

## ***Gut microbiome signatures of risk and prodromal markers of Parkinson's disease***

Sebastian Heinzel, PhD<sup>1,\*</sup>, Velma T. E. Aho, PhD<sup>2,3</sup>, Ulrike Suenkel, MD<sup>4</sup>, Anna-Katharina von Thaler, PhD<sup>4</sup>, Claudia Schulte, MSc<sup>4,5</sup>, Christian Deuschle<sup>4,5</sup>, Lars Paulin, MSc<sup>2</sup>, Sari Hantunen, PhD<sup>6</sup>, Kathrin Brockmann, MD<sup>4,5</sup>, Gerhard W. Eschweiler, MD<sup>7,8</sup>, Walter Maetzler, MD<sup>1</sup>, Daniela Berg, MD<sup>1,4</sup>, Petri Auvinen, PhD<sup>2</sup>, Filip Schepersjans, MD<sup>3</sup>

<sup>1</sup>Department of Neurology, Christian-Albrechts-University, Kiel, Germany

<sup>2</sup>Institute of Biotechnology, University of Helsinki, Helsinki, Finland

<sup>3</sup>Department of Neurology, Helsinki University Hospital and Department of Clinical Neurosciences (Neurology), University of Helsinki, Helsinki, Finland

<sup>4</sup>Department of Neurodegeneration, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

<sup>5</sup>German Center for Neurodegenerative Diseases, University of Tübingen, Tübingen, Germany

<sup>6</sup>Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland

<sup>7</sup>Department of Psychiatry and Psychotherapy, Geriatric Center, Tübingen University Hospital, Tübingen, Germany

<sup>8</sup>Geriatric Center at the University Hospital Tübingen, Tübingen, Germany

\*Correspondence to: Dr. Sebastian Heinzel, Department of Neurology, Christian-Albrecht-University, Kiel, Germany. Arnold-Heller-Straße 3, 24105 Kiel, Germany; Phone ++49 431 500-19010; Fax: +49 431 500-23849; s.heinzel@neurologie.uni-kiel.de

Email addresses of all authors: s.heinzel@neurologie.uni-kiel.de; velma.aho@helsinki.fi; ulrike.suenkel@uni-tuebingen.de; anna-katharina.von-thaler@uni-tuebingen.de; claudia.schulte@uni-tuebingen.de; christian.deuschle@medizin.uni-tuebingen.de; sari.hantunen@uef.fi, kathrin.brockmann@uni-tuebingen.de, gerhard.eschweiler@med.uni-tuebingen.de; lars.paulin@helsinki.fi, w.maetzler@neurologie.uni-kiel.de; daniela.berg@uksh.de; petri.auvinen@helsinki.fi; filip.schepersjans@hus.fi

***Running head:*** “Gut microbiome & risk and prodromal markers in PD”

***Keywords:*** Microbiome, microbiota, Parkinson's disease, prodromal markers, risk factors, age, physical activity, motor function, diet

## Abstract

**Objectives:** Alterations of the gut microbiome in Parkinson's disease (PD) have been repeatedly demonstrated. However, little is known about whether such alterations precede disease onset and how they may be related to risk and prodromal markers of PD. We investigated associations of these features with gut microbiome composition.

**Methods:** Established risk and prodromal markers of PD as well as factors related to diet/lifestyle, bowel function and medication were studied in relation to bacterial  $\alpha$ -/ $\beta$ -diversity, enterotypes, and taxonomic composition in stool samples of 666 elderly TREND study participants.

**Results:** Among risk and prodromal markers, physical inactivity, constipation and age showed associations with  $\alpha$ - and  $\beta$ -diversity, and for both measures subthreshold parkinsonism and physical inactivity showed interaction effects. Moreover, male sex, possible REM-sleep behavior disorder (RBD), smoking as well as body-mass-index, antidiabetic and urate-lowering medication were associated with  $\beta$ -diversity. Physical inactivity and constipation severity were increased in individuals with the *Firmicutes*-enriched enterotype. Subthreshold parkinsonism was least frequently observed in individuals with the *Prevotella*-enriched enterotype. Differentially abundant taxa were linked to constipation, physical inactivity, possible RBD, and subthreshold parkinsonism. Substantia nigra hyperechogenicity, olfactory loss, depression, orthostatic hypotension, urinary/erectile dysfunction, PD family history and the overall prodromal PD probability showed no significant microbiome associations.

**Interpretation:** Several risk and prodromal markers of PD are associated with changes in gut microbiome composition. However, the impact of the gut microbiome on PD risk and potential microbiome-dependent subtypes in the prodrome of PD need further investigation based on prospective clinical and (multi)omics data in incident PD cases.

## INTRODUCTION

The presence of gastrointestinal pathological  $\alpha$ -synuclein deposits and constipation in prodromal and clinically established Parkinson's disease (PD) suggests an integral role of the gut-brain-axis for the early pathogenesis of PD<sup>1-3</sup>. The synucleinopathy is hypothesized to ascend via the vagal nerve from peripheral neurons of the gastrointestinal tract to the brain<sup>4</sup>. Moreover, increased intestinal permeability<sup>5</sup>, elevated stool inflammatory cytokines<sup>6</sup>, and colonic wall inflammation<sup>7</sup> have been shown in PD patients, and may also represent key gastrointestinal processes in prodromal PD. Mice overexpressing  $\alpha$ -synuclein show aggravated motor dysfunction when colonized with intestinal microbiota from PD patients<sup>8</sup>. However, the specific role of gut microbiota for PD and the factors modulating such processes along the microbiota-gut-brain-axis are still largely unknown. The most consistently shown PD-related changes of gut microbial composition include an increase in the relative abundances of *Verrucomicrobiaceae* and *Akkermansia* and a decrease in *Prevotellaceae* and *Prevotella*<sup>9,10</sup>. The latter has also been

associated with progressive PD motor symptoms over two years<sup>11</sup> and with REM-sleep behavior disorder (RBD)<sup>12</sup>, a highly specific prodromal marker of PD. Prospective evidence of several PD risk markers (indicating an increased PD risk) and prodromal markers (indicating an already initiated neurodegenerative process) confirm the concept of a prodromal phase preceding clinically established PD by years or even decades<sup>13</sup>. Based on this evidence, it is possible to define specific predictive values for these markers and to calculate individual prodromal PD probabilities based on age and marker profiles<sup>14,15</sup>. Age<sup>16</sup> and several markers established in the International Parkinson and Movement Disorder Society (MDS) research criteria for prodromal PD<sup>14,15</sup> have also been associated with gut microbial changes; these include risk markers such as male sex<sup>17</sup>, diabetes<sup>18</sup>, non-smoking and non-use of caffeine<sup>16,17</sup> and prodromal markers, such as RBD<sup>12</sup>, constipation<sup>16</sup> and depression<sup>19</sup>. Moreover, physical inactivity, a risk marker for PD<sup>15</sup> and for most major chronic diseases occurring frequently in old age<sup>20</sup>, has been reported to affect microbial β-diversity in elderly males<sup>21</sup> as well as various inflammatory and immune processes<sup>22</sup>. However, these studies focused on microbial associations with single or few PD-related markers, or studied very small samples considering the multitude of other markers and confounding factors that may modulate the gut microbiome. For instance, factors related to diet, bowel function and disease/medication often exert effects on microbial composition<sup>16–18</sup>, and may thus bias findings of marker associations. The present study therefore assessed gut microbial diversity, enterotypes and taxonomic composition and investigated their associations with a comprehensive set of PD risk and prodromal markers, the overall prodromal PD probability and a wide range of other potential confounders in a large sample of elderly individuals.

## METHODS

All subjects were participants of the prospective Tübingen evaluation of Risk factors for Early detection of NeuroDegeneration (TREND) study. The cohort has been enriched regarding an increased PD risk by partly recruiting participants based on the presence of olfactory loss, depression and/or possible RBD. Biennial comprehensive assessments in 1202 individuals have been performed over the last 10 years ([www.trend-studie.de/english](http://www.trend-studie.de/english)). Stool samples were collected at the third follow-up of the study and associated with markers/factors assessed at the corresponding wave of assessments. Stool was sampled using collection tubes containing a DNA stabilizer (PSP Spin Stool DNA Plus Kit, STRATEC Molecular), provided using postal services and frozen and stored at -80 °C immediately upon arrival. Samples were available from 745 participants. After excluding individuals taking antibiotic medication (n = 47), patients with PD (n = 11) or atypical/secondary parkinsonism (n = 2), incident cases of PD (n = 3), and individuals with missing dietary and medication data (n = 16), data from 666 individuals were included in the analyses.

Risk and prodromal markers as defined by the recently updated MDS research criteria for prodromal PD were investigated.<sup>15</sup> We assessed the PD risk markers male sex, substantia nigra hyperechogenicity (transcranial ultrasound), non-smoking, positive PD family history (first degree relatives), physical inactivity (measured as hours of sport/week), type 2 diabetes

(diagnosis), and occupational solvent and pesticide exposure (self-reported). Prodromal markers comprised olfactory loss (Sniffin' Sticks, 16 SS), depression (acute or lifetime diagnosis), constipation (Rome-III criteria, sum score)<sup>23</sup>, possible RBD (RBD screening questionnaire), erectile and urinary dysfunction, symptomatic orthostatic hypotension (the latter three assessed using the UMSARS questionnaire as described previously)<sup>24</sup>. Moreover, motor deficits indicating subthreshold parkinsonism were assessed using the MDS-UPDRS-III<sup>25</sup> (score >6 after excluding action/postural tremor items; scores from 3 to 6 were defined as borderline motor deficit)<sup>14</sup>. Age and comprehensive individual marker profiles were used to calculate prodromal PD probabilities and to diagnose possible (>50% probability) and probable prodromal PD (>80%) according to the MDS research criteria for prodromal PD<sup>14</sup>. Moreover, body-mass-index (BMI), exhaustion from climbing 3 levels of stairs [no, yes, not capable], education level as well as irritable bowel syndrome, bloating, vomiting/diarrhea, gout, 15 diet-related variables (assessing weekly meat, vegetables, dairy, protein, carbohydrate, and alcohol consumption) and 51 variables related to medication were assessed. Overall, 108 variables were investigated for associations with microbial measures. For the full variable list, specific assessment methods, variable definitions and descriptive statistics, see Supporting material. The study was approved by the local ethical committee (Medical Faculty, University of Tübingen; 444/2019BO2). All participants provided written informed consent.

#### *Stool sample DNA extraction, library preparation, and sequencing*

DNA extraction, library preparation and sequencing were performed at the DNA Sequencing and Genomics Laboratory of the Institute of Biotechnology, University of Helsinki. All samples were stored at -80 °C and randomized in a -20 °C room before bulk DNA extraction with the PSP Spin Stool DNA Plus Kit (STRATEC Molecular). One blank was added per extraction batch to assess potential contamination. After extraction, DNA concentrations were measured with NanoDrop ND-1000 (Thermo Scientific). V3-V4 regions of the bacterial 16S ribosomal RNA gene were amplified in Verity 96-well Thermal Cyclers (Thermo Scientific) using a previously described PCR protocol<sup>11</sup>. The amount of template DNA used for the first PCR ranged between 11.25 ng and 1311.26 ng. Each PCR batch included blank samples for assessment of potential contamination. Dual-indexes were used in the second PCR; these had been selected using BARCOSEL<sup>26</sup> to allow pooling and sequencing of all samples in one pool run on three separate runs on a MiSeq (Illumina; v3 600 cycle kit, forward/reverse read length 328/278 bases). Thus, each sample was sequenced three times among all the other samples, reducing possible batch effects.

#### *Sequence analysis*

The raw sequence data contained 48,782,168 sequence reads (availability: ENA accession number PRJEB32920). We combined the sequence reads from the three sequencing runs, and then trimmed primers and low-quality sequences with cutadapt (v.1.8.3)<sup>27</sup> (parameters q=30 and m=160). Merging paired reads, alignment to a reference database (SILVA, v.132), chimera removal, taxonomic classification (reference database: RDP, v.16 PDS), and OTU clustering

(“cluster.split” approach) were run with mothur (v.1.40.0)<sup>28</sup>, following the Standard Operating Procedure for MiSeq data<sup>29</sup>. Parameters differing from the SOP were maxlen=500 and maxhomop=8 for the first “screen.seqs”, start=2 and end=17012 for the second “screen.seqs”, diff=4 for “pre.cluster”, cutoff=70 for “classify.seqs”, and keeping archaeal sequences in “remove.lineage”. Additionally, singleton sequences were removed prior to “classify.seqs” using “split.abund” with cutoff=1. After excluding data from blanks and all OTUs with  $\leq 10$  sequence reads, the final data set consisted of 25,390,744 sequence reads ( $34,082 \pm 4,785$  per sample).

### *Microbial measures*

All microbiota analyses were performed using genus-level data.  $\alpha$ -diversity was estimated with the inverse Simpson index (R package: phyloseq)<sup>30</sup>. This index summarizes richness (number of different taxonomic units) and evenness (abundance distribution of taxonomic units) for each sample. The measure chosen for  $\beta$ -diversity was Bray-Curtis (BC) dissimilarity (R package: vegan). It quantifies the inter-sample compositional dissimilarity based on both presence-absence of taxonomic units and their abundances. Microbial enterotypes<sup>31</sup> were determined using the algorithm provided at <http://enterotypes.org/>. All unclassified taxa (genus level) were excluded from alpha diversity, enterotype and differential abundance analyses. For differential abundance, analyses were also performed for the family-level and OTU-level.

### *Statistical analyses*

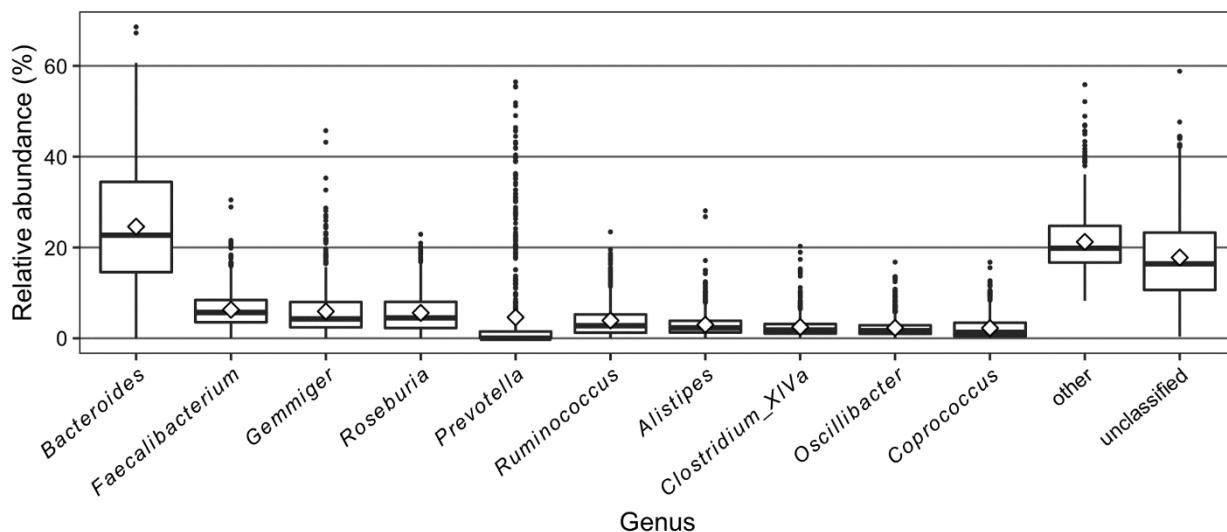
Each of the 108 single markers/factors was investigated for an association with the microbial measures. Variables showing a statistical trend ( $p < .1$ ) were selected for subsequent multifactorial statistical testing. Final multifactorial models comprised all markers/factors showing marginal effects with  $p < .1$ . Linear multiple regressions of  $\alpha$ -diversity were performed. Associations of markers/factors with  $\beta$ -diversity were tested using PERMANOVAs (single variable: *adonis*, multifactorial: *adonis2*, both commands from the R package vegan). Moreover,  $\alpha$ - and  $\beta$ -diversity were tested for interaction effects between physical inactivity and other risk and prodromal markers of PD. Marker/factor differences between enterotypes were tested using Fisher’s exact tests and Kruskal-Wallis tests, and subsequent multifactorial analyses using multinomial logistic regressions with enterotype as dependent variable. For differential abundance analysis we used DESeq2, a method based on generalized linear models with negative binomial distributions (sequence count data). The DESeq2 model included all covariates accounted for in the final PERMANOVA model for  $\beta$ -diversity, and p-values were adjusted for multiple testing using false-discovery-rate (FDR) corrections. Due to the exploratory nature of the study, for other statistical tests effects with an uncorrected  $p < .05$  were considered significant, while effects with  $.05 \leq p < .1$  are reported as non-significant statistical trends. We used R (v3.5.1) for all analyses and figures. TREND study data were collected and managed using REDCap electronic data capture tools hosted at the University of Tübingen.

## RESULTS

### Descriptive statistics

The TREND study sample ( $n = 666$ ) had a mean age ( $\pm$  SD, range) of  $68.4 \pm 6.3$  years (53 - 86). Risk and prodromal marker variables showed the following descriptive statistics: male sex (52.7%), substantia nigra hyperechogenicity (21.6%), positive PD family history (14.7%), physical inactivity (no activity (25.7%); <1 hrs/week (30.9%); 1-2 hrs/week (24.2%); 2-4 hrs/week (8.3%); >4 hrs/week (10.4%)), type 2 diabetes (8.3%), occupational solvent exposure (10.8%), occupational pesticide exposure (1.3%), olfactory loss (19.7%), depression (31.5%), constipation (severity sum score:  $3.1 \pm 3.9$ , range 0 – 25), possible RBD (12.6%), erectile dysfunction (in males: 21.5%), urinary dysfunction (5.1%), symptomatic orthostatic hypotension (4.7%), and regarding motor functions, no motor deficit (90.5%), borderline motor deficit (6.0%) and subthreshold parkinsonism (3.5%). Based on prodromal PD probabilities ( $3.8 \pm 11.4\%$ , 0.0 - 89.1%), 15 individuals (2.3%) were diagnosed with possible prodromal PD, 5 individuals (0.8%) with probable prodromal PD.

The most abundant bacterial genus in the subjects' stool samples was *Bacteroides*, followed by *Faecalibacterium*, *Gemmiger*, *Roseburia*, *Prevotella* and *Ruminococcus* (Fig 1).



**FIGURE 1** Relative abundances of the 10 most common bacterial genera.

### $\alpha$ -diversity

Alpha diversity on the genus level as indicated by the inverse Simpson index was significantly associated with age, two risk markers, two prodromal markers, and four additional factors (Table 1). Overall, these variables explained 8.0% of  $\alpha$ -diversity variance (adjusted  $R^2$ ).

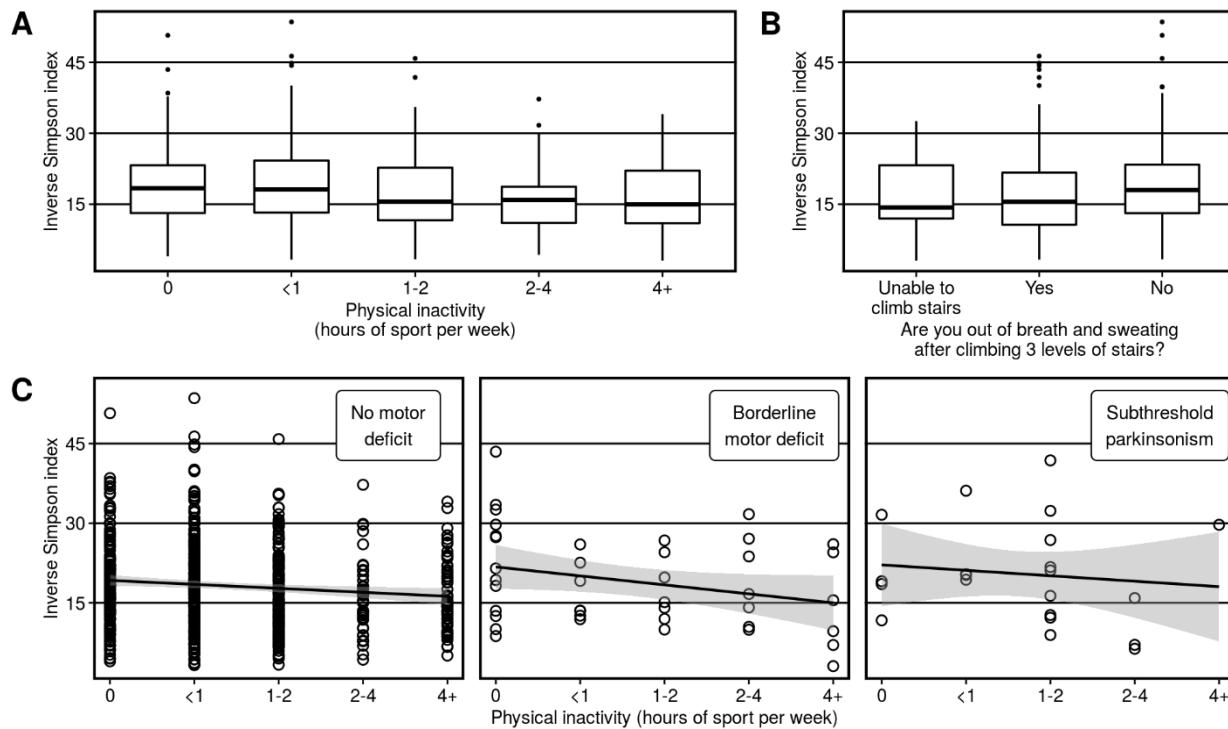
**TABLE 1.** Association of alpha diversity with risk and prodromal markers of PD and additional factors.

	Estimate	S.E.	p-value	
Intercept	4.05	1.24	0.00113	**
Age [years]	0.03	0.02	0.04757	*
Risk markers				
Physical inactivity [h sport/week]	-0.23	0.09	0.00765	**
Occupational solvent exposure [no/yes]	-0.91	0.32	0.00401	**
Prodromal markers				
Constipation [severity sum score]	0.23	0.12	0.04540	*
Subthreshold parkinsonism [no, borderline, yes]	1.06	0.56	0.05930	(*)
Physical inactivity * Subthreshold parkinsonism	-0.33	0.19	0.08769	(*)
Additional factors				
Exhaustion from climbing stairs [no, yes, not capable]	-0.75	0.21	0.00038	***
Urate-lowering medication [no/yes]	-1.25	0.47	0.00812	**
Thyroid medication [no/yes]	-0.61	0.24	0.00951	**
Total dairy consumption [score]	-0.20	0.11	0.07369	(*)

PD = Parkinson's disease; S.E. = Standard error; p<.1(\*); p<.05\*; p<.01\*\*; p<.001\*\*\*

Increasing age, constipation severity and physical inactivity, i.e. fewer hours of weekly sport (Fig 2A), were associated with an increase, and occupational solvent exposure with a decrease in  $\alpha$ -diversity. Moreover, intake of thyroid medication, urate-lowering medication, dairy consumption and exhaustion from walking stairs (Fig 2B) were associated with a decreased  $\alpha$ -diversity. None of the selected variables showed an interaction ( $p>.1$ ) with physical inactivity on alpha diversity. When entering motor deficits into the regression, no association with alpha diversity was observed ( $p>.1$ ); however, the interaction of motor deficits and physical inactivity showed a statistical trend ( $p=.088$ ; Fig 2C).

**FIGURE 2**  $\alpha$ -diversity and physical motor measures. Magnitude of  $\alpha$ -diversity by (A) levels of weekly physical activity, (B) physical exhaustion, and (C) by physical activity levels in different motor deficit groups.



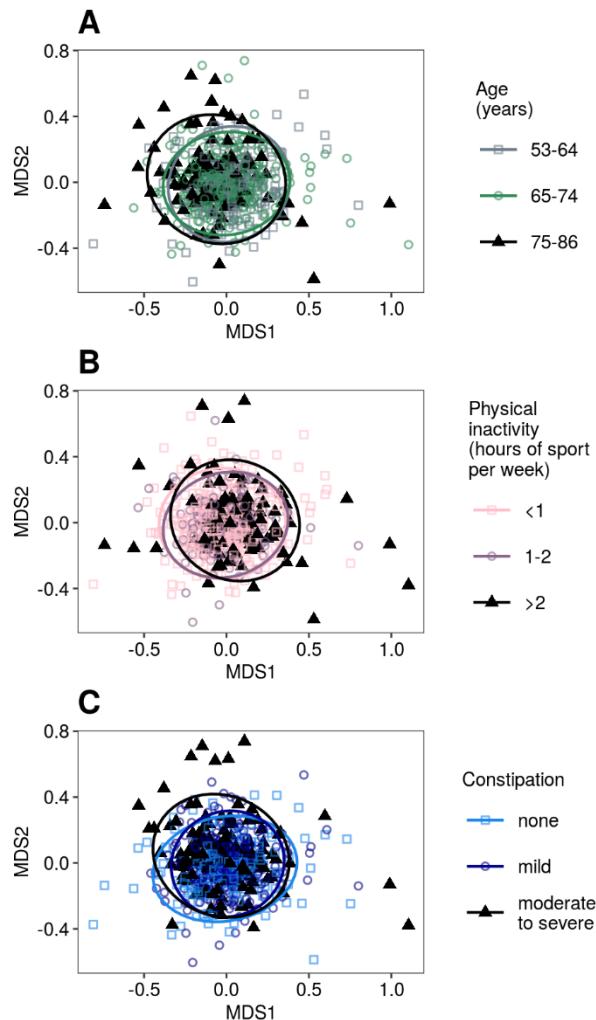
### $\beta$ -diversity

Significant inter-sample differences in microbial composition as indicated by  $\beta$ -diversity were associated with age (Fig 3A) as well as multiple risk and prodromal markers of PD (Table 2). Among these markers, physical inactivity (Fig 3B) and constipation (Fig 3C) explained most of the variance ( $R^2$ ). Moreover, BMI, different medications, and dark bread consumption showed associations with  $\beta$ -diversity. While motor deficits showed no effect, the interaction between physical inactivity and motor deficits significantly explained variance in  $\beta$ -diversity.

**TABLE 2.** Associations of  $\beta$ -diversity with risk and prodromal markers in PD and additional factors.

	R <sup>2</sup>	F-value	p-value	
Age [years]	0.008	4.88	0.001	***
Risk markers				
Physical inactivity [h sport/week]	0.014	9.09	0.001	***
Male sex [no/yes]	0.006	3.59	0.003	**
Smoking [pack-years]	0.004	2.36	0.020	*
Prodromal markers				
Constipation (severity sum score)	0.009	5.48	0.002	**
Possible RBD [no/yes]	0.003	2.14	0.037	*
Subthreshold parkinsonism	0.002	1.45	0.159	
Physical inactivity * Subthreshold parkinsonism	0.006	3.56	0.002	**
Additional factors				
Diabetes medication [no, yes]	0.004	2.66	0.011	*
Urate-lowering medication [no, yes]	0.004	2.43	0.016	*
Beta-blocker medication [no, yes]	0.003	1.91	0.060	(*)
Dark bread [no, yes]	0.003	2.19	0.034	*
BMI [kg/m <sup>2</sup> ]	0.013	8.41	0.001	***

BMI = body mass index; RBD = REM-sleep behavior disorder, PD = Parkinson's disease; p<.1(\*); p<.05\*; p<.01\*\*; p<.001\*\*\*



**FIGURE 3** Non-metric multidimensional scaling (NMDS) plots in 2-dimensional space (MDS1 and 2) of factors showing significant Bray-Curtis distance differences. Differences in microbial composition between (A) age groups, (B) groups with different physical activity levels, and (C) groups stratified by severity of constipation are shown. Circles indicate 95% confidence intervals.

### Enterotypes

*Bacteroides*-enriched (70.9% of samples) microbiomes were more frequent compared to *Prevotella*-enriched (21.5%) and *Firmicutes*-enriched (7.7%) microbiomes. Enterotypes differed significantly in several risk and prodromal markers and additional factors (Table 3). For instance, the lowest physical activity levels as well as most severe constipation were observed for the *Firmicutes*-enriched enterotype. Significantly higher physical activity levels were observed in individuals with the *Bacteroides*-enriched compared to other enterotypes. Subthreshold parkinsonism was least frequently observed in individuals with the *Prevotella*-enriched

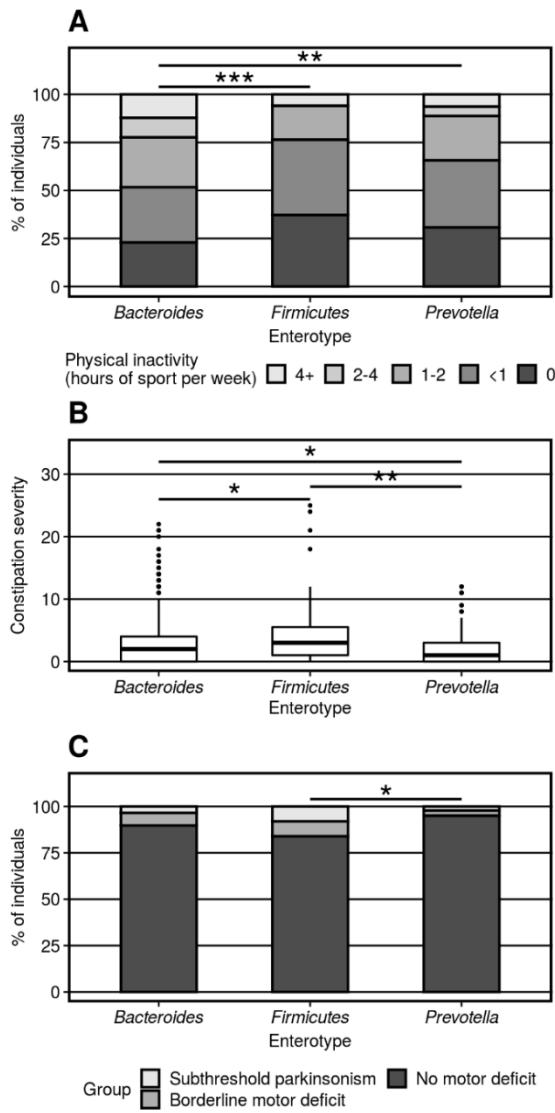
enterotype, with a significant difference in *Prevotella-Firmicutes* enterotype comparisons (Fig 4).

**TABLE 3.** Associations of enterotypes with risk and prodromal markers of PD and additional factors.

	<i>Bacteroides</i> versus <i>Firmicutes</i>			<i>Bacteroides</i> versus <i>Prevotella</i>			<i>Firmicutes</i> versus <i>Prevotella</i>				
	Estimate	S.E.	p-value	Estimate	S.E.	p-value	Estimate	S.E.	p-value		
Intercept	-1.44	0.44	0.00092	***	-0.28	0.26	0.28121	1.17	0.47	0.01299 *	
Risk markers											
Physical inactivity [h sport/week]	-0.58	0.16	0.00031	***	-0.25	0.09	0.00313	**	0.33	0.17	0.05763 (*)
Male sex [no, yes]	-0.60	0.34	0.08162	(*)	0.25	0.20	0.22016	0.85	0.38	0.02355 *	
Prodromal markers											
Constipation [severity sum score]	0.07	0.03	0.02404	*	-0.07	0.03	0.04502	*	-0.14	0.04	0.00133 **
Subthreshold parkinsonism [no, borderline, yes]	0.43	0.30	0.14722		-0.36	0.30	0.23427		-0.79	0.40	0.04628 *
Additional factors											
Legumes [score]	-1.08	0.36	0.00285	**	0.07	0.18	0.67751	1.16	0.39	0.00274 **	
Functional bloating [no, yes]	0.95	0.45	0.03235	*	-0.59	0.46	0.20614	-1.54	0.59	0.00906 **	
Vegetarian [no, yes]	1.05	0.51	0.04091	*	0.38	0.57	0.50650	-0.68	0.65	0.29823	

S.E. = Standard error

**FIGURE 4** Enterotype group differences regarding (A) proportions of levels of physical activity, (B) severity of constipation, and (C) proportions of individuals with motor deficits indicating subthreshold parkinsonism.



### Differential abundance

Several of the variables associated with  $\beta$ -diversity were also linked to the abundances of specific taxa (Supporting materials). The variable with the longest list of differentially abundant taxa was severity of constipation (Table 4). Out of taxa previously associated with PD on the genus level<sup>10</sup>, constipation severity was significantly associated with decreased abundance of *Faecalibacterium* (adjusted  $p = .022$ ), and physical exhaustion with a decrease in *Bifidobacterium* (adjusted  $p = .039$ ). Further, motor deficits indicating subthreshold parkinsonism were associated with a decrease in *Odoribacter* (adjusted  $p = .031$ ). Possible RBD was associated with a decrease in *Faecalicoccus*, *Lactobacillus* and *Victivallis*, and an increase

in the abundance of *Haemophilus*. The taxa *Prevotella* and *Akkermansia* were not statistically significant in any of the differential abundance comparisons.

**TABLE 4.** Number of differentially abundant taxa in association with risk and prodromal markers of PD and additional factors.

	Family-level	Genus-level	OTU-level
Age	0	2	5
Risk markers			
Physical inactivity	4	5	7
Smoking [pack-years]	1	1	12
Prodromal markers			
Constipation severity	10	25	51
Possible RBD	4	4	4
Subthreshold parkinsonism	0	1	4
Additional factors			
Exhaustion from climbing stairs	3	5	9
Urate-lowering medication	3	2	5
Antihypertensive medication	1	1	1
Diabetes medication	2	5	9

OTU = operational taxonomic unit; RBD = REM-sleep behavior disorder

## DISCUSSION

The present study investigated associations between gut microbial composition, risk markers and prodromal markers of PD, subthreshold parkinsonism as well as a wide range of potential confounders in 666 elderly individuals. Among these markers, particularly physical inactivity, constipation, possible RBD, and subthreshold parkinsonism were associated with significant alterations in microbial community composition. Moreover, age, male sex, smoking, and interactions of subthreshold parkinsonism and physical inactivity, showed significant associations with different microbial measures. As expected, effect sizes of individual markers and factors ( $\leq 1\%$  explained variance) and multifactorial models ( $\sim 8\%$ ) were small<sup>16,17</sup>. None of the microbial measures were associated with substantia nigra hyperechogenicity, olfactory loss, depression, orthostatic hypotension, urinary/erectile dysfunction, family history of PD or the overall prodromal PD probability calculated based on marker profiles of individuals.

The relative abundances of bacterial taxa were mostly similar to other samples of healthy elderly individuals<sup>32</sup>. Lower relative abundance of, e.g. *Bacteroides* compared to previous findings, may

be explained by differences in cohort composition, e.g. the TREND cohort enrichment with individuals showing prodromal PD markers.

Age, constipation severity, and physical inactivity were associated with increasing within-sample  $\alpha$ -diversity. Higher  $\alpha$ -diversity in PD patients compared to controls has been reported<sup>33-35</sup>, yet overall evidence is inconsistent<sup>10</sup> and it is unclear whether  $\alpha$ -diversity and PD are linked. Similar to our results, older age in an elderly population has been associated with higher  $\alpha$ -diversity<sup>36</sup>. However, statistical modeling of such effects is complicated by the diseases, medications and lifestyle changes often associated with advanced age. Regarding constipation severity, higher  $\alpha$ -diversity has been linked to harder stool consistency<sup>16</sup> and functional constipation<sup>37</sup>. Autonomous dysfunction related to prodromal PD may contribute to these  $\alpha$ -diversity effects, which however need to be disentangled from other potentially important factors, such as diet, stool water content and transit time, and bacterial growth rates<sup>38</sup>. Physical inactivity increases the risk of many chronic diseases<sup>20</sup> including PD<sup>15</sup>; conversely, being active may lead to lower prevalence of prodromal PD markers such as constipation<sup>39</sup>. The effects of inactivity could be related to the processes observed in PD, i.e. increased colonic inflammation<sup>7</sup>, immune responses<sup>6</sup> and intestinal barrier permeability<sup>5</sup>. In our comparisons, the inactivity-associated increase in  $\alpha$ -diversity was pronounced in individuals with borderline motor deficits and subthreshold parkinsonism. High  $\alpha$ -diversity in elderly, inactive and constipated individuals could therefore be hypothesized as a microbial indicator of impending PD. Exhaustion from climbing stairs, an indicator of low cardiorespiratory fitness, was associated with a reduction in  $\alpha$ -diversity, in line with earlier studies indicating lower diversity in lower fitness<sup>40</sup> and frailty<sup>41</sup>. This apparently contradictory result may be partly explained by differences between physical activity and capability. Moreover, the PD risk marker occupational solvent exposure<sup>15</sup> was associated with a decrease in  $\alpha$ -diversity. Strengths and directions of effects of different risk and prodromal markers on microbial composition may vary based on the underlying pathophysiological pathways, potentially explaining why the overall prodromal PD probability calculated based on comprehensive marker profiles showed no significant association with  $\alpha$ -diversity.

$\beta$ -diversity (between samples) has been consistently shown to differ between PD patients and controls<sup>10</sup>. After initial exploratory screening analyses of 108 variables, age and several PD risk and prodromal markers remained significantly associated with  $\beta$ -diversity even in multifactorial models. These findings suggest that microbial composition might already be altered in the prodromal phase of PD. Although  $\alpha$ - and  $\beta$ -diversity are not directly related, constipation and physical inactivity were again among the variables explaining most of the variance. As observed with  $\alpha$ -diversity, physical inactivity and motor deficits indicating subthreshold parkinsonism showed an interaction effect. It remains speculative to which degree these associations play a specific role in prodromal PD. Possible links between  $\beta$ -diversity and RBD and the PD risk markers male sex and smoking further support the concept of microbial changes preceding PD. While replicating effects of antidiabetic and beta-blocker medication<sup>17</sup>, dark bread was the only dietary factor associated with  $\beta$ -diversity. Dark bread can be seen as an indicator of high fiber consumption, and thus presumably better intestinal barrier integrity and short-chain fatty acid (SCFA) production<sup>42</sup>. Since both have been suggested to be impaired in PD<sup>5,43</sup> and to play a role for several processes along the microbiota-gut-brain axis<sup>42</sup>, dietary factors might thereby also be

important in prodromal PD. After physical inactivity, BMI explained most variance in  $\beta$ -diversity. For some dietary factors linked to BMI, independent effects might have been too small to reach the threshold for statistical significance.

Subthreshold parkinsonism was least frequently present in individuals with the *Prevotella*-enriched compared to the Firmicutes-enriched enterotype. This finding is in line with the reduced abundance of *Prevotella* in PD<sup>10</sup>, in more severely progressing PD<sup>11</sup>, and in RBD<sup>12</sup> compared to controls. Thus, the present study supports the relevance of *Prevotella* in prodromal PD. Microbial changes due to constipation are often argued to confound microbiome analysis of PD patients. The *Prevotella*-enriched enterotype was also the least common in individuals with high constipation severity scores. Subjects with high scores typically had the Firmicutes-enriched enterotype in accordance with previous findings<sup>44</sup>. Viewing constipation in PD as being linked to a disturbed enteric nervous system showing similar cellular changes as affected brain structures may suggest common relevance of *Prevotella* for prodromal dysautonomic and motor deficits. In this light, constipation may not be confounding, but reflecting a common pathogenic process. While constipation may exert effects on microbial composition via several plausible mechanisms, such effects may be different in diseased individuals as compared to healthy subjects, as recently demonstrated in PD for the association between  $\alpha$ -diversity and stool consistency<sup>11</sup>. Physical inactivity was more frequently observed in individuals with the Firmicutes-enriched enterotype, whereas the *Bacteroides*-enriched enterotype was more frequent in active subjects. While consistent with some evidence gained from human<sup>45</sup> and mouse studies<sup>46</sup>, processes underlying the links between physical activity, other prodromal markers in PD<sup>39</sup>, obesity, nutrition and microbial composition are complex and need to be further investigated.

The results of differential abundance analyses were less consistent with previous findings, as most taxa reported for PD<sup>10</sup> and RBD<sup>12</sup> showed no association with risk and prodromal markers in PD. In contrast with enterotype analyses, no significant differential abundance effect was observed for *Prevotella*, which may be partly explained by differences in covariates considered. Among other candidate genera, *Akkermansia* showed no significant association with any of the risk and prodromal markers. However, possible RBD showed several differentially abundant taxa, which have not been previously associated with PD<sup>10</sup> or RBD<sup>12</sup> (*Faecalibacterium*, *Victivallis* and *Haemophilus*). *Lactobacillus* was decreased in individuals with possible RBD, whereas an increase in PD patients has previously been shown repeatedly. It is possible that subtypes of prodromal PD exist with varying involvement of the gut (e.g., RBD representing a subtype with early autonomic denervation)<sup>47</sup>. Also, the microbiome may (differentially) change over time, i.e. from prodromal to clinically established PD. Such complexity may hamper the identification of (prodromal) PD-related microbiome signatures. However, these signatures may allow for early stratification of individuals based on their microbiome (and underlying or inherent pathologies), and early and targeted therapeutic interventions.

Constipation severity was significantly associated with decreased abundance of *Faecalibacterium*, a SCFA-producing taxon that can exert positive effects on the intestinal mucosa<sup>48</sup> and is decreased in PD<sup>10</sup>. The finding of a decrease in *Odoribacter* in individuals with

subthreshold parkinsonism has potential relevance for prodromal PD. *Odoribacter* is a taxon involved in SCFA production and tryptophan metabolism, and might be relevant for gastrointestinal integrity, and serotonergic bowel dysfunction (prolonged transit time) as well as central nervous dysfunction (anxiety) as suggested by findings in an autism mouse model<sup>49</sup>.

The present study has several limitations. (1) Some markers, and dietary and medication data were assessed using self-reports. While those assessments were structured and highly standardized, quantitative measures or medical records, such as weekly physical activity and dietary intakes, might be more accurate. (2) Interaction analyses were limited to those involving physical inactivity given its relevance as PD risk marker<sup>15</sup> and for prodromal PD markers<sup>39</sup>. Further research on marker interactions and clusters is needed to better model the complexity of microbial associations. (3) Stool samples were not frozen immediately after defecation but after postal delivery, and the delay may constitute a technical confounder. However, collection tubes contained a DNA stabilizer and the impact of delayed freezing on microbial composition should therefore be minimal<sup>11,50</sup>.

In conclusion, several risk and prodromal markers in PD, in particular markers related to motor aspects, were associated with altered microbial  $\alpha$ - and  $\beta$ -diversity, enterotypes, and bacterial abundance. Constipation, physical inactivity, possible RBD and subthreshold parkinsonism might be particularly relevant for the prodromal microbiome in PD. Yet, many other markers predictive of PD and overall prodromal PD probability values showed no significant association with any microbial measure. Prodromal microbial changes might only be observable in subgroups with specific marker constellations. The functional roles of these markers and associated microbiota for intestinal permeability, stool immune mediators, colonic inflammation, and the systemic interactions with the host organism need further investigation. The prodromal microbiome(s) in PD, temporal dynamics of microbiota towards PD diagnosis and etiological relevance in prodromal PD should be investigated based on prospective marker profiles and (multi)omics data in incident PD cases.

## Data availability

TREND data are available upon request. Microbiome sequence raw data is available under the ENA accession number PRJEB32920.

## Funding

The TREND study is being conducted at the University Hospital Tuebingen and has been, or is supported by the Hertie Institute for Clinical Brain Research, the German Centre for Neurodegenerative Diseases, the Geriatric Center of Tübingen, the Centre for Integrative Neuroscience, TEVA Pharmaceutical Industries Ltd., Union Chimique Belge (UCB), Janssen Pharmaceuticals, Inc. and the International Parkinson Fonds. This study was funded further by the Academy of Finland (295724, 310835) and the Finnish Medical Foundation. The supporting institutions had no influence on the design, conduct, or analysis of the study.

## Acknowledgements

We would like to thank the participants for their continued participation in the TREND study and providing their stool samples. We acknowledge the work of the numerous (doctoral) students and study nurses, who actively contributed to study organization, and data collection, entry and monitoring. The TREND organization team consists of Corina Maetzler, Susanne Nußbaum, Dr. Anna-Katharina von Thaler, Ulrike Sunkel, Christian Mychajliw and Ramona Täglich. We would also like to thank Albina Gashi, Ella Mustanoja, and Paula Collin for their work on biological sample processing at the DNA Sequencing Lab. Samples were obtained from the Neuro-Biobank of the University of Tuebingen, Germany (<https://www.hih-tuebingen.de/en/about-us/core-facilities/biobank/>). This biobank is supported by the local university, the Hertie Institute for Clinical Brain Research and the DZNE.

## Authors' contributions

SH: Study design, programming, statistical analyses, interpretation of results, drafting of manuscript.

VA: Bioinformatics pipeline, figures, statistical & programming support, interpretation of results, revision of manuscript.

US: TREND study organization and database, revision of manuscript.

A-KvT: TREND study organization and database, revision of manuscript.

LP: Training of technicians, organisation and coordination of DNA-extraction, DNA-libraries and DNA sequencing pipelines, development of DNA barcoding approach, revision of manuscript.

SHa: Conception of dietary factor assessment, revision of manuscript.

KB: TREND study conception and supervision, revision of manuscript.

GWE: TREND study conception and supervision, revision of manuscript.

WM: TREND study conception and supervision, revision of manuscript.

DB: TREND study conception and supervision, revision of manuscript.

PA: Study conception, interpretation of results, revision of manuscript.

FS: Study conception, interpretation of results, revision of manuscript.

## Disclosures

VTEA, LP, PA, and FS have patents FI127671B & US10139408B2 issued and patents US16/186,663 & EP3149205 pending.

US, A-KvT, CD, SHa CS, LP, GWE and PA have nothing to disclose.

SH receives a research grant from Lundbeck.

KB reports grants from MJFF, grants from BMBF, grants from University of Tuebingen, grants from DZNE, from UCB, Zambon, Abbvie.

WM receives or received funding from the European Union, the German Federal Ministry of Education of Research, Michael J. Fox Foundation, Robert Bosch Foundation, Neuroalliance, Lundbeck and Janssen. He received speaker honoraria from Abbvie, Bayer, GlaxoSmithKline, Licher MT, Rölke Pharma and UCB, was invited to Advisory Boards of Abbvie, Biogen, Lundbeck and Market Access & Pricing Strategy GmbH, and is an advisory board member of the Critical Path for Parkinson's Consortium. He serves as the co-chair of the MDS Technology Task Force.

DB received grants from Janssen Pharmaceutica N.V., German Parkinson's Disease Association (dPV), BMWi, BMBF, Parkinson Fonds Deutschland gGmbH, UCB Pharma GmbH, the European Union, Novartis Pharma GmbH, Lundbeck, Damp foundation, has been invited to advisory boards from Biogen, BIAL, Lundbeck, UCB Pharma GmbH; GE, and received speaker honoraria from AbbVie, Biogen, BIAL, Lundbeck, UCB Pharma GmbH Zambon, Desitin.

FS is founder and CEO of NeuroInnovation Oy, has received a grant from Renishaw, consulting fees from Herantis Pharma, UCB, and Orion, lecture fees from Abbvie, UCB, Zambon, Livanova, Axial Biotherapeutics, and Orion, travel support from Abbvie, Herantis Pharma,

Global Kinetics, UCB, NordicInfu Care, Zambon, Livanova, and Medtronic. He is member of the scientific advisory boards of Axial Biotherapeutics and LivaNova.

## References

1. Stokholm, M. G., Danielsen, E. H., Hamilton-Dutoit, S. J. & Borghammer, P. Pathological  $\alpha$ -synuclein in gastrointestinal tissues from prodromal Parkinson disease patients. *Ann. Neurol.* **79**, 940–949 (2016).
2. Schrag, A., Horsfall, L., Walters, K., Noyce, A. & Petersen, I. Prediagnostic presentations of Parkinson's disease in primary care: a case-control study. *Lancet Neurol.* **14**, 57–64 (2015).
3. Abbott, R. D., *et al.* Frequency of bowel movements and the future risk of Parkinson's disease. *Neurology* **57**, 456–462 (2001).
4. 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).
5. 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).
6. Houser, M. C., *et al.* Stool Immune Profiles Evince Gastrointestinal Inflammation in Parkinson's Disease. *Mov. Disord.* **33**, 793–804 (2018).
7. Devos, D., *et al.* Colonic inflammation in Parkinson's disease. *Neurobiol. Dis.* **50**, 42–48 (2013).
8. 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).
9. Schepersjans, F., *et al.* Gut microbiota are related to Parkinson's disease and clinical phenotype. *Mov. Disord.* **30**, 350–358 (2015).
10. Boertien, J. M., Pereira, P. A. B., Aho, V. T. E. & Schepersjans, F. Increasing Comparability and Utility of Gut Microbiome Studies in Parkinson's Disease: A Systematic Review. *Journal of Parkinson's Disease* **9**, S297–S312 (2019).
11. Aho, V. T. E., *et al.* Gut microbiota in Parkinson's disease: Temporal stability and relations to disease progression. *EBioMedicine* **44**, 691–707 (2019).
12. Heintz-Buschart, A., *et al.* The nasal and gut microbiome in Parkinson's disease and idiopathic rapid eye movement sleep behavior disorder. *Mov. Disord.* **33**, 88–98 (2018).
13. Mahlknecht, P., Seppi, K. & Poewe, W. The concept of prodromal Parkinson's disease. *Journal of Parkinson's Disease* **5**, 681–697 (2015).
14. Berg, D., *et al.* MDS research criteria for prodromal Parkinson's disease. *Movement Disorders* **30**, 1600–1611 (2015).

15. Heinzel, S., *et al.* Update of the MDS research criteria for prodromal Parkinson's disease. *Movement Disorders* **34**, 1464–1470 (2019).
16. Falony, G., *et al.* Population-level analysis of gut microbiome variation. *Science* **352**, 560–564 (2016).
17. Zhernakova, A., *et al.* Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352**, 565–569 (2016).
18. Forslund, K., *et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262–266 (2015).
19. Valles-Colomer, M., *et al.* The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat. Microbiol.* **4**, 623–632 (2019).
20. Booth, F. W., Roberts, C. K., Thyfault, J. P., Ruegsegger, G. N. & Toedebusch, R. G. Role of inactivity in chronic diseases: Evolutionary insight and pathophysiological mechanisms. *Physiological Reviews* **97**, 1351–1402 (2017).
21. Langsetmo, L., *et al.* The Association between Objectively Measured Physical Activity and the Gut Microbiome among Older Community Dwelling Men. *J. Nutr. Heal. Aging* **23**, 538–546 (2019).
22. Codella, R., Luzi, L. & Terruzzi, I. Exercise has the guts: How physical activity may positively modulate gut microbiota in chronic and immune-based diseases. *Dig. Liver Dis.* **50**, 331–341 (2018).
23. Longstreth, G. F., *et al.* Functional Bowel Disorders. *Gastroenterology* **130**, 1480–1491 (2006).
24. Pilotto, A., *et al.* Application of the movement disorder society prodromal Parkinson's disease research criteria in 2 independent prospective cohorts. *Mov. Disord.* **32**, 1025–1034 (2017).
25. Goetz, C. G., *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* **23**, 2129–2170 (2008).
26. Somervuo, P., *et al.* BARCOSEL: A tool for selecting an optimal barcode set for high-throughput sequencing. *BMC Bioinformatics* **19**, 257 (2018).
27. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
28. Schloss, P. D., *et al.* Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).
29. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).

30. McMurdie, P. J. & Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**, e61217 (2013).
31. Le Chatelier, E., *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541–546 (2013).
32. Claesson, M. J., *et al.* Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4586–4591 (2011).
33. Barichella, M., *et al.* Unraveling gut microbiota in Parkinson's disease and atypical parkinsonism. *Mov. Disord.* **34**, 396–405 (2019).
34. Qian, Y., *et al.* Alteration of the fecal microbiota in Chinese patients with Parkinson's disease. *Brain. Behav. Immun.* **70**, 194–202 (2018).
35. Keshavarzian, A., *et al.* Colonic bacterial composition in Parkinson's disease. *Mov. Disord.* **30**, 1351–1360 (2015).
36. Odamaki, T., *et al.* Age-related changes in gut microbiota composition from newborn to centenarian: A cross-sectional study. *BMC Microbiol.* **16**, 90 (2016).
37. Mancabelli, L., *et al.* Unveiling the gut microbiota composition and functionality associated with constipation through metagenomic analyses. *Sci. Rep.* **7**, 9879 (2017).
38. Vandepitte, D., *et al.* Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* **65**, 57–62 (2016).
39. Hughes, K. C., *et al.* Physical activity and prodromal features of Parkinson disease. *Neurology* (2019). doi:10.1212/WNL.0000000000008567
40. Estaki, M., *et al.* Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions. *Microbiome* **4**, 42 (2016).
41. Jackson, M., *et al.* Signatures of early frailty in the gut microbiota. *Genome Med.* **8**, 8 (2016).
42. Dalile, B., Van Oudenhove, L., Vervliet, B. & Verbeke, K. The role of short-chain fatty acids in microbiota–gut–brain communication. *Nature Reviews Gastroenterology and Hepatology* **16**, 461–478 (2019).
43. Unger, M. M., *et al.* Short chain fatty acids and gut microbiota differ between patients with Parkinson's disease and age-matched controls. *Park. Relat. Disord.* **32**, 66–72 (2016).
44. Zhu, L., *et al.* Structural changes in the gut microbiome of constipated patients. *Physiol. Genomics* **46**, 679–686 (2014).
45. Morita, E., *et al.* Aerobic exercise training with brisk walking increases intestinal bacteroides in healthy elderly women. *Nutrients* **11**, 868 (2019).
46. Ticinesi, A., *et al.* Exercise and immune system as modulators of intestinal microbiome: Implications for the gut-muscle axis hypothesis. *Exerc. Immunol. Rev.* **25**, 84–95 (2019).
47. Borghammer, P. & Van Den Berge, N. Brain-First versus Gut-First Parkinson's Disease:

A Hypothesis. *Journal of Parkinson's Disease* **9**, S281–S295 (2019).

48. Venegas, D. P., *et al.* Short chain fatty acids (SCFAs)mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. *Frontiers in Immunology* **10**, 277 (2019).
49. Golubeva, A. V, *et al.* Microbiota-related Changes in Bile Acid & Tryptophan Metabolism are Associated with Gastrointestinal Dysfunction in a Mouse Model of Autism. *EBioMedicine* **24**, 166–178 (2017).
50. Ilett, E. E., *et al.* Gut microbiome comparability of fresh-frozen versus stabilized-frozen samples from hospitalized patients using 16S rRNA gene and shotgun metagenomic sequencing. *Sci. Rep.* (2019). doi:10.1038/s41598-019-49956-7