Supplementary Online Content

Tkachev A, Stekolshchikova E, Vanyushkina A, et al. Lipid alteration signature in the blood plasma of individuals with schizophrenia, depression, and bipolar disorder. *JAMA Psychiatry*. Published online January 25, 2023. doi:10.1001/jamapsychiatry.2022.4350

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods

Study participants

Participants belonged to sample cohorts collected at three distinct geographical locations in China (CN cohort), Germany as well as Austria (DE-AT cohort), and Russia (RU cohort, fepRU datasets).

Chinese cohort (CN)

Subjects included in this cohort were outpatients and recruited at the psychiatric center of the First Affiliated Hospital at Chongqing Medical University. SCZ (n = 170; age 36.9 ± 11.6; 53% female) and MDD (n = 222; age 37.7 ± 12.3; 59% female) patients were diagnosed by the Structured Clinical Interview using the DSM-IV-TR criteria. Patient evaluation was performed by two senior psychiatrists. Clinical symptoms were assessed using the Positive and Negative Syndrome Scale (PANSS) [1]. During the same time period, healthy control subjects (n = 153; age 37.8 ± 11.3; 54% female) were recruited from the medical examination center of First Affiliated Hospital at Chongqing Medical University. Healthy controls were required to have no current or previous lifetime history of neurological, DSM-IV Axis I/II, or severe somatic illnesses that could be associated with psychiatric symptom manifestation. Written informed consent was acquired from all participants. The protocol of this study was reviewed and approved by the Ethics Committee of Chongqing Medical University (Reference Number: 2017013).

German/Austrian cohort (DE-AT)

German and Austrian participants were part of the PsyCourse study, a multi-center, naturalistic, longitudinal study of severe psychiatric disorders conducted at 18 in- and outpatient sites across Germany and Austria (www.psycourse.de)[2]. For the current lipidomics analyses, adult (≥18 years) participants with a DSM-IV lifetime diagnosis of SCZ (n = 184; age 39.2 ± 12.8; 30% female) or BPD (n = 148; age 43.3 ± 13.4; 48% female) as well as control participants without a psychiatric diagnosis (n = 187; age 38.3 ± 16.2; 57% female), for whom plasma samples were available, were selected. Diagnoses were assessed using parts of the Structured Clinical Interview for DSM-IV, Axis I Disorders (SCID-I) administered by trained raters. Clinical symptoms of SCZ patients were assessed by PANSS [1]. Exclusion criteria consisted of severe somatic illnesses that could be associated with psychiatric symptom manifestation. The study was approved by the ethics committees of the coordinating Universities of Munich and Göttingen (reference number 17-13)

as well as the local ethics committees of the participating centers, wherever required. Written informed consent was obtained from all participants. The entire study was conducted in line with the Declaration of Helsinki.

Russian cohort (RU)

Subjects included in this cohort were inpatients and recruited from of the Mental Health Research Center, Moscow. The cohort included adult (≥18 years) participants with a diagnosis of SCZ (n = 82; age 31.2 ± 8.4; 23% female), BPD (n = 36; age 28.2 ± 9.1; 67% female) or MDD (n = 34; age 31.9 ± 10.4; 71% female) according to the ICD-10 criteria. Sample collection was performed at the Neuroimmunology Laboratory of the Mental Health Research Center, Moscow. The control group consisted of healthy volunteers from Mental Health Research Centre who had no signs of psychiatric disorders and met the same exclusion criteria (age<18 years, family history of any psychiatric disorder, neurological illness or severe somatic illnesses that could be associated with psychiatric symptom manifestation, recent surgery, pregnancy, substance and alcohol abuse) (n = 138; age 29.5 ± 8.3; 22% female). Patient evaluation was performed by board-certified psychiatrists of the same center. Clinical symptoms of patients were assessed by PANSS [1]. The study was approved by the local ethics committee of the Mental Health Research Centre (Protocol No. 281; 05/05/2016). Written informed consent was obtained from all participants. The entire study was conducted in line with the World Medical Association Declaration of Helsinki formulating ethical principles for medical research involving human subjects.

Russian First Episode dataset (fepRU)

Participants were inpatients, diagnosed with a first episode of psychosis (n = 104; age 25.8 ± 5.9; 53% female) and recruited at N. A. Alekseev Psychiatric Hospital No. 1, Moscow, Russia. All participants underwent complete diagnostic evaluation by certified psychiatrists and individuals with intellectual disabilities, substance abuse and dependence and neurological disorder or severe somatic illnesses that could be associated with psychiatric symptom manifestation were excluded. Patients were treated for no more than several days prior to blood collection. Written informed consent was obtained from all participants. The control group (n = 94; age 27 ± 7; 31% female) consisted of healthy volunteers from Mental Health Research Centre who had no signs of psychiatric disorders and partially overlapped with control samples collected for RU sample cohort

(n = 69). The protocol of this study was approved by the Interdisciplinary Ethics Committee, Moscow (22/07/2017).

Sample collection

CN: Plasma was obtained from peripheral venous blood between 8 h and 18 h after various periods of fasting (4.6 ± 2.5 hours for SCZ, 4.1 ± 2.2 hours for MDD, overnight fast for CTL, eTable 4). Plasma samples were collected in 5 ml Vacutainer tubes containing the chelating agent ethylenediaminetetraacetic acid (EDTA) (BD Vacutainer, Franklin Lakes, NJ, USA). Tubes were centrifuged at 4°C at 1500g for 15 min. The supernatant was stored immediately in 500µl aliquots at -80°C.

DE-AT: Plasma was obtained from peripheral venous blood between 8 h and 18 h from non-fasting individuals. Plasma samples were collected in 4 ml Vacutainer tubes containing the chelating agent ethylenediaminetetraacetic acid (EDTA) (BD Vacutainer, Franklin Lakes, NJ, USA). Tubes were centrifuged at 4°C at 1100g for 15 minutes. The supernatant was stored immediately in 500µl aliquots at -80°C.

RU and fepRU: Plasma was obtained from peripheral venous blood in the morning from individuals that underwent an overnight fast. Plasma samples were collected in 4 ml Vacutainer tubes containing the chelating agent ethylenediaminetetraacetic acid (EDTA) (BD Vacutainer, Franklin Lakes, NJ, USA). Tubes were centrifuged at 4°C at 1100g for 15 min. The supernatant was stored immediately in 500µl aliquots at -80°C.

Experimental setup

Lipidome measurements were conducted in four temporally separate experimental runs. There were three primary experimental runs corresponding to the three geographical cohorts: CN, DE-AT, RU. Sample from each of these cohorts were processed separately using the same experimental conditions and the same computational pipeline. The lipid measurement for fepRU samples were produced in a fourth experimental run. These samples were processed using the same experimental conditions and the same computational pipeline.

Lipid extraction

Samples were randomized and split to 96-samples batches. Ten empty tubes without plasma (extractions blanks) were placed in the end of last batch in cohort and subjected to the same

analysis steps as plasma samples. Two different types of quality control samples – extraction (EQC) and technical (TQC) – were incorporated every 12th position to account for sample preparation and measurement variability. EQC and TQC samples consisted of aliquots of plasma, mixed before and after extraction, respectively.

Solvents and extraction buffers were prepared in stock. Mixture of isotopically labeled lipids (SPLASH LIPIDOMIX Mass Spec Standard, Avanti Polar Lipids) were added to methyl tert-butyl ether (MTBE) based on 1 vial for 100 mL of solvent. Internal standards were used for quality control and retention time alignment, but absolute concentrations were not assessed in this study. All tubes and solvents were pre-cooled down to 0°C and manipulations carried out on ice. Frozen plasma was thawed on ice for 2 hours. Then cold methanol (300 µL) was added to 40 µL of sample aliquot and vigorously vortex for 1 min. After, 1 mL of cold MTBE with spiked standards was added, mixture was sonicated for 10 min, incubated for 40 min at 4 °C in a shaker and then sonicated for another 10 min. Phase separation was induced by adding 250 µL of MS-grade water. Extract was vortexed for 1 min at 4 °C, then centrifuged for 10 min at 13,000 rpm, at 4 °C. A total of 1000 µL of the upper layer, containing most of the lipids, was collected to a fresh Eppendorf tube. 400 µL of re-extraction buffer (MeOH:MTBE:H_2O = 3:10:2,5) were added to lower phase. Sample was vortexed and centrifuged for 10 min at 13,000 rpm, at 4 °C. Additional 300 µL of upper phase were collected, and combined organic phases were evaporated to dryness in a Speed Vac concentrator at 30 °C. Dried lipid pellets were stored at -80°C before analysis. Then pellets were resuspended in 200 µl ice-cold acetonitrile:isopropanol mixture(7:3 (v:v)). After brief rigorous vortexing the samples were shaken for 10 min, sonicated in an ice-cooled sonication bath for 10 min, and centrifuged 5 min at 13,000 rpm. For mass spectrometry analysis, samples were diluted 1:5 and 1:2 with acetonitrile:isopropanol (7:3 (v:v)) for positive and negative ionization mode, respectively.

Liquid chromatography and mass spectrometry

The liquid chromatography coupled with mass spectrometry system consisted of a Waters Acquity UPLC system (Waters, Manchester, UK) and a Q Exactive orbitrap mass spectrometer (Thermo Fisher Scientific, USA) equipped with a heated electro-spray ionization (HESI) probe. Separation of lipids was performed at 60 °C using a reverse phase ACQUITY UPLC BEH C8 Column (2.1 × 100 mm, 1.7 µm, Waters co., Milford, MA, USA) coupled to Vanguard precolumn at a flow rate

of 0.4 mL/min. The mobile phases consisted of water containing 10 mM ammonium acetate (Buffer A), and a mixture of acetonitrile and isopropanol (7:3) containing 10 mM ammonium acetate (Buffer B). Both buffers contained 0.1 % formic (positive mode) or acetic acid (negative mode) by volume. Separation was carried out by gradient elution according to the following profile: 1 min 55 % B, 3 min linear gradient from 55 % to 80 % B, 8 min linear gradient from 80 % B to 85 % B, and 3 min linear gradient from 85 % B to 100 % B. After 4.5 min washing with 100 % B, the column was re-equilibrated with 55 % B for 4.5 min. The injection volume was 3 µL. Mass spectra were acquired both in positive and negative mode in different experiments with a mass range of m/z 100–1500 at a mass resolving power of 70,000 at m/z 200, AGC target: 1E6, Maximum IT: 100 ms. Ion source for scan mode was operated with following parameters: capillary temperature, 250 °C; aux gas heater temperature, 350 °C; spray voltage, 4.5 kV; S-lens RF level, 70; sheath gas (N2) flow rate, 45 arbitrary units (a.u.); auxiliary gas (N2) flow rate, 20 a.u.; sweep gas (N2) flow rate, 4 a.u. Data was acquired on the profile mode. External mass calibration (Pierce LTQ Velos ESI Positive Ion Calibration Solution, Pierce Negative Ion Calibration Solution) without the use of specific lock masses was employed.

For lipid MS2 measurements, mass spectra were acquired in data dependent acquisition mode with active inclusion list. Parameters for Full MS scan mode were set for both polarities as follow. Resolution: 70,000 at m/z 200; AGC target: 5E5; IT: 50 ms; mass range: 200−1800. All ions from inclusion list within 10 ppm range were subjected to fragmentation with following parameters. AGC target: 5E4; IT: 100 ms. Intensity threshold was kept at 8E2 and isolation width was set at 1.2 Da. Stepped normalized collision energy was set at 15, 20, 25, 30 %. Data was acquired in profile mode, peptide match option was off, and isotope exclusion was set to on. Acquired spectra were manually curated and identifications were based on MS2 spectra, with a MS1 mass error of <5 ppm and MS2 mass error of <8 ppm.

Lipidome data processing

Spectra were processed using XCMS software [3], using the “centWave” method for peak detection, the “obiwarp” method for retention time correction, and the “fillPeaks” method for missing value imputation (for more details see Appendix). Abundance values were log2 transformed. Missing values as reported by XCMS were replaced by random values sampled from a normal distribution with mean 11.7 and standard deviation 0.25, which corresponds to the

approximate noise level seen in our experimental method. The number of missing values was, on average, less than 1% across samples in the features that were used for statistical analysis. The number of features retained at each step of the data processing is denoted in Table S1. Duplicated features were deleted (Table S1, *Duplicates*). Features were also removed according to retention time thresholds 0<RT<19 for positive ionization mode and 0<RT<14.5 for negative ionization mode (Table S1, *RT*). Isotopes were removed as well (Table S1, *Isotopes*). Contaminants were filtered out using extraction blank samples according to the following rule: features for which (mean abundance in CTL samples < mean abundance in blank samples+1) AND (mean abundance in non-CTL samples < mean abundance in blank samples+1) (in log2 scale) were removed from the analysis (Table S1, *Contam.*). Technical variability was corrected for all samples, including QC and blank samples, by subtracting the median batch value, calculated excluding the QC and blank samples. The feature abundances were then returned to their original scale by adding the median value of abundances in each batch. Features retaining high technical variability after batch correction were removed using QC samples according to the following rule: features with standard deviation across QC samples > 1 (in log2 scale) were removed from the analysis (Table S1, *Variability*). Features representing clusters of co-eluting derivative products of the same compound were removed (Table S1, *Co-eluting*). For more details on these procedures see Appendix. Samples were not normalized to total lipid content.

Cohort	Ionization Mode	*total*	*Duplicates*	*RT*	*Isotopes*	*Contam.*	*Variability*	*Co-eluting*
CN	NEG	5761	5681	4875	3256	2607	2121	2004
	POS	12073	11897	10161	6957	4905	4044	2530
DE-AT	NEG	6271	6163	5216	3504	2936	2429	2192
	POS	17636	17045	14684	10574	9016	7841	5247
RU	NEG	6649	6585	4563	3283	2143	1987	1650
	POS	6620	6391	5775	3858	2195	2079	1654

Intersection	NEG	POS	NEG+POS
	722	639	1361

Table S1. The number of features retained at each step of the data processing.

After independent processing of each of the three primary datasets (CN, DE-AT, RU), lipids were matched between datasets by m/z values and retention time. Only features reproduced in all three cohorts were used for downstream statistical analysis (Table S1, *Intersection*). Additionally, features detected in the fepRU dataset with mass-to-charge values and retention times matching the three first ones were retained in downstream analysis of the fepRU samples. Specifically, retention times were aligned between datasets using linear interpolation between internal standards (PC, PE, PG, PI, LPC, LPE, MAG, DAG, SM, and TAG). Next, lipids detected in each pair of

datasets were matched by m/z according to a 10 ppm threshold and by retention time values according to a 0.2 minute threshold, and lipids matched across the three primary datasets were used in downstream anlaysis. For more details, see Appendix. The fepRU dataset was matched to the RU dataset in the same manner, and features matching across all four datasets (CN, DE-AT, RU, fepRU) were employed in downstream analysis of lipid measurements of the fepRU samples.

Lipid features were putatively annotated by m/z values and retention times. For each considered lipid class, lipid features were matched by m/z values with mass accuracy cutoff of ppm = 10 to a generated theoretical m/z list using one selected adduct of choice, and then matches with inappropriate retention times were deleted as false annotations. Appropriate retention times were determined based on in-house retention time for previously annotated lipid species, as well as chain length and double bond content of the presumed annotations. For more information, see Appendix. The lipid species detected in our analysis were reported with one adduct for each lipid class (eTable 5).

SCZ-associated lipid features were further validated in tandem mass spectrometry experiment on pooled plasma samples. To obtain comprehensive information about head groups and fatty acid composition, precursor ions for each lipid were subjected to the higher-energy collisional dissociation (HCD) in positive and negative modes. Depending on MS2-spectra information, lipids were determined on bulk composition level (number of total carbons and degrees of unsaturation) or on fatty acid-resolved level (chain length and number of double bound denoted for each fatty acid). According to the guidelines for shorthand notation, these two levels are attributed to “lipid species level” and “lipid subspecies level”, correspondingly [4] (eTable 6). In particular, sn-position of fatty acids, location and configuration of double bounds, and stereochemistry were not specified. Lipid class was established by present of specific fragment ions (or neutral losses) for head group or sphingoid base and variable m/z values of fragment (or neutral loss) of attached fatty acid(s). Ether-linked species (plasmanyl- / plasmenyl-phospholipids) were distinguished by their reverse-phase chromatographic behavior. The elution order validated with authentic standard of PC(P-18:0/18:1(9Z), Avanti Lipids, 852467C).

Statistical analysis

Fasting status assessment

Because not all of the blood samples were drawn from individuals after an over-night fast, lipids most affected by fasting status were removed in the following way. SCZ samples from CN cohort, for which information about fasting status (0-8 hours) was available, were used to calculate the Pearson correlation between fasting status and lipid abundance. Lipids with correlation p-values < 0.05 were excluded from downstream statistical analysis (eFig. 15). This procedure mostly affected free fatty acids (eTable 5, eFig. 1).

Definition of SCZ-associated features

To find statistical differences between SCZ and CTL, we used Mann–Whitney U tests (with continuity correction) in each CN, DE-AT and RU cohorts independently. First, for the DE-AT cohort for which body mass index (BMI) information was available, we used a subset of samples (eTable 4, eFig. 16) balanced for BMI between SCZ and CTL samples when performing the Mann–Whitney U test to exclude the influence of BMI on statistical results. Next, a linear regression model was employed to regress factors of age and sex on the abundance of each lipid feature and subtracted this model from the abundance of the lipid feature, thus, correcting for sex and age. Finally, Mann–Whitney U tests (with continuity correction) for SCZ vs CTL samples were used to calculate p-values. The linear regression model correcting for age and sex was estimated using only the samples that were used in the Mann-Whitney U test, and not all available samples. In each of the CN, DE-AT and RU cohorts, the resulting p-values were corrected for multiple testing using Benjamini-Hochberg correction with FDR 10%. Lipids passing this 10% FDR threshold in each of the CN, DE-AT and RU cohorts were defined as SCZ-associated lipids. Permutation p-value was calculated in each cohort by repeating the same procedure after sample label permutation 1000 times, and assessing how many times the number of statistically significant lipids was equal or higher than the true one.

Age and sex corrected fold-changes between SCZ and CTL

To calculate average lipid abundance log ratios (log2 fold-changes) between SCZ and CTL, we first corrected for age and sex as described above, and then calculated the difference between median abundances in SCZ samples and CTL samples, using the same subsets of samples as for the Mann–Whitney U test described above, or the particular sample subsets of SCZ if explicitly stated (non-medicated patients and heavy smoking/ non-smoking SCZ patients).

Smoking status assessment

To assess the effect of smoking on SCZ patients, we used SCZ patients from the DE-AT cohort with no history of smoking (n = 26) and heavy smokers (reporting more than a pack of cigarettes a day; n = 35), and assessed the log2 fold-change of SCZ-associated lipids between these groups of SCZ samples and CTL. For SCZ-associated lipids, the log2 fold-changes between SCZ and CTL defined based on the whole set of DE-AT samples correlated well with the log2 fold-changes between heavy smokers SCZ individuals and CTL, as well as the log2 fold-changes between non-smokers SCZ individuals and CTL (respectively, Pearson correlation r = 0.96, p = 4e-44, and r = 0.96, p = 1e-42; eFig. 17).

Definition of SCZ-associated lipid classes

To define lipid classes most affected by SCZ, we used two different methods. For the first method, we utilized a one-sided binomial test to assess whether the number of SCZ-associated lipids from each lipid class showing the same direction of effect was greater than by chance. The lipid classes with nominal p-values < 0.05 in all the of the CN, DE-AT and RU cohorts were declared as most affected in schizophrenia, according to the first method. For the second method, we used GSEA (Gene Set Enrichment Analysis, gsea R package [5]) and all computationally annotated lipids form each lipid class to assess whether the lipid class as a whole showed a consistent direction of effect in SCZ compared to CTL. T-test statistics were used to rank the lipids by strength of effect in SCZ, as opposed to the Mann–Whitney U test statistics to avoid tied values. Lipid classes with adjusted p-values (according to fgsea package) < 0.05 in all of the CN, DE-AT and RU cohorts were declared as most affected in SCZ, according to the second method.

Prediction model differentiating SCZ from CTL

To construct a model differentiating between SCZ and CTL, we used 395 computationally annotated lipids and logistic regression with Lasso regularization with the following additional steps. First, we normalized the data for each of the CN, DE-AT, RU and fepRU cohorts independently. This was required, because of the properties of mass spectrometry data, i. e. the relative and not absolute values of feature intensities. To do this, we subsampled 20 male SCZ patients (FEP patients in the case of fepRU cohort) under the age of 35, and 20 male CTL individuals under the age of 35 and calculated the mean value and standard deviation for each feature using these samples. We repeated this subsampling procedure 1000 times to calculate, for each feature, the corresponding average mean value and average standard deviation. Next, we used

these values to normalize the lipid intensities in each of the CN, DE-AT, RU and fepRU cohorts by subtracting from each feature the calculated mean value and dividing by the calculated standard deviation. Finally, we used these resulting normalized abundances in the Lasso Logisitic Regression model (inverse of regularization strength C = 0.1.). To estimate the performance of the logistic model in each of the CN, DE-AT and RU cohorts, we used pooled samples from the three cohorts to select a training set of 750 samples (375 SCZ and 375 CTL samples), and a test set of 50 SCZ and 50 CTL samples not included in the training set. from one of the cohorts. Finally, the training set was used to construct a logistic regression model (inverse of regularization strength C = 0.1) and AUROC (area under the *receiver operating characteristic* curve) was calculated for the test set. We performed 100 such train/test splits to estimate the average AUROC values as well as the standard deviations. To validate the results using the fepRU cohort, we used the same scheme to construct the training set using only the CN, DE-AT and RU cohorts, but used first episode patients from the fepRU cohort to construct the test set. Because part of the CTL individuals for the fepRU cohort were the same as the RU cohort, care was taken that the same individual was not included in both the training and test sets.

To assess the association between prediction probabilities and symptom severity, we averaged, for each sample, its prediction probabilities in the test sets of the 100 random train/test splits. Then, we calculated the Spearman correlation coefficients between these prediction probabilities and symptom severity according to the PANSS scale. For each cohort, these correlations were not significant, indicating that higher symptom severity did not yield a higher probability of SCZ: r = -0.15 and p = 0.17 for RU cohort, r = 0.06 and p = 0.47 for DE-AT cohort, r = 0.19 and p = 0.07 for CN cohort.

To assess the performance of the model on permutated labels, we performed random label permutations and repeated the same procedures as described above. The estimated performance, as expected, corresponded to random predictions: AUROC = 0.499+ 0.04 for RU dataset, 0.496 + 0.03 for DE-AT dataset, 0.496 + 0.035 for CN dataset, 0.496 + 0.032 for fepRU dataset.

To select best performing 20 lipids, we calculated how often each feature was chosen by the model, i. e. included in the Lasso Logistic Regression model with non-zero coefficient. The 20 lipids chosen 95% or more of the times were defined as best performing: CAR 10:1, CAR 12:2, CAR

18:2, FFA 10:1, FFA 12:2, FFA 13:0, FFA 27:3, FFA 8:0, LPC O-26:1, PC 34:4, PC 36:4, PC 36:5, PC 40:5, PC O-37:4, PC P-37:1, PE 40:7, PE P-40:4, TAG 55:6, Cer d36:2, SM 41:2.

For the model coefficients reported in eTable 5, we used the totality of samples of the three CN, DE-AT, and RU cohorts, as well as the same regularization parameter C=0.1 for the Lasso Logistic Regression.

For the additional analysis of the effect of the regularization coefficient on model performance, we used the parameters C = (0.005, 0.0075, 0.01, 0.015, 0.025, 0.05, 0.08, 0.12, 0.2, 0.35, 0.5, 1, 2.5, 10, 100, 1000), and preformed the same train/test spitting schema involving 750 training samples from the three cohorts, and 100 test samples not included in the training from one cohort. The average number of features selected by the model was also calculated for these parameters.

We further report the performance of the model for cross-cohort reproducibility (eFig 4; eTable 8). In this case, training samples were selected from one cohort, and test samples from another. The procedure was the same as described before, except for the number of samples used in training and testing: 15 SCZ and 15 CTL samples were used in testing dataset, and the number of samples in the training set were different for each cohort, as marked on the plots (eFig 4; eTable 8).

Analysis of the reproducibility of direction of alterations for SCZ

Estimation of the reproducibility of the SCZ log2 fold-changes across the cohorts was conducted as follows. First, we used one cohort to define features with statistically significant differences between SCZ and CTL (as described above, Mann–Whitney U test, BH FDR 10%). For another cohort, we calculated how many of these features showed the same direction of fold-change as in the original cohort. This resulted in 6 different pairwise comparisons, hence 6 different point estimates. Second, we used two cohorts to find lipids that were statistically significant in both of these cohorts. We then used the third cohort to calculate how many of these lipids showed consistent direction of fold-change for this cohort and one of the first two. Again, this resulted in 6 different pairwise comparisons and 6 different point estimates. Third, we used lipids statistically significant in all three cohorts (SCZ-associated lipids), and calculated for each pair of cohorts, how many of these lipids showed consistent direction of fold-change. Because there were only three pairwise comparisons, this resulted in 3 different point estimates. Fourth, we used the fepRU cohort and the same SCZ-associated lipids to calculate how many of these lipids showed consistent

direction of fold-change for the fepRU cohort and the CN, DE-AT and RU cohorts. This resulted in three comparisons and three point estimates.

BPD- and MDD- associated features

To define BPD-associated features, we used a similar procedure as described above for SCZ to find the features statistically significant for each of the available cohort, but because of the larger sex imbalance for the RU cohort, we used a subset of samples balanced for sex between BPD and CTL for this cohort. Lipids significant for BPD in the two available cohorts (DE-AT and RU) were defined as BPD-associated lipids. MDD-associated lipids were defined similarly.

Comparison of defined lipid alterations in SCZ with the available literature

To compare our results to the available literature, we first collected publications reporting differences in lipids at the level of individual lipid species (or "bulk" structure) for non-medicated and medicated SCZ patients, BPD, and MDD patients. Next, we transformed the reported fold-changes into log2 scale, when possible, so that the resulting values were comparable to our results. For two studies [6] [7], this was not possible, because correlation with symptom severity was reported instead of differences between disease and CTL. In this case, the scale of the reported values was not comparable to our results, however, the direction of effect was nevertheless comparable. Next, we matched the SCZ-associated lipids with the reported lipids by bulk structure (such as PC 36:1, for example). We used reported differences in lipid abundances, if available in the publication, even if the authors did not report them to be statistically significant. Additionally, because most authors do not distinguish ether phospholipids, such as PC-O and PC-P from each other, we considered each reported ether phospholipid as two possible lipids (such as, for example, PC(O-36:1) and PC(P-36:0) for the reported lipid PC(O-36:1)) when matching the reported bulk annotation to the SCZ-associated lipids we annotate in this study. We used Spearman correlations to assess how well the reported differences corresponded to the differences we find in this study. We calculated the differences found in our study as the average log2 fold-changes for the three CN, DE-AT, and RU sample cohorts. The publications used for comparison, as well as the concatenated reported differences, can be found in eTables 2, 3, and 12.

Disease-associated genes

We defined genes associated with SCZ, BPD, and MDD using the GRASP [8] (https://grasp.nhlbi.nih.gov/Search.aspx) search engine using the following categories: schizophrenia, depression, bipolar disorder. Next, each gene list was filtered by the following trair, respectively: schizophrenia, major depressive disorder, bipolar disorder.

Lipid enzymes and the over-representation of disease-associated genes

Enzymes directly linked to lipid classes were defined using KEGG API and KEGG LinkDB [9]. Lipid classes quantified in the lipid dataset were converted to the corresponding LIPID MAPS classes, defined by the first 8 symbols of the LIPID MAPS ID [10]. Next, LIPID MAPS IDs corresponding to these classes were linked to KEGG compound IDs. Links not present in the KEGG database were manually added: LMGP0102 C05212, LMGP0103 C00958, LMGP0203 C04756, LMGL0201 C00165, LMGL0101 C02112, LMFA0101/LMFA0103 C00638, LMGP0202 C04475 (eTable 13). KEGG compounds were then linked to KEGG reactions, and KEGG reactions to KEGG enzymes, producing a list of enzymes directly linked to the lipid classes quantified in the lipid dataset.

We used permutation tests to calculate the over-representation of SCZ, BPD, and MDD genes among the enzymes corresponding to each lipid class. Specifically, we calculated the number of genes associated with SCZ, BPD, or MDD, and simultaneously linked to one of the lipid classes. Then, from the genes linked to the other lipid classes, we subsampled a random set of genes of the same size, and calculated how many SCZ, BPD, or MDD associated genes intersected with this random subset. After performing the procedure 10,000 times, we calculated the permutation p-value as the number of random subsamples for which this intersection was equal or larger than the true one, as well as the enrichment ratio, defined as the number of genes in the true intersection divided by the average number of genes in the random intersections.

Lipid-associated genes and the over-representation of disease-associated genes

We used [11] and merged the two summaries of genes found by the authors and previously reported ones to define genes associated with plasma lipid concentrations (lipid-associated genes). The reported gene clusters were separate into genes (ex. FADS1-2-3 as FADS1, FADS2, FADS3). To calculate the over-representation of genes in the intersection of genes associated with SCZ, BPD and MDD, and the over-representation of genes in the intersection of these genes and lipid-associated genes, we used random subsamples of genes of the same sizes as these gene sets. We

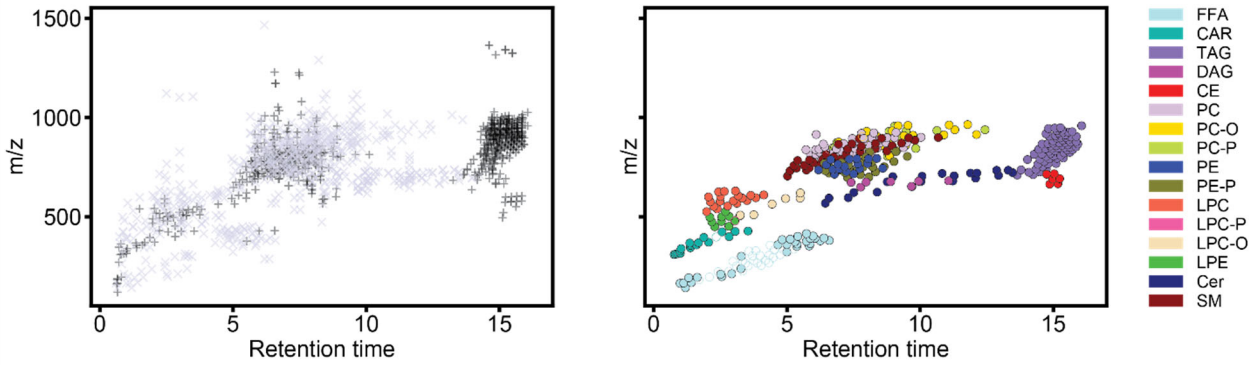
then calculated the average number of genes in the intersection of each of the Venn diagram sections and divided the true number by this average number.

To calculate the expected number of genes in the intersection of the disease-associated genes and these lipid-associated genes, we performed 10,000 random subsamples of genes of the same sizes and calculated the mean number of genes in the corresponding intersections.

Analysis of 13 non-medicated SCZ individuals

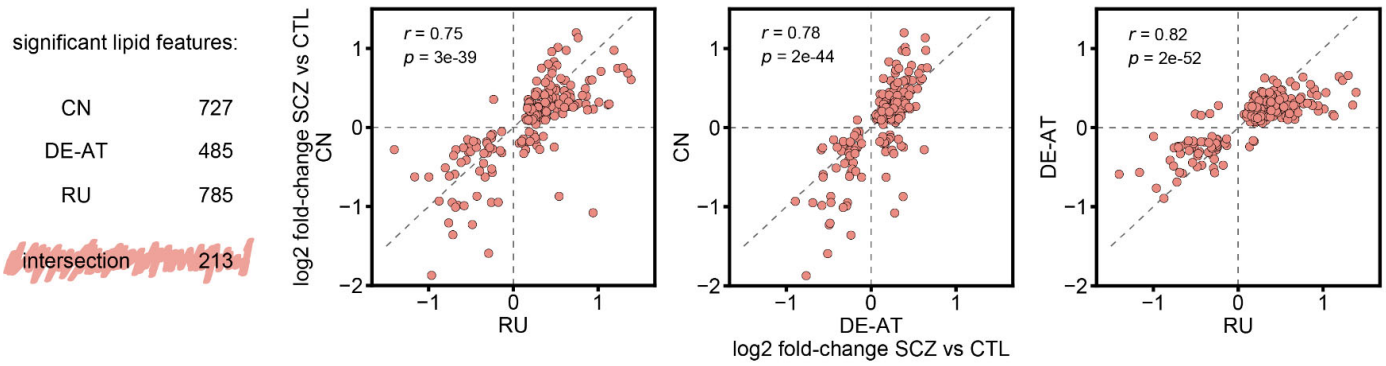
Analysis of 13 individuals contained in the DE-AT cohort who did not take SCZ-related medication for at least six months before blood sample collection largely due to non-compliance (eTable 4) replicated SCZ-associated differences identified using the complete sample set, with Pearson correlation 0.75 between lipid profiles of non-medicated individuals and those of SCZ in general (correlation of average lipid abundance log ratios between non-medicated SCZ and CTL, and total SCZ and CTL , $p < 0.00001$; Fig. 5A). Random subsampling of 13 CTL or medicated SCZ patients placed this correlation outside of CTLs and well within the SCZ patients' distributions (empirical $p = 0.018$ and 0.179, respectively; eFig. 13). Moreover, excluding triglycerides further shifted the correlation profile of non-medicated SCZ patients away from CTLs and into the interquartile range of the medicated SCZ subsample distribution (empirical $p = 0.001$ and 0.291 for random CTL and SCZ subsamples, respectively; eFig. 13).

Supplementary Figures



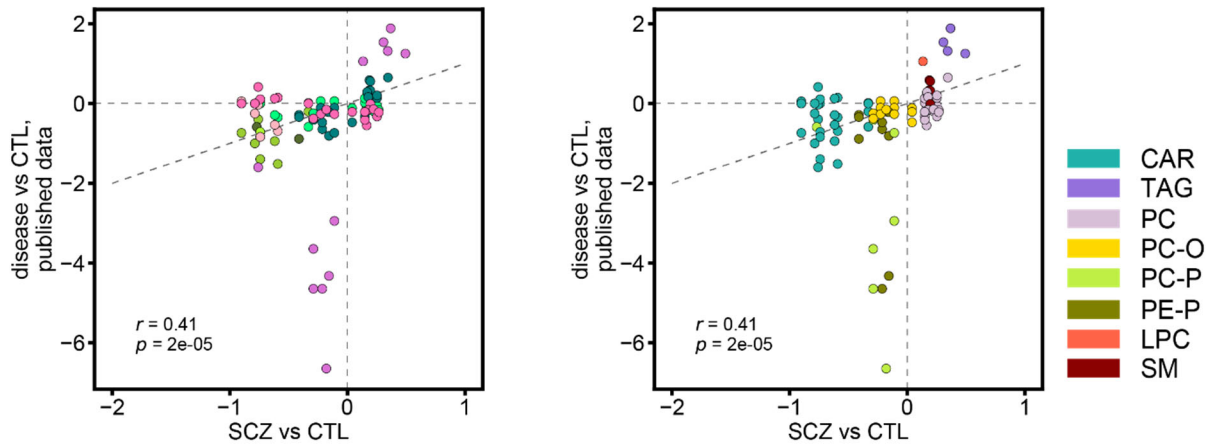
eFigure 1. Detected Features and Annotated Compounds

Left: The retention time and mass-to-charge values for all quantified 1,361 lipid features in positive (black, “plus” markers) and negative (grey, “x” markers) ionization modes. Right: same for 395 annotated lipids. Empty circles indicate lipids discarded from downstream analysis as related to fasting status (n = 30). Colors correspond to lipid classes, denoted to the right.



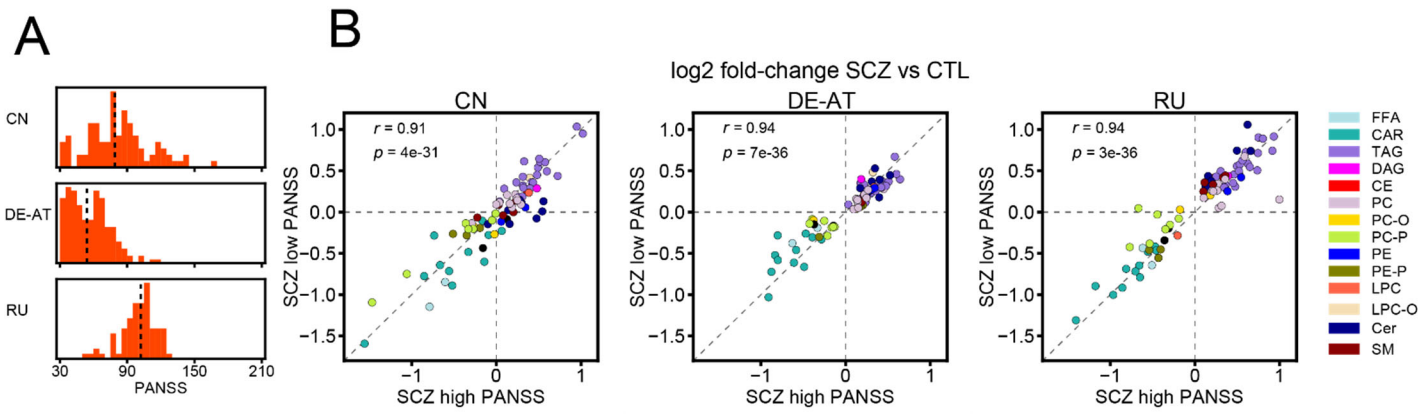
eFigure 2. SCZ-Associated Features and the Agreement of Their Abundace Alterations Between Cohorts

Left: number of lipid features showing statistical differences in each of the sample cohorts – CN, DE-AT, and RU – and their intersection (SCZ-associated lipid features). Right: pairwise comparisons of the lipid intensity differences between schizophrenia (SCZ) and control (CTL) samples between the cohorts. The intensity differences are displayed as the base two log-transformed fold change (log2 fold-change). Circles represent the 213 SCZ-associated lipid features. Pearson correlation coefficients and corresponding p-values are marked in the top left corner. The horizontal and vertical dashed lines indicate log2 fold-change = 0, the diagonal dashed line indicates the y-axis = x-axis line.



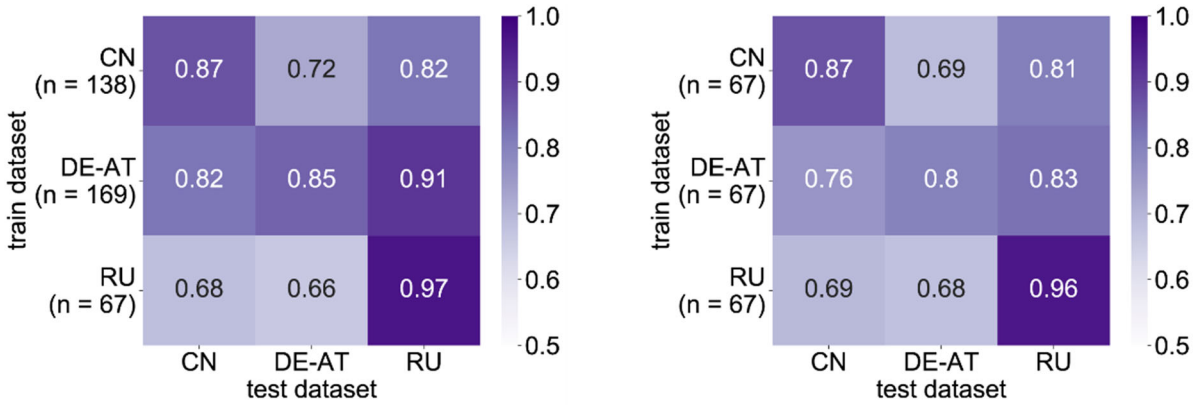
eFigure 3. Comparison of Detected and Published Average Lipid Abundance Log Ratios Between SCZ and CTL

The X-axis depicts base two log-transformed fold change (log2 fold-change) between SCZ and CTL (averaged for the three sample cohorts), y-axis depicts log2 fold-changes between SCZ and CLT as reported in the published studies (eTable 2). Each point (n = 101) corresponds to a particular lipid species reported in the literature and found among SCZ-associated lipids (n = 35 of the 77 SCZ-associated lipids). Left: colors correspond to published studies involving medicated SCZ (green) and non-medicated first psychotic episode or SCZ patients (pink). Shades of the corresponding color represent an individual publication. Spearman correlation coefficients and corresponding p-values are marked in the bottom left corner. The horizontal and vertical dashed lines indicate log2 fold-change = 0, the diagonal dashed line indicates the y-axis = x-axis line. Right: same, but colors correspond to lipid classes.



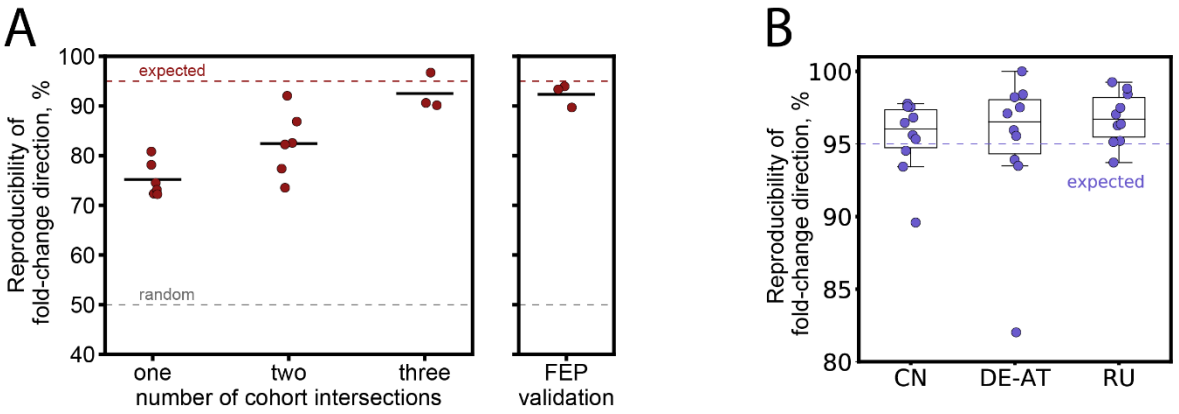
eFigure 4. Comparison of SCZ-Associated Lipid Differences Between Patients With High and Low PANSS Scores

(A) Distribution of PANSS scores for SCZ in each of the cohorts (CN, DE-AT, and RU). Dashed line demarcates the median value in each cohort. (B) Comparison of SCZ-associated lipid differences between patients with PANSS values lower than the median (low PANSS) and patients with PANSS values lower higher than the median (high PANSS) within each cohort. The differences correspond to the base two log-transformed fold change (log2 fold-change) between SCZ and CTL individuals. Pearson correlation coefficients and corresponding p-values are marked in the top left corner. Colors correspond to lipid classes. The horizontal and vertical dashed lines indicate log2 fold-change = 0, the diagonal dashed line indicates the y-axis = x-axis line.



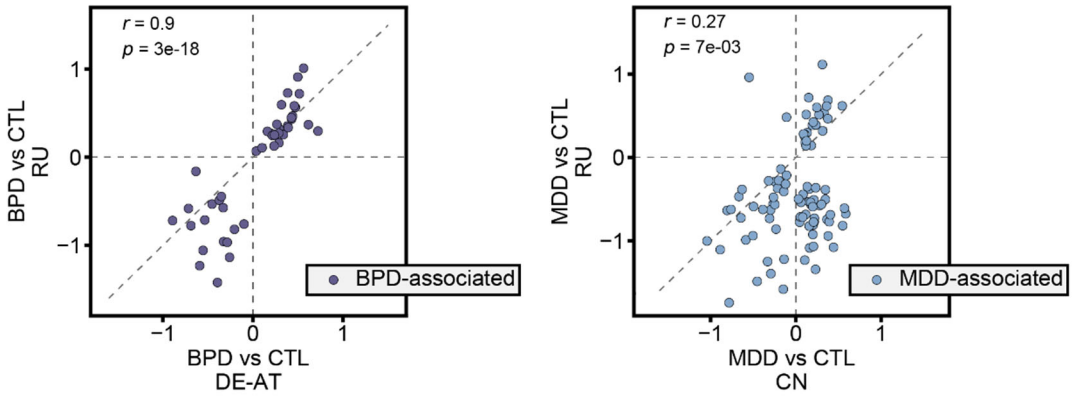
eFigure 5. Area Under the Receiver Operating Characteristic Curve (AUROC) Values for Lasso Logistic Regression Model Delineating SCZ and CTL Individuals

The model was trained and tested on samples from each of the possible pairs of CN, DE-AT, and RU cohorts. The average AUROC values for models trained on 100 random training/test subsets are depicted on the figure. Y-axis corresponds to the cohort used for training, and x-axis to the cohort used for testing. The number of SCZ and CTL samples used in the training is marked in parentheses on the left. Left: maximum number of samples was used for training. Right: equal number of samples for each cohort was used for training.



eFigure 6. Reproducibility of the Lipid Intensity Difference Direction Between SCZ and CTL Between and Within the Cohorts

(A) Reproducibility of the base two log-transformed fold change (log2FC) direction between SCZ and CTL individuals for lipids defined as SCZ-associated using varying number of independent cohorts. Each point represents the percentage of lipids with reproduced log2FC directions for a particular comparison between a pair of sample cohorts. One cohort: the log2FC of lipids statistically significant in one cohort were compared to the log2FC of the same lipids in another cohort. Two cohorts: lipids statistically significant in two cohorts were used to compare the log2FC in one of these cohorts and the third one. Three cohorts: lipids statistically significant in all three primary cohorts were used to compare the log2FC in a pair of cohorts. FEP validation: for lipids statistically significant in all three primary cohorts, we compared the log2FC direction between fepRU and each of the primary cohorts (CN, DE-AT, and RU). In the absence of unaccounted confounding factors between cohorts, the expected percentage of non-reproduced log2FC directions would be 5% with the 10% FDR cutoff for multiple testing correction. (B) We estimated reproducibility of the difference direction between SCZ and CTL by randomly splitting the SCZ and CTL samples within a cohort into two sets 10 times. We used one set to define statistically significant differences using 10% FDR threshold cutoff, and the second set – to assess the reproducibility of the difference direction between SCZ and CTL. Each dot depicts the percentage of lipid features with reproduced difference directions for one random splitting of the cohort. The boxplots illustrate the interquartile range of these estimated values. Dashed line demarcates the theoretically expected value when using 10% FDR threshold cutoff, which is 5%.



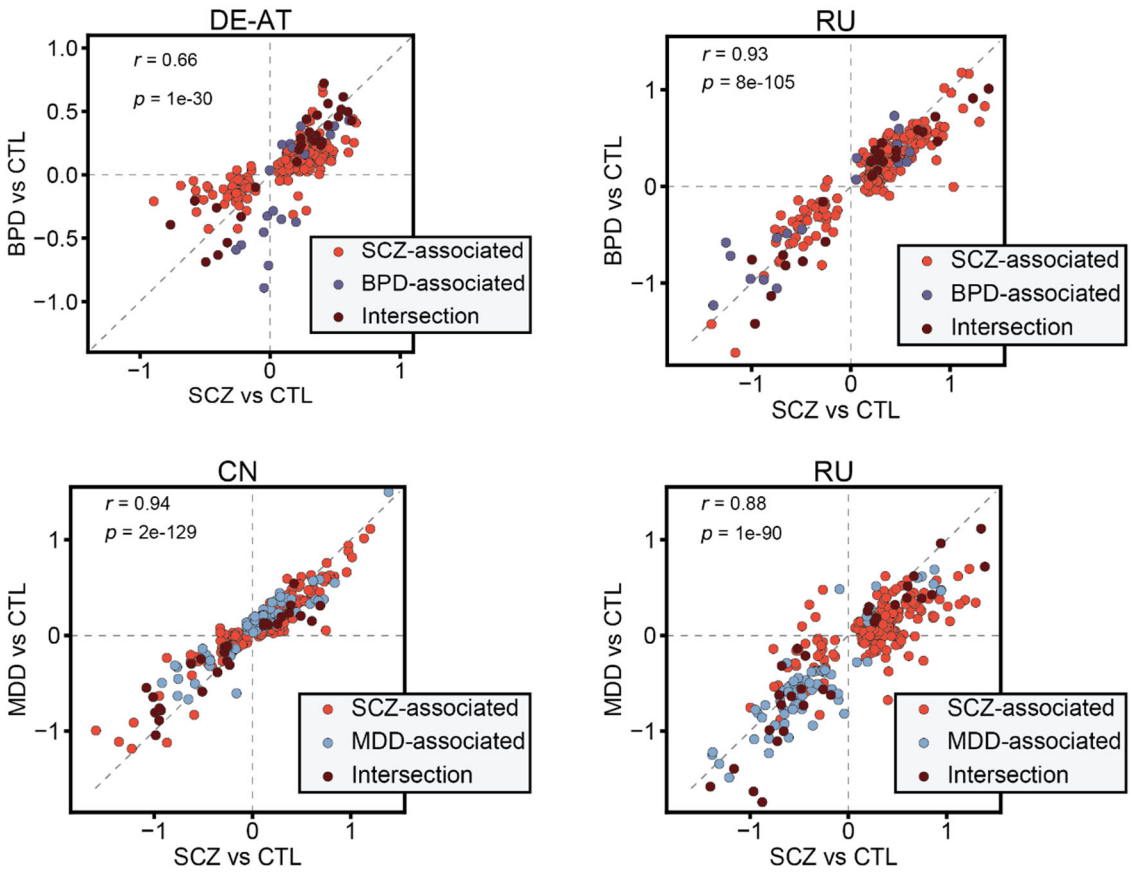
eFigure 7. Comparison of Statistically Significant BPD- and MDD-Associated Lipid Intensity Differences Between 2 Independent Cohorts

Comparison of the BPD-associated (left) and MDD-associated (right) lipid intensity differences between two available sample cohorts. The lipid intensity differences are displayed as the base two log-transformed fold change. Pearson correlation coefficients and corresponding p-values are marked in the top left corner of each plot. The horizontal and vertical dashed lines indicate log2 fold-change = 0, the diagonal dashed line indicates the y-axis = x-axis line.

SCZ-associated, total 213	BPD-associated, total 47	MDD-associated, total 97	num. lipids
+	-	-	162
-	+	-	15
-	-	+	63
+	+	-	21
+	-	+	23
-	+	+	4
+	+	+	7

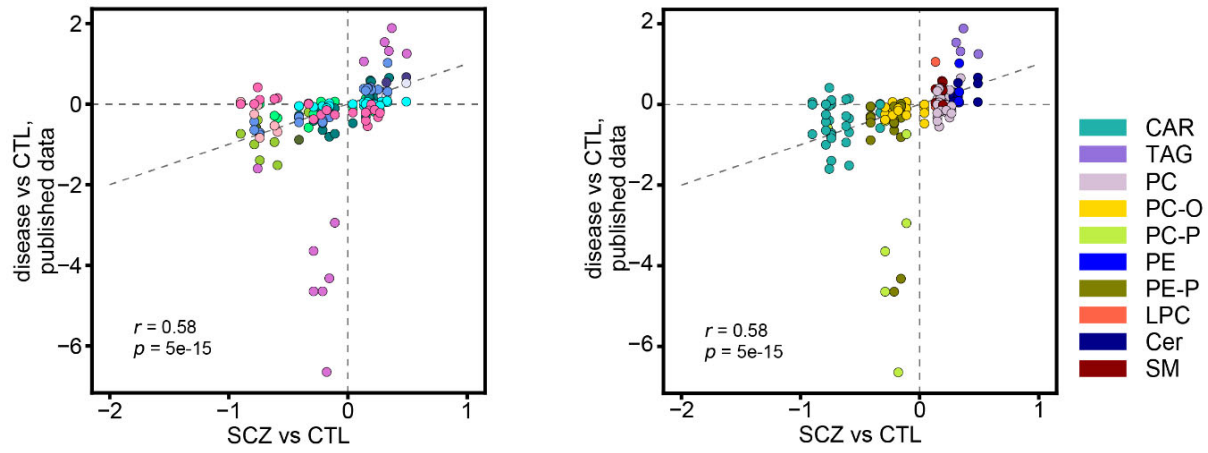
eFigure 8. Intersection of MDD-Associated, BPD-Associated, and SCZ-Associated Lipid Features

Number of lipids that are common and different for the lipid groups: SCZ-associated, BPD-associated, and MDD-associated: (+ + +) common for three disorders; (+ + -) (+ - +) (- + +) common for two disorders; (+ - -), (- + -), (- - +) found only for one disorders.



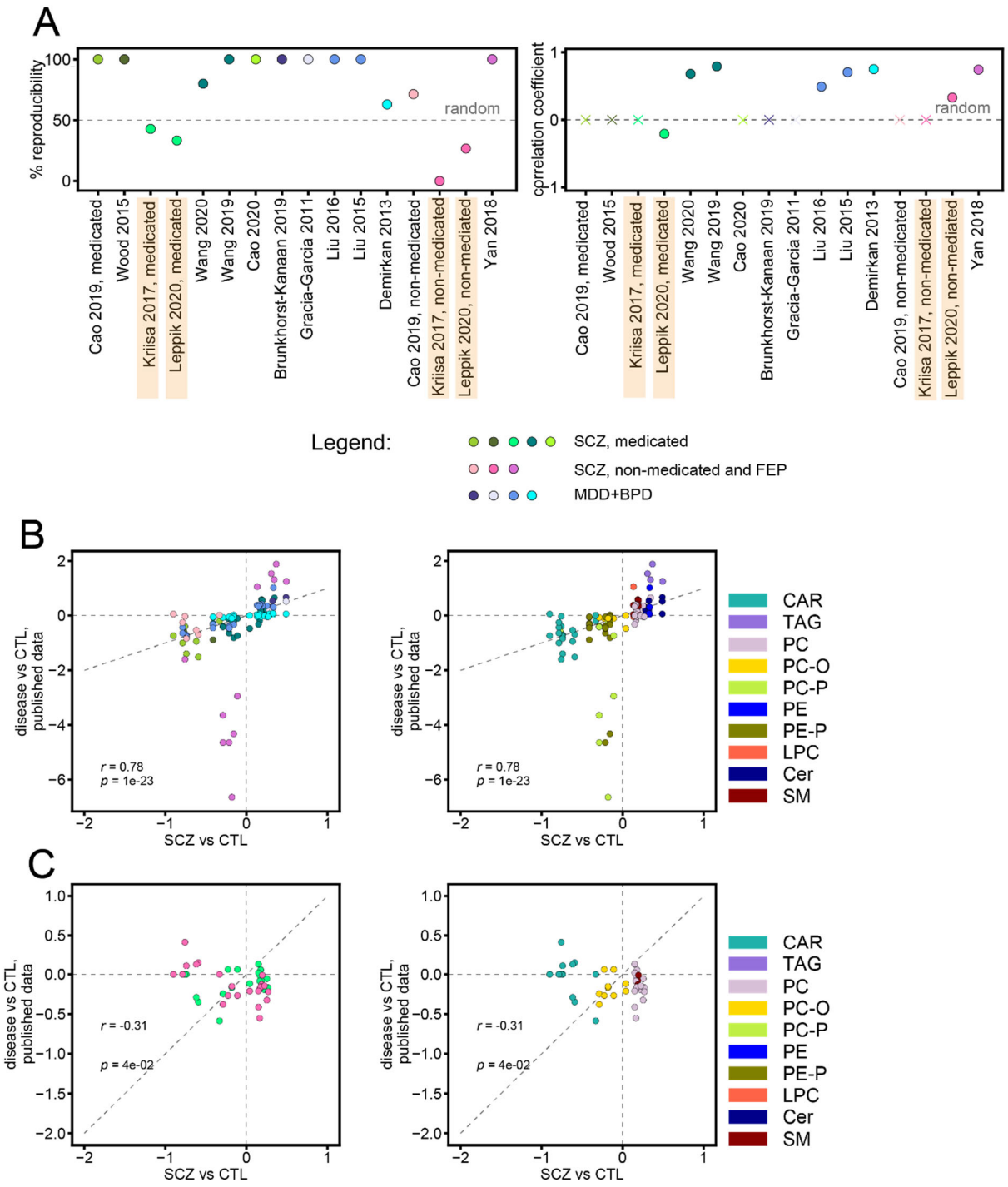
eFigure 9. Comparison Blood Plasma Lipidome Alterations Between Disorders Within Each of the Cohorts

Comparisons of the lipid intensity differences between SCZ and CTL, and BPD and CTL (top), or MDD and CTL (bottom), for each of the cohorts. The intensity differences are displayed as the base two log-transformed fold change (log2 fold-change). Circles represent the SCZ- (red), BPD- (dark blue), MDD- (light blue) associated lipid features, as well as their intersection (brown). Pearson correlation coefficients and corresponding p-values are marked in the top left corner of the panels. The horizontal and vertical dashed lines indicate log2 fold-change = 0, the diagonal dashed line indicates the y-axis = x-axis line



eFigure 10. Comparison of Detected and Published Average Lipid Abundance Log Ratios Between SCZ and CTL

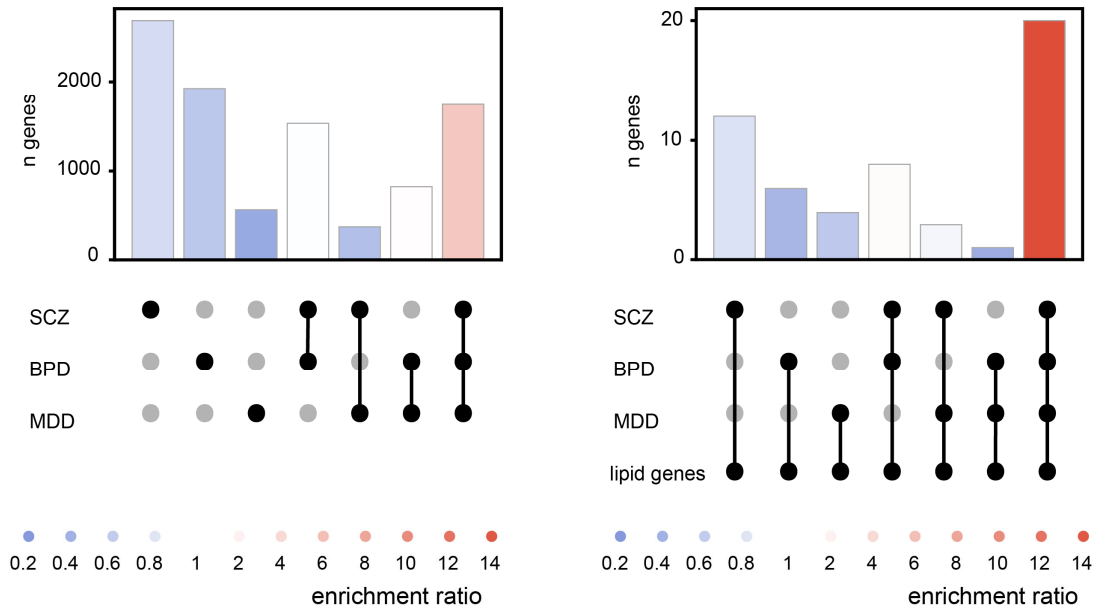
Comparison of the lipid intensity differences determined in our study and reported in the literature (eTable 2). Circles represent 38 SCZ-associated lipids with reported alterations. Top: colors represent diagnoses groups in the published data: medicated SCZ patients (green), non-medicated SCZ and FEP (pink), MDD and BPD (blue). Different shades of color mark individual studies (eTables 2,3). Bottom: colors correspond to lipid classes. The diagonal line indicates y = x.



eFigure 11. Comparison of Detected and Published Average Lipid Abundance Log Ratios Between SCZ and CTL by Individual Study

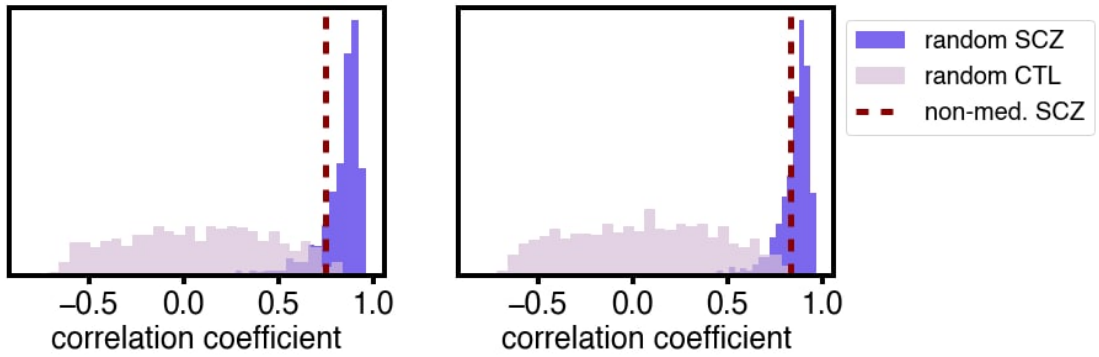
(A) Left: List of published studies (eTable 2) and the corresponding percentage of SCZ-associated lipids for which the direction of effect is in agreement with our results. Right: List of studies and the corresponding Spearman correlation between reported and detected differences between disease and CTL, x markers correspond to studies with less than 10 SCZ-associated lipids. Colors

represent diagnoses groups in the published data: medicated SCZ patients (green), non-medicated SCZ and FEP (pink), MDD and BPD (blue). Different shades of color mark individual studies (study references can be found in eTable 2), but studies using the same set of samples share the same shade. Orange color highlights two studies showing least agreement with our results. (B): Comparison of the lipid intensity differences determined in our study and reported in studies excluding Leppik2020 and Kriisa2017. The X-axis depicts base two log-transformed fold change (log2 fold-change) between SCZ and CTL (averaged for the three sample cohorts), y-axis depicts log2 fold-changes between disease and CLT as reported in the published studies. Each point corresponds to a particular lipid species reported in the literature and found among SCZ-associated lipids. The points are colored by study as in A (left) or by lipid class (right). Spearman correlation coefficients and corresponding p-values are marked in the bottom left corner. The horizontal and vertical dashed lines indicate log2 fold-change = 0, the diagonal dashed line indicates the y-axis = x-axis line. (C): Comparison of the lipid intensity differences determined in our study and reported by Leppik2020 and Kriisa2017.



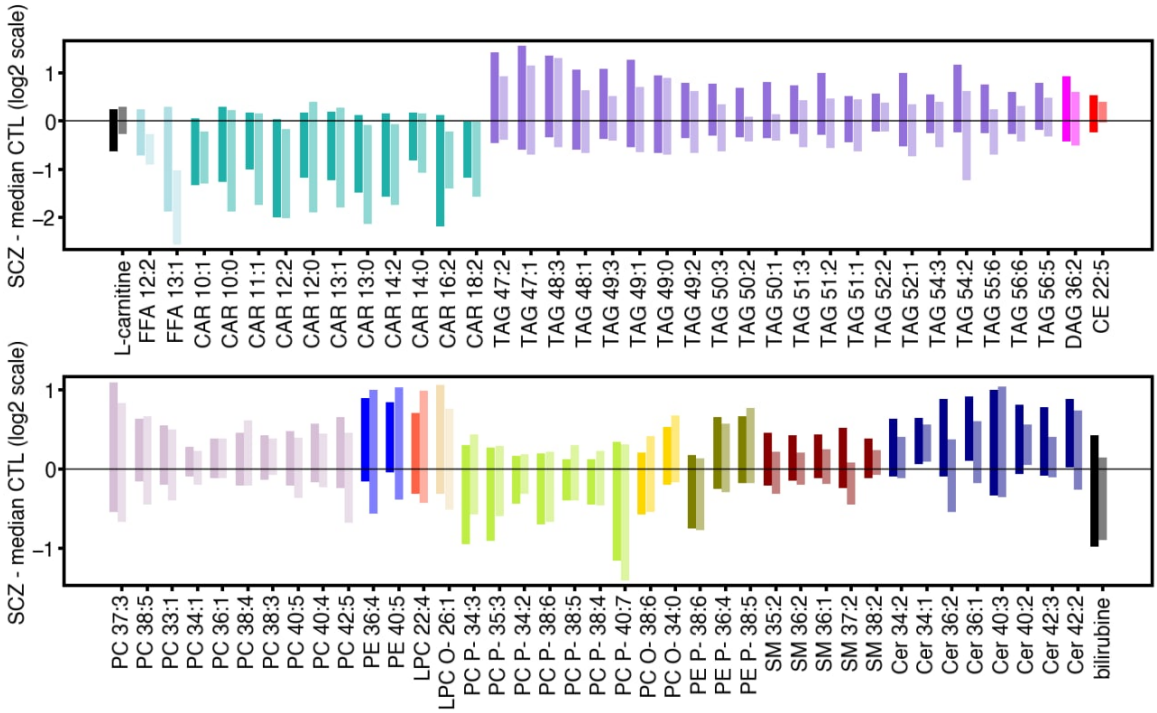
eFigure 12. Overlap of Genetic Variants Linked to Psychiatric Disorders and Blood Plasma Lipid Levels

Left: upset plot for genes associated with the three psychiatric disorders. Right: upset plot for genes associated with SCZ, BPD, and MDD, considering only the subset of 84 genes linked to the blood plasma lipid level variation. Colors correspond to the fold of enrichment, calculated for each group by subsampling random groups of genes of the same size 1,000 times.



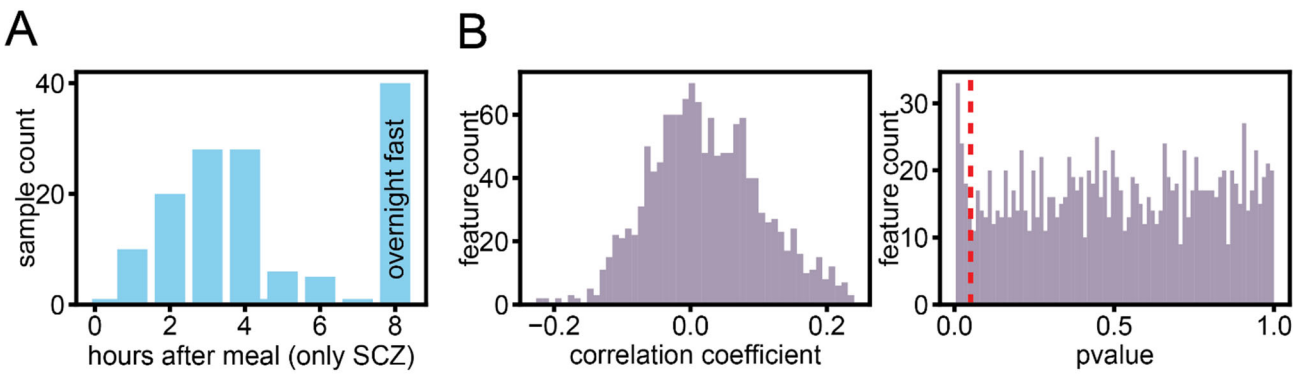
eFigure 13. Correlation Between Lipid Profiles of Nonmedicated Individuals and the SCZ Profile, Compared to the Correlation Between Random Lipid Profiles and the SCZ Profile

Distributions of Pearson correlation coefficients between SCZ lipid profile and lipid profiles of different subsamples (n = 100), were SCZ lipid profiles are defined as the base two log-transformed fold change (log2 FC) between SCZ and CTL in the DE-AT cohort. Grey histogram: distribution of correlation coefficients between SCZ lipid profile and the log2FC of 13 random CTL compared to all CTL samples. Blue histogram: distribution of correlation coefficients between SCZ lipid profile and the log2FC of 13 random SCZ compared to all CTL samples. Dashed red line: correlation coefficients between SCZ lipid profile and log2FC of 13 non-medicated SCZ compared to all CTL samples. On the left, all SCZ-associated lipids (n = 77) are considered, on the right, SCZ-associated lipids with the exclusion of TAGs (n = 56).



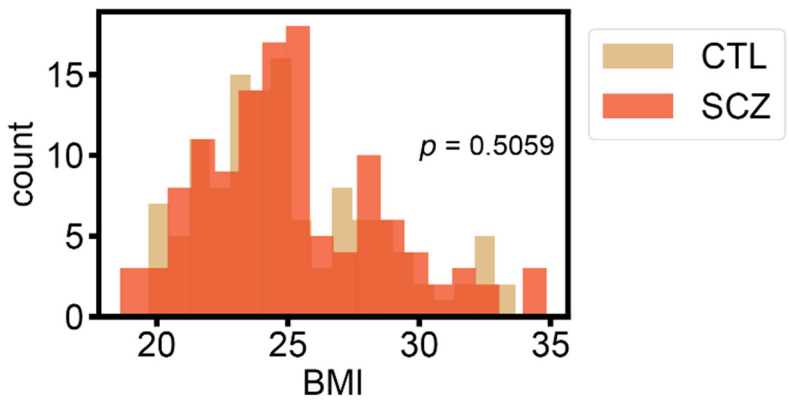
eFigure 14. Comparison of the Lipid Intensity Distributions in Medicated and Nonmedicated SCZ Samples

Distributions of the log2 lipid abundances in SCZ samples, centered by the median abundance in CTL, for all the SCZ-associated lipids. Each line corresponds to the inter-quartile range of the distribution. Darker shades correspond to the distributions of SCZ samples (DE-AT cohort, *n*=122), lighter shades to the 13 non-medicated SCZ samples. Colors delineate to lipid classes.

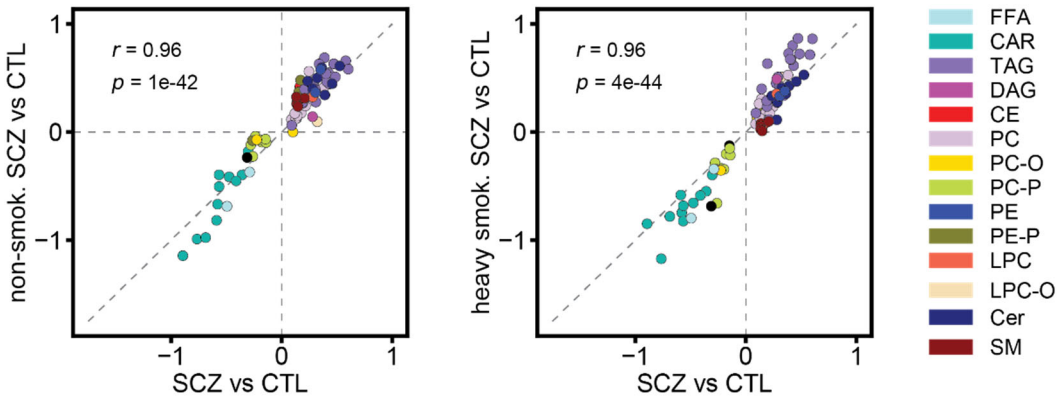


eFigure 15. Dependence of the Lipid Intensity Values on the Fasting Period Prior to Sample Collection

(A) Distribution of hours after the meal prior to blood sample collection in CN SCZ sample cohort. Overnight fast was imputed by the value = 8 hours. (B) Left: the Pearson correlation coefficients between lipid intensities and after the meal fasting period. Right: the corresponding p-values. The dashed red line shows nominal p-value cutoff = 0.05, chosen to define food-related lipids discarded from downstream analyses.



eFigure 16. The Distribution of the Body Mass Index (BMI) for Subsets of CTL and SCZ Samples Used in Analysis of DE-AT Cohort. The *P*-Value of the Mann–Whitney *U* Test is Marked Within the Plot



eFigure 17. Effect of Smoking on SCZ-Associated Lipid Intensity Differences

Comparison of SCZ-associated lipid intensity differences (log2-fold change) between non-smoking (left) or heavy smoking (reporting smoking at least a pack of cigarettes per day; right) individuals and lipid intensity differences defined using all samples. Pearson correlation coefficients and corresponding p-values are marked in the top left corner. Colors correspond to lipid classes, as indicated. The horizontal and vertical dashed lines indicate log2 fold-change = 0, the diagonal dashed line indicates the y-axis = x-axis line.

Supplementary Tables

eTable 1. Confounding Factors Assessment

	SCZ vs CTL		
	DE-AT	**RU**	**CN**
sex	substracted linear regression for sex	substracted linear regression for sex	substracted linear regression for sex
age	substracted linear regression for age	substracted linear regression for age	substracted linear regression for age
bmi	***subsampling** for balancing bmi	-	-
non-fasting vs fasting	discarded lipids affected by non-fasting status	discarded lipids affected by non-fasting status	discarded lipids affected by non-fasting status
smoking	comparison of heavy smokers and non smokers for SCZ-associated lipids	comparison of heavy smokers and non smokers for SCZ-associated lipids	comparison of heavy smokers and non smokers for SCZ-associated lipids

	BPD vs CTL	
	DE-AT	**RU**
sex	substracted linear regression for sex	substracted linear regression for sex + ***subsampling for balancing sex
age	substracted linear regression for age	substracted linear regression for age
bmi	**subsampling for balancing bmi	-
non-fasting vs fasting	discarded lipids affected by non-fasting status	discarded lipids affected by non-fasting status
smoking	-	-

	MDD vs CTL	
	RU	**CN**
sex	substracted linear regression for sex + ****subsampling for balancing sex	substracted linear regression for sex
age	substracted linear regression for age	substracted linear regression for age
bmi	-	-
non-fasting vs fasting	discarded lipids affected by non-fasting status	discarded lipids affected by non-fasting status
smoking	-	-

***Specific subsamples used in the analysis are included in the same table at https://drive.google.com/drive/folders/1qLhPD0Zmi3TkRjfFfpLrTJvOu0296xdF?usp=sharing

eTable 2. Summary of Studies Reporting Lipid Alterations in Psychiatric Disorders at the Level of Individual Lipid Species

Study		Disease	Medication status	Sample size	Lipid classes	Main findings	Comparable at lipid species level	Supports results of this study (Lipid class level)	Note
Cao 2019	(Cao et al., 2019)	SCZ	pre- and post-treatment vs control	SCZ=225, CTL=175	CAR	medium and high chain CAR decreased in pre-treatment SCZ compared to CTL, even lower post-treatment SCZ	yes	yes	
Wood 2015	(Wood et al., 2015)	SCZ	medicated	SCZ=23, CTL=27	LPE-P, LPC-P, PE-P, PC-P	PE-P, PC-P decreased in SCZ compared to CTL	yes	yes	
Kriisa 2017	(Kriisa et al., 2017)	SCZ	pre- and post-treatment vs control	SCZ=38, CTL=37	CAR	Long-chain CAR increased in FEP compared to CTL; short-chain CAR decreased in FEP compared to CTL; CAR decreased in FEP after medication	yes	no	
Leppik 2020	(Leppik et al., 2020)	SCZ	pre- and post-treatment vs control	SCZ=38, CTL=37	PC, PC-O, SM	PC, PC-O reduced in FEP compared to CTL; no difference medicated SCZ and CTL	yes	no	
Yan 2018	(Yan et al., 2018)	SCZ	pre-treatment vs control and pre- vs post-treatment	SCZ=20, CTL=29	different lipids	TAG, CE increased in FEP compared to CTL; LPC, PC-P, PE-P decreased in FEP compared to CTL; CAR 18:2 decreased in FEP compared to CTL	yes	yes	
Wang 2019	(Wang et al., 2019)	SCZ	medicated and non-medicated	SCZ=119 (28 FEP), CTL=109	PC, PE, LPC, LPE, SM, PC-O, PE-O	LPC, PC-O, PE-O decreased in SCZ compared to CTL; SM increased in SCZ compared to CTL; PC both increased and decreased depending on lipid species; no difference FEP vs recurrent SCZ	yes	partially (yes, except or PC lipid class)	
Wang 2020	(Wang et al., 2020)	SCZ	medicated and non-medicated	SCZ=119 (28 FEP), CTL=109	PC, PE, LPC, LPE, SM, PC-O, PE-O, FFA	similar to Wang 2019; FFA significantly altered	yes	partially (yes, except or PC lipid class)	
Cao 2020	(Cao et al., 2020)	SCZ	recently non-medicated (1 month)	SCZ=113, CTL=111	Car, LPC, LPC-O	LPC decreased in SCZ vs CTL; CAR significantly altered in SCZ	yes	Partially; most CAR species were increased in SCZ	

								(doesn't support our results); the CAR lipid species reported in this study and present among SCZ-associated were decreased (supports our results)	
He 2012	(He et al., 2012)	SCZ	medicated and non-medicated	SCZ=216, CTL=265	Only one lipid - PC-O 38:6	PC-O 38:6 decreased in SCZ compared to CTL	yes	yes	
Orešič 2012	(Orešič et al., 2012)	SCZ	medicated	SCZ=19, CTL=53; twins	LPC, PC, PE, PG, PS, PA, Cer, MAG, DAG, TAG	TAG increased in SCZ compared to CTL; LPC lipid species decreased in SCZ compared to CTL	no	not compared	twin study; difference between SCZ and CTL not included in table form
McEvoy 2013	(McEvoy et al., 2013)	SCZ	medicated and non-medicated	SCZ=20, FEP=20, CTL=29	PC, PE	decreased PUFA-containing PC and PE lipid species in SCZ compared to CTL	no	not compared	lipid species containing particular fatty acids, no bulk structure
Solberg 2016	(Solberg et al., 2016)	SCZ	medicated	SCZ=55	total trigylcerides, membrane PUFAs (blood cells, not blood plasma)	increased TAG in SCZ compared to CTL	no	not compared	no individual TAG species; esterified membrane (blood cells) PUFAs
Kaddurah-Daouk 2012	(Kaddurah-Daouk et al., 2012)	SCZ	recently non-medicated and drug naïve	SCZ=20+20 (naive and recently unmedicated), CTL=17	PC-P, PE-P	PC-P, PE-P decreased in both drug-naïve and medicated SCZ compared to CTL;	no	yes	lipid species containing particular fatty acids, no bulk structure
Gracia-Garcia 2011	(Gracia-Garcia et al., 2011)	MDD	medicated	MDD=23, CTL=23	Cer	Cer elevated in MDD compared to CTL	yes	yes	
Liu 2016	(Liu et al., 2016)	MDD	medicated and non-medicated	MDD=60, CTL=59+MDD=75 and CTL=52	different lipids	LPC, LPE, PC, PE, PI, TG increased with increase of symptome severity; PE-O decreased with increase of symptome severity	yes (only direction of effect)	yes	correlation to symptom severity (HAMD)
Liu 2015	(Liu et al., 2015)	MDD	medicated and non-medicated	MDD=60, CTL=59+MDD=75 and CTL=52	different lipids	LPC, LPE, PC, PE, PI, TG increased in MDD compared to CTL; PE-O, CAR decreased in MDD compared to CTL	yes	yes	
Demirkan 2013	(Demirkan et al., 2013)	MDD	medicated	samples=742,	different lipids	SM and PC-O significantly associated with	yes (only direction of effect)	partially (yes, except or	correlation to symptom

				family study		symptome severity		PC lipid class)	severity (HADS-D)
Kim 2018	(Kim et al., 2018)	MDD	recently non-medicated (8 weeks)	MDD=20, CTL=10 + 25*3 validation	different lipids	TAG levels significantly altered in MDD compared to CTL; Results not reproduced in validation	yes	not compared	results non-reproduced in validation; study not used for comparison
Cai 2019	(Cai et al., 2019)	MDD	medicated	MDD=30, CTL=30	PE, LPC	PE decreased in MDD compared to CTL; LPC increased in MDD compared to CTL	yes	no	none of the reported lipids among the SCZ-associated
Otoki 2017	(Otoki et al., 2017)	MDD	non-medicated	MDD=9	different lipids	Significant alterations between MDD and CTL	no	not compared	lipid species containing particular fatty acids, no bulk structure
Knowles 2017	(Knowles et al., 2017)	MDD	medicated	samples= 567, family study	Cer, SM, PC, PC-O, PC-P, LPC, LPC-O, PE, CE, DAG, TAG, phosphatidylinositol (PI), phosphatidylserine (PS), phosphatidylglycerol (PG), hexosylceramides (monohexosylceramide, dihexosylceramide, trihexosylcermide), GM3 gangliosi de	heritability: PC-P and PC-O exhibited the greatest degree of genetic overlap with MDD.	no	not compared	heritability study
Brunkhorst-Kanaan 2019	(Brunkhorst-Kanaan et al., 2019)	BPD, MDD	medicated	MDD+BPD=67, CTL=405	Cer, HexCer	Cer and HexCer elevated in MDD+BPD compared to CTL	yes	yes	
Scola 2018	(Scola et al., 2018)	BPD, MDD	medicated	BPD=87, MDD=34, CTL=31	FFA	ratio of FFA significantly altered in disease vs CTL	no	not compared	most FFA discarded in our analysis due to non-fasting status
Saunders 2015	(Saunders et al., 2015)	BPD	medicated and non-medicated	BPD=27, CTL=31	FFA	Eicosapentaenoic acid decreased in BPD vs CTL	no	not compared	most FFA discarded in our analysis due to non-fasting status
Pomponi 2013	(Pomponi et al., 2013)	BPD	medicated	BPD=42, CTL=57	FFA	FFA altered in BPD compared to CTL	no	not compared	most FFA discarded in our analysis

									due to non-fasting status
Ribeiro 2017	(Ribeiro et al., 2017)	BPD	medicated	BPD=14, CTL=21	different lipids	Significant alterations between BPD and CTL	no	not compared	annotation non-comparable, no bulk structure

Brunkhorst-Kanaan, N. et al. (2019) 'Targeted lipidomics reveal derangement of ceramides in major depression and bipolar disorder', Metabolism: Clinical and Experimental. doi: 10.1016/j.metabol.2019.04.002.

Cai, H. L. et al. (2019) 'Quantitative monitoring of a panel of stress-induced biomarkers in human plasma by liquid chromatography–tandem mass spectrometry: an application in a comparative study between depressive patients and healthy subjects', Analytical and Bioanalytical Chemistry. doi: 10.1007/s00216-019-01956-2.

Cao, B. et al. (2019) 'Characterizing acyl-carnitine biosignatures for schizophrenia: a longitudinal pre- and post-treatment study', Translational Psychiatry. doi: 10.1038/s41398-018-0353-x.

Cao, B. et al. (2020) 'Metabolic profiling for water-soluble metabolites in patients with schizophrenia and healthy controls in a Chinese population: A case-control study', World Journal of Biological Psychiatry. doi: 10.1080/15622975.2019.1615639.

Demirkan, A. et al. (2013) 'Plasma phosphatidylcholine and sphingomyelin concentrations are associated with depression and anxiety symptoms in a Dutch family-based lipidomics study', Journal of Psychiatric Research. doi: 10.1016/j.jpsychires.2012.11.001.

Gracia-Garcia, P. et al. (2011) 'Elevated plasma ceramides in depression', Journal of Neuropsychiatry and Clinical Neurosciences. doi: 10.1176/jnp.23.2.jnp215.

He, Y. et al. (2012) 'Schizophrenia shows a unique metabolomics signature in plasma', Translational Psychiatry. doi: 10.1038/tp.2012.76.

Kaddurah-Daouk, R. et al. (2012) 'Impaired plasmalogens in patients with schizophrenia', Psychiatry Research. doi: 10.1016/j.psychres.2012.02.019.

Kim, E. Y. et al. (2018) 'Serum lipidomic analysis for the discovery of biomarkers for major depressive disorder in drug-free patients', Psychiatry Research. doi: 10.1016/j.psychres.2018.04.029.

Knowles, E. E. M. et al. (2017) 'The lipidome in major depressive disorder: Shared genetic influence for ether-phosphatidylcholines, a plasma-based phenotype related to inflammation, and disease risk', European Psychiatry. doi: 10.1016/j.eurpsy.2017.02.479.

Kriisa, K. et al. (2017) 'Profiling of acylcarnitines in first episode psychosis before and after antipsychotic treatment', Journal of Proteome Research. doi: 10.1021/acs.jproteome.7b00279.

Leppik, L. et al. (2020) 'Profiling of lipidomics before and after antipsychotic treatment in first-episode psychosis', European Archives of Psychiatry and Clinical Neuroscience. doi: 10.1007/s00406-018-0971-6.

Liu, X. et al. (2015) 'Discovery and validation of plasma biomarkers for major depressive disorder classification based on liquid chromatography-mass spectrometry', Journal of Proteome Research. doi: 10.1021/acs.jproteome.5b00144.

Liu, X. et al. (2016) 'Plasma lipidomics reveals potential lipid markers of major depressive disorder', Analytical and Bioanalytical Chemistry. doi: 10.1007/s00216-016-9768-5.

McEvoy, J. et al. (2013) 'Lipidomics Reveals Early Metabolic Changes in Subjects with Schizophrenia: Effects of Atypical Antipsychotics', PLoS ONE. doi: 10.1371/journal.pone.0068717.

Orešič, M. et al. (2012) 'Phospholipids and insulin resistance in psychosis: A lipidomics study of twin pairs discordant for schizophrenia', Genome Medicine. doi: 10.1186/gm300.

Otoki, Y. et al. (2017) 'Plasma Phosphatidylethanolamine and Triacylglycerol Fatty Acid Concentrations are Altered in Major Depressive Disorder Patients with Seasonal Pattern', Lipids. doi: 10.1007/s11745-017-4254-1.

Pomponi, M. et al. (2013) 'Plasma levels of n-3 fatty acids in bipolar patients: Deficit restricted to DHA', Journal of Psychiatric Research. doi: 10.1016/j.jpsychires.2012.11.004.

Ribeiro, H. C. et al. (2017) 'A preliminary study of bipolar disorder type I by mass spectrometry-based serum lipidomics', Psychiatry Research. doi: 10.1016/j.psychres.2017.08.039.

Saunders, E. F. H. et al. (2015) ‘Low unesterified: Esterified eicosapentaenoic acid (EPA) plasma concentration ratio is associated with bipolar disorder episodes, and omega-3 plasma concentrations are altered by treatment’, Bipolar Disorders. doi: 10.1111/bdi.12337.

Scola, G. et al. (2018) ‘Alterations in peripheral fatty acid composition in bipolar and unipolar depression’, Journal of Affective Disorders. doi: 10.1016/j.jad.2017.12.025.

Solberg, D. K. et al. (2016) ‘Lipid profiles in schizophrenia associated with clinical traits: A five year follow-up study’, BMC Psychiatry. doi: 10.1186/s12888-016-1006-3.

Wang, D. et al. (2019) ‘Metabolic profiling identifies phospholipids as potential serum biomarkers for schizophrenia’, Psychiatry Research. doi: 10.1016/j.psychres.2018.12.008.

Wang, D. et al. (2020) ‘Characterising phospholipids and free fatty acids in patients with schizophrenia: A case-control study’, World Journal of Biological Psychiatry. doi: 10.1080/15622975.2020.1769188.

Wood, P. L. et al. (2015) ‘Dysfunctional plasmalogen dynamics in the plasma and platelets of patients with schizophrenia’, Schizophrenia Research. doi: 10.1016/j.schres.2014.11.032.

Yan, L. et al. (2018) ‘Unbiased lipidomic profiling reveals metabolomic changes during the onset and antipsychotics treatment of schizophrenia disease’, Metabolomics. doi: 10.1007/s11306-018-1375-3.

eTable 7. SCZ-Associated Lipid Classes Analysis

	DE-AT one-sided binomial test pvalue	CN one-sided binomial test pvalue	RU one-sided binomial test pvalue	SCZ-associated lipid classes, binomial test analysis (SCZ-associated lipids)
TAG	**4.8E-07**	**4.8E-07**	**4.8E-07**	yes
CAR	**0.00049**	**0.00049**	**0.00049**	yes
FFA	0.25	0.25	0.25	
CE	0.5	0.5	0.5	
DAG	0.5	0.5	0.5	
LPC	0.5	0.5	0.5	
LPC-O	0.5	0.5	0.5	
LPE	-	-	-	
PC	**0.00098**	**0.01074**	**0.00098**	yes
PC-O	0.75	0.25	0.75	
PC-P	**0.00781**	**0.00781**	**0.00781**	yes
PE	0.25	0.25	0.25	
PE-P	0.5	0.125	0.125	
Cer	**0.00391**	**0.03516**	**0.00391**	yes
SM	0.03125	0.1875	0.03125	

	DE-AT			CN			RU			SCZ-associated lipid classes, GSEA analysis (all annotated lipids)
	GSEA pvalue	GSEA adjusted pvalue	GSEA NES*	GSEA pvalue	GSEA adjusted	GSEA NES*	GSEA pvalue	GSEA adjusted	GSEA NES*	
TAG	1E-05	**8.9E-05**	2.04966	2.97E-05	**0.0001**	2.7398	3.23E-05	**7E-05**	2.47	yes
CAR	5.9E-05	**0.0003**	-3.4778	1.73E-05	**9E-05**	-2.463	1.72E-05	**6E-05**	-2.178	yes
FFA	0.44426	0.51261	-1.0025	1.64E-05	9E-05	-2.505	1.62E-05	6E-05	-2.391	
CE	0.99978	0.99978	0.38861	0.863426	0.8634	0.6732	0.444792	0.5132	1.024	
DAG	0.31335	0.45406	1.12734	0.094184	0.1413	1.4352	0.590731	0.6329	0.897	
LPC	0.58264	0.62426	0.92943	0.309015	0.3566	1.1067	1.71E-05	6E-05	-2.349	
LPC-O	0.04412	0.09453	1.46895	0.261255	0.3266	1.1899	0.012252	0.018	-1.692	
LPE	0.41487	0.51261	1.05093	0.001525	0.0045	1.9302	0.737791	0.7378	-0.787	
PC	0.01484	0.03711	1.52013	0.067974	0.1133	1.3518	0.0144	0.018	1.573	
PC-O	0.06099	0.11435	1.39942	0.004907	0.0092	-1.731	0.001682	0.0028	-1.856	
PC-P	0.00016	**0.00058**	-2.275	0.001817	**0.0045**	-1.842	0.000263	**0.0005**	-2.049	yes
PE	0.01243	**0.03711**	1.60966	0.00347	**0.0074**	1.8695	0.013855	**0.018**	1.694	yes
PE-P	0.27608	0.45406	1.15281	1.67E-05	9E-05	-2.402	1.64E-05	6E-05	-2.246	
Cer	1.2E-05	8.9E-05	2.49725	0.151648	0.2068	1.2753	2.44E-05	7E-05	2.541	
SM	0.33297	0.45406	1.10432	0.82516	0.8634	-0.767	2.65E-05	7E-05	2.349	

*NES: Normalized enrichment score

eTable 8. Intra- and Inter- Cohort Performance of the SCZ vs CTL Classifier

n = maximum possible in cohort		test cohort		
		CN	**DE-AT**	**RU**
train cohort	**CN (n=138)**	0.87 + - 0.06	0.72 + - 0.09	0.82 + - 0.07
	DE-AT (n=169)	0.82 + - 0.08	0.85 + - 0.07	0.91 + - 0.05
	RU (n=67)	0.68 + - 0.1	0.66 + - 0.1	0.97 + - 0.03

n = 67		test cohort		
		CN	**DE-AT**	**RU**
train cohort	**CN (n=67)**	0.87 + - 0.07	0.69 + - 0.09	0.81 + - 0.08
	DE-AT (n=67)	0.76 + - 0.09	0.8 + - 0.07	0.83 + - 0.09
	RU (n=67)	0.69 + - 0.1	0.68 + - 0.09	0.96 + - 0.03

n - number of SCZ samples and CTL samples used in training
columns: test cohort, rows: training cohort

eTable 9. Enrichment of Disease-Associated Genes Among Lipid Enzymes

	SCZ-associated genes		MDD-associated genes		BPD-associated genes		SCZ-ass. LC1	SCZ-ass. LC2
	perm. pv	perm. ratio	perm. pv	perm. ratio	perm. pv	perm. ratio		
CAR	0.1148	1.15542	0.4028	1.05955	0.7273	0.93918	yes	yes
TAG	0.819	0.92584	0.751	0.92199	0.9502	0.85522	yes	yes
PC-P	0.8455	0.83413	0.7673	0.83829	0.941	0.7322	yes	yes
PC	**0.0308**	1.11252	**0.0453**	1.15231	0.2405	1.04469	yes	
PE	**0.0236**	1.14595	**0.0226**	1.21798	0.2658	1.0489		yes
Cer	0.0887	1.12345	**0.0247**	1.26079	0.3574	1.03822	yes	
DAG	**0.0096**	1.05453	**0.0021**	1.09833	**0.0149**	1.05119		
FFA	0.8965	0.93312	0.8925	0.9062	0.9561	0.90526		
CE	0.7072	0.95444	0.8272	0.87654	0.8464	0.9056		
LPC	0.2512	1.073	0.5328	1.00093	0.883	0.8929		
LPE	0.2275	1.1106	0.6283	0.95622	0.975	0.76054		
PE-P	0.8511	0.83271	0.7742	0.83828	0.9384	0.73346		
PC-O	0.8498	0.83094	0.7779	0.83567	0.9427	0.73194		
LPC-O	0.8498	0.83505	0.7745	0.83882	0.9403	0.73264		
LPC-P	0.8498	0.83249	0.7647	0.84053	0.9396	0.73157		
SM	0.052	1.18177	**0.0179**	1.35234	0.2211	1.09085		

*

perm. pv: Permutation enrichment pvalue

perm. ratio: Permutation enrichment ratio

SCZ-ass. LC1: SCZ-associated lipid classes, binomial test analysis (only SCZ-associated

SCZ-ass. LC2: SCZ-associated lipid classes, GSEA analysis (all annotated lipids)

eAppendix

XCMS parameters

The following XCMS parameters and methods were used for data pre-processing:

xset <- xcmsSet (method = "centWave", peakwidth = c(10, 30), ppm = 15, noise = 50, snthresh = 3, mzdiff = 0.003, prefilter = c(3, 100), mzCenterFun = "wMean", integrate = 1, fitgauss = FALSE, verbose.columns = FALSE)

xset <- retcor (xset, method = "obiwarp", plottype = "none", distFunc = "cor_opt", profStep = 0.6352, center = 2, response = 1, gapInit = 0.448, gapExtend = 2.4, factorDiag = 2, factorGap = 1, localAlignment = 0)

xset <- group (xset, method = "density", bw = 3, mzwid = 0.008, minfrac = 0.05, minsamp = 10, max = 50)

xset <- fillPeaks(xset)
groups (xset,method="maxint",value="into")

Duplicated peaks

XCMS peak picking parameters that are not ideal for a given peak shape can result in one lipid signal being split into two peaks in the resulting feature table, identifiable as two highly correlated features with almost identical m/z values and retention time values. We dealt with this issue by grouping all features with pairwise difference in m/z values < 10 ppm and retention time difference < 0.05 minutes, keeping only one feature from each group.

Co-eluting peak clusters

Untargeted LCMS lipidomics experiments can produce large clusters of lipid features that are derived from the same lipid compound, identifiable as a large number of well-correlated lipid features with identical retention times. We dealt with this issue by deleting larger-than-average groups of lipid features that fall into a small retention time window as follows. First, we defined a window of 0.005 minutes and calculated the number of lipid features with retention times falling into this window, depending on the position of the window across the retention time gradient ("sliding histogram"). Next, we used kernel density estimation (python package KernelDensity, kernel='gaussian', bandwidth=0.5; this calculates the logarithm of the density) on the sliding histogram values, and calculated the smoothed (or averaged) sliding histogram by taking the

exponent of the kernel density estimation and normalizing it by the area under the sliding histogram. We defined a first cutoff function as this smoothed histogram. Next, we defined a spike in the sliding histogram as a group of adjacent values of the sliding histogram that are larger than the cutoff function. Each spike corresponds to a retention time boundary, and a set of lipid features that fall into this retention time boundary. The height of the spike is the maximum value of the sliding histogram in this retention time boundary. We aimed to delete abnormal spikes that correspond to co-eluting peak clusters. Because we expected co-eluting peak clusters to be a derivative from one lipid compound (which will have the highest intensity), we expected most peaks in a co-eluting cluster to have m/z values higher than this original lipid compound. However, this might not be true for all co-eluting peak clusters. For this reason, we defined two types of abnormal spikes. The first type of abnormal spikes was defined as spikes with height > 40. For the second type of abnormal spikes, we defined a second cutoff function as 4*smoothed histogram +10. We defined as abnormal any spike for which (1.) the height was larger than this cutoff function, and for which (2.) the percent of lipid features in this spike with m/z values > m/z value of the lipid feature with highest intensity in this spike was larger than 75%. We deleted from downstream analysis all lipid features that corresponded to an abnormal spike.

Isotope deletion

Isotopes were grouped based on differences in m/z and retention times of the lipid features. Additionally, we used pairwise correlations across samples, because we expect variability in isotope intensities to be explained only by unbiased random error in measurements (any biological variability or technical biases will be identically reflected in isotopes), and we expect these correlations to be quite high.

First, lipid features were separated into isotope pre-groups. Differences between m/z values (m/z_diff) and differences in retention times (rt_diff) were calculated for each pair of lipid features, and features with m/z_diff<0.01 and rt_diff< 10 seconds were considered to be connected. Next, we located all the connected components, each connected component defining an isotope pre-group. Because two features in a pre-group might be connected through a chain of connected pairs, the resulting m/z_diff and rt_diff could be much higher than the above-defined thresholds, and one pre-group could contain multiple isotope groups.

Next, for each pre-group we defined the isotope groups. Starting from lipid feature X with highest intensity in the pre-group, we identified the lipid features Y satisfying the following criteria: 1000000*(m/z of X - m/z of Y + C13 - C12)/ (m/z of X) < 10, difference in retention time < 1.5 seconds, correlation across samples < 0.8, and intensity of Y < intensity of X. If exactly one feature was found, this feature was added to the isotope group and the procedure was continued recursively. If no feature or multiple features were found, the procedure was stopped and the lipid features Y were not added to the isotope group. In all cases, lipid features Y were deleted from the pre-group, as well as feature X. Once the procedure stopped, if there were lipid features remaining in the pre-group, the algorithm was repeated recursively, starting with the remaining lipid feature with highest intensity.

This algorithm defined isotope groups. For each isotope group, only the highest intensity feature was retained in the data table, and all other lipid features (their higher level isotopes) were deleted from analysis.

Matching features from different datasets

After each of the three main dataset (CN, DE-AT+bisDE-AT, RU) was processed independently, lipids were matched between datasets by m/z values and retention time. Only features reproduced in all three cohorts were used for downstream statistical analysis. First, retention times were aligned between datasets using linear interpolation between internal standards (PC, PE, PG, PI, LPC, LPE, MAG, DAG, SM, and TAG). Next, lipids detected in each pair of datasets were matched by m/z according to 10 ppm threshold and by retention time values according to a 0.2 minute threshold. After this matching, if there were non-unique feature matches, only one of them was kept according to the following rule. If there were multiple features from one cohort that were matched to the same feature of another cohort with m/z difference less than 5 ppm, or if the differences in their ppm precision was less than 3 ppm, then the feature with smallest retention time difference was retained in the matching. Otherwise, the feature with smallest m/z difference was retained in the matching. Next, a consolidated table matching lipids from all three CN, DE-AT and RU dataset was constructed. If feature1 from cohort1 was matched to feature2 from cohort2, and feature2 was matched to feature 3 from cohort3, then feature1 from cohort1 had to match to feature3 from cohort3 for the features 1, 2 and 3 to be included in the final consolidated table as one single matched lipid feature. Additionally, the fepRU dataset was matched to the RU

dataset in the same manner, and the RU features in this consolidated table were used to make a table matching all four CN, RU, DE-AT and fepRU datasets.

Putative annotation

We used a semi-automated procedure based on m/z values and retention times for the putative annotation. For each considered lipid class, we generated a theoretical table of neutral masses corresponding to the lipid species with the various number of double bonds and chain lengths. Each lipid class was annotated in one ionization mode using one adduct of choice (eTable 2).

Starting from a lipid class and its corresponding adduct, we matched the theoretical m/z to the m/z values of the lipid features with a mass accuracy threshold of 10 ppm. Next, a model for the retention times of the lipid species was built, and only lipid species fitting into this model were retained in the putative annotation, according to the following procedure. We started with an in-house database of retention times of verified lipid species annotation, and retention time alignment between the current retention times and the database retention times based on linear interpolation between retention times of the internal standards. The retention times of the in-house database were matched to the current retention times to find which lipid species from the database were present in the current data. One of these matched lipid species was used to define an "anchor point" for the rest of the lipid species in the lipid class. The model for the retention time was defined by three main parameters: rt_delta1, rt_delta2, the anchor point, as well as a transformation retention time. First, the retention times were transformed according to this function. Different transformations were used for different lipid classes, specifically, base 2 log transformation for all lipids except triglycerides, and base 2 exponential transformation for triglycerides, because the corresponding elution region had a steeper slope. Parameter rt_delta1 corresponded to the shift in transformed retention time when one double bond was added to the lipid species, and rt_delta2 corresponded to the shift in transformed retention time when one carbon was added to the lipid species. Starting from the anchor point, this model defined the predicted retention time for any lipid species of the lipid class. Defining additional parameter rt_error, we used this model to annotate the lipid features by keeping the lipid features with retention times within +- rt_error boundary of the predicted retention times, and discarding the rest.

Next, the model was trained on the data to find optimal parameters rt_delta1 and rt_delta2. To do this, we defined the score function as the number of annotated lipid features. We performed a

simple grid search over an array of rt_delta1 and rt_delta2 to find the optimal parameters. Using these parameters, we annotated the lipid features as described above. This procedure was data-based, and the parameters were not necessarily ideal. For this reason, the resulting putative annotation based on m/z values and retention times were additionally manually curated.

For PC-O and PC-P lipid classes, as well as PE-O and PE-P classes, a modification of the procedure was used. Because these lipid classes have identical m/z values and similar retention times, the above described algorithm could not separate them. In the modified procedure, we similarly started with an anchor point for the -O lipid class and an anchor point for the -P lipid class, but used the optimal parameters rt_delta1 and rt_delta2 calculated for the corresponding PC or PE class.

eReferences

1. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull. 1987. 1987. https://doi.org/10.1093/schbul/13.2.261.
2. Budde M, Anderson-Schmidt H, Gade K, Reich-Erkelenz D, Adorjan K, Kalman JL, et al. A longitudinal approach to biological psychiatric research: The PsyCourse study. Am J Med Genet Part B Neuropsychiatr Genet. 2019. 2019. https://doi.org/10.1002/ajmg.b.32639.
3. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem. 2006. 2006. https://doi.org/10.1021/ac051437y.
4. Liebisch G, Vizcaíno JA, Köfeler H, Trötzmüller M, Griffiths WJ, Schmitz G, et al. Shorthand notation for lipid structures derived from mass spectrometry. J Lipid Res. 2013. 2013. https://doi.org/10.1194/jlr.M033506.
5. Sergushichev AA. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. BioRxiv. 2016. 2016. https://doi.org/10.1101/060012.
6. Liu X, Li J, Zheng P, Zhao X, Zhou C, Hu C, et al. Plasma lipidomics reveals potential lipid markers of major depressive disorder. Anal Bioanal Chem. 2016. 2016. https://doi.org/10.1007/s00216-016-9768-5.
7. Demirkan A, Isaacs A, Ugocsai P, Liebisch G, Struchalin M, Rudan I, et al. Plasma phosphatidylcholine and sphingomyelin concentrations are associated with depression and anxiety symptoms in a Dutch family-based lipidomics study. J Psychiatr Res. 2013. 2013. https://doi.org/10.1016/j.jpsychires.2012.11.001.
8. Eicher JD, Landowski C, Stackhouse B, Sloan A, Chen W, Jensen N, et al. GRASP v2.0: An update on the Genome-Wide Repository of Associations between SNPs and Phenotypes. Nucleic Acids Res. 2015. 2015. https://doi.org/10.1093/nar/gku1202.
9. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes - Release 72.1, December 1, 2014 . Nucleic Acids Res. 2000. 2000.
10. Sud M, Fahy E, Cotter D, Brown A, Dennis EA, Glass CK, et al. LMSD: LIPID MAPS structure database. Nucleic Acids Res. 2007;35:D527–D532.
11. Tabassum R, Rämö JT, Ripatti P, Koskela JT, Kurki M, Karjalainen J, et al. Genetic

architecture of human plasma lipidome and its link to cardiovascular disease. Nat Commun. 2019. 2019. https://doi.org/10.1038/s41467-019-11954-8.