npj Parkinson's Disease

https://doi.org/10.1038/s41531-026-01299-7

Article in Press

Gut ecosystem dysfunction in parkinson's disease: deciphering faecal metabolome-metagenome links for novel diagnostic panels

Received: 13 June 2025

Accepted: 12 February 2026

Cite this article as: Qian, Y., Xu, S., He, X. *et al.* Gut ecosystem dysfunction in parkinson's disease: deciphering faecal metabolome-metagenome links for novel diagnostic panels. *npj Parkinsons Dis.* (2026). https://doi.org/10.1038/s41531-026-01299-7

Yiwei Qian, Shaoqing Xu, Xiaoqin He, Yiqiu Lai, Yi Zhang, Chengjun Mo, Penghui Ai, Xiaodong Yang & Qin Xiao

Gut Ecosystem Dysfunction in Parkinson's Disease: Deciphering Faecal Metabolome-Metagenome Links for Novel Diagnostic Panels

Yiwei Qian[#1], Shaoqing Xu[#2], Xiaoqin He[3], Yiqiu Lai[1], Yi Zhang[1], Chengjun Mo[1], Penghui Ai[1], Xiaodong Yang[*1] and Qin Xiao[*1]

Affiliations:

[1] Department of Neurology and Institute of Neurology, Ruijin Hospital affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, P.R. China

[2] Department of Geriatrics, Ruijin Hospital affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, P.R. China

[3] Department of General Practice, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

[#] Yiwei Qian and Shaoqing Xu contributed equally to this work.

[*] Correspondence should be addressed to Xiaodong Yang and Qin Xiao.

Running Title: Gut Metabolomics and Metagenomics in PD patients

Corresponding authors:

Xiaodong Yang

Department of Neurology and Institute of Neurology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, P.R. China

Postal address: No. 197, Ruijin Er Road, Shanghai, P.R. China, 200025

Tel: +86 18521725657

Fax: +86-21-6445-4473

Email: yangxiaodong2011@yeah.net

Qin Xiao

Department of Neurology and Institute of Neurology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, P.R. China

Postal address: No. 197, Ruijin Er Road, Shanghai, P.R. China, 200025

Tel: +86 15921216530

Fax: +86-21-6445-4473

Email: qin_xiao2018@163.com

Abstract

Gut ecosystem dysfunction is implicated in Parkinson's disease (PD), but integrative faecal metabolome-metagenome links are undefined. We explored these interactions in Chinese PD patients to develop diagnostic panels. Targeted faecal metabolomics (LC–MS/MS) was performed on 132 PD and 113 healthy controls (HCs) and shotgun metagenomics was integrated for 39 PD/HC pairs. We identified 33 significantly altered faecal metabolites in PD (FDR-P<0.05). A novel 12-metabolite panel could distinguish PD from HCs. Multi-omic integration revealed gut ecosystem dysfunction manifests via co-disruptions in microbial genes (e.g., amino acid metabolism genes) and metabolites. Critically, a combinatorial diagnostic panel integrating faecal metabolites and microbial gene markers achieved exceptional PD detection (AUC=0.961, 95% CI=0.923-0.998). This study deciphers metabolome-metagenome links driving gut dysfunction in PD, identifying amino acid metabolism as a core perturbed pathway. The novel diagnostic panels provide mechanistic insights and clinical tools for PD precision diagnosis.

Introduction

Parkinson's disease (PD) is one of the most common neurodegenerative disorders and currently affects approximately 1.7% of people over 65 years of age in China[1]. The disease is primarily caused by progressive degeneration of dopaminergic neurons in the substantia nigra pars compacta and the inclusion of Lewy bodies (LBs) in the remaining neurons. Currently, the underlying molecular basis of PD is not completely understood. Moreover, because of the limited use of biomarkers, the diagnosis of PD largely depends on patient history and clinical interviews, which may lead to a high proportion of misdiagnoses[2]. Thus, identifying the molecular pathogenesis of PD and new diagnostic biomarkers for this disease are critical.

Cumulative evidence highlights the crucial contribution of the gut microbiota to PD, suggesting that microbial dysbiosis or imbalance could contribute to PD pathogenesis [3-5]. The metabolome, which comprises small-molecule metabolites, enables a comprehensive investigation of the microbial population and the interactions between the microbial ecosystem and host phenotypes[6,7]. Because faecal samples directly contain host-gut cometabolites, they have become one of the most widely used sample types for investigating the interactions between the host and the gut microbiota[8]. In recent years, a reduction in faecal short-chain fatty acids (SCFAs), including acetate, propionate and butyrate, has been identified in patients with PD compared to controls[9-11]. However, the complete profile of disturbed faecal metabolites associated with PD has not been fully elucidated.

Previous studies focusing on taxonomic information related to the faecal microbiome have proposed diagnostic models of PD using 16S sequencing or shotgun sequencing methods[4,12,13], but the results have not been consistent. To date, some of the existing integrative omics analyses of PD have been performed based on 16S sequencing data[14,15]; however, combined markers based on metabolomic and metagenomic data have been found to be more effective for other diseases, e.g., irritable bowel syndrome[16], hepatic cirrhosis[17], and depression[18]. Nevertheless, few studies have correlated the composition of the gut microbiota using the metagenome with the direct analysis of faecal metabolites in patients with PD, while Li et al found changes in the PD gut that

are consistent with a disruption in bile acid control based on the proteomic and transcriptomic analysis using intestinal tissue[19]. Previously, we identified a set of 25 gut microbial genes for identifying patients with PD and constructed a Parkinson's disease index (PDI) based on these 25 gene markers as a potentially reliable non-invasive diagnostic biomarker for PD [20]. However, it remains unclear whether the combination of the faecal metabolome and metagenome might be a potential reliable diagnostic marker for discriminating PD with greater accuracy.

Herein, we investigated the interactions between faecal metabolomic and microbial changes in Chinese PD patients. Furthermore, we developed a set of specific metabolites that could potentially be used as biomarkers to distinguish patients with PD from healthy controls and patients with other neurodegenerative diseases. In particular, the metagenomic data and PDI data from our previous study were reanalyzed[20] to characterize the reciprocal interaction between altered faecal metabolites and the gut microbiome in the gut ecosystem of PD patients and the potential use of these interactions as combined biomarkers for PD diagnosis to further uncover how these disturbed signatures modulate host metabolism.

Results

Demographic and clinical characteristics of the participants

A total of 132 patients with PD and 113 healthy controls were enrolled in this study (Figure. 1). The demographic and clinical characteristics of all participants and different subgroups are summarized in Table 1 and Supplementary Table 1. No differences in age, sex, body mass index (BMI), lifestyle factors or comorbidities were detected between the PD patients and healthy controls. Constipation and laxative use were more prevalent among the PD patients than among the healthy controls (87.1% vs. 3.5%, 50.8% vs. 2.7%, $P<0.001$).

Altered faecal metabolomics in Parkinson's disease patients

The global metabolomic characteristics of the faecal samples from 93 PD patients and 79 healthy controls (training set) were compared and are shown in Figure. 2. The

relative abundances and comparisons of the metabolite categories, including amino acids, fatty acids, SCFAs, organic acids, benzenoids, phenylpropanoic acids, phenols, pyridines, phenylpropanoids, indoles, carbohydrates, benzoic acids and carboxylic acids, in PD patients and healthy controls are shown in Figure. 2A and Supplementary Table 2. The levels of amino acids, phenols, pyridines, indoles, carbohydrates, benzoic acids and carboxylic acids were lower in PD patients, while the level of phenylpropanoids was higher in PD patients (false discovery rate (FDR)-P<0.05, Supplementary Table 2). The SCFA levels in PD patients seemed to be lower than those in healthy controls (FDR-P=0.0588; Supplementary Table 2). In particular, the acetic acid levels in PD patients were significantly lower than those in healthy controls (FDR-P<0.001; Supplementary Table 3), while the other SCFAs were not significantly different between the two groups. According to the partial least squares discriminant analysis (PLS-DA) score plots, the overall metabolomic profiles of PD patients were significantly different from those of healthy controls (P<0.001; Figure. 2B). Compared with healthy controls, PD patients had 28 downregulated faecal metabolites and 3 upregulated metabolites (methylglutaric acid, myristic acid and pentadecanoic acid) (FDR-P<0.05; Figure. 2C, Supplementary Table 3 & Supplementary Figure). The metabolites with decreased levels were involved mainly in amino acid metabolism, carbohydrate metabolism and energy metabolism. (FDR-P<0.05, Figure. 2D).

Associations between the altered metabolites and clinical features of PD patients

The associations between the altered metabolites and clinical features of PD patients were analysed and are shown in Figure. 3. After adjusting for age, sex, BMI, constipation and lifestyle factors (including smoking, drinking alcohol, drinking tea, drinking coffee and yogurt consumption), we found that pentadecanoic acid was positively associated with Unified Parkinson's Disease Rating Scale (UPDRS) part II scores (r=0.245, P=0.024; Figure. 3A), UPDRS part III scores (r=0.286, P=0.008; Figure. 3B) and total UPDRS scores (r=0.248, P=0.023; Figure. 3C) and negatively associated with Mini-Mental State Examination (MMSE) scores (r=-0.332, P=0.002; Figure. 3D) and Montreal Cognitive Assessment (MoCA) scores (r=-0.240, P=0.028;

Figure. 3E), suggesting that a higher level of pentadecanoic acid was associated with more severe motor symptoms and cognitive impairment in PD patients. After adjusting for age, sex, BMI, constipation and lifestyle factors, only dopamine agonist use was associated with a lower level of pentadecanoic acid (P=0.019; data not shown). In addition, the levels of tryptophan (r=-0.235, P=0.031), methionine and 4-hydroxybenzoic acid (r=-0.215, P=0.049) were negatively correlated with UPDRS part II scores (Supplementary Table 4). We also confirmed the correlations between faecal metabolites and nonmotor symptoms in patients with PD. (Supplementary Table 4).

Faecal metabolic markers for identifying Parkinson's disease

With the significantly different metabolites identified, the feature importance of the 33 altered faecal metabolites was calculated according to the Boruta method, and a panel of 12 confirmed metabolites was investigated to determine their potential diagnostic ole in PD (Figure. 4A). The combined 12 faecal metabolic markers, namely, citric acid, trans-aconitic acid, cis-aconitic acid, histidine, p-hydroxyphenylacetic acid, N-acetyltryptophan, acetic acid, docosahexaenoic acid (DHA), itaconic acid, aspartic acid, 5-hydroxy-tryptophan and glyceric acid, could discriminate 93 patients with PD from 79 healthy controls (training set) with an area under the receiver operating characteristic curve (AUC) of 0.831 (95% confidence interval (CI): 0.769-0.894, sensitivity=0.763, specificity=0.772, Figure. 4B). An AUC of 0.781 was validated in the test set for distinguishing 39 patients with PD from 34 healthy controls (95% CI=0.676-0.885, sensitivity=0.795, specificity=0.676; Figure. 4C). The numbers of subjects in the different cohorts are provided in the confusion matrix table (Supplemental Table 5).

Associations between the disturbed microbes in PD patients and faecal metabolites

As previously described in our study,[20] the potential value of gut metagenomic markers in PD diagnosis was investigated based on differential bacterial genes, and the PDI based on 25 bacterial gene markers was constructed in the same Chinese population, separating PD patients from healthy controls. To investigate the specific molecular links between altered microbial metabolism and host processes, faecal metabolomic

data were correlated with the metagenomic data from 39 PD patients and 39 healthy controls. A total of 227 correlations were identified between microbial gene markers and metabolites (P<0.05; Figure. 5A) via correlation analysis based on the faecal metabolomic and metagenomic data from the same cohort. After correction, 6 correlations were identified (FDR-P<0.05; Supplemental Table 6). The levels of 4-hydroxybenzoic acid, tryptophan, histidine and docosapentaenoic acid (DPA) were positively associated with gene 7, while the levels of β-alanine and serine were negatively associated with gene 23. As previously described, 25 gene markers were annotated using the NCBI nonredundant protein sequence, and 23 gene markers were annotated (Supplementary Table 7). Among these, 7 were annotated to the genus *Escherichia*, which was less common in PD patients; gene 23 was annotated to the species *Alistipes ihumii*, which was more common in PD patients, indicating some potential functional links between some species and faecal metabolites. A total of 156 correlations were also found between microbial species and faecal metabolites (P<0.05; Figure. 5B). After correction, 109 correlations were identified (FDR-P<0.05; Supplemental Table 8). The co-occurrence of the interaction networks of differential microbial species, microbial genes and faecal metabolites are also presented in Figure. 5C. For example, acetic acid and tryptophan were negatively associated with species such as s_Lachnospiraceae_bacterium_1_1_57FAA, s_Lachnospiraceae_bacterium_3_1_57FAA_CT1 and s_Clostridium_leptum and were positively associated with gene 5 and gene 6.

Combinatorial biomarkers for discriminating PD patients from healthy controls

Since we identified potential metabolic markers from the faecal metabolome, we validated the discriminatory power of the panel of 12 faecal metabolic markers and detected the PDI in the same cohort of 39 patients with PD and 39 healthy controls in a previous study (Figure. 5D, Supplemental Table 5). The panel of 12 faecal metabolomic markers discriminated patients with PD from healthy controls with an AUC of 0.854 (95% CI=0.774-0.934, sensitivity=0.821, specificity=0.744). The PDI could discriminate patients with PD from healthy controls with an AUC of 0.901 (95%

CI=0.838-0.965, sensitivity=0.923, specificity=0.744). The combination of the PDI and 12 faecal metabolomic biomarkers enabled the discrimination of PD patients from healthy controls with a high AUC of 0.961 (95% CI=0.923-0.998, sensitivity=0.923, specificity=0.872), indicating a greater discriminative performance than that of the individual microbial genes or metabolic markers.

Alterations in microbial function and the faecal metabolomic pathway associated with PD

To identify the biologically meaningful patterns based on the differentially abundant metabolites in faeces, pathway analysis was performed through the MePA database in Metaboanalyst, and Kyoto Encyclopedia of Genes and Genomes (KEGG) was also used. A total of 16 different metabolomic pathways were significantly disrupted in PD patients compared with healthy controls based on faecal metabolome analysis (FDR-P<0.05, Figure. 6A). The disturbed metabolomic pathway was mainly identified as the first-order pathway of metabolism. Specifically, β-alanine metabolism (β-alanine, histidine, aspartic acid and malonic acid), alanine, aspartate and glutamate metabolism (citric acid, aspartic acid and asparagine), glyoxylate and dicarboxylate metabolism (citric acid, acetic acid, glyceric acid, cis-aconitic acid, serine and tartaric acid) and biosynthesis of amino acids (citric acid, tryptophan, tyrosine, methionine, histidine, asparagine, serine, lysine and aspartic acid) were the top 4 pathways in metabolism. Additionally, "mineral absorption" (methionine, tryptophan, asparagine and serine) and "protein digestion and absorption" (tryptophan, acetic acid, methionine, histidine, asparagine, serine, β-alanine, lysine, aspartic acid and tyrosine) were the top 2 pathways in the first-order pathway of organismal systems (Figure. 6A).

Based on faecal metagenome analysis, a total of 284 differential KEGG orthologous genes (KO genes), which were involved mainly in metabolism, e.g., amino acid metabolism and carbohydrate, energy and lipid metabolism, were identified between PD patients and healthy controls[20]. Overall, amino acid metabolism was of particular interest in PD patients. We demonstrated that these differential faecal metabolites and

microbial genes were mainly related to the "glycine, serine and threonine metabolism" and "phenylalanine, tyrosine and tryptophan biosynthesis" pathways.

We demonstrated that the levels of the majority of genes (hprA, gckA, apgM, trpA, trpB and thrA) involved in the glycine, serine and threonine metabolism pathway were decreased in the PD group relative to healthy controls and that the levels of glyceric acid, aspartic acid, serine and the relative metabolite tryptophan were decreased in the PD group, while the gene trpB was increased in the PD group (Figure. 6B), indicating that the glycine, serine and threonine metabolism pathway was disrupted in patients with PD. Additionally, we found that the microbial species associated with glyceric acid, aspartic acid, serine and tryptophan, such as s_Clostridium_leptum, s_Lachnospiraceae_bacterium_1_1_57FAA and s_Lachnospiraceae_bacterium_3_1_57FAA_CT1, were increased in the PD group. In addition, we found that the phenylalanine, tyrosine and tryptophan biosynthesis pathways were disrupted in PD patients compared with healthy controls. Furthermore, we observed that the expression levels of the genes aroQ, aroK, trpE, trpC and trpA were decreased in PD patients (Figure. 6C), which might have resulted in decreased synthesis of tryptophan. Moreover, we observed downregulated expression levels of the aroQ, aroK, and aspB genes and lower levels of tyrosine. Additionally, we demonstrated that the increased levels of microbial species, such as s_Clostridium_leptum, s_Lachnospiraceae_bacterium_1_1_57FAA and s_Lachnospiraceae_bacterium_3_1_57FAA_CT1, in PD patients were associated with tryptophan and tyrosine, indicating the possible role of microbial species in faecal metabolism.

Discussion

To the best of our knowledge, this is the first report on the use of faecal metabolomics in a Chinese PD population and the first multiomic analysis for biomarker construction based on faecal metabolomics and metagenomics in patients with PD. In this work, we investigated metabolites using an liquid chromatography–tandem mass spectrometry (LC–MS) platform and found different metabolic profiles in faecal samples from PD

patients compared with those from healthy controls. A total of 33 altered faecal metabolites were annotated; these metabolites were involved mainly in amino acid metabolism, carbohydrate metabolism, energy metabolism and lipid metabolism. Moreover, exploiting the advantage of having paired metagenomic data, we found that some of the faecal metabolites were associated with gut microbes. The disturbed faecal metabolites and microbial genes were mostly and consistently mapped to amino acid pathways. More importantly, we constructed a combinatorial marker panel of faecal metabolites and microbial gene markers that could effectively distinguish PD patients from healthy controls and may facilitate the development of new biomarkers for PD diagnosis. Additionally, our findings contribute to understanding the roles of the gut ecosystem in PD pathogenesis.

Several studies have focused on faecal metabolites in PD patients (Supplemental Table 9). SCFAs are metabolites produced by the microbiota and play key roles in microbiota-gut-brain cross talk[21]. Research on gut microbial metabolism associated with PD patients has focused mainly on SCFAs (Supplemental Table 9), which include butyrate, propionate and acetate. We found that the acetic acid levels in PD patients were significantly lower than those in healthy controls, while the butyrate and propionate levels did not significantly differ between the two groups. The decrease in faecal SCFA levels, particularly acetic acid, is likely attributable to a combination of a reduced microbial capacity for SCFA production and an increase in microbes associated with SCFA consumption. This is supported by our previous metagenomic finding that PD patients exhibit a decrease in the abundance of microbial genes involved in SCFA biosynthesis pathways[20]. Furthermore, our novel integrated analysis of paired metagenomic and metabolomic data from the same cohort provides direct evidence: we identified significant negative correlations between acetic acid levels and several bacterial species increased in PD, including s_Clostridium_leptum, s_Lachnospiraceae_bacterium_3_1_57FAA_CT1, and s_Lachnospiraceae_bacterium_1_1_57FAA (Supplementary Table 8), suggesting these taxa may compete for or consume SCFAs, thereby contributing to their net

reduction. The lower level of faecal acetic acid in PD patients was also found in a previous study using gas chromatography[11]. Acetic acid is rapidly metabolized, contributes to gluconeogenesis, and enters the tricarboxylic acid (TCA) cycle[22]; these processes were also enriched in our research. Recently, we reported decreased concentrations of faecal acetic, propionic, and butyric acid and increased concentrations of plasma acetic and propionic acid in another cohort of PD patients, and the alterations in faecal and plasma SCFAs were associated with an impaired gut-blood barrier and might be aggravated by constipation in PD patients[9]. The decrease in faecal SCFA levels may be attributed mainly to the reduction in SCFA-producing bacteria in PD patients, as proven by previous studies[23,24].

In this study, we also detected a lower level of tyrosine in PD patients than in controls, which was consistent with the findings of the studies of Tan et al.[14] and Yan et al.[25]. Our group recently reported that plasma tyrosine levels were reduced in PD patients[26]. Tyrosine is a nonessential amino acid that originates from the diet or from another amino acid called phenylalanine in the body. A decrease in tyrosine may lead to a decrease in dopamine levels, especially in neurons. However, no evidence has demonstrated that tyrosine supplementation improves motor symptoms in patients with PD[27]. Levodopa and tyrosine also compete for absorption in the small intestine, which can interfere with the effectiveness of the drug[27]. The majority of patients included in our study were receiving levodopa medication. Our group discovered that the microbial tyrosine decarboxylases (tyrDCs) gene abundance could serve as a potential biomarker of levodopa responsiveness[28], which encoded by specific types of bacteria in the gut catalysing the conversion of tyrosine to tyramine and also catalysing the conversion of levodopa to dopamine[29]. The detailed mechanism by which tyrosine is associated with the gut microbiota in PD needs to be elucidated in the future. Moreover, lower levels of 5-hydroxy-tryptophan (5-HTP) and tryptophan were also detected in the PD samples. Consistent with our results, decreased serum levels of serine have been reported in PD patients[30]. The level of 5-HTP was found to be decreased in PD patients in plasma, CSF or even postmortem studies[31-33], which was also found to be decreased

in PD patients in the current study. Tryptophan is involved in the biosynthesis of proteins, is the precursor of 5-HTP, and tends to be present at lower concentrations in the serum of PD patients with psychiatric problems such as depression, panic attacks and anxiety[30]. Recently, Augustin A et al. reported that tryptophan deficiency was linked to longer colonic transit with brady/hypokinesia, tremor, sleep disorders and dysosmia in PD patients[34]. In addition, tryptophan may be broken down by the microbiota into neuroactive molecules, including indole derivatives and tryptamine and kynurenine metabolites.

Based on metabolomic findings in clinical and experimental models, the metabolic pathways that are strongly perturbed in PD patients were related to the metabolism of lipids, energy, fatty acids, bile acids, polyamine, and amino acids.[35], which was consistent with our study. Consistent with the correlation between the differentially abundant metabolites and microbiota, we determined a few pathways in which both the microbiota and metabolites were involved, mainly amino acid metabolism. As previously reported, the glycine, serine and threonine metabolism pathway was disturbed in PD[36], which was also mapped to altered metabolites, as well as to the consistent combination of microbial genes and metabolites. In addition, we demonstrated that the majority of genes involved in the phenylalanine, tyrosine and tryptophan biosynthesis pathways were downregulated in PD, which might have resulted in decreased levels of tryptophan and tyrosine. Regarding the identified microbe-metabolite associations, gut metagenomic data can exhibit strong multicollinearity between microbial taxa, meaning that correlations might be found between unrelated pairs of metabolites and microbes. More confidence in the microbe-metabolite associations could be obtained by checking if the identified microbial species have the metabolic capacity to produce or consume the associated metabolites.

Many metabolomic studies have identified potential biomarkers for PD through the use of blood, urine and CSF[37-39]; these biomarkers are summarized in Supplemental Table 8. Previous studies have found that some metabolites, e.g., SCFAs[9,10], can be used to

determine the diagnostic efficiency in PD; however, these were limited studies on the identification of diagnostic markers based on overall metabolism, which is more comprehensive than the use of partial metabolites in the diagnosis of PD. Moreover, faecal samples are highly interesting since collection of these samples is simple and clearly noninvasive. However, to date, less is known about the diagnostic performance of faecal metabolites in PD, which has been well tested for other diseases, such as irritable bowel syndrome, hepatic cirrhosis and depression[16-18]. In this study, we constructed a panel of 12 faecal metabolites, including citric acid, trans-aconitic acid, and acetic acid, which could discriminate patients with PD from healthy controls with an AUC of 0.831 in the training set and an AUC of 0.781 in the test set. These results indicated that the gut metabolomic markers identified here have important disease diagnostic value and considerable potential for clinical application. However, additional control groups with larger sample sizes might be needed in future studies. Age, sex, genetics, diet, stress, disease status and other factors largely affect metabolism. Previously, we reported the PDI based on a panel of 25 gene markers as a potentially valuable biomarker for diagnosing patients with PD. Particularly, combining signals of both types (microbes and metabolites) markedly improved the classification performance. We constructed a combined diagnostic model and found that the diagnostic efficiency of the combination therapy was better than that of the PDI or metabolites alone. Using paired microbiome and metabolomic data, we found that the panel of 12 faecal metabolomic markers could discriminate between the 39 patients with PD and 39 healthy controls, with an AUC of 0.854, and the PDI had an AUC of 0.901. The combination of the PDI and 12 faecal metabolomic biomarkers enabled the discrimination of PD patients from healthy controls with a high classification power of 0.961. Indeed, the faecal metabolome and microbiome have been combined for biomarker identification in diseases, such as irritable bowel syndrome[16], hepatic cirrhosis[17] and depression[18], and combined markers have been shown to yield a more robust discriminative performance than separate microbial genes or metabolic markers.

This study has several limitations. The main limitation of this study is its cross-sectional nature; all the observations are associative and cannot be used to determine the causal mechanism. In particular, analysing patients in the prodromal and de novo stages and performing a longitudinal study with a large sample size may provide further information on whether changes in the gut metabolome lead or follow the development of PD. Replication in a cohort from different populations is also needed to help investigate the stability of microbiome analysis for use as a diagnostic tool in the future. Additionally, the metagenome was available for only a subset of participants because of budgetary considerations; expanding this dataset could reveal additional microbe–metabolite relationships.

Taken together, our findings revealed metabolomic dysregulation in the faecal samples of PD patients, and a novel metabolic model was proposed for noninvasive diagnosis of PD. Moreover, using multiomic data, we discovered a combinatorial marker panel of faecal metabolites and microbial gene markers for distinguishing PD patients. We also outlined the interaction networks of differential microbial species, microbial genes and faecal metabolites. Together, our findings provide new clues for understanding the roles of the overall gut ecosystem in PD pathogenesis and may facilitate the development of novel diagnostic and treatment strategies for PD.

Methods

Study subjects

We recruited 132 patients with PD and 113 healthy controls from the Movement Disorders Clinic at the Department of Neurology of Ruijin Hospital. All patients were examined by at least two movement disorder specialists in the Department of Neurology of Ruijin Hospital. The inclusion criteria for the PD patients were as follows: (1) diagnosed with idiopathic PD according to the 2015 Movement Disorders society (MDS) clinical diagnostic criteria for PD[40]; (2) had no family history of PD extending to first-degree relatives; and (3) had no history of other serious chronic illnesses (e.g., heart failure, liver cirrhosis, malignancy, inflammatory bowel disease, or

haematological or autoimmune diseases). Healthy controls were recruited from healthy subjects exhibiting no significant disease symptoms. Individuals currently taking antibiotics or probiotic supplements in the three months prior to sample collection were excluded.

The study protocol was approved by the Research Ethics Committee of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine. All the subjects signed an informed consent form. Clinical data were collected through face-to-face interviews with movement disorder specialists.

Clinical data collection

The weight and height of each participant were measured and subsequently used to calculate BMI. Lifestyle factors (including smoking history, drinking history, drinking tea, drinking coffee and yogurt consumption), comorbidities (including hypertension and diabetes) and drug use were also recorded for each participant. All the individuals involved in our study were omnivorous. The following rating scales were used to assess the status of the patients. Constipation was assessed using the Rome III criteria. Clinical features, including the disease duration, the Hoehn and Yahr stage (H&Y stage), and the scores for UPDRS, Nonmotor Symptoms Questionnaire for PD (NMS-Quest), Hamilton Anxiety Scale (HAMA), Hamilton Depression Scale (HAMD-17), MMSE and MoCA, were evaluated. The levodopa equivalent doses (LEDs) for patients were calculated according to previous methods[41]. All of the assessments were performed when the patients were in the 'On' state. The antiparkinsonism medications included levodopa, dopamine agonists (pramipexole/piribedil), a monoamine oxidase B (MAO-B) inhibitor (selegiline), a catechol-O-methyltransferase (COMT) inhibitor (entacapone), benzhexol hydrochloride and amantadine. Motor complications were diagnosed according to the UPDRS Part IV (A) and IV (B) scores. All patients were assessed by well-trained doctors.

Sample preparation and LC–MS analysis for faecal metabolites

Each study participant was asked to collect a fecal sample in the morning using fecal collection containers. The containers were transferred on ice and stored at -80°C prior to processing. Faecal samples were freeze-dried, after which approximately 5 mg of each sample was carefully weighed. Then, 25 μL of ultrapure water was added, and 120 μL of methanol containing internal standard solution was added to extract the metabolites. After the samples were homogenized for 3 minutes and then centrifuged at 18000 × g for 20 minutes, 20 μL of the supernatant was carefully transferred to a 96-well plate for derivatization using a Biomek 4000 workstation (Biomek 4000, Beckman Coulter, Inc., Brea, CA). After derivatization, 400 μL of ice-cold 50% methanol solution was added to dilute the samples. Then, the plate was stored at -20 °C for 20 minutes and centrifuged at 4000 × g and 4 °C for 30 min. An aliquot of 135 μL of supernatant was transferred to a new 96-well plate for LC–MS analysis as described previously[42].

LC–MS analyses were performed using an ultra-performance liquid chromatography coupled to tandem mass spectrometry (UPLC–MS) system (ACQUITY UPLC-Xevo TQ-S, Waters Corp., Milford, MA, USA). The raw data files generated by these LC–MS analyses were processed using MassLynx software (v4.1; Waters, Milford, MA, USA) to perform peak integration, calibration, and quantitation for each metabolite. The raw data of the metabolomic data was quantitative, reliable analytical data on the metabolite concentrations. The targeted metabolite was detected using Q300 kit (Metabo-Profile, Shanghai, China).[42]

Comparisons of faecal metabolites

A total of 110 faecal metabolites were investigated and subjected to statistical analysis. A total of 132 PD patients and 113 healthy controls were randomly divided into a training dataset (70% of samples) and an independent test dataset (30% of samples), respectively. To explain the difference in global metabolic profiles between the PD patients and healthy controls, PLS-DA was performed using supervised multivariate regression techniques. No outlier value was found, so no sample was excluded. The normality of the distribution of the selected metabolites was tested utilizing the

Shapiro–Wilk test ($\alpha = 0.05$). The Mann–Whitney U test was used to compare the altered faecal metabolites between PD patients and healthy controls, followed by the Benjamini–Hochberg (BH) procedure to control for FDR and determine the adjusted P value. A volcano plot was drawn to show the differentially abundant metabolites screened based on one-dimensional statistical analysis.

Analysis of metabolome function

After obtaining the differentially abundant metabolites, KEGG metabolic pathways were analysed via the pathway analysis function of the Metaboanalyst 4.0 website (www.metaboanalys.ca). Pathway enrichment analysis was conducted using pathway data from *Homo sapiens* (human) pathway libraries. The FDRs were also calculated.

Identification of markers for PD based on faecal metabolites

Multivariate analysis was subsequently conducted based on the metabolomic data in the training dataset. Two machine learning methods, random forest (RF) and support vector machine (SVM), were used to screen potential biomarkers. The twelve most important metabolites selected by RF and SVM were combined, and the Boruta method was subsequently used to screen the final biomarkers for subsequent analysis. RF is a series of "weak learners" (such as multiple decision trees) are combined to obtain a "strong learner" with significantly improved performance through ensemble learning[43]. SVM is a powerful classifier that distinguishes different categories by finding the optimal boundaries between data points[44]. Boruta is a feature selection method based on RF, which selects key features with significantly better discriminative power than random permutation features[45]. Logistic regression analysis was subsequently performed for multivariate analysis, and receiver operating characteristic (ROC) curve analysis and 95% CIs were calculated to quantify the diagnostic performance of the selected metabolites. The diagnostic performance of potential metabolomic markers was investigated to discriminate PD patients from healthy controls.

Construction of the interaction network of the gut metabolome and microbiome

To elucidate the correlation between the faecal metabolome and microbiome in PD patients, we reanalysed our published metagenomic data for 40 PD patients and 40 age-matched healthy controls[20]. Among the subjects, the faecal metabolome was detected among 39 PD patients and 39 age-matched healthy controls (one sample in each group did not meet the requirements for metabolome detection). The set of 25 identified microbial gene markers and the 33 different metabolites from the same cohort was reanalysed, to perform the ROC analysis and the correlation among the 25 microbial genes, differential species and metabolites. On the basis of the microbial genes and metabolomic signatures of PD patients, logistic regression analysis was used to optimize the combinatorial biomarker panel (faecal metabolites and microbial gene markers), and the AUC with 95% CI was determined. In addition, the co-occurrence of bacteria, microbial genes and faecal metabolites was calculated on the basis of the relative abundances determined by Spearman's rank correlation coefficient ($P<0.05$). The network layout was calculated and visualized using a circular layout with Cytoscape software (version 3.1.1).

Statistical analysis

Statistical analysis was performed using the R studio package (ver. 3.1.0, the R Project for Statistical Computing) and SPSS software (ver. 21.0, SPSS, Inc., Chicago, IL, USA). To analyse the demographic and clinical characteristics, the $\chi 2$ test was used to analyse categorical variables, which are displayed as n (%). Student's t test was used to compare continuous variables, which are displayed as the means ± SDs. Differential abundance levels of metabolites were tested using the Wilcoxon rank sum test, and p values were corrected for multiple testing with the BH method for the FDR. Correlations between specific metabolites and clinical features in the PD group were evaluated using Spearman's partial correlation analysis correlation network analysis. The statistical significance level was set at $P < 0.05$.

Data Availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Acknowledgements

This study was funded by the National Natural Science Foundation of China (Grant Nos. 82171246, 81870998, 81901283 and 81801254), the Key Field Research and Development Program of Guangdong Province (Grant No. 2018B030337001), the Clinical Research Plan of SHDC (Grant No. SHDC2020CR3012A), the Shanghai Rising-Star Program (Grant No. 22QA1405700), the Shanghai Sailing Program (Grant No. 22YF1440200) and the National Key Research and Development Program of China (Grant No. 2022YFE0210100). The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Authors' contributions

YQ and SX performed the experiments, clinical analysis, and manuscript writing; XH and YL collected the samples; YZ, CM and PA helped to recruit the PD patients and controls; and XY performed the study design and manuscript revision. QX designed the study, recruited the PD patients and controls, managed the project, provided financial support and revised the manuscript. All the authors meet the qualifications for authorship and have reviewed and approved the final version of the manuscript.

Competing Interests

All authors declare no financial or non-financial competing interests.

Consent to publication

All the authors have approved the manuscript.

References

1. Zhang, Z.-X., *et al.* Parkinson's disease in China: prevalence in Beijing, Xian, and Shanghai.

The Lancet **365**, 595-597 (2005).

2. Armstrong, M.J. & Okun, M.S. Diagnosis and Treatment of Parkinson Disease. *Jama* **323**, 548 (2020).
3. Sampson, T.R., *et al.* Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. *Cell* **167**, 1469-1480 e1412 (2016).
4. Bedarf, J.R., *et al.* Functional implications of microbial and viral gut metagenome changes in early stage L-DOPA-naive Parkinson's disease patients. *Genome Med* **9**, 39 (2017).
5. Nishiwaki, H., *et al.* Meta-Analysis of Gut Dysbiosis in Parkinson's Disease. *Movement disorders : official journal of the Movement Disorder Society* **35**, 1626-1635 (2020).
6. Zierer, J., *et al.* The fecal metabolome as a functional readout of the gut microbiome. *Nat Genet* **50**, 790-795 (2018).
7. Tang, J. Microbial metabolomics. *Curr Genomics* **12**, 391-403 (2011).
8. Chen, M.X., Wang, S.Y., Kuo, C.H. & Tsai, I.L. Metabolome analysis for investigating host-gut microbiota interactions. *J Formos Med Assoc* **118 Suppl 1**, S10-S22 (2019).
9. Yang, X., *et al.* Parkinson's Disease Is Associated with Impaired Gut-Blood Barrier for Short-Chain Fatty Acids. *Movement disorders : official journal of the Movement Disorder Society* **37**, 1634-1643 (2022).
10. Chen, S.J., *et al.* Association of Fecal and Plasma Levels of Short-Chain Fatty Acids With Gut Microbiota and Clinical Severity in Patients With Parkinson Disease. *Neurology* **98**, e848-e858 (2022).
11. Unger, M.M., *et al.* Short chain fatty acids and gut microbiota differ between patients with Parkinson's disease and age-matched controls. *Parkinsonism Relat Disord* **32**, 66-72 (2016).
12. Lin, A., *et al.* Gut microbiota in patients with Parkinson's disease in southern China. *Parkinsonism Relat Disord* **53**, 82-88 (2018).
13. Qian, Y., *et al.* Alteration of the fecal microbiota in Chinese patients with Parkinson's disease. *Brain, Behavior, and Immunity* **70**, 194-202 (2018).
14. Tan, A.H., *et al.* Gut Microbial Ecosystem in Parkinson Disease: New Clinicobiological Insights from Multi-Omics. *Ann Neurol* **89**, 546-559 (2021).
15. Vascellari, S., *et al.* Gut Microbiota and Metabolome Alterations Associated with Parkinson's Disease. *mSystems* **5**(2020).
16. Jeffery, I.B., *et al.* Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. *Gastroenterology* **158**, 1016-1028 e1018 (2020).
17. Oh, T.G., *et al.* A Universal Gut-Microbiome-Derived Signature Predicts Cirrhosis. *Cell Metab* **32**, 878-888 e876 (2020).
18. Yu, M., *et al.* Variations in gut microbiota and fecal metabolic phenotype associated with depression by 16S rRNA gene sequencing and LC/MS-based metabolomics. *J Pharm Biomed Anal* **138**, 231-239 (2017).
19. Li, P., *et al.* Gut Microbiota Dysbiosis Is Associated with Elevated Bile Acids in Parkinson's Disease. *Metabolites* **11**(2021).
20. Qian, Y., *et al.* Gut metagenomics-derived genes as potential biomarkers of Parkinson's disease. *Brain* **143**, 2474-2489 (2020).
21. Dalile, B., Van Oudenhove, L., Vervliet, B. & Verbeke, K.A.-O. The role of short-chain fatty acids in microbiota-gut-brain communication.

22. Van Treuren, W. & Dodd, D. Microbial Contribution to the Human Metabolome: Implications for Health and Disease. *Annual Review of Pathology: Mechanisms of Disease* **15**, 345-369 (2020).
23. Cirstea, M.S., *et al.* Microbiota Composition and Metabolism Are Associated With Gut Function in Parkinson's Disease. *Movement disorders : official journal of the Movement Disorder Society* **35**, 1208-1217 (2020).
24. Keshavarzian, A., *et al.* Colonic bacterial composition in Parkinson's disease. *Movement disorders : official journal of the Movement Disorder Society* **30**, 1351-1360 (2015).
25. Yan, Z., *et al.* Alterations of gut microbiota and metabolome with Parkinson's disease. *Microb Pathog* **160**, 105187 (2021).
26. Zhang, Y., *et al.* Plasma branched-chain and aromatic amino acids correlate with the gut microbiota and severity of Parkinson's disease. *NPJ Parkinsons Dis* **8**, 48 (2022).
27. DiFrancisco-Donoghue, J., Rabin, E., Lamberg, E.M. & Werner, W.G. Effects of Tyrosine on Parkinson's Disease: A Randomized, Double-Blind, Placebo-Controlled Trial. *Mov Disord Clin Pract* **1**, 348-353 (2014).
28. Zhang, Y., *et al.* Association Between Microbial Tyrosine Decarboxylase Gene and Levodopa Responsiveness in Patients With Parkinson Disease. *Neurology* **99**, e2443-e2453 (2022).
29. Maini Rekdal, V., Bess, E.N., Bisanz, J.E., Turnbaugh, P.J. & Balskus, E.P. Discovery and inhibition of an interspecies gut bacterial pathway for Levodopa metabolism. *Science* **364**(2019).
30. Jimenez-Jimenez, F.J., Alonso-Navarro, H., Garcia-Martin, E. & Agundez, J.A.G. Cerebrospinal and blood levels of amino acids as potential biomarkers for Parkinson's disease: review and meta-analysis. *Eur J Neurol* **27**, 2336-2347 (2020).
31. Tong, Q., *et al.* Reduced plasma serotonin and 5-hydroxyindoleacetic acid levels in Parkinson's disease are associated with nonmotor symptoms. *Parkinsonism Relat Disord* **21**, 882-887 (2015).
32. Olivola, E., *et al.* Serotonin impairment in CSF of PD patients, without an apparent clinical counterpart. *PLoS One* **9**, e101763 (2014).
33. Shannak, K., *et al.* Noradrenaline, dopamine and serotonin levels and metabolism in the human hypothalamus: observations in Parkinson's disease and normal subjects.
34. Augustin, A., *et al.* Faecal metabolite deficit, gut inflammation and diet in Parkinson's disease: Integrative analysis indicates inflammatory response syndrome. *Clin Transl Med* **13**, e1152 (2023).
35. Shao, Y. & Le, W. Recent advances and perspectives of metabolomics-based investigations in Parkinson's disease. *Mol Neurodegener* **14**, 3 (2019).
36. Graham, S.F., *et al.* Biochemical Profiling of the Brain and Blood Metabolome in a Mouse Model of Prodromal Parkinson's Disease Reveals Distinct Metabolic Profiles. *J Proteome Res* **17**, 2460-2469 (2018).
37. Stoessel, D., *et al.* Promising Metabolite Profiles in the Plasma and CSF of Early Clinical Parkinson's Disease. *Front Aging Neurosci* **10**, 51 (2018).
38. Luan, H., *et al.* Elevated excretion of biopyrrin as a new marker for idiopathic Parkinson's disease. *Parkinsonism Relat Disord* **21**, 1371-1372 (2015).
39. Burte, F., *et al.* metabolic profiling of Parkinson's disease and mild cognitive impairment. *Movement disorders : official journal of the Movement Disorder Society* **32**, 927-932 (2017).
40. Postuma, R.B., *et al.* MDS clinical diagnostic criteria for Parkinson's disease. *Movement disorders : official journal of the Movement Disorder Society* **30**, 1591-1601 (2015).

41. Tomlinson, C.L., *et al*. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Movement disorders : official journal of the Movement Disorder Society* **25**, 2649-2653 (2010).
42. Xie, G., *et al*. A Metabolite Array Technology for Precision Medicine. *Anal Chem* **93**, 5709-5717 (2021).
43. Tin Kam, H. Random decision forests. in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Vol. 1 278-282 vol.271 (1995).
44. Cortes, C. & Vapnik, V. Support-vector networks. *Machine Learning* **20**, 273-297 (1995).
45. Kursa, M., Jankowski, A. & Rudnicki, W. Boruta - A System for Feature Selection. *Fundam. Inform.* **101**, 271-285 (2010).

Figure legends

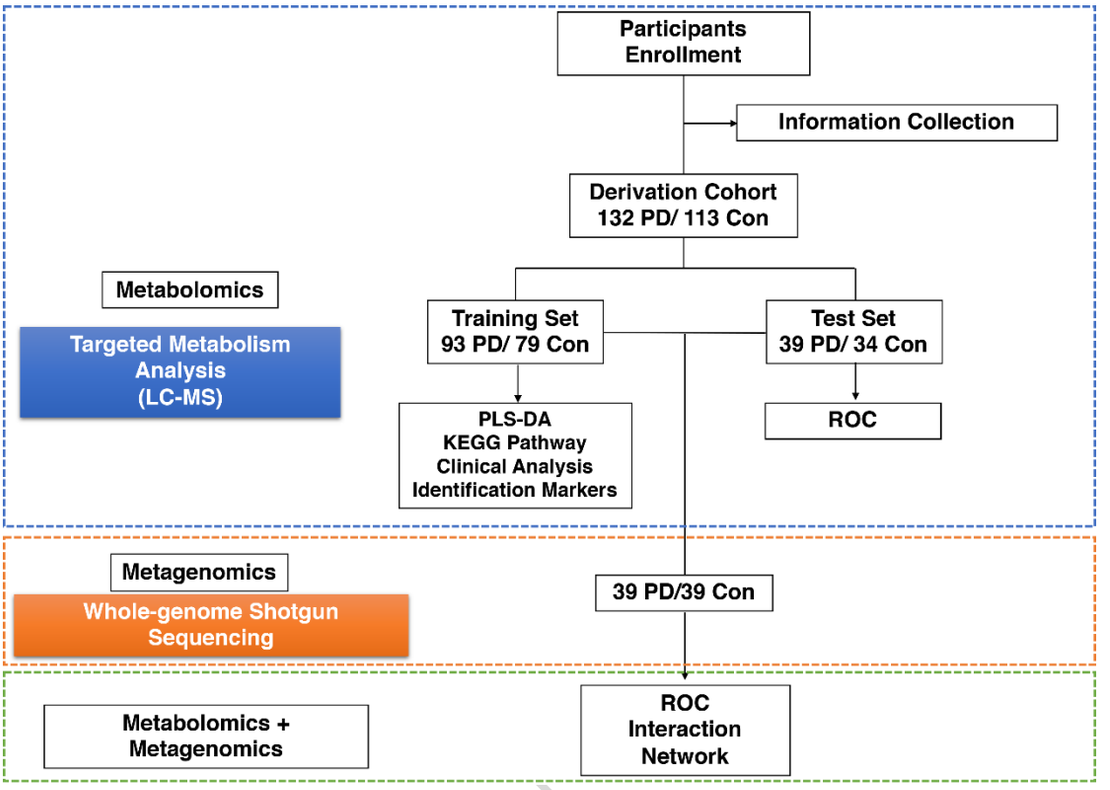

Figure. 1. Schematic diagram of the study design.

Con, control; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC–MS, liquid chromatography-tandem mass spectrometry; PD, Parkinson's disease; PLS-DA, partial least squares discriminant analysis; ROC, receiver operating characteristic.

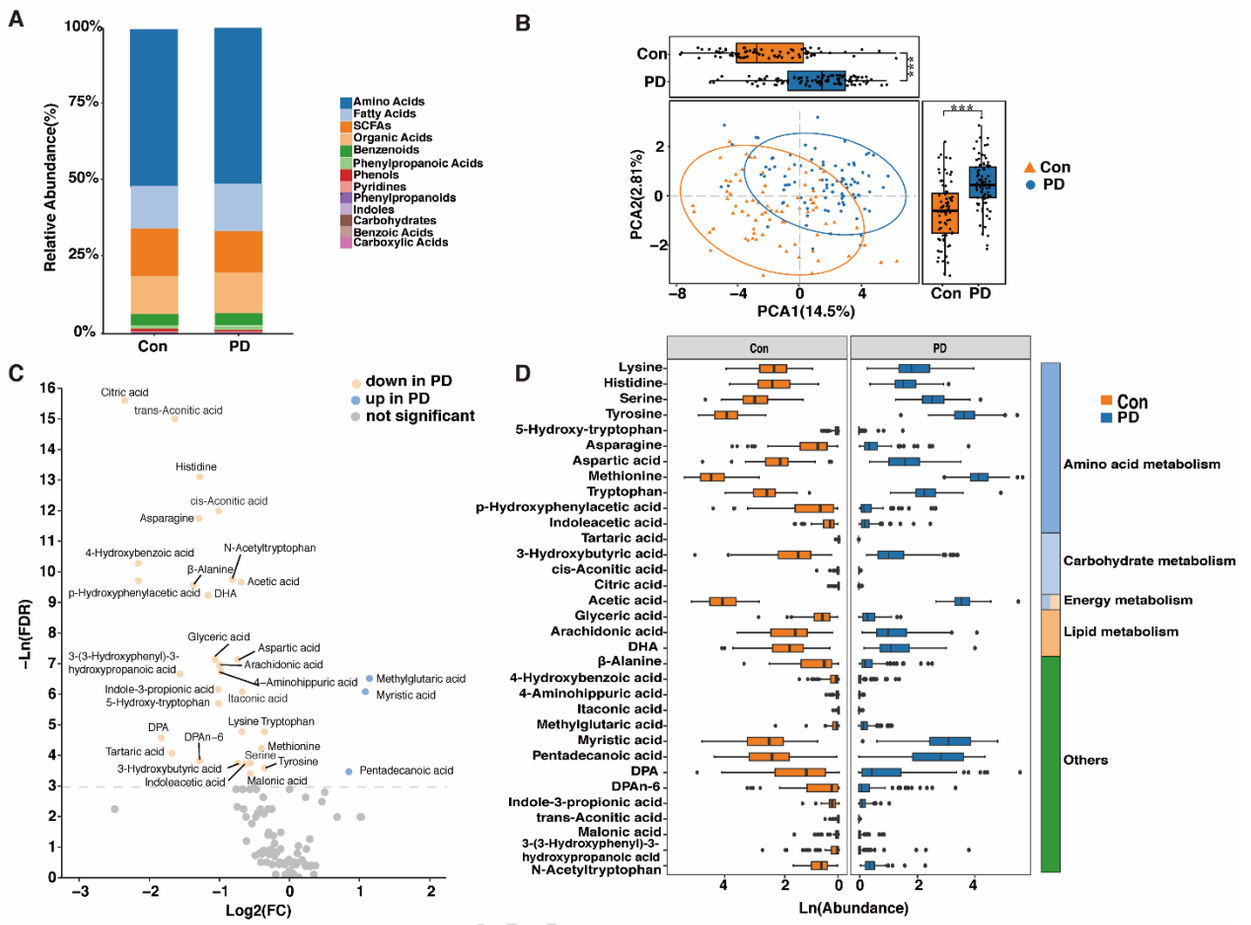

Figure. 2. Metabolomic characteristics of faecal metabolites.

(A) Relative abundances of metabolite categories identified between the two groups. (B) PLS-DA score plots of faecal metabolomic data were compared between patients with PD (blue) and healthy controls (orange). The two dimensions explained the greatest proportion of variance in the communities. Each symbol represents a sample. (C) Volcano plot comparing the fold change and p value of individual metabolites between the two groups. Light orange and light blue points indicate the significantly downregulated and upregulated metabolites, respectively, in the PD group. The grey points denote the metabolites with nonsignificant p values (FDR-P < 0.05). (D) Abundances of 33 differentiating faecal metabolites between PD patients (blue) and healthy controls (orange) were compared. These altered metabolites were involved mainly in amino acid, carbohydrate, energy, lipid and other metabolic processes. Each box represents the IQR, and the lines in the boxes indicate the median values. The whiskers show the lowest and highest values within the IQR from the first and third quartiles, respectively.

DHA, docosahexaenoic acid; DPA, docosapentaenoic acid; FC, fold change; FDR, false discovery rate; IQR, interquartile range; PCA, principal component analysis; PLS-DA, partial least squares discriminant analysis; SCFA, short-chain fatty acid.

***: P<0.001 PD vs. Con.

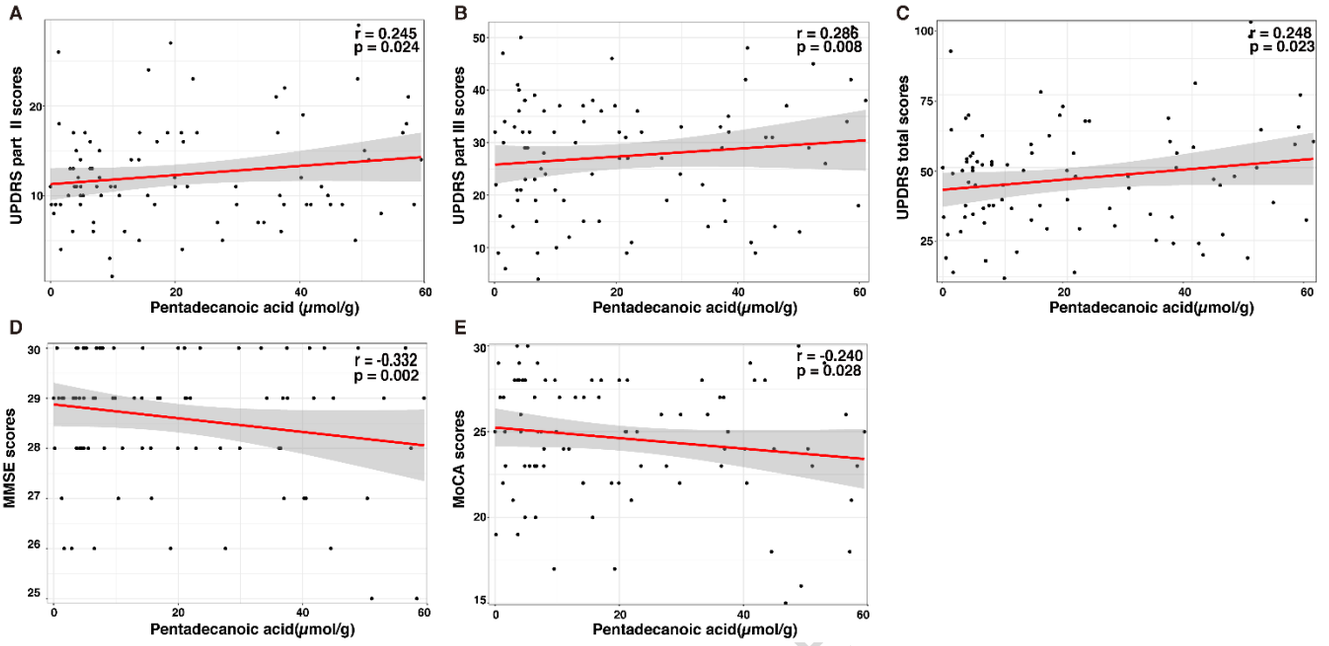

Figure. 3. Correlation between faecal metabolites and clinical features of PD patients.

(A) The associations between pentadecanoic acid levels and UPDRS part II scores were analysed by partial Spearman's correlation analysis after adjusting for age, sex, BMI, constipation and lifestyle factors. (B) The associations between pentadecanoic acid levels and UPDRS part III scores were analysed by partial Spearman's correlation analysis. (C) The associations between pentadecanoic acid concentrations and UPDRS total scores were analysed by partial Spearman's correlation analysis. (D) The associations between pentadecanoic acid concentrations and MMSE scores were analysed by partial Spearman's correlation analysis. (E) The associations between pentadecanoic acid concentrations and MoCA scores were analysed by partial Spearman's correlation analysis.

MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; UPDRS, Unified Parkinson's Disease Rating Scale.

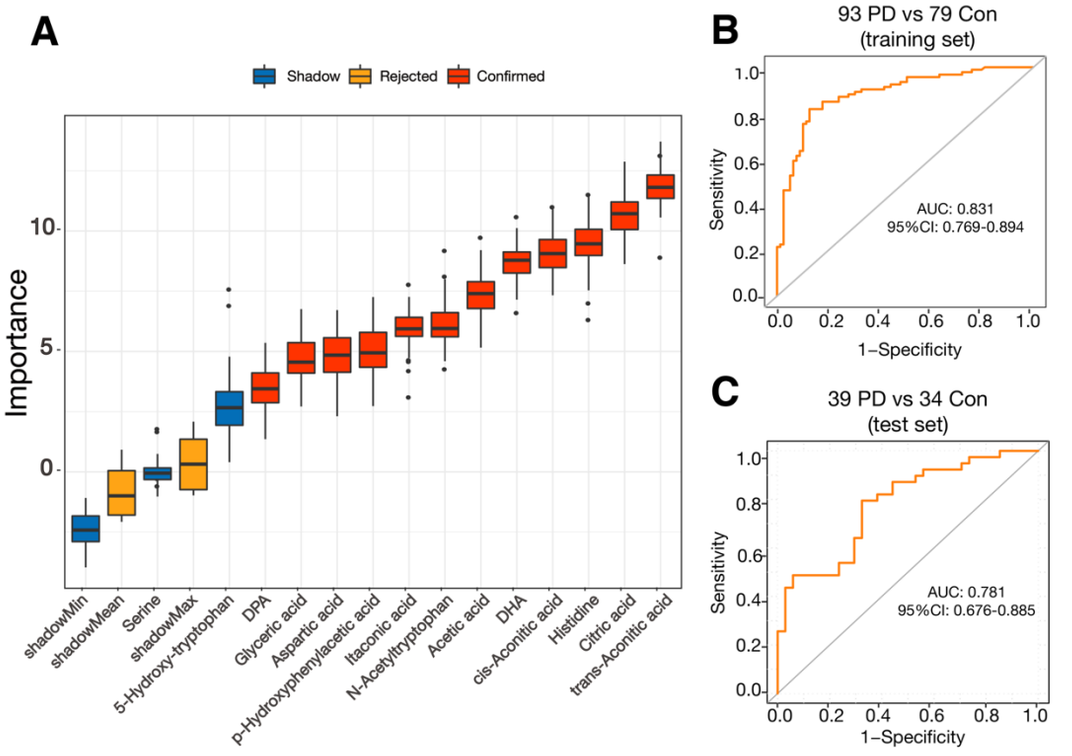

Figure. 4. Identification of potential biomarkers from the faecal metabolites for diagnosing PD patients.

(A) Feature importance calculated with the Boruta method; a panel of 12 confirmed metabolites (red) in the plot above served as biomarkers for subsequent modelling. The x-axis shows the feature variables coloured according to significantly different metabolites, and the y-axis shows importance in terms of Z scores. (B) ROC curve analysis of the logistic regression model showed that the panel of 12 metabolites could distinguish 93 PD patients from 79 healthy controls (in the training set). (C) ROC curve analysis was performed, and a logistic regression model was used to determine that the panel of 12 metabolites could distinguish 39 PD patients from 34 healthy controls (test set).

AUC, area under the ROC curve; CI, confidence interval; ROC, receiver operating characteristic.

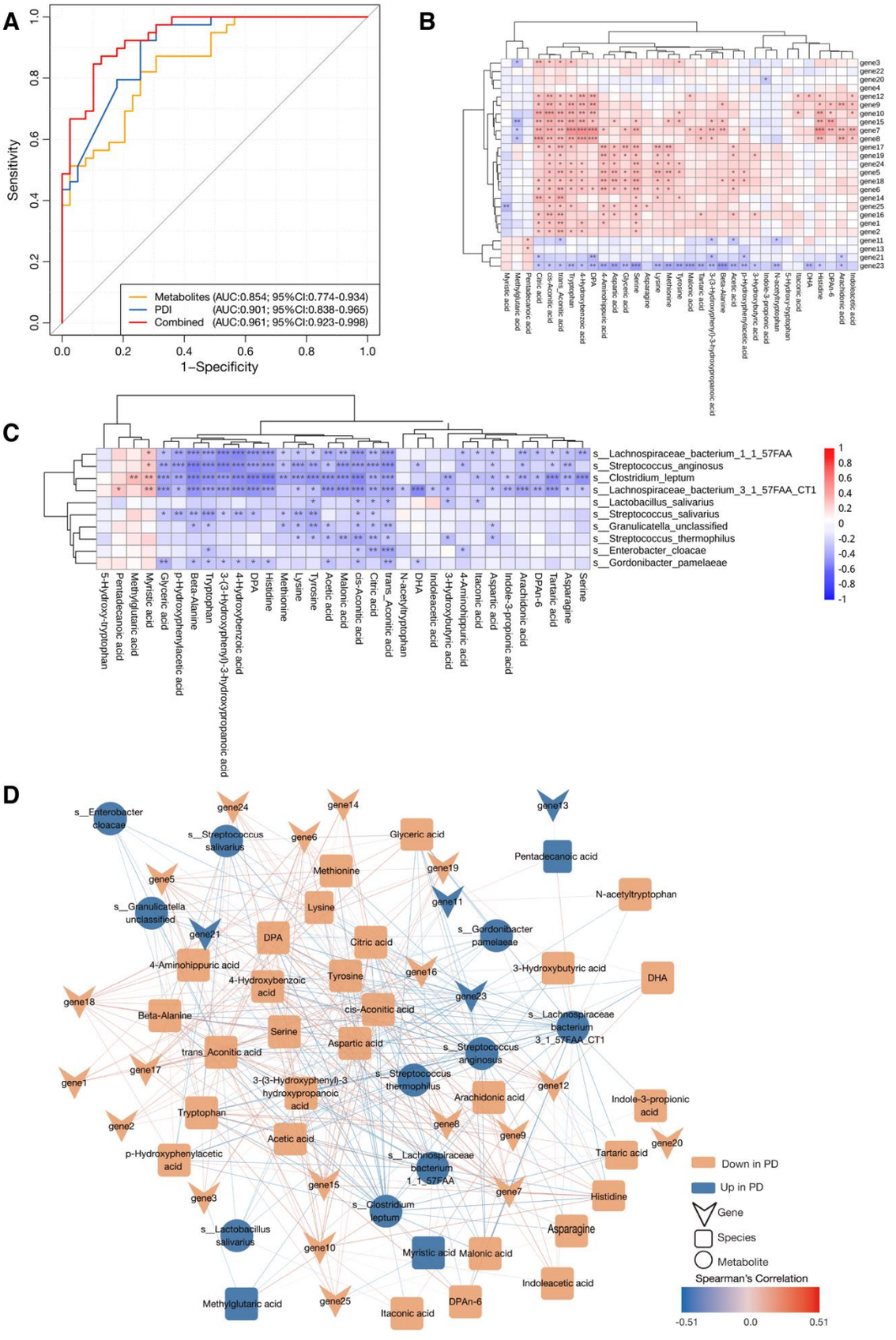

Figure. 5. Co-occurrence associations constructed from faecal metabolism and the metagenome of PD patients and healthy controls.

(A) The diagnostic performance of the 12 faecal metabolites and microbial gene markers was analysed in 39 PD patients and 39 healthy controls. ROC curve analyses of the panel of 12 metabolites (orange), the PDI (blue) and the combinatorial panel of 12 faecal metabolites and the PDI (red) were performed to distinguish the patients with PD from the healthy controls with AUC and 95% CI, respectively. (B) The correlations of 33 altered faecal metabolites and 25 bacterial gene markers were analysed by Spearman's correlation analysis in 39 PD patients and 39 healthy controls from the metagenome cohort. (C) The correlations between the altered faecal metabolites and disturbed species were analysed by Spearman's correlation analysis. (D) Visualization of the structure and co-occurrence network across faecal metabolites, microbial species and microbial genes. The arrow indicates the microbial genes, the square indicates the faecal metabolites, and the circle indicates the microbial species. Blue indicates increased levels of genes, microbial species or metabolites in PD patients, and orange indicates downregulated genes or metabolites in PD patients. The colour scale represents correlation coefficients. Red denotes positive correlations, and blue denotes negative correlations.

AUC, area under the ROC curve; CI, confidence interval; PDI, Parkinson's disease index.

*: P<0.05, **: P <0.01, ***: P <0.001

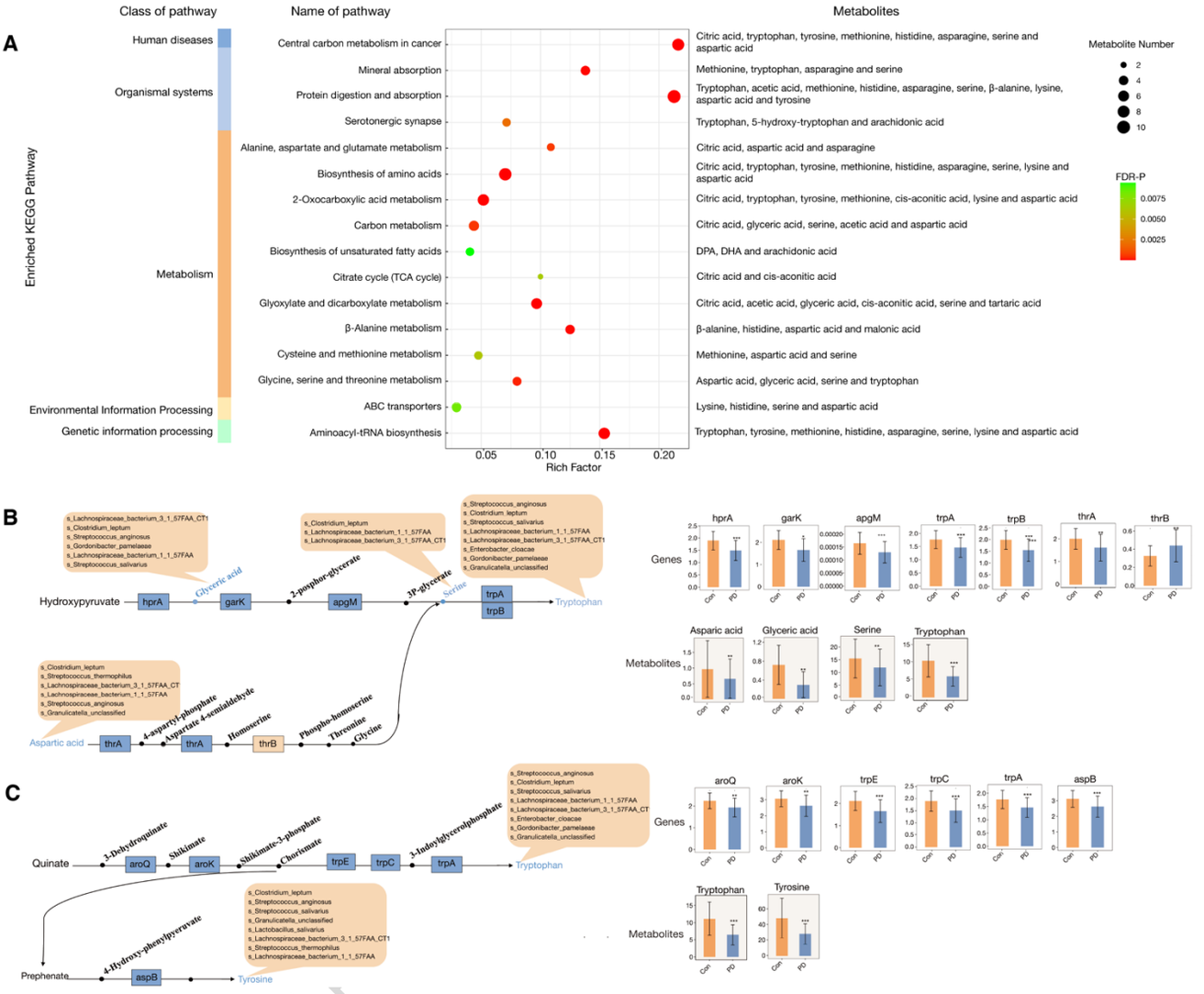

Figure. 6. Biological processes enriched in differential faecal metabolites and key amino acid metabolic pathways mapped to both the microbiota and faecal metabolites associated with PD.

(A) The altered faecal metabolites identified in the faecal metabolism cohort were subjected to pathway enrichment analysis using the KEGG database. The enrichment factor represents the ratio of the proportion of different metabolites in the pathway to the proportion of all metabolites in the pathway. The size of the bubbles indicates the number of metabolites enriched in the corresponding pathway, and the associated metabolites are listed on the right side. The disturbed metabolomic pathways were also identified and clustered into the different KEGG first-order pathways listed on the left side. The colour of the bubbles indicates the level of significance (FDR-P). (B) Based on the faecal metabolome and metagenome analysis, a disrupted glycine, serine and threonine metabolism pathway was identified in PD patients compared with healthy controls. The levels of most of the genes (hprA, gckA, apgM, trpA, trpB and thrA) involved and the relative levels of aspartic acid, glyceric acid, serine and tryptophan were decreased in the PD group, while the level of the thrB gene increased in the PD group. The microbial species associated with glyceric acid, aspartic acid, serine and

tryptophan were increased in the PD group. (C) Disturbed phenylalanine, tyrosine and tryptophan biosynthesis pathways were identified in patients with PD. The KEGG genes and different metabolites are coloured. Orange indicates increased genes, and blue indicates downregulated microbial genes or faecal metabolites in the PD group. The pathways were generated on the basis of KEGG pathway enrichment analysis.
ABC, ATP-binding cassette; DHA, docosahexaenoic acid; DPA, docosapentaenoic acid; FDR, false discovery rate; KEGG, Kyoto Encyclopedia of Genes and Genomes; TCA, tricarboxylic acid cycle.
*: P<0.05, **: P <0.01, ***: P <0.001 vs. con.

Table 1. Demographic and clinical characteristics of the participants in the study

	Con (n = 113)	PD (n = 132)	P
Age (years)	66.8 ± 7.5	68.0 ± 7.1	0.194
Female, n (%)	55 (48.7)	63 (47.7)	0.883
BMI (kg/m^2)	23.4 ± 3.4	22.9 ± 2.9	0.159
Smoking, n (%)	19 (16.8)	14 (10.6)	0.156
Drinking alcohol, n (%)	14 (12.4)	15 (11.4)	0.804
Drinking tea, n (%)	31 (27.4)	35 (26.5)	0.872
Drinking coffee, n (%)	5 (4.4)	13 (9.8)	0.105
Yogurt use, n (%)	16 (14.2)	18 (13.6)	0.906
Hypertension, n (%)	22 (19.5)	32 (24.2)	0.369
Diabetes, n (%)	9 (8.0)	7 (5.3)	0.401
Constipation, n (%)	4 (3.5)	115 (87.1)	< 0.001
Laxative use, n (%)	3 (2.7)	67 (50.8)	<0.001
Proton pump inhibitor use, n (%)	5 (4.4)	2 (1.5)	0.328
Acetylsalicylic acid use, n (%)	9 (8.0)	18 (13.6)	0.158
Statin use, n (%)	9 (8.0)	8 (6.1)	0.559
Disease duration (years)	/	6.3 ± 4.7	/
H&Y stage	/	2.4 ± 0.7	/
UPDRS part II scores	/	12.5 ± 5.5	/
UPDRS part III scores	/	28.4 ± 12.4	/
UPDRS total scores	/	46.5 ± 18.5	/
NMS scores	/	11.2 ± 5.4	/
HAMD-17 scores	/	7.2 ± 6.0	/
HAMA scores	/	12.6 ± 7.2	/
MMSE scores	/	28.2 ± 2.0	/
MoCA scores	/	24.6 ± 3.7	/
PD medication, n (%)	/	132 (100.0)	/
Levodopa use, n (%)	/	120 (90.9)	/
Dopamine agonist use, n (%)	/	83 (62.9)	/
COMT inhibitor use, n (%)	/	12 (9.1)	/
MAO-B inhibitor use, n (%)	/	38 (28.8)	/
Benzhexol hydrochloride use, n (%)	/	11 (8.3)	/
Amantadine use, n (%)	/	17 (12.9)	
LED (mg/day)	/	505.9 ± 309.3	/
Motor complications, n (%)	/	49 (37.1)	/

Categorical data are presented as number (percentage), continuous data are presented as mean ± standard deviation (SD). Differences between groups were assessed using the chi-square test for categorical data and t-test for continuous data.

P, where differences in the characteristics between patients with PD and Con were analysed.

Significance levels was set at 0.017 (2-tailed) after corrected by Bonferroni method for the multiple testing for the demographic and basic clinical characteristics (comorbidities).

BMI, Body Mass Index; COMT, catechol-O-methyltransferase; Con, controls; HAMA, Hamilton Anxiety Scale; HAMD-17, Hamilton Depression Scale; H&Y stage, Hoehn and Yahr stage; LED, levodopa equivalent dosage; MAO-B, monoamine oxidase B; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; NMS, non-motor symptoms; PD, Parkinson's disease; UPDRS, Unified Parkinson's Disease Rating Scale.