Microbiome-based biomarkers to guide personalized microbiome-based therapies for Parkinson's disease.

Haydeh Payami* [1,2], Timothy R Sampson [2,3], Charles F Murchison [1,2,4], Zachary D Wallen [1,2]

[1] Department of Neurology, University of Alabama at Birmingham, Birmingham, AL, 35233, USA
[2] Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, MD, 20815, USA
[3] Department of Cell Biology, Emory University, Atlanta, GA, 30329, USA
[4] Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, 35233, USA

***Senior and Corresponding author:**
Haydeh Payami, PhD
University of Alabama at Birmingham, Room 429, Civitan International Research Center, 1719 6th Ave S, Birmingham, AL 35233
Phone: 205-994-1909
Email: haydehpayami@uabmc.edu

Running title: Microbiome-based biomarkers

Word count: Abstract 191. Introduction 583. Method 1307. Results 616. Discussion 924. 35 references. 2 Tables. 3 Figures.

Abstract

We address an unmet challenge in Parkinson's disease: the lack of biomarkers to identify the right patients for the right therapy, which is a main reason clinical trials for disease modifying treatments have all failed. The gut microbiome is a new target for treatment of Parkinson's disease, with potential to halt disease progression. Our aim was to develop microbiome-based biomarkers to guide patient selection for microbiome-based clinical trials. We used microbial taxa that have been robustly associated with Parkinson's disease across studies and at high significance as dysbiotic features of Parkinson's disease. Using individual-level taxonomic relative abundance data, we classified patients according to their dysbiotic features, effectively defining microbiome-based subtypes of PD. We show that not all persons with Parkinson's disease have a dysbiotic microbiome, and not all dysbiotic Parkinson's disease microbiomes have the same features. Grounded in robust and reproducible data from differential abundance studies, we propose an intuitive and easily modifiable method to identify the optimal candidates for microbiome-based clinical trials, and subsequently, for treatments that are personalized for each individual's dysbiotic features. We demonstrate the method for Parkinson's disease. The concept, and the method, is generalizable for any disease with a microbiome component.

Introduction

Parkinson's disease (PD) is the fastest growing neurologic disease in the world [1]. PD is a progressively debilitating disease [2]. The earliest manifestations are often constipation, sleep disorder, and hyposmia, leading to the cardinal movement disorders, and as disease progresses, most patients develop psychosis and dementia. Treatments are symptomatic. There have been many clinical trials for disease modifying treatment aimed at stopping disease progression, and all have failed (clinicaltrials.gov). PD is a highly heterogenous disease. One treatment will not work for all patients. Biomarkers to guide selection of the right patients for the right drug has been a high-priority, unmet need.

It has long been known that constipation, inflammation in the gut, and permeable gut membrane are common in PD and precede motor signs [2]. Recent studies suggest some cases of PD start in the gut and spread to the brain, and in rodent models, PD pathologies are observed to spread

56 from gut to brain [3-7]. The gut origin raised the possibility that the gut microbiome could be
57 involved in the brain-gut-axis pathology of PD. A healthy gut microbiome controls metabolism of
58 drugs, toxicants and food, synthesis of vitamins and neurotransmitters, keeps the lining of the
59 gut intact, protects against pathogens, maintains the proper functioning of the immune and
60 nervous systems, and modulates brain-gut communication [8].
61
62 The gut microbiome is severely dysbiotic in PD, with characteristic features that are consistently
63 observed in large-scale high-resolution shotgun metagenomic studies and meta-analysis of
64 lower-resolution 16S studies [9-13]. Differential abundance analyses, conducted in different
65 geographic locations and with different methods, have converged on species and pathways that
66 are commonly dysbiotic in PD, among them are the following which we chose for the present
67 study: reduced levels of bacteria that degrade fiber and elevated levels of *Bifidobacterium*,
68 *Lactobacillus*, opportunistic pathogens (e.g., *Porphyromonas*), *Escherichia coli*, and certain
69 species of *Streptococcus* and *Actinomyces* [9-13]. Functional analyses of these data suggest
70 dysbiotic features of PD microbiome are relevant to multiple PD mechanisms, including
71 increased abundance of pathogens and immunogens, decreased production of neuroactive
72 molecules (dopamine, serotonin, glutamate and GABA), reduced capacity to degrade plant-
73 based fiber which leads to short chain fatty acids deficiency, increased inflammation and
74 compromised gut barrier, and elevated curli [9]. Curli is a bacterial amyloid produced by *E.coli*
75 which, in mice, induces alpha-synuclein aggregation, the hallmark of PD pathology [14,15]. Data
76 from genetic and toxicant-induced models of PD also suggest gut microbiome can trigger or
77 contribute to various PD pathologies and the manifestation of motor and non-motor phenotypes
78 [16-19], and that altering the microbiome via fecal microbiome transplantation (FMT)[20,21], antibiotics
79 treatments[22-25], or high fiber diet[26], each to some extent can reduce alpha-synuclein pathology,
80 inflammation, and motor and non-motor dysfunction.
81
82 The gut microbiome is an emerging target for PD therapies [27,28]. Results of early clinical trials
83 with fiber supplementation [29] and FMT [30-33] are promising. Next-generation targeted therapeutics
84 are forthcoming. One example (https://www.axialtx.com/our-programs/parkinsons-disease/) is a
85 small molecule that reduces curli and hence could potentially attenuate disease progression by
86 stopping curli-induced alpha-synuclein aggregation in the gut. The question is how to select
87 patients to optimize success of microbiome-based clinical trials. Do all PD patients have
88 elevated *E. coli* and curli? Is fiber supplementation sufficient to restore homeostasis to any
89 dysbiotic microbiome? Are all persons with PD candidates for FMT? Here, we show that not all

PD patients have a dysbiotic microbiome and not all dysbiotic microbiomes have the same features. We introduce microbiome-based biomarkers to guide patient selection for clinical trials, and upon success, as companion diagnostic for personalized treatment.

Method

Overview

We explored the possibility of subtyping PD patients based on the dysbiotic features of the microbiome; and using the microbiome profile of an individual to gauge their suitability for microbiome-based drug trials. The premise requires that disease associated features are already well established, i.e., reproduced across studies, using multiple differential abundance test methods controlling for confounders, at high statistical significance [9-13], and that the drug targeting a feature has undergone proper pre-clinical processes. The present paper describes a method to match patients to a treatment.

The overview of the method is shown in **Figure 1**. The concept and method are generalizable to any disorder with a microbiome component with well-established dysbiotic features, here, we focus on PD. The method requires a metagenomics reference dataset from a non-disease population. The reference dataset should be from the same population as the patients to be studied or treated (such reference datasets will soon be available for several populations). For demonstration, we select several features of PD as potential targets for treatment. (In reality, the feature can be a single gene or metabolite, a species, a cluster of related species, or multiple polymicrobial clusters that a drug has been developed to target.) We then calculate the mean relative abundance and confidence intervals (CI) of the features in the metagenomics dataset of non-disease population. We have now created the reference dataset. We define dysbiosis by setting thresholds for relative abundances in reference dataset (e.g., outside 95% CI). (We are cognoscente that “normal” and dysbiotic” microbiome are yet to be defined. Here, we use a working definition of dysbiotic to refer to abnormal levels of a well-established feature of PD microbiome). Patients can now be screened for dysbiotic features, which entails obtaining a stool sample, metagenomic sequencing and comparing the relative abundance of the features in the patient to the reference dataset and the thresholds that define dysbiosis.

Species-level data and a step-by-step protocol to reproduce the current work are in the Supplement as well as on ASAP-ZENODO Workplace (https://zenodo.org/records/11126860). The source data we used is public, available for download as raw data from SRA or processed data from Zenodo. Metagenomic sequences (raw data) are available at NCBI SRA under BioProject ID PRJNA834801. The full taxonomic profiling data (processed data) is available at Zenodo (https://zenodo.org/records/7246185).

Foundation of the method

Here, we use the data from the largest high-resolution metagenomic study of PD, published by Wallen et al. [9]. The Wallen et al. study was compliant with all relevant ethical regulations and was approved by Institutional Review Boards at University of Alabama at Birmingham and by the Human Research Protection Office of United States Department of Defense (funding agency). All subjects signed informed consent, including granting permission to share and use their de-identified data in subsequent studies.

The dataset is composed of deep-shotgun next-generation metagenome sequences of microbial DNA, extracted from stool samples of 490 individuals with PD and 234 neurologically healthy controls (NHC). Each metagenome (data point) represents one sample taken from one individual. Wallen et al. conducted differential abundance test via a metagenome-wide association study (MWAS) and identified species, genes and pathways that are depleted or overly abundant in PD. The feature we use here have been reproduced in other metagenomics and meta-analyses [9-13]. We also conducted correlations network analysis and identified polymicrobial clusters of species whose abundances grow or shrink together [9]. The visual of PD microbiome in Wallen et al data is depicted in correlation network (**Figure 2**). Species are plotted according to correlation in their abundances, forming clusters of correlated species, using SparCC (RRID:SCR_022734, https://web.mit.edu/almlab/sparcc.html).

The concept

The published studies have been microbiome-centric, i.e., they identified dysbiotic features in aggregates of patients as compared to controls. We now turn the approach around and make it individual-centric, i.e., what does the microbiome of one patient look like? Is the microbiome dysbiotic in every PD patient? Are all dysbiotic features present together, or are some features

present in a subset of patients, and if so, what percentage of patients has a given dysbiotic feature? The answers to these questions are fundamental to any microbiome-based therapeutic. To answer these questions, we had to (a) select features of interest that could potentially be a target of treatment and (b) create an operational definition of dysbiosis. The concept is generalizable to any disease and any microbiome in the body, and the method, both feature selection and definition of dysbiosis, can be modified to accommodate study design.

Features of PD gut microbiome

For feature selection, we used the data reported in Wallen et al [9] and selected species that were significantly associated with PD (i.e., their relative abundance was elevated or reduced in PD not only in Wallen et al but reproduced across studies [9-13]), plus the cluster of rare opportunistic pathogens were too rare to be tested individually but form a polymicrobial cluster that was shown in two independent datasets to be elevated in PD [9,35]. We grouped the selected PD-associated species into six features based on the correlation in their abundances ($|r|>0.2$, $P<0.05$) and taxonomic relatedness (members of same genus). The six dysbiotic features (**Table 1, Figure 2**) are as follows: (1) the cluster of opportunistic pathogens (elevated in PD), (2) species of *Bifidobacteria* (elevated in PD), (3) species of *Lactobacillus* (elevated in PD), (4) *Streptococcus* and *Actinomyces* species that are elevated in PD, (5) *E. coli* and *Klebsiella species* that are elevated in PD, and (6) fiber-degrading bacteria (reduced in PD). There is more to the dysbiosis of PD gut than these six features. Here, the aim being demonstrating the method, we strived for a balance between minimizing complexity while capturing the most robust and biologically relevant features.

Definition of dysbiotic and non-dysbiotic

Defining healthy and dysbiotic microbiome has been a challenge. For this project, we limited the definition of dysbiotic to a metagenome that has very high or very low relative abundance of a feature that is already well-established to be elevated or reduced in PD. To define very high and very low, we set extreme thresholds to maximize the odds that the patient being enrolled in trial is an optimal candidate. These thresholds can be changed to accommodate study design, or relaxed when a drug is approved, and the biomarker becomes a companion diagnostic for treatment.

For each feature, and per person, we calculated total relative abundance of each feature by summing the relative abundances of the species within it. We then calculated the mean relative abundance and its 95% CI in NHC (reference dataset). Next, we turned to individual-level PD data. Considering each feature separately, we called a PD person's metagenome dysbiotic if the relative abundance of the feature was outside the 95% CI of NHC. For features that are elevated in PD in aggregate data, a PD individual's gut was called dysbiotic if the relative abundance of the feature in that person was greater than the upper bound of 95% CI of NHC. For the feature that is reduced in PD, a PD individual's microbiome was called dysbiotic if the relative abundance of the feature in that person was less than the lower bound of 95% CI of NHC (**Table 2**).

A key question was whether all individuals with PD have a dysbiotic gut microbiome. To answer this question, we needed an operational definition of non-dysbiotic. We defined non-dysbiotic per feature as being well within the norm of NHC range, i.e., below the mean of NHC if the feature is elevated in PD, and above the mean if a feature is reduced in PD. When considering all six features, a PD person was designated as not having a dysbiotic microbiome only if their relative abundance for the five elevated features were all at or below the mean of NHC, and their relative abundance for reduced fiber degrading feature was at or greater than the mean in NHC.

Results

Not all PD metagenomes are dysbiotic, and not all dysbiotic microbiomes have the same features.

In Wallen et al data, among the 490 individuals with PD studied, 392 (80%) had at least one dysbiotic feature (**Figure 3a**). 98 (20%) did not have any dysbiotic feature defined as being outside the 95% CI of NHC. Using the more cautious definition of non-dysbiotic, i.e. elevated features being at or below the NHC mean and reduced features at or above the NHC mean, 76 (16%) of PD microbiomes were not dysbiotic. The patients without a dysbiotic microbiome can be identified and excluded from microbiome-based therapeutics; doing so will boost power and odds of success of clinical trials, and in clinical setting, will spare these patients unnecessary trial and error in finding the right treatment.

225 Severe reduction in fiber-degrading bacteria was seen in 333 of 490 of PD patients (68%)
226 (**Figure 3a**). Among the 333 low fiber-degraders, only 96 had no other dysbiotic features
227 (**Figure 3b**), hence 71% (=(333-96)/333) of patients with low fiber-degraders had high levels of
228 at least one other dysbiotic feature. Thus, while deficiency of fiber degraders is very common,
229 only 20% (=96/490) of PD patients had low fiber-degraders as their sole feature. This suggests
230 that while fiber treatment would be beneficial for majority of patients, a subset would stand to
231 benefit the most from fiber, while others may need treatment for their other dysbiotic features as
232 well.
233
234 Features that are elevated in PD were found in subsets of patients: high levels of opportunistic
235 pathogens was detected in 66 (13%) of the 490 PD, *Bifidobacterium* in 161 (33%), *Lactobacillus*
236 in 103 (21%), *Streptococcus* and *Actinomyces* in 94 (19%), and *E. coli* and *Klebsiella* in 86
237 (18%). To demonstrate utility of the biomarker, consider *E. coli* which has been linked to
238 pathobiology of PD. *E. coli* encodes curli, an amyloid that induces alpha-synuclein aggregation
239 in mice [14,15]. Alpha-synuclein aggregation is a pathogenic hallmark of PD. Moreover, *E. coli* and
240 genes that encode and regulate curli production are elevated in PD gut [9]. A drug developed
241 against *E. coli* or curli will likely have a better odds of success in clinical trial if patients were
242 chosen from the 18% that have high levels of *E. coli* and curli.
243
244 There is some overlap across the groups shown in **Figure 3a**, because these features are not
245 mutually exclusive. For example, as shown in **Figure 3b**, among the 66 patients with high
246 opportunistic pathogens, 33 (50%) also had high *Bifidobacteria*, 15 (23%) had high
247 *Lactobacillus*, 18 (27%) had high *Streptococcus* and *Actinomyces*, 13 (20%) had *E. coli* and
248 *Klebsiella*, and 57 (86%) had low fiber degraders. These data suggest that at the present time,
249 with fiber and FMT being the only microbiome-based treatments, the most rational option for
250 treating the patients who have a severely dysbiotic microbiome is cleansing the microbiome with
251 antibiotics followed by FMT, excluding those without dysbiosis and those with only low fiber-
252 degraders. We are not advocating for FMT or any treatment modality per se, rather, pointing out
253 the utility of biomarkers in designing a study as well as choosing the patients to study and treat.
254
255 **Application of the biomarker tool**
256
257 Defining the features of interest and setting abundance thresholds to define dysbiotic/non-
258 dysbiotic in a healthy population will constitute the reference dataset for this biomarker tool. To

259 determine the composition of the metagenome of a patient will entail obtaining a stool sample,
260 extracting the DNA, sequencing, bioinformatic processing and taxonomic assignment of
261 sequences, and calculating relative abundances of the microbial features in the individual's
262 metagenome. The patient's relative abundances for the features of interest are then compared
263 to the reference dataset to determine if they are dysbiotic (**Figure 1**).
264

265 Discussion

266

267 Meta-analyses and large metagenomic studies have converged on certain microbial features
268 that consistently and significantly are altered in PD microbiome, independent of geography,
269 confounding covariates, and methodology. These well-established disease-associated taxa are
270 pre-requisite to, and the foundation of biomarkers described here. Here, we used these robust
271 associations to define subtypes of PD based on dysbiotic features of microbiome. We
272 introduced a microbiome-based biomarker tool to guide patient selection for clinical trials and to
273 serve as companion diagnostic for personalized treatments. With this tool, we can "see" if and
274 what is dysbiotic in an individual and act accordingly. The method is applicable for any patient at
275 any stage of disease. The tool can be easily modified (by redefining features and adjusting
276 thresholds) as the field advances, and to meet varying needs of different studies. Here, we
277 demonstrated the development and the potential utility of the tool for identifying subtypes of PD
278 based on their microbiome features. The concept is generalizable to any disease with a
279 microbiome component.
280

281 Treating a dysbiotic microbiome can potentially be disease modifying, which has been an
282 elusive goal in PD field. Functional inference of human data and experiments in mice indicate
283 dysbiotic features of PD microbiome contribute to several disease mechanisms, including alpha-
284 synuclein aggregation which is at the core of PD pathology, disrupted neuro-signaling, and
285 inflammation [9,15,16]. It is therefore reasonable to work towards microbiome-based therapies that
286 could halt disease progression. As we learn which features are causal and which result from
287 disease (an active area of research), microbiome-based biomarkers will be useful for identifying
288 pathogenic dysbiosis early to prevent disease before onset.
289

290 Microbiome-based treatments under consideration for PD include prebiotics (e.g., fiber
291 supplements), probiotics (beneficial bacteria packaged and marketed direct to consumer), FMT,
292 and next generation therapeutics that are targeted to a specific feature (e.g., small molecules,

293 phage, CRSPR-Cas9). The microbiome-based biomarkers described here can serve as
294 companion diagnostic for all treatment modalities. Take fiber supplementation for example,
295 which is safe and likely beneficial for everyone. We show here that indeed, the majority of PD
296 patients have severely depleted levels of fiber-degrading bacteria. Increasing fiber intake is
297 effectively feeding these bacteria and causing them to grow and multiply, and it has shown
298 promise in an early clinical trial with newly diagnosed PD patients [29]. The value of coupling
299 biomarkers with fiber treatment is in identifying patients who, in addition to being low on fiber-
300 degraders, have elevated levels of harmful pathogenic bacteria, which is an important
301 consideration both for patient selections for clinical trials and for treating the individuals. As for
302 probiotics on the market, most are composed of "beneficial" species of *Bifidobacteria* and/or
303 *Lactobacillus*, which ironically, are highly elevated in PD. The value of the biomarker in this case
304 is to avoid supplementation that is counter indicated. FMT is showing promise for PD [30], but it
305 carries risks concerning safety and compatibility of donor sample with host environment. Using
306 biomarkers, unnecessary FMT can be avoided, sparing the patients who do not have a dysbiotic
307 microbiome and those whose only problem is low fiber-degraders. Targeted treatments could
308 replace the need for FMT, and they too would benefit from a companion diagnostic. For
309 example, for a treatment targeted to abolish the opportunistic pathogens, the biomarker can
310 identify the 13% of patients who would be the best candidates.
311
312 Currently, we do not know if alleviating one dysbiotic feature would restore homeostasis to the
313 rest of the microbiome, and what that central feature might be. Consumption of fiber is the
314 safest and simplest treatment. In a clinical trial performed on 20 newly diagnosed PD patients,
315 treatment with fiber increased the abundance of fiber-degrading bacteria and production of short
316 chain fatty acids and resulted in shifts in the microbiome community [29]. Fiber-degrading species
317 are in the center of the correlation network with tight positive and negative correlations to
318 several PD-associated species (**Figure 2**). It is therefore plausible that fiber treatment, aimed to
319 increase the abundance of fiber-degrading bacteria, may restore equilibrium to PD-associated
320 species that map to the center of network and are either positively or negatively correlated with
321 fiber-degraders. However, while possible, it is less likely that fiber treatment alone would
322 eliminate overabundance of opportunistic pathogens, for example, considering that opportunistic
323 pathogens are a dense isolated cluster with no connection to fiber-degrading species.
324
325 Profiling the patients' metagenome and comparing them to the reference dataset will show
326 where they fall in the spectrum of dysbiosis. This concept, and the proposed method, holds true

regardless of feature, disease, or dataset. However, the relative abundances of the features and therefore the range and thresholds for defining dysbiosis may differ across populations. Hence, it is important that the reference dataset is from the same source population as patients. Here, we used a cohort of neurologically healthy older adults from Birmingham, Alabama in United States for our reference dataset, as our patients are also from Birmingham. There are regional variation in microbiome, reflecting cultural, dietary, socioeconomic, and geographic influences. Despite regional variations, the main PD-associated features are robust and reproducible [9,11,12,35]. Even so, the relative abundances, and hence the thresholds for defining dysbiosis, may still vary across populations. Not every investigator can amass hundreds of controls to investigate population-specific abundances to set their thresholds. This impediment will be overcome soon by current efforts to collect all data generated and made public. These collaborative global efforts are aimed at meta-analyses, a by-product of which will be open-access to well-curated datasets from across the world that can be used as reference datasets to generate population-specific thresholds for biomarkers and companion diagnostics.

Data availability

The dataset used for biomarker development is publicly available, open access with no restrictions, as described in detail in Wallen et al 2022 [9]. The dataset is available on two repositories to enable the investigators to download the data in either the raw or processed form. The raw metagenomic sequences and accompanying metadata are on NCBI SRA under BioProject ID PRJNA834801 [https://www.ncbi.nlm.nih.gov/bioproject/834801]. The post-QC and post taxonomic profiling data are on Zenodo [https://zenodo.org/record/7246185]. Here we used species-level data (presented in Supplement) that was extracted from the taxonomic profiling data downloaded from Zenodo in January 2024.

Code and software

Algorithm was developed using Microsoft Excel v.16.84 (RRID:SCR_016137) https://www.microsoft.com/en-gb/. Numbers were double checked using R v4.3.3. (RRID:SCR_001905) https://www.r-project.org/. Both Excel and R code are provided in the Supplement, as well as in ASAP-ZENODO Workplace (https://zenodo.org/records/11126860).

Acknowledgements

This work was supported by The Strain Endowed Chair in Parkinson's Disease to the University of Alabama at Birmingham (HP) and Aligning Science Across Parkinson's [ASAP-020527]

through the Michael J. Fox Foundation for Parkinson's Research (MJFF) (HP, TRS, CFM, ZDW). For the purpose of open access, the authors have applied a CC BY public copyright license to all Author Accepted Manuscripts arising from this submission. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the funding agencies.

Author contribution

This is a collaborative team (genomics (HP), microbiology (TRS), statistics (CFM), bioinformatics (ZDW)). The lead/corresponding author (HP) conceptualized and developed the method. HP drafted the manuscript and all authors contributed to revision and approved the final version.

Competing interest

HP has intellectual property pending on Personalized Microbiome-Based Therapies, Serial No.: 63/571,787; Filing Date: March 29, 2024. TRS is a co-inventor on intellectual property related to a small molecule microbiome-based therapy - US Patent 11,707,493 and 11,147, 792.

References

1 Collaborators, G. B. D. P. s. D. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* **17**, 939-953 (2018). https://doi.org/10.1016/S1474-4422(18)30295-3

2 Obeso, J. A. *et al.* Past, present, and future of Parkinson's disease: A special essay on the 200th Anniversary of the Shaking Palsy. *Mov Disord* **32**, 1264-1310 (2017). https://doi.org/10.1002/mds.27115

3 Braak, H., Rub, U., Gai, W. P. & Del Tredici, K. Idiopathic Parkinson's disease: possible routes by which vulnerable neuronal types may be subject to neuroinvasion by an unknown pathogen. *J Neural Transm (Vienna)* **110**, 517-536 (2003). https://doi.org/10.1007/s00702-002-0808-2

4 Horsager, J. *et al.* Brain-first versus body-first Parkinson's disease: a multimodal imaging case-control study. *Brain* **143**, 3077-3088 (2020). https://doi.org/10.1093/brain/awaa238

5 Holmqvist, S. *et al.* Direct evidence of Parkinson pathology spread from the gastrointestinal tract to the brain in rats. *Acta Neuropathol* **128**, 805-820 (2014). https://doi.org/10.1007/s00401-014-1343-6

6 Breen, D. P., Halliday, G. M. & Lang, A. E. Gut-brain axis and the spread of alpha-synuclein pathology: Vagal highway or dead end? *Mov Disord* **34**, 307-316 (2019). https://doi.org/10.1002/mds.27556

7 Kim, S. *et al.* Transneuronal Propagation of Pathologic alpha-Synuclein from the Gut to the Brain Models Parkinson's Disease. *Neuron* **103**, 627-641 e627 (2019). https://doi.org/10.1016/j.neuron.2019.05.035

8 Fan, Y. & Pedersen, O. Gut microbiota in human metabolic health and disease. *Nat Rev Microbiol* **19**, 55-71 (2021). https://doi.org/10.1038/s41579-020-0433-9

9 Wallen, Z. D. *et al.* Metagenomics of Parkinson's disease implicates the gut microbiome in multiple disease mechanisms. *Nat Commun* **13**, 6958 (2022). https://doi.org/10.1038/s41467-022-34667-x

10 Palacios, N. *et al.* Metagenomics of the Gut Microbiome in Parkinson's Disease: Prodromal Changes. *Ann Neurol* **94**, 486-501 (2023). https://doi.org/10.1002/ana.26719

11 Romano, S. *et al.* Meta-analysis of the Parkinson's disease gut microbiome suggests alterations linked to intestinal inflammation. *NPJ Parkinsons Dis* **7**, 27 (2021). https://doi.org/10.1038/s41531-021-00156-z

12 Toh, T. S. *et al.* Gut microbiome in Parkinson's disease: New insights from meta-analysis. *Parkinsonism Relat Disord* **94**, 1-9 (2022). https://doi.org/10.1016/j.parkreldis.2021.11.017

13 Boktor, J. C. *et al.* Integrated Multi-Cohort Analysis of the Parkinson's Disease Gut Metagenome. *Mov Disord* **38**, 399-409 (2023). https://doi.org/10.1002/mds.29300

14 Chen, S. G. *et al.* Exposure to the Functional Bacterial Amyloid Protein Curli Enhances Alpha-Synuclein Aggregation in Aged Fischer 344 Rats and Caenorhabditis elegans. *Sci Rep* **6**, 34477 (2016). https://doi.org/10.1038/srep34477

15 Sampson, T. R. *et al.* A gut bacterial amyloid promotes alpha-synuclein aggregation and motor impairment in mice. *Elife* **9** (2020). https://doi.org/10.7554/eLife.53111

16 Sampson, T. R. *et al.* Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. *Cell* **167**, 1469-1480 e1412 (2016). https://doi.org/10.1016/j.cell.2016.11.018

17 Matheoud, D. *et al.* Intestinal infection triggers Parkinson's disease-like symptoms in Pink1(-/-) mice. *Nature* **571**, 565-569 (2019). https://doi.org/10.1038/s41586-019-1405-y

18 Radisavljevic, N. *et al.* Effects of Gut Microbiota Alterations on Motor, Gastrointestinal, and Behavioral Phenotype in a Mouse Model of Parkinson's Disease. *J Parkinsons Dis* **12**, 1479-1495 (2022). https://doi.org/10.3233/JPD-223165

429 19 Bhattarai, Y. *et al.* Role of gut microbiota in regulating gastrointestinal dysfunction and
430 motor symptoms in a mouse model of Parkinson's disease. *Gut Microbes* **13**, 1866974
431 (2021). https://doi.org/10.1080/19490976.2020.1866974
432 20 Sun, M. F. *et al.* Neuroprotective effects of fecal microbiota transplantation on MPTP-
433 induced Parkinson's disease mice: Gut microbiota, glial reaction and TLR4/TNF-alpha
434 signaling pathway. *Brain Behav Immun* **70**, 48-60 (2018).
435 https://doi.org/10.1016/j.bbi.2018.02.005
436 21 Zhao, Z. *et al.* Fecal microbiota transplantation protects rotenone-induced Parkinson's
437 disease mice via suppressing inflammation mediated by the lipopolysaccharide-TLR4
438 signaling pathway through the microbiota-gut-brain axis. *Microbiome* **9**, 226 (2021).
439 https://doi.org/10.1186/s40168-021-01107-9
440 22 Bisht, R., Kaur, B., Gupta, H. & Prakash, A. Ceftriaxone mediated rescue of nigral
441 oxidative damage and motor deficits in MPTP model of Parkinson's disease in rats.
442 *Neurotoxicology* **44**, 71-79 (2014). https://doi.org/10.1016/j.neuro.2014.05.009
443 23 Du, Y. *et al.* Minocycline prevents nigrostriatal dopaminergic neurodegeneration in the
444 MPTP model of Parkinson's disease. *Proc Natl Acad Sci U S A* **98**, 14669-14674 (2001).
445 https://doi.org/10.1073/pnas.251341998
446 24 Pu, Y. *et al.* Antibiotic-induced microbiome depletion protects against MPTP-induced
447 dopaminergic neurotoxicity in the brain. *Aging (Albany NY)* **11**, 6915-6929 (2019).
448 https://doi.org/10.18632/aging.102221
449 25 Kaur, B. & Prakash, A. Ceftriaxone attenuates glutamate-mediated neuro-inflammation
450 and restores BDNF in MPTP model of Parkinson's disease in rats. *Pathophysiology* **24**,
451 71-79 (2017). https://doi.org/10.1016/j.pathophys.2017.02.001
452 26 Abdel-Haq, R. *et al.* A prebiotic diet modulates microglial states and motor deficits in
453 alpha-synuclein overexpressing mice. *Elife* **11** (2022).
454 https://doi.org/10.7554/eLife.81453
455 27 Sorbara, M. T. & Pamer, E. G. Microbiome-based therapeutics. *Nat Rev Microbiol* **20**,
456 365-380 (2022). https://doi.org/10.1038/s41579-021-00667-9
457 28 Tan, A. H., Lim, S. Y. & Lang, A. E. The microbiome-gut-brain axis in Parkinson disease -
458 from basic research to the clinic. *Nat Rev Neurol* **18**, 476-495 (2022).
459 https://doi.org/10.1038/s41582-022-00681-2
460 29 Hall, D. A. *et al.* An open label, non-randomized study assessing a prebiotic fiber
461 intervention in a small cohort of Parkinson's disease participants. *Nat Commun* **14**, 926
462 (2023). https://doi.org/10.1038/s41467-023-36497-x

30 DuPont, H. L. *et al.* Fecal microbiota transplantation in Parkinson's disease-A randomized repeat-dose, placebo-controlled clinical pilot study. *Front Neurol* **14**, 1104759 (2023). https://doi.org/10.3389/fneur.2023.1104759

31 Cheng, Y. *et al.* Efficacy of fecal microbiota transplantation in patients with Parkinson's disease: clinical trial results from a randomized, placebo-controlled design. *Gut Microbes* **15**, 2284247 (2023). https://doi.org/10.1080/19490976.2023.2284247

32 Segal, A., Zlotnik, Y., Moyal-Atias, K., Abuhasira, R. & Ifergane, G. Fecal microbiota transplant as a potential treatment for Parkinson's disease - A case series. *Clin Neurol Neurosurg* **207**, 106791 (2021). https://doi.org/10.1016/j.clineuro.2021.106791

33 Xue, L. J. *et al.* Fecal microbiota transplantation therapy for Parkinson's disease: A preliminary study. *Medicine (Baltimore)* **99**, e22035 (2020). https://doi.org/10.1097/MD.0000000000022035

34 . (!!! INVALID CITATION !!! 9).

35 Wallen, Z. D. *et al.* Characterizing dysbiosis of gut microbiome in PD: evidence for overabundance of opportunistic pathogens. *NPJ Parkinsons Dis* **6**, 11 (2020). https://doi.org/10.1038/s41531-020-0112-6

Table 1. Six main dysbiotic features of PD gut microbiome

Feature	Species composition of feature	PD association	Number of connections
Opp pathogens	*Actinomyces turicensis*	Increased as cluster	8
Opp pathogens	*Anaerococcus octavius*	Increased as cluster	4
Opp pathogens	*Anaerococcus vaginalis*	Increased as cluster	5
Opp pathogens	*Atopobium minutum*	Increased as cluster	8
Opp pathogens	*Corynebacterium amycolatum*	Increased as cluster	2
Opp pathogens	*Corynebacterium aurimucosum*	Increased as cluster	2
Opp pathogens	*Finegoldia magna*	Increased as cluster	7
Opp pathogens	*Gemella asaccharolytica*	Increased as cluster	2
Opp pathogens	*Peptoniphilus harei*	Increased as cluster	1
Opp pathogens	*Peptoniphilus lacrimalis*	Increased as cluster	3
Opp pathogens	*Peptoniphilus sp HMSC062D09*	Increased as cluster	7
Opp pathogens	*Porphyromonas asaccharolytica*	Increased	6
Opp pathogens	*Porphyromonas sp HMSC065F10*	Increased as cluster	11
Opp pathogens	*Porphyromonas uenonis*	Increased as cluster	6
Opp pathogens	*Prevotella bergensis*	Increased as cluster	1
Opp pathogens	*Prevotella buccalis*	Increased as cluster	10
Opp pathogens	*Prevotella disiens*	Increased as cluster	1
Opp pathogens	*Prevotella timonensis*	Increased as cluster	8
Opp pathogens	*Varibaculum cambriense*	Increased as cluster	10
Bifidobacteria	*Bifidobacterium dentium*	Increased	12
Bifidobacteria	*Bifidobacterium bifidum*	Increased	3
Bifidobacteria	*Bifidobacterium breve*	Increased	24
Bifidobacteria	*Bifidobacterium gallinarum*	Increased	14
Bifidobacteria	*Bifidobacterium longum*	Increased	5
Bifidobacteria	*Bifidobacterium pullorum*	Increased	12
Bifidobacteria	*Bifidobacterium saeculare*	Increased	13
Lactobacillus	*Lactobacillus fermentum*	Increased	14
Lactobacillus	*Lactobacillus gasseri*	Increased	10
Lactobacillus	*Lactobacillus paragasseri*	Increased	9
Lactobacillus	*Lactobacillus salivarius*	Increased	9
Lactobacillus	*Lactobacillus reuteri*	Increased	1
Lactobacillus	*Lactobacillus rhamnosus*	Increased	11
Strepto / Actino	*Actinomyces naeslundii*	Increased	4
Strepto / Actino	*Actinomyces oris*	Increased	16
Strepto / Actino	*Actinomyces sp HPA0247*	Increased	3
Strepto / Actino	*Actinomyces sp oral taxon 448*	Increased	3
Strepto / Actino	*Streptococcus anginosus group*	Increased	10

Strepto / Actino	*Streptococcus mutans*	Increased	13
Strepto / Actino	*Streptococcus vestibularis*	Increased	15
Strepto / Actino	*Streptococcus lutetiensis*	Increased	3
E. coli / klebsiella	*Escherichia coli*	Increased	5
E. coli / klebsiella	*Klebsiella pneumoniae*	Increased	3
E. coli / klebsiella	*Klebsiella quasipneumoniae*	Increased	3
Fiber-degraders	*Blautia wexlerae*	Decreased	17
Fiber-degraders	*Blautia hansenii*	Decreased	17
Fiber-degraders	*Anaerostipes hadrus*	Decreased	48
Fiber-degraders	*Clostridium sp CAG 58*	Decreased	27
Fiber-degraders	*Eubacterium eligens*	Decreased	47
Fiber-degraders	*Eubacterium hallii*	Decreased	52
Fiber-degraders	*Eubacterium ramulus*	Decreased	25
Fiber-degraders	*Eubacterium rectale*	Decreased	21
Fiber-degraders	*Eubacterium sp CAG 38*	Decreased	22
Fiber-degraders	*Faecalibacterium prausnitzii*	Decreased	59
Fiber-degraders	*Fusicatenibacter saccharivorans*	Decreased	45
Fiber-degraders	*Roseburia faecis*	Decreased	21
Fiber-degraders	*Roseburia intestinalis*	Decreased	21
Fiber-degraders	*Roseburia inulinivorans*	Decreased	27
Fiber-degraders	*Ruminococcus bicirculans*	Decreased	33
Fiber-degraders	*Ruminococcus lactaris*	Decreased	12

We selected species that individually, or as a cluster, were significantly elevated or reduced in PD [9-13], and assigned them to a “feature” according to the correlation in their abundances (connectivity, see **Figure 2**) and their taxonomic relatedness (members of same genus). Here, our aim was to demonstrate the method, for which we strived for a balance between capturing the most robust association with PD and reducing the complexity. Features can be redefined to accommodate the aims of the investigation.

Table 2. **A working definition of dysbiotic for identifying microbiome-based subtypes of PD**

Feature	Relative abundance of feature in 234 NHC			Relative abundance of feature in 490 PD			N among 490 PD outside 95% CI of NHC
	Mean	Upper 95% CI	Lower 95% CI	Mean	Upper 95% CI	Lower 95% CI	
Opp. pathogens	0.01%	**0.01%**	0.00%	0.10%	0.20%	0.00%	66 (13%)
Bifidobacteria	2.06%	**2.66%**	1.47%	5.07%	5.98%	4.17%	161 (33%)
Lactobacillus	0.11%	**0.19%**	0.04%	0.55%	0.74%	0.36%	103 (21%)
Streptococcus / Actinomyces	0.19%	**0.33%**	0.06%	0.39%	0.54%	0.25%	94 (19%)
E. coli / Klebsiella	1.47%	**2.23%**	0.72%	1.94%	2.44%	1.43%	86 (18%)
Fiber-degraders	22.35%	23.88%	**20.81%**	15.29%	16.36%	14.22%	333 (68%)

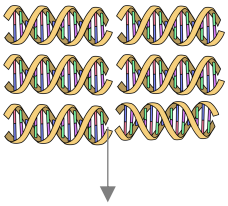
In this context, the term dysbiotic refers to a microbiome in which the relative abundance of a well-established PD-associated feature is very high (for features that are elevated in PD) or very low (if reduced in PD). Here, each feature is composed of several PD-associated species, as shown in **Table 1**. The relative abundance of each species in each individual was calculated as reported before [9] and can be found here https://zenodo.org/record/7246185. We calculated relative abundance of each feature in each neurologically healthy control (NHC) and individual with PD by summing the relative abundances of the species within the feature for that individual. We then calculated the mean and 95% confidence interval (CI) of the mean in the NHC and PD separately. We set the threshold for defining dysbiotic at the upper 95% CI of NHC for features that are elevated in PD, and at lower 95% CI of NHC for fiber-degraders that are reduced in PD. Last column is the number of PD patients, among 490 total, who were called dysbiotic for each feature. For example, in 161 of 490 PD patients, the relative abundance of *Bifidobacteria* was higher than 2.66% which was the upper 95% CI limit of NHC. Similarly, in 333 of 490 patients, the relative abundance of fiber-degrading bacteria was lower than 20.81% which was the lower 95% CI limit in NHC. See Figure 3 for overlaps.

Building the reference dataset

No Disease

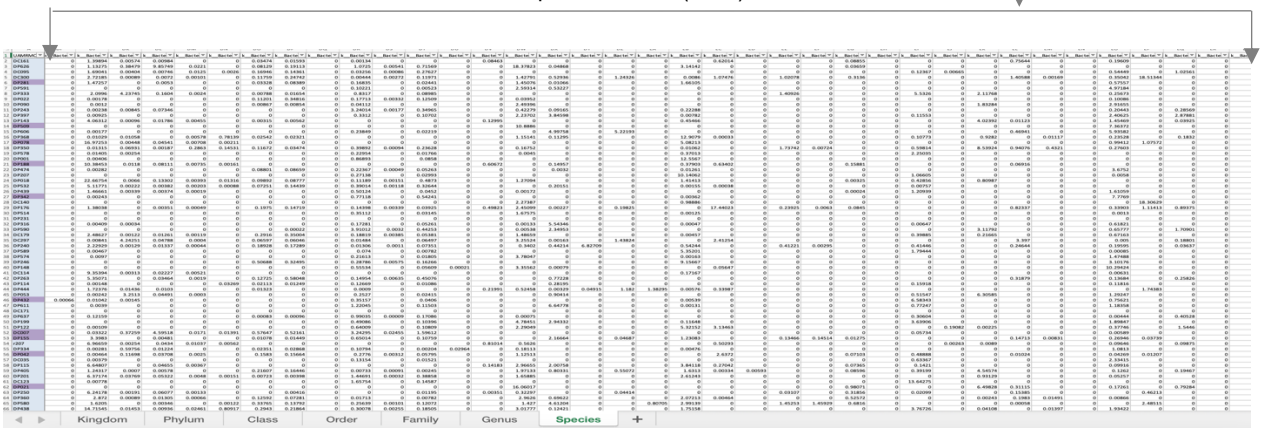
Sample collection for microbial DNA extraction

Metagenomic sequencing



Taxonomic assignment

Relative abundance of species (columns) per individual (rows)

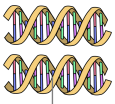


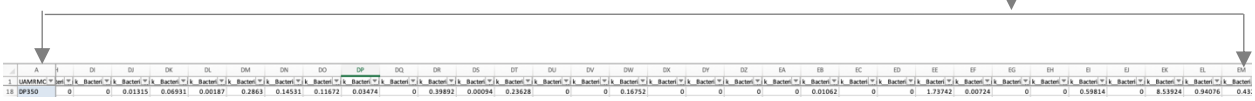
Kingdom | Phylum | Class | Order | Family | Genus | Species

Relative abundances
Range, mean
95% CI in
No Disease population

compare

Application





Relative abundances

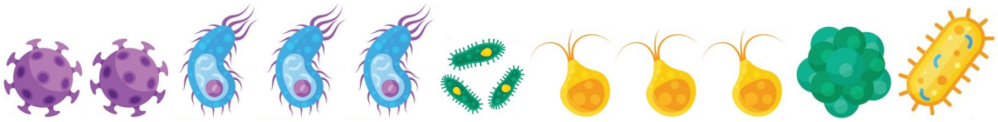


Figure 1. The method.

For Parkinson's disease, several disease-associated microbial features have already been well established, i.e., reproduced across studies at high statistical significance using multiple differential abundance test methods and including covariates and confounders [9-13]. Clinical trials of microbiome-based treatments are underway. The present paper describes a method to match patients to treatments, excluding patients who do not have the feature of interest, and enriching the trial with patients with high abundance of the feature being targeted.

A reference dataset is built by obtaining the metagenomic profiles of a non-disease cohort that represents the same population as the persons to be studied or treated. Once this infrastructure is in place, one can pick any feature (from a microbial gene to species to polymicrobial clusters), calculate its relative abundance and confidence interval (CI) in the non-disease group, and set thresholds for defining dysbiotic and non-dysbiotic. Application entails obtaining the metagenomic profile of the individual of interest and comparing it to the non-disease population. Here, we selected 6 PD-associated features (each a polymicrobial cluster), 5 of which are elevated in PD, and one reduced in PD [9]. We set the threshold at 95%CI of the non-disease reference dataset. A person would be marked as dysbiotic for a feature if their relative abundance of that feature is above (for features that are elevated in PD) or below (for feature that are reduced) the 95%CI in non-disease population.

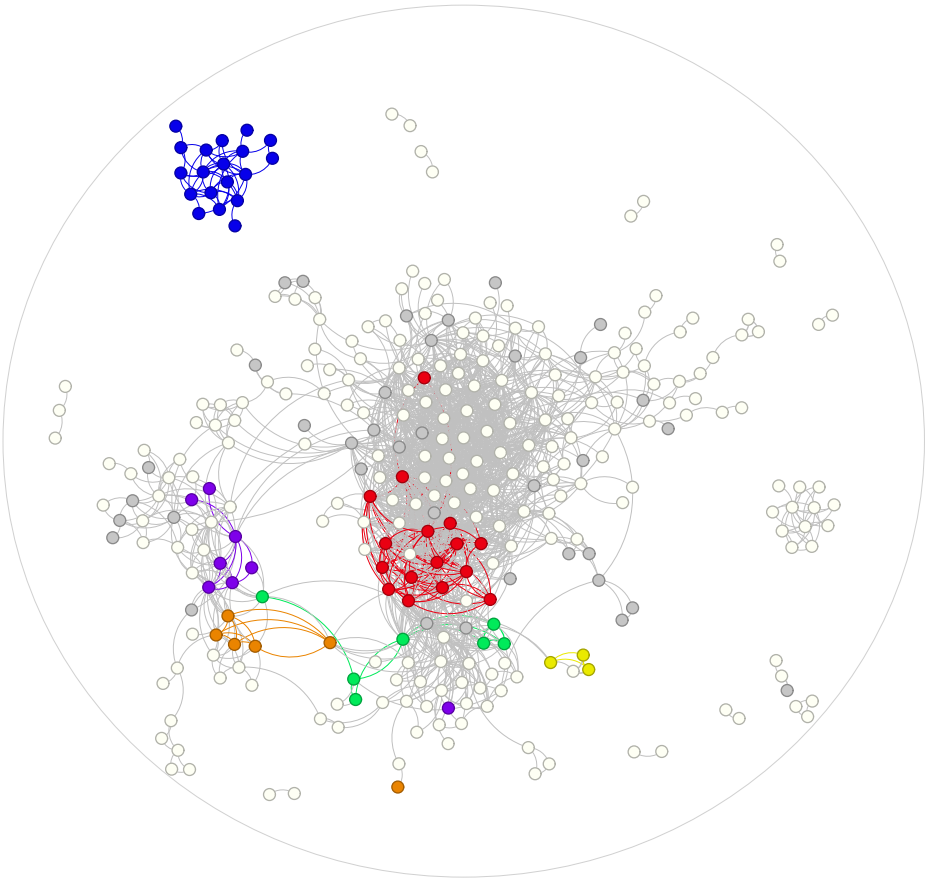


Figure 2. Polymicrobial clusters of species in PD gut microbiome.

Pairwise correlation in relative abundances of all species detected in 490 unique PD gut metagenomes (i.e. 490 individuals with PD) was calculated and plotted (methods are described in Wallen et al [9]). Each circle (node) represents one species and the curved lines (edges) connect species whose relative abundance correlate with each other (at correlation $|r|>0.2$, permuted $P<0.05$). Here, we have highlighted PD-associated species that were selected for this study, denoting by color the 6 dysbiotic features (species within each feature are listed in **Table 1**):

Blue: Opportunistic pathogens (elevated in PD)

Red: fiber-degraders (reduced in PD)

Yellow: *E. coli* and *Klebsiella* (elevated in PD)

Green: *Bifidobacteria* (elevated in PD)

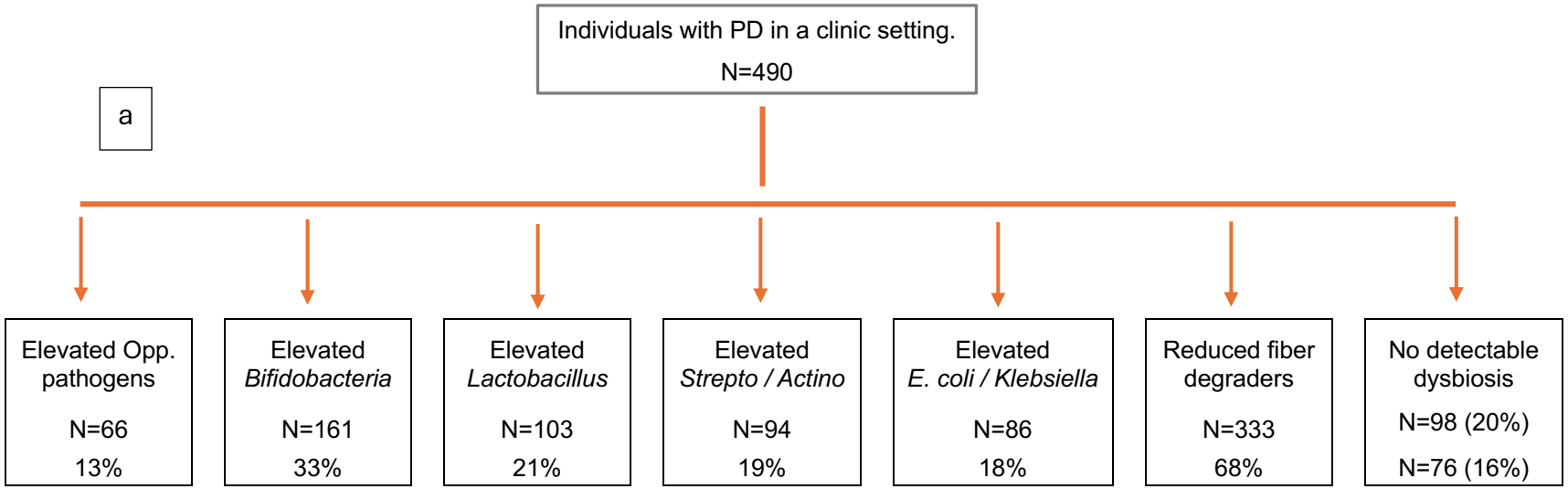
Orange: *Lactobacillus* (elevated in PD)

Purple: *Streptococcus* and *Actinomyces* (elevated in PD)

Grey: associated with PD, not selected as feature

Empty circle: not associated with PD

a



b

	Opp path	*Bifido-bacteria*	*Lacto-bacillus*	*Strepto /Actino*	*E. coli / Klebsiella*	Fiber-degraders
Opp Pathogens	66					
Bifidobacteria	33	161				
Lactobacillus	15	49	103			
Strepto / Actino	18	50	47	94		
E. coli/ Klebsiella	13	27	27	22	86	
Fiber-Degraders	57	135	88	81	69	333
No Others	5	18	5	5	9	96

c

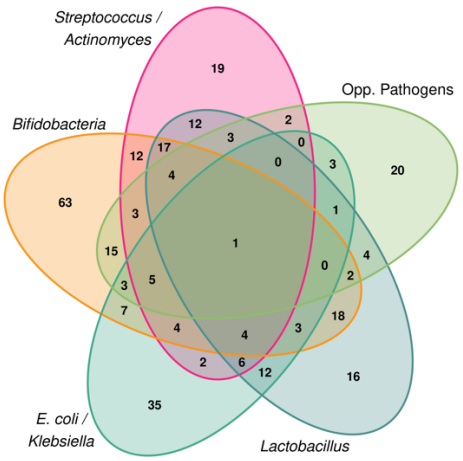


Figure 3. Subtyping PD by the dysbiotic features of the gut microbiome.

The total number of patients is 490. The total number of patients with a dysbiotic feature is 392 (=490-98). The number of patients with a given dysbiotic feature are given in a-c. Actual numbers are given so that percentages can be calculated with the desired denominator.

(a) Proportion of PD patients with different dysbiotic features.

- Elevated features: number of PD persons, among the 490, who had an abundance higher than upper limit of the 95% CI of the abundance in neurologically healthy control (NHC).
- Reduced feature: number of PD persons, among the 490, who had an abundance lower than the lower limit of 95% CI of the abundance in NHC.
- No detectable dysbiosis: 98 of 490 (20%) did not fall outside 95% CI of NHC for any feature. 76 of 490 (16%) of PD persons fit the more cautious definition of non-dysbiotic, whose abundances for every elevated feature (opportunistic pathogens, *Bifidobacteria*, *Lactobacillus*, *Streptococcus* and *Actinomyces* (Strepto/Actino), and *E. coli/Klebsiella*) was below the NHC mean, and their abundance for reduced fiber-degrading bacteria was above the NHC mean.

(b) Pairwise overlap in features, n=number of patients, from 490, who have the two specified features. The total for six features (the diagonal) is 843, nearly double the number of 490 subjects, because features do co-occur in an individual. For example, 66 of 490 subjects had elevated (>95%CI of NHC) levels of opportunistic pathogens, among them, 33 (50%) also had elevated Bifidobacteria. Or the example in the text, 71% (=(333-96)/333) of low fiber-degraders had high level of at least one other feature.

(c) Overlap across the 5 features that are elevated in PD (fiber-reducers not included for legibility). For example, 19 subjects have elevated levels (>95%CI of NHC) of *Streptococcus* and *Actinomyces* only, another 2 subjects have high *Streptococcus* and *Actinomyces* and high opportunistic pathogens, another 3 have high *Streptococcus* and *Actinomyces,* high opportunistic pathogens, and high *Lactobacillus*, etc.