npj | parkinson's disease

Article

Published in partnership with the Parkinson's Foundation

https://doi.org/10.1038/s41531-025-01040-w

Enrichment of gut-derived metabolites in a Parkinson's disease subtype with REM sleep behavior disorder

Sunjae Lee[1], Jihyun Kim[2], Jae Woo Baek[1], Ki-Young Jung[3], Yunjong Lee[4], Ara Koh[2] ✉ & Han-Joon Kim[3] ✉

Recent reports indicate that Parkinson's disease (PD)-like changes in gut microbial signatures develop in idiopathic REM sleep behavior disorder (iRBD), a prodrome of α-synucleinopathies. However, <50% of PD cases exhibit RBD at onset, underscoring PD's heterogeneity. Using untargeted metabolomics, plasma samples from PD patients with and without RBD, iRBD subjects, and healthy controls were analyzed to characterize the metabolic differences in PD patients with and without RBD. Metabolomic analysis revealed the enrichment of gut microbial-origin metabolites, such as secondary bile acids and p-cresol sulfate, in PD patients with RBD, while metabolites linked to neuropsychiatric diseases were elevated in PD patients without RBD. Additionally, our prediction that p-cresol sulfate, enriched in the RBD groups with a gut microbial origin, crosses the blood-brain barrier, suggests a potential mechanistic link between gut microbial dysbiosis, iRBD, and PD with RBD. Our study elucidates the heterogeneous nature of PD subtypes.

Abnormal α-synuclein aggregation and the loss of dopaminergic neurons in the substantia nigra contribute to the motor symptoms in Parkinson's disease (PD)[1]. PD patients also develop a diverse spectrum of non-motor symptoms even in the prodromal stages. The different origin and transmission sequence of α-synuclein pathology are supposed to be responsible for heterogeneous manifestation of diverse spectrum of PD symptoms. Indeed, accumulating evidence indicates that clinical and prodromal PD exhibit notable heterogeneity, prompting the identification of proposed subtypes[2].

Idiopathic REM sleep behavior disorder (iRBD), recognized as a prodromal manifestation of α-synucleinopathies, strongly correlates with PD development; but <50% of clinically manifest PD cases exhibit RBD at onset[2]. Recently, based on neuropathological and clinical evidence, two subtypes of PD have been proposed: brain-first (initiating in the brain) and body-first (originating in the enteric nervous system)[2–4]. PD with iRBD is linked to the body-first type, with initial α-synuclein aggregation in the gut affecting medulla and pontine structures before nigral dopamine neurons. In contrast, PD without RBD at onset is associated with the brain-first type, marked by the initial loss of nigrostriatal dopamine innervation but relatively intact sympathetic and parasympathetic innervation[2,4]. However, the robustness of these two subtype PD hypotheses requires further verification, and exploring alternative PD subtypes is critical for improved diagnosis and treatment based on distinct etiologies in this heterogeneous disorder.

A compelling association exists between gastrointestinal symptoms, notably constipation, and the onset of motor symptoms in PD, often manifesting many years prior[5]. Several studies emphasize the intricate link between PD and the gut microbiota[6,7]. A recent cross-sectional cohort study examining microbial signatures in healthy individuals, patients with iRBD, first-degree relatives (FDR) of patients with iRBD, and patients with PD and RBD indicates that PD-like changes in gut microbial signatures commence in iRBD, supporting body-first PD subtypes[8]. Given the potential different etiologies between PD and RBD, and PD without RBD at onset, further investigation is needed into microbial signatures between two subtypes of PD.

The gut microbiome can influence not only gut but also brain through the production of bioactive metabolites, either directly or indirectly[9,10]. Among these microbial metabolites, short-chain fatty acids (SCFAs) have been demonstrated to be capable of promoting microglia activation and motor deficits, potentially through an indirect pathway in germ-free α-synuclein-overexpressing PD mouse models[11]. Extensive metabolomic profiling of blood samples from PD patients has been conducted to identify metabolic biomarkers for better characterization or early-stage diagnosis of

[1]Department of Life Sciences, Gwangju Institute of Science and Technology, Gwangju, South Korea. [2]Department of Life Sciences, Pohang University of Science and Technology, Pohang, South Korea. [3]Department of Neurology, Seoul National University Hospital, College of Medicine, Seoul National University, Seoul, Korea. [4]Department of Pharmacology, Sungkyunkwan University School of Medicine, Suwon, Republic of Korea.
✉e-mail: ara.koh@postech.ac.kr; movement@snu.ac.kr

PD[12–16]. Some studies have provided insights into circulating blood metabolites of gut microbial origin, such as p-cresol sulfate and phenylacetylglutamine[14,16]. However, it is not clear whether PD subjects in these studies were examined for the presence of RBD at onset. To date, there have been no studies on untargeted metabolite profiling of PD subtypes in the early stages of disease.

In this study, we thus aim to dissect metabolic profiles of PD subtypes by performing untargeted metabolomics on neurologically healthy controls, patients with iRBD, PD with RBD, and PD without RBD in the early stages of the disease.

Results

Clinical characteristics and plasma metabolic features of the cohorts in this study

Clinical characteristics of subjects were shown in Table 1. A total of 101 plasma metabolomics samples were generated, comprising 24 PD patients without RBD at onset (PD_Only), 25 PD patients with RBD at onset (PD_RBD+), 25 patients with iRBD, and 27 healthy controls. All clinical groups showed no substantial differences in their demographic features, including age, gender, Hoehn and Yahr (HY) stage, and duration of disease at blood collection. Patients diagnosed with PD were in the early stage of disease, with a mean disease duration of 1.3 ± 1.1 years from symptom onset. In the PD_Only and PD_RBD+ groups, 18 (75%) and 19 (76%) patients, respectively, were not taking anti-Parkinson medications. Four patients in each group were taking levodopa.

Regarding metabolite analysis, a decrease in the total number of detected metabolites was observed in the diseased groups compared to healthy controls: Control vs. PD_Only ($P = 0.101$, mean diff = 8.1, 95% CI: −1.65 to 17.89); Control vs. iRBD ($P = 0.524$, mean diff = 3.2, 95% CI: −6.75 to 13.09); Control vs. PD_RBD+ ($P = 0.0333$, mean diff = 9.7, 95% CI: 0.80 to 18.65). (*P*-values by Student's *t*-tests) (Fig. 1A). To gain insight into the general similarity or dissimilarity in metabolite profiles between groups, a non-metric multi-dimensional scaling (NMDS) plot, an analytical technique to reduce complex, high-dimensional dataset into lower-dimensional spaces, from untargeted metabolomics analysis revealed a distinct shift in the metabolite profile in subjects with PD (its centroid shifted low in MDS2 axis) or RBD (its centroids shifted high in MDS2 axis) when compared to controls (Supplementary Figs. 1, and 1B, C). We examined the loadings of the given ordination plot and summarized its loading at the pathway level (Fig. 1D). We found that the top-left directions were associated with lipids and degraded products, including bile acids and bilirubin products, which are linked to iRBD and PD_RBD+ samples in their ordination plot. The top-right directions, opposite to PD_Only samples, were associated with nucleotides, cofactors/vitamins, energy metabolites, and xenobiotics, indicating a reduction in energy metabolism and xenobiotic levels in PD_Only samples (Fig. 1D).

Table 1 | Demographic summary of PD cohort in this study

	PD_Only	PD_RBD +	iRBD	Control	*P*
n	24	25	25	27	-
Men, *n* (%)	12 (50)	13 (52)	13 (52)	19 (70)	0.99
Age at sampling	62.8 ± 7.3	62.0 ± 5.7	63.5 ± 6.8	62.4 ± 7.1.	0.90
HY stage at sampling	1.2 ± 0.4	1.4 ± 0.5	NA	NA	0.15
PD duration at sampling	1.5 ± 1.4	1.2 ± 0.8	NA	NA	0.70

Data are presented as mean ± s.e.m or as proportions. *P* values were calculated using ANOVA for continuous variables and chi square test for categorical variables.
HY Hoehn and Yahr, *NA* not applicable, *PD* Parkinson's disease.

Plasma metabolic features of PD subtypes focusing on RBD

Considering that previous plasma metabolic profiling of PD patients reflects mixed profiles from PD patients without RBD (PD_Only) and PD patients with RBD (PD_RBD+) and acknowledging that RBD is a pivotal prodromal marker of PD, we have grouped PD_RBD+ and iRBD into the category "RBD group". This grouping is undertaken to more effectively distinguish between plasma metabolic features associated with PD_Only ("brain-first PD") and PD_RBD+ ("body-first PD") subtypes. To identify metabolites that could predict the RBD group and PD_Only subtypes, we conducted metabolome-wide associations with clinical groups using generalized fold changes and Wilcoxon rank-sum tests (Fig. 2A and Supplementary Data S1–S3). Comparing all significantly altered metabolites ($P < 0.05$) in any of the clinical groups (PD_Only, iRBD or PD_RBD+) versus healthy controls, we observed high correlations between iRBD and PD_RBD+ groups (Pearson's correlation coefficients = 0.510, $P < 0.001$, Fig. 2A), indicating similar enrichments of disease metabolites in both groups.

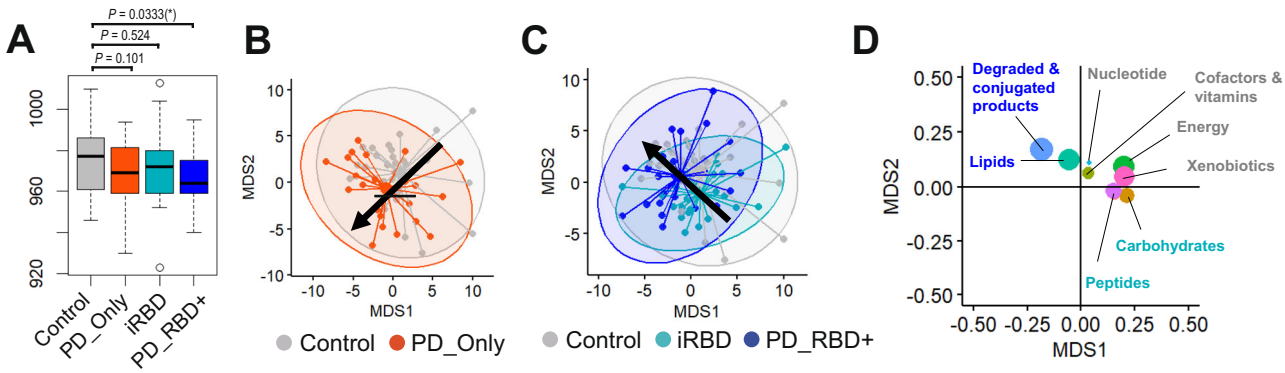

Fig. 1 | Overview of the plasma metabolic features of cohorts. A Number of annotated metabolites in each group based on untargeted metabolomics profiling. The control group ($n = 24$), Parkinson's disease (PD)_Only group ($n = 24$), idiopathic rapid eye movement sleep behavior disorder (iRBD) group ($n = 25$), and PD with RBD (PD_RBD+) group ($n = 25$) are represented in gray, scarlet, turquoise, and blue, respectively. A decreasing trend in the average number of detected metabolites was observed among control, PD_Only, iRBD, and PD_RBD+. * $P < 0.05$. *P*-values were determined by Student's *t*-tests. **B** Non-Metric Multi-dimensional Scaling (NMDS) plot comparing untargeted metabolomics profiles between the PD_Only and control groups. Centroids represent group means, with ellipses indicating group dispersion. PD_Only group exhibits a distinct metabolic shift, with low values both on Multidimensional Scaling (MDS)1 and MDS2 axes. **C** NMDS plot comparing untargeted metabolomics profiles between RBD groups (iRBD and PD_RBD+) and control groups. The PD_RBD+ group shows a distinct shift, characterized by higher values on the MDS2 axis and lower values on the MDS1 axis. **D** NMDS loading plot of metabolites, summarized by their respective metabolic pathways. Data are presented as box plots showing minimum, 25% quartile, median, 75% quartile, and maximum.

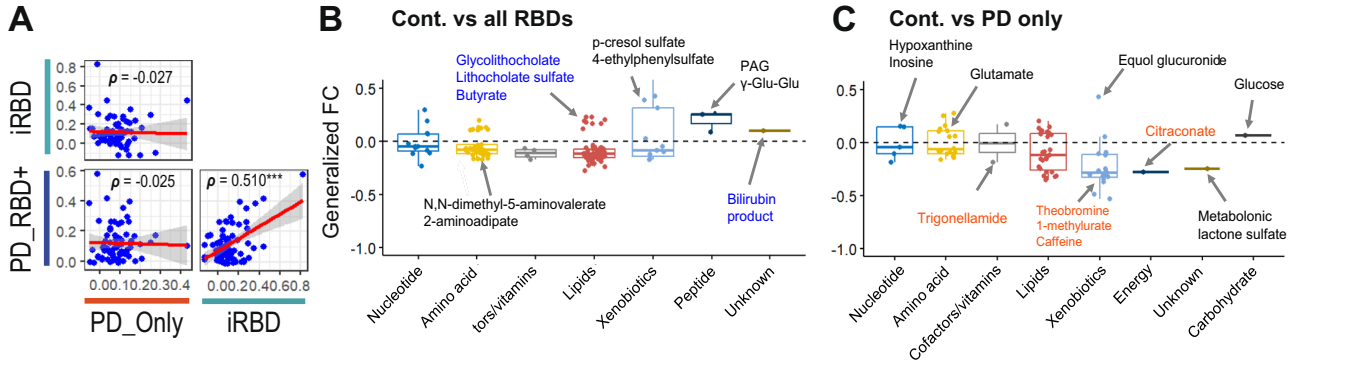

Fig. 2 | Metabolites discriminating patients with RBD and PD without RBD. **A** Comparison of generalized fold changes of significantly altered metabolites ($P < 0.05$) from the PD_Only group, iRBD group, and PD_RBD+ group versus controls (Cont.). Pearson's correlation coefficients (ρ) of generalized fold changes from the three groups versus controls were compared. Based on the correlation coefficients, we identified that metabolite alterations in iRBD group were highly correlated with those in PD_RBD+ group ($\rho = 0.510$; $P = 7.48 \times 10^{-6}$), which was not observed in other group comparisons. *** $P < 0.001$. **B** Generalized fold changes (FC) of metabolites significantly altered ($P < 0.05$) between RBD groups vs control, summarized by pathway classes (See details of statistics and P values in Supplementary Data S4). Metabolites corresponding to degraded/conjugated products and lipids are colored blue. **C** Generalized fold changes (FC) of metabolites significantly altered ($P < 0.05$) between PD_Only vs control, by pathway classes (See details of statistics and P values in Supplementary Data S1). Metabolites corresponding to xenobiotics, cofactors/vitamins, and energy metabolites are colored red.

Further comparisons between all RBD groups versus control (Fig. 2B, Supplementary Fig. 2A, and Supplementary Data S4) and the PD_Only group by generalized fold changes and Wilcoxon rank-sum tests (Fig. 2C and Supplementary Fig. 2B, and Supplementary Data S5) revealed RBD-specific enriched metabolites, including secondary bile acids (lithocholate sulfate, glycolithocholate), p-cresol sulfate, and phenylacetylglutamine (PAG) (Fig. 2B), which are of gut microbial origin[10]. Reanalysis of publicly available whole-genome shotgun sequencing datasets of PD patients[7] confirmed that subsets of PD patients had a higher abundance of *Clostridioides difficile*, a p-cresol producer, as well as increased levels of the enzymes *hpdB* and *hpdC* (which encode 4-hydroxyphenylacetate decarboxylase, responsible for producing p-cresol) in PD subjects compared to healthy controls (Supplementary Fig. 2C–F).

Enriched metabolites in PD_Only included glucose and cortisol, while caffeine levels were decreased (Fig. 2C). We also examined enriched metabolic pathways of significantly altered metabolites in all RBD groups (Supplementary Table 1) and PD_Only group (Supplementary Table 2), identifying benzoate, sterol, histidine and tyrosine metabolites as significantly altered in the RBD group, whereas xanthine metabolites were significantly altered in PD_Only group. This differential enrichment of metabolic pathways underscores the distinct metabolic signatures between RBD groups and PD_Only group.

Enrichment of feces-derived metabolites in the plasma of subjects with RBD

Given that pathologies are believed to originate from the gut in the body-first PD subtype, our investigation focused on identifying the potential origins of metabolites specific to PD_Only or RBD group. For this, we identified metabolites highly enriched in the PD only group, including cortisol, and in the RBD group, including p-cresol sulfate, resulting in the identification of many potential biomarkers to diagnose PD only group ($n = 24$) or RBD group ($n = 29$) (Wilcoxon rank-sum tests, $P < 0.05$; Fig. 3A–C, Supplementary Data S6). After excluding three metabolites enriched in both RBD and PD_Only groups (hypotaurine, hypoxanthine, and glutamate), we identified RBD group-specific metabolites, namely, RBD metabolites (RM) ($n = 26$), and PD_Only group-specific metabolites, namely, PD_Only metabolites (POM) ($n = 21$). We then examined their origins and associated bio-specimens (Fig. 3D, E and Supplementary Data S7, S8) based on the human metabolite resource database, HMDB (https://www.hmdb.ca/)[17]. Interestingly, among RM metabolites with HMDB annotations, 90% were predicted to originate from feces, while 55% were from feces in the POM metabolites (Fig. 3D and Supplementary Data S6). This was further supported by a stronger association of microbial origin with metabolites in the RBD group compared to those in PD_Only samples (Hypergeometric test, p value < 0.05, Fig. 3E). One of the prominent metabolites of fecal and microbial origins, p-cresol sulfate, was elevated in iRBD group and maintained its elevation in PD_RBD+ group but not in the PD_Only group (Fig. 3C), supporting PD_RBD+ as a body-first PD subtype. Moreover, metabolites enriched in the PD_Only group were associated with neuropsychiatric diseases (i.e., schizophrenia); however, those enriched in the RBD group were associated with colon cancer and inflammatory bowel disease, both of which are closely related to the gut microbiome dysbiosis (Fig. 3F, G and Supplementary Data S10, S11). This finding suggests that the RM metabolites, i.e., gut microbiome-associated metabolites, can also be risk factors for other gut-associated diseases, whereas POM metabolites, i.e., non-gut-associated Parkinson' disease metabolites, can be risk factors for other brain disorders. This suggests that not only the spatial progression of pathology but also the underlying pathogenic mechanisms are different between the two subtypes of PD, i.e., body-first PD and brain-first PD.

Prediction of metabolites of microbial origins with ability to penetrate blood-brain barrier

Even though the body-first PD subtype has been associated with pathologies originating in the gut and spreading to the CNS via the vagus nerve[2], other possible routes, such as direct penetration of the CNS by gut microbiome-derived metabolites may also be involved. To explore this possibility, we assess the probability of blood-brain barrier (BBB) permeability for metabolites of microbial origin using LightBBB[18] (Supplementary Data S9). RBD group-enriched p-cresol sulfate (Supplementary Fig. 3) with microbial origin was predicted to penetrate the BBB, consistent with a study showing elevated p-cresol sulfate levels in the cerebrospinal fluid (CSF) of PD patients[19]. Given that CSF represents the closest approximation of extracellular space of the CNS, these results suggest that RBD-enriched p-cresol sulfate as a potential mediator in the progression of PD in the body-first PD subtype.

Establishment and evaluation of prediction models of PD clinical groups

We proceeded to assess the biomarker potential of discriminative metabolites identified in PD clinical groups, specifically POMs and RMs (Supplementary Fig. S4 and Supplementary Data S12–S14), and machine learning models such as logistic regression (Fig. 4). First, we constructed Receiver Operating Characteristics (ROC) curves using

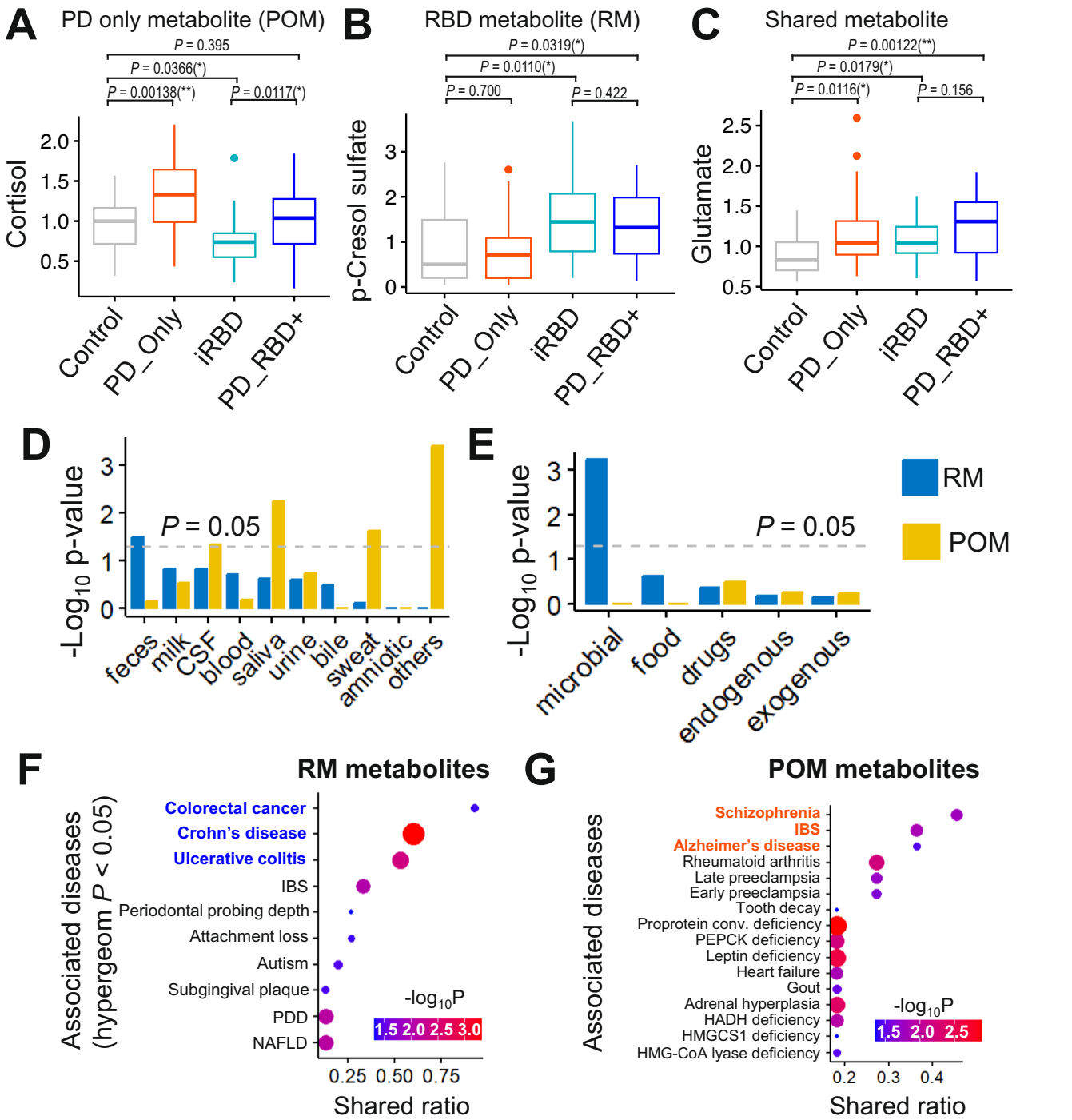

Fig. 3 | Characterizations of significantly enriched metabolites in RBD (iRBD and PD_RBD+) group and PD_Only group. A Relative abundance level of cortisol in different groups - a metabolite significantly enriched in PD_Only group (POM). **B** Relative abundance level of p-cresol level in different groups - a metabolite significantly enriched in RBD group (RM). **C** Relative abundance level of Glutamate level in different groups - a metabolite significantly enriched in both RBD group and PD_Only group. Statistical significances and confidence intervals were calculated based on student's *t*-test between two different clinical groups on this plot. **D** Significantly enriched known biospecimen type based on Human Metabolome Database (HMDB) annotations (Hypergeometric test *P* values < 0.05). RBD metabolites (RM) were significantly enriched among feces metabolites, whereas PD_Only metabolites (POM) were significantly enriched among saliva, sweat, and CSF metabolites (See *P*-values and other details in Supplementary Data S8). **E** Significantly enriched the origins of metabolites based on Human Metabolome Database (HMDB) annotations (Hypergeometric tests $P < 0.05$). RBD metabolites (RM) were significantly enriched among microbial metabolites (See *P*-values and other details in Supplementary Data S7). **F** Significantly enriched disease metabolites of given RBD metabolites (RM) ($P < 0.05$) (See *P* values and other details in Supplementary Data S11). Statistical significance (hypergeometric tests *P* values) and fraction of shared RM metabolites with disease metabolites (shared ratio) were shown by color scales and circle sizes. **G** Significantly enriched disease metabolites of given PD_Only metabolites (POM) ($P < 0.05$) (See *P* values and other details in Supplementary Data S10). Statistical significance (hypergeometric tests *P*-values) and fraction of shared RM metabolites with disease metabolites (shared ratio) were shown by color scales and circle sizes. $^{*}P < 0.05$, $^{**}P < 0.01$.

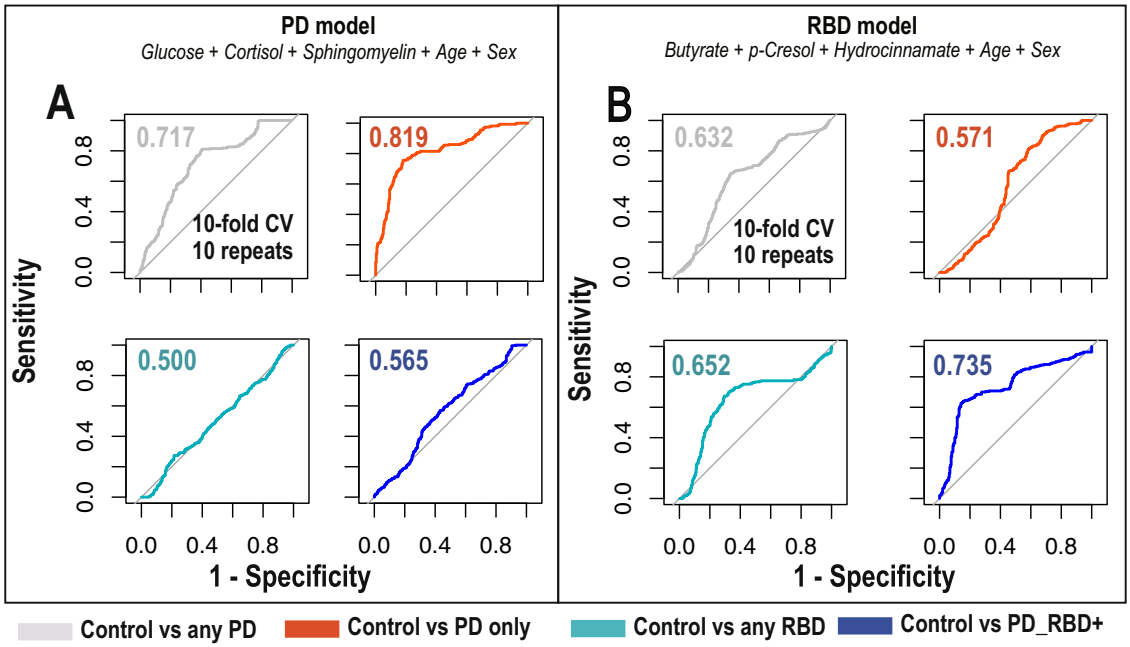

Fig. 4 | Evaluation of logistic regression models using PD_Only metabolites (POM) and RBD metabolites (RM) for the classification of clinical groups. Performance of logistic regression models using most discriminative markers of (**A**) three PD_Only metabolites (PM)—glucose, cortisol, sphingomyelin to classify between two different clinical groups as PD model and (**B**) three RBD metabolites (RM)—butyrate, p-cresol sulfate, and hydrocinnamate to classify between two different clinical groups as RBD model. In the model, age and sex were considered as covariates. Color coding for group comparisons: gray - control vs. any PD cases; scarlet - control vs. PD_Only; cyan - control vs. any RBD case; blue - control vs. PD_RBD + .

POM and RM metabolites to classify clinical groups versus healthy controls (Supplementary Fig. S4). Of note, among POM and RM metabolites, we excluded drug compounds, their derivatives, or xenobiotics compounds from the potential markers of classification models due to their exogenous nature. Based on the highest generalized fold changes, we identified three markers from POM metabolites—glucose, cortisol, and sphingomyelin (d18:1/20:2, d18:2/20:1, d16:1/22:2)—showing high Area Under the ROC curve (AUROC) values (>0.725) to classify the PD_Only group from healthy controls (Supplementary Fig. 4A, B, C). In addition, three markers from RM metabolites–butyrate/iso-butyrate, p-cresol sulfate, and hydrocinnamate—showed high AUROC values (>0.706) to classify iRBD/PD_RBD+ groups from healthy controls (Supplementary Fig. 4D, E, F). Furthermore, we included known markers from other metabolomics studies of PD, which lacked clinical clustering by RBD manifestation, such as indoleacetate, PAG, and free fatty acid (10:0). Although they showed high AUROC values to classify the PD_Only group from control or iRBD/PD_RBD+ group from control, their values were still lower than those from POM and RM metabolites (Supplementary Fig. 4G, H, I), underscoring the superior marker potential of the three metabolites we identified. Next, we established logistic regression models using PD markers—glucose, cortisol, and sphingomyelin (Fig. 4A–D), as well as RBD markers—butyrate, p-cresol sulfate, and hydrocinnamate (Fig. 4E–H). Interestingly, based on 10-fold cross-validation with 10 repetitions, the average AUROC values were substantially higher to classify the PD_Only group from control and RBD groups from control. Moreover, we assessed the prediction accuracy of the logistic regression models established (see confusion matrix to check prediction imbalances per class in Supplementary Fig. 5A—C) and found that the predicted class labels closely matched with the true class labels without biases in predicting clinical groups and controls. Consequently, our models achieved high accuracy for the diagnosis of clinical groups—80.4% accuracy to predict the PD_Only group, 69.2% accuracy to predict the PD_RBD+ group, and 74% accuracy to predict any RBD group from the controls.

Discussion

In this study, we investigated the plasma metabolic profiles of PD subtypes, specifically focusing on RBD as a prodromal marker (Fig. 5). Untargeted metabolomics analysis of plasma samples from the RBD group (comprising iRBD and PD with RBD) revealed distinct metabolic features compared to the PD_Only group (PD without RBD), supporting the notion of differential pathogenesis between these subtypes.

Our results also showed that plasma metabolites in the RBD group were enriched with gut microbial origins, aligning with the 'brain-first vs body-first' hypothesis[4,20]. One key finding was the differential presence of p-cresol sulfate across PD subtypes. Elevated levels of p-cresol sulfate in PD have been consistently reported across different cohorts[14,16]. However, in our study, we observed this metabolite only in the PD_RBD+ group, not in the PD_Only group. This discrepancy may be due to heterogeneous PD populations in previous studies. Supporting this, we confirmed that subsets of PD patients had a higher abundance of p-cresol producing enzymes and bacterium.

When aggregating both PD subtypes into PD, our findings aligned with previous metabolomics studies, showing altered levels of metabolites, including increased microbial metabolites (i.e., secondary bile acids, p-cresol sulfate, and PAG), cortisol, fatty acids, branched-chain amino acids, as well as caffeine in peripheral blood[21–26]. However, inconsistencies among studies[23,27,28] underscored the heterogeneous nature of PD pathogenesis, highlighting the need for more comprehensive subtyping. While differential pathogenesis between PD with RBD and without RBD is recognized, our study is among the first to consider the presence or absence of RBD at PD onset, which may partially explain the observed discrepancies between studies.

By segregating PD patients based on RBD, we identified both shared and distinct metabolic pathways. For instance, lysophospholipid metabolism and fatty acid metabolism were common across PD subtypes, while certain lysophospholipid species were uniquely enriched in PD patients with or without RBD (Supplementary Tables 1 and 2). Lysophospholipids, the smallest bioactive lipids, play diverse biological roles in the brain (i.e., synapse formation, synaptic transmission, and neuron survival[29]. Notably,

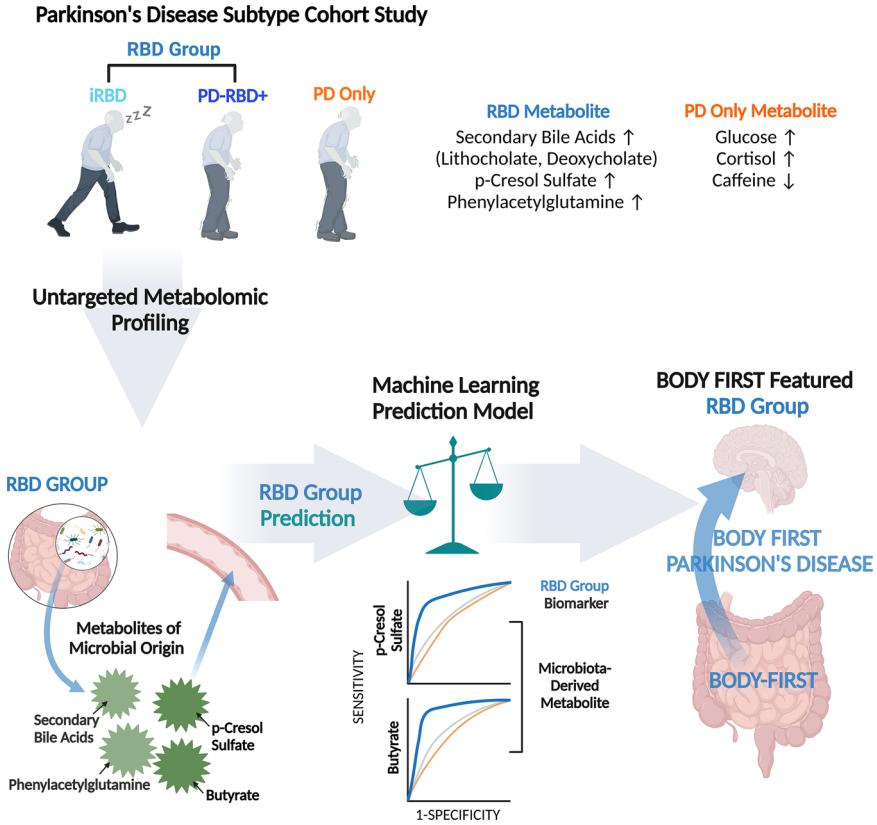

Fig. 5 | Schematic illustration of plasma metabolic profiles in PD subtypes based on RBD. This figure illustrates the plasma metabolic profiles observed in PD subtypes, categorized by the presence or absence of RBD, utilizing an untargeted metabolomics approach. Across PD cases with and without RBD, elevated levels of secondary bile acids, p-cresol sulfate, phenylacetylglutamine, and cortisol, alongside decreased caffeine, were identified, consistent with previously reported PD-associated metabolites. However, upon closer examination focusing on the presence of RBD, a significant divergence was noted. Specifically, secondary bile acids, p-cresol sulfate, and phenylacetylglutamine—metabolites originating from gut microbial activity—were distinctly evident in PD cases with RBD but absent in those without RBD. This finding suggests a potential association between PD subtypes characterized by RBD as a "body-first" type PD. Figure 5 was created using www.biorender.com.

alterations in lysophosphatidylcholines have been observed in the midbrain of 6-hydroxydopamine (6-OHDA)-induced rat model of PD[30]. Thus, the reduction in lysophospholipid species might be associated with mitochondrial impairment and neuronal degeneration in both subtypes of PD.

In the PD_Only group, we observed a depletion of metabolites associated with xanthine and purine metabolism, including caffeine, inosine, and urate (Supplementary Table 2), all of which have been previously associated with reduced PD risk[31–36]. However, recent phase 2 and phase 3 clinical trials of inosine in early PD (clinical trials.gov identifier # NCT00833690 and NCT02642393) did not yield significant differences in the rate of clinical disease progression. This underscores the importance of stratifying PD patients for this intervention, especially considering that inosine and urate levels were not affected in PD with RBD but only in the PD_Only group in our study. Additionally, reductions in caffeine metabolites, including caffeine, theobromine, and 1-methylurate, were consistent features of PD, aligning with previous studies[14,37–39] and reinforcing the notion that caffeine metabolism may be disrupted in PD patients.

Metabolic pathways overrepresented in RBD groups include the long-chain polyunsaturated fatty acid metabolic pathway, which is potentially associated with REM sleep[40–42]. Thus, it is feasible to speculate that the general repression of metabolites in the long-chain polyunsaturated fatty acid metabolic pathway in RBD might reflect dysfunctional biochemical production of sleep modulators due to neurodegeneration induced by α-synuclein pathology infiltration to the lower brainstem. Additionally, benzoate metabolism was overrepresented, with increased microbial metabolites like hydrocinnamate, 4-ethylphenylsulfate, and p-cresol sulfate, all of which have microbial origins. While some of these metabolites may be neuroprotective[43,44], others, such as 4-ethylphenylsulfate and p-cresol sulfate, have been associated with neurodegeneration and behavior disorders[14,16,45–48]. In our study, we, for the first time, showed that p-cresol sulfate is a feature of iRBD and PD with RBD but not in PD without RBD at onset. However, there might be a possibility of elevation of p-cresol sulfate in patients with progressed PD without RBD, which needs to be investigated in the future.

In summary, our study illuminates the metabolic differences between PD subtypes, particularly distinguishing between those with and without RBD at onset. Using untargeted metabolomics, we identified unique metabolic patterns associated with RBD, supporting the body-first hypothesis of PD pathogenesis and suggesting plasma metabolites, particularly of gut microbial origin, as potential biomarkers for subtype detection. Moreover, the resemblance in metabolic profiles between individuals with iRBD and those with RBD in PD hints at the potential utility of metabolic biomarkers in the early diagnosis of prodromal PD. This is significant as it presents an opportunity for the early identification of PD risk among individuals experiencing sleep disturbances, potentially enabling more effective treatment and prevention of ongoing midbrain dopaminergic neurodegeneration. While providing valuable insights, limitations such as sample size and focus on plasma metabolites require validation in larger cohorts and fecal microbiome samples. Furthermore, in confirming RBD, polysomnography was performed only in the iRBD group, whereas in the PD_RBD group, RBD was assessed using the RBD1Q questionnaire. Nonetheless, our findings emphasize the need for deep phenotyping for better PD diagnosis and potentially pave the way for tailored personalized treatment based on accurate subtype diagnosis.

Methods

Subjects

The subjects in this study were recruited from the Movement Disorders Clinic at Seoul National University Hospital (SNUH), comprising a total of 101 subjects (Table 1). Among them, 24 had PD without RBD, 25 had PD with RBD, 25 had iRBD without movement disorders or dementia, and 27 were controls (categorized as PD_Only, PD_RBD+, iRBD, and Control group, respectively). Patients with PD were followed for an average of 4.0 years after sampling to ensure the reliability of the diagnosis. iRBD was diagnosed using video-polysomnography in accordance with the third edition of the International Classification of Sleep Disorders. The diagnosis of PD was established following the clinical diagnostic criteria for PD by the

Movement Disorders Society[49]. PD subjects were further categorized based on the presence or absence of RBD at disease onset, resulting in two groups assessed using RBD1Q[50]: the PD_Only group and the PD_RBD+ group. The control group, free of PD or iRBD, was drawn from the SNUH biorepository, predominantly consisting of spouses sharing living environments with PD or iRBD patients. The study protocol was approved by the Institutional Review Board of the Seoul National University Hospital (IRB No. 2207-085-1340), and informed consent was waived by the board.

Plasma collection and metabolite analysis

Blood samples were collected from the participants' venous blood during visits. The samples underwent centrifugation, and the resulting plasma was stored at −80 °C until analysis. Global metabolite profiling analysis on the plasma samples was conducted by Metabolon (Durham, NC) using ultra-high-performance liquid chromatography coupled to tandem mass spectrometry (UPLC-MS/MS). In summary, automated sample preparation was carried out using the MicroLab STAR® system (Hamilton Company). A recovery standard was introduced before the initial step in the extraction process for quality control. Proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000), followed by centrifugation. The resulting extract was divided into five fractions: one for analysis by UPLC-MS/MS with positive ion mode electrospray ionization, one for analysis by UPLC-MS/MS with negative ion mode electrospray ionization, one for LC polar platform, one for analysis by GC-MS, and one sample was reserved for backup. Samples underwent a brief period on a TurboVap® (Zymark) to eliminate the organic solvent. For LC, the samples were stored overnight under nitrogen before preparation for analysis. For GC, each sample was vacuum-dried overnight before preparation for analysis.

Statistical analysis of plasma metabolomics

Using untargeted metabolomics profiles after zero-imputations for missing values, number of detected metabolites was checked and compared by groups, such as control, iRBD, PD_Only, and PD_RBD+, by Wilcoxon rank sum tests. Similarity of metabolomics samples was compared by non-metric multi-dimensional scaling (NMDS) by metaMDS function of R vegan package (version 2.6-4), and their associated loadings were also calculated by the envfit function of vegan library. Differentially abundant metabolite testing was performed by Wilcoxon rank sum tests, together with generalized fold change (R MicrobiotaProcess Package, version 1.11.5)[51,52], an estimator of log-2-fold change with capturing the variation of posterior distribution of log-2-fold change. Visualizations of boxplots, scatter plots, and bar plots were performed by R ggpubr package (version 0.6.0). Metabolites of given metabolomics profiles were annotated with Human Metabolome Database (HMDB)[17], and origins, associated biospecimen, and disease metabolite information were obtained from HMDB (https://hmdb.ca/) (as of 6th Feb 2024). In short, from all metabolites deposited in HMDB, metabolites annotated with their biospecimen (blood, urine, saliva, cerebrospinal fluid (CSF), feces, sweat, breast milk, bile, amniotic fluid, and other biospecimens), origin (exogenous, endogenous, food, microbial, and drug), cellular location (cytoplasm, nucleus, and mitochondria), and diseases associated with given metabolites were downloaded and compared with metabolites detected in this study (See biospecimen, origin, cellular location, and associated disease information of metabolites detected in this study on Supplementary Data S15–S18). Based on the annotated metabolite information, overlaps of significantly differential metabolites with above-mentioned metabolite classes (origins of metabolites, associated biospecimen, and associated diseases) were checked with hypergeometric tests to identify the significance of overlaps. For POM metabolites and RM metabolites, we checked BBB permeability by lightBBB prediction program[18]. AUROC based on the sensitivity and specificity was calculated with R pROC package (version 1.18.4). Logistic regression models using key metabolites, sex, and age were generated using glm function (family = binomial) of R stats package (version 4.2.1) from randomly assigned training and test datasets (ratio of data 2:1, respectively)

Clostridioides difficile abundance and *hpdB/C* abundance analyses

Publicly available metagenomic data of both Parkinson's disease patients ($n = 491$) and neurologically healthy elderly controls ($n = 234$) were obtained from BioProject ID PRJNA834801[7]. To ensure data quality, raw sequence reads were filtered using BBDuk and BBSplit. The Nextera XT adapter sequences and PhiX genome contamination were removed with BBDuk using the following parameters: "ftm = 5 tbo tpe qtrim = rl trimq = 25 minlen = 50 ref = adapters,phix". Subsequent reads were aligned to the human reference genome (GRCh38.p14) using BBSplit with default parameters to eliminate human contamination. First, to estimate the abundance of the authentic p-cresol producer *Clostridioides difficile*, quality-controlled reads underwent taxonomic profiling using MetaPhlAn 4 (v4.0.6, mpa_vOct22)[53]. Second, to estimate the abundance of the p-cresol-producing enzyme, quality-controlled reads were analyzed for species-level functional content (UniRef90 gene families) using HUMAnN3[54]. The reads per kilobase (RPK) abundances of gene families, normalized for gene length, were generated using default settings. To reduce the number of gene families being analyzed, the resulting counts based on UniRef90 (v201901b) and the ChocoPHLAn 3 database were converted into Kyoto Encyclopedia of Genes and Genomes (KEGG) orthologs using the HUMAnN utility tool humann_regroup_table. To quantify the abundance of genes encoding the p-cresol-producing enzyme hpdB/C, the specific gene identifiers K18427 (4-hydroxyphenylacetate decarboxylase large subunit) and K18428 (4-hydroxyphenylacetate decarboxylase small subunit) were targeted. The relative abundance of *C. difficile* and the RPK abundance of *hpdB/C* from 725 biologically independent samples (491 PD and 234 neurologically healthy controls) were visualized using the R ggplot2 package (version 3.5.1). Pearson's correlation between *hpdB* and *hpdC* counts and the relative abundance of *C. difficile* was conducted using the R ggpubr package (version 0.6.0).

Data availability

The datasets supporting the conclusions of this article are included within the article and its additional files. Metabolomics profile information was provided in Supplementary Data S19–S20.

Received: 5 August 2024; Accepted: 4 June 2025;
Published online: 01 July 2025

References

1. Morris, H. R., Spillantini, M. G., Sue, C. M. & Williams-Gray, C. H. The pathogenesis of Parkinson's disease. *Lancet* **403**, 293–304 (2024).
2. Berg, D. et al. Prodromal Parkinson disease subtypes - key to understanding heterogeneity. *Nat. Rev. Neurol.* **17**, 349–361 (2021).
3. Braak, H. et al. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* **24**, 197–211 (2003).
4. Horsager, J. et al. Brain-first versus body-first Parkinson's disease: a multimodal imaging case-control study. *Brain* **143**, 3077–3088 (2020).
5. Cersosimo, M. G. & Benarroch, E. E. Pathological correlates of gastrointestinal dysfunction in Parkinson's disease. *Neurobiol. Dis.* **46**, 559–564 (2012).
6. Romano, S. et al. Meta-analysis of the Parkinson's disease gut microbiome suggests alterations linked to intestinal inflammation. *Npj Parkinsons Dis.* **7**, https://doi.org/10.1038/s41531-021-00156-z (2021).
7. Wallen, Z. D. et al. Metagenomics of Parkinson's disease implicates the gut microbiome in multiple disease mechanisms. *Nat. Commun.* **13**, 6958 (2022).
8. Huang, B. et al. Gut microbiome dysbiosis across early Parkinson's disease, REM sleep behavior disorder and their first-degree relatives.

Nat. Commun. **14**, https://doi.org/10.1038/s41467-023-38248-4 (2023).
9. Cryan, J. F. & Dinan, T. G. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* **13**, 701–712 (2012).
10. Koh, A. & Backhed, F. From association to causality: the role of the gut microbiota and its functional products on host metabolism. *Mol. Cell* **78**, 584–596 (2020).
11. Sampson, T. R. et al. Gut microbiota regulate motor deficits and neuroinflammation in a model of Parkinson's disease. *Cell* **167**, 1469 (2016).
12. Roede, J. R. et al. Serum metabolomics of slow vs. rapid motor progression Parkinson's disease: a pilot study. *PLoS ONE* **8**, https://doi.org/10.1371/journal.pone.0077629 (2013).
13. Troisi, J. et al. Metabolomics in Parkinson's disease. *Adv. Clin. Chem.***104**, 107–149 (2021).
14. Shao, Y. P. et al. Comprehensive metabolic profiling of Parkinson's disease by liquid chromatography-mass spectrometry. *Mol. Neurodegener.* **16**, https://doi.org/10.1186/s13024-021-00425-8 (2021).
15. Shao, Y. P. & Le, W. D. Recent advances and perspectives of metabolomics-based investigations in Parkinson's disease. *Mol. Neurodegener.* **14**, https://doi.org/10.1186/s13024-018-0304-2 (2019).
16. Paul, K. C. et al. Untargeted serum metabolomics reveals novel metabolite associations and disruptions in amino acid and lipid metabolism in Parkinson's disease. *Mol. Neurodegener.* **18**, https://doi.org/10.1186/s13024-023-00694-5 (2023).
17. Wishart, D. S. et al. HMDB 5.0: the human metabolome database for 2022. *Nucleic Acids Res.* **50**, D622–D631 (2022).
18. Shaker, B. et al. LightBBB: computational prediction model of blood-brain-barrier penetration based on LightGBM. *Bioinformatics* **37**, 1135–1139 (2021).
19. Sankowski, B. et al. Higher cerebrospinal fluid to plasma ratio of p-cresol sulfate and indoxyl sulfate in patients with Parkinson's disease. *Clin. Chim. Acta* **501**, 165–173 (2020).
20. Borghammer, P. et al. Neuropathological evidence of body-first vs. brain-first Lewy body disease. *Neurobiol. Dis.* **161**, 105557 (2021).
21. Paul, K. C. et al. Untargeted serum metabolomics reveals novel metabolite associations and disruptions in amino acid and lipid metabolism in Parkinson's disease. *Mol. Neurodegener.* **18**, 100 (2023).
22. Breen, D. P. et al. Sleep and circadian rhythm regulation in early Parkinson disease. *JAMA Neurol.* **71**, 589–595 (2014).
23. Chen, S. J. et al. Alteration of gut microbial metabolites in the systemic circulation of patients with parkinson's disease. *J. Parkinson's Dis.* **12**, 1219–1230 (2022).
24. Fujimaki, M. et al. Serum caffeine and metabolites are reliable biomarkers of early Parkinson disease. *Neurology* **90**, e404–e411 (2018).
25. Nagesh Babu, G. et al. Serum metabolomics study in a group of Parkinson's disease patients from northern India. *Clin. Chim. Acta* **480**, 214–219 (2018).
26. Shao, Y. et al. Comprehensive metabolic profiling of Parkinson's disease by liquid chromatography-mass spectrometry. *Mol. Neurodegener.* **16**, 4 (2021).
27. Haglin, L. & Backman, L. Covariation between plasma phosphate and daytime cortisol in early Parkinson's disease. *Brain Behav.* **6**, e00556 (2016).
28. Shin, C., Lim, Y., Lim, H. & Ahn, T. B. plasma short-chain fatty acids in patients with Parkinson's disease. *Mov. Disord.* **35**, 1021–1027 (2020).
29. Fukushima, N., Weiner, J. A. & Chun, J. Lysophosphatidic acid (LPA) is a novel extracellular regulator of cortical neuroblast morphology. *Dev. Biol.* **228**, 6–18 (2000).
30. Farmer, K., Smith, C. A., Hayley, S. & Smith, J. Major Alterations of phosphatidylcholine and lysophosphotidylcholine lipids in the substantia nigra using an early stage model of Parkinson's disease. *Int. J. Mol. Sci.* **16**, 18865–18877 (2015).
31. Zhao, Y. et al. Association of coffee consumption and prediagnostic caffeine metabolites with incident parkinson disease in a population-based cohort. *Neurology* **102**, e209201 (2024).
32. LeWitt, P. A. et al. Metabolomic biomarkers as strong correlates of Parkinson disease progression. *Neurology* **88**, 862–869 (2017).
33. Annanmaki, T., Muuronen, A. & Murros, K. Low plasma uric acid level in Parkinson's disease. *Mov. Disord.* **22**, 1133–1137 (2007).
34. Andreadou, E. et al. Serum uric acid levels in patients with Parkinson's disease: their relationship to treatment and disease duration. *Clin. Neurol. Neurosurg.* **111**, 724–728 (2009).
35. Davis, J. W. et al. Observations on serum uric acid levels and the risk of idiopathic Parkinson's disease. *Am. J. Epidemiol.* **144**, 480–484 (1996).
36. Weisskopf, M. G., O'Reilly, E., Chen, H., Schwarzschild, M. A. & Ascherio, A. Plasma urate and risk of Parkinson's disease. *Am. J. Epidemiol.* **166**, 561–567 (2007).
37. Hatano, T., Saiki, S., Okuzumi, A., Mohney, R. P. & Hattori, N. Identification of novel biomarkers for Parkinson's disease by metabolomic technologies. *J. Neurol. Neurosurg. Psychiatry* **87**, 295–301 (2016).
38. Fujimaki, M. et al. Serum caffeine and metabolites are reliable biomarkers of early Parkinson disease. *Neurology* **90**, E404 (2018).
39. Takeshige-Amano, H. et al. Shared metabolic profile of caffeine in parkinsonian disorders. *Mov. Disord.* **35**, 1438–1447 (2020).
40. Dai, Y. & Liu, J. Omega-3 long-chain polyunsaturated fatty acid and sleep: a systematic review and meta-analysis of randomized controlled trials and longitudinal studies. *Nutr. Rev.* **79**, 847–868 (2021).
41. Irmisch, G., Schlafke, D., Gierow, W., Herpertz, S. & Richter, J. Fatty acids and sleep in depressed inpatients. *Prostaglandins Leukot. Ess. Fat. Acids* **76**, 1–7 (2007).
42. Avallone, R., Vitale, G. & Bertolotti, M. Omega-3 fatty acids and neurodegenerative diseases: new evidence in clinical trials. *Int. J. Mol. Sci.* **20**, https://doi.org/10.3390/ijms20174256 (2019).
43. Wilmanski, T. et al. Blood metabolome predicts gut microbiome α-diversity in humans. *Nat. Biotechnol.* **37**, 1217 (2019).
44. Prorok, T., Jana, M., Patel, D. & Pahan, K. Cinnamic acid protects the nigrostriatum in a mouse model of Parkinson's disease via peroxisome proliferator-activated receptoralpha. *Neurochem. Res.* **44**, 751–762 (2019).
45. Hsiao, E. Y. et al. Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* **155**, 1451–1463 (2013).
46. Needham, B. D. et al. A gut-derived metabolite alters brain activity and anxiety behaviour in mice. *Nature* **602**, 647 (2022).
47. Needham, B. D. et al. Plasma and fecal metabolite profiles in autism spectrum disorder. *Biol. Psychiatry* **89**, 451–462 (2021).
48. Sun, C. Y. et al. p-cresol sulfate caused behavior disorders and neurodegeneration in mice with unilateral nephrectomy involving oxidative stress and neuroinflammation. *Int. J. Mol. Sci.* **21**, https://doi.org/10.3390/ijms21186687 (2020).
49. Postuma, R. B. et al. Validation of the MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord.* **33**, 1601–1608 (2018).
50. Postuma, R. B. et al. A single-question screen for rapid eye movement sleep behavior disorder: a multicenter validation study. *Mov. Disord.* **27**, 913–916 (2012).
51. Feng, J. et al. GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics* **28**, 2782–2788 (2012).
52. Xu, S. et al. MicrobiotaProcess: a comprehensive R package for deep mining microbiome. *Innovation* **4**, 100388 (2023).

53. Blanco-Míguez, A. et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat. Biotechnol.* **41**, 1633–1644 (2023).
54. Beghini, F. et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* **10**, e65088 (2021).

Acknowledgements

The work is supported by the SNUH Research Fund (No. 0320220090 for H.J.K) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2022R1A2C2091254 for H.J.K). This work is also supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2023R1A2C1002876 for A.K). SL was also supported by grants of the Basic Science Research Program (2021R1C1C1006336) and the Bio & Medical Technology Development Program (2021M3A9G8022959 and RS-2024-00419699) of the Ministry of Science, ICT through the National Research Foundation; and by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare (HR22C141105), South Korea; and by a "National Institute of Health" research project **(**project no. 2024- ER2108-00 and 2024-ER0608-00**);** and also by a GIST Research Institute (GRI) GIST-MIT research Collaboration grant by the GIST in 2024. This work was also supported by the Korea Bio Data Station (K-BDS) with computing resources including technical support.

Author contributions

Concept and design: S.L., Y.L., A.K. and H.J.K. Acquisition, analysis, or interpretation of data: S.L., J.K., J.W.B., K.Y.J., Y.L., A.K. and H.J.K. Drafting of the manuscript: S.L., J.K., J.W.B., Y.L., A.K. and H.J.K. Critical revision of the manuscript for important intellectual content: JYK. Statistical analysis: S.L. and J.W.B. Obtained funding: S.L., A.K. and H.J.K. Administrative, technical, or material support: S.L., Y.L., A.K. and H.J.K. Supervision: A.K. and H.J.K.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41531-025-01040-w.

Correspondence and requests for materials should be addressed to Ara Koh or Han-Joon Kim.