

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Driuchina, Anastasiia; Hintikka, Jukka; Lehtonen, Marko; Keski-Rahkonen, Pekka; O'Connell, Thomas; Juvonen, Risto; Kuula, Juho; Hakkarainen, Antti; Laukkanen, Jari A.; Mäkinen, Elina; Lensu, Sanna; Pietiläinen, Kirsi H.; Pekkala, Satu

**Title:** Identification of Gut Microbial Lysine and Histidine Degradation and CYP-Dependent Metabolites as Biomarkers of Fatty Liver Disease

**Year:** 2023

**Version:** Published version

**Copyright:** © 2023 Driuchina et al.

**Rights:** CC BY 4.0

**Rights url:** <https://creativecommons.org/licenses/by/4.0/>

**Please cite the original version:**

Driuchina, A., Hintikka, J., Lehtonen, M., Keski-Rahkonen, P., O'Connell, T., Juvonen, R., Kuula, J., Hakkarainen, A., Laukkanen, J. A., Mäkinen, E., Lensu, S., Pietiläinen, K. H., & Pekkala, S. (2023). Identification of Gut Microbial Lysine and Histidine Degradation and CYP-Dependent Metabolites as Biomarkers of Fatty Liver Disease. *mBio*, 14(1), Article e02663-22. <https://doi.org/10.1128/mbio.02663-22>



# Identification of Gut Microbial Lysine and Histidine Degradation and CYP-Dependent Metabolites as Biomarkers of Fatty Liver Disease

Anastasiia Driuchina,<sup>a</sup> Jukka Hintikka,<sup>a</sup> Marko Lehtonen,<sup>b</sup> Pekka Keski-Rahkonen,<sup>c</sup> Thomas O'Connell,<sup>d</sup> Risto Juvonen,<sup>b</sup> Juho Kuula,<sup>e,j</sup> Antti Hakkarainen,<sup>e,j</sup> Jari A. Laukkanen,<sup>f,k</sup> Elina Mäkinen,<sup>a</sup> Sanna Lensu,<sup>a,g</sup> Kirsi H. Pietiläinen,<sup>h,l</sup>  Satu Pekkala<sup>a,i</sup>

<sup>a</sup>Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

<sup>b</sup>Faculty of Health Sciences, School of Pharmacy, University of Eastern Finland, Kuopio, Finland

<sup>c</sup>Nutrition and Metabolism Branch, International Agency for Cancer on Research, Lyon, France

<sup>d</sup>Indiana University School of Medicine, Indianapolis, Indiana, USA

<sup>e</sup>HUS Medical Imaging Center, Department of Radiology, University of Helsinki, Helsinki, Finland

<sup>f</sup>Department of Medicine, Central Finland Health Care District, Jyväskylä, Finland

<sup>g</sup>Faculty of Education and Psychology, University of Jyväskylä, Jyväskylä, Finland

<sup>h</sup>Obesity Research Unit, Research Program for Clinical and Molecular Medicine, Faculty of Medicine, University of Helsinki, Helsinki, Finland

<sup>i</sup>Department of Clinical Microbiology, Turku University Hospital, Turku, Finland

<sup>j</sup>Helsinki University Hospital, Population Health Research, Finnish Institute for Health and Welfare, Helsinki and Oulu, Finland

<sup>k</sup>Institute of Clinical Medicine, Department of Medicine, University of Eastern Finland, Kuopio, Finland

<sup>l</sup>HealthyWeightHub, Abdominal Center, Helsinki University Hospital, University of Helsinki, Helsinki, Finland

**ABSTRACT** Numerous studies have described specific metabolites as biomarkers of severe liver diseases, but very few have measured gut microbiota (GM)-produced metabolites in fatty liver disease. We aimed at finding GM signatures and metabolite markers in plasma and feces related to high liver fat content. Based on imaging, we divided study participants into low (<5%, LF,  $n = 25$ ) and high (>5%, HF,  $n = 39$ ) liver fat groups. Fecal (LF  $n = 14$ , HF  $n = 25$ ) and plasma (LF  $n = 11$ , HF  $n = 7$ ) metabolomes of subsets of participants were studied using liquid chromatography/high resolution mass spectrometry. The GM were analyzed using 16S rRNA gene sequencing. Additionally, blood clinical variables and diet were studied. Dyslipidemia, higher liver enzymes and insulin resistance characterized the HF group. No major differences in diet were found between the groups. In the GM, the HF group had lower abundance of *Bacteroides* and Prevotellaceae NK3B31 group than the LF group after adjusting for metformin use or obesity. In feces, the HF group had higher levels of lysine and histidine degradation products, while 6-hydroxytestosterone (metabolized by CYP3A4) was low. Higher plasma levels of caffeine and its metabolites in the HF group indicate that the activity of hepatic CYP1A2 was lower than in the LF group. Our results suggest, that low fecal Prevotellaceae NK3B31 and *Bacteroides* abundance, and increased lysine and histidine degradation may serve as GM biomarkers of high liver fat. Altered plasma caffeine metabolites and lowered testosterone metabolism may specify decreased CYP activities, and their potential utility, as biomarkers of fatty liver disease.

**IMPORTANCE** Because the high prevalence of nonalcoholic fatty liver disease sets diagnostic challenges to health care, identification of new biomarkers of the disease that in the future could have potential utility as diagnostic biomarkers of high liver fat content is important. Our results show that increased amino acid degradation products in the feces may be such biomarkers. In the blood, molecules that indicate defective hepatic metabolic enzyme activities were identified in individuals with high liver fat content.

**KEYWORDS** gut microbiota, insulin resistance, liver fat, metabolomics

**Invited Editor** Elizabeth K. Mallott, Washington University in St. Louis

**Editor** Avery August, Cornell University

**Copyright** © 2023 Driuchina et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Satu Pekkala, satu.p.pekkala@jyu.fi.

The authors declare no conflict of interest.

**Received** 17 November 2022

**Accepted** 4 January 2023

Nonalcoholic fatty liver disease (NAFLD), with a worldwide prevalence of up to 40% in adult population, is an acutely increasing burden on public health (1). NAFLD is characterized by hepatic steatosis, i.e., excess (>5%) fat accumulation in hepatocytes (2). Obesity and insulin resistance are associated with NAFLD (3). NAFLD is shown to be present in 40 to 80% of patients with type 2 diabetes (1, 4, 5), and in 30 to 90% of patients with obesity (6). Dietary factors, such as high fructose intake contribute to the development of NAFLD (7). In addition, defective hepatic CYP enzyme activities are well known molecular players in NAFLD (8).

The gut microbiota (GM) referring to the ~100 trillion microbes inhabiting the gastrointestinal tract (9) contribute to the pathophysiology of NAFLD. We have found that in humans high hepatic fat content is linked to lower abundance of *Faecalibacterium prausnitzii* (10), and further, in a mouse model we have been able to ameliorate NAFLD by administering *F. prausnitzii* (11). One previous study reported that NAFLD patients have lower abundance of Bacteroidetes (12), and another lower abundance of *Bacteroides* (13). Contrarily, others have shown higher abundance of *Bacteroides* in NAFLD (14, 15).

It is increasingly accepted that the abundance of certain members of the GM affect human physiology by processing the ingested food to bioactive metabolites. These molecules act as signaling messengers after absorption in extraintestinal tissues (16), including the liver, which is the first target organ that metabolizes them (17). Despite the importance of the gut-liver axis, studies describing the importance of fecal metabolites in human NAFLD are surprisingly scarce. One study reported that the levels of fecal bile acids and propionate are increased in NAFLD patients depending on the severity of hepatic fibrosis (18). Another study has shown elevated fecal propionate in NAFLD patients, accompanied with higher levels of acetate and butyrate (19). Contrarily, lower levels of acetate are found in obese pediatric NAFLD patients, while propionate is unaffected (20). These results suggest the role of GM in NAFLD, but more studies on the fecal metabolome are needed to assess whether short-chain fatty acids or other GM metabolites could be useful diagnostic biomarkers of NAFLD.

Metabolic profiling of plasma and serum samples have been used to study the pathophysiology of NAFLD, and has, for example linked bile acid homeostasis to NAFLD (21, 22). However, most studies have focused on diagnosing advanced hepatic fibrosis or differentiating between nonalcoholic steatohepatitis (NASH) and non-NASH (23, 24), whereas NAFLD have been less studied. One model proposes that NASH, fatty, or healthy liver status could be distinguished based on lipidomics, glycomics, and hormones (25). A few studies have identified serum lipid signatures, which might be used as diagnostic markers for NAFLD (26, 27). In addition, one recent study discovered 11 metabolites and three clinical parameters, which were associated with NAFLD (28).

Because the high prevalence of NAFLD sets diagnostic challenges to the health care system, we examined the GM, fecal and plasma metabolites, and diet as well as clinical parameters of humans with low (LF, <5%) and high liver (HF, >5%) fat content. The aim of this study was to identify new biomarkers of NAFLD that could have potential utility as diagnostic markers of high liver fat content.

## RESULTS

### Anthropometric and clinical variables differed between the liver fat groups.

The HF participants had higher body weight, BMI, waist circumference, and blood pressure than the LF participants ( $P < 0.05$  for all, Table 1). The HF group had 24 males and 15 females (61.5% and 38.5%, respectively), and the LF group 11 males and 14 females (44% and 56%, respectively), with no significant differences in gender between the groups.

The HF participants had ~2-times higher alanine aminotransferase (ALT), aspartate aminotransferase (AST) and gamma-glutamyltransferase (GGT) than the LF participants ( $P < 0.001$  for all, Table 2), yet only ALT was over the reference range. Compared to the LF, the HF group had ~19% lower HDL ( $P = 0.007$ ) and ~2-times higher triglycerides ( $P < 0.001$ ), and the latter was over the reference range (Table 2). At fasting, oral glucose tolerance test (OGTT) showed ~7% higher levels of glucose ( $P = 0.017$ ) and insulin ( $<0.001$ ) in the HF group compared to

**TABLE 1** Anthropometric measurements, liver fat % and blood pressure of the study participants<sup>a</sup>

Variable	LF group (n = 25)	HF group (n = 39)	P value
Age (yrs)	48 ± 12	51 ± 10	0.345
ht (cm)	171.4 ± 10.1	175.2 ± 9	0.115
wt (kg)	87 ± 17.2	105.3 ± 20.7	<b>&lt;0.001</b>
BMI (kg/m <sup>2</sup> )	29.8 ± 6.3	34.2 ± 5.8	<b>0.005</b>
Waist circumference (cm)	101.1 ± 14	117.8 ± 13.9	<b>&lt;0.001</b>
BP systolic (mm Hg)	127 ± 18 (n = 24)	138 ± 18	<b>0.042</b>
BP diastolic (mm Hg)	81 ± 11 (n = 24)	87 ± 9	<b>0.024</b>
Liver fat %	1.3 ± 1.3	17.2 ± 10	<b>&lt;0.001</b>

<sup>a</sup>The data are presented as mean ± SD. LF < 5% fat, HF > 5% fat in liver. Bolded P values indicate statistically significant difference between the groups. BMI, body mass index; BP, blood pressure; LF, low fiver fat; HF, high liver fat.

the LF group. Consequently, the HOMA-IR was higher in the HF group ( $P < 0.001$ ), showing significant insulin resistance (Table 2).

**The HF participants had lower GM diversity and different GM composition than the LF group.** By analyzing the GM alpha-diversity, we found that the HF group had lower Shannon index (i.e., species diversity) ( $P = 0.04$ , Fig. 1A), while no differences were seen in Chao1 (i.e., species richness, Fig. S1A). In addition, the HF group had lower GM phylogenetic diversity than the LF group ( $P = 0.03$ , Fig. 1B). Bray Curtis distance showed that the groups did not differ in beta-diversity (i.e., interindividual species diversity, Fig. S1B).

The average GM composition in the groups is shown in Fig. S2A. In group comparisons, we found differences in the abundance of several taxa (Fig. S2B). However, when adjusted for the use of metformin or obesity, only the difference in *Bacteroides* and Prevotellaceae NK3B31 between the groups remained significant, i.e., the abundance of these was lower in the HF group (Fig. 2A and B).

As diet affects the GM composition, the dietary intakes of macronutrients were analyzed. Notably, we found no significant differences in the energy and nutrient intakes between the groups, except for vitamin E ( $P = 0.035$ ) and sucrose ( $P = 0.038$ ) intakes that were higher in the HF than the LF group (Table S2).

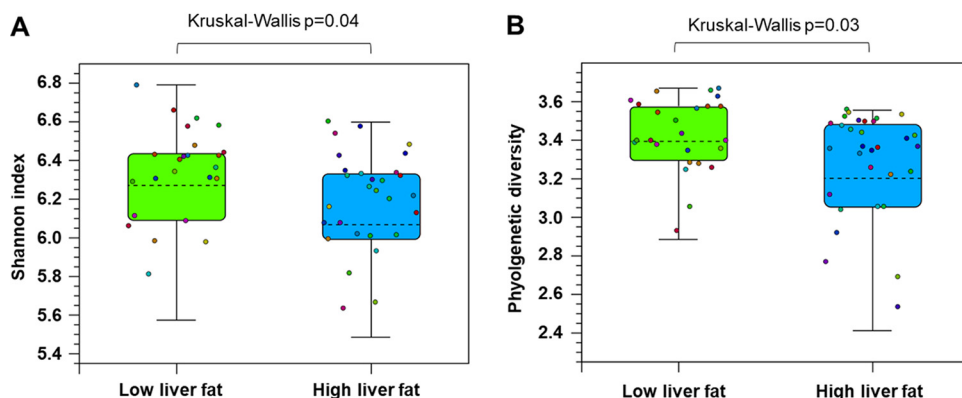
**The fecal metabolome of the HF group was characterized by high levels of lysine and histidine degradation products.** We found that the levels of 19 metabolites in feces differed between the groups (Wilcoxon rank-sum  $P < 0.05$  and  $\geq 2$ -fold change, Fig. 3, Table S3A). Identification of the metabolites is presented in Table S1A, and their correlation networks in Fig. S3. Levels of seven metabolites were higher in the HF group vary-

**TABLE 2** Serum clinical variables measured using clinical chemistry<sup>a</sup>

Variable	Reference range	LF group (n = 25)	HF group (n = 39)	P value
ALT (U/L)	male < 50, female < 35	25.3 ± 11	54 ± 41.1	<b>&lt;0.001</b>
AST (U/L)	15–45	23.5 ± 4.3	42.6 ± 35	<b>&lt;0.001</b>
GGT(U/L)	<61	27.1 ± 17.2	53.8 ± 32.5	<b>&lt;0.001</b>
Cholesterol (mmol/L)	<5	5.2 ± 1.4	5.5 ± 1.3	0.395
HDL (mmol/L)	male > 1, female > 1.2	1.6 ± 0.5	1.3 ± 0.3	<b>0.007</b>
LDL (mmol/L)	<3	3 ± 0.6	3.3 ± 0.9	0.282
Trigly (mmol/L)	<1.7	1.1 ± 0.5	2 ± 1.4	<b>&lt;0.001</b>
Glucose (mmol/L) <sup>b</sup>	4–6	5.7 ± 0.5	6.1 ± 0.7	<b>0.017</b>
Insulin (iU/L) <sup>b</sup>	2–20	9.1 ± 9.0	19.4 ± 13.1	<b>&lt;0.001</b>
HOMA-IR <sup>b</sup>	IS < 1, early IR < 1.9, significant IR > 2.9	2.3 ± 2.2	5.4 ± 3.8	<b>&lt;0.001</b>

<sup>a</sup>The data are presented as mean ± SD. LF < 5% fat, HF > 5% fat in liver. Bolded P values indicate statistically significant difference between the groups. The reference ranges are based on values obtained with standard methods in the accredited laboratory HUSLAB, Finland ([www.huslab.fi/ohjekirja](http://www.huslab.fi/ohjekirja)). ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transferase; HDL, high-density lipoprotein; LDL, low-density lipoprotein; Trigly, triglycerides; HOMA-IR, Homeostatic Model Assessment of Insulin Resistance; IS, insulin sensitive; IR, insulin resistance; LF, low fiver fat; HF, high liver fat.

<sup>b</sup>LF n = 23, HF n = 32. The lower number of participants in OGTT is because it was not done for diabetic participants. HOMA-IR was calculated as fasting insulin (mU/L) x fasting glucose (mmol/L)/22.5.



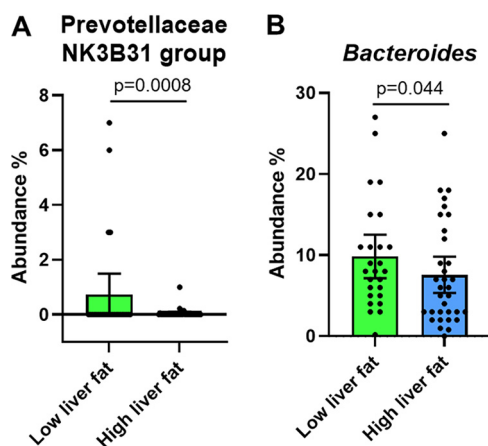
**FIG 1** The diversity of the gut microbiota in the low ( $n = 25$ ) and high ( $n = 37$ ) liver fat groups. (A) Shannon index, a measure of alpha-diversity indicating species diversity. (B) Phylogenetic diversity. The data are shown as mean  $\pm$  95 CI, with the dots indicating diversity in individual samples. Statistical differences between the groups are shown above the panels (Kruskal-Wallis test).

ing from 2- to 20-fold compared to LF, and the levels of 12 metabolites in the HF group were 0.16 to 0.50-fold lower from the levels of LF group (Fig. 3A and B, Table S3A). When adjusting for body weight and metformin use, seven metabolites remained significantly different, and the difference in glycyrrhetic acid became significant (Fig. 3B). The level of testosterone metabolism product, 6-betahydroxytestosterone was very low in the HF group compared to the LF group despite that the HF group included more males. The histidine degradation product N-omega-acetylhistamine, and lysine degradation product saccharopine were higher in the HF group than in the LF group.

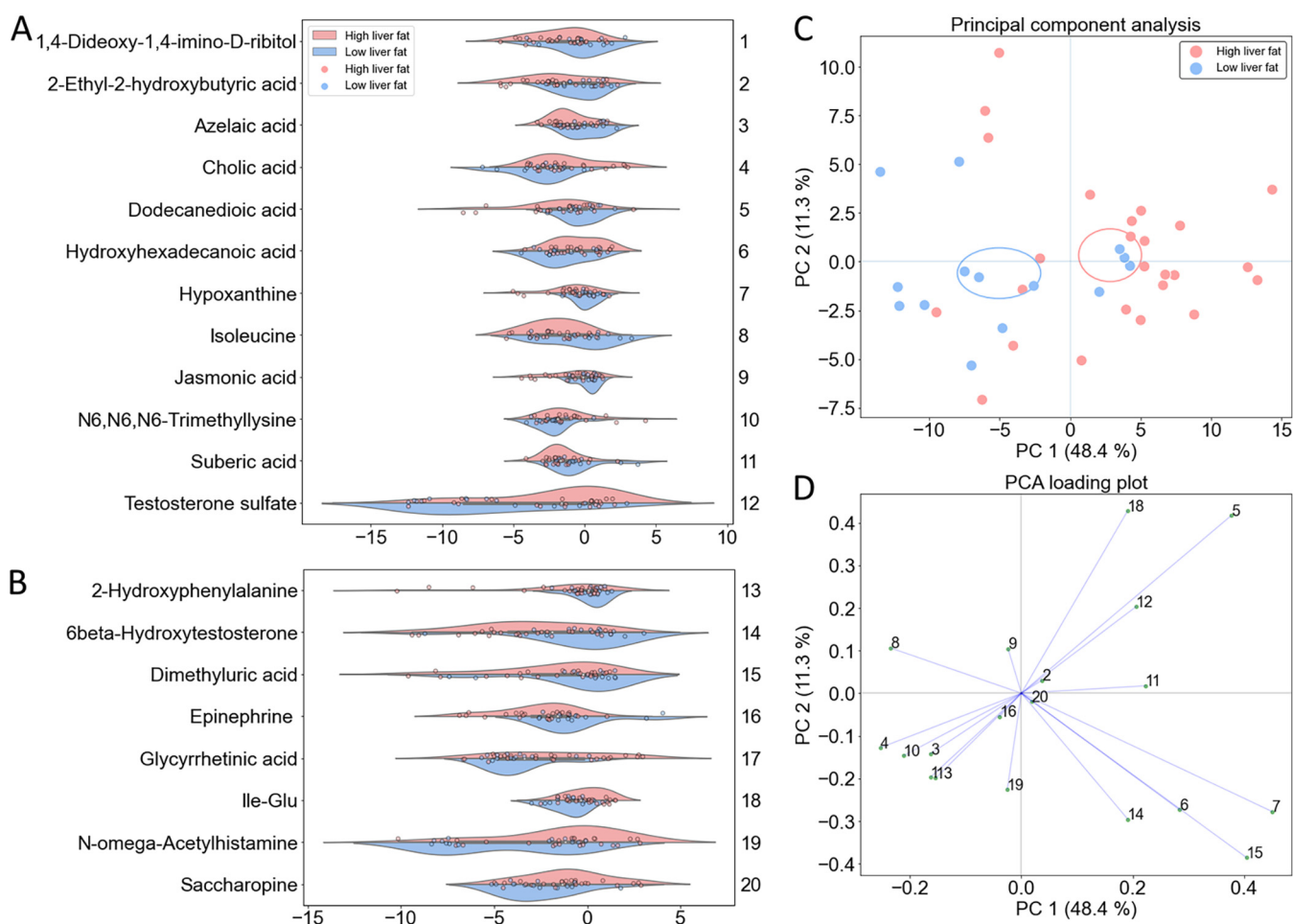
Next, we analyzed the associations between the fecal metabolites and liver fat % (Fig. 4A and B, Table S3B). Cholic acid and its precursor 7- $\alpha$ ,27-dihydroxycholesterol increased along with liver fat %. In addition, histidine metabolism products urocanic acid, methylhistidine, histamine, anserine, N-acetylhistamine, and methylimidazole acetic acid associated positively with liver fat %. The associations of the GM with fecal metabolites are visualized in Fig. S4.

**The HF group was characterized by higher plasma levels of caffeine metabolites, primary and secondary bile acids suggesting altered hepatic CYP activities.**

We found that 14 plasma metabolites differed between the groups (Wilcoxon rank-sum  $P < 0.05$  and  $\geq 2$ -fold change, Table S4A & Fig. 5). The identification of the metabolites is presented in Table S1B. The levels of 12 metabolites were 2- to 4.3-fold higher in the HF group compared to the LF group. The HF group had higher plasma levels of



**FIG 2** (A and B) The differential abundances of the gut microbiota between the low ( $n = 25$ ) and high ( $n = 37$ ) liver fat groups after adjusting for metformin use or obesity. The data are shown as mean  $\pm$  95 CI, with the dots indicating abundance % in individual samples. The group differences were analyzed with ANOVA-like test followed by FDR correction.

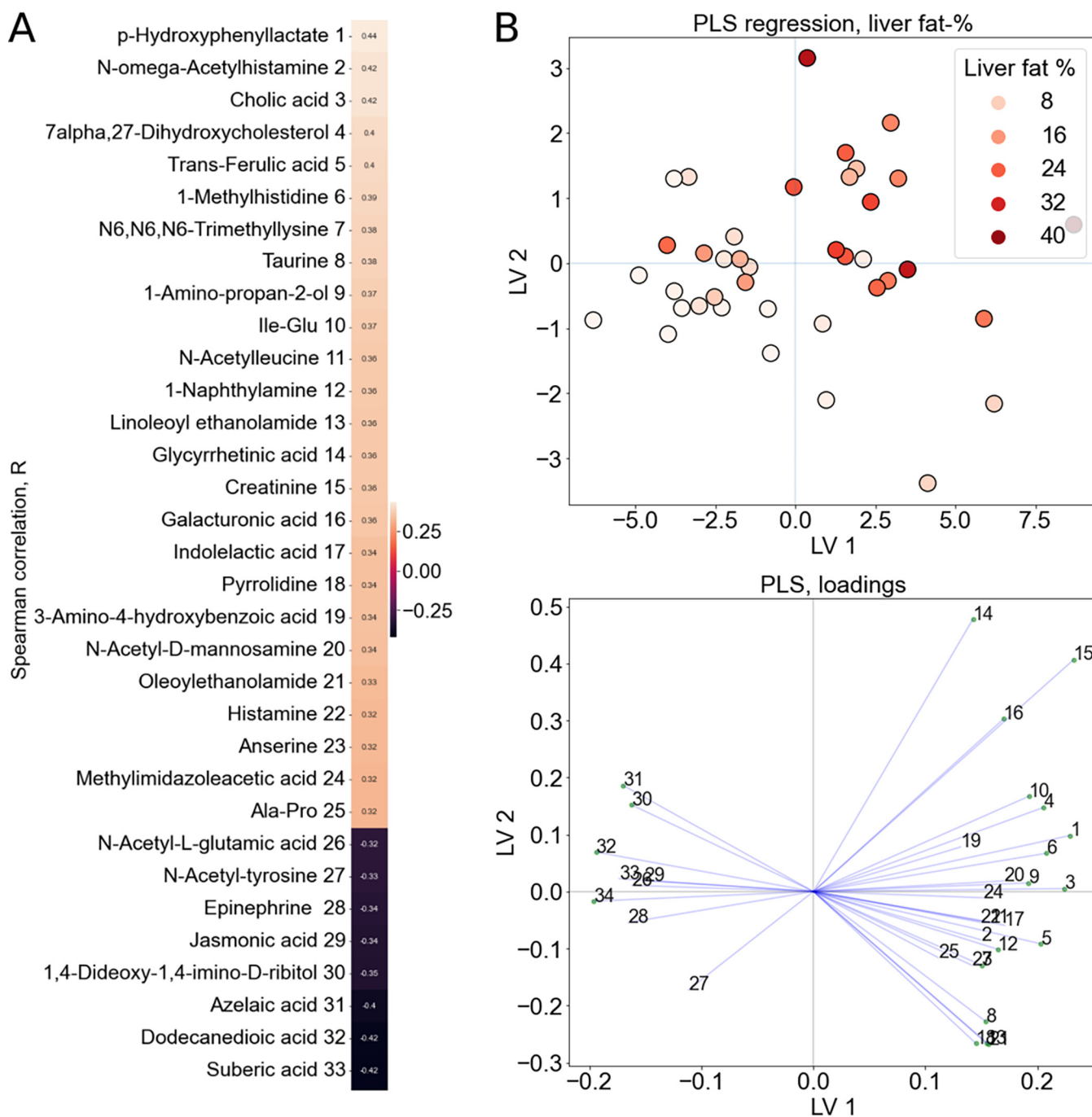


**FIG 3** Visualization of the fecal metabolites that (A) differed between the low ( $n = 14$ ) and high ( $n = 25$ ) liver fat groups without including confounding factors in the analysis, and (B) differed between the groups after adjusting for body weight and metformin use. The x axes indicate  $\log_2$ -fold differences. The group differences were analyzed using Mann-Whitney U-test. Features with a significant difference  $P < 0.05$  and fold difference of  $\geq 2$  were considered of interest. (C) Principal-component analysis (PCA) of the differing metabolites without adjustment. Circles denote a 95% CI of the scores in each dimension. (D) PCA loading plot of the differing metabolites showing covariances of the variables with the principal components.

several bile acids and their glycine conjugates, which are outcome of either hepatic CYP7A1 and metabolized by other hepatic CYPs or higher reabsorption of bile acids. However, we did not find difference in the downstream product of CYP7A1, 7 $\alpha$ -hydroxy-4-cholestelen-3-one, between the groups (fold change value HF/LF 0.67 in the plasma and 2.08 in the feces). Caffeine and its metabolites paraxanthine, theophylline, and theobromine as well as cyclo(leucylprolyl), were also higher in the HF group than in the LF group. Notably, there was no difference in coffee intake between the groups (Table S2). To estimate the activity of caffeine-metabolizing CYP1A2, we determined the ratio of paraxanthine/caffeine (29), which was  $\sim 2$ -fold lower in the HF group than in the LF group ( $P = 0.002$ , Fig. 5D). Further, the ratio associated negatively with liver fat % ( $R = -0.532$ ,  $P = 0.023$ , Fig. 5D).

We also analyzed the associations between plasma metabolites and liver fat %. Bile acids, their conjugates, cyclo(leucylprolyl) and amino acids proline, isoleucine and glutamic acid associated positively with liver fat %, while several phospholipids associated negatively with it (Fig. 6, Table S4B).

To further understand the metabolic derangements, a correlation-based network community analysis was carried out using the Girvan-Newman algorithm (30), where the metabolites are the network nodes and the connections between the nodes are the Pearson correlations. The network community analysis diagram was generated combining all 38 significant plasma metabolites. The algorithm detected four distinct communities (Fig. S5). The first community

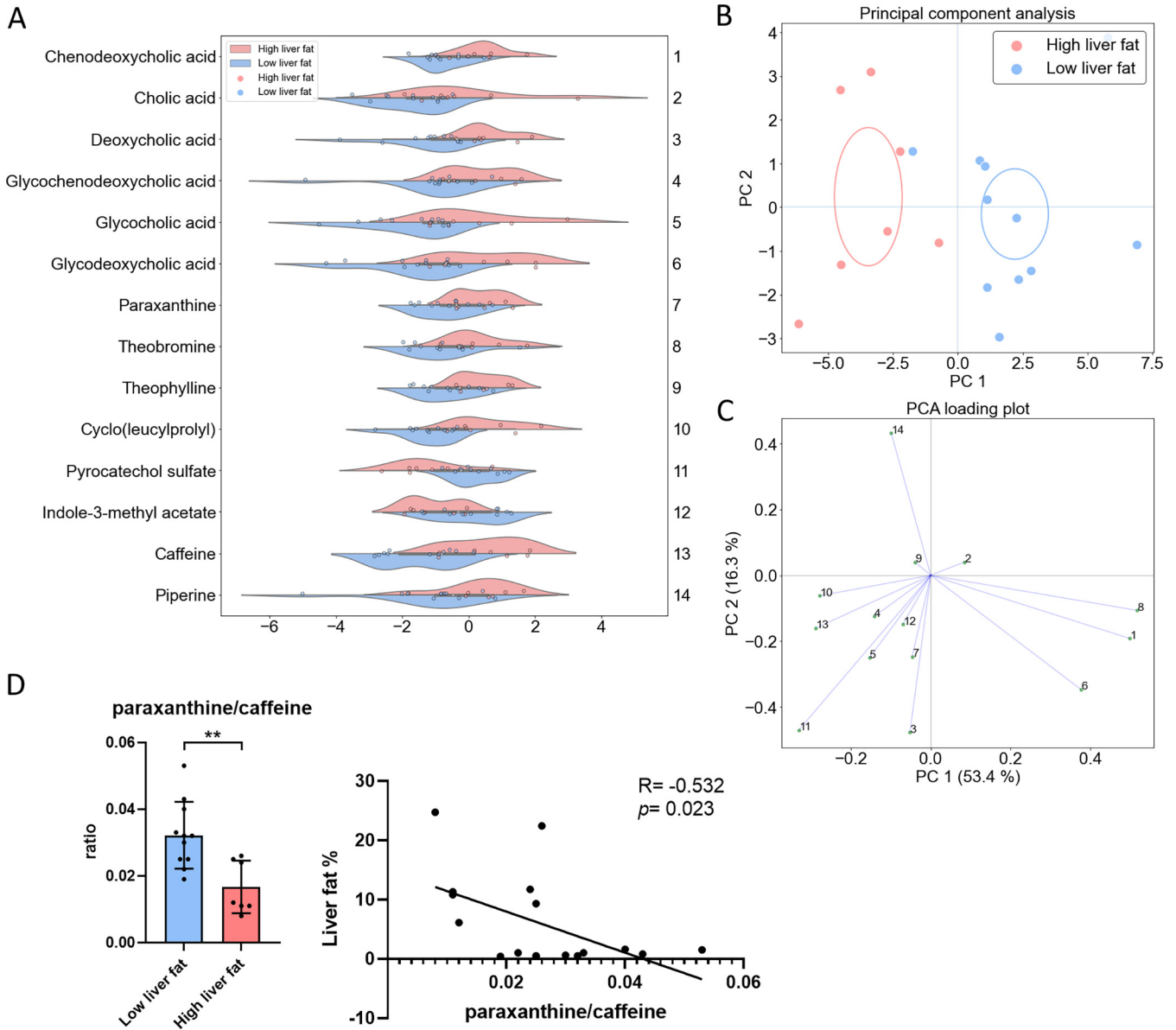


**FIG 4** The features with significant Spearman correlations ( $P < 0.05$ ) between the fecal metabolites and liver fat % visualized with (A) heatmap and (B) partial least-squares (PLS) plots and loadings. The color in the heatmap corresponds to the correlation coefficient between the metabolites and liver fat % as shown on the scale right side of the heatmap. LV, latent variable.

included pyrocatechol sulfate and indole-3-methyl acetate. The second community was composed of primary and secondary bile acids as well as two short-chain acylcarnitines. The third community was composed of diet-related metabolites, especially of caffeine metabolism. The last community was composed entirely of lipids, including lysophosphocholines and phosphocholines.

## DISCUSSION

Because the high prevalence of NAFLD sets diagnostic challenges to health care, we examined the GM, fecal and plasma metabolites as well as clinical parameters of



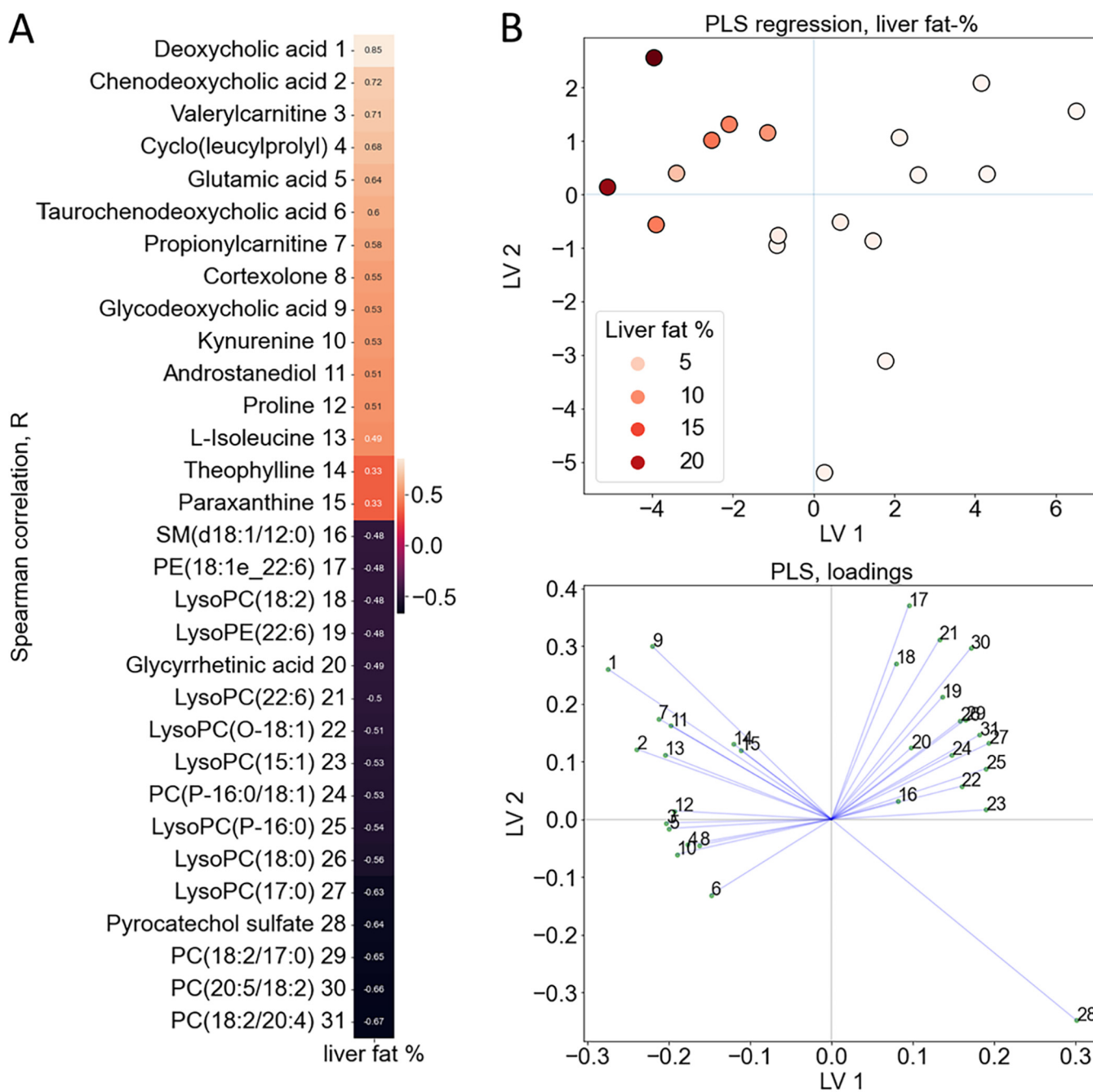
**FIG 5** (A) Visualization of the plasma metabolites that differed between the low ( $n = 11$ ) and high ( $n = 7$ ) liver fat groups. The X axis indicates  $\log_2$  fold differences. The group differences were analyzed using Mann-Whitney U-test. Features with a significant difference  $P < 0.05$  and fold change of  $\geq 2$  were considered of interest. (B) Principal component analysis (PCA) of the differing metabolites and (C) PCA loading plot of the differing metabolites, showing covariances of the variables with the principal components. (D) Paraxanthine/caffeine ratio, indicative of CYP1A2 activity and its association with liver fat %. In the graph, the data are presented as mean  $\pm$  SD, and the dots indicate individual data points. The group differences were analyzed with  $t$ -test and the association with Spearman correlation. \*\*,  $P < 0.01$ .

humans with low (LF,  $<5\%$ ) and high liver (HF,  $>5\%$ ) fat content in this study. We identified new potential microbial and plasma metabolic biomarkers of fatty liver disease.

Metabolically, especially higher levels of triglycerides and significant insulin resistance in the HF group suggest their utility as biomarkers of a fatty liver disease without biopsy specimen proven NAFLD. This is supported by earlier studies showing that dyslipidemia (31, 32), and the pathogenesis of insulin resistance (33–35) associated with NAFLD. Our previous study also showed insulin resistance in humans with high liver fat content ( $>5\%$  based on MRI) (10).

We found that the GM diversity was lower in the HF group, which is consistent, e.g., with the study by Astbury et al. showing lower GM diversity in NAFLD patients (36). When adjusting for obesity or the use of metformin, both of which are known to affect GM (37), the HF





**FIG 6** The features with significant Spearman correlations ( $P < 0.05$ ) between the plasma metabolites and liver fat % visualized with (A) heatmap and (B) partial least-squares (PLS) plot ( $R^2 = 0.74$ ) and loadings. The color in the heatmap corresponds to the correlation coefficient between the metabolites and liver fat % as shown on the scale right side of the heatmap. LV, latent variable.

group had lower abundance of Prevotellaceae NK3B31 than the LF group. Consistently, we have shown that when a rat model of NAFLD was supplemented with prebiotic Xylo-oligosaccharides, the abundance of Prevotellaceae NK3B31 increased along with reduction of liver fat (38). Thus, this genus might be an indicator of hepatic steatosis. Our finding on the lower abundance of *Bacteroides* agrees with a study by Testerman et al. in NAFLD patients (13), while opposite findings also exist (15). However, contrary to our results, Testerman and coworkers found that two GM pathways of histidine degradation were decreased in NAFLD patients (13). In contrast, elevated urinary methylhistidine has been described in NAFLD patients (39). In agreement with the latter, we report here that the fecal level of histidine

metabolism product, N-omega-acetylhistamine was markedly increased in the HF group. Further, other products of histidine degradation, such as anserine, positively associated with liver fat.

Of the GM metabolites, we also found that saccharopine, a degradation product of lysine was higher in the HF group compared to the LF. In agreement, Michail et al. found by studying GM metagenomes, that the lysine degradation pathway (lysine=>saccharopine=>acetoacetyl-CoA) was exclusively identified in NAFLD patients and not in healthy individuals (20). Together these findings suggest that increased histidine and lysine degradation by the GM may be HF biomarkers. Furthermore, in the feces of the HF group, 6-hydroxybetatestosterone was reduced. Because there were more males in the HF group than in the LF group, our finding further suggests that higher hepatic fat content could associate with lower activity of testosterone-metabolizing CYP3A4 (40).

We found elevated cyclo(leucylprolyl), glycine-conjugated bile acids and caffeine metabolites in the plasma of the HF participants. Interestingly, a previous study has shown that dipeptides, glycine-conjugated bile acids, and lower coffee intake were associated with the incidence of hepatocellular carcinoma and fatal liver disease in a prospective setting (41). Here, we report that paraxanthine, which is the major primary metabolite of caffeine (42), as well as minor primary metabolites theophylline and theobromine, were elevated in the plasma of HF participants, despite no difference in coffee intake between the groups. The results of our metabolite network analysis further supported these metabolic derangements. While, to our knowledge, elevated caffeine metabolites have not been reported in NAFLD, Holstege et al. have shown that the half-life of these metabolites increases as cirrhosis progresses (43). Whether the elevated plasma caffeine metabolites and cyclic dipeptides could be useful biomarkers in detecting NAFLD warrants further research in larger populations.

Caffeine metabolism is a CYP-dependent process in the liver, and alterations in CYP expression and activity have been described in NAFLD (44). We found that the activity of CYP1A2, as estimated from the increased paraxanthine/caffeine ratio (29), was reduced in the HF participants. Accordingly, CYP1A2 mRNA expression and enzyme activity are shown to decrease along with NAFLD progression from steatosis to NASH (8). Further, a decline in CYP1A2 enzymatic activity with exacerbating fibrosis has been observed (45). CYP1A2 is the most important CYP in metabolism of caffeine (29). Taken together, it seems that the low activity of CYP1A2 could be an early biomarker of NAFLD. In agreement with previous studies (46–49), we observed signs of altered bile acid metabolism in the plasma of the HF group, as increased bile acids levels were found both in feces and plasma. Our results therefore suggest that the activity of bile acids metabolizing CYP7A1 (50) is increased in the HF group. However, we did not find differences in the downstream product of CYP7A1, 7 $\alpha$ -hydroxy-4-cholestelen-3-one, between the groups which suggests that CYP7a1 activity is not decreased. This is in agreement with what has been shown in NAFLD and NASH (51).

Our study is not without limitations. The sample sizes for metabolite analyses can be considered rather small, and we could not find similar cohorts to reproduce our findings and to compare our results with others. Because the participants self-collected the fecal samples at their homes, the samples could not be immediately frozen at  $-80^{\circ}\text{C}$ . Thus, some microbial metabolism may have occurred during the 1 to 2 days when the samples were frozen at nonstandardized home freezers.

In conclusion, our study suggests that lower Prevotellaceae NK3B31 group and *Bacteroides* abundance as well as increased lysine and histidine degradation in feces could be microbial biomarkers of high liver fat. Further, higher plasma caffeine metabolites and signs of altered testosterone metabolism, suggesting lower hepatic CYP1A2 and CYP3A4 activities may be used as biomarkers of high liver fat. However, to establish the use of these variables as biomarkers in detecting the risk of NAFLD warrants further research in larger populations.

## MATERIALS AND METHODS

**The human study participants.** The study was approved by the Ethics Committee of the Hospital District of Southwest Finland (ETMK 72/2019) and by the Helsinki University Hospital (270/13/03/01/2008). The inclusion criteria were an age <75 years, being overweight (body mass index [BMI] >25), and

high waist circumference ( $>102$  cm for males,  $>88$  cm for females). The exclusion criteria were antibiotic treatment 1 month prior to the study, excessive alcohol consumption ( $>20$  g/day for females, 30 g/day for males), inflammatory bowel disease, celiac disease, major eating disorders, cardiovascular diseases, and hypothyroidism. Informed consent was collected from all participants prior to the study.

The participants were recruited by the University of Jyväskylä and the University of Helsinki, Finland, and data were analyzed between 2020 and 2022. The participants were divided into low (LF), and high (HF) liver fat groups based on magnetic resonance imaging (MRI) or magnetic resonance spectroscopy (MRS) quantification of the hepatic triglyceride content. Details of the imaging analyses are provided in Supplementary material (Text S1). Liver fat was measured with MRI for 46 eligible participants and with MRS for eligible 18 participants. Still, there is no consensus on the exact fat % threshold distinguishing between healthy and fatty liver. Based on the literature (2), we considered the participants as having fatty liver disease, when the liver had over 5% fat. Based on the measurements, 39 participants had high ( $>5\%$  fat, HF) and 25 had low ( $<5\%$  fat, LF) liver fat. The participants did not have clinical diagnosis of NAFLD, and thus, for clarity we refer to fatty liver disease.

**Oral glucose tolerance test and blood samples.** Oral glucose tolerance test (OGTT) was performed for nondiabetic participants ( $n = 59$ ) after overnight fasting. Blood samples were collected prior to commencing the OGTT and at 30, 60 and 120 min. Homeostatic Model Assessment for Insulin Resistance (HOMA-IR) was calculated as fasting insulin (mU/L)  $\times$  fasting glucose (mmol/L)/22.5 (52, 53). Serum AST, ALT, GGT, glucose, insulin and triglycerides were analyzed using standardized clinical methods at HUSLAB laboratories (Helsinki, Finland) or bioanalytical facility of the Faculty of Sport and Health Sciences, University of Jyväskylä, Finland.

**Body composition measurements.** Height was measured using a wall-fixed measuring device, and weight was measured using an electronic scale. BMI was calculated as weight (kg)/height (m)<sup>2</sup>. Waist circumference was measured midway between the lowest rib and the iliac crest twice with a tape measure and the mean value was used for the analyses.

**Fecal sample collection and 16S rRNA gene sequencing.** The participants self-collected the fecal samples, which were home-frozen immediately after collection, brought to laboratory frozen and stored at  $-80^{\circ}\text{C}$ . The bacterial DNA was extracted using Stool-kit v2 and semiautomated GenoXtract (Hain Lifescience, Nehren, Germany) accompanied with preceding bead-beating in 0.5/1.0 mm ceramic bead tubes. The 16S rRNA gene was amplified using primers targeting the V3-V4 regions. The protocol is described in Supplementary material (Text S1). 16S rRNA gene sequences were clustered into operational taxonomic units (OTUs) at 97% similarity using CLC Microbial Genomics Package (Qiagen, Hilden, Germany). The rRNA gene sequences were classified using SILVA SSU Reference database (v132, 99%).

**Analyses of fecal and plasma metabolites from subsets of participants.** Plasma sample preparation for metabolite analyses (LF  $n = 11$ , HF  $n = 7$ ) was done as previously (54). Briefly, samples were thawed on ice and a 100  $\mu\text{L}$  aliquot of plasma was dispensed into a 96-well filter plate (Captive ND, 0.2  $\mu\text{m}$  PP, Agilent Technologies) containing 400  $\mu\text{L}$  of ice-cold acetonitrile. Samples were mixed to thoroughly precipitate plasma proteins, and then centrifuged  $700 \times g$  for 5 min at  $4^{\circ}\text{C}$  and the supernatants were collected to a 96-well storage plate and stored at  $10^{\circ}\text{C}$ .

The fecal samples (LF  $n = 14$ , HF  $n = 25$ ) were suspended in phosphate-buffered saline with a ratio of 1:5 (w:v), vortexed for 10 min at full speed and centrifuged  $700 \times g$  for 5 min at  $4^{\circ}\text{C}$ . An aliquot of 100  $\mu\text{L}$  of the fecal extract was combined with 500  $\mu\text{L}$  of ice-cold methanol, vortexed for 30 sec at full speed and centrifuged  $700 \times g$  for 5 min at  $4^{\circ}\text{C}$ . The supernatant was transferred to a clean microcentrifuge tube through a 0.2  $\mu\text{m}$  syringe filter.

Nontargeted metabolic profiling was performed at the LC-MS metabolomics center (Biocenter Kuopio, University of Eastern Finland, Finland). The analysis was carried out using an ultra-high performance liquid chromatography (Vanquish Flex UHPLC system, Thermo Scientific, Bremen, Germany, and a 1290 LC system, Agilent Technologies, Waldbronn, Germany) coupled online to a high-resolution mass spectrometry (Q Exactive Focus, Thermo Scientific and 6540 qTOF-MS, Agilent Technologies). All samples were analyzed using two different chromatographic techniques, i.e., reversed phase (RP) and hydrophilic interaction chromatography (HILIC). Data were acquired in both electrospray ionization (ESI) polarities, i.e., ESI positive (ESI+) and ESI negative (ESI-). Data-dependent product ion spectrums (MS2 data) were acquired from pooled quality control (QC) samples at the beginning and end of the analysis for each mode. QC samples were injected in the beginning of the analysis and after every 12 samples. The LC-MS instrument set-ups and data acquisition parameters have been described previously (55). Detailed analyses of the samples and data are described in Supplemental material (Text S1). The identification information on the fecal and plasma metabolites are shown in Tables S1A and B, respectively.

**Diet analyses.** The diet was analyzed from 3-day self-reported food diaries (2 weekdays and 1 weekend day) using Aivodiet software (Aivodiet, Flow-team Oy, Oulu, Finland). An average daily intake of energy and nutrients were calculated from the 3 days. Fecal sampling and food diary filling were performed within the same week.

**Statistical analyses.** Statistical analyses of clinical variables and diet were performed using IBM SPSS Statistics 26 (Armonk, NY, USA). Normal distribution of the variables was analyzed using Shapiro Wilk's test. Differences between the groups were analyzed using independent samples T-test or Mann-Whitney U-test. Statistical significance was set at  $P < 0.05$ .

Statistical analyses of the GM were performed with CLC Microbial Genomics Package (Qiagen). The group differences in GM alpha-diversity measures and phylogenetic diversity were analyzed with Kruskal-Wallis test. GM beta-diversity analysis was based on Bray-Curtis distance and PERMANOVA (PERmutational Multivariate ANalysis Of Variance) between the groups. GM taxonomic differences between the groups were analyzed with ANOVA-like comparison, followed by Benjamini-Hochberg

correction for multiple testing. In the taxonomic comparisons, statistical significance was set at FDR (false discovery rate)  $P < 0.05$ . The group differences were adjusted for obesity and metformin use.

Statistical analyses of the metabolites were performed using Python 3 and the packages Scipy and Scikit Learn. The raw abundances of the molecular features were  $\log_2$ -transformed, the group differences were analyzed using Mann-Whitney U-test and adjusted for obesity and metformin use. Features with a significant difference  $P < 0.05$  and fold change of  $\geq 2$  between the groups were considered of interest.

**Data availability.** The access to the data is restricted due to personal information protection (General Data Protection Regulation (GDPR) 2016/679 and Directive 95/46/EC). However, it is possible to contact author to ask for a copy of the material. The metadata of the study can be found in <https://doi.org/10.17011/jyx/dataset/85068>.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**TEXT S1**, DOCX file, 0.02 MB.

**FIG S1**, DOCX file, 0.2 MB.

**FIG S2**, DOCX file, 0.4 MB.

**FIG S3**, DOCX file, 0.2 MB.

**FIG S4**, DOCX file, 0.5 MB.

**FIG S5**, DOCX file, 0.1 MB.

**TABLE S1**, DOCX file, 0.03 MB.

**TABLE S2**, DOCX file, 0.02 MB.

**TABLE S3**, DOCX file, 0.02 MB.

**TABLE S4**, DOCX file, 0.02 MB.

## ACKNOWLEDGMENTS

Satu Pekkala was funded by the Academy of Finland (grant numbers 308042 and 349264), and by the ERVA funding of the Southwest Finland Health Care District. The study was also supported by the Academy of Finland funded Profiling of University of Jyväskylä and the Faculty of Sport and Health Sciences: Physical activity through life span (PACTS2). Kirsi H. Pietiläinen was funded by the Academy of Finland (grant numbers 335443, 314383, 272376, 266286), Finnish Medical Foundation, Gyllenberg Foundation, Novo Nordisk Foundation (grant numbers NNF20OC0060547, NNF17OC0027232, NNF10OC1013354), Finnish Diabetes Research Foundation, University of Helsinki, and Helsinki University Hospital Government Research Funds. Sanna Lensu was funded from the projects by the Academy of Finland (grant ID 316966 and 321522), and grants from Jenny & Artturi Wihuri Foundation and the Central Finland Regional Fund of the Finnish Cultural Foundation.

We thank the laboratory personnel of the Universities of Jyväskylä and Helsinki for the excellent technical assistance. We also thank the assistance of the personnel at the Helsinki University Hospital and the radiology department of the Central Finland Health Care District. We appreciate Biocentre Finland and Biocentre Kuopio for supporting LC-MS laboratory facility. We thank consultant radiologist Eeva Elomaa from The Central Finland Health Care District for revising the magnetic resonance images for any abnormal findings.

We have no conflicts of interest to declare.

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

Anonymized research data will be available upon a reasonable request. The European personal data protection laws impede depositing sensitive data to public repositories.

S.P. and K.H.P. designed the study and provided resources; S.P., A.D., J.H., and S.L. did the measurements and sampling of the study participants; A.D., J.K., and A.H. measured and analyzed the liver fat of the study participants; S.P., A.D., J.H., M.L., and E.M. performed laboratory analyses; S.P., A.D., J.H., M.L., P.K.-R., and T.O. analyzed the data; M.L., S.L., P.K.-R., R.J., J.A.L., T.O., and K.H.P. contributed significantly to the interpretation of the data. S.P., A.D., and J.H. wrote the manuscript; all authors revised the manuscript.

## REFERENCES

1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. 2016. Global epidemiology of nonalcoholic fatty liver disease—Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 64: 73–84. <https://doi.org/10.1002/hep.28431>.
2. Brunt EM, Wong VW, Nobili V, Day CP, Sookoian S, Maher JJ, Bugianesi E, Sirlin CB, Neuschwander-Tetri BA, Rinella ME. 2015. Nonalcoholic fatty liver disease. *Nat Rev Dis Primers* 1. <https://doi.org/10.1038/nrdp.2015.80>.
3. Byrne CD, Targher G. 2015. NAFLD: a multisystem disease. *Journal of Hepatology* 62:S47–S64. <https://doi.org/10.1016/j.jhep.2014.12.012>.
4. Portillo-Sanchez P, Bril F, Maximov M, Lomonaco R, Biernacki D, Orsak B, Subbarayan S, Webb A, Hecht J, Cusi K. 2015. High prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus and normal plasma aminotransferase levels. *J Clin Endocrinol Metab* 100:2231–2238. <https://doi.org/10.1210/jc.2015-1966>.
5. Targher G, Bertolini L, Padovani R, Rodella S, Tessari R, Zenari L, Day C, Arcaro G. 2007. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. *Diabetes Care* 30:1212–1218. <https://doi.org/10.2337/dc06-2247>.
6. Polyzos SA, Kountouras J, Mantzoros CS. 2019. Obesity and nonalcoholic fatty liver disease: from pathophysiology to therapeutics. *Metabolism* 92: 82–97. <https://doi.org/10.1016/j.metabol.2018.11.014>.
7. Vos MB, Lavine JE. 2013. Dietary fructose in nonalcoholic fatty liver disease. *Hepatology* 57:2525–2531. <https://doi.org/10.1002/hep.26299>.
8. Fisher CD, Lickteig AJ, Augustine LM, Ranger-Moore J, Jackson JP, Ferguson SS, Cherrington NJ. 2009. Hepatic cytochrome P450 enzyme alterations in humans with progressive stages of non-alcoholic fatty liver disease. *Drug Metab Dispos* 37:2087–2094. <https://doi.org/10.1124/dmd.109.027466>.
9. Sender R, Fuchs S, Milo R. 2016. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol* 14:e1002533. <https://doi.org/10.1371/journal.pbio.1002533>.
10. Munukka E, Pekkala S, Wiklund P, Rasool O, Borra R, Kong L, Ojanen X, Cheng SM, Roos C, Tuomela S, Alen M, Lahesmaa R, Cheng S. 2014. Gut-adipose tissue axis in hepatic fat accumulation in humans. *c61*:132–138. <https://doi.org/10.1016/j.jhep.2014.02.020>.
11. Munukka E, Rintala A, Toivonen R, Nylund M, Yang B, Takanen A, Hänninen A, Vuopio J, Huovinen P, Jalkanen S, Pekkala S. 2017. Faecalibacterium prausnitzii treatment improves hepatic health and reduces adipose tissue inflammation in high-fat fed mice. *ISME J* 11:1667–1679. <https://doi.org/10.1038/ismej.2017.24>.
12. Mouzaki M, Comelli EM, Arendt BM, Bonengel J, Fung SK, Fischer SE, Mcgilvray ID, Allard JP. 2013. Intestinal microbiota in patients with nonalcoholic fatty liver disease. *Hepatology* 58:120–127. <https://doi.org/10.1002/hep.26319>.
13. Testerman T, Li Z, Galuppo B, Graf J, Santoro N. 2021. Insights from shotgun metagenomics into bacterial species and metabolic pathways associated with NAFLD in obese youth. *RS*. <https://doi.org/10.21203/rs.3.rs-1065165/v1>.
14. Ponziani FR, Bhoori S, Castelli C, Putignani L, Rivoltini L, del Chierico F, Sanguinetti M, Morelli D, Paroni Sterbini F, Petito V, Reddel S, Calvani R, Camisaschi C, Picca A, Tuccitto A, Gasbarrini A, Pompili M, Mazzaferro V. 2019. Hepatocellular carcinoma is associated with gut microbiota profile and inflammation in nonalcoholic fatty liver disease. *Hepatology* 69:107–120. <https://doi.org/10.1002/hep.30036>.
15. Boursier J, Mueller O, Barret M, Machado M, Fizanne L, Araujo-Perez F, Guy CD, Seed PC, Rawls JF, David LA, Hunault G, Oberti F, Calès P, Diehl AM. 2016. The severity of nonalcoholic fatty liver disease is associated with gut dysbiosis and shift in the metabolic function of the gut microbiota. *Hepatology* 63:764–775. <https://doi.org/10.1002/hep.28356>.
16. Koh A, Bäckhed F. 2020. From association to causality: the role of the gut microbiota and its functional products on host metabolism. *Molecular Cell* 78:584–596. <https://doi.org/10.1016/j.molcel.2020.03.005>.
17. Hines IN, Son G, Kremer M. 2010. Contribution of gut bacteria to liver pathobiology. *Gastroenterol Res Pract* 2010:453563. <https://doi.org/10.1155/2010/453563>.
18. Lee G, You HJ, Bajaj JS, Joo SK, Yu J, Park S, Kang H, Park JH, Kim JH, Lee DH, Lee S, Kim W, Ko GP. 2020. Distinct signatures of gut microbiome and metabolites associated with significant fibrosis in non-obese NAFLD. *Nat Commun* 4982:1–13. <https://pubmed.ncbi.nlm.nih.gov/33020474/>.
19. Rau M, Rehman A, Dittrich M, Groen AK, Hermanns HM, Seyfried F, Beyersdorf N, Dandekar T, Rosenstiel P, Geier A. 2018. Fecal SCFAs and SCFA-producing bacteria in gut microbiome of human NAFLD as a putative link to systemic T-cell activation and advanced disease. *United European Gastroenterol J* 6:1496–1507. <https://doi.org/10.1177/2050640618804444>.
20. Michail S, Lin M, Frey MR, Fanter R, Paliy O, Hilbush B, Reo NV. 2015. Altered gut microbial energy and metabolism in children with non-alcoholic fatty liver disease. *FEMS Microbiol Ecol* 91:1–9. <https://doi.org/10.1093/femsec/fiu002>.
21. Arab JP, Karpen SJ, Dawson PA, Arrese M, Trauner M. 2017. Bile acids and nonalcoholic fatty liver disease: molecular insights and therapeutic perspectives. *Hepatology* 65:350–362. <https://doi.org/10.1002/hep.28709>.
22. Canfora EE, Meex RCR, Venema K, Blaak EE. 2019. Gut microbial metabolites in obesity, NAFLD and T2DM. *Nature Reviews Endocrinology* 15: 261–273. <https://doi.org/10.1038/s41574-019-0156-z>.
23. Mayo R, Crespo J, Martínez-Arrián I, Banales JM, Arias M, Mincholé I, Aller de la Fuente R, Jimenez-Agüero R, Alonso C, de Luis DA, Vitek L, Stritesky J, Caballería J, Romero-Gómez M, Martín-Duce A, Muguéza Huguet JM, Busteros-Moraza JI, Idowu MO, Castro A, Martínez-Chantar ML, Ortiz P, Bruha R, Lu SC, Bedossa P, Noureddin M, Sanyal AJ, Mato JM. 2018. Metabolomic-based noninvasive serum test to diagnose nonalcoholic steatohepatitis: results from discovery and validation cohorts. *Hepatology* 67:807–820. <https://doi.org/10.1002/hep4.1188>.
24. Zhou Y, Orešić M, Leivonen M, Gopalacharyulu P, Hyysalo J, Arola J, Verrijken A, Francque S, Van Gaal L, Hyötyläinen T, Yki-Järvinen H. 2016. Noninvasive Detection of Nonalcoholic Steatohepatitis Using Clinical Markers and Circulating Levels of Lipids and Metabolites. *Clin Gastroenterol Hepatol* 14: 1463–1472.e6. <https://doi.org/10.1016/j.cgh.2016.05.046>.
25. Perakakis N, Polyzos SA, Yazdani A, Sala-Vila A, Kountouras J, Anastasilakis AD, Mantzoros CS. 2019. Non-invasive diagnosis of non-alcoholic steatohepatitis and fibrosis with the use of omics and supervised learning: a proof of concept study. *Metabolism: Clinical and Experimental* 101:154005. <https://doi.org/10.1016/j.metabol.2019.154005>.
26. Orešić M, Hyötyläinen T, Kotronen A, Gopalacharyulu P, Nygren H, Arola J, Castillo S, Mattila I, Hakkarainen A, Borra R, Honka MJ, Verrijken A, Francque S, Iozzo P, Leivonen M, Jaser N, Juuti A, Sørensen TIA, Nuutila P, van Gaal L, Yki-Järvinen H. 2013. Prediction of non-alcoholic fatty-liver disease and liver fat content by serum molecular lipids. *Diabetologia* 56: 2266–2274. <https://doi.org/10.1007/s00125-013-2981-2>.
27. Rodríguez-Gallego E, Guirro M, Riera-Borrull M, Hernández-Aguilera A, Mariné-Casadó R, Fernández-Arroyo S, Beltrán-Debón R, Sabench F, Hernández M, del Castillo D, Menendez JA, Camps J, Ras R, Arola L, Joven J. 2015. Mapping of the circulating metabolome reveals  $\alpha$ -ketoglutarate as a predictor of morbid obesity-associated non-alcoholic fatty liver disease. *Int J Obes (Lond)* 39:279–287. <https://doi.org/10.1038/ijo.2014.53>.
28. Khusli RD, Cioffi CE, Caltharp SA, Krasinskas AM, Alazraki A, Knight-Scott J, Cleeton R, Castillo-Leon E, Jones DP, Pierpont B, Caprio S, Santoro N, Akil A, Vos MB. 2019. Development of a plasma screening panel for pediatric nonalcoholic fatty liver disease using metabolomics. *Hepatology* 69:1311–1321. <https://doi.org/10.1002/hep4.1417>.
29. Hakooz N. 2009. Caffeine metabolic ratios for the in vivo evaluation of CYP1A2, N-acetyltransferase 2, xanthine oxidase and CYP2A6 enzymatic activities. *Curr Drug Metab* 10:329–338. <https://doi.org/10.2174/138920009788499003>.
30. Newman MEJ, Girvan M. 2003. Finding and evaluating community structure in networks. *Physical Rev E - Statistical, Nonlinear, and Soft Matter Physics* 026113:1–15.
31. Katsiki N, Mikhaillidis DP, Mantzoros CS. 2016. Non-alcoholic fatty liver disease and dyslipidemia: an update. *Metabolism: Clinical and Experimental* 65:1109–1123. <https://doi.org/10.1016/j.metabol.2016.05.003>.
32. Toledo FGS, Sniderman AD, Kelley DE. 2006. Influence of hepatic steatosis (fatty liver) on severity and composition of dyslipidemia in type 2 diabetes. *Diabetes Care* 29:1845–1850. <https://doi.org/10.2337/dc06-0455>.
33. Fabbrini E, Magkos F, Mohammed BS, Pietka T, Abumrad NA, Patterson BW, Okunade A, Klein S. 2009. Intrahepatic fat, not visceral fat, is linked with metabolic complications of obesity. *Proc Natl Acad Sci U S A* 106: 15430–15435. <https://doi.org/10.1073/pnas.0904944106>.
34. Gastaldelli A, Cusi K, Pettiti M, Hardies J, Miyazaki Y, Berria R, Buzzigoli E, Sironi AM, Cersosimo E, Ferrannini E, DeFronzo RA. 2007. Relationship Between hepatic/visceral fat and hepatic insulin resistance in nondiabetic and type 2 diabetic subjects. *Gastroenterology* 133:496–506. <https://doi.org/10.1053/j.gastro.2007.04.068>.
35. Yki-Järvinen H. 2010. Liver fat in the pathogenesis of insulin resistance and type 2 diabetes. *Dig Dis* 28:203–209. <https://doi.org/10.1159/000282087>.
36. Astbury S, Atallah E, Vijay A, Aithal GP, Grove JJ, Valdes AM. 2020. Lower gut microbiome diversity and higher abundance of proinflammatory genus *Collinsella* are associated with biopsy-proven nonalcoholic steatohepatitis. *Gut Microbes* 11:569–580. <https://doi.org/10.1080/19490976.2019.1681861>.
37. bin Lee C, Chae SU, Jo SJ, Jerng UM, Bae SK. 2021. The relationship between the gut microbiome and metformin as a key for treating type 2 diabetes mellitus. *Int J Molecular Sciences* 22:3566.
38. Lensu S, Pariyani R, Mäkinen E, Yang B, Saleem W, Munukka E, Lehti M, Driuchina A, Lindén J, Tirola M, Lahti L, Pekkala S. 2020. Prebiotic xylo-

- oligosaccharides ameliorate high-fat-diet-induced hepatic steatosis in rats. *Nutrients* 12:3225–3223. <https://doi.org/10.3390/nu12113225>.
39. Troisi J, Pierri L, Landolfi A, Marciano F, Bisogno A, Belmonte F, Palladino C, Guercio Nuzio S, Campiglia P, Vajro P. 2017. Urinary metabolomics in pediatric obesity and NAFLD identifies metabolic pathways/metabolites related to dietary habits and gut-liver axis perturbations. *Nutrients* 9:485. <https://doi.org/10.3390/nu9050485>.
40. Usmani KA, Tang J. 2004. Human Cytochrome P450: metabolism of Testosterone by CYP3A4 and Inhibition by Ketoconazole. *CP Toxicology* 20: 4.13.1–4.13.9. <https://doi.org/10.1002/0471140856.tx0413s20>.
41. Loftfield E, Rothwell JA, Sinha R, Keski-Rahkonen P, Robinot N, Albanes D, Weinstein SJ, Derkach A, Sampson J, Scalbert A, Freedman ND. 2020. Prospective investigation of serum metabolites, coffee drinking, liver cancer incidence, and liver disease mortality. *J Natl Cancer Inst* 112:286–294. <https://doi.org/10.1093/jnci/djz122>.
42. Benowitz NL, Jacob P, Mayan H, Denaro C. 1995. Sympathomimetic effects of paraxanthine and caffeine in humans. *Clin Pharmacol Ther* 58: 684–691. [https://doi.org/10.1016/0009-9236\(95\)90025-X](https://doi.org/10.1016/0009-9236(95)90025-X).
43. Holstege A, Staiger M, Haag K, Gerok W. 1989. Correlation of caffeine elimination and Child's classification in liver cirrhosis. *Klin Wochenschr* 67:6–15. <https://doi.org/10.1007/BF01736528>.
44. Jamwal R, Barlock BJ. 2020. Nonalcoholic fatty liver disease (Nafld) and hepatic cytochrome p450 (cyp) enzymes. *Pharmaceuticals Multidisciplinary Digital Publishing Institute* 13:222. <https://doi.org/10.3390/ph13090222>.
45. Wuensch T, Heucke N, Wizenty J, Quint J, Sinn B, Arsenic R, Jara M, Kaffarnik M, Pratschke J, Stockmann M. 2019. Hepatic CYP1A2 activity in liver tumors and the implications for preoperative volume-function analysis. *Am J Physiol Gastrointest Liver Physiol* 316:G608–G614. <https://doi.org/10.1152/ajpgi.00335.2018>.
46. Dong S, Zhan ZY, Cao HY, Wu C, Bian YQ, Li JY, Cheng GH, Liu P, Sun MY. 2017. Urinary metabolomics analysis identifies key biomarkers of different stages of nonalcoholic fatty liver disease. *World J Gastroenterol* 23: 2771–2784. <https://doi.org/10.3748/wjg.v23.i15.2771>.
47. Jiao N, Baker SS, Chapa-Rodriguez A, Liu W, Nugent CA, Tsompana M, Mastrandrea L, Buck MJ, Baker RD, Genco RJ, Zhu R, Zhu L. 2018. Suppressed hepatic bile acid signalling despite elevated production of primary and secondary bile acids in NAFLD. *Gut* 67:1881–1891. <https://doi.org/10.1136/gutjnl-2017-314307>.
48. Mouzaki M, Wang AY, Bandsma R, Comelli EM, Arendt BM, Zhang L, Fung S, Fischer SE, McGilvray IG, Allard JP. 2016. Bile acids and dysbiosis in non-alcoholic fatty liver disease. *PLoS One* 11:e0151829. <https://doi.org/10.1371/journal.pone.0151829>.
49. Zhang X, Deng R. 2019. Dysregulation of bile acids in patients with NAFLD nonalcoholic fatty liver disease—An update. IntechOpen.
50. Chiang JYL, Ferrell JM. 2020. Up to date on cholesterol 7 alpha-hydroxylase (CYP7A1) in bile acid synthesis. *Liver Res* 4:47–63. <https://doi.org/10.1016/j.livres.2020.05.001>.
51. Min HK, Kapoor A, Fuchs M, Mirshahi F, Zhou H, Maher J, Kellum J, Warnick R, Contos MJ, Sanyal AJ. 2012. Increased hepatic synthesis and dysregulation of cholesterol metabolism is associated with the severity of nonalcoholic fatty liver disease. *Cell Metab* 15:665–674. <https://doi.org/10.1016/j.cmet.2012.04.004>.
52. Levy JC, Matthews DR, Hermans MP. 1998. Correct homeostasis model assessment (HOMA) evaluation uses the computer program. *Diabetes Care*. American Diabetes Association <https://doi.org/10.2337/diacare.21.12.2191>.
53. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. 1985. Homeostasis model assessment: insulin resistance and  $\beta$ -cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412–419. <https://doi.org/10.1007/BF00280883>.
54. Klävus A, Kokla M, Noerman S, Koistinen VM, Tuomainen M, Zarei I, Meuronen T, Häkkinen MR, Rummukainen S, Babu AF, Sallinen T, Kärkkäinen O, Paananen J, Broadhurst D, Brunius C, Hanhineva K. 2020. “notame”: workflow for Non-Targeted LC-MS Metabolic Profiling. *Metabolites* 10:135. <https://doi.org/10.3390/metabo10040135>.
55. Noerman S, Kokla M, Koistinen VM, Lehtonen M, Tuomainen TP, Brunius C, Virtanen JK, Hanhineva K. 2021. Associations of the serum metabolite profile with a healthy Nordic diet and risk of coronary artery disease. *Clin Nutr* 40:3250–3262. <https://doi.org/10.1016/j.clnu.2020.10.051>.