% of patients' pathological symptoms disappeared completely. This was defined as pathological complete response (pCR) (Liedtke et al., 2008; Gluck et al., 2012). Patients whose pathological symptoms have not changed at all were defined to have pathological stable disease (pSD). Some patients' pathological symptoms were somewhere in between, and this was defined as pathological partial response (pPR). On the contrary, the cancer recurrence rate and metastasis rate of patients with residual disease after NAC have greatly increased (Liedtke et al., 2008; Masuda et al., 2013). After NAC differences in clinical response and survival tips, it is necessary to consider a more detailed classification in clinical TNBC. With the rise of metabolomics research, differences in metabolic phenotypes could provide us with a new idea of NAC for TNBC.

Metabolomics is the study of the multi-parametric metabolic response of living systems to pathophysiological stimuli or genetic modification (Nicholson et al., 1999). Metabolomics is a part of systems biology, which is downstream concerning the other -omic sciences (Vignoli et al., 2019). Metabolomics has a wide range of applications, including human health and diseases (Johnson et al., 2016), animals (Kirwan 2013), plants (Pontes et al., 2017), microorganisms (Ramirez-Gaona et al., 2017), and other areas (Kim H.-Y. et al., 2016; Munger et al., 2017). More and more researchers were using metabolomics technology to study tumor metabolism (Armitage and Ciborowski, 2017; Kumar and Misra, 2019). In the BC field, metabolomics has been fully applied (McCartney et al., 2018). Four metabolites of glutamine, isoleucine, threonine, and linolenic acid could be used as potential markers for predicting response to NAC for BC, by comprehensive use of nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) techniques (Wei et al., 2013). However, studies on the prognosis of TNBC surgery have not been performed.

On the contrary, metabolomics was also applied to individualize treatment (Jacob et al., 2019). Our early study used NMR-based metabolomics for finding the new biomarkers of colorectal cancer (Gu et al., 2019a). Mohammad *et al.* reviewed the application of metabolomics in the prognosis of acute coronary syndrome (Pouralijan Amiri et al., 2019). Similarly, metabolomics was also applied to bariatric surgery (Samczuk et al., 2018). In this study, we used metabolomics for the evaluation of NAC for TNBC. Our work looks forward to discovering metabolic phenotypes and differential metabolic pathways between the patients with pCR, pPR, and pSD.

MATERIALS AND METHODS

Chemical Reagents

Deuterated reagents of D_2O and sodium 3-(trimethylsilyl) propionate-2,2,3,3-d4 (DSS) were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA, United States). Chromatographic grade methanol was bought from Sigma-Aldrich (St. Louis, MO, United States). Other analytical grade reagents ($NaH_2PO_4{\cdot}2H_2O$ and $K_2HPO_4{\cdot}3H_2O$) were purchased from J&K Scientific Ltd. (Beijing, China). All ultra-pure water used in this study was produced by a Milli-Q IQ 7000 system.

Selection of TNBC Patients and Collection of Serum Samples

TNBC patients were recruited and treated at the Department of Breast Surgery, Zhejiang Cancer Hospital (tumor hospital affiliated to the University of Chinese Academy of Sciences). These female patients were enrolled in the study between 2019 and 2020. This study was performed in accordance with protocols approved by the Zhejiang Cancer Hospital Ethics Committee. The clinicopathological characteristics of participating subjects are summarized in **Supplementary Table S1**. There were 52 patients in our study, of which 8 had pCR, 16 had pSD, and 28 had pPR. There was no difference in age and BMI index of these patients. Based on T category, the classification of patients was mainly concentrated in III and IV stages. The criteria for patient selection included 1) pathologically confirmed primary TNBC; 2) being in line with NAC indications; 3) age of 20–65 years; and 4) performance status (PS) score 0–1. The criteria for patient exclusion included 1) non-primary TNBC; 2) combination with other malignant tumors; 3) not meeting NAC indications; 4) combination with blood system diseases and kidney diseases, including hemophilia, aplastic anemia and myelodysplastic syndromes, immune thrombocytopenia, sickle cell disease, sickle cell trait, and other hemoglobinopathies, diabetes, and thalassemias; 5) patients with advanced BC; 6) age >65°years or <20°years; and 7) those who cannot tolerate chemotherapy and surgery, or those who have a PS score >1.

The effect of NAC in the treatment of TNBC was comprehensively obtained by magnetic resonance imaging (MRI) and two-dimensional or three-dimensional ultrasound and mammography with histopathology. According to these test results, patients were divided into three groups, including pCR, pPR, and pSD. Here, pCR indicates all tumor tissue is disappeared, pPR indicates tumor volume is reduced by more than 30%, while the tumor volume is reduced by less than 30% or increased by not more than 20% in pSD (Neubauer et al., 2008).

Each patient had a light diet for 48 h before blood collection. After blood collection (5 ml), it was coagulated and centrifuged (4°C, 4,000 rcf, 15 min) to obtain serum. All serum samples were frozen and stored in the −80 C refrigerator until the NMR experiment.

Pretreatment of Serum Samples and Acquisition of NMR Spectra

Before NMR data acquisition, the serum samples were thawed on ice. 300 μL serum was mixed with 600 μL methanol (Tiziani et al., 2008). Then, the mixed samples were stored in the −20 C refrigerator for 30 min. The macromolecules in sera underwent denaturation and precipitation and were removed by

centrifugation (12,000 g, 4°C, 30 min). Then, all these supernatant solvents were removed by the lyophilizer. The lyophilized metabolites were redissolved in 450 μL of ultrapure water, and then 50 μL of phosphate buffer (1.5 M K_2HPO_4/NaH_2PO_4, pH 7.4, D_2O) was added for stabilizing the pH of serum and deuterium lock-in NMR measurements. All samples were analyzed in the BRUKER AVANCE III HD 600 MHz spectrometer (BRUKER BioSpin, Germany). The one-dimensional 1H spectra were operated in the TXI probe at 300 K by using a pulse sequence with water suppression (NOESYPR1D [RD-90°-t_1-90°-τ_m-90°-ACQ]) with 3s relaxation delay and 120 m mixing time. The detailed acquisition parameters were described in the following kinds of literature (Gu et al., 2019b; Rohnisch et al., 2018; Shao et al., 2014). Then, the metabolites were identified from the NMR spectra according to the following reference (Rohnisch et al., 2018) and the HMDB (http://www.hmdb.ca/) (Wishart et al., 2018). Meanwhile, the two-dimensional (2D) NMR spectrum named "13C-^{1}H HSQC" (heteronuclear single-quantum coherence spectroscopy) was used for the identification of metabolites (Bingo et al., 2016).

Multivariate Statistics

Data preprocessing including data organization, removal of undesired areas, and binning was performed with MATLAB 2015b (MathWorks, Inc., United States). Minor adjustments in peak alignment between different samples were performed using the *icoshift* algorithm in MATLAB 2015b (Savorani et al., 2010). At the same time, visualization of the data was also carried out in MATLAB. According to the identified metabolites, we developed and utilized a metabolite database in this study for metabolite quantification. Using the same method of metabolite quantification from the literature of Cuperlovic-Culf et al. (2012), the relative concentrations of identified metabolites were calculated, which were based on multivariable linear regression of spectra with properly aligned metabolite data. On the contrary, the calculation of the relative concentration of the identified metabolites is also referred to as the AQuA (Rohnisch et al., 2018). Before multivariate statistical analysis, all data are normalized and par scaled. Then, principal component analysis (PCA) was performed to show clusters among all samples (Trygg et al., 2007). The partial least squares-discrimination analysis (PLS-DA) was applied for distinguishing the metabolic phenotypes among three groups (Trygg et al., 2007), and the corresponding response permutation test (RPT) was used for verifying the robustness of PLS-DA models (Lin et al., 2019). The orthogonal PLS-DA (OPLS-DA) was applied for differential metabolite analysis by using the variable importance in projection (VIP) (Cloarec et al., 2005) and the correlation coefficients (r) for the variables that are related to the first predictive component (tp1) (Cho et al., 2008). Besides, probability p values of the Kruskal–Wallis test and fold changes were also calculated between the pSD group, the pPR group, and the pCR group for assessing the statistical significance of differential metabolites. These four parameters (VIP value, correlation coefficients (r), p value, and fold change) were employed in the enhanced volcano plots for visualizing the differential metabolites (Hur et al., 2013; Lin et al., 2019).

Identifying the Disturbed Metabolic Pathways

Metabolic pathway analysis was performed to identify significantly disturbed pathways associated with the three groups of TNBC patients in the Pathway Analysis module of MetaboAnalyst 5.0 (www.metaboanalyst.ca/) according to the relative concentration of the metabolites. Two parameters, statistical p value and pathway impact value, were used to evaluate the importance of the metabolic pathway. By matching the different metabolites with the metabolites in each metabolic pathway, the p value was calculated by the hyper-geometric test (Goeman and Bühlmann, 2007). At the same time, the pathway impact value was calculated from the topological analysis using the out-degree centrality algorithm through matched differential metabolites in metabolic pathways (Chong et al., 2018). According to the approaches described in other previous works (Gu et al., 2016; Gu et al., 2020a), we identified significantly disturbed metabolic pathways associated with p less than 0.05 and pathway impact values greater than 0.3.

Analyzing the Discriminative Ability of Disturbed Metabolic Pathways

Metabolomic analysis could help develop potential biomarkers for early diagnosis in multiple medical fields (Ni et al., 2014; Nobakht 2018; Gu et al., 2019a). In this study, the multi-receiver–operating characteristic (multi-ROC) curve analysis was operated on assessing discriminant capabilities of the metabolites involved in the significantly disturbed metabolic pathways (Zweig and Campbell, 1993; Gu et al., 2020b). In multi-ROC curve analysis, the logistic regression arithmetic was used for the classification of these three groups of patients, and the area under the ROC curve (AUC) value was used for evaluating the prediction performance of the metabolites in the disturbed metabolic pathway as the AUC was greater than 0.70 (Mandrekar, 2010).

RESULTS

Characteristics of Enrolled TNBC Patients

In this prospective study, detailed clinical characteristics of the participants are summarized in **Supplementary Table S1**. In our study, we used the TNM system to stage cancer, which is determined after cancer is assigned a letter to describe it, including tumor (T), node (N), and metastasis (M). On the contrary, a number after T (such as T1, T2, T3, or T4) might describe the tumor size. **Supplementary Table S1** shows that no significant differences were observed in age, BMI, and clinical T stage, between these three groups ($p > 0.05$).

Metabolic Profiles of Serum Samples

In the present study, all serum samples were collected from TNBC patients before neoadjuvant chemotherapy. A total of 63 metabolites were identified and relatively quantified from

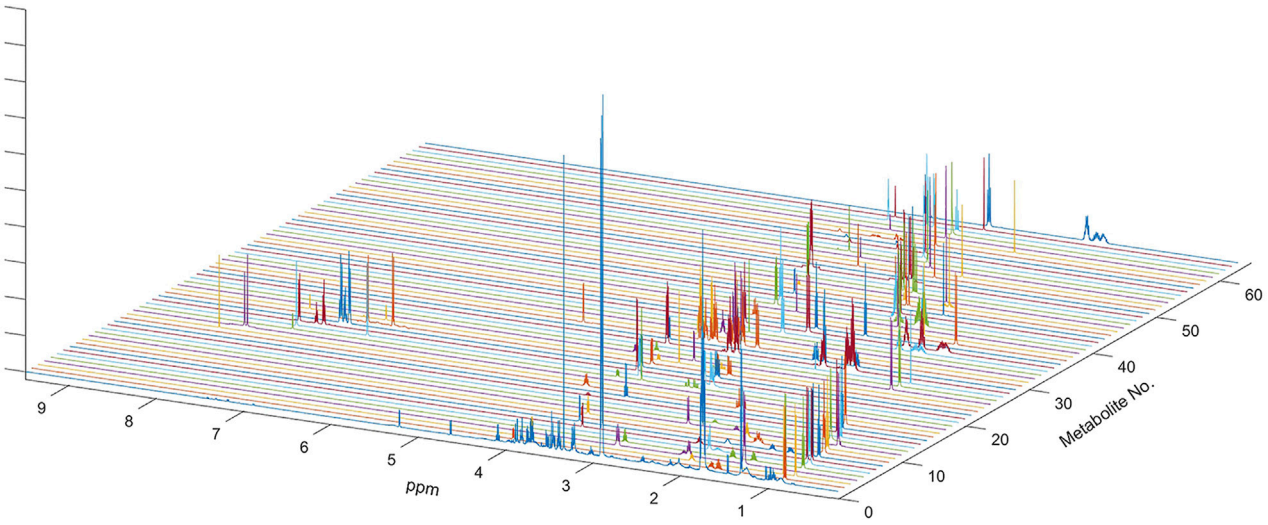

FIGURE 1 | NMR spectrum of metabolites used for multi-linear regression analysis of the global spectrum. Sixty-three metabolites used in the analysis included all metabolites previously determined in the HMDB and also confirmed on the Chenomx NMR Suite. One-dimensional spectra of all sixty-three metabolites are shown along with the outline of the average spectrum for the serum sample. Complete spectra of all metabolites were used in multivariate linear regression analysis.

the NMR spectra (**Supplementary Table S2**, **Figure 1**, and **Supplementary Figure S1**), which were calculated out by using the automated method based on multivariable linear regression (Cuperlovic-Culf et al., 2012) in MATLAB (version 2015b, MathWorks, Inc., United States). Then, the multivariate statistical analysis was utilized to analyze the quantitative data of metabolites. Using the first three components, the PCA score plots are shown in **Figure 2**. The metabolic phenotypes of the three groups could be roughly distinguished (**Figure 2A**). Overall, the pCR group was distinguished from the pSD group (**Figure 2B**), and the pPR group was roughly distinguished from the pSD group and pCR group (**Figures 2C,D**).

Furthermore, supervised multivariate statistical analysis was also applied to distinguish the metabolic profiles. These PLS-DA score plots and corresponding RPTs indicated the metabolic phenotypes of the three groups could be distinguishable (**Supplementary Figure S2**).

Comparison of the Relative Concentration of Different Metabolites in TNBC Patients

According to the relative quantitative value of 63 metabolites in serum samples, the mean and standard error of the mean (SEM) were calculated for each group (**Table 1**). Then, we performed the Kruskal–Wallis multiple-comparisons test to identify differential metabolites with $p < 0.05$ (**Table 1** and **Supplementary Figure S3**). By comparing the serum metabolites of the three groups of patients, it was found that a total of 26 metabolites changed in the three groups. In the pCR group, there were 10 metabolites with the highest relative concentration, including τ-methylhistidine, phenylalanine, π-methylhistidine, lactic acid, glucose, alanine, glutamic acid, citric acid, dimethylamine, and phosphocholine. In the pSD group, these were seven metabolites with the highest relative concentration, including valine, 2-aminobutanoic acid, propionic acid, ethanol, proline, asparagine, and N,N-dimethylglycine. In the pPR group, there were six metabolites with the highest relative concentration, including 2-hydroxyisovaleric acid, acetoacetate, trimethylamine, creatine, myo-inositol, and ornithine, and there were five metabolites with the lowest relative concentration, including isoleucine, phenylalanine, threonine, dimethylamine, and glycerophosphocholine.

Determination of Differential Metabolites Between Different TNBC Patients

For analyzing the differential metabolites, the four-dimensional enhanced volcano plots were used for data visualization (Lin et al., 2017). Based on OPLS-DA models, the VIP value and correlation coefficients (r) were calculated. The score plots of OPLS-DA models also showed that the metabolic profiles of different groups (pCR, pPR, and pSD) were differentiable (**Figures 3A–C**). And the corresponding RPTs demonstrated that the OPLS-DA models were not overfitting (**Figures 3D–F**). In the enhanced volcano plot (**Figure 4**), the differential metabolites were determined using the following four criteria: VIP value > 1, p value < 0.05, absolute $\log_2$ (fold change) > 0.2, and correlation coefficient (r) > corresponding threshold ($|r|>0.297$ in pPR vs. pSD; $|r|>0.329$ in pCR vs. pPR; $|r|>0.404$ in pCR vs. pSD). The differential metabolites are located at the upper-left and upper-right areas of the volcano plot with larger circular shapes and gradually warm colors. In **Figure 4A** (pPR vs. pSD), six metabolites were a significant difference. Compared with the pSD group, three of the metabolites (trimethylamine, glucose, and lactic acid) were increased and three metabolites (N,N-dimethylglycine, proline, and glycerophosphocholine) were decreased in the pPR group. The relevant statistical parameters in **Figure 4A** are shown in

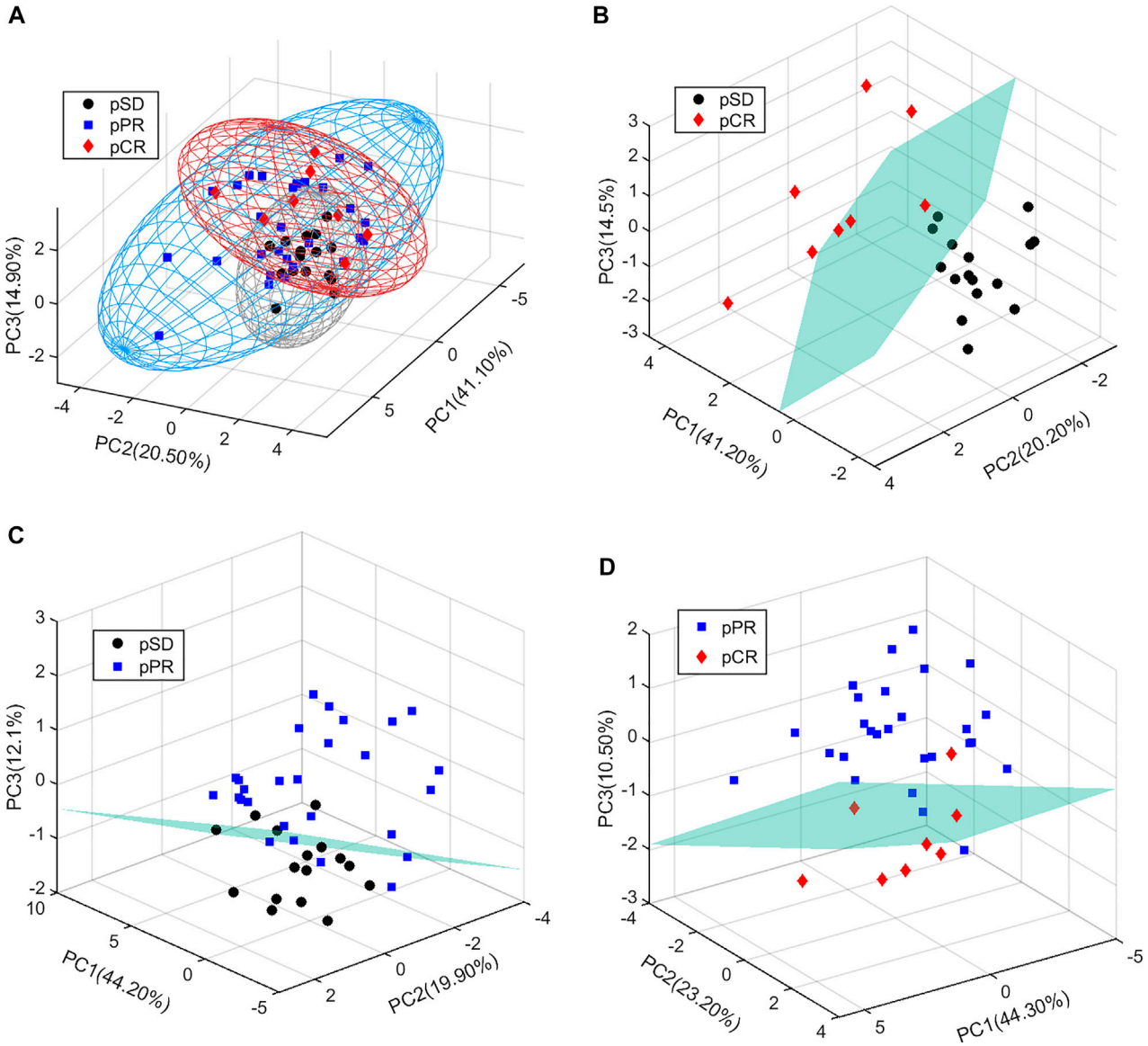

FIGURE 2 | PCA score plots of the relative concentration of metabolites' data from three groups of TNBC patients: **(A)** all patients; **(B)** pSD patients vs. pCR patients; **(C)** pSD patients vs. pPR patients; **(D)** pPR patients vs. pCR patients.

Supplementary Table S3. Compared with the pPR group, three of the metabolites (lactic acid, glutamic acid, and alanine) were increased and three metabolites (2-aminobutanoic acid, N,N-dimethylglycine, and ornithine) were decreased in the pCR group (**Figure 4B**). Similarly, the relevant statistical parameters in **Figure 4B** are shown in **Supplementary Table S4**. Compared with the pSD group, eight metabolites were a significant difference, including three of the metabolites (lactic acid, alanine, and glucose) which increased and five metabolites (2-aminobutanoic acid, N,N-dimethylglycine, asparagine, proline, and ornithine) which decreased in the pCR group (**Figure 4C**). The relevant statistical parameters in **Figure 4C** are shown in **Supplementary Table S5**.

Significantly Disturbed Metabolic Pathways in Different Groups

Based on the differential metabolites, we identified significantly disturbed metabolic pathways through pairwise comparison (**Figure 5**). On comparison between the pSD group and the pPR group, three metabolic pathways were changed, including glycine, serine, and threonine metabolism, valine, leucine and isoleucine biosynthesis, and alanine, aspartate, and glutamate metabolism (**Figure 5A**). Simultaneously, on comparison between the pCR group and the pSD group, more metabolic pathways were disturbed, including glycine, serine, and threonine metabolism, valine, leucine, and isoleucine biosynthesis, alanine, aspartate, and glutamate metabolism, glutamine and glutamate

TABLE 1 | Comparison of metabolite levels among the three groups on relative integrals calculated from 1D ^{1}H-NMR spectra of TNBC patients' serum samples.

	Mean ± SEM			Pairwise comparisons of Kruskal–Wallis test		
	pSD	pPR	pCR	pPR vs. pSD	pPR vs. pCR	pCR vs. pSD
2-Hydroxybutyric acid	1.859 ± 0.052	1.877 ± 0.059	1.827 ± 0.124	0.656	0.497	0.638
2-Hydroxyisovaleric acid	0.612 ± 0.028	0.558 ± 0.028	0.540 ± 0.045	0.012	0.573	0.019
Isocaproic acid	0.067 ± 0.014	0.063 ± 0.016	0.100 ± 0.045	0.703	0.139	0.215
3-Methyl-2-oxovaleric acid	0.690 ± 0.023	0.662 ± 0.016	0.636 ± 0.024	0.058	0.092	0.005
Isovaleric acid	0.454 ± 0.046	0.489 ± 0.042	0.385 ± 0.090	0.279	0.068	0.212
Valine	0.066 ± 0.016	0.106 ± 0.018	0.162 ± 0.038	0.002	0.024	0.001
Isoleucine	0.326 ± 0.023	0.280 ± 0.021	0.317 ± 0.055	0.006	0.247	0.794
Leucine	1.355 ± 0.032	1.421 ± 0.057	1.544 ± 0.184	0.058	0.247	0.089
2-Aminobutanoic acid	0.450 ± 0.067	0.385 ± 0.073	0.188 ± 0.072	0.204	0.001	0.000
2-Oxoisocaproate	0.264 ± 0.016	0.257 ± 0.014	0.284 ± 0.049	0.507	0.327	0.465
Isobutyric acid	0.034 ± 0.007	0.033 ± 0.006	0.025 ± 0.011	0.812	0.194	0.157
Propionic acid	0.107 ± 0.004	0.098 ± 0.004	0.102 ± 0.007	0.006	0.342	0.348
Isopropanol	0.079 ± 0.007	0.075 ± 0.008	0.081 ± 0.005	0.427	0.264	0.683
Ethanol	0.151 ± 0.008	0.140 ± 0.006	0.133 ± 0.013	0.045	0.354	0.027
3-Hydroxybutyric acid	0.798 ± 0.118	0.753 ± 0.082	0.692 ± 0.142	0.543	0.488	0.278
Formic acid	0.005 ± 0.001	0.005 ± 0.001	0.006 ± 0.001	0.285	0.468	0.198
Hypoxanthine	0.006 ± 0.002	0.005 ± 0.002	0.004 ± 0.003	0.526	0.598	0.347
τ-Methylhistidine	0.075 ± 0.035	0.123 ± 0.030	0.169 ± 0.070	0.045	0.248	0.036
Histidine	0.584 ± 0.032	0.592 ± 0.037	0.530 ± 0.076	0.721	0.187	0.224
Hippuric acid	0.063 ± 0.013	0.080 ± 0.016	0.075 ± 0.030	0.123	0.765	0.488
Phenylalanine	0.174 ± 0.013	0.173 ± 0.025	0.218 ± 0.035	0.931	0.058	0.047
Tyrosine	0.214 ± 0.031	0.215 ± 0.024	0.216 ± 0.044	0.955	0.964	0.933
π-Methylhistidine	0.532 ± 0.018	0.578 ± 0.025	0.646 ± 0.074	0.004	0.132	0.019
Threonine	0.298 ± 0.023	0.261 ± 0.021	0.296 ± 0.055	0.023	0.277	0.965
Lactic acid	2.844 ± 0.300	3.363 ± 0.420	4.821 ± 0.778	0.044	0.007	0.001
3-Hydroxyisovaleric acid	0.229 ± 0.018	0.205 ± 0.017	0.250 ± 0.043	0.065	0.086	0.412
Proline	3.285 ± 0.140	2.978 ± 0.130	2.995 ± 0.161	0.002	0.857	0.013
Pyroglutamic acid	0.362 ± 0.042	0.349 ± 0.028	0.375 ± 0.041	0.601	0.335	0.689
Glucose	6.449 ± 0.591	7.571 ± 1.089	7.970 ± 1.421	0.042	0.678	0.043
Serine	0.997 ± 0.044	1.070 ± 0.081	1.055 ± 0.087	0.125	0.812	0.267
Glycerol	0.300 ± 0.047	0.262 ± 0.039	0.273 ± 0.086	0.244	0.827	0.627
Glycine	0.198 ± 0.025	0.225 ± 0.030	0.225 ± 0.050	0.168	0.692	0.364
Arginine	3.505 ± 0.094	3.582 ± 0.110	3.564 ± 0.234	0.303	0.848	0.666
Lysine	3.229 ± 0.112	3.341 ± 0.136	3.354 ± 0.243	0.221	0.937	0.381
2-Oxoglutaric acid	0.010 ± 0.004	0.009 ± 0.003	0.013 ± 0.016	0.624	0.627	0.726
Alanine	0.242 ± 0.048	0.300 ± 0.050	0.513 ± 0.085	0.111	0.001	0.000
Acetic acid	0.088 ± 0.006	0.098 ± 0.007	0.094 ± 0.011	0.062	0.635	0.346
Acetoacetate	0.012 ± 0.006	0.032 ± 0.012	0.023 ± 0.011	0.005	0.327	0.134
Glutamic acid	0.384 ± 0.068	0.402 ± 0.064	0.512 ± 0.072	0.678	0.037	0.017
Glutamine	0.863 ± 0.080	0.828 ± 0.092	0.936 ± 0.158	0.603	0.287	0.448
Pyruvate	0.053 ± 0.003	0.053 ± 0.004	0.062 ± 0.008	0.855	0.081	0.065
N-Acetylglycine	0.132 ± 0.009	0.119 ± 0.014	0.108 ± 0.021	0.131	0.414	0.075
Citric acid	0.006 ± 0.002	0.008 ± 0.002	0.016 ± 0.008	0.196	0.128	0.043
Methionine	0.199 ± 0.008	0.203 ± 0.010	0.211 ± 0.017	0.531	0.415	0.232
Acetone	0.018 ± 0.001	0.018 ± 0.002	0.020 ± 0.002	0.355	0.063	0.167
Aspartic acid	0.136 ± 0.018	0.139 ± 0.014	0.155 ± 0.013	0.823	0.135	0.121
Methylguanidine	0.024 ± 0.002	0.022 ± 0.003	0.026 ± 0.006	0.302	0.223	0.441
Asparagine	1.565 ± 0.112	1.452 ± 0.078	1.220 ± 0.211	0.121	0.084	0.013
Trimethylamine	4.513 ± 0.212	5.007 ± 0.298	4.912 ± 0.576	0.012	0.784	0.265
Sarcosine	0.016 ± 0.004	0.015 ± 0.003	0.021 ± 0.009	0.788	0.276	0.375
Dimethylamine	0.010 ± 0.001	0.009 ± 0.001	0.013 ± 0.002	0.674	0.021	0.027
N,N-Dimethylglycine	6.994 ± 0.217	6.424 ± 0.219	5.423 ± 0.443	0.002	0.001	0.000
Creatine	0.046 ± 0.014	0.069 ± 0.012	0.062 ± 0.022	0.022	0.702	0.224
Dimethyl sulfone	0.034 ± 0.002	0.036 ± 0.002	0.040 ± 0.005	0.401	0.186	0.088
Choline	0.217 ± 0.023	0.216 ± 0.014	0.196 ± 0.028	0.944	0.335	0.309
Phosphocholine	0.245 ± 0.035	0.279 ± 0.032	0.390 ± 0.054	0.189	0.003	0.002
Glycerophosphocholine	0.407 ± 0.036	0.325 ± 0.051	0.395 ± 0.127	0.020	0.502	0.799
Succinic acid	0.062 ± 0.004	0.060 ± 0.004	0.060 ± 0.007	0.403	0.952	0.608
Betaine	0.116 ± 0.033	0.098 ± 0.031	0.076 ± 0.034	0.448	0.421	0.215
Trimethylamine N-oxide	0.236 ± 0.043	0.209 ± 0.030	0.216 ± 0.083	0.225	0.884	0.567
myo-Inositol	1.022 ± 0.029	1.077 ± 0.042	1.057 ± 0.065	0.027	0.587	0.343
Creatinine	0.484 ± 0.018	0.502 ± 0.022	0.472 ± 0.043	0.277	0.208	0.658
Ornithine	8.057 ± 0.245	8.080 ± 0.264	7.454 ± 0.476	0.912	0.037	0.046

SEM means the standard error of the mean, confidence interval.

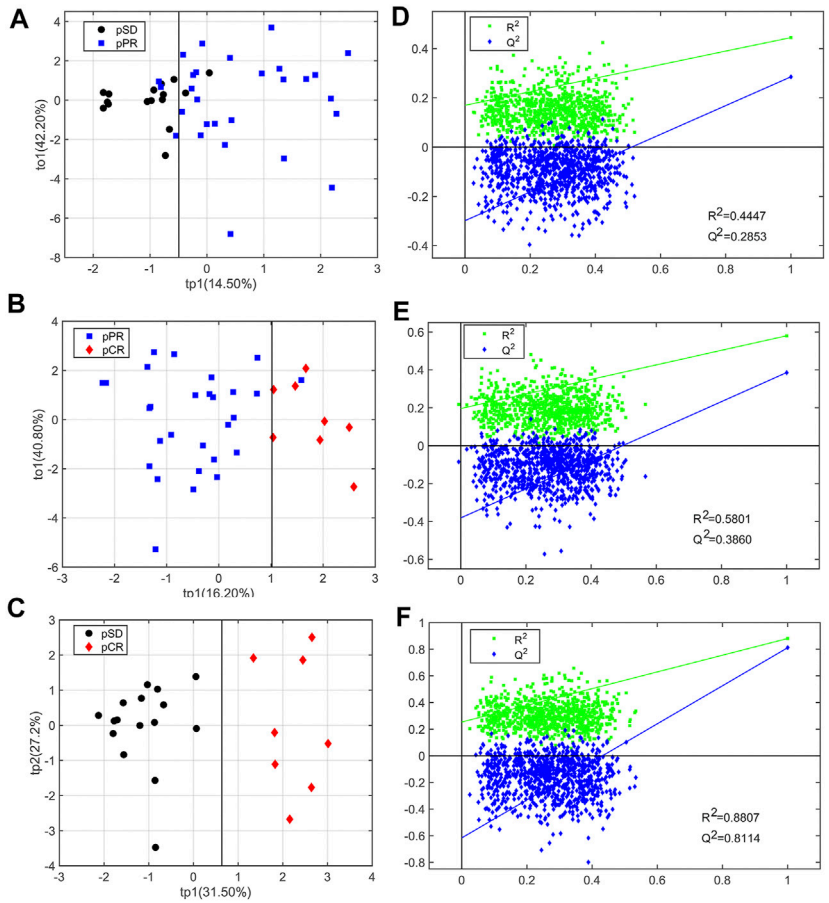

FIGURE 3 | OPLS-DA score plots and corresponding permutation tests of the relative concentration of metabolites' data from three groups of TNBC patients: **(A, D)** pSD patients vs. pPR patients; **(B, E)** pSD patients vs. pCR patients; **(C, F)** pPR patients vs. pCR patients. In the RPT plots, the green square is R^2 (cum), denoting the explained variance of the model. The blue diamond is Q^2 (cum), standing for the predictive ability of the model.

metabolism, histidine metabolism, and arginine biosynthesis (**Figure 5B**). On comparison between pCR and pPR groups, we found the disturbed metabolic pathways were the same as the metabolic pathways which were changed in comparison between the pSD and pPR groups (**Figure 5C**).

Potential Discriminant Analysis of Disturbed Metabolic Pathways in Different TNBC Patients

Based on the discriminant capabilities of the significant metabolites predicted from the multivariate ROC curve analysis, we analyzed the potential discriminative ability of disturbed metabolic pathways which could metabolically discriminate the different TNBC groups (**Figures 6–8**). Compared to the pSD group, three significant pathways displayed good discriminant capabilities in the pPR group with larger AUC values of 0.9129 for glycine, serine, and threonine metabolism, 0.8638 for valine, leucine, and isoleucine biosynthesis, and 0.8460 for alanine, aspartate, and glutamate metabolism (**Figure 6**). More significantly, the AUC values of N,N-dimethylglycine, valine, isoleucine, and creatine were higher than the threshold (0.7813, 0.7366, 0.7254, and 0.7009) in these pathways. These results showed that N,N-dimethylglycine, valine, isoleucine, and creatine could be used as potential biomarkers to distinguish between the pPR group and the pSD group.

Meanwhile, six significant pathways also showed good discriminant capabilities when the pCR group was compared with the pSD group (**Figure 7**), of which three of the same significant metabolic pathways showed excellent distinguishing ability (AUC values = 1) and the other three significant metabolic pathways also had a good distinguishing ability with a larger AUC value of 0.8047 for glutamine and glutamate metabolism, 0.8906 for arginine biosynthesis, and 0.9297 for histidine metabolism. The metabolites involved in these significant metabolic pathways also had a good ability to distinguish between pCR and pSD groups. The AUC values of N,N-dimethylglycine and pyruvate were higher than the threshold (0.9687 and 0.7734) in the metabolic pathway of glycine, serine, and threonine metabolism (**Figure 7A**). Valine and 3-methyl-2-oxovaleric acid had an excellent distinguishing ability with larger AUC

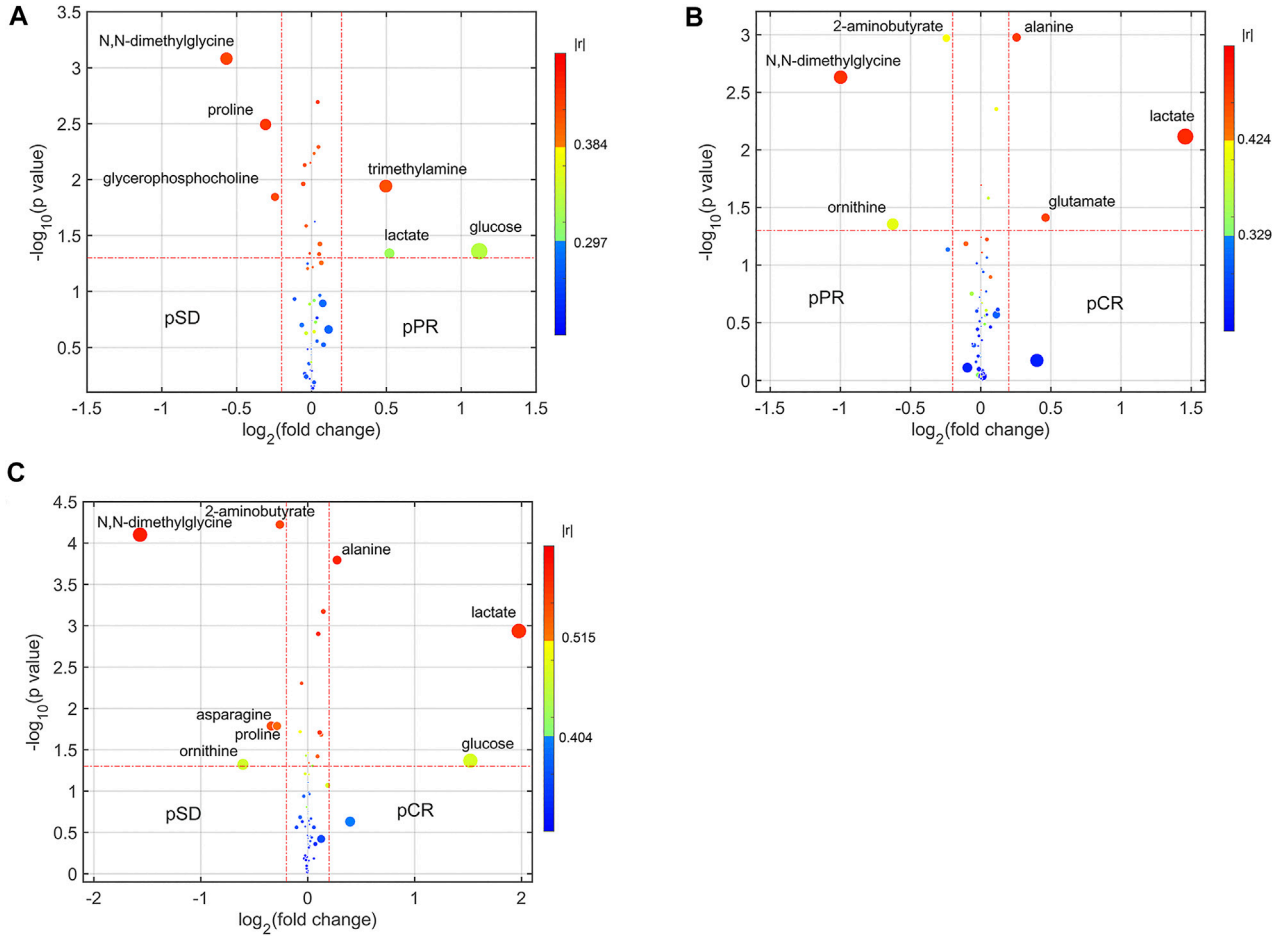

FIGURE 4 | Enhanced volcano plots showing significantly different metabolites: **(A)** pSD patients vs. pPR patients; **(B)** pPR patients vs. pCR patients; **(C)** pSD patients vs. pCR patients. The volcano plot shows $\log_2$ (fold change) on the *x*-axis and $-\log_{10}$ (*p* value) on the y-axis. Each point represents a metabolite. Circles' size and color are determined based on the variable importance in projection (VIP) and absolute correlation coefficient values (|r|), respectively. For each comparison, the larger the VIP value, the larger the size of the circle, and the warmer color corresponds to higher |r|; the gradient blue means |r| is less than 0.297; the gradual bright yellow means |r| is greater than 0.297 and is less than 0.384; the gradient red means |r| is greater than 0.384.

values (0.9844 and 0.8281, **Figure 7B**). In the metabolic pathway of alanine, aspartate, and glutamate metabolism, the metabolites of alanine, glutamic acid, pyruvate, citric acid, and asparagine had AUC values that exceed the threshold (0.9609, 0.7734, 0.7734, 0.7734, and 0.8438, **Figure 7C**). The same, the metabolite of glutamic acid showed the same discriminative ability in the metabolic pathway of glutamine and glutamate metabolism (**Figure 7D**). Similarly, ornithine had good discriminative ability in the metabolic pathway of arginine biosynthesis with 0.7856 AUC value (**Figure 7E**), and π-methylhistidine had good discriminative ability in the metabolic pathway of histidine metabolism with 0.8516 AUC value (**Figure 7F**). According to this multi-ROC curve analysis, the metabolites of N,N-dimethylglycine, pyruvate, valine, 3-methyl-2-oxovaleric acid, citric acid, asparagine, ornithine, and π-methylhistidine could be used as potential biomarkers to distinguish between the pCR group and the pSD group.

On comparison between the pCR group and pPR group patients, these three significant pathways also had a good discriminative ability with larger AUC values (**Figure 8**). In the metabolic pathway of glycine, serine, and threonine metabolism, the AUC values of N,N-dimethylglycine and pyruvate were higher than the threshold (0.8795 and 0.7366, **Figure 8A**). In the metabolic pathway of valine, leucine, and isoleucine biosynthesis, valine was the metabolite that mainly contributed to the discriminative ability with a larger AUC value (0.7902, **Figure 8B**). In the metabolic pathway of alanine, aspartate, and glutamate metabolism, only alanine and pyruvate contributed to the discriminative ability of this metabolic pathway (0.8928 and 0.7366, **Figure 8C**). According to this multi-ROC curve analysis, N,N-dimethylglycine, pyruvate, valine, and alanine could be used as potential biomarkers to distinguish between the pCR group and the pPR group.

According to the AUC values from the multi-ROC curves by differential metabolites, the AUC values of the pairwise comparison of N,N-dimethylglycine and valine were greater than 0.7. These two metabolites (N,N-dimethylglycine and valine) could be utilized as most potential biomarkers to distinguish among these three groups.

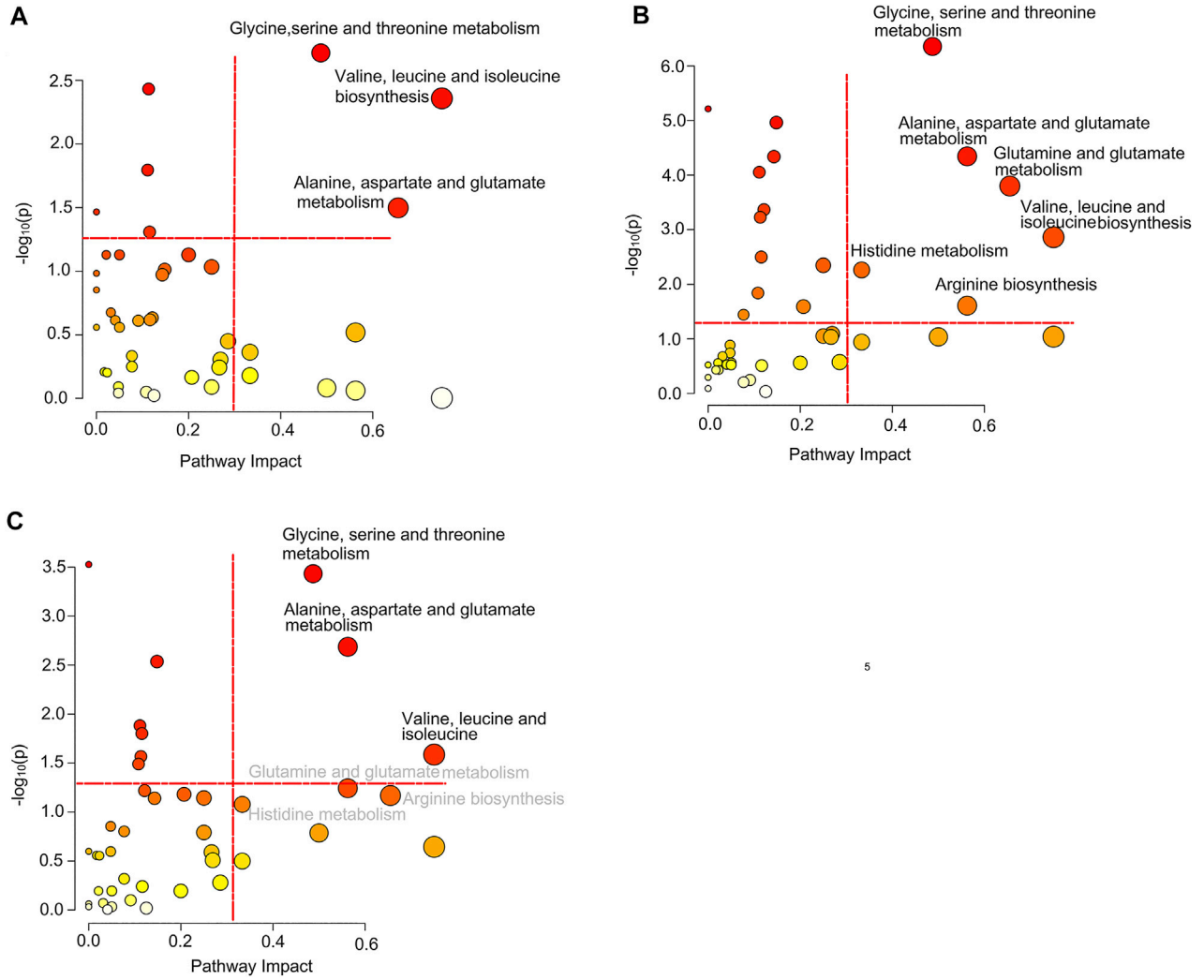

FIGURE 5 | Significantly disturbed metabolic pathways calculated in the comparison of three groups of TNBC patients: **(A)** pSD patients vs. pPR patients; **(B)** pSD patients vs. pCR patients; **(C)** pPR patients vs. pCR patients.

DISCUSSION

In this experimental study, we used the NMR-based metabolomics technology to predict the different sensitivities of NAC for TNBC patients. Our study found that the metabolic phenotype could classify the sensitivity of NAC, although each group of TNBC patients has different clinical and histopathological parameters. Some studies used transcriptomics to study the chemotherapy sensitivity of TNBC. Ozge Saatci *et al.* found lysyl-oxidase (LOX) to be a key inducer of chemoresistance in TNBC by whole-transcriptome sequencing (RNA-seq) (Saatci et al., 2020). In the same way, some research studies used metabolomics technology to analyze the predictive aspects of NAC of cancer (Wei et al., 2013; Hou et al., 2014; Yang et al., 2018). Meanwhile, our research was more focused on the differences in the overall metabolic pathways to distinguish the three groups of TNBC patients.

In our study, the resulting prediction models (OPLS-DA models) have high sensitivity and specificity. Through pairwise comparisons (pSD vs. pCR; pSD vs. pPR; pPR vs. pSD) by using the AUC values in multi-ROC curve analysis, three significant metabolic pathways of glycine, serine, and threonine metabolism, valine, leucine, and isoleucine biosynthesis, and alanine, aspartate, and glutamate metabolism could distinguish three groups in pairs. The metabolic pathway of glycine, serine, and threonine metabolism is linked to human BC invasion by comparing metabolic profiling of BC cells with different metastatic potentials (Kim S. et al., 2016). N,N-Dimethylglycine is involved in glycine, serine, and threonine metabolism, and also the methylation product of glycine. Biochemical methylation reaction mediates the transfer of methyl groups and regulates life activities. Ming Zhang *et al.* found the seven differentially methylated sites (DMSs) that were highly

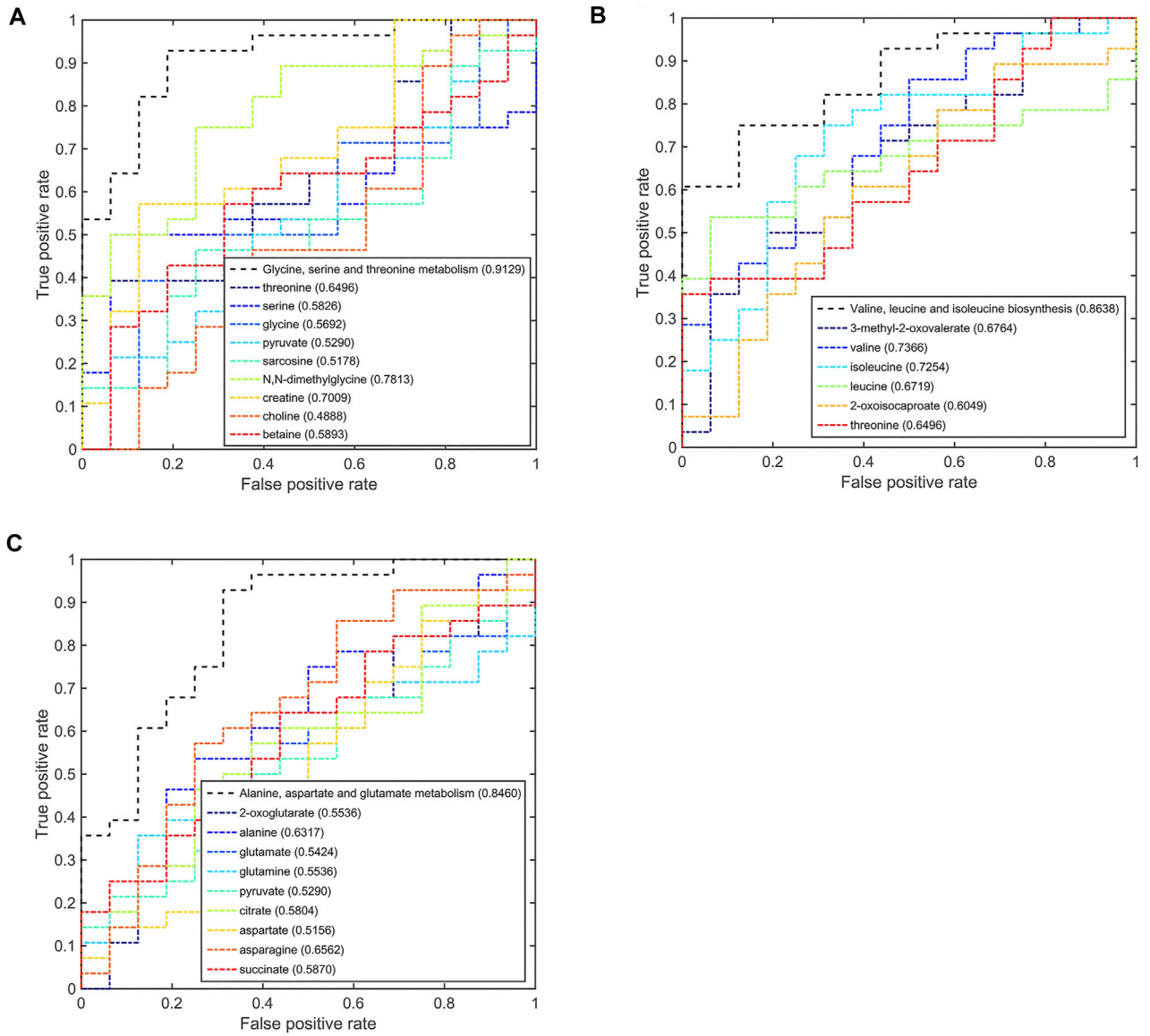

FIGURE 6 | Multi-ROC curves assessing discriminant capabilities of the significantly disturbed metabolic pathways in the pPR patients compared with the pSD patients. The AUC values shown in brackets are used to evaluate the performances of various biomarker models: **(A)** glycine, serine, and threonine metabolism; **(B)** valine, leucine, and isoleucine biosynthesis; **(C)** alanine, aspartate, and glutamate metabolism.

correlated with cell cycle as potential specific diagnostic biomarkers for BC patients (Zhang et al., 2020). Branched-chain amino acid metabolism was reprogrammed during tumorigenesis in many types of human cancers (Holecek 2018; Peng et al., 2020; Sivanand and Vander Heiden, 2020), including glioblastoma (Zhang et al., 2021), non-small-cell lung cancer (NSCLC) (Mayers et al., 2016), BC (Zhang, 2017), and ovarian cancer (Wang et al., 2015). In our work, we found the patients with different responses to NAC (pSD, pPR, and pCR) had different reprogramming metabolic pathways of valine, leucine, and isoleucine biosynthesis. Alanine, aspartate, and glutamate metabolism was reported to function as an alternative carbon source that fuels tumor metabolism (Sousa et al., 2016).

On the contrary, the other three metabolic pathways only could be used to distinguish between the pCR group and the pSD group. Glutamine and glutamate metabolism was perturbed in many types of cancers (Mates et al., 2019). Glutamine and glutamate metabolism has indispensable functions to provide amino acids, lipids, nucleotides, hexosamines, and polyamines, but also to render metabolic energy (ATP) (Mates et al., 2020). Meanwhile, glutamine and glutamate metabolism could regulate glutathione (GSH), the most important intracellular antioxidant molecule (Mates

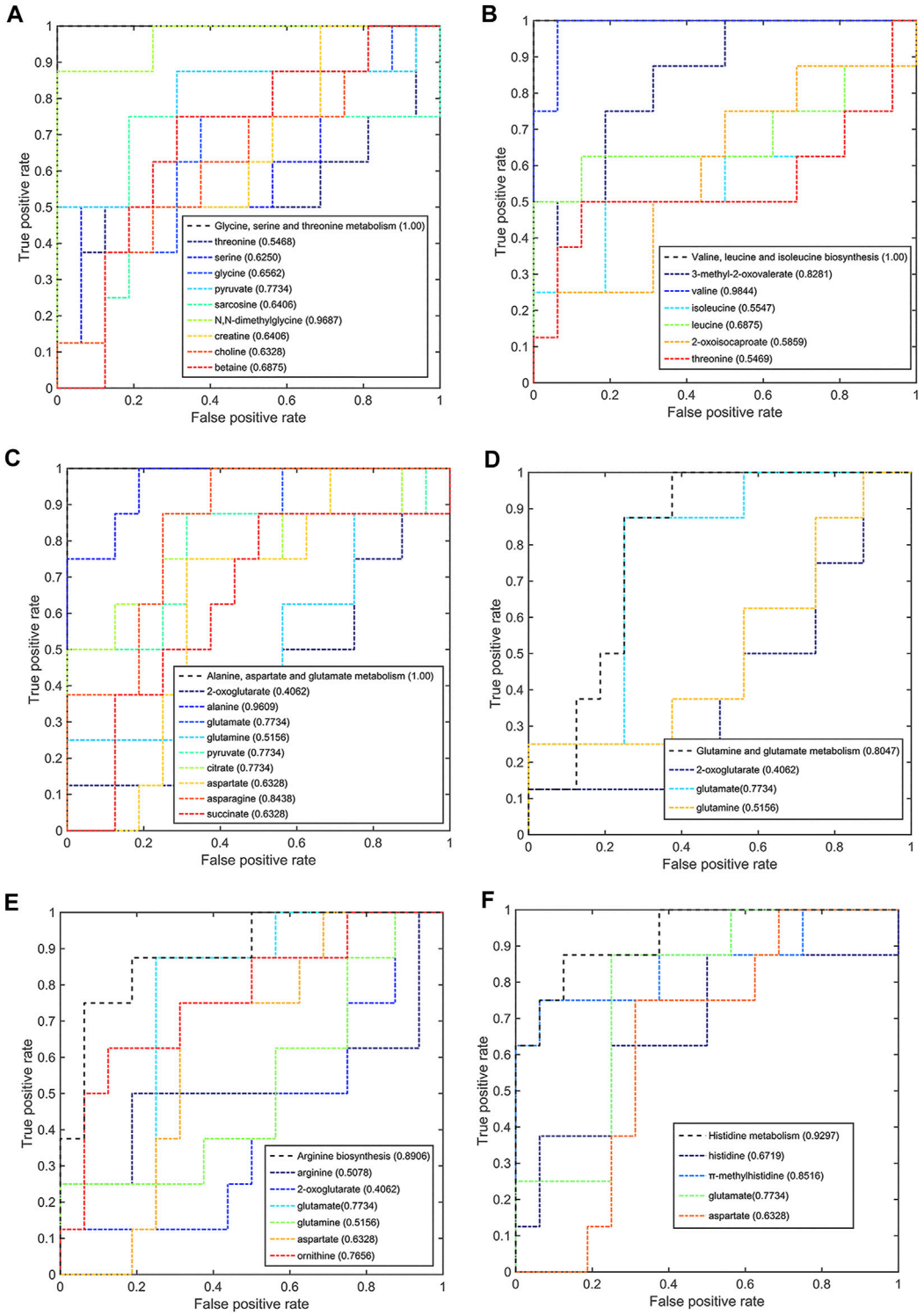

FIGURE 7 | Multi-ROC curves assessing discriminant capabilities of the significantly disturbed metabolic pathways in the pCR patients compared with the pSD patients. The AUC values shown in brackets are used to evaluate the performances of various biomarker models: **(A)** glycine, serine, and threonine metabolism; **(B)** valine, leucine, and isoleucine biosynthesis; **(C)** alanine, aspartate, and glutamate metabolism; **(D)** glutamine and glutamate metabolism; **(E)** arginine biosynthesis; **(F)** histidine metabolism.

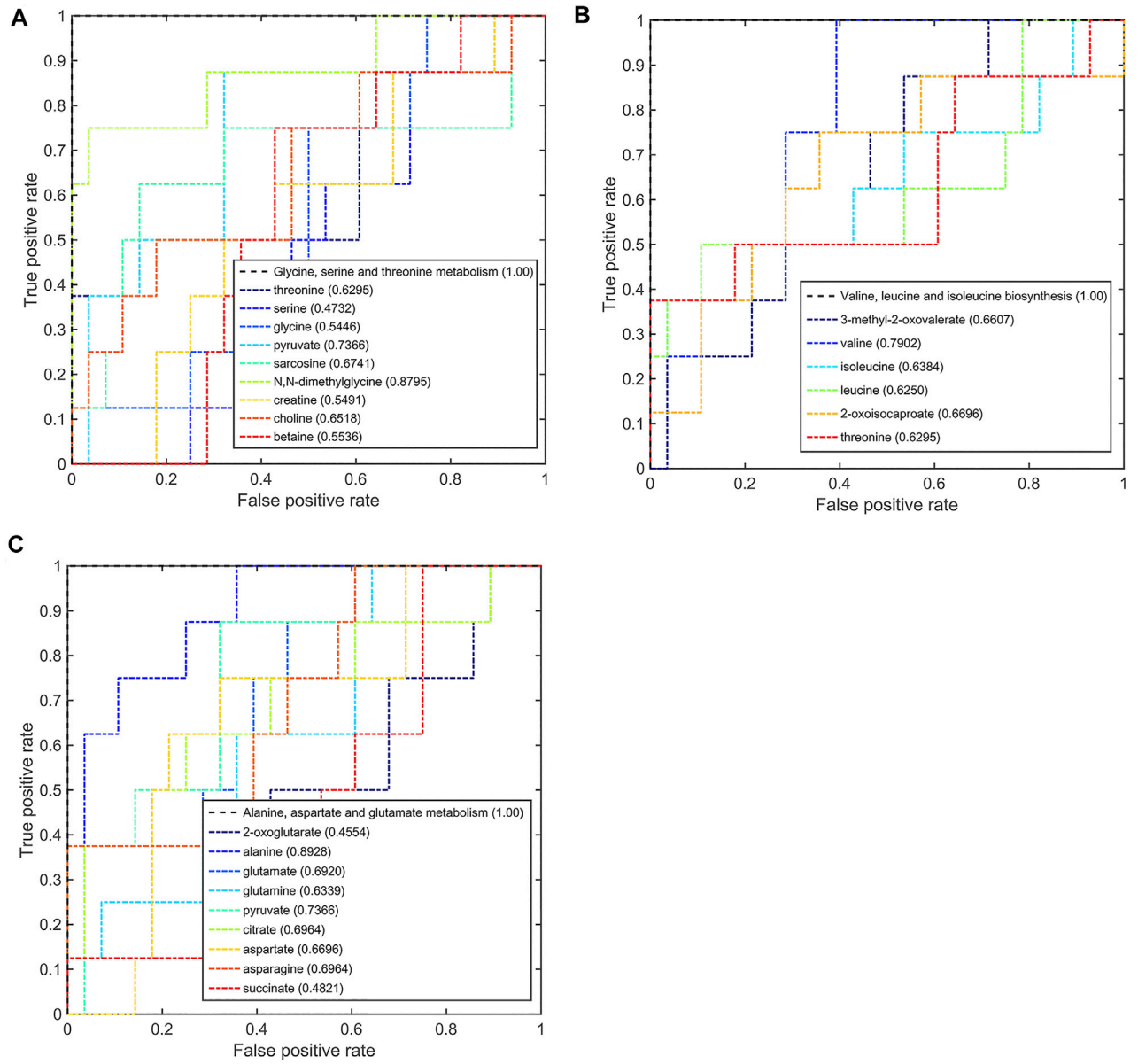

FIGURE 8 | Multi-ROC curves assessing discriminant capabilities of the significantly disturbed metabolic pathways in the pCR patients compared with the pPR patients. The AUC values shown in brackets are used to evaluate the performances of various biomarker models: **(A)** glycine, serine, and threonine metabolism; **(B)** valine, leucine, and isoleucine biosynthesis; **(C)** alanine, aspartate, and glutamate metabolism.

et al., 2018). Cancer cells frequently increase oxidative damage in response to changes in glutamine and glutamate metabolism (Mates et al., 2020). In our study, the significant metabolic pathway of glutamine and glutamate metabolism could distinguish different metabolic phenotypes between the pCR group and the pPR group. This result indicates that the oxidative stress state of the BC patients' microenvironment is different. Arginine biosynthesis was linked to the metabolic regulation of nitric oxide synthesis in cancer (Keshet and Erez, 2018). Paniz Jasbi *et al.* also found the arginine/proline metabolism was disturbed in the BC patients by using the targeted plasma metabolomics (Jasbi et al., 2019). Von Mach-Szczypiński *et al.* found histidine metabolism was abnormal in tissues of primary ductal BC (von Mach-Szczypinski et al., 2009a; von Mach-Szczypinski et al., 2009b). Similarly, histidine metabolism was abnormal in the serum of primary ductal BC (Sieja et al., 2005). Our research results also verify this result.

Since the chemotherapy response prediction for cancer remains challenging around the world, this promising metabolomics approach might open a new view for patients to select the promising treatment or even a truly "personalized

treatment" in clinical practice. Compared with other studies on single or multiple molecules as potential biomarkers, our research was more focused on the overall differences in metabolic pathways or metabolic phenotypes as potential biomarkers.

Our study analyzed differences in the metabolic phenotypes of TNBC patients with different sensitivity to neoadjuvant chemotherapy by using NMR-based metabolomics and then constructed a prediction model based on the metabolic phenotype. Three significant metabolic pathways of glycine, serine, and threonine metabolism, valine, leucine, and isoleucine biosynthesis, and alanine, aspartate, and glutamate metabolism could distinguish groups of patients with no, partial, or complete response. Additional three significant metabolic pathways of glutamine and glutamate metabolism, arginine biosynthesis, and histidine metabolism could distinguish groups of patients with no or complete response. Although this study only involved a small number of patient cohorts, the results have shown that these several metabolic pathways have good distinguishing ability for different patients with no, partial, or complete response. Of course, we need more clinical cohort samples for verifying these results. This method could be used as a preoperative choice for efficacy evaluation for patients with BC neoadjuvant chemotherapy.

LIMITATION

This study has the limitation that we lack of verification of other data of transcriptomics, proteomics, etc. Furthermore, given our relatively small sample size, our observation still remains to be verified in a large cohort. Thus, follow-up studies involving long-term studies of a large cohort of TNBC patients receiving NAC are required.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the Mendeley Data repository, accession number DOI: 10.17632/bbz8353xwh.1

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of Zhejiang Cancer Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XH and LT designed the project. JG, DZ, HY, YZ and YD performed experiments. JG analyzed the data. XH and JG prepared the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China Youth Science Fund Project (No. 21906147), the Natural Science Foundation of Zhejiang Province, China (No. LQ18B050003), and the Zhejiang Provincial Medical and Health Science and Technology Program (2020KY489).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.708052/full#supplementary-material

REFERENCES

Armitage, E. G., and Ciborowski, M. (2017). Applications of Metabolomics in Cancer Studies. *Adv. Exp. Med. Biol.* 965, 209–234. doi:10.1007/978-3-319-47656-8_9

Bingol, K., Li, D.-W., Zhang, B., and Brüschweiler, R. (2016). Comprehensive Metabolite Identification Strategy Using Multiple Two-Dimensional NMR Spectra of a Complex Mixture Implemented in the COLMARm Web Server. *Anal. Chem.* 88 (24), 12411–12418. doi:10.1021/acs.analchem.6b03724

Cho, H. W., Kim, S. B., Jeong, M. K., Park, Y., Miller, N. G., Ziegler, T. R., et al. (2008). Discovery of Metabolite Features for the Modelling and Analysis of High-Resolution NMR Spectra. *Int. J. Data Min Bioinform* 2 (2), 176–192. doi:10.1504/ijdmb.2008.019097

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 (W1), W486–W494. doi:10.1093/nar/gky310

Cloarec, O., Dumas, M. E., Trygg, J., Craig, A., Barton, R. H., Lindon, J. C., et al. (2005). Evaluation of the Orthogonal Projection on Latent Structure Model Limitations Caused by Chemical Shift Variability and Improved Visualization of Biomarker Changes in 1H NMR Spectroscopic Metabonomic Studies. *Anal. Chem.* 77 (2), 517–526. doi:10.1021/ac048803i

Cuperlovic-Culf, M., Ferguson, D., Culf, A., Morin, P., Jr., and Touaibia, M. (2012). 1H NMR Metabolomics Analysis of Glioblastoma Subtypes. *J. Biol. Chem.* 287 (24), 20164–20175. doi:10.1074/jbc.M111.337196

Glück, S., Ross, J. S., Royce, M., McKenna, E. F., Jr., Perou, C. M., Avisar, E., et al. (2012). TP53 Genomics Predict Higher Clinical and Pathologic Tumor Response in Operable Early-Stage Breast Cancer Treated with Docetaxel-Capecitabine ± Trastuzumab. *Breast Cancer Res. Treat.* 132 (3), 781–791. doi:10.1007/s10549-011-1412-7

Goeman, J. J., and Bühlmann, P. (2007). Analyzing Gene Expression Data in Terms of Gene Sets: Methodological Issues. *Bioinformatics* 23 (8), 980–987. doi:10.1093/bioinformatics/btm051

Gu, J., Hu, X., Shao, W., Ji, T., Yang, W., Zhuo, H., et al. (2016). Metabolomic Analysis Reveals Altered Metabolic Pathways in a Rat Model of Gastric Carcinogenesis. *Oncotarget* 7 (37), 60053–60073. doi:10.18632/oncotarget.11049

Gu, J., Huang, C., Hu, X., Xia, J., Shao, W., and Lin, D. (2020a). Nuclear Magnetic Resonance-based Tissue Metabolomic Analysis Clarifies Molecular Mechanisms of Gastric Carcinogenesis. *Cancer Sci.* 111 (9), 3195–3209. doi:10.1111/cas.14443

Gu, J., Huang, C., Hu, X., Xia, J., Shao, W., and Lin, D. (2020b). Nuclear Magnetic Resonance-based Tissue Metabolomic Analysis Clarifies Molecular Mechanisms of Gastric Carcinogenesis. *Cancer Sci.* 111 (9), 3195–3209. doi:10.1111/cas.14443

Gu, J., Su, F., Hong, P., Zhang, Q., and Zhao, M. (2019a). 1H NMR-Based Metabolomic Analysis of Nine Organophosphate Flame Retardants Metabolic Disturbance in Hep G2 Cell Line. *Sci. Total Environ.* 665, 162–170. doi:10.1016/j.scitotenv.2019.02.055

Gu, J., Xiao, Y., Shu, D., Liang, X., Hu, X., Xie, Y., et al. (2019b). Metabolomics Analysis in Serum from Patients with Colorectal Polyp and Colorectal Cancer by 1H-NMR Spectrometry. *Dis. Markers* 2019, 1–14. doi:10.1155/2019/3491852

Holeček, M. (2018). Branched-chain Amino Acids in Health and Disease: Metabolism, Alterations in Blood Plasma, and as Supplements. *Nutr. Metab. (Lond)* 15, 33. doi:10.1186/s12986-018-0271-1

Hou, Y., Yin, M., Sun, F., Zhang, T., Zhou, X., Li, H., et al. (2014). A Metabolomics Approach for Predicting the Response to Neoadjuvant Chemotherapy in Cervical Cancer Patients. *Mol. Biosyst.* 10 (8), 2126–2133. doi:10.1039/c4mb00054d

Hur, M., Campbell, A. A., Almeida-de-Macedo, M., Li, L., Ransom, N., Jose, A., et al. (2013). A Global Approach to Analysis and Interpretation of Metabolic Data for Plant Natural Product Discovery. *Nat. Prod. Rep.* 30 (4), 565–583. doi:10.1039/C3NP20111B

Jacob, M., Lopata, A. L., Dasouki, M., and Abdel Rahman, A. M. (2019). Metabolomics toward Personalized Medicine. *Mass. Spec. Rev.* 38 (3), 221–238. doi:10.1002/mas.21548

Jasbi, P., Wang, D., Cheng, S. L., Fei, Q., Cui, J. Y., Liu, L., et al. (2019). Breast Cancer Detection Using Targeted Plasma Metabolomics. *J. Chromatogr. B* 1105, 26–37. doi:10.1016/j.jchromb.2018.11.029

Johnson, C. H., Ivanisevic, J., and Siuzdak, G. (2016). Metabolomics: beyond Biomarkers and towards Mechanisms. *Nat. Rev. Mol. Cel Biol* 17 (7), 451–459. doi:10.1038/nrm.2016.25

Keshet, R., and Erez, A. (2018). Arginine and the Metabolic Regulation of Nitric Oxide Synthesis in Cancer. *Dis. Model. Mech.* 11 (8). doi:10.1242/dmm.033332

Kim, H.-Y., Lee, K.-M., Kim, S.-H., Kwon, Y.-J., Chun, Y.-J., and Choi, H.-K. (2016). Comparative Metabolic and Lipidomic Profiling of Human Breast Cancer Cells with Different Metastatic Potentials. *Oncotarget* 7 (41), 67111–67128. doi:10.18632/oncotarget.11560

Kim, S., Kim, J., Yun, E. J., and Kim, K. H. (2016). Food Metabolomics: from Farm to Human. *Curr. Opin. Biotechnol.* 37, 16–23. doi:10.1016/j.copbio.2015.09.004

Kirwan, J. (2013). Metabolomics for the Practising Vet. *Pract.* 35 (8), 438–445. doi:10.1136/inp.f5259

Kumar, A., and Misra, B. B. (2019). Challenges and Opportunities in Cancer Metabolomics. *Proteomics* 19 (21-22)–1900042. doi:10.1002/pmic.201900042

Kumar, P., and Aggarwal, R. (2016). An Overview of Triple-Negative Breast Cancer. *Arch. Gynecol. Obstet.* 293 (2), 247–269. doi:10.1007/s00404-015-3859-y

Liedtke, C., Mazouni, C., Hess, K. R., André, F., Tordai, A., Mejia, J. A., et al. (2008). Response to Neoadjuvant Therapy and Long-Term Survival in Patients with Triple-Negative Breast Cancer. *J. Clin. Oncol.* 26 (8), 1275–1281. doi:10.1200/jco.2007.14.4147

Lin, C., Chen, Z., Zhang, L., Wei, Z., Cheng, K.-K., Liu, Y., et al. (2019). Deciphering the Metabolic Perturbation in Hepatic Alveolar Echinococcosis: a 1H NMR-Based Metabolomics Study. *Parasites Vectors* 12 (1), 300. doi:10.1186/s13071-019-3554-0

Lin, C., Wei, Z., Cheng, K.-K., Xu, J., Shen, G., She, C., et al. (2017). 1H NMR-Based Investigation of Metabolic Response to Electro-Acupuncture Stimulation. *Sci. Rep.* 7 (1), 6820. doi:10.1038/s41598-017-07306-5

Mandrekar, J. N. (2010). Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *J. Thorac. Oncol.* 5 (9), 1315–1316. doi:10.1097/JTO.0b013e3181ec173d

Masuda, H., Baggerly, K. A., Wang, Y., Zhang, Y., Gonzalez-Angulo, A. M., Meric-Bernstam, F., et al. (2013). Differential Response to Neoadjuvant Chemotherapy Among 7 Triple-Negative Breast Cancer Molecular Subtypes. *Clin. Cancer Res.* 19 (19), 5533–5540. doi:10.1158/1078-0432.Ccr-13-0799

Masuda, N., Lee, S.-J., Ohtani, S., Im, Y.-H., Lee, E.-S., Yokota, I., et al. (2017). Adjuvant Capecitabine for Breast Cancer after Preoperative Chemotherapy. *N. Engl. J. Med.* 376 (22), 2147–2159. doi:10.1056/NEJMoa1612645

Matés, J. M., Campos-Sandoval, J. A., and Márquez, J. (2018). Glutaminase Isoenzymes in the Metabolic Therapy of Cancer. *Biochim. Biophys. Acta (Bba) - Rev. Cancer* 1870 (2), 158–164. doi:10.1016/j.bbcan.2018.07.007

Matés, J. M., Campos-Sandoval, J. A., Santos-Jiménez, J. d. l., and Márquez, J. (2019). Dysregulation of Glutaminase and Glutamine Synthetase in Cancer. *Cancer Lett.* 467, 29–39. doi:10.1016/j.canlet.2019.09.011

Matés, J. M., Di Paola, F. J., Campos-Sandoval, J. A., Mazurek, S., and Márquez, J. (2020). Therapeutic Targeting of Glutaminolysis as an Essential Strategy to Combat Cancer. *Semin. Cel Develop. Biol.* 98, 34–43. doi:10.1016/j.semcdb.2019.05.012

Mayers, J. R., Torrence, M. E., Danai, L. V., Papagiannakopoulos, T., Davidson, S. M., Bauer, M. R., et al. (2016). Tissue of Origin Dictates Branched-Chain Amino Acid Metabolism in Mutant Kras-Driven Cancers. *Science* 353 (6304), 1161–1165. doi:10.1126/science.aaf5171

McCartney, A., Vignoli, A., Biganzoli, L., Love, R., Tenori, L., Luchinat, C., et al. (2018). Metabolomics in Breast Cancer: A Decade in Review. *Cancer Treat. Rev.* 67, 88–96. doi:10.1016/j.ctrv.2018.04.012

Münger, L. H., Trimigno, A., Picone, G., Freiburghaus, C., Pimentel, G., Burton, K. J., et al. (2017). Identification of Urinary Food Intake Biomarkers for Milk, Cheese, and Soy-Based Drink by Untargeted GC-MS and NMR in Healthy Humans. *J. Proteome Res.* 16 (9), 3321–3335. doi:10.1021/acs.jproteome.7b00319

Neubauer, H., Gall, C., Vogel, U., Hornung, R., Wallwiener, D., Solomayer, E., et al. (2008). Changes in Tumour Biological Markers during Primary Systemic Chemotherapy (PST). *Anticancer Res.* 28 (3b), 1797–1804.

Ni, Y., Xie, G., and Jia, W. (2014). Metabonomics of Human Colorectal Cancer: New Approaches for Early Diagnosis and Biomarker Discovery. *J. Proteome Res.* 13 (9), 3857–3870. doi:10.1021/pr500443c

Nicholson, J. K., Lindon, J. C., Holmes, E., and 'Metabonomics' (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

Nobakht M. Gh, B. F. (2018). Application of Metabolomics to Preeclampsia Diagnosis. *Syst. Biol. Reprod. Med.* 64 (5), 324–339. doi:10.1080/19396368.2018.1482968

Peng, H., Wang, Y., and Luo, W. (2020). Multifaceted Role of Branched-Chain Amino Acid Metabolism in Cancer. *Oncogene* 39 (44), 6747–6756. doi:10.1038/s41388-020-01480-z

Pontes, J. G. M., Brasil, A. J. M., Cruz, G. C. F., de Souza, R. N., and Tasic, L. (2017). NMR-based Metabolomics Strategies: Plants, Animals and Humans. *Anal. Methods* 9 (7), 1078–1096. doi:10.1039/c6ay03102a

Pouralijan Amiri, M., Khoshkam, M., Salek, R. M., Madadi, R., Faghanzadeh Ganji, G., and Ramazani, A. (2019). Metabolomics in Early Detection and Prognosis of Acute Coronary Syndrome. *Clinica Chim. Acta* 495, 43–53. doi:10.1016/j.cca.2019.03.1632

Ramirez-Gaona, M., Marcu, A., Pon, A., Guo, A. C., Sajed, T., Wishart, N. A., et al. (2017). YMDB 2.0: a Significantly Expanded Version of the Yeast Metabolome Database. *Nucleic Acids Res.* 45 (D1), D440–d445. doi:10.1093/nar/gkw1058

Röhnisch, H. E., Eriksson, J., Müllner, E., Agback, P., Sandström, C., and Moazzami, A. A. (2018). AQuA: An Automated Quantification Algorithm for High-Throughput NMR-Based Metabolomics and its Application in Human Plasma. *Anal. Chem.* 90 (3), 2095–2102. doi:10.1021/acs.analchem.7b04324

Saatci, O., Kaymak, A., Raza, U., Ersan, P. G., Akbulut, O., Banister, C. E., et al. (2020). Targeting Lysyl Oxidase (LOX) Overcomes Chemotherapy Resistance in Triple Negative Breast Cancer. *Nat. Commun.* 11 (1), 2416. doi:10.1038/s41467-020-16199-4

Samczuk, P., Ciborowski, M., and Kretowski, A. (2018). Application of Metabolomics to Study Effects of Bariatric Surgery. *J. Diabetes Res.* 2018, 1–13. doi:10.1155/2018/6270875

Savorani, F., Tomasi, G., and Engelsen, S. B. (2010). Icoshift: A Versatile Tool for the Rapid Alignment of 1D NMR Spectra. *J. Magn. Reson.* 202 (2), 190–202. doi:10.1016/j.jmr.2009.11.012

Shao, W., Gu, J., Huang, C., Liu, D., Huang, H., Huang, Z., et al. (2014). Malignancy-associated Metabolic Profiling of Human Glioma Cell Lines Using ^{1}H NMR Spectroscopy. *Mol. Cancer* 13, 12. doi:10.1186/1476-4598-13-197

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer Statistics, 2020. *CA A. Cancer J. Clin.* 70 (1), 7–30. doi:10.3322/caac.21590

Sieja, K., Stanosz, S., von Mach-Szczypiński, J., Olewniczak, S., and Stanosz, M. (2005). Concentration of Histamine in Serum and Tissues of the Primary Ductal Breast Cancers in Women. *The Breast* 14 (3), 236–241. doi:10.1016/j.breast.2004.06.012

Sivanand, S., and Vander Heiden, M. G. (2020). Emerging Roles for Branched-Chain Amino Acid Metabolism in Cancer. *Cancer Cell* 37 (2), 147–156. doi:10.1016/j.ccell.2019.12.011

Sousa, C. M., Biancur, D. E., Wang, X., Halbrook, C. J., Sherman, M. H., Zhang, L., et al. (2016). Pancreatic Stellate Cells Support Tumour Metabolism through Autophagic Alanine Secretion. *Nature* 536 (7617), 479–483. doi:10.1038/nature19084

Tiziani, S., Emwas, A.-H., Lodi, A., Ludwig, C., Bunce, C. M., Viant, M. R., et al. (2008). Optimized Metabolite Extraction from Blood Serum for 1H Nuclear Magnetic Resonance Spectroscopy. *Anal. Biochem.* 377 (1), 16–23. doi:10.1016/j.ab.2008.01.037

Trygg, J., Holmes, E., and Lundstedt, T. (2007). Chemometrics in Metabonomics. *J. Proteome Res.* 6 (2), 469–479. doi:10.1021/pr060594q

Vignoli, A., Ghini, V., Meoni, G., Licari, C., Takis, P. G., Tenori, L., et al. (2019). High-Throughput Metabolomics by 1D NMR. *Angew. Chem. Int. Ed.* 58 (4), 968–994. doi:10.1002/anie.201804736

von Mach-Szczypiński, J., Stanosz, S., Sieja, K., and Stanosz, M. (2009a). Histamine and its Metabolizing Enzymes in Tissues of Primary Ductal Breast Cancer. *Eur. J. Gynaecol. Oncol.* 30 (5), 509–511.

von Mach-Szczypiński, J., Stanosz, S., Sieja, K., and Stanosz, M. (2009b). Metabolism of Histamine in Tissues of Primary Ductal Breast Cancer. *Metabolism* 58 (6), 867–870. doi:10.1016/j.metabol.2009.02.011

Wang, Z.-Q., Faddaoui, A., Bachvarova, M., Plante, M., Gregoire, J., Renaud, M.-C., et al. (2015). BCAT1 Expression Associates with Ovarian Cancer Progression: Possible Implications in Altered Disease Metabolism. *Oncotarget* 6 (31), 31522–31543. doi:10.18632/oncotarget.5159

Wei, S., Liu, L., Zhang, J., Bowers, J., Gowda, G. A. N., Seeger, H., et al. (2013). Metabolomics Approach for Predicting Response to Neoadjuvant Chemotherapy for Breast Cancer. *Mol. Oncol.* 7 (3), 297–307. doi:10.1016/j.molonc.2012.10.003

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 (D1), D608–d617. doi:10.1093/nar/gkx1089

Yang, K., Zhang, F., Han, P., Wang, Z.-z., Deng, K., Zhang, Y.-y., et al. (2018). Metabolomics Approach for Predicting Response to Neoadjuvant Chemotherapy for Colorectal Cancer. *Metabolomics* 14 (9), 110. doi:10.1007/s11306-018-1406-0

Yin, L., Duan, J.-J., Bian, X.-W., and Yu, S.-c. (2020). Triple-negative Breast Cancer Molecular Subtyping and Treatment Progress. *Breast Cancer Res.* 22 (1), 61. doi:10.1186/s13058-020-01296-5

Zhang, B., Chen, Y., Shi, X., Zhou, M., Bao, L., Hatanpaa, K. J., et al. (2021). Regulation of Branched-Chain Amino Acid Metabolism by Hypoxia-Inducible Factor in Glioblastoma. *Cell. Mol. Life Sci.* 78 (1), 195–206. doi:10.1007/s00018-020-03483-1

Zhang, L., and Han, J. (2017). Branched-chain Amino Acid Transaminase 1 (BCAT1) Promotes the Growth of Breast Cancer Cells through Improving mTOR-Mediated Mitochondrial Biogenesis and Function. *Biochem. Biophysical Res. Commun.* 486 (2), 224–231. doi:10.1016/j.bbrc.2017.02.101

Zhang, M., Wang, Y., Wang, Y., Jiang, L., Li, X., Gao, H., et al. (2020). Integrative Analysis of DNA Methylation and Gene Expression to Determine Specific Diagnostic Biomarkers and Prognostic Biomarkers of Breast Cancer. *Front. Cel Dev. Biol.* 8, 529386. doi:10.3389/fcell.2020.529386

Zweig, M. H., and Campbell, G. (1993). Receiver-operating Characteristic (ROC) Plots: a Fundamental Evaluation Tool in Clinical Medicine. *Clin. Chem.* 39 (4), 561–577. doi:10.1093/clinchem/39.4.561

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership