

RESEARCH

Open Access



Integrated multi-omics highlights alterations of gut microbiome functions in prodromal and idiopathic Parkinson's disease

Rémy Villette^{1†}, Júlia Ortís Sunyer^{1†}, Polina V. Novikova^{1†}, Velma T. E. Aho^{1†}, Viacheslav A. Petrov¹, Oskar Hickl¹, Susheel Bhanu Busi^{1,2}, Charlotte De Rudder¹, Benoit J. Kunath¹, Anna Heintz-Buschart^{1,3}, Jean-Pierre Trezzi¹, Rashi Halder¹, Christian Jäger¹, Laura A. Lebrun¹, Annegrät Daujeumont¹, Sebastian Schade⁵, Annette Janzen⁶, Nico Jehmlich⁷, Martin von Bergen⁷, Cédric C. Laczny¹, Patrick May¹, Claudia Trenkwalder⁸, Wolfgang Oertel⁶, Brit Mollenhauer^{4,5} and Paul Wilmes^{1,9*}

Abstract

Background Parkinson's disease (PD) is associated with gut microbiome shifts. These shifts are mainly described at taxonomic level, but the functional consequences remain unclear. To obtain insight into the functional disruptions of the gut microbiome in PD, we used an integrated multi-omics approach, comparing gut microbiomes of individuals with PD, prodromal PD, and healthy controls.

Results Meta-metabolomics, the most discriminatory and robust omics level, was selected to Guide the analysis. We identified 11 metabolites that were differentially abundant between the groups, among which β -glutamate was increased in PD and prodromal PD, and correlated with the transcriptional activities of *Methanobrevibacter smithii* and *Clostridium* spp. We identified decreases in transcripts, but not in gene abundances, related to glutamate metabolism, bile acids biosynthesis, chemotaxis, and flagellar assembly in PD, particularly in keystone genera such as *Roseburia*, *Agathobacter*, and *Blautia*. Our findings, integrated into the Expobiome map, reveal multifactorial microbiome alterations which converge with PD pathways.

Conclusion Our study highlights the apparent disruption of microbial gene expression in PD, particularly in genes associated to mobility. Moreover, we showcase the importance of investigating the gut microbiome's functional dimensions to better resolve microbiome-host interactions in health and disease.

Keywords Gut microbiome, Gene expression, Multi-omics, Flagella, Parkinson's disease, Bile acids

[†]Rémy Villette, Júlia Ortís Sunyer, Polina V. Novikova and Velma T. E. Aho contributed equally to this work.

*Correspondence:

Paul Wilmes

paul.wilmes@uni.lu

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

Parkinson's disease (PD), a neurodegenerative disease impacting movement due to dopaminergic neuron loss, is the second most prevalent neurodegenerative disease worldwide [1]. Individuals with PD are often characterized by an increase in gut permeability, inflammation, and constipation which, together, suggest a link between the gut microbiome and PD etiology [2–4]. This potential link is supported by numerous studies reporting differences in the gut microbiome structure of individuals with PD compared to healthy individuals [5–11]. These findings have been further confirmed by recent meta-analyses [12, 13]. Together, the studies highlight a decreased abundance for the genera *Roseburia*, *Blautia*, *Butyrivococcus*, and *Faecalibacterium* in PD while *Methanobrevibacter*, *Akkermansia*, *Lactobacillus*, *Bifidobacterium*, and *Hungatella* are typically enriched [5–13]. Similar changes in idiopathic REM sleep behavior disorder (iRBD), a prodromal stage of PD [14], have been reported [7, 10]. Moreover, the taxa decreased in PD are known producers of short-chain fatty acids (SCFAs), which correspondingly have also been found to be decreased in concentration in fecal samples of PD individuals [6, 15, 16].

In addition to SCFAs, several microbiome-derived metabolites such as bile acids (BAs), glycine, and glutamate have been associated with PD, either in plasma, serum [17, 18], or stool [19–21]. BAs are produced by the host and metabolized by the gut microbiome into secondary bile acids with different cytotoxic capacities but also immunomodulatory capacities [22, 23]. Glutamate is the major excitatory neurotransmitter and exerts toxic activity on neuronal cells [24]. Its levels in serum and cerebrospinal fluid have been reported as either increased [25, 26], not different [27], or decreased [28], but decreased in the gut in PD compared to healthy controls (HC) [29].

Altogether, alterations of the gut microbiome are linked to PD, but less is known about iRBD or other prodromal stages of the disease. Moreover, most of the associations between PD and the gut microbiome are based on taxonomic and metabolomic analyses. The resulting data, although insightful, lacks functional and systemic information that could better capture the complex cross-talk between the gut microbiome and the host in the context of PD. To obtain such information, we performed an integrated multi-omics study on a cross-sectional cohort comprised of individuals with iRBD and PD alongside HC. Metagenomics (MG), metatranscriptomics (MT), metaproteomics (MP), and meta-metabolomics (MM) were used to characterize taxonomic (taxMG, taxMT, taxMP) and functional (funMG, funMT, funMP, MM) differences between HC, iRBD, and PD gut microbiomes.

We identified substantial differences in gut microbiome functions and metabolites between the groups, including an increase in β -glutamate levels in PD that were related to a dysregulation of glutamate-related gene expression. Alterations in glutamate-related genes were linked with chemotaxis and flagellar assembly pathways, for which we identified strong and distinct taxonomic differences in transcription between PD and HC. Collectively, our data highlight the importance of multi-omics approaches for the identification of microbiome-mediated effects on neurological, and more broadly, complex human diseases involving host-microbiome interactions.

Results

Study cohort and integrated multi-omics pipeline

Our initial set of subjects consisted of 50 individuals with PD, diagnosed according to United Kingdom Parkinson's Disease Society Brain Bank (UKPDSBB) clinical diagnostic criteria [30], and 30 people with polysomnography-confirmed iRBD as well as 50 healthy control (HC) subjects. The data from 4 PD and 3 iRBD as well as 1 HC were subsequently excluded (see the "Methods" section), leading to a final data set of 46 PD, 27 iRBD, and 49 HCs (Fig. 1). The subjects in the three groups were of similar age but had slightly different gender distributions, with males overrepresented in the iRBD and PD groups (Table 1, Fisher test, $p=0.004$), as is typical for these conditions [31, 32]. Constipation, a prevalent non-motor symptom of PD [33], was also more common in the iRBD and PD groups compared to HC (Fisher test, $p<0.001$).

To allow integrated and paired sample analysis, we used our previously published Integrated Multi-omics Pipeline for data integration and analysis (Fig. 1) [34]. In short, metagenomic and metatranscriptomic reads were co-assembled in a first round, followed by a second round during which remaining reads that did not co-assemble were retained and mapped against the co-assembled contigs from the first round. Contigs were then used to call genes and perform functional annotations using multiple reference databases using the Mantis tool [35]. Predicted proteins from the hybrid assembled contigs were used to create theoretical mass spectra and matched against actually measured metaproteomic mass spectra. We chose to use the KEGG database to integrate metabolic pathways and connect the meta-metabolomic data with the other omics datasets.

Microbiome function is altered in PD

Alpha diversity comparisons revealed no statistically significant differences between the three groups when considering taxMG, taxMT, funMG, taxMP, and MM (Fig. 2A, Mann–Whitney test, $p>0.05$). However, funMT and funMP showed a statistically significant

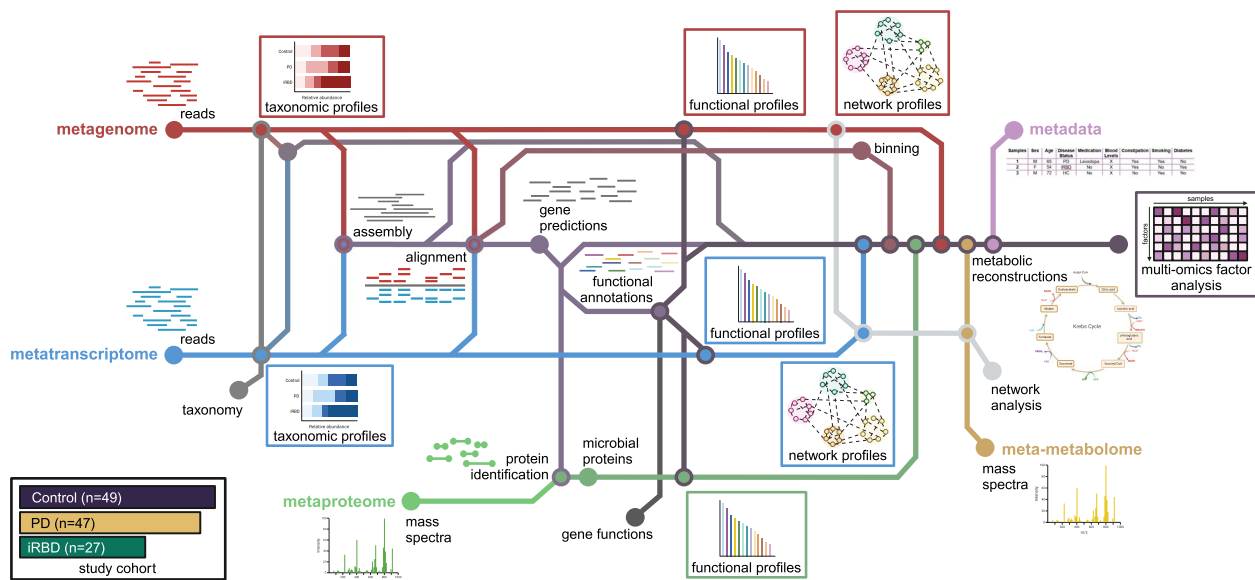


Fig. 1 Schematic representation of the analytical workflow. Metagenomic (MG), metatranscriptomic (MT), metaproteomic (MP), and meta-metabolomic (MM) data were generated for each sample. Pre-processed MG and MT reads were sample-wise assembled using the iterative hybrid assembly pipeline of the Integrated Meta-omics Pipeline (IMP). After assembly, taxonomic annotation was performed at the read and contig levels, followed by gene prediction and functional annotation on the assembled contigs. Expressed proteins (MP) were identified using the predicted genes from the MG/MT hybrid assembly. For these three omics levels, we generated taxonomic and functional profiles that are referred to as taxMG, taxMT, and taxMP for the taxonomic level, and funMG, funMT, and funMP for the functional level, respectively. Community-based networks were reconstructed from gene annotations. Finally, the meta-metabolome (MM) was integrated with the other omics data at the network level. The integrated multi-omics analysis was performed while accounting for the available clinical metadata

increase in alpha diversity in PD compared to HC and iRBD compared to PD, respectively (Fig. 2A, Mann–Whitney test, $p < 0.05$). We then analyzed beta diversity for all omics layers. TaxMG, funMG, and taxMP revealed no statistically significant differences between the three groups (Extended Fig. 1A, 1B, and 1C, PERMANOVA, $p = 0.2$, 0.8 and 0.35 , respectively), while taxMT, funMT, MM, and funMP showed a statistically significant separation of the groups, especially for funMT (Fig. 2B–D, Extended Fig. 1D, PERMANOVA $p = 0.001$, 0.001 , 0.005 , and $p = 0.008$, respectively). Permutation analysis revealed that funMT, funMP, and MM resulted in the best separation of the three groups while taxMG, funMG, and taxMP exhibited the lowest separation capacity (Extended Fig. 1E, PERMANOVA R^2 of 0.043 , 0.041 , 0.038 , 0.018 , 0.014 , and 0.017 , respectively). Pairwise comparisons using PERMANOVA demonstrated that most differences were found between either HC and PD or HC and iRBD, with only funMP and funMT showing statistically significant differences between PD and iRBD (Fig. 2E). Based on these differences, we next assessed how the confounding factors sex, age, and constipation may impact beta diversity. Sex and constipation were found to have a significant association with beta diversity for taxMT and funMT, while age was associated with taxMG

(Fig. 2F). Importantly, MM and funMP were not found to be associated with any confounders (Fig. 2F).

We next looked at differential abundances of taxa to highlight the compositional differences between the groups. No significant differences were found in taxMG at the genus and species levels between any of the three groups (Fig. 2G and Extended Fig. 2A $q > 0.05$, SIAMCAT). For taxMT, SIAMCAT highlighted an increase in *Alistipes obesi* and *Ruthenibacterium lactatiformans* in iRBD vs HC, *Roseburia incertae sedis*, *Blautia massiliensis*, *B. obeum*, and *Clostridium* sp. were decreased in PD vs HC, and no significant differences were found between PD and iRBD (Fig. 2G, $q < 0.05$, $q < 0.05$, and $q > 0.05$, respectively). We identified the genus *Eubacterium* as being decreased in PD compared to HC in taxMT (Extended Fig. 2A, $q < 0.05$, SIAMCAT). The ALDEx2 algorithm highlighted only *Roseburia incertae sedis* to be depleted in PD after FDR correction (Extended Fig. 2B).

Subsequently, we investigated overall differences in gene abundances and expressions of the microbial genes linked to the observed differences in metabolites using the KEGG database. We observed no statistically significant differences in gene abundances between HC vs PD and HC vs iRBD (funMG, Extended Fig. 2C and 2D, $q > 0.05$). Gene expression highlighted one Transcript

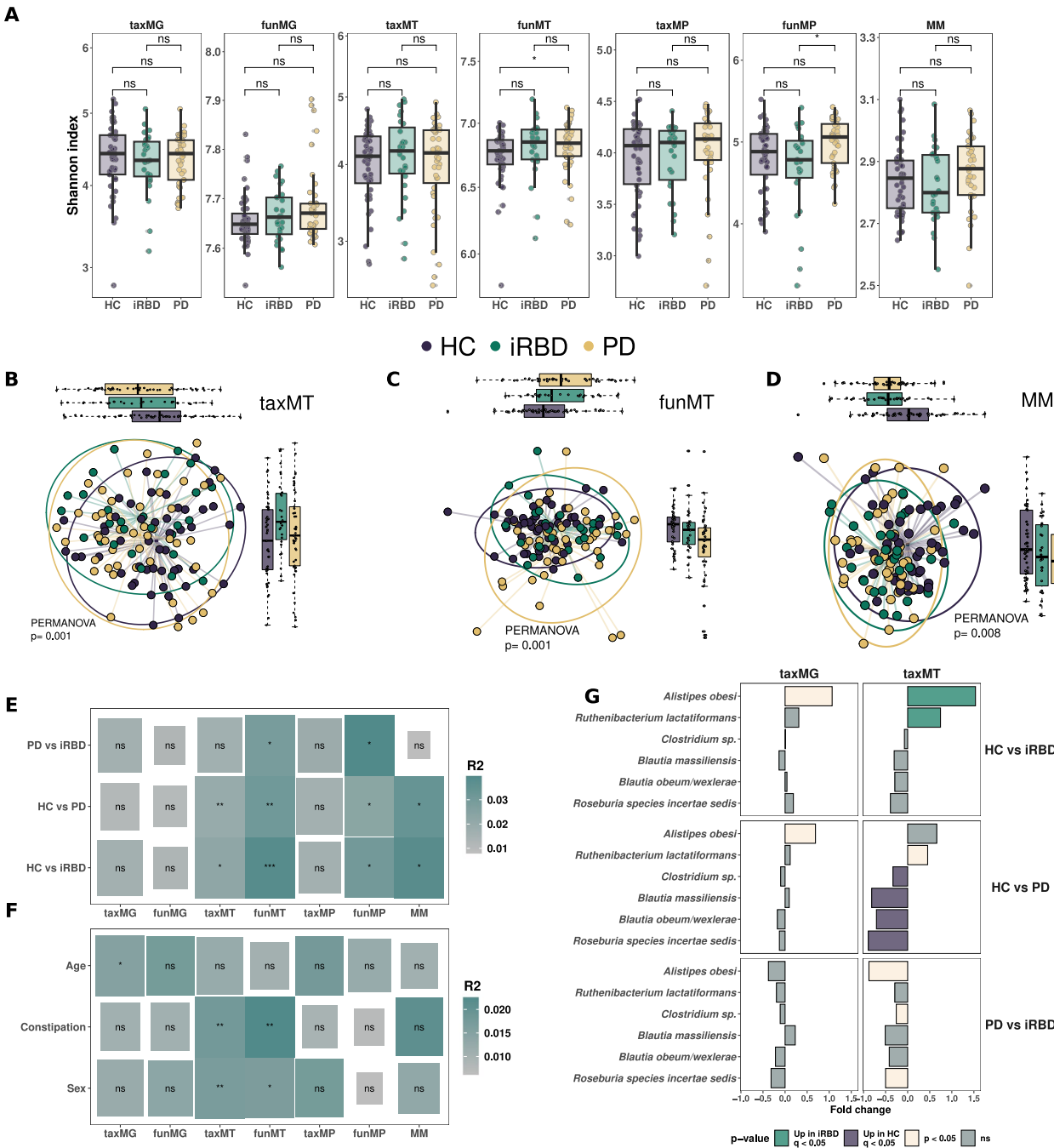


Fig. 2 Microbiome structure is altered in PD and iRBD vs HC. **A** Shannon indices for different omics levels comparing healthy controls (HC), idiopathic REM sleep behavior disorder (iRBD), and Parkinson's Disease (PD). *p*-values are based on pairwise Mann–Whitney tests. Non-metric multidimensional scaling (NMDS) based on Bray–Curtis dissimilarity for taxonomic annotation of metatranscriptomic (taxMT) (**B**) and functional metatranscriptomic (funMT) data (**C**). **D** Principal component analysis (PCA) for meta-metabolomic (MM) data based on untargeted, targeted SCFA, and targeted bile acid abundances. All three Quantifications have been sum-normalized before any merging. PCA was then computed on the merged matrix. All tests are based on PERMANOVA with 1000 permutations. **E** Pairwise PERMANOVA between groups for each omics level. **F** PERMANOVA analysis for “Age,” “Sex,” and “Constipation” for each omics level. Size of rectangle is based on $-\log_{10}(p\text{-value})$ and color on R^2 value. **G** Differential abundance analysis using SIAMCAT for taxMG and taxMT. Values are pseudo fold changes for pairwise comparisons between groups and size is based on $-\log_{10}(p\text{-value})$. Shape is based on significance before and after FDR correction

upregulated in HC vs iRBD, but 145 transcripts upregulated in HC vs PD (funMT, Extended Fig. 2E and 2F, $q < 0.05$ and $q > 0.05$, respectively). No genes or transcripts were significantly higher in iRBD or PD compared to HC (Extended Fig. 2C–F, $q > 0.05$). Finally, we found no statistically significant differences in gene or transcript abundances between PD and iRBD (data not shown).

Altered metabolome of PD patients is linked to microbial abundance and activity

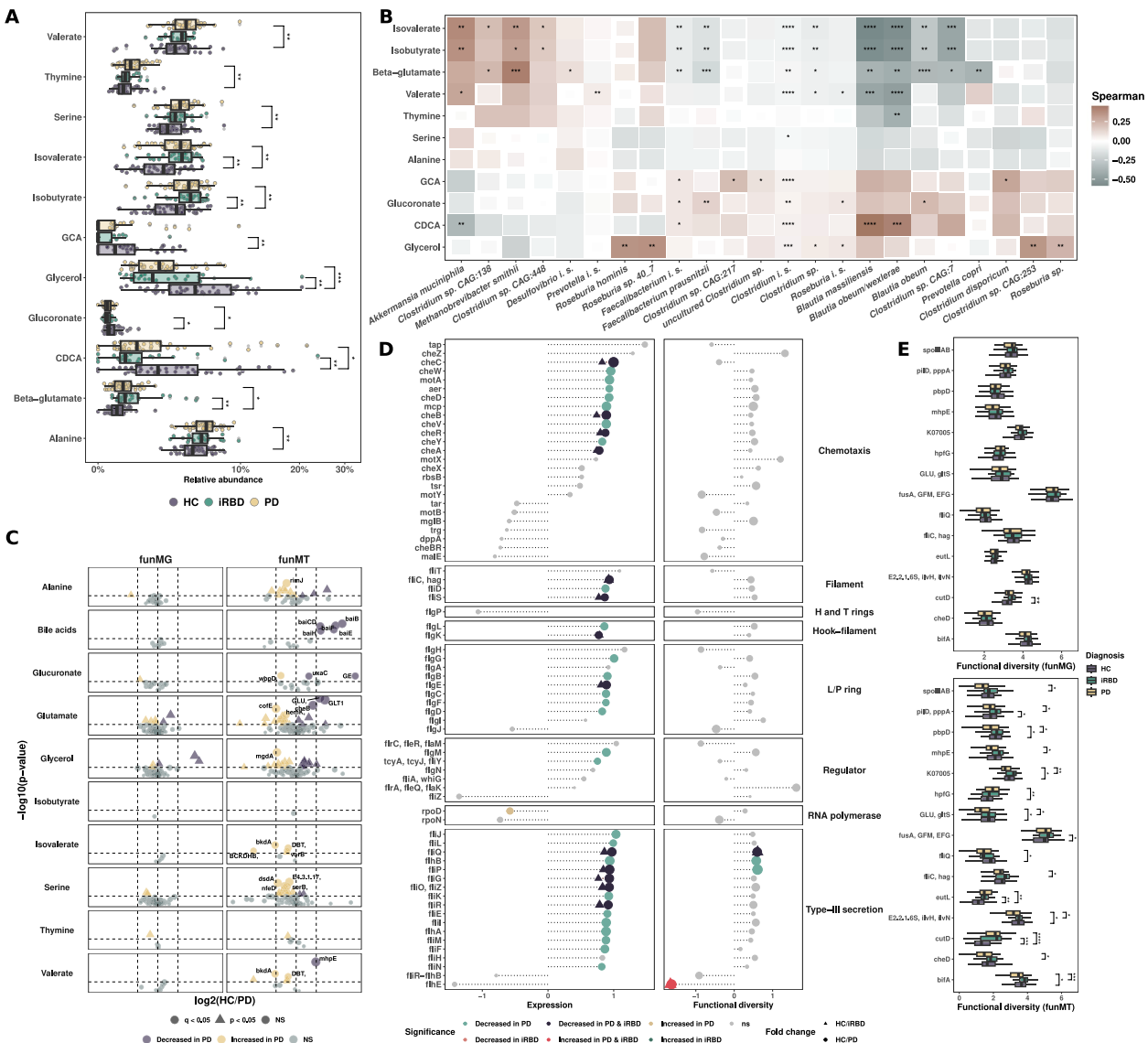
Considering that MM is associated with PD and iRBD but not with confounders, we chose MM as a robust guide for further statistical comparisons. In addition, MM can be considered as one of the final outputs of microbial activity and an important driver of microbial effects on the host. For this purpose, we removed unidentified compounds from the statistical testing, because we cannot link microbial genes with them. Our analyses revealed 11 statistically significantly different compounds between the groups, including alanine, β -glutamate, serine, and glycerol (Fig. 3A, $q < 0.05$). We found a significant increase in isovalerate, isobutyrate, and valerate in PD patients (Fig. 3A, $q < 0.05$) but no differences for butyrate, acetate, formate, propionate, or total SCFAs (data not shown). Primary bile acids glycocholic and chenodeoxycholic acids were decreased in PD and in iRBD and PD, respectively (Fig. 3A, $q < 0.05$). Glutamate was not differentially abundant between the three groups (data not shown). To unravel the effect of confounding factors, we measured the variance explained by each factor on the compounds' abundances. Diagnosis explained more variance than constipation or sex (Extended Fig. 3A–C, diagnosis mean = 4.18%; median = 3.33%, sex mean = 2.08%, median = 1.3%; constipation mean = 3.7%, median = 2.68%). We found only malic acid as being significantly associated with sex and no compound to be associated with constipation (Extended Fig. 3A–C, $q < 0.05$, $q > 0.05$). Based on the PERMANOVA results and variance analysis, the differences observed in MM were most strongly associated with the disease status and, importantly, not with the confounding factors.

We next checked correlations between metabolite abundances and microbial abundances alongside activities. Metabolite abundances were linked to microbial abundances for taxMG and taxMT. TaxMG exhibited fewer significant correlations with metabolite abundances compared to taxMT, except for *Methanobrevibacter smithii* and *Clostridium* sp. CAG:138, which were positively correlated with β -glutamate and isovalerate levels (Extended Fig. 3D, Spearman test, $q < 0.01$). Conversely, *Prevotella copri* and *Faecalibacterium prausnitzii* abundances were negatively correlated with β -glutamate levels, while *Roseburia hominis* was positively correlated

with glycerol levels (Extended Fig. 3D, Spearman test, $q < 0.001$). Additionally, for taxMT, we noted positive correlations with isovalerate and isobutyrate for *A. muciniphila* and *Clostridium* sp. CAG:448 (Fig. 3B, $q < 0.01$ and $q < 0.05$, respectively). *Methanobrevibacter smithii* was associated with β -glutamate, isovalerate, and isobutyrate (Fig. 3B, Spearman test, $q < 0.01$, $q < 0.001$, and $q < 0.05$, respectively), while *F. prausnitzii*, *Blautia massiliensis*, *Blautia obeum/wexlerae*, *Blautia obeum*, and *Clostridium* spp. were negatively correlated with β -glutamate ($q < 0.05$, Spearman test, Fig. 3B). In addition, *R. hominis*, *Roseburia* spp., and *Clostridium* spp. were positively correlated with glycerol (Fig. 3B, Spearman test, $q < 0.05$). Of note, while assessing the correlations at the genus level, we noted negative correlations between *Eubacterium* and isovalerate, isobutyrate, and valerate (taxMT, $q < 0.05$, Spearman test, data not shown). Strikingly, *A. muciniphila*, some *Clostridium* spp., and *M. smithii* were positively correlated with compounds found to be elevated in PD and iRBD (β -glutamate, isovalerate, and isobutyrate) while *R. hominis*, *F. prausnitzii*, *Blautia* species, and *P. copri* were negatively correlated with these same compounds and positively correlated with compounds increased in HC (glycerol and CDCA). This revealed groups of bacteria being either linked with individuals with PD and iRBD or HC (Fig. 3B).

Expression of genes linked to glutamate, bile acids, and flagella is dysregulated in the iRBD and PD gut microbiome

To reinforce the results of the differential analysis, we compiled all the orthologs related to the metabolites which were statistically significant between the three groups. More specifically, we used regular expression matching to retrieve all orthologous genes that are linked to the abovementioned metabolites in KEGG. Because the KEGG database only has one metabolite entry annotated as β -glutamate, we selected all glutamate-related genes instead (linked to both L- and D-glutamate). We found no statistically significant differences in gene abundances after correction between any of the three groups (funMG, Fig. 3C and Extended Fig. 4A, $q > 0.05$, Mann–Whitney test). However, we did find a decrease in transcripts in PD for the three known glutamate synthase genes (funMT, *GLT1*:K00264, *GLU*:K00284, and *glbB*:K00265, Fig. 3C, Mann–Whitney test, $q = 0.004$, $q = 0.004$, and $p = 0.006$, respectively) alongside a decrease in *cheB* in PD (K03412, protein-glutamate methylesterase/glutaminase, Fig. 3C, $q < 0.01$, Mann–Whitney test). Furthermore, we found an increase in *cofE*, mainly encoded by Archaea and involved in methanogenesis (K12234, coenzyme F420-0:L-glutamate ligase, Fig. 3C, $q < 0.05$,



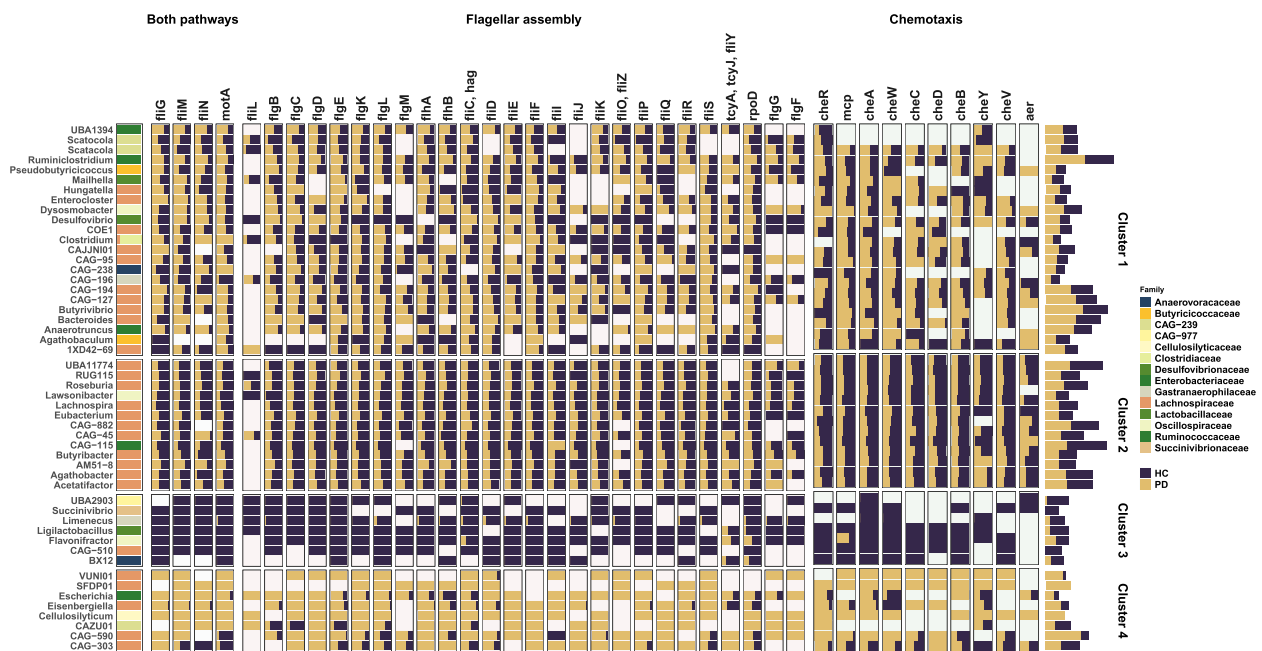


Fig. 4 Flagellar assembly and chemotaxis gene expressions according to taxonomy. Top 50 genera expressing flagellar assembly and/or chemotaxis-related genes are depicted by the mean of transcripts per million per disease status. Values are depicted for each genus and genes on the left panel and the sum of all genes per genus on the right panel with a square root transformation. Genera are clustered based on the $\log_2FC(HC/PD)$ for each gene using a Canberra distance and *hclust()* using “ward.D2” method. Family taxonomic level is depicted on the left

methylamine-glutamate N-methyltransferase, and glutaminase (Extended Fig. 4B, Spearman correlation test, $q < 0.01$).

Since *cheB* is part of the chemotaxis and flagellar assembly (FA) KEGG pathways, we further inspected these pathways to assess the microbial capacity for motility. We found no statistically significant differences in gene abundances (funMG) for either flagellar assembly or chemotaxis pathways between the three different groups (data not shown, $q > 0.05$). In contrast however, based on funMT, the chemotaxis pathway (14/26 Transcripts decreased in PD, 0/26 transcripts increased, Fig. 3D, $q < 0.05$) and the FA pathway (25/46 Transcripts decreased in PD, 1/46 transcripts increased in PD, Mann–Whitney, Fig. 3D, $q < 0.05$) were strongly downregulated in PD. In addition, we found transcripts decreased in both PD and iRBD for five transcripts belonging to chemotaxis and eight belonging to flagella assembly ($q < 0.05$, Mann–Whitney, Fig. 3D). Moreover, eight transcripts showed a decrease in alpha diversity for chemotaxis and flagellin assembly pathways ($p < 0.01$, Mann–Whitney, Fig. 3D). Finally, we observed a decrease in alpha diversity for *GLU*, flagellar assembly transcripts (*fliJ*, *fliQ*, and *fliC*), and *cheD* in PD compared to the other two groups (Fig. 3E, $q < 0.05$, Dunn test) while *pilD* and *aaaT* were elevated in iRBD (Fig. 3E, Dunn test, $p < 0.05$).

Flagellin and chemotaxis are differentially expressed depending on taxonomy and disease status

We next assessed the taxa expressing secondary bile acid biosynthesis (SBAS), FA, and chemotaxis pathway genes. While taxonomically resolving SBAS expression at the genus level, we found an interesting divergence with the differential analysis at the transcript level presented in Fig. 3C. We previously found that *baiCD*, *baiE*, *baiB*, *baiH*, and *baiF* were all significantly decreased in PD, while *baiA* and *baiN* were not (Fig. 3C). When we resolved taxonomic information of these functional genes, we found only statistically significant differences in transcripts for *baiA* and *baiN* (17 after FDR correction and 31 before correction), but not for the other transcripts previously found to be differentially expressed ($q < 0.05$ and $p < 0.05$, respectively, Extended Fig. 5A). We found a significant decrease in the expression for *baiN* in PD for *Blautia*, *Fusicatenibacter*, *Choladocola*, *Oli-verpabstia*, and *Ruminococcus* genera and an increase in expression for *baiA* encoded by the *Lawsonibacter* genus in PD ($q < 0.05$, Mann–Whitney test, Extended Fig. 5A).

When resolving the taxonomic expression of FA and chemotaxis genes, we noted a differential expression according to disease status and taxonomy. More specifically, we identified different clusters of taxa expressing FA and chemotaxis genes. The first cluster was composed of microbes expressing these genes principally in PD,

including *Ruminiclostridium*, *Enterocloster*, *Dysosmobacter*, and *Butyrvibrio*. A second cluster was composed of microbes expressing these genes principally in HC, including *Roseburia*, *Agathobacter*, and *Eubacterium* (Fig. 4). A third cluster was composed of taxa expressing FA and chemotaxis genes almost exclusively HC, including *Flavonifractor*, *Succinivibrio*, *Eisenbergiella*, or *CAG-603* (Fig. 4). Strikingly, we found a fourth cluster composed of taxa expressing flagellin or chemotaxis genes almost exclusively in PD, including *Escherichia*, *Cellulosilyticum*, *Citrobacter*, or *Eisenbergiella*. We subsequently investigated the expression levels of genes of the extracellular parts of flagella in the cluster wherein *Roseburia* and most *Lachnospiraceae* were located (cluster 2). Overall, we found a decrease in flagellin (*fliC*), filament cap (*fliD*, *fliS*), and hook-filament junction genes (*flgK*, *flgL*) in PD or iRBD for *Roseburia*, CAG-115, and *Agathobacter* genera (Extended Fig. 5B, $p < 0.05$, Dunn test).

Metabolites and metabolites-related genes associated with PD are central to the microbial ecosystemic metabolism

To quantify the importance of glutamate derivatives and related genes in microbial metabolism, we next reconstructed microbiome-wide metabolic networks as previously described [36]. We mapped genes related to the compounds identified earlier as significantly different

between the different groups. The metabolic network highlighted glutamate-related genes as compact and placed in the middle of the network while glycerol was more scattered across the network (Fig. 5A). Strikingly, glutamate-related genes formed a subnetwork central to the overall network with a betweenness centrality measure (BC) of 93.0 compared to 0.0009 for the whole community (Fig. 5B). Crucially, L-glutamate was the most central non-cofactor metabolite in the network and 2-oxoglutarate (L-glutamate derivative) was the second most central metabolite (Fig. 5C). Considering this, it is apparent that glutamate and glutamate-related genes are central to microbiome metabolism and that modifications in the levels of these metabolites or transcripts reflect profound modifications of microbial metabolism.

Multi-omics factor analysis validates β -glutamate and flagella links with PD

To validate our findings, we used an unsupervised method with the Multi-Omics Factor Analysis (MOFA). The resulting MOFA model included 10 factors (F1-10 hereafter) whereby F1 showed a strong association with disease status ($p < 0.05$, ANOVA, Fig. 6A). A complete description of the MOFA model is provided in the Extended information. We found that F1 mainly explained the variance of funMT, taxMT, and MM (17.6%, 9.6%, and 5.7%, respectively, Extended Fig. 6A–B) and showed separation of HC and PD,

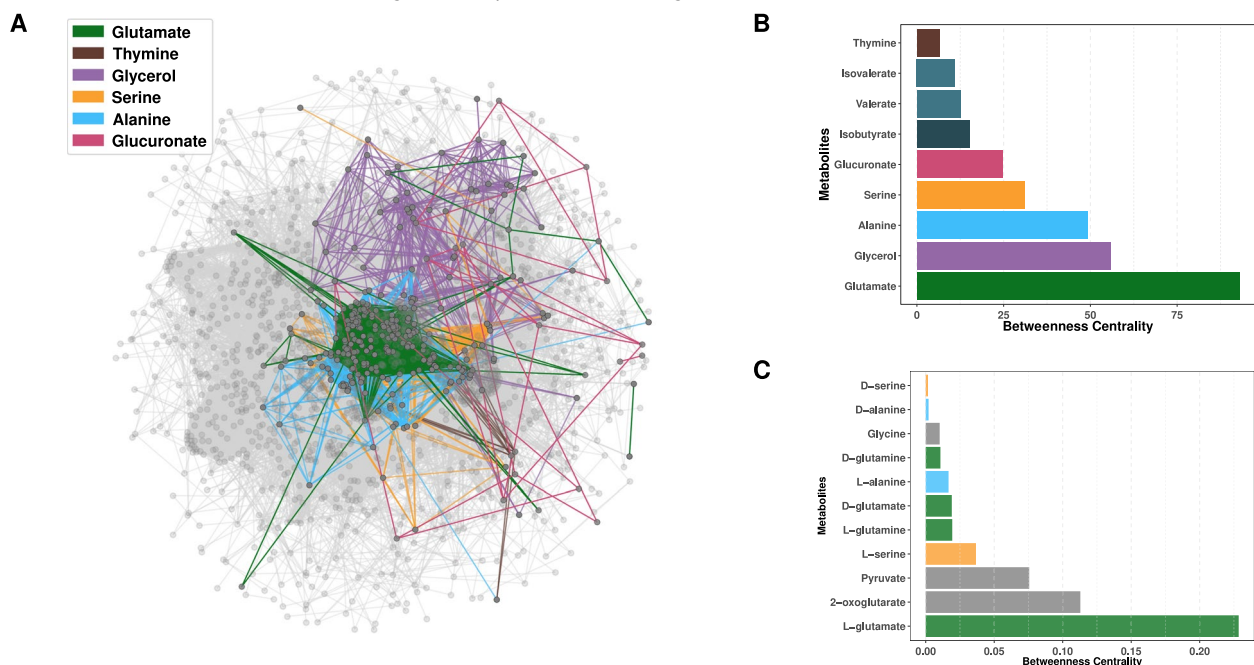


Fig. 5 Metabolic network of whole community interactions. **A** Metabolic network of whole community interactions, with KEGG KO's represented as nodes and associated metabolites as edges. Node sizes reflect MT/MG ratio of normalized read counts for each KO. Highlighted is the overlap between glutamate-, thymine-, glycerol-, serine-, alanine-, and glucuronate-associated subnetworks mapped on the whole community network. **B** Betweenness centrality calculated for the key metabolites highlighted in the whole-community network based on genes as nodes. **C** The network was inverted to calculate betweenness centrality for metabolites; here, metabolites are nodes and genes are edges

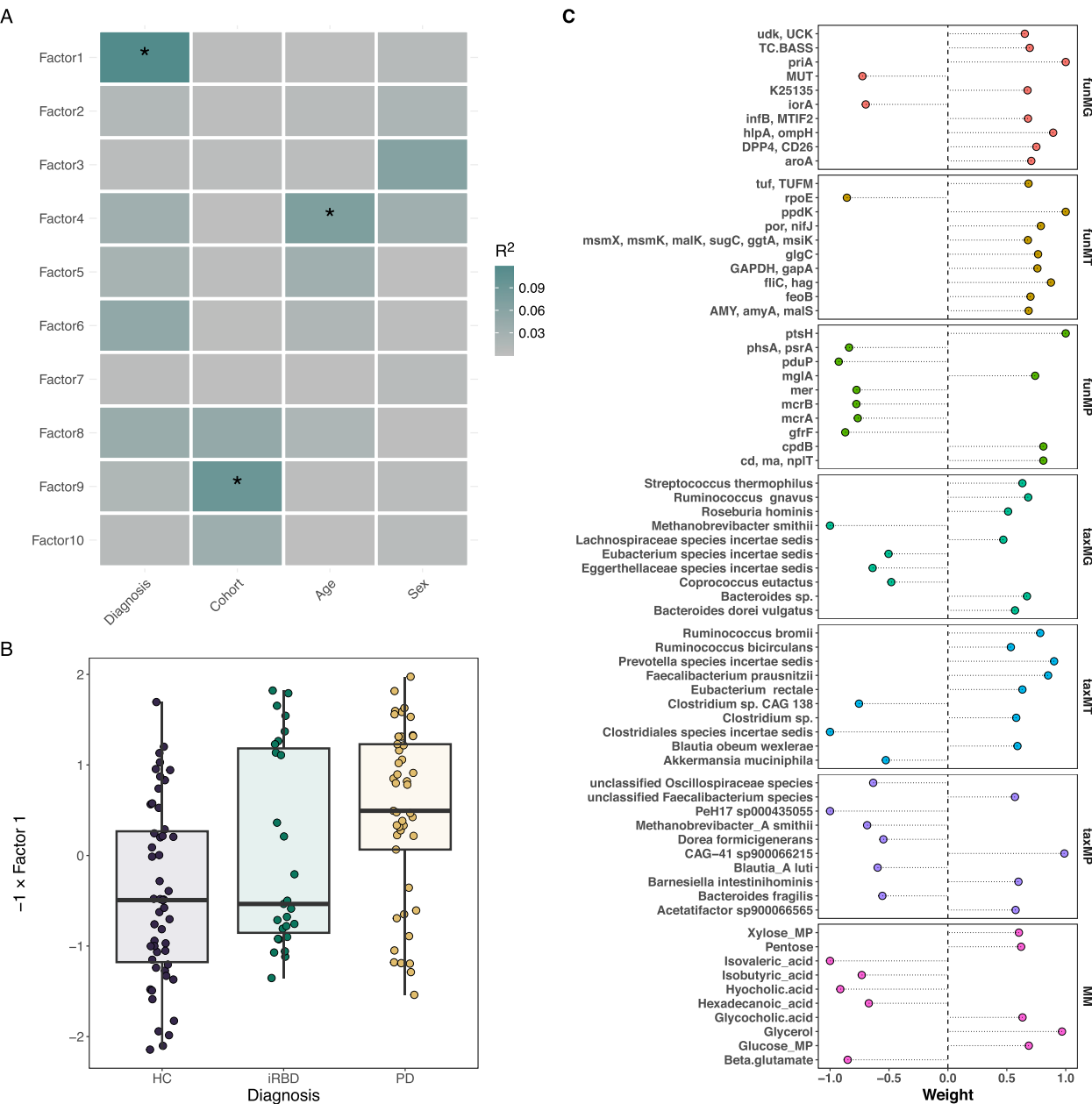


Fig. 6 MOFA analysis validates the findings of per omics layer analyses. **A** Associations of MOFA factors with diagnosis and confounders; the color of rectangles represents partial R^2 values, significant associations (FDR-adjusted p -value < 0.05) are marked with an asterisk. **B** Abundance of the factor 1 in studied groups; for graphical purpose, the Y axis was inverted. **C** Minimum–maximum scaled weights of top 10 features per omics layer contributing to factor 1. The sign of the weight indicates the direction of the effect, the abundance of features with positive weights is positively associated with the factor 1 level, and the abundance of features with negative weights is negatively associated with the factor 1 level

but not iRBD versus other groups ($p < 0.05$, ANOVA, Fig. 6B). The abundance of F1 was lower in the PD group compared to HC and iRBD groups, meaning omics features with a negative contribution to the F1 were enriched in PD, while features with a positive contribution were enriched in the HC and iRBD groups (Fig. 6C). Specifically, the microbiome of PD

was characterized by the joint increase in abundance of *M. smithii*, archaeal proteins and genes (based on MP and MG data), *A. muciniphila* (based on taxMT data), β -glutamate, isovalerate, isobutyrate, hexadecanoic, and hyocholic acids, whereas the abundance of *R. hominis* (taxMG data), flagellin (funMT), GCA, and glycerol were decreased in PD (Fig. 6C). Overall, the MOFA

results are strongly consistent with the per omic layer analysis and support our earlier findings.

From the microbiome to the host: expanding the Expobiome map

To contextualize our newly gained insights about the links between microbial taxa, their molecules, and mechanistic evidence from the literature on host physiology, we integrated our results into the Expobiome map [35] (<https://expobiome.lcsb.uni.lu>). Specifically, we added *R. intestinalis* and *hominis*, *B. wexlerae*, *coccoides* and *producta*, *Clostridium butyricum*, and *F. plautii* as new taxonomic elements linking flagellin, lipoteichoic acid (LTA), and SCFAs signaling to immune-relevant human proteins (Fig. 7). We also added mechanistic links on *F. prausnitzii* and its effects mediated by the short-chain fatty acid butyrate. Thereby, we integrated new mechanistic details linking the combinatorial functions from these taxa and pro-inflammatory component effectors such as TLR5 and its downstream activation of NLRP3 inflammasome and NF- κ B, IL-6, and TNF- α secretion. Thereby, these new additions to the Expobiome map highlight the convergent immunomodulatory functions of commensal taxa, perturbed in the context of PD.

To facilitate the investigation of potential relationships between microbial taxa, microbial compounds, and human proteins that have been identified as crucial players in PD, we have furthermore linked the Expobiome map to the Parkinson's disease map (<https://pdmap>.

[uni.lu](https://expobiome.lcsb.uni.lu)). Notable overlaps exist between host proteins that have been established to interact with microbial components in both the immune system and the central nervous system. For example, dendritic cells and macrophages have well characterized interactions with flagellin through TLR5 binding. TLR5 and other antigen sensing receptors are present on microglia, a cell type of critical importance in neuroinflammation and PD. This integration of information from both so far disconnected maps allows for a more comprehensive understanding of these interactions and helps to draw new hypotheses on the intricate links between the gut microbiome, the immune system, and PD.

Discussion

Here, we investigate the Links between the Gut microbiome and PD using an comprehensive integrated multi-omics approach based on standardized sample collection and extraction. We include individuals with iRBD as a prodrome of PD to compare early and later stages of the disease but do not find statistically significant differences between iRBD and PD, especially in comparison to the more pronounced differences found between HC and PD. Previous studies have shown differences in the different early stages of PD, but these studies were performed using 16S rRNA gene amplicon data [7, 10]. In contrast to the amplicon-based results, no differentiation between iRBD and PD is found using metagenomic data [5]. More specifically, and contrary to previous

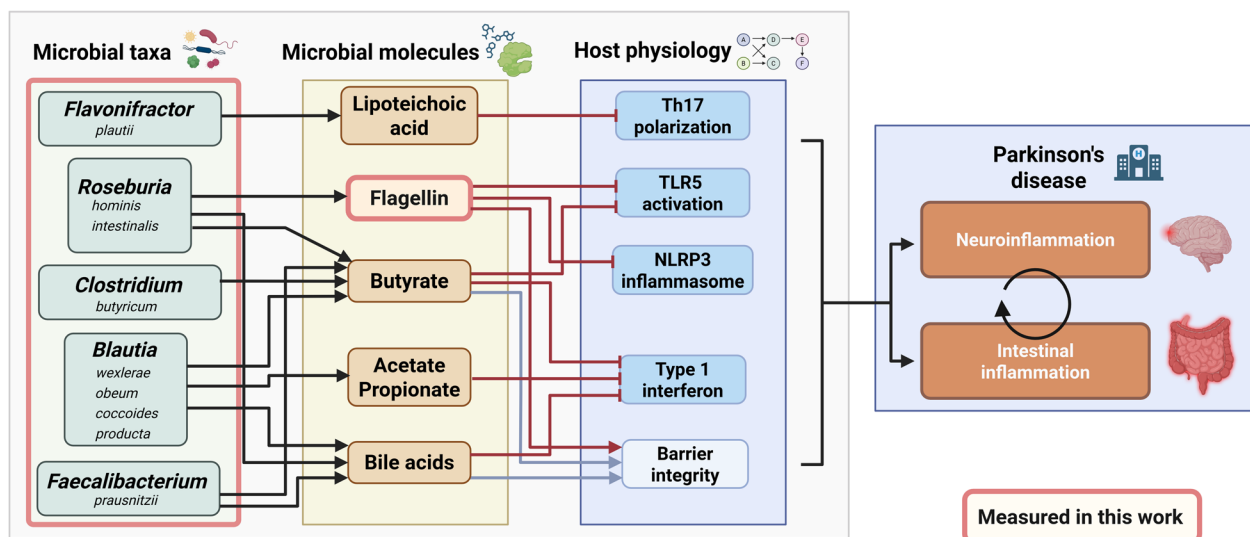


Fig. 7 Schematic representation of the new taxonomic and molecular additions to the Expobiome map. Additional literature highlighting mechanistic evidence linking microbial taxa to host physiology were collected. We specifically integrated *Roseburia*, *Blautia*, *Faecalibacterium*, *Flavonifractor*, and *Clostridium* species, taxa that we measured and found significant differences for in this work, as new taxonomic elements and linked their molecules to host physiology, in particular the immune system. The Expobiome map is now also connected to the Parkinson's disease map, allowing direct overlays of pathways between the maps. Parts of the scheme in pink are the elements we were able to measure in this work

findings [5] including our own work [7], we do not find significant differences between HC and PD individuals in the metagenome with respect to alpha or beta diversity. However, we show differences in transcriptional activity for *R. hominis*, *B. obeum*, *B. wexlerae*, and *P. copri* which are known to be found at lower abundance in PD [5, 6, 8–13]. Moreover, we find several species, including *A. muciniphila*, *M. smithii*, and *Clostridium* spp., to be positively correlated with the abundance of compounds increased in PD such as β -glutamate, isovalerate, or isobutyrate, while other species are inversely correlated with these compounds (species belonging to *Faecalibacterium*, *Blautia*, *Eubacterium*, and *Roseburia*). Collectively, these findings are in line with findings associating those genera with PD [5, 6, 8–13].

We identify metabolites that are associated with PD, among which BAs, alanine, serine, and β -glutamate showed differences in linked gene expression. The BA chenodeoxycholic acid (CDCA) is decreased in PD and iRBD whereas glycocholic acid (GCA) is decreased only in iRBD. We also detect a decrease in transcripts related to BAs in PD but not iRBD. BAs have a wide spectrum of effects on the immune system [22, 23], metabolism [37], hormones [38], and on the CNS [39]. GCA has been found to be increased in the CNS of PD and associated with disease duration [39], stressing the importance of analyzing BAs in PD and prodromal stages of PD. Interestingly, when we taxonomically resolve the differentially expressed SBAS genes, we find different signals than when we compare the genes by themselves. For the differential expression analysis, we find a significant decrease in PD for *baiB*, *baiCD*, *baiE*, *baiF*, and *baiH* but not for *baiA* and *baiN*. When we resolve the taxonomic expression of the SBAS genes we find differences only for *baiA* and *baiN*. This finding underscores the importance of considering the microbiome at the taxonomic and functional levels at once. In contrast to previous studies [6, 15], we do not find a decrease in butyrate, propionate, acetate, or overall levels of SCFAs. However, we find an increase in isovalerate, isobutyrate, and valerate as described previously [15]. Previous studies have highlighted the correlation of fecal concentrations of isobutyrate and isovalerate with PD severity [40] and concentrations of valerate with disease duration [41]. There currently is an apparent lack of knowledge regarding isobutyrate or isovalerate with respect to host-microbiome interactions and very few related genes annotated in the KEGG database, thereby hindering interpretations concerning potential links between these metabolites and PD.

β -glutamate and glutamate-related genes are of particular interest in the context of PD because of the reported toxic effects of L-glutamate on neurons [24] and

because of its association with microbial activity. In addition, glutamate levels have been reported to be increased in PD individuals' blood sera [17, 18]. β -glutamate has strikingly only one reaction described in the KEGG database; in contrast, L-glutamate has 213 reactions and 54 pathways, while D-glutamate has 12 reactions and 2 pathways. We find that glutamate and glutamate-related genes are central to microbial metabolism, which underpins the notion that the highlighted differences reflect a pronounced impact on gut microbiome function. Of note, glutamate likely has local effects on enteric neurons with subsequent influences on the CNS [42]. Although we show significant differences in β -glutamate and glutamate-related genes in the stool of PD patients, we detect no significant differences in the glutamate levels (both enantiomers) between the three groups. We highlight significant differences in β -glutamate for which we currently cannot evaluate the effects on the ENS and CNS. Overall, based on our results, microbiome-driven glutamate metabolism and its impact on the glutamatergic system must be comprehensively studied in the future to disentangle its link with PD.

Glutamate-related genes are further involved in chemotaxis and flagellar assembly pathways, highlighting a modification of the latter genes' expression with the majority being decreased in PD individuals and some of these genes in iRBD too. A decrease in flagellar assembly gene abundances has been previously reported in a metagenomic analyses of PD [9]. We do not see significant differences in the linked gene abundances, but their expression levels are significantly different, highlighting altered regulation of transcription in PD/iRBD compared to HC. Flagellar assembly and chemotaxis genes are also differentially expressed by specific microbes; the genera *Escherichia* and *Cellulosilyticum* for instance express flagellar assembly genes only in the context of PD without being statistically differentially abundant.

Flagellin is a known immunogenic molecule, a potent pro-inflammatory compound in pathogens [43–48] and is thus targeted by secretory IgA [49]. However, flagellin in commensals, especially in the *Lachnospiraceae* family, has been shown to be either “silently recognised” [50] or elicit anti-inflammatory effects [51–53]. Among the *Lachnospiraceae* family, the genus *Roseburia* shows a decrease in the transcription of flagellin in the gut microbiome of PD. This in turn may be linked to immune system dysregulation and exert indirect effect on the CNS, particularly in microglia as shown in a previous study using a murine model of PD [54]. Previous studies have shown an increased inflammatory state in PD, with increased pro-inflammatory circulating immune cells [55], cytokines [56], and activated microglia [57]. In addition, microglia can

be activated by α -synuclein via NLRP3 following TLR2 (lipopolysaccharide (LPS) sensing) and TLR5 (flagellin sensing) activation [58]. Therefore, the activation or inhibition of TLRs by distinct flagellins may modulate microglia activation by competing with α -synuclein and affect PD progression. It is however not known if flagellin can reach the CNS directly or if this potential effect would be mediated by an indirect effect through the immune system. Evidence exists that a local, microbiome-mediated effect on the enteric immune system can impact distant sites such as the brain [59] or tumors [60]. This provides solid grounds for the hypothesis that flagellin can regulate the CNS and modulate the neuroinflammation occurring in PD as resolved in the Expobiome map (Fig. 7).

Knowledge on microbial antigens is based on studies of pathogens, mainly *Salmonella typhimurium*, pathogenic *Escherichia coli*, *Staphylococcus aureus*, or *Bacillus subtilis*. The results of these studies have led the scientific community to consider bacterial antigens as pro-inflammatory and danger signals for the host [61]. Our findings associating flagellin expression with health, along with recent reports showing the anti-inflammatory effects of flagellin from *R. hominis* and *intestinalis* [50, 51], LTA from *F. plautii* [62], and LPS from *Bacteroides dorei* [63], call for a re-evaluation of the impact of common antigens on host physiology. Finally, the diverse compounds produced by the gut microbiome demand detailed investigation, with regard to the taxonomy and commensal/pathogen phenotype of the producing microbe, to fully understand their modulating effects on the immune and nervous systems alongside metabolism in health and disease.

In conclusion, our work clearly highlights the importance of studying microbiome functions rather than restricting microbiome analyses to taxonomic structure. Specifically, the combination of MT and MM provides clear insights into the activity of specific microbial taxa in relation to disease. In our present work, MT reveals that disease association is not solely determined by gene expression levels; the diversity of microbes capable of expressing a specific function and the specific taxa expressing those functions are also of immediate interest and relevance. While microbiome interventions are mainly designed to deal with dysbiosis at the taxonomic level, here, we propose to focus on restoring the lost functions from dysbiotic gut microbiomes. The future of microbiome research might lie in understanding how we can modulate, re-activate, or shut down specific microbial functions in vivo in order to improve knowledge and later improve or functionally tailor microbiome interventions.

Methods

Key resources table

Reagent or resource	Source	Identifier
MedAuxil Feces Collection Device	MecAuxil	Cat#Fe-Col
Sarstedt collection tubes	Sarstedt	Cat#80.623.022
Zymo DNA Clean&Concentrator-5	Zymobiomics	Cat#D4014
Zymo RNA Clean&Concentrator-5	Zymobiomics	Cat#R1014
Agilent 2100 Bioanalyzer	Agilent Technologies	G2939A
NanoDrop	Thermo Fisher Scientific	Cat#ND-1000
Qubit ds DNA BR Assay kit	Qubit	Cat#Q32850
Qubit RNA BR Assay kit	Qubit	Cat# Q10210
Bioruptor NGS	Diagenode	Cat#UCD-600
TruSeq Stranded mRNA	Illumina	Cat#20,020,595
RiboZero kit	Illumina	Cat#MRZB12424
NextSeq500	Illumina	Cat#SY-415-1001
Sera-Mag SpeedBeads	GE Healthcare	Cat#45,152,105,050,250; Cat#65,152,105,050,250
Q Exactive™ HF-X Hybrid Quadrupole-Orbitrap	Thermo Fisher Scientific	0726042
TriVersa NanoMate	Advion	Cat#STEM-LAHTPO-0097-ZJF
Acclaim PepMap 100 C18	Thermo Fisher Scientific	Cat#164,535
Tecan Freedom EVO 200	Tecan	https://lifesciences.tecan.com/freedom-evo-platform
Biological samples		
Fecal sample from PD and healthy individuals	Paracelsus-Elena Klinik, Kassel, Germany	
Fecal sample from iRBD individuals	Department of Neurology, Philipps-University, Marburg, Germany	
Chemicals, peptides, and recombinant proteins		
Ammonium bicarbonate	Sigma-Aldrich	Cat#09830
Iodoacetamide	Sigma-Aldrich	Cat#11149
Acetonitrile	Sigma-Aldrich	Cat#L010300
Ethanol	Sigma-Aldrich	Cat#459,844
Ammonium formate	Sigma-Aldrich	Cat#156,264
Trifluoroacetic acid	Sigma-Aldrich	Cat#900,518
Deposited data		
GTDB	Parks et al. 2022 [64]	https://github.com/ECogenomics/GtdbTk ; RRID:SCR_019136

Reagent or resource	Source	Identifier
SILVA 138.1	Quast et al. [65]	http://www.arb-silva.de/ ; RRID:SCR_006423
KEGG	Kanehisa et al. [66]	http://www.kegg.jp/ ; RRID:SCR_012773
UniProtKB	Schneider et al. [67]	http://www.uniprot.org/help/uniprotkb ; RRID:SCR_004426
cRAP	The Global Proteome Machine Organization	https://www.thegpm.org/crap/ ; RRID:SCR_018187
Software and algorithms		
Integrated Meta-omic Pipeline	Narayanasamy et al. [34]	https://git-r3lab.uni.lu/IMP/imp3
Python	Python Programming Language	http://www.python.org/ ; RRID:SCR_008394
R v4.4.1	R Core Team	https://www.r-project.org/ ; RRID:SCR_001905
RStudio v2024.4.2.764	Rstudio	https://posit.co/ ; RRID:SCR_000432
BioRender	BioRender	https://www.biorender.com/
MetaboliteDetector	Hiller et al. [68]	https://md.tu-bs.de
Thermo Xcalibur	Thermo Fisher Scientific	http://chemistry.unt.edu/~verbeck/LIMS/Manuals/XCAL_Quant.pdf ; RRID:SCR_014593
TaceFinder	Thermo Fisher Scientific	https://www.thermofisher.com/lu/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/lc-ms-data-acquisition-software/tracefinder-software.html RRID:SCR_023045
MEGAHIT	Li et al. [69]	https://github.com/voutcn/megahit ; RRID:SCR_018551
Prokka	Seeman et al. [70]	http://www.vicbioinformatics.com/software/prokka.shtml ; RRID:SCR_014732
featureCounts	Liao et al. [71]	http://bioinf.wehi.edu.au/featureCounts/ ; RRID:SCR_012919
Mantis	Queirós et al. [35]	https://github.com/PedroMTQ/mantis ; RRID:SCR_021001
MOTUs	Milanese et al. [72]	https://motu-tool.org/
Kraken	Davis et al. [73]	http://www.ebi.ac.uk/research/enright/software/kraken ; RRID:SCR_005484
Sipros	Guo et al. [74]	http://sipros.omicsbio.org
SIAMCAT	Wirbel et al. [75]	https://www.biocconductor.org/packages/release/bioc/html/SIAMCAT.html

Reagent or resource	Source	Identifier
NetworkX v3.3	Hagberg et al. [76]	https://networkx.org RRID:SCR_016864
ggplot2	Wickham et al. [77]	https://cran.r-project.org/web/packages/ggplot2/index.html ; RRID:SCR_014601
vegan	Oksanen et al. [78]	https://github.com/vegandevs/vegan RRID:SCR_011950
rstatix v0.7.2	Kassambara [79]	https://cran.r-project.org/web/packages/rstatix/index.html RRID:SCR_021240
ComplexHeatmap v2.20.0	Gu [80]	https://github.com/jokergoo/ComplexHeatmap
ggpubr v0.6.0	Kassambara [81]	https://cran.r-project.org/web/packages/ggpubr/index.html
DeSeq2 v1.42.1	Love et al. [82]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html RRID:SCR_015687
ALDEx2 v 1.36.0	Fernandes et al. [83]	https://github.com/ggloor/ALDEx2 ; RRID:SCR_003364
Multi-Omics Factor Analysis (MOFA) 2	Argelaguet et al. [84]	https://biofam.github.io/MOFA2/ RRID:SCR_022992
Other		

Experimental model and study participant details

All subjects from both cohorts provided informed written consent, and the sample analysis was approved by the Comité National d’Ethique de Recherche of Luxembourg (reference no.: 140174_ND).

Kassel cohort

The DeNoPa cohort represents a prospective, biannual follow-up study of (initially de novo) Parkinson’s disease (PD) patients at the Paracelsus-Elena Klinik, Kassel, Germany. Fecal samples from PD patients (46) and healthy controls (29) were collected during the 4-year follow-up visit for the cohort. Details on inclusion and exclusion criteria and ancillary investigations have been published previously [85, 86]. Subjects with idiopathic rapid-eye-movement sleep behavior disorder (iRBD, 13) were recruited at the same clinic, diagnosed according to consensus criteria of the International RBD study group [87] using video-assisted polysomnography, and were included only if they showed no signs of a neurodegenerative disorder. DeNoPa subjects were required to have a 4-week antibiotic free interval before fecal sample collection. As additional control subjects, we collected fecal samples from (20) neurologically healthy subjects living

in the same household as the DeNoPa participants. Samples of de novo PD patients from a cross-sectional cohort at the same clinic were included if subjects were recently diagnosed, drug-naïve, and met United Kingdom Parkinson's Disease Society Brain Bank (UKPDSBB) clinical diagnostic criteria [30]. All subjects except household HC were interviewed and examined by an expert in movement disorders. The study conformed to the Declaration of Helsinki and was approved by the ethics committee of the Physician's Board Hessen, Germany (FF 89/2008). The DeNoPa trial is registered at the German Register for Clinical trials (DRKS00000540).

Marburg cohort

We also added samples from 14 patients with polysomnography-confirmed iRBD which were recruited from the outpatient clinic of the Department of Neurology, Philipps-University, Marburg, Germany, between November 2015 and November 2016. iRBD was diagnosed according to the guidelines of the American Academy of Sleep Medicine (AASM ICSD-3) [88]. A detailed medical history was recorded, and a complete neurological examination performed to verify the subjects' suitability. Inclusion criteria were age above 18 years, no dopaminergic therapy, and no diagnosis of Parkinson's disease, multiple system atrophy, dementia with Lewy bodies, or progressive supranuclear palsy. Exclusion criteria were smoking, antibiotic therapy in the last 24 months, history of other neurological diseases, or disorders of the gastrointestinal tract. Non-motor and autonomic symptoms were evaluated with the SCOPA-AUT [89] and PD-NMS [90] questionnaires. Motor function was evaluated with the UPDRS [91]. Additionally, patients were asked to complete the RBD-Sleep questionnaire [92]. The study conformed to the Declaration of Helsinki and was approved by the ethics committee of the Medical Faculty of the Philipps-University, Marburg, Germany (46/14).

Method details

Fecal sample collection

Fecal samples were collected at the clinics via a Feces Collection Device (MedAuxil) and collection tubes (Sarstedt), as previously described [7]. Samples were immediately flash-frozen on dry ice after collection. Samples were subsequently stored at -80°C and shipped on dry ice.

Sample exclusions

The initial set of samples consisted of 50 PD, 30 iRBD, and 50 healthy control subjects (HC). Three PD and two iRBD cases were subsequently excluded for clinical reasons (adjusted diagnosis), one iRBD, and one PD subject

for logistical reasons, and one control due to a combination of microbiome-altering medications (metformin, antidepressants, statins, and proton pump inhibitors). Additional samples were excluded due to missing values (metabolomics) or a low amount of identified analytes (metaproteomics), leading to the final numbers of samples summarized below:

- Metagenomics (MG) and metatranscriptomics (MT): 49 HC, 27 iRBD, 46 PD
- Metaproteomics (MP): 42 HC, 22 iRBD, 40 PD
- Meta-metabolomics: 49 HC, 27 iRBD, 41 PD

Metagenomic and metatranscriptomic sequencing

Extractions from fecal samples were performed according to a previously published protocol [93], conducted on a customized robotic system (Tecan Freedom EVO 200). After extraction, DNA and RNA were purified prior the sequencing analysis by using the following commercial kits respectively: Zymo DNA Clean&Concentrator-5 (D4014) and Zymo RNA Clean&Concentrator-5 (R1014). RNA quality was assessed and Quantified with an Agilent 2100 Bioanalyser (Agilent Technologies) and the Agilent RNA 6000 Nano kit, and genomic DNA and RNA fractions with a NanoDrop Spectrophotometer 1000 (Thermo Scientific) as well as commercial kits from Qubit (Qubit ds DNA BR Assay kit, Q32850; Qubit RNA BR Assay kit, Q10210). All DNA samples were subjected to random shotgun sequencing. Following DNA isolation, 200–300 ng of DNA was sheared using a Bioruptor NGS (Diagenode) with 30 s ON and 30 s OFF for 20 cycles. Sequencing libraries were prepared using the TruSeq Nano DNA library preparation kit (Illumina) following the manufacturer's protocol, with 350 bp average insert size. For MT, 1 μg of isolated RNA was rRNA-depleted using the RiboZero kit (Illumina, MRZB12424). Library preparation was performed using the TruSeq Stranded mRNA library preparation kit (Illumina) following the manufacturer's protocol, apart from omitting the initial steps for mRNA pull down. MG and MT analyses, the qualities of the libraries were checked using a Bioanalyzer (Agilent) and quantified using Qubit (Invitrogen). Libraries were sequenced on an Illumina NextSeq500 instrument with 2×150 bp read length.

Metaproteomics

Twenty microliters protein extract was processed using the paramagnetic bead approach with SP3 carboxylate coated beads [94, 95]. Briefly, the protein samples were reduced with 2 μL 25 mM DTT in 20 mM ammonium bicarbonate (Sigma-Aldrich) for 1 h at 60°C . Subsequently, 4 μL 100 mM iodoacetamide (Merck) in 20 mM

ammonium bicarbonate was added and incubated for 30 min at 37 °C in the dark. Next, 5 µL of 10% formic acid was added as well as 70 µL 100% acetonitrile (ACN) to reach a final organic content higher than 50% (v/v). Two microliters SP3 beads per sample were washed with water three times with subsequent addition of the sample. After protein binding to the beads, the supernatant was discarded. The beads were washed twice with 200 µL 70% (v/v) ethanol, and once with 200 µL ACN. The protein lysates were proteolytically cleaved using trypsin (1:50) over night at 37 °C. Since trypsin is added in aqueous solution to the samples, the proteins are not bound to the beads during enzymatic cleavage. ACN was added to each sample to reach a final organic content higher than 95% (v/v). After peptide binding to the beads, the samples were washed with pure ACN on the magnetic rack. Finally, the peptides were eluted in two steps: first, with 200 µL 87% ACN (v/v) containing 10 mM ammonium formate (pH 10), and next with two times adding 50 µL water containing 2% (v/v) DMSO and combination of the two aqueous supernatants. Thus, two fractions of peptides were generated, which were evaporated and redissolved in water containing 0.1% formic acid (20 µL) and analyzed on a Q Exactive HF instrument (Thermo Fisher Scientific) equipped with a TriVersa NanoMate source (Advion) in LC chip coupling mode. Peptide lysates were injected on a trapping column (Acclaim PepMap 100 C18, 3 µm, nanoViper, 75 µm × 2 cm, Thermo Fisher Scientific) with 5 µL/min by using 98% water/2% ACN 0.5% trifluoroacetic acid, and separated on an analytical column (Acclaim PepMap 100 C18, 3 µm, nanoViper, 75 µm × 25 cm, Thermo Fisher Scientific) with a flow rate of 300 nL/min. Mobile phase was 0.1% formic acid in water (A) and 80% ACN/0.08% formic acid in water (B). Full MS spectra (350–1550 m/z) were acquired in the Orbitrap at a resolution of 120,000 with automatic gain control (AGC) target value of 3×10^6 ions.

Meta-metabolomics

Untargeted GC–MS as well as targeted measurements (SCFA GC–MS/MS and bile acids LC–MS/MS) from fecal samples were performed according to a previously published protocol [96]. All GC–MS chromatograms were processed using MetaboliteDetector, v3.220190704 [69], while LC–MS chromatogram were acquired with Thermo Xcalibur software (version 4.1.31.9) and analyzed with TaceFinder (Version 4.1). Compounds were initially annotated by retention time and mass spectrum using an in-house mass spectral library. Internal standards were added at the same concentration to every medium sample to correct for uncontrolled sample losses and analyte degradation during metabolite extraction. The data was normalized by using the response ratio of the integrated

peak area of the analyte and the integrated peak area of the internal standard.

Sequencing data processing and analysis

For all samples, MG and MT sequencing data were processed and hybrid-assembled using the Integrated Meta-omic Pipeline (IMP) [34] (<https://git-r3lab.uni.lu/IMP/imp3>, commit 8c1bd6fa443d064511909c9eede-20703f45e6c69). It includes steps for the trimming and quality filtering of the reads, the filtering of rRNA from the MT data, and the removal of human reads after mapping against the human genome (hg38). Pre-processed MG and MT reads were assembled using the IMP-based iterative hybrid-assembly pipeline using MEGAHIT [97] 1.0.3. After assembly, the prediction and annotation of structural features such as open-reading frames (ORFs) was performed using a modified version of Prokka [70] and followed by functional annotation of those using Mantis [35]. Structural features were quantified on MG and MT level using featureCounts [72]. Taxonomic annotation of reads and contigs was performed using Kraken2 [98] with a GTDB release207 database (http://ftp.tue.mpg.de/ebio/projects/struo2/GTDB_release207/kraken2) and a 0.5 confidence threshold. Additionally, taxon abundances were estimated using mOTUs 2.5.1 [73]. The mOTU abundances were used to generate abundance matrices for each taxonomic rank (phylum, class, order, family, genus, and species) by summing up taxon marker read counts at the respective levels.

Metaproteomics prediction and annotation

For each sample, the predicted proteins were concatenated with a cRAP database of contaminants and the human UniProtKB Reference Proteome prior to the MP search. In addition, reversed sequences of all protein entries were added to the databases for the estimation of false discovery rates. The search was performed using Sipsos v1.1 [75] as search engine with the following parameters: trypsin was used as the digestion enzyme and a maximum of two missed cleavages was allowed. The tolerance levels for matching to the database were 1 Da for MS1 and 0.01 Da for MS2. Peptides with large errors for parent ions were later filtered out by setting the Filter Mass Tolerance Parent Ion parameter to 0.05 Da. Carbamidomethylation of cysteine residues was set as a fixed modification and oxidation of methionines was allowed as a variable modification. Peptides with length between 7 and 60 amino acids, with a charge state composed between + 2 and + 4, and a maximum missed cleavages of 3 were considered for identification. The results from all identifications were filtered by Sipsos using at least one unique peptide per protein, and peptide false

discovery rate (FDR) was dynamically set to achieve a 1% of protein FDR.

Data analysis was performed on all samples with at least 2000 proteins identified. A summary matrix of all selected samples consisting of the KO annotations from the integrated MG and MT analysis and the spectral count from the MP identification was then generated and used for statistical analysis.

Quantification and statistical analysis

Dimensionality reduction and ordination

Beta diversity for MG and MT was assessed using the Bray–Curtis dissimilarity and subjected to a Non-Metric MultiDimensional Scaling (NMDS) for both the taxonomic and functional levels, using the *metaMDS()* function from the *vegan* package (2.6.2). Principal component analysis (PCA) was performed for MP and MM using the *rda()* function in the *vegan* package (2.6.4). PERMANOVA was used to assess statistical differences between groups using the Bray–Curtis dissimilarity and conducted in the *vegan* R package with the *adonis2()* function.

Differential abundance analysis and correlations

Differential abundance was done in two different approaches. The first approach consisted in using SIAMCAT [76] and ALDEx2 [83] algorithm to find all the taxa and genes differentially expressed between groups without prior assumption. We used two different algorithms to have a sensitive algorithm (SIAMCAT, less prone to have false negative) and a more conservative one (ALDEx2, less prone to have false positive). The second approach consisted in using the MM significant compounds to drive the analysis on the functional level for MG and MT. Therefore, differential abundance tests and multiple correlation tests were conducted with a classical approach. We used Mann–Whitney or Kruskal–Wallis followed by a Dunn test (depending on the number of groups) and Spearman correlation tests. We applied FDR correction using the Benjamini and Hochberg method [99]. We depicted both FDR corrected as *q*-values and non-FDR corrected *p*-values to represent most of the differences found in our datasets. All statistical tests were done using the *rstatix* package (0.7.2).

Variance analysis

Variance analysis was used to assess the importance of each clinical factor on MM. To verify the covariance of factors and to assess which factors explained the most variance in our datasets, we computed the total variance

for each clinical factor (removing the NAs for each factor) and the variance explained by each group within a clinical factor. Explained variance was calculated as follows: $var.explained = \frac{1 - variance.group}{variance.total}$.

Microbiome-wide metabolic network analysis

The microbiome-wide metabolic network analysis was conducted by establishing an association between KEGG KOs and corresponding ChEBI IDs. The networks were visualized utilizing the NetworkX package (release 3.3) [76]. In this network, the nodes were represented by KEGG KOs, while the edges were denoted by the corresponding metabolites (either products or reactants) [36]. Such compounds as water, energy Transporters, and cofactors were removed, to only consider main compounds of a given reaction. The analysis was restricted to genes that were present in a minimum of 50% of the samples. To construct metabolite-specific networks, we used KEGG KOs which have either a reactant or product in KEGG. Glutamate-, thymine-, glycerol-, serine-, alanine-, and glucuronate-specific subnetworks were composed of 146, 9, 66, 43, 70, and 18 genes, respectively. The network topology metric “Betweenness centrality” was used to underscore the importance of a metabolite in microbiome-wide metabolism [36].

Integrated multi-omics analysis using MOFA2

Integrative analysis for the seven omics layers was conducted with the Multi-Omics Factor Analysis (MOFA) 2 R package (version 1.10.0) [100]. Before the analysis, data were preprocessed as follows: (I) funMG, funMT, funMP, taxMG, taxMT, and taxMP data were filtered based on the number of non-zero features; a feature was kept if it was present in at least in 25% of samples in each group (PD, HC, iRBD) or at least in 75% in any of the groups; (II) funMG, funMT, taxMG, and taxMT count data were separately residualized in a linear model to remove variance explained by differences in sequencing depth; (III) funMP and taxMP data were residualized by the sum of protein counts per sample and information on the number of high-quality proteins recovered per sample; (IV) regression residuals were cubic-root transformed to account for heteroscedasticity; (V) MM data were Transformed using a centered log-ratio Transformation. Each dataset was then additionally filtered to retain the features with the largest variance for the subsequent analysis. For funMG and funMT, we included features with variances equal to or larger than 90% feature variance for a dataset; for the other datasets, we included features with variance equal to or larger than the median feature variance for a given dataset. In the results, the feature size

for omics layers was as follows: 759 for funMG, 657 for funMT, 410 for funMP, 109 for taxMG, 71 for taxMT, 115 for taxMP, and 34 for MM. MOFA analysis was run on scaled omics data with fifteen initial factors. All factors that explained less than 2% of the variance were excluded from the model. The remaining factors were tested for differential abundance between the groups studied using the linear regression followed by ANOVA type II controlling the participants' sex, age, and recruitment cohort.

Extended information

Multi-omics data overview

Using our previously developed methodological framework [104, 105], we performed a systematic multi-omic analysis of DNA, RNA, protein, and metabolite fractions isolated from flash-frozen fecal samples. We used MG, MT, MP and MM data to find biomarkers associated with the PD phenotype (Fig. 1). We generated a mean of 7.5 (std 1.7) Gbps and 7.5 (std 1.4) Gbps of sequencing data for MG and MT, respectively. After trimming and filtering, we retained a mean of 6.8 (std 1.7) Gbps and 3.2 (std 1.3) Gbps for MG and MT, respectively. The mean assembly size was 0.4 (std 0.1) Gbps, with on average 5.9×10^5 (std 1.7×10^5) genes predicted. Finally, protein databases contained a mean of 7.2×10^5 (std 1.8×10^5) proteins, an average of 4.1×10^4 (std 0.6×10^4) MS spectra per sample were acquired, and a mean of 3.4×10^3 (std 1.7×10^3) proteins were identified.

MOFA model description

MOFA is an unsupervised machine learning approach for the integration of multi-omics data sets [100]. It allows for the identification of highly informative features across multiple omics. It has previously been used in the study of the gut microbiome in several diseases, giving critical insights into the link between the gut microbiome, health, and disease [101–103]. The biggest proportion of variance was explained by funMG and funMT, followed by the taxMT and funMP datasets (Extended Fig. 6A). F1-2 incorporated most of the variance related to the funMT and taxMT, whereas funMG and taxMG variance was predominantly covered by F3-F5 (Extended Fig. 6B). The funMP variance was explained mostly by F6, and MM variance was explained by F1 and F6. MOFA factors were tested in a linear model followed by ANOVA with disease status, as well as confounders including patients' sex, age, and recruitment cohort. Among the MOFA factors, F1 showed an association with the disease status, whereas F4 and F9 were associated with patients' sex and recruitment cohort, respectively (Fig. 6A).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-025-02227-2>.

Supplementary Material 1: Extended figure 1. NMDS analysis of A. metagenomic taxonomic composition (taxMG), B. metagenomic functions (funMG) and C. meta-proteomic taxonomic composition (taxMP), using a Bray-Curtis dissimilarity matrix. D. PCA analysis of metaproteomic functions (funMP). E. PERMANOVA analysis for the three groups and all omics. Colour represents R^2 values and size is $-\log_{10}(p\text{-value})$. All PERMANOVA analysis were run using 1000 permutations using a Bray-Curtis dissimilarity matrix.

Supplementary Material 2: Extended figure 2. A. Differential abundance analysis at the genus level using SIAMCAT algorithm. B. Differential abundance analysis using ALDEx2 algorithm at the species level. Values are pseudo fold changes for HC/PD and size is based on $-\log_{10}(p\text{-value})$. Shape is referring to level of significance, triangular shape for p-value significance before and round shape for p-values < 0.05 after FDR correction. C to F ALDEx2 differential abundance analysis on funMG for HC vs PD (C) and HC vs iRBD (D); funMT for HC vs PD (E) and HC vs iRBD (F). All genes and transcripts are colored and shaped according to p-value significance before (triangle shape) and after FDR correction (round shape).

Supplementary Material 3: Extended figure 3. A. Percentage of variance explained for each metabolite for "Sex" and "Diagnosis" and B. Percentage of variance explained for each metabolite for "Constipation" and "Diagnosis" (right panel). Metabolites include untargeted meta-metabolomics, targeted SCFA and targeted bile acids, normalized by sum before merging and variance quantification. C. Variance for each metabolite associated to the clinical factors "Diagnosis", "Sex" and "Constipation". Metabolites include untargeted meta-metabolomics, targeted SCFA and targeted bile acids, normalized by sum before merging and variance quantification. D. Spearman correlation between metabolites and taxMG species. P-values are FDR corrected.

Supplementary Material 4: Extended figure 4. A. Absolute log2 fold change between HC and iRBD for funMG and funMT associated to significant compounds. Dots are scaled by the $-\log_{10}(p\text{-value})$, colored and shaped according to p-value significance before (triangle shape) and after FDR correction (round shape). B. Spearman correlation between beta-glutamate relative abundance and funMG-funMT KEGG orthologs related to glutamate species. Only genes with at least one significant correlation are plotted. All p-values are FDR corrected.

Supplementary Material 5: Extended figure 5. A. Bile acids transcripts found significantly different between the groups. P-values are corrected with FDR. B. Flagellar assembly transcripts encoding for extracellular component of the flagella for the genus present in Cluster 2. All tests are Wilcoxon tests.

Supplementary Material 6: Extended figure 6. Multiomics variance explained by MOFA factors. A. Variance explained by the MOFA factors across different omics layers, total. B. Variance explained by the MOFA factors across different omics layers, split by factors.

Supplementary Material 7.

Acknowledgements

We thank the staff of the Luxembourg Centre for Systems Biomedicine (LCSB), particularly the sequencing platform and the metabolomics platform, for running the sequencing and metabolomic analyses. The bioinformatics presented in this paper were carried out using the HPC facilities of the University of Luxembourg [106].

Rights retention statement

This research was funded in whole, or in part, by the Luxembourg National Research Fund (FNR), grant references CORE/16/BM/11333923 (MiBiPa), CORE/15/BM/10404093 (microCancer/MUST), PRIDE/11823097 (MICROH DTU), and CORE/19/BM/13684739 (metaPUF). For the purpose of open access, and in fulfillment of the obligations arising from the grant agreement, the author has applied a Creative Commons Attribution 4.0 International (CC

BY 4.0) license to any Author Accepted Manuscript version arising from this submission.

Authors' contributions

Conceptualization: J.-P.T., A.H.-B., W.O., B.M., P.W. Patient recruitment, clinical coordination, and sampling: S.S., A.J., C.T., W.O., B.M. Multi-omic data generation: J.-P.T., C.J., L.A.L., C.D.R., A.D., N.J., M.v.B. Bioinformatics, statistics and data visualization: R.V., J.O.S., P.N., V.T.E.A., V.P., O.H., B.K., P.M. Initial manuscript draft: R.V., J.O.S., P.N., V.T.E.A., V.P. Review and editing: C.C.L., S.B.B., P.M. and P.W. Funding acquisition: C.C.L., W.O., B.M., P.W. All authors read and approved of the submitted version.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 863664), and was further supported by the Luxembourg National Research Fund (FNR) CORE/16/BM/11333923 (MiBiPa), CORE/15/BM/10404093 (microCancer/MUST), PRIDE/11823097 (MICROH DTU), and CORE/19/BM/13684739 (metaPUF), the Michael J. Fox Foundation under grant IDs MJFF-14701 (MiBiPa-PLUS) and MJFF-019228 (PARKdiet), the Parkinson's Foundation (MiBiPa Saliva), the Institute for Advanced Studies of the University of Luxembourg through an AUDACITY grant (ref. no. MCI-BIOME_2019), and an "Espoir en tête" grant from the Rotary Club Luxembourg to P.W. This work was also supported by a Fulbright Research Scholarship from the Commission for Educational Exchange between the USA, Belgium, and Luxembourg to P.W. Additional funding was provided by the FNR under INTERMOBILITY/23/17856242. The MiBiPa project was co-funded by the German Research Foundation (DFG) under grant agreement MO 2088/5-1 to B.M. C.D.R. received funding from the European Union's Horizon 2020 Widening Fellowships (N° 101038088). The FNR supported PM as part of the National Center of Excellence in Research on Parkinson's disease (NCER-PD, FNR11264123) and the DFG Research Unit FOR2488 (INTER/DFG/19/14429377).

Data availability

Metagenomic and metatranscriptomic data are available on SRA, under the BioProject: PRJNA782492. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD031457 and <https://doi.org/10.6019/PXD031457>, with the following access details: **Username**: reviewer_pxd031457@ebi.ac.uk; **Password**: 63ZkH87p. Metadata and metabolomics data are available at <https://owncloud.lcsb.uni.lu/s/QWCbDhbpnpYLlSR>, with the following password: 7_drt4D;ZN=H"pC. The IMP pipeline, which was used for analysis of metagenomic and metatranscriptomic data, is available at <https://gitlab.lcsb.uni.lu/IMP/imp3>. The R and python code used for statistical analyses and visualizations is available at [https://gitlab.lcsb.uni.lu/ESB/\[TBA\]](https://gitlab.lcsb.uni.lu/ESB/[TBA]).

Declarations

Ethics approval and consent to participate

The study conformed to the Declaration of Helsinki and was approved by the ethics committee of the Physician's Board Hessen, Germany (FF 89/2008). The DeNoPa trial is registered at the German Register for Clinical trials (DRKS00000540).

Consent for publication

Not applicable.

Competing interest

The authors declare no competing interests.

Author details

¹Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-Sur-Alzette, Luxembourg. ²UK Centre for Ecology and Hydrology, Wallingford, Oxfordshire, UK. ³Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, The Netherlands. ⁴Department of Neurology, University Medical Center Göttingen, Göttingen, Germany. ⁵Paracelsus-Elena-Klinik, Kassel, Germany. ⁶Department of Neurology, Philipps-University Marburg, Marburg, Germany. ⁷Department of Molecular Toxicology, Helmholtz-Centre for Environmental Research GmbH – UFZ, Leipzig, Germany. ⁸Department of Neurosurgery, University Medical Center

Göttingen, Göttingen, Germany. ⁹Department of Life Sciences and Medicine, Faculty of Science, Technology and Medicine, University of Luxembourg, Belvaux, Luxembourg.

Received: 3 February 2025 Accepted: 11 August 2025

Published online: 07 October 2025

References

- Elbaz A, Carcaillon L, Kab S, Moisan F. Epidemiology of Parkinson's disease. *Rev Neurol (Paris)*. 2016;172:14–26.
- Schwartz A, et al. Fecal markers of intestinal inflammation and intestinal permeability are elevated in Parkinson's disease. *Parkinsonism Relat Disord*. 2018;50:104–7.
- Clairembault T, et al. Structural alterations of the intestinal epithelial barrier in Parkinson's disease. *Acta Neuropathol Commun*. 2015;3:12.
- Scd van Ij, PD. The intestinal barrier in Parkinson's disease: current state of knowledge. *Journal of Parkinson's disease*. 2019;9.
- Palacios N, et al. Metagenomics of the gut microbiome in Parkinson's disease: prodromal changes. *Ann Neurol*. 2023;94:486–501.
- Aho VTE, et al. Relationships of gut microbiota, short-chain fatty acids, inflammation, and the gut barrier in Parkinson's disease. *Mol Neurodegener*. 2021;16:6.
- Heintz-Buschart A, et al. The nasal and gut microbiome in Parkinson's disease and idiopathic rapid eye movement sleep behavior disorder. *Mov Disord*. 2018;33:88–98.
- Cirstea MS, et al. Microbiota composition and metabolism are associated with gut function in Parkinson's disease. *Mov Disord*. 2020;35:1208–17.
- Boktor JC, et al. Integrated multi-cohort analysis of the Parkinson's disease gut metagenome. *Mov Disord*. 2023;38:399–409.
- Huang B, et al. Gut microbiome dysbiosis across early Parkinson's disease, REM sleep behavior disorder and their first-degree relatives. *Nat Commun*. 2023;14:2501.
- Wallen ZD, et al. Metagenomics of Parkinson's disease implicates the gut microbiome in multiple disease mechanisms. *Nat Commun*. 2022;13:6958.
- Romano, S. et al. Meta-analysis of the Parkinson's disease gut microbiome suggests alterations linked to intestinal inflammation. *npj Parkinsons Dis*. 2021;7:1–13.
- Toh TS, et al. Gut microbiome in Parkinson's disease: new insights from meta-analysis. *Parkinsonism Relat Disord*. 2022;94:1–9.
- Howell MJ, Schenck CH. Rapid eye movement sleep behavior disorder and neurodegenerative disease. *JAMA Neurol*. 2015;72:707–12.
- Unger MM, et al. Short chain fatty acids and gut microbiota differ between patients with Parkinson's disease and age-matched controls. *Parkinsonism Relat Disord*. 2016;32:66–72.
- Chen S-J, Lin C-H. Gut microenvironmental changes as a potential trigger in Parkinson's disease through the gut–brain axis. *J Biomed Sci*. 2022;29:54.
- Hirayama M, Tsunoda M, Yamamoto M, Tsuda T, Ohno K. Serum tyrosine-to-phenylalanine ratio is low in Parkinson's disease. *J Parkinsons Dis*. 2016;6:423–31.
- Figura M, et al. Serum amino acid profile in patients with Parkinson's disease. *PLoS ONE*. 2018;13: e0191670.
- Vascellari S, et al. Gut microbiota and metabolome alterations associated with Parkinson's disease. *mSystems*. 2020. <https://doi.org/10.1128/mSystems.00561-20>.
- Li P, et al. Gut microbiota dysbiosis is associated with elevated bile acids in Parkinson's disease. *Metabolites*. 2021;11:29.
- Hertel J, et al. Integrated analyses of microbiome and longitudinal metabolome data reveal microbial-host interactions on sulfur metabolism in Parkinson's disease. *Cell Rep*. 2019;29:1767–1777.e8.
- Hang S, et al. Bile acid metabolites control T H 17 and T reg cell differentiation. *Nature*. 2019;576:143–8.
- Song X, et al. Microbial bile acid metabolites modulate gut RORγ + regulatory T cell homeostasis. *Nature*. 2020;577:410–5.
- Iovino L, Tremblay ME, Civerio L. Glutamate-induced excitotoxicity in Parkinson's disease: the role of glial cells. *J Pharmacol Sci*. 2020;144:151–64.

25. Iwasaki Y, Ikeda K, Shiojima T, Kinoshita M. Increased plasma concentrations of aspartate, glutamate and glycine in Parkinson's disease. *Neurosci Lett*. 1992;145:175–7.
26. Jiménez-Jiménez FJ, et al. Neurotransmitter amino acids in cerebrospinal fluid of patients with Parkinson's disease. *J Neurol Sci*. 1996;141:39–44.
27. Mally J, Szalai G, Stone TW. Changes in the concentration of amino acids in serum and cerebrospinal fluid of patients with Parkinson's disease. *J Neurol Sci*. 1997;151:159–62.
28. Yuan Y-S, et al. Change in plasma levels of amino acid neurotransmitters and its correlation with clinical heterogeneity in early Parkinson's disease patients. *CNS Neurosci Ther*. 2013;19:889–96.
29. Wang W, et al. Interactions between gut microbiota and Parkinson's disease: the role of microbiota-derived amino acid metabolism. *Front Aging Neurosci*. 2022;14: 976316.
30. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry*. 1992;55:181–4.
31. Beheshti I, Booth S, Ko JH. Differences in brain aging between sexes in Parkinson's disease. *NPJ Parkinsons Dis*. 2024;10:1–16.
32. Li X, et al. Sex differences in rapid eye movement sleep behavior disorder: a systematic review and meta-analysis. *Sleep Med Rev*. 2023;71: 101810.
33. Mertsalmi TH, et al. More than constipation – bowel symptoms in Parkinson's disease and their connection to gut microbiota. *Eur J Neurol*. 2017;24:1375–83.
34. Narayanasamy S, et al. IMP: a pipeline for reproducible reference-independent integrated metagenomic and metatranscriptomic analyses. *Genome Biol*. 2016;17:260.
35. Queirós P, Delogu F, Hickl O, May P, Wilmes P. Mantis: flexible and consensus-driven genome annotation. *Gigascience*. 2021;10: giab042.
36. Roume, H. et al. Comparative integrated omics: identification of key functionalities in microbial community-wide metabolic networks. *npj Biofilms Microbiomes* 1, 1–11 (2015).
37. Villette R, et al. Unraveling host-gut microbiota dialogue and its impact on cholesterol levels. *Front Pharmacol*. 2020. <https://doi.org/10.3389/fphar.2020.00278>.
38. Hurley MJ, Bates R, Macnaughtan J, Schapira AHV. Bile acids and neurological disease. *Pharmacol Ther*. 2022;240: 108311.
39. Kalecký K, Bottiglieri T. Targeted metabolomic analysis in Parkinson's disease brain frontal cortex and putamen with relation to cognitive impairment. *NPJ Parkinsons Dis*. 2023;9:1–16.
40. 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*. 2022;98:e848–58.
41. Baert, F. et al. Parkinson's disease patients' short chain fatty acids production capacity after in vitro fecal fiber fermentation. *npj Parkinsons Dis*. 2021;7:1–14.
42. Tan S, et al. Interaction between the gut microbiota and colonic enteroendocrine cells regulates host metabolism. *Nat Metab*. 2024;6:1076–91.
43. Qian F, et al. *Salmonella* flagellin is a potent carrier–adjuvant for peptide conjugate to induce peptide-specific antibody response in mice. *Vaccine*. 2015;33:2038–44.
44. Gram AM, et al. *Salmonella* flagellin activates NAIP/NLRC4 and canonical NLRP3 inflammasomes in human macrophages. *J Immunol*. 2021;206:631–40.
45. Tran HQ, Ley RE, Gewirtz AT, Chassaing B. Flagellin-elicited adaptive immunity suppresses flagellated microbiota and vaccines against chronic inflammatory diseases. *Nat Commun*. 2019;10:5650.
46. Gewirtz AT, Navas TA, Lyons S, Godowski PJ, Madara JL. Cutting edge: bacterial flagellin activates basolaterally expressed TLR5 to induce epithelial proinflammatory gene expression. *J Immunol*. 2001;167:1882–5.
47. Miao EA, Andersen-Nissen E, Warren SE, Aderem A. TLR5 and Ipaf: dual sensors of bacterial flagellin in the innate immune system. *Semin Immunopathol*. 2007;29:275–88.
48. Zeng H, et al. Flagellin/TLR5 responses in epithelia reveal intertwined activation of inflammatory and apoptotic pathways. *Am J Physiol Gastrointest Liver Physiol*. 2006;290:G96–108.
49. Cullender TC, et al. Innate and adaptive immunity interact to quench microbiome flagellar motility in the gut. *Cell Host Microbe*. 2013;14:571–81.
50. Clasen SJ, et al. Silent recognition of flagellins from human gut commensal bacteria by Toll-like receptor 5. *Sci Immunol*. 2023;8:eabq7001.
51. Quan Y, et al. *Roseburia intestinalis*-derived flagellin is a negative regulator of intestinal inflammation. *Biochem Biophys Res Commun*. 2018;501:791–9.
52. Shen Z, et al. *Roseburia intestinalis* stimulates TLR5-dependent intestinal immunity against Crohn's disease. *EBioMedicine*. 2022. <https://doi.org/10.1016/j.ebiom.2022.104285>.
53. Wu X, et al. *Roseburia intestinalis*-derived flagellin ameliorates colitis by targeting miR-223-3p-mediated activation of NLRP3 inflammasome and pyroptosis. *Mol Med Rep*. 2020;22:2695–704.
54. Liu J, et al. Microbiota-microglia crosstalk between *Blautia producta* and neuroinflammation of Parkinson's disease: a bench-to-bedside translational approach. *Brain Behav Immun*. 2024;117:270–82.
55. Tian J, et al. Specific immune status in Parkinson's disease at different ages of onset. *NPJ Parkinsons Dis*. 2022;8:5.
56. Tansey MG, et al. Inflammation and immune dysfunction in Parkinson disease. *Nat Rev Immunol*. 2022;22:657–73.
57. Ho MS. Microglia in Parkinson's disease. *Adv Exp Med Biol*. 2019;1175:335–53.
58. Scheiblich H, et al. Microglial NLRP3 inflammasome activation upon TLR2 and TLR5 ligation by distinct α -synuclein assemblies. *J Immunol*. 2021;207:2143–54.
59. Bostick JW, Schonhoff AM, Mazmanian SK. Gut microbiome-mediated regulation of neuroinflammation. *Curr Opin Immunol*. 2022;76:102177.
60. Gopalakrishnan V, et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science*. 2018;359:97–103.
61. Gallucci S, Matzinger P. Danger signals: SOS to the immune system. *Curr Opin Immunol*. 2001;13:114–9.
62. Mikami A, et al. Oral administration of *Flavonifractor plautii*, a bacteria increased with green tea consumption, promotes recovery from acute colitis in mice via suppression of IL-17. *Front Nutr*. 2020;7: 610946.
63. Vatanen T, et al. Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans. *Cell*. 2016;165:842–53.
64. Parks DH, et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res*. 2022;50:D785–94.
65. Quast C, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41:D590–6.
66. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30.
67. Schneider M, Bairoch A, Wu CH, Apweiler R. Plant protein annotation in the UniProt knowledgebase. *Plant Physiol*. 2005;138:59–66.
68. Hiller K, et al. MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. *Anal Chem*. 2009;81:3429–39.
69. Li D, et al. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*. 2016;102:3–11.
70. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9.
71. Liao Y, Smyth GK, Shi W. Featurecounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923–30.
72. Milanese A, et al. Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat Commun*. 2019;10:1014.
73. Davis MPA, van Dongen S, Abreu-Goodger C, Bartonicek N, Enright AJ. Kraken: a set of tools for quality control and analysis of high-throughput sequence data. *Methods*. 2013;63:41–9.
74. Guo X, et al. Sips ensemble improves database searching and filtering for complex metaproteomics. *Bioinformatics*. 2018;34:795–802.
75. Wirbel J, et al. Microbiome meta-analysis and cross-disease comparison enabled by the SIAMCAT machine learning toolbox. *Genome Biol*. 2021;22:1–27.
76. Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. 2008.

77. Wickham H. ggplot2: elegant graphics for data analysis. Springer-Verlag New York. 2016. ISBN 978-3-319-24277-4, <http://ggplot2.org>.
78. Oksanen J, Blanchet FG, Kindt R, et al. : vegan: community ecology package.R package version. 2016;2:3–5.
79. Kassambara, A. Pipe-friendly framework for basic statistical tests [R package rstatix version 0.6.0]. in 2020.
80. Gu Z. Complex heatmap visualization. iMeta. 2022;1: e43.
81. Kassambara, A. Ggpubr: 'ggplot2' based publication ready plots. 2023.
82. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
83. Fernandes AD, Macklaim JM, Linn TG, Reid G, Gloor GB. Anova-like differential expression (ALDEx) analysis for mixed population RNA-seq. *PLoS ONE.* 2013;8: e67019.
84. Argelaguet R, et al. Multi-omics factor analysis—a framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol.* 2018;14: e8124.
85. Mollenhauer B, et al. Nonmotor and diagnostic findings in subjects with de novo Parkinson disease of the DeNoPa cohort. *Neurology.* 2013;81:1226–34.
86. Mollenhauer B, et al. Monitoring of 30 marker candidates in early Parkinson disease as progression markers. *Neurology.* 2016;87:168–77.
87. Schenck CH, et al. Rapid eye movement sleep behavior disorder: devising controlled active treatment studies for symptomatic and neuroprotective therapy—a consensus statement from the International Rapid Eye Movement Sleep Behavior Disorder Study Group. *Sleep Med.* 2013;14:795–806.
88. Sateia MJ. International classification of sleep disorders-third edition. *Chest.* 2014;146:1387–94.
89. Visser M, Marinus J, Stiggelbout AM, Van Hilten JJ. Assessment of autonomic dysfunction in Parkinson's disease: the SCOPA-AUT. *Mov Disord.* 2004;19:1306–12.
90. Chaudhuri KR, et al. The metric properties of a novel non-motor symptoms scale for Parkinson's disease: results from an international pilot study. *Mov Disord.* 2007;22:1901–11.
91. Fahn S, Elton R. Members of the UPDRS Development Committee. Unified Parkinson's disease rating scale. Recent developments in Parkinson's disease. 1987;153–163.
92. Stiasny-Kolster K, et al. The REM sleep behavior disorder screening questionnaire—a new diagnostic instrument. *Mov Disord.* 2007. <https://doi.org/10.1002/mds.21740>.
93. Roume H, et al. A biomolecular isolation framework for eco-systems biology. *ISME J.* 2013;7:110–21.
94. Hughes CS, et al. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc.* 2019;14:68–85.
95. Bannuscher A, et al. A multi-omics approach reveals mechanisms of nanomaterial toxicity and structure–activity relationships in alveolar macrophages. *Nanotoxicology.* 2020;14:181–95.
96. De Saedeleer B, et al. Systematic characterization of human gut microbiome-secreted molecules by integrated multi-omics. *ISME Commun.* 2021;1:1–6.
97. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de bruijn graph. *Bioinformatics.* 2015;31:1674–6.
98. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with kraken 2. *Genome Biol.* 2019;20:257.
99. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc: Ser B (Methodol).* 1995;57:289–300.
100. Argelaguet, R. et al. Multi-omics factor analysis—a framework for unsupervised integration of multi-omics data sets | *Molecular Systems Biology*. EMBOpress <https://www.embopress.org/doi/full/https://doi.org/10.15252/msb.20178124> (2018).
101. Haak BW, et al. Integrative transkingdom analysis of the gut microbiome in antibiotic perturbation and critical illness. *mSystems.* 2021. <https://doi.org/10.1128/msystems.01148-20>.
102. Al Bataineh MT, et al. Uncovering the relationship between gut microbial dysbiosis, metabolomics, and dietary intake in type 2 diabetes mellitus and in healthy volunteers: a multi-omics analysis. *Sci Rep.* 2023;13: 17943.
103. Arkan, M. et al. Metaproteogenomic analysis of saliva samples from Parkinson's disease patients with cognitive impairment. *npj Biofilms Microbiomes.* 2023;9:1–10.
104. Wilmes P, Muller E, Heintz-Buschart A, Roume H, Lebrun L. The sequential isolation of metabolites, RNA, DNA, and proteins from a single, undivided mixed microbial community sample. Preprint at. 2014. <https://doi.org/10.1038/protex.2014.051>.
105. Heintz-Buschart A, et al. Integrated multi-omics of the human gut microbiome in a case study of familial type 1 diabetes. *Nat Microbiol.* 2016;2:1–13.
106. Varrette, S., Bouvry, P., Cartiaux, H. & Georgatos, F. Management of an academic HPC cluster: the UL experience. in 2014 International Conference on High Performance Computing & Simulation (HPCS). 2014;959–967. <https://doi.org/10.1109/HPCSim.2014.6903792>.

Publisher's Note

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