

RESEARCH PAPER



Microbiome signatures of virulence in the oral-gut-brain axis influence Parkinson's disease and cognitive decline pathophysiology

Frederick Clasen^a, Suleyman Yildirim^{b,c}, Muzaffer Arıkan^{b,c,*}, Fernando Garcia-Guevara^a, Lutfü Hanoğlu^e, Nesrin H. Yılmaz^f, Aysu Şen^g, Handan Kaya Celik^h, Alagoz Aybala Neslihan^h, Tuğçe Kahraman Demirⁱ, Zeynep Temel^j, Adil Mardinoglu^{a,d}, David L. Moyes^a, Mathias Uhlen^d, and Saeed Shoaie^{ib}

^aCentre for Host-Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London, London, UK; ^bDepartment of Medical Microbiology, Istanbul Medipol University International School of Medicine, Istanbul, Türkiye; ^cRegenerative and Restorative Medicine Research Center (REMER), Research Institute for Health Sciences and Technologies (SABITA), Istanbul Medipol University, Istanbul, Türkiye; ^dScience for Life Laboratory, KTH - Royal Institute of Technology, Stockholm, Sweden; ^eNeuroscience Graduate Program and Department of Neurology, Istanbul Medipol University School of Medicine, Istanbul, Türkiye; ^fDepartment of Neurology, Istanbul Medipol University School of Medicine, Istanbul, Türkiye; ^gDepartment of Neurology, Bakırköy Research and Training Hospital for Psychiatric and Neurological Diseases, Istanbul, Türkiye; ^hDepartment of Neurology, Kocaeli University Faculty of Medicine, Kocaeli, Türkiye; ⁱDepartment of Electroneurophysiology, Vocational School, Biruni University, Istanbul, Türkiye; ^jDepartment of Psychology, Faculty of Humanities and Social Sciences, Fatih Sultan Mehmet Vakıf University, Istanbul, Türkiye

ABSTRACT

The human microbiome is increasingly recognized for its crucial role in the development and progression of neurodegenerative diseases. While the gut-brain axis has been extensively studied, the contribution of the oral microbiome and gut-oral tropism in neurodegeneration has been largely overlooked. Cognitive impairment (CI) is common in neurodegenerative diseases and develops on a spectrum. In Parkinson's Disease (PD) patients, CI is one of the most common non-motor symptoms but its mechanistic development across the spectrum remains unclear, complicating early diagnosis of at-risk individuals. Here, we generated 228 shotgun metagenomics samples of the gut and oral microbiomes across PD patients with mild cognitive impairment (PD-MCI) or dementia (PDD), and a healthy cohort, to study the role of gut and oral microbiomes on CI in PD. In addition to revealing compositional and functional signatures, the role of pathobionts, and dysregulated metabolic pathways of the oral and gut microbiome in PD-MCI and PDD, we also revealed the importance of oral-gut translocation in increasing abundance of virulence factors in PD and CI. The oral-gut virulence was further integrated with saliva metaproteomics and demonstrated their potential role in dysfunction of host immunity and brain endothelial cells. Our findings highlight the significance of the oral-gut-brain axis and underscore its potential for discovering novel biomarkers for PD and CI.

ARTICLE HISTORY

Received 11 September 2024
Revised 25 March 2025
Accepted 7 May 2025

KEYWORDS

Oral-gut-brain axis;
parkinson's disease;
cognitive impairment;
virulence; microbiome

Introduction

Neurological disorders are the leading cause of physical and cognitive disability around the world, currently affecting approximately 15% of the worldwide population and expected to increase in future decades due to an aging population, industrialization and changes in environmental impacts.^{1–3} Parkinson's Disease (PD) is a complex neurodegenerative disease with the fastest growing prevalence worldwide.^{2,4} While it is primarily characterized by motor symptoms such as involuntary shaking, slow movements, and muscle stiffness,

one of its most common non-motor dysfunctions is cognitive impairment (CI). There is a high risk of dementia in patients with PD with nearly half of patients reaching the dementia stage within 10 years after diagnosis and virtually all patients develop full dementia within 20 years after diagnosis.⁵ CI develops on a spectrum that ranges from mild cognitive impairment (MCI) to full-scale dementia.^{4,6,7} Identification of the risk of developing CI and cognitive decline are important for clinical management of NDs.⁵ However, the evaluation of cognition remains challenging and

CONTACT Suleyman Yildirim  sulemanyildirim@medipol.edu.tr  Department of Medical Microbiology, Istanbul Medipol University International School of Medicine, Ekinçiler Cad. No. 19, Beykoz, Istanbul 34825, Türkiye; Saeed Shoaie  saeed.shoaie@kcl.ac.uk  Centre for Host-Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London SE1 9RT, UK

*Present address: Biotechnology Division, Department of Biology, Faculty of Science, Istanbul University, Istanbul, Türkiye.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19490976.2025.2506843>

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

there is currently an unmet need on whether patients with neurodegenerative diseases have CI or are at risk for further cognitive decline. Non-genetic factors, such as microbiome and environmental impacts, including diet, pollution, and drugs exposure, may have a significant role in this.^{8,9}

A growing body of evidence links the gastrointestinal (GI) tract with neurodegenerative diseases, including PD, and GI dysfunction is common in patients with PD.^{10,11} As such, several studies have investigated the role of the gut microbiome in PD for novel diagnostic and treatment avenues as well as a better understanding of the gut-brain axis.¹² Several studies consistently indicated an increased abundance in *Akkermansia*, *Bifidobacterium* and *Lactobacillus*, and a depletion in butyrate producers such as *Roseburia*, *Faecalibacterium* and *Blautia* in PD patients.² One of the key mediating factors of the gut microbiome composition is microbial metabolites and virulence that can have an impact on PD and CI onset and progression. This could be through induction of neuroinflammation and oxidative stress that exacerbate neurodegeneration.^{13,14} Among microbial metabolites, short chain fatty acids (SCFAs) production and especially butyrate has shown to have neuroprotective effects.¹⁵ At the same time, the secretion of bacterial endotoxins and cell components have been increasingly linked to the pathogenesis of NDs and, in particular PD.^{16,17} The presence of the lipopolysaccharides (LPS), a major component of gram-negative bacteria and indication of the gut-brain axis dysfunction, in blood can activate microglia and eventually leads to chronic neuroinflammation. LPS can also promote α -synuclein aggregation, a hallmark of PD and its progression, which could also lead to further neurodegeneration and CI.¹⁸ Release of gram-positive bacterial components such as peptidoglycan and lipoteichoic acid, could stimulate immune responses and promote the secretion of proinflammatory cytokines and contribute to the neuroinflammation.^{19,20}

Similarly, oral health of PD patients has also been shown to impact the course of disease.²¹ α -synuclein, the molecule that forms aggregates in neurons in PD, can be detected in saliva. The presence of α -synuclein in the oral cavity frequently results in reduced saliva production and difficulty

swallowing,^{22,23} and report the association of NDs with dysphagia. Oral bacteria contribute to chronic inflammation and neurodegeneration through various mechanisms. Opportunistic pathogens in the oral cavity, which proliferate due to dysbiosis within the oral ecosystem, can form biofilms leading to bacterial overgrowth.^{24–26} These biofilms, often associated with gingivitis and periodontitis,²⁷ enable bacteria to enter the bloodstream, potentially causing bacteremia and systemic inflammation.²⁸ *Porphyromonas gingivalis* is a well-studied oral pathogen and has been observed in Alzheimer's disease (AD) brains and active periodontitis have been reported to impact CI,^{29–31} and in bacteremia cases it can induce blood-brain permeability.³² The presence of inflammation, bacteremia, and dysfunction of the mucosal barriers can lead to spontaneous dissemination of bacteria across tissues.³³ Simultaneously, the use of specific drugs, such as proton pump inhibitors and antibiotics, to treat stomach reflux, gastritis, and ulcers, that are common conditions in PD patients,³⁴ modulate and accelerate microbial translocation.^{35,36} The presence of oral pathobionts and their overgrowth, exacerbates gut dysbiosis and systematic inflammation, as has been reported in several other diseases.^{37,38}

In this study we used metagenomics of feces and saliva in a cohort of PD patients that has no CI, has MCI (PD-MCI) or full dementia (PDD), together with a healthy control cohort. We hypothesize that compositional and functional differences in the microbiomes exist along the CI spectrum and that these differences, in turn, impact PD progression. We use a combination of machine learning approaches together with functional, correlative and network analyses to associate microbiome changes with CI. Through this, we aim to establish an oral-gut-brain axis in PD to bring forth a more mechanistic understanding of the human microbiome in ND.

Results

Gut and oral microbiome composition is associated with cognitive decline in Parkinson's disease

We performed shotgun metagenomics on 228 saliva and fecal samples taken from 41 Parkinson's

Disease (PD) patients with mild cognitive impairment (PD-MCI), 47 patients with full dementia (PDD) and 26 healthy controls (HC) (Figure 1(a), Supplementary Table S1). Additionally, we sequenced 20 fecal samples from PD patients

without CI. The age and gender distribution between the PD-MCI and PDD groups were similar (Table 1, Figure 1(b)), with a mean age of 67.27 years (SD = 8.75 years) and 70.89 years (SD = 7.34) for PD-MCI and PDD, respectively, and 34.15%

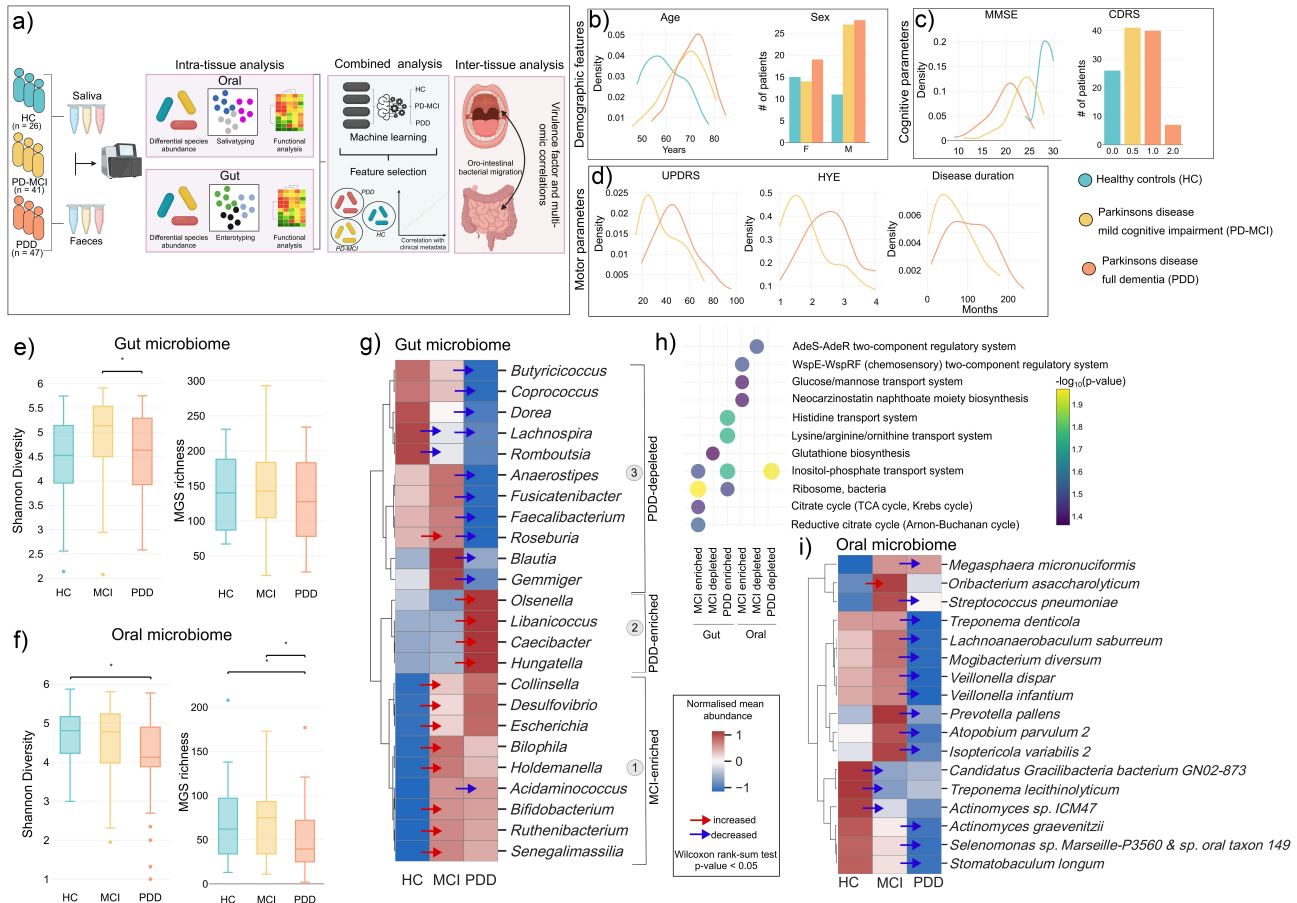


Figure 1. Gut and oral microbiome dysbiosis in PD patients with different degrees of cognitive impairment. (a) Study and methodology overview. A total of 114 individuals were included in the study. This included 41 PD patients with mild cognitive impairment (PD-MCI), 47 patients with full dementia (PDD) as well as 26 healthy controls (HC). Saliva and fecal samples were collected from all individuals and used for DNA extraction to perform shotgun metagenomics (Methods). We first performed intra-tissue analysis by investigating compositional and functional microbial changes in gut and oral separately. Thereafter, we combined gut and oral data to perform predictive modeling using machine learning. Finally, we investigated whether the translocation of oral species to the gut potentially impact disease. (b) Demographic features of study population. Distribution of age and gender for HC, PD-MCI and PDD patients. (c) Key cognitive features of study population. Distribution of scores for the mini mental state examination (MMSE) and CDRS scores. (d) Key motor parameters of study population. Distribution of UPDRS, HYE and disease duration (months). (e) Shannon diversity and MGS richness of the gut microbiome. Significance was calculated with a Wilcoxon rank-sum test with an asterisk (*) indicating p-value < 0.05. (f) Shannon diversity and MGS richness of the oral microbiome. Significance was calculated with a Wilcoxon rank-sum test with an asterisk (*) indicating p-value < 0.05. (g) Relative abundance changes of genera in the gut microbiome. MGS were mapped to their corresponding genus and differentially abundant genera were calculated using Wilcoxon rank-sum test with a p-value cutoff of 0.05. Significantly changing genera were visualized using normalized mean abundance by calculating Z-scores for each genus. Arrows indicate either a significant depletion (blue) or increase (red) of the abundance of a genus. (h) Metabolic pathway enrichment of different patient populations. The gene counts of all samples were used to map against the KEGG database to calculate genes counts for metabolic genes that were then used for enrichment analysis by first calculating differentially abundant genes using Wilcoxon rank-sum tests. Enrichment of KEGG modules were then performed using hypergeometric enrichment with a p-value cutoff of 0.05. (i) Relative abundance changes of genera in the oral microbiome. MGS were mapped to their corresponding species and differentially abundant species were calculated using Wilcoxon rank-sum test with a p-value cutoff of 0.05. Significantly changing species were visualized using normalized mean abundance by calculating Z-scores for each species. Arrows indicate either a significant depletion (blue) or increase (red) of the abundance of a species.

and 40.43% of females for PD-MCI and PDD, respectively. Cognitive assessment using the Mini-Mental State Examination (MMSE) and Clinical Dementia Rating Scale (CDRS) revealed a significant difference between PD-MCI and PDD patients (Table 1, Figure 1(c)). Additionally, motor function parameters, including the Unified Parkinson's Disease Rating Scale (UPDRS), Hoehn and Yahr Scale (HYE), and disease duration, also showed significant differences between the two groups (Table 1, Figure 1(d)). Overall, these findings indicate that patients with MCI exhibit distinct cognitive and motor characteristics compared to those with full dementia.

To study the in-depth compositional and functional changes of the gut and oral microbiomes on cognitive impairment and progression in PD, we performed deep sequencing with an average of 31.9 reads and 28.2 million reads per sample, respectively, and quantified microbial genes and metagenomic species (MGSs) (Figure 1(a)). Stool and saliva samples were mapped to respective gene catalogs^{39,40} to satisfactory standard with an average mapping rate of 67.50% and 39.32%, respectively, which allowed us to confidently proceed with downstream analysis to identify MGSs (Figure S1A, Supplementary Table S2, Supplementary Table S3, Supplementary Table S4). Our analysis revealed a significant decrease in gut microbiome diversity in PDD patients compared to those with PD-MCI (Figure 1(e)) and a decrease in both diversity and MGS richness in the oral microbiome of PDD patients (Figure 1(f), Method). Taxonomic profiling at the phylum level

in gut and oral microbiome showed Actinobacteria was increased in the gut of PD-MCI and PDD patients while Bacteroidetes was decreased (Figure S1B, Figure S1C, Supplementary Table S5). In the oral cavity we found a decrease in Actinobacteria, Bacteroidetes, Firmicutes, Proteobacteria and Spirochaetes, specifically in PDD patients. These findings suggest that global alterations in both gut and oral microbiomes are present and may be linked to cognitive decline in PD.

To further investigate how the composition of the microbiome changes at varying levels of cognitive impairment, we performed differential abundance testing of MGSs together with functional enrichment analysis. In the gut microbiome, we identified three clusters of signatures characterized by distinctly different genera (Figure 1g). The first genera cluster showed the enrichment of species for *Bifidobacterium longum*, *Bilophila wadsworthia*, *Ruthenibacterium lactatiformans* in PD-MCI patients (Figure S1D). Notably, consistent with a previous report, *Desulfovibrio* genus also increases with PD severity.⁴¹ These species were functionally enriched for energy generating metabolic pathways such as citrate cycle, as well as inositol-phosphate metabolism but depleted for glutathione biosynthesis (Figure 1(h)). The second cluster represents a significant enrichment in PDD patients with opportunistic pathogen species from genera such as *Olsenella sp. Marseille-P2912* and *Hungatella* (Figure S1D) and with functional enrichment like those of PD-MCI but additionally enriched for several amino acid transport systems (Figure 1(h)). The third cluster represents

Table 1. Demographic and clinical features of the study cohort.

	HC	PD	PD-MCI	PDD
Demographic features				
Age* (years, mean $\pm$ std)	59.62 $\pm$ 8.30	62.40 $\pm$ 10.03	67.27 $\ddagger$ $\pm$ 8.75	70.89 $\pm$ 7.34
Gender (Female)	15 (57.69%)	6 (30.00%)	14 (34.15%)	19 (40.43%)
Education (years mean $\pm$ std)	59.62 $\pm$ 8.30	–	67.27 $\pm$ 8.75	70.89 $\pm$ 7.34
Cognitive features				
MMSE† (mean $\pm$ std)	28.00 $\pm$ 1.85	28.35 $\pm$ 1.53	23.34 $\pm$ 3.42	19.66 $\pm$ 3.58
CDR (mean $\pm$ std)	0.00 $\pm$ 0.00	–	0.50 $\pm$ 0.00	1.15 $\pm$ 0.36
Motor features				
UPDRS‡ (mean $\pm$ std)	–	39.75 $\pm$ 19.02	35.59 $\pm$ 16.16	47.70 $\pm$ 17.60
HYES§ (mean $\pm$ std)	–	2.25 $\pm$ 0.98	1.97 $\pm$ 0.90	2.54 $\pm$ 0.86
PD duration (months, mean $\pm$ std)	–	63.60 $\pm$ 80.77	69.88 $\pm$ 49.67	105.20 $\pm$ 56.85

*.PDD vs. HC, PDD vs. PD, PD-MCI vs. HC.

†.PDD vs. PD-MCI, PDD vs. HC, PDD vs. PD, PD-MCI vs. HC, PD-MCI vs. PD.

‡.PDD vs. PD-MCI.

§.PDD vs. PD-MCI.

||.PDD vs. PD-MCI, PDD vs. PD.

||.Note: Pairwise Kruskal–Wallis tests was performed for all demographic and clinical features with significance comparisons indicated below.

a distinct depletion signature characteristic in PDD patients that predominantly consists of commensal and beneficial microbes. Several butyrate-producing microbes such as *Roseburia faecis*, *Faecalibacterium prausnitzii* together with several *Ruminococcus* species were all depleted in PDD compared to PD-MCI (Figure 1(g), Figure S1D). In addition, compared to HC patients, PDD patients also show enrichment of *B. longum*, *B. adolescentis*, *R. lactatiformans* (Figure S1E). We then reconstructed a correlation network using the gut microbiome genera of these three clusters which showed that cluster 1 and cluster 2 have an overall negative correlation with cluster 3 further supporting that the depletion signature in PDD identified in cluster 3 (Figure S1F). As expected, genera in cluster 2 had a stronger negative correlation with cluster 3 compared to cluster 1 and cluster 3, however, we interestingly also observed a strong negative correlation between *Bifidobacterium* in cluster 1 and other genera in cluster 3 suggesting that *Bifidobacterium* is a strong driver in gut microbiome community structures and that its increase leads to depletion of other species.

To identify whether observed species changes are a result of CI, we compared differentially abundant species in PD-MCI and PDD in the gut with differentially abundant species in PD patients without CI, termed PD (Figure S1H, Supplementary Table S5). Patients classified as PD had significantly different MMSE scores compared to PD-MCI and PDD patients but did not differ significantly from healthy controls (HC), suggesting that they likely did not have CI (Table 1). When comparing all three PD groups with HC, we identified a limited number of overlapping species. Specifically, eight enriched and three depleted species in PD overlapped with either PD-MCI or PDD (Figure S1H, Figure S1I). Notably, many species were uniquely enriched or depleted in PD-MCI or PDD, suggesting that these microbial changes are likely associated with CI.

Diagnosing CI in PD remains challenging due to its subjective nature the difficulty in distinguishing CI features from motor symptoms, such as those assessed by the UPDRS, and disease duration. To address this, we stratified patients into tertiles based on UPDRS scores and disease duration, respectively, and identified differentially abundant species

associated with these factors. We found that a total of 40 species differentially enriched or depleted because of UPDRS, while 65 species were associated with disease duration (Figure S1J). Interestingly, 43 species were uniquely enriched or depleted because of CI (Figure S1K). This group included key butyrate-producing species such as *Faecalibacterium prausnitzii*, *Bifidobacterium dentium*, *B. bifidum* and *Roseburia* sp. To further assess the relationship between microbial changes and PD-related factors we also constructed linear regression models to test if any of the enriched or depleted species were correlated with UPDRS or disease duration. We identified 12 gut species significantly correlated with either factor (Figure S2A, p-value <0.05). Among them, Butyrate-producing *Ruminococcus* sp., which was found to be depleted in PDD because of CI, was associated with both UPDRS and disease duration. Other species included, *Hungatella hathewayi*, *Klebsiella pneumoniae*, *Desulfovibrio fairfieldensis*, *Bacteroides stercoris*, *Senegalimassilia anaerobia*, *Parabacteroides merdae*, *Gemmiger formicilis*, *Anaerotruncus* sp., *Acidaminococcus* sp., *Clostridium* sp. In the oral microbiome, *Streptococcus pneumoniae*, *Megasphaera micronuciformis*, *Treponema denticola*, *Atopobium parvulum* were significantly associated with either of UPDRS and disease durations. Since these species strongly correlate with PD motor severity (UPDRS) and disease duration, their changes are likely driven by overall PD progression rather than being specific to cognitive impairment, as they are also influenced by broader PD-related factors.

In the oral cavity, we observed an overall depletion of several species in PD-MCI and PDD patients compared to HC, which could show the loss of diversity and commensalism in the oral cavity and opportunity for pathogens to triumph (Figure S1G). We did, however, find a significant increase in abundance of *Oribacterium asaccharolyticum* in PD-MCI (Figure 1(I)). Other potential pathogenic species such as *Streptococcus pneumoniae* and *Prevotella pallens* were also increased in PD-MCI compared to HC, albeit not significantly. Our functional enrichment showed that hexose sugar transport, chemosensory two component regulatory system and neocarzinostatin antibiotic biosynthesis are enriched in PD-MCI (Figure 1(h)).

Enterotypes and salivatypes revealed distinct functional features for Parkinson's disease and cognitive impairment

To understand the compositional changes of gut and oral communities we performed principal coordinate analysis (PCoA). Although there was a statistically significant separation in the gut (gut; PERMANOVA p-value = 0.017, oral; PERMANOVA p-value = 0.495), the clustering in both cases were discernible (Figure S2B, Figure S2C). Dirichlet multinomial mixture modeling has previously been shown to bring about hidden community structures in microbiome data that otherwise cannot clearly be distinguished with supervised clustering methods.⁴² Using this approach, we identified three clusters for gut microbiome, termed enterotypes, enriched for different bacterial genera (ENT1,2 and 3; Figure 2(a)). Our clinical study groups (PD-MCI and PDD) were enriched with different enterotypes (Figure 2(b)), and we found significant clustering of enterotypes using PCoA (Figure 2(c)). HC patients were enriched for ENT2 with a distinct signature of commensal *Prevotella* bacteria while PD-MCI and PDD showed a decrease in ENT2 instead. ENT2 was depleted for cytochrome C oxidase that potentially indicate deficient energy metabolism in the gut of PD patients (Figure 2(c)). PD-MCI patients were enriched for ENT1 that showed a signature for *Bacteroides* and *Alistipes*. PDD were enriched for *Bacteroides* of ENT1 and ENT3. ENT3 in PDD was depleted for aminoacyl-tRNA biosynthesis and ribosomal pathways suggests overall less translation of proteins and a reduction in citrate cycle pathways can potentially also indicate reduced production of SCFAs. In contrast, branch-chain amino acid (BCAA) production such as isoleucine was enriched which has been shown to be linked to different diseases including PD.^{43–46}

We then identified three clusters, termed salivatypes, in the oral cavity (SAL1, 2 and 3; Figure 2(e)). This result pointed out the PD-MCI and PDD groups were enriched to different salivatypes (Figure 2(f)), and salivatypes clustered significantly using PCoA (Figure 2(g)). SAL1, enriched in PDD patients, was increased in pathobionts such as *Streptococcus*, *Rothia* and *Veillonella* and showed

an enrichment for proline biosynthesis. PDD patients also showed a depletion in SAL2 that were dominated by *Neisseria*. Interestingly, SAL2 were enriched for multidrug resistance and its depletion in PDD potentially indicate a dysfunctional microbial community.

Gut and oral biomarkers accurately predict clinical phenotypes

The observation that patients can be stratified by their gut and oral microbiomes could reflect the potential to use the microbiome for prediction of clinical outcomes and further be extended to identify novel prognostic biomarkers. We therefore used the abundances of gut and oral microbial species together with clinical metadata as features for predicting clinical outcomes (PD-MCI and PDD) using two machine learning (ML) algorithms and then used SHapley Additive exPlanations (SHAP) scoring to interpret model predictions and explain the contribution of features, or species, toward model predictions.

In four different predictions, we used gut (SIM1), oral (SIM2), both (SIM3; gut and oral) and both together with age, gender, and education (SIM4) of these patients, as features for ML prediction. Using AUCROC and accuracy we showed that SIM4 performed the best compared to other models with an average AUCROC of 69.42% and average accuracy score of 66.91% (Supplementary Table S6, Methods). It was particularly interesting to see that the inclusion of clinical metadata (age, gender, and education) improved the AUCROC score. We therefore focussed further analysis and feature selection on outcomes of SIM4 that included gut and oral species abundances together with clinical metadata.

We first showed accurate prediction of PD-MCI compared to HC (AUCROC: SVC = 0.84, RF = 0.89; Accuracy: SVC = 0.88, RF = 0.77; Figure 3(a), Methods) as well as PDD compared to HC (AUCROC: SVC = 0.82, RF = 0.86; Accuracy: SVC = 0.73, RF = 0.78; Figure 3(b)). Of particular interest was to assess whether the microbiome can be used to distinguish different levels of cognitive decline by comparing PD-MCI and PDD patients. It was therefore supportive to also accurately predict PDD compared to PD-MCI (AUCROC: SVC

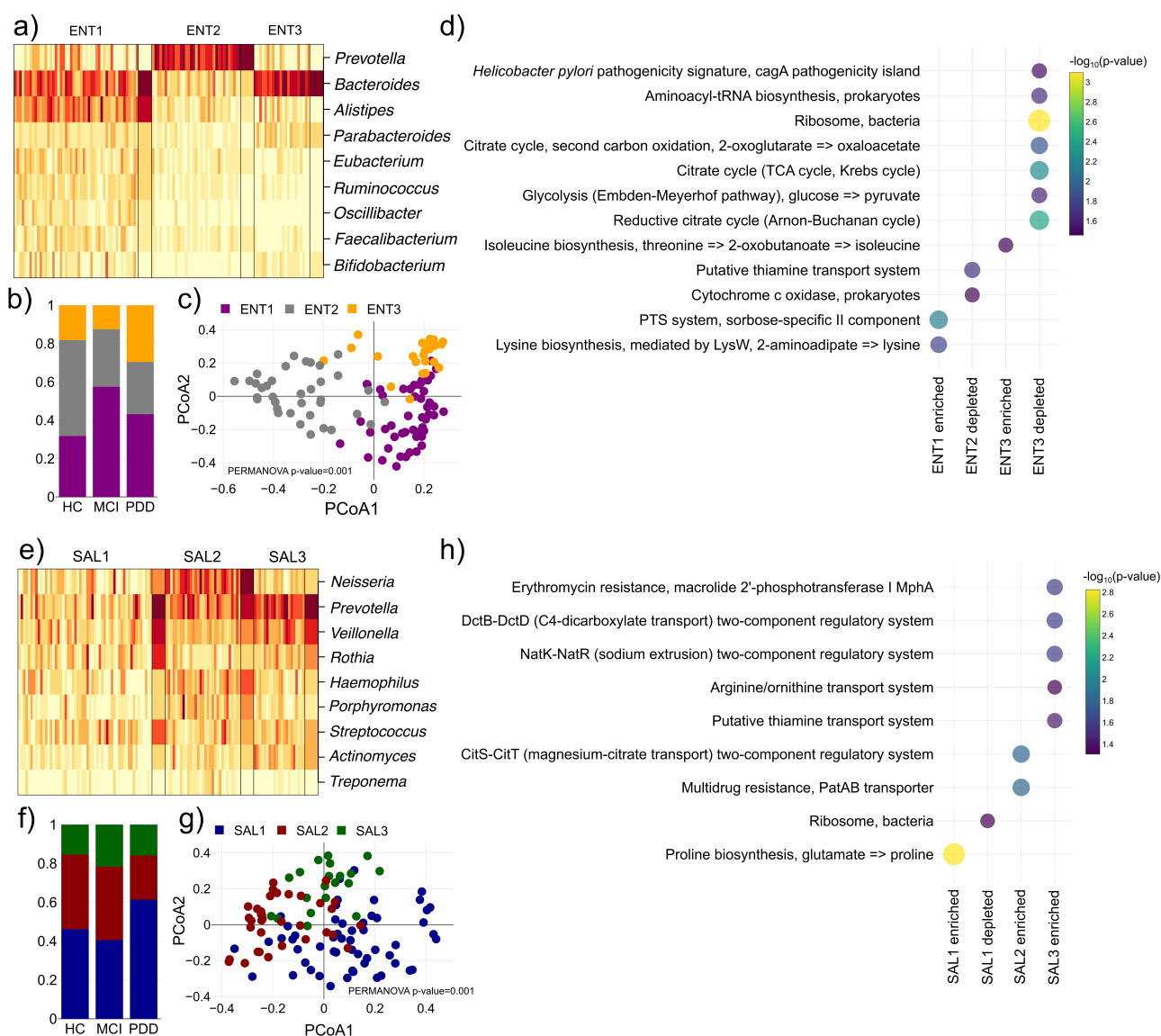


Figure 2. Enterotypes and salivatypes have distinct functional features. (a) Heatmap showing the genus abundance for three enterotypes (ET). Using dirichlet multinomial clustering we identified an optimal number of three clusters that differentiate the gut microbiome composition termed enterotype 1–3 (ENT1–3). Each cell in the heatmap depicts the relative abundance of a particular genus to each sample. (b) Fraction of samples classified as ENT1–3 for HC, PD-MCI and PDD, respectively. (c) PCoA of gut samples. The Bray-Curtis distance between all samples were calculated using species abundances and then used to perform PCoA. Each sample in the PCoA plot was colored according to its assigned enterotype. (d) Functional enrichment of KEGG modules for ENT1–3. (e) Heatmap showing the genus abundance for three salivatypes (SAL). Using dirichlet multinomial clustering we identified an optimal number of three clusters that differentiate the oral microbiome composition termed salivatype 1–3 (SAL1–3). Each cell in the heatmap depicts the relative abundance of a particular genus to each sample. (f) Fraction of samples that were classified as SAL1–3 for HC, PD-MCI and PDD, respectively. (g) PCoA of oral samples. The Bray-Curtis distance between all samples were calculated using species abundances and then used to perform PCoA. Each sample in the PCoA plot was colored according to its assigned salivatype. (h) Functional enrichment of KEGG modules for SAL1–3.

= 0.59, RF = 0.57; Accuracy: SVC = 0.45, RF = 0.5; Figure 3(c)). Furthermore, in predicting PDD versus PD-MCI the inclusion of clinical metadata in SIM4 improved the AUCROC and accuracy of the model compared to SIM3 where only gut and oral

metagenomics was used (AUCROC: SIM3 = 0.41, SIM4 = 0.58; Accuracy: SIM3 = 0.395, SIM4 = 0.475; Supplementary Table S6). This indicates that microbial species changes are sensitive enough to differentiate between PD with varying clinical

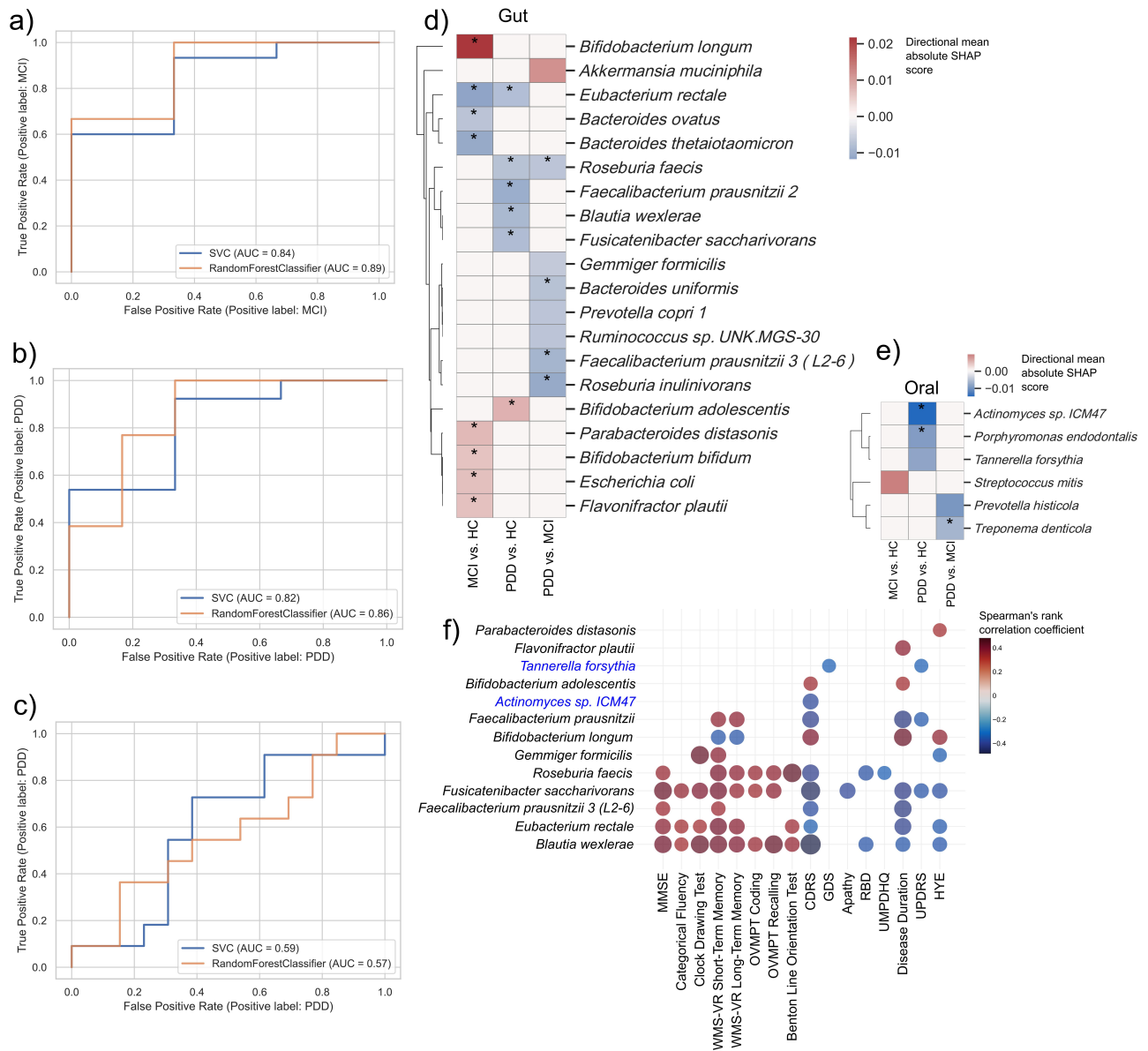


Figure 3. Machine learning models using gut and oral microbial species accurately predicts PD clinical phenotypes. a-c) random forest (RF) and support vector classification (SVC) models for prediction of clinical state. Using the species abundances of gut and oral metagenomes as features, two machine learning models were built to predict PD-MCI versus HC (a), PDD versus HC (b) and PDD versus MCI (c). The ROC curves show good AUC for all models. d-e) top features contributing to RF prediction in three different models. Using SHAP calculations we calculated the contribution of features (gut and oral microbial species, d and e, respectively) to the prediction of each model. The asterisks indicate whether these species were found to be significantly dysregulated using differential abundance analysis (table S5). (f) association of features to clinical metadata. Model features (species abundances) was correlated with clinical metadata using spearman rank correlation. The significant associations (p-value < 0.01) are indicated with asterisks (*) and the color is indicative of the correlation coefficient.

features, and the inclusion of additional features like age and gender can significantly improve model predictions.

We then applied SHAP interpretation on the ML outputs to determine the contribution of species to model predictions (Methods). Several species that were significantly dysregulated (Supplementary

Table S5) were also identified as important features for predicting PD-MCI and PDD. The decrease in beneficial bacteria such as *Faecalibacterium prausnitzii*, *Roseburia faecis*, *Roseburia inulinivorans*, *Eubacterium rectale* in the gut together with a decrease of *Treponema denticola*, *Porphyromonas endodontalis* and *Actinomyces* in the oral cavity

were the most important features in predicting PDD (Figures 3(d,e)) *Bacteroides uniformis* important for PDD prediction, have been shown to be associated with PD by increasing DAT (dopamine transporter) binding of dopamine and its decrease associated here with PDD might indicate poor dopamine metabolism that cause cognitive decline.⁴⁷ It was also interesting to see that the increase in *Akkermansia muciniphila* that has previously been associated with PD^{2,12} was an important feature for PDD prediction. The fact that the abundance of *A. muciniphila* did not significantly change in PDD using differential abundance analysis but with predictive modeling shows the value of using ML together with differential abundance. Another interesting observation was that species belonging to the *Bifidiobacterium* genus, also previously shown to be important in PD pathogenesis, were important features for predicting PD-MCI and PDD. However, *B. longum* and *B. bifidum* were associated with PD-MCI prediction while *B. adolescentis* was associated with PDD.

We finally associated top predictor species with clinical metadata and found several correlations (Figure 3(f)). For example, the abundance of *E. rectale*, *F. prausnitzii* and *R. faecis* showed a significant association with MMSE (Figure S3A). Other cognitive parameters also showed an overall positive correlation with these species. For example, *Fusicatenibacter saccharivorans* showed a positive correlation with categorical fluency tests of patients (Figure S3B). Motor parameters such as UPDRS, in addition, had correlation with oral pathogen *Tannerella forsythia* (Figure S3C).

Oral microbiome in the gut is enriched for virulence factors that contributes to PD pathophysiology and cognitive decline

The translocation of oral-specific microbial species or features to the gut lumen, a phenomenon termed oralization of the gut, has previously been associated with different diseases.^{48,49} As mentioned in the introduction, the migration of bacteria, and even fragments of their genomes to other body sites and tissues,⁵⁰ can increase the release of bacterial metabolites and cellular components causing systematic inflammation.⁵¹ To explore whether gut oralization is associated with cognitive

impairment (CI), we mapped gut metagenomes against a non-redundant oral microbial gene catalog to identify oral-specific genes in the gut (Methods).

After retrieving the gene counts from the cross-mapping gut samples, initially we performed gene richness analysis and observed PDD patients have significant enriched oral microbial genes in the gut (Figure 4(a)). To determine whether these genes potentially play a role in pathogenesis, we characterized potential virulence factors (VFs) and showed that PD-MCI and PDD patients have enriched oral-specific VFs in the gut compared to HC (Figure 4(b)). When calculating enriched or depleted VFs we found an overall enrichment of VFs in PD-MCI and PDD compared to HC and that 187 of enriched VFs overlap in PD-MCI and PDD (Figure 4(c), Supplementary Table S7).

To further understand how specific species contribute to enrichment of VFs we first identified which MGSs in both the gut and the oral contain these VFs within their pan-genomes (Figure S4). In both the gut and the oral, *Escherichia coli* contained the most VFs. In the oral cavity, other pathobionts such as *Klebsiella pneumoniae*, *Cronobacter sakazakii* and *Streptococcus salivarius* also showed several VFs as part of their genomes. Similarly, in the gut, several *Klebsiella* and *Enterobacter* species were found enriched for VFs.

To elucidate underlying community structures, we constructed an integrative correlation network between gut species, oral species, and oral VFs in the gut (Figure 4(d)). Using network analysis, we identified two clusters of co-occurring species in the gut. The first consisted of enriched species that all had a positive correlation with VFs while the second consisted of depleted species that all had a negative correlation with VFs. This was an interesting observation because enriched species were primarily pathobionts such as *E. coli*, *Egerthella* sp. and *Enterobacter xiangfangensis* which were most connected, while depleted species were all species associated with a healthier gut such as *Feacalibacterium* spp and *Roseburia intestinalis* (Figure 5(a)). Overall, this suggested that the enrichment of pathogenic species in PD, which are associated with increased VFs, potentially cause a depletion of commensal bacteria, reduces species diversity, and gut barrier dysfunction. We

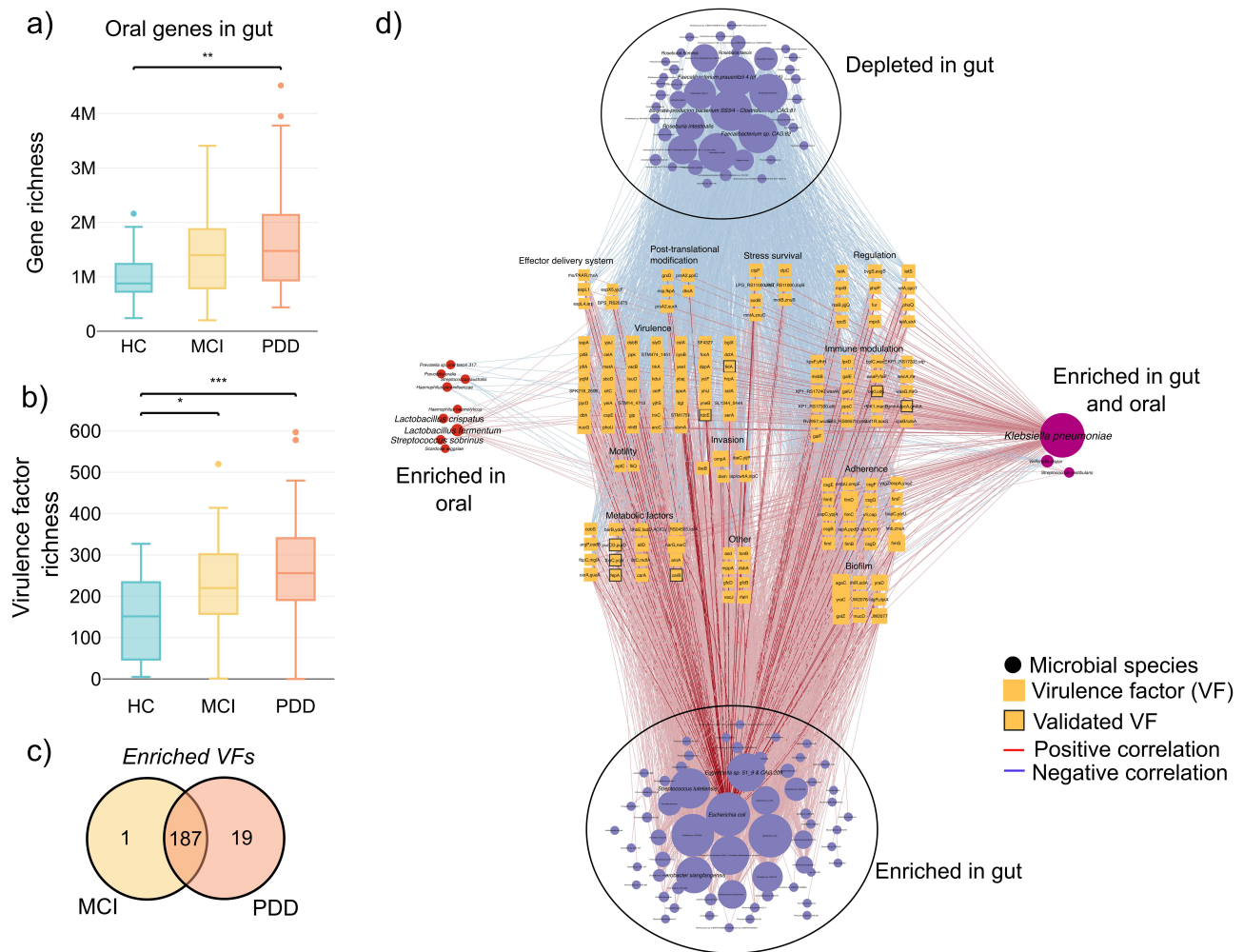


Figure 4. The sharing of gut and oral species is associated with increased virulence that contributes to disease progression. (a) Richness of oral-specific microbial genes in the gut. All gut samples were mapped against the gene catalog of oral microbial genes. The richness of genes was then calculated for different study groups. (b) Richness of oral-specific virulence factors (VFs) in the gut. VFs were identified using sequence alignment of the oral genes against the PATRIC database (methods). The richness of VFs was then calculated for different study groups. (c) Venn diagram of enriched VFs in PD-MCI and PDD patients. Differentially abundant VFs were calculated for PD-MCI and PDD patients against HC, respectively, using Wilcoxon rank-sum tests (p -value 0.05). (d) Network analysis of enriched VFs in PD patients. The 187 significantly enriched VFs were correlated with species in the gut and oral. Using network analysis, we then identified several enriched and depleted species in the gut that are positively and negatively correlated with these VFs, respectively. Virulence factors were aligned using BLASTP against saliva metaproteomics from the same subjects and significantly aligned VFs are annotated with black borders. Node sizes reflect the number of edges connected to the node.

also found that *K. pneumonia* was positively correlated in both the gut and oral with VFs and that *Streptococcus sobrinus*, *Lactobacillus fermentum* and *L. crispatus* were enriched in oral and positively correlated with VFs. The significant positive correlations between the oral and gut enriched species with virulence in the PDD and PD-MCI.

Enriched VFs had different functions related to stress, immune modulation, adherence, biofilm formation, invasion, and metabolism (Figure 4(d), Supplementary Table S8). The most connected VFs, *agaC*, *yraC*, *yraD*, *fimG*, *fimD*, *narG*, *narC*,

gatZ and *ompA*, were all involved in biofilm formation, adherence, and invasion. Outer membrane protein A (*ompA*) contributes to brain microvascular endothelial cells (BMECs) invasion via a ligand–receptor interaction. Further investigation revealed that *ibeB*, *ibeC* and *lmb* are also involved in brain microvascular endothelial cell invasion. It was also interesting to find that several VFs related to immune modulation are involved in LPS synthesis, for example *gmhA* (Figure 5(b)).

We then used saliva proteomics previously generated for the same cohort of patients²⁶ to

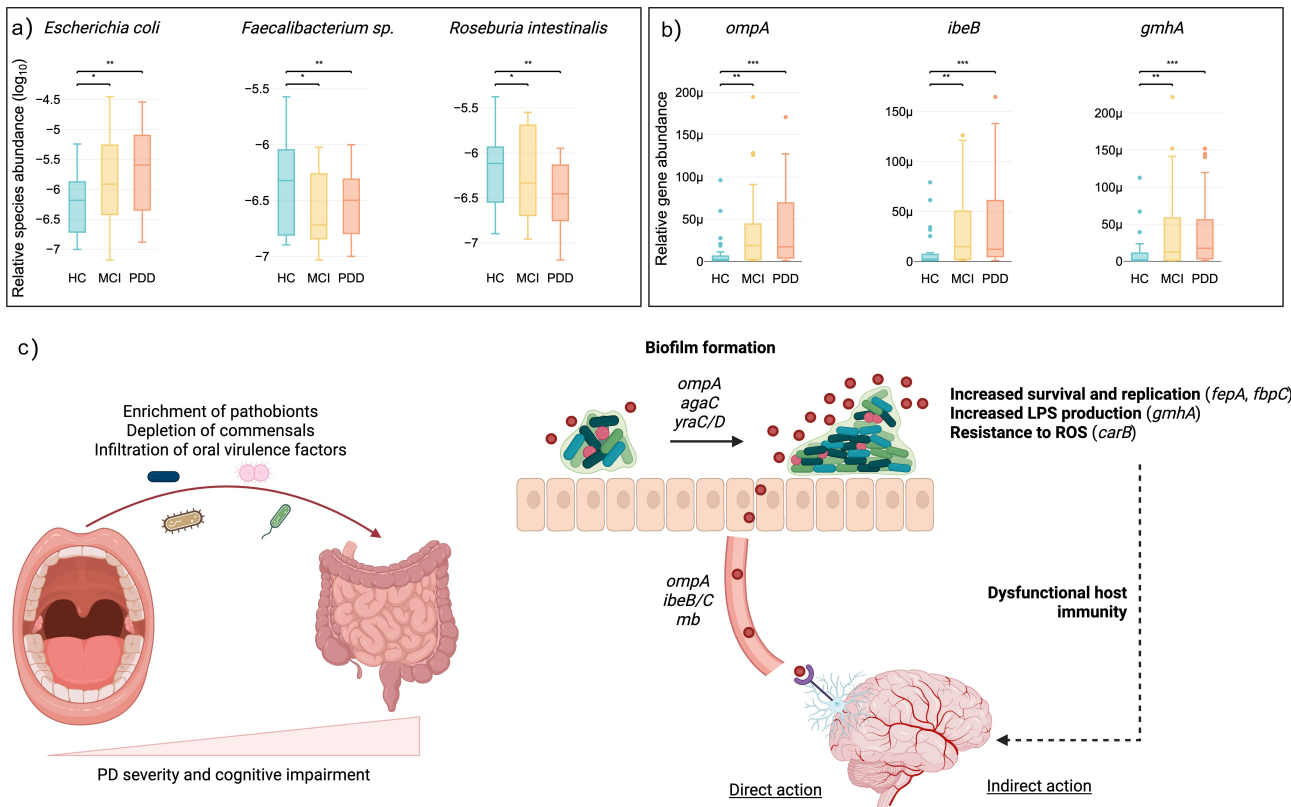


Figure 5. An oral-gut-brain axis established through the infiltration of oral virulence factors to the gut. Network analysis revealed key pathobiont species, such as *E. coli*, enriched in the gut while commensal species such as *faecalibacterium* and *R. intestinalis* are decreasing. Together with the increase of VFs such as *ompA*, *ibeB* and *gmhA* we propose a new oral-gut-brain axis where these compositional changes promote biofilm formation in the gut that, in turn, increase production of VFs that directly interact with the brain. We also propose an indirect mechanism where increased bacterial survival and replication, LPS production and bacterial protection from ROS, could lead to dysfunctional host immunity.

identify whether any of the discovered oral VFs in gut metagenomes were enriched in the oral cavity that could potentially suggest that these VFs have translocated to the gut and serve as a validation (Figure 4(d)). We were able to validate eight VFs using this approach. D-sedoheptulose 7-phosphate isomerase (*gmhA*), is a major immunogen involved in LPS biosynthesis and UDP-glucose 6-dehydrogenase (*udg*) also involved in immune modulation assist in the evasion of the host immune system by protecting bacteria from opsonophagocytosis and serum killing (Supplementary Table S8). Other validated VFs were involved in metabolism and two genes, outer membrane receptor for ferric enterobactin and colicins B, D (*fepA*) and ATP-binding protein (*fbpC*), were responsible for iron uptake by bacterial cells. Lastly, carbamoyl-phosphate synthase large chain (*carB*) mediates bacterial resistance to reactive oxygen species (ROS) and is important for phagosomal escape.

Discussion

Human gut and oral microbiomes have each been implicated to play a role in the pathology of PD and other neurodegenerative diseases.² Our study reveals a significant connection between the oral and gut microbiomes in PD patients, emphasizing the functional role of this oral-gut microbiome continuum in CI. Most notably, our functional analysis shows that infiltration of oral microbiome in the gut are responsible for increased virulence factors in the gut lumen of PD patients.

Microbial alterations in PD are modulated by cognitive impairment

The increased abundance of microbial species from the *Bifidobacterium*, *Lactobacillus* and *Akkermansia* genera in the gut have consistently been associated

with PD,^{2,11,12,25,48} while butyrate-producing taxa such as, *Roseburia*, *Faecalibacterium* and *Blautia* are consistently depleted in PD patients. Our findings provide the evidence that CI in PD further influences gut microbial composition when considering different stages of cognitive decline.

We found that *B. longum* is enriched in PD-MCI and PDD while *B. adolescentis* are only significantly enriched in PDD, suggesting that it is important to consider species specific changes with relation to CI. *Bifidobacterium adolescentis* can therefore be a potential marker for cognitive decline in PD. Similarly, *L. mucosae* was enriched in PD-MCI and PDD while *L. salivarius* and *L. gasseri* were only enriched in PDD and *L. johnsonii* was only enriched in MCI. Although not as frequently reported, *Ruthenibacterium lactatiformans*, enriched in PD-MCI and PDD, were also previously shown associated with PD.¹¹

In congruence with previous studies, we found several *Faecalibacterium* spp, *Roseburia* spp, and *Blautia* spp depleted in PDD.^{2,11} We also identified other butyrate producers such as *Ruminococcus* sp. to be depleted. Interestingly these species were depleted in PDD but not in PD-MCI compared to HC. Furthermore, we found some of these species, e.g., *Faecalibacterium prausnitzii* and *Blautia wexlerae*, were also significantly depleted in PDD compared to PD-MCI. Taken together, these results therefore suggest that depletion of species that are generally associated with a healthy gut environment decline as a function of CI and that they could potentially be used as markers of CI in PD.

There are fewer studies and less consensus on the compositional changes in the oral microbiome in PD. However, the recent discovery of *Porphyromonas gingivalis*, an oral pathogen causing periodontitis, in the brains of patients with Alzheimer's disease patients brought forth an interest on the functional role of the oral microbiome and CI²⁹. We found that *P. endodontalis* is significantly enriched in PDD. *Porphyromonas endodontalis* is also found in patients with chronic periodontitis and its abundance is correlated with *P. gingivalis*.⁵² We therefore show for the first time another species of *Porphyromonas* that are associated with CI and hypothesize that it could potentially play a functional role in PD pathogenesis.

A combination of unsupervised and supervised ML methodologies reveals functional features and associations with clinical data

Several studies have explored applying different ML methods using microbiome data.⁵³ We have used both an unsupervised and supervised approach and subsequently used model predictions for functional insights and feature selection.

Unsupervised clustering has previously been used in metagenomics to establish a defined compositional signature of microbiome, termed enterotypes (ENT).⁴² The concept of the three distinct enterotypes initially proposed has been challenged⁵⁴ and therefore we performed a *de novo* enterotype assignment for each sample. The depletion of a *Prevotella* enterotype and an enrichment of a *Bacteroides* enterotype in the gut of PD patients recapitulated our previous study.¹² Both PD-MCI and PDD showed a *Bacteroides* enterotype, however, PD-MCI also showed an *Alistipes* enterotype, a genus showed to be highly elevated in PD¹¹ and specifically enriched in PD-MCI.⁶ Therefore, our findings suggest that *Alistipes* could be a marker for differentiation of CI in PD. An interesting observation was the enrichment of isoleucine biosynthesis in PDD. Branch-chain amino acid (BCAA) metabolism in the gut has been linked to several diseases including correlation between BCAA levels and PD clinical symptoms.^{43,55} Using the approach of unsupervised clustering to identify salivatypes (SAL) has not been performed before in PD. Most notably, the enrichment of SAL1 in PDD because of increased *Prevotella*, *Veilonella*, *Rothia* and *Streptococcus* have enrichment of proline biosynthesis. Dysregulated proline metabolism has been shown in neuronal dysfunction and psychiatric disorders and in particular a recent study has functionally linked proline metabolism and gut microbiome in depression.^{56,57}

Several studies have explored using microbiome data as input for ML classifiers and more complex ML models.^{48,53} A major challenge, however, is development of robust interpreter methods to identify the most important features that contribute to model prediction. Here, we have used SHAP that is a state-of-the-art method for this

purpose that we have used in a previous study.⁴⁸ We showed that combining gut and oral microbiome data significantly improves model predictions, which is important for establishing clinically feasible biomarkers for PD. We found that the increase of *A. muciniphila*, commonly increased in PD, was an important feature for prediction of PDD against PD-MCI; however, *A. muciniphila* was not significantly increased in PDD when performing differential abundance analysis. This highlights the value that interpretable ML models can add to existing microbiome methodologies by identifying trends across the dataset that contribute to a specific disease phenotype where the species does not necessarily significantly change in abundance between groups. Given the significant influence of diet on microbiome, systematic investigation of dietary nutrient^{58–60} on microbiome composition and functions could elucidate a more mechanistic role of *A. muciniphila* in this context.

The depletion of *Bacteroides uniformis* was an important feature for differentiating between PDD and PD-MCI. The dopamine transporter (DAT) is responsible for transport of dopamine, the most common treatment for patients with PD. In a recent study, Hartstra *et al.* showed that fecal microbiota transplant of *B. uniformis* increased DAT binding and importantly that the gut-brain axis can be modulated.⁶¹ Our results here therefore shows that this species might be particularly important for cognition through dopamine metabolism and that its depletion leads to increased cognitive decline.

The oralization of the gut lumen correlates with increased virulence

There have been several studies that have shown that the translocation of oral bacterial species, often opportunistic pathogens, to the gut lumen drive disease phenotypes.^{36,48,62} Although this provides novel avenues for biomarkers, the underlying mechanisms and functionality in particular diseases remain largely unclear.

The enrichment of oral VFs in the gut of PD patients, in particular PDD, shows that these species potentially exert specific functions in the gut. The in-depth network analysis also highlighted

underlying microbiome community structures where enriched pathogenic species such as *E. coli*, *Enterobacter xiangfangensis* and *Egerthella* suppress commensal butyrate producing species such as *Faecalibacterium spp*, *R. intestinalis* and *R. faecis*. Competition between commensals and pathogenic bacteria and an imbalance are known to contribute to disease and here we show that it correlates with increased production of VFs by pathobionts providing mechanistic insights that could be exploited therapeutically.

Our integrative analysis suggests a potential mechanistic oral-gut-brain axis mediated by increased production of VFs in the gut (Figure 5(c)). In the first instance, several oral VFs are involved in biofilm formation and adherence. The formation of biofilms on the outer mucosal layer can lead to mucosal invasion by bringing bacteria close to the epithelium that contribute to a leaky gut. *OmpA*,⁶³ highly connected in our network, is a key VF that mediates the formation of bacterial biofilms but has interestingly also been shown to be a contributor to invasion of brain microvascular endothelial cells (BMECs) *via* ligand–receptor interaction. Together, with the increased connectivity of genes with similar function, for example *ibeB* that has been shown to invade BMECs,⁶⁴ our findings suggest a mechanism where infiltration of oral factors to the gut could cross the gut wall and interact with brain endothelial cells.

We hypothesize that the formation of biofilms in the gut contributes to overall virulence and dysregulation of the immune system through various mechanisms. Firstly, we validated two oral VFs, *fepA* and *fbpC*, involved in iron acquisition and iron metabolism that has been shown to be important in the pathogenicity and survival of pathogens such as *E. coli*.⁶⁵ In addition, reactive oxygen species (ROS) that are produced as by-products of metabolism can induce DNA damage and mediation through increased *carB* potentially protects biofilm formation. Finally, the production of lipopolysaccharide (LPS) has been shown to contribute to several diseases and a leaky gut. Apart from *gmhA* that was validated here, several other oral VFs in the gut were related to LPS metabolism. Together, we therefore hypothesize that the infiltration of the oral microbiome to the gut creates

a dysregulated microbial community structure, that promotes a leaky gut, increase pathogen survival which in turn increases LPS production and other VFs that can interact with BMECs either directly or indirectly to promote CI and PD (Figure 5(c)). It is, however, crucial that these findings are supported by experimental data in future studies to validate the proposed mechanism of action of VFs in the gut and its contribution to PD pathology.

Our study provides compelling evidence that the interplay between oral and gut microbiomes significantly influences PD pathology and CI. The translocation of oral microbial species to the gut, along with their associated VFs, highlights new avenues for understanding disease mechanisms and developing potential biomarkers and therapeutic avenues. The integration of machine learning techniques and microbiome data enhances our ability to identify key functional features and underscores the importance of a multifaceted approach in advancing our knowledge of PD and related neurodegenerative diseases.

Study limitations

Our analysis identified several microbiome features associated with CI in PD. However, disentangling cognitive features from other PD symptoms remains challenging. Factors such as disease duration and severity are known to influence the microbiome, and patients with increased disease severity often also experience CI. While the inclusion of gut microbiome samples from PD patients without CI and the correlation of species with disease duration and UPDRS partially addresses this limitation, it does not fully resolve the issue.

In addition, our study did not explicitly account for other host and environmental factors that can impact the gut and oral microbiome. For example, factors such as diet, stress, and sleep – known to influence microbiome composition – may differ in PD patients and could potentially amplify microbiome changes. Similarly, patients with more severe motor symptoms and CI may have compromised dental care, which could significantly affect the oral microbiome composition. The MDS-UPDRS stage II questionnaire, for example, contains a section on dental hygiene, and could be used

more explicitly to assess its impact on CI and PD. Another important factor that could impact microbiome composition related to PD severity and CI is physical activity or exercise that we also did not account for. There have been several studies that show how physical activity impacts gut microbiome composition and metabolism and future studies would benefit from more explicit assessment of this. Together, these limitations highlight the need for future studies to control for or measure several environmental variables to better understand their contributions to microbiome changes in CI of PD patients.

Methods

Study subjects, clinical characteristics and sample collection

The study was approved by the ethics committee at participating institutions (with the authorization number covering Medipol University Research and Training Hospital and Bakirkoy Research and Training Hospital for Psychiatric and Neurological Diseases 10,840,098–604.01.01-E.3958, respectively; and GOKAEK-2018/365/20.13 for Kocaeli University Research and Application Hospital), and informed consent was obtained from all participants prior to the study.

Between 2018 and 2022, a total of 140 subjects (HC = 26; PD = 20, PD-MCI = 41; PDD = 47) were recruited from the neurology clinics of these tertiary training hospitals. This case-control study used a subset of samples from a larger cohort ($N = 176$), recruited into an ongoing study on CI in PD. The subset was chosen based on the availability of both saliva and fecal samples. Healthy control subjects (HC) were recruited mainly from family members of patients, hospital staff, and students, and from individuals who responded to study advertisement. Saliva and fecal samples were placed on ice immediately upon collection and transferred to a -80°C freezer within 30 minutes without a preservation buffer, where they were stored until transport. At one-month intervals, samples were shipped to the research laboratory on dry ice, with delivery completed within one hour, and stored until further processing. Clinical data, including drug intake, and demographic

information, including age, gender, years of education were collected at clinic visits (Supplementary Table S1). The patients were examined by experienced neurologists and the diagnosis of PD was made within the framework of the “United Kingdom Parkinson’s Disease Society Brain Bank” criteria. Two well-trained interviewers from Department of Psychology conducted all assessments following a standardized protocol. The use of standardized administration and scoring procedures helped mitigate potential biases in measurement. Subjects with previous head trauma, stroke, or exposure to toxic substances, substance abuse, history of antibiotic or probiotic use within last one-month, chronic severe diseases (diabetes, cancer, kidney failure, etc.), autoimmune diseases, smokers, and those with symptoms suggestive of Parkinson’s plus syndromes were excluded from the study. The Hoehn-Yahr Stages Parkinson’s Staging Scale was utilized to assess the disease stage, while The Movement Disorder Society’s diagnostic criteria for Parkinson’s Disease Dementia were employed for evaluating dementia.⁶⁶ The diagnosis of Mild Cognitive Impairment (MCI) was established following the guidelines outlined by Litvan et al.,⁶⁷ employing level II criteria. This involved conducting a thorough cognitive assessment utilizing the MDS task force diagnostic criteria, which comprises neuropsychological evaluations covering two tests for each of the five cognitive domains.

Sample preparation and metagenomics sequencing

A total of 228 samples (114 stool and 114 saliva) (HC = 26, PD-MCI = 41, PDD = 47) were used for shotgun metagenomics sequencing. An additional 20 stool samples from PD patients without CI, termed PD, were also used for shotgun metagenomics sequencing. Microbial DNA was extracted from saliva samples using the DNeasy PowerSoil kit (Qiagen, Hilden, Germany) with previously described modifications.²⁶ For stool samples, the same extraction kit was used with adjustments to the manufacturer’s protocol. Stool samples were transferred to the PowerBead tube and homogenized by bead-beating using a Next Advance Bullet Blender (30 s at level 4, 30 s incubation on ice, and 30 s at level 4). Subsequently, the manufacturer’s

protocol was followed without further modification. The purity and concentration of the extracted genomic DNA were measured using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific, MA, USA). No human DNA depletion or microbial DNA enrichment was performed before sequencing. A DNA extraction negative control and a no-template PCR control were included and sequenced to identify potential microbiome contamination from reagents or the environment. Saliva and stool samples were processed in separate batches. To minimize batch effects, samples were randomized prior to DNA extraction, library preparation, and sequencing. Furthermore, we included technical replicates to enhance data robustness and reproducibility, with independent replicates at key stages, such as separate extractions from the same sample to assess nucleic acid consistency and duplicate sequencing to evaluate technical variability. Shotgun sequencing libraries were prepared according to Illumina’s Nextera XT library preparation protocol and sequenced using a NovaSeq600 platform with a 2 × 150 paired-end kit. All stool samples passed quality control and a total of 107 saliva samples passed quality control (HC = 26, PD-MCI = 37, PDD = 44).

Microbial gene and metagenomics species (MGS) quantification

Raw sequencing reads were mapped and counted using the METEOR pipeline (available at: https://github.com/sysbiomelab/meteor_pipeline). For gut samples, the IGC2 gene catalog³⁹ of human gut microbiome was used as reference and for oral samples the HS_8.4_oral gene catalog⁴⁰ was used. Mapping was performed using a >95% identity threshold to account for gene variability and the non-redundant nature of the catalog (Supplementary Table S2). This generated gene count matrices that were then subjected to downsizing and normalization (reads per kilo base per million mapped reads (RPKM method)) to generate the gene frequency matrix for downstream analysis. For the gut samples, downsizing was done at 5 million reads before normalization to correct for differences in sequencing depth. After downsizing a total of 126 samples were remaining

(HC = 22, PD = 20, PD-MCI = 40, PDD = 44) and for the oral samples, only normalization was performed (Figure S1A, Supplementary Table S2). The resulting gene matrices were then projected on previously reconstructed metagenomic species (MGSs) using the top 50 marker genes per MGS to calculate MGS abundances for each sample (Supplementary Table S3). Analysis was performed using the R package MetaOMineR (momr) designed to analyze large quantitative metagenomics datasets.⁶⁸

Functional gene annotation and analysis

The normalized gene count matrices were annotated for functional investigation. Amino acid sequences of the gut and oral catalogs were aligned to amino acid sequences of KEGG orthologs (KEGG database version 82)⁶⁹ using Diamond (version 0.9.22.123)⁷⁰ and best hit alignments with e-value $\leq 10^{-5}$ and bit scores ≥ 60 were considered. Amino acid sequences of the gut and oral catalogs were aligned to amino acid sequences of proteins in the PATRIC database⁷¹ using BLASTP and best hit alignments with e-value $\leq 10^{-10}$ and identity of $>80\%$ were considered. The virulence factor database (VFDB) incorporates the PATRIC database and gives more in depth-annotations and descriptions.⁷² We therefore enhanced the PATRIC databases annotations with that of VFDB. This was then used to calculate gene abundances for metabolic genes and virulence factors. For enriched and depleted KEGG modules we first identified differentially abundant metabolic genes using Wilcoxon rank-sum tests. For comparison of clinical study groups in Figure 1(h) we used a p-value cutoff of 0.05 and for comparison between enterotypes and salivatypes (Figure 2(d,h)) we used a p-value cutoff of 0.01. For both analyses only genes with a $\log_2$ foldchange of 2 or more were considered. We then used these genes to identify significantly enriched or depleted KEGG modules using hypergeometric tests adjusted for false discovery using the Benjamini-Hochberg procedure and considered p-adjusted <0.05 as significantly changing modules.

Richness, diversity analysis and differential abundance analysis

Richness was calculated as the sum of the number of MGS per sample and Shannon diversity was calculated using the *skbio* package in Python. Beta-diversity was done by first calculating the Bray-curtis distance between all samples using the *distance* function in *scipy* and then performing principal coordinate analysis using the *skbio* package. For differential abundance analysis the calculated abundances of MGSs were mapped to different taxonomical ranks (phylum, class, order, family, genus, or⁷³ species) and the sum of each taxon calculated per sample. We then used Wilcoxon rank-sum tests with false discovery rate adjustment using the Benjamini-Hochberg procedure to calculate differentially abundant microbes at these different taxonomical ranks. Details of statistical cutoffs can be found in the text.

Machine learning classification to predict clinical outcomes

We used the Scikit-learn python package to train random forest (RF) and support vector classification (SVC) models to predict different clinical outcomes.⁷⁴ Training and testing were performed on randomly selected samples split 70% and 30% of the full dataset, respectively, with a fixed random seed to ensure the reproducibility of the model. The following hyperparameters were set for the RF model: 'random_state': 1, 'n_estimators': 500, 'bootstrap': True and for the SVC model: 'random_state': 1. All other parameters were kept as their default. Model performance was measured using AUROC scoring and accuracy. Python implementation of the explainable AI algorithm, Shapley Additive Explanations (SHAP), was used to show the feature (species) contribution to disease classification the mean absolute SHAP score for each disease predictive model was determined using the sign of the Spearman rank correlation between the feature value and the SHAP score. Positive values indicate that a higher relative abundance is more likely to classify the disease than in healthy samples. Negative values indicate that a lower relative abundance is more likely to classify the disease than in healthy samples.

Integrative correlation network analysis of VFs and MGSs

We first mapped the gut metagenomes against the oral gene catalog to identify specific genes in the gut. Enriched and depleted VFs were calculated using Wilcoxon rank-sum tests (p -value <0.05) considering a positive and negative log-fold change, respectively. We then performed a Spearman's rank correlation between enriched VFs and MGSs in the gut and oral, respectively, using a p -value cutoff of 0.01. The protein sequences of enriched VFs were aligned to previously generated saliva metaproteomics using BLASTP with an e -value cutoff of 10^{-7} and percentage identity of higher than 50. The integrative network was then visualized in Cytoscape and used to calculate the degree of connections of each node.

Acknowledgments

This study was supported by grants from Engineering and Physical Sciences Research Council (EPSRC), EP/S001301/1, Biotechnology Biological Sciences Research Council (BBSRC) BB/S016899/1 and Science for Life Laboratory (SciLifeLab). Also this study was supported by Scientific and Technological Research Council of Turkey (TÜBİTAK) (grant no. 315S301) to Suleyman Yildirim. We acknowledge the Swedish National Infrastructure for Computing at SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project SNIC 2020-5-222, SNIC 2019/3-226, SNIC 2020/6-153 and King's College London computational infrastructure facility, CREATE, for high performance computing.

Disclosure statement

SS and AM are co-founders and shareholders of Trustlife Therapeutics. FC and SS are co-founders and shareholders of Gigabiome Ltd.

Funding

The work was supported by the Biotechnology and Biological Sciences Research Council [BB/S016899/1]; Engineering and Physical Sciences Research Council [EP/S001301/1]; Science for Life Laboratory (SciLifeLab); Türkiye Bilimsel ve Teknolojik Araştırma Kurumu [315S301].

ORCID

Saeed Shoaie  <http://orcid.org/0000-0001-5834-4533>

Author contributions

Conception and design, S.S., S.Y., and L.H., clinical metadata and sample collection and processing, T.K.D., Z.T., M.A., N. H.Y., A.S., and L.H.; data analysis, F.C.; data interpretation, F. C., S.Y., S.S. A.M., D.M., M.U; manuscript writing – original draft, F.C.; review and editing, F.C., S.Y., S.S. M.A., A.M., D. M., M.U. All authors read and approved the final manuscript.

Data availability statement

The metagenome data sequenced for this study from the total of 241 samples (134 stool and 107 saliva) can be found in the European Nucleotide Archive under the study accession PRJEB79944.

References

1. Global Burden of Disease Study C, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A, Degenhardt L, Dicker D, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013. *Lancet*. 2015;386(9995):743–800. doi: [10.1016/S0140-6736\(15\)60692-4](https://doi.org/10.1016/S0140-6736(15)60692-4).
2. Tan AH, Lim SY, Lang AE. The microbiome–gut–brain axis in parkinson disease — from basic research to the clinic. *Nat Rev Neurol*. 2022;18(8):476–495. doi: [10.1038/s41582-022-00681-2](https://doi.org/10.1038/s41582-022-00681-2).
3. Rosario D, Boren J, Uhlen M, Proctor G, Aarsland D, Mardinoglu A, Shoaie S. Systems biology approaches to understand the Host–microbiome interactions in neurodegenerative diseases. *Front Neurosci*. 2020;14:716. doi: [10.3389/fnins.2020.00716](https://doi.org/10.3389/fnins.2020.00716).
4. Kalia LV, Lang AE. Parkinson's disease. *Lancet*. 2015;386(9996):896–912. doi: [10.1016/S0140-6736\(14\)61393-3](https://doi.org/10.1016/S0140-6736(14)61393-3).
5. Aarsland D, Batzu L, Halliday GM, Geurtsen GJ, Ballard C, Ray Chaudhuri K, Weintraub D. Parkinson disease-associated cognitive impairment. *Nat Rev Dis Primers*. 2021;7(1):47. doi: [10.1038/s41572-021-00280-3](https://doi.org/10.1038/s41572-021-00280-3).
6. Ren T, Gao Y, Qiu Y, Jiang S, Zhang Q, Zhang J, Wang L, Zhang Y, Wang L, Nie K. Gut microbiota altered in mild cognitive impairment compared with normal cognition in sporadic Parkinson's disease. *Front Neurol*. 2020;11:137. doi: [10.3389/fneur.2020.00137](https://doi.org/10.3389/fneur.2020.00137).
7. O'Callaghan C, Lewis SJG. Cognition in Parkinson's disease. *Int Rev Neurobiol*. 2017;133:557–583. doi: [10.1016/bs.irn.2017.05.002](https://doi.org/10.1016/bs.irn.2017.05.002).
8. Shi H, Ge X, Ma X, Zheng M, Cui X, Pan W, Zheng P, Yang X, Zhang P, Hu M, et al. A fiber-deprived diet causes cognitive impairment and hippocampal microglia-mediated synaptic loss through the gut

- microbiota and metabolites. *Microbiome*. 2021;9(1):223. doi: [10.1186/s40168-021-01172-0](https://doi.org/10.1186/s40168-021-01172-0).
9. Olson CA, Iñiguez AJ, Yang GE, Fang P, Pronovost GN, Jameson KG, Rendon TK, Paramo J, Barlow JT, Ismagilov RF, et al. Alterations in the gut microbiota contribute to cognitive impairment induced by the ketogenic diet and hypoxia. *Cell Host & Microbe*. 2021;29(9):1378–1392 e1376. doi: [10.1016/j.chom.2021.07.004](https://doi.org/10.1016/j.chom.2021.07.004).
 10. Hill-Burns EM, Debelius JW, Morton JT, Wissemann WT, Lewis MR, Wallen ZD, Peddada SD, Factor SA, Molho E, Zabetian CP, et al. Parkinson's disease and Parkinson's disease medications have distinct signatures of the gut microbiome. *Mov Disord*. 2017;32(5):739–749. doi: [10.1002/mds.26942](https://doi.org/10.1002/mds.26942).
 11. Wallen ZD, Demirkan A, Twa G, Cohen G, Dean MN, Standaert DG, Sampson TR, Payami H. Metagenomics of Parkinson's disease implicates the gut microbiome in multiple disease mechanisms. *Nat Commun*. 2022;13(1):6958. doi: [10.1038/s41467-022-34667-x](https://doi.org/10.1038/s41467-022-34667-x).
 12. Rosario D, Bidkhorji G, Lee S, Bedarf J, Hildebrand F, Le Chatelier E, Uhlen M, Ehrlich SD, Proctor G, Wüllner U, et al. Systematic analysis of gut microbiome reveals the role of bacterial folate and homocysteine metabolism in Parkinson's disease. *Cell Rep*. 2021;34(9):108807. doi: [10.1016/j.celrep.2021.108807](https://doi.org/10.1016/j.celrep.2021.108807).
 13. Shandilya S, Kumar S, Kumar Jha N, Kumar Kesari K, Ruokolainen J. Interplay of gut microbiota and oxidative stress: perspective on neurodegeneration and neuroprotection. *J Adv Res*. 2022;38:223–244. doi: [10.1016/j.jare.2021.09.005](https://doi.org/10.1016/j.jare.2021.09.005).
 14. Loh JS, Mak WQ, Tan LKS, Ng CX, Chan HH, Yeow SH, Foo JB, Ong YS, How CW, Khaw KY. Microbiota–gut–brain axis and its therapeutic applications in neurodegenerative diseases. *Signal Transduct Targeted Ther*. 2024;9(1). doi: [10.1038/s41392-024-01743-1](https://doi.org/10.1038/s41392-024-01743-1).
 15. Alpino GCA, Pereira-Sol GA, Dias MME, Aguiar AS, Peluzio M. Beneficial effects of butyrate on brain functions: a view of epigenetic. *Crit Rev Food Sci Nutr*. 2024;64(12):3961–3970. doi: [10.1080/10408398.2022.2137776](https://doi.org/10.1080/10408398.2022.2137776).
 16. Tufekci KU, Genc S, Genc K. The endotoxin-induced neuroinflammation model of Parkinson's disease. *Parkinsons Dis*. 2011;2011:487450. doi: [10.4061/2011/487450](https://doi.org/10.4061/2011/487450).
 17. Zakaria R, Wan Yaacob WMH, Othman Z, Long I, Ahmad AH, AL-RAHBI B. Lipopolysaccharide-induced memory impairment in rats: a model of Alzheimer's disease. *Physiol Res*. 2017;66:553–565. doi: [10.33549/physiolres.933480](https://doi.org/10.33549/physiolres.933480).
 18. Couch Y, Alvarez-Erviti L, Sibson NR, Wood MJ, Anthony DC. The acute inflammatory response to intranigral α -synuclein differs significantly from intranigral lipopolysaccharide and is exacerbated by peripheral inflammation. *J Neuroinflammation*. 2011;8(1):166. doi: [10.1186/1742-2094-8-166](https://doi.org/10.1186/1742-2094-8-166).
 19. Mayerhofer R, Fröhlich EE, Reichmann F, Farzi A, Kogelnik N, Fröhlich E, Sattler W, Holzer P. Diverse action of lipoteichoic acid and lipopolysaccharide on neuroinflammation, blood-brain barrier disruption, and anxiety in mice. *Brain Behav Immun*. 2017;60:174–187. doi: [10.1016/j.bbi.2016.10.011](https://doi.org/10.1016/j.bbi.2016.10.011).
 20. Ginsburg I. Role of lipoteichoic acid in infection and inflammation. *Lancet Infect Dis*. 2002;2(3):171–179. doi: [10.1016/s1473-3099\(02\)00226-8](https://doi.org/10.1016/s1473-3099(02)00226-8).
 21. Figura M, Friedman A. In search of Parkinson's disease biomarkers — is the answer in our mouths? A systematic review of the literature on salivary biomarkers of Parkinson's disease. *Neurol Neurochir Pol*. 2020;54(1):14–20. doi: [10.5603/PJNNS.a2020.0011](https://doi.org/10.5603/PJNNS.a2020.0011).
 22. Adler CH, Beach TG. Neuropathological basis of non-motor manifestations of Parkinson's disease. *Mov Disord*. 2016;31(8):1114–1119. doi: [10.1002/mds.26605](https://doi.org/10.1002/mds.26605).
 23. Mu L, Chen J, Sobotka S, Nyirenda T, Benson B, Gupta F, Sanders I, Adler CH, Caviness JN, Shill HA, et al. Alpha-synuclein pathology in sensory nerve terminals of the upper aerodigestive tract of Parkinson's disease patients. *Dysphagia*. 2015;30(4):404–417. doi: [10.1007/s00455-015-9612-7](https://doi.org/10.1007/s00455-015-9612-7).
 24. Mihaila D, Donegan J, Barns S, LaRocca D, Du Q, Zheng D, Vidal M, Neville C, Uhlig R, Middleton FA, et al. The oral microbiome of early stage Parkinson's disease and its relationship with functional measures of motor and non-motor function. *PLOS ONE*. 2019;14(6):e0218252. doi: [10.1371/journal.pone.0218252](https://doi.org/10.1371/journal.pone.0218252).
 25. Jo S, Kang W, Hwang YS, Lee SH, Park KW, Kim MS, Lee H, Yoon HJ, Park YK, Chalita M, et al. Oral and gut dysbiosis leads to functional alterations in Parkinson's disease. *NPJ Parkinsons Dis*. 2022;8(1). doi: [10.1038/s41531-022-00351-6](https://doi.org/10.1038/s41531-022-00351-6).
 26. Arikan M, Demir TK, Yıldız Z, Nalbantoğlu ÖU, Korkmaz ND, Yılmaz NH, Şen A, Özcan M, Muth T, Hanoğlu L, et al. Metaproteogenomic analysis of saliva samples from Parkinson's disease patients with cognitive impairment. *NPJ Biofilms Microbiomes*. 2023;9(1). doi: [10.1038/s41522-023-00452-x](https://doi.org/10.1038/s41522-023-00452-x).
 27. Santamaria P, Jin Y, Ghuman M, Shoaie S, Spratt D, Troiano G, Nibali L. Microbiological and molecular profile of furcation defects in a population with untreated periodontitis. *J Clin Periodontol*. 2024;51(11):1421–1432. doi: [10.1111/jcpe.14034](https://doi.org/10.1111/jcpe.14034).
 28. Hajishengallis G, Chavakis T. Local and systemic mechanisms linking periodontal disease and inflammatory comorbidities. *Nat Rev Immunol*. 2021;21(7):426–440. doi: [10.1038/s41577-020-00488-6](https://doi.org/10.1038/s41577-020-00488-6).
 29. Dominy SS, Lynch C, Ermini F, Benedyk M, Marczyk A, Konradi A, Nguyen M, Haditsch U, Raha D, Griffin C, et al. *Porphyromonas gingivalis* in Alzheimer's disease brains: evidence for disease

- causation and treatment with small-molecule inhibitors. *Sci Adv.* **2019**;5(1):eaau3333. doi: [10.1126/sciadv.aau3333](https://doi.org/10.1126/sciadv.aau3333).
30. Sparks Stein P, Steffen MJ, Smith C, Jicha G, Ebersole JL, Abner E, Dawson D. Serum antibodies to periodontal pathogens are a risk factor for Alzheimer's disease. *Alzheimers & Dementia.* **2012**;8(3):196–203. doi: [10.1016/j.jalz.2011.04.006](https://doi.org/10.1016/j.jalz.2011.04.006).
 31. Iwasaki M, Kimura Y, Ogawa H, Yamaga T, Ansai T, Wada T, Sakamoto R, Ishimoto Y, Fujisawa M, Okumiya K, et al. Periodontitis, periodontal inflammation, and mild cognitive impairment: a 5-year cohort study. *J Periodontal Res.* **2019**;54(3):233–240. doi: [10.1111/jre.12623](https://doi.org/10.1111/jre.12623).
 32. Lei S, Li J, Yu J, Li F, Pan Y, Chen X, Ma C, Zhao W, Tang X. Porphyromonas gingivalis bacteremia increases the permeability of the blood-brain barrier via the Mfsd2a/Caveolin-1 mediated transcytosis pathway. *Int J Oral Sci.* **2023**;15(1). doi: [10.1038/s41368-022-00215-y](https://doi.org/10.1038/s41368-022-00215-y).
 33. Fine RL, Manfredo Vieira S, Gilmore MS, Kriegel MA. Mechanisms and consequences of gut commensal translocation in chronic diseases. *Gut Microbes.* **2020**;11(2):217–230. doi: [10.1080/19490976.2019.1629236](https://doi.org/10.1080/19490976.2019.1629236).
 34. Makaroff L, Gunn A, Gervasoni C, Richy F. Gastrointestinal disorders in Parkinson's disease: prevalence and health outcomes in a US claims database. *J Appl Psychol Parkinson's Disease.* **2011**;1(1):65–74. doi: [10.3233/JPD-2011-001](https://doi.org/10.3233/JPD-2011-001).
 35. Horvath A, Rainer F, Bashir M, Leber B, Schmerboeck B, Klymiuk I, Groselj-Strele A, Durdevic M, Freedberg DE, Abrams JA, et al. Biomarkers for oralization during long-term proton pump inhibitor therapy predict survival in cirrhosis. *Sci Rep.* **2019**;9(1):12000. doi: [10.1038/s41598-019-48352-5](https://doi.org/10.1038/s41598-019-48352-5).
 36. Lee S, Arefaine B, Begum N, Stamouli M, Witherden E, Mohamad M, Harzandi A, Cai H, Williams R, Curtis M, et al. Oral-gut microbiome interactions in advanced cirrhosis: characterisation of pathogenic enterotypes and salivatypes, virulence factors and antimicrobial resistance. *J Hepatol.* **2025**;82:622–633. doi: [10.1016/j.jhep.2024.09.046](https://doi.org/10.1016/j.jhep.2024.09.046).
 37. Patel VC, Lee S, McPhail MJW, Da Silva K, Guilly S, Zamalloa A, Witherden E, Støy S, Manakkat Vijay GK, Pons N, et al. Rifaximin- α reduces gut-derived inflammation and mucin degradation in cirrhosis and encephalopathy: RIFSYS randomised controlled trial. *J Hepatol.* **2022**;76(2):332–342. doi: [10.1016/j.jhep.2021.09.010](https://doi.org/10.1016/j.jhep.2021.09.010).
 38. Nagao JI, Kishikawa S, Tanaka H, Toyonaga K, Narita Y, Negoro-Yasumatsu K, Tasaki S, Arita-Morioka K-I, Nakayama J, Tanaka Y. Pathobiont-responsive Th17 cells in gut-mouth axis provoke inflammatory oral disease and are modulated by intestinal microbiome. *Cell Rep.* **2022**;40(10):111314. doi: [10.1016/j.celrep.2022.111314](https://doi.org/10.1016/j.celrep.2022.111314).
 39. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, Arumugam M, Kultima JR, Prifti E, Nielsen T, et al. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol.* **2014**;32(8):834–841. doi: [10.1038/nbt.2942](https://doi.org/10.1038/nbt.2942).
 40. Le Chatelier E, Almeida M, Plaza Oñate F, Pons N, Gauthier F, Ghoulane A, Ehrlich SD, Witherden E, Gomez-Cabrero D. A catalog of genes and species of the human oral microbiota. *Portail Data INRAE.* **2021**. doi: [10.15454/WQ4UTV](https://doi.org/10.15454/WQ4UTV).
 41. Murros KE, Huynh VA, Takala TM, Saris PEJ. Desulfovibrio bacteria are associated with Parkinson's disease. *Front Cell Infect Microbiol.* **2021**;11:652617. doi: [10.3389/fcimb.2021.652617](https://doi.org/10.3389/fcimb.2021.652617).
 42. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto J-M, et al. Enterotypes of the human gut microbiome. *Nature.* **2011**;473(7346):174–180. doi: [10.1038/nature09944](https://doi.org/10.1038/nature09944).
 43. Zhang Y, He X, Qian Y, Xu S, Mo C, Yan Z, Yang X, Xiao Q. Plasma branched-chain and aromatic amino acids correlate with the gut microbiota and severity of Parkinson's disease. *NPJ Parkinsons Dis.* **2022**;8(1). doi: [10.1038/s41531-022-00312-z](https://doi.org/10.1038/s41531-022-00312-z).
 44. Shoaie S, Ghaffari P, Kovatcheva-Datchary P, Mardinoglu A, Sen P, Pujos-Guillot E, de Wouters T, Juste C, Rizkalla S, Chilloux J, et al. Quantifying diet-induced metabolic changes of the human gut microbiome. *Cell Metab.* **2015**;22(2):320–331. doi: [10.1016/j.cmet.2015.07.001](https://doi.org/10.1016/j.cmet.2015.07.001).
 45. Proffitt C, Bidkhor G, Lee S, Tebani A, Mardinoglu A, Uhlen M, Moyes DL, Shoaie S. Genome-scale metabolic modelling of the human gut microbiome reveals changes in the glyoxylate and dicarboxylate metabolism in metabolic disorders. *iScience.* **2022**;25(7):104513. doi: [10.1016/j.isci.2022.104513](https://doi.org/10.1016/j.isci.2022.104513).
 46. Ezzamouri B, Rosario D, Bidkhor G, Lee S, Uhlen M, Shoaie S. Metabolic modelling of the human gut microbiome in type 2 diabetes patients in response to metformin treatment. *NPJ Syst Biol Appl.* **2023**;9(1). doi: [10.1038/s41540-022-00261-6](https://doi.org/10.1038/s41540-022-00261-6).
 47. Hamamah S, Aghazarian A, Nazaryan A, Hajnal A, Covasa M. Role of microbiota-Gut-brain Axis in regulating dopaminergic signaling. *Biomedicines.* **2022**;10(2):436. doi: [10.3390/biomedicines10020436](https://doi.org/10.3390/biomedicines10020436).
 48. Lee S, Portlock T, Le Chatelier E, Garcia-Guevara F, Clasen F, Oñate FP, Pons N, Begum N, Harzandi A, Proffitt C, et al. Global compositional and functional states of the human gut microbiome in health and disease. *Genome Res.* **2024**;34(6):967–978. doi: [10.1101/gr.278637.123](https://doi.org/10.1101/gr.278637.123).
 49. Kitamoto S, Nagao-Kitamoto H, Jiao Y, Gilliland MG, Hayashi A, Imai J, Sugihara K, Miyoshi M, Brazil JC, Kuffa P, et al. The intermucosal connection between the

- mouth and gut in commensal Pathobiont-driven colitis. *Cell*. 2020;182(2):447–462 e414. doi: [10.1016/j.cell.2020.05.048](https://doi.org/10.1016/j.cell.2020.05.048).
50. Kataria R, Shoaie S, Grigoriadis A, Wan JCM. Leveraging circulating microbial DNA for early cancer detection. *Trends Cancer*. 2023;9(11):879–882. doi: [10.1016/j.trecan.2023.08.001](https://doi.org/10.1016/j.trecan.2023.08.001).
 51. Omenetti S, Bussi C, Metidji A, Iseppon A, Lee S, Tolaini M, Li Y, Kelly G, Chakravarty P, Shoaie S, et al. The intestine harbors functionally distinct homeostatic tissue-resident and inflammatory Th17 cells. *Immunity*. 2019;51(1):77–89 e76. doi: [10.1016/j.immuni.2019.05.004](https://doi.org/10.1016/j.immuni.2019.05.004).
 52. Lombardo Bedran TB, Marcantonio RAC, Spin Neto R, Alves Mayer MP, Grenier D, Spolidorio LC, Spolidorio DP. Porphyromonas endodontalis in chronic periodontitis: a clinical and microbiological cross-sectional study. *J Oral Microbiol*. 2012;4(1):10123. doi: [10.3402/jom.v4i0.10123](https://doi.org/10.3402/jom.v4i0.10123).
 53. Asnicar F, Thomas AM, Passerini A, Waldron L, Segata N. Machine learning for microbiologists. *Nat Rev Microbiol*. 2024;22(4):191–205. doi: [10.1038/s41579-023-00984-1](https://doi.org/10.1038/s41579-023-00984-1).
 54. Knights D, Ward TL, McKinlay CE, Miller H, Gonzalez A, McDonald D, Knight R. Rethinking “enterotypes”. *Cell Host & Microbe*. 2014;16(4):433–437. doi: [10.1016/j.chom.2014.09.013](https://doi.org/10.1016/j.chom.2014.09.013).
 55. Wang W, Jiang S, Xu C, Tang L, Liang Y, Zhao Y, Zhu G. Interactions between gut microbiota and Parkinson’s disease: the role of microbiota-derived amino acid metabolism. *Front Aging Neurosci*. 2022;14:976316. doi: [10.3389/fnagi.2022.976316](https://doi.org/10.3389/fnagi.2022.976316).
 56. Yao Y, Han W. Proline metabolism in neurological and psychiatric disorders. *Mol Cells*. 2022;45(11):781–788. doi: [10.14348/molcells.2022.0115](https://doi.org/10.14348/molcells.2022.0115).
 57. Mayneris-Perxachs J, Castells-Nobau A, Arnoriaga-Rodríguez M, Martin M, de la Vega-Correa L, Zapata C, Burokas A, Blasco G, Coll C, Escrichs A, et al. Microbiota alterations in proline metabolism impact depression. *Cell Metab*. 2022;34(5):681–701 e610. doi: [10.1016/j.cmet.2022.04.001](https://doi.org/10.1016/j.cmet.2022.04.001).
 58. Clasen F, Nunes PM, Bidkhor G, Bah N, Boeing S, Shoaie S, Anastasiou D. Systematic diet composition swap in a mouse genome-scale metabolic model reveals determinants of obesogenic diet metabolism in liver cancer. *iScience*. 2023;26(2):106040. doi: [10.1016/j.isci.2023.106040](https://doi.org/10.1016/j.isci.2023.106040).
 59. Bidkhor G, Lee S, Edwards LA, Chatelier EL, Almeida M, Ezzamouri B, Onate FP, Ponte N, Shawcross DL, Proctor G, et al. The reactome unravels a new paradigm in human gut microbiome metabolism *Metabo*. 2021.
 60. Bidkhor G, Shoaie S. MIGRENE: the toolbox for microbial and individualized GEMs, reactome and community network modelling. *Metabolites*. 2024;14(3):132. doi: [10.3390/metabo14030132](https://doi.org/10.3390/metabo14030132).
 61. Hartstra AV, Schüppel V, Imangaliyev S, Schrantee A, Prodan A, Collard D, Levin E, Dallinga-Thie G, Ackermans MT, Winkelmeijer M, et al. Infusion of donor feces affects the gut–brain axis in humans with metabolic syndrome. *Mol Metab*. 2020;42:101076. doi: [10.1016/j.molmet.2020.101076](https://doi.org/10.1016/j.molmet.2020.101076).
 62. Lee S, Meslier V, Bidkhor G, Garcia-Guevara F, Etienne-Mesmin L, Clasen F, et al. Transient colonizing microbes promote gut dysbiosis and functional impairment. *NPJ Biofilms and Microbiomes*. 2024;10:80. doi: [10.1038/s41522-024-00561-1](https://doi.org/10.1038/s41522-024-00561-1).
 63. Nie D, Hu Y, Chen Z, Li M, Hou Z, Luo X, Mao X, Xue X. Outer membrane protein a (OmpA) as a potential therapeutic target for acinetobacter baumannii infection. *J Biomed Sci*. 2020;27(1):26. doi: [10.1186/s12929-020-0617-7](https://doi.org/10.1186/s12929-020-0617-7).
 64. Wang Y, Kim KS. Role of OmpA and IbeB in Escherichia coli K1 invasion of brain microvascular endothelial cells in vitro and in vivo. *Pediatr Res*. 2002;51(5):559–563. doi: [10.1203/00006450-200205000-00003](https://doi.org/10.1203/00006450-200205000-00003).
 65. Ratledge C, Dover LG. Iron metabolism in pathogenic bacteria. *Annu Rev Microbiol*. 2000;54(1):881–941. doi: [10.1146/annurev.micro.54.1.881](https://doi.org/10.1146/annurev.micro.54.1.881).
 66. Emre M, Aarsland D, Brown R, Burn DJ, Duyckaerts C, Mizuno Y, Broe GA, Cummings J, Dickson DW, Gauthier S, et al. Clinical diagnostic criteria for dementia associated with Parkinson’s disease. *Mov Disord*. 2007;22(12):1689–1707; quiz 1837 doi: [10.1002/mds.21507](https://doi.org/10.1002/mds.21507).
 67. Litvan I, Goldman JG, Tröster AI, Schmand BA, Weintraub D, Petersen RC, Mollenhauer B, Adler CH, Marder K, Williams-Gray CH, et al. Diagnostic criteria for mild cognitive impairment in Parkinson’s disease: movement disorder society task force guidelines. *Mov Disord*. 2012;27(3):349–356. doi: [10.1002/mds.24893](https://doi.org/10.1002/mds.24893).
 68. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, Almeida M, Arumugam M, Batto J-M, Kennedy S, et al. Richness of human gut microbiome correlates with metabolic markers. *Nature*. 2013;500(7464):541–546. doi: [10.1038/nature12506](https://doi.org/10.1038/nature12506).
 69. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res*. 2004;32(90001):D277–280. doi: [10.1093/nar/gkh063](https://doi.org/10.1093/nar/gkh063).
 70. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12(1):59–60. doi: [10.1038/nmeth.3176](https://doi.org/10.1038/nmeth.3176).
 71. Gillespie JJ, Wattam AR, Cammer SA, Gabbard JL, Shukla MP, Dalay O, Driscoll T, Hix D, Mane SP, Mao C, et al. PATRIC: the comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. *Infect Immun*. 2011;79(11):4286–4298. doi: [10.1128/IAI.00207-11](https://doi.org/10.1128/IAI.00207-11).
 72. Liu B, Zheng D, Zhou S, Chen L, Yang J. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res*. 2022;50(D1):D912–D917. doi: [10.1093/nar/gkab1107](https://doi.org/10.1093/nar/gkab1107).

73. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst.* 2017;30. <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>.
74. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, Gramfort A, Thirion B, Varoquaux G. Machine learning for neuroimaging with scikit-learn. *Front Neuroinform.* 2014;8(14). doi: 10.3389/fninf.2014.00014.