

ARTICLE OPEN



Association between Parkinson's disease and the faecal eukaryotic microbiota

Severin Weis¹, Alexandra Meisner¹, Andreas Schwiertz², Marcus M. Unger³, Anouck Becker³, Klaus Faßbender³, Sylvia Schnell⁴, Karl-Herbert Schäfer⁵ and Markus Egert¹✉

Parkinson's disease (PD) is one of the most common neurodegenerative disease, and is so far not considered curable. PD patients suffer from several motor and non-motor symptoms, including gastrointestinal dysfunctions and alterations of the enteric nervous system. Constipation and additional intestinal affections can precede the classical motor symptoms by several years. Recently, we reported effects of PD and related medications on the faecal bacterial community of 34 German PD patients and 25 age-matched controls. Here, we used the same collective and analysed the V6 and V7 hypervariable region of PCR-amplified, eukaryotic 18S rRNA genes using an Illumina MiSeq platform. In all, 53% (18) of the PD samples and 72% (18) of the control samples yielded sufficient amplicons for downstream community analyses. The PD samples showed a significantly lower alpha and a different beta eukaryotic diversity than the controls. Most strikingly, we observed a significantly higher relative abundance of sequence affiliated with the *Geotrichum* genus in the PD samples (39.7%), when compared to the control samples (0.05%). In addition, we observed lower relative abundances of sequences affiliated with *Aspergillus/Penicillium*, *Charophyta/Linum*, unidentified Opisthokonta and three genera of minor abundant zooflagellates in the PD samples. Our data add knowledge to the small body of data about the eukaryotic microbiota of PD patients and suggest a potential association of certain gut eukaryotes and PD.

npj Parkinson's Disease (2021)7:101; <https://doi.org/10.1038/s41531-021-00244-0>

INTRODUCTION

Parkinson's disease (PD) represents the second most common human neurodegenerative disorder. So far, no treatment to stop the neurodegenerative process is available¹. PD patients not only suffer from motor-associated and rare cognitive symptoms, but also from gastrointestinal (GI) symptoms². GI-associated symptoms include constipation, prolonged intestinal transit time or defecation-associated dysfunctions and can precede the classical motor symptoms by several years^{2–6}. Additionally, there is rising evidence that even the enteric nervous system (ENS) becomes compromised before the central nervous system^{7–9}. These findings, alongside the growing numbers of published reports in this field, support the hypothesis that PD might begin, at least in a subgroup of patients, in the GI tract and propagate to the central nervous system^{10,11}. The ENS is connected to the central nervous system via the sympathetic nervous system and the vagus nerve which form the so called gut–brain axis¹². Experimental data suggest that this neuronal chain between ENS and the brain allows pathological peptides to propagate in a prion-like fashion between the gut and the brain, thereby modulating the course of this neurological disease^{13–15}.

Gut-associated PD symptoms, such as constipation, are discussed to be dependent on changes of the intestinal microbiota composition and their metabolic activity¹⁶. Recent publications suggest that especially bacterial dysbiosis might play an important role in PD pathogenesis^{17–20}. Other studies even indicate that the pathological process of PD alongside the gut–brain axis might be modulated or initiated by the gut microbiota^{21,22}. Increased inflammation in PD, which is indicated by increased levels of faecal markers for inflammation and gut permeability, is suggested to be linked to dysbiosis in the gut

environment^{23,24}. Additionally, elevated levels of pathogenic or opportunistic pathogenic bacteria were found to be independent of medication²⁵. The changes in relative abundance of some bacterial genera regarded as health promoting was found to be at least partially due to the medication²⁵. To better understand the influence of the gut microbiota on the pathobiology of PD, and the influence of the disease on the microbiota itself, a deeper understanding of compositional changes in the gut microbiota is needed. So far, most PD microbiota studies focused on intestinal bacteria, and there is indeed growing evidence that bacterial dysbiosis may play an important role in PD pathogenesis^{17–19}. However, the cause–effect relationship between PD pathogenesis and the structure and function of the gut microbiota is still obscure.

Besides bacteria, the gut microbiota also comprises archaea, viruses and various eukaryotic taxa including fungi. For bacteriophages, associations with shifts in the phagobiota in PD patients were reported, suggesting a closer look at the non-bacterial microbiota associated with this disease²⁶. With estimated 0.01–0.1% of genes in stool samples, the gut mycobiome represents only a small fraction when compared to the faecal bacterial community²⁷. Nevertheless, *Candida* species were reported to be present in about 70% of healthy adults²⁸. The majority of intestinal eukaryotes still await cultivation and are difficult to identify²⁹. However, the gut mycobiome was also reported to influence the gut–brain axis by secretion of neurotransmitters, modulation of cytokine production and production of short-chain fatty acids³⁰. Additionally, reports showed a connection between gut eukaryotes and anorexia nervosa, a central nervous disease³¹, as well as autism spectrum disorders³².

¹Faculty of Medical and Life Sciences, Institute of Precision Medicine, Microbiology and Hygiene Group, Furtwangen University, Villingen-Schwenningen, Germany. ²MVZ Institute of Microecology, Herborn, Germany. ³Department of Neurology, Saarland University, Homburg, Germany. ⁴Institute of Applied Microbiology, Justus-Liebig-University, Giessen, Germany. ⁵Working Group Enteric Nervous System (AGENS), University of Applied Sciences Kaiserslautern, Zweibrücken, Germany. ✉email: Markus.Egert@hs-furtwangen.de

In this study, we used next-generation sequencing to study differences in the faecal eukaryotic microbiota composition between a previously well studied group of PD patients and matched controls. To create a starting point and as only approximately 0.01% of human gut metagenomic sequences generated can be mapped to fungal genomes, a targeted, amplicon-based approach was chosen²⁷. We analysed the same set of samples previously used by Unger et al. (2016) and Weis et al. (2019). Here, we used Illumina MiSeq-based next-generation sequencing to sequence the V6 and V7 region of the eukaryotic 18S rRNA gene^{20,33}. Monitoring differences in the composition of faecal eukaryotes between PD patients and suitable healthy controls is a first step to elucidate whether gut eukaryotes might play any role in PD. Our study might help identifying organisms showing an association with PD and thereby increase our knowledge on the etiopathogenesis of PD.

RESULTS

Sequencing and bioinformatics

In all, 18 (53%) of the PD patient samples (6 females and 12 males) and 18 (72%) of the control group samples (10 females and 8 males) yielded sufficient 18S rRNA gene amplicons for downstream applications. Additional information about those samples is provided in Supplementary Table 1. The sequencing dataset comprised 493,979 partial 18S rRNA gene sequences after quality control, denoising, paired-end merging and chimera removal. The mean number of sequences per sample was 13,350 (min: 3895; max: 23,080). Assignment of taxonomy was done with the seven taxonomic ranks provided by SILVA, where the sixth rank equals genera for most listed taxa, the fifth lists mainly families, the fourth mainly orders, and so forth. Following rarefaction to even depth of sequences per sample, 1612 Amplicon Sequence Variants (ASV) affiliated with 122 genera-equivalent ranks in the SILVA database, 95 family-equivalent ranks and 24 order equivalent ranks were assigned. ASV rarefaction curves (Fig. 1) showed that the sequencing depth was sufficient to detect the vast majority of taxa in all samples.

Structural diversity measures

The three applied indices for alpha diversity (Fig. 2a), Observed, Shannon and Simpson, revealed a significantly lower eukaryotic diversity in PD patients compared to the controls (Ctrl) ($p_{\text{Observed}} < 0.002$, $p_{\text{Shannon}} < 0.001$, $p_{\text{Simpson}} < 0.002$).

For beta-diversity, non-parametric multivariate analysis of variance (ADONIS) was calculated for the different groups (Fig. 2b, c). Between the PD and the control group, unity of fraction (UniFrac) measures revealed significant differences regarding the

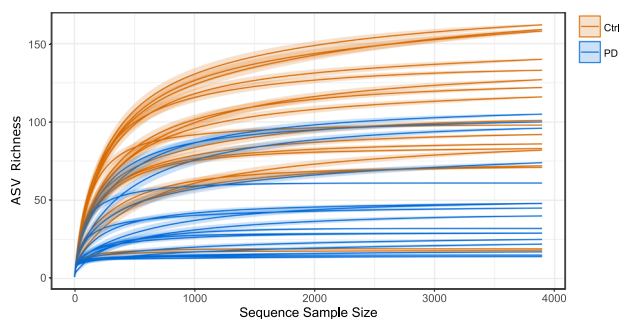


Fig. 1 Alpha rarefaction curves to visualize the dependence of eukaryotic richness and sequencing depth. Shown are amplicon sequence variation (ASV) richness rarefaction curves for the sequencing sample depth for all samples. Blue: PD samples, orange: controls. Lines represent the mean of 50 sampling steps, each, with the opaque outlines as standard errors.

weighted ($p_{\text{wUniFrac}} = 0.0073$) and the unweighted ($p_{\text{UniFrac}} < 0.0128$) UniFrac measure. For the unweighted UniFrac measure, a significant difference was found between controls and PD patients treated with entacapone ($p_{\text{UniFrac}} < 0.0483$) or L-dopa ($p_{\text{UniFrac}} < 0.0051$) as well as between controls and PD patients without entacapone ($p_{\text{UniFrac}} < 0.0012$) or L-dopa ($p_{\text{UniFrac}} < 0.0068$) treatment. For the weighted UniFrac measure, a significant difference was found between PD patients without entacapone treatment and the controls ($p_{\text{wUniFrac}} < 0.0003$) as well as PD

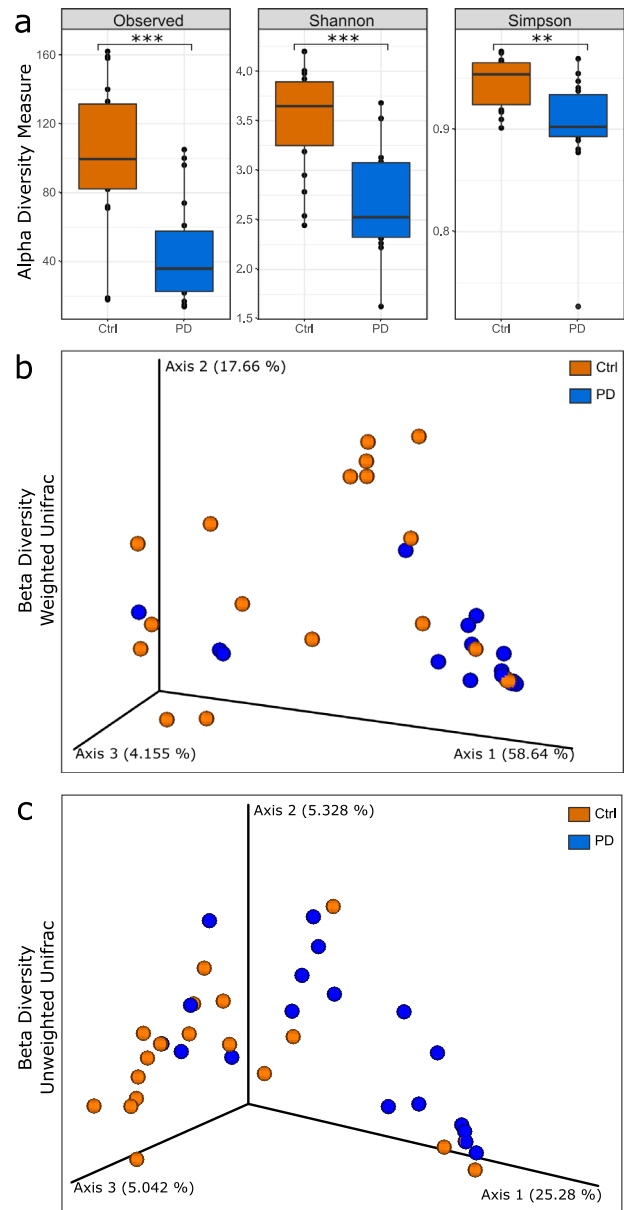


Fig. 2 Alpha and beta-diversity plots for visualization of the eukaryotic community structure and according to differences between the control group and PD patients. PD samples are displayed in blue and control samples in orange. With Observed, Shannon and Simpson, the three most common used indices for alpha diversity are shown in (a). Alpha-diversity boxplots consist of the median and lower and upper quartiles while the whiskers represent the minimal and maximal spread of individual samples. Significant differences between the groups are indicated with asterisks ($***p < 0.001$, $**p < 0.01$). For beta-diversity, PCoA plots of the weighted (b) and unweighted (c) UniFrac measures are shown. Both beta-diversity measures show significant differences between controls and PD.

patients with entacapone treatment ($p_{wUniFrac} < 0.0129$). For the medication with L-Dopa, significant differences in the beta diversity were found between controls and both PD patients without treatment ($p_{wUniFrac} < 0.0006$) and those with treatment ($p_{wUniFrac} < 0.0353$).

Compositional differences

Hypothesis testing revealed seven taxa on the genus-equivalent rank, assigned using the SILVA database, to be significantly different in relative abundance between PD and controls. The genera *Aspergillus* ($p_{FDR} < 0.031$), *Cercomonas* ($p_{FDR} < 0.037$) and *Heteromita* ($p_{FDR} < 0.022$), as well as three unknown genus-equivalent features of the phylum division Charophyta ($p_{FDR} < 0.019$), the order Chromulinales ($p_{FDR} < 0.013$) and the clade Opisthokonta ($p_{FDR} < 0.013$), were found to be significantly lower in relative abundance in the group of PD patients when compared to the controls. In contrast, ASVs affiliated with the genus *Geotrichum* ($p_{FDR} < 0.017$) were significantly higher in relative abundance in the group of PD patients. As displayed in Table 1, the taxonomy assignment of the representative sequence with nucleotide-based basic local alignment search tool (BLASTN) for the nucleotide database provided by the National Center for Biotechnology Information (NCBI) revealed different species for those features. The genera *Geotrichum*, *Aspergillus*, *Cercomonas* and *Heteromita* were identified as being closely related (sequence identity >99%) to *Geotrichum candidum*, *Penicillium roqueforti*, *Paracercomonas* sp. and *Cercozoa* sp. B134, respectively. The highest degree of sequence identity (97–99%) for the unknown Charophyta taxon was found with *Linum usitatissimum* and with *Poterioochromonas malhamensis* in case of the unknown Chromulinales taxon. For Opisthokonta no sequences with identity >90% were found.

Using analysis of composition of microbiomes (ANCOM-II³⁴) with the non-condensed ASV table, all 12 ASVs assigned to *Geotrichum* were found to be significantly different between controls and PD patients ($W > 0.8$, $p_{FDR} < 0.05$). All of those ASV were then identified as *Geotrichum candidum* (sequence identity >98.7%) using BLASTN with the NCBI database. However, no other ASV was found to be significantly differing between PD patients and controls according to ANCOM-II, even with a W -cut-off at 0.6.

Samples from PD patients with elevated levels of faecal calprotectin ($n = 6$) showed no significant differences in community composition when compared to the controls or PD patients with normal calprotectin levels ($n = 12$). PD patients treated with L-dopa ($n = 12$) showed significantly higher relative abundances of *Geotrichum* ($p_{FDR} = 0.0138$). However, there were no significant differences between PD patients who were not treated with L-dopa ($n = 6$) and the control group or the PD patients treated with L-dopa. PD patients treated with Entacapone in addition to L-dopa ($n = 7$) also showed significant higher shares of *Geotrichum* when compared to the controls ($p_{FDR} = 0.0069$). Finally, no significant differences were found between PD patients without Entacapone treatment ($n = 11$) and the control group, or between PD patients without Entacapone treatment and PD patients treated with Entacapone.

Correlation analysis using Spearman rho and Pearson rho did not reveal any significant correlation between the count data of the *Geotrichum*-affiliated sequences and the count data of any bacterial genera from the same samples, investigated in our previous study²⁰.

DISCUSSION

Several publications addressed the composition of the intestinal bacterial microbiota in PD patients by means of qPCR³³ or sequencing of different regions of 16S rRNA gene amplicons, using pyrosequencing^{21,35}, Ion Torrent²⁰ and Illumina MiSeq

technology^{20,36,37}. However, only little is known about associations and potential functional implications of the eukaryotic gut microbiota with PD. Recently, sequencing of fungal-specific internal transcribed spacer (ITS)-2 amplicons, derived from faecal samples of a cohort of PD patients, did not reveal any genetic differences in relative abundance or diversity between PD and controls³⁸. The authors also reported a very low fungal load for their samples and significantly lower amounts of fungal DNA in PD patients³⁸. This finding might partly explain why 28% (controls) and 47% (PD) of our samples did not yield sufficient amplicons for downstream analyses. However, we found significant differences in relative abundance for several eukaryotic taxa, including fungi, between the remaining PD and control samples.

Our data indicate a significantly decreased richness of eukaryotes according to the observed richness, Simpson metric and Shannon metric (Fig. 2a). This is in contrast to Cirstea et al.³⁸ who did not report changes in richness, which might be due to fundamental differences in the sample groups, such as dietary habits, geographic origin or medical history and status. However, our finding matches previously published decreases in bacterial richness^{20,36,39}. The beta diversity measures also showed significant differences between PD and control samples for the unweighted and weighted UniFrac metric (Fig. 2b, c), suggesting substantial differences in community composition between the eukaryotic microbiota of PD patients and the healthy control group. Therefore, our expectation was to find significant differences in the relative abundance of several taxa.

For the bacterial community of this study cohort, we previously reported a significantly lower observed alpha diversity under PD, too²⁰. However, in contrast to the eukaryotic data presented here, indicators assuming uniformity of the distribution (Shannon and Simpson) and all beta-diversity parameters did not show significant differences between controls and PD patients²⁰, suggesting that eukaryotic and prokaryotic diversity might behave differently under PD. However, these differences might also be due to the smaller sample size analysed here or the use of different sequencing methods.

Even though we found seven genus-equivalent taxa being significantly different in their relative abundance between control and PD samples, only *Geotrichum* was found to be significantly different when using ANCOM-II, i.e., when using a compositional approach³⁴. It might be speculated that the observed depletion of the other, minor abundant groups is just due to the overgrowth of *Geotrichum* in the PD samples rather than actual depletion. Further studies might reveal whether there are significant differences in other groups, such as protozoa too. Our data at least suggest that besides moulds like *Geotrichum* also other eukaryotic genera might differ between PD patients and controls. Although bacterivorous gliding zooflaellates such as *Cercomonas* (*Paracercomonas*), *Heteromita* (*Cercozoa* sp. B134) and *Poterioochromonas* are not regarded as typical representatives of the human gut microbiota, grazing of protists is often selective and might influence the virulence, metabolism and morphology of intestinal bacteria^{40–43}. *Penicillium* and *Aspergillus* are commonly found in the healthy human gut⁴⁴. Moulds are known to produce various mycotoxins and to influence the mucosal cytokine response, which affect GI homeostasis and the composition of the bacterial gut microbiota^{45,46}. Finally, some differences in relative abundance of eukaryotic taxa such as *Linum usitatissimum* or *Penicillium roqueforti* might just result from differential nutrition.

In contrast to the other genera, the family Dipodascaceae and its affiliated genus *Geotrichum* showed a markedly higher relative abundance in the PD samples compared to the controls. Dipodascaceae are commonly found in human stool samples and are potentially able to colonize the gut⁴⁴. Few species are regarded as a health issue and capable of causing geotrichosis, including *Geotrichum candidum*, which produces several toxins⁴⁷. However, this species is also used in dairy production⁴⁸, where it is

Table 1. Taxa differing significantly between PD samples and controls.

SILVA 132	Opisthokonta	Archaeplastida	Opisthokonta	Opisthokonta	Opisthokonta	SAR	SAR
Assignment	Nucleotmycea	Chloroplastida	Nucleotmycea	Nucleotmycea	Rhizaria	Stramenopiles	Rhizaria
"Order"	Fungi	Charophyta	Fungi	NA	Cercozoa	Ochrophyta	Cercozoa
"Family"	Dipodascaceae	NA	Aspergillaceae	NA	Cercomonadidae	Chromulinales	Glissomonadida
"Genus"	<i>Geotrichum</i>	NA	<i>Aspergillus</i>	NA	<i>Cercomonas</i>	NA	<i>Heteromita</i>
NCBI nucleotide	<i>Geotrichum candidum</i>	<i>Linum usitatissimum</i>	<i>Penicillium roqueforti</i>	<i>Paracercomonas</i> sp.	<i>Poterioochromonas malhamensis</i>	<i>Poterioochromonas malhamensis</i>	<i>Cercozoa</i> sp. B134
MegaBLAST	573	555	564	No sequences	569	560	566
Total score	573	1108	564	With identity	569	560	566
Query cover	100%	100%	100%	>90%	99%	100%	100%
E value	3.00E-159	1.00E-153	2.00E-156		3.00E-158	2.00E-155	4.00E-157
% Identity	99.37%	98.72%	99.36%		99.68%	98.73%	99.05%
Acc. length	1556	18299719	1775		1803	2517	1511
Accession #	KU899094.1	CP027631.1	MT544459.1	FP929054.1	MG775632.1	MH536660.1	HQ918175.1
Mean (Ctrl)	0.05%	12.24%	1.83%	2.97%	0.32%	0.81%	0.31%
SE (Ctrl)	0.03%	2.36%	0.32%	0.40%	0.05%	0.10%	0.05%
Mean (PD)	39.65%	3.07%	1.62%	0.61%	0.09%	0.08%	0.07%
SE (PD)	5.57%	0.99%	0.50%	0.12%	0.02%	0.02%	0.02%
P value	0.0004	0.0006	0.0015	0.0002	0.0021	0.0001	0.0009
P (FDR)	0.0161	0.0182	0.0303	0.0121	0.0363	0.0121	0.0218
ANCOM-II	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

Shown are taxa that differed significantly between control (Ctrl) and Parkinson's disease (PD) samples together with the annotation of the representative sequences based on the SILVA database entry in QIIME2 (SILVA) and the NCBI database using megablast. For each taxon, mean relative abundance (abun.), standard deviation (SD) and prevalence per subgroup are displayed, too. P values for hypothesis testing between PD and control samples with the two-sided Wilcoxon-Mann-Whitney test for unpaired and non-normally distributed samples (p value Ctrl vs. PD) were subjected to false discovery correction for multiple testing (FDR). NA not assigned.

also regarded as spoilage organism⁴⁸ causing opportunistic infections in immunocompromised patients⁴⁹. An increased abundance of lactococcal phages, which are also present in dairy products, was suggested to cause depletion of *Lactococcus* in PD patients²⁶. Together with our data, such observations suggest a special influence of dietary habits, in particular dairy products, on the intestinal microbiota of PD patients. It might be also be speculated that PD patients offer more suitable intestinal conditions for this species than healthy patients do. Due to its marked and significant differences in relative abundance between controls and PD patients, *Geotrichum* was used for a correlation analysis with our previously published bacterial 16S rRNA gene data²⁰. However, after false discovery rate correction no significant correlation with any bacterial genus was found, suggesting no direct fungus–bacterium associations here. However, also technological reasons (e.g. the differing sequencing technologies) or the time gap between the studies might at least partially explain these missing correlations.

Clearly, the respective subsample sizes in our study were small, also due to the discarded samples, so that all hypotheses need to be verified. Additionally, some findings might also be caused by different food preferences of the investigated patients, although all of them reported an omnivorous European diet. Clearly, further analyses are needed to verify whether the observed differences in community composition might be of biological/medical relevance in the case of PD. To do so, more complex analyses using multiple variable regions of the 18S rRNA gene sequence and ITS region sequences or even a metagenomics approach will be needed⁵⁰. Since only 0.01% of metagenomics sequences generated from human gut samples can be aligned to fungal genomes, such metagenomics approaches require sufficient sequencing depth²⁷.

We were able to show that PD faecal samples contained a eukaryotic microbial community of lower diversity and that several eukaryotic taxa differed in relative abundance between PD and control faecal samples, suggesting an association of PD and the eukaryotic microbiota. Within the PD group, we particularly observed a very high relative abundance of *Geotrichum*, a fungal genus commonly found in the human gut and usually not regarded as an intestinal pathogen. In addition, difference in a few minor abundant groups were detected. Future studies will have to show, whether these findings are reproducible, of functional relevance or diagnostic value. Despite sampling issues (sample age, small sample size), environmental factors, such as geographic and cultural background of the studied cohort and the used sequencing techniques might explain differences to previous studies. Clearly, studies with larger patient and control cohorts are needed for deeper insights into the potential links between gut eukaryotes and PD, ideally combining metagenomic, metatranscriptomic and/or metabolomic approaches.

METHODS

Cohort, sample collection and DNA isolation

All 34 PD patients (10 females, 24 males), whose faecal samples were used in this work, were diagnosed according to the UK PD Society Brain Bank Clinical Diagnostic Criteria⁵¹. The 25 control persons (14 females, 11 males) were age matched and did not report any pre-existing medical conditions or medication. All subjects followed an omnivorous diet without special dietary habits, and did not report on acute or chronic gastrointestinal disorders or intake of antibiotics, pro- or prebiotics up to 3 months prior to sampling. For more sample details, sample collection and DNA isolation, the reader is referred to our previous publication on the bacterial community composition of the same study cohort²⁰. In short, subjects were provided with sterile faecal collector sets (MED AUXIL, Süsse Labortechnik, Gudensberg, Germany) for home sampling. Samples were stored and transported at least at -20°C and DNA was isolated using the FastDNA SPIN kit for faeces (MP Biomedicals, Heidelberg, Germany). The study was approved by the ethics committee of the medical association of the German federal state of Saarland and is recorded therein with the

identification number 111/12. All enrolled subjects provided written informed consent for their participation.

Library preparation and sequencing

Sequencing library preparation of the V6 and V7 region of eukaryotic 18S rRNA genes were performed using the 18S-specific primers 1152F (5'-TGAACTTAAAGRAATTGACGGA-3') and 1428R (5'-GGRCATMACDGCCT-GYTAT-3')²⁷ with additional adapter sequences for the Illumina Nextera indexing to produce amplicons of approximately 250–380 bp length. Replicate PCR amplification per sample, amplicon verification and purification were performed twice using the KAPA HiFi Polymerase-Kit (Roche, Mannheim, Germany) as described elsewhere⁵². Fifty microliters PCR were set up with a 5 μl DNA template, 30 amplification cycles (initial denaturation: 3 min, 95°C ; 30 cycles: 30 s, 95°C , 30 s, 55°C , 45 s, 72°C ; final elongation: 5 min, 72°C), *Saccharomyces cerevisiae* DNA as positive controls, and water-template controls. Verification of amplification by agarose gel electrophoresis using Midori Green as DNA-dye (Biozym, Olderdorf, Germany) showed that 18 control samples and 18 PD samples produced sufficient amplicons for downstream analyses. Metadata for these remaining samples is provided in Supplementary Table 1. Successful replicate PCR reactions of the same samples were pooled and purified using Agencourt AMPure beads (Beckman Coulter, Krefeld, Germany).

Unique index barcodes with sequencing adaptors were added to the amplicon targets as described previously⁵² using the unique indexing primer combinations of the Nextera XT index kit v2 set D (Illumina), 5 μl of purified template, and the PCR profile detailed above, albeit with 8 cycles and a total volume of 25 μl . After another purification step, quality checks, checks for library size and DNA concentration using Agilent DNA 1000 chips in the Agilent bioanalyzer (Agilent Technologies, Waldbronn, Germany) and the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Schwerte, Germany) were carried out as described before⁵²; libraries were normalized to 4 nM and pooled for sequencing. By following the manufacturer's instructions, the pooled libraries were sequenced in a final concentration of 6 pM on an Illumina MiSeq platform (Illumina). For the targeted amplicon sequencing procedure, a 600 cycle paired-end format ($2 \times 300 \text{ bp} + 2 \times 8 \text{ bp}$ index cycles) was chosen using the MiSeq Reagent Kit v3 with 20% phiX spike in in the pooled library.

Bioinformatics and statistics

Sequence data were processed using QIIME 2 version 2019.7 (ref. ⁵³). Quality cut-offs were performed with the median at $Q \geq 25$. Minimum and maximum sequence lengths of 200 and 1000 bp were used (QIIME2 default). Denoising, merging paired-ends and removal of chimeras were performed using the DADA2 pipeline of QIIME2. Amplicon sequence variations (ASV) were chosen within 99% sequence identity. A primer-fitted taxonomy classifier, trained with the SILVA database release 132 using the classify-sklearn algorithm in QIIME2, was used to assign taxonomy and align sequences⁵⁴.

Using a QIIME2-generated ASV table, normalized by rarefaction to even depth, alpha-diversity indices for Observed, Shannon and Simpson metrics, as well as beta-diversity indices for weighted and unweighted UniFrac were calculated with the according QIIME2 plugins. PCoA plots were also produced with QIIME2 using the previously calculated distance matrices. Hypothesis testing of the alpha- and beta-diversity measures were also performed in QIIME2 using Kruskal–Wallis statistics for alpha diversity and PERMANOVA with 10,000-fold permutations for beta-diversity measures.

Further statistical analyses were calculated within R (version 4.0.2). The phyloseq package (version 1.32.0) was used as interface between QIIME2 and R to prepare the previously rarefied table provided by QIIME2 by condensing ASVs onto the genus-equivalent rank and calculating alpha diversity boxplots with the QIIME2 provided rarefied table⁵⁵. To find significant differences in feature counts between PD and controls, a two-sided Wilcoxon–Mann–Whitney test for unpaired and non-normally distributed samples in a 10,000-fold Monte-Carlo simulation within the R package coin (version 1.3.1) was used⁵⁶. For compositional data analysis, the statistical framework “Analysis of Composition of Microbiomes” (ANCOM⁵⁷) in the zero inflation avoiding version ANCOM-II³⁴ was used in an ASV-wise manner with the non-condensed ASV table to compare compositional differences in the the Control and PD groups. Zero- and out-cut-off were kept at the default of 0.90 and 0.05, respectively. The FDR adjusted alpha-value was set as 0.05 and for the *W*-statistic cut-off the arbitrary numbers 0.7 and 0.8 were used. All *p* values from hypothesis

testing were corrected for the false discovery rate (FDR) in their corresponding group⁵⁸.

Correlation analysis for Spearman correlation index and Pearson correlation of the merged count data of the *Geotrichum* genus with all bacterial genera published before²⁰ of the same samples successfully sequenced in this work. Correlations were calculated using the cor.test function of the stats package in R⁵⁹. Additionally, correlations between *Geotrichum* and bacterial genera were calculated for the PD group only.

For all ASVs or ASV clusters differing significantly in relative abundance in the performed hypothesis testing, the representative sequence was used for further comparison with the National Center for Biotechnology Information (NCBI) nucleotide database. Here, the nucleotide-based basic local alignment search tool (BLAST) with the megablast option without further exclusions or limitations against the standard databases provided by NCBI was used.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

Sequences generated and analysed here are accessible at the European Nucleotide Archive (ENA) under the accession number PRJEB45549. Subject metadata is included in the Supplementary Files. Complete, rarefied feature and taxonomy tables with provenance (qza-format), as well as other datasets are available from the corresponding author on reasonable request.

CODE AVAILABILITY

Only open source code from QIIME2 (2019.7) and the cited R (4.0.2) packages (phyloseq 1.32.0, coin 1.3.1) were used, using either the default settings or the settings stated in the “Methods” section. Further information for clarification is available from the corresponding author on reasonable request.

Received: 3 February 2021; Accepted: 21 October 2021;

Published online: 18 November 2021

REFERENCES

- Perez-Pardo, P. et al. The gut-brain axis in Parkinson's disease: possibilities for food-based therapies. *Eur. J. Pharmacol.* **817**, 86–95 (2017).
- Savica, R. et al. Medical records documentation of constipation preceding Parkinson disease: a case-control study. *Neurology* **73**, 1752–1758 (2009).
- Fasano, A. et al. The role of small intestinal bacterial overgrowth in Parkinson's disease. *Mov. Disord.* **28**, 1241–1249 (2013).
- Jost, W. H. Gastrointestinal dysfunction in Parkinson's disease. *J. Neurol. Sci.* **289**, 69–73 (2010).
- Pfeiffer, R. F. Gastrointestinal dysfunction in Parkinson's disease. *Parkinsonism Relat. Disord.* **17**, 10–15 (2011).
- Cersosimo, M. G. & Benarroch, E. E. Pathological correlates of gastrointestinal dysfunction in Parkinson's disease. *Neurobiol. Dis.* **46**, 559–564 (2012).
- Chen, H. et al. Meta-analyses on prevalence of selected Parkinson's nonmotor symptoms before and after diagnosis. *Transl. Neurodegener.* **4**, 1 (2015).
- Abbott, R. D. et al. Frequency of bowel movements and the future risk of Parkinson's disease. *Neurology* **57**, 456–462 (2001).
- Gao, X., Chen, H., Schwarzschild, M. A. & Ascherio, A. A prospective study of bowel movement frequency and risk of Parkinson's disease. *Am. J. Epidemiol.* **174**, 546–551 (2011).
- Hawkes, C. H., Del Tredici, K. & Braak, H. Parkinson's disease: a dual-hit hypothesis. *Neuropathol. Appl. Neurobiol.* **33**, 599–614 (2007).
- Horsager, J. et al. Brain-first versus body-first Parkinson's disease: a multimodal imaging case-control study. *Brain* **143**, 3077–3088 (2020).
- Cryan, J. F. & Dinan, T. G. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* **13**, 701–712 (2012).
- Braak, H. & Del Tredici, K. Neuropathological staging of brain pathology in sporadic Parkinson's disease: separating the wheat from the chaff. *J. Parkinsons Dis.* **7**, S71–S85 (2017).
- Pan-Montojo, F. et al. Progression of Parkinson's disease pathology is reproduced by intragastric administration of rotenone in mice. *PLoS ONE* **5**, e8762 (2010).
- Pan-Montojo, F. et al. Environmental toxins trigger PD-like progression via increased alpha-synuclein release from enteric neurons in mice. *Sci. Rep.* **2**, 898 (2012).
- Zhao, Y. & Yu, Y.-B. Intestinal microbiota and chronic constipation. *Springerplus* **5**, 1130 (2016).
- Sampson, T. R. et al. Gut microbiota regulate motor deficits and neuroinflammation in a model of Parkinson's disease. *Cell* **167**, 1469–1480.e12 (2016).
- Choi, J. G. et al. Oral administration of *Proteus mirabilis* damages dopaminergic neurons and motor functions in mice. *Sci. Rep.* **8**, 1275 (2018).
- Sun, M.-F. et al. Neuroprotective effects of fecal microbiota transplantation on MPTP-induced Parkinson's disease mice: gut microbiota, glial reaction and TLR4/TNF- α signaling pathway. *Brain Behav. Immun.* **70**, 48–60 (2018).
- Weis, S. et al. Effect of Parkinson's disease and related medications on the composition of the fecal bacterial microbiota. *NPJ Parkinsons Dis.* **5**, 28 (2019).
- Keshavarzian, A. et al. Colonic bacterial composition in Parkinson's disease. *Mov. Disord.* **30**, 1351–1360 (2015).
- Forsyth, C. B. et al. Increased intestinal permeability correlates with sigmoid mucosa alpha-synuclein staining and endotoxin exposure markers in early Parkinson's disease. *PLoS ONE* **6**, e28032 (2011).
- Schwartz, A. et al. Fecal markers of intestinal inflammation and intestinal permeability are elevated in Parkinson's disease. *Parkinsonism Relat. Disord.*, <https://doi.org/10.1016/j.parkreldis.2018.02.022> (2018).
- 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).
- Wallen, Z. D. et al. Characterizing dysbiosis of gut microbiome in PD: evidence for overabundance of opportunistic pathogens. *NPJ Parkinsons Dis.* **6**, 11 (2020).
- Tetz, G., Brown, S. M., Hao, Y. & Tetz, V. Parkinson's disease and bacteriophages as its overlooked contributors. *Sci. Rep.* **8**, 10812 (2018).
- Nash, A. K. et al. The gut mycobiome of the Human Microbiome Project healthy cohort. *Microbiome* **5**, 153 (2017).
- Schulze, J. & Sonnenborn, U. Yeasts in the gut: from commensals to infectious agents. *Dtsch. Arzteblatt Int.* **106**, 837–842 (2009).
- Huffnagle, G. B. & Noverr, M. C. The emerging world of the fungal microbiome. *Trends Microbiol.* **21**, 334–341 (2013).
- Enaud, R. et al. The mycobiome: a neglected component in the microbiota-gut-brain axis. *Microorganisms* **6**, <https://doi.org/10.3390/microorganisms6010022> (2018).
- Gouba, N., Raoult, D. & Drancourt, M. Gut microeukaryotes during anorexia nervosa: a case report. *BMC Res. Notes* **7**, 33 (2014).
- Strati, F. et al. New evidences on the altered gut microbiota in autism spectrum disorders. *Microbiome* **5**, 24 (2017).
- Unger, M. M. et al. Short chain fatty acids and gut microbiota differ between patients with Parkinson's disease and age-matched controls. *Parkinsonism Relat. Disord.* **32**, 66–72 (2016).
- Kaul, A., Mandal, S., Davidov, O. & Peddada, S. D. Analysis of microbiome data in the presence of excess zeros. *Front. Microbiol.* **8**, 2114 (2017).
- Scheperjans, F. et al. Gut microbiota are related to Parkinson's disease and clinical phenotype. *Mov. Disord.* **30**, 350–358 (2015).
- Hopfner, F. et al. Gut microbiota in Parkinson disease in a northern German cohort. *Brain Res.* **1667**, 41–45 (2017).
- Hill-Burns, E. M. et al. Parkinson's disease and Parkinson's disease medications have distinct signatures of the gut microbiome. *Mov. Disord.* **32**, 739–749 (2017).
- Cirstea, M. S. et al. The gut mycobiome in Parkinson's disease. *J. Parkinsons Dis.*, <https://doi.org/10.3233/JPD-202237> (2020).
- Li, W. et al. Structural changes of gut microbiota in Parkinson's disease and its correlation with clinical features. *Sci. China Life Sci.* **60**, 1223–1233 (2017).
- Adiba, S., Nizak, C., van Baalen, M., Denamur, E. & Depaulis, F. From grazing resistance to pathogenesis: the coincidental evolution of virulence factors. *PLoS ONE* **5**, e11882 (2010).
- Hahn, M. W. & Höfle, M. G. Grazing of protozoa and its effect on populations of aquatic bacteria. *FEMS Microbiol. Ecol.* **35**, 113–121 (2001).
- Sun, S., Noorian, P. & McDougald, D. Dual role of mechanisms involved in resistance to predation by protozoa and virulence to humans. *Front. Microbiol.* **9**, 1017 (2018).
- Matz, C. & Kjelleberg, S. Off the hook-how bacteria survive protozoan grazing. *Trends Microbiol.* **13**, 302–307 (2005).
- Hallen-Adams, H. E. & Suhr, M. J. Fungi in the healthy human gastrointestinal tract. *Virulence* **8**, 352–358 (2017).
- Kulas, J. et al. Pulmonary *Aspergillus fumigatus* infection in rats affects gastrointestinal homeostasis. *Immunobiology* **224**, 116–123 (2019).
- Dubey, M. K. et al. PR toxin—biosynthesis, genetic regulation, toxicological potential, prevention and control measures: overview and challenges. *Front. Pharmacol.* **9**, <https://doi.org/10.3389/fphar.2018.00288> (2018).
- Rippon, J. W. *Medical Mycology: The Pathogenic Fungi and the Pathogenic Actinomycetes*. 3rd edn (Saunders, Philadelphia, 1988).
- Eliskases-Lechner, F., Guéguen, M. & Panoff, J. M. In Fuquay, J. W. (ed.) *Encyclopedia of Dairy Sciences* Vol. 102, 765–771 (Elsevier, 2011).
- Pottier, I., Gente, S., Vernoux, J.-P. & Guéguen, M. Safety assessment of dairy microorganisms: *Geotrichum candidum*. *Int. J. Food Microbiol.* **126**, 327–332 (2008).

50. Bass, D. et al. Phylogeny and classification of Cercomonadida (Protozoa, Cercozoa): *Cercomonas*, *Eocercomonas*, *Paracercomonas*, and *Cavernomonas* gen. nov. *Protist* **160**, 483–521 (2009).
51. Massano, J. & Bhatia, K. P. Clinical approach to Parkinson's disease: features, diagnosis, and principles of management. *Cold Spring Harb. Perspect. Med.* **2**, a008870 (2012).
52. Hantschel, J. et al. Effect of endometriosis on the fecal bacteriota composition of mice during the acute phase of lesion formation. *PLoS ONE* **14**, e0226835 (2019).
53. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
54. Pruesse, E. et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35**, 7188–7196 (2007).
55. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
56. Hothorn, T., Hornik, K., van de Wiel, Mark, A. & Zeileis, A. A Lego system for conditional inference. *Am. Stat.* **60**, 257–263 (2006).
57. Mandal, S. et al. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb. Ecol. Health Dis.* **26**, 27663 (2015).
58. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. A Method* **57**, 289–300 (1995).
59. Best, D. J. & Roberts, D. E. Algorithm AS 89: the upper tail probabilities of Spearman's Rho. *Appl. Stat.* **24**, 377 (1975).

ACKNOWLEDGEMENTS

The authors wish to thank the German Federal Ministry of Education and Research (BMBF) for funding this research (Project immENS, grant number 03FH036PB5).

AUTHOR CONTRIBUTIONS

A.S., M.M.U., A.B. and K.F. conceived the clinical setup and executed the sampling, the storage, and the distribution of the samples. S.W. performed the DNA isolation. A.M. performed the quality control and sequencing procedure. S.W. performed the bioinformatics procedures. S.W., A.M., S.S., A.S., M.M.U., K.-H.S. and M.E. contributed to the manuscript. All authors approved the final version of this article.

FUNDING

Open Access funding enabled and organized by Projekt DEAL.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41531-021-00244-0>.

Correspondence and requests for materials should be addressed to Markus Egert.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2021