Title: Interplay of Human Metabolome and Gut Microbiome in Major Depression

Najaf Amin[1*], Jun Liu[1*], Bruno Bonnechere[1,2], Siamak MehmoudianDehkordi[3], Matthias Arnold[3,4], Richa Batra[5], Yu-Jie Chiou,[1,6] Marco Fernandes[7], M. Arfan Ikram[8], Robert Kraaij[9], Jan Krumsiek[5], Danielle Newby[7], Kwangsik Nho[10], Djawad Radjabzadeh[9], Andrew J Saykin[10], Liu Shi[7], William Sproviero[7], Laura Winchester[7], Yang Yang,[1,11] Alejo J Nevado-Holgado[7], Gabi Kastenmüller[4], Rima F Kaddurah-Daouk[3*], Cornelia M van Duijn[1*]

* Contributed equally

1. Nuffield Department of Population Health, University of Oxford, Oxford, UK

2. REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, UHasselt, Belgium

3. Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA.

4. Institute of Computational Biology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

5. Department of Physiology and Biophysics, Institute for Computational Biomedicine, Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY 10021, USA

6. Department of Psychiatry, Kaohsiung Chang Gung Memorial Hospital, Chang Gung University College of Medicine, Kaohsiung, Taiwan

7. Department of Psychiatry, University of Oxford, Oxford, UK

8. Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands

9. Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands

10. Center for Neuroimaging, Department of Radiology and Imaging Sciences and Indiana Alzheimer's Disease Research Center, Indiana University School of Medicine, Indianapolis, IN, USA

11. Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Keywords: Major depression, metabolome, Krebs/TCA cycle, gut microbiome, lipids

1

Abstract:

The pathogenesis of depression is complex involving the interplay of genetic and environmental risk factors including diet, lifestyle and the gut microbiome. Metabolomics studies may shed light on the interplay of these factors. We study over 63,000 individuals including 8462 cases with a lifetime major depression and 5403 cases with recurrent major depression from the UK Biobank profiled for nuclear magnetic resonance (NMR) spectroscopy based metabolites with the Nightingale platform. We identify 124 metabolites that are associated with major depressive disorder (MDD), including 49 novel associations. No differences were seen between the metabolic profiles of lifetime and recurrent MDD. We find that metabolites involved in the tricarboxylic acid (TCA) cycle are significantly altered in patients with MDD. Integrating the metabolic signatures of major depression and the gut microbiome, we find that the gut microbiome might play an important role in the relationship between these metabolites, lipoproteins in particular, and MDD. The order *Clostridiales*, and the phyla *Proteobacteria* and *Bacteroidetes* were the most important taxa, which link the lipoprotein particles to MDD. Our study shows that at the molecular level energy metabolism is disturbed in patients with MDD and that the interplay between the gut microbiome and blood metabolome may play a key role in the pathogenesis of MDD.

Introduction

Major depression is an important determinant of population health, affecting people across the life span from adolescence to old age. The disease is associated with a plethora of debilitating symptoms beyond emotional dysregulation[1], spanning from cognition, motoric function, neurovegetative symptoms to inflammation and disturbances of the immune system. MDD is further linked to increased risks of cardiometabolic disorders and mortality[2]. Currently, most antidepressant therapies modulate the monoamine pathway, but evidence is increasing for a more complex interplay of multiple pathways involving a wide range of metabolic alterations spanning energy metabolism[3] lipid metabolism[4] and inflammation[5].

There is increased interest in energy metabolism and mitochondrial dysregulation as a contributor to the pathogenesis of major depression[6,7]. The brain has high aerobic activity requiring 20 times more energy than the rest of the body[8] and is vulnerable to impaired energy production[6]. A decreased brain energy production has been found in depressed patients [9]. An excess of structural mutations in the mitochondrial DNA have been reported in individuals diagnosed with depression compared to controls[6]. There is a high prevalence (54%) of depression in patients with mitochondrial diseases[10]. Symptoms of depression such as loss of energy and appetite, tiredness, weakness, cognitive impairments, and sleep disturbance are also frequent in mitochondrial diseases[9,11] Mitochondria are cell organelles that provide energy in the form of adenosine triphosphate via oxidative phosphorylation. Simultaneously, mitochondria are also responsible for generating reactive oxygen species (ROS) and anti-oxidants such as creatine, coenzyme Q10, niconitamide and glutathione, which protect cells from various deleterious effects of ROS[12]. The imbalance in the generation of ROS and antioxidants leads to oxidative stress, damages to lipids and proteins, inflammation and apoptosis [13,14].

Disturbed plasma lipid concentrations have been implicated in the development of mitochondrial dysfunction[13], with high density lipoprotein (HDL) inversely correlating with mitochondrial DNA damage[15]. A recent study using Nightingale's proton nuclear magnetic resonance (NMR) metabolomics platform in nine Dutch cohorts (5,283 patients with depression and 10,145 controls) showed a shift towards decreased levels of high density

lipoprotein (HDL)-cholesterol and increased levels of very low density lipoprotein (VLDL)-cholesterol and tri- and diglycerides particles in patients with depression [16]. The relevance and molecular mechanisms underlying this shift are not understood. Gut microbiome has been shown to be a major determinant of the circulating lipids, specifically triglycerides and HDL[17] and is also known to regulate mitochondrial function through the production of microbial metabolites including short chain fatty acids (SCFA), lipids, vitamins and amino acids.[18] Disruptions to the gut microbiome have been found in patients with major depressive disorder[19,20]. Metabolic signatures of the gut microbiome can be found in feces and blood[17,21,22]. In our recent study connecting the gut microbiome with the Nightingale's metabolites[21] we observed association of metabolites in plasma with 32 gut microbial groups. A higher abundance of family *Christensenellaceae*, genera *ChristensenellaceaeR7 group*, *Ruminococcaceae* (*UCG002, UCG003, UCG005, UCG010, UCG014*), *Coprococcus*, *RuminococcaceaeNK4A214group* and *Ruminiclostridium6* associated with a favourable lipid profile, i.e., decreased VLDLs and increased HDLs. Interestingly, decreased abundance of the same groups of gut microbiota we also find associated with higher scores on depressive symptoms in our study of gut microbiome and depression[20]. This raises the questions 1) whether the gut microbiome explains part of the shift in VLDL and HDL levels seen in patients with depression [16] and 2) can we use the metabolic signatures of the disease based on Nightingale's metabolites as a tool to infer the association between gut microbiome and the disease.

In this study, we harness the power of the UK biobank (UKB) to study over 63000 individuals including 8462 patients with MDD and 5403 patients with recurrent MDD, who are profiled for NMR spectroscopy based metabolites with the Nightingale platform. The largest study conducted to date discovers dysregulation of metabolites involved in the mitochondrial functioning in patients with MDD. When integrating the data with those of the Rotterdam Study to understand the interplay between the blood metabolome, gut microbiome and MDD, we find evidence that the altered gut microbiome in patients is a key player in the shift of the VLDL/HDL axis in MDD patients.

Methods

Study design and participants

The study was performed in the UK Biobank dataset, which comprises of more than 500,000 participants aged from 37 to 73 years during recruitment (2006 to 2010) for whom blood sampling was performed[23]. A random subset of 118,466 individuals was profiled for metabolites using a high-throughput ^{1}H-NMR metabolomics (Nightingale Health, Helsinki, Finland) platform. The participants were registered with the UK National Health service and from 22 assessment centres across England, Wales, and Scotland using standardised procedures for data collection which included a wide range of questionnaires, anthropological measurement, clinical biomarkers, genotype data, etc. All participants provided electronically signed informed consent. UK Biobank has approval from the Northwest Multi-centre Research Ethics Committee, the Patient Information Advisory Group, and the Community Health Index Advisory Group. Further detail on the rationale, study design, survey methods, data collection are available elsewhere[23]. The current study is a part of UK Biobank projects 30418 and 54520.

Definition of traits

For the initial analyses we considered two phenotypes including 1) lifetime major depressive disorder (MDD) and 2) recurrent major depressive disorder. Both lifetime and recurrent major depressive disorder were defined using the UK Biobank field code 20126, ICD10 codes F32 (single episode) and F33 (recurrent) or if participants were on antidepressant therapy at the baseline. Individuals who reported any other mental illnesses (e.g., bipolar disorder, schizophrenia, psychosis, etc.) were excluded from the study. Controls included individuals who had not reported depression at the baseline.

Definition of covariates

The covariates considered in the analysis included baseline age, sex, ethnicity, fasting time, assessment center, lifestyle factors including body mass index (BMI), smoking status, alcohol intake frequency, education and status regarding multiple medication from touchscreen or verbal interview and technical variables during the NMR measurement, i.e., batch and

spectrometer. Fasting time was defined as the time interval between the consumption of food or drink and blood sampling and natural log-transformed. Ethnicity was categorized to White, Asian (excluding Chinese), Black, Chinese, mixed and others. Smoking status was categorized to never, previous and current. Alcohol intake frequency was categorized to 1) daily or almost daily, 2) three to four times a week, 3) once or twice a week, 4) less than once a week. Education was categorized to 1) College or University degree, 2) A levels, advanced subsidiary (AS) levels or equivalent, 3) Certificated of secondary education (CSEs) or equivalent, 4) National vocational qualification (NVQ) or higher national diploma (HND) or higher national certificate (HNC) or equivalent, 5) O levels, general certificate of secondary education (GCSEs) or equivalent, 6) Other professional qualifications, and 7) none of the above based on the highest qualification. Information for those who chose "prefer not to answer", was put as missing. Medication status was based on the medication codes collected from the verbal interview which were further coded to Anatomical Therapeutic Chemical (ATC) codes[9]. The medications considered in the covariates were selected based on our previous publication[24], including five anti-hypertensives (C08, C09, C07, C03 and C02), anti-diabetes (metformin and other anti-diabetes under A10), lipid-lowering drugs (C10), digoxin (C01AA), anti-thrombotic (B01AC06), proton pump inhibitors (PPI, A02BC), hypnotics and sedatives (N05) and antidepressants (N06).

Imputation of missing values in the covariates

Fast imputation of missing values by chained random forests was performed through the R package *missRanger* to impute the missing values for the shared covariates, including smoking status, BMI, alcohol intake frequency, education, and ethnicity. The information used in the imputation included baseline age, sex, smoking status, pack-years of smoking, alcohol intake frequency, physical activity from International Physical Activity Questionnaire (IPAQ) groups, ethnicity, BMI, education, blood pressure and waist-hip ratio. In brief, the large matrix was imputed with maximum of ten chaining interactions and 200 trees and weighted by the number of non-missing values; three candidate non-missing values were selected from in the predictive mean matching steps.

Metabolite profiling

The metabolites were measured in plasma using the targeted high-throughput ^{1}H-NMR metabolomics platform of Nightingale (Nightingale Health Ltd; biomarker quantification version 2020)[25,26] which includes 249 metabolites. They include clinical lipids, lipoprotein subclass profiling with lipid concentrations within 14 subclasses, fatty acid composition, and various low-molecular weight metabolites such as amino acids, ketone bodies and glycolysis metabolites quantified in molar concentration units. The technology is based a standardized protocol of sample quality control and sample preparation, data storage and automated spectral analyses.

The data obtained from the baseline sampling was used[25]. For the samples with repeated measurements of the metabolites, one of the values was extracted at random. The metabolite values which were suggested to be technical errors in the quality control provided by Nightingale Health during the measurement procedure were treated as missing. A natural logarithm transformation of each metabolite was performed for the analysis. The zero values were replaced by the lowest value except for zero. Finally the transformed values were scaled to standard deviation units.

Replication

For replication we considered the results from the previously published study performed in Dutch cohorts by the BBMRI-NL consortium[16]. The study included 5,283 patients with depression and 10,145 controls, who were characterized using the Nightingale platform. We further looked up the association of metabolites in the Predictors of Remission in Depression to Individual and Combined Treatments (PReDICT) study. The design and clinical outcomes of PReDICT have been detailed previously [27]. Briefly, the PReDICT study aimed to identify predictors and moderators of response to 12 weeks of randomly-assigned treatment with duloxetine (30-60 mg/day), escitalopram (10-20 mg/day) or cognitive behaviour therapy (CBT, 16 one-hour individual sessions). Eligible participants were adults aged 18-65 years with an active untreated major depressive episode without psychotic features. Severity of depression at the randomization visit was assessed with the 17-item Hamilton Depression Rating Scale (HRSD$_{17}$) [28]. Eligibility required an HRSD$_{17}$ score ≥18 at the screening visit and ≥15

at the randomization visit, indicative of moderate-to-severe depression. Active significant suicide risk, current illicit drug use (assessed with urine drug screen) or a history of substance abuse in the three months prior to randomization, pregnancy, lactation, and uncontrolled general medical conditions were all excluded. Details on metabolomics profiling and statistical analysis are provided in the **Supplemental text**.

Statistical analysis

All analyses were performed in R statistical software. Descriptive analysis was performed using the 'CBCgrps' [29] library of R.

Metabolome-wide association analysis

We used logistic regression to test the association of the metabolite levels with lifetime and recurrent depression. We considered four models with increasing number of covariates in the subsequent models to identify the effects of most known confounders in the regression analysis. Model 1 was adjusted for age, sex, fasting time, ethnicity, assessment centre, and technical variables during the NMR measurement, i.e., batch and spectrometer; model 2 was additionally adjusted for BMI; model 3 further for antidepressant use and model 4 additionally adjusted for most known lifestyle factors including smoking status, alcohol intake frequency, physical activity, level of education and medication use for most chronic diseases. The current analysis included all 249 metabolites measured by Nightingale. False discovery rate (FDR) of 0.05 was used to identify significantly associated metabolites. We further performed a sensitivity analysis by removing those on antidepressant therapy. Finally, we performed forward regression analysis in R on model 4 adjusted residuals of depression-associated metabolites to identify independent metabolites. A multivariate regression analysis was performed including all metabolites selected from forward regression analysis. Multiple testing correction was performed using false discovery rate (FDR).

Integration of metabolic signatures of human gut microbiome and MDD

To identify patterns of correlation in the metabolic signatures of MDD and the human gut microbiome we estimated correlation coefficients from the effect estimates from linear

mixed regression of 1) $Z\text{-score}_{MDD}$ on $Z\text{-score}_{Microbe}$ and 2) $Z\text{-score}_{Microbe}$ on $Z\text{-score}_{MDD}$. $Z\text{-score}_{MDD}$ are the results of association analysis of MDD with Nightingale metabolites from the current study while $Z\text{-score}_{Microbe}$ are the results of association analysis of the gut microbiome ascertained with 16S RNA sequencing and with the Nightingale metabolites published earlier by Vojinovic et al.[21]. Metabolites were clustered to 20 groups based on the method of Li & Ji [30]. These groups were used as random effects in the linear mixed regression. This step was performed to control for inflation in the statistic because of high correlation between the metabolites. Correlation between the metabolic profiles of the 361 gut microbial taxa and MDD was estimated as the square root of the product of the effect estimates from the two regressions described above. Significance of the correlation was tested using the Student's T test. FDR was applied to correct for multiple testing. Next, we compared these T-Statistics (proxy association for MDD-Microbiome based on the MDD-metabolome) as obtained in the present study with the Z-scores (direct association of MDD-Microbiome) from Djawad et al. [20].

Mendelian Randomisation (MR)

To elucidate the causal relationships between depression and the associated metabolites, we performed bi-directional two-sample MR using the R package of *TwoSampleMR* for the inverse variance weighted MR, heterogeneity test and pleiotropy test from MR-Egger regression[31]. The default pipelines in the packages were used. In brief, for the analysis using *TwoSampleMR*, the genetic score was based on the top SNPs (P-value < 10^{-6}) with linkage disequilibrium $R^2 < 0.001$ within 10,000kbps clumping distance. The overlapping SNPs were used without seeking proxy SNPs as we assumed that each meaningful locus should have multiple SNPs significant and overlapped. The metabolite GWAS were obtained from the MRC IEU OpenGWAS in MRbase[31] using all the UK Biobank participants who were profiled for Nightingale metabolites (n=118,000). For MDD we used the publicly available results of the largest GWAS by Howard et al. [32].

Results

Baseline characteristics of the studied samples are provided in **Table 1**. Cases consisted of 8462 individuals with a lifetime major depression and 5403 patients with recurrent major

depression and the controls included > 55420 participants. Patients are significantly younger, are more often female, smokers, have higher education and have a higher body mass index, are less physically active, consume less alcohol and have a black/mixed ethnic background compared to the controls (**Table 1**). Of the patients with history of depression, 1958 (23%) were using anti-depressants at the time blood was drawn for the metabolomic characterization. The patients were found to use more often medication related to gastric diseases, pain and addiction (**Table 1**).

Metabolome-wide association analysis for major depression using the Nightingale platform
Results of the analysis are shown in Figure 1 and Supplementary Figures 1-4. Adjusting for age, sex and technical covariates, 178 (71.0%) metabolites of the 249 metabolites tested were significantly (false discovery rate (FDR) < 0.05) associated with MDD in the basic model 1. When adjusting for BMI (model 2), 163 (65.5%) metabolites remained significantly associated with MDD and further adjusting for antidepressant use (model 3) yielded 132 (53.0%) metabolites significantly associated (FDR<0.05). In the full model further adjusted for lifestyle factors including physical activity, alcohol consumption, smoking, education and medication use for cardio-vascular morbidity (model 4), a total of 124 (49.8%) metabolites remained significantly associated with MDD (**Figure 1**, **Supplementary Table 1**). These include 27 small to extremely large VLDL particles (chylomicrons) that were increased in individuals with MDD patients, 18 medium to very large sized HDL particles all of which were decreased in MDD, except the **triglycerides** in the small and medium HDL particles, and 5 intermediate density lipoprotein (IDL) particles all of which, except the **triglyceride content in IDL**, were decreased in MDD patients. Among fatty acids, total monounsaturated fatty acids (MUFA) and its ratio to total fatty acids was significantly increased in MDD while the ratios of linoleic acid (LA), omega 6 and polyunsaturated fatty acid (PUFA) to total fatty acids were significantly decreased in MDD patients. Further, apolipoprotein A1 (ApoA1), cholesteryl esters, citrate and sphingomyelins were significantly decreased in MDD while alanine and pyruvate were significantly increased in MDD patients (**Figure 1**, **Supplementary Table 1**). Findings were very similar when excluding those with antidepressant use instead of adjusting for antidepressants (**Supplementary Figure 5, Supplementary Table 2**). Comparing the results of lifetime MDD with that of the recurrent depression, the metabolic profiles were found to be highly correlated (**Supplementary Figures 1-4**).

10

Comparing our results with those of the BBMRI-NL [16], 113 of the metabolites that we find associated to major depression were also studied by BBMRI-NL. Supplementary Figure 6 shows that the effect direction of the identified metabolites from model 4 in our study is consistent with that seen in the previous study by the BBMRI-NL consortium (**Supplementary Figure 6, Supplementary Table 3**. In the present study, we identified 49 metabolites that were FDR significant that were not reported in the BBMRI-NL study (**Supplementary Table 4)**. These include metabolites involved in mitochondrial functioning, including alanine, citrate, pyruvate, fatty acids including PUFA%, LA% and omega6%, sphingomyelins, IDL subfractions in addition to some VLDL and HDL subfractions that were not associated earlier. The association of depression to omega 6, PUFA, citrate and pyruvate was replicated in the PReDICT study (**Supplementary Table 5**). We could not replicate the finding on alanine and no data were available on sphingomyelins, IDL or other lipid fractions, although most of these show consistency in direction of effects in the BBMRI-NL study (**Supplementary Figure 6**).

Mendelian Randomization (MR) analysis

Results of bidirectional MR are provided in the **Figure 2** and **Supplementary Table 6**. No significant pleiotropy was observed (**Supplementary Table 7)** in the MR analysis. However, significant heterogeneity was observed for most metabolites (**Supplementary Table 8**) suggesting violation of assumptions of MR. We therefore considered the results of the weighted median method as these have shown to produce consistent results in the presence of heterogeneity[33]. Significant MR results were obtained after multiple testing correction using FDR < 0.05 (**Figure 2**) when MDD was used as the exposure and metabolites as the outcome. Changes in all metabolites except ApoA1, citrate, pyruvate, alanine, total cholesteryl ester, sphingomyelins, medium and large HDL particles, HDL cholesterol, total concentration of HDL particles, total lipids in HDL and average diameter of HDL particles, and a few VLDL subfractions **Figure 2**) appeared to be associated to the genes that are associated to depression.

Integration of human gut microbiome and metabolic signatures

We compared the pattern of association of MDD with the metabolites (metabolic signature of MDD) to the pattern of association of human gut microbial taxa with the metabolites

(metabolic signatures of the gut microbial taxa) (**Supplementary Table 9**, **Figures 3 & 4**). The association of the gut microbiota and the Nightingale metabolites is based on an independent study published earlier[21]. **Figure 3** shows all taxa whose metabolic signatures show significant positive correlation (r > 0.50 & FDR < 0.05) with that of MDD and **Figure 4** illustrates all taxa whose metabolic signatures show significant negative correlation (r < -0.50 & FDR < 0.05) with the metabolic signature of MDD. We henceforth refer to this correlation of the metabolic signatures of MDD and gut microbial taxa as 'proxy association' between MDD and gut microbiome. When comparing this proxy association with the direct association of depression with the gut microbiome from a previous independent study[20], we find a highly significant correlation (r=0.58, p-value < 10^{-16})(**Figure 5)**. This finding suggests that the Nightingale platform can be used as a proxy measure for the microbiome and that the correlation of the microbiome and MDD are mainly driven by lipoproteins, especially VLDL and HDL particles. Of note is that the bacteria which were associated with a healthy lipid profile, i.e., increased levels of HDL subfractions and decreased VLDL lipid levels, were found to be decreased in MDD (**Figure 3**). Conversely, bacteria which were associated with an unhealthy lipid profile, i.e., decreased levels of lipids in HDL particles and increased levels of lipids in VLDL, were found to be increased in patients with MDD (**Figure 4**).

Overall, we observed 223 bacterial taxa significantly associated with MDD using proxy association (FDR$_{corr}$ < 0.05) (**Figure 6**, **Supplementary Table 9**). Figure 6 shows the complete gut microbiome profile of MDD at hierarchical levels. Family *Ruminococcaceae* (r=-0.60, FDR=2.9*10^{-13}) and most of its genera were significantly negatively correlated with MDD patients. Several other families belonging to the order *Clostridiales* (*Clostridiaceae* (r=-0.59, FDR=8.8*10^{-13}), *Christensenellaceae* (-0.68, FDR=1.8*10^{-17}), *Peptostreptococcaceae* (r=-0.50, FDR=2.1*10^{-09}), *Defluviitaleaceae* (r=-0.45, FDR=9.4*10^{-08}) and *Peptococcaceae* (r=-0.41, FDR=2.3*10^{-06})) also showed significant negative correlation with MDD. Families *Lachnospiraceae* (r=0.43, FDR= 6.3*10^{-07}) and *Eubacteriaceae* (r=0.38, FDR=1.3*10^{-05}) were significantly positively correlated with MDD, however, some genera belonging to these families were significantly negatively correlated with MDD (**Figure 6**). Further, several families belonging to the phylum *Proteobacteria* including *Methanobacteriaceae*, *Rhodospirillaceae*, *Desulfovibrionaceae*, *Pasteurellaceae*, *Neisseriaceae* and *Oxalobacteraceae* and phylum *Bacteroidetes* including *Porphyromonadaceae*, *Rikenellaceae*

and *Prevotellaceae* were significantly negatively correlated with MDD (**Figure 6**,

Supplementary Table 9).

Discussion

We have identified 124 metabolites associated with major depression. No differences were

observed between the metabolic profiles of lifetime MDD and recurrent MDD. The findings

of our study corroborate well with those of the previously published large study of the

BBMRI-NL consortium[16] in that most metabolites (90%) show consistency in the direction of

association. Compared to the study of the BBMRI consortium, we find 49 metabolites that

were not significantly associated to MDD earlier, including the amino acid alanine as well as

citrate and pyruvate, which are two key metabolites of the energy metabolism pathway.

The shift of the metabolites observed in depression was associated with bacterial taxa

belonging to order *Clostridiales*, and phyla *Proteobacteria* and *Bacteroidetes*.

Our findings extend previous reports on the major changes in lipid metabolism seen in

patients with MDD. These include novel associations with polyunsaturated fatty acids

(PUFA%, LA%, omega6%), sphingomyelins and various lipoprotein subfractions. In this study

we replicate the shift in the VLDL / HDL axis in patients with MDD[16]. VLDLs are produced by

the liver and are rich in triglyceride. The removal of triglycerides from VLDL by muscle and

adipose tissue results in the formation of IDL particles, which we also find significantly

associated with MDD in the present study (not reported earlier in BBMRI-NL). VLDL, IDL and

LDL particles are atherogenic and are found to be increased in cardio-vascular morbidities[34].

HDL, on the other hand is synthesized in the liver and the intestine[35]. Apolipoprotein A1

(ApoA1) is the major structural protein of HDL accounting for 70% of HDL protein[35]. HDL

particles are responsible for reverse cholesterol transport and have anti-oxidant, anti-

inflammatory, anti-thrombotic, and anti-apoptotic properties[34]. HDLs have been shown to

be protective against mitochondrial dysfunction and positively associated with

mitochondrial oxidative function[15].

The finding that various VLDL particles and MUFA are increased and HDL particles and

APOA1 are decreased in depression are consistent with those of previous studies[16,36]. Our

study provides evidence suggesting that the microbiome is a key player in the shift in

13

VLDL/HDL axis in MDD. Integrating metabolic profiles of MDD and the gut microbiome, we find that the shift in most metabolites, particularly the HDL and VLDL fractions are associated to several families of gut microbiota. Those belonging to the order *Clostridiales*, and phylum *Proteobacteria* and *Bacteroidetes* are decreased in MDD patients and are associated with high HDL lipid levels and low VLDL lipid levels in blood while families *Lachnospiracea* and *Eubacteriaceae* are predicted to be increased in MDD patients and are associated with low levels of HDL subfractions and high levels of VLDL subfractions. Earlier studies have shown that gut microbiome is a major determinant of the circulating lipids and has a bidirectional relationship with mitochondrial function[17,18]. Members of microbiota from order *Clostridiales* are known to provide transforming growth factor *β* enriched environment for promotion and accumulation of regulatory T cells in the gut[37]. These regulatory T cells have been found to be reduced in mood disorders[38]. It is interesting to note that we find several families belonging to the order *Clostridiales* associated with MDD in our proxy association. Our findings are in line with the previous study where *Clostridiales* were found to be the predominant microbes mediating psychiatric disorders including depression[39].

This significant link between the microbiome and MDD may also shed light on the interpretation of the Mendelian Randomization experiments that did not yield significant evidence that any of the human genes associated to the metabolites investigated here. However, we find evidence that changes in lipoproteins and their subfractions and fatty acids associate with genes involved in MDD. We hypothesise that the metabolic change we observe is part of the disease process and that the genes that have been implicated in MDD explain the shift towards increased VLDL and decreased HDL subfractions (except medium and large HDL particles) in MDD. This may be explained by the pleiotropic effects of the genes determining lipid metabolism or by the changes in diet and physical activity that are consequences of the changes in a patient's mood. Alternatively, we hypothesise that the change in these metabolites may be due to altered composition of the gut microbiome in depressed individuals, which may be driven by the genes that determine MDD. Such a genetically driven shift in microbiome is also seen in transgenic models for Alzheimer's disease (AD) in which the introduction of the major genes involved in AD in mice, resulted in a shift of the microbiome[40]. Although we cannot exclude that the VLDL levels in the blood

drive the gut microbiome and MDD, this mechanism is not supported by the Mendelian Randomization. Based on the Mendelian Randomization analysis, we hypothesize that MDD genes drive the gut microbiome, which determines the metabolic spectrum (VLDLs and small HDLs) in the blood[17]. The exact mechanism is to be determined in future experiments.

The second major finding is the disruption in the mitochondrial metabolism, more specifically the tricarboxylic acid (TCA)/Krebs cycle. We find significantly increased levels of pyruvate and decreased levels of citrate, which are major components of the mitochondrial Krebs cycle[41]. To our knowledge, there is no evidence in human data for an association of these metabolites to MDD. Decreased citrate levels were found in the urine of rats in chronic unpredictable mild stress depression model[42,43]. Further, increased mitochondrial activity generating citrate and reduced oxidative stress parameters has been observed in patients with bipolar depression treated with lithium compared to untreated patients[44]. In the brain, citrate contributes to the regulation of neuronal excitability chelating and controlling the availability of divalent ions such as Ca^{+2} and Mg^{+2} [41,45]. Extracellular citrate is also used for neurotransmitter synthesis[41,45]. In animal models of induced oxidative stress, oral administration of citrate decreased brain lipid peroxidation and inflammation, liver damage, and DNA fragmentation[46]. 90% of the total citrate in the body is localized in bones[47,48]. Osteoblast metabolic production of citrate provides the source of citrate in bone in the presence of zinc, which inhibits the oxidation of citrate in mitochondria[48,49]. Citrate is released in the plasma during bone resorption[47]. Decreased levels of blood citrate (hypocitricemia) causes loss of bone citrate and osteoporosis. The relationship of depression and osteoporosis is well established[50], however, Mendelian randomisation study showed no causal effect of depression on osteoporosis[51]. Interestingly, administration of vitamin D increases the plasma and bone citrate concentrations by inhibiting mitochondrial citrate oxidation[47]. This implies that, if plasma citrate deficiency has a causal influence on major depression, the subgroup of individuals with MDD exhibiting low citrate levels may very well be treated with zinc and/or vitamin D supplements, both of which have shown to ameliorate symptoms of depression[52,53].

Low citrate levels may also be a result of intestinal dysbiosis or impairment of the pyruvate dehydrogenase complex that catalyses the conversion of pyruvate into acetyl-CoA[54]. One of

the questions to answer in future studies is whether acetyl-CoA links the findings on lipids and amino acids. Cholesterol synthesis initiates from acetyl coenzyme A (acetyl-CoA)[55]. Acetyl-CoA is synthesized in mitochondria from the oxidative decarboxylation of pyruvate, oxidation of fatty acids or oxidative degradation of some amino acids (e.g., phenylalanine, tyrosine, leucine, lysine, and tryptophan)[56]. Acetyl-CoA is transported out of the mitochondria after being converted into citrate[56]. Cytosolic acetyl-CoA is used in fatty acid metabolism and lipid biosynthesis [57]. We hypothesize that citrate, which we found to be decreased in depression may be the key metabolite connecting both the lipid and energy metabolism pathways via Acetyl-CoA.

In conclusion, we have performed the largest and most comprehensive study investigating the association of NMR metabolites with major depression. We find that metabolites involved in the tricarboxylic acid (TCA) cycle are significantly altered in patients with MDD, suggesting perturbations in metabolites involved in energy metabolism in patients with MDD. Our finding that the interplay between the gut microbiome and the blood metabolome may play a key role in MDD and suggests that the gut microbiome may be a target for novel preventive and therapeutic interventions for MDD.

References

1. Chesney, E., Goodwin, G.M. & Fazel, S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. *World Psychiatry* **13**, 153-60 (2014).
2. Friedrich, M.J. Depression Is the Leading Cause of Disability Around the World. *JAMA* **317**, 1517 (2017).
3. Rezin, G.T., Amboni, G., Zugno, A.I., Quevedo, J. & Streck, E.L. Mitochondrial dysfunction and psychiatric disorders. *Neurochem Res* **34**, 1021-9 (2009).
4. Parekh, A., Smeeth, D., Milner, Y. & Thure, S. The Role of Lipid Biomarkers in Major Depression. *Healthcare (Basel)* **5**(2017).
5. Maeng, S.H. & Hong, H. Inflammation as the Potential Basis in Depression. *Int Neurourol J* **23**, S63-71 (2019).
6. Allen, J., Romay-Tallon, R., Brymer, K.J., Caruncho, H.J. & Kalynchuk, L.E. Mitochondria and Mood: Mitochondrial Dysfunction as a Key Player in the Manifestation of Depression. *Frontiers in Neuroscience* **12**(2018).
7. Bansal, Y. & Kuhad, A. Mitochondrial Dysfunction in Depression. *Current Neuropharmacology* **14**, 610-618 (2016).
8. Kety, S.S. Blood flow and metabolism of the human brain in health and disease. *Trans Stud Coll Physicians Phila* **18**, 103-8 (1950).
9. Manji, H. *et al.* Impaired mitochondrial function in psychiatric disorders. *Nature Reviews Neuroscience* **13**, 293-307 (2012).
10. Fattal, O., Link, J., Quinn, K., Cohen, B.H. & Franco, K. Psychiatric comorbidity in 36 adults with mitochondrial cytopathies. *CNS Spectr* **12**, 429-38 (2007).
11. Gardner, A. & Boles, R.G. Beyond the serotonin hypothesis: Mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* **35**, 730-743 (2011).
12. Bansal, Y. & Kuhad, A. Mitochondrial Dysfunction in Depression. *Curr Neuropharmacol* **14**, 610-8 (2016).
13. Madamanchi, N.R. & Runge, M.S. Mitochondrial dysfunction in atherosclerosis. *Circ Res* **100**, 460-73 (2007).
14. Allen, J., Romay-Tallon, R., Brymer, K.J., Caruncho, H.J. & Kalynchuk, L.E. Mitochondria and Mood: Mitochondrial Dysfunction as a Key Player in the Manifestation of Depression. *Front Neurosci* **12**, 386 (2018).
15. Parikh, N.I. *et al.* Lipoprotein concentration, particle number, size and cholesterol efflux capacity are associated with mitochondrial oxidative stress and function in an HIV positive cohort. *Atherosclerosis* **239**, 50-4 (2015).
16. Bot, M. *et al.* Metabolomics Profile in Depression: A Pooled Analysis of 230 Metabolic Markers in 5283 Cases With Depression and 10,145 Controls. *Biol Psychiatry* **87**, 409-418 (2020).
17. Fu, J. *et al.* The Gut Microbiome Contributes to a Substantial Proportion of the Variation in Blood Lipids. *Circ Res* **117**, 817-24 (2015).
18. Franco-Obregon, A. & Gilbert, J.A. The Microbiome-Mitochondrion Connection: Common Ancestries, Common Mechanisms, Common Goals. *mSystems* **2**(2017).
19. Simpson, C.A. *et al.* The gut microbiota in anxiety and depression - A systematic review. *Clin Psychol Rev* **83**, 101943 (2021).
20. Djawad, R. *et al. Nature Portfolio* (2021).

21.   Vojinovic, D. *et al.* Relationship between gut microbiota and circulating metabolites in population-based cohorts. *Nat Commun* **10**, 5813 (2019).

22.   Zierer, J. *et al.* The fecal metabolome as a functional readout of the gut microbiome. *Nat Genet* **50**, 790-795 (2018).

23.   Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).

24.   Liu, J. *et al.* Integration of epidemiologic, pharmacologic, genetic and gut microbiome data in a drug-metabolite atlas. *Nat Med* **26**, 110-117 (2020).

25.   Ritchie, S.C. *et al.* Quality control and removal of technical variation of NMR metabolic biomarker data in ~120,000 UK Biobank participants. *medRxiv*, 2021.09.24.21264079 (2021).

26.   Julkunen, H., Cichonska, A., Slagboom, P.E., Wurtz, P. & Nightingale Health, U.K.B.I. Metabolic biomarker profiling for identification of susceptibility to severe pneumonia and COVID-19 in the general population. *Elife* **10**(2021).

27.   Dunlop, B.W. *et al.* Predictors of remission in depression to individual and combined treatments (PReDICT): study protocol for a randomized controlled trial. *Trials* **13**, 106 (2012).

28.   Hamilton, M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* **23**, 56-62 (1960).

29.   Zhang, Z., Gayle, A.A., Wang, J., Zhang, H. & Cardinal-Fernandez, P. Comparing baseline characteristics between groups: an introduction to the CBCgrps package. *Ann Transl Med* **5**, 484 (2017).

30.   Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)* **95**, 221-7 (2005).

31.   Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**(2018).

32.   Howard, D.M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci* **22**, 343-352 (2019).

33.   Bowden, J., Davey Smith, G., Haycock, P.C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* **40**, 304-14 (2016).

34.   Feingold, K.R. Introduction to Lipids and Lipoproteins. in *Endotext* (eds. Feingold, K.R. *et al.*) (South Dartmouth (MA), 2000).

35.   Mahley, R.W., Innerarity, T.L., Rall, S.C., Jr. & Weisgraber, K.H. Plasma lipoproteins: apolipoprotein structure and function. *J Lipid Res* **25**, 1277-94 (1984).

36.   Segoviano-Mendoza, M. *et al.* Hypocholesterolemia is an independent risk factor for depression disorder and suicide attempt in Northern Mexican population. *Bmc Psychiatry* **18**(2018).

37.   Imdad, S., Lim, W., Kim, J.H. & Kang, C. Intertwined Relationship of Mitochondrial Metabolism, Gut Microbiome and Exercise Potential. *Int J Mol Sci* **23**(2022).

38.   Bauer, M.E. & Teixeira, A.L. Neuroinflammation in Mood Disorders: Role of Regulatory Immune Cells. *Neuroimmunomodulation* **28**, 99-107 (2021).

39.   Li, J.J. *et al.* Clostridiales are predominant microbes that mediate psychiatric disorders. *Journal of Psychiatric Research* **130**, 48-56 (2020).

40. Łuc, M. *et al.* Gut microbiota in dementia. Critical review of novel findings and their potential application. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **104**, 110039 (2021).

41. Mycielska, M.E., Milenkovic, V.M., Wetzel, C.H., Rummele, P. & Geissler, E.K. Extracellular Citrate in Health and Disease. *Current Molecular Medicine* **15**, 884-891 (2015).

42. Liu, X.J. *et al.* Urinary metabonomic study using a CUMS rat model of depression. *Magnetic Resonance in Chemistry* **50**, 187-192 (2012).

43. Zheng, S.N. *et al.* Urinary metabonomic study on biochemical changes in chronic unpredictable mild stress model of depression. *Clinica Chimica Acta* **411**, 204-209 (2010).

44. Machado-Vieira, R. *et al.* Oxidative stress parameters in unmedicated and treated bipolar subjects during initial manic episode: A possible role for lithium antioxidant effects. *Neuroscience Letters* **421**, 33-36 (2007).

45. Westergaard, N. *et al.* Uptake, Release, and Metabolism of Citrate in Neurons and Astrocytes in Primary Cultures. *Journal of Neurochemistry* **62**, 1727-1733 (1994).

46. Abdel-Salam, O.M. *et al.* Citric acid effects on brain and liver oxidative stress in lipopolysaccharide-treated mice. *J Med Food* **17**, 588-98 (2014).

47. Costello, L.C. & Franklin, R.B. Plasma Citrate Homeostasis: How It Is Regulated; And Its Physiological and Clinical Implications. An Important, But Neglected, Relationship in Medicine. *HSOA J Hum Endocrinol* **1**(2016).

48. Costello, L.C., Franklin, R.B., Reynolds, M.A. & Chellaiah, M. The Important Role of Osteoblasts and Citrate Production in Bone Formation: "Osteoblast Citration" as a New Concept for an Old Relationship. *Open Bone J* **4**(2012).

49. Franklin, R.B., Chellaiah, M., Zou, J., Reynolds, M.A. & Costello, L.C. Evidence that Osteoblasts are Specialized Citrate-producing Cells that Provide the Citrate for Incorporation into the Structure of Bone. *Open Bone J* **6**, 1-7 (2014).

50. Cizza, G., Primma, S., Coyle, M., Gourgiotis, L. & Csako, G. Depression and osteoporosis: a research synthesis with meta-analysis. *Horm Metab Res* **42**, 467-82 (2010).

51. He, B. *et al.* Depression and Osteoporosis: A Mendelian Randomization Study. *Calcif Tissue Int* **109**, 675-684 (2021).

52. Petrilli, M.A. *et al.* The Emerging Role for Zinc in Depression and Psychosis. *Front Pharmacol* **8**, 414 (2017).

53. Shaffer, J.A. *et al.* Vitamin D supplementation for depressive symptoms: a systematic review and meta-analysis of randomized controlled trials. *Psychosom Med* **76**, 190-6 (2014).

54. Thomas, G.W., Mains, C.W., Slone, D.S., Craun, M.L. & Bar-Or, D. Potential Dysregulation of the Pyruvate Dehydrogenase Complex by Bacterial Toxins and Insulin. *Journal of Trauma-Injury Infection and Critical Care* **67**, 628-633 (2009).

55. Lent-Schochet, D. & Jialal, I. Biochemistry, Lipoprotein Metabolism. in *StatPearls* (Treasure Island (FL), 2021).

56. Bhagavan, N.V. & Ha, C.-E. Chapter 16 - Lipids I: Fatty Acids and Eicosanoids. in *Essentials of Medical Biochemistry (Second Edition)* (eds. Bhagavan, N.V. & Ha, C.-E.) 269-297 (Academic Press, San Diego, 2015).

57. Mycielska, M.E. *et al.* Citrate transport and metabolism in mammalian cells: prostate epithelial cells and prostate cancer. *Bioessays* **31**, 10-20 (2009).
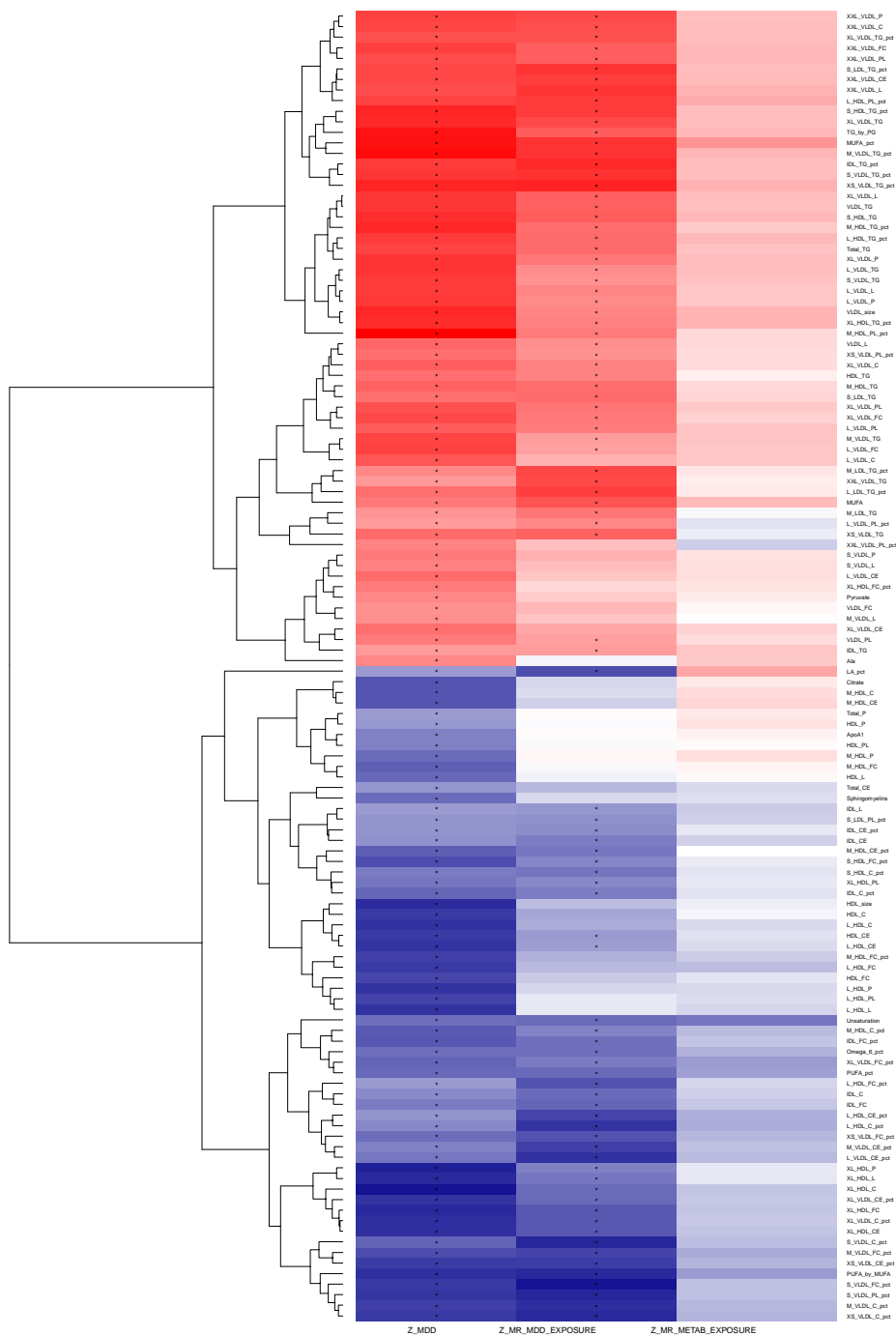
Acknowledgements

Conflicts of interest

Table 1: Descriptive Statistics for the patients with MDD and controls and recurrent depression

	Lifetime MDD			Recurrent MDD		
Variables	0 (n = 55420)	1 (n = 8462)	p	0 (n = 67805)	1 (n = 5403)	p
age0, Median (Q1,Q3)	59 (51, 64)	57 (49, 62)	< 1e-10	58 (50, 64)	57 (49, 62)	< 1e-10
sex, n (%)			< 1e-10			< 1e-10
Female	25987 (47)	5474 (65)		32638 (48)	3529 (65)	
Male	29433 (53)	2988 (35)		35167 (52)	1874 (35)	
BMI, Median (Q1,Q3)	26.67 (24.17, 29.65)	26.81 (24.11, 30.32)	1.31E-06	26.68 (24.15, 29.72)	26.93 (24.1, 30.55)	1.44E-07
Fasting_time, Median (Q1,Q3)	1.1 (0.69, 1.39)	1.1 (1.1, 1.39)	< 1e-10	1.1 (0.69, 1.39)	1.1 (1.1, 1.39)	2.30E-09
Ethnicity, n (%)			2.00E-10			3.13E-06
Asian	992 (2)	153 (2)		1420 (2)	111 (2)	
Black	751 (1)	140 (2)		1078 (2)	102 (2)	
Chinese	182 (0)	11 (0)		248 (0)	8 (0)	
Mixed	247 (0)	81 (1)		312 (0)	51 (1)	
Others	461 (1)	83 (1)		670 (1)	55 (1)	
White	52787 (95)	7994 (94)		64077 (95)	5076 (94)	
Education, n (%)			< 1e-10			< 1e-10
A_levels_AS_levels_or_equivalent	6127 (11)	1042 (12)		7413 (11)	655 (12)	
CSEs_or_equivalent	2816 (5)	480 (6)		3636 (5)	304 (6)	
College_or_University_degree	18209 (33)	2954 (35)		21836 (32)	1918 (35)	
NVQ_or_HND_or_HNC_or_equivalent	4054 (7)	547 (6)		4921 (7)	346 (6)	
None	9374 (17)	1117 (13)		11925 (18)	700 (13)	
O_levels_GCSEs_or_equivalent	11742 (21)	1884 (22)		14414 (21)	1194 (22)	
Other_professional_qualifications	3098 (6)	438 (5)		3660 (5)	286 (5)	
Smoking status, n (%)			< 1e-10			< 1e-10
current	4752 (9)	1101 (13)		6217 (9)	749 (14)	
never	31917 (58)	4232 (50)		38827 (57)	2673 (49)	

|						
previous	18751 (34)	3129 (37)		22761 (34)	1981 (37)	
alcohol_freq, n (%)			< 1e-10			< 1e-10
Daily_or_almost_daily	11791 (21)	1661 (20)		14071 (21)	1070 (20)	
Less_than_once_a_week	14831 (27)	2938 (35)		19037 (28)	1917 (35)	
Once_or_twice_a_week	14867 (27)	2038 (24)		18101 (27)	1265 (23)	
Three_or_four_times_a_week	13931 (25)	1825 (22)		16596 (24)	1151 (21)	
physical_activity, n (%)			8.48E-06			0.06170695
high	22422 (40)	3299 (39)		27078 (40)	2100 (39)	
low	9806 (18)	1678 (20)		12358 (18)	1051 (19)	
moderate	23192 (42)	3485 (41)		28369 (42)	2252 (42)	
Medication						
A02BC_Proton_pump_inhibitors, n (%)	4176 (8)	1107 (13)	< 1e-10	5561 (8)	735 (14)	< 1e-10
C01AA_Digitalis_glycosides, n (%)	134 (0)	15 (0)	0.30530126	164 (0)	6 (0)	0.07576164
C10_LIPID_MODIFYING_AGENTS, n (%)	11083 (20)	1583 (19)	0.00578375	13527 (20)	1080 (20)	0.9589938
A10BA02_metformin, n (%)	1456 (3)	268 (3)	0.00482339	1840 (3)	207 (4)	2.01E-06
A10.excl.A10BA02_Anti_diabetes_excl.Metformin, n (%)	469 (1)	62 (1)	0.31366624	614 (1)	47 (1)	0.84782497
B01AC06_acetylsalicylic_acid, n (%)	7719 (14)	970 (11)	8.00E-10	9503 (14)	643 (12)	1.64E-05
C07_BETA_BLOCKING_AGENTS, n (%)	3617 (7)	561 (7)	0.73857387	4516 (7)	365 (7)	0.80899618
C08_CALCIUM_CHANNEL_BLOCKERS, n (%)	4137 (7)	583 (7)	0.06264847	5058 (7)	385 (7)	0.38234663
C03_DIURETICS, n (%)	4220 (8)	608 (7)	0.17061941	5133 (8)	392 (7)	0.41400884
C02_ANTIHYPERTENSIVES, n (%)	833 (2)	103 (1)	0.04660797	1010 (1)	69 (1)	0.2345284
C09_AGENTS_ACTING_ON_THE_RENIN_ANGIOTENSIN_SYSTEM, n (%)	7883 (14)	1100 (13)	0.00268185	9550 (14)	712 (13)	0.06769179
N02A_OPIOIDS, n (%)	1651 (3)	627 (7)	< 1e-10	2475 (4)	416 (8)	< 1e-10
N02B_OTHER_ANALGESICS_AND_ANTIPYRETICS, n (%)	14312 (26)	2770 (33)	< 1e-10	18370 (27)	1839 (34)	< 1e-10
N02C_ANTIMIGRAINE_PREPARATIONS, n (%)	429 (1)	147 (2)	< 1e-10	579 (1)	101 (2)	< 1e-10
N03A_ANTIEPILEPTICS, n (%)	545 (1)	253 (3)	< 1e-10	816 (1)	134 (2)	< 1e-10
N04A_ANTICHOLINERGIC_AGENTS, n (%)	9 (0)	11 (0)	1.15E-05	14 (0)	10 (0)	3.55E-06
N04B_DOPAMINERGIC_AGENTS, n (%)	110 (0)	26 (0)	0.05804231	147 (0)	12 (0)	1

N05A_ANTIPSYCHOTICS, n (%)	101 (0)	212 (3)	< 1e-10	127 (0)	111 (2)	< 1e-10
N05B_ANXIOLYTICS, n (%)	54 (0)	88 (1)	< 1e-10	81 (0)	64 (1)	< 1e-10
N05C_HYPNOTICS_AND_SEDATIVES, n (%)	90 (0)	155 (2)	< 1e-10	155 (0)	90 (2)	< 1e-10
N06A_ANTIDEPRESSANTS, n (%)	46 (0)	1958 (23)	< 1e-10	880 (1)	1297 (24)	< 1e-10
N06B_PSYCHOSTIMULANTS_AGENTS, n (%)	4 (0)	3 (0)	0.05378305	5 (0)	1 (0)	0.36873461
N06C_PSYCHOLEPTICS_AND_PSYCHOANALEPTICS, n (%)	0 (0)	4 (0)	0.00030769	5 (0)	1 (0)	0.36873461
N06D_ANTI_DEMENTIA_DRUGS, n (%)	290 (1)	27 (0)	0.01609039	353 (1)	16 (0)	0.03214806
N07A_PARASYMPATHOMIMETICS, n (%)	14 (0)	5 (0)	0.09632632	25 (0)	3 (0)	0.45870756
N07B_DRUGS_USED_IN_ADDICTIVE_DISORDERS, n (%)	30 (0)	17 (0)	9.76E-06	43 (0)	9 (0)	0.01309514
N07C_ANTIVERTIGO_PREPARATIONS, n (%)	125 (0)	39 (0)	0.00010914	161 (0)	24 (0)	0.00556512
N01A_ANESTHETICS_GENERAL, n (%)	4 (0)	3 (0)	0.05378305	14 (0)	2 (0)	0.33285557

Figure 1: Results of metabolome-wide association analysis

Left Y-axis shows effect estimates and right Y-axis shows the strength of association. Red dashed line is FDR=0.05. Points denote the -log10(FDR).
Lines denote the effect estimates from the four models.
Model 1: MDD ~ age + sex + ethnicity + fasting time + technical covariates
Model 2: MDD ~ age + sex + BMI + ethnicity + fasting time + technical covariates
Model 3: MDD ~ age + sex + BMI + antidepressants + ethnicity + fasting time + technical covariates
Model 4: MDD ~ age + sex + BMI + antidepressants + education + smoking + alcohol intake + physical activity + PPI + antidiabetics + antihypertensives + lipid modifying agents + sleep medication + ethnicity + fasting time + technical covariates

Figure 2: Results of bi-directional Mendelian Randomization

The first column from the left shows the direction of association of 124 significantly associated metabolites with MDD (Z-scores). Blue is negative association and red is positive association. The second column shows the results of the Mendelian randomization (MR) analysis when MDD is the exposure and microbiome is the outcome. The third column depicts the results of MR when metabolites were used as exposure and MDD outcome. Black stars represent significant ones after correcting for multiple testing using FDR.

Figure 3: Association of metabolic profiles (z-scores) of 'healthy' gut microbiota and MDD
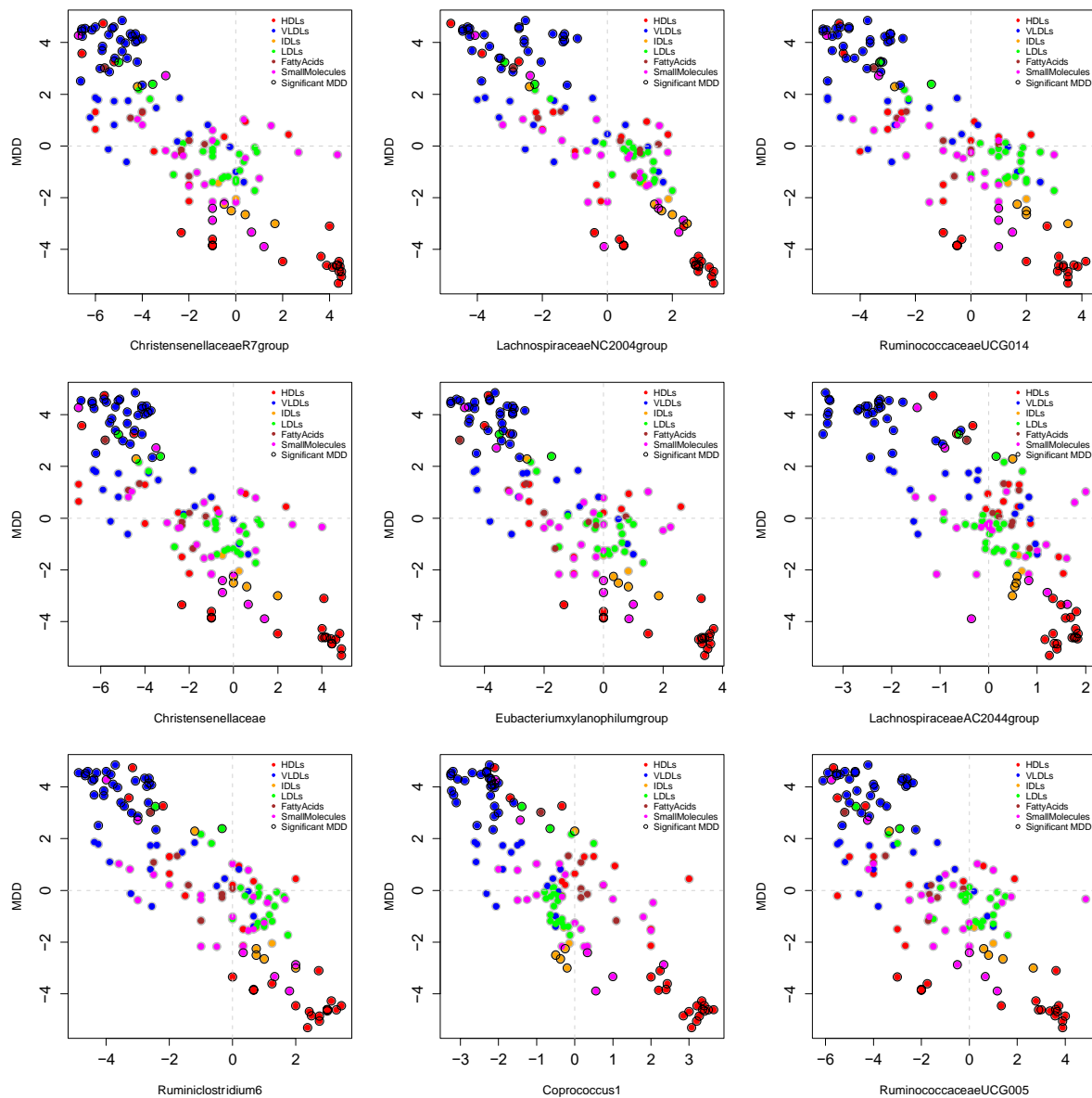
A scatter plot showing the correlation between the metabolic profiles of microbial taxa that are negatively associated the metabolic profile of MDD. Each dot represents a metabolite. X-axis shows the association of the metabolite to microbial taxa and Y-axis shows the association of the metabolite to MDD. Different colours of the dots represent the class of the metabolite and the dots highlighted with black circles are the significantly associated metabolites with MDD in model 4.

Figure 4: Association of metabolic profiles (z-scores) of unhealthy microbiota and MDD
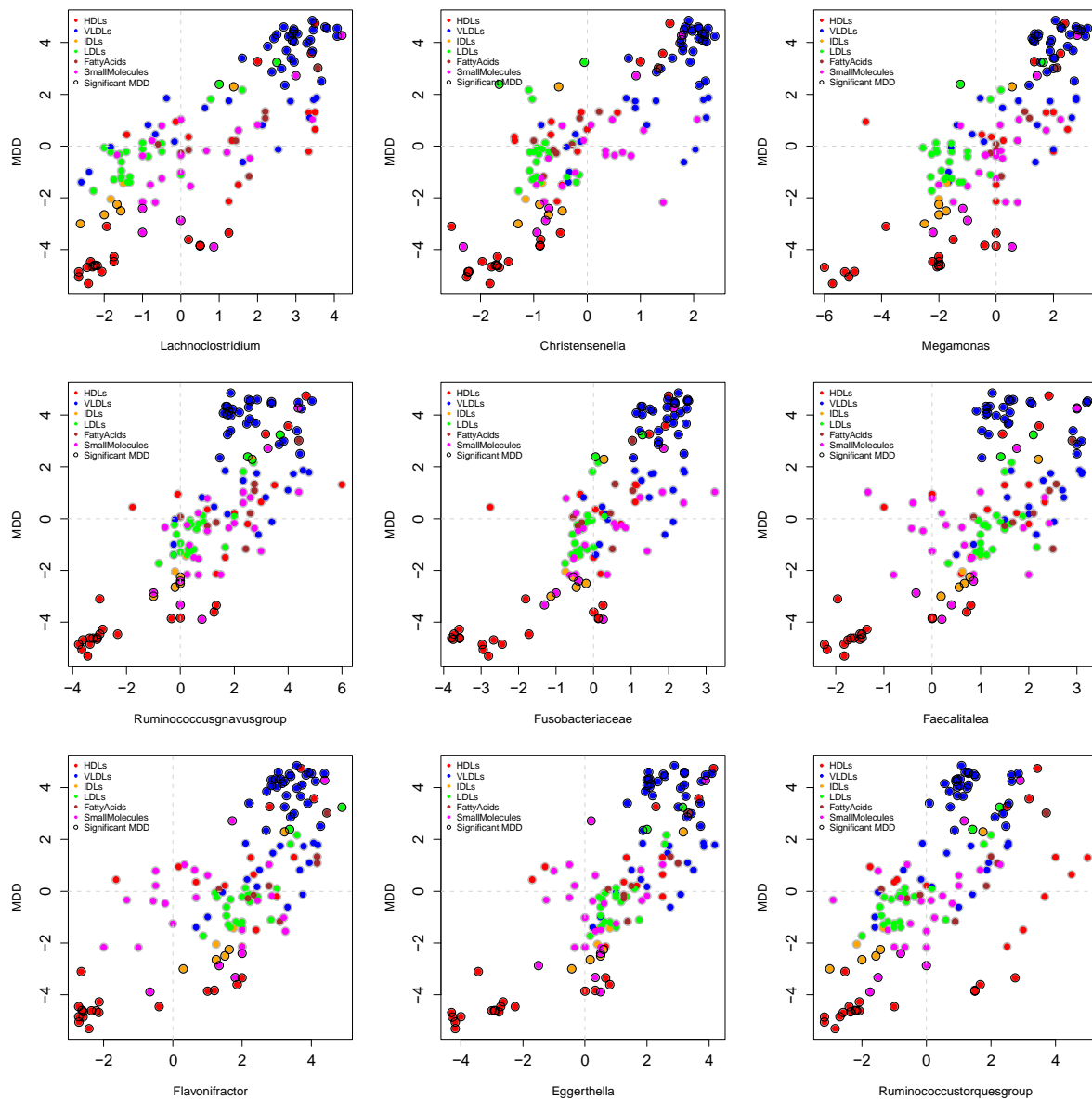
A scatter plot showing the correlation between the metabolic profiles of microbial taxa that are positively associated the metabolic profile of MDD. Each dot represents a metabolite. X-axis shows the association of the metabolite to microbial taxa and Y-axis shows the association of the metabolite to MDD. Different colours of the dots represent the class of the metabolite and the dots highlighted with black circles are the significantly associated metabolites with MDD in model 4.

Figure 5: Scatter plot of direct and proxy association of microbiome with major depression. Taxa that have a p-value < 0.05 in both proxy and direct association are annotated.
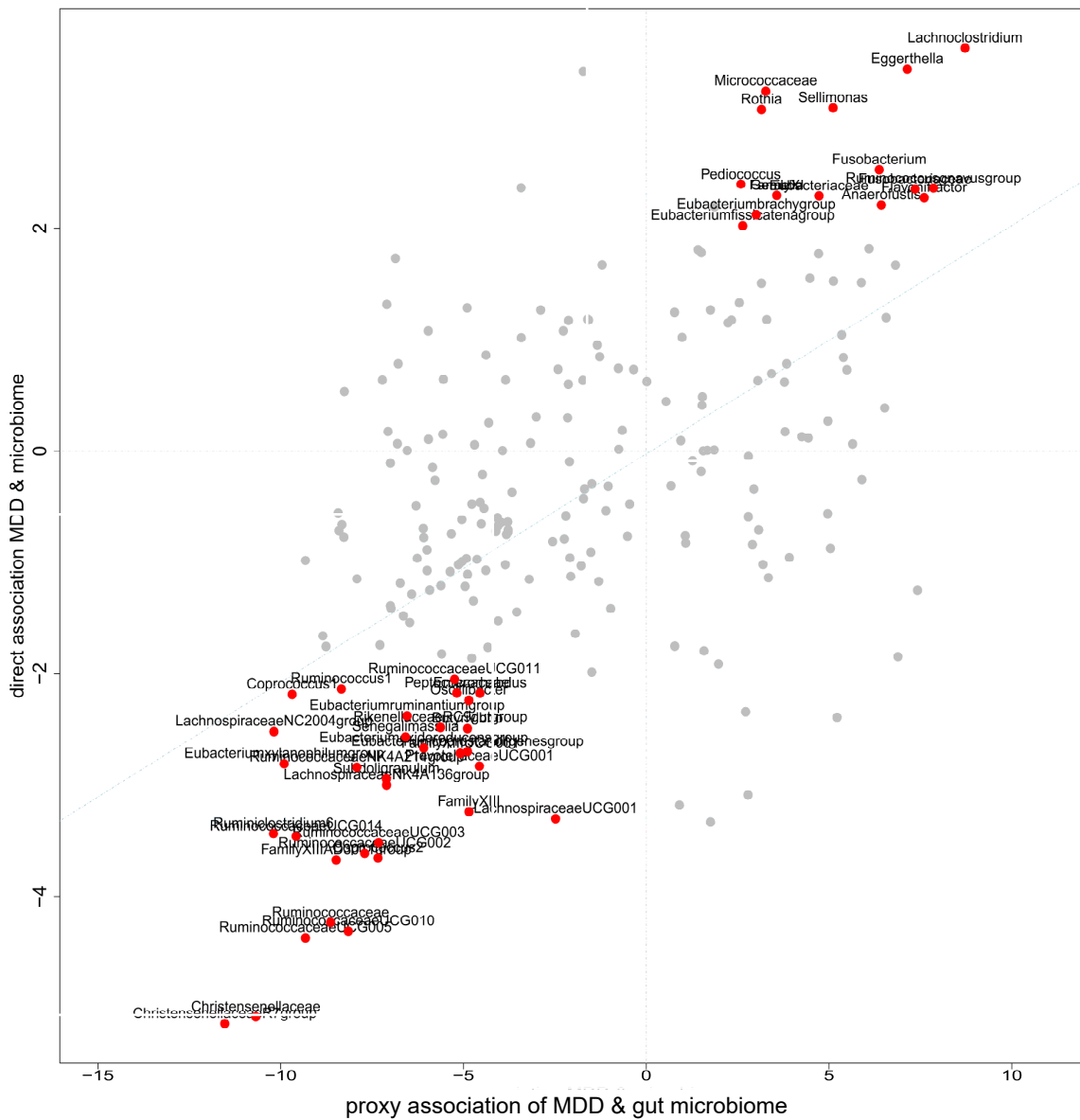
Each dot represents a microbial taxon. X-axis depicts the proxy association (T-scores) between microbial taxa and MDD inferred through their metabolic signatures and Y-axis depicts the direct association of microbial taxa with MDD (Z-scores) from the Rotterdam study performed in Radjabzadeh et al[20]. Taxa that are nominally significant in both proxy and direct associations are annotated.

Figure 6: Hierarchical illustration of all healthy and pathogenic bacteria that showed significant correlation (r > 0.3 & FDR < 0.05) with MDD metabolic profile
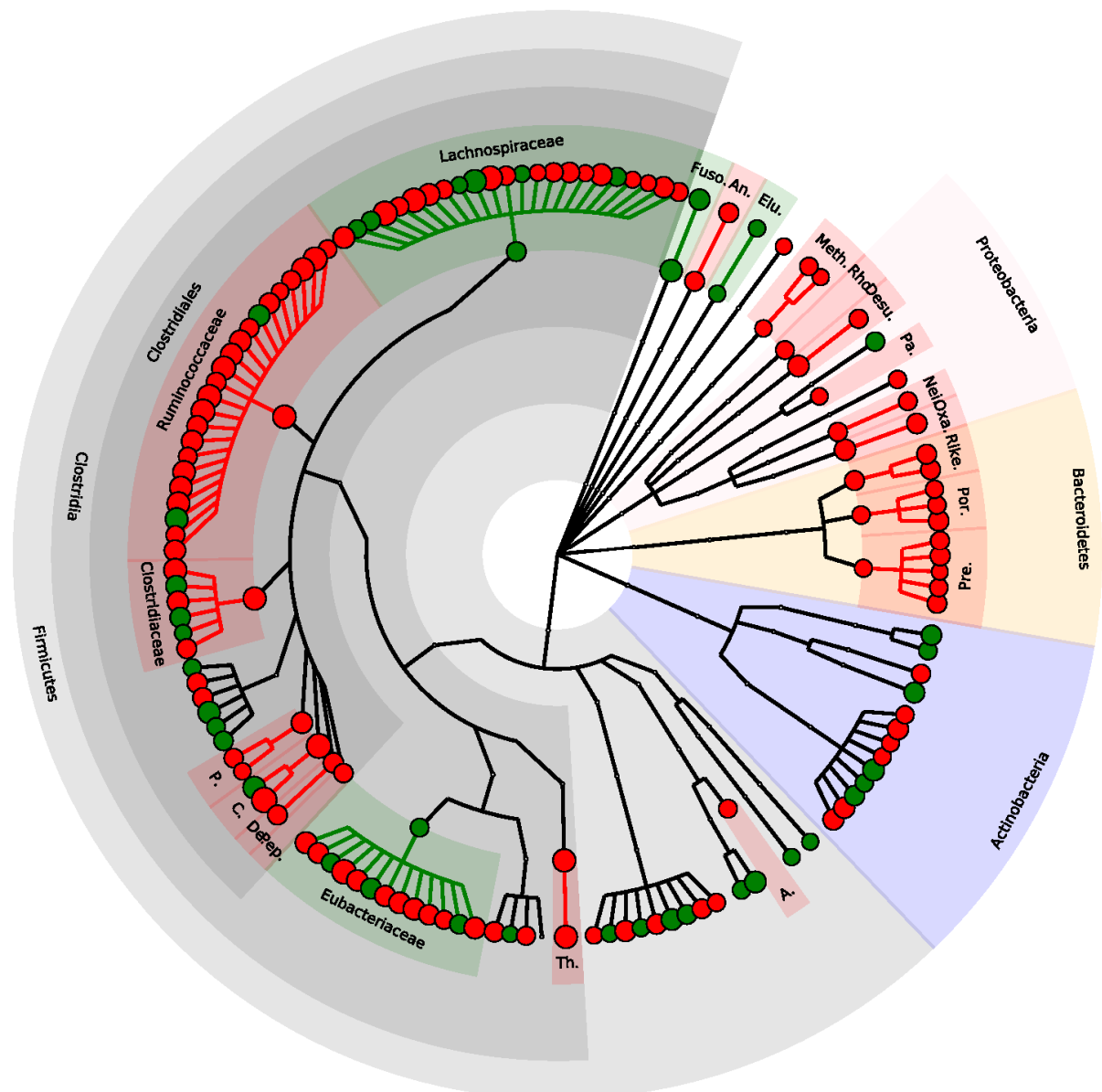
A.: Acidaminococcaceae
An.: Anaeroplasmataceae
C.: Christensenellaceae
De.: Defluviitaleaceae
Desu.: Desulfovibrionaceae
Elu.: Elusimicrobiacea
Fuso.: Fusobacteriaceae
Meth.: Methanobacteriaceae
Nei.: Neisseriaceae
Oxa.: Oxalobacteraceae
P.: Peptostreptococcaceae
Pa.: Pasteurellaceae
Pep.: Peptococcaceae
Por.: Porphyromonadaceae
Pre.: Prevotellaceae
Rho.: Rhodospirillaceae
Rike.: Rikenellaceae
Th.: Thermoanaerobacteraceae

Red dots represent significant negative proxy association between microbial taxa and MDD and green dots represent significant positive proxy association between microbial taxa and MDD. The outermost layer depicts the phylum followed by, class, order, family and genus.