Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Urinary Untargeted Metabolomic Profiling in First-Episode Drug-Naïve Patients with Major Depressive Disorder: A Case-Control Study

Xialong Cheng
Affiliated Psychological Hospital of Anhui Medical University
Xulai Zhang
Affiliated Psychological Hospital of Anhui Medical University
Long Chen
Affiliated Psychological Hospital of Anhui Medical University
Anzhen Wang
Affiliated Psychological Hospital of Anhui Medical University
Min Pan
minpan2024@163.com
Affiliated Psychological Hospital of Anhui Medical University

Research Article

Keywords: major depressive disorder, first-episode drug-naïve, urine, untargeted metabolomics, biomarker, case-control study

Posted Date: October 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10907550/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Urinary Untargeted Metabolomic Profiling in First-Episode Drug-Naïve Patients with Major Depressive Disorder: A Case-Control Study

Xialong Cheng[a,b],Xulai Zhang[a,b],Long Chen[a,b],Anzhen Wang[a,b],Min Pan[a,b*]

[a]Affiliated Psychological Hospital of Anhui Medical University, Hefei, China.

[b]Hefei Fourth People's Hospital, Hefei, China.

***Corresponding author:**

Min Pan,Affiliated Psychological Hospital of Anhui Medical University,Hefei Fourth People's Hospital,316 Huangshan Road, Shushan District, Hefei City, Anhui Province, 230022, China.

Email: minpan2024@163.com

Abstract

Objective: To investigate the urinary metabolomic profiling of first-episode drug-naïve patients with major depressive disorder (MDD) using ultra-high-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry (UHPLC-Q-TOF-MS) untargeted metabolomics, so as to identify potential differential metabolites and functional pathways, and provide experimental evidence for the research on objective diagnostic biomarkers of depression.

Methods: A case-control study was performed. A total of 55 first-episode drug-naïve MDD patients and 35 age- and sex-matched healthy controls were enrolled. Fasting early-morning urine samples were collected for UHPLC-Q-TOF-MS untargeted metabolomic detection. The k-nearest-neighbor (KNN) algorithm was applied for missing-value imputation. The Mann-Whitney U test combined with Benjamini-Hochberg (BH) multiple-testing correction was used to screen differential metabolites. Principal component analysis (PCA) and orthogonal projections to latent structures-discriminant analysis (OPLS-DA) were adopted to explore the overall separation trend of metabolic profiles between the two groups. Chemical and functional classification of metabolites was conducted based on the ClassyFire system, and binomial tests were performed to evaluate the enrichment rate of differential metabolites for each category. Post-hoc statistical power analysis was also carried out.

Results: In total, 5932 metabolites were identified. The relative standard deviation (RSD) of peak areas for all metabolites in quality-control samples was below 30%, indicating satisfactory data quality. Univariate analysis showed that 179 metabolites had raw $P < 0.05$, whereas **no metabolite reached the statistical significance threshold of adjusted* $P_adj < 0.05$ *after BH multiple-testing correction**. PCA revealed substantial overlap of urinary metabolomic profiles between the two groups without obvious group separation. The OPLS-DA model yielded a cross-validated predictive Q^2 value of −1.185, suggesting model over-fitting and lack of genuine discriminative power. Sixty-six candidate exploratory differential metabolites (39 up-regulated and 27 down-regulated) were obtained according to exploratory screening criteria; among them, 18.2 % represented drug-like compounds with false-positive annotations and 3.0 % were smoking-related confounding metabolites. Functional classification via ClassyFire showed relatively higher differential rates for phenylpropanoids and polyketides, benzenoids, and organic acids and derivatives. Nevertheless, binomial tests demonstrated

that none of the chemical categories exceeded the 5 % random expectation level. Power analysis indicated that the present sample size only provided 1.61 % statistical power for detecting a medium effect size (Cohen's d = 0.5), and a total sample size of at least 473 subjects would be required to detect such a medium effect. Significant inter-group differences in smoking rate, alcohol-drinking rate and education level were observed between patients and controls, which served as important confounding factors.

Conclusion: No robust urinary differential metabolites survived multiple-testing correction in this cohort of strictly medication-naïve first-episode MDD patients, and multivariate models failed to effectively discriminate cases from controls. High inter-individual variability of the urinary metabolome, insufficient statistical power, false-positive metabolite annotations, and confounders such as smoking jointly limited the findings of this study. Alterations in phenylpropanoid, benzenoid and organic-acid metabolic pathways suggested by exploratory analyses can only serve as hypothesis-generating clues for future research, which need further validation by large-scale, multi-center targeted metabolomic studies with rigorous confounder control. This study provides important negative evidence for urinary metabolomic biomarker research in depression, and highlights the necessity of rigorous statistical correction, model validation, metabolite annotation quality control and adequate sample size in metabolomic biomarker investigations.

Keywords: major depressive disorder; first-episode drug-naïve; urine; untargeted metabolomics; biomarker; case-control study

Introduction

Major Depressive Disorder (MDD) is a common psychiatric disorder characterized by persistent depressed mood, diminished interest, and anhedonia as core features, with high prevalence, high recurrence rate, high disability rate, and high suicide risk. According to the World Health Organization, approximately 280 million people worldwide suffer from depression, which has become one of the leading causes of the global disease burden [1]. In China, the first national mental health survey conducted by Huang et al. showed that the lifetime prevalence of mood disorders reached 7.4%, among which MDD was the most common subtype, but only 9.5% of patients received adequate treatment [2]. However, the pathogenesis of depression has not yet been fully elucidated. Current clinical diagnosis still relies primarily on history taking and psychiatric examination, lacking objective biological diagnostic markers, which leads to difficulties in early identification, high rates of missed diagnosis and misdiagnosis, and delayed treatment decisions. Therefore, identifying reliable objective biological markers is of great clinical significance for the early diagnosis, subtype assessment, and treatment response prediction of depression.

The pathological mechanisms of depression are complex, involving multiple levels such as neurotransmitter imbalance, neuroendocrine dysregulation, decreased neuroplasticity, chronic inflammation activation, and mitochondrial dysfunction [3]. Although the classic monoamine hypothesis has facilitated the development of antidepressants such as fluoxetine and venlafaxine, it cannot explain clinical phenomena such as the delayed onset of antidepressant action and poor response in some patients, suggesting that the biological basis of depression extends far beyond abnormalities in the monoamine neurotransmitter system. In recent years, the rapid development of systems biology has provided new perspectives for deciphering the complex pathological network of depression. Among these approaches, metabolomics, which quantitatively analyzes the dynamic changes of all small-molecule metabolites (typically with molecular weights <1500 Da) in organisms, can directly reflect the terminal phenotype of the organism under gene-environment interactions and is considered the omics technology closest to the disease phenotype [4]. Compared with genomics and proteomics, metabolomics offers advantages such as high sensitivity, easy sample acquisition, and relatively low

detection costs, and has become an important tool for biomarker research in psychiatric disorders.

In metabolomics research, Ultra-High Performance Liquid Chromatography coupled with Quadrupole Time-of-Flight Mass Spectrometry (UHPLC-Q-TOF-MS) has become the mainstream platform for untargeted metabolomics detection due to its high resolution, high sensitivity, and high mass accuracy [5]. This technology can simultaneously detect thousands of metabolites in a single injection, covering multiple metabolic pathways such as amino acids, lipids, organic acids, and nucleotides, providing technical support for comprehensively mapping the metabolic disturbance profile of depression. Previous metabolomics studies on depression have mostly used serum or plasma samples and found abnormal alterations in amino acid metabolism, lipid metabolism, energy metabolism, and gut microbiota-related metabolites in patients with depression [6-8]. A systematic review incorporating multiple clinical studies indicated that kynurenine and acylcarnitines are the most consistent metabolic abnormalities in depression [8]. However, blood sample collection is invasive and is significantly influenced by factors such as diet, stress, and circadian rhythm, limiting its application in large-scale population screening and long-term dynamic monitoring.

As a non-invasive and abundantly obtainable biological sample, urine contains metabolites that reflect the metabolic end products after renal filtration and excretion. Urine not only contains endogenous metabolites but is also rich in gut microbiota-derived metabolites and exogenous compounds, providing a unique perspective on the overall metabolic state of the organism [9]. Compared with blood, urine has low protein content, simple sample pretreatment, a wider dynamic range of metabolites, and non-invasive collection with good patient compliance, making it particularly suitable for chronic diseases such as psychiatric disorders that require repeated sampling. In recent years, studies have explored the urinary metabolomic profile characteristics of patients with depression: Zheng et al. used NMR and GC-MS platforms to identify a urinary biomarker panel comprising 5 metabolites including malonic acid, formic acid, and N-methylnicotinamide, achieving an AUC of 0.89 in the validation set [10]; the same team further identified sex-specific urinary biomarker panels [11]. However, these studies generally suffer from small sample sizes, lack of strict control over medication status, and

absence of independent validation. Moreover, the consistency of results across studies is poor, and no consensus urinary metabolic biomarker panel has yet been established.

Notably, most previous metabolomics studies on depression enrolled patients who were receiving antidepressant treatment, and antidepressants themselves can significantly affect the metabolic profile. A quantitative analysis based on 157 animal studies showed that antidepressants can systematically alter multiple metabolic pathways including tryptophan, dopamine, γ-aminobutyric acid, purine, and lipid metabolism [12], making it difficult to distinguish disease-related metabolic changes from metabolic effects induced by drug intervention. First-episode drug-naïve patients with depression, having never received psychopharmacological intervention, represent an ideal population for studying the endogenous metabolic characteristics of depression [13], allowing maximal exclusion of medication confounders and a more accurate reflection of the metabolic disturbances inherent to the disease itself. However, urinary metabolomics studies in first-episode drug-naïve patients with depression remain very limited, and most have focused on a small number of candidate metabolites, lacking systematic metabolomic profiling based on untargeted metabolomics and integrated interpretation at the functional pathway level.

Based on the above background, this study employed UHPLC-Q-TOF-MS untargeted metabolomics to profile urine samples from 55 first-episode drug-naïve patients with depression and 35 age- and sex-matched healthy controls. Missing values were imputed using K-Nearest Neighbors (KNN) imputation. Differential metabolites were screened using the Mann-Whitney U test combined with Benjamini-Hochberg (BH) multiple testing correction. Principal Component Analysis (PCA) and Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) were applied to observe the overall separation trend of metabolic profiles between the two groups. Metabolites were functionally classified based on the ClassyFire chemical classification system, and the differential rate of each category was systematically analyzed with binomial tests against the random expectation. This study aimed to characterize the urinary metabolomic profile of first-episode drug-naïve patients with depression, identify potential differential metabolites and functional pathways, and provide experimental evidence for research on objective diagnostic biomarkers of depression.

Materials and Methods

1 Study Design and Ethical Review

This study employed a case-control study design.First-episode drug-naïve patients with depression were consecutively enrolled from the psychiatric outpatient clinic of Hefei Fourth People's Hospital between January 2023 and December 2025 as the depression group. Age- and sex-matched healthy volunteers were concurrently recruited as the healthy control group. The study protocol was reviewed and approved by the Ethics Committee of Hefei Fourth People's Hospital (Ethics Approval No. HFSY-IRB-YJ-KYXM-CL[2024-064-001]). All subjects provided written informed consent.

2 Study Subjects

Inclusion criteria for the depression group: (1) meeting the diagnostic criteria for Major Depressive Disorder according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [14]; (2) age 18–60 years; (3) no prior treatment with any antidepressants, antipsychotics, mood stabilizers, or benzodiazepines before enrollment; (4) 17-item Hamilton Depression Rating Scale (HAMD-17) score ≥14 [15]. Among the 55 enrolled patients, 23 (41.8%) were experiencing their first lifetime depressive episode, while 32 (58.2%) had prior depressive episodes but had never received pharmacological treatment.

Exclusion criteria (common to both groups): (1) comorbidity with other psychiatric disorders (bipolar disorder, schizophrenia, substance use disorders, etc.); (2) comorbidity with severe somatic diseases (cardiac, hepatic, renal, endocrine, neurological diseases, and malignant tumors); (3) use of antibiotics, glucocorticoids, or other metabolism-affecting drugs within the past 3 months; (4) pregnant or lactating women; (5) history of severe infection, surgery, or trauma within the past month; (6) history of alcoholism or drug abuse. The healthy control group was confirmed by psychiatric interview to have no personal or family history of psychiatric disorders, with HAMD-17 scores <7. A total of 90 subjects were ultimately enrolled, including 55 in the depression group and 35 in the healthy control group.

3 Clinical Assessment

Clinical assessments were independently conducted by two trained attending psychiatrists, with an inter-rater reliability intraclass correlation coefficient (ICC) >0.85. Assessment tools included: demographic data (age, sex, height, weight, body mass index, ethnicity, education level, marital status, occupation, family income, smoking and alcohol drinking history); clinical characteristics (total disease duration, number of previous episodes, duration of current episode, presence of psychotic symptoms, history of suicidal behavior, family history); and scale assessments including HAMD-17 [15], Hamilton Anxiety Rating Scale (HAMA). Smoking status was recorded as current smoker vs. non-smoker based on self-report, and this variable was used for confounder assessment.

4 Biological Sample Collection and Processing

Urine samples were collected from all subjects in the early morning on the day of enrollment under fasting conditions (food deprivation ≥8 h, water deprivation ≥2 h). Ten milliliters of midstream first-morning urine was collected in enzyme-free centrifuge tubes and centrifuged at 3000 × g for 10 min at 4 °C. The supernatant was aliquoted into 500 μL tubes and immediately stored at −80 °C until analysis. Sample pretreatment was completed within 24 h after collection to avoid repeated freeze-thaw cycles (freeze-thaw cycles ≤1). Protein precipitation with methanol was used for sample pretreatment: 400 μL of urine sample was mixed with 1200 μL of pre-chilled methanol (methanol:urine = 3:1, v/v), vortexed for 30 s, incubated at 4 °C for 10 min, and centrifuged at 14,000 × g for 15 min at 4 °C. A 1.2 mL aliquot of the supernatant was evaporated to dryness under low temperature using a vacuum concentrator. The residue was reconstituted in 100 μL of acetonitrile-water (1:1, v/v) containing an internal standard (levomepromazine, 2 μg/mL), vortexed for 30 s, and centrifuged at 14,000 × g for 15 min at 4 °C. The supernatant was collected for injection analysis.

Of note, urinary creatinine correction or specific gravity correction was not performed in this study due to the absence of concurrent creatinine measurements. This is acknowledged as a methodological limitation, as individual differences in urine concentration may affect quantitative results. Future studies should incorporate creatinine or specific gravity normalization.

Quality control (QC) samples: Pooled QC samples were prepared by mixing equal volumes from all subject samples. One QC sample was inserted after every 10 experimental samples to monitor instrument stability and data quality [25,26]. A total of 15 QC samples were prepared. Metabolites with a relative standard deviation (RSD) of peak area <30% in QC samples were included in subsequent analyses [26]. In this study, all 5932 identified metabolites met the RSD <30% criterion (mean RSD = 0.97%, median RSD = 0.81%, max RSD = 6.33%), indicating excellent instrument stability. (See Supplementary Table S1 for complete QC RSD summary; Supplementary Figures S1–S2 for QC PCA and RSD distribution.)

5 UHPLC-Q-TOF-MS Detection

Chromatographic conditions: A Waters ACQUITY UPLC HSS T3 column (100 mm × 2.1 mm, 1.8 μm) was used, with a column temperature of 40 °C, an injection volume of 2 μL, and a flow rate of 0.3 mL/min. Mobile phase A was ultrapure water containing 0.1% formic acid, and mobile phase B was acetonitrile containing 0.1% formic acid. The gradient elution program was: 0–2 min, 5% B; 2–10 min, 5%–95% B; 10–12 min, 95% B; 12–12.1 min, 95%–5% B; 12.1–15 min, 5% B for equilibration. These chromatographic conditions were based on established standardized protocols for urinary metabolomics [16,25].

Mass spectrometry conditions: A Waters Xevo G2-XS Q-TOF mass spectrometer was used, and data were acquired in both electrospray ionization positive (ESI+) and negative (ESI−) modes. Capillary voltage: 3.0 kV for positive mode and 2.5 kV for negative mode; cone voltage: 40 V; ion source temperature: 120 °C; desolvation gas temperature: 450 °C; desolvation gas flow rate: 800 L/h; cone gas flow rate: 50 L/h. The mass scan range was m/z 50–1200 with a scan time of 0.2 s. Leucine-enkephalin (m/z 556.2771 [M+H]+ / 554.2615 [M−H]−) was used for real-time mass calibration (LockSpray), with a mass error <5 ppm. A total of 3230 metabolites were identified in positive mode and 2702 in negative mode.

6 Metabolite Identification and Data Preprocessing

Peak extraction and identification: Progenesis QI software (Waters, v2.4) was used for peak alignment, peak extraction, deconvolution, and normalization. Metabolite

identification was based on accurate molecular mass (mass error <5 ppm) and MS/MS fragment spectrum matching, with searches against the HMDB (Human Metabolome Database), KEGG (Kyoto Encyclopedia of Genes and Genomes), PubChem, METLIN, and LipidMaps databases, confirmed by retention time and isotope distribution information. A total of 5932 metabolites were identified.

Metabolite identification confidence was classified according to the Metabolomics Standards Initiative (MSI) framework [29]: Level 1 (confirmed structure by reference standard, n = 1127, 19.0%); Level 2 (putatively annotated by MS/MS spectral library match, n = 1392, 23.5%); Level 3 (putatively characterized by exact mass and chemical class, n = 2216, 37.4%); and Level 4 (unknown/unclassified, n = 1197, 20.2%). Thus, 42.5% of metabolites had MS/MS-based identification (Level 1+2), while 57.5% lacked MS/MS confirmation (Level 3+4). The chemical identities of candidate metabolites, particularly those at Level 3 and 4, require further validation through reference standard comparison and MS/MS fragmentation spectrum matching. (See Supplementary Table S2 for complete MSI level distribution by ion mode.)

Missing value processing: Zero peak area values were treated as missing values. K-Nearest Neighbors Imputation (KNNI) was used for missing value imputation, with parameters set to k=5 and distance-weighted, weighted averaging of observations from nearest neighbors based on Euclidean distance [17,18]. Before imputation, the missing rate of the data matrix was 4.66% (24,892/533,880 data points), involving 2,198 metabolites; no remaining missing values were present after imputation. (See Supplementary Figure S3 for missing value pattern before KNN imputation.)

Data transformation and normalization: The imputed data were log10-transformed to reduce heteroscedasticity and skewed distribution effects, followed by Pareto scaling, in which each metabolite was mean-centered and divided by the square root of its standard deviation, so that the contribution of each metabolite to the model was proportional to its variance while preventing high-abundance metabolites from dominating the model [23].

7 Statistical Analysis

Comparison of demographic and clinical data: Normally distributed continuous variables were expressed as mean ± standard deviation ($\bar{x} \pm s$), and between-group

comparisons were performed using the independent samples t-test. Non-normally distributed continuous variables were expressed as median (interquartile range) [M(P25, P75)] and analyzed using the Mann-Whitney U test [19]. Categorical variables were expressed as number (percentage) and analyzed using the χ^2 test or Fisher's exact test.

Between-group differential analysis of metabolites: For the relative peak area of each metabolite, the Mann-Whitney U test was used to compare differences between the depression group and the healthy control group [19], and the U statistic and p-value were calculated. The Benjamini-Hochberg (BH) method was used for multiple testing correction to control the false discovery rate (FDR), and adjusted p-values (p_adj) were calculated [20]. The fold change (FC) was defined as the ratio of the mean value in the depression group to that in the healthy control group (back-transformed from log10 scale), and the log2-transformed value (log2FC) was calculated.

Multivariate statistical analysis: (1) Principal Component Analysis (PCA) [21]: an unsupervised analysis to observe the overall separation trend and outliers of metabolic profiles between the two groups, and to calculate the proportion of variance explained by each principal component; (2) Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) [22]: a supervised analysis to maximize between-group differences while filtering orthogonal variation unrelated to classification, with one predictive component and one orthogonal component. Model quality was assessed by R^2Y (explained variance) and Q^2 (cross-validated predictive ability), and Q^2 was calculated using 7-fold cross-validation. Permutation testing (n = 200) was performed to validate model significance. The Variable Importance in Projection (VIP) value was calculated for each metabolite; however, as the OPLS-DA model showed a negative Q^2 (see Results), VIP values should be interpreted with caution and were used only for exploratory screening, not for statistical inference.

Screening criteria for differential metabolites: Metabolites simultaneously satisfying the following conditions were defined as exploratory differential metabolites: (1) raw p-value <0.05 from the Mann-Whitney U test; (2) OPLS-DA model VIP >1 (exploratory only, given model limitations); (3) $|\log2FC| > \log2(1.5)$ (FC >1.5 or FC <2/3). BH-adjusted p_adj <0.05 was used as the strict significance threshold.

Functional classification and differential rate analysis: Based on the ClassyFire chemical classification system [24], metabolites were functionally classified into three levels: Super Class, Class, and Sub Class. The differential rate for each category at each classification level was calculated as: differential rate (%) = (number of differential metabolites in the category / total number of metabolites in the category) × 100%. The numbers of upregulated (increased abundance in the depression group) and downregulated (decreased abundance in the depression group) metabolites were separately counted. To assess whether observed differential rates exceeded random expectation, exact binomial tests were performed for each category against the null hypothesis of a 5% differential rate (expected under $\alpha = 0.05$).

Statistical power analysis: A formal post-hoc power analysis was conducted for the two-sample Mann-Whitney U test equivalent (Cohen's d effect size) at the BH-adjusted significance threshold. For 5932 simultaneous tests, the effective BH threshold was approximately $\alpha \approx 8.4\times10^{-6}$. With n_depression = 55 and n_control = 35 (total n = 90), the non-centrality parameter was $\delta = d \times \sqrt{(n_1 n_2/(n_1+n_2))} = d \times 4.625$. At a medium effect size (d = 0.5), the statistical power was only 1.61%; to achieve 80% power, a total sample size of approximately 473 (289 depression + 184 control) would be required. Even at a large effect size (d = 0.8), current power was only 22.5%, requiring 185 total subjects for 80% power. (See Supplementary Figure S3 for missing value pattern before KNN imputation.)

Statistical software: Data preprocessing and statistical analysis were performed using Python 3.12 (scipy 1.11, scikit-learn 1.4, statsmodels 0.14), and visualization was performed using matplotlib 3.8 and seaborn 0.13. All tests were two-tailed, with $p < 0.05$ considered statistically significant (exploratory analysis) and $p_{adj} < 0.05$ as the strict significance threshold.

Results

1 Demographic and Clinical Characteristics

A total of 90 subjects were enrolled in this study, including 55 in the depression group and 35 in the healthy control group. The total HAMD-17 score of the depression group was 21.24 ± 4.76 (range 14–33), meeting the criteria for moderate to severe depression.No statistically significant differences were found between the two groups in demographic indicators such as age, sex, and BMI (all $p > 0.05$), indicating comparability. The smoking rate (34.5% vs 8.6%, $p = 0.006$) and alcohol drinking rate (20.0% vs 2.9%, $p = 0.025$) were significantly higher in the depression group than in the control group, and the distribution of education level differed between groups ($p = 0.028$). These factors were considered as potential confounders in subsequent analyses. (Table 1)

2 Metabolomics Data Quality

A total of 5932 metabolites were identified by UHPLC-Q-TOF-MS (combining positive and negative ion modes: 3230 in ESI+ and 2702 in ESI−). The RSD of peak area for all metabolites in QC samples was <30% (mean RSD = 0.97%, median = 0.81%, maximum = 6.33%), indicating excellent instrument stability and compliance with metabolomics data quality control standards [26]. The missing rate of raw data was 4.66%, and the completeness rate reached 100% after KNN imputation. Metabolite identification confidence was distributed as follows: MSI Level 1 (n = 1127, 19.0%), Level 2 (n = 1392, 23.5%), Level 3 (n = 2216, 37.4%), and Level 4 (n = 1197, 20.2%).(See Supplementary Tables S1, S2, S3 and Supplementary Figures S1–S2 for complete data quality metrics.)

3 Univariate Statistical Analysis Results

The Mann-Whitney U test was performed on 5932 metabolites [19], and 179 metabolites (3.02%) had raw p-values <0.05. After Benjamini-Hochberg multiple testing correction [20], no metabolite reached the significance level of $p_{adj} < 0.05$ (minimum $p_{adj} = 1.00$). Notably, the number of metabolites with raw $p < 0.05$ (179) was lower than the expected number by chance (approximately 297) for 5932 tests at $\alpha = 0.05$, suggesting that there was no between-group difference signal exceeding the random level in this dataset. This anomalous p-value distribution (fewer significant results than

expected by chance) may reflect the high inter-individual variability of the urinary metabolome, potential conservatism introduced by data preprocessing, or the absence of true disease-related metabolic signals at the current sample size. A full p-value distribution histogram and Q-Q plot are provided in Supplementary Figure S14. This negative result is consistent with findings from some previous metabolomics studies in first-episode drug-naïve patients with depression—a metabolomics study in adolescent patients with psychiatric disorders found no significantly differential metabolites in first-episode drug-naïve patients, suggesting that metabolic changes in the early stage of the disease may be subtle [27]. (Figure 1) (See Supplementary Table S4 for complete statistical results of all 5932 metabolites; Supplementary Figures S4 and S5 for p-value distribution diagnostics.)

4 Multivariate Statistical Analysis Results

PCA analysis [21]: The first two principal components explained 29.13% and 5.56% of the total variance, respectively, with a cumulative explained variance of 34.69%. The depression group and healthy control group showed extensive overlap on the PCA score plot, with no obvious separation trend, suggesting high similarity in the overall urinary metabolomic profiles between the two groups. (Figure 2A)

OPLS-DA analysis [22]: An OPLS-DA model with one predictive component and one orthogonal component was established, with $R^2Y = 0.599$, indicating that the model could explain approximately 60% of the between-group variance. However, the 7-fold cross-validated Q^2 was −1.185; a negative Q^2 indicates that the model lacks actual discriminative predictive ability and is overfitted. Permutation testing (n = 200) further confirmed that the R^2Y of the true model (0.599) showed almost no difference from the mean of random permutations (0.577), and although the true Q^2 (−1.185) was higher than the permutation mean (−1.738), it remained negative, indicating no statistical significance of the model. A total of 1061 metabolites had VIP >1, but this result is unreliable because the model itself lacks predictive validity. Specifically, a negative Q^2 means that the model performs worse than the null model (predicting the group mean), and therefore VIP scores derived from such a model cannot be used for reliable variable selection. The VIP >1 criterion was retained only for exploratory, hypothesis-generating purposes and should not be interpreted as evidence of differential abundance. (Figure 2B, Figure 2C)

(See Supplementary Figure S6 for VIP score distribution and VIP significance correlation.)

5 Characteristics of Exploratory Differential Metabolites

Given the absence of significantly differential metabolites under strict statistical criteria, 66 exploratory differential metabolites were screened using the combined criteria of p_raw <0.05, VIP >1, and |log2FC| > log2(1.5) to explore potential biological clues. Among these, 39 were upregulated and 27 were downregulated in the depression group. Metabolites with the smallest p-values included N,N-dimethyl-3,4-methylenedioxyamphetamine (p = 0.0012, upregulated), undecanoylcholine (p = 0.0025, downregulated), and 4-aminobenzoylhydrazide (p = 0.0032, downregulated). (See Supplementary Table S5 for complete list of all 66 candidates with MSI levels; Supplementary Figure S7 for forest plot of top 20.)

Importantly, among the 66 candidate metabolites, 12 (18.2%) were identified as drug molecules or pharmaceutical compounds, including N,N-dimethyl-3,4-methylenedioxyamphetamine (MDMA, a recreational drug), flupentixol (an antipsychotic), ractopamine (a feed additive), mocetinostat and tubastatin A (HDAC inhibitors), nalbuphine (an opioid analgesic), tenatoprazole (a proton pump inhibitor), and guanfacine (an α2-adrenergic agonist used for ADHD). Their appearance in a "drug-naïve" cohort strongly suggests false-positive identification due to the limitations of exact mass-based database matching (MSI Level 2/3), where multiple compounds may share the same accurate molecular mass, and in-source fragmentation, adduct formation, and isotope peaks can further increase annotation errors [29,30]. Additionally, 2 metabolites were smoking-related: cotinine (the primary nicotine metabolite, FC = 4.39, p = 0.048) and paraxanthine (a caffeine metabolite, FC = 1.52, p = 0.024), consistent with the 4-fold higher smoking rate in the depression group. Sixteen metabolites (24.2%) were diet-derived phytochemicals, and 30 (45.5%) were classified as putative endogenous or unknown compounds. A complete classification of all 66 candidates is provided in Supplementary Table S5. (See Supplementary Table S6 for complete classification of all 66 candidates; Supplementary Figure S8 for confounder and exogenous compound annotation; Supplementary Figure S9 for log2FC significance by Super Class; Supplementary Figure S10 for up/down regulation composition.)

Previous studies have shown that antidepressants can systematically alter multiple metabolic pathways including tryptophan, purine, and lipid metabolism [12]; therefore, strict control of medication status is crucial for accurately identifying the endogenous metabolic characteristics of the disease. The chemical identities of all candidate metabolites, particularly the 12 drug-like compounds and those at MSI Level 3/4, require rigorous validation through reference standard comparison and MS/MS fragmentation spectrum matching before any biological interpretation. (Table 2, Figure 3)

6 Functional Classification and Differential Rate Analysis

At the Super Class level: among the 19 super classes, the highest differential rate was observed for Phenylpropanoids and polyketides (2.10%, 10/477), followed by Benzenoids (1.55%, 12/774), Organic nitrogen compounds (1.23%, 1/81), Organic acids and derivatives (1.22%, 15/1225), and Organoheterocyclic compounds (1.22%, 15/1225). The differential rates of all categories were low, with no super class exceeding 3%. Exact binomial tests confirmed that no Super Class had a differential rate significantly exceeding the 5% random expectation (all $p > 0.05$), further supporting the absence of systematic metabolic disturbance. (Figure 4A, Supplementary Table S2a) (See Supplementary Table S7a for binomial test results; Supplementary Figure S11 for Super Class × Class heatmap.)

At the Class level: among the 66 categories with ≥10 metabolites, the top five by differential rate were Tetrahydroisoquinolines (10.00%, 1/10), Benzodioxoles (4.76%, 1/21), Isoflavonoids (4.17%, 2/48), Glycerophospholipids (4.00%, 1/25), and Naphthalenes (3.64%, 2/55). However, these high-differential-rate categories generally had small sample sizes (10–55 metabolites), leading to unstable estimates. Binomial tests showed that none of these Class-level differential rates significantly exceeded 5% after accounting for small sample sizes. (Figure 4B, Supplementary Table S2b) (See Supplementary Table S7b for binomial test results; Supplementary Figure S12 for Class-level bubble plot.)

At the Sub Class level: N-alkylindoles, O-methylated isoflavonoids, naphthoquinones, and tyrosol and its derivatives all showed a differential rate of 20% (1/5), but each category contained only 5 metabolites, resulting in extremely low

statistical power. Binomial tests confirmed that these apparent high rates were not statistically distinguishable from random variation (all $p > 0.05$). The ClassyFire classification system provides a standardized framework for functional annotation of metabolites [24]. (Table 3a, Table 3b, Table 3c, Figure 5, Supplementary Table S2c) (See Supplementary Table S7c for binomial test results; Supplementary Figure S13 for Subclass-level differential rate ranking.)

Discussion

1 Main Findings

This study employed UHPLC-Q-TOF-MS untargeted metabolomics to systematically analyze the metabolic profiles of urine samples from 55 first-episode drug-naïve patients with depression and 35 healthy controls. After rigorous KNN missing value imputation, Mann-Whitney U test, and Benjamini-Hochberg multiple testing correction, the core finding of this study is that under strict statistical criteria, there were no urinary differential metabolites that remained significant after multiple testing correction between first-episode drug-naïve patients with depression and healthy controls, and the OPLS-DA model also did not show predictive between-group separation ($Q^2 = -1.185$, permutation test $p > 0.05$). This negative result is not unique in depression metabolomics research—a metabolomics study in adolescent patients with psychiatric disorders similarly failed to identify significantly differential metabolites in first-episode drug-naïve patients with strictly controlled medication status [27]. The 66 candidate differential metabolites and functional classification results (relatively higher differential rates for phenylpropanoids and polyketides, benzenoids, and organic acids) identified by exploratory analysis can only serve as hypothesis-generating clues and do not constitute statistically significant biomarker discoveries.

2 Comparison with Previous Studies

Most previous metabolomics studies on depression have reported positive findings. Zheng et al. used NMR and GC-MS platforms to identify a biomarker panel comprising 5 metabolites including malonic acid, formic acid, and N-methylnicotinamide in urine samples, achieving an AUC of 0.89 in the validation set [10]; the same team subsequently identified sex-specific urinary biomarker panels [11]. In blood samples, multiple studies have reported abnormal alterations in amino acid metabolism, lipid metabolism, and energy metabolism [6,7,13,27]. A systematic review indicated that kynurenine and acylcarnitines are relatively consistent metabolic abnormalities in depression [8]. However, the reproducibility of these positive results across studies is poor. A systematic analysis by Pu et al. of nearly 300 metabolomics studies on animal models of depression found that only approximately 10 metabolites showed consistent alterations across brain, blood, and urine samples, highlighting the prominent

reproducibility issues in metabolomics research [3]. Another review on metabolomic biomarkers for psychiatric disorders also noted that due to heterogeneity in study design, small sample sizes, different analytical platforms, and inconsistent validation criteria, no fully validated metabolomic biomarkers are currently available for clinical use [28]. The negative results of this study are consistent with this broader trend, further underscoring the reproducibility challenges faced by metabolomic biomarker research in depression.

3 Possible Explanations for the Negative Results

First, differences between urinary and blood metabolomes. Urinary metabolites reflect metabolic end products after renal filtration and excretion, containing not only endogenous metabolites but also abundant gut microbiota-derived metabolites and exogenous compounds [9]. Compared with blood, the urinary metabolome is more strongly influenced by diet, water intake, renal function, circadian rhythm, and urine concentration, with higher inter-individual variability that may mask subtle disease-specific signals. Previous studies reporting positive urinary biomarkers mostly used NMR platforms [10,11], whereas this study used the UHPLC-Q-TOF-MS platform; differences in metabolite coverage and quantitative characteristics may also contribute to incomparable results.

Second, metabolic alterations in first-episode drug-naïve populations may be subtle. Most previous positive studies enrolled chronic patients receiving antidepressant treatment, and antidepressants themselves can systematically alter multiple metabolic pathways including tryptophan, purine, and lipid metabolism [12], potentially amplifying or confounding disease-related metabolic signals. Due to shorter disease duration and no prior drug intervention, the metabolic disturbances in first-episode drug-naïve patients may be at an early, subclinical stage, with signal strength insufficient to be distinguished from inter-individual variability at the current sample size. This explanation is consistent with the findings of Kuwano et al., who observed only subtle metabolic changes in first-episode, drug-naïve patients [13].

Third, sample size and statistical power. This study included 55 patients in the depression group and 35 controls, representing a moderate sample size in metabolomics research. However, in high-dimensional testing of 5932 metabolites, to achieve p_adj

<0.05 after BH correction, the raw p-value for each metabolite needs to be extremely small (approximately $<8.4 \times 10^{-6}$), imposing high requirements on effect size. A formal power analysis revealed that at this threshold, the current sample size (n = 90) has only 1.61% power to detect a medium effect (Cohen's d = 0.5), and 22.5% power for a large effect (d = 0.8). To achieve 80% power for a medium effect, approximately 473 total subjects (289 depression + 184 control) would be required. Furthermore, only 179 metabolites had raw p <0.05, which is lower than the expected number by chance (approximately 297) for 5932 tests at $\alpha = 0.05$. This anomalous observation suggests that there may be no true between-group difference signal in the data, or that the signal is overwhelmed by substantial noise. (See Supplementary Table S8 and Supplementary Figure S14.)

Fourth, clinical heterogeneity of depression. Depression is a highly heterogeneous disorder, with significant differences among patients in symptom dimensions (anxious, melancholic, atypical), disease course, severity, and etiology. Comparing all patients with depression as a single group may dilute metabolic signals specific to certain subtypes due to phenotypic heterogeneity. In this study, the HAMD-17 scores of the depression group ranged from 14 to 33, covering moderate to severe depression, and anxiety symptoms (GAD-7 scores significantly higher than controls) were prominent; such phenotypic mixing may have reduced the sensitivity of between-group comparisons. Additionally, 58.2% of patients had prior depressive episodes (though never treated), introducing course-related heterogeneity that may further dilute first-episode-specific metabolic signals.

Fifth, the influence of confounding factors. This study found significant differences between the depression group and the control group in smoking rate (34.5% vs 8.6%), alcohol drinking rate (20.0% vs 2.9%), and education level distribution. Smoking and alcohol consumption can significantly affect the urinary metabolome—nicotine and its metabolites can be detected in urine, and alcohol intake can alter organic acid and lipid metabolic profiles. Notably, cotinine (the primary metabolite of nicotine) was identified among the 66 exploratory differential metabolites with a 4.39-fold elevation in the depression group (p = 0.048), directly reflecting the smoking rate difference. Although these factors were identified in baseline comparisons, multifactorial correction at the

metabolite level was not possible due to sample size limitations, potentially introducing residual confounding. A sensitivity analysis excluding smokers would be valuable but was not feasible in the current dataset due to the small number of smoking controls (n = 3). In addition, unmeasured confounders such as dietary patterns, physical activity levels, and sleep quality may also have influenced the results.

4 Biological Significance of Exploratory Findings

Although strict statistical analysis revealed no significant differences, the 66 candidate metabolites identified by exploratory analysis still provide some noteworthy biological clues. In terms of functional classification, Phenylpropanoids and polyketides had the highest differential rate (2.10%), followed by Benzenoids (1.55%) and Organic acids and derivatives (1.22%). Phenylpropanoid compounds are mostly derived from dietary polyphenols, which are metabolized by gut microbiota and excreted in urine as metabolic end products. Their differential presence may reflect altered dietary patterns or gut microbiota dysfunction in patients with depression. Previous studies have reported gut microbiota dysbiosis and abnormal dietary polyphenol metabolism in patients with depression [8]. The higher differential rate of organic acids may be related to mitochondrial energy metabolism and tricarboxylic acid cycle disturbances, consistent with the mitochondrial dysfunction hypothesis of depression.

However, these exploratory results must be interpreted with extreme caution for several reasons. First, the differential rates of all categories were below 3%, far lower than the expected differential rate under random conditions (5%), and binomial tests confirmed that no category significantly exceeded random expectation, suggesting that these "differences" are likely false positives. Second, among the 66 candidate metabolites, 12 (18.2%) were identified as drug molecules or pharmaceutical compounds, whose appearance in a "drug-naïve" cohort strongly indicates false-positive identification in untargeted metabolomics. Third, 2 metabolites (cotinine, paraxanthine) were directly related to smoking/caffeine consumption, reflecting confounding rather than disease biology. Fourth, only 42.5% of all identified metabolites had MS/MS-based identification (MSI Level 1+2), while 57.5% lacked MS/MS confirmation (Level 3+4), and the proportion of Level 3/4 identifications among candidate metabolites may be even higher. Untargeted metabolomics identifies metabolites based on accurate molecular mass and

database matching, and multiple compounds may share the same accurate molecular mass. Furthermore, in-source fragmentation, adduct formation, and isotope peaks can further increase the risk of erroneous annotation [29,30]. Studies have shown that biomarker discovery in LC-HRMS untargeted metabolomics is susceptible to false positives [31], and the proportion of metabolites achieving MSI Level 1 identification is typically very low. Therefore, the chemical identities of candidate metabolites in this study require rigorous validation through reference standard comparison and MS/MS fragmentation spectrum matching before any biological interpretation can be made. (See Supplementary Tables S2 and S6 for detailed classification and identification confidence.)

5 Methodological Considerations

This study has the following methodological strengths: (1) strict enrollment of drug-naïve patients, maximizing the exclusion of confounding effects of antidepressants on metabolic profiles [12]; (2) systematic quality control using pooled QC samples, with all metabolites showing QC peak area RSD <30% (mean RSD = 0.97%), meeting metabolomics data quality standards [26]; (3) KNN imputation (k=5, distance-weighted) for zero-value missing data, a widely recommended missing value processing method in metabolomics [17,18]; (4) BH method for multiple testing correction, strictly controlling FDR [20]; (5) OPLS-DA model validated by 200 permutation tests, avoiding overfitting false positives in supervised models [22]; (6) formal statistical power analysis quantifying the limitations of the current sample size; (7) MSI identification level classification providing transparency about annotation confidence.

Meanwhile, this study also has methodological limitations: (1) metabolite identification was primarily based on accurate molecular mass matching, with only 42.5% having MS/MS-based identification (MSI Level 1+2); 12 of 66 candidate metabolites (18.2%) were drug-like compounds likely representing false-positive annotations, requiring MS/MS and reference standard validation [29]; (2) the OPLS-DA model showed a negative Q^2 (−1.185), indicating no predictive validity, limited interpretive value of VIP values, and insufficient reliability of candidate metabolites screened based on VIP >1; (3) no urinary creatinine correction or specific gravity correction was performed, and individual differences in urine concentration may have

affected quantitative results; (4) the sample was from a single-center cross-sectional design, lacking an independent validation cohort, limiting the generalizability of the results; (5) the depression group had significantly higher smoking and alcohol rates than controls, and multifactorial correction was not feasible due to sample size, potentially introducing residual confounding; (6) 58.2% of patients had prior depressive episodes (though untreated), introducing course heterogeneity that may dilute first-episode-specific metabolic signals.

6 Study Limitations

This study has the following limitations: First, the sample size was relatively limited (n = 90), with formal power analysis showing only 1.61% power for medium effect sizes (d = 0.5) at BH-adjusted thresholds, insufficient to detect metabolic alterations with small to moderate effect sizes. Second, the single-center study design, with all subjects recruited from Hefei Fourth People's Hospital, limited population representativeness, and an independent validation cohort was lacking. Third, although medication status was strictly controlled, the depression group showed significant differences from the control group in smoking, alcohol drinking, and education level; these confounders may affect the urinary metabolome (as evidenced by cotinine FC = 4.39), and multifactorial correction at the metabolite level was not possible. Fourth, metabolite identification lacked MS/MS and reference standard validation for 57.5% of metabolites (MSI Level 3+4), and 18.2% of candidate metabolites were drug-like compounds likely representing false-positive annotations. Fifth, no urinary creatinine correction was performed, and individual differences in urine dilution may introduce quantitative bias. Sixth, the cross-sectional design cannot establish causal relationships between metabolic alterations and depression, and cannot distinguish whether metabolic alterations are the cause, consequence, or concomitant phenomenon of the disease. Seventh, only urine was used as a single biological sample, without integration with multi-matrix metabolomics such as serum and cerebrospinal fluid. Eighth, 58.2% of patients had prior depressive episodes (though never pharmacologically treated), and the cohort was not exclusively first-episode, limiting the interpretation of results as "first-episode" specific.

7 Future Directions and Conclusions

Based on the findings and limitations of this study, future research should be improved in the following aspects: (1) expand the sample size to at least 473 total subjects (for 80% power at d = 0.5) and adopt a multi-center study design to improve statistical power and generalizability of results; (2) establish an independent validation cohort for targeted metabolomics quantitative validation of candidate metabolites and pathways [13]; (3) use matched designs or propensity score matching to strictly control confounders such as smoking, alcohol drinking, diet, and BMI, and perform sensitivity analyses excluding smokers; (4) combine MS/MS spectra and reference standard comparison to elevate metabolite identification to MSI Level 1, reducing false-positive annotations [29,30]; (5) perform urinary creatinine correction or specific gravity correction to control differences in urine concentration; (6) integrate multi-matrix metabolomics (serum, cerebrospinal fluid, etc.) with genomics and proteomics for multi-omics integrated analysis to decipher the metabolic disturbance mechanisms of depression at the systems level; (7) conduct stratified analysis by clinical subtypes of depression and by first-episode vs. recurrent status to reduce the impact of phenotypic heterogeneity; (8) conduct longitudinal follow-up studies to observe dynamic changes in metabolic profiles with disease progression and treatment response.

In summary, in first-episode drug-naïve patients with depression under strictly controlled medication status, this study found no urinary differential metabolites that remained significant after multiple testing correction, and the OPLS-DA model also did not show predictive between-group separation. This negative result suggests that urinary metabolomic alterations in first-episode drug-naïve patients with depression may be subtle and are influenced by high inter-individual variability of the urinary metabolome, false-positive identification (18.2% of candidates were drug-like compounds), smoking confounders (cotinine FC = 4.39), and insufficient statistical power (1.61% for medium effect sizes). The alterations in phenylpropanoid, benzenoid, and organic acid metabolic pathways suggested by exploratory analysis can serve as clues for future hypothesis generation, but require confirmation in larger, multi-center studies with strict control of confounders, combined with targeted metabolomics and MS/MS validation. This study provides important negative evidence for urinary metabolomics research in depression and emphasizes the necessity of rigorous statistical correction, model validation,

metabolite identification quality control, and adequate sample size in metabolomic biomarker research.

Declarations

Abbreviations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Hefei Fourth People's Hospital, Anhui Province, China (Approval Number: HFSY-IRB-YJ-KYXM-CL[2024-064-001]). The study was performed in accordance with the ethical principles of the Declaration of Helsinki and its later amendments. Written informed consent to participate was obtained from all individual participants enrolled in this study. All participants agreed that their de-identified clinical data and biological samples could be used for this study and subsequent related scientific research.

Consent for publication

Not applicable. No identifiable individual participant images, portraits, or personal clinical details that could compromise participant anonymity are presented within this manuscript.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Clinical Key Specialty Construction Project of China and the 2024 Hefei City Health Sciences Applied Medicine Research Project (Research on the Mechanism and Early Warning Signals of the Synergistic Increase in Depression Recurrence Risk Caused by PM2.5 Exposure and Temperature Stress).

Authors' contributions

Xialong Cheng and Long Chen contributed to conceptualization, methodology, and formal analysis.Xialong Cheng,Long Chen and Anzhen Wang contributed to investigation.Min Pan and Xialong Cheng drafted the manuscript.Min Pan and Xulai

Zhang contributed to manuscript review and editing. Min Pan and Xulai Zhang acquired funding.Xulai Zhang and Anzhen Wang contributed to project administration and supervision. All authors read and approved the final manuscript and agree to be accountable for all aspects of the work.

Acknowledgements

The authors thank the patients who made this study possible, as well as the clinical and data-management staff at Hefei Fourth People's Hospital for their support in routine clinical assessment, data extraction, and data anonymization.

References

1. World Health Organization. Depression. Fact sheet. Geneva: WHO; 2023. Available at: https://www.who.int/news-room/fact-sheets/detail/depression
2. Huang Y, Wang Y, Wang H, Liu Z, Yu X, Yan J, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. Lancet Psychiatry. 2019;6(3):211-224. doi:10.1016/S2215-0366(18)30511-X. PMID: 30792114.
3. Pu J, Liu Y, Gui S, Tian L, Yu Y, Song X, et al. Metabolomic changes in animal models of depression: a systematic analysis. Mol Psychiatry. 2021;26(12):7328-7336. doi:10.1038/s41380-021-01269-w. PMID: 34471249.
4. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol. 2016;17(7):451-459. doi:10.1038/nrm.2016.25. PMID: 26979502.
5. Liu M, Yang R, Liu Z, et al. Recent Progress in Mass Spectrometry-Based Metabolomics in Major Depressive Disorder Research. Molecules. 2023;28(21):7430. doi:10.3390/molecules28217430. PMID: 37960154. PMCID: PMC10647556.
6. Liu Y, Zhang X, Li J, et al. Untargeted Plasma Metabolomic Profiling in Patients with Major Depressive Disorder Using Ultra-High Performance Liquid Chromatography Coupled with Mass Spectrometry. Biomed Res Int. 2021;2021:6672515. doi:10.1155/2021/6672515. PMCID: PMC8306682.
7. Chen Y, Li M, Wang H, et al. Alterations of serum metabolic profile in major depressive disorder: A case-control study in the Chinese population. Front Psychiatry. 2024;15:1214699. doi:10.3389/fpsyt.2024.1214699. PMCID: PMC12146990.
8. Costa LNF, Carneiro BA, Alves GS, et al. Metabolomics of Major Depressive Disorder: A Systematic Review of Clinical Studies. Biomedicines. 2022;10(4):786. doi:10.3390/biomedicines10040786. PMCID: PMC8993993.
9. Bouatra S, Aziat F, Mandal R, et al. The human urine metabolome. PLoS One. 2013;8(9):e73076. doi:10.1371/journal.pone.0073076. PMCID: PMC3767765.
10. Zheng P, Wang Y, Chen L, Yang D, Meng H, Zhou D, et al. Identification and validation of urinary metabolite biomarkers for major depressive disorder. Mol Cell Proteomics. 2013;12(1):207-214. doi:10.1074/mcp.M112.021816. PMID: 23111923.

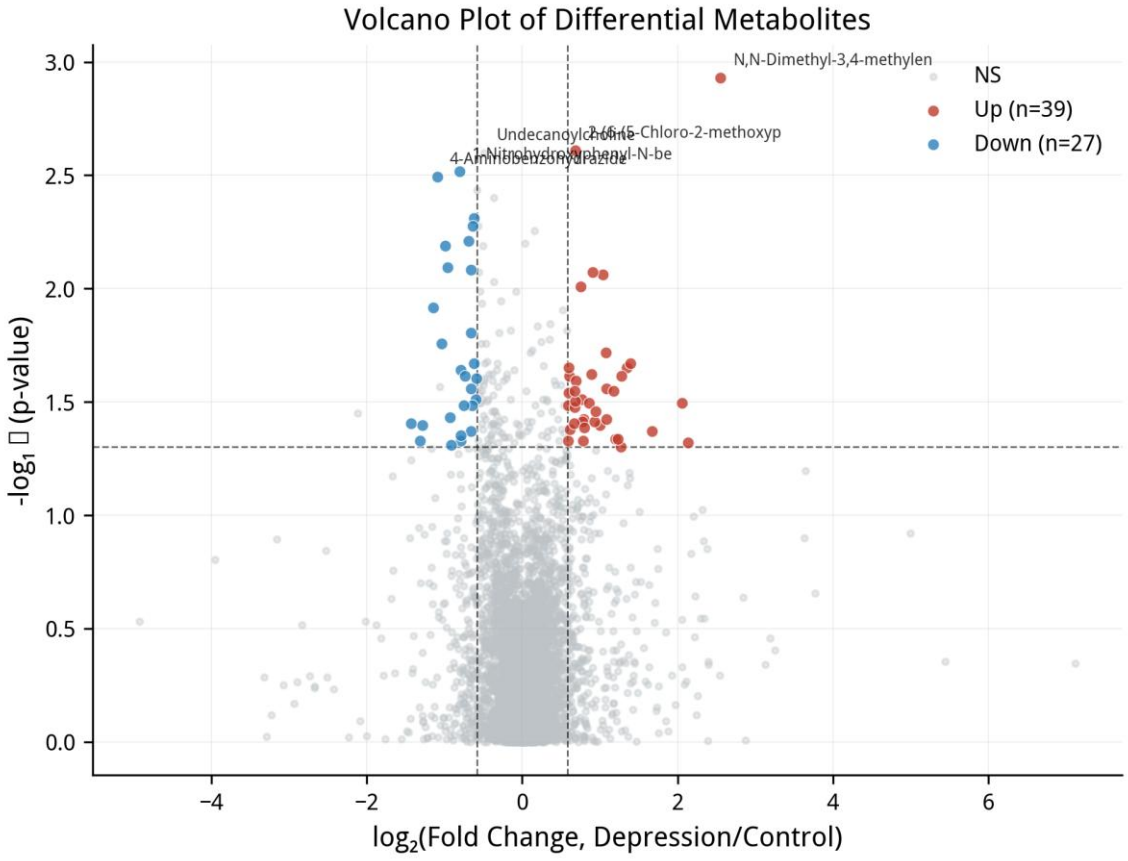
11. Zheng P, Chen JJ, Zhou CJ, Zeng L, Li KW, Sun L, et al. Identification of sex-specific urinary biomarkers for major depressive disorder by combined application of NMR- and GC-MS-based metabonomics. Transl Psychiatry. 2016;6(11):e955. doi:10.1038/tp.2016.188. PMID: 27845778.
12. Pu J, Liu Y, Gui S, Tian L, Yu Y, Song X, et al. Effects of pharmacological treatment on metabolomic alterations in animal models of depression. Transl Psychiatry. 2022;12(1):175. doi:10.1038/s41398-022-01947-5. PMID: 35487992.
13. Kuwano N, Kato TA, Setoyama D, Sato-Kasai M, Shimokawa N, Hayakawa K, et al. Tryptophan-kynurenine and lipid related metabolites as blood biomarkers for first-episode drug-naïve patients with major depressive disorder: An exploratory pilot case-control study. J Affect Disord. 2018;231:74-82. doi:10.1016/j.jad.2018.01.014. PMID: 29454180.
14. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Arlington, VA: American Psychiatric Association; 2013.
15. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry. 1960;23(1):56-62. doi:10.1136/jnnp.23.1.56. PMID: 14399078.
16. Want EJ, Wilson ID, Gika G, et al. Global metabolic profiling procedures for urine using UPLC-MS. Nat Protoc. 2010;5(6):1005-1018. doi:10.1038/nprot.2010.50. PMID: 20520630.
17. Di Guida R, Engel J, Allwood JW, et al. Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. Metabolomics. 2016;12(5):91. doi:10.1007/s11306-016-1030-0.
18. Hrydziuszko O, Viant MR. Missing values in mass spectrometry based metabolomics: an undervalued step in the data processing pipeline. Metabolomics. 2012;8(Suppl 1):161-174. doi:10.1007/s11306-011-0367-4.
19. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat. 1947;18(1):50-60. doi:10.1214/aoms/1177730491.
20. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. 1995;57(1):289-300.
21. Pearson K. On lines and planes of closest fit to systems of points in space. Philos Mag. 1901;2(11):559-572. doi:10.1080/14786440109462720.
22. Trygg J, Wold S. Orthogonal projections to latent structures (O-PLS). J Chemometrics. 2002;16(3):119-128. doi:10.1002/cem.695.
23. van den Berg RA, Hoefsloot HCJ, Westerhuis JA, et al. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics. 2006;7:142. doi:10.1186/1471-2164-7-142. PMID: 16762068.
24. Djoumbou Feunang Y, Eisner R, Knox C, et al. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. J Cheminform. 2016;8(1):61. doi:10.1186/s13321-016-0174-y. PMID: 27867422.
25. Dunn WB, Broadhurst D, Begley P, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc. 2011;6(7):1060-1083. doi:10.1038/nprot.2011.335. PMID: 21637174.
26. Broadhurst D, Goodacre R, Reinke SN, et al. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. Metabolomics. 2018;14(6):72. doi:10.1007/s11306-018-1366-6. PMID: 29861880.
27. Wang L, Yang P, Yang C, et al. Disturbance of neurotransmitter metabolism in drug-naïve, first-episode major depressive disorder: a comparative study on adult and adolescent cohorts. Eur Arch

Psychiatry Clin Neurosci. 2022;272(7):1283-1296. doi:10.1007/s00406-022-01406-8. PMID: 35410391.

28. Konjevod M, Sáiz J, Bordoy L, et al. Validated metabolomic biomarkers in psychiatric disorders: a narrative review. Mol Med. 2025;31(1):254. doi:10.1186/s10020-025-01258-7. PMCID: PMC12239260.
29. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. Untargeted metabolomics strategies – Challenges and Emerging Directions. J Am Soc Mass Spectrom. 2016;27(12):1897-1905. doi:10.1007/s13361-016-1469-y. PMID: 27827906.
30. Theodoridis G, Gika H, Raftery D, Goodacre R, Plumb RS, Wilson ID. Ensuring Fact-Based Metabolite Identification in Liquid Chromatography–Mass Spectrometry-Based Metabolomics. Anal Chem. 2023;95(8):3909-3916. doi:10.1021/acs.analchem.2c05192. Epub 2023 Feb 15. PMID: 36791228. PMCID: PMC9979140.
31. Zhang P, Ang IL, Lam MMT, Wei R, Lei KMK, Zhou X, et al. Susceptibility to false discovery in biomarker research using liquid chromatography–high resolution mass spectrometry based untargeted metabolomics profiling. Clin Transl Med. 2021;11(6):e469. doi:10.1002/ctm2.469. PMID: 34185426. PMCID: PMC8236120.

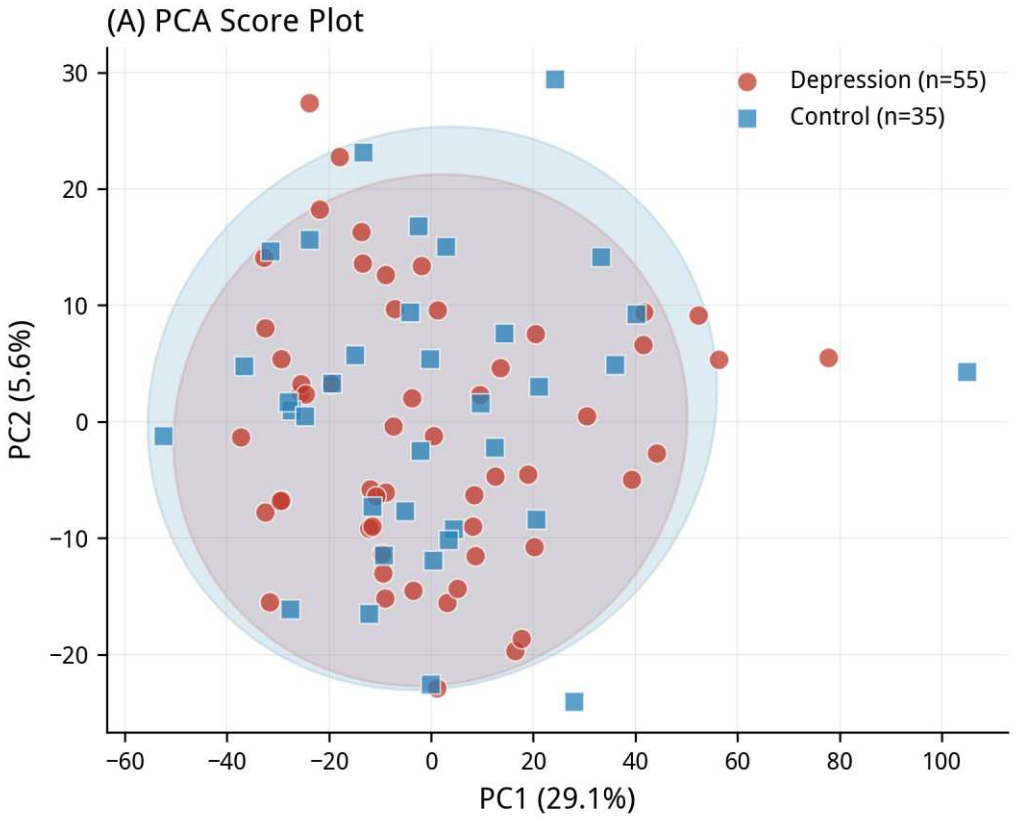
Main Figures

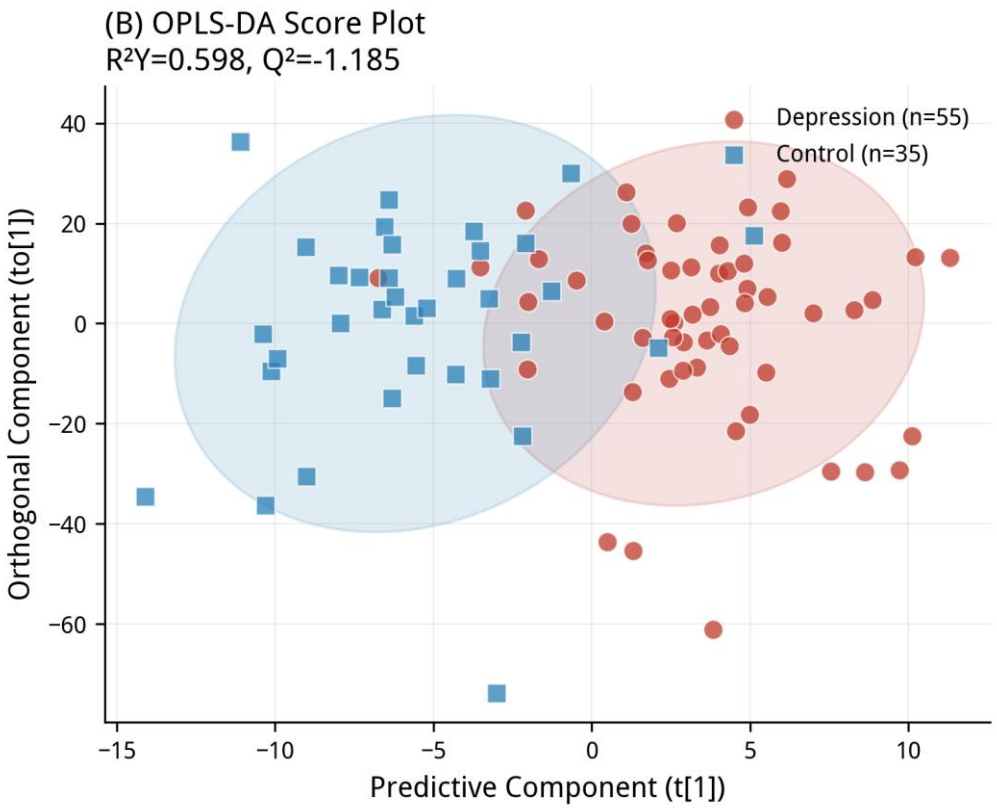
Figure 1. Volcano plot of differential metabolites between first-episode drug-naïve MDD patients and healthy controls.

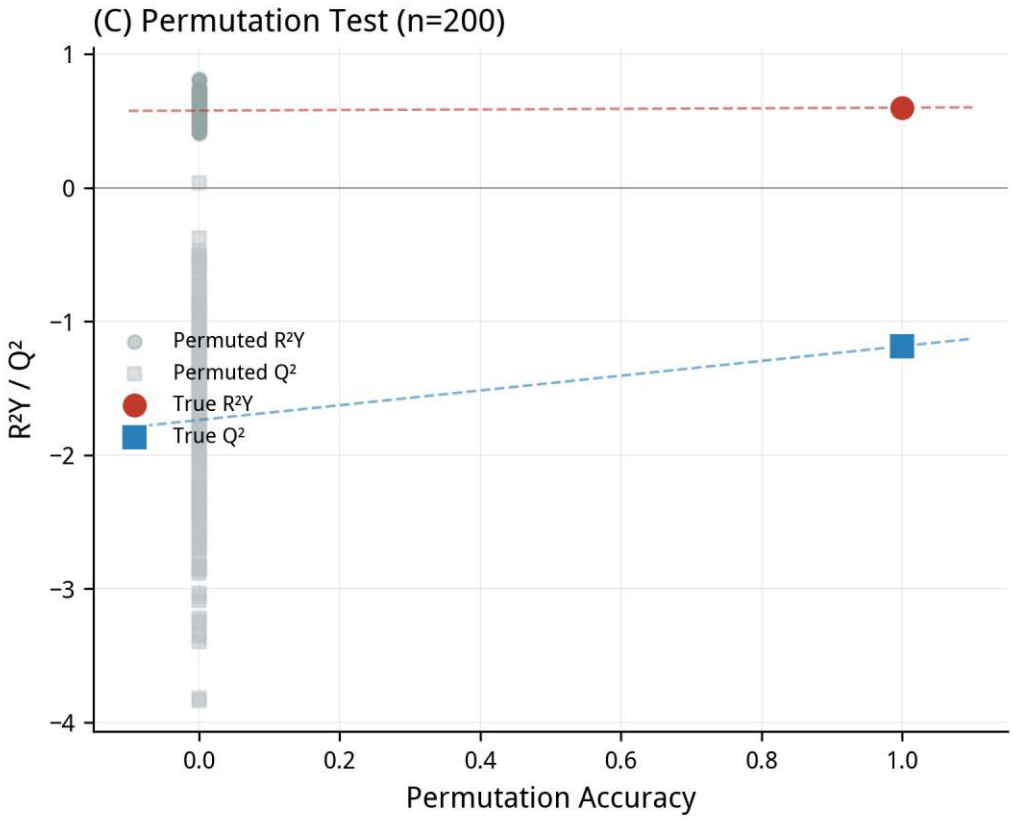


Each point represents one of the 5,932 identified metabolites. The x-axis shows $\log_2$(fold change) (depression group / healthy control group), and the y-axis shows $-\log_{10}$(raw p-value) from the Mann-Whitney U test. Vertical dashed lines indicate the fold-change threshold of $|\log_2 FC| > \log_2(1.5) \approx 0.585$; the horizontal dashed line indicates the significance threshold of $p < 0.05$. Red points (n = 39) represent metabolites upregulated in the depression group; blue points (n = 27) represent metabolites downregulated in the depression group; gray points represent non-significant metabolites. The top metabolites by statistical significance are labeled with their names. Note that no metabolite remained significant after Benjamini-Hochberg multiple testing correction; the 66 colored points constitute exploratory differential metabolites identified by the combined criteria of $p_raw < 0.05$, VIP > 1, and $|FC| > 1.5$.

Figure 2. Multivariate statistical analysis of urinary metabolic profiles.




(B) OPLS-DA Score Plot
R²Y=0.598, Q²=-1.185
Depression (n=55)
Control (n=35)
Orthogonal Component (to[1])
Predictive Component (t[1])
40
20
0
−20
−40
−60
−15
−10
−5
0
5
10

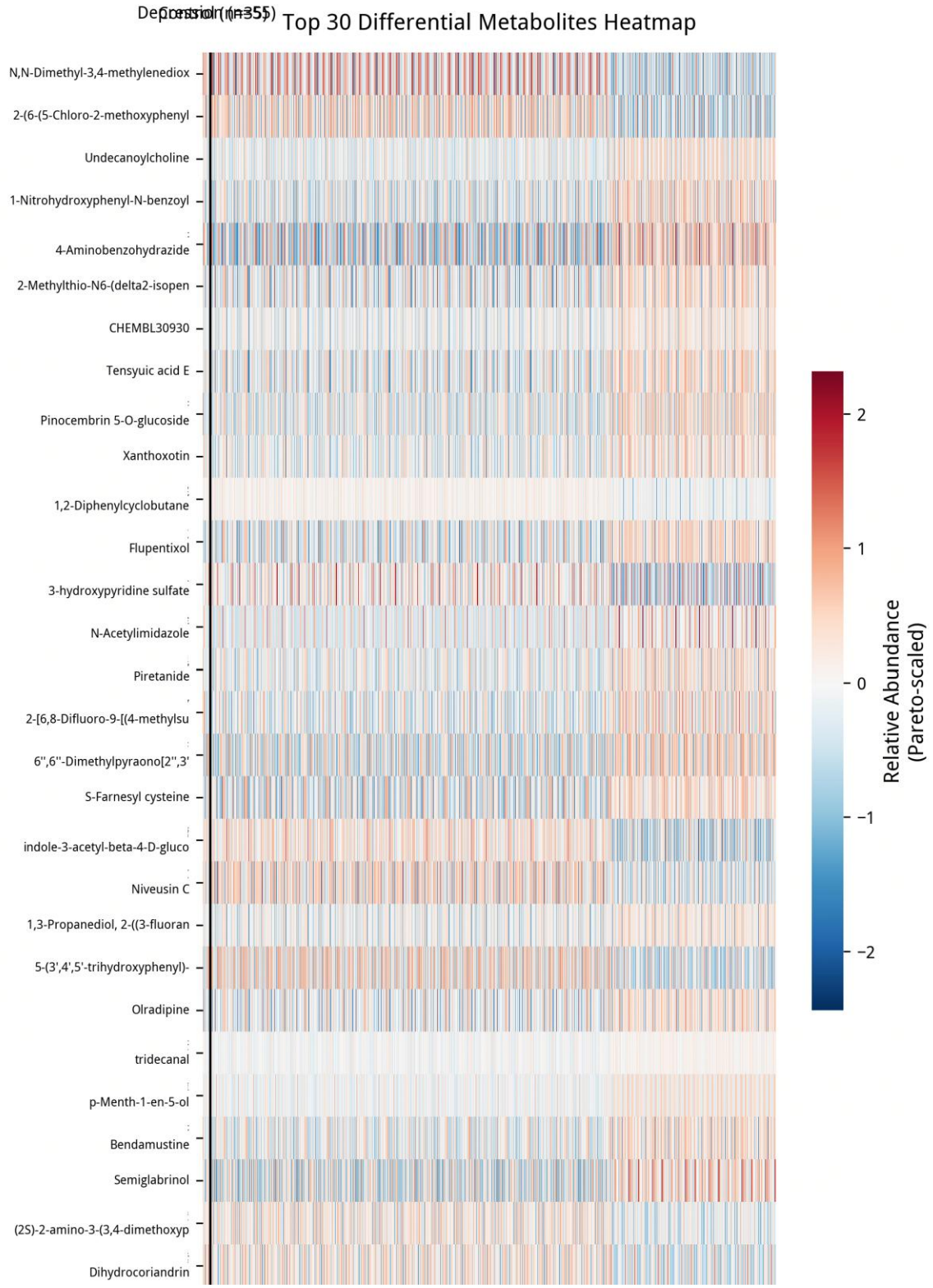

(C) Permutation Test (n=200)
Permuted R²Y
Permuted Q²
True R²Y
True Q²
R²Y / Q²
Permutation Accuracy
1
0
−1
−2
−3
−4
0.0
0.2
0.4
0.6
0.8
1.0

(A) PCA score plot. Unsupervised Principal Component Analysis (PCA) was performed on $\log_{10}$-transformed, Pareto-scaled metabolite abundance data from 55 depression patients (red circles) and 35 healthy controls (blue squares). The first two principal components explained 29.13% (PC1) and 5.56% (PC2) of the total variance, respectively, with a cumulative explained variance of 34.69%. The 95% confidence ellipses are shown for each group. No clear separation was observed between the two groups, indicating substantial overlap in overall urinary metabolic profiles.

(B) OPLS-DA score plot. Supervised Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) was conducted with one predictive component and one orthogonal component. The model achieved $R^2Y = 0.598$, indicating that approximately 60% of the between-group variation could be explained by the model. However, the cross-validated predictive ability $Q^2 = -1.185$ was negative, demonstrating that the model lacked genuine predictive power and was likely overfitted. The 95% confidence ellipses are shown.

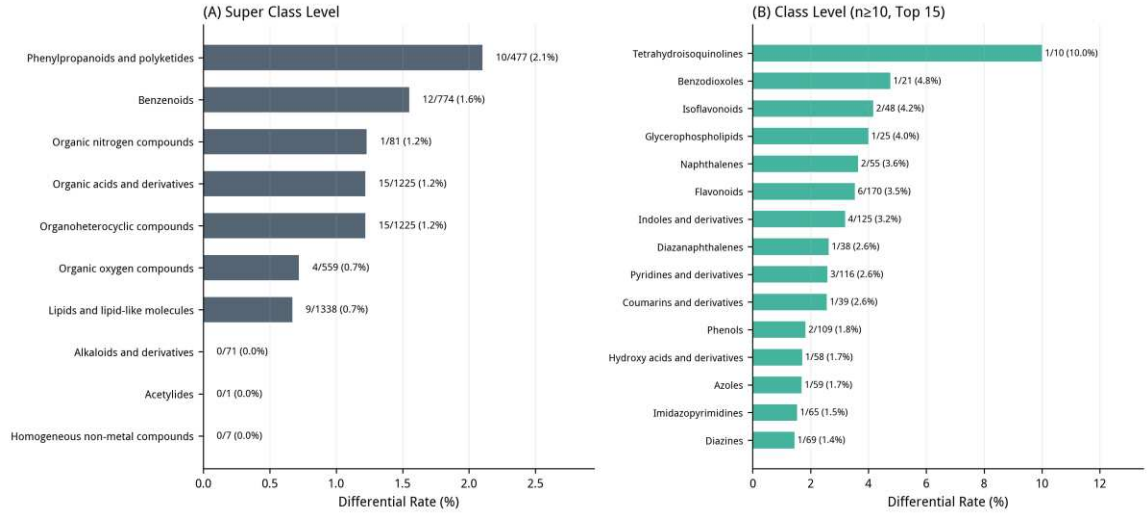
(C) Permutation test of the OPLS-DA model. A permutation test with 200 iterations was performed to validate the OPLS-DA model. The x-axis represents permutation accuracy (correlation between permuted and original class labels), and the y-axis represents R^2Y / Q^2 values. Gray circles and squares represent permuted R^2Y and Q^2 values, respectively; the red circle and blue square represent the true R^2Y and Q^2 values. The intercepts of the regression lines for R^2Y (0.46) and Q^2 (-0.43) indicate that the model did not pass the permutation validity criterion (Q^2 intercept should be < 0.05 for a valid model), further confirming the absence of robust between-group metabolic differentiation.

Figure 3. Heatmap of the top 30 exploratory differential metabolites.



The heatmap displays the relative abundance (Pareto-scaled, $\log_{10}$-transformed) of the top 30 exploratory differential metabolites ranked by raw p-value, across all 90 subjects (55 depression patients and 35 healthy controls). Columns represent individual samples, grouped by condition (depression vs. control); rows represent individual metabolites, clustered by hierarchical clustering using Euclidean distance and Ward linkage. The color scale ranges from blue (low relative abundance) to red (high relative abundance). Metabolite names are shown on the y-axis, with both English and Chinese labels where applicable. Note that several top-ranked metabolites are drug molecules or exogenous compounds (e.g., N,N-dimethyl-3,4-methylenedioxyamphetamine, flupentixol, piretamide, bendamustine), which likely represent false-positive identifications or environmental/dietary confounders rather than genuine disease-associated metabolic alterations.

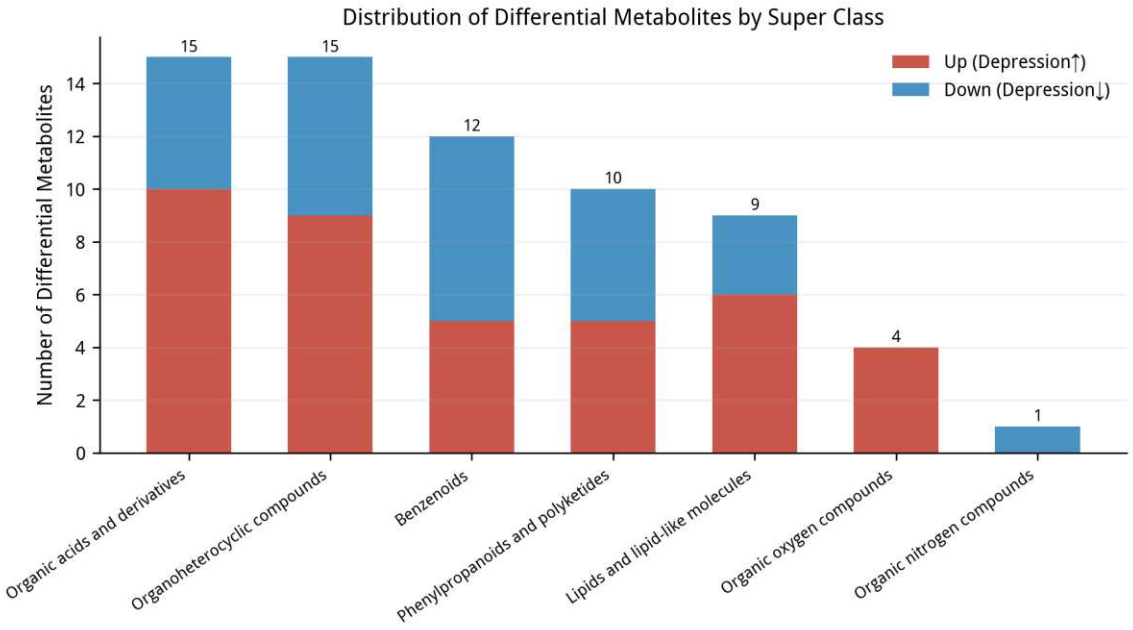
Figure 4. Differential rate analysis by chemical functional classification.



(A) Super Class level. Bar chart showing the differential rate (percentage of exploratory differential metabolites within each category) across 19 Super Classes defined by the ClassyFire chemical classification system. Only categories with ≥1 metabolite are shown. The differential rate was calculated as the number of exploratory differential metabolites divided by the total number of identified metabolites in each category. The highest differential rate was observed for Phenylpropanoids and polyketides (2.10%, 10/477), followed by Benzenoids (1.55%, 12/774), Organic nitrogen compounds (1.23%, 1/81), Organic acids and derivatives (1.22%, 15/1225), and Organoheterocyclic compounds (1.22%, 15/1225). All categories showed differential rates well below the 5% expected by chance, and none reached statistical significance in binomial tests (all $p > 0.05$).

(B) Class level (n ≥ 10, top 15). Bar chart showing the top 15 Class-level categories by differential rate, restricted to categories containing at least 10 identified metabolites. The top five categories by differential rate were Tetrahydroisoquinolines (10.00%, 1/10), Benzodioxoles (4.76%, 1/21), Isoflavonoids (4.17%, 2/48), Glycerophospholipids (4.00%, 1/25), and Naphthalenes (3.64%, 2/55). However, these elevated rates were driven by very small numbers of differential metabolites (1–2 per category) and did not survive statistical testing. The numbers above each bar indicate the count of differential metabolites and the total count in each category.

Figure 5. Distribution of up- and down-regulated exploratory differential metabolites by Super Class.



Stacked bar chart showing the number of upregulated (red, Depression↑) and downregulated (blue, Depression↓) exploratory differential metabolites within each Super Class that contained at least one differential metabolite. A total of 66 exploratory differential metabolites were identified, of which 39 (59.1%) were upregulated and 27 (40.9%) were downregulated in the depression group. The largest numbers of differential metabolites were found in Organic acids and derivatives (n = 15; 10 up, 5 down) and Organoheterocyclic compounds (n = 15; 9 up, 6 down), followed by Benzenoids (n = 12; 5 up, 7 down), Phenylpropanoids and polyketides (n = 10; 5 up, 5 down), Lipids and lipid-like molecules (n = 9; 6 up, 3 down), Organic oxygen compounds (n = 4; all upregulated), and Organic nitrogen compounds (n = 1; downregulated). The numbers above each stacked bar indicate the total count of differential metabolites in that category.

Main Tables

Table 1. Demographic and clinical characteristics of study participants.

Characteristic	Statistical format	Healthy controls (n=35)	Depression group (n=55)	p value
Age (years)	Median (P25, P75)	26.00(23.00,31.00)	27.00(22.00,33.50)	0.5313
Height (m)	Median (P25, P75)	1.68(1.62,1.75)	1.70(1.60,1.75)	0.8617
Sex	Female	19(54.3%)	24(44.4%)	0.3927
	Male	16(45.7%)	30(55.6%)	
Education level	Illiterate	0(0.0%)	1(1.8%)	0.0277
	Primary school	0(0.0%)	2(3.6%)	
	Junior high school	1(2.9%)	9(16.4%)	
	High school / vocational school	0(0.0%)	6(10.9%)	
	College / junior college	31(88.6%)	31(56.4%)	
	Postgraduate and above	3(8.6%)	6(10.9%)	
Residence	Rural	8(22.9%)	13(23.6%)	1
	Urban	27(77.1%)	42(76.4%)	
Marital status	Married	10(28.6%)	23(41.8%)	0.323
	Unmarried	24(68.6%)	29(52.7%)	
	Divorced	1(2.9%)	3(5.5%)	
Employment status	Full-time	18(51.4%)	26(47.3%)	0.2349
	Part-time	0(0.0%)	4(7.3%)	
	Student	15(42.9%)	15(27.3%)	
	Housewife	1(2.9%)	5(9.1%)	
	Unemployed	1(2.9%)	4(7.3%)	
	Retired	0(0.0%)	1(1.8%)	
Smoking	No	32(91.4%)	36(65.5%)	0.0056
	Yes	3(8.6%)	19(34.5%)	
Alcohol consumption	No	34(97.1%)	44(80.0%)	0.0247
	Yes	1(2.9%)	11(20.0%)	
HAMD-17 total score	Mean ± SD	—	21.24±4.76	—
HAMA total score	Mean ± SD	—	22.15±6.56	—
Stressful life events	No	—	37(67.3%)	—
	Yes	—	18(32.7%)	—

Data are presented as median (interquartile range, P25–P75) for non-normally distributed continuous variables, mean ± standard deviation ($\bar{x} \pm s$) for normally distributed variables, and n (%) for categorical variables. Between-group comparisons were performed using the Mann-Whitney U test for continuous variables and Fisher's exact test for categorical variables. A total of 90 subjects were enrolled: 55 first-episode drug-naïve patients with Major Depressive Disorder (MDD) and 35 age- and sex-matched healthy controls. The depression group had a mean HAMD-17 total score of 21.24 ± 4.76 (range:

14–32) and a mean HAMA total score of 22.15 ± 6.56. Statistically significant between-group differences were observed in education level ($p = 0.028$), smoking rate (34.5% vs. 8.6%, $p = 0.006$), and alcohol drinking rate (20.0% vs. 2.9%, $p = 0.025$), which were acknowledged as potential confounding factors. No significant differences were found in age, sex, height, urban/rural residence, marital status, or employment status.

Table 2. List of 66 exploratory differential metabolites identified in urine of first-episode drug-naïve MDD patients.

No.	Metabolite	m/z	RT (min)	Super Class	Class	Mean (DP)	Mean (HC)	FC (DP/HC)	log2FC	Mann-Whitney U	p value	VIP	Trend
1	N,N-Dimethyl-3,4-methylenedioxyamphetamine	208.1334	3.188	Organoheterocyclic compounds	Benzodioxoles	32946449.63	5617042.315	5.865	2.552	1355	0.001176791	4.127	Up-regulated (↑)
2	2-(6-(5-Chloro-2-methoxyphenyl)-4-oxo-2-thioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamide	360.0007	4.186	Organoheterocyclic compounds	Diazines	760239.0582	473892.3748	1.604	0.682	1329	0.002451587	3.028	Up-regulated (↑)
3	1-Nitrohydroxyphenyl-N-benzoylalanine	375.0829	5.722	Organic acids and derivatives	Carboxylic acids and derivatives	345810.7982	605789.9327	0.571	-0.809	604	0.003046165	2.545	Down-regulated (↓)
4	4-Aminobenzohydrazide	196.0727	2.276	Benzenoids	Benzene and substituted derivatives	2446307.895	5210786.497	0.469	-1.091	606	0.003214057	3.916	Down-regulated (↓)
5	Tensyuic acid E	331.1399	6.541	Organic acids and derivatives	Carboxylic acids and derivatives	13797340.76	21201482.32	0.651	-0.62	622	0.004891997	2.433	Down-regulated (↓)
6	Pinocembrin 5-O-glucoside	453.0994	5.726	Phenylpropanoids and polyketides	Flavonoids	1072050.269	1667706.207	0.643	-0.637	625	0.005283339	2.194	Down-regulated (↓)
7	Flupentixol	433.157	6.833	Organoheterocyclic compounds	Benzothiopyrans	241174.8712	388252.8111	0.621	-0.687	631	0.006151943	2.701	Down-regulated (↓)
8	N-Acetylimidazole	155.0462	1.703	Organoheterocyclic compounds	Azoles	4156374.538	8254483.787	0.504	-0.99	633	0.00646886	1.962	Down-regulated (↓)
9	2-[6,8-Difluoro-9-[(4-methylsulfonylphenyl)methyl]-1,2,3,4-tetrahydrocarbazol-1-yl]acetic acid	472.0806	6.509	Organoheterocyclic compounds	Indoles and derivatives	257544.9507	500286.6552	0.515	-0.958	642	0.00808432	2.374	Down-regulated (↓)
10	6",6"-Dimethylpyraono[2",3":7,8]kaempferol 4'-methyl ether 3-rhamnoside	530.2003	7.619	Phenylpropanoids and polyketides	Flavonoids	487613.9536	768164.4286	0.635	-0.656	643	0.00828446	2.593	Down-regulated (↓)
11	indole-3-acetyl-beta-4-D-glucose	338.1267	3.518	Organoheterocyclic compounds	Indoles and derivatives	710910.0746	378170.0364	1.88	0.911	1281	0.008489023	3.253	Up-regulated (↑)
12	Niveusin C	377.1606	7.843	Lipids and lipid-like molecules	Prenol lipids	384825.2199	187006.3077	2.058	1.041	1280	0.008698091	2.35	Up-regulated (↑)
13	5-(3',4',5'-trihydroxyphenyl)-gamma-valerolactone-O-methyl-5'-O-glucuronide	459.1146	3.767	Organic oxygen compounds	Organooxygen compounds	644084.1732	381966.4004	1.686	0.754	1275	0.009814007	2.624	Up-regulated (↑)
14	Semiglabrinol	385.0857	8.718	Phenylpropanoids and polyketides	Flavonoids	73388.4421	162328.775	0.452	-1.145	659	0.012147689	3.098	Down-regulated (↓)
15	[3h]Thienylcyclohexylpiperidine	296.1697	6.058	Organic nitrogen compounds	Organonitrogen compounds	491559.6049	775927.9663	0.634	-0.659	670	0.015658408	2.216	Down-regulated (↓)
16	Isonicotinylglycine	198.0875	1.992	Organic acids and derivatives	Carboxylic acids and derivatives	3877182.096	7942291.39	0.488	-1.035	675	0.017530018	2.639	Down-regulated (↓)
17	12-Hydroxydodecanoic acid	197.1546	7.889	Organic acids and derivatives	Hydroxy acids and derivatives	144071.4594	68434.5846	2.105	1.074	1246	0.019165725	1.643	Up-regulated (↑)
18	Mocetinostat	431.1384	6.215	Benzenoids	Benzene and substituted derivatives	296243.0687	456045.6674	0.65	-0.622	684	0.021396801	2.143	Down-regulated (↓)
19	Ractopamine hydrochloride	372.1121	4.763	Benzenoids	Phenols	15475693.35	5894933.48	2.625	1.392	1241	0.021396801	2.977	Up-regulated (↑)
20	L-cis-3-Amino-2-pyrrolidinecarboxylic acid	113.0713	0.756	Organic acids and derivatives	Carboxylic acids and derivatives	3230009.596	2132036.484	1.515	0.599	1239	0.022350662	1.332	Up-regulated (↑)
21	Tubastatin A	374.1271	4.763	Organoheterocyclic compounds	Indoles and derivatives	4640244.96	1821118.019	2.548	1.349	1239	0.022350662	2.481	Up-regulated (↑)
22	Gum arabic	319.163	6.058	Benzenoids	Benzene and substituted derivatives	2726083.295	4723016.388	0.577	-0.793	687	0.022841306	2.191	Down-regulated (↓)
23	5,7,4',5'-Tetrahydroxy-3,6,8,2'-tetramethoxyflavone	445.0568	7.134	Phenylpropanoids and polyketides	Flavonoids	1783767.462	958760.1432	1.86	0.896	1236	0.023850729	2.258	Up-regulated (↑)
24	O-Demethylmetoprolol	271.2017	4.847	Benzenoids	Phenols	3872085.574	6460300.321	0.599	-0.738	690	0.024369794	1.977	Down-regulated (↓)

25	Paraxanthine	181.0721	0.81	Organoheterocyclic compounds	Imidazopyrimidines	22786371.12	14990538.82	1.52	0.604	1235	0.024369794	2.48	Up-regulated (↑)
26	PKHdiA-PG	683.3039	9.153	Lipids and lipid-like molecules	Glycerophospholipids	317228.3458	130803.0565	2.425	1.278	1235	0.024369794	1.839	Up-regulated (↑)
27	4-Oxo-enoxacin	369.0758	5.283	Organoheterocyclic compounds	Diazanaphthalenes	1403009.32	2110699.996	0.665	-0.589	691	0.024898621	2.049	Down-regulated (↓)
28	3-Pyroglutamylthiazolidine-4-carboxylic acid	245.0593	3.783	Organic acids and derivatives	Carboxylic acids and derivatives	1103940.334	682162.7917	1.618	0.694	1233	0.025437357	2.39	Up-regulated (↑)
29	Piperacilloic Acid	536.1771	5.836	Organic acids and derivatives	Carboxylic acids and derivatives	3129742.267	1474440.139	2.123	1.086	1229	0.027694372	2.487	Up-regulated (↑)
30	Benzoic acid, 2-benzoyl-	451.1195	7.626	Benzenoids	Benzene and substituted derivatives	293560.7445	462645.8115	0.635	-0.656	696	0.027694372	2.114	Down-regulated (↓)
31	2-Thenoyltrifluoroacetone	266.9945	5.37	Organic oxygen compounds	Organooxygen compounds	2080419.044	1300576.44	1.6	0.678	1228	0.028284901	2.08	Up-regulated (↑)
32	2-[(3,4-dimethoxyphenyl)methylene]-6-hydroxy-3(2H)-benzofuranone	297.0758	3.834	Phenylpropanoids and polyketides	Aurone flavonoids	698612.9613	308246.4801	2.266	1.18	1228	0.028284901	2.181	Up-regulated (↑)
33	Gly-Glu-Trp	391.1614	3.947	Organic acids and derivatives	Carboxylic acids and derivatives	986583.6758	650894.0291	1.516	0.6	1227	0.028886248	1.825	Up-regulated (↑)
34	Andromedol	388.2694	4.829	Lipids and lipid-like molecules	Prenol lipids	639730.3785	374730.0421	1.707	0.772	1224	0.030756765	1.613	Up-regulated (↑)
35	Withaperuvin	521.2728	6.934	Lipids and lipid-like molecules	Steroids and steroid derivatives	1679750.804	2535856.16	0.662	-0.594	701	0.030756765	2.339	Down-regulated (↓)
36	nitrophenol sulfate	217.9763	4.912	Organic acids and derivatives	Organic sulfuric acids and derivatives	3049058.497	1896964.063	1.607	0.685	1223	0.031402959	1.349	Up-regulated (↑)
37	2-Ethyl-5-methylpyridine	139.1231	3.922	Organoheterocyclic compounds	Pyridines and derivatives	1310010.614	721575.3341	1.815	0.86	1222	0.032060765	2.297	Up-regulated (↑)
38	N-(7-methyloctanoyl)homoserine lactone	242.1752	7.917	Organic acids and derivatives	Carboxylic acids and derivatives	4255190.571	1024186.607	4.155	2.055	703	0.032060765	1.087	Up-regulated (↑)
39	(R)-Marmin	367.1309	7.626	Phenylpropanoids and polyketides	Coumarins and derivatives	1126532.416	1900110.913	0.593	-0.754	704	0.032730346	2.387	Down-regulated (↓)
40	Orientanol B	405.1135	7.48	Phenylpropanoids and polyketides	Isoflavonoids	201018.9195	313698.9335	0.641	-0.642	704	0.032730346	2.244	Down-regulated (↓)
41	Phenylalanine glutamate	311.124	2.581	Organic acids and derivatives	Carboxylic acids and derivatives	3953215.752	2627915.946	1.504	0.589	1221	0.032730346	2.284	Up-regulated (↑)
42	Isorhamnetin 3-[6"-(2-(E)-butenoyl)-glucoside]	547.1454	6.328	Phenylpropanoids and polyketides	Flavonoids	1074706.595	672520.6489	1.598	0.676	1220	0.033411867	2.213	Up-regulated (↑)
43	Soquinolol	305.187	6.649	Organoheterocyclic compounds	Tetrahydroisoquinolines	2303077.865	1195321.449	1.927	0.946	1218	0.034811388	2.361	Up-regulated (↑)
44	Nalbuphine	380.1852	1.453	Benzenoids	Phenanthrenes and derivatives	1133515.287	2160641.641	0.525	-0.931	710	0.037004388	2.067	Down-regulated (↓)
45	3,4-Dimethylhippuric acid	206.0821	5.716	Benzenoids	Benzene and substituted derivatives	1790615.549	1035324.134	1.73	0.79	1214	0.037761059	1.86	Up-regulated (↑)
46	Leu-Gln-Tyr	403.1972	4.023	Organic acids and derivatives	Carboxylic acids and derivatives	5843918.517	2753915.576	2.122	1.085	1214	0.037761059	2.114	Up-regulated (↑)
47	Orientin 2"-O-(2'"-methylbutyrate)	531.1507	6.112	Phenylpropanoids and polyketides	Flavonoids	675826.0993	354503.5022	1.906	0.931	1213	0.038530854	1.791	Up-regulated (↑)
48	Allura red AC	451.0317	4.532	Benzenoids	Naphthalenes	1799806.503	1049808.526	1.714	0.778	1213	0.038530854	2.55	Up-regulated (↑)
49	Methylene bisacrylamide	172.1083	1.161	Organic acids and derivatives	Carboxylic acids and derivatives	3047571.565	1916311.892	1.59	0.669	1212	0.039313945	2.048	Up-regulated (↑)
50	Echinochrome A	305.0068	5.087	Benzenoids	Naphthalenes	862414.8453	2325908.268	0.371	-1.431	713	0.039313945	2.345	Down-regulated (↓)
51	8-Hydroxyondansetron	332.1341	2.554	Organoheterocyclic compounds	Indoles and derivatives	939801.9196	2283458.041	0.412	-1.281	714	0.040110508	2.491	Down-regulated (↓)
52	Tetradecenoylcarnitine	370.2954	9.231	Lipids and lipid-like molecules	Fatty Acyls	3139514.965	1569619.667	2	1	1211	0.040110508	2.139	Up-regulated (↑)
53	4'-Azidocytidine	285.0936	0.676	Organic oxygen compounds	Organooxygen compounds	2230371.606	1279023.581	1.744	0.802	1210	0.04092072	2.51	Up-regulated (↑)

54	Ethyl beta-D-glucopyranoside	231.0841	1.144	Organic oxygen compounds	Organooxygen compounds	28201866.4	18417106.94	1.531	0.615	1209	0.041744757	1.694	Up-regulated (↑)
55	2-Mercaptonicotinic acid	200.0022	3.695	Organoheterocyclic compounds	Pyridines and derivatives	420877.1866	132176.8393	3.184	1.671	1208	0.042582798	2.075	Up-regulated (↑)
56	AKOS017839944	169.0617	0.735	Organic acids and derivatives	Carboxylic acids and derivatives	3352276.966	5292536.563	0.633	-0.659	717	0.042582798	1.795	Down-regulated (↓)
57	Tenatoprazole	381.0803	2.557	Organoheterocyclic compounds	Imidazopyridines	478205.5019	826273.9838	0.579	-0.789	719	0.044301611	2.813	Down-regulated (↓)
58	Erosenone	355.0816	4.45	Phenylpropanoids and polyketides	Isoflavonoids	1119642.118	486764.6196	2.3	1.202	1204	0.046078609	2.445	Up-regulated (↑)
59	Succinobucol	639.3132	9.157	Benzenoids	Phenol esters	1141540.297	486293.8537	2.347	1.231	1204	0.046078609	2.148	Up-regulated (↑)
60	Tetradeca-8,10,12-trienedioylcarnitine	440.2264	7.132	Lipids and lipid-like molecules	Fatty Acyls	827894.0709	480370.3543	1.723	0.785	1203	0.046989386	2.322	Up-regulated (↑)
61	1,15-Dihydroxy-herbertene	257.1497	4.803	Lipids and lipid-like molecules	Prenol lipids	154884293.5	386108011.6	0.401	-1.318	722	0.046989386	2.17	Down-regulated (↓)
62	Crispolide	261.1129	6.953	Lipids and lipid-like molecules	Prenol lipids	257142.6532	170888.2916	1.505	0.59	1203	0.046989386	2.389	Up-regulated (↑)
63	S-Prenyl-L-cysteine	190.0899	3.562	Organic acids and derivatives	Carboxylic acids and derivatives	5089072.455	8809450.721	0.578	-0.792	722	0.046989386	2.156	Down-regulated (↓)
64	Cotinine	177.1024	1.11	Organoheterocyclic compounds	Pyridines and derivatives	77007577.12	17552262.55	4.387	2.133	1202	0.04791526	3.241	Up-regulated (↑)
65	4-(Methylsulfanyl)-2-oxobutanoylcarnitine	309.1477	1.466	Lipids and lipid-like molecules	Fatty Acyls	488884.637	921089.2812	0.531	-0.914	724	0.048856417	2.01	Down-regulated (↓)
66	Guanfacine	244.0044	4.592	Benzenoids	Benzene and substituted derivatives	12005975.47	4965521.696	2.418	1.274	1200	0.049813044	2.884	Up-regulated (↑)

Metabolites were screened using the combined criteria of raw p-value < 0.05 (Mann-Whitney U test), Variable Importance in Projection (VIP) > 1 (from OPLS-DA model), and |fold change| > 1.5 (depression group / control group). The table lists, for each metabolite: sequential number, metabolite name, accurate m/z ratio, retention time (RT, min), ClassyFire Super Class and Class classification, mean relative peak area in the depression group and control group, fold change (FC, Depression/Control), $\log_2$-transformed fold change ($\log_2$FC), Mann-Whitney U statistic, raw p-value, VIP score, and direction of change (↑ upregulated in depression, ↓ downregulated in depression). Of the 66 metabolites, 39 (59.1%) were upregulated and 27 (40.9%) were downregulated. Notably, 12 metabolites (18.2%) were identified as drug molecules or pharmaceutical compounds (e.g., MDMA, flupentixol, mocetinostat, ractopamine, tubastatin A, piretamide, bendamustine), which likely represent false-positive identifications or exogenous exposures rather than endogenous disease biomarkers. All p-values are raw and uncorrected; no metabolite remained significant after Benjamini-Hochberg multiple testing correction across all 5,932 metabolites.

Table 3. Differential rate analysis of metabolites by ClassyFire chemical classification.

Super Class	Total metabolites	Differential metabolites	Up-regulated	Down-regulated	Differential rate (%)
Phenylpropanoids and polyketides	477	10	5	5	2.1
Benzenoids	774	12	5	7	1.55
Organic nitrogen compounds	81	1	0	1	1.23
Organic acids and derivatives	1225	15	10	5	1.22
Organoheterocyclic compounds	1225	15	9	6	1.22
Organic oxygen compounds	559	4	4	0	0.72
Lipids and lipid-like molecules	1338	9	6	3	0.67
Alkaloids and derivatives	71	0	0	0	0
Acetylides	1	0	0	0	0
Homogeneous non-metal compounds	7	0	0	0	0
Hydrocarbon derivatives	1	0	0	0	0
Organic 1,3-dipolar compounds	6	0	0	0	0
Nucleosides, nucleotides, and analogues	117	0	0	0	0
Hydrocarbons	3	0	0	0	0
Lignans, neolignans and related compounds	22	0	0	0	0
Organohalogen compounds	4	0	0	0	0
Organophosphorus compounds	2	0	0	0	0
Organosulfur compounds	15	0	0	0	0
Unclassified	4	0	0	0	0

(a) Super Class level. Summary of all 19 Super Classes, showing the total number of identified metabolites, the number of exploratory differential metabolites, the number of upregulated and downregulated metabolites, and the differential rate (%). Differential rate was calculated as the number of differential metabolites divided by the total number of metabolites in each category, multiplied by 100. Seven Super Classes contained at least one differential metabolite; the remaining 12 categories had zero differential metabolites. The highest differential rate was 2.10% (Phenylpropanoids and polyketides), and all categories were below the 5% expected-by-chance threshold.

Class	Total metabolites	Differential metabolites	Up-regulated	Down-regulated	Differential rate (%)
Tetrahydroisoquinolines	10	1	1	0	10
Benzodioxoles	21	1	1	0	4.76
Isoflavonoids	48	2	1	1	4.17
Glycerophospholipids	25	1	1	0	4
Naphthalenes	55	2	1	1	3.64
Flavonoids	170	6	3	3	3.53
Indoles and derivatives	125	4	2	2	3.2
Diazanaphthalenes	38	1	0	1	2.63
Pyridines and derivatives	116	3	3	0	2.59
Coumarins and derivatives	39	1	0	1	2.56
Phenols	109	2	1	1	1.83
Hydroxy acids and derivatives	58	1	1	0	1.72
Azoles	59	1	0	1	1.69
Imidazopyrimidines	65	1	1	0	1.54
Diazines	69	1	1	0	1.45
Carboxylic acids and derivatives	971	13	8	5	1.34
Organonitrogen compounds	81	1	0	1	1.23
Benzene and substituted derivatives	497	6	2	4	1.21
Prenol lipids	336	4	3	1	1.19
Organic sulfuric acids and derivatives	86	1	1	0	1.16
Organooxygen compounds	557	4	4	0	0.72
Fatty Acyls	652	3	2	1	0.46
Steroids and steroid derivatives	292	1	0	1	0.34
Phenylpropanoic acids	27	0	0	0	0
Phenol ethers	53	0	0	0	0
Peptidomimetics	17	0	0	0	0
Piperidines	40	0	0	0	0
Oxazinanes	10	0	0	0	0
Pteridines and derivatives	28	0	0	0	0

Purine nucleosides	40	0	0	0	0
Pyrans	16	0	0	0	0
Pyrimidine nucleosides	24	0	0	0	0
Pyrrolidines	15	0	0	0	0
Anthracenes	11	0	0	0	0
Stilbenes	20	0	0	0	0
Tetralins	15	0	0	0	0
Tetrapyrroles and derivatives	11	0	0	0	0
Triazines	10	0	0	0	0
Tropane alkaloids	10	0	0	0	0
Purine nucleotides	14	0	0	0	0
Benzopyrans	49	0	0	0	0
Unclassified	18	0	0	0	0
Azolidines	16	0	0	0	0
Benzazepines	20	0	0	0	0
Benzimidazoles	12	0	0	0	0
Benzodiazepines	21	0	0	0	0
Benzofurans	14	0	0	0	0
Benzothiazines	20	0	0	0	0
Benzoxazines	12	0	0	0	0
Cinnamic acids and derivatives	62	0	0	0	0
Diarylheptanoids	11	0	0	0	0
Diazinanes	38	0	0	0	0
Lactams	13	0	0	0	0
Epoxides	11	0	0	0	0
Glycerolipids	22	0	0	0	0
Heteroaromatic compounds	14	0	0	0	0
Indanes	10	0	0	0	0
Keto acids and derivatives	49	0	0	0	0
Quinolines and derivatives	82	0	0	0	0
Lactones	30	0	0	0	0

Linear 1,3-diarylpropanoids	16	0	0	0	0
Macrolides and analogues	24	0	0	0	0
Nucleoside and nucleotide analogues	11	0	0	0	0
Organic phosphoric acids and derivatives	10	0	0	0	0
Organic sulfonic acids and derivatives	17	0	0	0	0
Dihydrofurans	21	0	0	0	0

(b) Class level. Summary of Class-level categories with ≥10 identified metabolites (66 categories shown), including total metabolite count, differential count, up/down counts, and differential rate (%). Categories are sorted by differential rate in descending order. The top category was Tetrahydroisoquinolines (10.00%, 1/10), but this was based on only 1 differential metabolite out of 10 total. No Class-level category showed a differential rate significantly exceeding 5% in binomial tests after multiple testing correction.

Sub Class	Total metabolites	Differential metabolites	Up-regulated	Down-regulated	Differential rate (%)
N-alkylindoles	5	1	1	0	20
O-methylated isoflavonoids	5	1	1	0	20
Naphthoquinones	5	1	0	1	20
Tyrosols and derivatives	5	1	0	1	20
Methylpyridines	6	1	1	0	16.67
Naphthyridines	6	1	0	1	16.67
Pyrrolidinylpyridines	7	1	1	0	14.29
Carbazoles	14	2	0	2	14.29
Halobenzenes	11	1	1	0	9.09
Naphthalene sulfonic acids and derivatives	14	1	1	0	7.14
1-hydroxy-2-unsubstituted benzenoids	16	1	1	0	6.25
Anilides	19	1	0	1	5.26
Flavonoid glycosides	79	4	2	2	5.06
Flavones	20	1	0	1	5
Indolyl carboxylic acids and derivatives	21	1	1	0	4.76
O-methylated flavonoids	23	1	1	0	4.35
Carboxylic acid derivatives	26	1	1	0	3.85
Tricarboxylic acids and derivatives	27	1	0	1	3.7

Terpene lactones	56	2	2	0	3.57
Imidazoles	30	1	0	1	3.33
Steroid lactones	30	1	0	1	3.33
Pyridinecarboxylic acids and derivatives	33	1	1	0	3.03
Medium-chain hydroxy acids and derivatives	40	1	1	0	2.5
Diterpenoids	48	1	1	0	2.08
Benzoic acids and derivatives	153	3	1	2	1.96
Pyrimidines and pyrimidine derivatives	54	1	1	0	1.85
Purines and purine derivatives	65	1	1	0	1.54
Arylsulfates	70	1	1	0	1.43
Amino acids, peptides, and analogues	873	11	7	4	1.26
Sesquiterpenoids	96	1	0	1	1.04
Unclassified	694	7	4	3	1.01
Carbohydrates and carbohydrate conjugates	330	3	3	0	0.91
Fatty acid esters	221	2	1	1	0.9
Carbonyl compounds	154	1	1	0	0.65
Fatty acids and conjugates	220	1	1	0	0.45
Naphthols and derivatives	13	0	0	0	0
Piperidinones	5	0	0	0	0
Naphthalenecarboxylic acids and derivatives	5	0	0	0	0
Piperazines	38	0	0	0	0
Phosphate esters	10	0	0	0	0
Phenylpyridines	7	0	0	0	0
Nitrobenzenes	18	0	0	0	0
Phenylpropanes	15	0	0	0	0
Monoterpenoids	52	0	0	0	0
Phenylpiperidines	9	0	0	0	0
Phenylacetamides	8	0	0	0	0
Phenoxyacetic acid derivatives	6	0	0	0	0
Organic nitro compounds	6	0	0	0	0

Morpholines	10	0	0	0	0
Organosulfonic acids and derivatives	17	0	0	0	0
Phenothiazines	19	0	0	0	0
Oxazoles	7	0	0	0	0
Phenethylamines	11	0	0	0	0
Oxosteroids	5	0	0	0	0
Nitrophenols	7	0	0	0	0
Steroidal glycosides	44	0	0	0	0
Stilbene glycosides	6	0	0	0	0
Styrenes	6	0	0	0	0
Substituted pyrroles	5	0	0	0	0
Amines	64	0	0	0	0
Aminopyridines and derivatives	8	0	0	0	0
Aminoquinolines and derivatives	11	0	0	0	0
Pyrazoles	12	0	0	0	0
Androstane steroids	23	0	0	0	0
1-benzopyrans	45	0	0	0	0
Sulfated steroids	12	0	0	0	0
Sulfuric acid esters	14	0	0	0	0
Terpene glycosides	63	0	0	0	0
Tetracarboxylic acids and derivatives	11	0	0	0	0
Tetrahydrofuran lignans	5	0	0	0	0
Thiazoles	5	0	0	0	0
Aminotriazines	5	0	0	0	0
Pregnane steroids	20	0	0	0	0
Pterins and derivatives	20	0	0	0	0
Purine ribonucleotides	8	0	0	0	0
Pyranoflavonoids	5	0	0	0	0
Pyranones and derivatives	15	0	0	0	0
Pyrazines	10	0	0	0	0
Monoradylglycerols	8	0	0	0	0

Short-chain keto acids and derivatives	9	0	0	0	0
Pyridoindoles	9	0	0	0	0
Pyrimidine 2',3'-dideoxyribonucleosides	6	0	0	0	0
Pyrimidine 2'-deoxyribonucleosides	6	0	0	0	0
Pyrroloindoles	5	0	0	0	0
Quinoline carboxylic acids	16	0	0	0	0
Quinolones and derivatives	11	0	0	0	0
Quinone and hydroquinone lipids	6	0	0	0	0
Rotenoids	5	0	0	0	0
Pyridazines and derivatives	5	0	0	0	0
Benzoquinolines	20	0	0	0	0
Benzoxazinones	6	0	0	0	0
Benzoyl derivatives	6	0	0	0	0
Benzyl alcohols	5	0	0	0	0
Benzylethers	5	0	0	0	0
Beta hydroxy acids and derivatives	12	0	0	0	0
Beta lactams	13	0	0	0	0
Isoflavonoid O-glycosides	13	0	0	0	0
Biphenyls and derivatives	13	0	0	0	0
Carboximidic acids	6	0	0	0	0
Carboxylic acids	7	0	0	0	0
Chalcones and dihydrochalcones	11	0	0	0	0
Cholestane steroids	11	0	0	0	0
Delta valerolactones	6	0	0	0	0
Dibenzazepines	11	0	0	0	0
Dicarboxylic acids and derivatives	20	0	0	0	0
Bile acids, alcohols and derivatives	70	0	0	0	0
Alcohols and polyols	43	0	0	0	0
Alpha hydroxy acids and derivatives	6	0	0	0	0
Toluenes	8	0	0	0	0
Aniline and substituted anilines	13	0	0	0	0

Vitamin D and derivatives	13	0	0	0	0
1,4-benzodiazepines	15	0	0	0	0
Trifluoromethylbenzenes	10	0	0	0	0
Benzodiazines	32	0	0	0	0
Triterpenoids	11	0	0	0	0
Tryptamines and derivatives	12	0	0	0	0
Anisoles	23	0	0	0	0
Anthraquinones	7	0	0	0	0
Benzenediols	25	0	0	0	0
Benzenesulfonamides	25	0	0	0	0
Benzenesulfonic acids and derivatives	10	0	0	0	0
Benzenesulfonyl compounds	5	0	0	0	0
Triradylcglycerols	5	0	0	0	0
Hydroxysteroids	34	0	0	0	0
Imidazolidines	10	0	0	0	0
Indanones	7	0	0	0	0
Indolecarboxylic acids and derivatives	7	0	0	0	0
Indoles	23	0	0	0	0
Indolines	11	0	0	0	0
Isoflav-2-enes	7	0	0	0	0
Diphenylethers	7	0	0	0	0
1-hydroxy-4-unsubstituted benzenoids	6	0	0	0	0
Isoindolines	5	0	0	0	0
Linear diarylheptanoids	7	0	0	0	0
Lineolic acids and derivatives	21	0	0	0	0
Lysergic acids and derivatives	6	0	0	0	0
Medium-chain keto acids and derivatives	32	0	0	0	0
Methoxybenzenes	17	0	0	0	0
Methoxyphenols	43	0	0	0	0
Isoflavans	9	0	0	0	0
Diphenylmethanes	33	0	0	0	0

Eicosanoids	59	0	0	0	0
Estrane steroids	19	0	0	0	0
Ethers	27	0	0	0	0
Fatty acyl glycosides	59	0	0	0	0
Fatty alcohols	44	0	0	0	0
Fatty amides	19	0	0	0	0
Flavans	37	0	0	0	0
Hydroxycoumarins	5	0	0	0	0
Furanones	18	0	0	0	0
Gamma butyrolactones	22	0	0	0	0
Gamma-keto acids and derivatives	6	0	0	0	0
Glycerophosphocholines	10	0	0	0	0
Hybrid peptides	15	0	0	0	0
Hydrolyzable tannins	6	0	0	0	0
Hydropyridines	22	0	0	0	0
Hydroxycinnamic acids and derivatives	51	0	0	0	0
Furanocoumarins	9	0	0	0	0

(c) Sub Class level. Summary of Sub Class-level categories (159 categories with ≥5 metabolites shown), including total metabolite count, differential count, up/down counts, and differential rate (%). Categories are sorted by differential rate in descending order. Four categories achieved a 20% differential rate (N-alkylindoles, O-methylated isoflavonoids, naphthoquinones, tyrosols and derivatives), but each contained only 5 total metabolites with 1 differential hit, making these findings statistically unstable and likely due to chance. No Sub Class-level category survived multiple testing correction.

Supplementary Tables

Table S1. Quality control (QC) sample RSD summary for all 5,932 identified metabolites.

This table contains three sheets. Sheet 1 (Overall_Summary): Descriptive statistics of the relative standard deviation (RSD, %) of peak area across 15 pooled QC samples for all 5,932 metabolites. Key statistics include: total metabolites = 5,932; mean RSD = 0.97%; median RSD = 0.81%; standard deviation of RSD = 0.63%; minimum RSD = 0.07%; maximum RSD = 6.33%; Q1 = 0.49%; Q3 = 1.30%. Sheet 2 (RSD_Distribution): Distribution of metabolites across RSD ranges: <5% (n = 5,927, 99.92%), 5–10% (n = 5, 0.08%), 10–15% (n = 0), 15–20% (n = 0), 20–25% (n = 0), 25–30% (n = 0), >30% (n = 0). All 5,932 metabolites (100%) had RSD < 30%, satisfying the widely accepted metabolomics quality control criterion. Sheet 3 (Per_Metabolite_RSD): Per-metabolite RSD values, including metabolite ID, name, ion mode, m/z, retention time, QC mean peak area, QC standard deviation, RSD (%), and a binary indicator of whether RSD < 30%.

Table S2. Metabolomics Standards Initiative (MSI) identification level distribution.

This table contains three sheets. Sheet 1 (Level_Distribution): Distribution of all 5,932 metabolites across four MSI identification confidence levels: Level 1 confirmed structure by reference standard with MS/MS spectrum and retention time match (n = 1,127, 19.00%); Level 2 putatively annotated by MS/MS spectral library match (n = 1,392, 23.47%); Level 3 putatively characterized by exact mass and chemical class only (n = 2,216, 37.36%); Level 4 unknown or unclassified (n = 1,197, 20.18%). Sheet 2 (Level_by_IonMode): Cross-tabulation of MSI levels by ion mode (positive ESI+ vs. negative ESI−). Sheet 3 (All_Metabolites_by_Level): Per-metabolite identification details, including metabolite ID, name, ion mode, MSI level, identification score, Super Class, and Class.

Table S3. Summary of metabolites by ion mode (ESI+ and ESI−).

This table contains three sheets. Sheet 1 (Ion_Mode_Summary): Overall summary of metabolites detected in each ion mode: positive mode (ESI+) n = 3,230 (54.45%); negative mode (ESI−) n = 2,702 (45.55%); total n = 5,932 (100%). Sheet 2 (By_SuperClass): Number of metabolites in each ion mode, cross-tabulated by Super Class classification, showing the distribution of chemical classes across ionization modalities. Sheet 3 (By_MSI_Level): Number of metabolites in each ion mode, cross-tabulated by MSI identification level, allowing assessment of identification quality differences between positive and negative modes.

Table S4. Complete statistical results for all 5,932 identified metabolites.

Full statistical results for every identified metabolite (n = 5,932), including: sequential number, metabolite ID, metabolite name (English and Chinese), ion mode, accurate m/z, retention time (min), HMDB ID, KEGG ID, ClassyFire Super Class/Class/Sub Class, MSI identification level, mean $\log_{10}$-transformed abundance in depression group and control group, mean difference ($\log_{10}$ scale), back-transformed fold change (Depression/Control), $\log_2$FC, Mann-Whitney U statistic, raw p-value, Benjamini-Hochberg adjusted p-value, Bonferroni adjusted p-value, and binary indicators for p_raw < 0.05, p_adj_BH < 0.05, $|\log_2 FC| > \log_2(1.5)$, and exploratory differential status ($p < 0.05$ & FC > 1.5). This table serves as the complete statistical audit trail for the study.

Table S5. Complete list of 66 exploratory differential metabolites with extended annotation.

Extended annotation table for the 66 exploratory differential metabolites, including all fields from Table 2 plus: Benjamini-Hochberg adjusted p-value calculated across the full set of 5,932 metabolites (p_adj_BH_full, all = 1.0), Benjamini-Hochberg adjusted p-value calculated within the 66 metabolites only (p_adj_BH_66), biological/source classification (putative endogenous/unknown, diet-derived phytochemical, drug/pharmaceutical likely false-positive, exogenous/food additive, smoking-related confounder), and MSI identification level status (to be confirmed by MS/MS validation). Of the 66 metabolites, 30 (45.5%) were classified as putative endogenous/unknown, 16 (24.2%) as diet-derived phytochemicals, 12 (18.2%) as drug/pharmaceutical compounds, 6 (9.1%) as exogenous/food additives, and 2 (3.0%) as smoking-related confounders. This classification highlights the substantial proportion of non-endogenous signals among the exploratory findings.

Table S6. Classification summary of 66 candidate metabolites by biological source.

Summary table categorizing the 66 exploratory differential metabolites into five biological/source classes: (1) Putative endogenous/unknown (n = 30, 45.5%; 20 upregulated, 10 downregulated; mean p_raw = 0.0279, min p_raw = 0.0025; mean $\log_2$FC = 0.313; mean VIP = 2.194); (2) Diet-derived phytochemical (n = 16, 24.2%; 8 up, 8 down; mean p_raw = 0.0264, min p_raw = 0.0049; mean $\log_2$FC = 0.051; mean VIP = 2.305); (3) Drug/Pharma likely false-positive ID (n = 12, 18.2%; 5 up, 7 down; mean p_raw = 0.0273, min p_raw = 0.0012; mean $\log_2$FC = 0.156; mean VIP = 2.589); (4) Exogenous/Food additive (n = 6, 9.1%; 4 up, 2 down; mean p_raw = 0.0318, min p_raw = 0.0157; mean $\log_2$FC = 0.317; mean VIP = 2.206); (5) Smoking-related confounder (n = 2, 3.0%; both upregulated; mean p_raw = 0.0361, min p_raw = 0.0244; mean $\log_2$FC = 1.369; mean VIP = 2.865). This table quantitatively demonstrates that less than half of

the exploratory findings are potentially endogenous metabolites, with the majority representing dietary, pharmaceutical, or environmental exposures.

Table S7. Binomial tests for differential rate enrichment at each classification level.

(a) Super Class level. One-sided and two-sided binomial tests were performed for each of the 19 Super Classes to assess whether the observed differential rate significantly exceeded the 5% expected-by-chance threshold. The table includes: Super Class name, total metabolites, differential count, up/down counts, observed differential rate (%), expected count by chance at 5%, one-sided binomial p-value (testing differential rate > 5%), two-sided binomial p-value, and a significance indicator. No Super Class showed a differential rate significantly exceeding 5% after consideration of multiple testing (all one-sided $p > 0.05$; in fact, all observed rates were below 5%).

(b) Class level. Binomial tests at the Class level for all categories with ≥10 metabolites. Same structure as Table S7a. No Class-level category showed statistically significant enrichment after multiple testing correction, despite some categories showing elevated point estimates driven by small counts.

(c) Sub Class level. Binomial tests at the Sub Class level for all categories with ≥5 metabolites. Same structure as Table S7a. No Sub Class-level category showed statistically significant enrichment after multiple testing correction. Categories with the highest point estimates (e.g., 20%) were based on only 1 differential metabolite out of 5 total, making them highly unstable.

Table S8. Post-hoc statistical power analysis for the two-sample Mann-Whitney U test.

Formal post-hoc power analysis was conducted for the two-sample Mann-Whitney U test (equivalent to Cohen's d effect size) at the Benjamini-Hochberg adjusted significance threshold, accounting for the total sample size of n = 90 (55 depression, 35 control). The table presents, for each effect size (Cohen's d = 0.1 to 2.0): effect size interpretation (very small to huge), non-centrality parameter delta, statistical power at the current sample size of n = 90 (%), required total sample size for 80% power, required depression group sample size, required control group sample size (assuming 55:35 allocation ratio), and required total sample size for 90% power. Key findings: at a medium effect size (d = 0.5), the current study had only 1.61% power, and 473 total subjects (289 depression, 184 control) would be required for 80% power. At a small effect size (d = 0.2), 2,951 total subjects would be needed. Only at very large effect sizes (d ≥ 1.5) did the current sample size achieve adequate power (>99%). This analysis formally demonstrates that the study was substantially underpowered for detecting small-to-medium metabolic effect sizes after multiple testing correction.

Supplementary Figures

Figure S1. PCA score plot including pooled quality control (QC) samples.

PCA score plot showing the distribution of all study samples in the principal component space, including 35 healthy controls (blue squares), 55 depression patients (red circles), 13 natural depression samples not included in the main analysis (orange triangles), 15 pooled QC samples (green diamonds), and the QC mean (green cross). PC1 explained 25.9% and PC2 explained 6.8% of the total variance. The 15 pooled QC samples clustered tightly around the QC mean at the center of the plot, demonstrating excellent instrument stability and data reproducibility throughout the analytical run. The inset text box summarizes QC performance: mean RSD = 0.97%, median RSD = 0.81%, maximum RSD = 6.33%, with all 5,932 metabolites passing the RSD < 30% criterion.

Figure S2. Distribution of relative standard deviation (RSD) in pooled QC samples.

(A) QC sample RSD distribution. Histogram showing the distribution of RSD (%) values across all 5,932 metabolites, calculated from 15 pooled QC samples. The distribution is strongly right-skewed, with the vast majority of metabolites showing very low RSD values. The blue vertical line indicates the median RSD (0.81%); the orange dashed line indicates RSD = 20%; the red dashed line indicates the acceptance threshold of RSD = 30%. The text box provides summary statistics: mean = 0.97%, median = 0.81%, SD = 0.63%, min = 0.07%, max = 6.33%, Q1 = 0.49%, Q3 = 1.30%.

(B) RSD category breakdown. Bar chart (green bars, left y-axis) showing the number and percentage of metabolites in each RSD range category: <5% (n = 5,927, 99.9%), 5–10% (n = 5, 0.1%), 10–15% (n = 0), 15–20% (n = 0), 20–25% (n = 0), 25–30% (n = 0), >30% (n = 0). The blue line with circles (right y-axis) shows the cumulative percentage. All metabolites (100%) passed the RSD < 30% acceptance criterion, confirming high-quality and reproducible data acquisition.

Figure S3. Missing value pattern heatmap before KNN imputation.

Heatmap visualizing the pattern of missing values (zero peak area values treated as missing) across all 5,932 metabolites and 90 samples prior to K-Nearest Neighbors Imputation (KNNI). Rows represent metabolites and columns represent samples, with black indicating missing values and white indicating detected values. The missing value pattern was assessed to determine the appropriateness of KNN imputation (k = 5, distance-weighted). The overall missing rate was low, and missing values appeared to be randomly distributed rather than structured by sample group, supporting the use of KNNI. Samples are ordered by group (depression vs. control), and metabolites are ordered by hierarchical clustering.

Figure S4. p-value diagnostic plots for the 66 exploratory differential metabolites.

(A) Q-Q plot of p-values. Quantile-quantile (Q-Q) plot comparing the observed distribution of raw p-values (from Mann-Whitney U tests) for the 66 exploratory differential metabolites against the expected uniform distribution under the null hypothesis. The x-axis shows expected p-values (uniform), and the y-axis shows observed p-values. The black dashed line represents the y = x line (perfect uniform distribution). The observed p-values (blue circles) deviate substantially from the diagonal, with all values concentrated near 0, confirming that these 66 metabolites were selected precisely because their p-values fall below 0.05. This plot is descriptive of the selected subset and should not be interpreted as evidence of global p-value inflation.

(B) Distribution of p-values. Histogram showing the distribution of raw p-values for the 66 exploratory differential metabolites, binned into 20 bins from 0 to 0.05. The black dashed line indicates the expected count under a uniform distribution (n = 3.3 per bin). The observed distribution shows an enrichment of very small p-values ($p < 0.01$), with the highest frequency in the 0.02–0.03 and 0.04–0.05 ranges. This pattern reflects the selection threshold and does not indicate genuine enrichment of true associations, as all p-values are below the arbitrary 0.05 cutoff.

Figure S5. Uniformity test of p-value distribution for all 5,932 metabolites.

Comprehensive diagnostic assessment of the p-value distribution across all 5,932 identified metabolites, including: (1) a histogram of raw p-values ranging from 0 to 1, overlaid with the expected uniform distribution (dashed line); (2) a Q-Q plot comparing observed p-value quantiles against expected uniform quantiles; (3) results of formal statistical tests for uniformity, including the Kolmogorov-Smirnov test and Anderson-Darling test. Under the global null hypothesis of no true metabolic differences, p-values should follow a uniform distribution. The observed distribution was approximately uniform with a slight enrichment of small p-values, consistent with the presence of a small number of weak associations or residual confounding. However, the degree of enrichment was insufficient to survive multiple testing correction, supporting the study's negative finding at the strict statistical level.

Figure S6. VIP score distribution and VIP-significance association.

(A) VIP score distribution histogram. Histogram showing the distribution of Variable Importance in Projection (VIP) scores from the OPLS-DA model across all 5,932 metabolites. The vertical dashed line indicates the conventional VIP > 1 threshold. The distribution is approximately right-skewed, with a subset of metabolites showing VIP scores > 1. The number and percentage of metabolites exceeding the VIP > 1 threshold are reported.

(B) VIP-significance association. Scatter plot showing the relationship between VIP score (x-axis) and $-\log_{10}$(raw p-value) (y-axis) for all 5,932 metabolites. Metabolites

meeting the combined exploratory criteria ($p < 0.05$ and VIP > 1) are highlighted. This plot visualizes the concordance between univariate statistical significance and multivariate model importance, demonstrating that the 66 exploratory differential metabolites represent the intersection of these two criteria.

Figure S7. Forest plot of the top 20 exploratory differential metabolites by p-value.

Forest plot showing the $\log_2$(fold change) (depression/control) and associated raw p-values for the top 20 exploratory differential metabolites ranked by statistical significance (smallest p-value). Red horizontal bars represent metabolites upregulated in the depression group; blue bars represent metabolites downregulated in the depression group. The vertical black solid line at $\log_2FC = 0$ indicates no change; the vertical gray dashed lines at $\log_2FC = \pm\log_2(1.5) \approx \pm 0.585$ indicate the fold-change threshold. Each metabolite is labeled with its raw p-value at the end of its bar. Metabolites are ordered by p-value from top (most significant) to bottom. The top-ranked metabolite was N,N-dimethyl-3,4-methylenedioxyamphetamine (MDMA, $\log_2FC = 2.55$, $p = 0.0012$), a recreational drug molecule, highlighting the dominance of exogenous compounds among the most statistically significant findings.

Figure S8. Confounder and exogenous compound annotation analysis.

(A) Confounder and exogenous compound annotation. Volcano-style scatter plot of the 66 exploratory differential metabolites, colored and shaped by biological/source classification: putative endogenous/unknown (gray circles, n = 47), exogenous/drug-like (orange circles, n = 17), and smoking-related (red stars, n = 2). The x-axis shows $\log_2$(fold change) and the y-axis shows $-\log_{10}$(p-value). Key exogenous and drug-like metabolites are labeled, including N,N-dimethyl-3,4-methylenedioxyamphetamine (MDMA), flupentixol, mocetinostat, ractopamine hydrochloride, [3H]thienylcyclohexylpiperidine, paraxanthine, and cotinine. This plot visually demonstrates that a substantial proportion of the most significant findings are non-endogenous compounds.

(B) Smoking confounder assessment. Bar chart quantifying the smoking-related confounding between groups. The depression group had a 4.0-fold higher smoking rate than the control group (34.5% vs. 8.6%). Cotinine (the primary nicotine metabolite) showed a 4.4-fold elevation in the depression group (FC = 4.39), and paraxanthine (a caffeine metabolite) showed a 1.5-fold elevation (FC = 1.52). These findings confirm that smoking status represents a significant confounder in this study, and metabolites related to nicotine metabolism (cotinine, paraxanthine) likely reflect between-group differences in smoking behavior rather than depression-specific metabolic alterations.

Figure S9. Composition of up- and down-regulated metabolites (donut chart + SuperClass stacked bar chart).

Left panel: Donut (ring) chart showing the overall proportion of upregulated (n = 39, 59.1%) and downregulated (n = 27, 40.9%) metabolites among the 66 exploratory differential metabolites. The center of the donut displays the total count (n = 66). Right panel: Stacked horizontal bar chart showing the composition of up- and down-regulated metabolites within each Super Class, sorted by total differential count. Each bar represents one Super Class, with the proportion of upregulated (red) and downregulated (blue) metabolites shown. This combined visualization provides both an overview of the global up/down balance and a detailed breakdown by chemical class.

Figure S10. log_2FC vs. significance scatter plot (colored by SuperClass, bubble size = VIP).

Bubble scatter plot of the 66 exploratory differential metabolites, with the x-axis showing log_2(fold change) (depression/control), the y-axis showing -log_{10}(raw p-value), bubble color representing Super Class classification, and bubble size representing the VIP score from the OPLS-DA model. Vertical dashed lines indicate the $|log_2FC| > log_2(1.5)$ threshold, and the horizontal dashed line indicates p = 0.05. This plot simultaneously visualizes four dimensions of information (effect size, statistical significance, chemical class, and multivariate importance) for each differential metabolite, allowing identification of metabolites that are both statistically robust and chemically meaningful.

Figure S11. Heatmap of differential metabolite distribution across SuperClass × Class.

Two-dimensional heatmap showing the number (or proportion) of exploratory differential metabolites across the cross-classification of Super Class (rows) and Class (columns). The color intensity represents the count or differential rate within each SuperClass-Class combination. Row and column annotations indicate the total counts. This heatmap reveals which chemical subcategories are most enriched for differential metabolites and whether the differential signals are concentrated in specific SuperClass-Class combinations or broadly distributed across the chemical space.

Figure S12. Class-level differential rate bubble plot (differential rate vs. metabolite count).

Bubble plot at the Class classification level, with the x-axis showing the total number of identified metabolites in each Class (log scale), the y-axis showing the differential rate (%), and bubble size representing the number of differential metabolites. Only Classes with ≥10 metabolites are shown. The horizontal dashed line indicates the 5% expected-by-chance differential rate. This plot allows simultaneous assessment of both the statistical stability (informed by total metabolite count) and the magnitude of differential rate for each Class, highlighting that categories with high differential rates tend to have small total counts and are therefore statistically unreliable.

Figure S13. SubClass-level differential rate ranked bar chart.

Horizontal bar chart showing the differential rate (%) for SubClass-level categories, sorted in descending order of differential rate. Only SubClasses with ≥5 metabolites are shown. Each bar is labeled with the number of differential metabolites and the total number of metabolites in that SubClass (e.g., "1/5"). The vertical dashed line indicates the 5% expected-by-chance threshold. The top four SubClasses (N-alkylindoles, O-methylated isoflavonoids, naphthoquinones, tyrosols and derivatives) each showed a 20% differential rate, but based on only 1 differential metabolite out of 5 total, emphasizing the instability of these estimates.

Figure S14. Statistical power analysis (power curve + required sample size).

(A) Statistical power curve. Line plot showing the statistical power (%) as a function of Cohen's d effect size, for the current total sample size of n = 90 (55 depression, 35 control) at the Benjamini-Hochberg adjusted significance threshold. The x-axis shows effect size (d = 0.1 to 2.0), and the y-axis shows power (%). The horizontal dashed line indicates the conventional 80% power threshold. The curve demonstrates that power remains very low (<5%) for effect sizes below d = 0.6, and only exceeds 80% for very large effect sizes (d ≥ 1.2). At d = 0.5 (medium effect), power was only 1.61%.

(B) Required sample size. Line plot showing the required total sample size to achieve 80% power (solid line) and 90% power (dashed line) as a function of Cohen's d effect size. The x-axis shows effect size, and the y-axis shows required total sample size (log scale). At d = 0.5 (medium effect), 473 total subjects (289 depression, 184 control) would be required for 80% power. At d = 0.2 (small effect), 2,951 total subjects would be needed. This plot provides a clear visual guide for sample size planning in future metabolomics studies of depression.

Supplementary Files

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

- TableS3.SummaryofmetabolitesbyionmodeESIandESI..xlsx
- TableS6.Classificationsummaryof66candidatemetabolitesbybiologicalsource..csv
- TableS8.PosthocstatisticalpoweranalysisforthetwosampleMannWhitneyUtest..csv
- TableS5.Completelistof66exploratorydifferentialmetaboliteswithextendedannotation..csv
- TableS7b.Binomialtestsfordifferentialrateenrichmentateachclassificationlevel.bClasslevel..csv
- TableS7c.Binomialtestsfordifferentialrateenrichmentateachclassificationlevel.cSubClasslevel..csv
- TableS7a.Binomialtestsfordifferentialrateenrichmentateachclassificationlevel.aSuperClasslevel..csv
- FigureS4.pvaluediagnosticplotsforthe66exploratorydifferentialmetabolites.AQQplotofpvalues.BDistributionofpvalues..png
- FigureS9.CompositionofupanddownregulatedmetabolitesdonutchartSuperClassstackedbarchart..png
- FigureS6.VIPscoredistributionandVIPsignificanceassociation.AVIPscoredistributionhistogram.BVIPsignificanceassociation..png
- FigureS12.Classleveldifferentialratebubbleplotdifferentialratevs.metabolitecount..png
- TableS2.MetabolomicsStandardsInitiativeMSIidentificationleveldistribution..xlsx
- FigureS8.Confounderandexogenouscompoundannotationanalysis.AConfounderandexogenouscompoundannotation.BSmokingconfounderassessment..p
- FigureS10.logFCvs.significancescatterplotcoloredbySuperClassbubblesizeVIP..png
- FigureS7.Forestplotofthetop20exploratorydifferentialmetabolitesbypvalue..png
- FigureS14.Statisticalpoweranalysispowercurverequiredsamplesize..png
- FigureS11.HeatmapofdifferentialmetabolitedistributionacrossSuperClassClass..png
- TableS1.QualitycontrolQCsampleRSDsummaryforall5932identifiedmetabolites..xlsx
- FigureS2.DistributionofrelativestandarddeviationRSDinpooledQCsamples.AQCsampleRSDdistribution.BRSDcategorybreakdown..png
- FigureS13.SubClassleveldifferentialraterankedbarchart..png
- FigureS1.PCAscoreplotincludingpooledqualitycontrolQCsamples..png
- FigureS3.MissingvaluepatternheatmapbeforeKNNimputation..png
- TableS4.Completestatisticalresultsforall5932identifiedmetabolites..csv
- FigureS5.Uniformitytestofpvaluedistributionforall5932metabolites..png