A pilot study optimizing metabolomic and lipidomic acquisition in serum for biomarker discovery in nonalcoholic fatty liver disease

Dandan He [a,1], Yang Su [b,1], Duanyue Meng [b], Xinmiao Wang [c], Jun Wang [d,*], Hui Ye [c,*]

[a] Experimental Center of Molecular and Cellular Biology, China Pharmaceutical University, Tongjiaxiang #24, Nanjing 210009, China
[b] School of Pharmacy, China Pharmaceutical University, Tongjiaxiang #24, Nanjing 210009, China
[c] Jiangsu Provincial Key Laboratory of Drug Metabolism and Pharmacokinetics, State Key Laboratory of Natural Medicines, China Pharmaceutical University, Tongjiaxiang #24, Nanjing 210009, China
[d] Nanjing First Hospital, Qinhuai District, Nanjing, Jiangsu Province 210001, China

ABSTRACT

Background: The worldwide prevalence of non-alcoholic fatty liver disease (NAFLD) has stimulated work to identify biomarkers and develop effective treatments. Metabolomics is an emerging tool that has been widely applied to discover biomarkers and simultaneously uncover pathological mechanisms. Here, we aim to optimize metabolomic acquisition with the goal of obtaining a systemic metabolic profile to unravel the potential link between dysregulated metabolism and NAFLD.

Methods: We analyzed serum samples collected from healthy subjects (n = 8) and NAFLD patients (n = 8) via an integrative analytical workflow using two orthogonal separation modes with T3 and amide columns and two ionization polarity modes on a UPLC-ESI-Q/TOF. Data dependent acquisition was employed for data acquisition. Differentially expressed metabolites and lipids were identified by comparing the collected metabolic and lipidomic profiles between the healthy subjects and NAFLD patients.

Results: The integrative LC-MS/MS analytical workflow employed here features an improved coverage of metabolites and lipids, which leads to the identification of 20 potential biomarkers of NAFLD, including lipids, acylcarnitines, and organic acids.

Conclusions: This pilot study has identified potential biomarkers for NAFLD and revealed corresponding dysregulated metabolic pathways related to NAFLD's occurrence and progression, establishing a molecular basis for NAFLD diagnosis and therapeutic intervention.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is a clinicopathologic syndrome. It is usually caused by excessive fat deposition in hepatocytes. NAFLD has been associated with diabetes, hypertension, dyslipidemia, and obesity [1]. It covers a broad spectrum of hepatopathy ranging from simple steatosis to non-alcoholic steatohepatitis (NASH) with varying degrees of fibrosis. It can develop into cirrhosis and, ultimately, hepatocellular carcinoma (HCC) [2]. Currently, NAFLD has become the most common chronic liver disease in developed countries and affects about a quarter of the world's population [3].

Hepatic ultrasonography and biopsy [4,5] are the conventional detection approaches for definitive diagnosis. However, liver biopsy is an invasive, painful, and costly procedure [6]. In contrast,

ultrasonography can obtain liver histomorphology images that indicate the severity and disease stages for NAFLD. Nevertheless, both measures cannot provide clues regarding the underlying mechanisms of NAFLD occurrence and progression, and, subsequently, suggest pharmacological interventions [7]. Therefore, a broad, non-invasive medical method that can provide a molecular basis for NAFLD diagnosis and therapeutic intervention is urgently needed.

Metabolomics is a relatively new field that complements genomics, transcriptomics, and proteomics. It collects comprehensive information on identities and dynamic abundance levels of metabolites and lipids in cells, tissues, organisms, and body fluids, allowing the obtainment of an overall metabolic profile from different subjects [8]. Thus, metabolomics studies that aim to identify potential therapeutic biomarkers for diseases are emerging [9,10]. Specifically, metabolomics has been applied to characterize various biomarkers for NAFLD. For instance, in a high-fat diet (HFD)-induced NAFLD mouse model, 30 potential biomarkers including serum glucose, total cholesterol, and hepatic triglyceride were identified, and their abundance was found to be intimately associated with the development of NAFLD [11]. Besides animal models, mass spectrometry (MS)-empowered metabolomics has been clinically applied to identify various metabolite and lipid biomarkers to distinguish NASH and steatosis patients, and reveal candidates, such as glycerophospholipids, sphingolipids, and sterols, as potential NAFLD biomarkers [5]. Moreover, additional LC-MS/MS-based metabolomics studies reported elevated sphingomyelins and phosphatidylcholine in the serum of NAFLD patients when compared to healthy controls [12]. Collectively, these studies demonstrate that metabolism is significantly dysregulated between the NAFLD and control groups, hence, identifying the differential metabolic and lipidomic markers are anticipated to advance clinical diagnosis and inspire the development of treatment for NAFLD patients.

Nevertheless, these non-targeted metabolomics studies that aim to characterize the biomarkers for NALFD usually only employ one specific separation mode and one specific polarity mode for data acquisition. Thus, it is almost impossible to achieve comprehensive retention of all hydrophilic and hydrophobic substances by solely employing either hydrophilic interaction chromatography (HILIC) or reversed-phase liquid chromatography (RPLC). Similarly, only using positive ion mode or negative ion mode cannot obtain all substances with different ionization characteristics. Given the complementarity between distinct chromatographic columns and ionization modes, we have attempted to combine different chromatographic separation columns and ionization modes to optimize the metabolome coverage for biomarker discovery.

We designed the metabolome study by integrating four analytical modes, including the T3 column-based reverse phase separation plus the positive ion mode (T3-pos), the T3 column-based reverse phase separation plus the negative ion mode (T3-neg), the amide column-based HILIC separation plus the positive ion mode (amide-pos), and the amide column-based HILIC separation plus the negative ion mode (amide-neg). Database search-assisted metabolic identification shows that each of the four modes is capable of detecting unique metabolites that cannot be addressed by the other analytical modes. Accordingly, the optimal metabolome coverage should be able to be achieved by the integration of the four modes. A total of 20 potential metabolic biomarkers associated with NAFLD were identified by the integrative analytical mode using a limited number of serum samples collected from clinical NAFLD patients and healthy subjects in this pilot study. Notably, our metabolomic analysis revealed the significantly elevated concentration of selected lipids, acylcarnitines, and organic acids in NAFLD patients. By analyzing the small pool of patients recruited for this pilot study, we have demonstrated that metabolome coverage improvement can be achieved using the integrative analytical modes (i.e., T3-pos, T3-neg, amide-pos, amide-neg), whereas a single analytical method was not adequate for comprehensive coverage of metabolomic/lipidomic biomarkers. We anticipate that the integrative analytical mode we tested for biomarker discovery will contribute to translational research in future metabolomics research.

Material and methods

Patients

We recruited NAFLD patients from the Nanjing First Hospital (Nanjing, Jiangsu Province, China). Eight subjects with NAFLD and eight healthy controls participated in the metabolomic study. Hepatic ultrasonography was performed on all human subjects by the same investigator to confirm the presence or absence of steatosis. None of the patients were on any medications known to cause hepatic steatosis, nor were they taking vitamin supplements. Dietary advice and standardized overnight fasting were implemented to ensure the collected human serum samples' uniformity and consistency. All participants provided informed consent, and the local ethics committee approved the protocol of this study. Fasting blood samples were drawn via venipuncture by clinical nurses. The blood sample was collected and then centrifuged at 3,500 g for 10 min at 4 °C, and the clear serum was stored at − 80 °C.

Reagents

Formic acid (FA), ammonia solution, and ammonium formate were purchased from Sigma-Aldrich (Saint Louis, MO, USA). Ultra-pure water was obtained from the Milli-Q Pure Water System (Bedford, MA, USA). LC-MS grade methanol and acetonitrile (ACN) were purchased from Merck (Darmstadt, Germany).

Sample preparation for metabolomic analysis

Frozen human serum samples were thawed at 4 ℃ and processed for metabolomic analysis as previously reported. Briefly, metabolites were extracted by adding 400 μL of ice-cold methanol to a 100 μL aliquot of human serum followed by 2 min vortexing. The mixture was then centrifuged at 4 °C and 15,000 g for 5 min, and the supernatant was collected. The remaining pellet was repetitively extracted with ice-cold methanol. All the supernatants were pooled and evaporated to dryness. The extracts were subsequently reconstituted in 100 μL of 50% methanol containing 0.1% FA, centrifuged at 18,000 g for 10 min, and then collected. For quality control, aliquots of each sample were pooled and injected throughout the injection queue to monitor the reproducibility of instrumental performance.

Chromatography

Chromatographic separation of serum metabolites was performed on a UPLC system (I Class Plus, Waters, Milford, MA, USA). Two chromatographic separation modes were employed for improved metabolome coverage using an ACQUITY UPLC HSS T3 column (2.1 × 100 mm, 1.8 μm, Waters) for RP separation and an XBridge BEH Amide column (4.6 × 100 mm, 3.5 μm, Waters) for HILIC mode separation. During separation, the T3 column temperature was set at 40 ℃, and the flow rate was 0.4 mL/min. Mobile phase A consisted of 95 % 10 mM ammonium formate buffer, and mobile phase B consisted of ACN. The gradient was set as follows: 0–1.00 min, 0% B; 1.00–4.00 min, 0–35% B; 4.00–15.50 min, 35–95% B; 15.50–18.00 min, 95% B; 18.10–23.00 min, 0% B. The temperature of the BEH amide column was set at 40 ℃, and the flow rate was 0.4 mL/min. Mobile phase A consisted of 95 % 10 mM ammonium acetate buffer and 5% ACN (pH adjusted to 9), and mobile phase B consisted of ACN. The gradient was set as follows: 0–3 min, 85 % B; 3–6 min, 85–30% B; 6–15 min, 30–2% B; 15–18 min, 2% B; 18–19 min, 2–85% B; and 19–26 min, 85% B.

Mass spectrometry

Metabolites and lipids were detected on a UPLC-QTOF-MS/MS mass

spectrometer (Waters, Milford, MA, USA) equipped with both ionization polarity modes (ESI − and ESI +). The detection parameters were set based on previous literature as follows: ion spray voltage at + 3.0 kV (ESI +) and − 2.5 kV (ESI −), sampling cone at 40, source offset at 80, source temperature at 100 ℃, desolvation temperature at 450 ℃, cone gas at 50 L/h, desolvation gas at 600 L/h. MS parameters were set as follows: MS scan *m/z* range at 50–700 Da, MS survey scan time at 0.20 s, MS/MS scan time at 0.15 s, MS survey switching threshold at 8000, MS/MS collision energy at 10–20 V for the low-end and 24–50 V for the high-end. All acquisition was accompanied by lockspray infusion of 200 ng/mL leucine encephalin (*m/z* 556.2771 at positive ion mode and *m/z* 554.2615 at negative ion mode). Metabolomics data acquisition was conducted using the software MassLynx 4.1 (Waters).

Data processing

The control and NAFLD subjects' metabolomics and lipidomics profiles of serum samples were compared using Progenesis QI (version 2.0, Waters) for potential biomarker discovery. The metabolic and lipidomic identification process in Progenesis QI involved the following steps: create a new experiment, import data, review alignment, experiment design setup, peak picking, review deconvolution, identify compounds, review compounds, and compound statistics. Specifically, $[M + H]^+$, $[M + Na]^+$, $[M + K]^+$, $[M + NH_4]^+$, $[M + H-H_2O]^+$ were set for possible positive adducts and $[M−H]^-$, $[M−H_2O−H]^-$, $[M + FA-H]^-$ were selected for possible negative adducts in creating a new experiment. Leucine enkephalin (LE) lock mass *m/z* 556.2766 was set in positive mode, and *m/z* 554.2620 was set in the negative mode. For compound identification, MS and MS/MS fragment ions generated from compounds of interest were searched against those documented in databases including the Human Metabolome Database (http://www.hmdb.ca/), MONA (http://mona.fiehnlab.ucdavis.edu/), and METLIN (https://metlin.scripps.edu/). The identification criteria were set as follows: a mass error of 10 ppm was allowed for precursor ion matching and 20 ppm for fragment ion matching; the error of isotope distribution was set as<20%; the library score of precursor ions must be higher than 40 and the fragmentation spectrum of each putative biomarker was manually inspected to verify the assigned identity. When comparing the abundance levels of lipids/metabolites between the two groups for biomarker discovery, the abundance of each assigned compound was calculated by integrating the corresponding chromatographic peak followed by normalization against the total ion chromatogram (TIC) of each sample using Progenesis QI. Manual inspection of extracted ion chromatograms (EICs) of each identified compound was also performed to ensure data integrity.

Statistical analysis

The metabolomics data were processed by QI and further exported to EZinfo software (version 3.0, Waters) for orthogonal partial least square discriminant analysis (OPLS-DA). Upon OPLS-DA, compounds with variable importance in the projection calculated based on S-plots (VIP) > 1, P-value < 0.05 (Student's *t*-test) was considered statistically significant difference and fold change (FC) > 1.5 were summarized and assigned as a potential biomarker of NAFLD. Specifically, for identical compounds yet detected in different analytical modes (chromatography and/or ion polarity), we adopted the peak displaying the smallest FC for further data analysis. After metabolic identification, bioinformatic analysis of assigned metabolomic and lipidomic identities, including pathway analysis and receiver operating characteristic (ROC) curve were plotted on MetaboAnalyst 3.0 (http://www.metaboanalyst.ca/).

Results and discussion

Metabolomics analysis of NAFLD using UPLCQ-TOF/MS

Chromatographic separation and MS acquisition conditions are of paramount importance for comprehensively detecting metabolic and lipidomic components in biological samples. Therefore, different chromatographic columns and MS acquisition modes were combinatorialy employed to achieve increased detection coverage compared to single separation and MS polarity conditions. The integration of the HSS T3 RP column with a BEH amide column and the use of both positive and negative MS acquisition modes allows for detection of hydrophilic and hydrophobic metabolites. Specifically, we extracted characteristic peaks from the serum samples using the T3 column and amide column in both positive and negative modes. Of the extracted peaks, 1163, 697, 459, and 396 compounds were detected when T3-pos, T3-neg, amide-pos, and amide-neg mode were employed, respectively (Fig. 1A). Subsequently, metabolic identification was tentatively made by searching these compounds against MS and MS/MS spectra in public databases exemplified by METLIN, MONA, and HMDB. Under the matching scores threshold of 40, we retrieved 303, 103, 142, and 51 MS and MS/MS metabolites as high-confidence identifications by using the four analytical modes, respectively (Fig. 1B).

Next, we asked whether the integration of the four analytical modes truly benefits the detection coverage. Notably, for high-confidence metabolites detected by the T3 column-based chromatographic separation in the positive and negative ion modes, a total of 10 compounds were shared between the T3-pos and T3-neg modes, with 293 metabolites being exclusively detected in the T3-pos mode *vs.* 93 in the T3-neg mode (Fig. 1C). Only 2 compounds were shared between the amide-pos and amide-neg data, with 140 metabolites being exclusively detected in the amide-pos mode *vs.* 49 in the amide-neg mode. Our results highlight the necessity of conducting metabolomic analysis in both positive and negative ionization modes.

We next explored whether chromatographic behaviour would markedly impact the detection coverage for metabolites of similar ionization properties. We found 53 overlapping metabolites were detected in both T3-pos and amide-pos modes, whereas 242 metabolites were only detected in the T3-pos mode and 86 metabolites were exclusively identified in the amide-pos mode. The analysis of T3-neg and amide-neg data has led us to assign 8 overlapping metabolites in both analytical modes. In contrast, only 85 metabolites were exclusively detected in the T3-neg mode and 40 metabolites were identified in the amide-neg mode (Fig. 1C).

We then asked whether the metabolites detected by different analytical modes belong to distinct metabolic pathways (Fig. 1D). We found metabolites detected in the T3-pos mode are involved in tryptophan metabolism, arginine biosynthesis, caffeine metabolism, aminoacyl-tRNA biosynthesis, and riboflavin metabolism. In contrast, metabolites detected using the same column, but in the negative mode, are involved in the ascorbate and alternate metabolism and lysine degradation. The enriched pathways under the amide-pos mode include arginine biosynthesis, histidine metabolism, D-glutamine, D-glutamate metabolism, arginine, and proline phenylalanine metabolism, and porphyrin and chlorophyll metabolism. Metabolic pathways including pentose and glucuronate interconversions and lipid metabolism pathways were enriched in the amide-neg analytical mode. Collectively, we summarized that it is necessary to employ different chromatography separation and polarity modes for achieving comprehensive metabolome and lipidome coverage due to their variant physicochemical properties. The improved coverage is anticipated to allow us to glean information regarding the altered biological pathways in response to physiological and pathological changes.
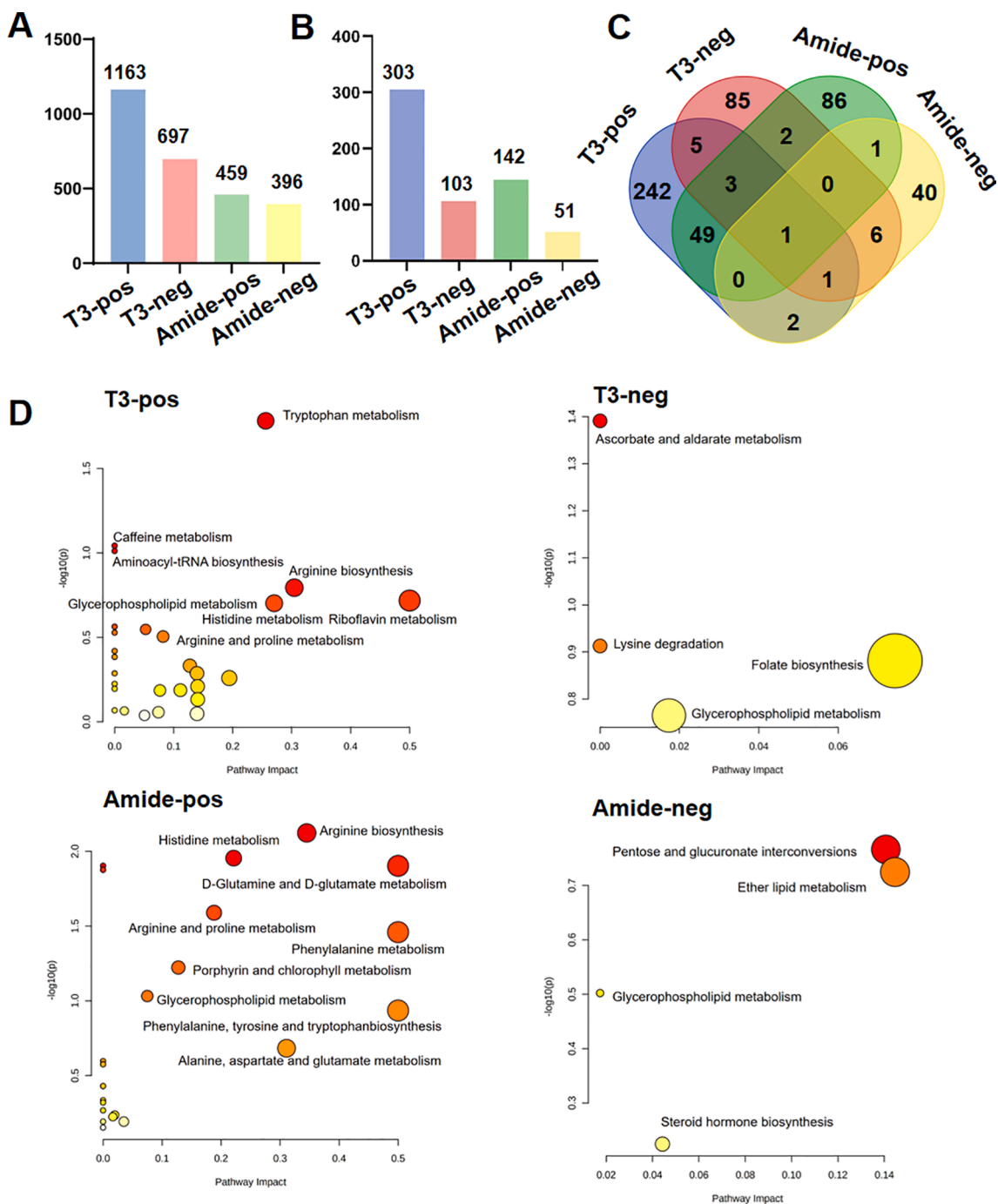
Fig. 1. Identification of metabolites and pathway analysis through different acquisition modes. (A) Histogram showing the distribution of the detected compounds among the T3-pos, T3-neg, amide-pos, and amide-neg analytical modes. (B) Histogram showing the distribution of the identified compounds among the T3-pos, T3-neg, amide-pos, and amide-neg analytical modes. (C) Venn diagram showing the distribution of the identified compounds among the T3-pos, T3-neg, amide-pos, and amide-neg analytical modes. (D) Metabolic pathways involving the identified metabolites in the control and NAFLD subjects were collected by the T3-pos, T3-neg, amide-pos, and amide-neg analytical modes. The larger the representative circle's size and the deeper the red color intensity, the greater the impact on the respective pathway. A total of 16 serum samples from NAFLD patients and healthy controls were acquired on LC-Q/TOF under different analytical modes, and datasets acquired under different analytical modes were subjected to data analysis and comparison. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Metabolic and lipidomic changes in NAFLD patients sera

We then employed the four analytical modes to obtain a holistic profile from the sera of NAFLD patients and healthy subjects. As shown in Fig. 2, we found the Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) plots showed that the healthy subjects and NAFLD patients all exhibited different patterns under the four analytical modes ($R^2X = 98\%$ $Q^2 = 97\%$ for the T3-pos mode, $R^2X = 99\%$ $Q^2 = 99\%$ for the T3-neg mode, $R^2X = 99\%$ $Q^2 = 98\%$ for the amide-pos mode, $R^2X = 99\%$ $Q^2 = 98\%$ for the amide-neg mode). These scoring plots signify the distinct metabolic and lipidomic patterns between NAFLD and healthy subjects.

To elaborate on which metabolites and lipids show significantly aberrant levels in NAFLD patients, a student's *t*-test followed by Tukey's
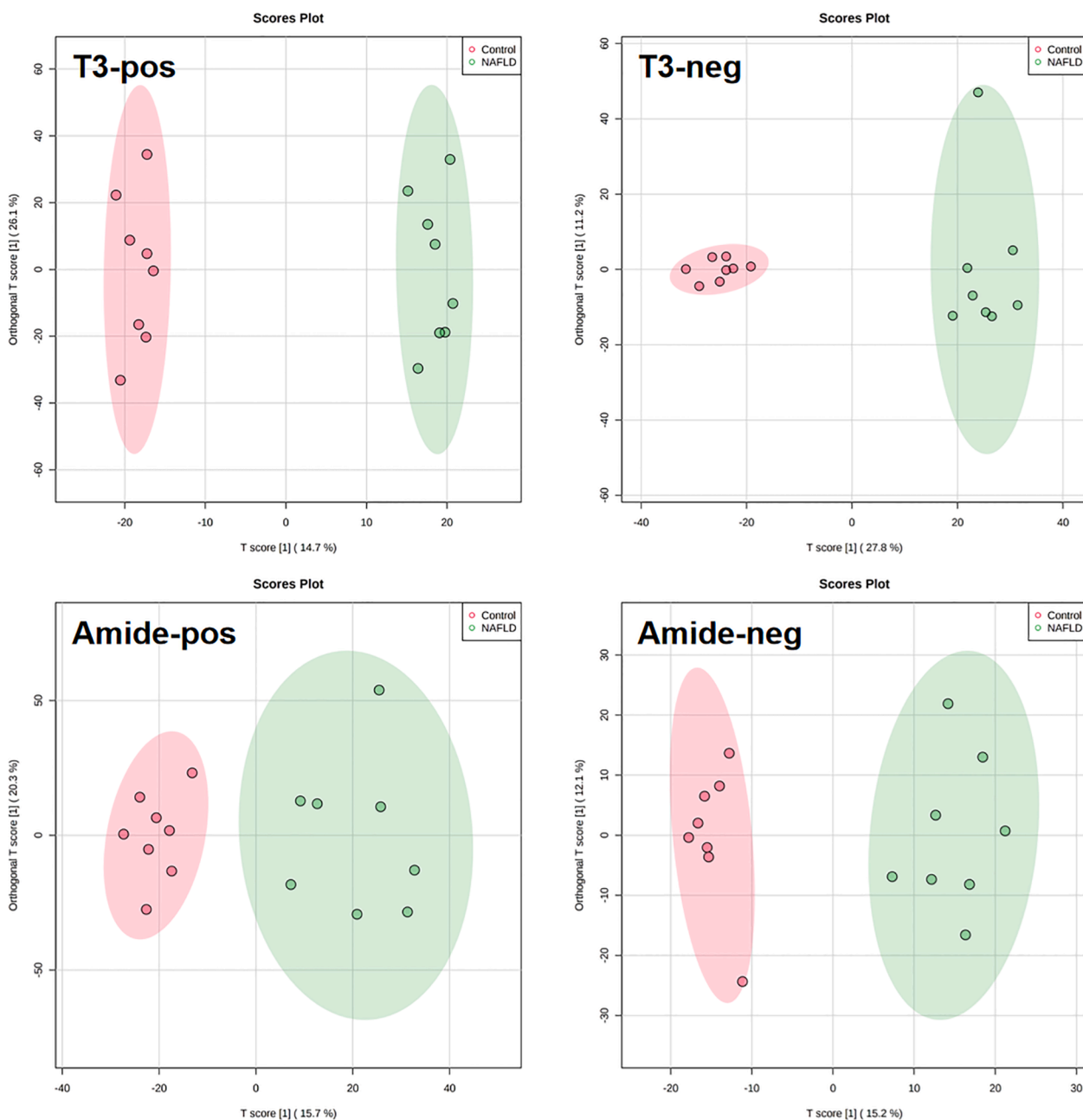
Fig. 2. OPLS-DA score plots derived from the serum metabolomics datasets collected from the control and NAFLD subjects using LC-Q/TOF via the T3-pos, T3-neg, amide-pos, and amide-neg analytical mode (n = 8/group).

test for post-hoc (*p* value < 0.05) was performed for metabolites with VIP > 1 based on the established OPLS-DA models. An FC cutoff of 1.5 was further calculated to identify the potential metabolic and lipidomic biomarkers that display differentiation in NAFLD patients and healthy subjects. A total of 20 potential metabolic biomarkers were combinatorially identified using the T3-pos, T3-neg, amide-pos, and amide-neg analytical modes (summarized in Table 1), and their distribution in each analytical mode is illustrated in Fig. 3A. Specifically, 9, 4, 7, and 4 putative metabolic biomarkers were detected in the T3-pos, T3-neg, amide-pos, and amide-neg modes, respectively. Specifically, lysophosphatidylcholine (LPC) 20:3 was detected in T3-pos, T3-neg, and amide-pos modes, whereas valerylcarnitine was detected in both T3-pos and amide-pos modes, and retinyl-beta glucuronide was detected in both T3-neg and amide-neg modes (Fig. 3B). These findings confirmed the necessity of adopting multiple chromatographic separation and polarity modes for metabolic biomarker discovery. Expectedly, we found that

metabolites detected in both analytical modes exhibited consistent trends in abundance changes.

ROC curve analysis of the potential biomarkers

We then created ROC curves to explore whether these putative biomarkers have potential for clinically diagnosing NAFLD. The metabolic biomarkers identified to differentiate serum profiles collected from the healthy and NAFLD subjects were analyzed using binary logistic regression on Metaboanalyst to obtain the predictive probability. Although the patient cohort was small and there may have been overfitting, the AUC of putative biomarkers reached 0.984 (Fig. 3C), which appears to indicate that the identified metabolites can differentiate healthy subjects from NAFLD patients based on serum metabolic and lipidomic profiles. However, we also realize that the small size of the cohort enrolled in this study could have easily resulted in overfitting of

Table 1

Potential metabolic biomarkers differentiating the NAFLD and control subjects using different analytical modes (T3-pos, T3-neg, amide-pos, amide-neg). All metabolic identifications have been confirmed by MS and MS/MS. Fold change was calculated by comparing the abundance level of the metabolite in the NAFLD group *vs.* the control group (NAFLD/Control). *P* value was calculated using the Student's *t*-test.

Mode	Metabolite	Formula	Mass (Da)	Retention time (min)	Adducts[a]	Observed (*m/z*)	Mass error (*ppm*)	Fold change	*P* value
T3-pos	Valerylcarnitine	$C_{12}H_{23}NO_4$	245.1627	3.31	$[M + H]^+$	246.1694	−2.25	1.92	0.0386
	Octanoylcarnitine	$C_{15}H_{29}NO_4$	287.2097	4.47	$[M + H]^+$	288.2164	−1.74	4.65	0.0124
	Hydroxyundecanoyl carnitine	$C_{18}H_{35}NO_5$	345.2515	5.08	$[M + H-H_2O]^+$	328.2480	−0.59	2.01	0.0197
	Hydroxydodecanedioic acid	$C_{12}H_{22}O_5$	246.1461	4.50	$[M + H-H_2O]^+$	229.1428	−2.39	2.11	0.0045
	PE(17:0/0:0)	$C_{22}H_{46}NO_7P$	467.3018	7.40	$[M + H]^+$	468.3091	1.26	1.76	0.0121
	LPC(18:3)	$C_{26}H_{48}NO_7P$	517.3168	7.81	$[M + H]^+$	518.3242	0.22	1.55	0.0123
	LPC(20:3)	$C_{28}H_{52}NO_7P$	545.3481	8.91	$[M + H]^+$	546.3561	1.27	1.64	0.0300
	25-Hydroxyvitamin D3-26,23-lactol	$C_{27}H_{42}O_4$	430.3083	9.38	$[M + H]^+$	431.3155	−0.20	1.60	0.0203
	MG(18:1)	$C_{21}H_{40}O_4$	356.2925	13.77	$[M + H]^+$	357.3016	−0.53	3.65	0.0250
T3-neg	Deoxycholic acid 3-glucuronide	$C_{30}H_{48}O_{10}$	568.3247	5.83	$[M−H]^-$	567.3165	−1.70	4.76	0.0084
	Tuftsin	$C_{21}H_{40}N_8O_6$	500.3071	5.34	$[M−H]^-$	499.3007	1.75	3.89	0.0211
	LPC(20:3)	$C_{28}H_{52}NO_7P$	545.3481	8.91	$[M + FA-H]^-$	590.3411	−9.61	1.77	0.0210
	Retinyl beta-glucuronide	$C_{26}H_{38}O_7$	462.2618	5.45	$[M−H]^-$	461.2554	1.93	0.64	0.0427
Amide-pos	LPC(14:0)	$C_{22}H_{46}NO_7P$	467.3012	3.39	$[M + H]^+$	468.3097	2.65	2.73	0.0026
	DG(20:3/20:4)	$C_{43}H_{70}O_5$	666.5234	2.43	$[M + H]^+$	684.5542	1.62	7.22	0.0261
	Valerylcarnitine	$C_{12}H_{23}NO_4$	245.1627	7.36	$[M + H]^+$	246.1697	−0.99	1.99	0.0045
	3-Hydroxy-*cis*-5-tetradecenoylcarnitine	$C_{21}H_{39}NO_5$	385.2828	3.34	$[M + H-H_2O]^+$	368.2798	0.58	1.69	0.0057
	LPC(20:3)	$C_{28}H_{52}NO_7P$	545.3491	3.31	$[M + H]^+$	546.3564	1.83	1.67	0.0224
	Hydroxylauroylcarnitine	$C_{19}H_{37}NO_5$	359.2672	5.40	$[M + H-H_2O]^+$	342.2641	0.69	1.55	0.0251
	DG(18:1)/18:2)	$C_{39}H_{70}O_5$	618.5191	2.45	$[M + H-H_2O]^+$	636.5558	−5.16	2.74	0.0423
Amide-neg	Cortolone-3-glucuronide	$C_{27}H_{42}O_{11}$	542.2727	7.33	$[M−H]^-$	541.2697	7.93	2.58	0.0004
	Tetrahydroaldosterone-3-glucuronide	$C_{27}H_{40}O_{11}$	540.2571	7.22	$[M−H]^-$	539.2548	9.24	2.02	0.0007
	LPC(18:0)	$C_{26}H_{54}NO_7P$	523.3638	3.34	$[M + FA-H]^-$	568.3603	−3.31	2.13	0.0323
	Retinyl beta-glucuronide	$C_{26}H_{38}O_7$	462.2618	3.11	$[M−H]^-$	461.2564	4.17	0.57	0.0228

[a] The metabolite adducts of the highest intensity were chosen to represent the abundance of given metabolites.

the model. As with most pilot studies, it is clear that a large population of healthy subjects and NAFLD patients will be necessary to verify the diagnostic value of these putatively assigned biomarkers of NAFLD.

Metabolomic and lipidomic biomarkers of NAFLD and their pathological roles

We created a heatmap to visualize the 20 potential biomarkers consisting of lipids, acylcarnitines, and organic acids that we used to differentiate between NAFLD and healthy subjects (Fig. 4). PE(17:0/0:0), LPC 14:0, LPC 18:0, LPC 20:3, LPC 18:3, DG(18:1/18:2), DG(20:3/20:4), and MG (18:1) all displayed increased abundance in the NAFLD group as shown in Fig. 5A. These lipids can be classified as phosphatidylethanolamines (PEs), phosphorylcholines (PCs), lysophosphatidylcholines (LPCs), diacylglycerols (DGs), and monoacylglycerols (MGs). Indeed, the increased PCs and PEs in NAFLD patients have been previously reported and such increase might be attributed to their biological roles. PCs and PEs are the most abundant glycerophospholipids in all eukaryotic cell membranes [13], whereas the biosynthesis of PCs in the human liver is mainly through the methylation of PEs [14]. Upon formation, hepatic PCs are specifically used to form monolayers of lipid droplets and can simultaneously function as a surfactant to prevent the monolayer from undergoing coalescence. Thus, larger lipid droplets are less likely to go through lipolysis in presence of abundant PCs, which eventually leads to liver steatosis [15]. Besides NAFLD, abnormal levels of PC and PE have also been observed in serum samples of patients with Alzheimer's disease and Parkinson's disease [16].

LPC is a monoglycerophospholipid in which a phosphorylcholine moiety occupies a glycerol substitution site. An elevated level of LPCs was also noted in the sera of NAFLD patients compared with the healthy subjects. Previously, LPCs and DGs have been considered lipotoxic

species and associated with hepatic injury [17]. Notably, LPC levels were also found elevated in a methionine-choline-deficient (MCD) diet-induced advanced NASH mouse model [18]. Thus, our results substantiate a possible link between increased LPC levels and NAFLD and further progression to NASH [19]. Additionally, the pathogenic role of DGs has been associated with its function as the activator of protein kinase Cϵ (PKCϵ), which leads to impaired insulin signaling. Thus, the high level of DGs in NAFLD patients, as we observed here, potentially explains why many NAFLD patients exhibit hepatic insulin resistance and type 2 diabetes [20]. Further, a previous study has recorded insulin-mediated deactivated expression of lipid hydrolysis enzymes in NAFLD patients [21], explaining the accumulation of DGs and MGs in sera of NAFLD patients, as detected in this study. Accordingly, we hypothesize that preventing the generation and accumulation of these lipotoxic lipids appears to serve as a therapeutic strategy for NAFLD treatment [24].

Besides lipids, our preliminary results also draw attention to acylcarnitines, such as octanoylcarnitine (C8), 3-hydroxy-*cis*-5-tetradecenoylcarnitine (C14), valerylcarnitine (C5), and two medium-chain acylcarnitines, hydroxyundecanoyl carnitine and hydroxylauroylcarnitine, that all exhibited marked increases in the NAFLD group compared to the healthy control group (Fig. 5B). We, thus, wondered why acylcarnitines were abnormally increased in NAFLD patients. Previous biochemical studies have shown that acylcarnitines derive from the esterification of L-carnitine in mammalian cells, a process responsible for transporting fatty acids into the mitochondria for subsequent β-oxidation [22]. Moreover, acylcarnitines, together with cardiolipin and ubiquinone, are lipid classes associated with mitochondrial function [23]. Thus, abnormal acylcarnitines levels in NAFLD patients could reflect the inability of mitochondria to properly metabolize fatty acids [24]. In agreement with this hypothesis, mitochondrial dysfunction has
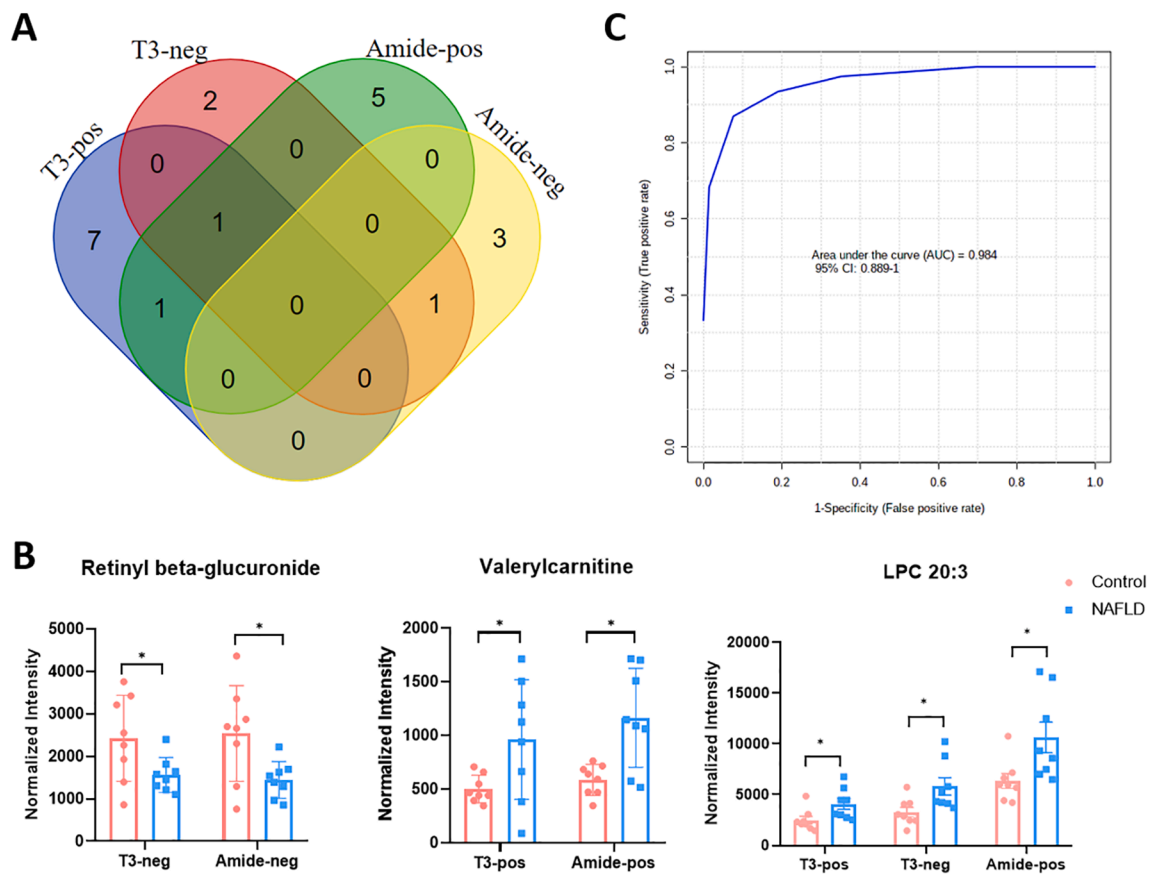
Fig. 3. Identification of potential metabolic biomarkers under different acquisition modes. (A) Venn diagram shows the potential biomarkers' distribution among the T3-pos, T3-neg, amide-pos, and amide-neg modes between NAFLD and control subjects. (B) Abundance changes of retinyl beta-glucuronide, valerylcarnitine, and LPC (20:3) in the NAFLD and control subjects collected in different modes. Bar plot represents the normalized intensity (normalized by total ion chromatogram, TICs) of selected metabolites in serum. Data are expressed as normalized intensity ± SEM (n = 8/group), and fold change is indicated as FC. A comparison was made between NAFLD and control subjects using the Student's t-test. *, significantly different between the control and NAFLD subjects at $p < 0.05$; **, significantly different between the control and NAFLD subjects at $p < 0.01$. (C) ROC curve analysis of the potential biomarkers for NAFLD.
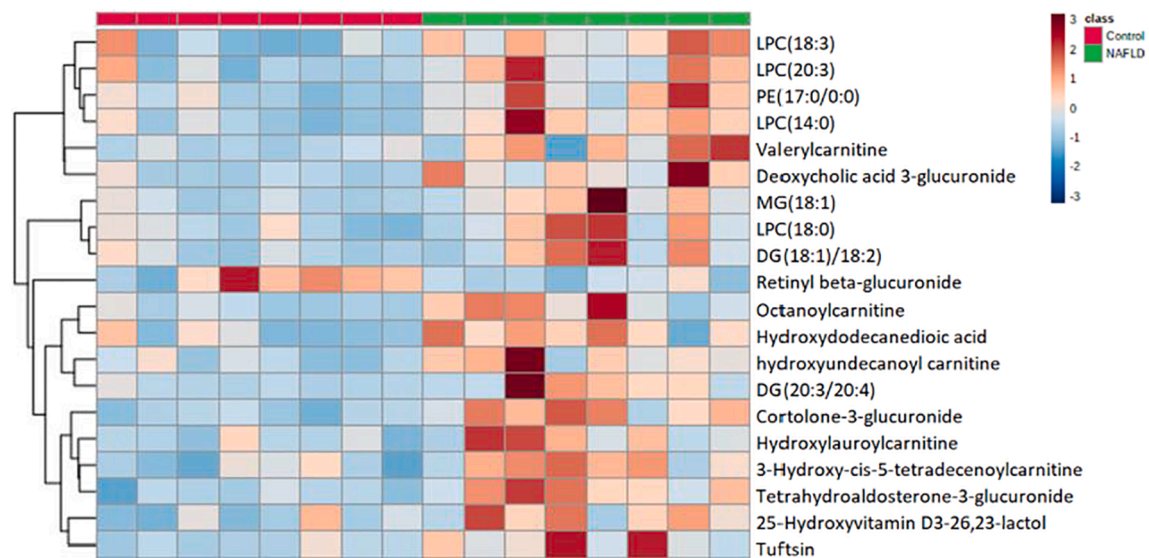

Fig. 4. Heat map of hierarchical clustering of the serum samples from the control and NAFLD subjects. Red blocks indicate up-regulation, whereas blue indicates down-regulation (n = 8/group). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

been suggested to play a pivotal role in the pathogenesis of NAFLD. The dysregulated homeostasis of acylcarnitines, as we detected, is supported by a recent study that records elevated octenoylcarnitine (C8), 3-

hydroxy-*cis*-5-tetradecenoylcarnitine (C14), decanoylcarnitine (C10), dodecenoylcarnitine (C12) in serum samples of overweight NAFLD individuals [25]. Further, excessive total carnitine, including free
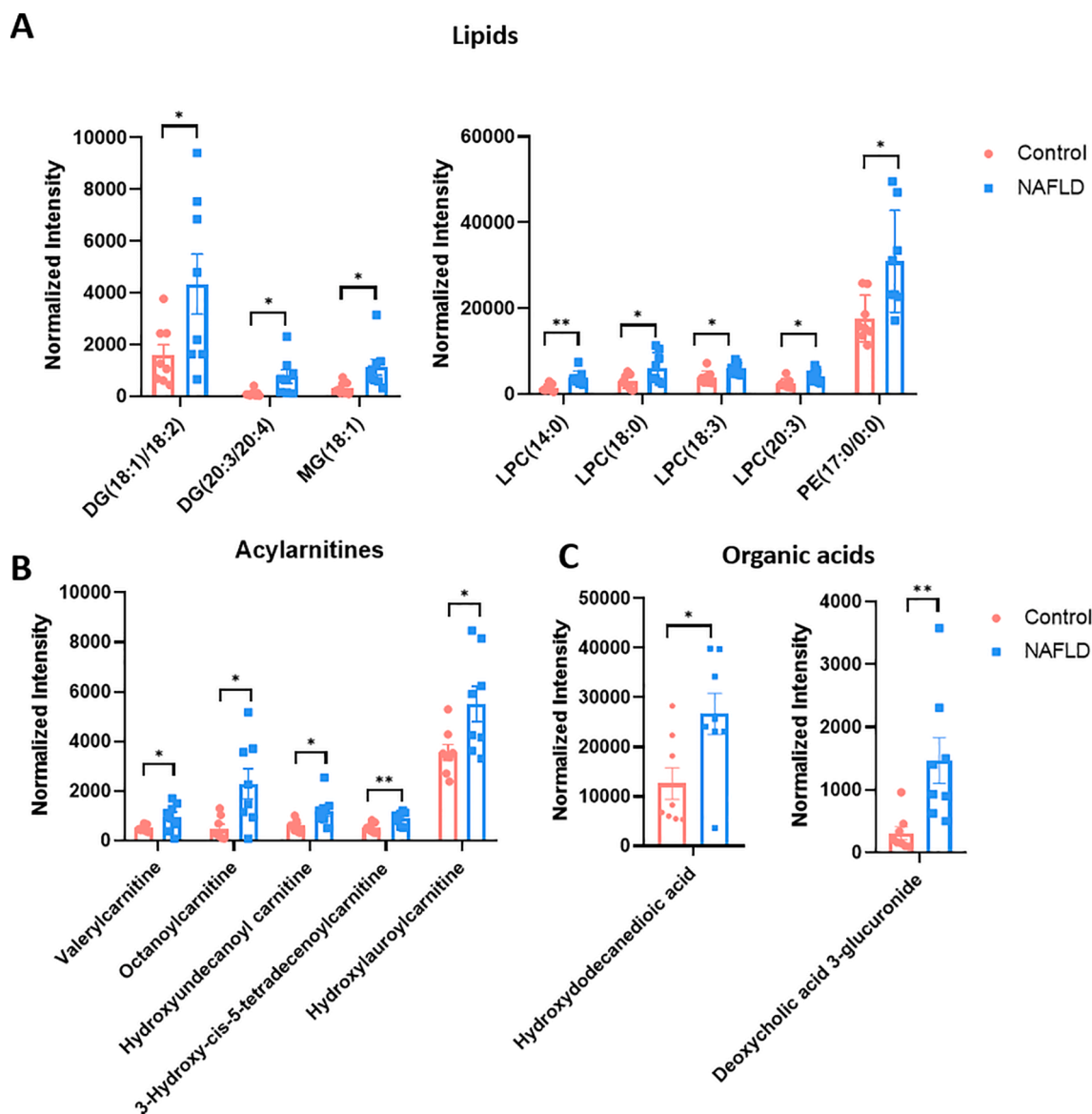
Fig. 5. Abundance changes of representative serum metabolic and lipidomic biomarkers that differentiate the NAFLD and healthy subjects. (A) Comparison of lipids in the control and NAFLD subjects. (B) Comparison of acylcarnitines in the control and NAFLD subjects. (C) Comparison of organic acids in the control and NAFLD subjects. Bar plot represents the normalized intensity (normalized to TIC) of metabolites detected from each serum sample. Data are represented as normalized intensity $\pm$ SEM (n = 8/group). A comparison was made between the NAFLD and control subjects by Student's t-test. *, significantly different between the control and NAFLD subjects at $p < 0.05$; **, significantly different between the control and NAFLD subjects at $p < 0.01$.

carnitine, as well as short- and long-chain acylcarnitines, were also found in cirrhotic patients [26]. Together, our results confirm acylcarnitines as potential metabolic biomarkers of NAFLD, but how mitochondria dysfunction, as suggested by aberrant acylcarnitines levels, is correlated with NAFLD occurrence and progression warrants future investigations [27,29].

Besides lipids and acylcarnitines, we find organic acids hydroxydodecanedioic acid and deoxycholic acid 3-glucuronide are also useful metabolites in differentiating NAFLD and health subjects (Fig. 5C). Of note, deoxycholic acid 3-glucuronide in human serum is a natural metabolite of deoxycholic acid generated in the liver by uridine 5′-diphospho- glucuonyltransferase. Deoxycholic acid is a secondary bile acid produced by intestinal microbiota and its function is to emulsify fats for intestinal absorption. Interestingly, a recent study reported elevated serum concentrations of secondary bile acids in NAFLD patients [28], which may correlate with the increased deoxycholic acid 3-glucuronide, as we detected here. Together, the putative metabolic and lipidomic

biomarkers we report in this pilot study suggest the potential to discover promising non-invasive histopathological markers for improving NAFLD diagnosis and stimulating therapeutic strategies.

Conclusions

Untargeted metabolomic analysis aims to comprehensively measure changes in endogenous metabolites. We have optimized a non-targeted serum metabolomic acquisition workflow using both T3 and HILIC columns and adopted both positive and negative ionization modes. Comparison of the four analytical modes has highlighted the necessity to include all modes of analysis in order to optimize coverage of the metabolome. Integration of the metabolomic data collected under the four combinatorial modes contributed to detecting 20 putative NAFLD biomarkers in this pilot study. The significant changes in lipid species, acylcarnitines, and organic acid levels from NAFLD patients revealed in this study agree with previous reports and reinforces the foundation for

continued investigation of NAFLD biomarkers for translatation into diagnostics and therapeutic interventions in clinical practice.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmsacl.2021.10.001.

References

[1] C.D. Byrne, G. Targher, NAFLD: a multisystem disease, J. Hepatol. 62 (1) (2015) S47–S64.

[2] S.L. Friedman, B.A. Neuschwander-Tetri, M. Rinella, A.J. Sanyal, Mechanisms of NAFLD development and therapeutic strategies, Nat. Med. 24 (7) (2018) 908–922.

[3] Z. Younossi, F. Tacke, M. Arrese, B. Chander Sharma, I. Mostafa, E. Bugianesi, V. Wai-Sun Wong, Y. Yilmaz, J. George, J. Fan, M.B. Vos, Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis, Hepatology 69 (6) (2019) 2672–2682.

[4] V.-S. Wong, L.A. Adams, V. de Lédinghen, G.-H. Wong, S. Sookoian, Noninvasive biomarkers in NAFLD and NASH - current progress and future promise, Nat. Rev. Gastroenterol. Hepatol. 15 (8) (2018) 461–478.

[5] D.L. Gorden, D.S. Myers, P.T. Ivanova, E. Fahy, M.R. Maurya, S. Gupta, J. Min, N. J. Spann, J.G. McDonald, S.L. Kelly, J. Duan, M.C. Sullards, T.J. Leiker, R. M. Barkley, O. Quehenberger, A.M. Armando, S.B. Milne, T.P. Mathews, M. D. Armstrong, C. Li, W.V. Melvin, R.H. Clements, M.K. Washington, A. M. Mendonsa, J.L. Witztum, Z. Guan, C.K. Glass, R.C. Murphy, E.A. Dennis, A. H. Merrill, D.W. Russell, S. Subramaniam, H.A. Brown, Biomarkers of NAFLD progression: a lipidomics approach to an epidemic, J. Lipid Res. 56 (3) (2015) 722–736.

[6] K. Tokushige, E. Hashimoto, K. Kodama, M. Tobari, N. Matsushita, T. Kogiso, M. Taniai, N. Torii, K. Shiratori, Y. Nishizaki, T. Ohga, Y. Ohashi, T. Sato, Serum metabolomic profile and potential biomarkers for severity of fibrosis in nonalcoholic fatty liver disease, J. Gastroenterol. 48 (12) (2013) 1392–1400.

[7] H. Morikawa, K. Mano, H. Horinaka, T. Matsunaka, Y. Matsumoto, T. Ida, Y. Kawaguchi, K. Wada, N. Kawada, Development of non-invasive method for assessment of hepatic steatosis, Ultrasonics 72 (2016) 195–200.

[8] G.N. Ioannou, G.A. Nagana Gowda, D. Djukovic, D. Raftery, Distinguishing NASH Histological Severity Using a Multiplatform Metabolomics Approach, Metabolites 10 (4) (2020) 168, https://doi.org/10.3390/metabo10040168.

[9] S.-S. Zhu, R. Long, T. Song, L.i. Zhang, Y.-L. Dai, S.-W. Liu, P. Zhang, UPLC-Q-TOF/MS Based Metabolomics Approach to Study the Hepatotoxicity of Cantharidin on Mice, Chem. Res. Toxicol. 32 (11) (2019) 2204–2213.

[10] S.C. Kalhan, L. Guo, J. Edmison, S. Dasarathy, A.J. McCullough, R.W. Hanson, M. Milburn, Plasma metabolomic profile in nonalcoholic fatty liver disease, Metabolism 60 (3) (2011) 404–413.

[11] Y.-S. Lai, W.-C. Chen, T.-C. Kuo, C.-T. Ho, C.-H. Kuo, Y.J. Tseng, K.-H. Lu, S.-H. Lin, S. Panyod, L.-Y. Sheen, Mass-Spectrometry-Based Serum Metabolomics of a C57BL/6J Mouse Model of High-Fat-Diet-Induced Non-alcoholic Fatty Liver Disease Development, J. Agric. Food Chem. 63 (35) (2015) 7873–7884.

[12] S. Tiwari-Heckler, H. Gan-Schreier, W. Stremmel, W. Chamulitrat, A. Pathil, Circulating Phospholipid Patterns in NAFLD Patients Associated with a Combination of Metabolic Risk Factors, Nutrients 10 (5) (2018) 649, https://doi.org/10.3390/nu10050649.

[13] N. Perakakis, K. Stefanakis, C.S. Mantzoros, The role of omics in the pathophysiology, diagnosis and treatment of non-alcoholic fatty liver disease, Metabolism 111 (2020) 154320, https://doi.org/10.1016/j.metabol.2020.154320.

[14] F. Gibellini, T.K. Smith, The Kennedy pathway–De novo synthesis of phosphatidylethanolamine and phosphatidylcholine, IUBMB Life 62 (6) (2010) 414–428.

[15] N. Krahmer, Y.i. Guo, F. Wilfling, M. Hilger, S. Lingrell, K. Heger, H. Newman, M. Schmidt-Supprian, D. Vance, M. Mann, R. Farese, T. Walther, Phosphatidylcholine synthesis for lipid droplet expansion is mediated by localized activation of CTP:phosphocholine cytidylyltransferase, Cell Metab. 14 (4) (2011) 504–515.

[16] E. Calzada, O. Onguka, S.M. Claypool, Phosphatidylethanolamine Metabolism in Health and Disease, Int. Rev. Cell Mol. Biol. 321 (2016) 29–88.

[17] E.E. Kershaw, J.K. Hamm, L.A.W. Verhagen, O. Peroni, M. Katic, J.S. Flier, Adipose triglyceride lipase: function, regulation by insulin, and comparison with adiponutrin, Diabetes 55 (1) (2006) 148–157.

[18] A. Pathil, J. Mueller, A. Warth, W. Chamulitrat, W. Stremmel, Ursodeoxycholyl lysophosphatidylethanolamide improves steatosis and inflammation in murine models of nonalcoholic fatty liver disease, Hepatology 55 (5) (2012) 1369–1378.

[19] S.R. Nagarajan, L.M. Butler, A.J. Hoy, The diversity and breadth of cancer cell fatty acid metabolism, Cancer Metab 9 (1) (2021) 2.

[20] M.C. Petersen, G.I. Shulman, Roles of Diacylglycerols and Ceramides in Hepatic Insulin Resistance, Trends Pharmacol. Sci. 38 (7) (2017) 649–665.

[21] R. Zechner, R. Zimmermann, T. Eichmann, S. Kohlwein, G. Haemmerle, A. Lass, F. Madeo, FAT SIGNALS–lipases and lipolysis in lipid metabolism and signaling, Cell Metab. 15 (3) (2012) 279–291.

[22] A. Gnoni, S. Longo, G.V. Gnoni, A.M. Giudetti, Carnitine in Human Muscle Bioenergetics: Can Carnitine Supplementation Improve Physical Exercise? Molecules 25 (1) (2020) 182, https://doi.org/10.3390/molecules25010182.

[23] K.-Y. Peng, M.J. Watt, S. Rensen, J.W. Greve, K. Huynh, K.S. Jayawardana, P. J. Meikle, R.C.R. Meex, Mitochondrial dysfunction-related lipid changes occur in nonalcoholic fatty liver disease progression, J. Lipid Res. 59 (10) (2018) 1977–1986.

[24] B. Bjørndal, et al., Associations between fatty acid oxidation, hepatic mitochondrial function, and plasma acylcarnitine levels in mice, Nutr Metab (Lond) 15 (2018) 10.

[25] M. Kim, M. Kim, M. Kang, H.J. Yoo, M.S. Kim, Y.-T. Ahn, J.-H. Sim, S.H. Jee, J. H. Lee, Effects of weight loss using supplementation with Lactobacillus strains on body fat and medium-chain acylcarnitines in overweight individuals, Food Funct. 8 (1) (2017) 250–261.

[26] D.G. Johnston, K.G. Alberti, Carbohydrate metabolism in liver disease, Clin. Endocrinol. Metab. 5 (3) (1976) 675–702.

[27] J. Li, et al., Serum metabolomic analysis of the effect of exercise on nonalcoholic fatty liver disease, Endocr. Connect 8 (4) (2019) 299–308.

[28] N. Jiao, et al., Suppressed hepatic bile acid signalling despite elevated production of primary and secondary bile acids in NAFLD, Gut 67 (10) (2018) 1881–1891.