Preprints.org

Article

Not peer-reviewed version

Gut Microbiome Networks Across Schizophrenia, Major Depressive Disorder, and Bipolar Disorder: A Systematic Reanalysis and Machine Learning Study

Parsa Bazdar * , Alireza Daneshmand , Amirhosein Kazemi , Amir Jalilian , Parto Ghahramani

Posted Date: 19 November 2026

Keywords: schizophrenia; bipolar disorder; major depressive disorder; psychiatry; microbiome; 16S; networks; bioinformatics; machine learning

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Gut Microbiome Networks Across Schizophrenia, Major Depressive Disorder, and Bipolar Disorder: A Systematic Reanalysis and Machine Learning Study

Parsa Bazdar [1,*], Alireza Daneshmand [2], Amirhosein Kazemi [3], Amir Jalilian [3] and Parto Ghahramani [3]

[1] Independent Researcher, Ilam, Iran
[2] Independent Researcher, Razavi Khorasan, Iran
[3] Independent Researcher, Kermanshah, Iran
* Correspondence: parsa.bazdar1379@gmail.com

Abstract

Psychiatric disorders represent a leading cause of global disease burden, yet their accurate and timely diagnosis remains challenging due to incomplete understanding of etiology and a lack of objective pathognomonic features. These factors leave diagnosis and treatment vulnerable to diagnostic errors. This diagnostic uncertainty, alongside emerging evidence linking the microbiota-gut-brain axis to psychiatric pathophysiology, has fueled the ongoing search for objective microbial biomarkers. We systematically searched PubMed and Scopus for gut microbiome 16S sequencing studies involving patients with schizophrenia (SCZ), major depressive disorder (MDD), and bipolar disorder (BD) that had publicly available sequencing read data. Data were extracted and analyzed through a harmonized bioinformatics pipeline to assess differences in α- and β-diversity and microbial composition. Co-occurrence networks were constructed, and differences in network topology, hubs, and taxon centrality were investigated. Finally, a random forest (RF) and a convolutional neural network ensemble (CNN) classifier were trained, and their performance in predicting diagnosis from microbiome composition was evaluated using two different cross-validation schemes. Analysis of 1402 samples from 17 studies revealed reduced α-diversity in MDD ($q \leq 0.0027$) and significant β-diversity differences across groups ($q \leq 0.025$), though these were largely accompanied by dispersion differences and extensive group overlap. MDD and SCZ exhibited depletion of butyrate-producing Lachnospiraceae, with SCZ additionally showing Actinobacteriota enrichment; BD exhibited no robust differential abundance. Network analysis showed higher modularity and lower density in SCZ—with MDD showing the inverse pattern—alongside increased centrality of taxa associated with neurotransmitter metabolism or inflammation across disorder groups, and decreased centrality of butyrate producers in MDD and SCZ. Classifiers performed well under 10-fold cross-validation (RF AUC = 0.856; CNN AUC = 0.862) but only modestly under leave-one-study-out cross-validation (RF AUC = 0.762; CNN AUC = 0.656). We identified a broad transdiagnostic signature of diminished butyrate-producer influence and increased prominence of taxa associated with neurotransmitter metabolism or inflammation. While machine learning demonstrates diagnostic potential, it is not yet ready for routine clinical application. Nevertheless, despite statistical significance, effect sizes were small, and group overlap was extensive, with study origin explaining far more variance than diagnosis. Echoing broader inconsistencies in the literature, these findings underscore the need to contextualize results and address criticisms of categorical diagnoses to avoid hyperbolic claims about microbial biomarkers or machine learning in psychiatry.

Keywords: schizophrenia; bipolar disorder; major depressive disorder; psychiatry; microbiome; 16S; networks; bioinformatics; machine learning

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

1. Introduction

Psychiatric disorders represent a leading cause of global disease burden, with mood disorders contributing the most to cumulative disability-adjusted life years. In 2023, approximately 236 million individuals were estimated to be living with major depressive disorder (MDD), while 36 million experienced bipolar disorder (BD). Although schizophrenia (SCZ) is less prevalent than depressive disorders, it ranks third among mental disorders in terms of associated health burden [1,2]. Notably, individuals with SCZ have an average life expectancy nine years shorter than the general population [3]. Consequently, the accurate and timely diagnosis and treatment of psychiatric disorders are of critical importance. However, this task remains challenging due to insufficient understanding of mental disorder etiology, the subjective and self-reported nature of symptoms, significant symptom overlap, and the lack of pathognomonic features distinguishing specific disorders. These factors render psychiatric diagnosis and subsequent treatment vulnerable to various errors, including underdiagnosis, overdiagnosis, and misdiagnosis [4]. For example, studies indicate that between 20.8% and 61.5% of BD patients are misdiagnosed as having MDD [5–8]. Administering antidepressants appropriate for MDD to these patients may induce mood switches and foster treatment resistance, leading to increased healthcare and economic burdens [9]. Similarly, patients with mood disorders exhibiting psychotic features are frequently misdiagnosed with SCZ, resulting in delayed treatment and poorer prognoses. Subsequent evaluations reveal that the initial diagnosis of SCZ changes in 36-51% of cases [10–12]. Additionally, research demonstrates that Black patients are more likely than White patients to be diagnosed with SCZ, highlighting the impact of structural and intersectional biases on the diagnostic process [13–15]. These and other challenges have fueled the ongoing search for objective biomarkers in psychiatry.

In recent years, the association between gut microbiota and mental disorders has garnered significant attention. Numerous mechanisms have been proposed to explain the involvement of gut microbiota in the pathophysiology of mental disorders [16]. These include the modulation of neuroinflammation by microbially derived short-chain fatty acids [17], the regulatory effects of gut microbiota on tryptophan and serotonin metabolism [18], and the influence on vagal signaling [19], among others. Beyond clarifying the causal role of gut microbiota in disorders such as SCZ [20], and MDD [21], a growing body of research has explored the gut microbiota of patients for diagnostic and therapeutic purposes, aiming to identify objective biomarkers and develop novel treatment strategies [22,23].

Accordingly, we performed a systematic search to identify 17 studies that examined the gut microbiome composition of individuals with BD, MDD, and SCZ using publicly available sequencing data. The collected data were aggregated, analyzed, and classified according to psychiatric diagnosis. Co-occurrence networks were constructed for BD, MDD, SCZ, and healthy control (HC) groups, then compared to reveal structural differences in microbiota composition across these conditions. Furthermore, we trained machine learning models on the compositional data to differentiate samples from each group, aiming to assess the potential of gut microbiota as an objective diagnostic biomarker for psychiatric disorders.

2. Materials and Methods

2.1. Dataset Collection

We conducted keyword searches in PubMed and Scopus to identify 16S sequencing studies, employing the following query:

((”major depressive disorder” OR MDD OR depression) OR (bipolar OR B1D OR B2D OR BD OR BMD) OR (schizophrenia OR SCZ OR schizoaffective OR schizophreniform)) AND (microbiome OR metagenome OR MWAS OR ”metagenome-wide” OR microbiota OR microbiomics OR 16S OR ”shotgun sequencing”)

The screening process was undertaken by two researchers, blinded to each other’s decisions on Rayyan [24]. Cases of decision conflicts were handled by a separate third researcher. We included studies that provided publicly available raw sequencing data and metadata indicating whether each

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

sample was from a case or control group. All data were downloaded from online databases such as SRA, ENA, and CNGB. We excluded studies that required additional committee approvals or access authorizations. When studies sampled multiple body sites, we included only those samples related to the gut microbiome and disregarded all others. Additionally, we excluded studies that examined patients with comorbid psychiatric disorders (e.g., GAD, PTSD) and conditions known to alter gut microbiota composition (e.g., IBD) [25,26].

2.2. 16S Processing

All data processing and analysis steps were performed using the QIIME 2 Python API v2024.5.0 [27]. Multiplexed raw sequences were demultiplexed by finding exact matches to the provided barcodes [28,29]. Primer sequences (and their reverse complements) were trimmed using Cutadapt [30]; for Illumina reads, the universal adapter sequence listed in the FastQC adapter list [31] was also removed, allowing a maximum error rate of 0.1. Paired-end sequences were merged using VSEARCH [32]. The resulting contigs were then filtered by quality score (Q) using the quality filter plugin with default parameters [33]. Denoising was performed on the filtered contigs using Deblur [34], chosen for its superior interstudy consistency and stability [35]. Where possible, all contigs were trimmed to 400 bp. In studies where this cutoff discarded the majority of contigs or allowed bases with $Q < 25$, the threshold was lowered to 380, 250, or 130 bp as appropriate. Taxonomic classification was conducted using a scikit-learn classifier [36–38] pretrained on the SILVA dataset [39,40], and resulting taxonomic feature tables were merged [41]. Additionally, fragment insertion was performed using the SEPP algorithm to generate a phylogenetic tree [42–45], and all identified OTUs were collapsed to the genus level to facilitate comparison across studies with different methodologies.

2.3. Microbiome Community Analysis

Non-collapsed abundance data were used to generate α rarefaction plots, which allowed us to evaluate the sufficiency of sequencing depth and sample diversity. After visually inspecting these plots, we selected a sampling depth of 6,000 sequences per sample to achieve a balance between sample retention and depth of coverage. At this standardized depth, we calculated α and β-diversity indices [46].

Following proposed guidelines [47], we measured observed features, Pielou's evenness [48], and Faith's phylogenetic diversity [49,50] and Shannon diversity metrics [51]. Each of these metrics represents a different aspect of α-diversity, namely richness, dominance, phylogenetic diversity, and information content. This careful selection of metrics enhances the interpretability of the α-diversity results and allows for a thorough assessment of within-sample diversity [47]. The significance of the observed intergroup differences in these calculated metrics was evaluated using the Kruskal-Wallis test [52].

For β-diversity, we computed pairwise dissimilarity metrics including Jaccard [53], Bray-Curtis [54,55], and both weighted and unweighted UniFrac distance metrics [56–60]. The resulting matrices were then subjected to dimensionality reduction through principal coordinate analysis (PCoA) [61,62], and the resulting plots were analyzed visually. We used the PERMANOVA [63], ANOSIM [64], and PERMDISP [65,66] methods to assess the intergroup significance of differences in the calculated measures.

Differential abundance analysis of microbial taxa was performed using the QIIME 2 implementation of ANCOM-BC [67], with statistical significance set at a threshold of 0.001. To control for multiple hypothesis testing and reduce the false discovery rate, we applied the Benjamini–Hochberg method [68].

2.3.1. Sequencing Context Sensitivity Analysis

To assess the effect of sequencing context on community analysis results, we repeated the previously described analyses in a subgroup of samples with a restricted sequencing context, including

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

only samples from studies that profiled the V3–V4 hypervariable regions using the Illumina MiSeq sequencer.

α-diversity sensitivity was evaluated by comparing the ordering of group medians for each diversity measure and by comparing the statistical significance of group differences.

For β-diversity, we calculated absolute and relative differences in PERMANOVA R^2, ANOSIM R, and PERMDISP F, and compared the statistical significance of differences across all β-diversity measures.

To assess the stability of differential abundance findings, we calculated Spearman rank correlations for taxon log-fold abundance changes between the primary and subgroup analyses. We also assessed sign concordance of log-fold changes separately for positive and negative associations using Jaccard similarity. Finally, we tested the stability of statistically significant results by calculating Jaccard similarity between the sets of statistically significant changes in the two analyses and by assessing the retention of taxa with significant difference in abundance in the subgroup analysis.

2.3.2. Study Effect Sensitivity Analysis

To adjust for study effects, samples were filtered so that each disorder was compared only with its corresponding healthy controls (BD, MDD, and SCZ vs. HC) and analyzed separately; because diagnosis was nearly completely confounded with study, no cross-disorder comparisons were made. All three analyses used fixed-effects models with study and diagnosis as predictors.

α-diversity metrics were modeled with distributions appropriate to their scale: negative binomial GLMs for observed features, β-regression with a precision submodel for Pielou's evenness, and linear models with HC3 heteroscedasticity-robust standard errors for log-transformed Faith's PD and Shannon entropy; influential outliers were excluded (Supplementary Table 1).

β-diversity was assessed in vegan by PERMANOVA with permutations stratified by study, and homogeneity of dispersion was tested with PERMDISP.

Differential abundance testing was performed using ANCOM-BC with the same significance threshold ($p < 0.001$) as the primary analysis. All p-values were Benjamini–Hochberg corrected across the three disorder comparisons.

2.4. Network Analysis

Association matrices (R) for each group were calculated using FastSpar [69]. The resulting matrices were sparsified as follows:

$$r_{ij}^{*} = \begin{cases} 0 & \text{if } |r_{ij}| \leq r_{\theta} = 0.3 \\ r_{ij} & \text{if } |r_{ij}| > r_{\theta} = 0.3 \end{cases} \tag{1}$$

The sparsified association matrices (R^{*}) were then transformed into dissimilarity (D) and subsequently similarity matrices (S), following the method outlined by Van Dongen and Enright [70], using the same formulation as NetCoMi package [71]:

$$d_{ij} = \begin{cases} \sqrt{0.5(1 - r_{ij}^{*})} & \text{if } r_{ij}^{*} \neq 0 \\ \infty & \text{if } r_{ij}^{*} = 0 \end{cases} \tag{2}$$

Then,

$$s_{ij} = 1 - d_{ij} \tag{3}$$

This process ensured that uncorrelated taxa remained disconnected in the final network (N).

Using NetworkX v3.4.2 [72], we generated microbial association networks from the similarity matrices. For each network, we calculated degree, eigenvector, closeness, and betweenness centrality measures for every node. Hub taxa were identified using the PageRank algorithm [73], which previous studies have shown to be effective for hub detection in microbial networks [74,75]. The Louvain

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

algorithm was employed for community detection [76]. We identified the largest connected component (LCC) and calculated the $\frac{|LCC|}{|N|}$ ratio. Additionally, we computed the average clustering, modularity, degree assortativity, density, positive edge ratio, diameter, radius, node connectivity, edge connectivity, and natural connectivity of the networks. Natural connectivity ($\overline{\lambda}$) serves as a measure of network robustness, defined as the natural logarithm of the arithmetic mean of the exponentials of the network's eigenvalues (λ) [77]:

$$\overline{\lambda} = \ln\left(\frac{1}{|LCC|}\sum_{i=1}^{|LCC|} e^{\lambda_i}\right) \tag{4}$$

To assess clustering and hub similarity between network pairs, we utilized scikit-learn v1.6.1 [36] to calculate the adjusted Rand index [78] and Jaccard score [53], respectively. We also applied the ORCA algorithm [79] to compute graphlet correlation distance (GCD) to compare the topological similarity between networks [80]. To assess the statistical significance of the uncovered differences in network features, we computed p-values through permutation testing, conducting 1000 permutations for each network pair. We used SciPy v1.13.0 [81] to apply the Benjamini–Hochberg procedure for controlling the false discovery rate.

2.4.1. Sparsification Threshold Sensitivity Analysis

Networks were reconstructed using alternative sparsification thresholds $r_\theta \in \{0.2, 0.25, 0.35, 0.4\}$, with $r_\theta = 0.3$ retained as the primary threshold. Sensitivity to threshold selection was assessed by comparing each alternative network with the primary network. For taxon-level centrality measures, Spearman rank correlations quantified concordance between thresholds; the same approach was applied to between-group centrality differences. Stability of identified hubs and statistically significant between-group differences in taxon centrality was quantified using Jaccard similarity and retention rates. For continuous global network characteristics, relative changes between thresholds were calculated. Finally, directional concordance of between-group centrality differences was assessed separately for positive and negative associations using Jaccard similarity.

2.4.2. Sequencing Context Sensitivity Analysis

To assess the effect of sequencing context on network analysis results, we repeated the previously described network analysis on data from studies with restricted sequencing context, as defined in the corresponding microbiome community sensitivity analysis methods subsection, using the primary sparsification threshold $r_\theta = 0.3$. Results were compared with those of the primary analysis, and their stability was assessed using the same methods employed in the sparsification threshold sensitivity analysis.

2.5. Machine Learning Modeling

Genus-level absolute counts were exported. Taxa found in less than 1% of samples, those with a maximum frequency of less than 1% across samples, and those with an unidentified genus were removed. For the remaining data, each taxon's abundance value was transformed to a Z-score by subtracting the mean and dividing by the standard deviation of its abundance across all samples, then clipped to the range $[-5, 5]$. Taxa were classified into four clades based on their phyla: Firmicutes, Proteobacteria, Bacteroidota, and others. For taxa in each clade, a Spearman's rank correlation matrix (P) was calculated, and for each taxon (i), a cumulative correlation coefficient (ρ^*) was calculated according to the taxoNN$_{\text{corr}}$ algorithm [82]:

$$\rho_i^* = \sqrt[k]{\prod_{j=1}^{k} |\rho_{ij}|} \tag{5}$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

The resulting set (P^*) was ordered in a descending manner:

$$P^* = \{\rho_1^*, \rho_2^*, \ldots, \rho_k^* \mid \rho_l^* \geq \rho_{l+1}^* \text{ for } l = 1, 2, \ldots, k-1\} \quad (6)$$

Taxa within each clade were ordered based on their positions in P^*, ensuring consistent input structure across samples. To address the substantial class imbalance in the dataset—particularly for underrepresented groups such as BD—the SVM-SMOTE algorithm was applied to oversample minority classes in the training data, producing a more balanced training set and reducing potential biases during model training [83].

We developed a deep neural network ensemble model [84] using TensorFlow v2.21.0 [85], with the sorted abundance vectors from each clade serving as input features. The model architecture comprised two sequential 1-D convolutional layers [86] per clade, each followed by max pooling, batch normalization, and dropout to mitigate overfitting and stabilize training [87,88]. Batch normalization was applied after each convolutional layer to standardize the feature distributions to zero mean and unit variance within each mini-batch [87]. The first convolutional layer utilized 32 filters with a kernel size of 2 and a stride of 1, while the second convolutional layer employed 64 filters, also with a kernel size of 2 and a stride of 1. Rectified Linear Unit (ReLU) was selected as the activation function for both convolutional and dense layers to introduce non-linearity [89]. Max pooling layers used a pool size of 2 and a stride of 2, and dropout layers were applied to further reduce overfitting. Feature vectors extracted from each clade were then flattened and concatenated to form a comprehensive feature representation for each sample. This unified feature vector was passed through two fully connected dense layers: the first with 32 nodes and ReLU activation, and the final output layer with 4 nodes and softmax activation to classify samples into one of four diagnostic categories (SCZ, BD, MDD, or HC) [90]. The overall architecture, which builds upon the taxoNN framework [82] with several modifications for improved performance, is depicted in Figure 7. The model was compiled using the Adam optimizer [91] and employed sparse categorical cross-entropy as the loss metric. It underwent training over 100 epochs, incorporating an early stopping strategy to prevent overfitting.

As our second model, we developed a random forest classifier [92] using scikit-learn with 250 estimators, unlimited tree depth, a minimum split size of 6 samples, a minimum leaf size of 20 samples, and bootstrapping enabled [36].

To ensure robust estimation of predictive accuracy, we assessed model performance using both 10-fold cross-validation and leave-one-study-out cross-validation. All preprocessing steps—including filtering, normalization, ordering, and augmentation—were performed independently within each fold to prevent data leakage.

For the leave-one-study-out cross-validation, data from Jiang et al. (2020), Reininghaus et al. (2020), Xiang et al. (2022), and one additional study (selected iteratively) were used as the held-out test set, excluding Hu et al. (2019). Models were trained on the remaining studies. This configuration ensured that all four classes were present in both training and test sets, allowing the model to learn from every class while still enabling class-specific performance evaluation.

For each model and cross-validation scheme, we generated receiver operating characteristic (ROC) curves for both one-vs-one and one-vs-rest classification tasks. We also calculated balanced accuracy, F1-score, and Matthews correlation coefficient (MCC) as evaluation metrics [93].

3. Results

3.1. Dataset Collection

In April 2024, searches were conducted in the PubMed and Scopus databases, resulting in a total of 10,502 records. Duplicates, non-original journal articles, publications in languages other than English, and articles published before 2014 were automatically excluded. This left 4041 records, which underwent abstract screening. During this screening, studies were excluded if they exclusively involved in vitro or animal research, did not perform microbiome sequencing, or did not specifically

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

target patients with bipolar disorder, major depressive disorder, or schizophrenia. This process narrowed the pool down to 128 records for full-text screening.

The full texts of these articles were retrieved successfully, with the exception of 6 articles, which were subsequently excluded from the study. The remaining 122 articles underwent thorough full-text screening. Articles lacking publicly available sequencing data were excluded, as were those that had publicly available datasets but did not include metadata to indicate whether the samples belonged to the case or control groups. Additionally, studies that did not perform 16S sequencing on gut microbiome samples or included subjects with other psychiatric disorders or conditions known to affect gut microbiota were also excluded.

Ultimately, 17 studies were included in this analysis. The PRISMA flow diagram in Figure 1 illustrates the screening process in detail [94]. Characteristics of these studies, including sample sizes, conditions, sequencer model, 16S hypervariable region, data accession number, and references, are shown in Table 1.

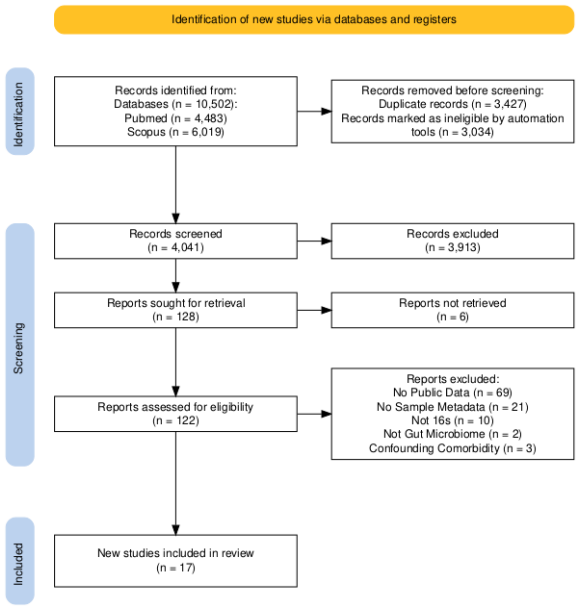

Figure 1. The PRISMA flow diagram of the study identification and screening process.

Table 1. Overview of included studies. **MDD**: Major depressive disorder; **BD**: Bipolar disorder; **SCZ**: Schizophrenia.

Study	Condition	Cases	Controls	Sequencer	16S Region	Raw Data	Reference
Reininghaus et al. (2020)	MDD	61	0	Illumina MiSeq	V3	ENA:PRJEB40986	[95]
Chen et al. (2022)	MDD	10	10	Illumina MiSeq	V3–V4	SRA:PRJNA890246	[96]
Liu et al. (2020)	MDD	43	47	Illumina MiSeq	V4	SRA:PRJNA591924	[97]
Jiang et al. (2020)	BD/MDD	10/14	16	Illumina MiSeq	V3–V4	SRA:PRJNA565097	[98]
Dong et al. (2022)	MDD	63	30	Illumina MiSeq	V3–V4	SRA:PRJNA778934	[99]
Liu et al. (2022)	MDD	66	43	Illumina HiSeq 1500	V3–V4	SRA:PRJNA776170	[100]
Dong et al. (2021)	MDD	23	10	Illumina MiSeq	V3–V4	SRA:PRJNA647236	[101]
Zhang et al. (2022)	MDD	53	37	Illumina HiSeq 1500	V3–V4	SRA:PRJNA830325	[102]
Ye et al. (2021)	MDD	26	28	Illumina HiSeq 2500	V3–V4	SRA:PRJNA685916	[103]
Hu et al. (2019)	BD	52	45	Illumina MiSeq	V3–V4	ENA:PRJEB23500	[104]
O'Donnell et al. (2022)	SCZ	22	22	Illumina MiSeq	V3–V4	SRA:PRJNA896324	[105]
Nguyen et al. (2021)	SCZ	48	48	Illumina MiSeq/HiSeq 4000	V4	ENA:ERP107975	[106]
Ling et al. (2022)	SCZ	90	71	Illumina MiSeq	V3–V4	SRA:PRJNA807473	[107]
Xiang et al. (2022)	SCZ	25	0	Illumina MiSeq	V3–V4	SRA:PRJNA821673	[108]
Shen et al. (2018)	SCZ	64	53	Illumina HiSeq 2500	V3–V4	GSA:CRA000653	[109]
Pan et al. (2020)	SCZ	29	29	Illumina HiSeq 2000	V3–V4	SRA:PRJNA559773	[110]
Tsai et al. (2022)	SCZ	40	0	Illumina MiSeq	V3–V4	SRA:PRJNA790710	[111]

The synthesized demographic information from the included studies, including mean age, BMI, and the ratio of female subjects for each condition, is presented in Table 2. Studies that did not report demographic measures were excluded when calculating the averages for those measures. The results indicate highly heterogeneous demographics and significant differences in these measures between groups.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

All differences in mean age and BMI between groups were statistically significant ($p < 0.05$), except for the difference between the BD and MDD groups. In terms of the ratio of female subjects, the only statistically significant differences were found between the MDD and SCZ groups, as well as between the MDD and HC groups.

Ultimately, the available sequencing data from each study were downloaded and pooled, resulting in a total of 1402 samples, which included 81 BD, 412 MDD, 397 SCZ, and 512 HC samples. These samples underwent subsequent analysis. The difference between these statistics and those in Table 2 reflects the original discrepancy in the statistics reported in the papers and the associated datasets.

Table 2. Aggregated baseline demographic data of included studies. Studies with missing demographics were excluded from calculations.

Condition	Number	Female (%)	Mean Age ± SD	Mean BMI ± SD
BD	62	47.70	26.25 ± 10.32	21.90 ± 4.33
MDD	359	70.30	29.15 ± 11.13	22.31 ± 4.91
SCZ	318	42.10	45.25 ± 12.41	26.04 ± 6.24
HC	489	52.87	33.42 ± 13.70	23.38 ± 4.31

MDD: Major depressive disorder; **BD**: Bipolar disorder; **SCZ**: Schizophrenia; **HC**: Healthy control.

3.2. Microbiome Community Analysis

3.2.1. α-Diversity Analysis

The rarefied genus-level abundance data were subjected to further analysis for diversity estimation. α-diversity metrics, including the number of observed features, Pielou's evenness, Shannon's entropy, and Faith's phylogenetic diversity were calculated to provide a comprehensive view of within-sample diversity (Figure 2; Supplementary Table 2).

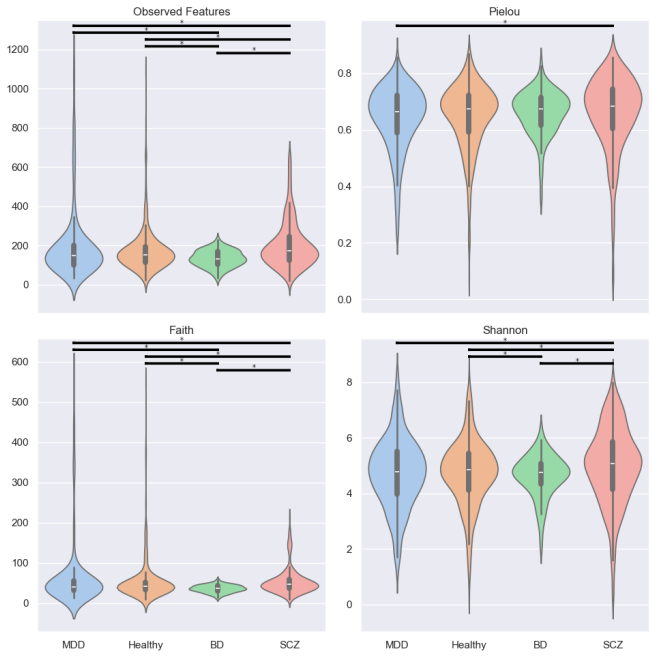

Figure 2. Comparison of α-diversity metrics. *: $q < 0.05$; **MDD**: Major depressive disorder; **BD**: Bipolar disorder; **SCZ**: Schizophrenia.

The Kruskal-Wallis test revealed a statistically significant difference ($p = 3.063 \times 10^{-8}$) in the number of observed features across all groups, demonstrating an increase in diversity from BD to HC, MDD, and SCZ, in that order. Post-hoc pairwise analysis indicated significant differences (Benjamini–Hochberg corrected $q < 0.05$) between every group pair except for the comparison between the MDD and HC groups.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

The multiclass Kruskal-Wallis test did not find a statistically significant difference in Pielou's evenness among the groups. Post-hoc analysis revealed no significant differences between group pairs, aside from the SCZ group, which exhibited a significantly increased diversity compared to the MDD group ($q = 0.0224$).

Evaluating Shannon's entropy showed a statistically significant difference ($p = 0.0022$) in diversity among groups, reflecting a similar pattern to that observed in the analysis of the number of observed features. Pairwise testing revealed significant differences between every group pair, except for the MDD-HC and BD-MDD comparisons.

Multiclass testing of Faith's phylogenetic diversity also replicated the results observed with the number of observed features, yielding a statistically significant result in both multiclass ($p = 2.258 \times 10^{-8}$) and pairwise analyses, except for the comparison between the MDD and HC groups.

3.2.2. β-Diversity Analysis

To evaluate β-diversity, we calculated Bray-Curtis, Jaccard, weighted UniFrac, and unweighted UniFrac distance matrices and applied PERMANOVA, PERMDISP, and ANOSIM to each. Omnibus PERMANOVA was significant for every distance matrix ($p < 0.001$). Weighted UniFrac produced the largest omnibus pseudo-F (15.463), followed by Bray-Curtis (11.549), unweighted UniFrac (10.317), and Jaccard (9.284); PERMDISP F was lowest for weighted UniFrac (14.093) and highest for Jaccard (126.962). Omnibus ANOSIM R values were uniformly low (0.019-0.040), indicating weak rank-based separation.

Post-hoc pairwise PERMANOVA detected significant compositional differences for all group pairs in every matrix ($q < 0.05$). However, pairwise PERMDISP was significant for most pairs, so most pairwise PERMANOVA results may reflect differences in centroid location, dispersion, or both. Only pairs with non-significant PERMDISP—weighted UniFrac BD-MDD; unweighted UniFrac HC-MDD and MDD-SCZ; Bray-Curtis and Jaccard HC-SCZ—are more compatible with a location-based interpretation. Pairwise ANOSIM detected significant but weak separation only for a subset of pairs (all $R \leq 0.151$), mainly involving HC-MDD, HC-SCZ, and MDD-SCZ, with weighted UniFrac additionally showing BD-MDD separation. The observed discrepancies between PERMANOVA and ANOSIM results may be attributed to ANOSIM's lower statistical power in the presence of strong data gradients [112]. Exact pairwise PERMANOVA, ANOSIM, and PERMDISP statistics are summarized in Table 3 (see Supplementary Tables 4–6 for the full results).

Table 3. Summary of β-diversity analyses across distance matrices. Omnibus statistics are reported as PERMANOVA R^2; ANOSIM R; PERMDISP F.

Matrix	Omnibus	PERMANOVA significant pairs (R^2)	ANOSIM significant pairs (R)	PERMDISP non-significant pairs
Weighted UniFrac	$R^2 = 0.034$ $R = 0.040$ $F = 14.093$	All pairs $q < 0.05$ Max: BD–MDD (0.059) Min: HC–MDD (0.011)	BD–MDD (0.151) HC–MDD (0.013) HC–SCZ (0.059) MDD–SCZ (0.053)	BD–MDD
Unweighted UniFrac	$R^2 = 0.023$ $R = 0.033$ $F = 16.937$	All pairs $q < 0.05$ Max: BD–SCZ (0.029) Min: HC–MDD (0.008)	HC–MDD (0.033) HC–SCZ (0.073) MDD–SCZ (0.068)	HC–MDD MDD–SCZ
Bray–Curtis	$R^2 = 0.025$ $R = 0.030$ $F = 103.309$	All pairs $q < 0.05$ Max: BD–MDD (0.030) Min: HC–SCZ (0.008)	HC–MDD (0.028) HC–SCZ (0.034) MDD–SCZ (0.102)	HC–SCZ
Jaccard	$R^2 = 0.021$ $R = 0.019$ $F = 126.962$	All pairs $q < 0.05$ Max: BD–SCZ (0.027) Min: HC–SCZ (0.007) HC–MDD (0.007)	HC–MDD (0.022) HC–SCZ (0.034) MDD–SCZ (0.088)	HC–SCZ

A visual inspection of the PCoA results revealed a trend consistent with previously observed patterns (Figure 3). The PCoA based on the weighted UniFrac distance matrix accounted for the

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

highest proportion of observed variance (PCo1: 18.36%, PCo2: 17.07%, PCo3: 10.17%). This was followed by the unweighted UniFrac matrix (PCo1: 9.16%, PCo2: 7.83%, PCo3: 4.13%), the Jaccard distance matrix (PCo1: 10.04%, PCo2: 4.79%, PCo3: 3.58%), and finally, the Bray-Curtis distance matrix (PCo1: 7.53%, PCo2: 4.46%, PCo3: 3.38%).

All PCoA plots exhibited significant overlap between the various groups. Notably, the BD group was clearly distinguishable from all other groups across all distance measures, despite a clear difference in dispersion. Conversely, the HC group was less distinct when compared to the SCZ and MDD groups, as all three occupied similar areas in the variable space, except for specific regions exclusively filled with HC samples. This pattern was consistent across all distance measures. The MDD and SCZ groups presented a more intricate pattern; in the UniFrac PCoA plots, samples from these two groups occupied similar areas yet displayed contrasting densities. Furthermore, the Jaccard and Bray-Curtis PCoA plots demonstrated a greater degree of separability, occupying more mutually exclusive regions within the variable space.

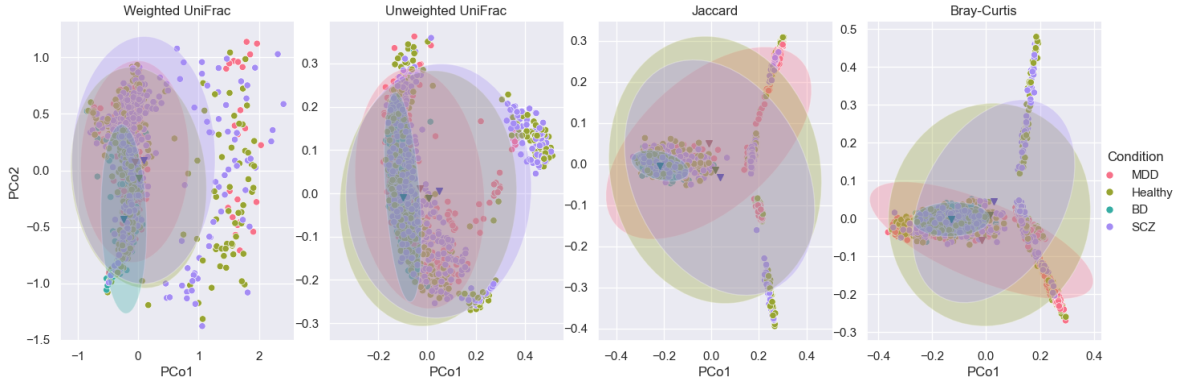

Figure 3. Principal coordinate analysis plots of weighted UniFrac, unweighted UniFrac, Bray-Curtis and Jaccard distance measures, with 95% confidence ellipsoids. For the sake of brevity, only the first two principal coordinates accounting for the highest proportion of observed variation are plotted. **MDD**: Major depressive disorder; **BD**: Bipolar disorder; **SCZ**: Schizophrenia.

3.2.3. Differential Abundance Analysis

In this study, we conducted differential abundance testing utilizing the ANCOM-BC method to identify taxa with statistically significant differences in abundance ($q < 0.001$) across the various groups assessed. Results are presented in Figure 4 and Supplementary Table 7.

The comparison between the HC group and the BD group revealed notable shifts in microbiome composition. The BD group exhibited an enrichment of genera associated with the order Bacteroidales, including *Bacteroides, Parabacteroides, Butyricimonas*, and *Odoribacter*. Conversely, the BD group displayed a depletion of genera within the *Lachnospiraceae* family, specifically *Blautia, Agathobacter*, and the *Lachnospiraceae FCS020 group*.

The microbiome profile of the MDD group presented a more intricate pattern regarding the abundance of Firmicutes. Notably, there was an enrichment of genera such as *Eubacterium hallii group, Blautia, Streptococcus*, and *Enterococcus*, while *Lachnospira, Eubacterium ruminantium group, Roseburia*, and *Lachnoclostridium* were found to be depleted. Additionally, the MDD group was characterized by an enrichment of *Bifidobacterium, Eggerthella, Acinetobacter*, and *Saccharimonadales*, alongside a reduction of *Sutterella* and *Bacteroides*.

In contrast, the SCZ group exhibited a widespread depletion of Firmicutes, including key genera such as *Blautia, Roseburia, Faecalibacterium, Lachnoclostridium, Coprococcus, Ruminococcus*, and the *Lachnospira* genus. There was also a notable reduction in Bacteroidales taxa (e.g., *Bacteroides, Alistipes, Parabacteroides*), while this group showed an enrichment of Actinobacteriota, including genera such as *Senegalimassilia, Actinomyces, Slackia*, and *Corynebacterium*.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

When comparing the BD group to the other disorder groups, we observed significant enrichment of Bacteroidales and a depletion of Actinobacteriota. Furthermore, relative to the SCZ group, the BD group exhibited widespread enrichment of Firmicutes. Notably, when comparing the BD and MDD groups, the BD group showed an enrichment of *Lachnospira*, *Lachnoclostridium*, *Lachnospiraceae NK4A136 group*, *Sutterella*, and *Haemophilus*, while demonstrating a depletion of *Blautia*, *Eubacterium hallii group*, *Enterococcus*, *Dorea*, and *Streptococcus*.

The comparison between the SCZ and MDD groups revealed a more complex pattern of microbial composition. The SCZ group demonstrated profound depletion of Firmicutes, with the exception of a few genera such as the *Eubacterium ruminantium group* and *Megasphaera*. Among the Bacteroidota phylum, the SCZ group exhibited enrichment of *Prevotella* and *Alloprevotella*, alongside the depletion of *Coprobacter* and *Alistipes*. With respect to Proteobacteria, the SCZ group was enriched in *Enterobacter* while exhibiting a depletion of *Acinetobacter* and *Parasutterella*. Among the Actinobacteria, the SCZ group showed an increase in *Senegalimassilia*, *Actinomyces*, and *Slackia*, along with a decrease in *Eggerthella*, *Adlercreutzia*, and *Bifidobacterium*. Additionally, the SCZ group exhibited enrichment of *Fusobacterium*, further contributing to the complexity of the microbiome profiles associated with these conditions.

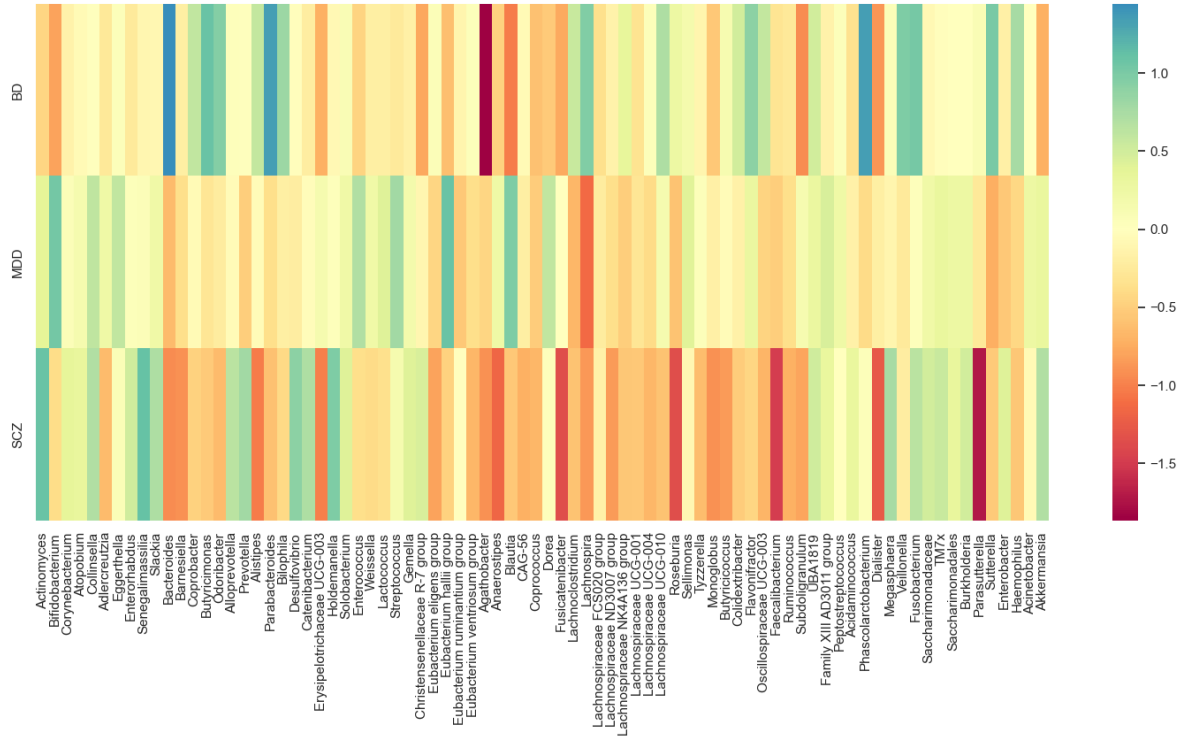

Figure 4. Differential abundance of taxa relative to the HC group, expressed as log-fold changes. Only taxa that show a statistically significant ($q < 0.001$) difference in abundance between any two groups are depicted. **MDD**: Major depressive disorder; **BD**: Bipolar disorder; **SCZ**: Schizophrenia; **HC**: Healthy Control.

3.2.4. Sequencing Context Sensitivity Analysis

Recalculating α-diversity metrics in the restricted-sequencing-context subgroup reproduced the existing pattern of higher SCZ diversity. The statistical significance of differences between the SCZ and other groups was insensitive to sequencing context, except for Pielou's evenness; differences among the remaining groups were not reproduced (Supplementary Table 2).

PERMANOVA R^2 results were generally stable, with absolute differences < 0.034 across Bray-Curtis, Jaccard, and unweighted UniFrac. Weighted UniFrac was more sensitive, with larger differences for BD-SCZ and BD-MDD (0.096 and 0.078, respectively). ANOSIM R was considerably more sensitive, with absolute differences reaching approximately 0.30 and multiple sign changes. Relative differences suggested greater instability, but these should not be overinterpreted because primary effect sizes

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

close to zero inflate relative values. Across group contrasts, PERMANOVA results were fairly stable, while ANOSIM results were more stable for HC-MDD and MDD-SCZ but less stable for BD-related contrasts. The general pattern—prominent SCZ contrasts in Bray-Curtis, Jaccard, and unweighted UniFrac, and prominent BD contrasts in weighted UniFrac—was preserved. PERMDISP results showed absolute changes comparable in scale to PERMANOVA, and the ordering of primary results was relatively preserved, although weighted UniFrac remained more sensitive to sequencing context. Regarding statistical significance, HC-MDD and HC-SCZ dispersion differences were not significant in the subgroup analysis, and ANOSIM revealed a significant BD-MDD difference in intersample diversity that was not observed in the primary analysis (Supplementary Table 4–6).

For differential abundance sensitivity analyses, Spearman rank correlations between log-fold abundance changes in the primary and subgroup analyses were relatively strong, ranging from $\rho = 0.78$ for HC-MDD to 0.96 for BD-SCZ. Sign concordance Jaccard scores ranged from 0.64 to 0.79. Overall, SCZ and BD contrasts were less sensitive to sequencing context than HC and MDD contrasts. Statistical significance was considerably more sensitive, with Jaccard scores ranging from 0.31 for HC-MDD to 0.65 for BD-SCZ. Retention of statistically significant abundance changes showed a similar pattern: HC-MDD was least retained (0.36) and BD-SCZ most retained (0.94), again indicating that SCZ and BD contrasts were less sensitive than HC and MDD contrasts (Supplementary Table 7).

3.2.5. Study Effect Sensitivity Analysis

Among the four α-diversity metrics, only Pielou's evenness and Shannon entropy differed significantly between MDD and HCs, while no metric differed in the BD or SCZ comparisons; all models included study and diagnosis as fixed effects, and Study was a significant source of variation in every comparison. All p-values were corrected for multiple testing using the Benjamini–Hochberg procedure. Observed features showed no significant differences (BD ratio $= 0.91$, $q = 0.120$; MDD ratio $= 0.96$, $q = 0.260$; SCZ ratio $= 1.06$, $q = 0.120$). Pielou's evenness was significantly lower in MDD (mean difference $= -0.031$, $q = 0.0006$), whereas BD (mean difference $+0.001$, $q = 0.962$) and SCZ (mean difference $+0.001$, $q = 0.962$) showed no difference; the precision submodel was not significant in any comparison ($p > 0.2$). Shannon entropy was also significantly lower in MDD (mean difference $= -0.281$, $q = 0.0027$), with no difference in BD (mean difference -0.078, $q = 0.534$) or SCZ (mean difference $+0.093$, $q = 0.456$). Faith's phylogenetic diversity met no threshold: BD showed nominally lower PD (ratio $= 0.91$, $q = 0.098$), SCZ showed a nonsignificant trend toward higher PD (ratio $= 1.05$, $q = 0.098$), and MDD showed no difference (ratio $= 0.99$, $q = 0.596$). Complete results are included in Supplementary Table 3.

β-diversity was assessed using PERMANOVA and PERMDISP, with study and diagnosis as predictors and permutations stratified by study. Diagnosis was significantly associated with community composition in all comparisons and across all four distance metrics (all $q \leq 0.025$), but explained only a small fraction of variance ($R^2 = 0.003$-0.019), whereas study explained substantially more ($R^2 = 0.014$-0.377). PERMDISP revealed significant heterogeneity of dispersion for BD (Jaccard $F = 9.23$, $q = 0.021$; unweighted UniFrac $F = 6.48$, $q = 0.020$) and MDD (unweighted UniFrac $F = 20.97$, $q = 0.003$; weighted UniFrac $F = 8.70$, $q = 0.015$), but not SCZ (all $q \geq 0.359$). Consequently, the BD and MDD PERMANOVA results may partly reflect dispersion differences rather than centroid shifts, whereas the SCZ result, though significant, was driven by a small effect ($R^2 = 0.005$-0.007) with homogeneous dispersion (Supplementary Table 4–6).

Regarding differential abundance testing, correlations of log-fold changes between primary and sensitivity analyses were weak in BD ($\rho = 0.491$), moderate in MDD ($\rho = 0.725$), and strong in SCZ ($\rho = 0.834$). Sign concordance was more uniform across groups, with Jaccard scores from 0.454 to 0.652, highest in SCZ. Taxa with statistically significant differences between disorder and HCs were much less stable, with Jaccard scores and retention ratios ranging from 0 in BD to 0.324 in SCZ.

In BD, no taxa showed significant abundance differences from HCs. MDD showed depletion of several Firmicutes taxa, mostly from Lachnospiraceae, replicating the depletion of *Lachnospira*, *Roseburia*, *Lachnoclostridium*, and non-*hallii Eubacteria*, as well as *Romboutsia*, *Haemophilus*, *Barnesiella*, and

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

Faecalibacterium. The SCZ-HC contrast revealed a more complex pattern: enrichment of Actinobacteriota, including *Senegalimassilia*, *Actinomyces*, *Collinsella*, *Slackia*, and *Corynebacterium*; within Firmicutes, Lachnospiraceae were generally depleted, while *Megasphaera*, *Lactobacillus*, and *Turicibacter* were enriched. Notably, *Akkermansia*, *Desulfovibrio*, *Fusobacterium*, *Pelagibacterium*, and *Saccharimonadaceae TM7x* were also enriched (Supplementary Table 7).

3.3. Network Analysis

Genus-level taxonomic abundance data for each group were transformed into co-occurrence networks, as illustrated in Figure 6. Hub nodes were identified as described in Table 4. The genera *Flavonifractor, Fusobacterium, Oscillospiraceae UCG-003,* and *Erysipelatoclostridium* were exclusive to the BD group. In contrast, the genera *Parabacteroides, Eubacterium hallii group, Streptococcus, Eggerthella, Enterococcus,* and *Lachnoclostridium* were exclusive to the MDD group. Additionally, the genera *Dorea* and *Oscillibacter* were exclusive to the SCZ and HC groups, respectively.

Table 4. Hub taxa identified in each group's co-occurrence network using the PageRank algorithm.

Hub Rank	BD	MDD	SCZ	HC
1	***Oscillospiraceae UCG-002****	***Bacteroides****	***Actinomyces****	***Alistipes****
2	***Flavonifractor****	***Parabacteroides****	***Christensenellaceae R-7 group****	*Ruminococcus gnavus group*
3	***Ruminococcus gnavus group***	*Eubacterium hallii group**	***Alistipes****	***Oscillospiraceae UCG-005***
4	***Eubacterium coprostanoligenes group****	***Alistipes****	***Oscillospiraceae UCG-005****	***Oscillospiraceae UCG-002****
5	*Ruminococcus**	*Actinomyces*	*Bacteroides**	***Christensenellaceae R-7 group***
6	***Oscillospiraceae UCG-005****	*Streptococcus*	*Ruminococcus gnavus group*	*Eubacterium coprostanoligenes group*
7	***Alistipes****	***Oscillospiraceae UCG-005***	*Eubacterium coprostanoligenes group*	*Lachnospiraceae NK4A136 group**
8	*Oscillospiraceae UCG-003**	*Eggerthella*	*Oscillospiraceae NK4A214 group**	*Oscillibacter*
9	*Fusobacterium*	*Enterococcus**	*Oscillospiraceae UCG-002**	*Ruminococcus*
10	*Erysipelatoclostridium*	*Lachnoclostridium*	*Dorea*	***Bacteroides****

Bold font indicates hubs insensitive to changes in r_θ threshold; * indicates hubs insensitive to sequencing context. **MDD**: Major depressive disorder; **BD**: Bipolar disorder; **SCZ**: Schizophrenia; **HC**: Healthy control.

Centrality indices for each node in the network were calculated, allowing for the identification of taxa with significant between-group differences in at least one centrality measure, as detailed in Figure 5, and Supplementary Table 8–9. In the BD group, the taxa *Fusobacterium* and *Eubacterium coprostanoligenes group* exhibited significantly higher centrality compared to all other groups and disorder groups, respectively. Conversely, the genus *Actinomyces* demonstrated reduced centrality when compared to other disorder groups. In the MDD group, the taxa *Parabacteroides*, *Lachnospira*, and the *Eubacterium hallii group* showed a significant increase in centrality compared to all other groups. Additionally, the genus *Enterococcus* exhibited increased centrality within the MDD group compared to other disorder groups. However, the taxon *Oscillospiraceae NK4A214 group* showed lower centrality in the MDD group than in all other groups. In the SCZ group, the genus *Intestinibacter* demonstrated greater centrality than the BD and HC groups, while the genus *Collinsella* showed greater centrality than both the MDD and HC groups. Finally, in the HC group, when compared to the MDD and SCZ groups, the taxon *Oscillospiraceae UCG-003* exhibited greater centrality. In contrast, the genus *Intestinibacter* displayed reduced centrality.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1



Figure 5. Hub taxa and their network centrality across diagnostic groups. Panels show degree, eigenvector, betweenness, and closeness centrality for the 23 taxa that were hubs in at least one group's co-occurrence network. Each measure was min–max scaled to $[0-1]$ before plotting. **BD**: bipolar disorder; **MDD**: major depressive disorder; **SCZ**: schizophrenia; **HC**: healthy control.

Ultimately, network-wide topological measures were computed for each group, and statistically significant differences in these measures were assessed among the groups, as described in Table 5. No statistically significant differences were observed among the groups, except for the MDD group, which exhibited reduced modularity compared to the SCZ group, a lower positive edge ratio, and increased graph and largest connected component (LCC) density compared to both the SCZ and HC groups. More sophisticated inter-network distance measures—such as the adjusted Rand index, Jaccard distance, and graphlet correlation distance—did not reveal significant differences among the groups, failing to detect statistically significant variations in community structure, hub class composition, and topological architecture, respectively.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

Table 5. Network topological properties for each diagnostic group.

Index	BD	MDD	SCZ	HC
LCC Ratio	0.979	0.931	0.898	0.855
Clustering	0.206	0.254	0.267	0.236
Modularity	0.383	0.463$^{\diamond}$	0.594$^{\diamond}$	0.546
Assortativity	0.154	0.035	0.207	0.092
Positive Edge Ratio	0.631	0.679$^{*\diamond}$	0.783$^{\diamond}$	0.752*
Graph Density	0.097	0.087$^{*\diamond}$	0.058$^{\diamond}$	0.062*
LCC Density	0.103	0.099$^{*\diamond}$	0.069$^{\diamond}$	0.082*
Diameter	4.478	3.361	4.756	3.946
Radius	2.509	1.877	2.525	2.455
Natural Connectivity	2.093	1.242	1.145	0.849
Node Connectivity	1.000	1.000	1.000	1.000
Edge Connectivity	1.000	1.000	1.000	1.000

$^{\diamond}$, *: $q < 0.05$; **LCC**: Largest connected component; **MDD**: Major depressive disorder; **BD**: Bipolar disorder; **SCZ**: Schizophrenia; **HC**: Healthy control.

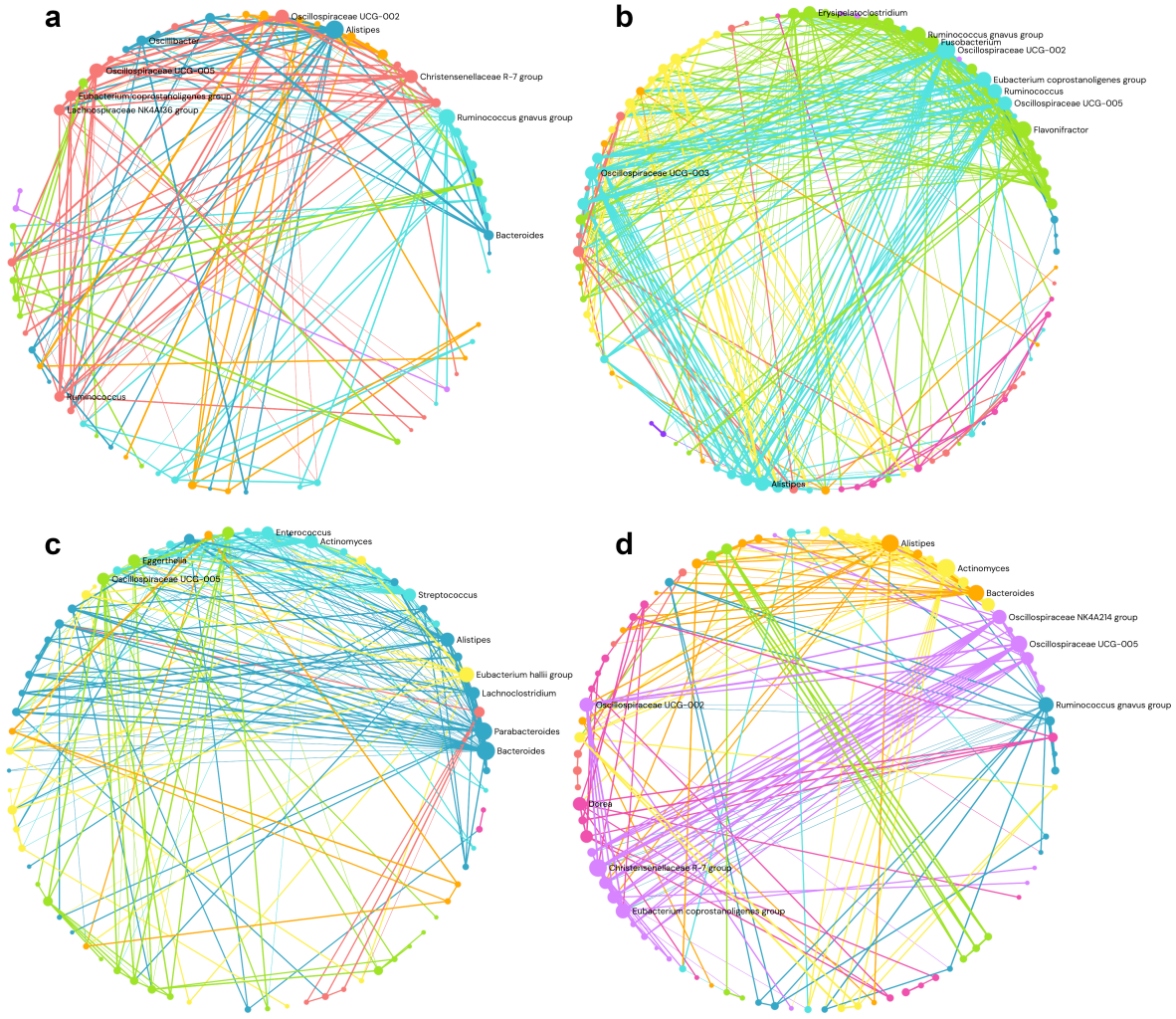

Figure 6. Largest connected components of co-occurrence networks of microbial genera for **a)** healthy control group, **b)** Bipolar disorder group, **c)** Major depressive disorder group, and **d)** Schizophrenia group. Each color designates a distinct community, and node size corresponds to PageRank scores. Hub taxa are noted within the networks.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

3.3.1. Sparsification Threshold Sensitivity Analysis

Across alternative r_θ thresholds, hub selection showed moderate to good stability, with median retention of primary hubs ranging from 0.60 to 0.90 and median Jaccard similarity from 0.428 to 0.742. Hubs in HC and BD groups were more stable than those in MDD and SCZ groups; taxa identified as hubs in every run are shown in bold in Table 4. Centrality measures showed high concordance for closeness, degree, and eigenvector ($\rho > 0.80$), but weaker concordance for betweenness. The same pattern held for between-group centrality contrasts (degree and eigenvector $\rho > 0.75$; betweenness weakest). Directional stability was greatest for degree and eigenvector (median Jaccard > 0.50 for both positive and negative changes) and least for betweenness. Among statistically significant centrality differences, degree centrality had the highest median retention (0.47). Further details are included in Supplementary Table 10–11.

For topological network measures, relative changes from the primary analysis were calculated across thresholds. Using median relative-change ranges (values ≤ 0.5 were considered stable, and values ≥ 1 were considered unstable), LCC ratio, clustering, positive edge ratio, graph diameter and radius, and node and edge connectivity were relatively stable, whereas assortativity, LCC density, and natural connectivity were unstable. The primary ordering of groups was best preserved for modularity (3 of 4 thresholds), followed by assortativity and natural connectivity (2 of 4 each), and least for positive edge ratio and graph radius (0 of 4). For intergroup contrasts, Jaccard and adjusted Rand indices were relatively stable (median relative-change < 1); large ranges for other measures should be interpreted cautiously because relative change is inflated when the primary contrast is near zero.

Significant differences in modularity between MDD and SCZ were reproduced at 3 of 4 thresholds. Significant differences in positive edge ratio for MDD-SCZ and MDD-HC were reproduced at 1 of 4 thresholds, as was the MDD-SCZ difference in graph density. The MDD-HC difference in graph density and all significant LCC density differences were not reproduced. Statistically nonsignificant results were generally stable across thresholds (Supplementary Table 12–13).

3.3.2. Sequencing Context Sensitivity Analysis

Hubs that reappeared in the sensitivity analysis are marked with * in Table 4. Within each group, centrality indices showed strong Spearman rank correlations between the primary and sensitivity analyses ($\rho > 0.80$), except for betweenness centrality ($\rho = 0.682$) and the HC group ($\rho = 0.732$). Intergroup centrality contrasts showed much weaker correlations (median ρ: 0.41–0.54 across centrality indices; 0.28–0.74 across disorder-pair contrasts), with the lowest values consistently involving the HC group. The same pattern emerged for positive and negative sign concordance (median Jaccard: 0.38–0.70 and 0.32–0.63, respectively), and median Jaccard scores across centrality indices were similar for both directions (approximately 0.37–0.51). For taxa with statistically significant centrality differences, median Jaccard scores ranged from 0 in contrasts involving the HC group to 0.8 in the BD-MDD contrast, with median retention ratios showing the same pattern (0 and 1, respectively). Complete results are shown in Supplementary Table 14–15.

Compared with primary results, all measures had relative changes below 0.5 except assortativity, graph density, LCC density, and natural connectivity, with HC samples showing the greatest variability. Among contrasts, median unsigned relative differences exceeded 1 for LCC ratio, clustering, graph density, LCC density, diameter, radius, and natural connectivity, but remained below 0.5 for assortativity, node and edge connectivity, Jaccard similarity, adjusted Rand index, and graphlet correlation distance. No significant between-group differences were found except for LCC ratio (SCZ>BD) and LCC density (HC>SCZ). Despite the non-reproducibility of significance, some biological patterns were preserved: SCZ showed the highest modularity and lowest density, and BD and SCZ networks had larger diameters (Supplementary Table 16–17).

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

3.4. Machine Learning Classification

A random forest model and a convolutional neural network ensemble model were designed and constructed. The ensemble model was built using a phylogenetically informed architecture, as illustrated in Figure 7. Both models were trained on augmented genus-level taxonomic compositional data to distinguish among samples from individuals with BD, MDD, SCZ, and HCs.

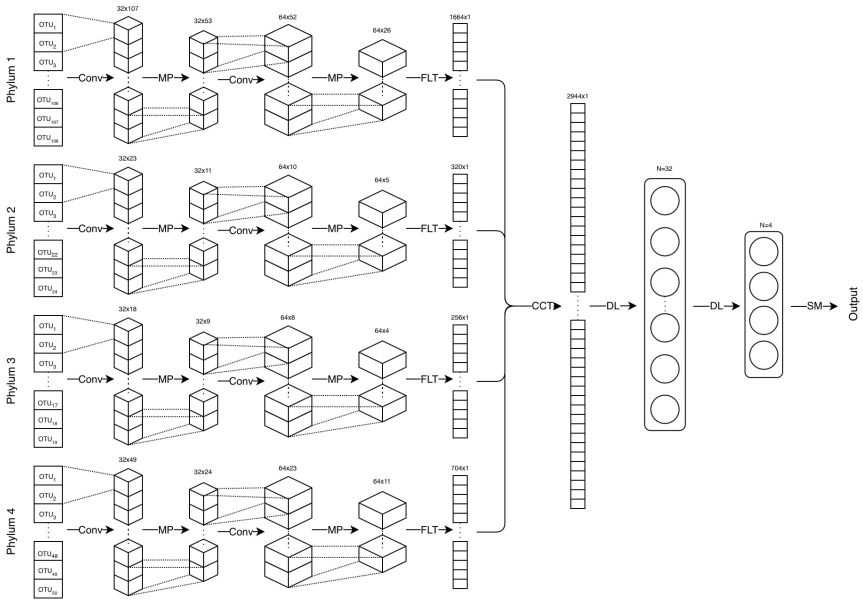

Figure 7. The architecture of the deep convolutional neural network ensemble model. Batch normalization and dropout layers were omitted for simplicity. **Conv**: convolution; **MP**: max-pool; **FLT**: flatten; **CCT**: concatenate; **DL**: dense layer; **SM**: soft-max.

Model performance was evaluated using 10-fold cross-validation and leave-one-study-out cross-validation, with held-out datasets from separate studies serving as test sets. Benchmark results are reported in Table 6. In addition, the multiclass classification problem was decomposed into several binary classification tasks, and receiver operating characteristic (ROC) curves were plotted for each binary task, as shown in Figure 8.

Table 6. Machine learning model benchmarking results. Values are reported as mean ± SD across cross-validation folds and studies.

Model	**Acc_b** (mean ± SD)	**F1** (mean ± SD)	**MCC** (mean ± SD)	**AUC_{OvO}** (mean ± SD)	**AUC_{OvR}** (mean ± SD)
10-fold CV					
Random Forest	0.634 ± 0.06	0.557 ± 0.08	0.443 ± 0.08	0.856 ± 0.03	0.844 ± 0.03
CNN Ensemble	0.609 ± 0.07	0.598 ± 0.06	0.472 ± 0.06	0.862 ± 0.03	0.859 ± 0.03
LOSO CV					
Random Forest	0.521 ± 0.06	0.313 ± 0.08	0.263 ± 0.08	0.762 ± 0.07	0.739 ± 0.07
CNN Ensemble	0.386 ± 0.08	0.298 ± 0.04	0.120 ± 0.05	0.656 ± 0.06	0.643 ± 0.04

Acc_b: Balanced accuracy; **F1**: Macro-F1 score; **MCC**: Matthews correlation coefficient; **AUC_{OvO}**: Macro-average area under the one-versus-one ROC curve; **AUC_{OvR}**: Macro-average area under the one-versus-rest ROC curve; **CV**: Cross-validation; **LOSO**: Leave-one-study-out; **CNN**: Convolutional neural network; **SD**: Standard deviation.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

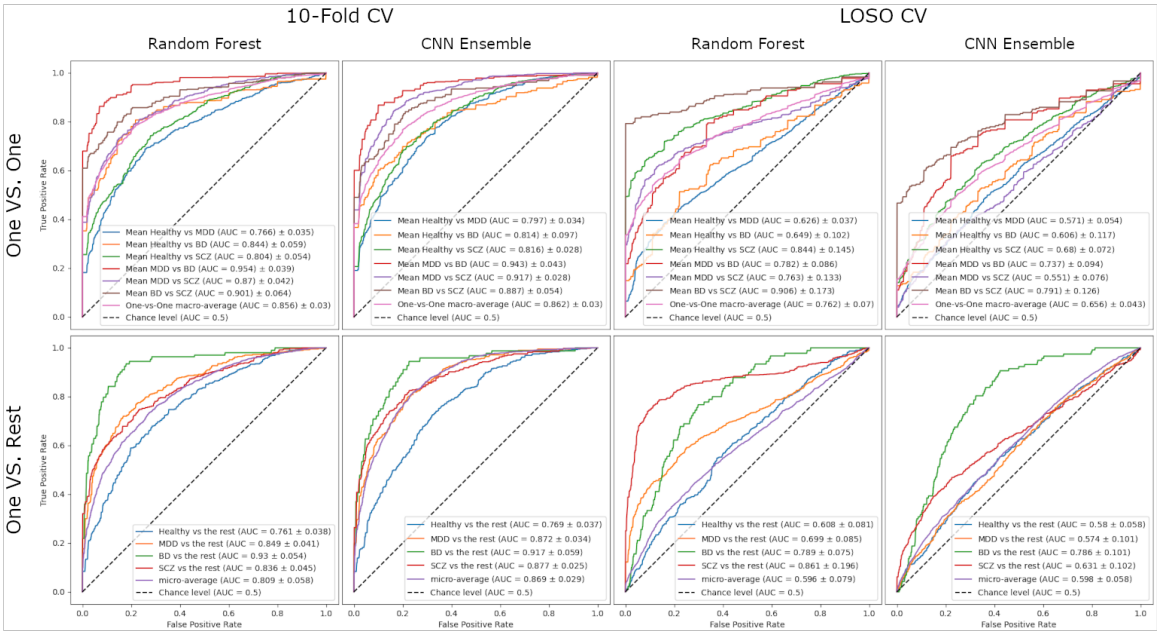

Figure 8. Receiver operating characteristic (ROC) curves for the convolutional neural network ensemble and random forest models, assessed through 10-fold cross-validation and leave-one-study-out cross-validation schemes. **CNN:** Convolutional Neural Network; **CV:** Cross-Validation; **LOSO:** Leave-One-Study-Out; **BD**: Bipolar Disorder; **MDD**: Major Depressive Disorder; **SCZ**: Schizophrenia.

4. Discussion

This study pooled 1402 samples from 17 studies to investigate gut microbiome differences across psychiatric diagnoses and their diagnostic potential. Four findings emerged. First, SCZ was characterized by reduced *Lachnospiraceae* and other Firmicutes and increased Actinobacteriota, while MDD showed depletion of *Lachnospira*, *Roseburia*, and *Lachnoclostridium*. Second, MDD exhibited reduced α-diversity with small effect sizes (Figure 2), while β-diversity differed across all group pairs. Third, network analyses revealed reduced centrality of butyrate producers in MDD and SCZ and increased centrality of taxa associated with neurotransmitter metabolism or pro-inflammatory potential. Finally, machine-learning classifiers achieved modest discrimination under LOSO-CV, indicating a detectable but insufficient diagnostic signal for clinical application (Table 6).

Regarding diversity, the primary analyses detected many statistically significant differences; however, after adjusting for study-level effects, only reductions in Shannon entropy and Pielou's evenness in the MDD group relative to HCs proved to be significant. This pattern suggests increased dominance by certain taxa in the MDD group rather than reduced overall richness. However, effect sizes were small, study origin explained substantially more variance than diagnosis, and considerable residual variance remained. Accordingly, we conclude that α-diversity has little or uncertain diagnostic value for distinguishing the aforementioned disorders from one another and HCs [113,114].

In contrast, β-diversity showed significant pairwise differences across all four distance measures. While most of these differences were accompanied by dispersion effects, comparisons between disorder groups and HCs were a relative exception, suggesting shifts in centroid locations. Despite differential clustering, substantial overlap remained. BD appeared most distinct, but its smaller sample size and dispersion differences warrant caution. Study-stratified analyses confirmed modest associations between disorder status and β-diversity, while study origin explained considerably more variance. This pattern is consistent with psychiatric studies reporting case-control β-diversity differences but limited evidence for separation between diagnoses [114,115]. Ioannou et al. (2024) also reported a case-control difference, but did not find a significant difference in β-diversity across BD and schizophrenia spectrum disorder (SSD) [116]. The greater contribution of study origin is replicated beyond psychiatric disorders, consistent with multicohort metabolic and neurological microbiome studies [117,118]. Compatible

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

with shared, modest case-control compositional differences, our results do not establish a common transdiagnostic cluster or clear separation of every diagnosis.

The differential abundance analysis revealed both shared and disorder-associated taxonomic alterations across the three psychiatric conditions. SCZ exhibited the most extensive changes, characterized by a depletion of the saccharolytic, butyrate-producing *Lachnospiraceae* and *Faecalibacterium*. This was accompanied by an expansion of Actinobacteriota—a phylum containing immuneomodulatory taxa [119] capable of influencing host neuroactive [120] and epigenetic pathways [121], with context-dependent pro-inflammatory potential [122]—as well as the pro-inflammatory *Fusobacterium* [123]. These patterns were directionally replicated after adjustment for study effects.

MDD showed a more heterogeneous profile. It shared the depletion of *Lachnospira*, *Lachnoclostridium*, and the key butyrate producers *Roseburia*, *Coprococcus*, and *Faecalibacterium*. Alongside the enrichment of pro-inflammatory *Eggerthella* [124], the concurrent enrichment of *Enterococcus* and *Streptococcus* suggested additional alterations in microbial metabolic potential, including amino-acid metabolism and, for some taxa, GABA-related pathways [125,126]. Unlike these enrichments, the *Lachnospiraceae* depletion survived study adjustment.

BD was initially characterized by an enrichment of Bacteroidales, including *Bacteroides* and *Parabacteroides*, and a depletion of *Lachnospiraceae*, but no taxon survived adjustment. This limits confidence, given that the BD samples were derived from only two small studies.

Despite these apparent differences, the findings should be interpreted within the broader evidence for transdiagnostic microbial alterations in psychiatric disorders. Nikolova et al. (2021) found little evidence of disorder specificity, reporting shared depletion of *Faecalibacterium* and *Coprococcus* and enrichment of *Eggerthella* across MDD, BD, psychosis, SCZ, and anxiety [114]. Similarly, McGuinness et al. (2022), in a systematic review of 44 studies, identified increased *Eggerthella* and *Lactobacillus* and decreased *Coprococcus* across all three disorders, with depleted *Faecalibacterium* shared by MDD and BD and depleted *Roseburia* in BD and SCZ. They also reported directionally heterogeneous results within *Lachnospiraceae* [113].

Our findings partially align with this literature. *Lachnospira*, *Roseburia*, and *Lachnoclostridium* were depleted in both MDD and SCZ, while *Faecalibacterium* depletion was observed in the primary SCZ analysis and in the study-adjusted MDD analysis. Furthermore, the Actinobacteriota enrichment and Firmicutes depletion in SCZ mirror the phylum-level patterns reported by McGuinness et al. (2022) [113].

Conversely, *Eggerthella* enrichment was observed in MDD but did not replicate after adjustment and was absent in SCZ. Similarly, while our primary analysis revealed a strong enrichment of *Bacteroides* and *Parabacteroides* in the BD group—alongside reductions in these genera within MDD and SCZ—these differences did not survive adjustment for study effects.

This finding contradicts Majie et al. (2026), who proposed *Parabacteroides* abundance as a discriminatory biomarker for distinguishing BD from SCZ, based on their observation of greater *Parabacteroides* enrichment in SCZ than in BD [127]. Furthermore, contrasting with our results, a shotgun metagenomic study of 103 BD and SSD patients and 128 matched controls reported enrichment of *Eggerthella* and *Lachnoclostridium* across two cohorts and proposed *Eggerthella* as a transdiagnostic marker [116].

These differences may stem from variations in population characteristics, medication exposure, sequencing methods (16S versus shotgun), taxonomic resolution, and other study-specific factors. The observed partial convergence across diagnostic boundaries, together with the limited reproducibility of individual taxa after controlling for study effects, indicates that psychiatric microbiome alterations may in part reflect shared biological or clinical features. Nevertheless, shared confounders must not be overlooked.

Network analysis revealed distinct hub taxa for each group, though there was substantial overlap in inventories—especially between SCZ and HC, reflecting our PCoA results. All networks shared a stable, non-differential core consisting of the indole metabolizer *Alistipes* and the butyrate producers *Oscillospiraceae UCG-005* and less robustly, *UCG-002*. The generally insignificant differences in eigen-

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

vector centrality further suggest the presence of a preserved topological anchor, around which less stable, disorder-associated peripheral rewiring takes place. Among taxa with differential centrality, BD exhibited increased centrality of the pro-inflammatory genera *Flavonifractor* and *Fusobacterium*, aligning with previous reports of *Flavonifractor* enrichment in newly diagnosed BD and their relatives [128] and the established role of *Fusobacterium* in inflammation [123]. In MDD, there was increased centrality of taxa characterized by potentially opportunistic or neuroactive metabolic properties, such as GABA-associated *Bacteroides* and *Parabacteroides* [129,130], as well as the opportunistic gut commensal *Eggerthella* [131]. Unlike BD, MDD also showed reduced centrality of butyrate producers, consistent with a recent microbiome-wide association study [132]. Additionally, the higher centrality of *Parabacteroides* in MDD compared to BD provides additional grounds for further scrutiny of the aforementioned biomarker proposal by Majie et al. (2026) [127]. Notably, SCZ was least differentiated from HC, aside from increased centrality of oral-associated *Actinomyces* and decreased centrality of butyrate-associated *Butyricicoccus* and *Oscillospiraceae UCG-005* [133,134]. Overall, these findings suggest that the healthy network is more centrally organized around butyrate producers compared to MDD and SCZ.

Beyond hub designation, differences in network-wide topological measures were mostly non-significant and highly sensitive to modeling parameters; nonetheless, SCZ showed a trend toward increased modularity and reduced density. Alongside the reduced centrality of *Ruminococcaceae* and *Lachnospiraceae*, this aligns with Wang et al. (2023) [135], who linked the differential abundance of these taxa to weakened stabilizing competitive interactions, thereby promoting a more fragmented, compartmentalized network.

It is worth noting that, due to design constraints, adjustment for study effects was not undertaken, and we propose that all interpretations of network results be considered exploratory rather than robust evidence of functional or causal insights into the microbiome's role in psychopathology.

Consistent with Grinsztajn et al.'s (2022) observation that tree-based models tend to outperform deep learning models on tabular data [136], both of our models exhibited modest performance, with the random forest model outperforming the CNN ensemble. This was expected given the heterogeneous, multiclass, highly imbalanced, and high-dimensional nature of our dataset. The considerable drop in performance during LOSO-CV suggests overfitting to study-specific background features—an issue that, despite the heavy use of dropout and regularization, continued to affect the CNN ensemble.

Both models performed best when classifying SCZ and BD samples and worst when classifying HCs. However, we suspect that the small BD test sample size and the specific design of our LOSO-CV scheme may have inflated model performance for BD, and these results should not be overinterpreted. These findings suggest that microbiome data preserve some diagnostic signal, particularly in the SCZ group. This aligns with the intricate patterns observed in the differential abundance results—though contrary to the diversity and network analysis findings for this group. Crucially, it is also consistent with the notably stronger causal evidence for the microbiome's role in SCZ compared to other disorders [20]. Ultimately, we believe the current performance of these machine learning models remains inadequate for routine clinical application.

The motivation for this work stems from the pervasive burden of diagnostic uncertainty in day-to-day clinical psychiatric practice, which fundamentally undermines the quality of patient care. This challenge, coupled with the increasing availability of omics data and recent advances in artificial intelligence, has prompted many research groups to investigate omics signatures in psychiatric disorders with the ambitious goal of resolving psychiatry's core diagnostic dilemmas. Yet, despite significant progress in understanding psychopathology, this elusive goal remains out of reach. In psychiatric microbiome research, inconsistent results, pronounced heterogeneity, extensive overlap, small effect sizes, and modest classification performance appear to be the rule rather than the exception [137,138]. This situation is mirrored in parallel fields [139].

While we anticipate gaining significant insights into the pathophysiology of psychiatric disorders and improving classifier performance through the integration of multi-omics datasets, increased se-

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

quencing depth, and much larger cohorts, we believe the current problem cannot be reduced solely to a lack of sufficient data. Rather, we posit that the inadequate construct validity of existing diagnostic categories also contributes to the present impasse [140,141]. Integrating multi-omics research with transdiagnostic computational endophenotyping approaches—such as the B-SNIP project [142]—which seek to replace consensus-based, sociohistorically contingent, mechanism-agnostic diagnostic categories with evidence-based, data-driven, and biologically grounded systems, offers a promising path forward.

Although microbiome-derived biomarkers do not yet achieve sufficient discriminatory performance for realistic utilization in routine clinical practice—and their contribution remains largely confined to generating and exploring hypotheses regarding the pathophysiology of psychiatric disorders—their integration with other contemporary research trends may yield profound insights into psychopathology and, eventually, useful clinical decision support systems in the future.

4.1. Limitations and Strengths

These results are bounded by several key constraints. First, substantial heterogeneity across the included studies, coupled with a lack of sample metadata on critical confounders (e.g., BMI, sex, age, pharmacological regimen), precluded explicit adjustment for these variables. Second, due to the near-complete confounding of disorder status with study origin, adjusting for study effects in cross-diagnostic comparisons and network analyses was not possible; these results are therefore reported as exploratory. This heterogeneity also necessitated reducing taxonomic resolution to the genus level to abstract away study-context noise.

For machine learning classification specifically, excessive class imbalance in the training data proved to be a major limitation, despite our use of data augmentation to ameliorate this constraint. Furthermore, while this study recruited a large sample size compared to similar studies, we believe that properly training a neural network of such complexity requires a substantially larger dataset. Additionally, the LOSO-CV scheme—which held out a fixed set of three studies plus one rotating study to ensure the presence of every label group in the test dataset—may have caused an underestimation of standard errors and a possible overestimation of performance, particularly for the BD group. Finally, given the cross-sectional design and lack of direct functional data, any mechanistic or causal inferences generated here must be interpreted as speculative and hypothesis-generating.

Strengths of the current study include a comparatively large sample size assembled through the harmonized reprocessing of raw sequences from multiple cohorts using a single pipeline. This provides substantially greater statistical power than most individual case-control studies in the field, while still acknowledging that it remains smaller than the largest aggregated repositories. Additional strengths include: a transdiagnostic design covering three disorders and HCs; sensitivity analyses for sparsification threshold, sequencing context, and study effects; and validation on held-out studies. Furthermore, the integration of co-occurrence network analysis, machine learning classification, multiple intersample distance metrics, α-diversity indices, and network topological measures provided a more comprehensive view of the data.

4.2. Directions for Future Research

Since large interstudy heterogeneity has been a major limiting factor in interpreting microbiome results, large cohort studies incorporating standardized sample processing, detailed confounder profiling, and standardized clinical questionnaires alongside deep metagenomic sequencing will be highly informative. Furthermore, multi-omics approaches—by integrating diverse biological perspectives on mental disorders—may uncover novel functional associations, enhance machine learning model accuracy, and enable causal inference through methods such as Mendelian randomization.

Given the aforementioned limitations of the current psychiatric diagnostic system, incorporating endophenotyping and metagenomic sequencing shows immense potential. In particular, computational psychiatry approaches that link microbiome alterations to computational endophenotypes represent promising avenues for investigation [143]. Additionally, integrating dimensional nosological

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

systems such as HiTOP with metagenomics [144]—and empirically comparing their explanatory power against categorical systems—could help determine whether dimensional approaches better capture the variation observed in microbiome composition and diversity.

Finally, it will be worthwhile to explore common transdiagnostic patterns across different psychiatric diagnoses relative to HCs, with the hope of identifying shared mechanistic pathways in the emergence of psychiatric disorders or discovering how psychiatric disorders commonly affect microbiome composition.

5. Conclusions

This study suggests a transdiagnostic pattern: reduced centrality and depletion of anti-inflammatory butyrate producers in MDD and SCZ, contrasted with increased influence of taxa associated with neurotransmitter metabolism or pro-inflammatory potential across BD, MDD, and SCZ. While the microbiome retains a discriminatory signal that machine learning models can capture, their current modest performance precludes clinical application. These findings are tempered by extensive group overlap and study heterogeneity, underscoring the strong influence of confounders. To avoid hyperbolic claims regarding disorder-specific biomarkers or machine learning capabilities, researchers must contextualize their results and address ongoing criticisms of categorical psychiatric diagnoses. Ultimately, integrating multi-omics with computational endophenotyping offers a promising path forward.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Acknowledgments: The authors thank Pardis Ghahramani for providing the computational infrastructure essential to this research. We acknowledge the use of AI-assisted tools (Grammarly, DeepSeek, and ChatGPT) solely for proofreading, language polishing, and readability enhancement of the manuscript. The authors affirm that they take full responsibility for the content and integrity of this article.

Funding: This paper is based on independent research funded solely by the authors. The authors declare that they have not received any external commissions or funding from academic institutions or business enterprises.

Conflicts of Interest: The authors declare no competing interests.

Ethics Approval and Consent to Participate: This secondary analysis used previously collected, publicly available 16S rRNA sequencing data and metadata from 17 published human studies, retrieved from public repositories using the accession numbers in Table 1; studies requiring controlled access or additional approvals were excluded. The present study involved no participant recruitment, contact, intervention, new specimen collection, or reporting of identifiable information. Ethical oversight and informed consent for the original data collection and deposition remain with the primary investigations cited in Table 1. As only openly available data were analyzed, no additional ethics approval or participant consent was sought.

Data Availability Statement: Raw sequencing data for each study can be accessed as described in Table 1. The code used for this study, and associated QIIME 2 artifacts, are available at GitHub.com/AVXPANAOR/Schizomachines [145]. All other relevant data supporting the study's findings are either included in the supplementary materials or available upon request from the corresponding author.

References

1. Fan Y, Fan A, Yang Z, Fan D. Global burden of mental disorders in 204 countries and territories, 1990–2021: results from the global burden of disease study 2021. BMC Psychiatry. 2025;25(1):486.
2. Institute for Health Metrics and Evaluation (IHME). GBD Results. Seattle, WA: Institute for Health Metrics and Evaluation, University of Washington; 2025. Online. Available from: https://vizhub.healthdata.org/gbd-results/.
3. Yung NCL, Wong CSM, Chan JKN, Chen EYH, Chang WC. Excess Mortality and Life-Years Lost in People With Schizophrenia and Other Non-affective Psychoses: An 11-Year Population-Based Cohort Study. Schizophrenia Bulletin. 2020 Oct;47(2):474-84.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

4. Bradford A, Meyer AND, Khan S, Giardina TD, Singh H. Diagnostic error in mental health: a review. BMJ Quality & Safety. 2024 Apr;33(10):663-72.
5. Vöhringer PA, Perlis RH. Discriminating Between Bipolar Disorder and Major Depressive Disorder. Psychiatric Clinics of North America. 2016 Mar;39(1):1-10.
6. Xiang Y, Zhang L, Wang G, Hu C, Ungvari GS, Dickerson FB, et al. Sociodemographic and clinical features of bipolar disorder patients misdiagnosed with major depressive disorder in China. Bipolar Disorders. 2013 Feb;15(2):199-205.
7. Hu C, Xiang YT, Ungvari GS, Dickerson FB, Kilbourne AM, Si TM, et al. Undiagnosed bipolar disorder in patients treated for major depression in China. Journal of Affective Disorders. 2012 Oct;140(2):181-6.
8. Altamura AC, Buoli M, Caldiroli A, Caron L, Cumerlato Melter C, Dobrea C, et al. Misdiagnosis, duration of untreated illness (DUI) and outcome in bipolar patients with psychotic symptoms: A naturalistic study. Journal of Affective Disorders. 2015 Aug;182:70-5.
9. McIntyre RS, Laliberté F, Germain G, MacKnight SD, Gillard P, Harrington A. The real-world health resource use and costs of misdiagnosing bipolar I disorder. Journal of Affective Disorders. 2022;316:26-33.
10. Maung K, Ohnmar H, Than W. Assessment of documentation of DSM-IV-TR Criteria A for diagnosis of schizophrenia in psychiatric unit, tertiary hospital, Malaysia. La Clinica Terapeutica. 2015 Feb:87-90.
11. Coulter C, Baker KK, Margolis RL. Specialized Consultation for Suspected Recent-onset Schizophrenia: Diagnostic Clarity and the Distorting Impact of Anxiety and Reported Auditory Hallucinations. Journal of Psychiatric Practice. 2019 Mar;25(2):76-81.
12. Lopez-Castroman J, Leiva-Murillo JM, Cegla-Schvartzman F, Blasco-Fontecilla H, Garcia-Nieto R, Artes-Rodriguez A, et al. Onset of schizophrenia diagnoses in a large clinical cohort. Scientific Reports. 2019 Jul;9(1).
13. Barnes A. Race and Schizophrenia Diagnoses in Four Types of Hospitals. Journal of Black Studies. 2013;44(6):665-81.
14. Olbert CM, Nagendra A, Buck B. Meta-analysis of Black vs. White racial disparity in schizophrenia diagnosis in the United States: Do structured assessments attenuate racial disparities? Journal of Abnormal Psychology. 2018 Jan;127(1):104-15.
15. Strakowski SM, Keck PE, Arnold LM, Collins J, Wilson RM, Fleck DE, et al. Ethnicity and Diagnosis in Patients With Affective Disorders. The Journal of Clinical Psychiatry. 2003 Jul;64(7):747-54.
16. Kamath S, Sokolenko E, Clark SR, Cross CB, Scott J, Wardill HR, et al. Distinguishing the causative, correlative and bidirectional roles of the gut microbiota in mental health. Nature Mental Health. 2025 Sep;3(10):1137-51.
17. Wu M, Tian T, Mao Q, Zou T, Zhou Cj, Xie J, et al. Associations between disordered gut microbiota and changes of neurotransmitters and short-chain fatty acids in depressed mice. Translational Psychiatry. 2020 Oct;10(1).
18. Zhou M, Fan Y, Xu L, Yu Z, Wang S, Xu H, et al. Microbiome and tryptophan metabolomics analysis in adolescent depression: roles of the gut microbiota in the regulation of tryptophan-derived neurotransmitters and behaviors in human and mice. Microbiome. 2023 Jun;11(1).
19. Siopi E, Galerne M, Rivagorda M, Saha S, Moigneu C, Moriceau S, et al. Gut microbiota changes require vagus nerve integrity to promote depressive-like behaviors in mice. Molecular Psychiatry. 2023 May;28(7):3002-12.
20. Wei N, Ju M, Su X, Zhang Y, Huang Y, Rao X, et al. Transplantation of gut microbiota derived from patients with schizophrenia induces schizophrenia-like behaviors and dysregulated brain transcript response in mice. Schizophrenia. 2024 Apr;10(1).
21. Knudsen JK, Michaelsen TY, Bundgaard-Nielsen C, Nielsen RE, Hjerrild S, Leutscher P, et al. Faecal microbiota transplantation from patients with depression or healthy individuals into rats modulates mood-related behaviour. Scientific Reports. 2021 Nov;11(1).
22. Delanote J, Correa Rojo A, Wells PM, Steves CJ, Ertaylan G. Systematic identification of the role of gut microbiota in mental disorders: a TwinsUK cohort study. Scientific Reports. 2024 Feb;14(1). Available from: http://dx.doi.org/10.1038/s41598-024-53929-w.
23. O'Riordan KJ, Moloney GM, Keane L, Clarke G, Cryan JF. The gut microbiota-immune-brain axis: Therapeutic implications. Cell Reports Medicine. 2025 Mar;6(3):101982. Available from: http://dx.doi.org/10.1016/j.xcrm.2025.101982.
24. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016 Dec;5(1). Available from: http://dx.doi.org/10.1186/s13643-016-0384-4.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

25. Amos GCA, Sergaki C, Logan A, Iriarte R, Bannaga A, Chandrapalan S, et al. Exploring how microbiome signatures change across inflammatory bowel disease conditions and disease locations. Scientific Reports. 2021 Sep;11(1). Available from: http://dx.doi.org/10.1038/s41598-021-96942-z.
26. Zheng J, Sun Q, Zhang M, Liu C, Su Q, Zhang L, et al. Noninvasive, microbiome-based diagnosis of inflammatory bowel disease. Nature Medicine. 2024 Oct;30(12):3555-67. Available from: http://dx.doi.org/10.1038/s41591-024-03280-4.
27. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology. 2019 Jul;37(8):852-7. Available from: http://dx.doi.org/10.1038/s41587-019-0209-9.
28. Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods. 2008 Feb;5(3):235-7. Available from: http://dx.doi.org/10.1038/nmeth.1184.
29. Hamady M, Knight R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. Genome Research. 2009 Apr;19(7):1141-52. Available from: http://dx.doi.org/10.1101/gr.085464.108.
30. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal. 2011 May;17(1):10. Available from: http://dx.doi.org/10.14806/ej.17.1.200.
31. Andrews S, Krueger F, Segonds-Pichon A, Biggins L, Krueger C, Wingett S. FastQC. Babraham, UK; 2012. Babraham Institute. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc.
32. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016 Oct;4:e2584. Available from: http://dx.doi.org/10.7717/peerj.2584.
33. Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nature Methods. 2013 Jan;10(1):57-9. Available from: http://dx.doi.org/10.1038/nmeth.2276.
34. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems. 2017 Apr;2(2). Available from: http://dx.doi.org/10.1128/msystems.00191-16.
35. Liu G, Li T, Zhu X, Zhang X, Wang J. An independent evaluation in a CRC patient cohort of microbiome 16S rRNA sequence analysis methods: OTU clustering, DADA2, and Deblur. Frontiers in Microbiology. 2023 Jul;14. Available from: http://dx.doi.org/10.3389/fmicb.2023.1178744.
36. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825-30.
37. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome. 2018;6(1):90. Available from: https://doi.org/10.1186/s40168-018-0470-z.
38. Robeson MS II, O'Rourke DR, Kaehler BD, Ziemski M, Dillon MR, Foster JT, et al. RESCRIPt: Reproducible sequence taxonomy reference database management. PLOS Computational Biology. 2021 11;17(11):1-37. Available from: https://doi.org/10.1371/journal.pcbi.1009581.
39. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research. 2012 Nov;41(D1):D590-6. Available from: http://dx.doi.org/10.1093/nar/gks1219.
40. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Research. 2013 Nov;42(D1):D643-8. Available from: http://dx.doi.org/10.1093/nar/gkt1209.
41. McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D, et al. The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. GigaScience. 2012;1(1):7.
42. Eddy SR. Accelerated Profile HMM Searches. PLoS Computational Biology. 2011 Oct;7(10):e1002195. Available from: http://dx.doi.org/10.1371/journal.pcbi.1002195.
43. Janssen S, McDonald D, Gonzalez A, Navas-Molina JA, Jiang L, Xu ZZ, et al. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. mSystems. 2018 Jun;3(3). Available from: http://dx.doi.org/10.1128/mSystems.00021-18.
44. Matsen FA, Kodner RB, Armbrust EV. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinformatics. 2010 Oct;11(1). Available from: http://dx.doi.org/10.1186/1471-2105-11-538.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

45. Matsen FA, Hoffman NG, Gallagher A, Stamatakis A. A Format for Phylogenetic Placements. PLoS ONE. 2012 Feb;7(2):e31009. Available from: http://dx.doi.org/10.1371/journal.pone.0031009.
46. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, et al. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome. 2017 Mar;5(1):27.
47. Cassol I, Ibañez M, Bustamante JP. Key features and guidelines for the application of microbial alpha diversity metrics. Scientific Reports. 2025 Jan;15(1). Available from: http://dx.doi.org/10.1038/s41598-024-77864-y.
48. Pielou EC. The measurement of diversity in different types of biological collections. Journal of theoretical biology. 1966;13:131-44.
49. Faith DP. Conservation evaluation and phylogenetic diversity. Biological conservation. 1992;61(1):1-10.
50. Armstrong G, Cantrell K, Huang S, McDonald D, Haiminen N, Carrieri AP, et al. Efficient computation of Faith's phylogenetic diversity with applications in characterizing microbiomes. Genome Research. 2021;31(11):2131-7. Available from: http://genome.cshlp.org/content/31/11/2131.abstract.
51. Shannon CE. A mathematical theory of communication. The Bell System Technical Journal. 1948;27:379-423, 623-56.
52. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. Journal of the American statistical Association. 1952;47(260):583-621.
53. Jaccard P. Nouvelles recherches sur la distribution florale. Bull Soc Vard Sci Nat. 1908;44:223-70.
54. Sørensen T. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Biol Skr. 1948;5:1-34.
55. Bray JR, Curtis JT. An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecological Monographs. 1957;27(4):325-49. Available from: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1942268.
56. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. The ISME Journal. 2011;5:169-72.
57. Lozupone C, Knight R. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. Applied and Environmental Microbiology. 2005;71(12):8228-35. Available from: https://aem.asm.org/content/71/12/8228.
58. McDonald D, Vázquez-Baeza Y, Koslicki D, McClelland J, Reeve N, Xu Z, et al. Striped UniFrac: enabling microbiome analysis at unprecedented scale. Nature Methods. 2018;15:847-8.
59. Lozupone CA, Hamady M, Kelley ST, Knight R. Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. Applied and Environmental Microbiology. 2007;73(5):1576-85. Available from: https://aem.asm.org/content/73/5/1576.
60. Hamady M, Lozupone C, Knight R. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. The ISME Journal. 2010;4:17-27.
61. Legendre P, Legendre L. In: Numerical Ecology. 3rd ed. Elsevier; 2012. p. 499.
62. Halko N, Martinsson PG, Shkolnisky Y, Tygert M. An Algorithm for the Principal Component Analysis of Large Data Sets. SIAM Journal on Scientific Computing. 2011 oct. Publisher: Society for Industrial and Applied Mathematics. Available from: https://epubs.siam.org/doi/abs/10.1137/100804139.
63. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral ecology. 2001;26(1):32-46.
64. Somerfield PJ, Clarke KR, Gorley RN. A generalised analysis of similarities (ANOSIM) statistic for designs with ordered factors. Austral Ecology. 2021 Jun;46(6):901-10. Available from: http://dx.doi.org/10.1111/aec.13043.
65. Anderson MJ. Distance-Based Tests for Homogeneity of Multivariate Dispersions. Biometrics. 2005 Oct;62(1):245-53. Available from: http://dx.doi.org/10.1111/j.1541-0420.2005.00440.x.
66. Anderson MJ, Ellingsen KE, McArdle BH. Multivariate dispersion as a measure of beta diversity. Ecology Letters. 2006 Apr;9(6):683-93. Available from: http://dx.doi.org/10.1111/j.1461-0248.2006.00926.x.
67. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. Nature communications. 2020;11(1):3514. Available from: https://www.nature.com/articles/s41467-020-17041-7.
68. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1995 Jan;57(1):289-300. Available from: http://dx.doi.org/10.1111/j.2517-6161.1995.tb02031.x.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

69. Watts SC, Ritchie SC, Inouye M, Holt KE. FastSpar: rapid and scalable correlation estimation for compositional data. Bioinformatics. 2018 Aug;35(6):1064-6. Available from: http://dx.doi.org/10.1093/bioinformatics/bty734.
70. van Dongen S, Enright AJ. Metric distances derived from cosine similarity and Pearson and Spearman correlations; 2012. Available from: https://arxiv.org/abs/1208.3145.
71. Peschel S, Müller CL, von Mutius E, Boulesteix AL, Depner M. NetCoMi: network construction and comparison for microbiome data in R. Briefings in Bioinformatics. 2020 Dec;22(4). Available from: http://dx.doi.org/10.1093/bib/bbaa290.
72. Hagberg AA, Schult DA, Swart PJ. Exploring Network Structure, Dynamics, and Function using NetworkX. In: Proceedings of the 7th Python in Science Conference. SciPy. SciPy; 2008. p. 11-5. Available from: http://dx.doi.org/10.25080/TCWV9851.
73. Page L, Brin S, Motwani R, Winograd T. The PageRank Citation Ranking: Bringing Order to the Web. Stanford InfoLab; 1999. 1999-66. Previous number = SIDL-WP-1999-0120. Available from: http://ilpubs.stanford.edu:8090/422/.
74. Bhar A, Gierse LC, Meene A, Wang H, Karte C, Schwaiger T, et al. Application of a maximal-clique based community detection algorithm to gut microbiome data reveals driver microbes during influenza A virus infection. Frontiers in Microbiology. 2022;Volume 13 - 2022. Available from: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2022.979320.
75. Layeghifard M, Li H, Wang PW, Donaldson SL, Coburn B, Clark ST, et al. Microbiome networks and change-point analysis reveal key community changes associated with cystic fibrosis pulmonary exacerbations. npj Biofilms and Microbiomes. 2019 Jan;5(1). Available from: http://dx.doi.org/10.1038/s41522-018-0077-y.
76. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment. 2008 oct;2008(10):P10008. Available from: https://doi.org/10.1088/1742-5468/2008/10/P10008.
77. Jun W, Barahona M, Yue-Jin T, Hong-Zhong D. Natural Connectivity of Complex Networks. Chinese Physics Letters. 2010;27(7):078902. Available from: https://doi.org/10.1088/0256-307X/27/7/078902.
78. Hubert L, Arabie P. Comparing partitions. Journal of Classification. 1985 Dec;2(1):193-218. Available from: http://dx.doi.org/10.1007/BF01908075.
79. Hočevar T, Demšar J. A combinatorial approach to graphlet counting. Bioinformatics. 2014 Feb;30(4):559-65. Available from: http://dx.doi.org/10.1093/bioinformatics/btt717.
80. Yaveroğlu ÖN, Malod-Dognin N, Davis D, Levnajic Z, Janjic V, Karapandza R, et al. Revealing the Hidden Language of Complex Networks. Scientific Reports. 2014 Apr;4(1). Available from: http://dx.doi.org/10.1038/srep04547.
81. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods. 2020;17:261-72.
82. Sharma D, Paterson AD, Xu W. TaxoNN: ensemble of neural networks on stratified microbiome data for disease prediction. Bioinformatics. 2020 May;36(17):4544-50. Available from: http://dx.doi.org/10.1093/bioinformatics/btaa542.
83. Nguyen HM, Cooper EW, Kamei K. Borderline over-sampling for imbalanced data classification. International Journal of Knowledge Engineering and Soft Data Paradigms. 2011;3(1):4. Available from: http://dx.doi.org/10.1504/IJKESDP.2011.039875.
84. Hansen LK, Salamon P. Neural network ensembles. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1990;12(10):993-1001.
85. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al.. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems; 2015. Software available from tensorflow.org. Available from: https://www.tensorflow.org/.
86. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burges CJ, Bottou L, Weinberger KQ, editors. Advances in Neural Information Processing Systems. vol. 25. Curran Associates, Inc.; 2012. Available from: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.
87. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. CoRR. 2015;abs/1502.03167. Available from: http://arxiv.org/abs/1502.03167.
88. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research. 2014;15(56):1929-58. Available from: http://jmlr.org/papers/v15/srivastava14a.html.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

89. Glorot X, Bordes A, Bengio Y. Deep Sparse Rectifier Neural Networks. In: Gordon G, Dunson D, Dudík M, editors. Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. vol. 15 of Proceedings of Machine Learning Research. Fort Lauderdale, FL, USA: PMLR; 2011. p. 315-23. Available from: https://proceedings.mlr.press/v15/glorot11a.html.
90. Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016. http://www.deeplearningbook.org.
91. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. arXiv; 2014. Available from: https://arxiv.org/abs/1412.6980.
92. Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32. Available from: https://doi.org/10.1023/A:1010933404324.
93. Grandini M, Bagli E, Visani G. Metrics for Multi-Class Classification: an Overview; 2020. Available from: https://arxiv.org/abs/2008.05756.
94. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. Campbell Systematic Reviews. 2022;18(2):e1230. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/cl2.1230.
95. Reininghaus EZ, Platzer M, Kohlhammer-Dohr A, Hamm C, Mörkl S, Bengesser SA, et al. PROVIT: Supplementary Probiotic Treatment and Vitamin B7 in Depression—A Randomized Controlled Trial. Nutrients. 2020 Nov;12(11):3422. Available from: http://dx.doi.org/10.3390/nu12113422.
96. Chen HM, Chung YCE, Chen HC, Liu YW, Chen IM, Lu ML, et al. Exploration of the relationship between gut microbiota and fecal microRNAs in patients with major depressive disorder. Scientific Reports. 2022 Dec;12(1). Available from: http://dx.doi.org/10.1038/s41598-022-24773-7.
97. Liu RT, Rowan-Nash AD, Sheehan AE, Walsh RFL, Sanzari CM, Korry BJ, et al. Reductions in anti-inflammatory gut bacteria are associated with depression in a sample of young adults. Brain, Behavior, and Immunity. 2020 Aug;88:308-24. Available from: http://dx.doi.org/10.1016/j.bbi.2020.03.026.
98. Jiang H, Pan L, Zhang X, Zhang Z, Zhou Y, Ruan B. Altered gut bacterial–fungal interkingdom networks in patients with current depressive episode. Brain and Behavior. 2020 Jun;10(8). Available from: http://dx.doi.org/10.1002/brb3.1677.
99. Dong Z, Shen X, Hao Y, Li J, Xu H, Yin L, et al. Gut microbiome: A potential indicator for predicting treatment outcomes in major depressive disorder. Frontiers in Neuroscience. 2022 Jul;16. Available from: http://dx.doi.org/10.3389/fnins.2022.813075.
100. Liu P, Gao M, Liu Z, Zhang Y, Tu H, Lei L, et al. Gut Microbiome Composition Linked to Inflammatory Factors and Cognitive Functions in First-Episode, Drug-Naive Major Depressive Disorder Patients. Frontiers in Neuroscience. 2022 Jan;15. Available from: http://dx.doi.org/10.3389/fnins.2021.800764.
101. Dong Z, Shen X, Hao Y, Li J, Li H, Xu H, et al. Gut Microbiome: A Potential Indicator for Differential Diagnosis of Major Depressive Disorder and General Anxiety Disorder. Frontiers in Psychiatry. 2021 Sep;12. Available from: http://dx.doi.org/10.3389/fpsyt.2021.651536.
102. Zhang Y, Zhang R, Liu P, Wang J, Gao M, Zhang J, et al. Characteristics and Mediating Effect of Gut Microbiota With Experience of Childhood Maltreatment in Major Depressive Disorder. Frontiers in Neuroscience. 2022 Jun;16. Available from: http://dx.doi.org/10.3389/fnins.2022.926450.
103. Ye X, Wang D, Zhu H, Wang D, Li J, Tang Y, et al. Gut Microbiota Changes in Patients With Major Depressive Disorder Treated With Vortioxetine. Frontiers in Psychiatry. 2021 May;12. Available from: http://dx.doi.org/10.3389/fpsyt.2021.641491.
104. Hu S, Li A, Huang T, Lai J, Li J, Sublette ME, et al. Gut Microbiota Changes in Patients with Bipolar Depression. Advanced Science. 2019 May;6(14). Available from: http://dx.doi.org/10.1002/advs.201900752.
105. O'Donnell M, Teasdale SB, Chua XY, Hardman J, Wu N, Curtis J, et al. The Role of the Microbiome in the Metabolic Health of People with Schizophrenia and Related Psychoses: Cross-Sectional and Pre-Post Lifestyle Intervention Analyses. Pathogens. 2022 Nov;11(11):1279. Available from: http://dx.doi.org/10.3390/pathogens11111279.
106. Nguyen TT, Kosciolek T, Daly RE, Vázquez-Baeza Y, Swafford A, Knight R, et al. Gut microbiome in Schizophrenia: Altered functional pathways related to immune modulation and atherosclerotic risk. Brain, Behavior, and Immunity. 2021 Jan;91:245-56. Available from: http://dx.doi.org/10.1016/j.bbi.2020.10.003.
107. Ling Z, Jin G, Yan X, Cheng Y, Shao L, Song Q, et al. Fecal Dysbiosis and Immune Dysfunction in Chinese Elderly Patients With Schizophrenia: An Observational Study. Frontiers in Cellular and Infection Microbiology. 2022 Jun;12. Available from: http://dx.doi.org/10.3389/fcimb.2022.886872.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

108. Xiang M, Zheng L, Pu D, Lin F, Ma X, Ye H, et al. Intestinal Microbes in Patients With Schizophrenia Undergoing Short-Term Treatment: Core Species Identification Based on Co-Occurrence Networks and Regression Analysis. Frontiers in Microbiology. 2022 Jun;13. Available from: http://dx.doi.org/10.3389/fmicb.2022.909729.
109. Shen Y, Xu J, Li Z, Huang Y, Yuan Y, Wang J, et al. Analysis of gut microbiota diversity and auxiliary diagnosis as a biomarker in patients with schizophrenia: A cross-sectional study. Schizophrenia Research. 2018 Jul;197:470-7. Available from: http://dx.doi.org/10.1016/j.schres.2018.01.002.
110. Pan R, Zhang X, Gao J, Yi W, Wei Q, Su H. Analysis of the diversity of intestinal microbiome and its potential value as a biomarker in patients with schizophrenia: A cohort study. Psychiatry Research. 2020 Sep;291:113260. Available from: http://dx.doi.org/10.1016/j.psychres.2020.113260.
111. Tsai YL, Liu YW, Wang PN, Lin CY, Lan TH. Gender Differences in Gut Microbiome Composition Between Schizophrenia Patients With Normal Body Weight and Central Obesity. Frontiers in Psychiatry. 2022 Mar;13. Available from: http://dx.doi.org/10.3389/fpsyt.2022.836896.
112. Gotelli NJ, Ellison AM. A Primer of Ecological Statistics. 2nd ed. Oxford University Press; 2012.
113. McGuinness AJ, Davis JA, Dawson SL, Loughman A, Collier F, O'Hely M, et al. A systematic review of gut microbiota composition in observational studies of major depressive disorder, bipolar disorder and schizophrenia. Molecular Psychiatry. 2022;27(4):1920-35. Available from: https://doi.org/10.1038/s41380-022-01456-3.
114. Nikolova VL, Smith MRB, Hall LJ, Cleare AJ, Stone JM, Young AH. Perturbations in Gut Microbiota Composition in Psychiatric Disorders: A Review and Meta-analysis. JAMA Psychiatry. 2021 Dec;78(12):1343-54.
115. Knudsen JK, Bundgaard-Nielsen C, Hjerrild S, Nielsen R, Leutscher P, Sørensen S. Gut microbiota variations in patients diagnosed with major depressive disorder-A systematic review. Brain Behav. 2021 Jul;11(7):e02177.
116. Ioannou M, Borkent J, Andreu-Sánchez S, Wu J, Fu J, Sommer IEC, et al. Reproducible gut microbial signatures in bipolar and schizophrenia spectrum disorders: A metagenome-wide study. Brain Behav Immun. 2024 Oct;121:165-75.
117. Mistry D, Kurmi S, Chouhan U. Integrated in silico approaches for gut microbiome variations across coexisting diabetes and obesity. In Silico Research in Biomedicine. 2026;2:100380. Available from: https://www.sciencedirect.com/science/article/pii/S305078712600209X.
118. Toh TS, Chong CW, Lim SY, Bowman J, Cirstea M, Lin CH, et al. Gut microbiome in Parkinson's disease: New insights from meta-analysis. Parkinsonism & Related Disorders. 2022;94:1-9. Available from: https://www.sciencedirect.com/science/article/pii/S1353802021004211.
119. Brown RL, Larkinson MLY, Clarke TB. Immunological design of commensal communities to treat intestinal infection and inflammation. PLOS Pathogens. 2021 01;17(1):1-31. Available from: https://doi.org/10.1371/journal.ppat.1009191.
120. Kaur H, Bose C, Mande SS. Tryptophan Metabolism by Gut Microbiome and Gut-Brain-Axis: An in silico Analysis. Frontiers in Neuroscience. 2019;Volume 13 - 2019. Available from: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2019.01365.
121. D'Addario C, Pucci M, Bellia F, Girella A, Sabatucci A, Fanti F, et al. Regulation of oxytocin receptor gene expression in obsessive–compulsive disorder: a possible role for the microbiota-host epigenetic axis. Clinical Epigenetics. 2022;14(1):47. Available from: https://doi.org/10.1186/s13148-022-01264-0.
122. Li J, Li Y, Zhou Y, Wang C, Wu B, Wan J. Actinomyces and Alimentary Tract Diseases: A Review of Its Biological Functions and Pathology. BioMed Research International. 2018;2018(1):3820215. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1155/2018/3820215.
123. Li Z, Liu J, Li J, Zhou Z, Huang X, Gopinath D, et al. Fusobacterium in the microbiome: from health to disease across the oral–gut axis and beyond. npj Biofilms and Microbiomes. 2025;11(1):200. Available from: https://doi.org/10.1038/s41522-025-00838-z.
124. Eicher TP, Mohajeri MH. Overlapping Mechanisms of Action of Brain-Active Bacteria and Bacterial Metabolites in the Pathogenesis of Common Brain Diseases. Nutrients. 2022;14(13). Available from: https://www.mdpi.com/2072-6643/14/13/2661.
125. Zhang Q, Xu B, Min Y, Liu J, Shang Y, Lan X, et al. Evaluation of the safety and probiotic properties of GABA-producing Enterococcus faecium AB157 based on whole genome and phenotype analysis. LWT. 2025;215:117242. Available from: https://www.sciencedirect.com/science/article/pii/S0023643824015251.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

126. Mousavi R, Mottawea W, Audet MC, Hammami R. Survival and Interplay of γ-Aminobutyric Acid-Producing Psychobiotic Candidates with the Gut Microbiota in a Continuous Model of the Human Colon. Biology. 2022;11(9). Available from: https://www.mdpi.com/2079-7737/11/9/1311.
127. Majie W, Yifang H, Yuncui H, Yaping C, Longhui L, Qianyan J, et al. Parabacteroides distinguishes bipolar disorder from schizophrenia: toward a microbial biomarker for differential diagnosis. Frontiers in Microbiology. 2026;Volume 16 - 2025. Available from: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2025.1735998.
128. Coello K, Hansen TH, Sørensen N, Munkholm K, Kessing LV, Pedersen O, et al. Gut microbiota composition in patients with newly diagnosed bipolar disorder and their unaffected first-degree relatives. Brain, Behavior, and Immunity. 2019;75:112-8. Available from: https://www.sciencedirect.com/science/article/pii/S0889159118306287.
129. Otaru N, Ye K, Mujezinovic D, Berchtold L, Constancias F, Cornejo FA, et al. GABA Production by Human Intestinal Bacteroides spp.: Prevalence, Regulation, and Role in Acid Stress Tolerance. Frontiers in Microbiology. 2021;Volume 12 - 2021. Available from: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2021.656895.
130. Yousuf B, Mottawea W, Esmail GA, Nazemof N, Bouhlel NE, Njoku E, et al. Multi-omics unveils strain-specific neuroactive metabolite production linked to inflammation modulation by Bacteroides and their extracellular vesicles. Current Research in Microbial Sciences. 2025;8:100358. Available from: https://www.sciencedirect.com/science/article/pii/S2666517425000203.
131. Ugarte-Torres A, Gillrie MR, Griener TP, Church DL. Eggerthella lenta Bloodstream Infections Are Associated With Increased Mortality Following Empiric Piperacillin-Tazobactam (TZP) Monotherapy: A Population-based Cohort Study. Clinical Infectious Diseases. 2018 07;67(2):221-8. Available from: https://doi.org/10.1093/cid/ciy057.
132. Radjabzadeh D, Bosch JA, Uitterlinden A, Zwinderman AH, Ikram MA, van Meurs JBJ, et al. Gut microbiome-wide association study of depressive symptoms. Nature Communications. 2022;13(1):7128. Available from: https://doi.org/10.1038/s41467-022-34502-3.
133. Singh R, Stogios N, Smith E, Lee J, Maksyutynsk K, Au E, et al. Gut microbiome in schizophrenia and antipsychotic-induced metabolic alterations: a scoping review. Therapeutic Advances in Psychopharmacology. 2022;12:20451253221096525. PMID: 35600753. Available from: https://doi.org/10.1177/20451253221096525.
134. Kong L, Wang X, Chen G, Zhu Y, Wang L, Yan M, et al. Gut microbiome characteristics in individuals across different stages of schizophrenia spectrum disorders: A systematic review and meta-analysis. Neuroscience & Biobehavioral Reviews. 2025;173:106167. Available from: https://www.sciencedirect.com/science/article/pii/S0149763425001678.
135. Wang D, Russel WA, Sun Y, Belanger KD, Ay A. Machine learning and network analysis of the gut microbiome from patients with schizophrenia and non-psychiatric subject controls reveal behavioral risk factors and bacterial interactions. Schizophrenia Research. 2023;251:49-58. Available from: https://www.sciencedirect.com/science/article/pii/S0920996422004625.
136. Grinsztajn L, Oyallon E, Varoquaux G. Why do tree-based models still outperform deep learning on typical tabular data? In: Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K, Oh A, editors. Advances in Neural Information Processing Systems. vol. 35. Curran Associates, Inc.; 2022. p. 507-20. Available from: https://proceedings.neurips.cc/paper_files/paper/2022/file/0378c7692da36807bdec87ab043cdadc-Paper-Datasets_and_Benchmarks.pdf.
137. Almarzooqi S, Yassin LK, Alnuaimi F, Alketbi S, Skrabulyte-Barbulescu J, AlAhbabi N, et al. From association to causality: mechanistic pathways linking the gut microbiome to psychiatric disorders. Translational Psychiatry. 2026;16(1):561. Available from: https://doi.org/10.1038/s41398-026-04305-x.
138. Walter J, Armet AM, Finlay BB, Shanahan F. Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. Cell. 2020;180(2):221-32. Available from: https://doi.org/10.1016/j.cell.2019.12.025.
139. Grotzinger AD, Werme J, Peyrot WJ, Frei O, de Leeuw C, Bicks LK, et al. Mapping the genetic landscape across 14 psychiatric disorders. Nature. 2026;649(8096):406-15. Available from: https://doi.org/10.1038/s41586-025-09820-3.
140. Ruffalo ML, Pies RW. The Need for a Validity-Based Diagnostic Hierarchy in DSM-6. Psychiatric Services. 2026;0(0):appi.ps.20260333. Available from: https://psychiatryonline.org/doi/abs/10.1176/appi.ps.20260333.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 19 November 2026 https://doi.org/10.20944/preprints202611.1974.v1

141. Phillips J. The Conceptual Status of DSM-5 Diagnoses. In: Paris J, Phillips J, editors. Making the DSM-5: Concepts and Controversies. New York, NY: Springer New York; 2013. p. 143-57. Available from: https://doi.org/10.1007/978-1-4614-6504-1_10.
142. Clementz BA, Parker DA, Trotti RL, McDowell JE, Keedy SK, Keshavan MS, et al. Psychosis Biotypes: Replication and Validation from the B-SNIP Consortium. Schizophrenia Bulletin. 2022 01;48(1):56-68. Available from: https://doi.org/10.1093/schbul/sbab090.
143. Friston K. Computational psychiatry: from synapses to sentience. Molecular Psychiatry. 2023;28(1):256-68. Available from: https://doi.org/10.1038/s41380-022-01743-z.
144. Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. J Abnorm Psychol. 2017 May;126(4):454-77.
145. Bazdar P, Daneshmand A, Kazemi A, Jalilian A, Ghahramani P. Schizomachines Qiime2 Artifacts. Zenodo; 2026. Available from: https://doi.org/10.5281/zenodo.23598736.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.