

Metabolomic signatures associated with depression and predictors of antidepressant response in humans: A CAN-BIND-1 report

Giorgia Caspani ¹, Gustavo Turecki ^{2,3}, Raymond W. Lam ⁴, Roumen V. Milev ^{5,6}, Benicio N. Frey ^{7,8}, Glenda M. MacQueen⁹, Daniel J. Müller^{10,11}, Susan Rotzinger^{11,12,13}, Sidney H. Kennedy^{11,12,13}, Jane A. Foster ^{7,8,13} & Jonathan R. Swann ^{1,14}✉

One of the biggest challenges in treating depression is the heterogeneous and qualitative nature of its clinical presentations. This highlights the need to find quantitative molecular markers to tailor existing treatment strategies to the individual's biological system. In this study, high-resolution metabolic phenotyping of urine and plasma samples from the CAN-BIND study collected before treatment with two common pharmacological strategies, escitalopram and aripiprazole, was performed. Here we show that a panel of LDL and HDL subfractions were negatively correlated with depression in males. For treatment response, lower baseline concentrations of apolipoprotein A1 and HDL were predictive of escitalopram response in males, while higher baseline concentrations of apolipoprotein A2, HDL and VLDL subfractions were predictive of aripiprazole response in females. These findings support the potential of metabolomics in precision medicine and the possibility of identifying personalized interventions for depression.

¹ Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK. ² Department of Psychiatry, McGill University, Montreal, QC, Canada. ³ Douglas Mental Health University Institute Frank B. Common (FBC), Montreal, QC, Canada. ⁴ Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada. ⁵ Department of Psychiatry, Queen's University, Kingston, ON, Canada. ⁶ Providence Care Hospital, Kingston, ON, Canada. ⁷ Department of Psychiatry & Behavioral Neurosciences, McMaster University, Hamilton, ON, Canada. ⁸ St. Joseph's Healthcare, Hamilton, ON, Canada. ⁹ Hotchkiss Brain Institute, University of Calgary, Calgary, AL, Canada. ¹⁰ Pharmacogenetics Research Clinic, Centre for Addiction and Mental Health, Toronto, ON, Canada. ¹¹ Department of Psychiatry, University of Toronto, Toronto, ON, Canada. ¹² Department of Psychiatry, Centre for Mental Health, University Health Network, Krembil Research Institute, University of Toronto, Toronto, ON, Canada. ¹³ Centre for Depression and Suicide Studies, St Michael's Hospital, Li Ka Shing Knowledge Institute, Toronto, ON, Canada. ¹⁴ School of Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK. ✉email: j.swann@soton.ac.uk

Major depressive disorder (MDD) is a complex, debilitating psychiatric disorder. Not only is it the most prevalent mood disorder, but it is also the leading cause of disability worldwide, with an estimated 322 million people affected globally¹. MDD is a heterogeneous disease characterized by highly variable symptomatic manifestations². Our ability to elicit the biology of MDD is impacted by the lack of association between clinical symptoms and underlying biology. Currently, treatment selection is based on a trial-and-error strategy³. As a result, the proportion of patients who achieve complete remission after a first pharmacological therapy is limited to one third⁴, while ~30% of patients have not achieved remission after four consecutive treatments^{5,6}. These observations highlight the urgent need to identify biomarkers of MDD and predictors of treatment response, to refine diagnosis, and personalization of treatment strategies to enhance their effectiveness.

Metabolic phenotyping involves the comprehensive measurement of metabolites in biological samples to identify disease-associated metabolic signatures and to characterize biochemical responses to stimuli. This powerful technology enables a broad range of biochemical events occurring in the whole organism to be studied simultaneously⁷. Previous metabolomic studies have identified metabolic signatures of MDD in both urine and plasma, including disruptions to amino acid and energy metabolism as well as alterations in ketone bodies and gut microbial metabolites^{8,9}. Pharmacometabolomics has emerged as a promising approach that can predict an individual's response to pharmacological therapy based on their pre-dose metabolic phenotype. For example, individuals who responded to sertraline or placebo treatment could be discriminated from those who did not by investigating their serum metabolic profile at baseline¹⁰ and its change over time¹¹. Baseline plasma lipids such as phosphatidylcholines and non-hydroxylated sphingomyelins¹², as well as metabolites from the tryptophan, tyrosine, and purine pathways¹³, were predictive of response to citalopram/escitalopram and to the combination of bupropion and escitalopram. Similarly, changes in glutamate and circulating phospholipids correlated with responses to ketamine and its enantiomer esketamine¹⁴, while glycine was negatively associated with escitalopram treatment outcome¹⁵. Both sex and BMI are likely to exert a strong influence on pharmacokinetics. There have been reports of females being more responsive to selective serotonin reuptake inhibitors (SSRIs) than males¹⁶, possibly due to hormones like progesterone and estradiol, lower rates of gastric emptying (leading to quicker absorption¹⁷), and higher adiposity (leading to greater distribution^{18,19}). For this reason, the present study investigated sex differences in metabolic associations with treatment outcome, while accounting for the confounding effects of BMI, as well as age and inflammation. Although there is extensive evidence supporting a link between dyslipidemia and MDD^{20–24}, to the best of our knowledge, only one metabolomic study to date has defined differences in lipoprotein profile between MDD and healthy control (HC) participants²⁵, and none has tested the potential of lipoproteins to predict treatment outcome.

Canadian Biomarker Integration Network in Depression (CAN-BIND) is a national network of researchers across eight Canadian universities that focuses on identifying biomarkers of depression and their potential to predict treatment response²⁶. In this study, high-resolution metabolic phenotyping of urine and plasma samples from the CAN-BIND study collected pre-treatment with two common pharmacological strategies, escitalopram and aripiprazole, was performed. Significant, sex-specific metabolic signatures of depression and of antidepressant treatment response were identified in plasma, but not in urine. In males, depression was significantly correlated with low plasma

concentrations of a number of LDL and HDL subfractions. Additionally, low baseline concentrations of apolipoprotein A1 and HDL were predictive of escitalopram response in males, while high baseline concentrations of apolipoprotein A2, HDL, and VLDL subfractions were predictive of aripiprazole augmentation response in females.

Results

Clinical characteristics of the CAN-BIND 1 cohort. A total of 323 participants were recruited from 6 outpatient centers across Canada between August 2013 and December 2016^{26–28}. These consisted of 112 HCs (Montgomery–Åsberg Depression Rating Scale [MADRS] = 0.8 ± 1.7) and 211 depressed (MDD) participants (MADRS = 29.9 ± 5.6). The groups were matched for age ($N = 323$, Mann–Whitney $U = 17,229.5$, $p = 0.252$) and sex ($N = 323$, $\chi(1) = 0.0041$; $p = 0.9491$). Participants with clinical depression underwent a 16-week two-phase treatment trial: escitalopram (10–20 mg) was administered during phase I (baseline to week 8), and escitalopram either alone (in responders) or augmented with aripiprazole (2–10 mg, in non-responders) was administered during phase II (week 9–16), as shown in Fig. 1. Plasma and urine samples were collected pre-treatment and analyzed by ¹H nuclear magnetic resonance (NMR), giving rise to a panel of 112 lipoproteins (Supplementary Data 1) and 50 urinary metabolites (Supplementary Data 2). PCA models constructed on the urine metabolic profiles and plasma lipoprotein profiles of all participants showed that no clustering was apparent based on the site that the samples were collected from, demonstrating that the metabolic profile was not affected by study location (Supplementary Fig. 1). Plasma data from 12 participants and urine data from 9 participants were excluded as they did not pass QC checks. In addition, 31 participants dropped out of the study before the termination of the clinical trial. The N given in the following sections reflects the number of samples that were included in the analysis, excluding removed observations and dropouts. Of the remaining 180 participants, 83 (46%) were given escitalopram monotherapy for the entire duration of the study, while 97 (54%) were augmented with aripiprazole in phase II. Further demographic and clinical information on the participants, including response rates, can be found in Table 1.

Biochemical signature of MDD. To determine the biochemical variation associated with depression, weighted correlation networks were constructed on the metabolic phenotypes of MDD and HC participants. Correlation network analysis allows the relationship of clusters of highly associated features with an outcome of interest to be studied. In this section, separate networks were built for the plasma and urinary data and their metabolic associations with MADRS score at baseline, as well as with the covariates age, BMI, and CRP. The analytical workflow is depicted in Fig. 1.

In the plasma data ($N = 310$), 10 distinct modules were identified with high topological overlap ($\beta = 7$). Of these, only module 1, containing six highly correlated metabolites, was positively associated with MADRS score at baseline ($p = 0.004$), as well as with age ($p < 0.001$) and BMI ($p < 0.001$), (Supplementary Fig. 2a). This included low-density lipoprotein (LDL)₆ and its subfractions (triglycerides, cholesterol, free cholesterol, phospholipids, apolipoprotein B).

In total, seven modules were positively associated with age, indicating a general age-related increase in lipoprotein concentrations. For CRP, four modules (LDL of high size and very-low-density lipoproteins [VLDL]-predominant) were positively associated and four (high-density lipoproteins [HDL]-predominant, and LDL of low size) were negatively associated, indicating the

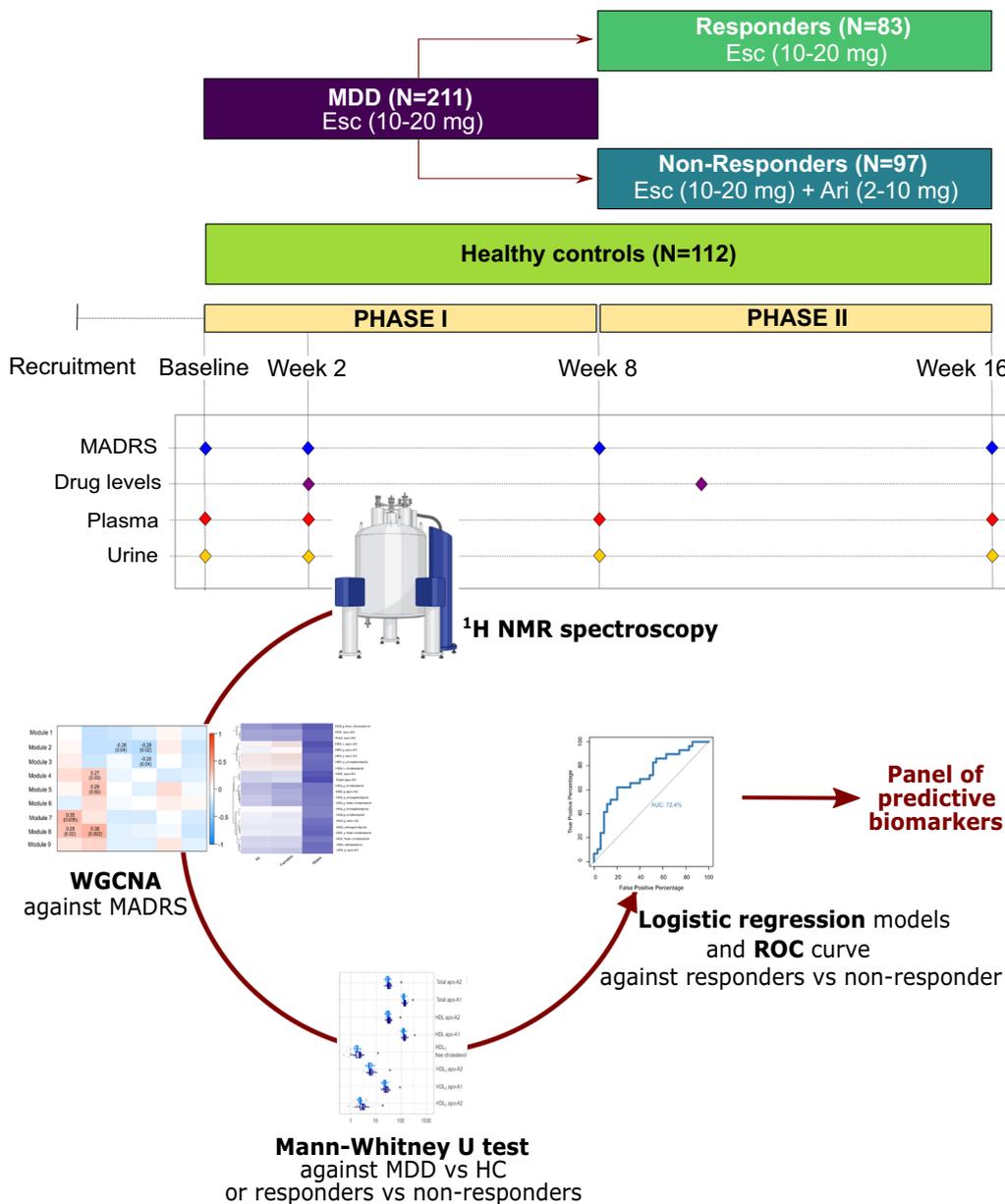


Fig. 1 Design of the CAN-BIND 1 study and analytical workflow. In phase I (week 0–8), MDD participants ($N = 211$) received escitalopram (ESC). MDD participants who responded to therapy continued on this treatment ($N = 83$), while non-responders received an augmentation with aripiprazole (ARI) during phase II (week 9–16, $N = 97$). Red diamonds show the timings of data collection for assessment of depressive symptoms (via Montgomery–Åsberg Depression Rating Scale [MADRS] at baseline, week 2, week 8, week 16), circulating drug levels for both ESC and ARI (week 2 and week 10) and blood and urine samples (only baseline data considered). Samples were analyzed by proton nuclear magnetic resonance (¹H NMR) spectroscopy (spectrometer image created by www.bioRender.com) and the resulting plasma and urine data were studied by weighted gene correlation network analysis (WGCNA), and significant features were validated by univariate statistics (Mann–Whitney U test) and by the generation of receiver operating characteristic (ROC) curves. Significant features met the following criteria: (1) exhibited WGCNA $p < 0.05$, (2) have a $p < 0.05$ in Mann–Whitney U test, (3) pass the Benjamini–Hochberg correction at an false detection rate of 20%.

contrasting relationships between high-density lipoproteins (HDLs) and LDLs and the immune system and inflammation. This information is summarized in Supplementary Fig. 2a.

Considering males and females independently, additional significant associations were highlighted (Supplementary Fig. 2b, c). In males ($N = 115$, 1 outlier removed, $\beta = 2$), module 4 ($p = 0.015$) and module 6 ($p < 0.001$) were negatively correlated with MADRS at baseline. The former was also negatively associated with BMI ($p = 0.004$) and CRP ($p = 0.036$). In females ($\beta = 5$), module 7 was positively associated with baseline MADRS score ($p = 0.016$) and with age ($p = 0.004$), BMI ($p < 0.001$), and CRP ($p = 0.014$). In both

sexes, additional significant correlations were found between modules and the covariates, as shown in Supplementary Fig. 2b, c. The correlations between each metabolite and the baseline MADRS score for all participants, as well as for the male and female participants, is depicted in the heatmap in Fig. 2a. Based on the Mann–Whitney U test after the FDR correction, baseline concentrations of HDL-cholesterol ($p = 0.017$), HDL₂-cholesterol ($p = 0.027$), LDL₂ ($p = 0.003$), LDL₂-cholesterol ($p = 0.001$), LDL₂-free cholesterol ($p = 0.003$), LDL₃-free cholesterol ($p = 0.037$), LDL₂-phospholipids ($p = 0.001$), LDL₂-apolipoprotein B ($p = 0.003$) were found to be significantly lower in males with depression compared

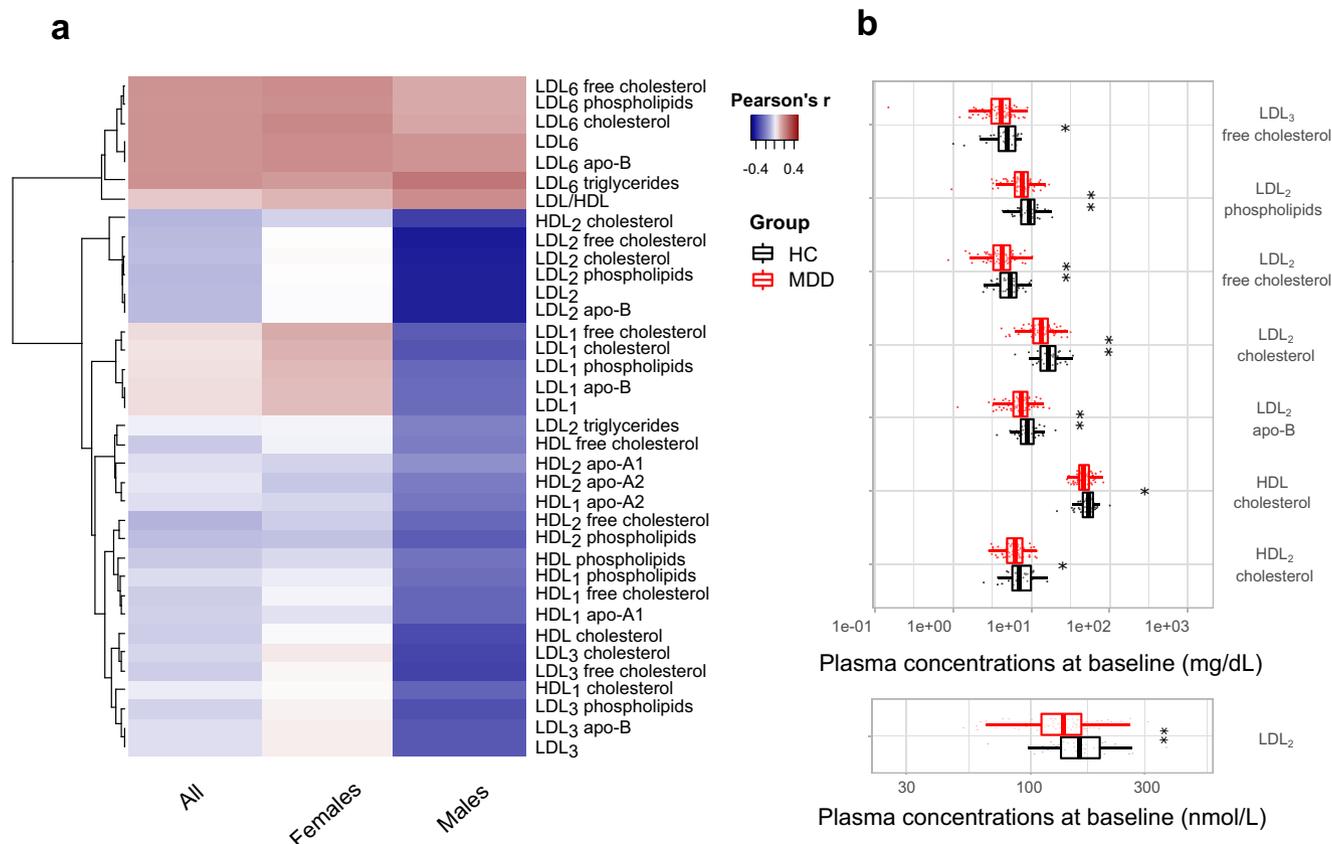


Fig. 2 Biochemical signature of major depressive disorder (MDD). **a** Heatmap depicting the correlation coefficients between the baseline concentrations of plasma metabolites in the significant weighted correlation network analysis (WGCNA) modules and baseline MADRS score for all (*N* = 310), female (*N* = 190) and male participants (*N* = 115). Positive correlations are shown in red and negative correlations are shown in blue. **b** Lipoproteins from the WGCNA analysis that are statistically significantly different between MDD and healthy control (HC) males. Boxplots show median (with first and third quartile), bars indicate 95% confidence interval of the median. Statistical significance determined by Mann-Whitney *U* test after false discovery rate correction (**p* < 0.05, ***p* < 0.01, ****p* < 0.001).

(IDL)-triglycerides (*p* = 0.050), HDL-apolipoprotein A2 (*p* = 0.003), VLDL₂-triglycerides (*p* = 0.034), VLDL₃-triglycerides (*p* = 0.032), VLDL₂-cholesterol (*p* = 0.021), VLDL₃-cholesterol (*p* = 0.023), VLDL₄-cholesterol (*p* = 0.019), VLDL₄-free cholesterol (*p* = 0.020), VLDL₂-phospholipids (*p* = 0.038), VLDL₃-phospholipids (*p* = 0.040), LDL₅-cholesterol (*p* = 0.046), LDL₅-free cholesterol (*p* = 0.032), LDL₅-phospholipids (*p* = 0.042), LDL₅-apolipoprotein B (*p* = 0.030), HDL₃-cholesterol (*p* = 0.034), HDL₄-cholesterol (*p* = 0.043), HDL₃-free cholesterol (*p* = 0.015), HDL₄-free cholesterol (*p* = 0.002), HDL₃-phospholipids (*p* = 0.026), HDL₄-phospholipids (*p* = 0.006), HDL₃-apolipoprotein A1 (*p* = 0.036), HDL₄-apolipoprotein A1 (*p* = 0.038), HDL₃-apolipoprotein A2 (*p* = 0.010), and HDL₄-apolipoprotein A2 (*p* = 0.012). Individual fold changes in these lipoprotein features from week 16 to baseline are reported in Supplementary Data 4. Baseline concentrations of LDL₅ (Rho = 0.406, *p* = 0.009), as well as its phospholipid (Rho = 0.373, *p* = 0.017), cholesterol (Rho = 0.374, *p* = 0.016), and apolipoprotein B (Rho = 0.406, *p* = 0.009) subfractions, were found to be positively correlated with dehydroaripiprazole/aripiprazole ratio at week 16 (Fig. 4c), suggesting more extensive aripiprazole metabolism in the responders. The top 10 lipoproteins (by *p* value) associated with response to the combination of escitalopram and aripiprazole were used to construct a logistic regression model. The resulting ROC curve showed that this panel of baseline metabolites (HDL₄-free cholesterol, HDL-apolipoprotein A2, apolipoprotein A2, HDL₄-phospholipids, HDL₃-apolipoprotein A2, HDL₄-apolipoprotein

A2, HDL₃-free cholesterol, VLDL₄-cholesterol, VLDL₄-free cholesterol, VLDL₂-cholesterol) at baseline could discriminate female responders from non-responders to adjunctive aripiprazole therapy at week 16 with an accuracy of 88.5%, CI: 0.7922–0.9776 (Fig. 4d). Furthermore, apolipoprotein A2 (Rho = 0.360, *p* = 0.040), HDL-apolipoprotein A2 (Rho = 0.383, *p* = 0.040), HDL₃-free cholesterol (Rho = 0.361, *p* = 0.040) and HDL₃-apolipoprotein A2 (Rho = 0.448, *p* = 0.026) were positively correlated with the percentage reduction in MADRS score occurring in phase II after controlling for age, BMI and CRP.

Discussion

With the increasing burden of MDD, its heterogenous symptomology, and inter-individual variation in treatment responses, there is a pressing need to identify clinically relevant non-invasive biomarkers for diagnosing MDD and informing personalized pharmacological strategies. In this work, we have identified sex-specific biochemical signatures of depression in the plasma lipoprotein profiles as well as predictive markers of escitalopram and aripiprazole response. For MDD, several lipoprotein species were significantly lower in the circulation of male MDD participants compared to their control equivalents. This included the overall number of LDL₂ particles and their main fractions, cholesterol and free cholesterol, apolipoprotein B, and phospholipids. In addition, HDL-cholesterol and HDL₂-cholesterol were less abundant in the circulation of MDD males. This is consistent with previous reports of a link between low cholesterol and

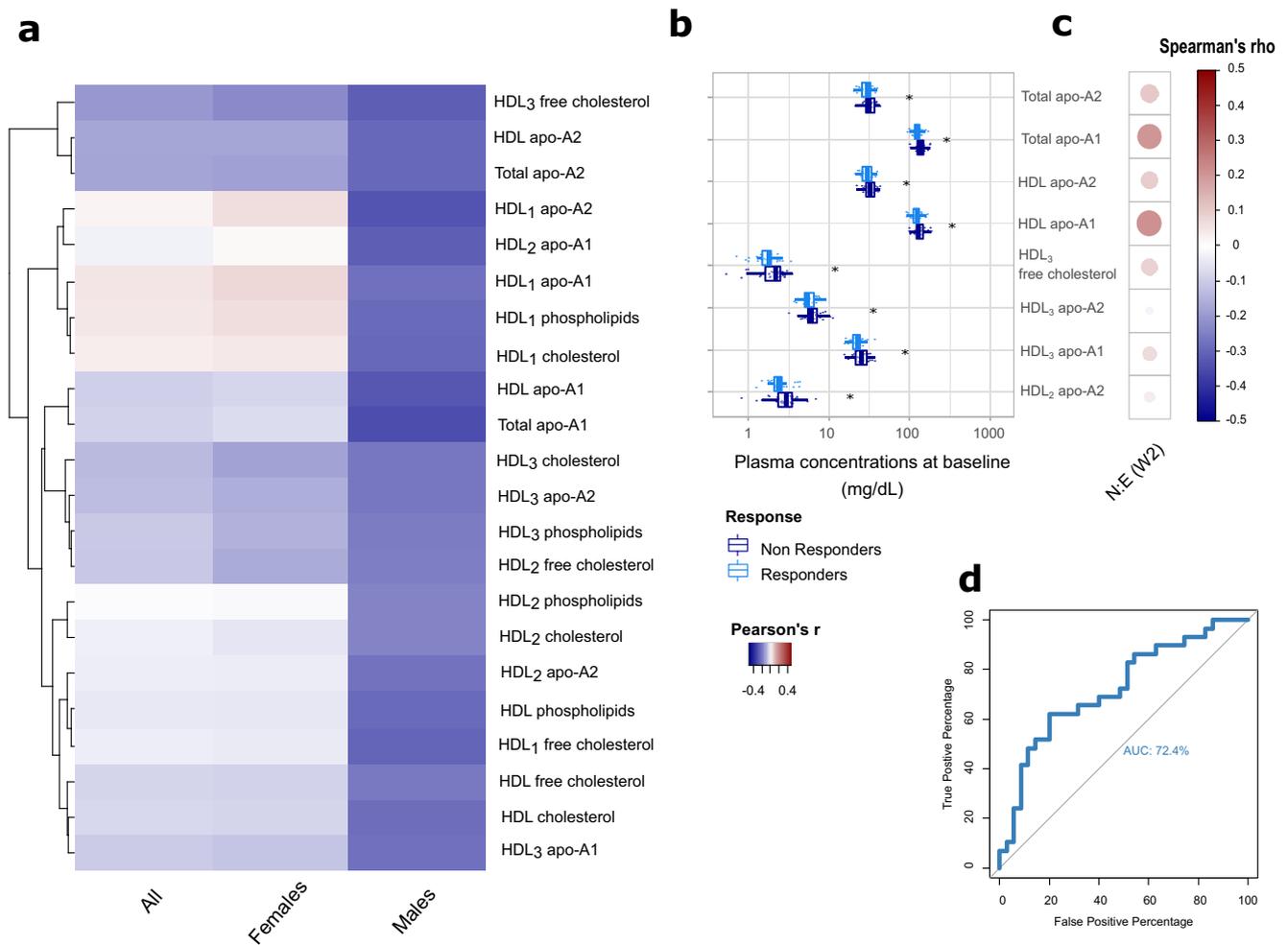


Fig. 3 Biochemical predictors of response to escitalopram monotherapy. **a** Heatmap depicting the correlation coefficients between the baseline concentrations of plasma metabolites in the WGCNA modules significantly associated with the reduction in MADRS score during phase I in all ($N = 166$), female ($N = 104$) and male participants ($N = 64$). In red, metabolites that display a positive association with MADRS reduction, in blue those that display a negative association (i.e. low baseline concentrations predictive of improvement (greater reduction)). **b** Baseline lipoproteins that are significantly different between responders and non-responders to escitalopram at week 8 in males (Mann-Whitney U test with false discovery rate; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Boxplots show median (with first and third quartile), bars indicate 95% confidence interval of the median. **c** Correlogram showing the correlation coefficients (positive in red; negative in blue; Spearman correlation) between baseline lipoproteins and circulating noescitalopram/escitalopram ratio at week 2. No significant associations were detected. **d** Receiver operating characteristic (ROC) curve evaluating the ability of apolipoprotein A1, apolipoprotein A2, HDL apolipoprotein A1, HDL apolipoprotein A2, HDL₃ free cholesterol, HDL₃ apolipoprotein A1, HDL₂ apolipoprotein A2, HDL₃ apolipoprotein A2 at baseline to discriminate between male escitalopram responders and non-responders at week 8 (AUC = 72.4%).

depression^{20–22}. Cholesterol depletion in the CNS affects central nerve terminal structure and function and modulates receptor responsiveness to serotonin. However, it remains unclear if peripheral cholesterol can impact on central serotonergic neurotransmission²⁹. The present study revealed that LDL can have opposing relationships depending on size and density. LDL₆, which is small (21.8–23.2 nm) and dense, was positively associated with MADRS, while LDL₂, which is larger (25.5–26.0 nm) and less dense, was negatively associated with MADRS. Size and density confer LDL lipoproteins with specific electrical, immunoreactive, and biochemical properties³⁰. Small dense LDL (known as ‘pattern B’), as opposed to larger less dense LDL particles (‘pattern A’), is related to cardiovascular risk³¹ due to their lower affinity for the LDL receptor, resulting in slower LDL clearance from the circulation³². Interestingly, there are strong links between MDD and coronary heart disease and the prognosis for both worsens when present together^{33,34}. This relationship was not observed in the female participants. Inherent sex differences in lipoprotein metabolism and the wide age range including

both pre- and post-menopausal females (a process known to alter lipid homeostasis³⁵) may underlie these observations.

Clear sex differences were also observed in relation to lipoprotein signatures of antidepressant response. In males, but not females, several HDL subfractions were predictive of treatment response to escitalopram. Conversely, a broad range of lipoproteins (VLDL, LDL, IDL, HDL) was predictive of response to aripiprazole augmentation in females but not males. For escitalopram, male responders had low baseline concentrations of HDL subfractions. These subfractions were also noted to be negatively correlated with MADRS in the males suggesting that individuals with severe depressive symptoms had a greater response to escitalopram. This is consistent with the observation that antidepressant drugs are less effective for patients with mild depression compared to those with moderate to severe symptoms³⁶. Low baseline concentrations of apolipoprotein A1, HDL-apolipoprotein A1, and HDL₃-free cholesterol were also predictive of escitalopram response 8 weeks later even after controlling for age, BMI, and inflammation (determined by CRP).

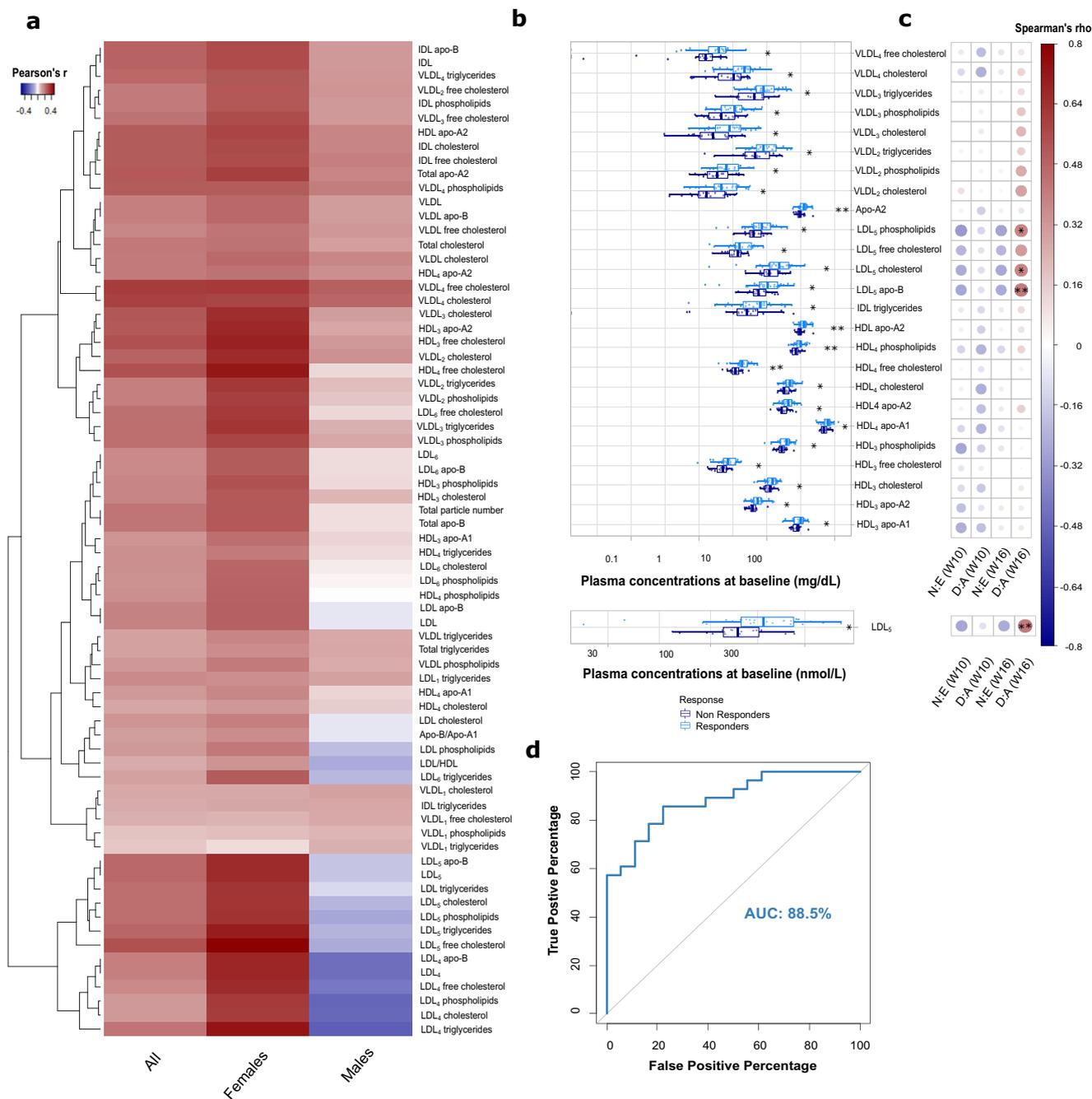


Fig. 4 Biochemical predictors of response to escitalopram and aripiprazole combined therapy. **a** Heatmap identifying the correlation coefficients between the baseline concentrations of plasma metabolites in the significant WGCNA modules and the reduction in MADRS score during phase II in all ($N = 73$), female ($N = 43$) and male participants ($N = 32$). In red, metabolites that display a positive association with MADRS reduction, in blue those that display a negative association (e.g. low baseline concentrations predictive of improvement). **b** Baseline lipoproteins that are significantly different between responders and non-responders to aripiprazole at week 16 in females (Mann-Whitney U test with false discovery rate; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Boxplots show median (with first and third quartile), bars indicate 95% confidence interval of the median. **c** Correlogram depicting the correlation coefficients (positive in red; negative in blue; Spearman correlation) between baseline lipoproteins and circulating norescitalopram/escitalopram (N:E) and dehydroaripiprazole/aripiprazole (D:A) ratios at weeks 2, 10, and 16. **d** Receiver operating characteristic (ROC) curve evaluating the ability of baseline HDL₄ free cholesterol, HDL apolipoprotein-A2, total apolipoprotein-A2, HDL₄ phospholipids, HDL₃ apolipoprotein-A2, HDL₄ apolipoprotein-A2, HDL₃ free cholesterol, VLDL₄ cholesterol, VLDL₄ free cholesterol, VLDL₂ cholesterol to discriminate between female aripiprazole responders and non-responders at week 16 (AUC = 88.5%).

These associations are likely to be independent of variations in drug metabolism since the concentrations of these metabolites were not correlated with the circulating norescitalopram/escitalopram ratio. Collectively, the panel of lipoproteins comprising apolipoproteins A1 and A2, HDL-apolipoproteins A1 and A2,

HDL₃-free cholesterol, HDL₃-apolipoprotein A1, and HDL₂₋₃-apolipoprotein A2 showed a moderate ability (72.4% accuracy) to discriminate male escitalopram responders from non-responders.

In females, apolipoprotein A2, the apolipoprotein A2 component of HDL, and the apolipoprotein A2 and free cholesterol

components of HDL₃ were predictive of response to aripiprazole augmentation independent of age, BMI, and inflammation. At baseline, prior to escitalopram and aripiprazole treatment, LDL₅ subfractions were significantly higher in responders to the combination of escitalopram and aripiprazole compared to the non-responders. These features were also positively correlated with the dehydroaripiprazole/aripiprazole ratio at week 16, suggesting that individuals with inherently higher circulating amounts of LDL₅ subfractions experience a greater breakdown of aripiprazole to its major active metabolite dehydroaripiprazole. Both the parent drug and its active metabolite are partial agonists at dopamine D₂ receptors and in this study, the dehydroaripiprazole/aripiprazole ratio was found to be positively associated with treatment response. At present it is unclear if LDL₅ can influence the distribution or metabolism of aripiprazole, however, lipophilic drugs like escitalopram ($\log P = 3.8$) and aripiprazole ($\log P = 4.9$) bind to lipoproteins with moderate affinity³⁷. Drug–lipoprotein binding alters the pharmacokinetic properties (e.g. distribution and clearance) of the drug and affects concentration–effect relationships by promoting drug uptake into tissues expressing specific lipoprotein receptors (30). Overall, a baseline lipoprotein signature comprising 10 HDL- and VLDL-related components was found to accurately predict (88.5% accuracy) female responders to aripiprazole augmentation.

With the current dataset, it was not possible to isolate the effect of food intake on the associations found. Future work should take dietary intake and fasting status into account. The lack of an independent validation dataset is also a limitation preventing these candidate biomarkers from being validated. Future research should explore this, controlling for other factors that can affect lipid status, such as smoking, menopausal status, and medication use (including hormonal medication). This is, to the best of our knowledge, the first time that individual lipoprotein subclasses (of varying size and density) have been investigated for their ability to predict antidepressant response, allowing us to identify associations that may have been masked in previous studies. Several methods can be used for lipoprotein subclass analysis including density-gradient ultracentrifugation, gel electrophoresis, ion mobility analysis, and NMR spectroscopy. Refinement of these platforms could lead to the development of a cost-effective, easy-to-use test capable of rapidly evaluating an individual's potential to respond to an antidepressant in the clinic. One example is an FDA-approved polyacrylamide tube gel electrophoresis system that is able to separate and quantify seven subclasses of LDL for clinical use³⁸.

This exploratory, hypothesis-generating study revealed the existence of a sex-dependent plasma signature of depression and treatment response to two commonly used pharmacological strategies. In this study, urinary markers were not identified, suggesting that blood may be more informative than urine for identifying metabolic predictors of response to escitalopram and aripiprazole.

Based on the identified markers, accurate predictions were made for treatment responses to escitalopram (72.4% accuracy) and aripiprazole (88.5% accuracy). In 2017, an estimated 26 million individuals were prescribed escitalopram, and 7 million were prescribed aripiprazole, with response rates of 66% and 61%, respectively^{28,39}. Therefore, accurate prediction of treatment responses has the potential to impact the quality of life of millions of people and have wide economic benefits to healthcare systems. The integration of these predictive signatures with other markers identified via different modalities by the CAN-BIND network will facilitate the construction of a multi-level, personalized signature of MDD and of antidepressant response useful for clinical practice. Efforts to optimize the treatment pipeline are essential to counter the pressing, growing burden of MDD toward accurate personalized treatments.

Methods

Participants. A total of 323 (204 females and 119 males) MDD and HC participants were recruited from six outpatient centers across Canada between August 2013 and December 2016^{26–28}: Vancouver (Djavad Mowafaghian Centre for Brain Health), Calgary (Hotchkiss Brain Institute), Toronto (2 sites: University Health Network and Centre for Addiction and Mental Health), Hamilton (St. Joseph's Healthcare Hamilton), and Kingston (Providence Care, Mental Health Services). Recruitment draws upon outpatient-referral networks, community-based advertising, and dedicated knowledge translation activities. Participants were enrolled based on meeting eligibility criteria for either group. Individuals with no history of MDD or other psychiatric disorders were classified as healthy, while individuals with a MADRS score of ≥ 24 (indicating the presence of moderate-to-severe depressive symptoms) and a diagnosis of MDD confirmed with the Mini-International Neuropsychiatric Interview version 6.1, in a current major depressive episode of 3 months or longer were classified as 'depressed'. A complete list of exclusion and inclusion criteria can be found in²⁷. Medication use was included as an exclusion criterion. Using these eligibility criteria, 112 HC participants (MADRS = 0.8 ± 1.7) and 211 MDD participants (MADRS = 29.9 ± 5.6) were identified. Participants provided written informed consent, and the study was conducted in compliance with the principles of Good Clinical Practice and relevant institutional research ethics boards. Blood and urine samples were collected as part of the Canadian Biomarker Integration Network in Depression Study-1 (CAN-BIND-1), a multi-center discovery study designed to identify predictors of MDD treatment response (ClinicalTrials.gov identifier: NCT01655706)^{27,28}.

Study design. The total duration of the study was 16 weeks. In phase I (baseline-week 8), the MDD group received escitalopram (an SSRI with a well-established antidepressant profile) at a dose of 10–20 mg daily. At the end of phase I (week 8), subjects were classified as 'responders' to escitalopram monotherapy if they achieved a 50% or greater reduction in MADRS score relative to baseline. In phase II (week 9–week 16), responders were maintained on the same escitalopram dose, while non-responders received escitalopram augmented with aripiprazole (2–10 mg). Response to the combination of escitalopram and adjunctive aripiprazole was defined as a 50% or higher reduction in MADRS score from baseline to week 16. A placebo group was not included in this study design as the focus of this study was to identify biomolecular predictors of treatment response (e.g. baseline versus week 8; baseline versus week 16) rather than the metabolic changes induced by escitalopram and aripiprazole (e.g. placebo group versus escitalopram group at week 8). Urine and blood samples were collected at baseline, week 2, week 8 and week 16 from both HC and MDD groups (Fig. 1) and matched with MADRS scores from the same visits. MADRS scores and plasma CRP concentrations (mean \pm s.d.) of responders and non-responders to the two antidepressant drugs are reported in Supplementary Table 1 for each individual timepoint and sex. In addition, concentrations of escitalopram and aripiprazole, and their respective breakdown products norescitalopram and dehydroaripiprazole were measured in plasma by the Clinical Laboratory and Diagnostic Services at CAMH (Toronto, ON) at week 2, week 10 and week 16. Standard operating procedures (SOPs) were in place to ensure maximal consistency and reproducibility across recruitment sites, and sample collection protocols have been published in²⁷. All data produced in this study are available from Brain-CODE (<https://www.braincode.ca/>)⁴⁰.

Metabolic phenotyping. Plasma and urine samples were analyzed by ¹H NMR spectroscopy at the National Phenome Centre, Imperial College London. All steps of the analysis, from sample preparation and NMR data acquisition, were performed according to published protocols⁴¹. Upon collection, urine samples were centrifuged at 1200 \times g at 4 °C for 5 min, and blood samples were centrifuged at 1600 \times g for 15 min. The supernatant was collected and stored at –80 °C. Prior to analysis, urine samples were centrifuged at 12,000 \times g at 4 °C for 5 min, and 540 μ L of the supernatant was combined with 60 μ L of potassium phosphate buffer containing the internal standard trimethylsilylpropionate (TSP). Samples were vortexed to mix and transferred to 5 mm NMR tubes. Similarly, plasma samples were centrifuged at 12,000 \times g at 4 °C for 5 min. Plasma supernatants (350 μ L) were combined with 350 μ L of disodium phosphate buffer containing TSP, vortexed, and then transferred to 5 mm NMR tubes. To monitor technical variation potentially introduced at the sample preparation stage or due to the stability of the analytical approach, quality control (QC) samples were measured with the study samples. QC samples were prepared by pooling equal parts of each study sample and analyzed every 40 samples. Experiments were performed on a Bruker Advance III HD 600 MHz spectrometer operating at 14.1 T. Three separate experiments were performed on each plasma sample: standard 1D spectra using the 1D-NOESY presat pulse sequence, relaxation edited spin-echo using the 1D-Carr-Purcell-Meiboom-Gill (CPMG) presat pulse sequence and pseudo-2D spectra using a *J*-resolved sequence. For each urine sample, 1-D-NOESY presat and *J*-resolved pulse sequences were run. Automated processing of the spectra was performed using TopSpin 3.6 (Bruker Corporation, Germany) including spectral calibration, phase and baseline correction. The Bruker B.I.-LISA method (Bruker BioSpin 08/2019 T165319) was used to quantify lipoprotein main class (very low-density lipoprotein [VLDL], intermediate-density lipoprotein [IDL], low-density lipoprotein [LDL], and high-density lipoprotein [HDL]), their particle size and density (referred to as "subclass": VLDL_{1–6}, LDL_{1–6}, HDL_{1–4} sorted according to increasing density and decreasing

size), and their building components (referred to as “subfraction”: apolipoprotein, cholesterol, free cholesterol, phospholipids, triglycerides) in the plasma samples directly from their NMR spectra⁴². A list of all 112 lipoproteins measured can be found in Supplementary Data 1. For the urine samples, the B.I.QUANT-UR method was used (Bruker BioSpin 08/2019 T165319) to quantify 50 known compounds from their NMR spectra (Supplementary Data 2). PCA models were built on the urinary metabolic profiles and the plasma lipoprotein profiles of all study samples and the QC samples. The scores plots from these models showed that the profiles of the QC samples clustered together at the center of the plots (Supplementary Fig. 4) confirming low intra-study variation in these samples. This demonstrates the reliability and reproducibility of this analytical approach. The high reproducibility of the B.I.-LISA method for quantifying lipoprotein subclasses has been previously shown⁴², supporting its potential applicability for clinical purposes.

Following the acquisition of the urine profiles, the data was normalized using a probabilistic quotient approach to account for variability in sample dilutions. The urine variables were then log-transformed to account for a left-skewed distribution. As plasma is homeostatically controlled no normalization steps were applied to the lipoprotein profiles. No data imputation was necessary as there were no missing values. Metabolic profiles were filtered to remove variables with zero variance across the samples. Hierarchical clustering, based on Euclidean distances and the average linkage method, was used to identify samples that were outliers.

To quantify serum concentrations of escitalopram and aripiprazole and of their respective major metabolites, a mix of deuterated internal standards (Cerilliant) was added to 100 μ L of each serum specimen, calibrator (Cerilliant) and QC (MassCheck Antidepressants/Neuroleptics, Level 1 and Level 2). Proteins were precipitated with 300 μ L of 9:1 acetonitrile:methanol (Sigma-Aldrich) followed by 5 min centrifugation at 9000 rpm. 25 μ L of each supernatant was diluted with 200 μ L of 0.1% formic acid (Sigma-Aldrich) and then analyzed on the LC-MS/MS platform consisting of ThermoFisher TSQ Quantum Ultra mass spectrometer coupled with ThermoFisher Surveyor LC pump and HTC PAL autosampler fitted with a Kinetex F5 2.6 μ m, 100 \times 2.1 mm column (Phenomenex). Collision energies ranged from 14 to 20. Mass transitions (M + H) were monitored in SIM mode as follows: aripiprazole (448 \rightarrow 285); dehydroaripiprazole (446 \rightarrow 285); escitalopram (325 \rightarrow 109); desmethyl-escitalopram (314 \rightarrow 109) Quantification was performed against a calibration curve ranging from 10 to 1000 ng/mL for each analyte.

Statistics and reproducibility. A schematic representation of the analytical workflow is shown in Fig. 1. Statistical analysis was carried out in R (version 4.0.1). Weighted correlation network analysis was performed with the R package WGCNA (version 1.69)⁴³. While originally developed for the study of gene co-expression analysis, the suitability of WGCNA for the analysis of metabolomic data has been demonstrated⁴⁴. Correlation networks allow for system-level investigations of the metabolites (i.e. nodes) and their relationships (i.e. edges) in complex biological systems. They represent a powerful approach for analyzing multicol-linear data while preserving the information about the correlation structure of the data. For the current WGCNA analysis, the metabolites’ correlation matrix was converted to a co-expression similarity matrix, calculated as the absolute value of the (Pearson) correlation coefficient between variable pairs (unsigned network). An unsigned network was chosen to allow for the clustering of metabolites with both positive and negative associations. The adjacency matrix was obtained by raising the similarity measure to the power β , where β is the soft threshold (weighted networks). The value of β was chosen by taking into account mean connectivity and the scale-free topology criterion, that is, the value for which the network approximates to a scale-free topology. Values of β were adjusted for each model and referred to when reporting the results of the WGCNA analysis⁴⁵. The final network was constructed on the topological overlap measure (TOM), computed as the normalized matrix product of the adjacency matrix with itself. Clusters of highly connected metabolites, referred to as “modules”, were identified by average-linkage hierarchical clustering⁴⁵, with the dendrogram cut height for module merging (mergeCutHeight) set at 0.25. The deep split was set to 2, and the minimum cluster size was set to 3 for lipoproteins and 2 for urinary metabolites. The significance of the module eigengene (or “eigenmetabolite”) with the outcome was calculated by univariate regression. Weighted correlation networks were constructed to identify clusters of highly connected metabolites explaining the variability in (i) MADRS score at baseline, (ii) the percentage MADRS reduction during phase I (between baseline and week 8), phase II (between week 8 and week 16) and during the entire trial (between baseline and week 16), as well as in the covariates of interest (i.e. age, BMI and CRP). Percentage change in MADRS score was chosen over absolute change to obtain a measure of improvement in depressive symptoms independent of the initial severity at baseline. Networks were built on all metabolic data from the entire cohort, as well as separately for males and females, to account for the known sex differences in metabolic signatures of depression and antidepressant response. Metabolites belonging to clusters exhibiting a significant association ($p < 0.05$) with the clinical variables were evaluated by a two-sided Mann-Whitney U test to identify features discriminating MDD from HC groups or treatment responders from non-responders. The p values of the WGCNA modules and of the Mann-Whitney U test comparisons were corrected for multiple comparisons using the Benjamini-Hochberg correction with a false discovery rate

(FDR) of 20%. Only the metabolites that passed the FDR correction for multiple comparisons are reported. For these metabolites, two-tailed logistic regression models were built against a binary output (i.e. MDD vs. HC or responders vs. non-responders) and the receiver operating characteristic (ROC) curves were generated to evaluate the predictive ability of the selected metabolite panels. The 95% confidence intervals (CI) of the resulting area under the curve (AUC) were calculated using the DeLong method. A two-tailed partial correlation against MADRS was also performed to correct for the confounding effect of age, BMI and CRP. Additionally, for metabolites predictive of antidepressant response, Spearman correlation was carried out to investigate the relation between the baseline concentrations of the metabolites and the circulating levels of the antidepressant drugs, calculated as the ratio between norecitalopram/escitalopram and dehydroaripiprazole/aripiprazole. Norecitalopram and dehydroaripiprazole are the metabolic products of escitalopram and aripiprazole, respectively.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data produced and used in this study are supported by The Canadian Biomarker Integration Network in Depression (CAN-BIND, <https://canbind.ca/>), which is an Integrative Discovery Program funded by the Ontario Brain Institute (OBI, <https://braininstitute.ca/>). In accordance with the Research Activity Agreements between CAN-BIND and OBI, data produced in this study must be submitted to Brain-CODE (<https://www.braincode.ca/>), an informatic platform created by the OBI to facilitate open-access of data generated from research funded by the OBI. Researchers requesting Data will provide OBI with written documentation of the proposed use of the Data in the form of an Research Ethics Board (REB) approval package which includes the full REB submission package and REB approval letter from their local Institutional REB. Should the REB determine that ethics review is not required, an exemption letter from the REB will be required. If the External Researcher is from an Institution that does not have a local REB, OBI will work with the External Researcher to identify a mutually acceptable REB for review. All documents not in English or French will require a certified translation copy. For detailed data access policy and procedure, please refer <https://www.braininstitute.ca/research-data-sharing/brain-code> or contact help@braincode.ca. Otherwise, data points underlying Figs. 2b, 3b, and 4b are presented in Supplementary Data 5–7.

Received: 23 October 2020; Accepted: 7 June 2021;

Published online: 22 July 2021

References

- World Health Organization. Depression and other common mental disorders: global health estimates. World Health Organization 1–24 (2017) CC BY-NC-SA 3.0 IGO.
- Goldberg, D. The heterogeneity of ‘major depression’. *World Psychiatry* **10**, 226–228 (2011).
- Young, J. J. et al. Is there progress? An overview of selecting biomarker candidates for major depressive disorder. *Front. Psychiatry* **7**, 1–15 (2016).
- Gaynes, B. N. et al. What did STAR*D teach us? Results from a large-scale, practical, clinical trial for patients with depression. *Psychiatr. Serv.* **60**, 1439–1445 (2014).
- Fava, M. Diagnosis and definition of treatment-resistant depression. *Biol. Psychiatry* **53**, 649–659 (2003).
- Labermaier, C., Masana, M. & Müller, M. B. Biomarkers predicting antidepressant treatment response: how can we advance the field? *Dis. Markers* **35**, 23–31 (2013).
- Nicholson, J. K., Lindon, J. C. & Holmes, E. ‘Metabonomics’: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **29**, 1181–1189 (1999).
- Martins-de-Souza, D. Proteomics and metabolomics in psychiatry. *Adv. Biol. Psychiatry* **29**, I-VII (2014).
- Quinones, M. P. & Kaddurah-Daouk, R. Metabolomics tools for identifying biomarkers for neuropsychiatric diseases. *Neurobiol. Dis.* **35**, 165–176 (2009).
- Kaddurah-Daouk, R. et al. Pretreatment metabolite as a predictor of response to sertraline or placebo in depressed outpatients: a proof of concept. *Transl. Psychiatry* **1**, e26 (2011).
- Kaddurah-Daouk, R. et al. Pharmacometabolomic mapping of early biochemical changes induced by sertraline and placebo. *Transl. Psychiatry* **3**, e223 (2013).
- Czys, A. H. et al. Can targeted metabolomics predict depression recovery? Results from the CO-MED trial. *Transl. Psychiatry* **9**, 11 (2019).

13. Bhattacharyya, S. et al. Metabolomic signature of exposure and response to citalopram/escitalopram in depressed outpatients. *Transl. Psychiatry* **9**, 173 (2019).
14. Rotroff, D. M. et al. Metabolomic signatures of drug response phenotypes for ketamine and esketamine in subjects with refractory major depressive disorder: New mechanistic insights for rapid acting antidepressants. *Transl. Psychiatry* **6**, e894 (2016).
15. Y., J. et al. Glycine and a glycine dehydrogenase (GLDC) SNP as citalopram/escitalopram response biomarkers in depression: pharmacometabolomics-informed pharmacogenomics. *Clin. Pharmacol. Ther.* **89**, 97–104 (2011).
16. Kornstein, S. G. et al. Gender differences in treatment response to sertraline versus imipramine in chronic depression. *Am. J. Psychiatry* **157**, 1445–1452 (2000).
17. Datz, F. L., Christian, P. E. & Moore, J. Gender-related differences in gastric emptying. *J. Nucl. Med.* **28**, 1204–1207 (1987).
18. Greenblatt, D. J. et al. Trazodone kinetics: effect of age, gender, and obesity. *Clin. Pharm. Ther.* **42**, 193–200 (1987).
19. Yonkers, K. A., Kando, J. C., Cole, J. O. & Blumenthal, S. Gender differences in pharmacokinetics and pharmacodynamics of psychotropic medication. *Am. J. Psychiatry* **149**, 587–595 (1992).
20. Terao, T. et al. Low serum cholesterol levels and depressive state in human dock visitors. *Acta Psychiatr. Scand.* **101**, 231–234 (2000).
21. Horsten, M., Wamala, S. P., Vingerhoets, A. & Orth-Gomer, K. Depressive symptoms, social support, and lipid profile in healthy middle-aged women. *Psychosom. Med.* **59**, 521–528 (1997).
22. Rabe-Jabłońska, J. & Poprawska, I. Levels of serum total cholesterol and LDL-cholesterol in patients with major depression in acute period and remission. *Med. Sci. Monit.* **6**, 539–547 (2000).
23. Gupta, A., Petkar, S. B., Jadhav, A. A. & Dubey, V. Study of lipid derangement in psychiatric disorder. *Ind. Med. Gaz.* **147**, 253–256 (2013).
24. Olusi, S. O. & Fido, A. A. Serum lipid concentrations in patients with major depressive disorder. *Biol. Psychiatry* **40**, 1128–1131 (1996).
25. 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 (2019).
26. Rotzinger, S. & Kennedy, S. The Canadian Biomarker Integration Network for Depression (CAN-BIND): looking deeper into major depressive disorder. *Mood Anxiety Disord. Rounds* **2**, (2013).
27. Lam, R. W. et al. Discovering biomarkers for antidepressant response: Protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. *BMC Psychiatry* **16**, 1–13 (2016).
28. Kennedy, S. H. et al. Symptomatic and functional outcomes and early prediction of response to escitalopram monotherapy and sequential adjunctive aripiprazole therapy in patients with major depressive disorder: a can-bind-1 report. *J. Clin. Psychiatry* **80**, 18m12202 (2019).
29. Persons, J. E. & Fiedorowicz, J. G. Depression and serum low-density lipoprotein: a systematic review and meta-analysis. *J. Affect. Disord.* **206**, 55–67 (2016).
30. Chapman, M. J. et al. Further resolution of the low density lipoprotein spectrum in normal human plasma: physicochemical characteristics of discrete subspecies separated by density gradient ultracentrifugation. *J. Lipid Res.* **29**, 442–458 (1988).
31. Austin, M. A. et al. Low-density lipoprotein subclass patterns and risk of myocardial infarction. *JAMA* **260**, 1917–1921 (1988).
32. Nigon, F., Lesnik, P., Rouis, M. & Chapman, M. J. Discrete subspecies of human low density lipoproteins are heterogeneous in their interaction with the cellular LDL receptor. *J. Lipid Res.* **32**, 1741–1753 (1991).
33. Musselman, D. L., Evans, D. L. & Nemeroff, C. B. The relationship of depression to cardiovascular disease: epidemiology, biology, and treatment. *Arch. Gen. Psychiatry* **55**, 580–592 (1998).
34. Rosengren, A. et al. Association of psychosocial risk factors with risk of acute myocardial infarction in 11 119 cases and 13 648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* **364**, 953–962 (2004).
35. Reddy Kilim, S. & Rao Chandala, S. A comparative study of lipid profile and oestradiol in pre- and post-menopausal women. *J. Clin. Diagn. Res.* **7**, 1596–1598 (2013).
36. Fournier, J. C. et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* **303**, 47–53 (2010).
37. Wasan, K. M., Brocks, D. R., Lee, S. D., Sachs-Barrable, K. & Thornton, S. J. Impact of lipoproteins on the biological activity and disposition of hydrophobic drugs: Implications for drug discovery. *Nat. Rev. Drug Discov.* **7**, 84–99 (2008).
38. Hoefner, D. M. et al. Development of a rapid, quantitative method for LDL subfractionation with use of the quantimetrix lipoprint LDL system. *Clin. Chem.* **47**, 266–274 (2001).
39. Kennedy, S. H., Andersen, H. F. & Lam, R. W. Efficacy of escitalopram in the treatment of major depressive disorder compared with conventional selective serotonin reuptake inhibitors and venlafaxine XR: a meta-analysis. *J. Psychiatry Neurosci.* **31**, 122–131 (2006).
40. Ontario Brain Institute. *Brain-CODE*. <https://www.braincode.ca/>, Accessed on 21/11/2020.
41. Dona, A. C. et al. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal. Chem.* **86**, 9887–9894 (2014).
42. Jiménez, B. et al. Quantitative lipoprotein subclass and low molecular weight metabolite analysis in human serum and plasma by ¹H NMR spectroscopy in a multilaboratory trial. *Anal. Chem.* **90**, 11962–11971 (2018).
43. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
44. Pei, G., Chen, L. & Zhang, W. WGCNA Application to proteomic and metabolomic data analysis. *Methods Enzymol.* **585**, 135–158 (2017).
45. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, Article17 (2005).

Acknowledgements

CAN-BIND is an Integrated Discovery Program carried out in partnership with, and financial support from, the Ontario Brain Institute, an independent non-profit corporation, funded partially by the Ontario government. The opinions, results and conclusions are those of the authors and no endorsement by the Ontario Brain Institute is intended or should be inferred. Additional funding is provided by the Canadian Institutes of Health Research (CIHR), Lundbeck, Bristol-Myers Squibb, Pfizer, and Servier. Funding and/or in kind support is also provided by the investigators' universities and academic institutions. All study medications are independently purchased at wholesale market values. This work was also supported by the Medical Research Council and National Institute for Health Research [grant number MC_PC_12025] and infrastructure support was provided by the National Institute for Health Research (NIHR) Imperial Biomedical Research Centre (BRC). G.C. is supported by the MRC (grant number MR/N014103/1).

Author contributions

G.C. performed the analysis and drafted the manuscript under the direct supervision of J.A.F. and J.R.S. B.N.F., R.W.L., G.M.M., R.V.M., S.H.K., S.R., D.J.M., G.T., J.A.F. are the clinical site leads and/or principal investigators responsible for the conception and design of the study. They were involved in participant recruitment and data collection. All authors participated in the discussion regarding the interpretation of the data and provided comments on the manuscript.

Competing interests

R.W.L. has received honoraria for ad hoc speaking or advising/consulting, or received research funds, from: Allergan, Asia-Pacific Economic Cooperation, BC Leading Edge Foundation, Canadian Institutes of Health Research, Canadian Network for Mood and Anxiety Treatments, Canadian Psychiatric Association, Hansoh, Healthy Minds Canada, Janssen, Lundbeck, Lundbeck Institute, Michael Smith Foundation for Health Research, MITACS, Myriad Neuroscience, Ontario Brain Institute, Otsuka, Pfizer, St. Jude Medical, University Health Network Foundation, and VGH-UBCH Foundation. R.V.M. has received consulting and speaking honoraria from Allergan, Janssen, KYE, Lundbeck, Otsuka, and Sunovion, and research grants from CAN-BIND, CIHR, Janssen, Lallemand, Lundbeck, Nubiyota, OBI and OMHF. B.N.F. received a research grant from Pfizer. SHK has received research funding or honoraria from the following sources: Abbott, Alkermes, Allergan, Boehringer-Ingelheim, Brain Canada, Canadian Institutes of Health Research (CIHR), Janssen, Lundbeck, Lundbeck Institute, Ontario Mental Health Foundation (OMHF), Ontario Brain Institute, Ontario Research Fund (ORF), Otsuka, Pfizer, Sanofi, Servier, St. Jude Medical, Sun Pharmaceuticals, and Sunovion. The remaining authors declare that there is no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-021-02421-6>.

Correspondence and requests for materials should be addressed to J.R.S.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available. Primary handling editors: Melanie Bahlo, George Inglis.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021