의학박사 학위논문

Plasma metabolomic signatures of nonalcoholic fatty liver disease and weight loss interventions in pediatric population with obesity

소아청소년 비만 환자에서 비알코올성 지방간질환 및 체중감량 중재 연관 혈장 대사체의 특성 연구

2023년 2월

서울대학교 대학원

의과학과 의과학전공

채 우 리

Ph.D. Dissertation of Medical Science

소아청소년 비만 환자에서 비알코올성 지방간질환 및 체중감량 중재 연관 혈장 대사체의 특성 연구

Plasma metabolomic signatures of nonalcoholic fatty liver disease and weight loss interventions in pediatric population with obesity

February 2023

Seoul National University Graduate School
Department of Biomedical Sciences
Major in Biomedical Sciences

Woori Chae

소아청소년 비만 환자에서 비알코올성 지방간질환 및 체중감량 중재 연관 혈장 대사체의 특성 연구

지도교수 조 주 연

이 논문을 의학박사 학위논문으로 제출함

2022년 10월

서울대학교 대학원

의과학과 의과학전공

채 우 리

채우리의 의학박사 학위논문을 인준함

2023년 1월

위 원 장 _____ (인)

부 위 원 장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

Plasma metabolomic signatures of nonalcoholic fatty liver disease and weight loss interventions in pediatric population with obesity

Joo-Youn Cho

Submitting a Ph.D. Dissertation of Medical Science

October 2022

Seoul National University Graduate School
Department of Biomedical Sciences
Major in Biomedical Sciences

Woori Chae

Confirming the Ph.D. Dissertation written by Woori Chae

January 2023

Chair	(Seal)
Vice Chair	(Seal)
Examiner	(Seal)
Examiner	(Seal)
Examiner	(Seal)

Abstract

Plasma metabolomic signatures of nonalcoholic fatty liver disease and weight loss interventions in pediatric population with obesity

Woori Chae

Major in Biomedical Sciences

Department of Biomedical Sciences

Seoul National University Graduate School

INTRODUCTION: The prevalence of pediatric obesity in Korea is increasing rapidly from 8.4% in 2008 to 25% in 2018, as it is worldwide. Pediatric obesity is closely related to various metabolic diseases, such as insulin resistance and diabetes, cardiovascular diseases, and hepatic diseases. Nonalcoholic fatty liver disease (NAFLD) is one of the most common liver diseases in the pediatric population. The most effective prevention and treatment methods for obesity and related diseases in children are weight loss interventions, including regular physical activity and dietary interventions. In this study, metabolic

differences according to the occurrence of NAFLD and weight loss interventions were investigated in plasma samples using a metabolomics approach by leveraging the pediatric NAFLD and the obesity intervention cohorts.

METHODS: 165 children and adolescents from a pediatric NAFLD cohort were classified into four groups based on steatosis grade and body mass index z-score to discover the NAFLD-specific metabolic biomarkers in plasma. In addition, plasma samples of selected 40 children and adolescents in the obesity intervention cohort were compared at baseline, 6 months post-intervention, and 18 months post-intervention to explore significant changes in plasma metabolites according to the weight loss intervention.

RESULTS: I discovered 18 NAFLD-specific metabolic biomarkers by investigating the pediatric NAFLD cohort. These metabolites were related to glutathione-related metabolism, lipid metabolism, and branched-chain amino acid metabolism. This study also demonstrated that the metabolites can be used as ancillary biomarkers for the diagnosis of NAFLD in the pediatric population. In the obesity intervention cohort, a time-series change in plasma metabolites was observed according to the intervention period regardless of the intervention response. More metabolites were changed as the intervention period increased, of which changes in the

TCA cycle, urea cycle, and amino acid metabolism were particularly prominent.

CONCLUSION: By using two pediatric cohorts and metabolomics methodology, I suggested the clinical implications of metabolic changes by NAFLD and a weight loss intervention in the pediatric population with obesity. This study proposes the pathophysiology of NAFLD and conceptualizes metabolically healthy obesity. The results help establish appropriate treatment for childhood and adolescent obesity and related diseases.

* Part of this work has been published as two original articles in *Metabolites* (Chae W, Lee KJ, et al. Association of Metabolic Signatures with Nonalcoholic Fatty Liver Disease in Pediatric Population. Metabolites. 2022;12(9); Sohn MJ, Chae W, et al. Metabolomic Signatures for the Effects of Weight Loss Interventions on Severe Obesity in Children and Adolescents. Metabolites. 2022;12(1)).

KEYWORDS: pediatric obesity, metabolomics, nonalcoholic fatty liver disease, weight loss intervention, plasma metabolome

STUDENT NUMBER: 2017-39919

Table of Contents

List of Figures

List of Tables

Introduction

Child and adolescent obesity rates have rapidly escalated worldwide (1). The World Health Organization reported that in just 40 years, the pediatric population with obesity has risen more than 10-fold. This translates to a surge from 11 million to 124 million (2016 estimates) (2). Within one decade, from 2008 to 2018, Korean childhood obesity rates burgeoned from 8.4% to 25% (3). In addition, COIVD-19 pandemic resulted in heightened stress, weight gain, and fewer physical activity opportunities. As a consequence, childhood obesity risk has multiplied, nearly doubling the average body mass index (BMI) between January and November 2020 (4). Childhood obesity prevention and early treatment methods are vital as as obesity correlates with various medical complications, including endocrine, cardiovascular, musculoskeletal, and gastrointestinal diseases and metabolic syndromes, such as dyslipidemia, hypertension, and insulin resistance (5).

Since first reporting pediatric nonalcoholic fatty liver disease (NAFLD) in 1983, it has become one of the most prevalent hepatic disorders in children and adolescents with obesity (6). The prevalence of fatty liver is also increasing in proportion to the increase in childhood and adolescent obesity (7). The risk factors for NAFLD include dietary, environmental, and genetic factors, which show

complex interactions resulting in insulin resistance and obesity (8). These factors cause hepatic triglyceride (TG) accumulation, lipotoxicity due to the high levels of free fatty acids, and oxidative stress, which are involved in hepatic inflammation. Despite the increasing prevalence of NAFLD worldwide over the last few decades, no licensed drugs have been approved for its treatment. Developing a single effective drug may be complicated by the complex pathophysiology of NAFLD.

Current first-line treatments for controlling obesity and related diseases are physical activity, dietary regulation and other weight-management lifestyle modifications (9). However, unmet medical needs remain due to complex pathophysiology, inadequate medications, and insufficient obesity and NAFLD studies of the pediatric population. Thus, thoroughly understanding pathophysiological characteristics and identifying efficient pediatric obesity and obesity-related surrogate biomarkers, particularly for NAFLD, is necessary to establish appropriate treatment strategies.

Metabolomics is a technical tool that aims to detect and measure minute changes in molecular (< 1,500 Da) cell, tissue, organ, or whole organism levels due to genetic variation or physiological or pathological conditions (10). Metabolomics also assesses metabolic changes from obesity and obesity-related diseases at cellular and body fluid levels (11). Plasma metabolomics is advantageous in determining disease-related biomarkers because it is easy to collect iteratively, less

invasive, inexpensive, and reflective of systemic changes. Some research has reported metabolic signatures based on elements such as branched-chain amino acids (BCAAs), aromatic amino acids, and other lipidomic profiles associated with NAFLD. However, most of these studies were conducted on adults and individuals with obesity (12-14). In addition, increasing studies have revealed the metabolomic signatures and pathophysiological changes associated with obesity such as inflammation or oxidative stress (11). However, few studies have investigated childhood obesity, especially involving weight loss interventions (15-19). In this regard, a metabolomics approach may address many unmet medical needs.

In this study, I aimed to identify NAFLD-specific or obesity intervention-related metabolic biomarkers and associated pathways to elucidate metabolic changes induced by obesity-related diseases and interventions. The study involved leveraging two study cohorts, the pediatric NAFLD cohort and the obesity intervention cohort, and investigating the plasma metabolome signatures in pediatric patients with obesity and NAFLD.

Methods[A]

1. Study population

1.1. Pediatric NAFLD cohort[B]

This study was approved by the Institutional Review Board (IRB) of each hospital (IRB No. 2018-10-015 by the Hallym University Sacred Heart Hospital and 1811-149-98 by Seoul National University Children's Hospital) and conducted per the Declaration of Helsinki. Children and adolescent participants who visited the pediatric departments of the Hallym University Sacred Heart Hospital and Seoul National University Children's Hospital from January 2019 to May 2020 were recruited, after obtaining informed consent from the subjects and their parents.

The presence and grade of the participants' fatty livers were evaluated by ultrasonography. The steatosis grade was assessed by comparing hepatic echogenicity to kidney parenchyma and graded as normal, 0; mild, 1; moderate, 2; and severe, 3 (20, 21). The participants were categorized into four groups according to the steatosis grade as

---

[A] Study design and analysis workflows of the pediatric cohort and the obesity intervention cohort are abridged in **Figures 1 and 2**.

[B] Patient recruitment, sample collection, and anthropometric and laboratory assessment were conducted by Prof. Kyung Jae Lee and Prof. Jae Sung Ko.

determined by abdominal ultrasonography and BMI z-score based on the 2017 Korean National Growth Chart for children and adolescents (22): healthy control (HC), steatosis grade = 0 and BMI z-score $\leq$ 1; lean NAFLD (LN), steatosis grade $\geq$ 1 and BMI z-score $\leq$ 1; overweight control (OC), steatosis grade = 0 and BMI z-score > 1; and overweight NAFLD (ON), steatosis grade $\geq$ 1 and BMI z-score > 1. Participants who were taking alcohol or medications known to affect the results of liver function test results. Participants with viral hepatitis, such as hepatitis A, B, or C, or with Epstein–Barr virus, Wilson's disease, autoimmune hepatitis, or muscular disease were also excluded. Thus, 165 subjects were included in this study.

Anthropometric characteristics such as height, weight, and BMI z-score were evaluated. After overnight fasting, 4 mL peripheral blood samples were collected to assess the insulin, hemoglobin A1c (HbA1c), and platelet count levels. Serum samples were also collected by clotting blood for 30 min, followed by centrifugation at 4 °C. Laboratory assessments, including the measurement of the serum levels for fasting glucose, TGs, and cholesterol, and liver function tests, including serum aspartate transaminase (AST), alanine transaminase (ALT), gamma-glutamyl transferase (GGT), and alkaline phosphatase (ALP) activity measurements were performed. The homeostatic model assessment for insulin resistance (HOMA-IR) was calculated as fasting glucose (mg/dL) multiplied by fasting insulin (mU/L) divided by 405. For clinical variables with missing data, such as GGT, fasting glucose,

insulin, HbA1c, and HOMA-IR, I chose appropriate statistical methods according to the number of data points for each variable following the exclusion of missing values.

1.2. Obesity intervention cohort[C]

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Hallym University Sacred Heart Hospital (IRB No.: 2015-1134 and 2016-1135) and Seoul National University Hospital (IRB No.: 1907-195-1054 and 1102-099-357). Subjects with obesity (BMI > 97th percentile for age and sex) were selected from 242 patients in the Intervention for Children and Adolescent Obesity via Activity and Nutrition (ICAAN) cohort, a 24-month post-intervention follow-up study, which was designed as a multidisciplinary intervention test to prevent excessive weight gain and improve several health indices in children and adolescents (aged 6–17 years) with obesity in Korea (23-25).

Participants without obesity-related heredity or other underlying diseases were randomly divided into three groups. They received a 24-month intervention varying by group: usual care, extensive exercise, or nutrition feedback. Each group was similar in size (usual care group

[C] Patient recruitment, intervention, sample collection, and anthropometric and laboratory assessment were conducted by Prof. Kyung-Hee Park.

= 10, 25%; extensive exercise group = 15, 37.5%; nutrition feedback group = 15, 37.5%), and all groups experienced five categories of intervention: nutrition, physical activity, group activity, parental education, and self-monitoring. The exercise group incorporated the usual care group components with weekly exercise classes and activity feedback. The nutrition feedback group received individual nutrition feedback and the usual care group components.

Their intervention response was evaluated based on changes in BMI z-scores at 18 months post-intervention, compared to those at baseline (ΔZ_{BMI}). Responders were defined by $\Delta Z_{BMI} < -0.45$ and non-responders were defined by $\Delta Z_{BMI} > -0.1$. As no patients gained weight because of the intervention, patients with the least weight change were selected as the non-responder group. Random sampling was not possible owing to the limited sample size. Therefore, the portion and number of samples in each intervention group were considered.

Blood samples were collected at baseline, 6 months, and 18 months post-intervention in both groups, resulting in a total of 120 collected samples for anthropometric data such as age, sex, and BMI z-score, and laboratory assessment data such as AST, ALT, TGs, high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) levels.

2. Mass spectrometry-based analyses in plasma

2.1. Pediatric NAFLD cohort

4 mL of blood was collected from each participant after overnight fasting and centrifuged at 4 °C for metabolomic analysis. Separated plasma samples were collected and stored until use at −80 °C. Plasma metabolites, including 21 amino acids, 21 biogenic amines, 55 acylcarnitines (ACs), 18 diglycerides (DGs), 42 TGs, 172 phosphatidylcholines (PCs), 24 lysophosphatidylcholines (LPCs), 31 sphingomyelins (SMs), 9 ceramides (Cers), and 14 cholesteryl esters (CEs) and hexoses were analyzed using the AbsoluteIDQ™ p400 HR kit (Biocrates Life Sciences AG, Innsbruck, Austria). Samples and reagents were prepared according to the manufacturer's instructions. Additionally, three pooled plasma samples per plate were added for quality control and prepared in the same way as the analytical samples to normalize the batch-to-batch effect.

Briefly, frozen plasma samples were thawed on ice and vortexed, followed by centrifugation at 2750× g, 4 ℃ for 5 min before the samples were loaded onto a 96-well plate with a filter. After 10 μL of analytical and pooled plasma samples and calibration standards were added to each well, the plates were dried with a nitrogen evaporator and derivatized with phenyl isothiocyanate. Then, dried samples were extracted with an ammonium acetate solution in methanol and

aliquoted into two deep-well plates for liquid chromatography mode and flow injection analysis (FIA) mode (described in the manual), followed by dilution with water and an FIA solvent, respectively. Both deep-well plates were placed in an autosampler of Ultimate 3000 ultra-performance liquid chromatography coupled with a Q Exactive Plus hybrid quadrupole-orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) and analyzed using the validated method.

2.2. Obesity intervention cohort

The 120 plasma samples were transported to Human Metabolome Technologies Inc. (HMT) to measure the metabolome. The samples were stored in a deep freezer below −80℃ until analysis. 200 μL of methanol containing internal standards (L-methionine sulfone and D-camphor-10-sulfonic acid, 20 μM) was added to each 50 μL of sample, diluted with 150 μL of distilled water, and mixed thoroughly. Each mixture (300 μL) was filtered through a 5-kDa cut-off filter (Ultrafree-MC-PHCC, HMT, Yamagata, Japan) to remove macromolecules.

The filtrate was centrifugally concentrated and resuspended in 50 μL of distilled water immediately before analysis. The compounds were measured in the cation and anion modes using a capillary electrophoresis-time-of-flight mass spectrometer system (Agilent Technologies Inc., Santa Clara, CA, USA) and a fused silica capillary

(50 μm i.d. × 80 cm total length) with the mass range from m/z 50-1000. Cationic metabolites were analyzed with Cation Buffer Solution (Human Metabolome Technologies, solution ID: H3301-1001) as the electrolyte. The sample was injected at a pressure of 50 mbar for 5 sec with an applied voltage set at 30 kV. The capillary voltage of the electrospray ionization source was set at 4000 V. Anionic metabolites were analyzed with Anion Buffer Solution (Human Metabolome Technologies, solution ID: I3302-1023) as the electrolyte. The sample was injected at a pressure of 50 mbar for 10 sec with the applied voltage set at 30 kV. The capillary voltage was set at 3500 V. The relative standard deviation of internal standards added to the sample was monitored (less than 10%) to guarantee analytical reproducibility.

3. Raw data process and metabolite quantitation

3.1. Pediatric NAFLD cohort

Raw data were processed using the Xcalibur Software (Thermo Fisher Scientific, Waltham, MA, USA) and MetIDQ Oxygen (Biocrates Life Sciences AG, Innsbruck, Austria) to calculate the metabolite concentrations in each sample per the manufacturer's instruction. The final quantitative results were exported micromolar values with pooled quality control normalization by the median. Subsequently, values under the lower limit of detection were imputed by one-fifth of the minimum positive values of their corresponding variables. Finally, 342 metabolites were reliably detected and selected for further analysis.

3.2. Obesity intervention cohort

Detected peaks were extracted using an automatic integration software (MasterHands ver. 2.17.4.19, Keio University; Tokyo, Japan) to obtain peak information, including m/z values, migration time, and peak area. The peak area was then converted and normalized to a relative peak area by dividing it by the sample amount and internal standard peak area. The peak detection limit was determined based on a signal-to-noise ratio of 3. Relative peak areas under the peak detection limit were imputed by the K-nearest neighbor method.

Putative metabolites were then assigned from HMT's standard library and the Known-Unknown peak library based on m/z and migration time with tolerances of 10 ppm and 0.5 min, respectively. A candidate was assigned a branch number if several peaks were assigned to it.

4. Statistical analyses

4.1. Pediatric NAFLD cohort

A Kruskal–Wallis test, followed by a post-hoc Dunn's multiple comparisons test with Prism 7 (GraphPad Software, San Diego, CA, USA) was used to compare the anthropometric and laboratory data between study groups. MetaboAnalyst 5.0 (26) was used to perform principal component analysis with Pareto-scaled metabolome data to visualize the metabolite profile distribution of each group. Wilcoxon rank-sum test was used to determine significant plasma metabolites between HC and LN, LN and ON, HC and OC, or OC and ON. Significance was defined as a false discovery rate (FDR)-adjusted p-value < 0.05 and fold change > 1.1. The obesity markers that were statistically significant between the normal weight (defined as BMI z-score ≤ 1) and overweight (defined as BMI z-score > 1) groups or the NAFLD markers that were statistically significant between the control (defined as steatosis grade = 0) and NAFLD (defined as steatosis grade ≥ 1) groups were determined by student t-test. Significance was defined as FDR-adjusted p-value < 0.05 and fold change > 1.2. The markers were clustered by hierarchical cluster analysis with the Ward algorithm and visualized by heatmaps with MetaboAnalyst 5.0. The distance of markers was calculated by Euclidean distance method.

NAFLD-specific metabolic biomarkers (significant metabolites between OC and ON) were illustrated by a volcano plot. Significance was defined as FDR-adjusted *p*-value < 0.05 and fold change > 1.1. Concentrations of significant metabolites between OC and ON were standardized by autoscaling the features, followed by hierarchical metabolite clustering using the Ward method and Euclidean distance. The Kruskal–Wallis test was used to calculate the significance of NAFLD-specific metabolic biomarkers in the four groups, followed by the two-stage step-up method proposed by Benjamini, Krieger, and Yekutieli to correct multiple comparisons by controlling FDR (27), in which the number of multiple comparisons per metabolites was four (HC vs. OC, HC vs. LN, OC vs. ON, and LN vs. ON). Multiple linear regression analysis (metabolites ~ BMI z-score + group (OC or ON)) was used to determine the significance (raw *p*-value) of NAFLD-specific metabolic biomarkers after BMI adjustment. Raw *p*-values were adjusted by controlling FDR using the Benjamini–Hochberg method. Spearman correlation analyses were used to determine correlation coefficients and two-tailed *p*-values of the NAFLD-specific metabolic biomarkers and HOMA-IR.

Enriched metabolite sets between the OC and ON groups were identified by querying significant metabolites in an SMPDB-based database provided by MetaboAnalyst 5.0. I excluded significant metabolites without a Human Metabolite Database (HMDB) ID or PubChem CID in this analysis. Chemical and biochemical

relationships of significant metabolites were mapped onto MetaMapp (28) and visualized with Cytoscape 3.8.2 (29).

4.2. Obesity intervention cohort

Longitudinal changes (from baseline to 6 months to 18 months post-intervention) in participants' clinical characteristics were analyzed through One-Way Repeated Measures ANOVA with the Greenhouse-Geisser correction (p-value < 0.05), followed by FDR-adjusted multiple comparisons proposed by Benjamini, Krieger, and Yekutieli (FDR-adjusted p-value < 0.05). Two-Way Repeated Measures ANOVA with Sidak's multiple comparisons test were used to determine a significant interaction effect between clinical characteristic changes (from baseline to 18 months post-intervention) and intervention responsiveness (p-value < 0.05, 95% confidence interval).

To explore the metabolome distribution by time series, interactive principal component analysis with Pareto-scaled data was performed using MetaboAnalyst 5.0. Metabolites were then hierarchically clustered by the Ward method with Euclidean distance, following autoscale feature standardization. Significant metabolic changes according to the intervention duration were analyzed by repeated measures one-way ANOVA (FDR-adjusted p-value < 0.05). Differences in longitudinal metabolic changes between responders and non-

responders were determined by repeated measures two-way ANOVA (FDR-adjusted p-value < 0.05).

Significant metabolites between baseline and 6 months post-intervention or baseline and 18 months post-intervention were determined by paired t-test (FDR-adjusted p-value < 0.05 and fold change > 1.2). For the significant metabolites, I identified enriched metabolite sets based on KEGG supported by MetaboAnalyst 5.0 and mapped the metabolites according to chemical and biochemical relationships by MetaMapp. Cytoscape 3.8.2 was used to visualize the metabolite networks.

5. Development of diagnostic models for NAFLD[D]

Diagnostic models for NAFLD were developed and validated using machine learning techniques. The total dataset was split into training/validation and test datasets at a ratio of 4:1 and repeated 100 times (nested cross-validation). Additionally, each step was repeated three times for recurrent cross-validation. Variables were separated into NAFLD-specific metabolic biomarkers and clinical and genetic variables[E] including age, sex, BMI z-score, AST, ALT, GGT, ALP, and three significant genetic variants (*PNPLA3* rs738409, *SAMM50* rs2073080, and rs3761472). The diagnostic models using NAFLD-specific metabolic biomarkers were not adjusted for clinical factors during their development, as I wanted to create models that do not require clinical factor inputs and could be compared with models using clinical and genetic variables.

The following four machine learning models, which were previously used in the diagnosis of NAFLD, were evaluated (30): logistic regression, the generalized linear model with an elastic net penalty (ElasticNet) (31), random forest (32), and extreme gradient

---

[D] Establishing and validating models for diagnosis of NAFLD were performed by Ki Young Huh.

[E] Genotyping of NAFLD-related genetic variants had been performed by Prof. Kyung Jae Lee and Prof. Jae Sung Ko and previously reported by Woori Chae and Prof. Kyung Jae Lee (Chae W, Lee KJ, et al. Association of Metabolic Signatures with Nonalcoholic Fatty Liver Disease in Pediatric Population. Metabolites. 2022;12(9)).

boosting (XGBoost) (33). The model hyperparameters, except the logistic regression model, were tuned using grid searching. The model performance for each repeated test set was evaluated by measuring the area under the receiver operating characteristic curve (AUROC), accuracy, sensitivity, specificity, and F1 score on the test. For the logistic regression model with a median AUROC, the regression coefficient, standard error, and z- and *p*-values of the selected variables were obtained. The variable importance scores of 18 NAFLD-specific metabolic biomarkers were calculated for ElasticNet, random forest, and XGBoost models with a median AUROC. Model building and validation were conducted using R version 4.1.0 (34) and R package *caret* (35).

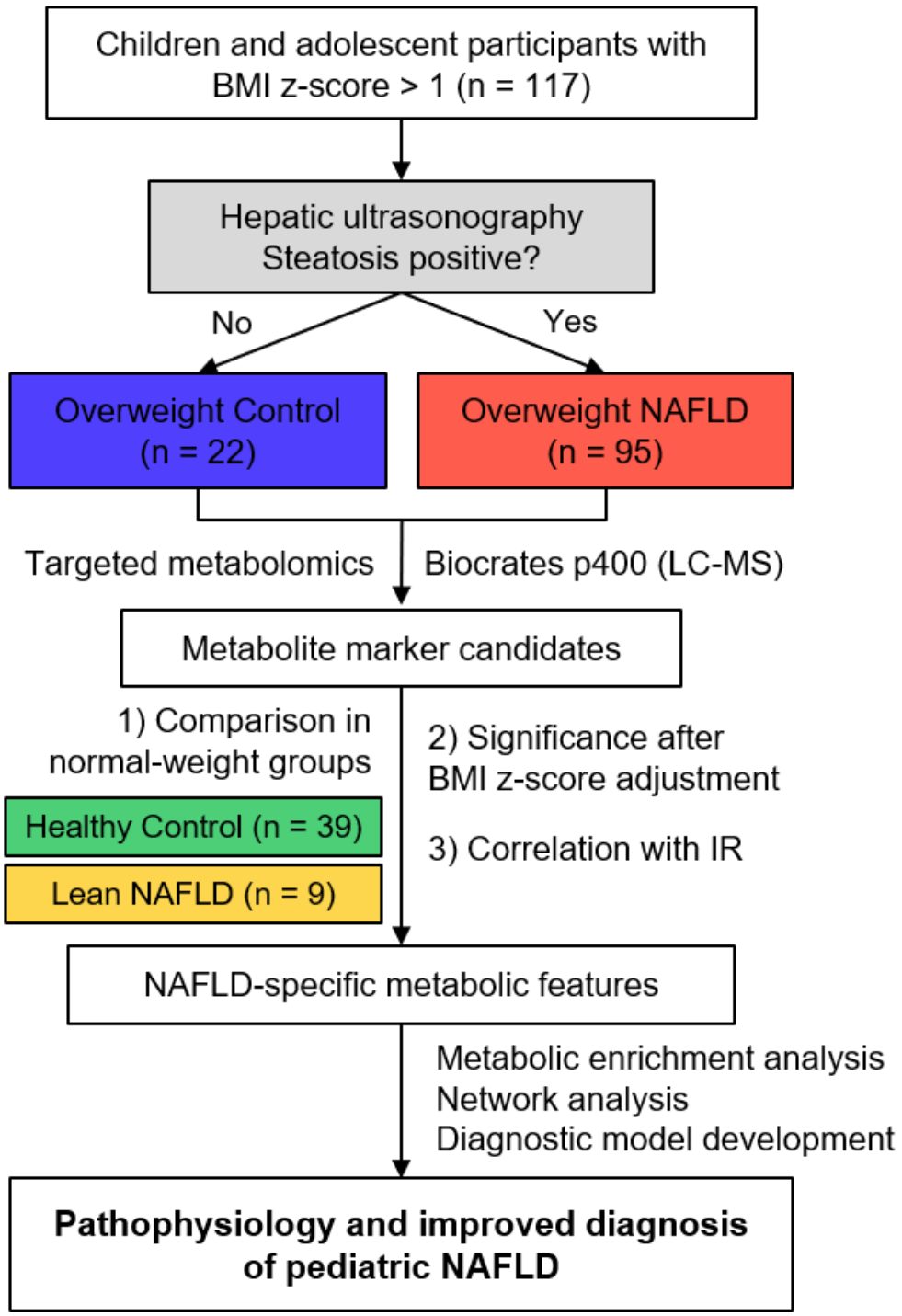


Figure 1. Schematic study design and analysis workflow of the pediatric NAFLD cohort

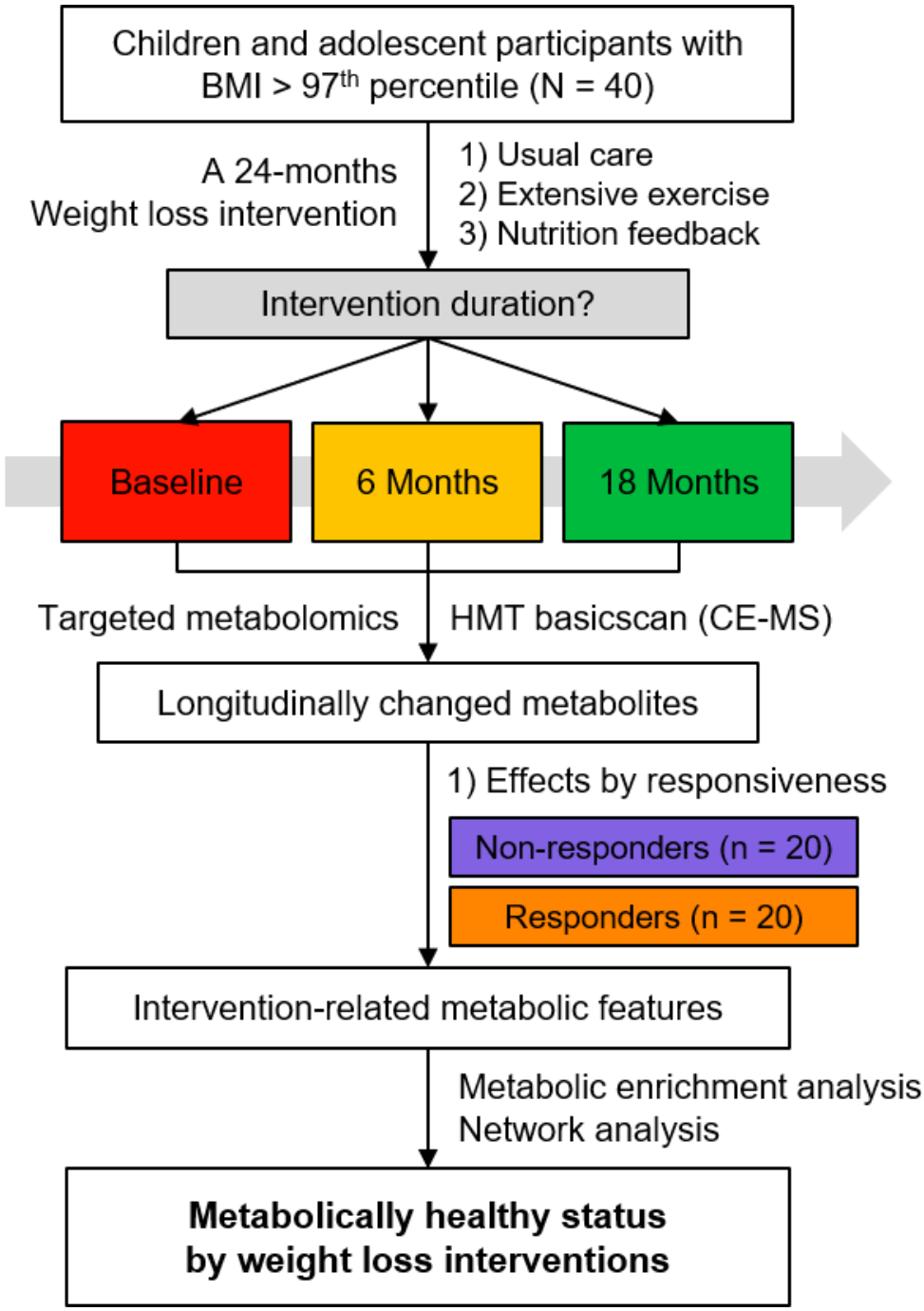


Figure 2. Schematic study design and analysis workflow of the obesity intervention cohort

Results

1. Clinical characteristics of the pediatric NAFLD cohort

I performed plasma metabolomics on a pediatric cohort with NAFLD to clarify the characteristics of pediatric NAFLD. In this study, 165 Korean children and adolescent participants aged 6 to 19 years were selected from the cohort. Their demographic features, including age, sex, BMI z-score, steatosis grade, liver function test results, and insulin resistance-related parameters, are summarized in **Table 1** and **Figure 3**. As described in the Methods section, steatosis grade and BMI z-score were used to classify the participants into four groups. The AST, ALT, and GGT levels were abnormally elevated in the NAFLD group, whereas those in the control group were in the normal range (36). In contrast, the between-group difference in the ALP level was not significant. While insulin and HOMA-IR levels were significantly increased in the ON group compared to the OC group, no differences in HbA1c levels were observed between the HC, LN, OC, and ON groups.

Table 1. Clinical characteristics of the study population according to the occurrence of obesity and NAFLD

	Healthy Control (HC)	Lean NAFLD (LN)	Overweight Control (OC)	Overweight NAFLD (ON)	Significance *
The number of subjects	39	9	22	95	-
Age (year)	14.3 (8.7–18.6)	10.6 (9.6–17.4)	14.3 (6.6–17.6)	12.5 (6.4–18.9)	0.3546
Sex (male/female)	27/12	9/0	12/10	74/21	-
BMI z-score	−0.45 (−2.37–0.96)	0.88 (0.74–1.00)	1.63 (1.11–3.04)	2.42 (1.08–5.94)	<0.0001
Steatosis grade	0 (0)	2 (1.5–2.5)	0 (0)	2 (1–3)	-
AST (IU/L)	21 [17–25]	51 [30–57]	20 [16–25]	46 [29–76]	<0.0001
ALT (IU/L)	12 [10–17]	73 [52–91]	18 [14–25]	84 [40–144]	<0.0001
GGT (IU/L)	11 [9–13]	28 [19–42]	16 [13–19]	34 [22–57] †	<0.0001
ALP (IU/L)	202 [141–290]	267 [227–370]	137 [90–305]	256 [128–371]	0.05

Fasting glucose (mg/dL)	97 [91–102]	96 [94–105]	101 [99–104]	100 [95–108] ‡	0.0175
Insulin (mU/L) §	7.3 [4.5–16]	10.9 [8.1–48]	9.0 [6.3–12]	17.5 [12–23]	0.0012
HOMA-IR §	1.74 [1.16–4.89]	2.48 [1.96–12.2]	2.35 [1.55–2.89]	4.27 [3.01–5.66]	0.0028
HbA1c (%) ¶	5.3 [4.8–5.9]	5.3 [5.0–5.8]	5.2 [5.0–5.4]	5.4 [5.1–5.7]	0.2264

* Significance by the Kruskal–Wallis test; † n = 93; ‡ n = 94; § HC (n = 4), LN (n = 5), OC (n = 12), ON (n = 67); ¶ HC (n = 4), LN (n = 3), OC (n = 15), and ON (n = 77). Continuous variables are given as the median (min-max) or median [25th–75th percentile].

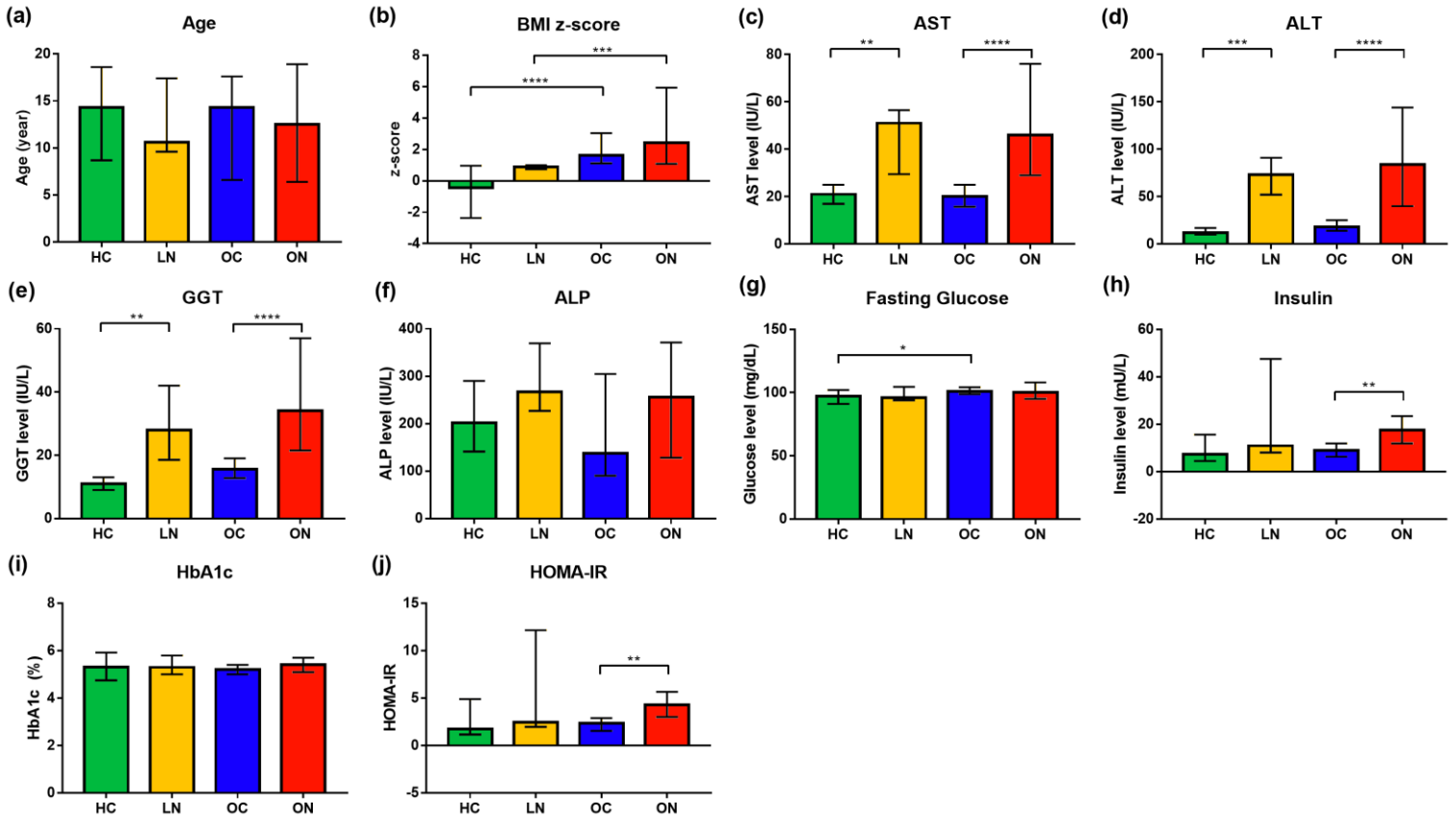


Figure 3. Clinical characteristics of the study population according to the occurrence of obesity and NAFLD

(a)-(b) Median bar charts with error bars indicating the range. (c)-(j) Median bar charts with error bars indicating the interquartile range. P-values: < 0.0332(*), < 0.0021 (**), < 0.0002 (***), < 0.0001 (****) by post-hoc Dunn's multiple comparison test following Kruskal–Wallis test.

2. Establishment of a strategy to find NAFLD-specific metabolic biomarkers

I sought to establish a strategy to find NAFLD-specific metabolic biomarkers by performing exploratory metabolomic analyses, such as multivariate and univariate analyses (**Figure 4A**). The metabolic distribution of the study population showed that the NAFLD groups (LN and ON) had relatively high intra-group variability compared to the control groups (HC and OC) (**Figure 4B**). **Figure 4C** shows the number of significant metabolites (FDR-adjusted p-value < 0.05, fold change > 1.1) between HC and LN, HC and OC, LN and ON, or OC and ON by Wilcoxon's rank-sum test. These findings imply that more metabolites were changed by NAFLD than by obesity, as none of the metabolites were even significantly changed in HC vs. OC.

I compared the metabolic profiles of the subgroups, control versus NAFLD group, or normal-weight versus overweight group (**Figure 5A**). A greater number of significant metabolites (FDR-adjusted p-value < 0.05, fold change > 1.2) were observed in the comparison between the control and NAFLD groups (84 metabolites) than in the comparison between the normal-weight and overweight groups (48 metabolites), as illustrated by Venn diagram (**Figure 5B**). In addition, most of the plasma TG, DG, and PC levels were significantly elevated in the overweight group, irrespective of NAFLD presence, but

they were also simultaneously selected as NAFLD markers (**Figures 5C and 5D**) which may act as concomitant variables.

Considering these findings, this study focused on a subpopulation with BMI z-scores > 1 (OC and ON groups) to identify promising candidates (**Figure 4C**, black arrow), then verified these by comparing the normal-weight group and adjusting the BMI and performing a correlation analysis with insulin resistance.

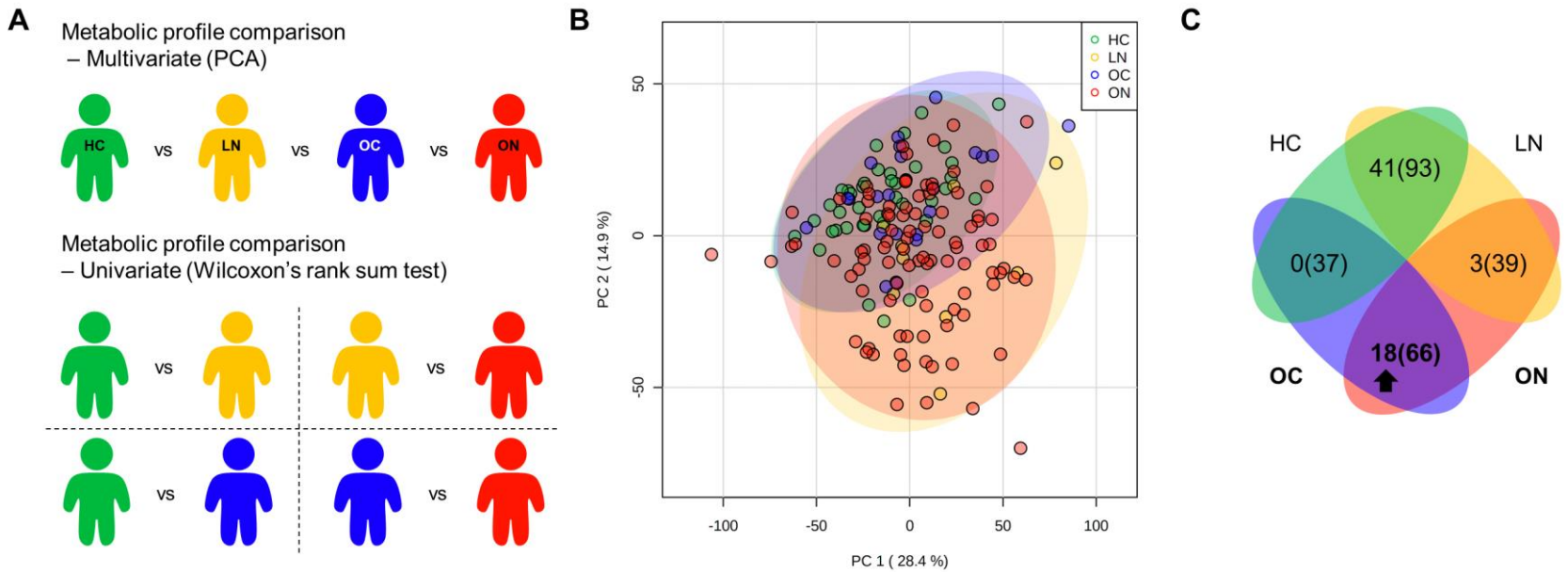
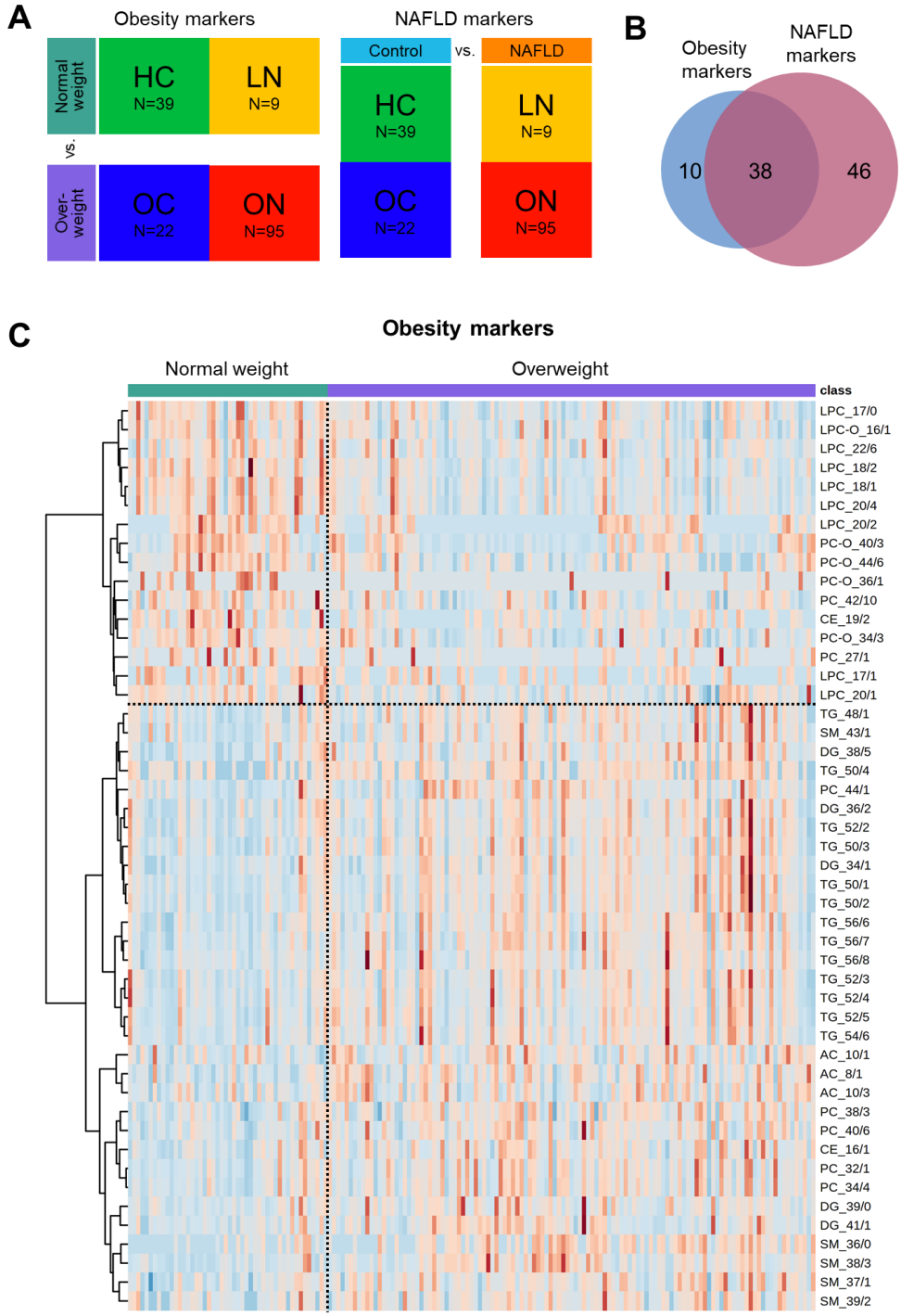


Figure 4. Exploratory metabolomic analyses using multivariate and univariate analysis of the pediatric NAFLD cohort

(A) Multivariate and univariate analyses of the exploratory comparison metabolic profiles by grouping. (B) Participant metabolic distribution through a principal component analysis (PCA) score plot that portrays relatively high intra-group variability in the NAFLD groups (LN and ON) compared to the control groups (HC and OC). (C) Significant metabolite quantities (without parentheses, FDR-adjusted *p*-value < 0.05, fold change > 1.1; in parentheses, raw *p*-value < 0.05, fold change > 1.1) between HC and LN, HC and OC, LN and ON, or OC and ON by Wilcoxon's rank-sum test.


A
Obesity markers
Normal weight
HC
N=39
LN
N=9
vs.
Over-weight
OC
N=22
ON
N=95
NAFLD markers
Control
vs.
NAFLD
HC
N=39
LN
N=9
OC
N=22
ON
N=95
B
Obesity markers
NAFLD markers
10
38
46
C
Obesity markers
Normal weight
Overweight
class
LPC_17/0
LPC-O_16/1
LPC_22/6
LPC_18/2
LPC_18/1
LPC_20/4
LPC_20/2
PC-O_40/3
PC-O_44/6
PC-O_36/1
PC_42/10
CE_19/2
PC-O_34/3
PC_27/1
LPC_17/1
LPC_20/1
TG_48/1
SM_43/1
DG_38/5
TG_50/4
PC_44/1
DG_36/2
TG_52/2
TG_50/3
DG_34/1
TG_50/1
TG_50/2
TG_56/6
TG_56/7
TG_56/8
TG_52/3
TG_52/4
TG_52/5
TG_54/6
AC_10/1
AC_8/1
AC_10/3
PC_38/3
PC_40/6
CE_16/1
PC_32/1
PC_34/4
DG_39/0
DG_41/1
SM_36/0
SM_38/3
SM_37/1
SM_39/2

(continued)

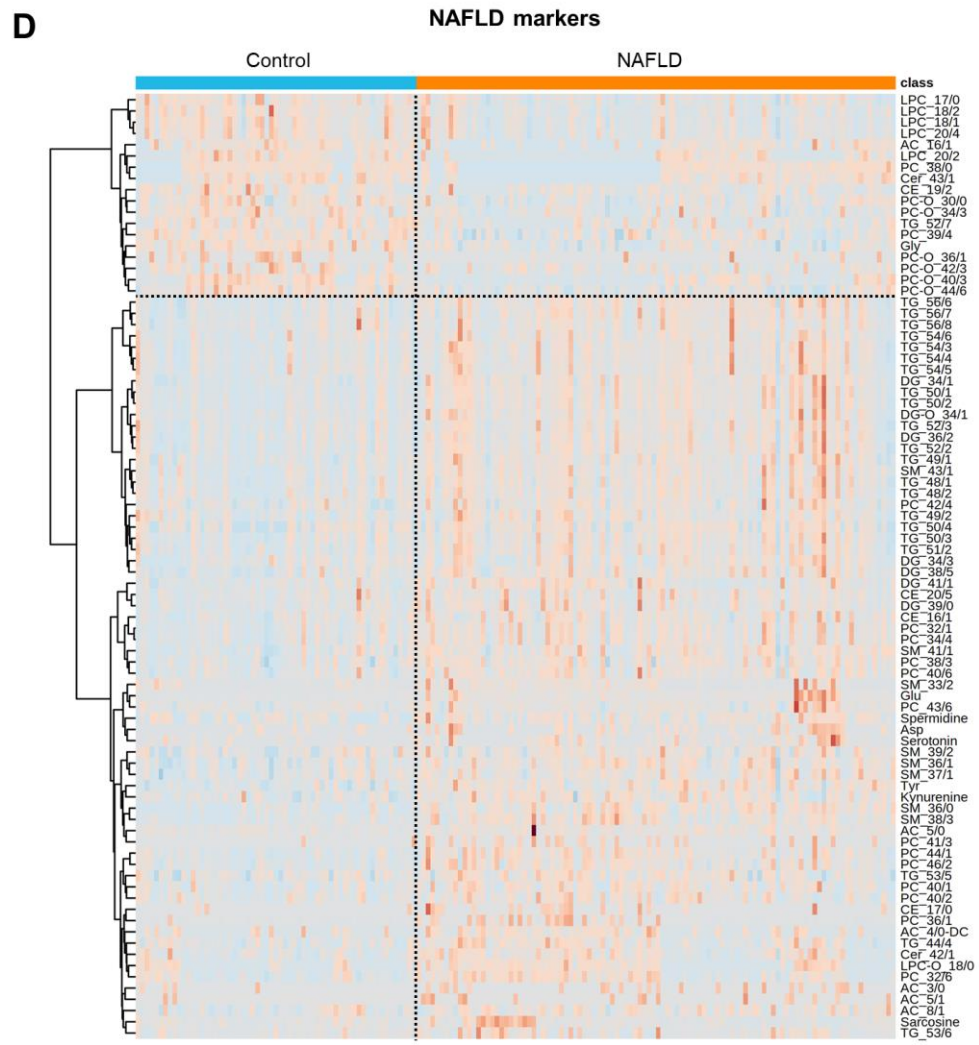


Figure 5. Differences in metabolic profiles of the pediatric NAFLD cohort subgroups

(A) Metabolic profile comparisons of normal-weight versus overweight groups and control versus NAFLD groups. (B) A Venn diagram illustrating significant metabolite quantities (FDR-adjusted *p*-value < 0.05, fold change > 1.2) observed between the control and NAFLD comparison and the normal-weight and overweight comparison, simultaneously selected as obesity and NAFLD biomarkers (38 metabolites). (C) Obesity and (D) NAFLD biomarker heatmaps.

3. Finding NAFLD-specific metabolic biomarkers

3.1. Significant metabolites between OC and ON groups

As previously mentioned in the Results section, 18 metabolites were significantly different (FDR-adjusted *p*-value < 0.05, fold change > 1.1, 14 up and 4 down) between the OC and ON groups (**Figure 6A**). The levels of multiple amino acids, including BCAAs (valine, leucine, and isoleucine), lysine, tyrosine, and glutamic acid, were significantly higher in the ON group than in the OC group, whereas the glycine level was lower in the ON group (**Figure 6B**). The levels of glycerolipids, phospholipids, and sphingolipids, including TG (50:1), TG (54:3), DG (34:1), PC (46:2), PC (44:1), SM (36:0), and SM (38:3), were also elevated in the ON group, while TG (52:7), LPC (18:2), and PC-O (30:0) levels were reduced. The valerylcarnitine (AC (5:0)) level was higher in the ON than in the OC group.

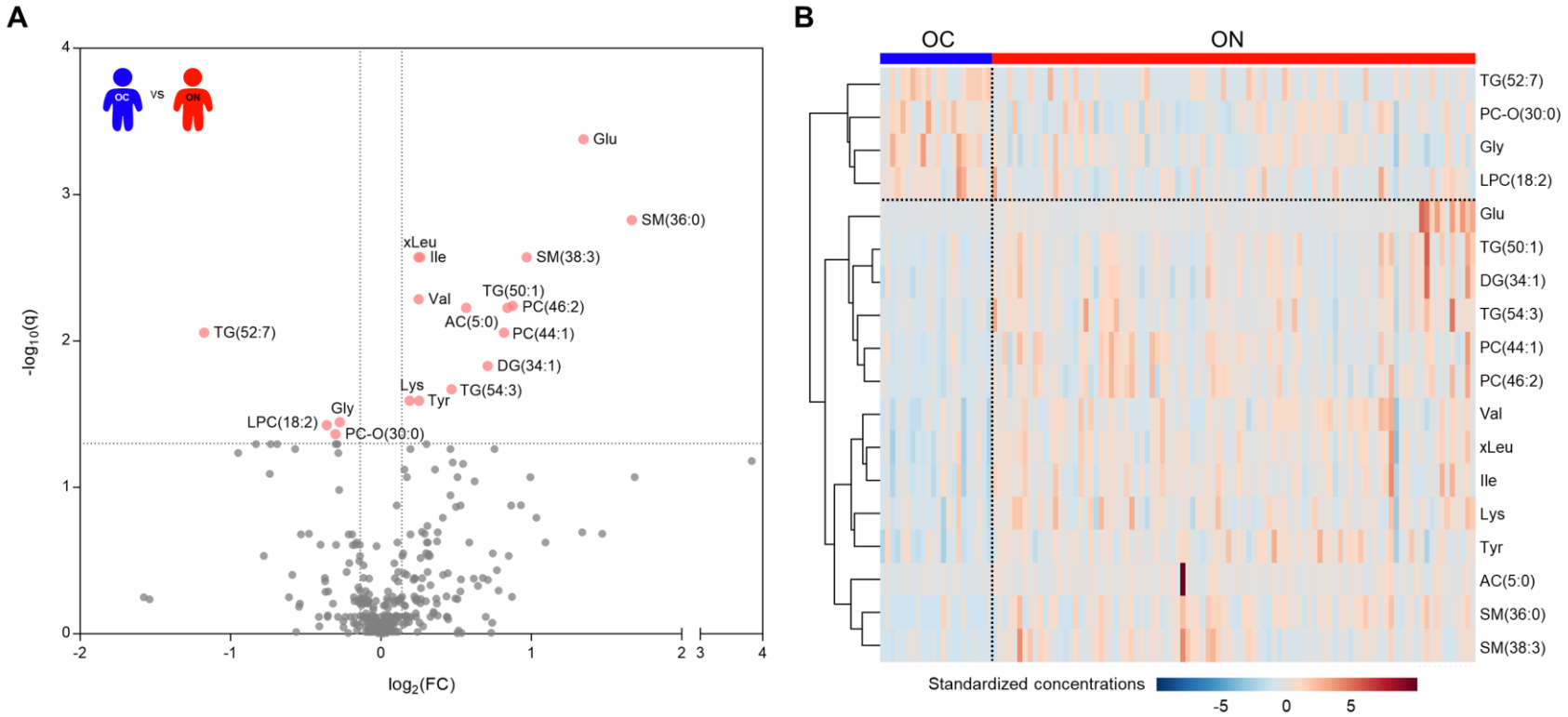


Figure 6. Metabolic features of overweight NAFLD (ON) compared overweight controls (OC)

(A) Volcano plot and (B) heat map of significant plasma metabolites (NAFLD-specific metabolic biomarkers, FDR adjusted p-value < 0.05 by Wilcoxon rank-sum test, fold change > 1.1) between OC and ON groups. In the volcano plot, significant metabolites are labeled with coral red. Each metabolite's concentration was standardized in the heat map by feature autoscaling, followed by metabolite clustering using the Ward method with Euclidean distance.

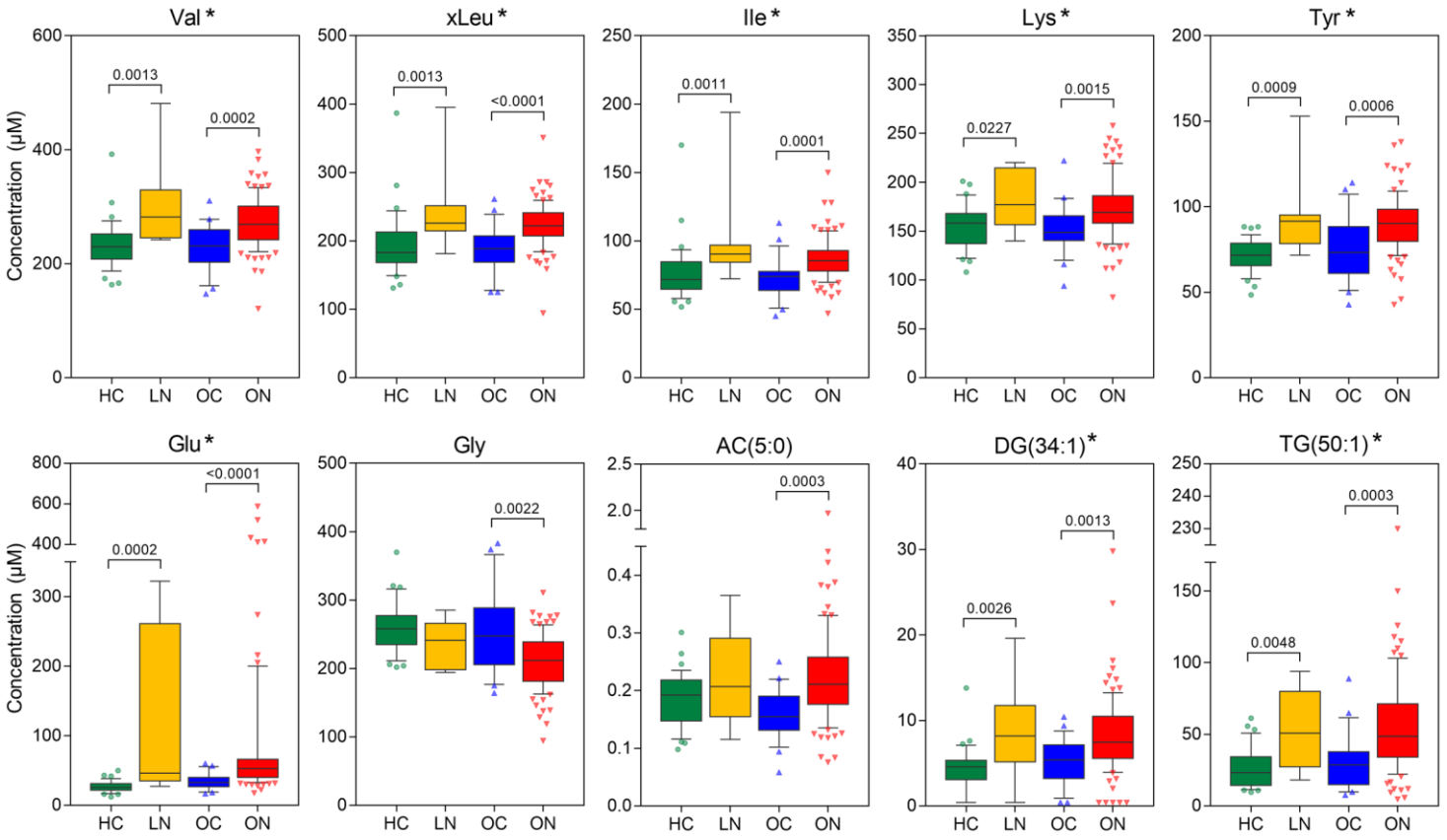
3.2. Verification of the significant metabolites: are they really "NAFLD-specific"?

The 18 significant metabolites were evaluated in the normal-weight group to determine if they were NAFLD-specific. In total, 11 of these 18 metabolites showed statistically significant differences, with the same direction of change as observed in the overweight population (**Figure 7**, metabolites marked with an asterisk). The other 7 metabolites (**Figure 7**, metabolites without an asterisk) also showed the same direction of change, although changes were not significant between the HC and LN groups due to the small number of participants.

I compared the metabolic features to the clinical characteristics to demonstrate if these significant metabolites were specifically correlated to NAFLD and performed multiple linear regression analyses for each of the 18 metabolites, with the BMI z-score as a confounding factor, to investigate the effect of BMI. I found that 16 of the metabolites, excluding AC (5:0) and glutamate, were significantly different between the OC and ON groups, even after controlling for the FDR (Benjamini–Hochberg method) (**Table 2**).

I also performed a correlation analysis to investigate the potential effect of insulin resistance on the significant metabolites and found that 13 of the 18 significant metabolites were weakly correlated with

insulin resistance (**Figure 8**, Spearman r > 0.35, $p < 0.05$), except Tyr, Lys, Gly, LPC (18:2), and PC-O (30:0). However, no difference in HOMA-IR level was observed between the HC and the LN groups in the normal-weight population, whereas a higher HOMA-IR level was observed in the ON group than in the OC group ($p = 0.0035$ by post-hoc Dunn's multiple comparison test following the Kruskal–Wallis test). Hence, these 18 metabolites were regarded as "NAFLD-specific" metabolic biomarkers.


Val *
xLeu *
Ile *
Lys *
Tyr *
Glu *
Gly
AC(5:0)
DG(34:1)*
TG(50:1) *
Concentration (μM)
HC
LN
OC
ON
0.0013
0.0002
<0.0001
0.0011
0.0001
0.0227
0.0015
0.0009
0.0006
0.0022
0.0003
0.0026
0.0048

(continued)

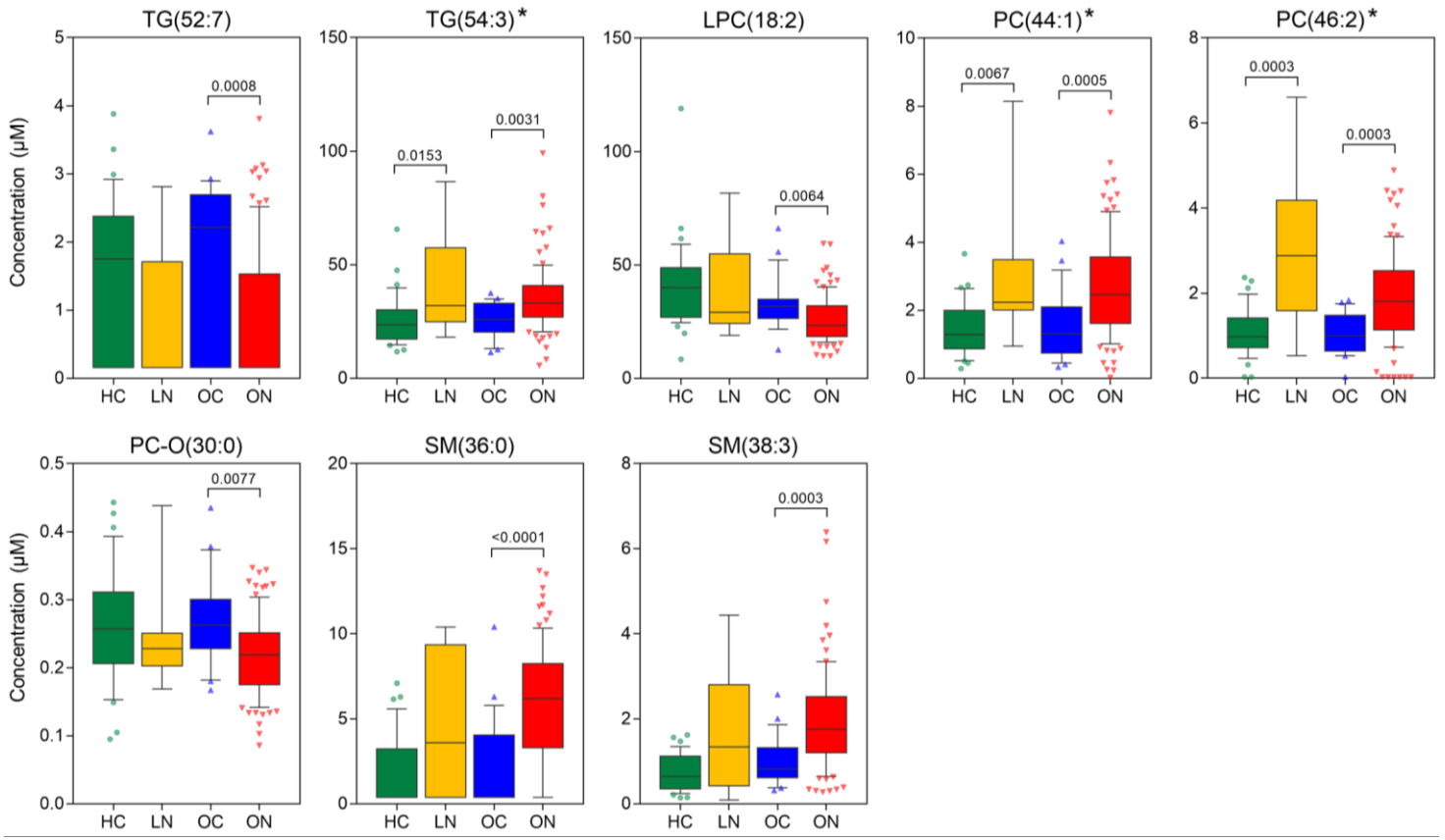


Figure 7. Concentrations of NAFLD-specific metabolic biomarkers in four pediatric NAFLD cohort groups

The box-and-whiskers plots indicate the median with the 10^{th} and 90^{th} percentile whiskers, whereas the X-axis distinguishes the study group. Significance (*q*-value) was calculated from the Kruskal–Wallis test, followed by the two-stage step-up method of Benjamini, Krieger, and Yekutieli for multiple comparison correction by controlling FDR. Multiple comparison per metabolite: HC vs. OC, HC vs. LN, OC vs. ON, and LN vs. ON. *q*-values < 0.05 were denoted on the panels. Metabolites ordered according to structural classes: BCAAs (Val, xLeu, and Ile), other amino acids (Lys, Tyr, Glu, and Gly), acylcarnitines and glycerolipids (AC (5:0), DG (34:1), TG (50:1), TG (52:7), and TG (54:3)), phosphatidylcholines (LPC (18:2), PC (44:1), PC (46:2), and PC-O (30:0)), and sphingomyelins (SM (36:0) and SM (38:3)). Metabolites marked with an asterisk (*) showed statistically significant differences in HC vs. LN, with the same direction of change as observed in the overweight population.

Table 2. Significant differences in 18 NAFLD-specific metabolic biomarkers after BMI z-score adjustment

Metabolite	Raw *p*-value*	FDR**-adjusted *p*-value
AC (5:0)	0.12682	0.1268
Glu	0.10228	0.1083
Gly	0.00115	0.0036
Ile	0.00159	0.0036
Lys	0.01402	0.0168
Tyr	0.00157	0.0036
Val	0.00145	0.0036
xLeu	0.00058	0.0035
(34:1)	0.01595	0.0179
TG (50:1)	0.01134	0.0146
TG (52:7)	0.00119	0.0036
TG (54:3)	0.00810	0.0121
LPC (18:2)	0.00910	0.0126
PC (44:1)	0.00724	0.0118
PC (46:2)	0.00309	0.0056
PC-O (30:0)	0.00057	0.0035
SM (36:0)	0.00007	0.0013
SM (38:3)	0.00179	0.0036

* Raw *p*-values were calculated from multiple linear regression analyses (metabolite ~ BMI z-score + Phenotype (OC or ON)). ** False discovery rate (FDR) was controlled for using the Benjamini–Hochberg method.

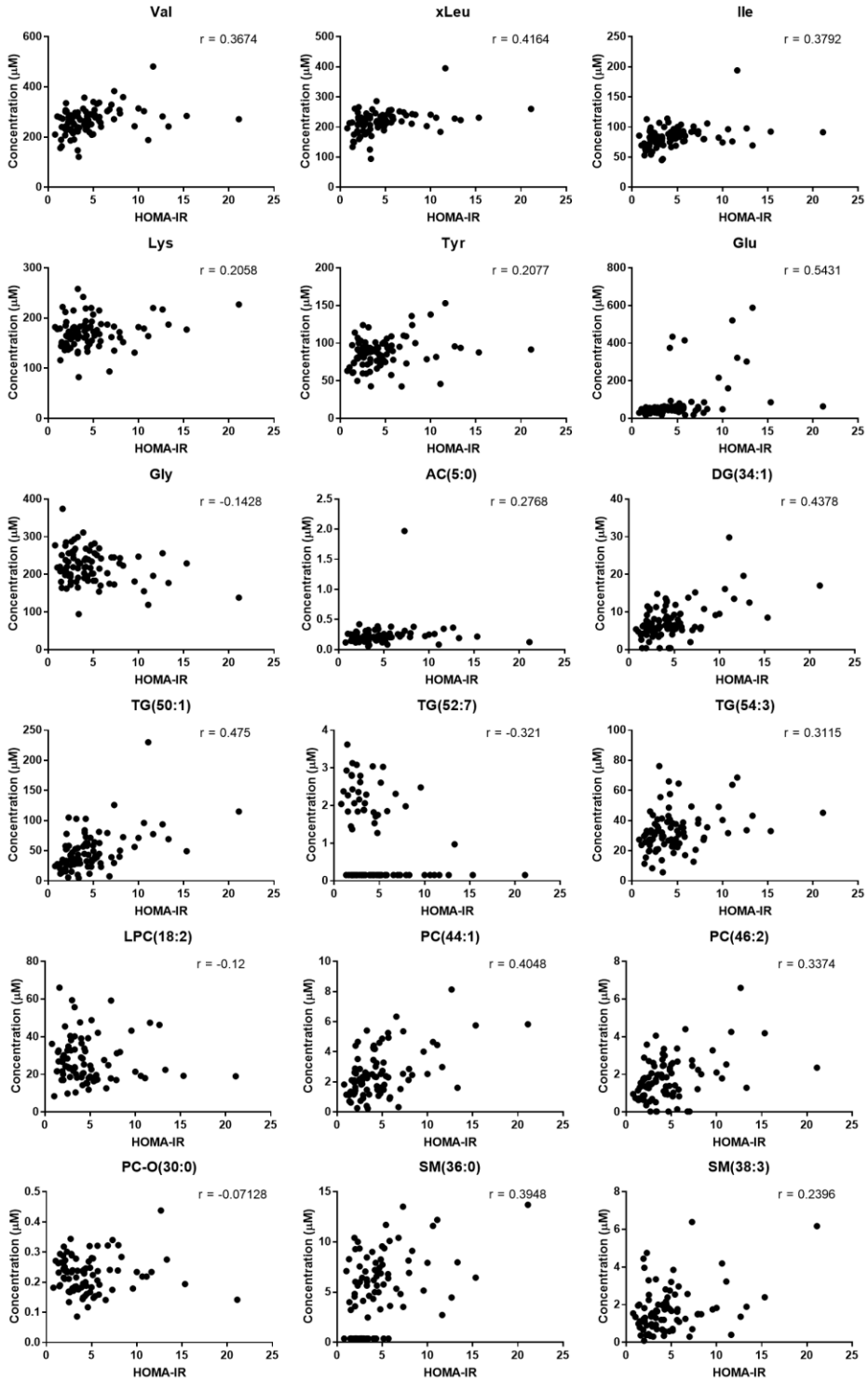


Figure 8. Spearman correlations between NAFLD-specific metabolic biomarkers and insulin resistance (HOMA-IR)

4. Longitudinal changes in the obesity intervention cohort

4.1. Changes in clinical characteristics by weight loss intervention

From the ICAAN cohort, I selected 40 patients with obesity. In detail, 131 subjects among 242 participants dropped out during the intervention due to busy schedules, no response, lack of willingness, or their parents; busy schedules. 163 and 111 participants were followed up at 6 and 18 months after the intervention, respectively. The clinical characteristics of the study population are presented in **Table 3** and **Figure 9**. BMI z-score was significantly decreased by weight loss intervention compared to baseline (repeated measures ANOVA $p = 0.0401$; post-hoc FDR-adjusted $p = 0.0162$ (baseline-6M and baseline-18M)), whereas no longitudinal differences were observed in other clinical characteristics including AST, ALT, TG, HDL-C, and LDL-C.

Table 3. Demographic characteristics of the obesity intervention cohort

	Baseline	6M post-intervention	18M post-intervention
The number of datapoints	40	30	40
Type of intervention (U/E/N)	10/15/15	-	-
Intervention response (R/NR)	20/20	-	-
Sex (female/male)	19/21	-	-
Age at baseline (year)	10.8 [9.91-12.2]	-	-
BMI z-score	2.935 [2.397-3.585]	2.595 [2.124-3.423]	2.514 [2.026-3.539]
Weight (kg)	64.0 [50.1-79.3]	71.7 [52.4-86.1]	68.25 [60.0-86.8]
AST (IU/L)	21.0 [18.3-26.0]	19.5 [17.0-24.3]	19.0 [14.3-24.8]
ALT (IU/L)	18.5 [13.3-30.0]	21.0 [14.8-26.0]	19.0 [13.0-26.8]
TGs (mg/dL)	90.5 [61.0-133]	108 [77.5-141]	96.5 [67.5-122]
HDL-C (mg/dL)	50 [41-58]	47 [40-57]	53 [43-59]
LDL-C (mg/dL)	113 [101-130]	113 [95.8-128]	112 [97.0-124]

Continuous variables are given as the median [25th–75th percentile]. Abbreviations: U, usual care group; E, extensive exercise group; N, nutrition feedback group; R, responder; NR, non-responder.

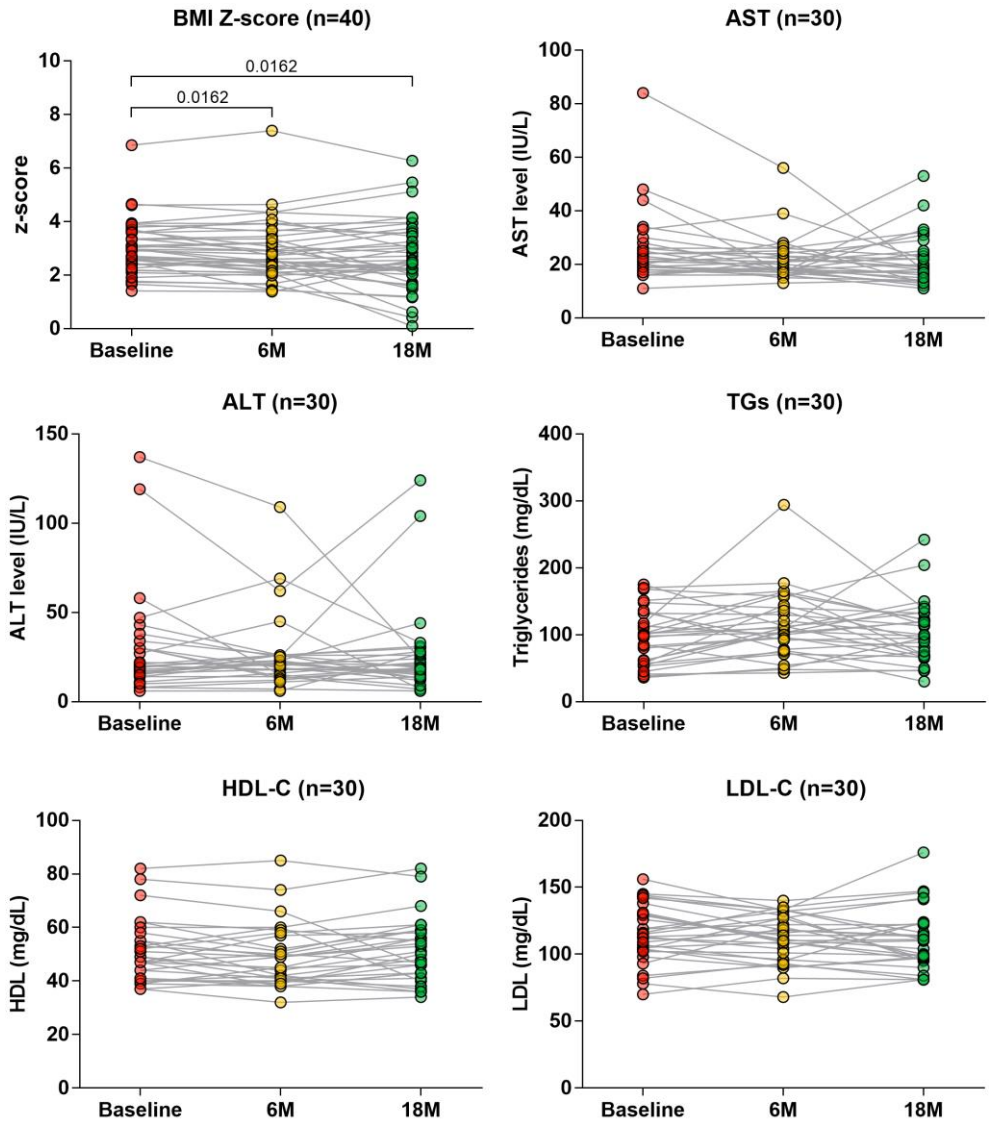
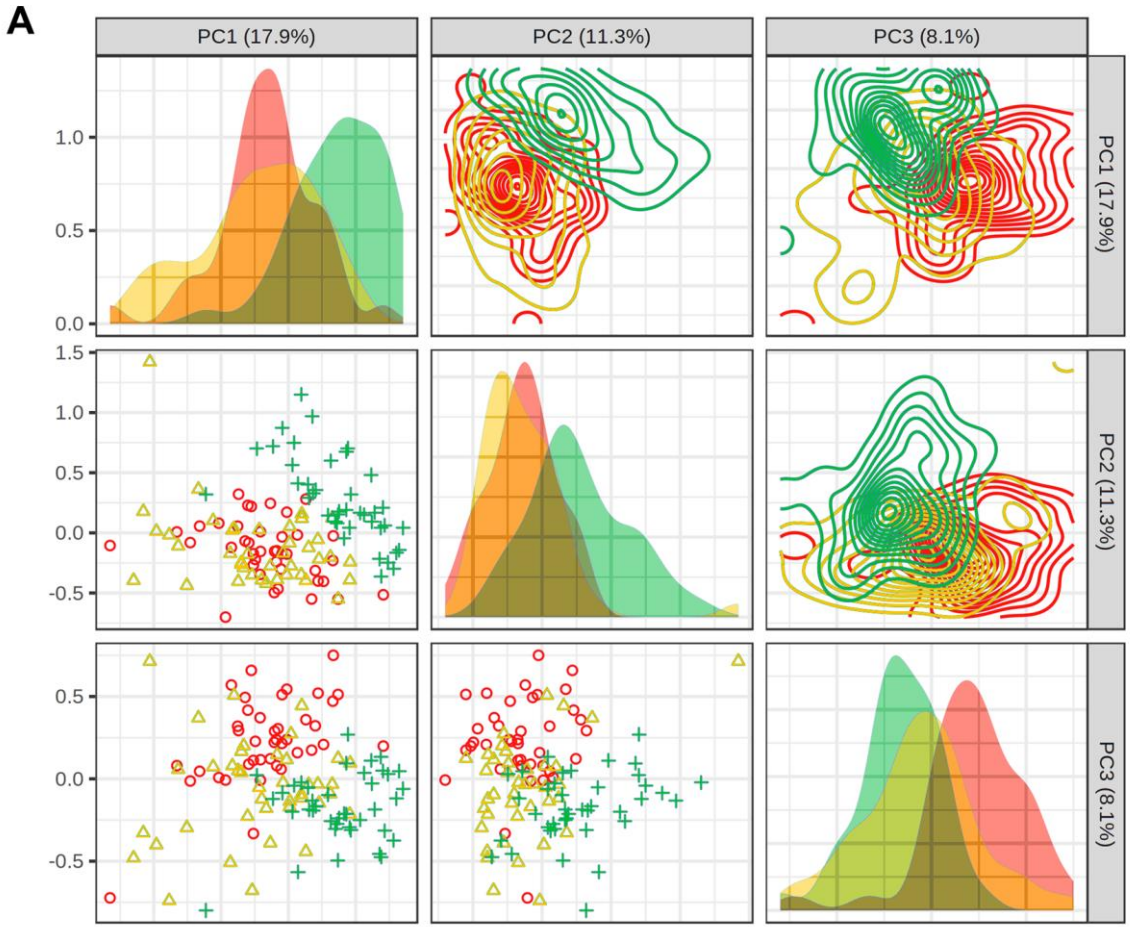


Figure 9. Longitudinal changes of the individual study population's clinical characteristics by weight loss intervention

Longitudinal change significance calculated through One-Way Repeated Measures ANOVA with Greenhouse-Geisser correction, followed by the FDR-adjusted multiple comparisons (baseline-6M and baseline-18M, denoted in the panels) proposed by Benjamini, Krieger, and Yekutieli.

4.2. Changes in plasma metabolites by weight loss intervention

A total of 194 plasma metabolites were successfully identified from capillary electrophoresis time-of-flight mass spectrometry measurements based on HMT's standard library and Known-Unknown peak library. I selected 185 metabolites for further analyses after filtering and observed the distinct metabolomic distribution of 18 months post-intervention (green crosses) compared to baseline (red circles) or 6 months post-intervention (yellow triangles) data on the score plots by interactive principal component analysis (**Figure 10A**). These remarkable changes were also revealed by hierarchical clusters of metabolites on the heatmap (**Figure 10B**) and by repeated measures of one-way ANOVA (**Figure 10C**, 77 metabolites labeled with blue circles, FDR-adjusted p-value < 0.05).


A
PC1 (17.9%)
PC2 (11.3%)
PC3 (8.1%)
PC1 (17.9%)
PC2 (11.3%)
PC3 (8.1%)
1.0
0.5
0.0
1.5
1.0
0.5
0.0
-0.5
0.5
0.0
-0.5

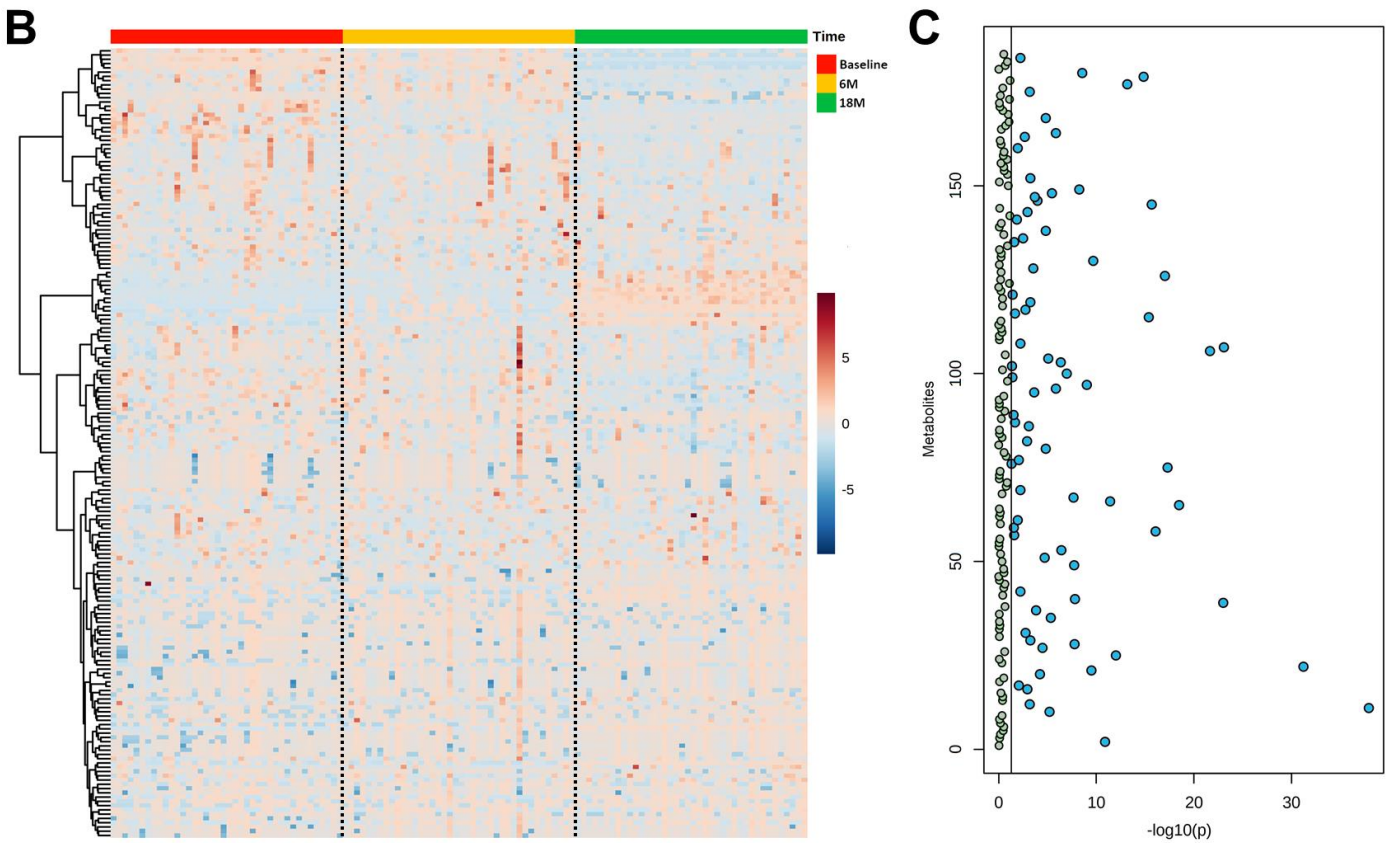


Figure 10. Longitudinal metabolic changes by weight loss intervention in the study cohort's plasma

(A) Interactive principal component analysis score plot with Pareto-scaled metabolite data and (B) hierarchically clustered heat map through Ward method with Euclidean distance showing longitudinal metabolic changes after weight loss interventions (red = baseline; yellow = 6 months post-intervention; green = 18 months post-intervention). (C) Significantly changed metabolites calculated through One-Way Repeated Measures ANOVA (blue = FDR-adjusted p-value < 0.05)

4.3. The effects of intervention responsiveness on clinical and metabolic changes following a weight loss intervention

Clinical and metabolic change comparisons between responder and non-responder groups determined weight loss intervention effects (**Figure 11**). Significant interaction effects between longitudinal change in HDL-C or LDL-C levels and intervention responsiveness were observed (interaction, $p = 0.0152$ for HDL-C; $p = 0.0055$ for LDL-C levels). However, simple main effects, such as baseline to 18 months post-intervention (intervention duration) longitudinal changes and intervention responsiveness at baseline or 18 months post-intervention (intervention response), were not statistically significant (**Table 4**). Other clinical characteristics include AST, ALT, and TGs levels, and TG/HDL-C ratio did not show any significant interaction effects or simple main effects.

In addition, the principal component analysis revealed that no significant plasma metabolites differ between responder and non-responder groups at baseline, 6 months post-intervention, and 18 months post-intervention, respectively, whereas plasma samples at 18 months post-intervention were separated from those at baseline or 6 months post-intervention (**Figure 12A**). Repeated measures two-way ANOVA analysis also showed that plasma metabolites were

longitudinally changed (time, 77 metabolites), irrespective of the intervention effect (response, none of the metabolites) (**Figure 12B**). These results suggest that a long-term weight loss intervention may induce metabolic changes regardless of the intervention response in children and adolescents.

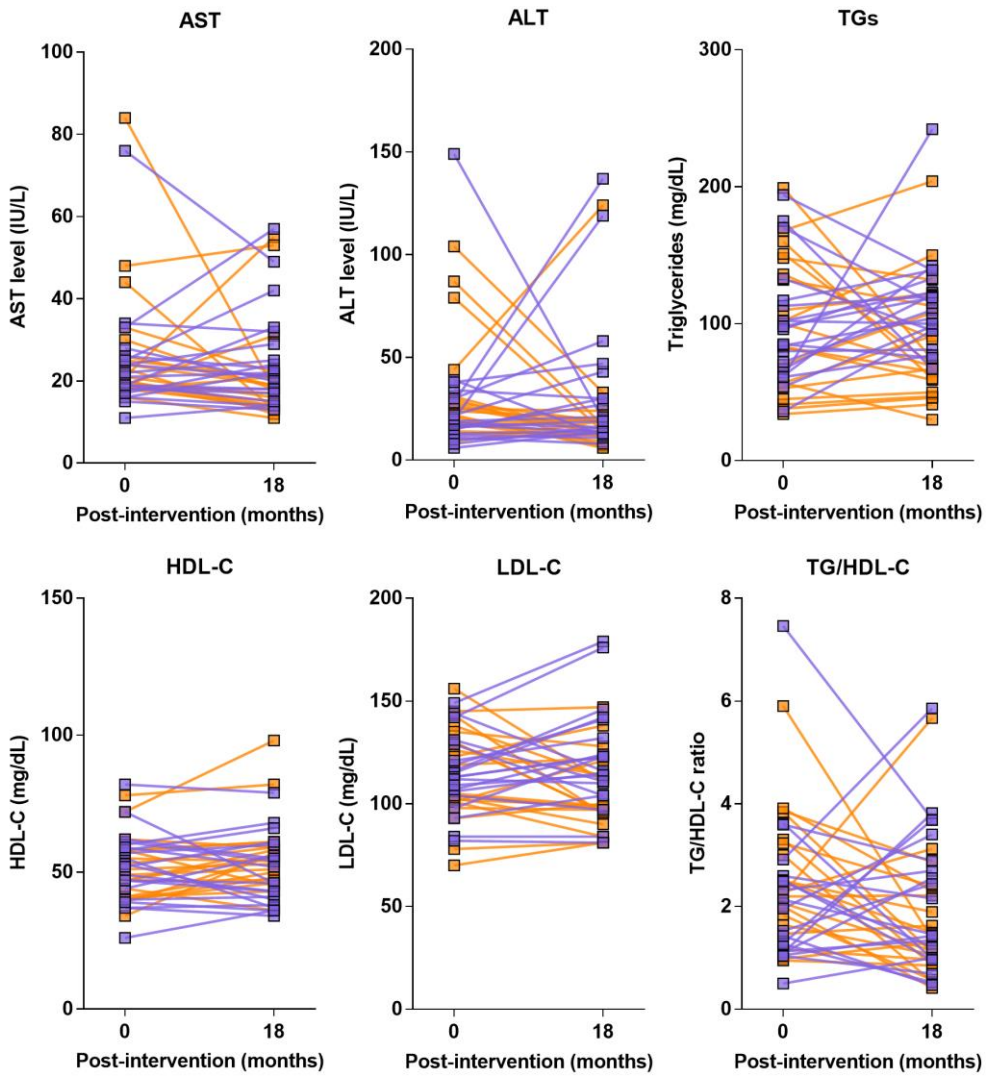


Figure 11. Weight loss intervention-induced individual changes (n=40) in clinical characteristics grouped by intervention responsiveness

Spaghetti plots to compare significant metabolic changes from baseline to 18 months post-intervention between responders (orange) and non-responders (purple) determined by repeated measures two-way ANOVA and Sidak's multiple comparisons test.

Table 4. Significant interaction effects between clinical characteristic changes and intervention responsiveness

P-values*	AST	ALT	TGs	HDL-C	LDL-C	TG/HDL-C
Interaction	0.1002	0.1303	0.075	**0.0152**	**0.0055**	0.106
Intervention duration	0.2728	0.7182	0.8751	0.3218	0.7879	0.1158
Intervention response	0.8504	0.6689	0.174	0.4797	0.1972	0.9057

* *P*-values were calculated by the repeated measures two-way ANOVA. Significance was defined as p-value < 0.05.

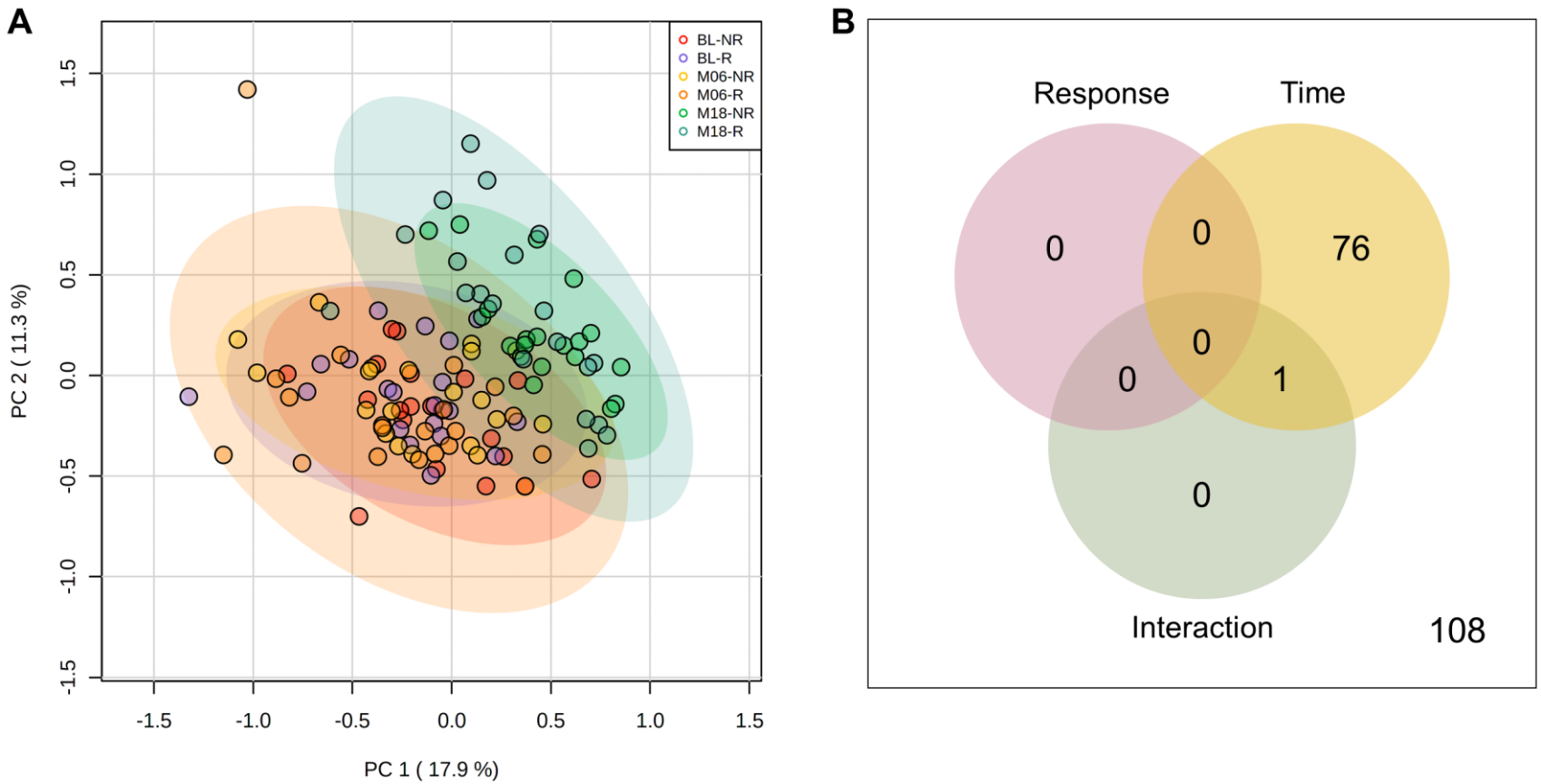


Figure 12. Response-independent longitudinal metabolic changes by the weight loss intervention

(A) Principal component analysis score plot showing longitudinal metabolic changes without differences between intervention responders (R) and non-responders (NR) at each timepoints (BL = baseline; M06 = 6 months post-intervention; M18 = 18 months post-intervention). (B) Two-Way Repeated Measures ANOVA analysis displaying longitudinally changed metabolites (Time, 77 metabolites), irrespective of intervention effect (Response and Interaction, none of the metabolites).

5. Biological and clinical implications of NAFLD-specific and intervention-related biomarkers

5.1. Relevance of NAFLD-specific metabolic biomarkers in metabolic pathways

I performed metabolite set enrichment analysis based on SMPDB with the metabolic features to identify the metabolic pathways dysregulated by NAFLD. Several metabolite sets, including valine, leucine, and isoleucine degradation, alanine metabolism, glutathione metabolism, and carnitine synthesis were altered (enrichment ratio > 4, raw $p < 0.05$) in the ON group in comparison with the OC group (**Figure 13A** and **Table 5**).

Next, I mapped NAFLD-specific metabolic biomarkers and other selected metabolites based on MetaMapp and visualized the network to explore the features' biochemical and chemical relationships (**Figure 13B**). In this network, nodes with gradient color by FDR-adjusted *p*-value are NAFLD-specific metabolic biomarkers. Gray nodes indicate other selected metabolites with a raw *p*-value < 0.05 but no significance after FDR adjustment. The network can be divided into three main clusters: (i) lipids, (ii) glutathione metabolism-related metabolites, and (iii) BCAA-related metabolites. Most of the metabolic lipid biomarkers, including PCs, SMs, TGs, and DGs, were upregulated

in the plasma of ON patients. Interestingly, I observed upregulation of glutamic acid and tyrosine and downregulation of glycine in ON patients, which are key metabolites in glutathione metabolism. Moreover, plasma BCAAs, which are reportedly altered in adult NAFLD, were also upregulated in pediatric NAFLD.

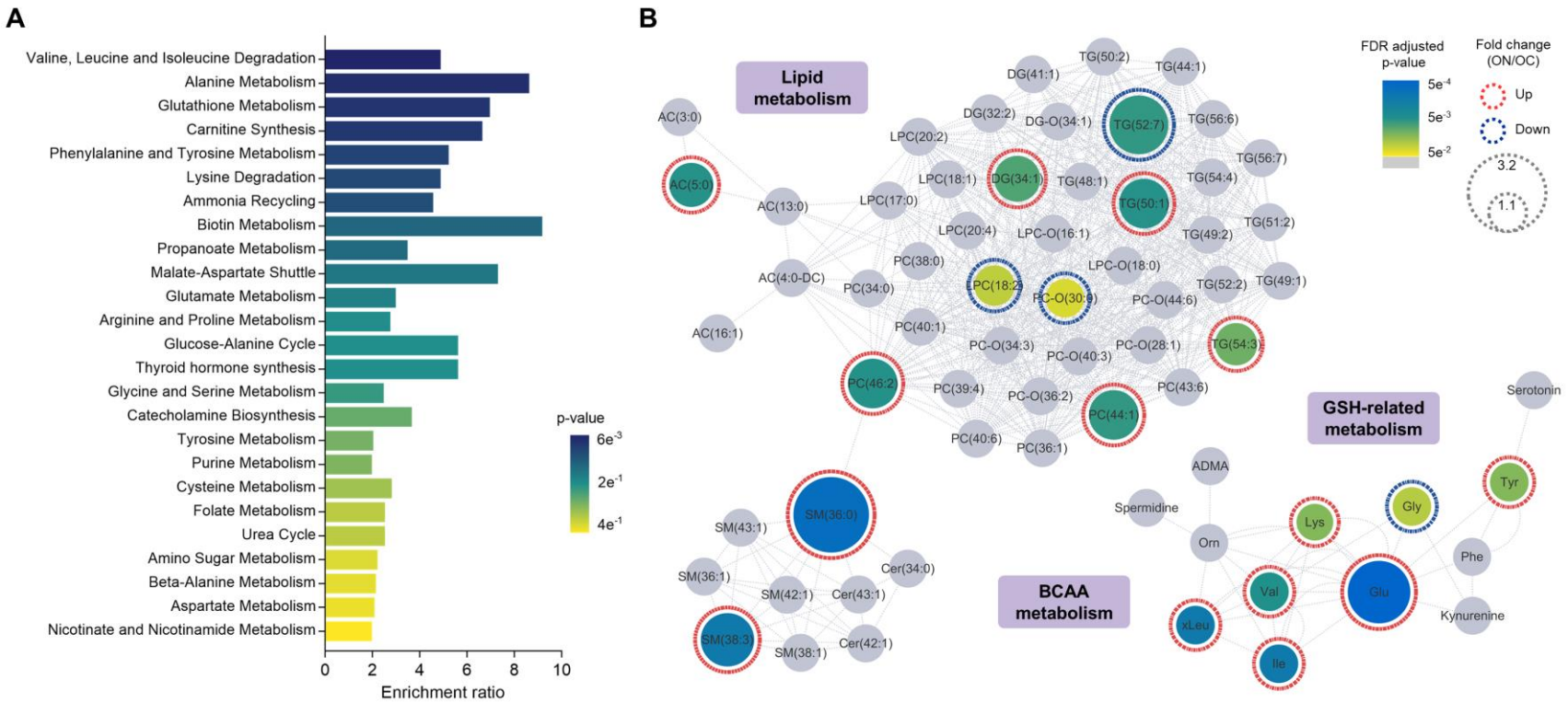


Figure 13. Relevance of the NAFLD-specific metabolic biomarkers in the metabolic pathways

(A) NAFLD-specific metabolic biomarkers' enriched metabolite sets and (B) selected significant metabolites' biochemical and chemical networks indicated the pertinent metabolic pathways in NAFLD status, such as branched-chain amino acid (BCAA),

glutathione (GSH)-related, and lipid metabolism. In the network, gray nodes represent metabolites with raw *p*-values < 0.05 that were not significant after FDR adjustment.

Table 5. Enriched significant metabolite sets in the overweight control and overweight NAFLD groups based on SMPDB by metabolite set enrichment analysis

Metabolite sets	Total†	Expected hits	Observed hits	Enrichment ratio	Raw p	Holm p	FDR
Valine, Leucine, and Isoleucine Degradation	60	0.820	4	4.88	6.86E-03	0.672	0.672
Alanine Metabolism	17	0.232	2	8.62	2.10E-02	1	0.841
Glutathione Metabolism	21	0.287	2	6.97	3.14E-02	1	0.841
Carnitine Synthesis	22	0.301	2	6.64	3.43E-02	1	0.841
Phenylalanine and Tyrosine Metabolism	28	0.383	2	5.22	5.36E-02	1	0.954
Lysine Degradation	30	0.410	2	4.88	6.07E-02	1	0.954
Ammonia Recycling	32	0.438	2	4.57	6.81E-02	1	0.954
Biotin Metabolism	8	0.109	1	9.17	1.05E-01	1	1
Propanoate Metabolism	42	0.574	2	3.48	1.09E-01	1	1
Malate-Aspartate Shuttle	10	0.137	1	7.30	1.29E-01	1	1
Glutamate Metabolism	49	0.670	2	2.99	1.42E-01	1	1
Arginine and Proline Metabolism	53	0.725	2	2.76	1.61E-01	1	1
Glucose-Alanine Cycle	13	0.178	1	5.62	1.65E-01	1	1

Thyroid Hormone Synthesis	13	0.178	1	5.62	1.65E-01	1	1
Glycine and Serine Metabolism	59	0.807	2	2.48	1.91E-01	1	1
Catecholamine Biosynthesis	20	0.273	1	3.66	2.43E-01	1	1
Tyrosine Metabolism	72	0.984	2	2.03	2.58E-01	1	1
Purine Metabolism	74	1.010	2	1.98	2.68E-01	1	1
Cysteine Metabolism	26	0.355	1	2.82	3.04E-01	1	1
Folate Metabolism	29	0.396	1	2.53	3.33E-01	1	1
Urea Cycle	29	0.396	1	2.53	3.33E-01	1	1
Amino Sugar Metabolism	33	0.451	1	2.22	3.70E-01	1	1
Beta-Alanine Metabolism	34	0.465	1	2.15	3.79E-01	1	1
Aspartate Metabolism	35	0.479	1	2.09	3.87E-01	1	1
Nicotinate and Nicotinamide Metabolism	37	0.506	1	1.98	4.05E-01	1	1
Porphyrin Metabolism	40	0.547	1	1.83	4.30E-01	1	1
Methionine Metabolism	43	0.588	1	1.70	4.54E-01	1	1
Histidine Metabolism	43	0.588	1	1.70	4.54E-01	1	1
Warburg Effect	58	0.793	1	1.26	5.60E-01	1	1
Tryptophan Metabolism	60	0.820	1	1.22	5.73E-01	1	1

Bile Acid Biosynthesis	65	0.889	1	1.12	6.03E-01	1	1
Arachidonic Acid Metabolism	69	0.943	1	1.06	6.26E-01	1	1

† The number of metabolites in a metabolite set

5.2. Application of the NAFLD-specific metabolic biomarkers to the diagnosis of pediatric NAFLD

Diagnostic models were established based on machine learning approaches using the 18 NAFLD-specific metabolic biomarkers found in this study to suggest a pathophysiology-based complementary method for biopsy-proven diagnosis (**Table 6**). Based on coefficients of the logistic regression model and variable importance scores in other models, I identified the following metabolites as significant features: valine, tyrosine, glutamic acid, glycine, and SM (38:3) (**Tables 7 and 8**). All four diagnostic models using NAFLD-specific metabolic biomarkers demonstrated excellent predictive performances, with median AUROC values of 0.95 (ElasticNet and random forest) and 0.94 (logistic regression and XGBoost) without significant differences between the models (**Figure 14**, colored line).

A logistic regression model using clinical and genetic variables was also developed. As the number of risk alleles of rs738409 (*PNPLA3*), rs2073080 (*SAMM50*), and rs3761472 (*SAMM50*) was positively associated with the presence of NAFLD, and the proportions of homozygous risk alleles were significantly higher in the NAFLD group than in the control group (37), I chose these three variants as the genetic variables. The model also showed comparable performance (**Figure 12**, black dashed line) to the metabolic feature-based models,

with the BMI z-score and ALT levels having a critical influence on the model. Among the diagnostic models using NAFLD-specific metabolic biomarkers, the ElasticNet model outperformed the other models with the highest median AUROC, yielding a sensitivity of 0.75 and specificity of 0.95.

Table 6. Performance metrics summary from 100 repeated runs of the diagnostic model using four machine learning methods

	NAFLD-specific metabolic biomarkers				Clinical and genetic variables			
	Logistic regression	**ElasticNet**	**Random forest**	**XGBoost**	**Logistic regression**	**ElasticNet**	**Random forest**	**XGBoost**
AUROC	0.94 (0.76-1.00)	0.95 (0.85-1.00)	0.95 (0.80-1.00)	0.94 (0.78-1.00)	0.95 (0.80-1.00)	0.95 (0.86-1.00)	0.96 (0.88-1.00)	0.95 (0.84-1.00)
Accuracy	0.88 (0.69-0.97)	0.88 (0.75-1.00)	0.88 (0.72-0.97)	0.84 (0.72-0.94)	0.88 (0.72-0.97)	0.88 (0.81-1.00)	0.88 (0.75-1.00)	0.88 (0.75-0.97)
Sensitivity	0.83 (0.50-1.00)	0.75 (0.58-1.00)	0.83 (0.50-1.00)	0.75 (0.42-1.00)	0.83 (0.58-1.00)	0.83 (0.50-1.00)	0.83 (0.50-1.00)	0.83 (0.50-1.00)
Specificity	0.90 (0.70-1.00)	0.95 (0.75-1.00)	0.90 (0.65-1.00)	0.90 (0.70-1.00)	0.90 (0.70-1.00)	0.95 (0.75-1.00)	0.90 (0.75-1.00)	0.90 (0.75-1.00)
F1 score	0.81 (0.55-0.96)	0.82 (0.67-1.00)	0.82 (0.63-0.96)	0.78 (0.53-0.92)	0.86 (0.67-0.96)	0.86 (0.67-1.00)	0.84 (0.63-1.00)	0.82 (0.60-0.96)

Values are given as the median (minimum-maximum).

Table 7. Multiple logistic regression model using NAFLD-specific metabolic biomarkers and clinical and genetic variables

Variables	Coefficient	SE	z-value	*p*-value
NAFLD-specific metabolic biomarkers				
(Intercept)	4.395	1.180	3.725	0.0002
Val	5.131	2.252	2.278	0.0227
Ile	-3.652	2.460	-1.485	0.1376
Lys	3.081	2.065	1.492	0.1356
Tyr	7.123	2.104	3.385	0.0007
Glu	28.869	9.109	3.169	0.0015
Gly	-4.840	1.896	-2.552	0.0107
TG (52:7)	-2.081	1.248	-1.668	0.0954
PC-O (30:0)	-2.910	1.527	-1.906	0.0566
SM (38:3)	5.852	2.637	2.220	0.0264
Clinical and genetic variables				
(Intercept)	3.374	0.975	3.460	0.0005
BMI z-score	7.379	1.833	4.027	< 0.0001
Sex (female)	-2.433	0.951	-2.560	0.0105
ALT	15.717	4.413	3.561	0.0004
PNPLA3 rs738409	2.585	1.105	2.339	0.0193

Table 8. Variable importance of three diagnostic models using NAFLD-specific metabolic biomarkers

ElasticNet (glmnet)		Random forest (ranger)		XGBoost (xgbTree)	
Metabolite	**Score**	**Metabolite**	**Score**	**Metabolite**	**Score**
Tyr	100	SM(38:3)	100	Tyr	100
SM(38:3)	85.872	Tyr	85.905	SM(38:3)	71.746
Glu	78.406	xLeu	73.956	Gly	53.489
Gly	73.254	Val	73.506	Glu	44.524
PC:O(30:0)	63.572	SM(36:0)	63.616	Val	26.3
Val	54.924	Gly	48.47	PC:O(30:0)	21.545
AC(5:0)	53.196	Ile	43.735	xLeu	21.016
LPC(18:2)	52.006	LPC(18:2)	40.95	TG(52:7)	20.137
TG(50:1)	41.624	TG(50:1)	37.987	PC(46:2)	18.839
SM(36:0)	35.029	Glu	32.414	Lys	16.355
PC(46:2)	34.573	PC(46:2)	28.593	LPC(18:2)	13.294
PC(44:1)	34.356	PC:O(30:0)	23.944	SM(36:0)	10.795
DG(34:1)	31.008	DG(34:1)	23.745	TG(50:1)	7.421
Ile	30.839	TG(54:3)	16.24	Ile	6.543
xLeu	23.012	PC(44:1)	15.903	PC(44:1)	4.163
TG(52:7)	15.234	AC(5:0)	10.108	DG(34:1)	1.194
TG(54:3)	7.187	Lys	6.875	TG(54:3)	1.164
Lys	0	TG(52:7)	0	AC(5:0)	0

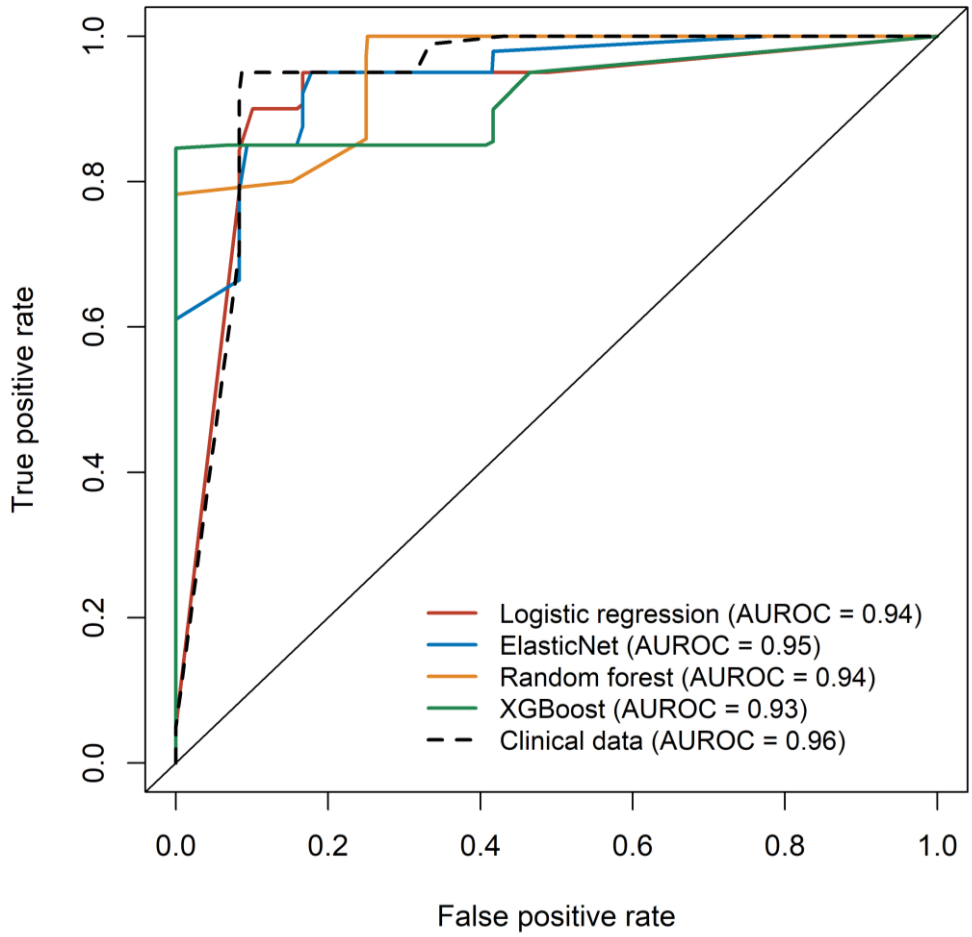


Figure 14. The receiver operating characteristic (ROC) curves of the NAFLD diagnostic models developed by machine learning approaches

The solid black line represents the line of unity. The area under ROC curve (AUROC) values present median values obtained from 100 repeated runs. All four diagnostic models using NAFLD-specific metabolic biomarkers demonstrated excellent predictive performances, comparable to the clinical data model.

5.3. Relevance of intervention-related metabolites in Metabolic Pathways

I performed metabolite set enrichment analysis with the significant metabolites between baseline and 18 months post-intervention to clarify which metabolic pathways are altered by long-term interventions (**Figure 15** and **Table 9**). D-glutamine and D-glutamate metabolism and arginine biosynthesis were significantly modified (enrichment ratio > 9.0, FDR-adjusted $p < 0.05$), and other metabolite sets, including alanine, aspartate, and glutamate metabolism, tricarboxylic acid (TCA) cycle, and valine, leucine, and isoleucine biosynthesis, were also enriched (enrichment ratio > 4.5, raw $p < 0.05$). I also constructed chemical and biochemical networks using these significant metabolites (**Figure 16**).

Compared to metabolic changes at 6 months post-intervention (**Figure 16A**), carnitines and many organic acids, including o-acetyl carnitine, octanoylcarnitine, azelaic acid, hydroxyoctanoic acid, and alpha-ketooctanoic acid were upregulated whereas galactaric acid, threonic acid, caproic acid, and alpha-ketoisovaleric acid levels, were downregulated at 18 months post-intervention (**Figure 16B**). TCA cycle intermediates, such as succinic acid, oxoglutaric acid, isocitric acid, malic acid, and phosphocholines, were also significantly downregulated at 18 months post-intervention. Levels of metabolites related to the methionine–cysteine pathway (cystine, methionine, S-

methylcysteine, and methionine sulfoxide), urea-cycle-related metabolites, and amino acids (arginine, ornithine, N-acetylornithine, lysine, N-acetyllysine, 5-oxoproline, glutamine, and glutamic acid) were also significantly changed.

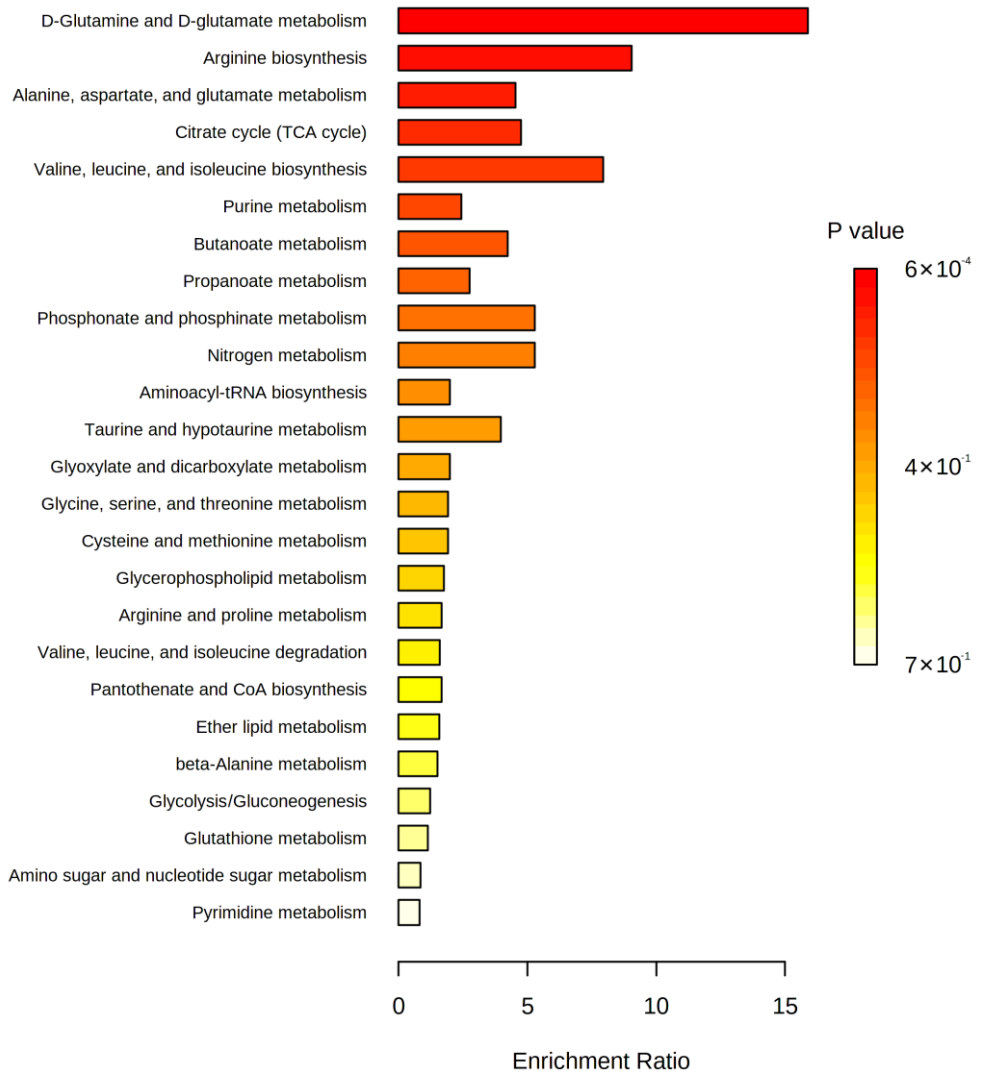
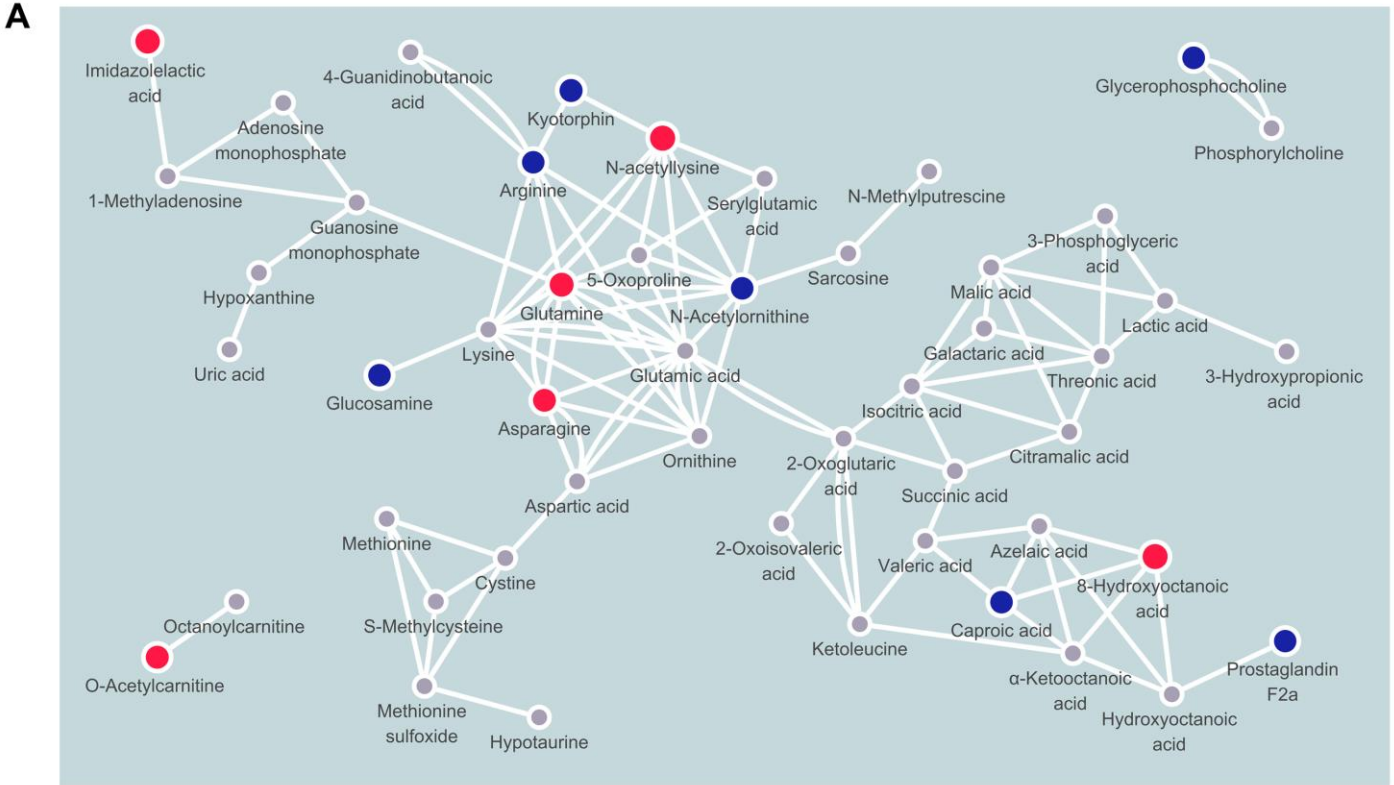


Figure 15. Enriched metabolite sets of significant metabolites between baseline and 18 months post-intervention based on the KEGG database

Table 9. Metabolite set enrichment analysis of significantly changed metabolites by intervention after 18 months based on the SMPDB (hits $\geq$ 2)

Metabolite set	total	expected	hits	Raw *p*	FDR
Aspartate Metabolism	35	1.64	5	0.0205	1
Urea Cycle	29	1.36	4	0.0425	1
Ammonia Recycling	32	1.5	4	0.0581	1
Phenylacetate Metabolism	9	0.422	2	0.0628	1
Glutamate Metabolism	49	2.3	5	0.0737	1
Malate-Aspartate Shuttle	10	0.469	2	0.0762	1
Carnitine Synthesis	22	1.03	3	0.0796	1
Oxidation of Branched Chain Fatty Acids	26	1.22	3	0.118	1
Mitochondrial Beta-Oxidation of Short Chain Saturated Fatty Acids	27	1.27	3	0.129	1
Warburg Effect	58	2.72	5	0.129	1
Glycine and Serine Metabolism	59	2.77	5	0.136	1
Citric Acid Cycle	32	1.5	3	0.186	1
Beta Oxidation of Very Long Chain Fatty Acids	17	0.797	2	0.187	1
Alanine Metabolism	17	0.797	2	0.187	1
Butyrate Metabolism	19	0.891	2	0.222	1
Gluconeogenesis	35	1.64	3	0.223	1
Arginine and Proline Metabolism	53	2.48	4	0.233	1
Purine Metabolism	74	3.47	5	0.262	1
Valine, Leucine and Isoleucine Degradation	60	2.81	4	0.308	1
Propanoate Metabolism	42	1.97	3	0.314	1
Methionine Metabolism	43	2.02	3	0.327	1

Cysteine Metabolism	26	1.22	2	0.347	1
Phytanic Acid Peroxisomal Oxidation	26	1.22	2	0.347	1
Phenylalanine and Tyrosine Metabolism	28	1.31	2	0.382	1
Lysine Degradation	30	1.41	2	0.416	1
Amino Sugar Metabolism	33	1.55	2	0.465	1
Nicotinate and Nicotinamide Metabolism	37	1.73	2	0.527	1


A
Imidazolelactic acid
4-Guanidinobutanoic acid
Kyotorphin
Glycerophosphocholine
Phosphorylcholine
Adenosine monophosphate
N-acetyllysine
1-Methyladenosine
Arginine
Serylglutamic acid
N-Methylputrescine
Guanosine monophosphate
3-Phosphoglyceric acid
5-Oxoproline
Sarcosine
Malic acid
Hypoxanthine
Glutamine
N-Acetylornithine
Lactic acid
Lysine
Galactaric acid
Uric acid
Glucosamine
Glutamic acid
Threonic acid
3-Hydroxypropionic acid
Isocitric acid
Asparagine
Citramalic acid
Ornithine
2-Oxoglutaric acid
Succinic acid
Aspartic acid
Methionine
2-Oxoisovaleric acid
Azelaic acid
Valeric acid
Cystine
8-Hydroxyoctanoic acid
Octanoylcarnitine
S-Methylcysteine
Caproic acid
Ketoleucine
Prostaglandin F2a
O-Acetylcarnitine
α-Ketooctanoic acid
Methionine sulfoxide
Hydroxyoctanoic acid
Hypotaurine

(continued)

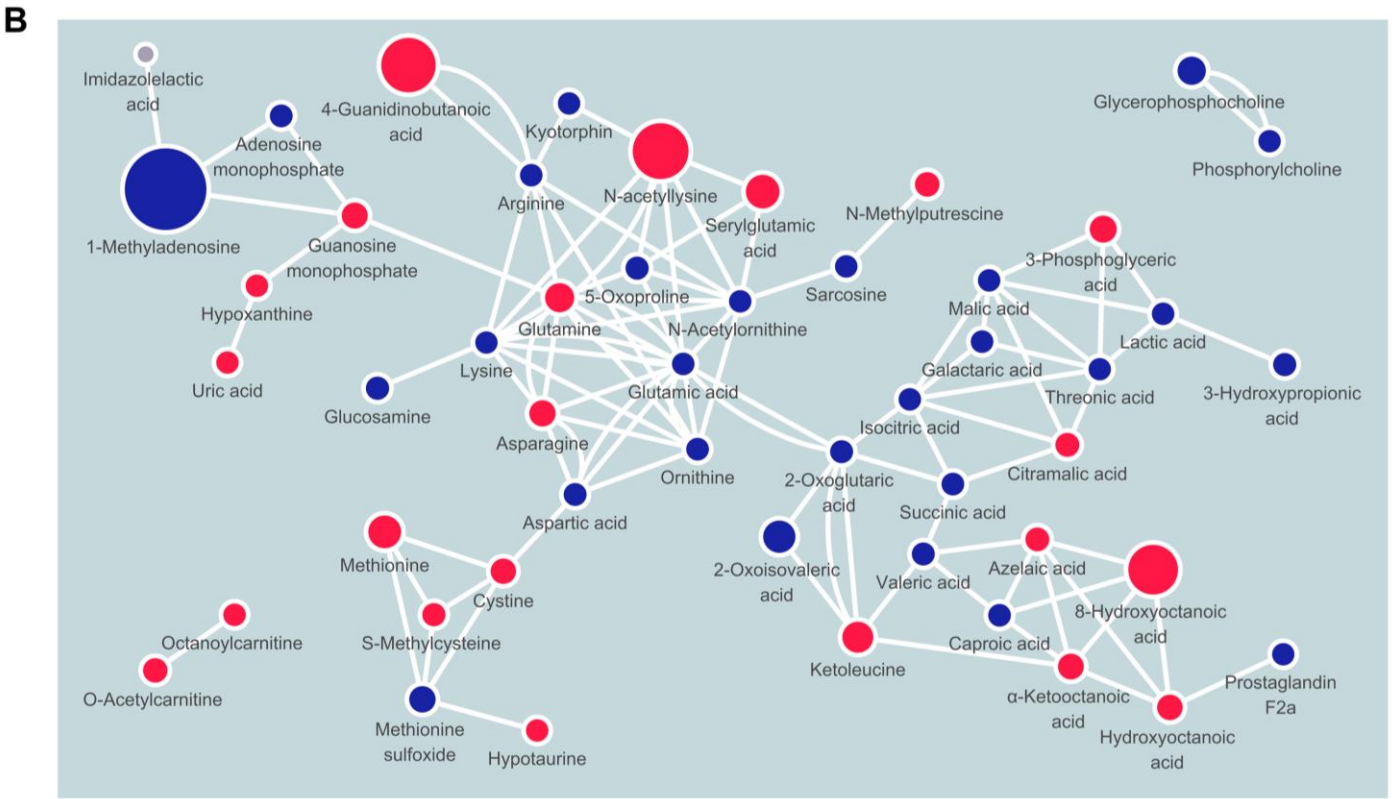


Figure 16. Relevance of the weight loss intervention-related metabolites in the metabolic pathways

Network mapping of significantly changed metabolic pathways from (A) 6 months and (B) 18 months post-intervention compared to baseline. Node color indicates the direction of change (red = up; blue = down) or non-significance (gray). Node size indicates each metabolite’s fold change.

Discussion

1. Metabolic pathway alterations reflecting NAFLD pathophysiology or weight loss interventions

In the pediatric NAFLD cohort study, plasma metabolomic data revealed that circulating metabolite levels related to glutathione-related amino acid metabolism, lipid metabolism, and BCAA metabolism were remarkably altered in the diseased state. Based on these results, the potential effects of altered metabolism on NAFLD in pediatric patients are proposed in **Figure 17**.

Glutathione metabolism-related metabolite levels, including those of glutamate and glycine, were significantly changed in the blood and may be important therapeutic targets, as excessive ROS-induced oxidative stress affects the progression of NAFLD. Increased circulating glutamic acid and decreased glycine levels in NAFLD have been consistently reported in both pediatric (38) and adult populations (12, 39-41). However, the causes for their changes remain unclear. Interestingly, Oren et al. suggested that impaired glycine metabolism might play a causative role in NAFLD (42). Several clinical studies have evaluated the effect of glycine supplementation (43-45), as glycine is a limiting substrate in the de novo synthesis of endogenous glutathione (41), which may have therapeutic potential.

White et al. observed that the impairment of BCAA metabolism in obesity can also affect the decreased level of circulating glycine (46).

The elevated blood levels of DG and TG in patients with NAFLD shown in this study may also be associated with increased oxidative stress and lipid peroxidation in hepatocytes. Once circulating DGs and TGs are transferred by hepatic uptake, they can accumulate as lipid droplets or convert into free fatty acids in the liver. Mitochondrial oxidation of excessive hepatic free fatty acids induced oxidative stress and lipid peroxidation, resulting in hepatocellular apoptosis (47, 48), which was reflected in the markedly increased serum AST and ALT levels of the NAFLD groups in this study. Lipid accumulation and excessive oxidative stress are directly associated with hepatic cell damage and prohibit hepatic uptake or lipid accumulation, including TGs and DGs in hepatocytes which are crucial therapeutic targets.

Similarly, modifications in sphingolipid and phospholipid metabolism are also associated with metabolic disease and NAFLD (49-52). In young adults with obesity, serum SMs with saturated acyl chains are reportedly associated with obesity, insulin resistance, and decreased liver function (53). Some lipidomic studies have suggested that the plasma PC/PE ratio is associated with obesity (54). However, the mechanisms linking SMs, PCs, and LPCs with liver steatosis and NAFLD are unclear (51) and there are inconsistencies in the level and pattern of the reported sphingolipids and phospholipids in both pediatric and adult NAFLD patients.

The concentration of systemic BCAAs is altered in various metabolic diseases such as diabetes, insulin resistance, and obesity, which are well-known etiologic factors in NAFLD (55-57). This study observed elevated circulating BCAAs levels in the pediatric population with overweight NAFLD, which has been reported by several in vitro and in vivo studies in pediatric (39) and adult populations (58-61). In addition to serving as substrates for protein synthesis and energy production, BCAAs also stimulate protein synthesis, inhibit proteolysis, and affect glucose metabolism and oxidative stress (55, 62), indicating that the homeostatic regulation of BCAA levels is crucial to maintaining physiological status. Excessive systemic levels of BCAAs can increase abnormal adipocyte lipolysis and suppress hepatic lipogenesis, resulting in hyperlipidemia and hepatic lipotoxicity (59), which is also supported by increased blood DGs and TGs in the overweight NAFLD group in this study. BCAA-based metabolic signatures may predict liver steatosis and NAFLD in children and adolescents with obesity (13, 63).

In the obesity intervention cohort study, significant metabolic pathway changes by the time in weight loss intervention in childhood obesity were identified, regardless of the intervention type or response (**Figure 18**). The metabolic pathways are included in the urea and TCA cycles and several amino acid metabolic pathways (i.e., glutamine, glutamate, arginine, cysteine, and methionine). Several studies have shown that glutamate is increased in children with

obesity and glutamine is decreased in patients with obesity not undergoing weight loss interventions (16, 64). These previous results are consistent with this study's results that plasma glutamine was increased after a weight loss intervention, whereas glutamate was decreased. In particular, glutamate was the metabolite with the highest bivariate correlation with body fat and insulin sensitivity in an American-Indian adolescent with obesity (65).

Obesity is due to an inflammation mechanism that impacts various immune cells. In patients with obesity and diabetes, changes in macrophage polarization can induce decreased ketoglutarate production and increased succinate in the TCA cycle. Glutamine metabolism is essential for this process (66). Succinate, an intermediate product of the TCA cycle, is released from adipose tissue in obesity and helps induce inflammation by macrophage activation (67). Consequently, macrophage polarization involving glutamine metabolism and the TCA cycle in patients with obesity and diabetes might play a key role in obesity pathology. This study showed that levels of TCA intermediates, including isocitrate, malate, oxoglutarate (ketoglutarate), and succinate were decreased by the weight loss intervention. This implies the interventions' anti-inflammatory effects by minimizing macrophage polarizations.

Arginine, aspartate, and ornithine are associated with urea cycle metabolism and play an important role in ammonia detoxification. This study showed decreased urea cycle intermediates, including

aspartate, ornithine, and arginine, in the plasma, which implies downregulated production of ammonia by the weight loss intervention. Aspartate, like pyruvic acid, is an amino acid associated with the TCA cycle. The metabolic shift in pyruvic acid decreased after weight control intervention in overweight pre-adolescents and women with obesity (68, 69). From the POUND LOST and DIRECT trials, Zheng et al. confirmed that a two-year diet-based weight loss intervention induced amino acid profile changes in adult patients with obesity (70). These studies showed correlation between weight loss intervention and simultaneous reduction of BCAAs, aromatic amino acids (tyrosine and phenylalanine), and other amino acids (alanine, sarcosine, hydroxyproline, and methionine) (70). Glutamine and glutamate were positively correlated with a HOMA-IR in a study of overweight-to-obese adult with dyslipidemia (71).

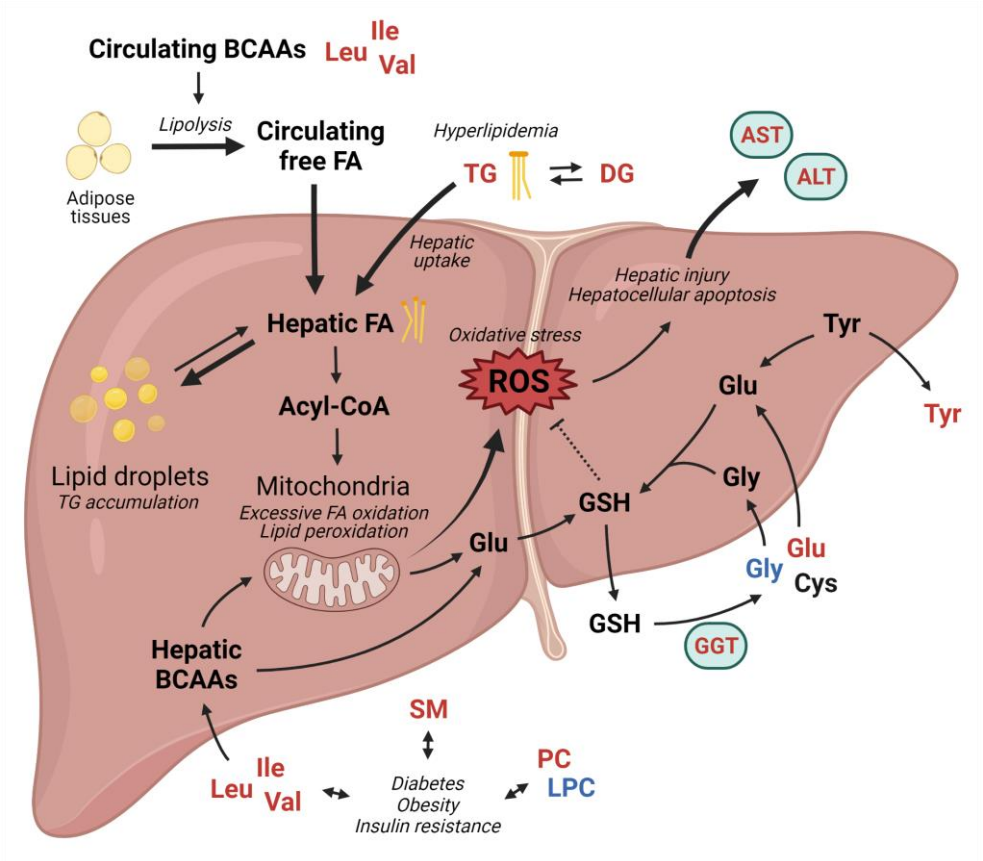


Figure 17. Metabolic pathway alterations reflecting the pathophysiology of NAFLD in overweight pediatric patients

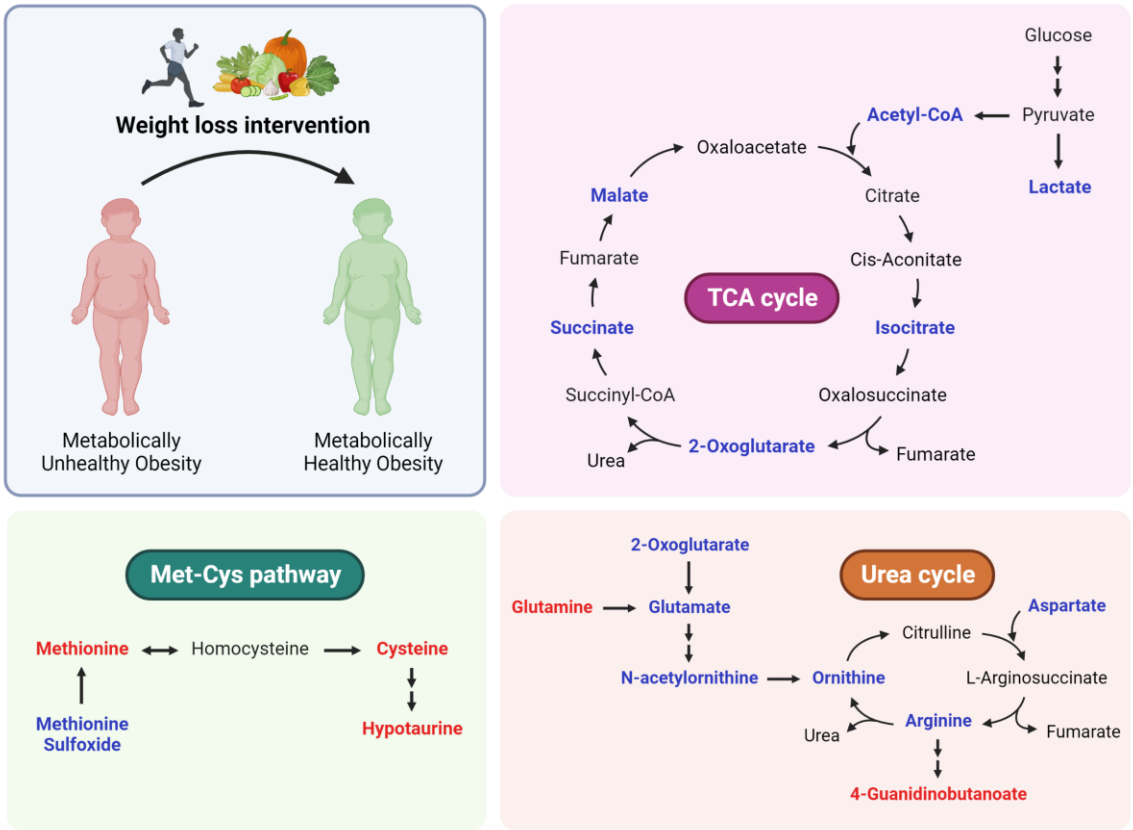


Figure 18. Metabolic pathway changes by weight loss intervention

2. Clinical implications of two cohort studies

First, this pediatric NAFLD cohort study investigated the metabolomic distributions of the HC, LN, OC, and ON groups and found that both the LN and ON groups showed metabolic heterogeneity, which may be closely related to the complex pathophysiology of NAFLD. Metabolic changes were also identified to elucidate the characteristics and mechanisms of pediatric NAFLD and NAFLD-specific metabolic biomarkers by comparing the metabolomic signatures of the overweight and normal-weight groups. Additionally, I revealed that these metabolic changes, which can be observed in the adult population, emerged during the adolescent period. It is difficult to evaluate whether the significant metabolites are NAFLD-specific, as most of the significant metabolites between the OC and ON groups, including BCAAs, are related to insulin resistance and obesity. Despite partially missing data, such as fasting insulin and HOMA-IR levels, these metabolites are NAFLD-specific, regardless of insulin resistance and obesity.

Using the NAFLD-specific metabolic biomarkers, this study successfully suggested cross-validated diagnostic models based on pathophysiology that might be easily applied in clinical practice. These models performed better than other diagnostic models based on metabolomics (72-74). These encouraging results demonstrate that a NAFLD diagnosis with only a small volume of plasma may

ameliorate the limitations of a liver biopsy and allow for the high-throughput pediatric NAFLD screening, even in schools.

This study also found that clinical features, such as the BMI z-score and liver function test results (especially ALT), and metabolic features were directly associated with NAFLD development in pediatric patients, which is reflected in the outstanding machine learning model performance using clinical and genetic variables. The fact that the increased level of liver function tests in adults resulted from complex interactions between the disease pathophysiology and external factors, such as smoking, alcohol, and stress, whereas those in children mainly resulted from hepatic inflammation itself, can be one of the explanations why the models using clinical and genetic variables developed in this pediatric population study show excellent performance, comparing with other diagnostic models that were previously reported in the adult population (74-76).

Metabolic changes from weight loss intervention might become metabolically healthy (MHO) and metabolically unhealthy obesity (MUO) phenotypic characteristics. Though the World Health Organization defines obesity as “abnormal or excessive fat accumulation that may impair health” (77), some patients, often young and physically active with a good nutritional status, show a healthier phenotype than others (78). There is no universally accepted MHO concept yet; however, MHO can reinforce the existing BMI-based single definition by reflecting obesity characteristics in various

aspects. Many clinical and behavioral MHO characteristics, such as anthropometric and clinical parameters, lifestyle factors, and comorbidities, have been suggested: fasted serum TGs and HDL-C levels, fasting blood glucose, blood pressure, no drug treatment for dyslipidemia, diabetes, or hypertension (79, 80). In contrast, only a few studies have focused on the metabolomic signatures between MHO and MUO (81). Dihe Cheng et al.'s review explains that BCAAs, aromatic amino acids, long-chain fatty acids, and propionyl carnitine levels might be reduced in MHO (81). Metabolic changes observed in this study, including TCA cycle, urea cycle, and amino acid and organic acid alteration, suggest that weight loss intervention can ameliorate pediatric patients toward a metabolically healthy status. These changes can be additional MHO signatures.

Regardless of clinical changes, metabolic change from weight loss intervention might become an efficient monitoring biomarker for participants' compliance of future obesity intervention studies. Although participants' compliance is an essential factor affecting intervention effects, ways to monitor and evaluate their compliance are limited. The intervention-related biomarkers presented in this study are more sensitive to weight loss interventions than clinical parameters such as BMI; therefore, they can be utilized as more effective biomarkers for evaluating and selecting participants with poor compliance.

3. Study limitations

The pediatric NAFLD study has some limitations. Although liver biopsies are deemed the gold standard for diagnosing liver steatosis and fibrosis, obtaining them in this study cohort was challenging. Due to the invasiveness and effectiveness of biopsies and the ethical considerations related to performing biopsies in pediatric patients, this study examined steatosis grades using hepatic ultrasonography instead of liver biopsies. The targeted metabolomics used in this study excluded exogenous compounds and unwanted analytical noise. However, this may only partially explain the issue in comparison with untargeted approaches, which implies there may be unrevealed metabolic signatures.

In addition, the diagnostic models based on machine learning approaches in this study showed excellent performance. Further evaluation of the models using external validation cohorts is required. Although the changes in the levels of circulating plasma metabolites directly reflect metabolic changes in the liver, the lack of observed metabolic changes in the hepatocytes may have influenced the interpretation of the results. For example, the antioxidant effects of significant metabolic markers, or free fatty acid accumulation in the liver, could not be confirmed in this study. Lastly, NAFLD was already progressing in these subjects, making it difficult to assess whether the differences in the metabolite levels were a reflection of their etiologic

roles, or if these were a consequence of the disease, as this causality problem is a common limitation of observational studies. Although the associations between NAFLD and plasma metabolites were observable in this study, further research including in vitro/in vivo studies using cell lines or model organisms is required to confirm the disease mechanism-related functions of selected metabolites and the direction of causality.

The obesity intervention study was limited by the availability of subjects for longitudinal studies owing to many dropouts from the childhood obesity intervention cohort. In addition, there was a possibility of selection bias in the results. Lower family functioning, exercise group, lower initial attendance rate, and non-self-referral pathways were significantly associated with early dropouts, and lower family functioning and lower initial attendance rates were associated with late dropouts in our cohort study (82). These factors are key to reducing the dropout rate in further intervention-based cohort research. Poor family function was associated with high levels of depressive symptoms in the childhood obesity cohort (83). Further research is needed to determine how these factors affect the intervention outcomes in children.

Also, this study did not analyze the factors related to adolescent age in the long-term follow-up during the intervention. In the heatmap, metabolic changes were most pronounced after 18 months, but the probability that hormonal changes and other factors affected the

metabolites was not considered. Another limitation is that the patients' hospital visit times and participation levels were not considered in the study. Factors such as diet could not be limited to homogeneity, which may have influenced the metabolite results. In particular, metabolites are affected by sampling time or what is consumed at the time of sampling. It is important to control these factors and follow the metabolite trends.

Conclusion

I explored the metabolic profiles of pediatric subjects with NAFLD and found NAFLD-specific metabolic biomarkers. These biomarkers demonstrated that dysregulated glutathione, lipid, and BCAA metabolism were linked to the pathophysiological conditions underlying NAFLD. Despite the restricted sample accessibility, these findings provide evidence for the pathophysiology of pediatric NAFLD that may guide potential therapeutic targets for new drug development and as ancillary diagnostic biomarkers to help alleviate NAFLD in pediatric patients.

In addition, regardless of intervention response, this study determined weight loss intervention-related metabolic biomarkers for pediatric obesity and the intervention-induced serial changes in metabolic pathways. This includes the urea cycle, TCA cycle, and specific amino acid metabolism, which may influence other obesity mechanisms. These findings can redefine obesity from a metabolic perspective (i.e., metabolically healthy-unhealthy obesity), to manage pediatric obesity based on its pathophysiology. This also highlights the importance of regularly visiting a clinic for obesity treatment and how tracking metabolite changes guides disease monitoring.

There are unmet medical needs for the pediatric population with obesity and other medical complications, including NAFLD and

weight loss intervention treatments. This dissertation demonstrates that plasma metabolomics succeeds as a powerful biomarker discovery tool for fulfilling current unmet medical needs and can be efficiently used to identify pathophysiology, improve diagnosis, and establish a new understanding of obesity and obesity-related mechanism (**Figure 19**).

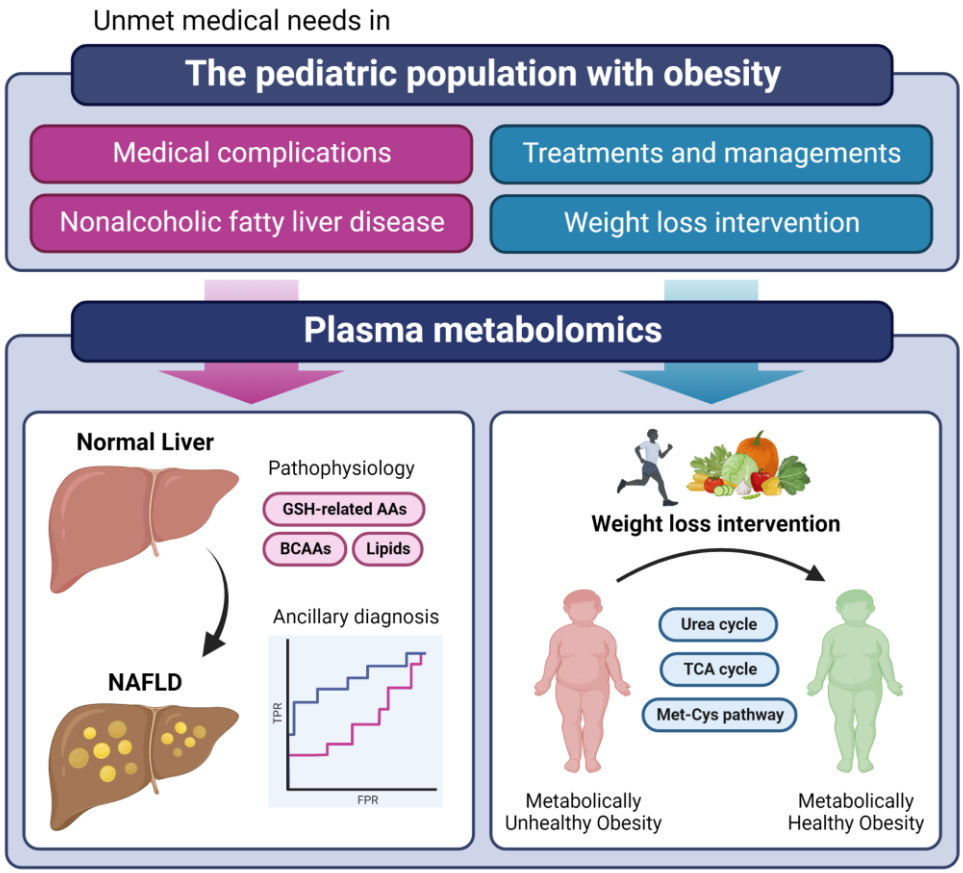


Figure 19. Summary of the dissertation

References

1. Abarca-Gómez L, Abdeen ZA, Hamid ZA, Abu-Rmeileh NM, Acosta-Cazares B, Acuin C, et al. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. The Lancet. 2017;390(10113):2627-42.
2. Organization WH. Taking action on childhood obesity. World Health Organization; 2018.
3. Yi DY, Kim SC, Lee JH, Lee EH, Kim JY, Kim YJ, et al. Clinical Practice Guideline for the Diagnosis and Treatment of Pediatric Obesity: Recommendations from the Committee on Pediatric Obesity of the Korean Society of Pediatric Gastroenterology Hepatology and Nutrition. Pediatr Gastroenterol Hepatol Nutr. 2019;22(1):1-27.
4. Lange SJ, Kompaniyets L, Freedman DS, Kraus EM, Porter R, Dnp, et al. Longitudinal Trends in Body Mass Index Before and During the COVID-19 Pandemic Among Persons Aged 2-19 Years - United States, 2018-2020. MMWR Morb Mortal Wkly Rep. 2021;70(37):1278-83.
5. De Spiegeleer M, De Paepe E, Van Meulebroek L, Gies I, De Schepper J, Vanhaecke L. Paediatric obesity: a systematic review and pathway mapping of metabolic alterations underlying early disease processes. Mol Med. 2021;27(1):145.
6. Moran JR, Ghishan FK, Halter SA, Greene HL. Steatohepatitis in obese children: a cause of chronic liver dysfunction. Am J Gastroenterol. 1983;78(6):374-7.

7. Nobili V, Alisi A, Valenti L, Miele L, Feldstein AE, Alkhouri N. NAFLD in children: new genes, new diagnostic modalities and new drugs. Nat Rev Gastroenterol Hepatol. 2019;16(9):517-30.
8. Buzzetti E, Pinzani M, Tsochatzis EA. The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). Metabolism. 2016;65(8):1038-48.
9. Han JC, Lawlor DA, Kimm SYS. Childhood obesity. The Lancet. 2010;375(9727):1737-48.
10. Gibney MJ, Walsh M, Brennan L, Roche HM, German B, van Ommen B. Metabolomics in human nutrition: opportunities and challenges. Am J Clin Nutr. 2005;82(3):497-503.
11. Rangel-Huerta OD, Pastor-Villaescusa B, Gil A. Are we close to defining a metabolomic signature of human obesity? A systematic review of metabolomics studies. Metabolomics. 2019;15(6):93.
12. Gaggini M, Carli F, Rosso C, Buzzigoli E, Marietti M, Della Latta V, et al. Altered amino acid concentrations in NAFLD: Impact of obesity and insulin resistance. Hepatology. 2018;67(1):145-58.
13. Goffredo M, Santoro N, Trico D, Giannini C, D'Adamo E, Zhao H, et al. A Branched-Chain Amino Acid-Related Metabolic Signature Characterizes Obese Adolescents with Non-Alcoholic Fatty Liver Disease. Nutrients. 2017;9(7).
14. Puri P, Baillie RA, Wiest MM, Mirshahi F, Choudhury J, Cheung O, et al. A lipidomic analysis of nonalcoholic fatty liver disease. Hepatology. 2007;46(4):1081-90.
15. Sorrow P, Maguire R, Murphy SK, Belcher SM, Hoyo C. Elevated metabolites of acetaminophen in cord blood of children with obesity. Pediatr Obes. 2019;14(1).

16. Butte NF, Liu Y, Zakeri IF, Mohney RP, Mehta N, Voruganti VS, et al. Global metabolomic profiling targeting childhood obesity in the Hispanic population. Am J Clin Nutr. 2015;102(2):256-67.
17. Wahl S, Yu Z, Kleber M, Singmann P, Holzapfel C, He Y, et al. Childhood obesity is associated with changes in the serum metabolite profile. Obes Facts. 2012;5(5):660-70.
18. Gawlik A, Shmoish M, Hartmann MF, Malecka-Tendera E, Wudy SA, Hochberg Z. Steroid Metabolomic Disease Signature of Nonsyndromic Childhood Obesity. J Clin Endocrinol Metab. 2016;101(11):4329-37.
19. Cho K, Moon JS, Kang JH, Jang HB, Lee HJ, Park SI, et al. Combined untargeted and targeted metabolomic profiling reveals urinary biomarkers for discriminating obese from normal-weight adolescents. Pediatr Obes. 2017;12(2):93-101.
20. Charatcharoenwitthaya P, Lindor KD. Role of radiologic modalities in the management of non-alcoholic steatohepatitis. Clin Liver Dis. 2007;11(1):37-54, viii.
21. Ferraioli G, Soares Monteiro LB. Ultrasound-based techniques for the diagnosis of liver steatosis. World J Gastroenterol. 2019;25(40):6053-62.
22. Kim JH, Yun S, Hwang SS, Shim JO, Chae HW, Lee YJ, et al. The 2017 Korean National Growth Charts for children and adolescents: development, improvement, and prospects. Korean J Pediatr. 2018;61(5):135-49.
23. Moon JS, Lee SY, Nam CM, Choi J-M, Choe B-K, Seo J-W, et al. 2007 Korean National Growth Charts: review of developmental process and an outlook. Korean Journal of Pediatrics. 2008;51(1).
24. Seo YG, Kim JH, Kim Y, Lim H, Ju YS, Kang MJ, et al. Validation of body composition using bioelectrical impedance analysis in

children according to the degree of obesity. Scand J Med Sci Sports. 2018;28(10):2207-15.
25. Seo YG, Lim H, Kim Y, Ju YS, Lee HJ, Jang HB, et al. The Effect of a Multidisciplinary Lifestyle Intervention on Obesity Status, Body Composition, Physical Fitness, and Cardiometabolic Risk Markers in Children and Adolescents with Obesity. Nutrients. 2019;11(1).
26. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 2021;49(W1):W388-W96.
27. Benjamini Y, Krieger AM, Yekutieli D. Adaptive Linear Step-up Procedures That Control the False Discovery Rate. Biometrika. 2006;93(3):491-507.
28. Barupal DK, Haldiya PK, Wohlgemuth G, Kind T, Kothari SL, Pinkerton KE, et al. MetaMapp: mapping and visualizing metabolomic data by integrating information from biochemical pathways and chemical and mass spectral similarity. BMC Bioinformatics. 2012;13:99.
29. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-504.
30. Dinani AM, Kowdley KV, Noureddin M. Application of Artificial Intelligence for Diagnosis and Risk Stratification in NAFLD and NASH: The State of the Art. Hepatology. 2021;74(4):2233-40.
31. Zou H, Hastie T. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2005;67(2):301-20.

32. Breiman RF. Vaccines as tools for advancing more than public health: perspectives of a former director of the National Vaccine Program office. Clin Infect Dis. 2001;32(2):283-8.
33. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. Xgboost: extreme gradient boosting. R package version 04-2. 2015;1(4):1-4.
34. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2021.
35. Kuhn M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software. 2008;28(5).
36. Park SH, Park HY, Kang JW, Park J, Shin KJ. Aminotransferase upper reference limits and the prevalence of elevated aminotransferases in the Korean adolescent population. J Pediatr Gastroenterol Nutr. 2012;55(6):668-72.
37. Chae W, Lee KJ, Huh KY, Moon JS, Ko JS, Cho JY. Association of Metabolic Signatures with Nonalcoholic Fatty Liver Disease in Pediatric Population. Metabolites. 2022;12(9).
38. Jin R, Banton S, Tran VT, Konomi JV, Li S, Jones DP, et al. Amino Acid Metabolism is Altered in Adolescents with Nonalcoholic Fatty Liver Disease-An Untargeted, High Resolution Metabolomics Study. J Pediatr. 2016;172:14-9 e5.
39. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab. 2009;9(4):311-26.
40. Bianchi G. Impaired insulin-mediated amino acid plasma disappearance in non-alcoholic fatty liver disease: a feature of insulin resistance. Digestive and Liver Disease. 2003;35(10):722-7.
41. Mardinoglu A, Bjornson E, Zhang C, Klevstig M, Soderlund S, Stahlman M, et al. Personal model-assisted identification of

NAD(+) and glutathione metabolism as intervention target in NAFLD. Mol Syst Biol. 2017;13(3):916.

42. Rom O, Liu Y, Liu Z, Zhao Y, Wu J, Ghrayeb A, et al. Glycine-based treatment ameliorates NAFLD by modulating fatty acid oxidation, glutathione synthesis, and the gut microbiome. Sci Transl Med. 2020;12(572).
43. Gonzalez-Ortiz M, Medina-Santillan R, Martinez-Abundis E, von Drateln CR. Effect of glycine on insulin secretion and action in healthy first-degree relatives of type 2 diabetes mellitus patients. Horm Metab Res. 2001;33(6):358-60.
44. Nguyen D, Hsu JW, Jahoor F, Sekhar RV. Effect of increasing glutathione with cysteine and glycine supplementation on mitochondrial fuel oxidation, insulin sensitivity, and body composition in older HIV-infected patients. J Clin Endocrinol Metab. 2014;99(1):169-77.
45. Diaz-Flores M, Cruz M, Duran-Reyes G, Munguia-Miranda C, Loza-Rodriguez H, Pulido-Casas E, et al. Oral supplementation with glycine reduces oxidative stress in patients with metabolic syndrome, improving their systolic blood pressure. Can J Physiol Pharmacol. 2013;91(10):855-60.
46. White PJ, Lapworth AL, An J, Wang L, McGarrah RW, Stevens RD, et al. Branched-chain amino acid restriction in Zucker-fatty rats improves muscle insulin sensitivity by enhancing efficiency of fatty acid oxidation and acyl-glycine export. Mol Metab. 2016;5(7):538-51.
47. Luedde T, Kaplowitz N, Schwabe RF. Cell death and cell death responses in liver disease: mechanisms and clinical relevance. Gastroenterology. 2014;147(4):765-83 e4.
48. Fuchs M, Sanyal AJ. Lipotoxicity in NASH. J Hepatol. 2012;56(1):291-3.

49. Tiwari-Heckler S, Gan-Schreier H, Stremmel W, Chamulitrat W, Pathil A. Circulating Phospholipid Patterns in NAFLD Patients Associated with a Combination of Metabolic Risk Factors. Nutrients. 2018;10(5).
50. Regnier M, Polizzi A, Guillou H, Loiseau N. Sphingolipid metabolism in non-alcoholic fatty liver diseases. Biochimie. 2019;159:9-22.
51. Meikle PJ, Summers SA. Sphingolipids and phospholipids in insulin resistance and related metabolic disorders. Nat Rev Endocrinol. 2017;13(2):79-91.
52. Holland WL, Summers SA. Sphingolipids, insulin resistance, and metabolic disease: new insights from in vivo manipulation of sphingolipid metabolism. Endocr Rev. 2008;29(4):381-402.
53. Hanamatsu H, Ohnishi S, Sakai S, Yuyama K, Mitsutake S, Takeda H, et al. Altered levels of serum sphingomyelin and ceramide containing distinct acyl chains in young obese adults. Nutr Diabetes. 2014;4:e141.
54. Weir JM, Wong G, Barlow CK, Greeve MA, Kowalczyk A, Almasy L, et al. Plasma lipid profiling in a large population-based cohort. J Lipid Res. 2013;54(10):2898-908.
55. Holecek M. Branched-chain amino acids in health and disease: metabolism, alterations in blood plasma, and as supplements. Nutr Metab (Lond). 2018;15:33.
56. Lynch CJ, Adams SH. Branched-chain amino acids in metabolic signalling and insulin resistance. Nat Rev Endocrinol. 2014;10(12):723-36.
57. Zhao X, Han Q, Liu Y, Sun C, Gang X, Wang G. The Relationship between Branched-Chain Amino Acid Related Metabolomic Signature and Insulin Resistance: A Systematic Review. J Diabetes Res. 2016;2016:2794591.

58. Lake AD, Novak P, Shipkova P, Aranibar N, Robertson DG, Reily MD, et al. Branched chain amino acid metabolism profiles in progressive human nonalcoholic fatty liver disease. Amino Acids. 2015;47(3):603-15.
59. Zhang F, Zhao S, Yan W, Xia Y, Chen X, Wang W, et al. Branched Chain Amino Acids Cause Liver Injury in Obese/Diabetic Mice by Promoting Adipocyte Lipolysis and Inhibiting Hepatic Autophagy. EBioMedicine. 2016;13:157-67.
60. Sunny NE, Kalavalapalli S, Bril F, Garrett TJ, Nautiyal M, Mathew JT, et al. Cross-talk between branched-chain amino acids and hepatic mitochondria is compromised in nonalcoholic fatty liver disease. Am J Physiol Endocrinol Metab. 2015;309(4):E311-9.
61. Harris RA, Joshi M, Jeoung NH, Obayashi M. Overview of the molecular and biochemical basis of branched-chain amino acid catabolism. J Nutr. 2005;135(6 Suppl):1527S-30S.
62. Zhang S, Zeng X, Ren M, Mao X, Qiao S. Novel metabolic and physiological functions of branched chain amino acids: a review. J Anim Sci Biotechnol. 2017;8:10.
63. Lischka J, Schanzer A, Hojreh A, Ba Ssalamah A, Item CB, de Gier C, et al. A branched-chain amino acid-based metabolic score can predict liver fat in children and adolescents with severe obesity. Pediatr Obes. 2021;16(4):e12739.
64. McCormack SE, Shaham O, McCarthy MA, Deik AA, Wang TJ, Gerszten RE, et al. Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. Pediatr Obes. 2013;8(1):52-61.
65. Short KR, Chadwick JQ, Teague AM, Tullier MA, Wolbert L, Coleman C, et al. Effect of Obesity and Exercise Training on

Plasma Amino Acids and Amino Metabolites in American Indian Adolescents. J Clin Endocrinol Metab. 2019;104(8):3249-61.

66. Ren W, Xia Y, Chen S, Wu G, Bazer FW, Zhou B, et al. Glutamine Metabolism in Macrophages: A Novel Target for Obesity/Type 2 Diabetes. Adv Nutr. 2019;10(2):321-30.
67. van Diepen JA, Robben JH, Hooiveld GJ, Carmone C, Alsady M, Boutens L, et al. SUCNR1-mediated chemotaxis of macrophages aggravates obesity-induced inflammation and diabetes. Diabetologia. 2017;60(7):1304-13.
68. Campbell C, Grapov D, Fiehn O, Chandler CJ, Burnett DJ, Souza EC, et al. Improved metabolic health alters host metabolism in parallel with changes in systemic xeno-metabolites of gut origin. PLoS One. 2014;9(1):e84260.
69. Meucci M, Baldari C, Guidetti L, Alley JR, Cook C, Collier SR. Metabolomic Shifts Following Play-Based Activity in Overweight Preadolescents. Curr Pediatr Rev. 2017;13(2):144-51.
70. Zheng Y, Ceglarek U, Huang T, Li L, Rood J, Ryan DH, et al. Weight-loss diets and 2-y changes in circulating amino acids in 2 randomized intervention trials. Am J Clin Nutr. 2016;103(2):505-11.
71. Huffman KM, Shah SH, Stevens RD, Bain JR, Muehlbauer M, Slentz CA, et al. Relationships between circulating metabolic intermediates and insulin action in overweight to obese, inactive men and women. Diabetes Care. 2009;32(9):1678-83.
72. Khusial RD, Cioffi CE, Caltharp SA, Krasinskas AM, Alazraki A, Knight-Scott J, et al. Development of a Plasma Screening Panel for Pediatric Nonalcoholic Fatty Liver Disease Using Metabolomics. Hepatol Commun. 2019;3(10):1311-21.
73. Chashmniam S, Ghafourpour M, Rezaei Farimani A, Gholami A, Nobakht Motlagh Ghoochani BF. Metabolomic Biomarkers in the

Diagnosis of Non-Alcoholic Fatty Liver Disease. Hepatitis Monthly. 2019;19(9).

74. Mayo R, Crespo J, Martinez-Arranz I, Banales JM, Arias M, Minchole I, et al. Metabolomic-based noninvasive serum test to diagnose nonalcoholic steatohepatitis: Results from discovery and validation cohorts. Hepatol Commun. 2018;2(7):807-20.
75. Zhou Y, Oresic M, Leivonen M, Gopalacharyulu P, Hyysalo J, Arola J, et al. Noninvasive Detection of Nonalcoholic Steatohepatitis Using Clinical Markers and Circulating Levels of Lipids and Metabolites. Clin Gastroenterol Hepatol. 2016;14(10):1463-72 e6.
76. Barr J, Caballeria J, Martinez-Arranz I, Dominguez-Diez A, Alonso C, Muntane J, et al. Obesity-dependent metabolic signatures associated with nonalcoholic fatty liver disease progression. J Proteome Res. 2012;11(4):2521-32.
77. Obesity: World Health Organization; 2022 [Available from: https://www.who.int/health-topics/obesity#tab=tab_1.
78. Despres JP. Body fat distribution and risk of cardiovascular disease: an update. Circulation. 2012;126(10):1301-13.
79. Bluher M. Metabolically Healthy Obesity. Endocr Rev. 2020;41(3).
80. Iacobini C, Pugliese G, Blasetti Fantauzzi C, Federici M, Menini S. Metabolically healthy versus metabolically unhealthy obesity. Metabolism. 2019;92:51-60.
81. Cheng D, Zhao X, Yang S, Cui H, Wang G. Metabolomic Signature Between Metabolically Healthy Overweight/Obese and Metabolically Unhealthy Overweight/Obese: A Systematic Review. Diabetes Metab Syndr Obes. 2021;14:991-1010.
82. Park J, Woo S, Ju YS, Seo YG, Lim HJ, Kim YM, et al. Factors associated with dropout in a lifestyle modification program for

weight management in children and adolescents. Obes Res Clin Pract. 2020;14(6):566-72.

83. Noh H-M, Park J, Sung E-J, Ju YS, Lee H-J, Jeong Y-K, et al. Family Factors and Obesity in Relation to Mental Health Among Korean Children and Adolescents. Journal of Child and Family Studies. 2019;29(5):1284-92.

국문 초록

서 론: 소아청소년 비만은 유병률이 전 세계적으로 크게 증가하고 있으며, 국내 소아청소년의 비만 유병률 또한 2008년 기준 8.4%에서 2018년 기준 25%로 급격히 상승하고 있다. 소아청소년 비만은 인슐린 저항성, 당뇨 등 다양한 대사성 질환, 심혈관계질환 및 간질환과 밀접한 연관이 있으며, 특히 비알코올성지방간질환은 소아청소년에서 가장 흔한 간질환 중 하나로 소아청소년 비만의 증가와 비례하여 비알코올성지방간질환의 유병률 또한 증가하고 있다. 한편 소아청소년 집단에서 비만 및 관련 질환의 가장 효과적인 예방 및 치료방법은 규칙적인 신체활동과 식이요법 등의 체중감량 중재를 병행하는 것이다. 본 연구에서는 소아청소년 비만 환자에서 흔히 발생하는 비알코올성지방간질환의 유무에 따른 혈장 대사체 차이와 체중감량 중재에 따른 혈장 대사체 차이를 비교하고 병태생리를 규명하고자 하였다.

방 법: 소아청소년 비알코올성지방간질환 코호트 환자 165 명에 대해 지방증 등급과 체질량지수 z 점수를 바탕으로 4 개의 군으로 분류한 후 과체중 그룹에서 비알코올성지방간질환 유무에 따라 혈장 대사체 차이를 탐색하여 비알코올성지방간 특이적 대사체를 발굴하고자 하였다. 또한 소아청소년 체중감량 중재 코호트 환자 40 명에 대해 중재 전, 중재 후 6 개월, 중재 후 18 개월 시점의 혈장 대사체 차이를 비교하여 체중감량 중재 기간에 따른 혈장 대사체의 변화가 갖는 의미를 탐색하고자 하였다.

결 과: 소아청소년 비알코올성지방간질환 코호트 환자 연구를 통해 18개의 비알코올성지방간 특이적 대사체를 발굴하였다. 이 대사체들은 대사 경로

네트워크 분석을 통해 글루타티온 연관 대사, 지질 대사, 분지쇄 아미노산 대사와 관련이 있음을 보였으며, 비알코올성지방간 진단에도 보조적으로 활용할 수 있음을 제시하였다. 소아청소년 체중감량 중재 코호트 환자 연구를 통해 중재 반응 여부와 상관없이 중재 기간에 따른 시계열적 혈장 대사체 변화가 있음을 확인하였다. 또한 중재 기간이 길어질수록 더 많은 종류의 대사체가 변화하며, 대사 경로 네트워크 분석을 통해 시트르산 회로와 요소 회로, 아미노산 관련 대사 변화가 두드러지게 나타남을 보였다.

결 론: 두 개의 소아청소년 코호트와 대사체학 방법론을 활용하여 소아청소년 비만과 관련된 비알코올성지방간질환과 체중감량 중재에 따른 혈장 대사체 변화가 대사 경로 상에서 어떠한 의미가 있는지 밝혔다. 본 연구 결과를 바탕으로 비알코올성지방간질환의 병태생리와 대사적으로 건강한 비만에 대한 개념을 정립할 수 있는 근거를 제시하였으며, 궁극적으로 소아청소년 비만 및 대사성 질환의 치료를 위한 근간을 마련하는데 기여할 수 있을 것으로 기대한다.

* 본 내용의 일부는 Metabolites 학술지 (Chae W, Lee KJ, et al. Association of Metabolic Signatures with Nonalcoholic Fatty Liver Disease in Pediatric Population. Metabolites. 2022;12(9); Sohn MJ, Chae W, et al. Metabolomic Signatures for the Effects of Weight Loss Interventions on Severe Obesity in Children and Adolescents. Metabolites. 2022;12(1))에 두 편의 원저로 출판 완료된 내용임.

주요어: 소아청소년 비만, 대사체학, 비알코올성지방간질환, 체중감량 중재, 혈장 대사체

학 번: 2017-39919