Supplementary

Materials and methods

Histology

For histology, each biopsy specimen was fixed in formalin and embedded in paraffin. A minimum of seven step sections were stained with hematoxylin and eosin, Prussian blue, periodic acid-Schiff stain with and without prior diastase digestion, reticuline and chromotrope- aniline blue staining

Supplementary Table 1S: Overview of the clinical indices and frequencies of genetic risk factors for NC, HO, NAFL and NASH patient groups [(1)]

Total (n=337)	**NC (49)** Normal control	**HO (51)** Healthy obese	**NAFL (143)** Simple steatosis	**NASH (94)** Steatohepatitis	**P value** between all 4 groups	**P value** between HO/NAFL/NASH
NAS score	0.00 (0)	0.00 (0)	1.46 (0.06)	3.80 (0.16)	<0.0001*	<0.0001*
Age	67.7 (1.6)	42.3 (1.5)	45.6 (1.2)	45.4 (1.4)	<0.0001*	0.317*
BMI	24.1 (0.5)	46.9 (1.1)	47.1 (0.9)	49.0 (1.2)	<0.0001*	0.340*
Sex, % female	51 %	92 %	66 %	64 %	<0.0001**	0.0006**
Diabetes %	19 %	11 %	27 %	40 %	0.001**	0.001**
ALT (IU/L)	24.8 (1.9)	25.1 (2.6)	30.0 (1.8)	54.5 (5.6)	<0.0001*	<0.0001*
AST (IU/L)	30.9 (3.9)	23.0 (1.5)	26.4 (1.0)	44.5 (4.0)	<0.0001*	<0.0001*
Bilirubin (mg/dl)	0.51 (0.05)	0.40 (0.03)	0.46 (0.02)	0.84 (0.27)	0.157*	0.107*
*PNPLA3*_rs738409 (MAF)	0.20	0.15	0.23	0.31	0.046**	0.043**
*TM6SF2*_rs58542926 (MAF)	0.07	0.06	0.05	0.09	0.699**	0.513**
*MBOAT7*_rs641738 (MAF)	0.40	0.43	0.43	0.41	0.883**	0.851**
*HSD17B13*_rs72613567 (MAF)	0.38	0.23	0.25	0.30	0.050**	0.048**
*SERPINA*_rs28929474 (MAF)	0.05	0.03	0.04	0.03	0.841**	0.834**
*SERPINA*_rs17580 (MAF)	0.03	0.03	0.01	0.04	0.215**	0.186**

(1) Patients were grouped according to NAS[1]. Values are given as counts (with percentage) or as mean with standard error (in parentheses). MAF: frequency of the minor allele; Statistical tests: *ANOVA; ** Fisher exact test.

Common chemicals and lipid standards

All solvents were of LC-MS or better grade. 700 µL of the internal standard mix for human samples contained: 1778 pmol of D7-cholesterol; 2215 pmol of D7- CE 16:0; 1041 pmol of D5-TG 50:0; 595 pmol of D7-DG 33:1; 1376 pmol of PC 25:0; 386 pmol of LPC 13:0; 589 pmol of PE 25:0; 85 pmol of LPE 13:0; 480 pmol of PI 25:0, 140 pmol of PG 25:0; 73 pmol of Cer 30:1,2; 127 pmol of D7-PA 33:1; 78 pmol of U-13C LPA 16:0; 271 pmol of

d18-SM 30:1,2; 64 pmol of LPI 13:0, 137 pmol of PS 25:0, 74 pmol of LPS 13:0; 75 pmol of LPG 13:0, and 137 pmol of CL 56:4.

Sample preparation

Sample preparation is described in the main text.

Pilot experiment and making the Quality Control (QC) sample

First, selected samples were chosen for the pilot shotgun lipidomics run as previously described in [2]. The goal of this selection was to identify and analyze samples with extreme lipid content. To consider all features, including meta-data, a regression analysis was performed. Specifically, a gradient boosted tree was applied since it is well suited for ranking models [3]. The initial data were numerically formatted by converting categorical data into values, and by removing incomplete entries; additionally, a feature was added based on the concurrence of multiple factors, *e.g.* fat content, inflammation and fibrosis. Samples ranked 1-19 and 349-367 representing the extreme fatty and non-fatty samples were analyzed in the pilot experiment. 36 samples from the pilot experiment were pooled together in order to create a single quality control (QC) sample. During pilot experiment we consistently quantified 255 lipid species from 19 lipid classes, including energy storage lipids CE, TG and DG; membrane lipids: Cer, Chol, CL, PA, PC, PE, PG, PI, PS and SM; and lysolipids: LPA, LPC, LPE, LPG, LPI and LPS.

After the pilot we performed the full cohort run, adding the pooled QC sample to each measurement batch to control for inter-batch variation. All samples were pipetted in 96 well plates, each row had one well with a blank extraction sample and one well with a QC sample. Biological samples and QC were measured as technical duplicates by splitting them to separate wells.

MS^2 validation

To validate the assignment of SMs, we isolated corresponding molecular ions at the unit *m/z* range and subjected them to high resolution HCD FT MS/MS. In MS^2 spectra of $[M+H]^+$ precursors we detected characteristic phosphocholine head group fragment *m/z* 184.074. Fragmenting $[M+ HCOO]^-$ precursors at low normalized collision energy (nCE)

of 15% produced demethylation fragments [M- CH_3]$^-$ as neutral loss of Δ*m/z* = - 60.021 (losing one methyl group (CH_3) and the formate adduct (HCOO)). Taken together, this confirmed their elemental composition and assignment of the molecular ions to SM class.

PE O- 38:7 was identified by HCD FT MS/MS as a plasmalogen PE O-16:1 / 22:6 by its accurate *m/z*; neutral loss of phosphoethanolamine group Δ*m/z* = 141.019 in positive and acyl anion of FA 22:6 (*m/z* = 327.233) together with subsequent CO_2 loss (Δ*m/z*= 43.990) in negative mode.

Identification of SM molecular species by LC-MS^n

Pooled lipid extracts (500 µL in methanol / chloroform 1:1) were dried and re-suspended in 100 µL 2-propanol : chloroform : methanol (90:5:5 v/v/v). Data acquisition was performed on an LTQ Orbitrap Velos Pro instrument (Thermo Fisher Scientific) coupled to an Ultimate 3000 UHPLC (Thermo Fisher Scientific) as described [4-6]. Chromatographic separation was performed on BEH C8 column (100 × 1 mm, 1.7 µm) from Waters (Milford, MA) at 50 °C. Mobile phase A was deionized water containing 10 mM ammonium formate and 0.1 vol% of formic acid. Mobile Phase B was a mixture of acetonitrile/isopropanol 5:2 (v/v) with the same formate buffer. Gradient elution started at 50 % of B rising to 100 % B over 40 minutes; 100 % B were held for 10 minutes and then the column was re-equilibrated with 50 % B for 8 minutes before the next injection. The injection volume was 2 µL; flow rate of 150 µL/min was infused into HESI II ion source. In the autosampler samples were kept at 8 °C. The mass spectrometer was operated in Data Dependent Acquisition mode. Every sample was measured once in positive polarity and five times in negative polarity. Full scan profile spectra were acquired in the Orbitrap mass analyzer at the target mass resolution of $R_{m/z=400}$ = 100,000. For MS/MS experiments, the precursor masses of SM 34:0;2, SM 41:3;2, SM 42:4;2, SM 43:1;2 and SM 43:3;2 were added to the inclusion list and were fragmented in parallel mode by CID using an NCE of 50%. Subsequently, an additional MS^3 scan event targeted neutral loss fragments of Δ*m/z* = 60. Data analysis was performed by Lipid Data Analyzer [6–8].

Raw data processing

The LipidXplorer output was processed in several ways, including different filtering steps (implemented in Python programming language) (**Supplementary Figure 1S)**. In the first step we averaged over the technical replicates of each patient sample and separated the QC probes from the rest of the cohort. Subsequently, we discarded all datapoints for which the calculated standard deviation (SD) exceeded 40% by setting them to NaN in a first filtering (SD filter, 40%). To avoid loss of datapoints close to the experimental detection threshold, we did not filter datapoints smaller or equal to two times the minimum detected for the respective lipid species (i.e., DP <= 2*min(species)). Filtering was not applied to cholesterol since it was quantified by PRM.

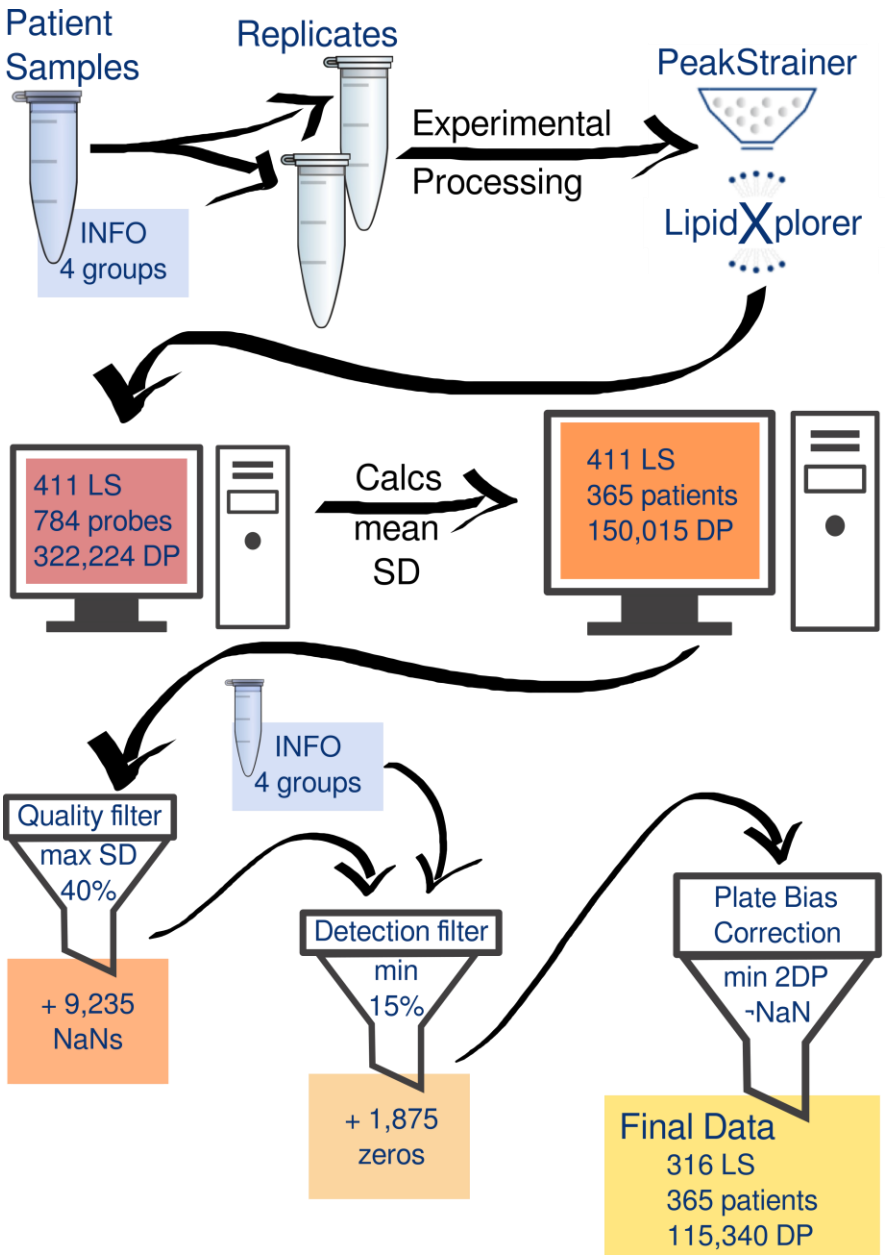
The obtained dataset is composed of 365 patient samples (compared to 784 probes in the raw data set) and still 411 lipid species from 22 lipid classes. The total number of 150,015 datapoints includes 61,759 numeric zeros and 9,334 NaNs (**Supplementary Figure 1S**).

In a second preprocessing step, we applied a group-wise detection filter of 15%. To this end, we grouped the patient samples according to their disease status (NC, HO, NAFL, NASH and none). Subsequently, a lipid species that were not detected in at least 15% of all patient samples of one status group was set to zero in this status group. This introduced 1,875 additional numeric zeros into our dataset (**Supplementary Figure 1S**).

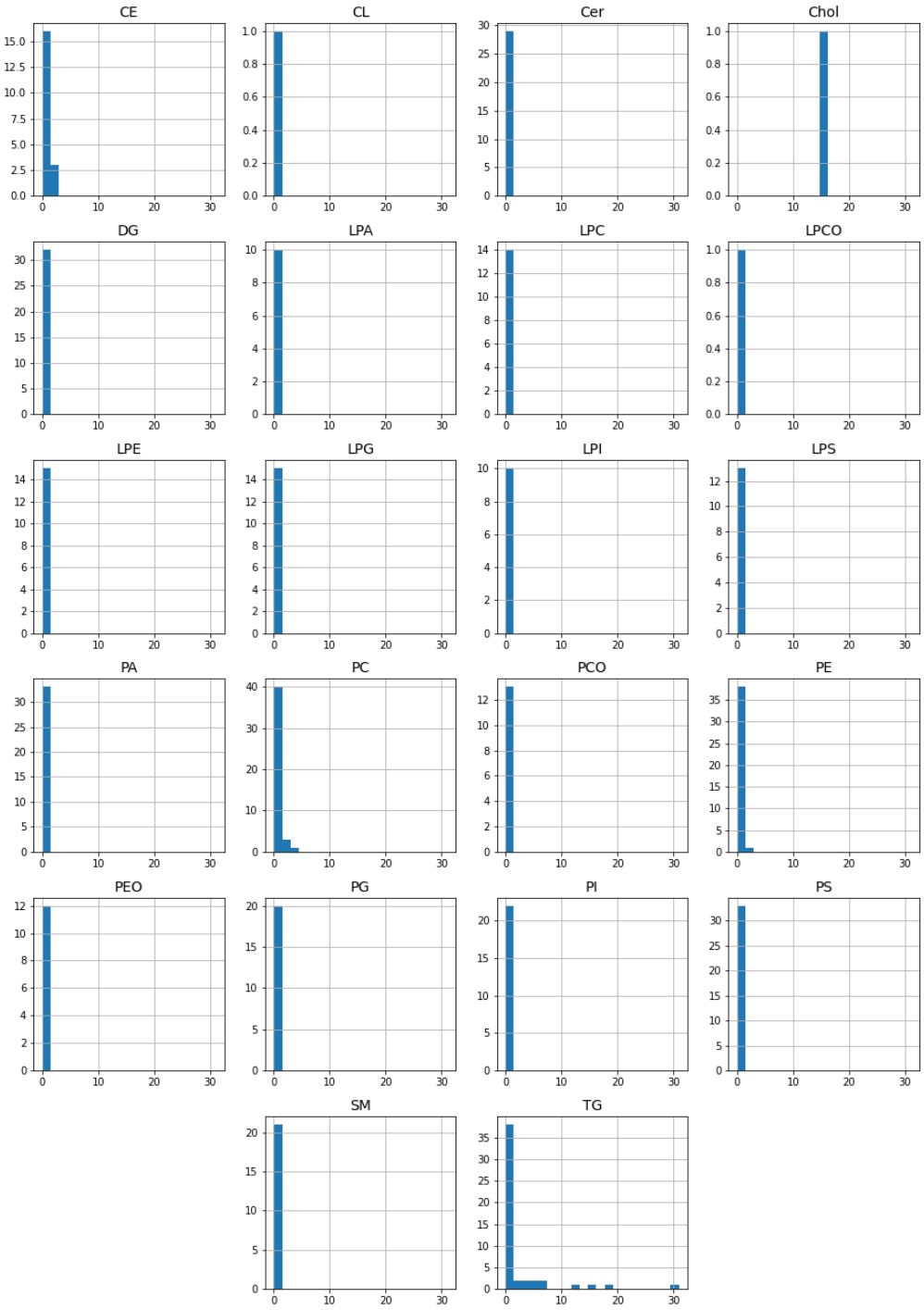
We applied a plate bias correction using the ComBat approach [9] as detailed below. The extraction date was used as a batch covariate. Before adjusting, all lipid species with less than 2 numeric values >0 were discarded and the data was transformed into log-2 space. The correction was performed with non-parametric priors.

In summary, the preprocessing provided us with a final dataset (**Supplementary Figure 1S**) with a total number of 115,340 datapoints, each representing a measurement for one out of 316 lipid species in one patient (out of 365). Those datapoints sum 101.013 numeric

(non-zero) values plus 6,381 NaNs (**Supplementary Figure 1S**). This dataset was used for all further analyses.



Supplementary Figure 1S: Preprocessing of LipidXplorer output dataset. LipidXplorer output was processed in several steps: elimination of not trustworthy datapoints (standard deviation (SD) threshold 40%), patient group specific detection filter (threshold 15%), and plate bias correction have been applied. A detailed explanation of filtering steps is in the main text (Materials and Methods) and Supplementary, Raw data processing.



Supplementary Figure 2S: QC overview for each lipid class. Histograms of standard deviation (SD) between QC probes for all species of a lipid class. Standard deviation in

QC was below 5% for 98% of lipid species except cholesterol and a few TG species (**Supplementary Figure 2S**.)

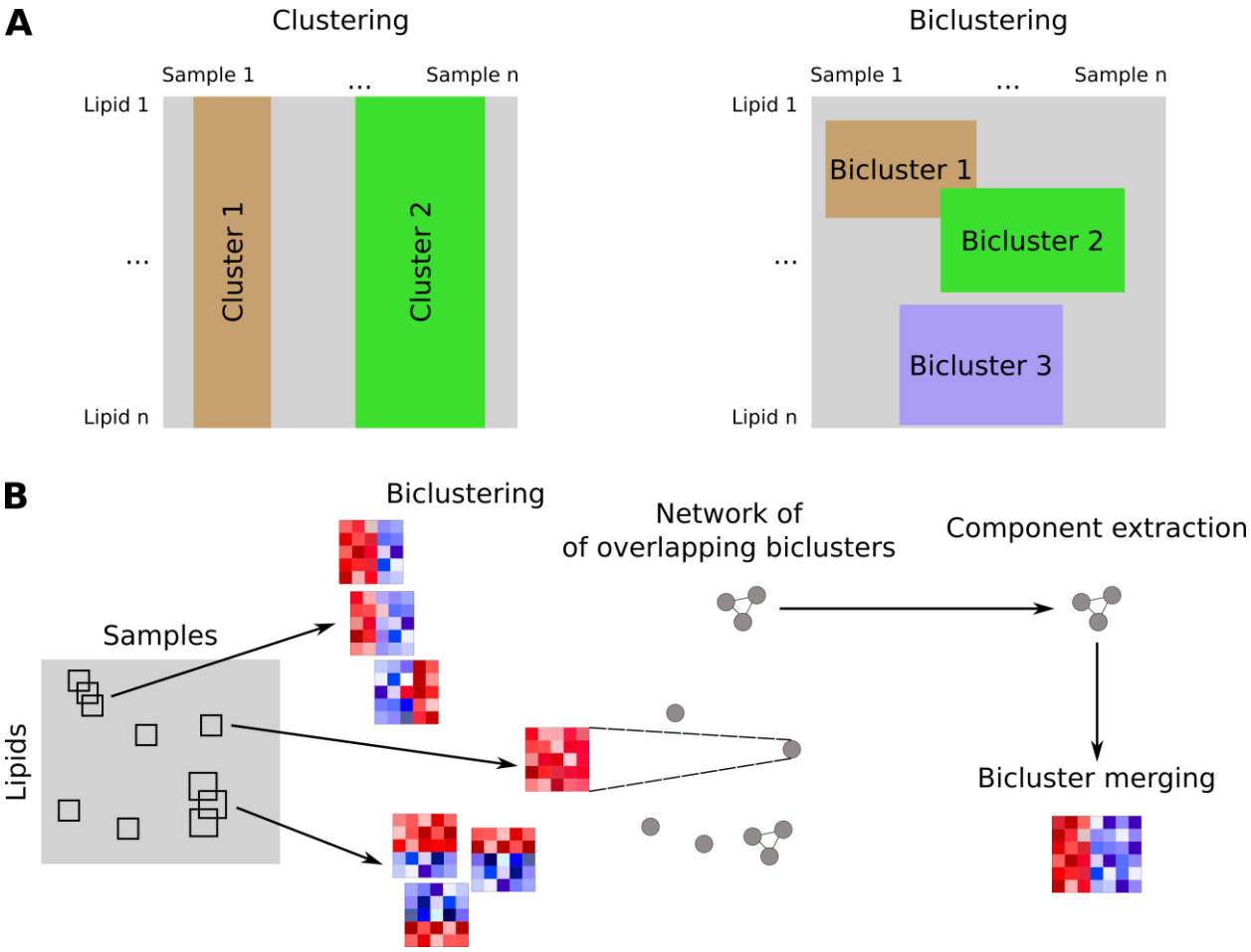
Bioinformatics and data mining

After batch correction lipid-wise z-scores were computed and biclustering was performed. Biclustering (simultaneous clustering of samples and features in a data matrix) allows to simultaneously extract patient subgroups with corresponding lipid signatures to analyze the differences between the groups. Accordingly, biclusters are two-dimensional matrices that consist of a subset of patients and lipid species (details in **Supplementary, Bioinformatics and data mining and Supplementary Figure 3S A**).

Since the lipidomic data contained missing values, we selected two biclustering algorithms; isa [10] and QUBIC [11] were applied under default parameters. Together, they identified 462 biclusters. To address redundancies, biclusters that were fully contained in other biclusters were removed resulting in a total set of 380 biclusters. To deal with this quantity of biclusters, we developed a consensus approach inspired by Hanczar and Nadif [12]. We used a custom two-dimensional similarity function and generated a network of similar biclusters for a better visualization and a parameter for the number of clusters (**See Supplementary Figure 3S B**). This network comprises highly overlapping biclusters containing very similar sets of patients and lipid species. In these networks, biclusters were colored using pie-charts that represent the distribution across the five groups of patients. Interesting components were marked with blue/red outlines corresponding to healthy and disease endotypes. Finally, the subset of patients as well as the subset of lipids were extracted from each connected component representing consistency in any cofactor with a general focus on disease states. Biclustering analysis was performed using the R programming language.

Classification was performed using random forest classifiers from the “randomForest" library in the R programming language. Lipidomic data were split into training (60%), validation (20%) and testing (20%) sets using bootstrapping. To increase robustness,

bootstrapping was repeated 100 times for each classification scenario. Per repetition, 100 trees were trained and in case of class imbalance an equal number of samples from each group were drawn for each tree. The resulting 100 bootstrapped models were then evaluated using their corresponding test sets using precision-recall (PR) and receiver-operator-characteristics (ROC) area under the curve (AUC) values. Feature importance was computed using the average Gini index over all bootstrapped models.



Supplementary Figure 3S: Biclustering analysis. (A) Comparison of traditional clustering to biclustering, where samples and features (lipids) are clustered simultaneously. (B) Schematic description of the biclustering workflow. Biclusters are calculated using different algorithms, afterwards biclusters are represented as nodes in a network with edges between biclusters indicating a similarity exceeding a certain threshold. Connected components of biclusters are then extracted and merged as ensemble biclusters.

Genotyping

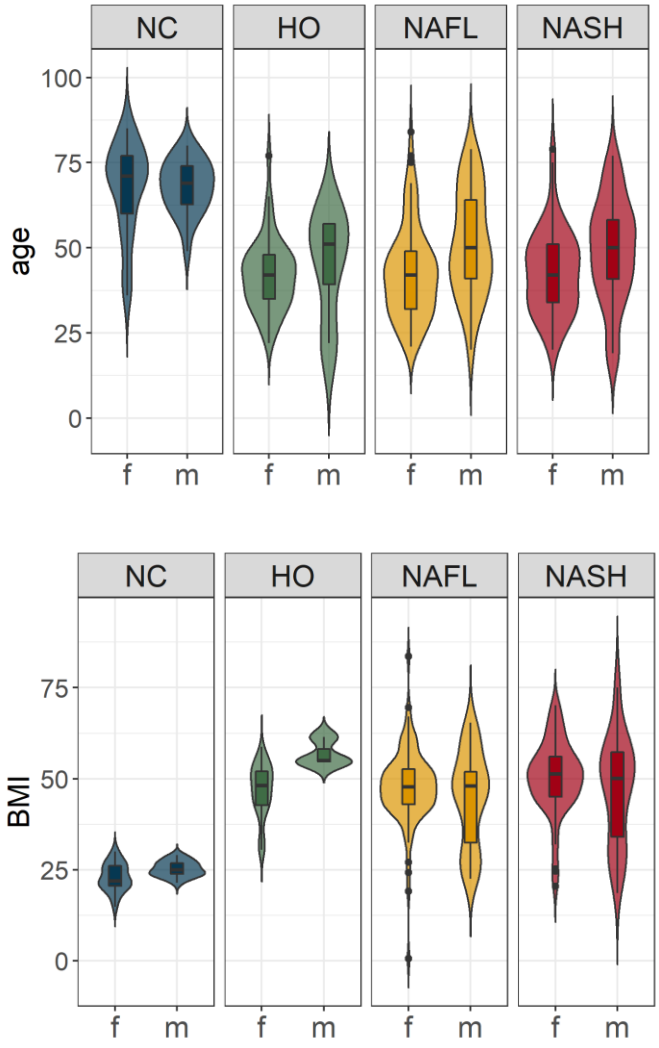
Mutation variants analyzed in this paper:
TM6SF2 variant rs58542926 (hcv89463510);
PNPLA3 variant rs738409 (hcv7241);
MBOAT7 variant rs641738 (hcv8716820);
SERPINA1 variant PiZ (Glu342Lys) rs28929474 (hcv34508510);
SERPINA1 variant PiS (Glu264Val) rs17580 (hcv594695);
HSD17B13 rs72613567 (primer and probe sets manufactured through custom TaqMan Assay design).

Genotyping for the presence of the two most relevant AAT mutations, the Pi*Z variant (rs28929474, also known as p.E342K or Glu342Lys) and the Pi*S variant (rs17580, also known as p.E264V or Glu288Val) was performed. For the NAFLD cohort, genomic DNA from all cases and controls was genotyped using TaqMan chemistry (Applied Biosystems) as described before [2]. Moreover, genotyping was performed for the PNPLA3 variant rs738409 (hcv7241), the TM6SF2 variant rs58542926 (hcv89463510) and the MBOAT7 variant rs641738 (hcv8716820) on an automated platform with TECAN Freedom EVO and 384well TEMO liquid handling robots (TECAN, Männedorf, Switzerland). Reactions were completed and read in a 7900 HT TaqMan sequence detector system (Applied Biosystems, Foster City, CA, USA). All process data were logged and administered using a database-driven LIMS [2,14]. Genotyping of the single-nucleotide polymorphisms (SNP) of interest HSD17B13 rs72613567 (primer and probe sets manufactured through custom TaqMan Assay design) was performed using TaqMan SNP genotyping assays and chemistries (Applied Biosystems, Waltham, MA) on an automated platform with Tecan Freedom EVO and 384-well TeMO liquid handling robots (Tecan, Männedorf, Switzerland) as described previously [14].

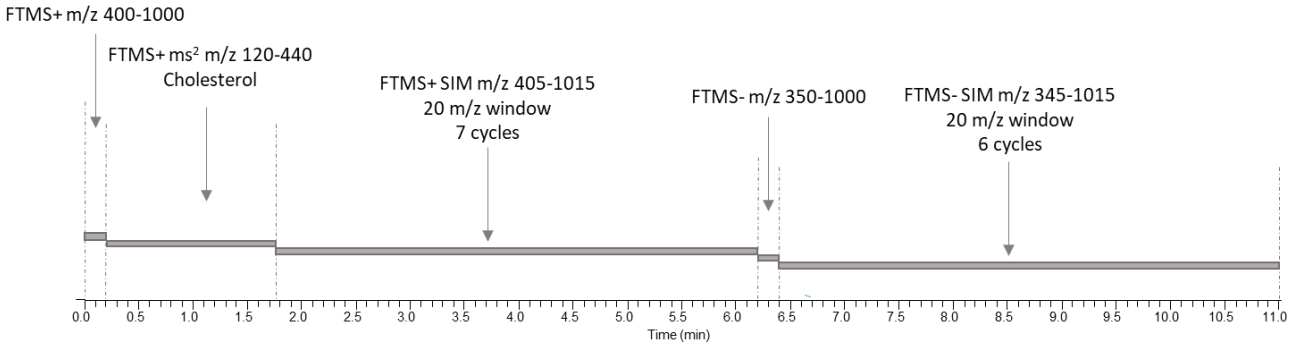
Statistical analysis of mutation-correlated lipidome changes

Mutations were tested within each group, where the lipidomes of homozygotes, heterozygotes and normal genotypes were compared pairwise. The adopted procedure consisted of the following steps: a) log2-scaling of the dataset pruned of the samples with no information on the mutation; b) Shapiro test and correction of the p-values for multiple feature comparisons using Benjamini-Hochberg; c) analysis of variance using either ANOVA or the Levene's test according to the results of the Shapiro test; d) the lipids displaying a significantly larger between-group variance were subjected to either Kruskal-Wallis or ANOVA according to the results of the Shapiro test; e) when significant, Dunn's or Tukey tests were respectively used post-hoc. At steps c) to e), p-values were adjusted for multiple group and feature comparisons using Benjamini-Hochberg.

Results



Supplementary Figure 4S: Age and BMI characteristics of the cohort. Upper panel distribution of age in the four groups of patients split by gender. Lower panel distribution of BMI in the four groups of patients split by gender.

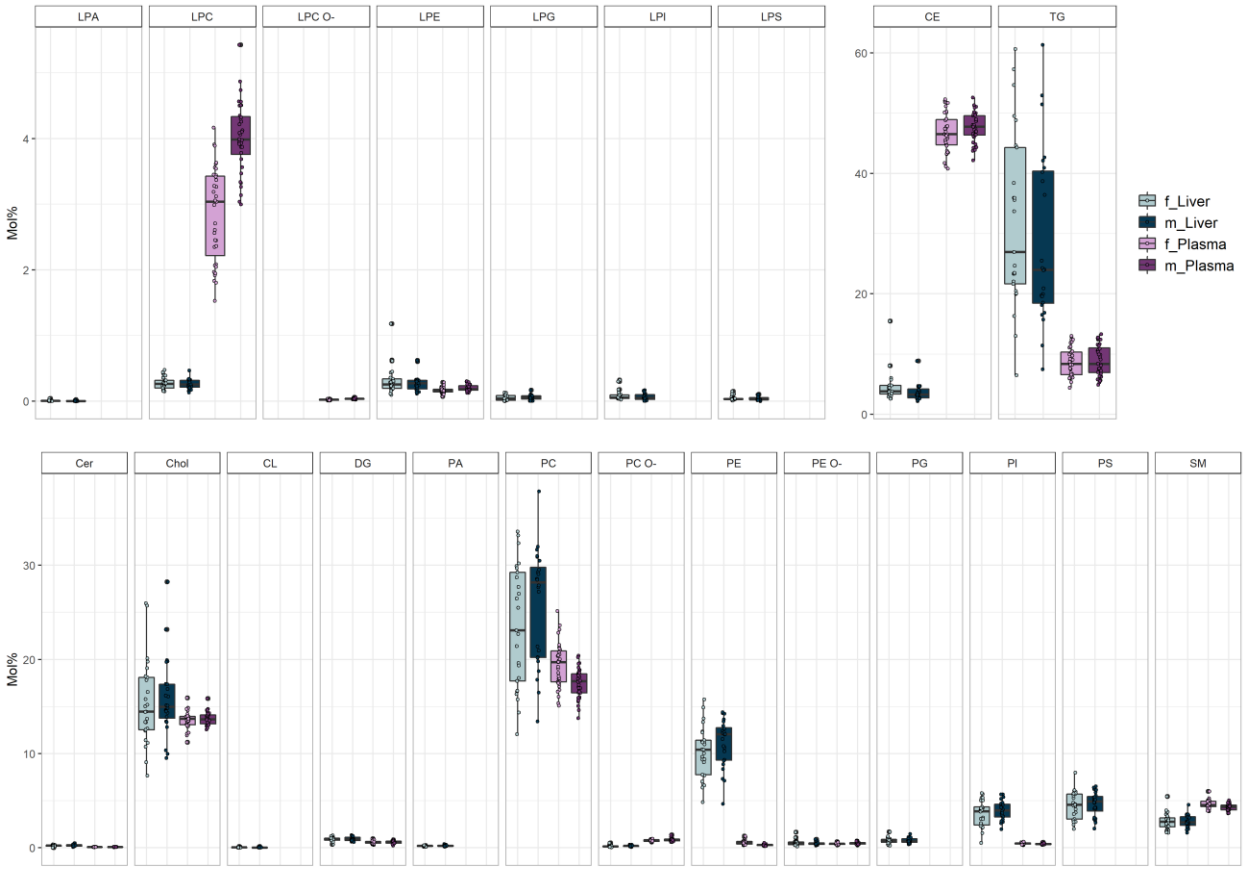


Supplementary Figure 5S: Typical *t*-SIM acquisition scheme used in this study.

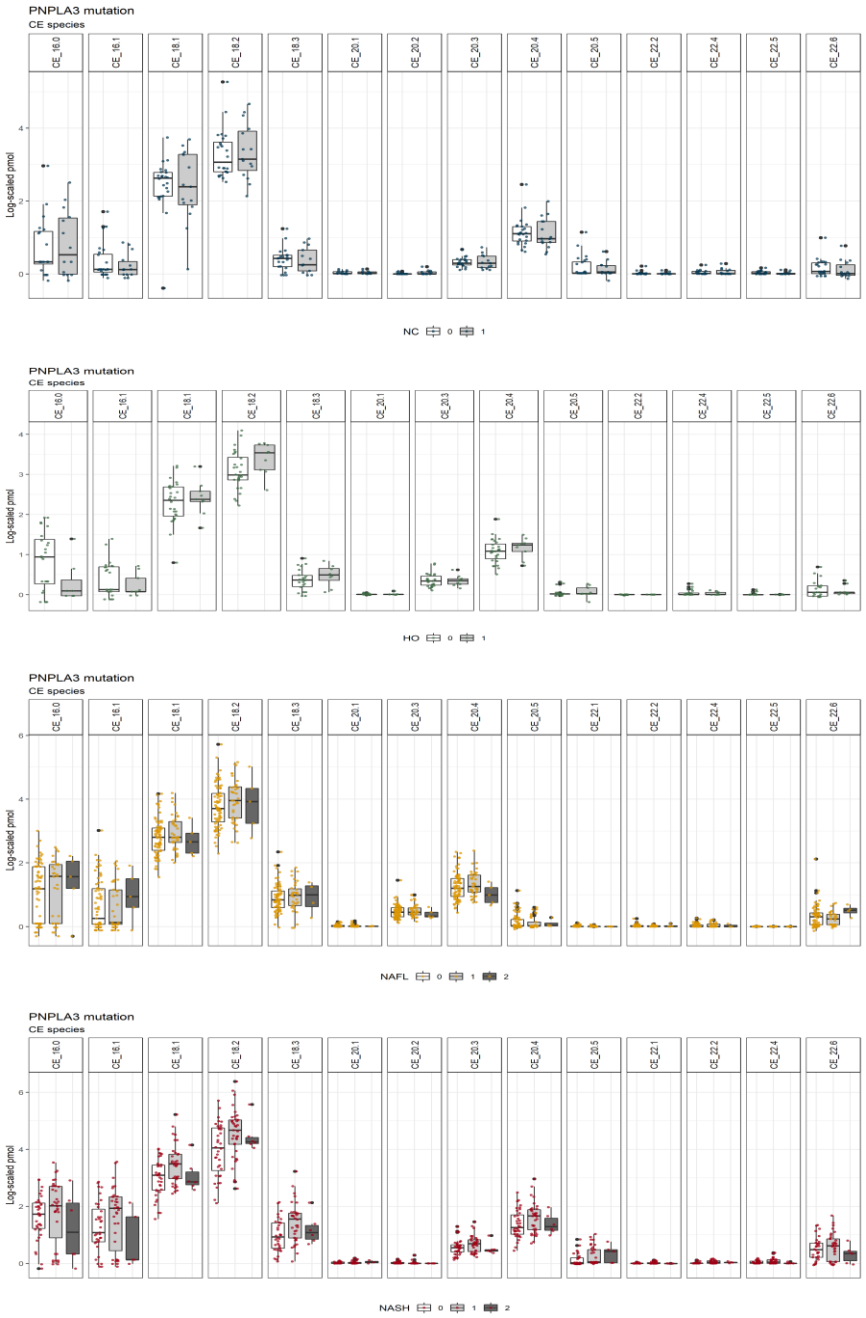
X-axes presents the time-line of a single acquisition. The sequence starts by applying spraying voltage (t=0) followed by a short spraying stabilizing period. Next, cholesterol is quantified using positive mode PRM followed by t-SIM acquisition of full mass range positive and negative mode spectra. To acquire a technical replicate, equal volume of the same lipid extract was taken from another well and spectra acquisition cycle was repeated.

Supplementary Table 2S: Comparison lipid class composition of human liver biopsies (in mol%) from this work and from [15]**.**

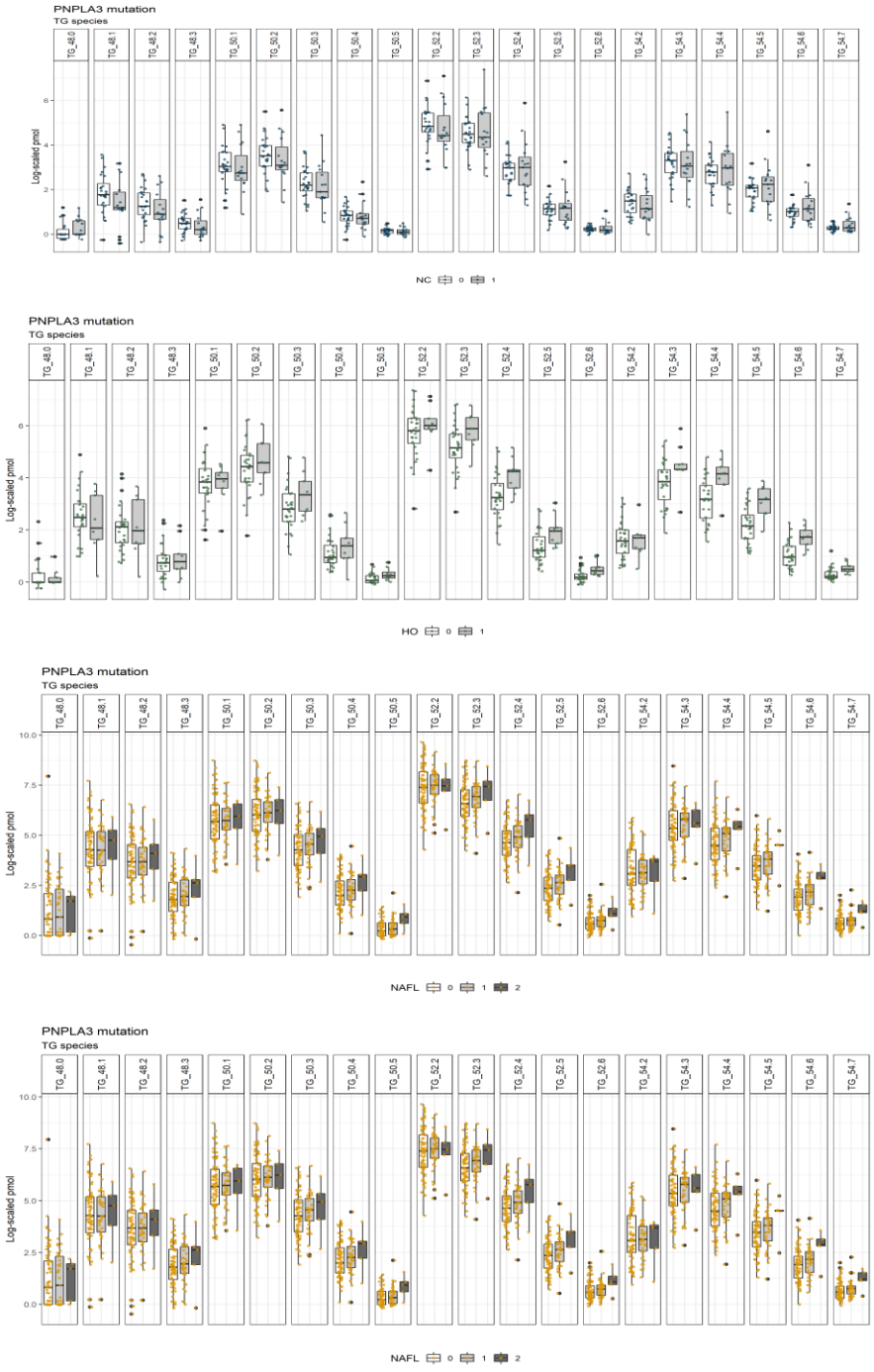
	Vvedenskaya & Rose *et a*l	Gorden *et al*
CE	5.11	3.90
Cer	0.28	0.28
CL	0.03	0.27
DG	1.13	16.23
LPC	0.33	0.30
LPE	0.37	0.27
LPG	0.06	0.15
LPI	0.09	0.06
PA	0.24	2.73
PC	30.58	23.60
PC O-	0.21	0.62
PE	13.19	7.36
PE O-	0.59	1.14
PG	0.93	0.14
PS	5.67	2.33
SM	3.48	1.86
TG	37.71	38.76



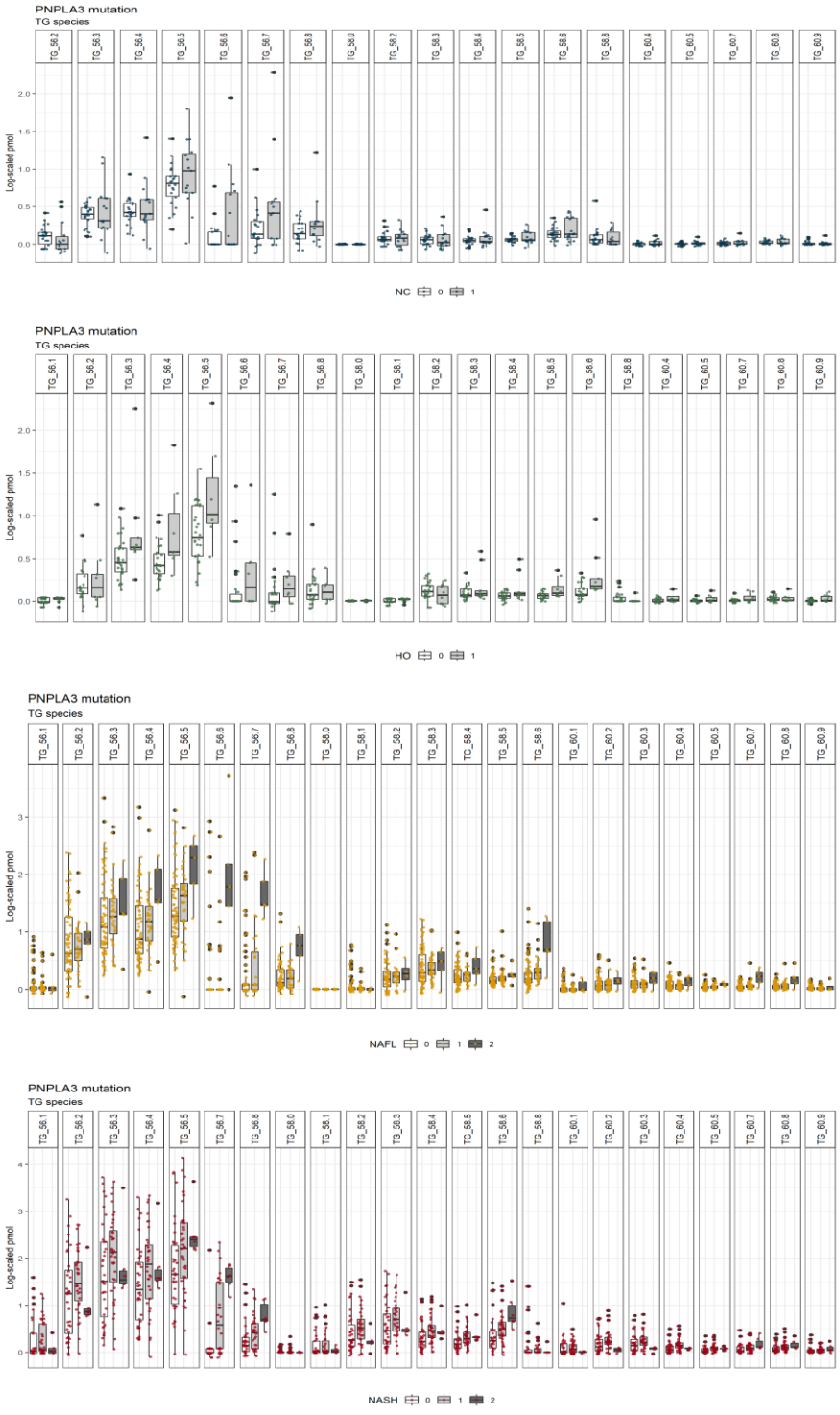
Supplementary Figure 6S: Comparison of plasma and liver lipidomes. Lipid abundances (in mol%) of normal control liver in females (f) and males (m) and of plasma in healthy females (f) and males (m). Plasma values were taken from [16].



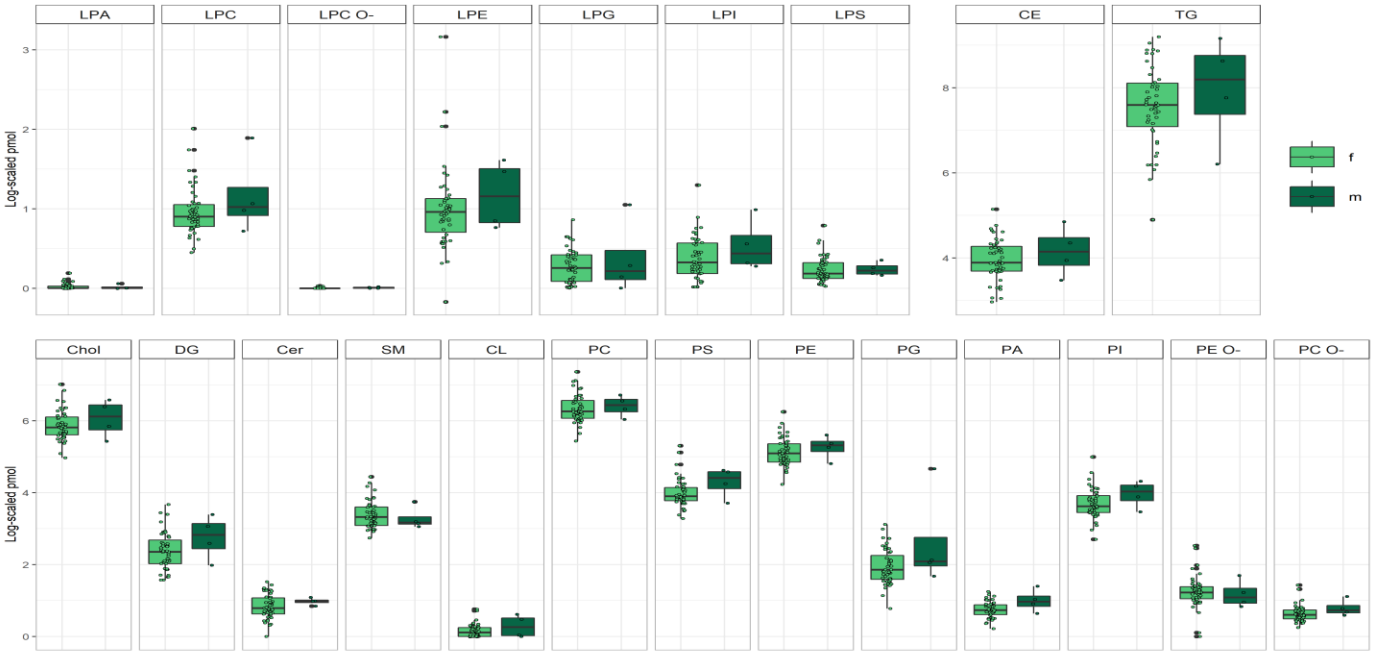
Supplementary Figure 7S. Changes of CE in samples with *PNPLA3* mutation. Boxplots of CE species abundance (log2 scaled pmol per µg of total protein) of the different *PNPLA3* mutation statuses (0 – no mutation, 1 – heterozygote, 2 – homozygote) in each patient group.



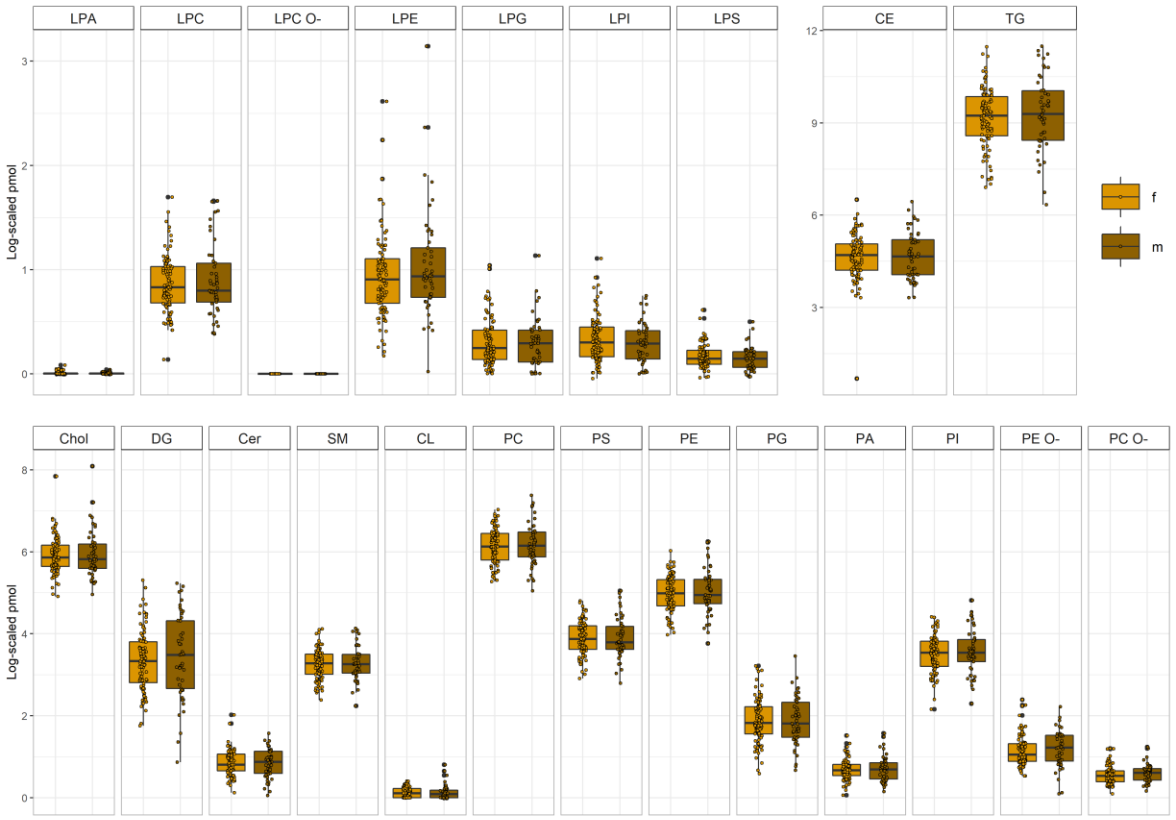
Supplementary Figure 8S A. **Changes of TG in the samples with *PNPLA3* mutation.** Boxplots of TG 48:0 – TG 54:7 (log2 scaled pmol per µg of total protein) for the four patient groups for different *PNPLA3* mutation statuses (0 – no mutation, 1 – heterozygote, 2 – homozygote).



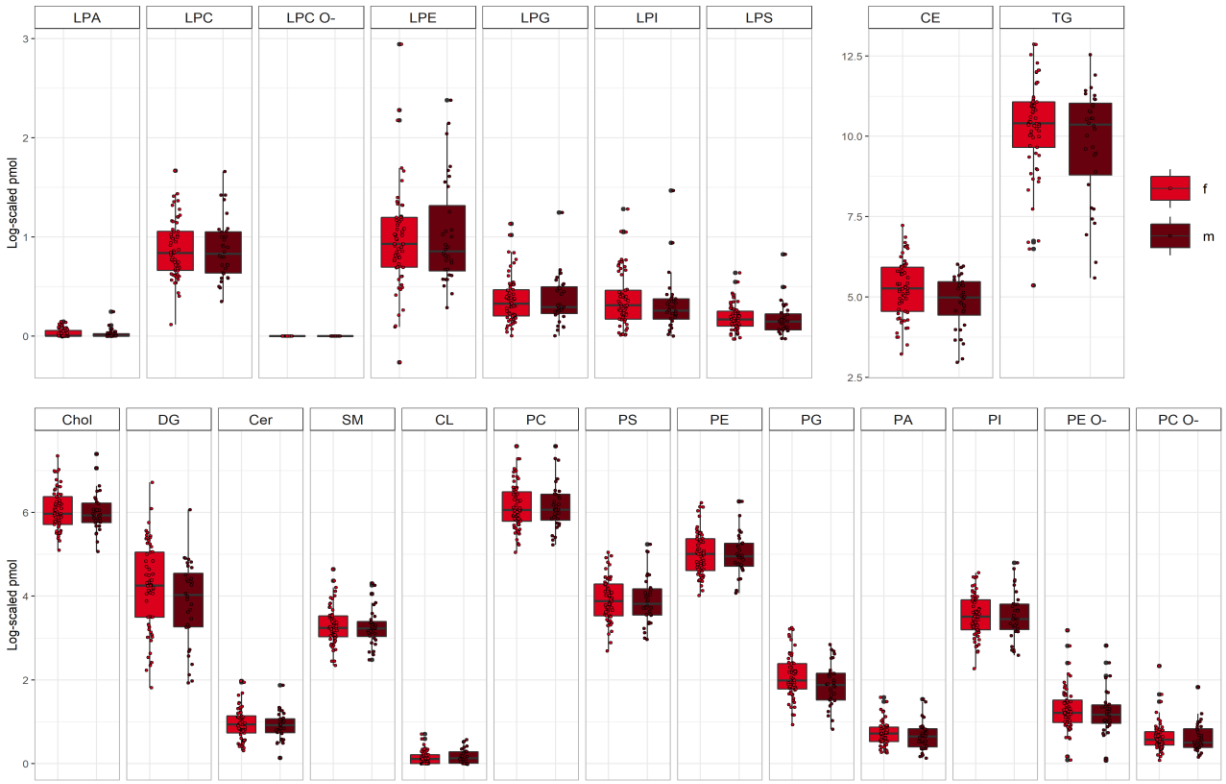
Supplementary Figure 8S B. Changes of TG in the samples with *PNPLA3* mutation. Boxplots of TG 56:1 – TG 60:9 abundances (log2 scaled pmol per µg of total protein) of the different *PNPLA3* mutation statuses (0 – no mutation, 1 – heterozygote, 2 – homozygote) in each patient group.



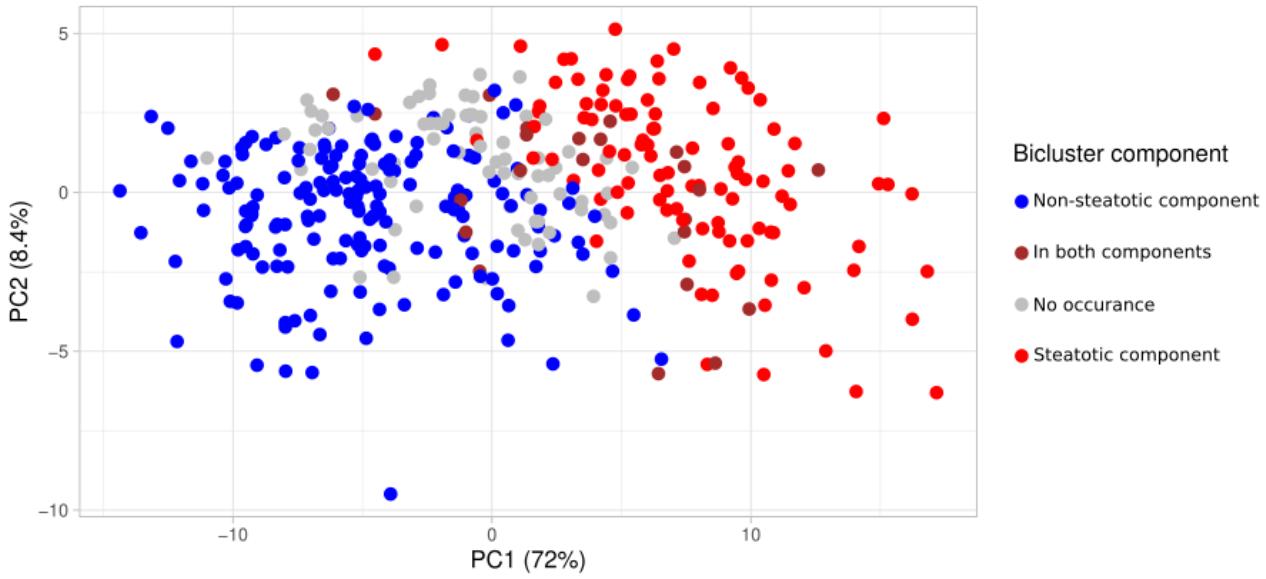
Supplementary figure 9S: Lipid classes abundance (log2-scaled pmol per µg of total protein) in HO (n=51) **females (f)** (n=47) **and males (m)** (n=4)**.**



Supplementary Figure 10S: Lipid classes abundance (log2-scaled pmol per µg of total protein) in NAFL (n=143) **females (f) (**n=94) **and males (m)** (n=49)**.**



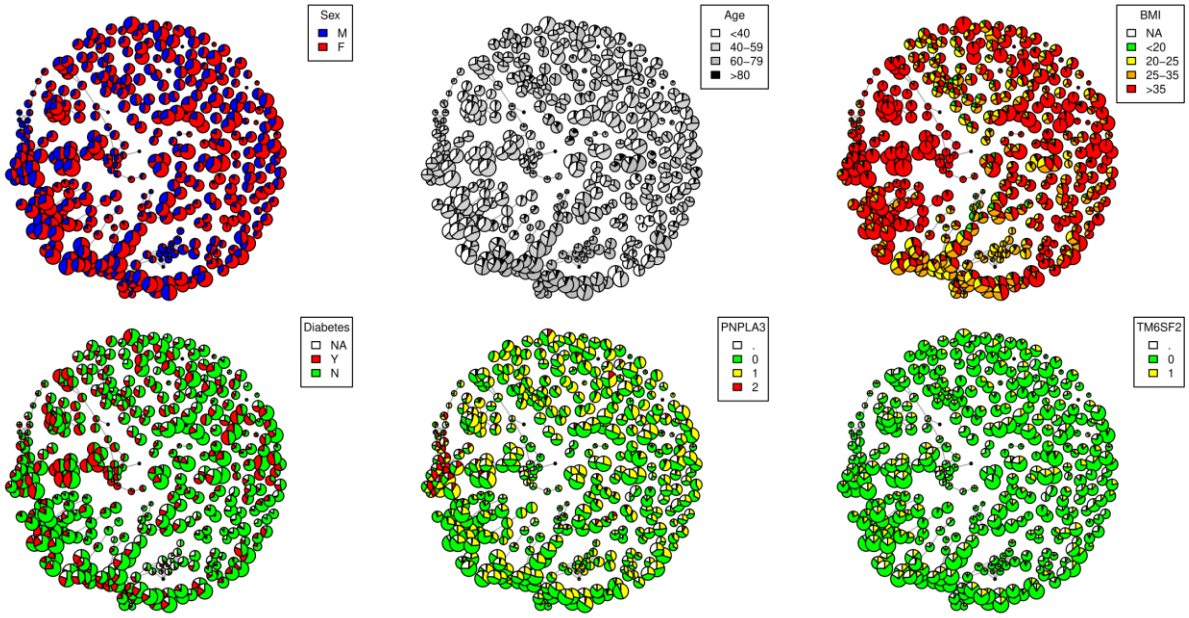
Supplementary Figure 11S: Lipid classes abundance (log2-scaled pmol per µg of total protein) in NASH (n=94) **females (f)** (n=62) **and males (m)** (n=32)**.**



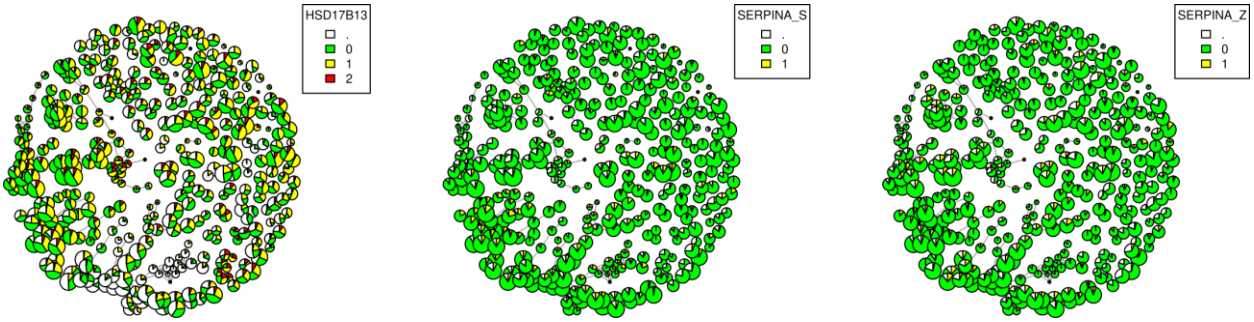
Supplementary Figure 12S Association of samples to bicluster components and DG and TG abundances in sample groups: PCA of the lipidomics data. Color-coding scheme is in the inset. Biclustering was able to separate most of the patients into disease (steatotic) and healthy (non-steatotic) groups.

Supplementary Table 3S: **The ten most abundant TG species in human liver from the four groups of the liver cohort, human white adipose tissue and human plasma from obese patients.** TG species from NC are labeled with different colors, and species overlapping in other liver conditions and other tissues are highlighted with the same color. Species shared between liver and plasma are in yellow and orange. *[17]; **[18]

Liver, NC	Liver, HO	Liver, NAFL	Liver, NASH	Plasma, obesity*	WAT, visceral, obesity **	WAT, breast, obesity **	WAT, lower back, obesity**	WAT, thighs, obesity**
TG 52:3	TG 52:2	TG 52:2	TG 52:2	TG 48:0	TG 52:2	TG 52:2	TG 52:2	TG 52:2
TG 52:2	TG 52:3	TG 50:1	TG 50:2	TG 46:0	TG 54:2	TG 54:2	TG 54:2	TG 54:3
TG 52:4	TG 50:2	TG 52:3	TG 52:3	TG 50:1	TG 54:3	TG 54:3	TG 54:3	TG 54:2
TG 50:2	TG 50:1	TG 50:2	TG 50:1	TG 48:1	TG 52:3	TG 52:3	TG 52:3	TG 52:3
TG 54:4	TG 54:3	TG 54:3	TG 54:3	TG 46:1	TG 54:4	TG 54:4	TG 54:4	TG 54:4
TG 54:3	TG 52:4	TG 48:0	TG 48:1	TG 46:2	TG 52:4	TG 52:1	TG 52:4	TG 52:4
TG 50:1	TG 54:4	TG 48:1	TG 50:3	TG 54:1	TG 52:1	TG 52:4	TG 52:1	TG 50:2
TG 54:5	TG 48:1	TG 54:4	TG 48:2	TG 48:2	TG 50:2	TG 50:1	TG 50:2	TG 52:1
TG 50:3	TG 50:3	TG 52:4	TG 54:4	TG 48:3	TG 50:1	TG 50:2	TG 50:1	TG 50:1
TG 48:1	TG 48:2	TG 50:3	TG 52:4	TG 51:2	TG 50:3	TG 54:1	TG 50:3	TG 56:3



Supplementary Figure 13S: Visualization of clinical confounding factors and mutations on the bicluster network from Figure 2 C. Lipidome compositions are clustered into networks of connected components comprising lipid compositions specific for several patients. Sex: M – male, F – female. Diabetes (type 2): NA – data not available, Y – yes, N – no. Risk factor mutations in *PNPLA3* and *TM6SF2*: unfilled – no data available; 0 - no mutation; 1 – heterozygote; 2 – homozygote.



Supplementary Figure 14S: Visualization of clinical confounding factors and mutations on the bicluster network from Figure 2 C. Lipidome compositions are clustered into networks of connected components comprising lipid compositions specific for several patients. Risk factor mutations in *HSD17B13*, *SERPINA S* and *SERPINA Z* mutations: unfilled – data not available, 0 - no mutation, 1 – heterozygote, 2 – homozygote.

Supplementary Table 4S: Number of covered samples and lipid signatures of selected bicluster-network components.

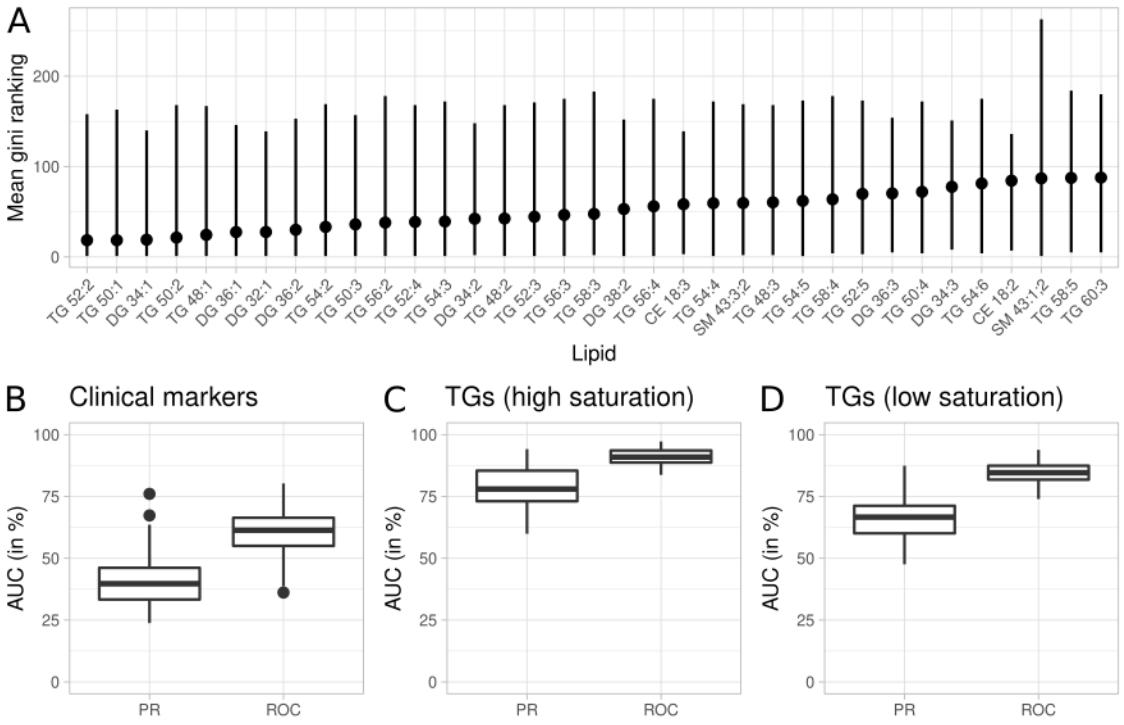
	Non-steatotic component 1	**Non-steatotic component 2**	**Steatotic component 1**	**Steatotic component 2**
Lipid class	178 samples	35 samples	48 samples	122 samples
CE	20:1, 20:2, 22:5	22:5	16:0, 16:1, 18:1, 18:2, 18:3, 20:3, 20:4	16:0, 16:1, 18:1, 18:2, 18:3, 20:2, 20:3, 20:4, 20:5, 22:1, 22:2, 22:6
Cer	34:1;2, 36:1;2, 38:1;2, 42:1;2, 44:2;2			34:1;2, 36:1;2
CL	74:8			
DG	38:6, 40:6		30:1, 32:0, 32:1, 32:2, 34:1, 34:2, 36:1, 36:2, 36:3, 36:4	30:0, 30:1, 32:0, 32:1, 32:2, 34:0, 34:1, 34:2, 34:3, 34:4, 36:1, 36:2, 36:3, 36:4, 36:5, 36:6, 38:1, 38:2, 38:3, 38:6, 40:2, 40:5
LPA	20:4			16:1, 18:1, 18:2

LPC	16:1, 18:0, 18:2, 20:3, 22:4, 22:6	16:0		18:3, 22:3
LPCO		16:0		
LPE	18:1, 20:0, 20:1, 20:2, 22:6	16:0, 18:0		
LPG	18:2, 22:6			22:6
LPI	18:0, 18:1, 22:6			
LPS	18:0, 22:1		20:5	20:1, 20:3, 22:1
PA	34:1, 34:2, 36:1, 36:4, 38:4, 38:5, 40:4, 40:6, 42:4	32:2	32:2	32:0, 32:2, 36:3
PC	30:0, 32:0, 32:1, 34:1, 34:2, 34:3, 34:4, 36:1, 36:2, 36:3, 36:4, 36:5, 36:6, 38:3, 38:4, 38:5, 38:6, 38:7, 40:4, 40:5, 40:6, 40:7, 42:2, 42:5	42:2		30:0, 36:1, 42:4, 44:4, 44:5
PC O-	32:0, 34:1, 36:1, 36:2,			

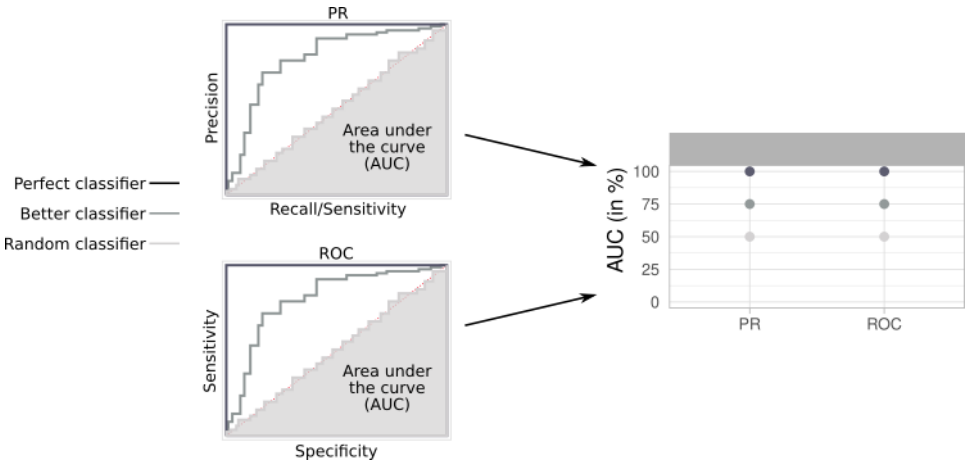
	36:3, 36:4, 36:5, 38:4, 38:5, 38:6, 40:5, 40:6			
PE	34:1, 34:2, 34:3, 36:1, 36:2, 36:3, 36:4, 36:5, 36:6, 38:1, 38:2, 38:4, 38:5, 38:6, 38:7, 40:4, 40:5, 40:6, 40:7, 40:8, 42:7			
PE O-	34:2, 34:3, 36:2, 36:3, 36:4, 36:5, 38:4, 38:5, 38:6, 38:7, 40:4, 40:5			
PG	34:0, 36:1, 40:8			
PI	32:0, 34:0, 34:1, 34:2, 36:1, 36:2, 36:3, 36:4, 38:3, 38:4, 38:5, 38:6, 40:4, 40:5, 40:6, 40:7			32:1, 34:1, 36:1

PS	32:1, 32:2, 34:1, 34:2, 34:4, 34:5, 36:0, 36:1, 36:2, 36:3, 36:4, 36:5, 38:0, 38:1, 38:2, 38:3, 38:4, 38:5, 38:6, 40:1, 40:2, 40:3, 40:4, 40:6, 40:7, 40:8, 42:3, 42:4, 42:9	32:1, 32:2		
SM	32:1;2, 33:1;2, 34:0;2, 34:1;2, 35:1;2, 36:1;2, 36:2;2, 38:1;2, 39:1;2, 40:1;2, 40:2;2, 41:1;2, 42:1;2, 42:2;2, 42:3;2, 43:1;2, 43:2;2		41:3;2, 42:4;2	41:3;2, 42:4;2, 43:3;2
TG	56:7, 56:8		50:4, 52:5, 52:6, 54:2, 54:3, 54:4, 54:5, 54:6, 54:7, 56:1, 56:2, 56:3, 56:4, 56:5,	48:0, 48:1, 48:2, 48:3, 50:1, 50:2, 50:3, 50:4, 50:5, 52:2, 52:3, 52:4, 52:5, 52:6,

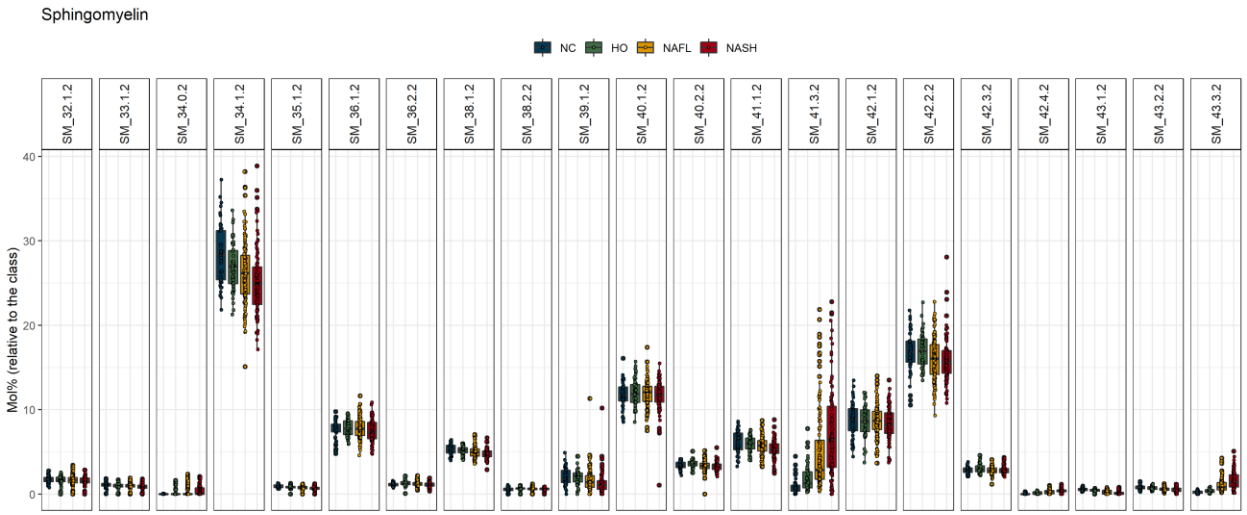
			56:7, 56:8, 58:1, 58:3, 58:4, 58:5, 60:2, 60:3	54:2, 54:3, 54:4, 54:5, 54:6, 54:7, 56:1, 56:2, 56:3, 56:4, 56:5, 56:6, 56:7, 56:8, 58:0, 58:1, 58:2, 58:3, 58:4, 58:5, 58:6, 58:8, 60:1, 60:2, 60:3, 60:4, 60:5, 60:7, 60:8, 60:9



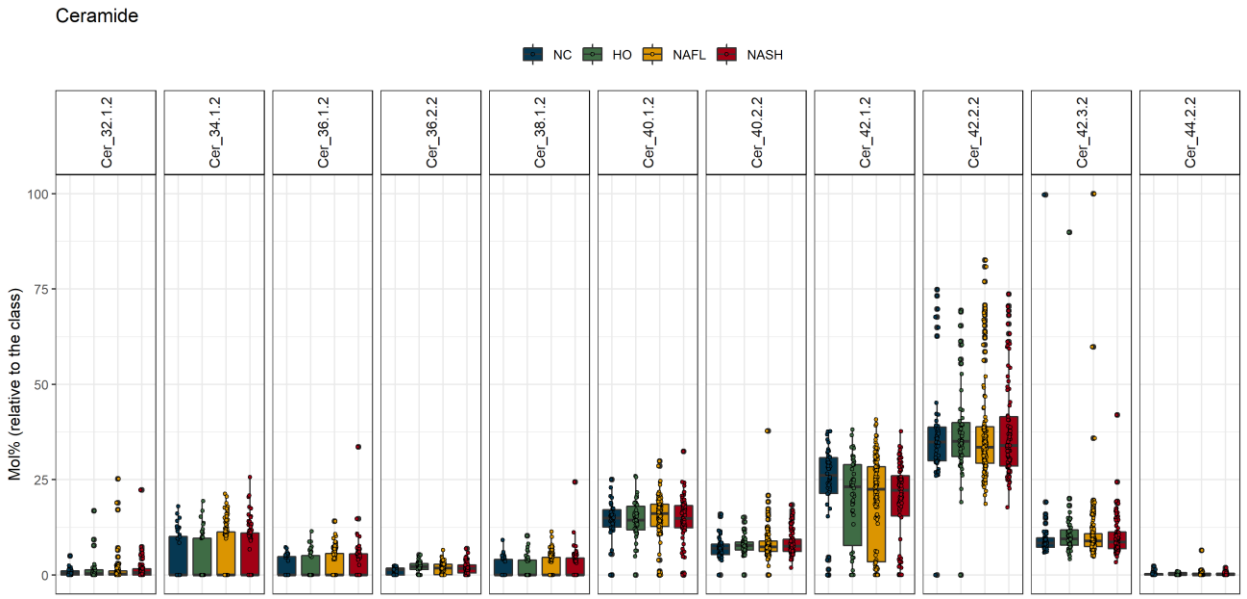
Supplementary Figure 15S: Classifiers trained to differentiate non-steatotic (NC & HO) from steatotic (NAFL & NASH) samples. (A) Feature importance defined by the mean Gini ranking for the classifications on bootstrapped lipidomics data. Lower values indicate higher importance. Error bars show the lowest and highest ranking achieved in different runs. The dot stands for the average. (B) Classification performance based on clinical markers (ALT, AST, AP, GGT, total bilirubin). (C) Classification performance based on highly saturated TGs (1-4 double bonds). (D) Classification performance based on low saturated TGs (5-8 double bonds).



Supplementary Figure 16S: Explanation of classification performance boxplots. Precision recall (PR) and receiver operator characteristic (ROC) area under the curve (AUC) are common measures of classification performance. A perfect classifier has a value of 100%, while a completely random classifier has a value of 50%. To ensure robust classification, boxplots showing the distribution of AUC values are used with multiple repetitions of the trained classification of bootstrapped data.



Supplementary Figure 17S: Profile of SM species (in mol%) in the four patient groups.



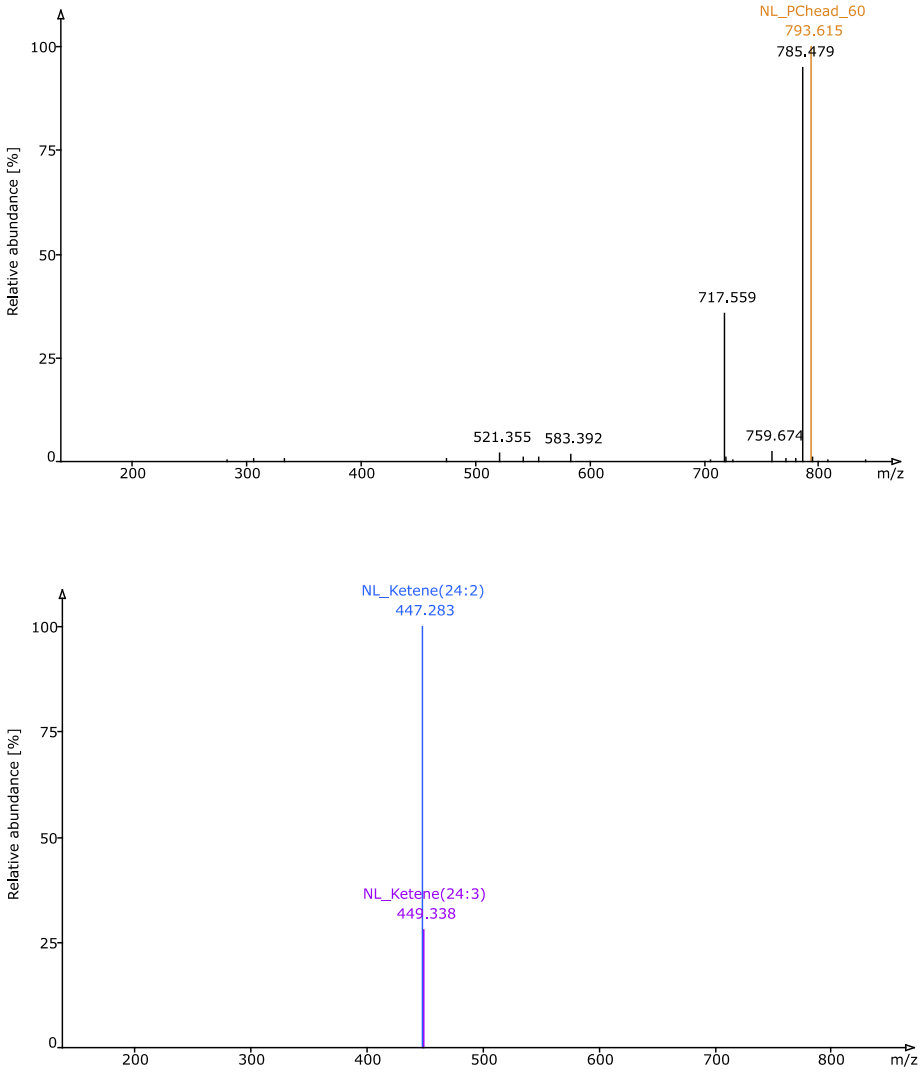
Supplementary Figure 18S: Profile of Cer species (in mol%) in four patient groups.

Supplementary Table 5S: Characterization of molecular composition of SM species by targeted LC-MS^n

SM species	Precursor *m/z* [2]	Major component[1]	Minor component[1]
SM 34:0;2	749.581	SM d18:0 / n16:0	SM d16:0 / n18:0
SM 38:2;2	801.613	SM d18:2 / n20:0	
SM 43:3;2	869.675	*inc.* [3]	
SM 41:3;2	841.644	*inc.* [3]	
SM 42:4;2	853.640	SM d18:2 / n24:2	SM d18:1 / n24:3
SM 43:1;2	873.707	SM d19:1 / n24:0	SM d18:1 / n25:0

[1] Semi-quantitative assessment based on the relative abundance of ketene loss fragments as exemplified in **Supplementary Figure 19S**
[2] adducts with formate anions $[M+HCOO]^-$
[3] identification of molecular species was inconclusive

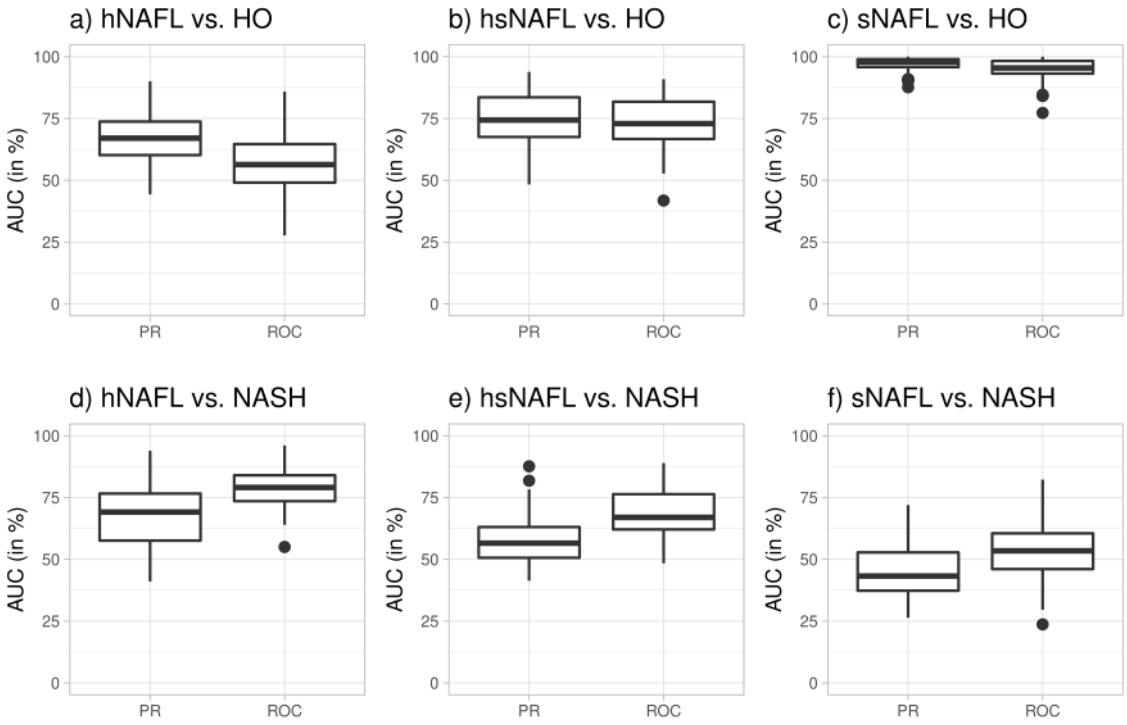
Molecular species: SM d18:1 / n24:3 and SM d18:2 / n24:2



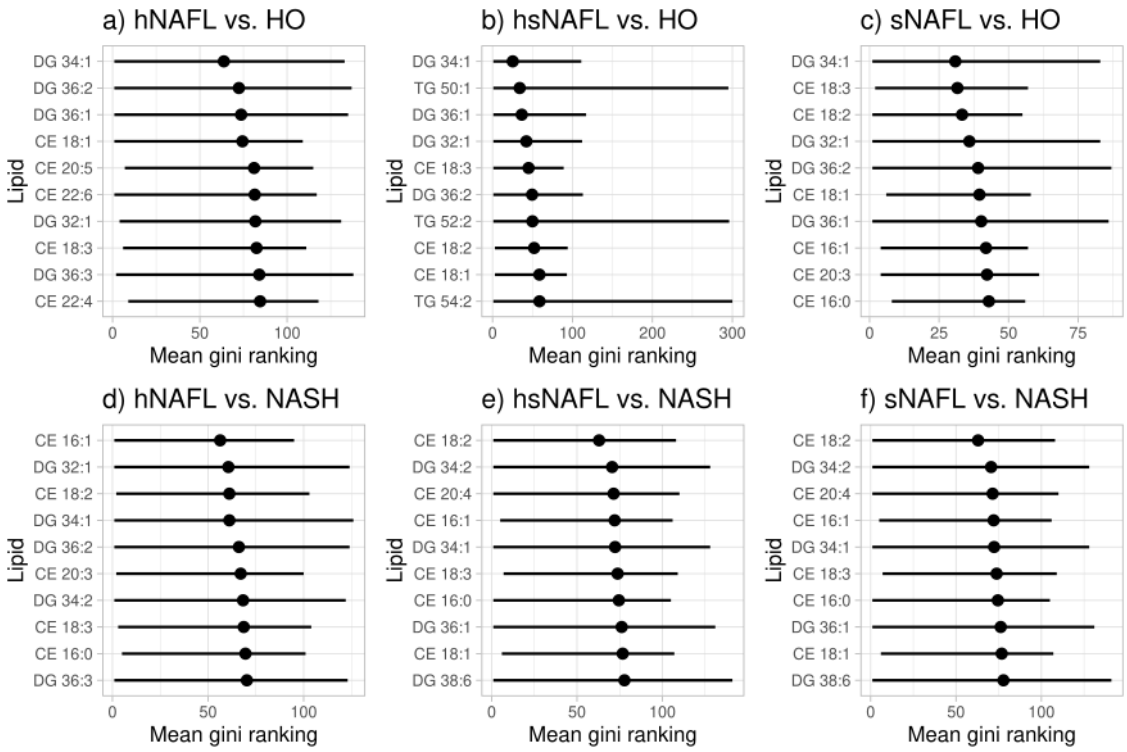
Supplementary Figure 19S: Molecular species of SM 42:4;2 identified by LC-MS^n.

Upper panel: MS^2 spectrum of [M+HCOO]$^-$ (precursor *m/z*: 853.6401; retention time: 25.00 min; file: Method_2_S2_1). MS^2 spectrum confirmed that this precursor belongs to SM class, yet revealed no further structural details.

Lower panel: MS^3 spectrum of precursor *m/z*: 793.615; retention time: 25.00 min; file: Method_2_S2_1) identified the fatty acid moieties of n24:2 and n24:3 by their ketene loss fragments



Supplementary Figure 20S: Classification performances. a) to f) precision recall (PR) and receiver operator characteristic (ROC) area under the curve (AUC) measures of the classification performance for NAFL subgroups (hNAFL, hsNAFL, sNAFL) against HO or NASH groups. Classification is based on the full lipidome.

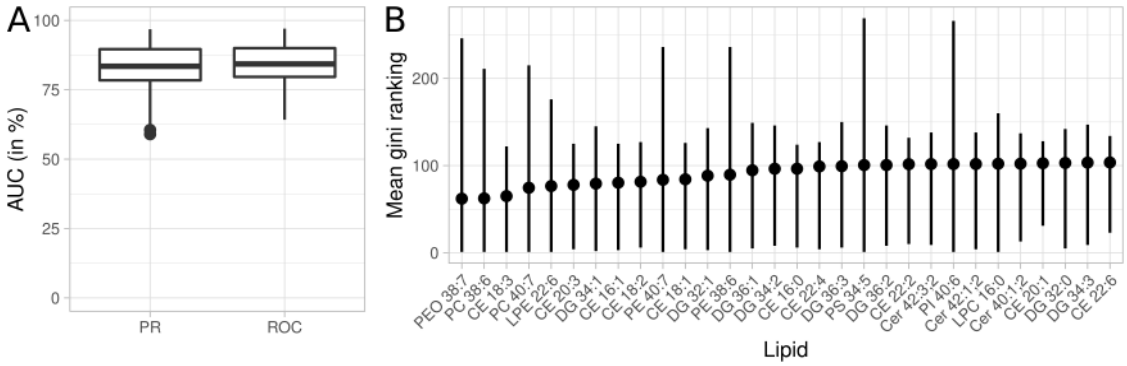


Supplementary Figure 21S: Lipids differentiating NAFL subgroups (hNAFL, hsNAFL, sNAFL) against HO or NASH groups based on the full lipidome composition. Measures are expressed as mean Gini ranking for the classifications on bootstrapped lipidomics data. Lower values indicate higher importance. Error bars show the lowest and highest achieved ranking in different runs. The dot shows the average.

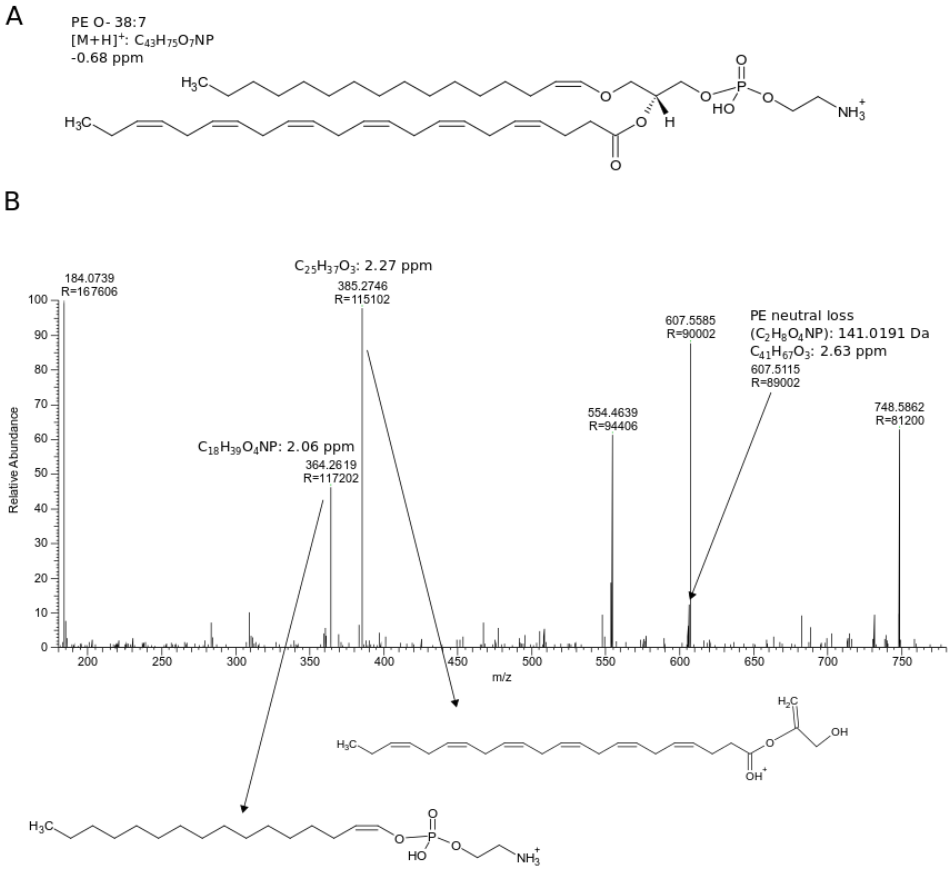
Supplementary Table 6S: Median values of indices related to NAS and fibrosis [1] for each condition including sub-groups of NAFL.

Condition	NAS total	NAS fat	NAS ballooning	NAS inflammation	Fibrosis	BMI
HO	0.0	0.0	0.0	0.0	0	47.4
hNAFL	1.2	1.1	0.1	0.0	0	45.7
hsNAFL	1.4	1.3	0.1	0.0	0	47.2
sNAFL	1.9	1.7	0.2	0.0	0	48.1
NASH	3.8	2.1	0.6	1.1	1	49.0

[1] Histological indices used for calculating NAS (NAS fat, NAS ballooning, NAS inflammation) and fibrosis according to Kleiner et al. [1]



Supplementary Figure 22S: Classification of samples from the sNAFL subgroup against other NAFL samples. (A) Classification performance as PR and ROC AUC. (B) Measures are expressed as mean Gini ranking for the classifications on bootstrapped lipidomics data. Lower values indicate higher importance. Error bars show the lowest and highest achieved ranking in different runs. The dot shows the average.



Supplementary Figure 23S: Confirming PE O- 38:7 by MS2 spectrum acquired from $[M+H]^+$ molecular ion. (A) Chemical structure of PE O- 38:7; (B) MS2 spectrum of PE O- 38:7 acquired from its $[M+H]^+$ molecular ion.

References

1. Kleiner, D. E., Brunt, E. M., Van Natta, M., Behling, C., Contos, M.J., Cummings, O. W., Ferrell, L. D., Liu, Y., Torbenson, M. S., Unalp-Arida, A. *et al* (2005) Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology.* **41**, 1313–1321

2. Thangapandi VR, Knittelfelder O, Brosch M, Patsenker E, Vvedenskaya O, Buch S, Hinz S, Hendricks A, Nati M, Herrmann A, *et al*. (2021) Loss of hepatic Mboat7 leads to liver fibrosis. *Gut* **70**, 940–50.

3. Li P, Wu Q, Burges CJ. (2008) McRank: learning to rank using multiple classification and gradient boosting. *Adv. Neural Inf. Process. Syst*., 897–904.

4. Triebl A, Trötzmüller M, Hartler J, Stojakovic T, Köfeler HC. (2017) Lipidomics by ultrahigh performance liquid chromatography - high resolution mass spectrometry and its application to complex biological samples. *J. Chromatogr. B*. **1053**, 72–80.

5. Fauland A, Köfeler HC, Trötzmüller M, Knopf A, Hartler JJ, Eberl A, Chitraju C, Lankmayr E, Spener F. (2011) A comprehensive method for lipid profiling by liquid chromatography-ion cyclotron resonance mass spectrometry. *J. Lipid Res*. **52**, 2314-2322.

6. Hartler J, Armando AM, Trötzmüller M, Dennis EA, Köfeler HC, Quehenberger O. (2020) Automated annotation of sphingolipids including accurate identification of hydroxylation sites using MSn data. *Anal Chem.* **92**, 14054–14062.

7. Hartler J, Triebl A, Ziegl A, Trötzmüller M, Rechberger GN, Zeleznik OA, Zierler KA, Torta F, Cazenave-Gassiot A, Wenk MR. (2017) Deciphering lipid structures based on platform-independent decision rules. *Nat Methods*.**14**, 1171–1174.

8. Hartler J, Trötzmüller M, Chitraju C, Spener F, Köfeler HC, Thallinger GG. (2011) Lipid Data Analyzer: unattended identification and quantitation of lipids in LC-MS data. *Bioinformatics* **27**, 572–577.

9. Johnson WE, Li C, Rabinovic A. (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**,118–127.

10. Bergmann S, Ihmels J, Barkai N. (2003) Iterative signature algorithm for the analysis of large-scale gene expression data. *Phys Rev E - Stat Physics, Plasmas, Fluids, Relat Interdiscip Top*. **67**, 031902.

11. Li G, Ma Q, Tang H, Paterson AH, Xu Y. (2009) QUBIC: a qualitative biclustering algorithm for analyses of gene expression data. *Nucleic Acids Res*. **37**, e101.

12. Hanczar B, Nadif M. (2011) Using the bagging approach for biclustering of gene expression data. *Neurocomputing* **74**, 1595–1605.

13. Strnad P, Buch S, Hamesch K, Fischer J, Rosendahl J, Schmelz R, Brueckner S, Brosch M, Heimes CV, Woditsch V, et al. (2019) Heterozygous carriage of the alpha1-antitrypsin Pi* Z variant increases the risk to develop liver cirrhosis. *Gut*. **68**,1099–1107.

14. Stickel F, Buch S, Nischalke HD, Weiss KH, Gotthardt D, Fischer J, Rosendahl J, Marot A, Elamly M, Casper M. (2018) Genetic variants in *PNPLA3* and *TM6SF2* predispose to the development of hepatocellular carcinoma in individuals with alcohol-related cirrhosis. *Am J Gastroenterol*. **113**, 1475–1483.

15. Gorden DL, Myers DS, Ivanova PT, Fahy E, Maurya MR, Gupta S, Min J, Spann NJ, McDonald JG, Kelly SL, Duan J, Sullards MC, Leiker TJ, Barkley RM, Quehenberger O, Armando AM, Milne SB, Mathews TP, Armstrong MD, Li C, Melvin WV., Clements RH, Washington MK, Mendonsa AM, Witztum JL, Guan Z, Glass CK, Murphy RC, Dennis EA, Merrill AH, Russell DW, Subramaniam S, Brown HA. (2015) Biomarkers of NAFLD progression: a lipidomics approach to an epidemic. *J Lipid Res*. **56**, 722–736.

16. Sales S, Graessler J, Ciucci S, Al-Atrib R, Vihervaara T, Schuhmann K, Kauhanen D, Sysi-Aho M, Bornstein SR, Bickle M, Cannistraci CV, Ekroos K, Shevchenko A. (2016) Gender, contraceptives and individual metabolic predisposition shape a healthy plasma lipidome. *Sci Rep*. **6**, 1–14.

17. Graessler J, Schwudke D, Schwarz PEH, Herzog R, Shevchenko A, Bornstein SR. Top-down lipidomics reveals ether lipid deficiency in blood plasma of hypertensive patients. PLoS One. 2009;4(7):e6261.

18. Al-Sari N, Suvitaival T, Mattila I, Ali A, Ahonen L, Trost K, Henriksen TF, Pociot F, Dragsted LO, Legido-Quigley C. Lipidomics of human adipose tissue reveals diversity between body areas. PLoS One. 2020;15(6):e0228521.