Metabolic analysis of early nonalcoholic fatty liver disease in humans using liquid chromatography-mass spectrometry

Cheng Hu1†, Tao Wang2†, Xiaoyu Zhuang1†, Qiaoli Sun3, Xiaochun Wang2, Hui Lin2, Mingli Feng2, Jiaqi Zhang1,4*, Qin Cao2* and Yuanye Jiang2*

Abstract

Background: Nonalcoholic fatty liver disease (NAFLD) is a common metabolic disease that affects 20–30% of individuals worldwide. Liver puncture remains the gold standard for the diagnosis of liver diseases despite limitations regarding invasive nature and sample variability. It is of great clinical significance to find noninvasive biomarkers to detect and predict NAFLD.

Objective: The aims of this study were to identify potential serum markers in individuals with early-stage NAFLD and to advance the mechanistic understanding of this disease using a high-throughput mass spectrometry-based untargeted metabolomics approach.

Methods: One hundred and twelve patients with early-stage NAFLD aged 18–55 were recruited according to the guidelines. The control group included 112 healthy participants. The demographic, anthropometric, clinical and laboratory data of all participants were systematically collected. Serum samples were obtained after an overnight fast. The comprehensive serum metabolomic analysis was performed by ultra-performance liquid chromatography-Orbitrap mass spectrometry. The resultant data was processed by Compound Discover and SIMCA-P software to validate the potential biomarkers. Significantly altered metabolites were evaluated by variable importance in projection value (VIP > 1) and ANOVA (p < 0.01). Pathway analysis was performed using MetaboAnalyst 4.0.

Results: The liver function test of early NAFLD patients showed no statistical differences to control group (p > 0.05). However, obvious differences in blood lipids were observed between subjects with NAFLD and controls (p < 0.001). In total, 55 metabolites showed significant changes in experimental group were identified. The area under curve (AUC) values deduced by receiver operating curve (ROC) analysis indicated that these newly identified biomarkers have high predictability and reliability. Of these, 15 metabolites with AUC greater than 0.9 were of great diagnostic value in early NAFLD patients.
Conclusion: In this study, a total of 15 serum metabolites were found to strongly associate with early NAFLD. These biomarkers may have great clinical significance in the early diagnosis of NAFLD, as well as to follow response to therapeutic interventions.

Keywords: NAFLD, Early diagnosis, Metabolomics, LC–MS/MS, Biomarkers

Background
In recent years, the trend in liver diseases has changed from traditional infectious diseases to metabolic disorders [1]. Liver disease is a high incidence disease in Asia [2–4]. There is strong evidence that a sedentary lifestyle and unhealthy dietary habits (especially those with high fat and high salt) are setting the stage for the prevalence of obesity and nonalcoholic fatty liver disease (NAFLD) in many urbanized Asian countries [5, 6]. NAFLD has also become a burgeoning health problem in developed country [2, 7]. The prevalence of NAFLD is highly underestimated because it often presents with minor to no symptoms in patients at the early stages. Given the increasing trends of obesity and metabolic syndrome, the two basic risk factors for NAFLD, incidence rates are expected to further rise in the next decades. Approximately 20–30% of patients with NAFLD progress to steatohepatitis and fibrosis that may progress to cirrhosis in extreme cases [8]. Considering the possible association between NAFLD-related cirrhosis and hepatocellular cancer (HCC), NAFLD is becoming an increasingly important problem in China where it is currently the primary/most common condition leading to cirrhosis) [9]. HCC has also been found to rise in patients with NAFLD in the absence of cirrhosis [10].

Currently, the histological examination of liver biopsy specimen remains the gold standard for NAFLD diagnosis despite well-acknowledged disadvantages, such as its invasive nature, inevitable sampling error, poor short-term repeatability and subjective differences among observers. Thus liver biopsy is unlikely to be carried out widely as a routine examination method in clinical practice [11]. Efficient diagnosis methods are needed for the facile identification of NAFLD patients, disease progression risk assessment, and monitoring the response to potential new treatment strategies. Radiologic technique including ultrasonography and magnetic resonance imaging (MRI) are the main representatives of non-invasive diagnostic modalities. The assessment of serum biomarkers that considered to be a more convenient and promising approaches for monitoring NAFLD has also been introduced [12, 13]. Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) are the most common blood indicators; however, such liver enzymes may not elevate until histological injury of the liver occurs [14]. A number of patients with normal ALT levels may also have NAFLD and even advanced fibrosis. Therefore, the exploration and identification of novel biomarkers involved in the early stage of NAFLD is of crucial significance. As a high-throughput technology, metabolomics allows thousands of serum metabolites to be measured and identified simultaneously [15–17], and is ideally suited for the discovery of new biomarkers [18, 19]. In this study, the metabolic characteristics of blood samples from NAFLD patients were analyzed to explore more accurate and specific biomarkers.

Materials and methods
Chemicals and reagents
Methanol and acetonitrile (HPLC grade) were purchased from Fisher Chemicals (Waltham, USA); formic acid was purchased from Merck (Darmstadt, Germany); 2-chloro-DL-phenylalanine was from Merck (Darmstadt, Germany).

Participants, inclusion & exclusion criteria
A total of 3802 participants who were hospitalized by the ward and outpatient of digestive department of Putuo Hospital Affiliated to Shanghai University of Traditional Chinese Medicine were enrolled consecutively and separately between January 2019 and December 2020.

According to the guidelines for diagnosis and treatment of nonalcoholic fatty liver disease (2018) [20] formulated by the National Workship on Fatty Liver and Alcoholic Liver Disease and Chinese Society on Hepatology, the diagnosis of early-stage NAFLD was based on the detection of steatosis by abdominal ultrasonography. Only if three deputy director physicians make the same diagnosis could the patient be included.

The following exclusion criteria were used: (1) had a history of liver diseases other than NAFLD, including viral hepatitis, cirrhosis, liver cancer, autoimmune liver disease, alcoholic liver disease, hereditary liver disease, etc.; (2) excessive alcohol consumption (≥ 210 g/week for men, ≥ 140 g/week for women); (3) were taking medications that can affect metabolism or cause liver damage; (4) had been diagnosed with diabetes, had received or were undergoing hypoglycemic drugs or insulin treatment; (5) had severe heart disease (myocardial infarction, heart failure and / or severe arrhythmia); (6) had severe infections and severe trauma; (7)
pregnant or breastfeeding women; (8) had thyroid dis-
eases, including hyperthyroidism and hypothyroidism.

In the end, 112 subjects (68 males and 44 females,
aged ≥ 18 years) met the diagnosis criteria of the
research (Fig. 1). The control group included 112
healthy people. This research was conducted in accord-
ance with the Declaration of Helsinki to protect the
health and rights of the participants. Written informed
consent was obtained from each participant. The study
protocol was approved by the Medical Ethics Commit-
tee of the Putuo Hospital Affiliated to Shanghai Uni-
versity of traditional Chinese Medicine (Ethics approval
number was PTEC-A-2018–49-1).

Data collection
The medical history and physical examination of the
participants were collected by the full-time doctors in
the outpatient or ward of the department of gastroen-
terology, and the general conditions of the participants
were recorded in detail, including name, gender, age,
medical history, smoking and drinking history, etc.
The total amount of alcohol consumed per week was
classified into three grades: nondrinkers, light drinkers(< 70 g/week), or moderate drinkers
(7–210 g/week for male and 70–140 g/week for female)
and defined nondrinkers as did not drink any alcohol in
the past 12 months [21].

Anthropometric and biomedical measurements
The height, weight, waist circumference and blood
pressure were measured in the morning by a specially
assigned physician. The waist was measured at the mid-
dle point of the line between the lower edge of the arch
and the iliac spine, measured twice for an average. Blood
pressure was measured in quiet state for times, 10 min
apart each time, and an average was taken by three times
measurements. The body mass index (BMI) is a person’s
weight in kilograms divided by the square of height in
meters.

After overnight fasting for 12 h, fasting blood was col-
clected from the veins early in the morning. Measure-
ments of blood glucose, ALT, AST, TB, DB, TP, ALB, Che,
ALP, γ—GT, TG, TC, HDL-C, LDL-C and other serum
indicators were performed on an automated chemistry
analyzer (Hitachi 7600d–210, Japan).

Sample preparation for metabolomics
One hundred microliters of serum was mixed with 400
μL methanol and 5 μL of 2-chloro-DL-phenylalanine
(0.3 g/L, internal standard) for extraction. The superna-
tant of each sample was collected for LC–MS analysis
after centrifugation at 15,000 rpm at 4 °C for 10 min.

UPLC-Orbitrap MS condition
The serum metabolites profiling was performed on
Ultimate 3000 UPLC system (Thermo Fisher Scien-
tific) coupled with an Orbitrap Elite mass spectrometer
(Thermo Fisher Scientific). Samples were eluted through
an ACQUITY UPLC column (HSS T3, 100 × 2.1 mm,
1.8 μm, Waters Corp.) with a 20 min gradient (mobile
phase A was 0.1% formic acid in water and mobile phase
B was acetonitrile) at a flow rate of 300 nL/min. The sol-
vent gradient was as follows: 0–2 min, 95% A; 2–12 min,
95–5% A; 12–15 min, 5% A; 15–20 min, 5–95% A.

The mass spectrometer was equipped with an electro-
spray ionization source and operated in both positive and
negative ion modes. The source parameters were as fol-
lows: heater temperature, 300 °C; sheath gas flow, 45 psi;
 auxiliary gas flow, 5 L/min; tail gas flow, 0.3 L/min; elec-
trospray voltage, 3.0 kV for positive ion mode and 3.2 kV
for negative ion mode; capillary temperature, 350 °C;
S-lens RF level, 30 and 60 for positive and negative ion
mode, respectively.

Data processing, biomarker identification
and metabolomic pathway analysis
The raw LC–MS data were first processed with Com-
pound Discover 2.0 software (Thermo Fisher Scientific).
The Compound Discover software finds components that
have reproducible differences across multiple sample
groups. The resultant data matrix including \( m/z \), RT and intensity was imported into the SIMCA-P 14.0 (Umetrics, Umea, Sweden) software for multivariate statistical analysis. PCA and OPLS-DA analyses were performed, and the variable importance projection (VIP) value was used to screen potential biomarkers. Metabolites of interest (candidate biomarkers) were identified based on their accurate masses and/or MS/MS spectra information in both positive and negative ion mode. HMDB, KEGG and mzCloud databases were searched to assist with metabolite identification. Pathway analysis of the significant altered metabolites was performed with MetaboAnalyst 4.0.

**Statistical analysis**

SPSS 25.0 software (Chicago, United States) was used to perform two-way ANOVAs, receiver operating characteristic (ROC) curve and logistic regression analyses among two groups. The area under the ROC curve (AUC) was used to evaluate the diagnostic power of each potential biomarker in NAFLD. Data were presented as percentages for categorical variables and as means±SD for continuous variables. Differences in categorical and continuous variables between groups were assessed with the \( \chi^2 \) test and the independent samples t-test, while the non-normal distribution was expressed by M (p25-p75), the Wilcoxon nonparametric test was used between the two groups and the non-normal distribution was expressed by M (p25-p75), the Wilcoxon nonparametric test was used between the two groups. In all cases, \( p < 0.05 \) was considered as significant.

**Results**

**Demographic characteristics**

From Table 1, we found that among the 224 subjects, the NAFLD group had significant differences in body weight, BMI, and waist statistics compared with the control group (\( p < 0.001 \)), while there were no obvious differences in age, gender, height, systolic blood pressure, diastolic blood pressure, smoking and drinking (\( p > 0.05 \)).

**Serum index**

From Table 1, we found that there was no remarkable difference in biochemical indexes (including TBA, CHE, DB, TP, ALB, \( \gamma \)-GT, ALP, AST, ALT and GLU) between the NAFLD group and the control group (\( p > 0.05 \)). However, conspicuous differences in all blood lipid indexes (HDL, LDL, TC and TG) between the two groups (\( p < 0.001, p < 0.01, p < 0.01, p < 0.01 \)) were observed.

**PCA of serum samples in NAFLDs**

PCA was performed for both positive and negative ionization modes. Quality control samples were determined for instrument precision assessments and the results confirmed the stability of the UPLC-MS/MS system. As can be seen in Fig. 2, NAFLDs compared to controls showed distinct separations in the PCA score plots, indicating global changes to serum metabolite composition in NAFLD. The cumulative values of \( R^2_X \) and \( Q^2_Y \) in all cases, \( p < 0.05 \) was considered as significant.

**OPLS-DA and metabolites identification in serum from NAFLD patients**

OPLS-DA was employed in NAFLD and control groups to identify potential metabolic biomarkers. The \( R^2_Y \) and \( Q^2_Y \) values were 0.986 and 0.895 in positive ion mode, respectively, and 0.957 and 0.877 in negative ion mode, respectively. The data indicated that the degree of the method’s discrimination and predictability met the analysis requirements. The VIP and \( p \)-values were used to
screen potential metabolic biomarkers. Fifty-five metabolites in serum met the retrieval requirements ($VIP > 1$ and $p < 0.01$) and were identified by using the databases (Table 2).

**Pathway enrichment analysis**
Pathway enrichment was achieved using MetaboAnalyst 4.0 and the results are showed in Fig. 3. The pathways with significant differences ($p < 0.05$) are: (1) phenylalanine metabolism; (2) aminoacyl-tRNA biosynthesis; (3) glycerophospholipid metabolism; (4) ether lipid metabolism; (5) fatty acid biosynthesis; and (6) the tricarboxylic acid cycle (TCA cycle).

**Diagnostic accuracy of the candidate biomarkers for early NAFLD**
The ROC curves of the four lipids (TC, TG, HDL and LDL) were 0.770, 0.728, 0.706 and 0.711, respectively (Fig. 4). A detailed summary of the AUCs, 95% CI lower and upper limit, sensitivities and specificities of the identified serum metabolites are shown in Table 2. The AUCs for 15 metabolites in serum were above 0.9, indicating that they were of high diagnostic value (Fig. 5).

**Discussion**
Nowadays, NAFLD has risen as the most common chronic liver disease in China [22, 23]. Many studies have shown that the prevalence of NAFLD increases in parallel with components of the metabolic syndrome such as obesity, type 2 diabetes mellitus (T2DM), hyperlipidemia and hypertension [24, 25]. The prevalence of NAFLD in industrialized countries is considered to be between 40 and 50%, even higher in patients with T2DM, and the prevalence is up to 90% in morbidly obese patients [26, 27]. The relation between NAFLD and T2DM is considered bidirectional. In addition to the evidence that insulin resistance may contribute to progressive liver disease, NAFLD may also predispose to developing or worsening insulin resistance and metabolic syndrome [28]. Although the natural history of NAFLD is not fully understood, participants with T2DM were excluded in this study.

In addition to the type and frequency of alcohol consumed, it is also unsettled whether moderate alcohol intake plays a role in the development of NAFLD. Many studies suggested that alcohol consumption below safe limits can promote lipid metabolism and reduce insulin resistance, thereby reducing the prevalence of NAFLD, whereas others have reported deterioration of steatohepatitis and fibrosis [29–31]. Nondrinkers and subjects with moderate alcohol consumption that compatible with the diagnosis of NAFLD were included in this study. No clear association between NAFLD and the total alcohol intake per week was observed.

Liver enzymes, especially ALT and AST, are the first laboratory tests every clinician will consider worth evaluating in a patient with liver diseases. In most cases, the higher the ALT and AST levels, the more severe the liver damage [32]. However, all patients included in this study were early NAFLD patients showed normal liver chemistry, thus there were no statistically significant differences in serum concentrations of ALT, AST, ALP, γ-GT, TB, DB, TP, ALB and CHE between the two groups (Additional file 1: Fig. S1).

In the current study, we found that TG, TC, LDL-C and HDL-C were closely related to NAFLD disease. NAFLD patients tend to have high TG, high TC, high LDL-C and low HDL-C levels which in consistence with a previous study reported by Malik and coworkers [33]. However, Abdul et al. [34] found that there was no obvious relationship between NAFLD and LDL-C, TC, but a significant relationship with TG. In addition, Fang [35] used the
| No | Name                                                        | AUC  | 95% CI | Sensitivity | Specificity |
|----|-------------------------------------------------------------|------|--------|-------------|-------------|
|    |                                                             |      |        | Lower limit | Upper limit |
| 1  | LysoPC(20:3(Z2,11Z,14Z))                                    | 0.97 | 0.933  | 1           | 85.19       |
| 2  | Succinic acid                                               | 0.963| 0.917  | 1           | 85.19       |
| 3  | LysoPC(22:5(Z2,10Z,13Z,16Z,19Z))                            | 0.949| 0.896  | 1           | 85.19       |
| 4  | Indole                                                      | 0.947| 0.881  | 1           | 85.19       |
| 5  | LysoPC(22:4(Z2,10Z,13Z,16Z))                                | 0.942| 0.887  | 0.998       | 81.48       |
| 6  | Oleic acid                                                  | 0.938| 0.878  | 0.998       | 85.19       |
| 7  | Desaminotyrosine                                            | 0.926| 0.849  | 1           | 81.48       |
| 8  | L-Phenylalanine                                             | 0.918| 0.838  | 0.997       | 82.59       |
| 9  | L-Tryptophan                                                | 0.915| 0.825  | 1           | 85.19       |
| 10 | LysoPE(22:1(Z2,13Z,16Z)/0:0)                                | 0.915| 0.837  | 0.993       | 85.19       |
| 11 | Leukotriene C5                                              | 0.909| 0.831  | 0.988       | 82.59       |
| 12 | 1-Alkyl-2-acylglycerophosphoethanolamine                     | 0.905| 0.829  | 0.981       | 81.48       |
| 13 | LysoPE(0:0/20:3(5Z,8Z,11Z))                                 | 0.905| 0.83   | 0.981       | 77.78       |
| 14 | L-Lysine                                                    | 0.905| 0.815  | 0.995       | 85.19       |
| 15 | LysoPE(0:0/22:4(Z2,10Z,13Z,16Z))                            | 0.9   | 0.82   | 0.979       | 96.30       |
| 16 | Homovanillic acid                                           | 0.9   | 0.801  | 0.999       | 88.89       |
| 17 | Sulfuric acid                                               | 0.894| 0.81   | 0.979       | 77.78       |
| 18 | LysoPE(16:1(9Z)/0:0)                                        | 0.893| 0.806  | 0.98        | 92.59       |
| 19 | Coumarone                                                   | 0.886| 0.797  | 0.975       | 77.78       |
| 20 | LysoPC(22:6(Z2,7Z,10Z,13Z,16Z))                             | 0.885| 0.796  | 0.974       | 77.78       |
| 21 | LysoPC(20:4(Z2,11Z,14Z,17Z))                                | 0.879| 0.784  | 0.975       | 81.48       |
| 22 | LysoPE(0:0/22:5(Z2,10Z,13Z,16Z))                            | 0.878| 0.78   | 0.976       | 74.07       |
| 23 | Benzoic acid                                                | 0.877| 0.784  | 0.969       | 77.78       |
| 24 | LysoPC(20:2(Z2,11Z,14Z))                                    | 0.87 | 0.768  | 0.972       | 74.07       |
| 25 | LysoPE(0:0/24:6(Z2,9Z,12Z,15Z,18Z,21Z))                     | 0.87 | 0.761  | 0.978       | 85.19       |
| 26 | LysoPC(20:4(Z2,8Z,11Z,14Z))                                 | 0.868| 0.771  | 0.965       | 81.48       |
| 27 | 1-arachidonoyl-sn-glycero-3-phosphoethanolamine             | 0.867| 0.774  | 0.96        | 88.89       |
| 28 | 1-Acyl-sn-glycero-3-phosphoethanolamine                     | 0.863| 0.745  | 0.98        | 74.07       |
| 29 | L-methionine                                                | 0.86 | 0.758  | 0.963       | 85.19       |
| 30 | 4-Hydroxycinnamic acid                                      | 0.855| 0.753  | 0.956       | 70.37       |
| 31 | Arachidonic acid                                            | 0.85 | 0.75   | 0.951       | 85.19       |
| 32 | L-TYROSINE                                                 | 0.846| 0.741  | 0.952       | 70.37       |
| 33 | L-Lactic Acid                                               | 0.845| 0.738  | 0.952       | 92.59       |
| 34 | Palmitic acid                                               | 0.844| 0.72   | 0.967       | 85.19       |
| 35 | LysoPE(20:1(Z2,11Z))/0:0                                   | 0.841| 0.73   | 0.952       | 85.19       |
| 36 | LysoPC(18:2(Z9Z,12Z))                                       | 0.833| 0.724  | 0.941       | 70.37       |
| 37 | LysoPC(18:1(Z9Z))                                           | 0.826| 0.711  | 0.94        | 70.37       |
| 38 | Ethyl acetate                                               | 0.824| 0.709  | 0.94        | 74.07       |
| 39 | 1-(9Z)-Hexadecenoyl-sn-glycero-3-phosphocholine             | 0.822| 0.708  | 0.935       | 96.30       |
| 40 | LysoPC(15:0)                                                | 0.82 | 0.694  | 0.947       | 77.78       |
| 41 | Glycerophosphorylcholine                                    | 0.818| 0.688  | 0.947       | 77.78       |
| 42 | LysoPC(18:3(Z9Z,12Z,15Z))                                   | 0.816| 0.706  | 0.926       | 74.07       |
| 43 | Uric Acid                                                  | 0.815| 0.69   | 0.939       | 66.67       |
| 44 | 2-Acyl-sn-glycero-3-phosphoethanolamine                     | 0.811| 0.694  | 0.928       | 81.48       |
| 45 | Stearic acid                                                | 0.793| 0.664  | 0.922       | 70.37       |
| 46 | LysoPE(0:0/20:0)                                            | 0.785| 0.65   | 0.92        | 77.78       |
| 47 | L-Palmitoylcarnitine                                        | 0.761| 0.628  | 0.895       | 70.37       |
| 48 | LysoPE(0:0/16:0)                                            | 0.76 | 0.626  | 0.894       | 77.78       |
TG / HDL-C ratio as a predictor of NAFLD. From this, it can be found that whether lipid level can be used as a predictor of the progression of NAFLD is still controversial. It has been reported that liver lipotoxicity of free fatty acids, cholesterol, ceramide and lysophosphatidylcholine is the main reason for the progression of NAFLD. The simple accumulation of triglycerides may not lead to NAFLD [36–38], but rather, the type of accumulated lipids may determine the severity and development trend of NAFLD. The results of metabolomic analysis showed that the levels of lysoPC, lysoPE, phenylalanine, oleic acid and tryptophan were obviously increased in NAFLD patients with hepatitis [39–42]. These findings suggested that these serum metabolites play important roles in the development of NAFLD, as well as their considerable clinical value. An issue that was not addressed in the present study was the lead-time bias. Since there is a period of time before the detection of NAFLD in which NAFLD has developed with no clinical manifestations, the effectiveness of these biomarkers might be changed whenever estimated during this period.

An overview of the altered pathways is shown in Fig. 3. The contents of all metabolites in phenylalanine metabolism pathway were increased. Phenylalanine and its related metabolites are mainly metabolized in the liver [43–45]. Some studies have shown that the increased levels of phenylalanine are highly correlated with obesity and liver steatosis [46–48]. Other studies have found that phenylalanine levels in T2DM patients are significantly increased, especially after a normal diet [49]. In addition, through logistic regression analysis from 72 high and 75 low insulin sensitivity subjects, Palmer et al. [50] observed significantly decreased glycine and increased valine, leucine, phenylalanine, and combined glutamine and glutamate in insulin-resistant subjects. Therefore, deterioration of liver function in NAFLD patients may cause the decline in phenylalanine metabolism, and ultimately

Table 2 (continued)

| No | Name                               | AUC  | 95% CI     | Sensitivity | Specificity |
|----|------------------------------------|------|------------|-------------|-------------|
|    |                                    |      | Lower limit| Upper limit |             |
| 49 | 2-linoleoyl-sn-glycero-3-phosphoethanolamine | 0.749 | 0.618 0.88 | 66.67 77.78 |
| 50 | Platelet-activating factor          | 0.708 | 0.563 0.852 | 51.85 96.30 |
| 51 | L-Carnitine                        | 0.705 | 0.568 0.842 | 59.26 74.07 |
| 52 | 1-heptadecanoyl-sn-glycero-3-phosphocholine | 0.672 | 0.52 0.824 | 55.56 85.19 |
| 53 | L-Valine                           | 0.236 | 0.102 0.37 | 14.81 33.33 |
| 54 | L-Pyroglutamic acid                | 0.091 | 0.007 0.174 | 22.22 7.41  |
| 55 | Citric acid                        | 0.056 | 0 0.131 | 3.70 7.41  |

Fig. 3 Pathway analysis of significant altered metabolites

Fig. 4 ROC curve of clinical indicators
lead to the accumulation of phenylalanine and its related metabolites in liver and serum.

**Conclusion**

By high-throughput LC–MS-based metabolomics analysis, a total of 55 metabolites significantly associate with early-stage NAFLD were identified. Of these, 15 potential biomarkers showed high diagnosis value with AUC greater than 0.9. Moreover, our results provide comprehensive insights into the metabolic pathways involved in NAFLD. These combined serum metabolites could be the basis of a clinically feasible method of non-invasive NAFLD screening, as well as to follow response to therapeutic interventions.

**Abbreviations**

NAFLD: Nonalcoholic Fatty Liver Disease; TBA: Total bile acid; CHE: Cholinesterase; TB: Total Bilirubin; DB: Direct Bilirubin; TP: Total Protein; ALB: Albumin; γ-GT: γ-Glutamyltransferase; ALP: Alkaline Phosphatase; AST: Aspartate Aminotransferase; ALT: Alanine Aminotransferase; HDL: High Density Lipoprotein; LDL: Low Density Lipoprotein; TC: Cholesterol; TG: Triglyceride; GLU: Glucose in Urine; PCA: Principal Component Analysis; OPLS-DA: Orthogonal Partial Least Squares Discriminant Analysis; VIP: Variable Importance in Projection; LC–MS: Liquid Chromatography Mass Spectrometry; AUC: Area under Curve; ROC: Receiver Operating Curve.

**Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1186/s12967-021-02820-7.

**Acknowledgements**

This work was financially supported by National Natural Science Foundation of China (81703879 and 21904022); Shanghai Municipal Commission of Health and Family Planning general project for clinical research of health industry (201840377, 201940449); Putuo District of Shanghai Science and Technology Commission Research Project (ptkwss201813). Key specialties of Putuo Hospital Affiliated to Shanghai University of Traditional Chinese Medicine (2016103A).

**Authors’ contributions**

Study concept and design: YJ and XZ. Acquisition of data: CH, TW, QS and XW. Analysis and interpretation of data: HL and XP. Drafting of the manuscript: ML and JZ. Critical revision of the manuscript for important intellectual content: all the authors. Statistical analysis: QC.

**Funding**

This work was financially supported by National Natural Science Foundation of China (81703879 and 21904022); Shanghai Municipal Commission of Health and Family Planning general project for clinical research of health industry (201840377, 201940449); Putuo District of Shanghai Science and Technology Commission Research Project (ptkwss201813).

**Availability of data and materials**

All data are included in this article.

**Declarations**

**Ethics approval and consent to participate**

The experimental protocol was approved by the Human Ethics Committee of Putuo Hospital. Written informed consent was obtained from individual or guardian participants.

**Consent for publication**

Consent for publication has been obtained from all authors.

**Competing interests**

The authors declare that they have no conflict of interest.

**Fig. 5 ROC curve of the new biomarkers**
References

1. Zhou J, Zhou F, Wang W, et al. Epidemiological Features of NAFLD From 1999 to 2018 in China. Hepatology. 2020;71(5):1851–64.

2. Liu CJ. Prevalence and risk factors for non-alcoholic fatty liver disease in Asian people who are not obese. J Gastroenterol Hepatol. 2012;27(10):1555–60.

3. Hu W, Shao X, Guo D, et al. Relationship of serum betatrophin with nonalcoholic fatty liver in a Chinese population. PLoS ONE. 2017;12(1):e0170758.

4. Khaw KB, Choi RH, Kam JH, et al. Interval increase in the prevalence of symptomatic cholelithiasis-associated non-alcoholic fatty liver disease over a ten-year period in an Asian population. Singap Med J. 2017;58(2):703–7.

5. Stringhini S, Carmeli C, Jokela M, et al. Socioeconomic status and the 25 x 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. Lancet. 2017;389(10073):1299–307.

6. Fan JG, Kim SU, Wong VW, et al. New trends on obesity and NAFLD in Asia. J Hepatol. 2017;67(4):862–73.

7. Lu FB, Hu ED, Xu LM, et al. The relationship between obesity and the severity of non-alcoholic fatty liver disease: systematic review and meta-analysis. Exp Rev Gastroenterol Hepatol. 2018;12(5):491–502.

8. Vernon G, Baranova A, YounoszI ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther. 2011;34(3):274–85.

9. White DL, Kanwali F, El-Serag HB. Association between nonalcoholic fatty liver disease and risk for hepatocellular cancer, based on systematic review. Clin Gastroenterol Hepatol. 2012;10(12):1342-1359 e2.

10. Mittal S, El-Serag HB, Sada YH, et al. Hepatocellular carcinoma in the absence of cirrhosis in United States veterans is associated with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol. 2016;14(1):124-131 e1.

11. Castella L, Pinzani M. Non-invasive assessment of liver fibrosis: are we ready? Lancet. 2010;375(9724):1419–20.

12. Alkhouri N, Fieldstein AE. Noninvasive diagnosis of nonalcoholic fatty liver disease: are we there yet? Metab Clin Exp. 2016;65(8):1087–95.

13. Wong VW, Adams LA, De Ledshingh V, et al. Noninvasive biomarkers in NAFLD and NASH—current progress and future promise. Nat Rev Gastroenterol Hepatol. 2018;15(8):461–78.

14. Fitzpatrick E, Dhawan A. Noninvasive biomarkers in non-alcoholic fatty liver disease: current status and a glimpse of the future. World J Gastroenterol. 2014;20(3):1085–63.

15. Soininen P, Kangas AJ, Wurtz P, et al. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. Circ Cardiovasc Genet. 2015;8(1):192–206.

16. Tzoulaki I, Ebels TM, Valdes A, Elliott P, Ioannidis JP. Design and analysis of metabolomics studies in epidemiologic research: a primer on omic technologies. Am J Epidemiol. 2014;180(2):129–39.

17. Zhao L, Zhang J, Pan L, et al. Protective effect of 7,3′,4′-Flavon-3-ol (fisetin) on acetaminophen-induced hepatotoxicity in vitro and in vivo. Phytother Res J Phytother Phytopharmacol. 2019;58:152865.

18. Pirola CJ, Sookias CA. Multimarkers for the prediction of nonalcoholic fatty liver disease severity. World J Gastroenterol. 2018;24(15):1601–15.

19. Procopet B, Fischer P, Farcau O, et al. Metabolomics: From liver chiro-mancy to personalized precision medicine in advanced chronic liver disease. World J Hepatol. 2018;10(3):371–8.

20. National Workshop on Fatty Liver and Alcoholic Liver Disease, Chinese Society of Hepatology, Chinese Medical Association, et al. Guideline of prevention and treatment for nonalcoholic fatty liver disease: a 2018 update. J Mod Hepatol Med Health. 2018;34(5):644–645.

21. Xu X, Lee J, Chen S, et al. Light-to-moderate alcohol consumption is associated with increased risk of type 2 diabetes in individuals with non-alcoholic fatty liver disease: a nine-year cohort study. Am J Gastroenterol. 2020;115(6):876–84.

22. Fan JG, Farrell GC. Epidemiology of non-alcoholic fatty liver disease in China. J Hepatol. 2009;50(1):204–10.

23. Loomba R, Sanyal AJ. The global NAFLD epidemic. Nat Rev Gastroenterol Hepatol. 2013;10(11):686–90.

24. Leite NC, Salles GF, Araujo AL, et al. Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus. Liver Int. 2009;29(1):113–9.

25. Almobarak AD, Barakat S, Khalifa MH, et al. Nonalcoholic fatty liver disease (NAFLD) in a Sudanese population: what is the prevalence and risk factors? Arab J Gastroenterol. 2014;15(1):12–5.

26. Chavez-Tapia NC, Tellez-Avilia FI, Barrientos-Gutierrez T, et al. Bariatric surgery for non-alcoholic steatohepatitis in obese patients. Cochrane Database Syst Rev. 2010;1(1):CD007340.

27. Musso G, Gambino R, Cassader M. Obesity, diabetes, and gut microbiota: the hygiene hypothesis expanded. Diabetes Care. 2010;33(10):2277–84.

28. Pillai AA, Rinella ME. Non-alcoholic fatty liver disease is a bariatric surgery answer? Clin Liver Dis. 2009;13(4):689–710.

29. Mitchell T, Jeffrey GP, de Boer B, et al. Type and pattern of alcohol consumption is associated with liver fibrosis in patients with non-alcoholic fatty liver disease. Am J Gastroenterol. 2018;113(10):1484–93.

30. Chang Y, Cho YK, Kim Y, et al. Nonheavy drinking and worsening of noninvasive fibrosis markers in nonalcoholic fatty liver disease: a cohort study. Hepatology. 2019;69(1):64–75.

31. Hajifathalian K, Torabi Sagivand B, McCullough AJ. Effect of alcohol consumption on survival in nonalcoholic fatty liver disease: a national prospective cohort study. Hepatology. 2019;70(2):511–21.

32. Nguyen LH, Chao D, Lim JK, et al. Histologic changes in liver tissue from patients with chronic hepatitis B and minimal increases in levels of alanine aminotransferase: a meta-analysis and systematic review. Clin Gastroenterol Hepatol. 2014;12(8):1262–6.

33. Malik A, Cheah PL, Hillim IN, et al. Non-alcoholic fatty liver disease in Malaysia: a demographic, anthropometric, metabolic and histological study. J Dig Dis. 2007;8(1):58–64.

34. Chelangui I, Morais LF, Ribeiro V, et al. The role of exercise on nonalcoholic fatty liver disease: a systematic review of randomized controlled trials. Int J Mol Sci. 2019;20(5):1220–34.

35. Fan X, Liu EY, Hoffman VP, et al. Triglyceride/high-density lipoprotein cholesterol ratio: a surrogate to predict insulin resistance and low-density lipoprotein cholesterol particle size in non-diabetic patients with schizophrenia. J Clin Psychol. 2010;72(6):806–12.

36. Alkhouri N, Dixon LJ, Fieldstein AE. Lipotoxicity in nonalcoholic fatty liver disease: not all lipids are created equal. Exp Rev Gastroenterol Hepatol. 2009;3(4):445–51.

37. Cusi K. Role of obesity and lipotoxicity in the development of nonalcoholic steatohepatitis: pathophysiology and clinical implications. Gastroenterology. 2012;142(4):711–725 e6.

38. Gentile CL, Pagliassotti MJ. The role of fatty acids in the development and progression of nonalcoholic fatty liver disease. J Nutr Biochem. 2008;19(9):567–76.

39. Xu Y, Han J, Dong J, et al. Metabolomics characterizes the effects and mechanisms of quercetin in nonalcoholic fatty liver disease development. Int J Mol Sci. 2019;20(5):1220–34.

40. Li J, Zhao Y, Huang C, et al. Serum metabolomics analysis of the effect of exercise on nonalcoholic fatty liver disease. Endocr Connect. 2019;8(4):299–308.

41. Alkhouri N, Dixon LJ, Fieldstein AE. Lipotoxicity in nonalcoholic fatty liver disease: what is the prevalence and risk factors? Hepatology. 2010;51(1):227–34.

42. Carvalho MB, Costa EP, Sabino MT, et al. The role of exercise on non-alcoholic fatty liver disease: preliminary results. Rev Bras Med Esportiva. 2017;23(3):679–82.

43. Sun X, Li Y, Chen L, et al. The role of physical activity on nonalcoholic fatty liver disease: a systematic review. Tex Rep Med Sci. 2017;27(1):6120.

44. Li J, Zhao Y, Huang C, et al. Serum metabolomics analysis of the effect of exercise on nonalcoholic fatty liver disease. Endocr Connect. 2019;8(4):299–308.

45. Tu LN, Shoveller MR, Cajka T, et al. Metabolomic characteristics of cholesteryl ester-enriched non-obese nonalcoholic fatty liver disease in mice. Sci Rep. 2017;7(1):6120.

46. Zhu H, Wang Z, Wu Y, et al. Untargeted metabolomics reveals intervention effects of chiorocy polyacetylene in a rat model of non-alcoholic fatty liver disease. Int J Biol Macromol. 2019;128:363–75.
43. Zello GA, Pencharz PB, Ball RO. Phenylalanine flux, oxidation, and conversion to tyrosine in humans studied with L-[1–13C]phenylalanine. Am J Physiol. 1990;259(6 Pt 1):E835–43.
44. Burke PA, Stack JA, Wagner D, et al. L-[1–13C] Phenylalanine oxidation as a measure of hepatocyte functional capacity in end-stage liver disease. Am J Surg. 1997;173(4):270–3 (discussion 3–4).
45. Ishii Y, Suzuki S, Kohno T, et al. Patients with severe liver cirrhosis followed up by L-[1–13C] phenylalanine breath test. J Gastroenterol. 2003;38(11):1086–90.
46. Chang Y, Ryu S, Sung E, et al. Higher concentrations of alanine aminotransferase within the reference interval predict nonalcoholic fatty liver disease. Clin Chem. 2007;53(4):686–92.
47. Felig P, Mark E, Cahill GF. Plasma amino acid levels and insulin secretion in obesity. N Engl J Med. 1969;281(15):811–6.
48. Moretto M, Kupski C, Mottin CC, et al. Hepatic steatosis in patients undergoing bariatric surgery and its relationship to body mass index and co-morbidities. Obesity Surg. 2003;13(4):622–4.

49. Mook-Kanamori DO, de Mutsert R, Rensen PC, et al. Type 2 diabetes is associated with postprandial amino acid measures. Arch Biochem Biophys. 2016;589:138–44.
50. Palmer NC, Stevens RD, Antinozzi PA, et al. Metabolomic profile associated with insulin resistance and conversion to diabetes in the Insulin Resistance Atherosclerosis Study. J Clin Endocrinol Metab. 2015;100(3):E463–68.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.