Parkinson's disease-associated alterations of the gut microbiome can invoke disease-relevant metabolic changes

Federico Baldini[1]*, Johannes Hertel[2,3]*, Estelle Sandt[4], Cyrille C. Thinnes[1,2], Lorieza Neuberger-Castillo[4], Lukas Pavelka[1,5], Fay Betsou[4], Rejko Krüger[1,5,6], Ines Thiele[1,2,7 #] on behalf of the NCER-PD Consortium

* These authors contributed equally.

[1] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Campus Belval, Esch-sur-Alzette, Luxembourg.

[2] School of Medicine, National University of Ireland, Galway, Ireland.

[3] Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany.

[4] Integrated BioBank of Luxembourg, 1, Rue Louis Rech, 3555 Dudelange, Luxembourg

[5] Parkinson Research Clinic, Centre Hospitalier de Luxembourg (CHL), Luxembourg

[6] Transversal Translational Medicine, Luxembourg Institute of Health (LIH), Luxembourg

[7] Discipline of Microbiology, School of Natural Sciences, National University of Ireland, Galway, Ireland.

[#] email: ines.thiele@nuigalway.ie

ABSTRACT

Parkinson's disease (PD) is a systemic disease clinically defined by the degeneration of dopaminergic neurons in the brain. While alterations in the gut microbiome composition have been reported in PD, their functional consequences remain unclear. Herein, we first analysed the gut microbiome of patients and healthy controls by 16S rRNA gene sequencing of stool samples from the Luxembourg Parkinson's study (n=147 typical PD cases, n=162 controls). All individuals underwent detailed clinical assessment, including neurological examinations and neuropsychological tests followed by self-reporting questionnaires. Second, we predicted the potential secretion for 129 microbial metabolites through personalised metabolic modelling using the microbiome data and genome-scale metabolic reconstructions of human gut microbes. Our key results include: 1. eight genera and nine species changed significantly in their relative abundances between PD patients and healthy controls. 2. PD-associated microbial patterns statistically depended on sex, age, BMI, and constipation. The relative abundances of *Bilophila* and *Paraprevotella* were significantly associated with the Hoehn and Yahr staging after controlling for the disease duration. In contrast, dopaminergic medication had no detectable effect on the PD microbiome composition. 3. Personalised metabolic modelling of the gut microbiomes revealed PD-associated metabolic patterns in secretion potential of nine microbial metabolites in PD, including increased methionine and cysteinylglycine. The microbial pantothenic acid production potential was linked to the presence of specific non-motor symptoms and attributed to individual bacteria, such as *Akkermansia muciniphila* and *Bilophila wardswarthia*. Our results suggest that PD-associated alterations of gut microbiome could translate into functional differences affecting host metabolism and disease phenotype.

INTRODUCTION

Parkinson's Disease (PD) is a complex multifactorial disease, with both genetic and environmental factors contributing to the evolution and progression of the disease (Kalia et al. 2015). While several studies have elucidated the role of genetic factors in the pathogenesis of the disease (Kitada et al. 1998; Bonifati et al. 2003; Paisan-Ruiz et al. 2004; Di Fonzo et al. 2009), the role and the contribution of various environmental and lifestyle factors are still not completely understood (Gatto et al. 2010). Importantly, about 60% of the PD patients suffer from constipation (Fasano et al. 2015), which can start up to 20 years before the diagnosis and is one of the prodromal syndromes (Savica et al. 2009; Cersosimo et al. 2013).

The human being is considered to be a superorganism recognising a complex interplay between the host and microbes (Sleator 2010). For instance, the human gut microbiome has been shown to complement the host with essential functions (trophic, metabolic, protective) and to influence the host's central nervous system (CNS) via the gut-brain axis through the modulation of neural pathways and GABAergic and serotoninergic signalling systems (Carabotti et al. 2015).

Recent studies have reported an altered gut composition in PD (Hasegawa et al. 2015; Keshavarzian et al. 2015; Scheperjans et al. 2015; Bedarf et al. 2017; Hill-Burns et al. 2017; Hopfner et al. 2017; Petrov et al. 2017; Heintz-Buschart et al. 2018; Barichella et al. 2019). One of these studies has been conducted using samples from recently diagnosed, drug-naive patients (Bedarf et al. 2017). These studies have demonstrated that PD patients have an altered microbiome composition, compared to age-matched controls. However, the functional implications of the altered microbiome remain to be elucidated, e.g., using animal models (Sampson et al. 2016). A complementary approach is computational modelling, or constraint-based reconstruction and analyses (COBRA) (Orth et al. 2010), of microbiome-level metabolism. In this approach, metabolic reconstructions for hundreds of gut microbes

(Magnusdottir et al. 2017) are combined based on microbiome data (Baldini et al. 2018; Heirendt et al. 2019)). Flux balance analysis (FBA) (Orth et al. 2010) is then used to compute, e.g., possible metabolite uptake or secretion flux rates of each microbiome model (microbiome metabolic profile) (Heinken et al. 2019) or to study of microbial metabolic interactions (cross-feedings) (Klitgord and Segre 2010; Heinken and Thiele 2015). This approach has been applied to various microbiome data sets to gain functional insights (Thiele et al. 2018; Heinken et al. 2019; Hertel et al. in revision), including for PD where we propose that microbial sulphur metabolism could contribute to changes in the blood metabolome of PD patients (Hertel et al. in revision).

In the present study, we aim at investigating microbial changes associated with PD while focusing on possible covariates influencing microbial composition and at proposing functional, i.e., metabolic, consequences arising from the microbiome changes. First, we analysed the faecal microbial composition of PD patients and controls from the Luxembourg Parkinson's study (Hipp et al. 2018) (Figure 1). Second, based on the observed significant differences in the composition of microbial communities between PD patients and controls, we created and interrogated personalised computational models representing the metabolism of each individual's microbial community. We demonstrate that the combined microbial composition and functional metabolite analysis provides novel hypotheses on microbial changes associated with PD and disease severity, enabling future mechanism-based experiments.

RESULTS

The Luxembourg Parkinson's Disease study includes patients with typical PD and atypical parkinsonism, as well as matched healthy control subjects from Luxembourg and its neighbouring regions from a broad age-range (Hipp et al. 2018). For the present study, we

focused on typical PD patients and healthy controls over the age of 50 (Table 1, Methods). Stool samples were analysed for 147 PD patients and 162 controls using 16S rRNA gene sequences (Methods: Analysis of the microbial composition with 16S rRNA gene sequencing).

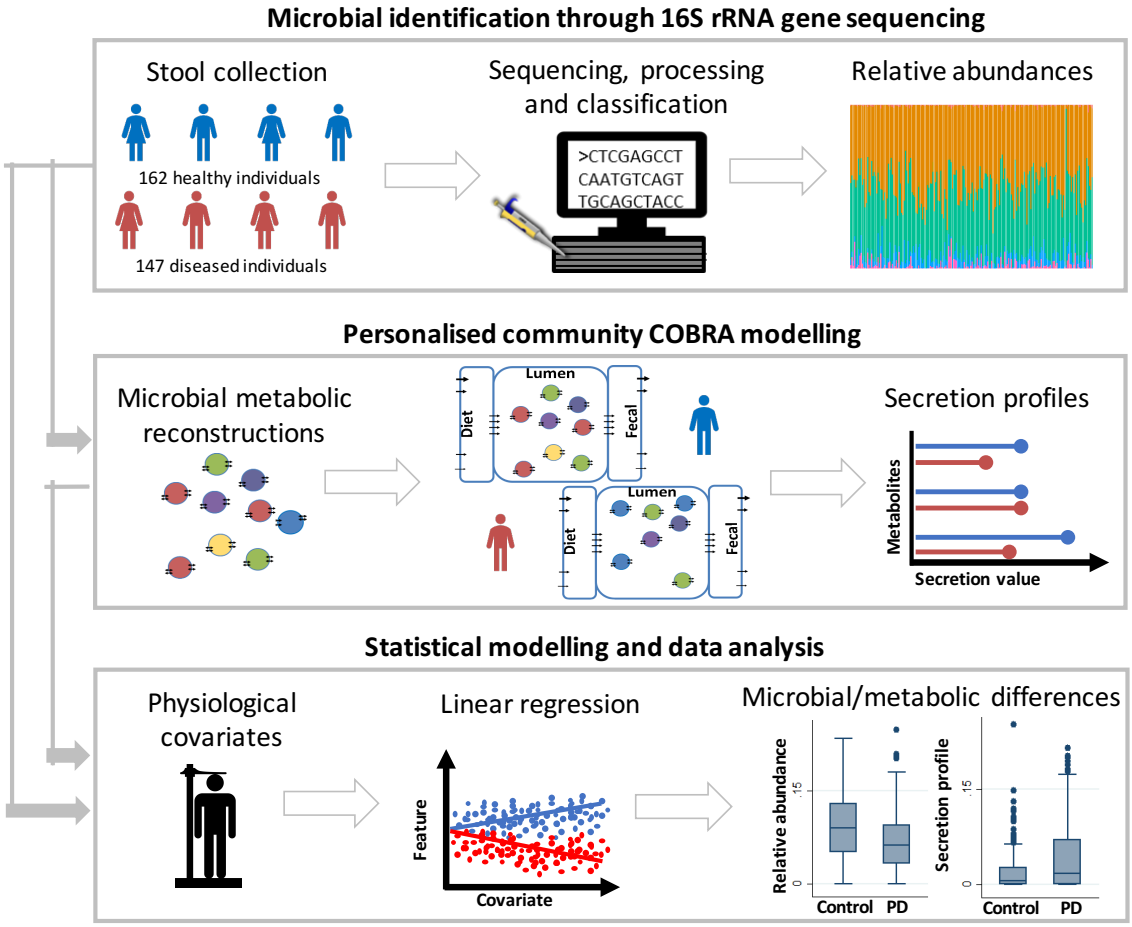


Figure 1: Overview of the study approach and the key methods used. Relative abundances were derived from 16S rRNA gene sequences (Methods: Analysis of the microbial composition with 16S rRNA gene sequencing) and used as input for the personalised community modelling to simulate metabolites secretion profiles. Relative abundances and secretion profiles were statistically analysed to identify microbial or metabolic differences between PD patients and controls.

Variable	PD	Control	Missing values, in %		Genera influenced by PD-covariate interaction effects	Genera associated with trait	NMPCs associated with trait
			PD	Control	(FDR<0.05)	(up/down, FDR<0.05)	(up/down, FDR<0.05)
Cases vs. Controls	147	162	0%	0%	--	*Anaerotruncus*, *Christensenella*, *Lactobacillus*, *Streptococcus*, *Akkermansia*, *Bilophila*, *Turicibacter*	D-alanine, Oxalate, D-Mannitol, Cysteinylglycine, L-Methionine, L-alanine, D-Ribose, 4Hydroxybenzoic acid, Uracil,
Sex (female subjects)	31.5 %	35.8%	0%	0%	*Paraprevotella*	--	--
Age at basic assessment[a]	69.3 ± 8.6	63.3 ± 8.3	0%	0%	*Anaerotruncus*, *Roseburia*	--	Phosphate, Glycine
Body mass index[a]	27.3 ± 4.5	27.9 ± 4.8	0.7%	0%	*Paraprevotella*	*Victivallis*	--
Sniff score[a]	7.1 ± 3.4	12.7 ± 2.1	0%	0%		--	
Metabolic diabetes	4.1 %	3.1%	0%	0%	--	--	--
Non-motor symptoms questionnaire score[a]	9.3 ± 5.1	3.9 ± 3.9	9.5%	3.7%	--	---	Pantothenate
Constipated	36.7%	6.2 %	0%	0%	*Bifidobacterium*	*Bifidobacterium*	Xanthine, D-Alanine, Pantothenate, L-Lactate, D-Ribose
PD disease duration	5.9 ± 5.7	--	6.1%	--	--	*Lactobacillus*	--
UPDRS-part I	10.0 ± 5.9	4.5 ± 4.4	3.4%	3.1%	--	--	--
UPDRS-part II	11.8 ± 8.1	1.3 ± 2.8	1.4%	2.4%	--	--	--
UPDRS-part III	34.6± 16.1	2.3± 2.9	1.4%	0%	--	*Peptococcus*, *Flavonifractor*, *Paraprevotella*	--
UPDRS-part IV	1.7 ± 3.2	--	1.4%	--	--	--	--
Hoehn and Yahr	2.2 ± 0.6	--	0%	--	--	*Bilophila*, *Paraprevotella*	--
L-DOPA intake	66.7%	0%	0%	0%	--	--	--
Dopamine agonist intake	56.5%	0%	0%	0%	--	--	--
MAO-B COMT inhibitors intake	41.5%	0%	0%	0%	--	--	--

Table 1: Descriptive statistics of the analyses sample from the Luxembourg Parkinson's Disease study and overview over associations. A red label means increased in PD, blue decreased in PD, while -- "nothing to report". PD disease duration refers to time since diagnosis at the date of stool sampling. UPDRS=Unified Parkinson Rating Scale, L-DOPA=levodopa, MAO-B=monoaminooxidase B, COMT=Catecholamine-Methyl-Transferase, NMPC=Net maximal production capability.

Species and genus level changes in PD microbiomes

We investigated disease-associated microbial changes at the species level. We found that the mean species diversity (i.e., the alpha-diversity) did not significantly differ between PD cases and controls (b=-0.04351, 95%-CI:(-.107;0.177), p=0.177), in agreement with earlier studies (Scheperjans et al. 2015; Bedarf et al. 2017) (Hopfner et al. 2017), but in disagreement

with two other studies (Keshavarzian et al. 2015; Heintz-Buschart et al. 2018). However, seven species were significantly altered in PD (FDR<0.05, Figure 2). Note that when comparing results between different taxonomic levels, changes observed for *Ruminococcus* and *Roseburia* species were not significant on the genus level but only on the species level, highlighting the importance of species-level resolution. The highest effect size was associated with *Akkermansia muciniphila* (Odds ratio (OR)=1.80, 95%-CI=(1.29, 2.51), p=6.02e-04, FDR<0.05; Supplementary Table 1) in agreement with the previously reported higher abundance of *A. muciniphila* in PD patients (Bedarf et al. 2017; Heintz-Buschart et al. 2018)). Subsequently, we examined possible differences at the genus level by performing semiparametric fractional regressions while adjusting for age, sex, the body mass index (BMI), batch, and total read counts. We identified eight genera to be significantly increased in PD (FDR<0.05; Figure 3A, Table 1), with *Lactobacillus* showing the highest effect size (Odds ratio (OR)=5.75, 95%-CI=(2.29, 14.45), p=1.96e-04, FDR<0.05; Supplementary Table 2). In contrast, the genera *Turicibacter* decreased significantly in PD cases (FDR<0.05). To summarise, significant changes could be observed on the species and genus level.

PD modifies the effects of basic covariates on the microbiome

Furthermore, we investigated whether the genus level alterations in PD were affected by basic confounding factors. This interaction analyses uncovered rich effect modifications, revealing that microbiome changes in PD have to be considered in the context of age, BMI, and gender. Our analyses demonstrate that the effects of PD are not homogeneous among important sub-groups of patients. For example, *Paraprevotella* was exclusively reduced in female patients but not in female controls (Figure 3B), highlighting gender-dependent alterations of microbial communities in PD. In addition, the effects of BMI and age were modified in PD cases. The PD cases had increased *Anaerotruncus* abundance with age, while

non-linear, overall decreasing abundances of *Roseburia* and *Paraprevotella* were observed with age and BMI, respectively (Figure 3C). Taken together, these analyses suggest that microbial abundances are shifted in PD cases and that also the effects of important covariates were altered in PD, reflecting the systemic and complex nature of PD.

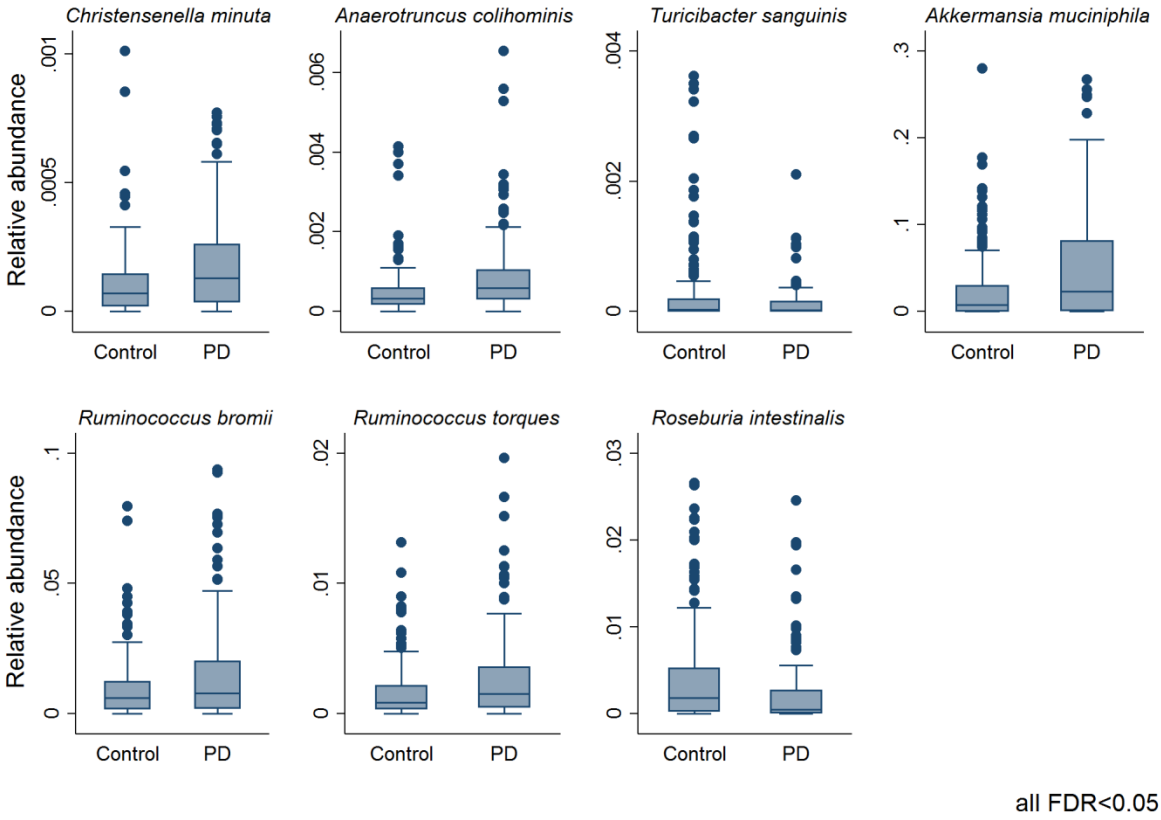


Figure 2: Boxplots of seven significantly changed species in PD versus controls (FDR<0.05). Significance levels were determined using multivariable semi-parametrical fractional regressions with the group variable (PD vs. control) as predictor of interest, including age, gender, BMI, and technical variables (i.e., total read-counts and batch effect) as covariates. FDR=false discovery rate.

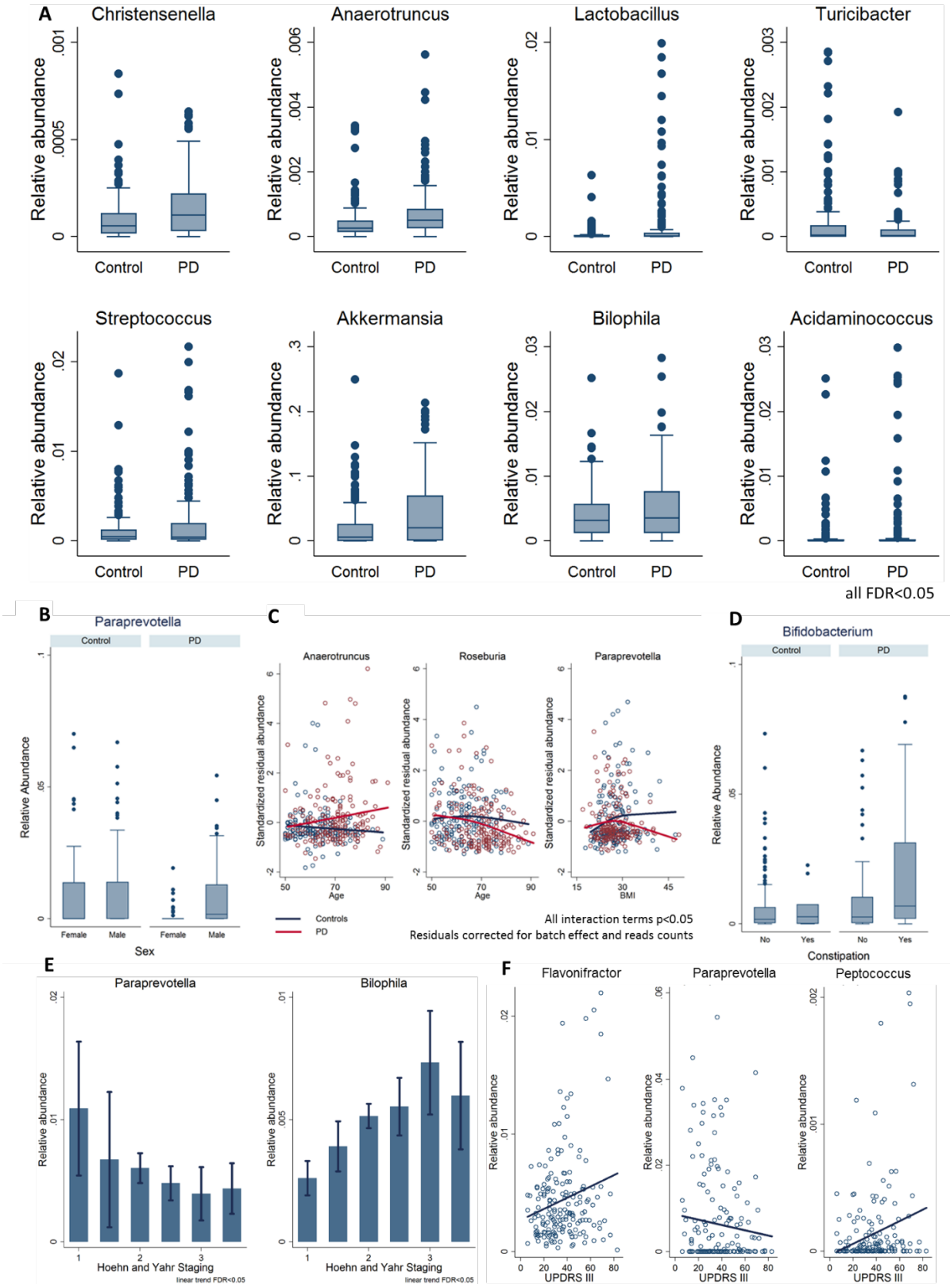


Figure 3: Genus alterations in PDs due to interactions with basic covariates. Relative abundance is given on a logarithmic scale. **A.** Boxplots of the seven significant species (FDR<0.05). **B.** Female PD patients have a reduced abundance of *Paraprevotella* (FDR<0.05).

C. Genus abundance age and BMI dependencies of *Anaerotruncus*, *Roseburia*, and *Paraprevotella* (global test on all interaction terms, FDR<0.05). For graphical assessment of the interaction terms the z-transformed residual abundances are displayed after correction for technical covariates (batch and read counts). **D.** The genus relative abundance of Bifidobacterium was increased in patients reporting to be constipated (FDR<0.05). **E.** Genus association with disease staging showed a decrease of relative abundance of *Paraprevotella* and an increase of *Bilophila* genus over increasing Hoehn and Yahr scale values (FDR<0.05). **F.** An increased score in motor symptoms (UPDRS III) was associated with an increased trend in abundances of *Flavonifractor* and *Peptococcus* and a decreased trend in *Paraprevotella* abundance (FDR<0.05). UPDRS=Unified Parkinson Rating Scale, BMI=body mass index, FDR=false discovery rate.

Microbial abundances, medication intake, and constipation in PD

The Luxembourg Parkinson's study enrols patients of all stages of PD. Therefore, the patients have considerable inter-individual variance in PD-related features, such as constipation and intake of medication (Table 1). We analysed whether these features had an impact the microbiome composition in PD. In our data, we could not find any evidence for an effect of the three medication types on the microbiome, i.e., levodopa, COMT inhibitors, or MAO-B inhibitors, when correcting for multiple testing (Supplementary Table 2). In contrast, constipation, a prevalent non-motor symptom in PD patients (Lesser 2002), was associated with an increased abundance of *Bifidobacterium*, with a clear effect in constipated PD cases (Figure 3D). However, since there were only ten constipated controls (Table 1), these results must be confirmed in larger cohorts.

Genus association with the disease severity

We next investigated whether the stage of the disease, i.e., defined by Hoehn and Yahr staging, NMS, and UPDRS (Unified Parkinson Rating Scale) scores, and its subscales, was associated with altered genus abundance. For the Hoehn and Yahr staging, *Paraprevotella* showed a negative association and *Bilophila* showed a positive association, both of which were significant after multiple testing (Figure 3E). For the UPDRS III subscale score (i.e., motor symptoms, Table 1), three genera, being *Peptococcus*, *Flavonifractor,* and *Paraprevotella*, survived correction for multiple testing (Figure 3F). In contrast, the other UPDRS subscales and the NMS were not significantly associated with microbial changes, after correction for multiple testing. Note that these analyses were performed while adjusting for disease duration. When analysing the association pattern of disease duration, we found *Lactobacillus* positively correlated with the disease duration (FDR<0.05, Supplementary Figure S1). In conclusion, our data suggest that the microbial composition may be utilised as a correlate of disease severity.

Metabolic modelling reveals distinct metabolic secretion capabilities of PD microbiomes

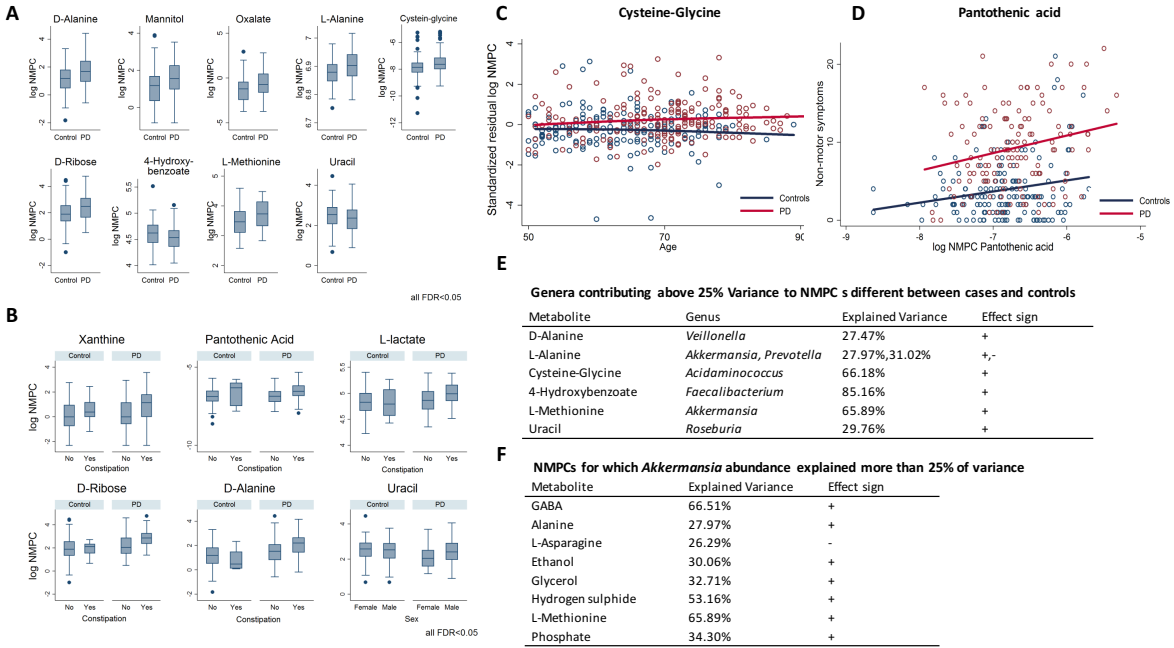
To obtain insight into the possible functional consequence of observed microbiome changes in PD, we used metabolic modelling (cf. Methods). Briefly, we mapped each of the 309 microbiome samples on the generic microbial community model consisting of 819 gut microbial reconstructions (Magnusdottir et al. 2017; Heinken et al. 2019) (cf. Supplementary Material) to derived personalised microbiome models (Baldini et al. 2018). We then computed a net maximal production capability (NMPC) for 129 different metabolites that could be secreted by each microbial community model (cf. Methods), providing thereby a characterisation of the differential microbial metabolic capabilities in PDs and controls. The secretion of nine metabolites had differential NMPCs in PD (Figure 4A, all FDR<0.05) as determined by multivariable regressions adjusting for age, sex, BMI, and technical covariates.

Moreover, although less dominant in comparison to the abundance data, PD-covariate interactions were also prevalent, with the uracil secretion potential showing a sex-specific effect and cysteine-glycine showing a BMI-dependent PD-effect (Figure 4B, 4C). In subsequent analyses, we tested for associations of the NMPCs with constipation, medication, disease duration, Hoehn-Yahr staging, NMS, and UPDRS III scores, complementing thereby the analyses on the abundance level. Notably, we found xanthine, D-alanine, L-lactic acid, D-ribose, and pantothenic acid positively associated with constipation (Figure 4B), while no NMPC was associated with medication or with disease duration. However, the pantothenic acid secretion potential was positively associated with higher NMS scores, interestingly both in PD and in controls (Figure 4D), while no NMPC survived correction for multiple testing regarding associations with the UPDRS III score and Hoehn-Yahr staging. To conclude, these results suggest that the altered microbial composition in PD could result in broad changes in metabolic capabilities, which manifested themselves additionally in non-motor symptoms and constipation.

PD specific secretion profiles were altered due to changed community structure and species abundances

Next, we analysed which microbes contributed to the differential secretion profiles by correlating the NMPCs to the abundance data (Figure 4E/F, Supplementary Table 3). Six metabolite NMPCs had strong contribution or where even dominated by single genera (Figure 4D), while for the other four NMPCs no single dominant genus could be identified. We then computed the contribution value of each genus to the production of each secreted metabolite (NMPC). From the aforementioned genera, which were associated on genus or species level with PD, only *Akkermansia*, *Acidaminococcus*, and *Roseburia* had substantial metabolic contributions (over 25%). *Acidaminococcus* was responsible for 64% of the variance in

240 cysteine-glycine production and *Roseburia* for 30% of the variance in uracil production
241 potential. *Akkermansia* impacted the secretion profiles the most and substantially contributed
242 to the metabolism of nine metabolites (Figure 4F), including the neurotransmitter gamma-
243 aminobutyric acid (GABA) and two sulphur species, being hydrogen sulphide and methionine.
244 GABA was also significantly altered between PD and controls on a nominal level missing FDR
245 corrected significance narrowly (b=0.18, 95%-CI:(0.06;0.30), p=0.003, FDR=0.0501). These
246 analyses demonstrate the added value of metabolic modelling to investigate altered metabolic
247 functions from the whole microbial composition.
248



Metabolite	Genus	Explained Variance	Effect sign
D-Alanine	*Veillonella*	27.47%	+
L-Alanine	*Akkermansia, Prevotella*	27.97%,31.02%	+,-
Cysteine-Glycine	*Acidaminococcus*	66.18%	+
4-Hydroxybenzoate	*Faecalibacterium*	85.16%	+
L-Methionine	*Akkermansia*	65.89%	+
Uracil	*Roseburia*	29.76%	+

Metabolite	Explained Variance	Effect sign
GABA	66.51%	+
Alanine	27.97%	+
L-Asparagine	26.29%	-
Ethanol	30.06%	+
Glycerol	32.71%	+
Hydrogen sulphide	53.16%	+
L-Methionine	65.89%	+
Phosphate	34.30%	+

249

250 **Figure 4: Result of analysing secretion profiles of microbial communities. A.** Box plots for
251 NMPCs differential between cases and controls with FDR<0.05. **B.** NMPCs with sex-specific
252 PD signature or constipation effects (all FDR<0.05). **C.** Differential age trajectory between
253 cases and controls for cysteine-glycine (p<0.05). **D.** Association of pantothenic acid with non-
254 motor symptoms. **E.** Genera contributing more than 25% to NMPCs different between cases
255 and controls. **F.** *Akkermansia* contribution to community production of 12 metabolites
256 expressed as a percentage of total production for each compound. Metabolites highlighted in

257 red were significantly increased in PD (FDR<0.05). NMPC=net maximal production capacity,
258 GABA=gamma-aminobutyrate, H2S=hydrogen sulphide, FDR=false discovery rate. Effect
259 sign "–": negative correlation. Effect sign "+": positive correlation.
260
261 **DISCUSSION**
262 In this study, we aimed to elucidate compositional and functional changes in the faecal
263 microbiome of PD patients. Therefore, we analysed 16S rRNA data from a cohort of typical
264 PD patients (n=147) and controls (n=162), and performed personalised microbial
265 computational modelling. We identified i) eight genera and nine species that changed
266 significantly in their relative abundances between PD patients and healthy controls. ii) PD-
267 associated microbial patterns that were dependent on sex, age, BMI, constipation, and iii) in
268 PD patients altered secretion potentials, particularly in sulphur metabolism, using metabolic
269 modelling of microbial communities. Overall, our work demonstrated compositional and
270 functional differences in the gut microbial communities of Parkinson's disease patients
271 providing novel experimentally testable hypothesis related to PD pathogenesis.
272 The microbial compositional analyses of our cohort identified significantly different
273 microbial abundance distributions between PD patients and healthy controls (Table 1). Up to
274 date, 13 studies have described altered colonic microbial compositions associated with PD and
275 an overall picture starts to arise (Figure 5). For instance, the microbial the families of
276 Verrucomicrobiaceae and Lactobacillaceae have been consistently found to have an increased
277 abundance in PD (Figure 5). In accordance, our study also reports increased abundance in PD
278 of Akkermansia, Christensenella, and Lactobacillus. Similarly, Bifidobacteria has also been
279 repeatedly associated with PD (Figure 5) but in our study, we could show that the
280 Bifidobacteria association dependent on constipation (Figure 3) highlighting the need for
281 incorporating disease-specific phenotypes as covariates into the statistical design.

		Bedarf, 2017	Barichella, 2019	Hasegawa, 2015	Heintz-Buschart, 2018	Hill-Burns, 2017	Hopfner, 2017	Keshavarzian, 2015	Li, 2017	Lin, 2018	Petrov, 2017	Qian, 2018	Scheperjans, 2015	Unger, 2016	Current study
Study size	PD cases	31*	193	52	76	197	29	38	14	75	89	45	72	34	147
	Controls	28	191	36	78	130	29	34	24	45	66	45	72	34	162
Method	Shotgun	x													
	16S rRNA (region)		V3-V4		V4	x	V1-V2	V4	V3-V4;V4-V5	x	V3-V4	V3-V4	V1-V3		V3-V4
	Quantitative PCR			x										x	
Family	Bacteroidaceae			b											
	Bifidobacteriaceae										c				
	Christensenellaceae										d				d
	Clostridiaceae			e							f	g			h
	Comamonadaceae											i			
	Coriobacteriaceae														
	Enterobacteriaceae														
	Enterococcaceae														
	Erysipelotrichaceae											j			
	Lachnospiraceae														
	Lactobacillaceae			k							k				k
	Pasteurellaceae														
	Prevotellaceae														
	Ruminococcaceae										l				m
	Sphingomonadaceae											n			
	Streptococcaceae														o
	Verrucomicrobiaceae	p			p										p

Figure 5: Reported microbial changes at the family level associated with PD in different studies. Only those bacterial families are shown, for which significant associations with species or genera have been reported in at least two studies comparing stool samples from patients and controls. Red - increased in PD, Blue - decreased in PD. a: *Actinomycetales*, b: *Bacteroides fragilis*, c: *Bifidobacterium*, d: *Christensenella*, e: *Clostridium coccoides/ leptum*, f: *Faecalibacterium*, *Dorea*, g: *Clostridium* IV/XVIII, *Butyricicoccus*, *Anaerotruncus*, h: *Anaerotruncus*, i: *Aquabacterium*, j: *Holdemania*, k: *Lactobacillus*, l: *Oscillospira*, m: *Ruminococcus romii, Ruminococcus torques,* n: *Sphingomonas*, o: *Streptococcus*, p: *Akkermansia*. * Drug-naive, de novo PD patients only. Based on (Barichella et al. 2019).

At the same time, inconsistencies between the studies remain and they may be due to differences in study design, inclusion criteria, faecal sampling, RNA extraction protocols, and metagenomic and statistical methods. For instance, we used a relatively large, PD cohort while

Bedarf and colleagues (Bedarf et al. 2017) studied a small cohort of drug-naïve, male PD patients and male controls (Figure 5). Three studies included individuals of Chinese descent (Li et al. 2017; Lin et al. 2018; Qian et al. 2018) while the other studies focused on Caucasian individuals. It has been shown that microbial composition is associated with ethnic background, geography, and dietary habits (Turnbaugh et al. 2008; De Filippo et al. 2010; Yatsunenko et al. 2012), which may explain some of the discrepancies. The differences between the studies hence highlight the importance of performing meta-analysis to identify global microbial signatures, as it has been done for, e.g., colorectal cancer (Wirbel et al. 2019). Such meta-analysis may also permit to investigate subgroups of PD, as the number of cases and controls would be substantially increased and thus provide higher statistical power. For instance, we observed various effect modulators that were not reported before in humans (Table 1), such as *Paraprevotella* abundance reduction being specific to women. This result is apparently in contradiction with findings from Bedarf and colleagues (Bedarf et al. 2017) who reported decreased levels of Prevotellaceae in a cohort of only male PD patients. However, once again, differences might be explained by different inclusion criteria, methodologies, and related possible sex-specific effects. Interestingly, a recent study reported a higher abundance of *Paraprevotella* in male mice compared to female mice (Huang et al. 2018). Despite the lack of extensive studies on gender-specific differences in microbiome composition, we suggest that machine learning procedures on microbiome data should be performed in a sex-stratified manner. Larger cohorts, e.g., through meta-analysis of published cohorts would allow the identification of generalizable microbial differences in PD patients and also, specific microbial changes associated with certain traits and physiological characteristics, as suggested by our data.

We could not detect an effect of the dopaminergic, PD specific medication on the microbiome composition, after correction for multiple testing. Also the fact that key findings

from the study of Bedarf and colleagues were reproduced in other cohorts of PD patients under medication, including ours, support that notion. Nonetheless, in previous studies, Dorea and Phascolarctobacterium genera have been negatively associated with levodopa equivalent doses (Qian et al. 2018) and members of the family of *Bacillaceae* have been correlated with levodopa treatment (Heintz-Buschart et al. 2018). Consequently, it cannot be excluded that medication is associated with microbial changes, albeit the association may be weaker than the effects of other covariates. As PD drugs are often taken in combinations, it would require a larger sample size than used in our study to permit the investigation of all possible drug combinations. The lack of clear association is somewhat expected as levodopa is absorbed in the upper part of the small intestine (Streubel et al. 2006) and thus small intestinal rather than large intestinal microbes may play a more prominent role in levodopa bioavailability. Consistently, a recent study showed that bacterial tyrosine decarboxylases restrict the bioavailability of levodopa (van Kessel et al. 2019). Interestingly, 193/818 reconstructed microbes (Magnusdottir et al. 2017), commonly found in the human gut, carry genes encoding for proteins that convert levodopa into dopamine (Noronha et al. 2019). Levodopa is always given with decarboxylase inhibitors, such as carbidopa or benserazide, targeting the human decarboxylases, but it cannot be excluded that they also act on the microbial counterpart. However, Van Kessel et al. have shown that carbidopa as well as benserazide is only a weak inhibitor of the microbial tyrosine decarboxylase (van Kessel et al. 2019).

We identified a positive association of *Bilophila* abundance with the Hoehn and Yahr staging, which captures motor impairment and disability independent of disease duration. Indeed, the abundance of *Bilophila* was not associate with disease duration, indicating mainly dependency on the progression of symptoms. This finding is consistent with experimental mice studies demonstrating the pro-inflammatory effect of *Bilophila* overgrowth (Devkota et al. 2012; Natividad et al. 2018). Notably, the Hoehn and Yahr staging was also positively

346 associated on a nominal level with the predicted pyruvate secretion profile (Supplementary File
347 4), which was accordingly significantly increased in PD patients on a nominal level alongside
348 with L- and D-alanine. *Bilophila* has the rare capability to use taurine, an inhibitory
349 neurotransmitter with neuroprotective effects (Saransaari and Oja 2007; Wu et al. 2009), as an
350 energy source (Laue and Cook 2000). This pathway is initiated by the taurine: pyruvate
351 aminotransferase (Laue and Cook 2000), converting pyruvate and taurine into L-alanine and
352 sulfoacetaldehyde. The only microbe of the 818 species in our AGORA collection encoding
353 the corresponding gene was *Bilophila*, which was significantly increased (FDR<0.05) and
354 hence, the corresponding reaction (VMH ID: TAURPYRAT) was increased in abundance in
355 PD microbiomes as well. In a previous study (Hertel et al. in revision), we have shown that
356 blood taurine conjugated bile acids were positively associated with motor symptoms. *Bilophila*
357 may be a marker of disease progression in PD, and it could modulate human sulphur
358 metabolism through its taurine degradation capabilities. Alterations in sulphur metabolism
359 have been already described when using computational modelling of microbiomes from a
360 cohort of early diagnosed and levodopa naive PD patients (Bedarf et al. 2017; Hertel et al. in
361 revision) as well as an increased concentration of methionine and derived metabolites in blood
362 samples (Hertel et al. in revision). Furthermore, we and others have reported alterations in bile
363 acids and taurine-conjugated bile acids in PD patients (Graham et al. 2018; Hertel et al. in
364 revision). Our present study suggests again a key role of *Bilophila* in host-microbiome sulphur
365 co-metabolism, which may link with bile acid metabolism.

366 Interestingly, an increased abundance of *B. wadsworthia* has been linked to constipation
367 (Vandeputte et al. 2017). *B. wadsworthia* is the only microbe in the AGORA collection capable
368 of the metabolic reaction converting pyruvate and taurine to L-alanine and sulphoacetaldehyde
369 (VMH ID: TAURPYRAT). Therefore, an increased production of L-alanine might be due to
370 the increased *B. wadsworthia* abundance. This resulting higher production rate of L-alanine

371 could then lead to an increased conversion into D-alanine via the alanine racemase (VMH ID:
372 ALAR), which was present in 808/818 gut microbes in the AGORA collection. Accordingly,
373 D-alanine was one of the three metabolite secretion profiles increased in constipated PD
374 patients (Figure 4E). This hypothesis of *B. wadsworthia* playing a role in constipation of PD
375 patients would need to be experimentally validated, especially since we could not find
376 statistically significant changes in the association between the abundance of *B. wadsworthia*
377 and constipated individuals. In contrast, we found an increase in *Bifidobacteria* abundance in
378 constipated individuals and particularly in constipated PD patients. This result disagreed with
379 an earlier study on individuals with chronic constipation, which reported a decrease in
380 *Bifidobacteria* abundance (Khalif et al. 2005). Overall, the available data suggest that complex
381 alterations in microbial composition are associated with constipation but may differ between
382 diseases.

383 The mucin degrading microbe, *A. muciniphila*, represents about 1-4% of the faecal
384 microbiome in humans (Naito et al. 2018). Numerous diseases have been associated with a
385 decrease in *A. muciniphila* abundance (Schneeberger et al. 2015; Grander et al. 2018), while
386 an increase has been consistently reported in PD patients (Figure 5). The *A. muciniphila*
387 abundance had the largest contribution to the significantly altered metabolite secretion profiles
388 (Figure 4E), including the neurotransmitter gamma-aminobutyric acid (GABA). While its
389 predicted secretion potential was only nominally increased in PD patients the present study,
390 higher GABA secretions rates have also been predicted based on microbiome data from early
391 stage levodopa naive PD patients (Hertel et al. in revision). Importantly, GABA receptors have
392 been found in the enteric nervous system, gut muscle, gut epithelial layers, and endocrine-like
393 cells (Hyland and Cryan 2010) and its gut receptors are thought to be related to gastric motility
394 (peristalsis), gastric emptying, and acid secretion (Hyland and Cryan 2010). Experiments with
395 the GABAb agonist baclofen have shown that GABAb receptors can reduce gastric mobility

396 in the colon of rabbits (via cholinergic modulation) (Tonini et al. 1989). Interestingly, *A.*
397 *muciniphila* has been shown to be positively associated with gastrointestinal transit time
398 (Gobert et al. 2016; Vandeputte et al. 2016). GABA could reach the CNS via blood stream as
399 a lipophilic compound, being able to pass the blood brain barrier. Additionally, microbial
400 GABA could affect the brain-gut axis by contributing the human GABA pools, especially as it
401 has been shown that the microbiome can affect GABA receptor density in the CNS via the
402 vagus nerve (Bravo et al. 2011). To establish whether and which role *A. muciniphila* and GABA
403 may play a role in prodomal PD, further experimental studies will be required.

404 In order to move beyond mere cataloguing of microbial changes associated with
405 diseases, pathway-based tools (Abubucker et al. 2012) have been developed, in which
406 microbial sequences (or reads) are mapped, e.g., onto KEGG ontologies present in the KEGG
407 database (Kanehisa et al. 2017). Using such tools, Bedarf et al reported decreased glucuronate
408 degradation and an increase in tryptophan degradation and formate conversion (Bedarf et al.
409 2017). Similarly, Heinz-Buschart et al. reported 26 KEGG pathways to be altered in PD
410 microbiomes (Heintz-Buschart et al. 2018). In our study, we complemented the compositional
411 analysis with computational modelling to gain insight into potential functional, i.e., metabolic,
412 consequences of changed microbe abundances in PD. The advantage of our approach is that
413 the functional assignments may be more comprehensive than more canonical methods, such as
414 KEGG ontologies because (1) the underlying genome-scale metabolic reconstructions have
415 been assembled based on refined genome annotations and have been manually curated to
416 ensure that the reaction and gene content is consistent with current knowledge about the
417 microbe's physiology, and (2) each of these reconstructions, alone or in combinations, are
418 amenable to metabolic modelling and thus functional and metabolic consequences of a changed
419 environment (e.g., nutrients or other microbes in the models) can be computed. These

simulations are thus allowing to predict functional consequences and not only pathway or reaction enrichment, as typically done.

Strength and limitation

Here, we present microbiome analyses in a large population-based, monocentric case-control study on PD from a defined area (Figure 5). Capitalising on the overall clinical spectrum of PD of the LuxPark cohort, which reflects a representative sample of PD patients of different disease stages from a defined geographical area, we demonstrated that microbial composition is not only altered in PD but also that the observed associations of PD with changes in the composition of the microbiome should be interpreted in the context of age, sex, BMI, and constipation. This information is of importance for clinical translation, highlighting the need for both, (i) a personalised and (ii) a holistic approach, to understand the role of microbial communities in PD pathogenesis. In a second step targeting the potential functional changes related to PD-associated microbiomes, we performed metabolic modelling based on the AGORA collection (Magnusdottir et al. 2017) of genome-scale metabolic reconstructions, allowing for the predictions of metabolite secretion profiles. Thus, our analyses facilitated a detailed investigation of the altered metabolism of PD-related microbial communities in the gut, pointing towards a role of the known pro-inflammatory species *B. wadsworthia* interacting with the host on sulphur metabolism. Hence, metabolic modelling provides a valuable tool for deciphering the metabolic activity of microbial communities in PD.

However, despite the partial confirmation of previous results by our study (Table 5), several limitations should be kept in mind. First, certain covariates were not investigated, such as diet, exercise, and smoking. Whether these covariates alter the PD-specific signature is yet to be analysed. Although our study belongs to the three largest studies performed yet on PD, our sample size was still too small to deliver insights on combinations of drugs. Furthermore,

16S RNA sequencing, as applied in our study, is not allowing analyses on the strain level and may lead to misclassifications (Janda and Abbott 2007), and follow-up studies based on shotgun sequencing are needed to further corroborate our results. However, our results are notably well aligned with a previous shotgun sequencing study (Bedarf et al. 2017), which would further support a role of 16S RNA sequencing as a cost-efficient screening method. Being cross-sectional in nature, causal inference is not possible. Consequently, although metabolic modelling has been numerous times been shown to correctly predict attributes of living systems (Oberhardt et al. 2009; Aurich and Thiele 2016; Nielsen 2017), our hypothesis on the role of *B. wadsworthia* in PD interlinking sulphur metabolism with disease severity requires experimental validation. To conclude, by combining metabolic modelling with comprehensive statistical analyses, we identified a promising research target in PD and refined the understanding of PD-related microbial changes.

METHODS

Description of the Luxembourg Parkinson's study

For this study, data and biospecimen of the LuxPark cohort were utilised (Hipp et al. 2018). The Luxembourg Parkinson's study includes a variegated group of patients with typical PD and atypical parkinsonism, and controls from Luxembourg and its neighbouring regions (Hipp et al. 2018). Controls were partly sampled among relatives of patients. The corresponding information on the family relation between controls and cases was not available. Cancer diagnosis with ongoing treatment, pregnancy, and secondary parkinsonism (drug-induced parkinsonism and parkinsonism in the frame of normotensive hydrocephalus) were exclusion criteria for enrolling in the patient or healthy control group. For 454 individuals (controls: n=248, PD: n=206) from the LuxPark cohort, stool samples were available and used for 16S RNA gene sequencing data (see below). Within LuxPark, controls were selected among

spouses of chosen patients and volunteers and individuals from other independent Luxembourgish studies (Crichton and Alkerwi 2014; Ruiz-Castell et al. 2016). As we aimed to target specifically typical PD (IPD), we excluded all individuals with age below 50 (controls: n=47, PD: n=9) and all individuals with an unclear status of PD diagnosis or an atypical PD diagnosis (PD: n=47). PD patients were defined as typical PD, according to the inclusion criteria by the United Kingdom Parkinson's Disease Society Brain Bank Clinical Diagnostic Criteria (Hughes et al. 1992). Furthermore, we excluded control patients with a United Parkinson's Disease Rating Scale (UPDRS) III score above ten, except for one control where the high UPDRS III score was caused by an arm injury. Furthermore, we excluded control persons who took dopaminergic medications (n=5), and individuals who reported to have taken antibiotics in the last six months (controls: n=20, PD: n=13). Note that excluded observations behave sub-additive, because of overlap between the exclusion criteria (i.e. individuals below age 50 and taking antibiotics). Finally, 309 individuals (controls: n=162, cases: n=147) were included in the statistical analyses.

All study participants gave written informed consents, and the study was performed in accordance with the Declaration of Helsinki. The LuxPark study (Hipp et al. 2018) was approved by the National Ethics Board (CNER Ref: 201407/13) and Data Protection Committee (CNPD Ref: 446/2017).

Measurements and neuropsychiatric testing

All patients and healthy controls were assessed by a neurologist, neuropsychologist or trained study nurse during the comprehensive battery of clinical assessment. Olfaction testing was conducted using the Sniffin' Sticks 16-item version (SS) within the LuxPark cohort (Hipp et al. 2018). Antibiotics usage was defined as intake of antibiotic within the previous six months to stool collection. For assessing PD-related motor and non-motor symptoms, the UPDRS

rating scales I-IV were used (Goetz et al. 2008). The severity of the disease was reflected by the Hoehn and Yahr staging (Hoehn and Yahr 1967). Non-motor symptoms were measured via the NMS questionnaire (Romenets et al. 2012). The use of medication was recorded, and PD-specific medication was classified into three classes, 1) levodopa, 2) dopamine receptor agonist, and 3) MAO-B/COMT inhibitors.

Collection and processing of stool samples

All samples were processed following standard operating procedures (Lehmann et al. 2012; Mathay et al. 2015): stool samples were collected at home by patients using the OMNIgene.GUT stool tubes (DNA Genotek) and sent to the Integrated Biobank Luxembourg (IBBL) where one aliquot of 1 ml was used for DNA extraction. For the DNA extraction, a modified Chemagic DNA blood protocol was used with the MSM I instrument (PerkinElmer), the Chemagic Blood kit special 4 ml (Ref. CMG-1074) with a lysis buffer for faecal samples, and MSM I software. Samples were lysed using the SEB lysis buffer (included in the kit) and vortexed to obtain a homogenous suspension that was incubated for 10min at 70°C, then 5min at 95°C. Lysates (1.5mL) were centrifuged for five minutes at 10,000 g at RT. Supernatants were transferred to a 24XL deep-well plate. Plates were processed using the MSM I automated protocol.

Analysis of the microbial composition with 16S rRNA gene sequencing

The V3-V4 regions of the 16S rRNA were sequenced at IBBL using an Illumina Platform (Illumina MiSeq) using 2x300bp paired-end reads (Hipp et al. 2018). The gene-specific primers targeted the V3 - V4 regions of the 16S rRNA gene. These primers were designed with Illumina overhang adapters and used to amplify templates from genomic DNA. Amplicons were generated, cleaned, indexed, and sequenced according to the Illumina-

520 demonstrated 16S Metagenomic Sequencing Library Preparation Protocol with certain
521 modifications. In brief, an initial PCR reaction contained at least 12.5 ng of DNA. A subsequent
522 limited-cycle amplification step was performed to add multiplexing indices and Illumina
523 sequencing adapters. Libraries were normalised, pooled, and sequenced on the Illumina MiSeq
524 system using 2x300 bp paired-end reads.

525 The demultiplexed samples were processed merging forward and reverse reads and
526 quality filtered using the dedicated pipeline "Merging and Filtering tool (MeFit)" (Parikh et al.
527 2016) with default parameters. To obtain a reliable microbial identification, identification to
528 both genus and species taxonomic level was obtained using the SPINGO (SPecies level
529 IdentificatioN of metaGenOmic amplicons) classifier (Allard et al. 2015) with default
530 parameters. Relative abundances were computed, for each sample, using an R (R Foundation
531 for Statistical Computing, Vienna, Austria) (Ihaka and Gentleman 1996) custom script. Briefly,
532 for each sample, the counts of each genera/species were retrieved, and then the sum of the
533 counts of all the genera/species was used to normalise to a total value of 1 each genera/species
534 count.

535

536 **Personalised constraint-based modelling of microbial communities**

537 AGORA consists of a set of 819 strains of microbes commonly found in the human gut
538 (Magnusdottir et al. 2017; Noronha et al. 2019). To match species taxonomic resolution, we
539 combined strain models of the same species in one species model ('panSpeciesModel.m') using
540 the function 'createPanModels.m' of the microbiome modelling toolbox (Baldini et al. 2018)**.**
541 Briefly, reactions of multiple strains are combined into one pan-reconstruction. The pan-
542 biomass reaction is built from the average of all strain-specific biomass reactions. Microbial
543 abundances were mapped onto a set of 646 species performing an automatic name matching
544 between SPINGO species taxonomic assignment and panSpecies names. A threshold for

545 assessing the bacterial presence of a relative abundance value of 0.0001 was used to reduce the
546 time of computations while limiting the order of magnitude simulations results of
547 stoichiometric coefficients to ten. A total of 259 species overlapped between our set of species
548 models and SPINGO species assignment when considering species identified at least in 10 %
549 of samples (Supplementary Material). The retrieved microbial abundance information for each
550 sample was integrated into a community modelling setup obtaining personalised microbiome
551 models using the automated module of the microbiome modelling toolbox (Baldini et al. 2018)
552 called mgPipe within the COBRA toolbox (Heirendt et al. 2019) (commit:
553 b097185b641fc783fa6fea4900bdd303643a6a7e). Briefly, the metabolic models of the
554 community members are connected by a common compartment, where each model can
555 secrete/uptake metabolites. An average European diet was set as input for each microbiome
556 model (Noronha et al. 2019). A community objective function was formulated based on the
557 sum of each microbial model objective function and constrained to a lower bound of 0.4 per
558 day and upper bound of one per day. A set of exchange reactions connects the shared
559 compartment to the environment enabling to predict metabolite uptake and secretion flux rates
560 (metabolic profiles/NMPCs) consistent with the applied constraints. The personalisation of
561 each microbiome model was achieved by adjusting stoichiometric coefficients in the
562 community biomass reactions to each sample's relative microbial abundance and removing
563 species undetected from the community models.

564 Relative reactions abundances were calculated by summing the number of species
565 having the reaction in a microbiome model and scaling the sum by the respective species
566 relative abundance. Community metabolic profiles of these microbial communities were
567 assessed using flux variability analysis on the exchange reactions (Gudmundsson and Thiele
568 2010). AGORA microbial metabolic reconstructions used for the construction of the
569 community models were downloaded from the VMH (www.vmh.life, (Noronha et al. 2019)).

570 All computations were performed in MATLAB version 2018a (Mathworks, Inc.), using the
571 IBM CPLEX (IBM, Inc.) solver through the Tomlab (Tomlab, Inc.) interface.
572
573 **Analyses of relative abundances**
574 For descriptive statistics, metric variables were described by means and standard
575 deviations, while nominal variables were described by proportions. Missing values were not
576 imputed, and the pattern of missing values was not assessable via the ADA platform (Hipp et
577 al. 2018). The read counts for each metagenomic feature (e.g., genera and species) were divided
578 by total read counts such that relative abundances were retrieved. Relative abundances were
579 checked for outliers. Observations with more than four standard deviations from the mean were
580 excluded from analyses. Only genera and species detected in more than 50% of all samples
581 were included in the analyses, resulting in 62 genera and 127 species.
582
583 The metagenomic data was analysed using fractional regressions as developed by
584 (Papke and Wooldridge 1996). Fractional regressions, first applied to econometric problems,
585 are semiparametric methods designed to model fractional data without the need of specifying
586 the distribution of the response variable. Fractional regressions are further inherently robust
587 against heteroscedasticity and can be parametrised in odds ratios, delivering convenient
588 interpretations of the regression coefficients. All statistical models included technical
589 covariates, batch, total read counts, and unclassified read counts (reads for which a taxonomic
590 assignment was not possible independently from any threshold of confidence estimate value
591 used). The read count variables were included into the statistical model, as it has been shown
592 that normalisation by division can introduce bias if certain statistical assumptions implied by
593 the application of division are not fulfilled (Hertel et al. 2018). In the case of metagenomic
594 data, the effect of read counts would be removed by division if the observations would be

595 sampled from a multinomial distribution. However, this is not a given as species and genera
596 correlate amongst each other, violating the assumptions needed to construct multinomial
597 distributions. In consequence, read count normalisation by division is prone to introduce a bias
598 into metagenomic data; a potential bias, we corrected for by including the read counts as
599 covariates into the model.

600 Before fitting the final statistical models, we explored the associations of basic
601 covariates (age, sex, and BMI) with metagenomic features using fractional regressions as
602 described above to avoid misspecifications of the statistical models. Since the data showed a
603 high range in age and BMI, we checked for potential non-linear associations by including these
604 variables into the models as restricted cubic splines (Harrell 2001) using three knots defined
605 by the 5%-percentile, the median, and the 95%-percentile. As in the case for age, we found
606 species with indications of non-linear age-associations with $p<0.01$, age was modelled in all
607 analyses via restricted cubic splines.

608 All p-values are reported two-tailed. Statistical analyses were performed in STATA
609 14/MP (College Station, Texas, USA). Summary statistics of the performed analyses are given
610 in the Supplementary files 'Supplementary Tables' 1-4.

611

612 **Differences between PD and controls in microbial composition and the influence of**
613 **covariates**

614 To analyse difference between genus abundances between PD and controls, fractional
615 regressions were carried out with the relative abundance of the genus as the response variable,
616 while including technical covariates, age (restricted cubic splines), sex, and BMI into the
617 statistical modelling. The predictor of interest was the study group indicator variable. We
618 corrected for multiple testing using the Benjamini-Hochberg procedure (Benjamini 2010) by
619 setting the false discovery rate (FDR) to 0.05. Consequently, we corrected for 62 tests when

reporting genera results. These analyses were repeated analogously for the taxonomic level of species, while correcting for multiple testing via the FDR.

Next, we explored the possibility of statistical interactions between basic covariates (age, sex, and BMI) and the group indicator. For these analyses, we once again modelled age and BMI via restricted cubic splines allowing for non-linear interaction terms. We only tested two-way interaction terms. All interaction terms were introduced simultaneously into the statistical model and tested on significance via a Wald test (Harrell 2001), correcting for multiple testing via the FDR. For the globally significant test, the single interaction terms were investigated to explore which covariate-group interaction contributed to the overall significance. For interpretation, the interaction terms were visually inspected by plotting the predictions conditional on technical covariates. These analyses were then rerun with species abundances as response variable instead of genus abundances.

We assessed the influence of constipation on the microbial composition. We introduced the binary predictor constipation (yes/no) as additional predictor into the model and the corresponding group-constipation interaction term. Both terms were tested simultaneously on zero with a Wald test. The analyses were once again adjusted for technical covariates, age (restricted cubic splines), sex, and BMI, and we corrected for multiple testing via the FDR.

Analyses of within PD phenotypes in relation to microbial composition

We investigated the association pattern of medication and clinical features regarding the microbial composition. These analyses were only performed on the IPD cases, while controls were excluded from the analyses. First, we analysed the disease duration as measured in years between the date of the stool sampling and the year of the diagnosis. The analyses were conducted as before via fractional regressions with the genus abundances as the response variable, while adjusting for technical covariates, age (restricted cubic splines), sex, and BMI.

Then, we assessed in separate analyses the UPDRS III score as an indicator for motor symptoms, the non-motor symptoms as measured by the NMS, the Hoehn-Yahr staging of the disease as a global measure of disease progression, and the sniff-score. All these analyses were performed adjusted for technical covariates, age (restricted cubic splines), sex, BMI, and disease duration. Each of these series of regression represents 62 test, which was accounted for using the FDR. The impact of medication was analysed by examining three classes of medication, a) levodopa, b) mono-amino oxidase/catechol-O-methyltransferase inhibitors, and c) dopamine receptor agonists. We generated three corresponding binary phenotypes (intake/no intake) and added these three variables simultaneously to the statistical model determining the significance of this add-on via a Wald test. We then tested each medication-class in separate analyses, strictly correcting for multiple testing via the FDR (186 tests in total). The analyses were performed adjusted for technical covariates, age (restricted cubic splines), sex, BMI, and disease duration.

Statistical analyses of fluxes

The NMPCs were log transformed such that the skewness of the distribution was minimised (Box and Cox 1964)). This type of transformation was applied because of the very differently skewed distributions of the single NMPCs. Then, outliers were excluded using the 4-SD outlier rule as before. Only fluxes with more than 50% non-zero values were retained in analyses. Furthermore, NMPCs with distributions not suitable for statistical analyses (e.g., distributions with a high number of observations with exact the same numerical value) were excluded resulting in 129 NMPCs included into analyses.

The NMPCs were analysed with mixed linear regressions including the batch as random effects. Including the batch variable as a random effect has a higher statistical power in comparison to the fixed effect approach, but relies on more restrictive assumptions. We tested

the corresponding random effect assumption by Hausman specification tests and found no indications of violations of the Hausman specification test. Note that this possibility to account for batch effects via random effects is not available with fractional regressions where batch effects were corrected via fixed effects.

We performed the same analyses as with the metagenomic data, with the sole exception of replacing the fractional regression model with the linear mixed model. In all other aspects, the analyses followed the same scheme.

Analyses of species contribution to fluxes

To investigate the contribution of species and genera, we calculated for all included genera and all analysed fluxes the pairwise correlation and the corresponding variance contribution (the squared correlation). We classified every correlation above 0.5 (equal to 25% of variance contribution) as a strong correlation in accordance with classical classifications of effect size (Cohen 1988).

Material availability

All 16S rRNA sequences can be requested from I.T. (ines.thiele@nuigalway.ie). The mgPipe pipeline is available within the COBRA toolbox (https://github.com/opencobra/cobratoolbox), and the custom scripts with related documentation are available at the GitHub repository: https://github.com/ThieleLab/CodeBase/ND_collect.

Acknowledgment

This study was funded by grants from the Luxembourg National Research Fund (FNR) within the National Centre of Excellence in Research (NCER) on Parkinson's disease (FNR/

NCER13/BM/11264123) and the PEARL programme (FNR/P13/6682797 to RK), by European Union's Horizon 2020 research and innovation programme grant agreement No 692320 to RK, and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 757922) to IT.

References

Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B et al. 2012. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol* **8**: e1002358.

Allard G, Ryan FJ, Jeffery IB, Claesson MJ. 2015. SPINGO: a rapid species-classifier for microbial amplicon sequences. *BMC Bioinformatics* **16**: 324.

Aurich MK, Thiele I. 2016. Computational Modeling of Human Metabolism and Its Application to Systems Biomedicine. *Methods in molecular biology (Clifton, NJ* **1386**: 253-281.

Baldini F, Heinken A, Heirendt L, Magnusdottir S, Fleming RMT, Thiele I. 2018. The Microbiome Modeling Toolbox: from microbial interactions to personalized microbial communities. *Bioinformatics (Oxford, England)* doi:10.1093/bioinformatics/bty941.

Barichella M, Severgnini M, Cilia R, Cassani E, Bolliri C, Caronni S, Ferri V, Cancello R, Ceccarani C, Faierman S et al. 2019. Unraveling gut microbiota in Parkinson's disease and atypical parkinsonism. *Movement disorders : official journal of the Movement Disorder Society* **34**: 396-405.

Bedarf JR, Hildebrand F, Coelho LP, Sunagawa S, Bahram M, Goeser F, Bork P, Wullner U. 2017. Functional implications of microbial and viral gut metagenome changes in early stage L-DOPA-naive Parkinson's disease patients. *Genome Med* **9**: 39.

Benjamini Y. 2010. Discovering the false discovery rate. *Journal of the Royal Statistical Society, Statistical Methodology, Series B* **72**: 405-416.

Bonifati V, Rizzu P, Squitieri F, Krieger E, Vanacore N, van Swieten JC, Brice A, van Duijn CM, Oostra B, Meco G et al. 2003. DJ-1(PARK7), a novel gene for autosomal recessive, early onset parkinsonism. *Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology* **24**: 159-160.

Box GEP, Cox DR. 1964. An Analysis of Transformations. *Statistical Methodology Series B* **26**: 211-243.

Bravo JA, Forsythe P, Chew MV, Escaravage E, Savignac HM, Dinan TG, Bienenstock J, Cryan JF. 2011. Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 16050-16055.

Carabotti M, Scirocco A, Maselli MA, Severi C. 2015. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Ann Gastroenterol* **28**: 203-209.

Cersosimo MG, Raina GB, Pecci C, Pellene A, Calandra CR, Gutierrez C, Micheli FE, Benarroch EE. 2013. Gastrointestinal manifestations in Parkinson's disease: prevalence and occurrence before motor symptoms. *J Neurol* **260**: 1332-1338.

Cohen J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Routledge.

Crichton GE, Alkerwi A. 2014. Association of sedentary behavior time with ideal cardiovascular health: the ORISCAV-LUX study. *PLoS One* **9**: e99829.

De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, Collini S, Pieraccini G, Lionetti P. 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 14691-14696.

Devkota S, Wang Y, Musch MW, Leone V, Fehlner-Peach H, Nadimpalli A, Antonopoulos DA, Jabri B, Chang EB. 2012. Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in Il10-/- mice. *Nature* **487**: 104-108.

Di Fonzo A, Dekker MC, Montagna P, Baruzzi A, Yonova EH, Correia Guedes L, Szczerbinska A, Zhao T, Dubbel-Hulsman LO, Wouters CH et al. 2009. FBXO7 mutations cause autosomal recessive, early-onset parkinsonian-pyramidal syndrome. *Neurology* **72**: 240-245.

Fasano A, Visanji NP, Liu LW, Lang AE, Pfeiffer RF. 2015. Gastrointestinal dysfunction in Parkinson's disease. *Lancet Neurol* **14**: 625-639.

Gatto NM, Rhodes SL, Manthripragada AD, Bronstein J, Cockburn M, Farrer M, Ritz B. 2010. alpha-Synuclein gene may interact with environmental factors in increasing risk of Parkinson's disease. *Neuroepidemiology* **35**: 191-195.

Gobert AP, Sagrestani G, Delmas E, Wilson KT, Verriere TG, Dapoigny M, Del'homme C, Bernalier-Donadille A. 2016. The human intestinal microbiota of constipated-predominant irritable bowel syndrome patients exhibits anti-inflammatory properties. *Scientific reports* **6**: 39399.

Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stern MB, Dodel R et al. 2008. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Movement disorders : official journal of the Movement Disorder Society* **23**: 2129-2170.

Graham SF, Rey NL, Ugur Z, Yilmaz A, Sherman E, Maddens M, Bahado-Singh RO, Becker K, Schulz E, Meyerdirk LK et al. 2018. Metabolomic Profiling of Bile Acids in an Experimental Model of Prodromal Parkinson's Disease. *Metabolites* **8**.

Grander C, Adolph TE, Wieser V, Lowe P, Wrzosek L, Gyongyosi B, Ward DV, Grabherr F, Gerner RR, Pfister A et al. 2018. Recovery of ethanol-induced Akkermansia muciniphila depletion ameliorates alcoholic liver disease. *Gut* **67**: 891-901.

Gudmundsson S, Thiele I. 2010. Computationally efficient flux variability analysis. *BMC Bioinformatics* **11**: 489.

Harrell FE. 2001. In *Regression modeling strategies*, (ed. FE Harrell). Springer, New York, USA.

Hasegawa S, Goto S, Tsuji H, Okuno T, Asahara T, Nomoto K, Shibata A, Fujisawa Y, Minato T, Okamoto A et al. 2015. Intestinal Dysbiosis and Lowered Serum Lipopolysaccharide-Binding Protein in Parkinson's Disease. *PLoS One* **10**: e0142164.

Heinken A, Ravcheev DA, Baldini F, Heirendt L, Fleming RMT, Thiele I. 2019. Personalized modeling of the human gut microbiome reveals distinct bile acid deconjugation and biotransformation potential in healthy and IBD individuals. *Microbiome* **7**: 75.

Heinken A, Thiele I. 2015. Anoxic Conditions Promote Species-Specific Mutualism between Gut Microbes In Silico. *Applied and environmental microbiology* **81**: 4049-4061.

Heintz-Buschart A, Pandey U, Wicke T, Sixel-Doring F, Janzen A, Sittig-Wiegand E, Trenkwalder C, Oertel WH, Mollenhauer B, Wilmes P. 2018. The nasal and gut microbiome in Parkinson's disease and idiopathic rapid eye movement sleep behavior disorder. *Movement disorders : official journal of the Movement Disorder Society* **33**: 88-98.

Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, Haraldsdottir HS, Wachowiak J, Keating SM, Vlasov V et al. 2019. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nature protocols* **14**: 639-702.

Hertel J, Harms AC, Heinken A, Baldini F, Thinnes CC, Glaab E, Vasco D, Trenkwalder C, Krüger R, Hankemeier T et al. in revision. Integrated Analyses of Microbiome and

797 Longitudinal Metabolome Data Reveal Microbial-Host Interactions on Sulfur
798 Metabolism in Parkinson's Disease.
799 Hertel J, Rotter M, Frenzel S, Zacharias HU, Krumsiek J, Rathkolb B, Hrabe de Angelis M,
800 Rabstein S, Pallapies D, Bruning T et al. 2018. Dilution correction for dynamically
801 influenced urinary analyte data. *Anal Chim Acta* **1032**: 18-31.
802 Hill-Burns EM, Debelius JW, Morton JT, Wissemann WT, Lewis MR, Wallen ZD, Peddada
803 SD, Factor SA, Molho E, Zabetian CP et al. 2017. Parkinson's disease and
804 Parkinson's disease medications have distinct signatures of the gut microbiome.
805 *Movement disorders : official journal of the Movement Disorder Society* **32**: 739-749.
806 Hipp G, Vaillant M, Diederich NJ, Roomp K, Satagopam VP, Banda P, Sandt E, Mommaerts
807 K, Schmitz SK, Longhino L et al. 2018. The Luxembourg Parkinson's Study: A
808 Comprehensive Approach for Stratification and Early Diagnosis. *Front Aging*
809 *Neurosci* **10**: 326.
810 Hoehn MM, Yahr MD. 1967. Parkinsonism: onset, progression and mortality. *Neurology* **17**:
811 427-442.
812 Hopfner F, Kunstner A, Muller SH, Kunzel S, Zeuner KE, Margraf NG, Deuschl G, Baines
813 JF, Kuhlenbaumer G. 2017. Gut microbiota in Parkinson disease in a northern
814 German cohort. *Brain research* **1667**: 41-45.
815 Huang R, Li T, Ni J, Bai X, Gao Y, Li Y, Zhang P, Gong Y. 2018. Different Sex-Based
816 Responses of Gut Microbiota During the Development of Hepatocellular Carcinoma
817 in Liver-Specific Tsc1-Knockout Mice. *Front Microbiol* **9**: 1008.
818 Hughes AJ, Daniel SE, Kilford L, Lees AJ. 1992. Accuracy of clinical diagnosis of idiopathic
819 Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg*
820 *Psychiatry* **55**: 181-184.
821 Hyland NP, Cryan JF. 2010. A Gut Feeling about GABA: Focus on GABA(B) Receptors.
822 *Front Pharmacol* **1**: 124.
823 Ihaka R, Gentleman R. 1996. R: A Language for Data Analysis and Graphics. *Journal of*
824 *Computational and Graphical Statistics* **5**: 299-314.
825 Janda JM, Abbott SL. 2007. 16S rRNA gene sequencing for bacterial identification in the
826 diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* **45**: 2761-2764.
827 Kalia LV, Lang AE, Hazrati LN, Fujioka S, Wszolek ZK, Dickson DW, Ross OA, Van Deerlin
828 VM, Trojanowski JQ, Hurtig HI et al. 2015. Clinical correlations with Lewy body
829 pathology in LRRK2-related Parkinson disease. *JAMA Neurol* **72**: 100-105.
830 Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives
831 on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353-D361.
832 Keshavarzian A, Green SJ, Engen PA, Voigt RM, Naqib A, Forsyth CB, Mutlu E, Shannon
833 KM. 2015. Colonic bacterial composition in Parkinson's disease. *Movement disorders*
834 *: official journal of the Movement Disorder Society* **30**: 1351-1360.
835 Khalif IL, Quigley EM, Konovitch EA, Maximova ID. 2005. Alterations in the colonic flora and
836 intestinal permeability and evidence of immune activation in chronic constipation. *Dig*
837 *Liver Dis* **37**: 838-849.
838 Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, Yokochi M,
839 Mizuno Y, Shimizu N. 1998. Mutations in the parkin gene cause autosomal recessive
840 juvenile parkinsonism. *Nature* **392**: 605-608.
841 Klitgord N, Segre D. 2010. Environments that induce synthetic microbial ecosystems. *PLoS*
842 *Comput Biol* **6**: e1001002.
843 Laue H, Cook AM. 2000. Biochemical and molecular characterization of taurine:pyruvate
844 aminotransferase from the anaerobe Bilophila wadsworthia. *Eur J Biochem* **267**:
845 6841-6848.
846 Lehmann S, Guadagni F, Moore H, Ashton G, Barnes M, Benson E, Clements J, Koppandi I,
847 Coppola D, Demiroglu SY et al. 2012. Standard preanalytical coding for
848 biospecimens: review and implementation of the Sample PREanalytical Code
849 (SPREC). *Biopreserv Biobank* **10**: 366-374.
850 Lesser GT. 2002. Frequency of bowel movements and future risk of Parkinson's disease.
851 *Neurology* **58**: 838; author reply 838-839.

852 Li W, Wu X, Hu X, Wang T, Liang S, Duan Y, Jin F, Qin B. 2017. Structural changes of gut
853 microbiota in Parkinson's disease and its correlation with clinical features. *Sci China*
854 *Life Sci* **60**: 1223-1233.
855 Lin A, Zheng W, He Y, Tang W, Wei X, He R, Huang W, Su Y, Huang Y, Zhou H et al. 2018.
856 Gut microbiota in patients with Parkinson's disease in southern China. *Parkinsonism*
857 *Relat Disord* **53**: 82-88.
858 Magnusdottir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, Greenhalgh K, Jager
859 C, Baginska J, Wilmes P et al. 2017. Generation of genome-scale metabolic
860 reconstructions for 773 members of the human gut microbiota. *Nat Biotechnol* **35**: 81-
861 89.
862 Mathay C, Hamot G, Henry E, Georges L, Bellora C, Lebrun L, de Witt B, Ammerlaan W,
863 Buschart A, Wilmes P et al. 2015. Method optimization for fecal sample collection
864 and fecal DNA extraction. *Biopreserv Biobank* **13**: 79-93.
865 Naito Y, Uchiyama K, Takagi T. 2018. A next-generation beneficial microbe: Akkermansia
866 muciniphila. *J Clin Biochem Nutr* **63**: 33-35.
867 Natividad JM, Lamas B, Pham HP, Michel ML, Rainteau D, Bridonneau C, da Costa G, van
868 Hylckama Vlieg J, Sovran B, Chamignon C et al. 2018. Bilophila wadsworthia
869 aggravates high fat diet induced metabolic dysfunctions in mice. *Nat Commun* **9**:
870 2802.
871 Nielsen J. 2017. Systems Biology of Metabolism: A Driver for Developing Personalized and
872 Precision Medicine. *Cell metabolism* **25**: 572-579.
873 Noronha A, Modamio J, Jarosz Y, Guerard E, Sompairac N, Preciat G, Danielsdottir AD,
874 Krecke M, Merten D, Haraldsdottir HS et al. 2019. The Virtual Metabolic Human
875 database: integrating human and gut microbiome metabolism with nutrition and
876 disease. *Nucleic Acids Res* **47**: D614-D624.
877 Oberhardt MA, Palsson BO, Papin JA. 2009. Applications of genome-scale metabolic
878 reconstructions. *Molecular systems biology* **5**: 320.
879 Orth JD, Thiele I, Palsson BO. 2010. What is flux balance analysis? *Nat Biotechnol* **28**: 245-
880 248.
881 Paisan-Ruiz C, Jain S, Evans EW, Gilks WP, Simon J, van der Brug M, Lopez de Munain A,
882 Aparicio S, Gil AM, Khan N et al. 2004. Cloning of the gene containing mutations that
883 cause PARK8-linked Parkinson's disease. *Neuron* **44**: 595-600.
884 Papke LE, Wooldridge JM. 1996. Econometric methods for fractional response variables
885 with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*
886 **11**: 619-632.
887 Parikh HI, Koparde VN, Bradley SP, Buck GA, Sheth NU. 2016. MeFiT: merging and filtering
888 tool for illumina paired-end reads for 16S rRNA amplicon sequencing. *BMC*
889 *Bioinformatics* **17**: 491.
890 Petrov VA, Saltykova IV, Zhukova IA, Alifirova VM, Zhukova NG, Dorofeeva YB, Tyakht AV,
891 Kovarsky BA, Alekseev DG, Kostryukova ES et al. 2017. Analysis of Gut Microbiota
892 in Patients with Parkinson's Disease. *Bull Exp Biol Med* **162**: 734-737.
893 Qian Y, Yang X, Xu S, Wu C, Song Y, Qin N, Chen SD, Xiao Q. 2018. Alteration of the fecal
894 microbiota in Chinese patients with Parkinson's disease. *Brain Behav Immun* **70**:
895 194-202.
896 Romenets SR, Wolfson C, Galatas C, Pelletier A, Altman R, Wadup L, Postuma RB. 2012.
897 Validation of the non-motor symptoms questionnaire (NMS-Quest). *Parkinsonism*
898 *Relat Disord* **18**: 54-58.
899 Ruiz-Castell M, Kandala NB, Kuemmerle A, Schritz A, Barre J, Delagardelle C, Krippler S,
900 Schmit JC, Stranges S. 2016. Hypertension burden in Luxembourg: Individual risk
901 factors and geographic variations, 2013 to 2015 European Health Examination
902 Survey. *Medicine (Baltimore)* **95**: e4758.
903 Sampson TR, Debelius JW, Thron T, Janssen S, Shastri GG, Ilhan ZE, Challis C, Schretter
904 CE, Rocha S, Gradinaru V et al. 2016. Gut Microbiota Regulate Motor Deficits and
905 Neuroinflammation in a Model of Parkinson's Disease. *Cell* **167**: 1469-1480 e1412.

906 Saransaari P, Oja SS. 2007. Taurine release in mouse brain stem slices under cell-
907 damaging conditions. *Amino acids* **32**: 439-446.
908 Savica R, Carlin JM, Grossardt BR, Bower JH, Ahlskog JE, Maraganore DM, Bharucha AE,
909 Rocca WA. 2009. Medical records documentation of constipation preceding
910 Parkinson disease: A case-control study. *Neurology* **73**: 1752-1758.
911 Scheperjans F, Aho V, Pereira PA, Koskinen K, Paulin L, Pekkonen E, Haapaniemi E,
912 Kaakkola S, Eerola-Rautio J, Pohja M et al. 2015. Gut microbiota are related to
913 Parkinson's disease and clinical phenotype. *Movement disorders : official journal of*
914 *the Movement Disorder Society* **30**: 350-358.
915 Schneeberger M, Everard A, Gomez-Valades AG, Matamoros S, Ramirez S, Delzenne NM,
916 Gomis R, Claret M, Cani PD. 2015. Akkermansia muciniphila inversely correlates
917 with the onset of inflammation, altered adipose tissue metabolism and metabolic
918 disorders during obesity in mice. *Scientific reports* **5**: 16643.
919 Sleator RD. 2010. The human superorganism - of microbes and men. *Med Hypotheses* **74**:
920 214-215.
921 Streubel A, Siepmann J, Bodmeier R. 2006. Drug delivery to the upper small intestine
922 window using gastroretentive technologies. *Current opinion in pharmacology* **6**: 501-
923 508.
924 Thiele I, Sahoo S, Heinken A, Heirendt L, Aurich MK, Noronha A, Fleming RMT. 2018.
925 Personalized whole-body models integrate metabolism, physiology, and the gut
926 microbiome. *BioRxiv preprint* doi:10.11.01/255885.
927 Tonini M, Crema A, Frigo GM, Rizzi CA, Manzo L, Candura SM, Onori L. 1989. An in vitro
928 study of the relationship between GABA receptor function and propulsive motility in
929 the distal colon of the rabbit. *British journal of pharmacology* **98**: 1109-1118.
930 Turnbaugh PJ, Backhed F, Fulton L, Gordon JI. 2008. Diet-induced obesity is linked to
931 marked but reversible alterations in the mouse distal gut microbiome. *Cell Host*
932 *Microbe* **3**: 213-223.
933 van Kessel SP, Frye AK, El-Gendy AO, Castejon M, Keshavarzian A, van Dijk G, El Aidy S.
934 2019. Gut bacterial tyrosine decarboxylases restrict levels of levodopa in the
935 treatment of Parkinson's disease. *Nat Commun* **10**: 310.
936 Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. 2016. Stool
937 consistency is strongly associated with gut microbiota richness and composition,
938 enterotypes and bacterial growth rates. *Gut* **65**: 57-62.
939 Vandeputte D, Falony G, Vieira-Silva S, Wang J, Sailer M, Theis S, Verbeke K, Raes J.
940 2017. Prebiotic inulin-type fructans induce specific changes in the human gut
941 microbiota. *Gut* **66**: 1968-1974.
942 Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, Fleck JS, Voigt AY, Palleja A,
943 Ponnudurai R et al. 2019. Meta-analysis of fecal metagenomes reveals global
944 microbial signatures that are specific for colorectal cancer. *Nat Med* **25**: 679-689.
945 Wu JY, Wu H, Jin Y, Wei J, Sha D, Prentice H, Lee HH, Lin CH, Lee YH, Yang LL. 2009.
946 Mechanism of neuroprotective function of taurine. *Adv Exp Med Biol* **643**: 169-179.
947 Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M,
948 Hidalgo G, Baldassano RN, Anokhin AP et al. 2012. Human gut microbiome viewed
949 across age and geography. *Nature* **486**: 222-227.
950