

Serum lipidomics as diagnostic potential for metabolic-associated hepatocellular carcinoma

Monika Lewinska

University of Copenhagen <https://orcid.org/0000-0002-5275-992X>

Alvaro Santos-Laso

Donostia University Hospital

Enara Arretxe

OWL Metabolomics, Bizkaia Technology Park, Derio

Cristina Alonso

<https://orcid.org/0000-0002-2019-678X>

Raul Jimenez-Aguero

Donostia University Hospital

Emma Eizaguirre

Donostia University Hospital

María Pareja

Hospital Juan Ramón Jiménez

Manuel Romero-Gomez

Virgen Macarena-Virgen del Rocío University Hospitals

Marco Arrese

Pontificia Universidad Católica de Chile

Malte Suppli

University of Copenhagen, Center for Clinical Metabolic Research, Gentofte Hospital

Filip Knop

Gentofte Hospital, University of Copenhagen <https://orcid.org/0000-0002-2495-5034>

Stine Oversoe

Aarhus University Hospital <https://orcid.org/0000-0001-9762-1679>

Gerda Villadsen

Aarhus University Hospital

Thomas Decaens

Université Grenoble Alpes

Flair Carrilho

Universidade de São Paulo

Claudia Oliveira

Universidade de São Paulo

Bruno Sangro

National Institute of Health, Clinica Universidad de Navarra

Rocio Macias

National Institute of Health, University of Salamanca <https://orcid.org/0000-0002-4748-0326>

Jesus Banales

National Institute of Health, Donostia University Hospital

Jesper Andersen (✉ jesper.andersen@bric.ku.dk)

University of Copenhagen <https://orcid.org/0000-0003-1760-5244>

Article

Keywords: Metabolic-associated fatty liver disease (MAFLD), metabolic dysfunctions, hepatocellular carcinoma (MAFLD-HCC).

Posted Date: January 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-140974/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Metabolic-associated fatty liver disease (MAFLD) is affecting more people globally. Indeed, MAFLD is associated with a spectrum of metabolic dysfunctions that can progress to hepatocellular carcinoma (MAFLD-HCC). This development can occur in a non-cirrhotic liver and thus, often lack clinical surveillance. Using comprehensive ultra-high-performance liquid chromatography mass-spectrometry, we investigated 1,295 metabolites in serum from 249 Caucasian liver patients. Here we show that MAFLD-HCC is characterized by a complete rearrangement of the serum lipidome which distinguishes MAFLD-HCC from other HCC patients. We used machine learning to build a diagnostic model for MAFLD to MAFLD-HCC. We quantified predictive metabolites and adjusted the model into the MAFLD-HCC Diagnostic Score, presenting superior diagnostic potential compared to current alpha-fetoprotein (AFP). The metabolic landscape shows a progressive depletion in unsaturated fatty acids and acylcarnitines during transformation. Therefore, serum metabolomics may provide valuable insight to monitor patients at risk, including morbidly obese, diabetics, and MAFLD patients.

Introduction

Non-alcoholic fatty liver disease (NAFLD) involves a spectrum of metabolic diseases affecting 24% of the population globally¹. NAFLD is associated with obesity and type 2 diabetes², and represents a spectrum of non-malignant conditions that range from hepatic steatosis to non-alcoholic steatohepatitis (NASH) with fibrosis and ultimately cirrhosis³. NAFLD is emerging as a leading risk factor of hepatocellular carcinoma (HCC) in both men and women (hazard ratio ~17)⁴, but tools to detect progression of NAFLD to HCC remain inadequate. Recently, the definition of NAFLD was suggested to change into metabolic-associated fatty liver disease (MAFLD), recognizing that this disease is associated with metabolic dysfunction⁵, and acknowledging the multiple overlapping risk factors of the disease.

HCC is the fourth most common cause of cancer-related death worldwide⁶ with rapidly increasing incidence and mortality rates^{7,8}. The incidence of MAFLD-related HCC (MAFLD-HCC) varies between 0.5% and 19.5%, depending on the disease state (presence or absence of cirrhosis) and geography⁸⁻¹⁰. Indeed, the prevalence of MAFLD-HCC is increasing compared to alcohol- and viral-related HCC (AV-HCC)^{9,11}. Importantly, accumulating evidence suggests that MAFLD HCC may develop in a non-cirrhotic background¹⁰. This presents a major clinical challenge as non-cirrhotic MAFLD patients currently are not under surveillance for HCC⁸. Therefore, non-invasive biomarkers are urgently needed for the monitoring of MAFLD and its progression to HCC.

Serum metabolomics provides insight into holistic metabolic changes and can potentially predict hepatic pathological lipid abundance¹². However, metabolomic profiles are highly dependent on ethnicity and risk factors such as obesity and diabetes¹³⁻¹⁶. The understanding of the biology underlying MAFLD-to-HCC progression is still lacking. Therefore, inquiry of the metabolic shifts in the liver during malignant progression is crucial for the development of new diagnostics and therapeutic modalities¹⁷.

Here, we used a comprehensive serum metabolomics-based approach to identify unique MAFLD-HCC biomarkers, allowing us to distinguish MAFLD-HCC from patients with MAFLD or HCC on a background of alcohol or viral hepatitis (AV-HCC) in the Caucasian population.

Patients And Methods

Study population: A total of 249 patient samples were collected in this international and multicenter study. Our study cohort was divided into a discovery and a validation sets. The discovery set included serum samples from 196 patients from Spain and France: 27 patients with MAFLD-HCC, 32 patients with AV-HCC, 93 morbidly obese MAFLD patients undergoing bariatric surgery (OB-MAFLD), and 44 (CTRLs) that included 35 healthy subjects described previously¹⁸ and 9 obese bariatric surgery patients with a non-alcoholic fatty liver disease (NAFLD)-activity score (NAS) less than 3 and liver fibrosis score <2. The validation set included serum samples from 37 MAFLD patients from Chile and Spain and plasma sampled from 16 MAFLD-HCC patients in Brazil and Denmark. MAFLD patients from the validation set were overweight (body mass index (BMI)>25) or obese (BMI) >30), however significantly leaner compared to patients in the OB-MAFLD group ($p<0.0001$). All MAFLD patients had biopsy-proven non-alcoholic fatty liver disease (NAFLD). The MAFLD-HCC group consisted of patients who were diagnosed non-alcoholic steatohepatitis (NASH) based on liver biopsy, self-reported alcohol consumption (<20 g/day) and hepatitis B or C serology (hepatitis B surface antigen, hepatitis B surface antibody, hepatitis B core antibody, and hepatitis C antibody). All patients were qualified for curative liver resection. The clinical and biochemical representation of the study population is presented in Table 1. The study was performed following individual patient consent, local institutional review board (IRB), and approval from the Committees on Health and Research Ethics for the Capital Region of Denmark for use of archival material no. 17029679. All patient data sets were anonymized.

Metabolomic and statistical analyses: Comprehensive metabolomics including 1,295 metabolites was performed on serum samples with three platforms as described in detail in supplementary materials. The absolute quantification of selected metabolites was performed using the calibration curves and addition of heavy isotope-labeled standards. Metabolites with missing signal in >50% of samples were excluded from analysis. Otherwise, missing values were estimated by the k-nearest neighbor method with Metaboanalyst 4.0¹⁹. Data were quantile-normalized and log-transformed before analysis (SPSS 25.0.0). Additional details, including statistical analyses, can be found in the supplementary material.

Results

Clinical and biochemical features of studied populations

We analyzed the metabolic composition of serum obtained from 249 Caucasian patients divided into a discovery set comprising patients with MAFLD-HCC with pre-diagnosed NASH (n=27), patients with alcohol- or viral-related HCC (AV-HCC) (n=32), and non-cancerous controls (n=137). These controls included 44 healthy individuals (35 healthy subjects reported previously¹⁸ and 9 bariatric surgery patients with NAS<3, and fibrosis score <2) and 93 morbidly obese MAFLD patients awaiting bariatric surgery (OB-MAFLD). Furthermore, we independently validated our findings in serum from 37 patients with MAFLD and in plasma

from 16 patients with MAFLD-HCC. Importantly, all MAFLD-HCC patients had no prior history of viral hepatitis or excessive alcohol consumption. The clinical, pathological, and biochemical features of all patients are summarized in Table 1 and Supp. Fig. 1. We observed a significant difference in mean age (MAFLD-HCC patients were significantly older compared to CTRL, OB-MAFLD and AV-HCC), BMI (significantly higher in OB-MAFLD compared to CTRL, AV-HCC and MAFLD-HCC), as well as male:female ratio (significantly higher prevalence of HCC among males, but equal between MAFLD-HCC and AV-HCC). Nevertheless, unsupervised principal component analysis (PCA) showed that the metabolic profiles grouped independently of these covariates (Supp. Fig. 2). Additionally, the presence of underlying diabetes, cirrhosis, and level of fibrosis were not confounders. Further, the analysis of alpha-fetoprotein (AFP) or other liver biochemical features (alanine aminotransferase (ALT), gamma-glutamyl transferase (GGT), alkaline phosphatase (AP), bilirubin, albumin and prothrombin activity) to follow the liver function showed extensive variability already in non-cancerous patients, but generally remained within the reference ranges (Supp. Fig. 1). As such, these markers presented limited potential for diagnosing MAFLD-HCC (Table 2). To investigate whether metabolomic changes in MAFLD-HCC are etiology-specific, we compared the profiles of MAFLD-HCC patients and 32 patients with AV-HCC. As such, the metabolite profiles of patients with AV-HCC showed significant overlap, but a clear separation from the profiles of patients with MAFLD-HCC (Supp. Fig. 2). Compared to patients with AV-HCC, MAFLD-HCC patients were older ($p < 0.05$) and less likely to develop HCC on a cirrhotic background (respectively 90% compared to 30%, $p = 0.035$). However, no significant differences were observed between MAFLD-HCC and AV-HCC in measurements of the liver function (AFP, AP, ALT, GGT, and bilirubin) or diabetes. Interestingly, tumors obtained from MAFLD-HCC patients were larger in size compared to tumors from patients with AV-HCC ($p = 0.0006$) and more frequently displayed microvascular invasion (41% MAFLD-HCC compared to 20% AV-HCC, $p = 0.016$).

MAFLD-HCC patients present a disparate serum metabolome

To establish a serum-based metabolomic landscape of MAFLD-HCC, we performed detailed metabolomics using a comprehensive library of 1,295 metabolites covering amino acids (AA), glycerophospholipids, fatty acyls, sterols, sphingolipids, and glycerolipids. In total, we detected 470 metabolites, of which 43 were excluded after correcting for age, gender, and BMI, before univariate and multivariate analyses.

As such, sparse partial least squares discriminant analysis (sPLS-DA) revealed that MAFLD-HCC patients metabolically are the most distinct group compared to controls (CTRL and OB-MAFLD) and patients with AV-HCC (Fig. 1A). Besides, MAFLD-HCC patients were the most dissimilar using unsupervised hierarchical clustering (Fig. 1B). Interestingly, AV-HCC and MAFLD-HCC subgroups showed low similarity and clustered far apart (Fig. 1A), suggesting that unique metabolic programs can be driven by different etiologies.

Unique metabolomic profile of MAFLD-HCC patients

To investigate differences in the metabolic profiles between MAFLD-HCC patients and the other groups, we performed a series of pair-wise tests. As such, orthogonal partial least squares discriminant analysis (oPLS-DA) showed substantial separation of MAFLD-HCC patients from CTRL ($R^2X = 0.178$, $R^2Y = 0.751$, $Q^2 = 0.723$) (Fig. 2A). The differential expression analysis identified 274 significantly different (false-discover rate (FDR)

corrected $p < 0.05$) metabolites (DEMs), among which 152 metabolites were depleted and 122 metabolites were increased (Fig. 2B, Supp. Table 1A). Next, we performed pathway overrepresentation analysis of the depleted metabolites using integrated molecular pathway level analysis (IMPALA)²⁰ and identified linoleic acid metabolism and G protein-coupled receptor (GPCR) signaling as the most impaired networks (Supp. Table 1B). Contrary, cholesterol synthesis, membrane fluidity and trafficking, and glycerophospholipid metabolism were among the most upregulated pathways (Supp. Table 1C). In addition to relate individual metabolites to processes, we compared unique classes of metabolites with the same chemical characteristics. As such, we defined a unique depletion of acylcarnitines (AC), sterol lipids (ST), and fatty acids (FA; especially oxidized fatty acids (oxFA) and omega-6 FA) in MAFLD-HCC, while saturated triglycerides (TG) were upregulated (Fig. 2C, Supp. Fig. 3). Furthermore, we utilized Bioinformatics Methodology For Pathway Analysis (BioPAN)²¹ for lipid pathway enrichment analysis. We observed a significant activation of reactions converting sphingomyelins (SM) to ceramides (Cer), a process catalyzed by sphingomyelin phosphodiesterases (*SMPD2* and *SMPD3*), as well as phosphatidylcholines (PC) to diglycerides (DG), which is catalyzed by the sphingomyelin synthases (*SGMS1* and *SGMS2*) (Supp. Fig. 4A). Similarly, significant alterations were revealed in the activity of FA desaturases and elongases with specific activation of fatty acid desaturase 1 (*FADS1*) and impairment in *FADS2*, stearoyl-CoA desaturase 1 (*SCD1*) and elongation of very long chain fatty acid (ELOVL) elongases (*ELOVL2*, and *ELOVL5*) (Supp. Fig. 6A).

Next, we compared OB-MAFLD and MAFLD-HCC patients. Indeed, oPLS-DA analysis showed moderate separation ($R^2X=0.173$, $R^2Y=0.715$, $Q^2=0.697$) (Fig. 2D). The differential expression analysis identified 316 DEMs with 154 metabolites upregulated and 162 downregulated (Fig. 2E, Supp. Table 2A). The overrepresentation analysis identified cholesterol and BA metabolism, aminoacyl-tRNA biosynthesis, as well as protein and glucose metabolism as significantly upregulated in MAFLD-HCC compared to OB-MAFLD. Contrary, FA biosynthesis and GPCR signaling were downregulated (Supp. Table 2B-C). Furthermore, OB-MAFLD significantly differed from MAFLD-HCC in specific classes of metabolites. As such, AA, BA, Cer, TG, PC, and phosphatidylethanolamines (PE) were all significantly upregulated in MAFLD-HCC compared to OB-MAFLD. Contrary, FA and AC were reduced in MAFLD-HCC (Fig. 2C, Supp. Fig. 3). Moreover, the pathway enrichment analysis identified two reaction paths (SM@Cer and DG@PE) as the most dynamic, whereas the opposite direction (Cer@SM and PE@PC@DG) was repressed (Supp. Fig. 4B). Moreover, *ELOVL3* and *ELOVL6* presented an increased activity, whereas *ELOVL2*, *SCD1*, and *SCD3* were suppressed (Supp. Fig. 5B). Contrary, reactions catalyzed by *FADS1*, *FADS2*, and *ELOVL5* displayed a mixed activity.

Lastly, we compared the metabolomes of MAFLD-HCC to HCCs with alcohol and/or viral etiology (AV-HCC). The oPLS-DA model showed moderate separation ($R^2X=0.171$, $R^2Y=0.739$, $Q^2=0.694$) (Fig. 2F). As expected, the AV-HCC group showed greater heterogeneity compared to MAFLD-HCC, which partly can be due to the mixed etiology. Differential expression analysis identified 257 DEMs (125 downregulated and 132 upregulated) in MAFLD-HCC compared to AV-HCC (Fig. 2G, Supp. Table 3A), showing an overall impairment in the free FA (FFA) biosynthesis. Contrary, insulin resistance, glycerophospholipid, and choline metabolism were upregulated in MAFLD-HCC (Supp. Table 3B-C). Next, we compared metabolic classes between MAFLD-HCC and AV-HCC patients and observed significant depletion of AC, BA, FA, ST, lysophosphatidylethanolamines (LPE), and phosphatidylinositols (PI) in MAFLD-HCC. Contrary, cholesteryl

esters (ChoE), SM, Cer, and TG were all increased in MAFLD-HCC compared to AV-HCC (Fig. 2C, Supp. Fig. 3). Moreover, lipid pathway enrichment analysis showed an elevated reaction activity in the path SM@Cer (Supp. Fig. 4C) with the enzyme activities of *FADS2*, *ELOVL5*, and *ELOVL2* significantly reduced (Supp. Fig. 5C).

Taken together, the serum of MAFLD-HCC patients is characterized by a significant depletion of FA reflective of a significantly lower FA biosynthesis with decreased FA desaturase and elongase activities. A depletion in both AC and ST with concurrent higher TG and Cer abundance is suggestive of a unique metabolic reprogramming in MAFLD-HCC patients. The altered SM: Cer ratio could be the result of an increased activity in the enzymes *SMPD2* and *SMPD3* or reduced activity of *SGMS1*, *SGMS2*, *CERT1* in MAFLD-HCC patients. A simplified association between the lipid classes and their deregulation in MAFLD-HCC is presented in Fig. 2H.

Diagnostic potential of serum metabolomics

Serum metabolomics has been successfully used as a diagnostic tool to discriminate liver diseases^{17,18}. Here, we investigated the potential of distinguishing MAFLD-HCC not only from healthy individuals and MAFLD patients, but also from AV-HCC. Thus, to generate a predictive metabolite signature, we first used receiver operating characteristic (ROC) curves and calculated area under the curve (AUC) for each metabolite as a contrast test between MAFLD-HCC and the respective comparative groups (CTRL, OB-MAFLD, and AV-HCC). As such, 89 metabolites presented an AUC>0.75 distinguishing MAFLD-HCC patients from the other control patients (healthy or disease). Among the metabolites in the DEM signature, 14 metabolites presented a superior AUC>0.9 in all contrast tests (Table 2). These metabolites and their fold change compared to CTRL are presented in Fig. 3A. Importantly, these 14 metabolites individually present AUCs markedly superior to alpha-fetoprotein (AFP: $AUC^{\text{MAFLD-HCC vs CTRL}}=0.791$ and $AUC^{\text{MAFLD-HCC vs AV-HCC}}=0.614$) and all biochemical measurements for the liver function (ALT: $AUC^{\text{MAFLD-HCC vs CTRL}}=0.776$, $AUC^{\text{MAFLD-HCC vs MAFLD}}=0.733$, and $AUC^{\text{MAFLD-HCC vs AV-HCC}}=0.570$ or GGT: $AUC^{\text{MAFLD-HCC vs CTRL}}=0.927$, $AUC^{\text{MAFLD-HCC vs MAFLD}}=0.865$, and $AUC^{\text{MAFLD-HCC vs AV-HCC}}=0.562$), discriminating MAFLD-HCCs from all controls and HCCs with a different etiology.

Next, we assessed if a combination of the 14 metabolites would increase the diagnostic potential. As such, we employed support vector machine (SVM) modeling and determined that a panel of 5 metabolites yielded the optimal predictive accuracy (Supp. Fig. 6A). Indeed, the model based on the 5 metabolites reached an AUC>0.98 (Fig. 3B) for any of the contrasts (compared to all controls and AV-HCC) with a predictive accuracy greater than 90% (Fig. 3C). Importantly, the accuracy of the diagnostic panel was confirmed in the validation set and a model performance with an AUC of 0.91, including matching MAFLD-HCC patients according to BMI (Supp. Fig. 6B).

Validation and quantification of diagnostic metabolites

To reinforce the clinical relevance of the metabolite panel in the diagnosis of MAFLD-HCC patients, we wanted to validate and quantify the abundance of each metabolite. Also, we established their reference concentration range. Among the 14 metabolites with the greatest diagnostic value, only 10 of them have commercial standards available. Also, each metabolite needs to be detected within the linear range, which allows for absolute quantification. As such, we established the abundance of metabolites in the validation set (Fig. 4A). All metabolites except PC(0:0/22:5) showed a similar trend as in the discovery set, but overall with great variability. Moreover, we utilized the quantification to generate linear regression models estimating the concentration levels in the discovery set. The linear dependence and estimated concentrations are presented in Supp. Table 4. Lastly, among the validated metabolites in the panel, we selected an optimal set (linoleic acid, osbond acid, monounsaturated fatty acid MUFA (14:1n-5trans), and PC(18:2/0:0)) to build the MAFLD-HCC Diagnostic Score (MHDS). The MHDS is dependent on the absolute concentration level of each metabolite in the score. The MHDS performed with an AUC>0.75 for any given contrast (Fig. 4B-C). Finally, in the combined (discovery and validation) cohort, we established the odds ratio and relative risk for the MHDS (at cut-off value 0) based on the clinical and biochemical features. Taken together, the MDHS>0 was predictive of MAFLD-HCC with the highest odds ratio OR=97.55 (95% CI=16.40 to 999.7) (Fig. 4D) and relative risk RR=1.67 (95% CI=1.45 to 1.99) (Fig. 4E).

The serum lipidome landscape reflects the progression to MAFLD-HCC

We investigated whether the serum lipidome reflects the progressive nature of MAFLD-HCC development. To that end, we first performed pair-wise contrasts between each group (Fig. 5A) to establish the trajectory of disease progression. Then, we employed a pattern hunter approach with Pearson correlation to designate the metabolic rearrangements in the development of the disease (Fig. 5B). A total of 412 lipids were detected in the discovery and validation MAFLD samples. As such, we independently compared OB-MAFLD (bariatric surgery) and MAFLD (non-bariatric surgery) patients with healthy CTRLs, resulting in 257 DEMs distinguishing OB-MAFLD from CTRLs (Fig. 5A). Likewise, in MAFLD, we defined a total of 266 DEMs compared to the CTRL group (Fig. 5B), suggesting that both OB-MAFLD and MAFLD patients experience a significant metabolic rearrangement compared to healthy individuals. Indeed, a total of 138 metabolites were significantly different from CTRL and shared similar directionality, showing a significant depletion of AC, ChoE, and LPC (Fig. 5C). Contrary, DG and TG are among metabolites progressively upregulated in MAFLD patient groups. Interestingly, omega-6 FA and Cer were among deregulated metabolite classes.

Consequently, we next compared OB-MAFLD and MAFLD patients with MAFLD-HCC detecting 297 and 136 DEMs, respectively (Fig. 5D-E). As such, we noticed a second significant metabolic shift from MAFLD to MAFLD-HCC. Notably, OB-MAFLD samples were significantly different, suggesting that the disease progression is CTRL→OB-MAFLD→MAFLD→MAFLD-HCC. Following this disease progression, we defined a total of 83 metabolites significantly different from MAFLD-HCC that shared the same directionality. This metabolic shift included an increase in PC and BA levels as well as depletion of AC, FA, and SM levels (Fig. 5F). Interestingly, TG were higher in the serum of MAFLD-HCC when compared to OB-MAFLD (patients awaiting bariatric surgery), but diminished when compared to MAFLD patients, who were significantly leaner.

To further explore the OB-MAFLD and MAFLD differences, we directly compared these groups. First, the MAFLD patients were significantly older and have a lower BMI compared to OB-MAFLD (Supp. Fig. 1). Also, 76% of MAFLD patients were classified as NASH (biopsy proven) compared to only 12% among OB-MAFLD (Table 1, Fisher's test $p < 0.0001$). As such, we detected 323 DEMs reflecting a complete change in the lipidome landscape of these patient groups (Supp. Fig. 7). FA, AC, and ST metabolites were significantly diminished in MAFLD compared to OB-MAFLD. Contrary, SM, Cer, DG, TG, PE, and PC were all significantly higher in MAFLD. Interestingly, phosphatidylethanolamine N-methyltransferase (PEMT), the only enzyme catalyzing the reaction chain $PE \rightarrow PC \rightarrow DG$, and choline/ethanolamine phosphotransferase 1 (CEPT1), catalyzing the conversion of DG to PE are the only enzymes deregulated between the major lipid subclasses. As expected, we also observed significant differences in the FA reaction chain (Supp. Fig. 8).

Finally, we used pattern hunting to identify metabolites significantly associated with the full progression from healthy individuals to HCC (CTRL \rightarrow OB-MAFLD \rightarrow MAFLD \rightarrow MAFLD-HCC axis) (Fig. 5G). As such, we found a total of 169 lipids progressively altered (FDR $p < 0.05$, $r > 0.3$; 100 DEMs negatively, and 69 positively) following this axis (Supp. Table 5). The negatively correlated metabolites include AC, ChoE, PUFA, LPC, and ST subclasses. Among the metabolites positively correlated, we found an overrepresentation of TG (47 out of 69 metabolites). Importantly, 18 TG significantly correlate with increasing NAS and fibrosis scores in OB-MAFLD patients, suggesting their importance in the progressive deterioration of the liver (Fig. 5H). Lastly, we tested whether any of these metabolites correlated with tumor size, recurrence, microvascular invasion or liver cirrhosis in MAFLD-HCC patients, however, none of the metabolites reached statistical significance. This suggests that these metabolites are associated with MAFLD-HCC, but not directly involved the progression axis.

Discussion

A sedentary lifestyle and overnutrition have led to an epidemic of obesity, diabetes, and MAFLD that soon may become leading causes of HCC development. Metabolic reprogramming is at the core of MAFLD progression to HCC¹⁷. However, due to the need of invasive techniques for MAFLD diagnosis, our understanding of MAFLD-HCC is limited and the underlying metabolomic landscape remains elusive. The present study provides the first comprehensive analysis covering 22 classes of metabolites and identifying complex relationships in the gradual progression of MAFLD to MAFLD-HCC. Importantly, this study was performed in patients with biopsy-proven liver histology before therapy. To our knowledge, it is also the first study aiming to differentially diagnose MAFLD-HCC from HCC of other etiologies.

The diagnosis of MAFLD-HCC is challenging as patients with metabolic syndrome are not screened for HCC⁸. Indeed, often only patients with chronic liver disease of viral etiology are offered HCC screening. Increasing levels of liver enzymes are suggestive of liver damage but are not specific to hepatocarcinogenesis. Additionally, the diagnostic capacity of AFP²² for HCC is limited and was particularly low for MAFLD-HCC patients (Fig. 3G, H, Supp. Fig. 1). As such, we have applied machine learning approaches to pin-point a combination of metabolites that offers the best diagnostic potential for MAFLD-HCC patients. We found that a combination of 5 metabolites accurately distinguishes MAFLD-HCC patients

from healthy individuals (AUC=0.989), OB-MAFLD patients (AUC=0.997), and patients with AV-HCC (AUC=0.999). This model performed well against a validation set of MAFLD patients (AUC=0.905). Furthermore, to bring metabolomics to clinical practice, it is crucial to establish methods for absolute quantification of metabolites as well as their reference ranges. Here, we were able to validate 10 metabolites using commercially available internal standards and measure their concentrations. As such, we built a MHDS based on the absolute concentrations of 4 metabolites (3 unsaturated FA and 1 phosphatidylcholine) that can be applied for both serum and plasma measurement and performs superior to AFP and GGT (Fig. 3E-H).

The unsaturated fatty acids (mono- (MUFA) and polyunsaturated (PUFA)) significantly differed between MAFLD-HCC and other groups were (Fig. 1C, 2B, D, F & Fig. 3A) with lower levels in MAFLD-HCC. The essential, omega-6 FA, linoleic acid, is a precursor to long-chain metabolites including arachidonic acid, as a substrate in prostaglandin synthesis and thus, their depletion may cause altered signaling and inflammatory response²³. The role of linoleic acid in carcinogenesis remains controversial. On the one hand, linoleic acid and its derivatives have shown a tumor-suppressive role in colorectal cancer²⁴. Conversely, linoleic acid accumulation (in a murine model) increases oxidative stress and causes a selective loss of intrahepatic CD4+ cells leading to MAFLD-mediated hepatocarcinogenesis²⁵. Interestingly, the serum linoleic acid was heavily depleted in MAFLD-HCC patients (Fig. 3A), suggesting that the murine model was not able to fully mimic human disease and further emphasizing the potential limitation of murine studies in this field. Indeed, increased metabolism of linoleic acid and arachidonic acid have been implicated in viral-associated HCC in the Korean population²⁶. However, the different ethnicity (Korean compared to European) might affect serum metabolome and ethnicity or culturally driven dietary differences should be considered^{15,27}. As such, sphingolipid and acylcarnitine profiles differ between European, African, and South Asian patients with diabetes, showing an increase in acylcarnitines in individuals from Suriname¹⁵. Targeted acylcarnitine profiles were recently reported as upregulated in Japanese MAFLD-HCC individuals²⁸, which is opposite to what we observed in our study (Fig. 2C). Furthermore, acylcarnitine profiles might be influenced by the prevalence of diabetes, as the authors did not control for this covariate²⁸. As such, our study was limited to patients of Caucasian origin, who are overrepresented in countries significantly affected by the prevalence of MAFLD.

HCC is a progressive disease developing over decades. We observed a gradual attenuation of acylcarnitines, lysophosphatidylcholines, and unsaturated FA and an increase in phosphatidylcholines and triglycerides during the progression of liver disease. Interestingly, some of the triglycerides that were found to be augmented during progression to MAFLD-HCC were previously associated with NAS and fibrosis scores in NASH²⁹ (Fig. 5H) but were not associated with tumor size. Although, we have presented a successful metric, demonstrating the ability of the metabolite panel in diagnosing CRTL®MAFLD®MAFLD-HCC patients, and distinguishing these patients from AV-HCC, the utility of the MHDS needs to be shown in an appropriate clinical assay, in a large cohort, and in ethnically diverse patients. However, stratification of HCC based on their metabolome may in the future be an approach to evaluate cryptogenic HCC patients. It is notable that MAFLD-HCC patients were significantly older compared to other groups which limited the power to detect metabolic differences.

In conclusion, the depletion of unsaturated FA, and the increase of triglycerides are at the core of deregulated metabolic networks in MAFLD-HCC, leading to altered signaling and likely different nutrient utilization by cancer cells. These changes can be exploited for non-invasive surveillance of the 'at risk' population for early HCC detection in the background of metabolic syndrome.

Abbreviations

acylcarnitines (AC), alcohol- and viral-associated HCC (AV-HCC), alpha-fetoprotein (AFP), amino acids (AA), aromatic amino acids (ArAA), bile acids (BA), body mass index (BMI), branch-chain amino acids (BCAAs), ceramides (Cer), cholesteryl esters (ChoE), control (CTRL), differentially expressed metabolites (DEM), diglycerides (DG), lysophosphatidylcholines (LPC), lysophosphatidylethanolamines (LPE), lysophosphatidylinositols (LPI), metabolic-associated fatty liver disease (MAFLD), monohexylceramides (CMH), monounsaturated fatty acids (MUFA), morbidly obese bariatric surgery MAFLD (OB-MAFLD), non-alcoholic fatty liver disease (NAFLD), N-acyl ethanolamines (NAE), oxidized fatty acids (oxFA), phosphatidylcholines (PC), phosphatidylethanolamines (PE), phosphatidylinositols (PI), polyunsaturated fatty acids (PUFA), saturated fatty acids (SFA), sphingomyelins (SM), sphingosine-1-phosphate (S1P), triglycerides (TG)

Declarations

Conflicts of interest: Drs. Alonso and Arretxe are employed by OWL Metabolomics (One Way Liver, S.L). Dr. Banales is a member of the scientific advisory board of OWL Metabolomics.

Grant support: The laboratory of JBA is supported by the Novo Nordisk Foundation (14040), Danish Cancer Society (R167-A10784), and the Danish Medical Research Council (4183-00118A). JMB is funded by the Spanish Ministry of Economy and Competitiveness and 'Instituto de Salud Carlos III' grants (PI18/01075, Miguel Servet Programme CON14/00129 and CPII19/00008) co-financed by 'Fondo Europeo de Desarrollo Regional' (FEDER); CIBERehd, Spain; IKERBASQUE, Basque foundation for Science, Spain; BIOEF (Basque Foundation for Innovation and Health Research: EiTB Maratoia BIO15/CA/016/BD); Department of Health of the Basque Country (2017111010), Euskadi RIS3 (2019222054, 2020333010); Department of Industry of the Basque Country (Elkartek: KK-2020/00008), AECC Scientific Foundation and European Commission Horizon 2020 program (ESCALON project no.: 825510).

Author contribution:

Conceptualization: ML, JBA. Data curation: ML, ASL, RJA, EE, MJP, MPS, FKK, MRG, MAJ, SK, GEV, TD, CPO, BS, RIR, JMB, JBA. Formal Analysis: ML, EA, CA. Writing – original draft: ML, JBA.

Data availability

The data used in this study will according to IRBs be made available upon request if the requestor has an approved protocol.

Acknowledgment

We thank the participating patients and their families as well as clinical staff at all contributing centers. We thank the French Liver Biobanks network – INCa (BB-0033-00085) for the contribution of MAFLD-HCC samples (participating hospitals: Beaujon, Bordeaux, Grenoble, Henri-Mondor, Paris-SUD, and Rennes).

References

- 1 Younossi, Z. *et al.* Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nat Rev Gastroenterol Hepatol* **15**, 11-20, doi:10.1038/nrgastro.2017.109 (2018).
- 2 Younossi, Z. M. *et al.* The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. *Hepatology* **64**, 1577-1586, doi:10.1002/hep.28785 (2016).
- 3 Hardy, T., Oakley, F., Anstee, Q. M. & Day, C. P. Nonalcoholic Fatty Liver Disease: Pathogenesis and Disease Spectrum. *Annu Rev Pathol* **11**, 451-496, doi:10.1146/annurev-pathol-012615-044224 (2016).
- 4 Kim, G. A. *et al.* Association between non-alcoholic fatty liver disease and cancer incidence rate. *J Hepatol*, doi:10.1016/j.jhep.2017.09.012 (2017).
- 5 Eslam, M., Sanyal, A. J., George, J. & an international consensus, p. MAFLD: A consensus-driven proposed nomenclature for metabolic associated fatty liver disease. *Gastroenterology*, doi:10.1053/j.gastro.2019.11.312 (2020).
- 6 Yang, J. D. *et al.* A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol* **16**, 589-604, doi:10.1038/s41575-019-0186-y (2019).
- 7 Sayiner, M., Golabi, P. & Younossi, Z. M. Disease Burden of Hepatocellular Carcinoma: A Global Perspective. *Dig Dis Sci* **64**, 910-917, doi:10.1007/s10620-019-05537-2 (2019).
- 8 Nouredin, M. & Rinella, M. E. Nonalcoholic Fatty liver disease, diabetes, obesity, and hepatocellular carcinoma. *Clin Liver Dis* **19**, 361-379, doi:10.1016/j.cld.2015.01.012 (2015).
- 9 Pais, R. *et al.* Temporal trends, clinical patterns and outcomes of NAFLD-related HCC in patients undergoing liver resection over a 20-year period. *Aliment Pharmacol Ther* **46**, 856-863, doi:10.1111/apt.14261 (2017).
- 10 Anstee, Q. M., Reeves, H. L., Kotsiliti, E., Govaere, O. & Heikenwalder, M. From NASH to HCC: current concepts and future challenges. *Nat Rev Gastroenterol Hepatol*, doi:10.1038/s41575-019-0145-7 (2019).
- 11 Bertot, L. C. & Adams, L. A. Trends in hepatocellular carcinoma due to non-alcoholic fatty liver disease. *Expert Rev Gastroenterol Hepatol* **13**, 179-187, doi:10.1080/17474124.2019.1549989 (2019).
- 12 Parker, B. L. *et al.* An integrative systems genetic analysis of mammalian lipid metabolism. *Nature* **567**, 187-193, doi:10.1038/s41586-019-0984-y (2019).

- 13 Vos, T. *et al.* Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2163-2196, doi:10.1016/S0140-6736(12)61729-2 (2012).
- 14 Cirulli, E. T. *et al.* Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk. *Cell Metab* **29**, 488-500 e482, doi:10.1016/j.cmet.2018.09.022 (2019).
- 15 van Valkengoed, I. G. M. *et al.* Ethnic differences in metabolite signatures and type 2 diabetes: a nested case-control analysis among people of South Asian, African and European origin. *Nutr Diabetes* **7**, 300, doi:10.1038/s41387-017-0003-z (2017).
- 16 Floegel, A. *et al.* Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* **62**, 639-648, doi:10.2337/db12-0495 (2013).
- 17 Satriano, L., Lewinska, M., Rodrigues, P. M., Banales, J. M. & Andersen, J. B. Metabolic rearrangements in primary liver cancers: cause and consequences. *Nat Rev Gastroenterol Hepatol* **16**, 748-766, doi:10.1038/s41575-019-0217-8 (2019).
- 18 Banales, J. M. *et al.* Serum metabolites as diagnostic biomarkers for cholangiocarcinoma, hepatocellular carcinoma and primary sclerosing cholangitis. *Hepatology*, doi:10.1002/hep.30319 (2018).
- 19 Chong, J. *et al.* MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res* **46**, W486-W494, doi:10.1093/nar/gky310 (2018).
- 20 Cavill, R. *et al.* Consensus-phenotype integration of transcriptomic and metabolomic data implies a role for metabolism in the chemosensitivity of tumour cells. *PLoS Comput Biol* **7**, e1001113, doi:10.1371/journal.pcbi.1001113 (2011).
- 21 Nguyen, A., Rudge, S. A., Zhang, Q. & Wakelam, M. J. Using lipidomics analysis to determine signalling and metabolic changes in cells. *Curr Opin Biotechnol* **43**, 96-103, doi:10.1016/j.copbio.2016.10.003 (2017).
- 22 Muehleman, M., Miller, K. D., Dauphinee, M. & Mizejewski, G. J. Review of Growth Inhibitory Peptide as a biotherapeutic agent for tumor growth, adhesion, and metastasis. *Cancer Metastasis Rev* **24**, 441-467, doi:10.1007/s10555-005-5135-2 (2005).
- 23 Gomes, R. N., Felipe da Costa, S. & Colquhoun, A. Eicosanoids and cancer. *Clinics (Sao Paulo)* **73**, e530s, doi:10.6061/clinics/2018/e530s (2018).
- 24 Lu, X., Yu, H., Ma, Q., Shen, S. & Das, U. N. Linoleic acid suppresses colorectal cancer cell growth by inducing oxidant stress and mitochondrial dysfunction. *Lipids Health Dis* **9**, 106, doi:10.1186/1476-511X-9-106 (2010).
- 25 Ma, C. *et al.* NAFLD causes selective CD4(+) T lymphocyte loss and promotes hepatocarcinogenesis. *Nature* **531**, 253-257, doi:10.1038/nature16969 (2016).

- 26 Jee, S. H. *et al.* Metabolomics Profiles of Hepatocellular Carcinoma in a Korean Prospective Cohort: The Korean Cancer Prevention Study-II. *Cancer Prev Res (Phila)* **11**, 303-312, doi:10.1158/1940-6207.CAPR-17-0249 (2018).
- 27 Tillin, T. *et al.* Diabetes risk and amino acid profiles: cross-sectional and prospective analyses of ethnicity, amino acids and diabetes in a South Asian and European cohort from the SABRE (Southall And Brent REvisited) Study. *Diabetologia* **58**, 968-979, doi:10.1007/s00125-015-3517-8 (2015).
- 28 Enooku, K. *et al.* Altered serum acylcarnitine profile is associated with the status of nonalcoholic fatty liver disease (NAFLD) and NAFLD-related hepatocellular carcinoma. *Sci Rep* **9**, 10663, doi:10.1038/s41598-019-47216-2 (2019).
- 29 Mayo, R. *et al.* Metabolomic-based noninvasive serum test to diagnose nonalcoholic steatohepatitis: Results from discovery and validation cohorts. *Hepatol Commun* **2**, 807-820, doi:10.1002/hep4.1188 (2018).

Tables

Table 1. Clinical characteristics of discovery and validation sets

	P val	DISCOVERY				VALIDATION	
		CTRL	OB-MAFLD	MAFLD-HCC	AV-HCC	MAFLD	MAFLD-HCC
n		44	93	27	32	37	16
Females (%)	<0.0001	59%	71%	7%	6%	65%	17%
Age (years)	<0.0001	53.2±12.9	47.4±10.8	69.7±6.9	59.5±8.1	55.6±9.2	64.4±10.4
BMI	<0.0001	28.9±7.9	45.6±6.7	30.2±2.7	26.4±3.5	30.9±4.5	29.1±4.7
AFP (ng/mL)	<0.0001	2.6±2	nd	129±546	172±371	nd	1904±3575
ALT (U/L)	<0.0001	35±53	29±14	70±106	46±30	nd	56±49
GGT (U/L)	<0.0001	41±56	46±54	161±133	126±80	nd	178±148
AP (U/L)	0.0009	70±32	79.5±39	97±56	99±34	nd	173±82
Bilirubin (mg/dL)	<0.0001	0.63±0.25	0.46±0.22	0.8±0.4	1±1.3	nd	nd
Diabetes	<0.001	14%	31%	52%	31%	nd	81%
Fibrosis 0-2	<0.0001		88%	44%		24% [#]	6%
Fibrosis 3-4			12%	41%		76% [#]	94%
Tumor size (mm)	<0.001			59.5±29.4	39.4±31.6		nd
Cirrhosis	<0.0001		nd	30%	90%	nd	nd

Statistical significance for the categorical traits was established with Fisher's exact test, the significant differences for continuous traits are reported in Suppl. Fig. 1. All measurements are given as mean ± standard deviation. [#]Described by the pathologist as 'steatosis' and 'NASH', respectively. Abbreviations: body mass index (BMI), alpha-fetoprotein (AFP), alanine aminotransferase (ALT), gamma-glutamyl transferase (GGT) and alkaline phosphatase (AP). Not determined (nd)

Table 2 The area under the curve of the receiver operating characteristic curves for the model, biochemical parameters, and individual metabolites

Model	AUC		
	CTRL	OB-MAFLD	AV-HCC
	0.989	0.997	0.999
Biochemical features			
AFP	0.786	NA	0.613
ALT	0.776	0.733	0.570
GGT	0.927	0.865	0.562
AP	0.781	0.606	0.515
Bilirubin	0.619	0.859	0.539
Individual metabolites			
AC(18:2n-6)	0.957	0.950	0.961
PC(16:0/17:0)	0.992	0.991	0.942
Linoleic acid	0.910	0.990	0.968
Osbond acid	0.944	0.964	0.965
Palmitolinoleic acid	0.915	0.977	0.948
MUFA (14:1n-5trans)	0.934	0.962	0.979
PUFA (14:2n-x)	0.978	0.971	0.998
PUFA (16:3n-x)	0.963	0.964	1.000
Hydroxy-octadecadienoic acid	0.931	0.982	0.966
PC(0:0/18:2)	0.924	0.909	0.925
PC(18:2/0:0)	0.900	0.925	0.973
PC(0:0/22:5)	0.914	0.930	0.987
TG(47:0)	0.978	0.968	0.933
TG(45:1)	0.955	0.933	0.903

Area under the curve (AUC) for biochemical features and individual metabolites calculated as AUC of the receiver operating characteristic (ROC) curves and AUC from support vector machine (SVM)-driven modelling. Abbreviations: alpha-fetoprotein (AFP), alanine aminotransferase (ALT), gamma-glutamyl transferase (GGT), alkaline phosphatase (AP), acylcarnitine (AC), phosphatidylcholines (PC), monounsaturated fatty acids (MUFA), polyunsaturated fatty acids (PUFA) and triglycerides (TG).

Figures

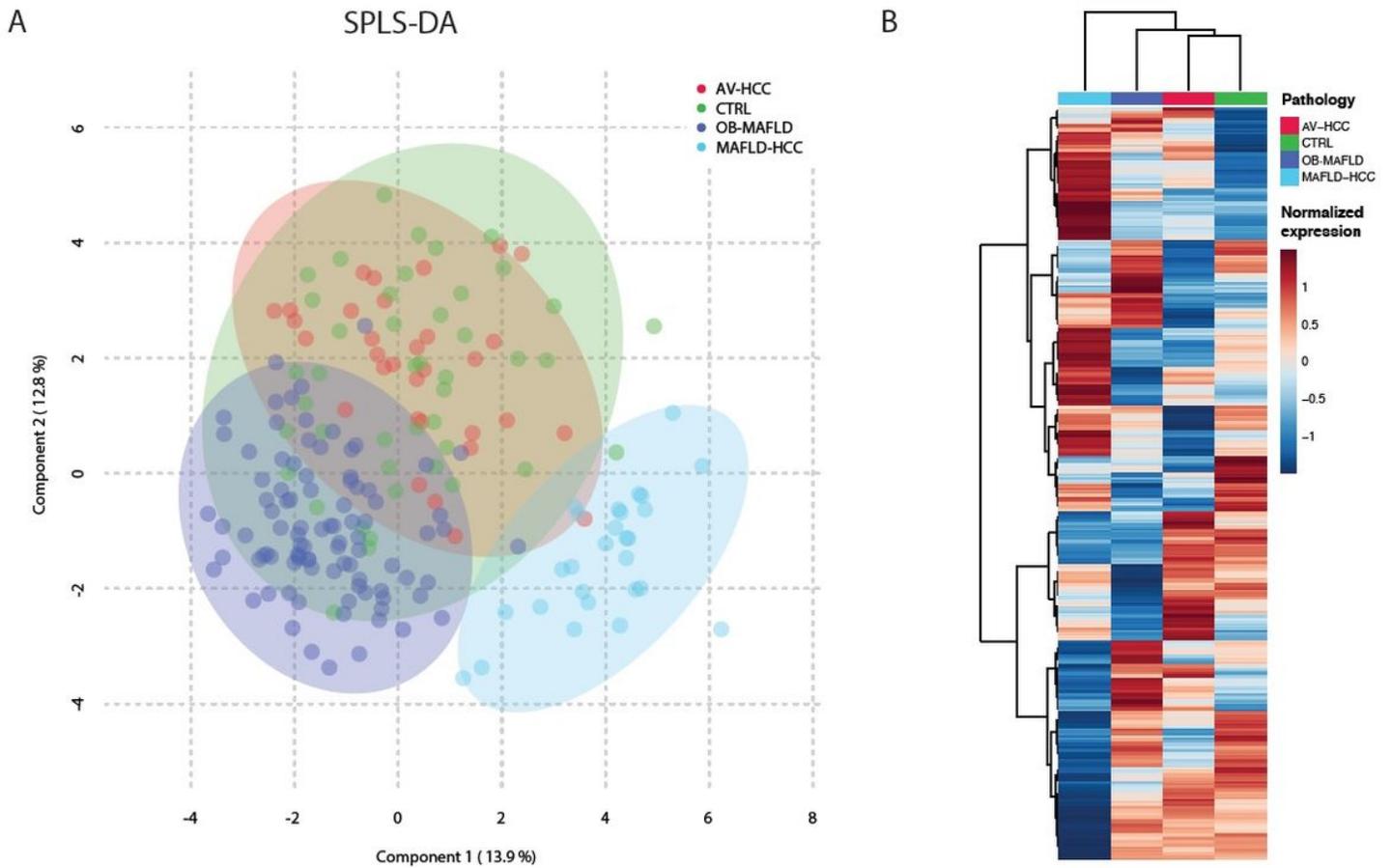


Figure 1

Multivariate analysis between the patient groups. A. The 3 components based on sparse partial least squares discriminant analysis (sPLS-DA model) of all detected metabolites showing samples separation. B. Unsupervised hierarchical clustering of patient's groups (group average) across all metabolites (Ward's method clustering).

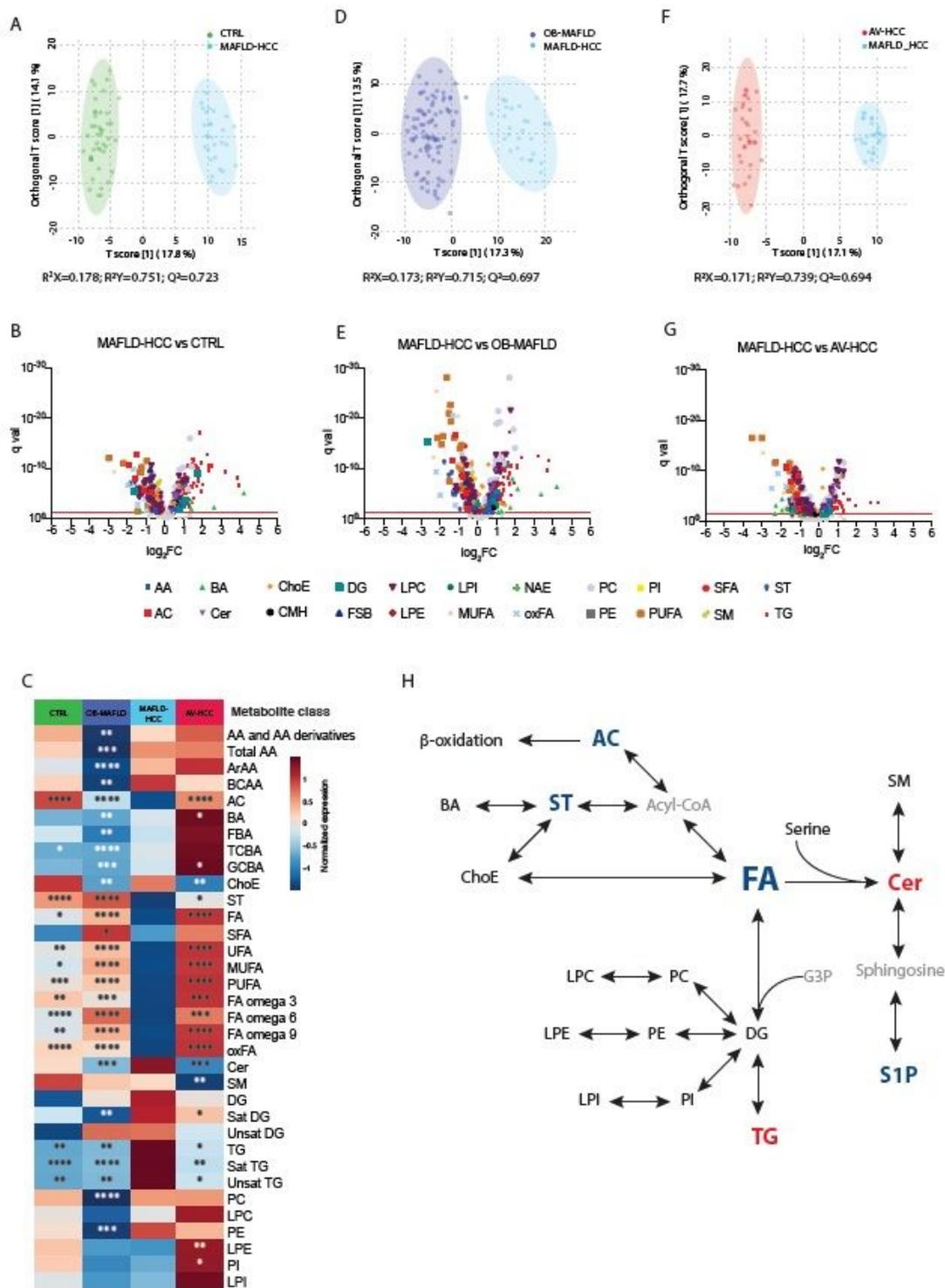


Figure 2

Pair-wise comparison between MAFLD-HCC and other groups. A. The orthogonal partial least squares discriminant analysis (oPLS-DA) showing separation between MAFLD-HCC and CTRL. B. Volcano plot representing DEMs between MAFLD-HCC and CTRL samples. The FDR $p=0.05$ is marked with a red dashed line. C. Differences between MAFLD-HCC and other groups across major metabolic classes. Statistical significance was between MAFLD-HCC and other groups * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$. D. The oPLS-DA model showing separation between MAFLD-HCC and OB-MAFLD. E. Volcano plot representing

DEMs between MAFLD-HCC and OB-MAFLD. F. The oPLS-DA model showing separation between MAFLD-HCC and AV-HCC. G. Volcano plot representing DEMs between MAFLD-HCC and AV-HCC. H. Simplified relationship between major metabolic classes. Classes marked in blue are downregulated and red upregulated in MAFLD-HCC patients. Abbreviations: acylcarnitines (AC), amino acids (AA), aromatic amino acids (ArAAs), bile acids (BA), branched-chain amino acids (BCAAs), ceramides (Cer), cholesteryl esters (ChoE), diacylglycerols (DG), diacylglycerophosphocholine (DAPC), diacylglycerophosphoethanolamine (DAPE), fatty acids (FA), fatty acids omega 3 (FA omega 3), fatty acids omega 6 (FA omega 6), fatty acids omega 9 (FA omega 9), free bile acids (FBA), free sphingoid bases (FSB), glycine-conjugated bile acids (GCBA), lysophosphatidylcholines (LPC), lysophosphatidylethanolamines (LPE), lysophosphatidylinositols (LPI), monohexosylceramides (CMH), monounsaturated fatty acids (MUFA), N-acyl ethanolamines (NAE), oxidized fatty acids (oxFA), phosphatidylcholines (PC), phosphatidylethanolamines (PE), phosphatidylinositols (PI), polyunsaturated fatty acids (PUFA), saturated diacylglycerols (Sat.DG), saturated fatty acids (SFA), saturated triacylglycerols (Sat.TG), sphingomyelins (SM), steroids (ST), taurine-conjugated bile acids (TCBA), triacylglycerols (TG), unsaturated fatty acids (UFA).

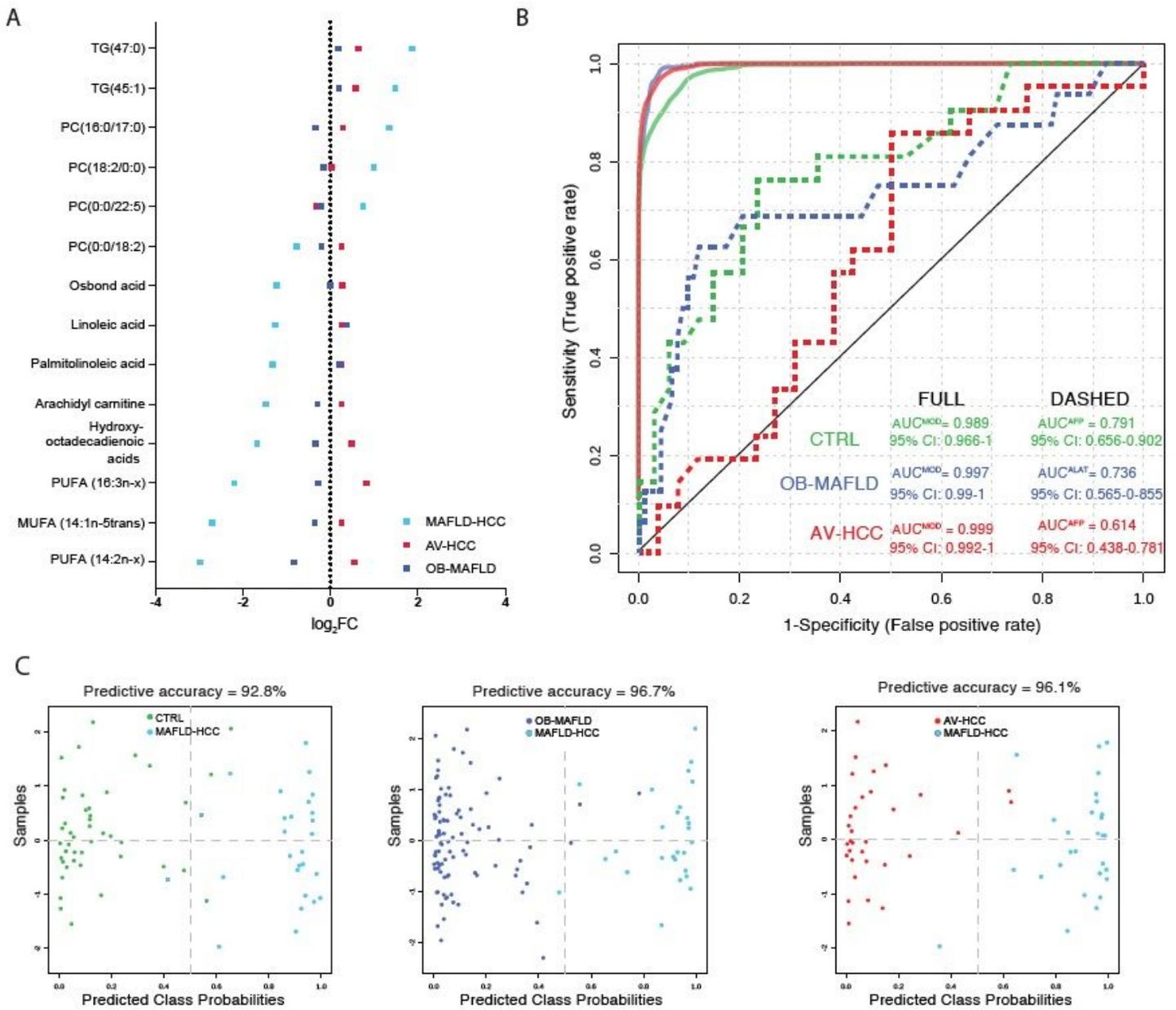


Figure 3

The diagnostic potential of lipidomics. A. The waterfall plot presenting the 14 metabolites signature distinguishing MAFLD-HCC patients. Metabolites are presented as the fold change to CTRL. B. ROC curves of 5-metabolite based model distinguishing MAFLD-HCC patients from other groups in compresence to biochemical tests. C. The predicted class probabilities (average of the cross-validation) for each sample using the best classifier (based on AUC).

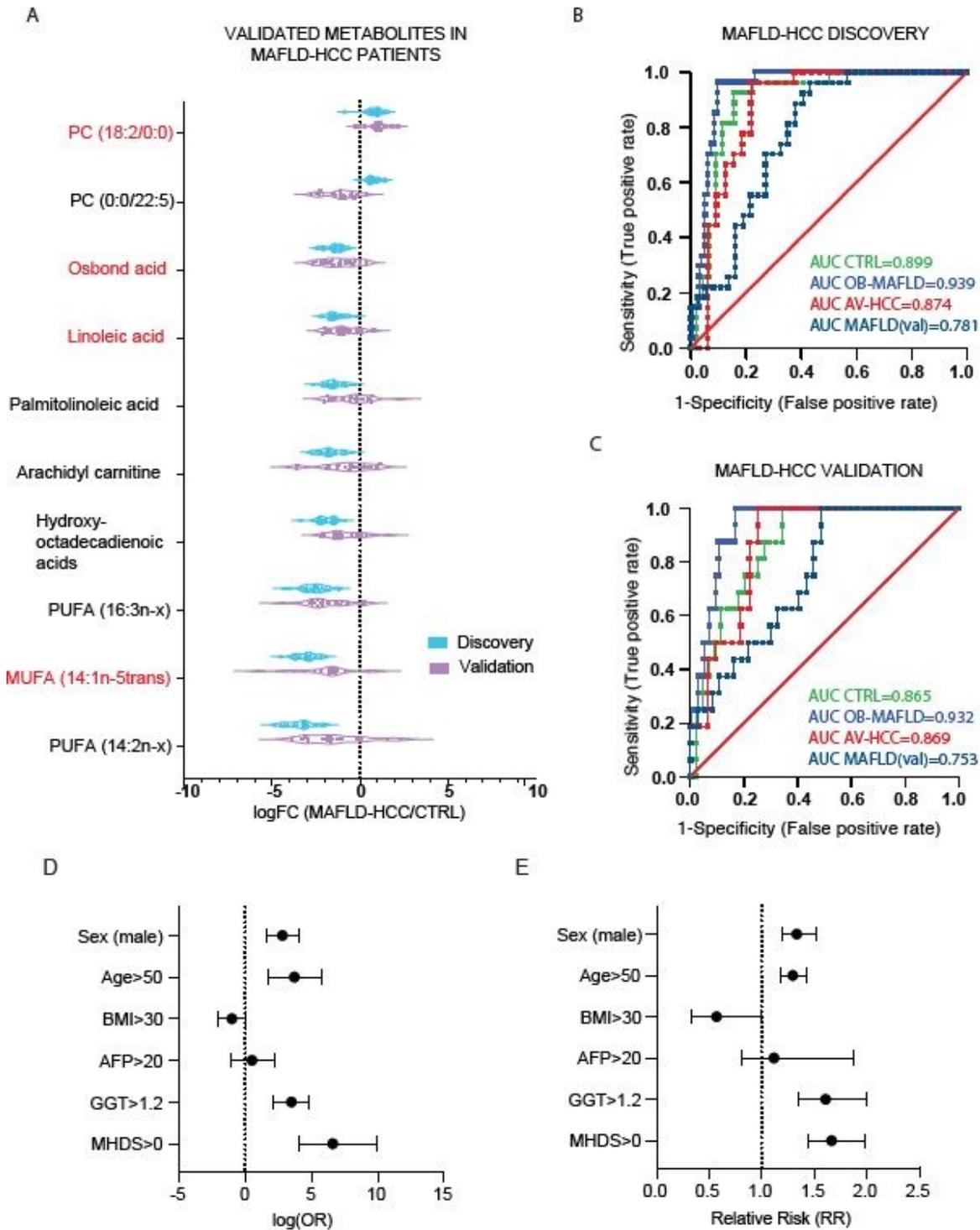


Figure 4

Validation of metabolites and MAFLD-HCC Diagnostic Score (MHDS). A. The relative abundances of 10 validated metabolites in discovery (blue) and validation (purple) sets are expressed as a fold change against CTRL. Metabolites included in MHDS are marked in red. B. ROC curves of MHDS in the discovery set. C. ROC curves of MHDS in the MAFLD-HCC validation set. D. The forest plot presenting odds ratios of MHDS at the established cut-off value as well as clinical and biochemical characteristics for pulled (discovery and validation) sets. AFP and GGT were available only in the discovery set. E. The forest plot

presenting a relative risk of MHDS at the established cut-off value as well as clinical and biochemical characteristics for pulled (discovery and validation) sets. AFP and GGT were available only in the discovery set.

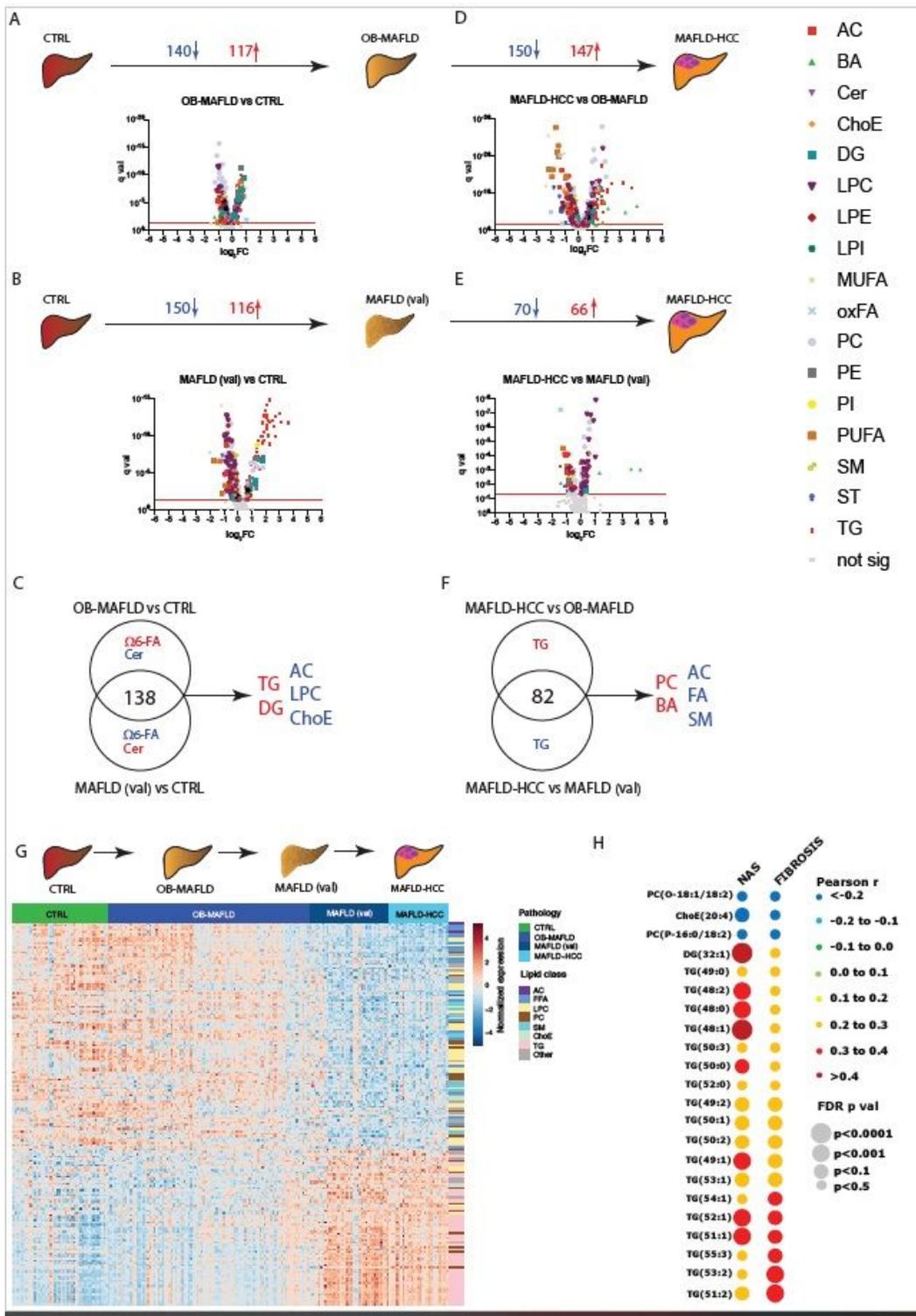


Figure 5

The progressive metabolic perturbations in CTRL to MAFLD to MAFLD-HCC trajectory. A. The pair-wise compresence between OB-MAFLD and CTRL. B. The pair-wise compresence between MAFLD patients in the validation set (MAFLD(val)) and CTRL. C. The Venn diagram presenting common (significant and same

directionality) DEMs between CTRL and MAFLD sets. D. The pair-wise compresence between MAFLD-HCC and OB-MAFLD. E. The pair-wise compresence between MAFLD-HCC and MAFLD(val). F. The Venn diagram presenting common (significant and same directionality) DEMs between MAFLD-HCC and MAFLD patient groups. G. The heatmap presenting the expression of metabolites statistically significantly correlated with disease progression. H. The bubble plot presenting correlation of metabolites with NAS and fibrosis level in OB-MAFLD patients.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [LewinskaEtalSupplementaryData.docx](#)
- [LewinskaEtalFig.S1.pdf](#)
- [LewinskaEtalFig.S2.pdf](#)
- [LewinskaEtalFig.S3.pdf](#)
- [LewinskaEtalFig.S4.pdf](#)
- [LewinskaEtalFig.S5.pdf](#)
- [LewinskaEtalFig.S6.pdf](#)
- [LewinskaEtalFig.S7.pdf](#)
- [LewinskaEtalFig.S8.pdf](#)
- [LewinskaEtalSupportingTables.xlsx](#)