(19) SUOMI - FINLAND
(FI)

PATENTTI- JA REKISTERIHALLITUS
PATENT- OCH REGISTERSTYRELSEN
FINNISH PATENT AND REGISTRATION OFFICE

(10) **FI 127671 B**
(12) **PATENTTIJULKAISU**
PATENTSKRIFT
PATENT SPECIFICATION

(45) Patentti myönnetty - Patent beviljats - Patent granted **30.11.2018**

(51) Kansainvälinen patenttiluokitus - Internationell patentklassifikation - International patent classification
C12Q 1/04 (2006.01)
C12Q 1/68 (2018.01)
G01N 33/569 (2006.01)
G01N 33/68 (2006.01)
A61K 35/74 (2015.01)
A61P 25/16 (2006.01)

(21) Patenttihakemus - Patentansökning - Patent application 20145492

(22) Tekemispäivä - Ingivningsdag - Filing date **28.05.2014**

(23) Saapumispäivä - Ankomstdag - Reception date **28.05.2014**

(43) Tullut julkiseksi - Blivit offentlig - Available to the public **29.11.2015**

(73) Haltija - Innehavare - Proprietor
1 •NeuroInnovation Oy, Pellonperäntie 1 B, 00830 HELSINKI, SUOMI - FINLAND, (FI)

(72) Keksijä - Uppfinnare - Inventor
1 •Scheperjans, Filip, HELSINKI, SUOMI - FINLAND, (FI)
2 •Auvinen, Petri Olli Viljami, HELSINKI, SUOMI - FINLAND, (FI)
3 •Aho, Velma Tea Essi, HELSINKI, SUOMI - FINLAND, (FI)
4 •Pereira, Pedro Alexandre Bento, HELSINKI, SUOMI - FINLAND, (FI)
5 •Koskinen, Kaisa Marjaana, HELSINKI, SUOMI - FINLAND, (FI)
6 •Paulin, Lars Göran, HELSINKI, SUOMI - FINLAND, (FI)

(74) Asiamies - Ombud - Agent
Kolster Oy Ab, Salmisaarenaukio 1, 00180 Helsinki

(54) Keksinnön nimitys - Uppfinningens benämning - Title of the invention
Menetelmä Parkinsonin taudin diagnostiikkaan
Förfarande för diagnostik av Parkinsons sjukdom
Method for diagnostics of Parkinson's disease

(56) Viitejulkaisut - Anförda publikationer - References cited
WO 2013056222 A1, WO 2013176774 A1, WO 2013053836 A1, WO 2013036290 A1, US 8580522 B2, WO 2011053653 A2, WO 2013190068 A1,
Aiheuttavatko bakteerit Parkinsonin taudin? [online] 9 April 2014 (09.04.2014), [retrieved 08.12.2014]. Retrieved from the Internet: <URL: http://yle.fi/aihe/artikkeli/2014/04/09/aiheuttavatko-bakteerit-parkinsonin-taudin> the whole document, esp. page 1, last two lines - page 2, first two lines, page 2, paragraph 4, page 3, paragraph 2, , Tottey, William et al. The human gut chip "HuGChip", an explorative phylogenetic microarray for determining gut microbiome diversity at family level. PloS One, 2013, Vol. 8, No. 5, e62544, pages 1-12, Scheperjans, F. et al. Gut microbiota are associated with Parkinson's disease and clinical phenotype - a case-control study. European Journal of Neurology, May 2014, Vol. 21, Suppl. 1, Sp. Iss. SI, page 759, first published online 28 May 2014, DOI: 10.1111/ene.12500 the whole document for information, Scheperjans, F. et al. Gut microbiota are related to Parkinson's disease and clinical phenotype. Movement Disorders, 2014, p. 1-9, published online 5 December 2014 (05.12.2014) [Epub ahead of print] & database Medline [online] Accession no. 25476529 [retrieved 10.12.2014] the whole document

(57) Tiivistelmä - Sammandrag - Abstract

Esillä oleva keksintö liittyy lääketieteen alaan ja erityisesti Parkinsonin tautiin (PD). Erityisesti esillä oleva keksintö liittyy menetelmiin ja välineisiin PD:n varhaiseen toteamiseen. Keksintö liittyy myös menetelmiin ja keinoihin PD:n hoitamiseksi ja ehkäisemiseksi. Keksinnön menetelmän avulla mitataan todennäköisyyttä, jolla tutkittava sairastaa Parkinsonin tautia (PD), mittaamalla yhden tai useamman mikrobitaksan suhteelliset runsaudet tutkittavalta otetussa näytteessä; ja arvioimalla todennäköisyyttä, jolla tutkittava sairastaa PD:a mitattujen runsauksien perusteella. Esillä oleva keksintö tarjoaa uuden lähestymistavan PD:n diagnosoimiseen.

The present invention relates to the field of medicine and in particular to Parkinson's disease (PD). Specifically the present invention relates to methods and means for early detection of PD. The invention relates also to methods and means for treatment or prophylaxis of PD. In the method of the invention a probability of a subject having Parkinson's disease (PD) is determined by measuring the relative abundances of one or multiple microbial taxa in a sample from a subject; and the probability of the subject having PD is determined based on the measured abundances. The present invention provides a novel approach for the diagnostics of PD.

Method for diagnostics of Parkinson's disease

Field of the invention

The present invention relates to the field of medicine and in particular to Parkinson's disease (PD). Specifically the present invention relates to methods and means for early detection of PD. Methods and means for treatment or prophylaxis of PD are also disclosed.

Background of the invention

Parkinson's disease is the most frequent movement disorder. The prevalence in people above 65 years is approximately 1% and it has a great negative effect on quality of life. The cause of Parkinson's disease (PD) is unknown and there are no disease modifying treatments available. In Parkinson's disease (PD), the cardinal motor symptoms are mainly related to the loss of dopaminergic neurons in the substantia nigra {{60 Dickson,D.W. 2009}}. However, neuropathologic changes are much more widespread involving the autonomic nervous system, olfactory structures, lower brainstem and cerebral cortex.{{265 Cersosimo,M.G. 2012; 214 Shannon,K.M. 2012; 44 Braak,H. 2006; 287 Bloch,A. 2006}} Extranigral pathology is related to a broad spectrum of non-motor symptoms (NMS) that have been increasingly recognized as an important feature of PD.{{157 Poewe,W. 2008; 211 Muller,B. 2013}} Gastrointestinal dysfunction, in particular constipation, affects up to 80% of PD-patients and may precede the onset of motor symptoms by years.{{242 Korczyn,A.D. 1990; 170 Savica,R. 2009; 265 Cersosimo,M.G. 2012; 212 Jost,W.H. 1997; 157 Poewe,W. 2008; 213 Noyce,A.J. 2012}}. Idiopathic constipation is one of the strongest risk-factors for PD.{{213 Noyce,A.J. 2012}}

It is not known what factors initiate the pathophysiological cascade leading to neurodegeneration in PD, but an environmental factor likely plays a key role in PD pathogenesis probably against a background of genetic vulnerability {{216 Kieburtz,K. 2013; 217 Wirdefeldt,K. 2011}}. The early involvement of the gastrointestinal tract in PD lends support to the hypothesis that this environmental factor exerts its influences primarily via the gut {{216 Kieburtz,K. 2013; 44 Braak,H. 2006}}. However, recent studies have revealed that, changes of the complex equilibrium of the entire microbiome may be related to human disease {{218 Kim,B.S. 2013; 219 de Vos,W.M. 2012}}

Intestinal microbiota have gained a lot of attention in research in recent years and dysequilibrium of the gut microbiome has been associated with

20145492 PRH 07-11-2018

several diseases, including autism, bowel disease and cancer, rheumatoid arthritis, diabetes, and obesity. The gut microbiome in PD has not been previously investigated.

Currently there is no method available for an early diagnostics of PD. Biomarkers for PD, especially for the premotor phase, are urgently needed since future disease modifying therapies should be initiated as early as possible in the disease process to maximize their effect. Moreover, there is no disease modifying treatment available for PD.

Brief description of the invention

An object of the present invention is thus to provide methods and means for early detection of PD. A further object is to provide methods and means for treatment or prophylaxis of PD.

The objects of the application are achieved by a method for measuring the probability of a subject developing or having Parkinson's disease (PD), wherein the definition of PD includes the premotor period of the disorder in which the subject has no motor symptoms, the method comprising

a. obtaining a sample from a subject;
b. determining the relative abundances of one or multiple microbial taxa in the sample; and
c. determining the probability of the subject having PD based on the abundances measured in b,

wherein a high relative abundance of said taxa indicates a low probability of the subject developing or having PD.

Further, it is disclosed a method for treatment of patients with PD preferably before the appearance of motor symptoms of PD, or for prophylaxis of PD in subjects with conditions or symptoms indicating an increased risk for PD, the method comprising administering an effective amount of a composition that increases the relative abundance of *Prevotellaceae* in the intestines to a subject in need of such treatment.

Additionally, it is disclosed a kit for detection of PD and a composition to be used in the treatment of PD.

The preferred embodiments of the invention are disclosed in the dependent claims.

The inventors of the present application surprisingly noticed an association of gut microbiota and neurodegenerative disease. The inventors found that PD patients did not only have an altered microbiome when com-

20145492 PRH 07-11-2018

pared to matched control subjects, but microbiota were associated also with the motor phenotype of PD. These findings were independent of controlled confounders such as degree of constipation, disease duration, medication, comorbidities, gender, or age.

The present invention provides a novel approach for the diagnostics of PD. The method is rapid, non-invasive and easy to use. The present invention provides a considerable advantage of enabling the individuals having a risk of PD being diagnosed at an early stage preferably before motor symptoms appear e.g. in patients with severe constipation of unknown cause or irritable bowel syndrome. Once diagnosed at an early stage as belonging to the risk group, the onset of PD in the individual can be prevented by modifying the community structure of the gut microbiota. The present disclosure specifically provides novel means for preventing PD in a subject and also means for slowing the disease process or even stopping it. The combination of the method and the composition of the disclosure enable development of personalized treatment and possibly personalized dietary guidance. A further advantage of the invention is that a method to measure the risk of a person to develop PD provides a new tool for investigation of patients with irritable bowel syndrome (IBS) that currently is a diagnose of exclusion (other possible sources of symptoms are excluded). A considerable fraction of these patients do have premotor PD and the present invention the first test that can be used to identify PD at this stage. IBS symptoms are frequently severe and impairing quality of life (pain, constipation, diarrhea) and treatment is purely symptomatic. Diagnosing PD in these patients using the methods of the present invention opens a way for disease modifying treatment. The same applies to patients with idiopathic constipation.

Still a further advantage of the disclosure is that selecting donors for fecal transplantation to PD subjects based on high Prevotellaceae abundance ensures effective treatment and is a clear improvement to a transplantation of feces with unknown microbiome structure.

The present disclosure may improve gut function in PD patients leading to relief of gastrointestinal symptoms including constipation, improved nutritional status It may increase levels of thiamine, folate, short chain fatty acids, and gut hormoes like ghrelin in the body of PD patients leading to alleviaton of symptoms and slowing of disease progression Prevotella enterotype is associated with high prodcution of these molecules and those have been

20145492 PRH 07-11-2018

shown to be decreased in PD and effective in PD treatment

The present invention may provide a tool for differentiating patients with PD from patients with another disease mimicking PD (such as Progressive supranuclear palsy, Multiple system atrophy, Corticobasal degeneration, vascular parkinsonism, Alzheimer's disease, Essential tremor, Normal pressure hydrocephalus)

The disclosed method improves barrier function of the gut mucosa reducing exposure to bacterial endotoxin, inflammation and oxidative stress in the patient organism leading to better health.

Since microbiota composition is associated with certain medications e.g. COMT-inhibitors, analysis of microbiota in PD patients can be useful to select the optimal medications for a patient, minimizing adverse side effects.

Brief description of the drawings

In the following the invention will be described in greater detail by means of preferred embodiments with reference to the attached drawings, in which

Figure 1 shows relative abundances of the 10 most abundant bacterial families in the subsampled dataset;

Figure 2 shows distributions of *Prevotellaceae* abundance in both study groups

Box plots show the distributions of *Prevotellaceae* abundance in both study groups. Black horizontal lines indicate the median values and the boxes around them delineate the IQR. Whiskers extend to the highest value within 1.5 IQR of the upper quartile. Circles represent outliers beyond the whisker limit and asterisks represent extreme outliers beyond 3 IQR of the upper quartile. Median [IQR]: Parkinson 0.16% [0.00%-1.66%]; Control: 0.77% [0.00%-18.18%];

Figure 3 shows ROC curves of discriminators of study group membership

The ROC curve analysis of *Prevotellaceae* abundance was somewhat hampered by the considerable number of samples in both groups being devoid of this family (AUC = 0.591 [95% CI 0.496-0.685], P=0.060). Since samples devoid of Prevotellaceae were equally frequent in PD and control groups, we performed a second ROC curve analysis only including subjects positive for Prevotellaceae (n=103; AUC = 0.664 [95% CI 0.556-0.771]; P=0.004; not shown). Logistic regression classifier including abundances of

20145492 PRH 07-11-2018

Prevotellaceae, Lactobacillaceae, Bradyrhizobiaceae and Clostridiales Incertae Sedis IV: AUC = 0.722 [95% CI 0.641-0.804], P<0.001. Wexner score: 0.747 [95% CI 0.666-0.828], P<0.001. Logistic regression classifier as above, but including also Wexner constipation score: AUC = 0.832 [95% CI 0.766-0.897], P<0.001;

Figure 4 shows distribution of Enterobacteriaceae in phenotypic subgroups of PD

Box plots showing the distributions of Enterobacteriaceae abundances in tremor dominant (TD) and postural instability / gait difficulty (PIGD) phenotypes. Black horizontal lines indicate the median values and the boxes around them delineate the IQR. Whiskers extend to the highest value within 1.5 IQR of the upper quartile. Asterisks represent extreme outliers beyond 3 IQR of the upper quartile. Median [IQR]: TD (0.04 [0.00-0.27]; n=23); PIGD (0.46 [0.07-1.84]; n=40);

Figure 5 shows microbial taxa that can be used as biomarkers for the diagnosis of PD. We chose 10 PD fecal samples with lowest abundances of Prevotellaceae and highest abundances of Ruminococcaceae. The other 10 were control samples with high abundances of Prevotellaceae. For identification of taxa, DNA from fecal samples was shotgun sequenced using Illumina MiSeq. Sequences equal or above 75 nucleotides were taxonomically identified using MG-Rast. Biomarkers for PD were selected using the LefSe method Segata N et al, 2011 . The plots shows the biomarkers found by LEfSe ranking them accordingly to their effect size and associating them with the class with the highest median. Green columns indicate taxa more abundant in PD patients and red columns indicate taxa more abundant in control subjcts.

Figure 6 shows plots and data generated using the partitioning around the medoid (PAM) and Dirichlet multinomial mixture (DMM) methods to classify samples into groups based on similarity of their microbiome community structure on bacterial family level (fecal samples of 72 PD patients and 72 control subjects with microbiome community structure determined as described previously. A: Results of principal component analysis demonstrating that samples are classified into 3 groups (P, R, B). B and D: Only 5 PD subjects, but 19 control subjects were assigned to group P. Chi Square test confirming significant association between study group and classification group. C Non-metric Multidimensional Scaling arranging individual subject samples according to similarity of microbiome community structure. The closer two samples,

20145492 PRH 07-11-2018

the more similar the microbiome. Dots=control subjects Triangles=PD patients. The plot also shows the dominating families. E and F Relative abundance of the five most abundant families in the three groups based on the DMM (E) and PAM (F) methods. Both methods show that the P group shows particularly high abundance of Prevotellaceae, whereas the R group is dominated by Ruminococcaceae and the B group mainly by Bacteroidaceae

Figure 7 shows plots and data generated using the same techniques and samples as for figure 6, but for the bacterial genus level. A: Results of principal component analysis demonstrating that samples are classified into 3 groups. B and D: Only 1 PD subject, but 16 control subjects were assigned to group P (= group 1 from DMM) whereas 45 PD subjects, but only 28 subjects were assigned to group O (= group 3 from DMM). Chi Square test confirming significant association between study group and classification group. C Non-metric Multidimensional Scaling arranging individual subject samples according to similarity of microbiome community structure. The closer two samples, the more similar the microbiome. Dots=control subjects Triangles=PD patients. E and F Relative abundance of the ten most abundant genus in the three groups based on the DMM (E) and PAM (F) methods. The P group (= group 2 in E and group 1 in F) shows particularly high abundance of *Prevotella,* whereas the B group (= group 1 in E and group 2 in F) is dominated by *Bacteroides* and Lachnospiraceae incertae sedis. Group O (= group 3) shows relatively high abundance of *Oscillibacter* and *Saccharofermentans*.

Figure 8 shows microbial taxa that can be used as biomarkers for the diagnosis of PD based on analysis of 72 PD and 72 control fecal samples. Biomarkers for PD were selected using the LefSe method. The plots shows the biomarkers found by LEfSe ranking them accordingly to their effect size and associating them with the class with the highest median. Light grey columns indicate taxa more abundant in PD patients and dark grey columns indicate taxa more abundant in control subjcts.

Detailed description of the invention

The present invention relates to a method, wherein the probability of a subject having Parkinson's disease (PD) is determined by a method wherein a sample is obtained from a subject; the relative abundances of one or more microbial taxa in the sample are measured; and the probability of the subject having PD based is determined based on the measured relative abundances of one or multiple microbial taxa in the sample.

20145492 PRH 07-11-2018

The present invention concerns generally Parkinsons's disease and also Parkinsons's disease where the definition of PD includes the premotor period of the disorder in which the subject has no motor symptoms. According to the present invention PD is suspected based on the subject having one or more of the following conditions that are associated with PD: family history of PD, family history of tremor, constipation, irritable bowel syndrome, hyposmia, REM-sleep behavior disorder, mood disorder, hyperechogenicity of the substantia nigra on transcranial ultrasound examination.

In the present invention a sample is a gut microbiota sample obtained from lower gastrointestinal tract, preferably it is e.g. feces sample.

A microbial taxon as used herein refers to a taxonomic unit, whether named or not: i.e. a population, or group of populations of organisms which are usually inferred to be phylogenetically related and which have characters in common which differentiate the unit (e.g. a genus, a family) from other such units. A taxon encompasses all included taxa of lower rank and individual organisms. The term as used herein furthermore includes termed species-level phylotypes or operational taxonomic units that are identified only by their complete 16S rRNA sequence and usually defined as sharing 97% or less sequence identity with other entries in the ribosomal databases. In the present invention the measurement of relative abundances of microbial taxa may include all taxa that can be identified in the sample.

Preferably the microbial taxa is *Prevotellaceae*, and most preferably *Prevotella*. According to one embodiment of the invention the taxon is one or more of the following: *Prevotella amnii, Prevotella bergensis, Prevotella bivia, Prevotella bryantii, Prevotella buccae, Prevotella buccalis, Prevotella copri, Prevotella disiens, Prevotella marshii, Prevotella melaninogenica, Prevotella oralis, Prevotella oris, Prevotella ruminicola, Prevotella salivae, Prevotella sp oral taxon 299, Prevotella sp oral taxon 317, Prevotella sp oral taxon 472, Prevotella tannerae, Prevotella timonensis* and *Prevotella veroralis*.

The basic idea behind the present invention is the finding of the reduced abundance of *Prevotellaceae* in PD patients. According to the present study, a person with a relative abundance of *Prevotellaceae* of more than 6.5% is very unlikely to have PD (86.1% sensitivity).

20145492 PRH 07-11-2018

Prevotella is a genus of Gram-negative bacteria. It is a commensal microbe in the human large intestine and has the ability to degrade a broad spectrum of plant polysaccharides {{262 Kang,D.W. 2013}}{{268 Wu,G.D. 2011}}. It not only plays a key role in digesting carbohydrate-rich food and degrading of mucin glycoproteins in the mucosal layer of the gut, but may also interact with the immune system.{{15 Arumugam,M. 2011; 288 Scher,J.U. 2013}}{{15 Arumugam,M. 2011}}. *Prevotella* is the main contributor of one of the recently suggested gut microbiome enterotypes {{15 Arumugam,M. 2011}}.

In addition to Prevotellaceae, also the abundances of *Lactobacillaceae, Verrucomicrobiaceae, Bradyrhizobiaceae, and Clostridiales Incertae Sedis IV* were independently associated with PD in the studies of the present invention. Thus, in one aspect of the invention one or more of the following taxa may be included in the measurement: *Lactobacillaceae, Verrucomicrobiaceae, Bradyrhizobiaceae, Clostridiales Incertae Sedis IV* and *Ruminococcaceae*.

In one embodiment of the invention low relative abundance of *Prevotellaceae* and high relative abundances of *Lactobacillaceae, Verrucomicrobiaceae, Bradyrhizobiaceae, Clostridiales Incertae Sedis IV* and/or *Ruminococcaceae* indicates a high probability of the subject having PD.

According to another embodiment one or more of the following taxa are included in the determination in addition to *Prevotella*: *Sutterella, Saccharofermentans, Mahella, Lactobacillus, Phaeovibrio, Agromonas,* and/or *Anaerotruncus*. In addition to low relative abundance of *Prevotella* low relative abundance of the taxon *Sutterella* and high relative abundance of the taxa *Saccharofermentans, Mahella, Lactobacillus, Phaeovibrio, Agromonas,* and/or *Anaerotruncus* indicates a high probability of the subject having PD.

According to another embodiment one or more of the following taxa are included in the determination in addition to Prevotellaceae and *Prevotella*: *Bacteroidia, Clostridiaceae, Clostridium sensu stricto, Faecalibacterium, Mogibacterium, Oscillibacter, Prevotella, Prevotellaceae, Roseburia, Acetivibrio, Actinobacteria, Allobaculum, Anaerotruncus, Blautia, Butyricicoccus, Clostridium IV, Collinsella, Coriobacteriaceae, Coriobacteriales, Eggerthella, Firmicutes, Parabacteroides, Porphyromonadaceae, Ruminococcaceae, Sporobacterium.* In addition to low relative abundance of *Prevotella* and/or Prevotellaceae low relative abundance of the taxa *Bacteroidia, Clostridiaceae, Clostridium sensu stricto, Faecalibacterium, Mogibacterium, Oscillibacter, and/or Roseburia* and

20145492 PRH 07-11-2018

high relative abundance of the taxa *Acetivibrio, Actinobacteria, Allobaculum, Anaerotruncus, Blautia, Butyricicoccus, Clostridium IV, Collinsella, Coriobacteriaceae, Coriobacteriales, Eggerthella, Firmicutes, Parabacteroides, Porphyromonadaceae, Ruminococcaceae, and/or Sporobacterium* indicates a high probability of the subject having PD.

According to one embodiment of the invention in addition to the low relative abundance of *Prevotella* and the taxa from one or more of the further bacterial species, which may be selected from the group of *Bacteroides barnesiae, Bacteroides coprocola, Bacteroides coprophilus, Clostridium argentinense, Dialister invisus, Dialister micraerophilus, Fibrobacter succinogenes, Haemophilus influenzae, Haemophilus parainfluenzae, Leadbetterella byssophila, Shigelladysenteriae, Shigella* and *Sporanaerobacteracetigenes,* and the high relative abundance of one or more of the taxa, which may be selected from the following group: *Achromobacter denitrificans, Alistipes putredinis, Alistipes shahii, Alistipes sp HGB5, Alkaliphilus metalliredigens, Bacteroides dorei, Bacteroides eggerthii, Bacteroides fragilis, Bacteroides intestinalis, Bacteroides, Bacteroides sp 1 1 14, Bacteroides sp 2 1 16, Bacteroides sp 2 1 7, Bacteroides sp 2 2 4, Bacteroides sp 3 1 33FAA, Bacteroides sp 3 2 5, Bacteroides sp 4 1 36, Bacteroides sp 9 1 42FAA, Bacteroides sp D20, Bacteroides uniformis, Bifidobacterium longum, Bifidobacterium scardovii, Bilophila wadsworthia, Caldicellulosiruptor saccharolyticus, Clostridium cellulolyticum, Clostridium hathewayi, Clostridium methylpentosum, Clostridium symbiosum, Coprococcus catus, Desulfovibrio, Desulfovibrio vulgaris, Enterococcus faecalis, Enterococcus faecium, Erysipelotrichaceae bacterium 3 1 53, Erysipelotrichaceae bacterium 5 2 54FAA, Ethanoligenens harbinense, Eubacterium cylindroides, Eubacterium siraeum, Eubacterium ventriosum, Holdemania filiformis, Lachnospiraceae bacterium 5 1 63FAA, Lachnospiraceae bacterium 8 1 57FAA, Paenibacillus sp TS12, Paracoccus aminophilus, Ruminococcaceae bacterium D16, Thermoanaerobacter sp X514* indicates a high probability of the subject having PD.

In the present invention the probability of the subject having PD based is determined based on the measured relative abundances of one or multiple microbial taxa in the sample. For example, in one embodiment of the present method high relative abundance of Prevotellacae indicates a low probability of the subject having PD. The high relative abundance is defined as being higher as a reference value and vice versa.

20145492 PRH 07-11-2018

In the method of the invention taxa abundances in reference samples including samples from PD patients are used to generate a logistic regression model, wherein the probability of having PD is modelled as a function of taxa abundances. The value of a logistic regression classifier is calculated based on said logistic regression model and the values of the subject. The value of said logistic regression classifier is then used to determine the probability of the subject having PD.

The method of the present invention may include a quantitative measurement of constipation, wherein values indicating constipation indicate a high probability of the subject having PD. The constipation may be measured e.g by using a questionnaire or an objective assessment of colonic transit time.

In the present invention clinical measures of patient symptoms that may include, but are not restricted to quantitative measurements of constipation, are included in the logistic regression model. Also information on the presence of PD risk alleles in the genome of the subject is included in the model.

Specifically, in the method of the present invention taxa abundances from reference samples including samples from PD patients and healthy controls and abundances from the sample of the subject are subjected to a statistical classification method such as, but not restricted to "partitioning around the medoid" and "Dirichlet multinomial mixture", that assigns samples into groups based on their taxonomic composition. If the method from a statistical classification assigns the subject's sample to a group containing few PD samples, the probability of the subject having PD is determined to be low and vice versa. Also if the method from a statistical classification assigns the subject's sample to a group where the samples have higher abundances of one or more of the taxa *Prevotellaceae* or *Prevotella* as compared to the other groups the probability of the subject having PD is determined to be low.

PD is a clinically heterogenic disorder and it has been suggested that different pathophysiologic mechanisms underlie the observed differences in expression of tremor and non-tremor symptoms between patients. In comparison to tremor dominant (TD) patients, patients with a non-TD phenotype progress faster, have a worse prognosis, and show more severe alpha-synuclein pathology in the colonic enteric nervous system.

20145492 PRH 07-11-2018

It is also disclosed a method for determining the motor subtype of a PD patient. The method comprises obtaining a sample from a subject; determining the abundance of one or multiple of the following taxa: *Enterobacteriaceae, Clostridium XVIII, Anaerofilum, Papillibacter, Succiniclasticum, Klebsiella, Escherichia/Shigella* and *Paludibacter* in said sample; and by using statistical methods determining the motor subtype based on the measured abundances, wherein low relative abundance of the said taxon or taxa indicates a tremor dominant subtype and high relative abundance indicates a non-tremor subtype.

According to the present invention the relative abundance of the microbial taxa can be measured using techniques based on DNA sequencing, quantitative PCR, DNA microarray, by using test beds utilizing arrays or other suitable platforms including microfluidistic solutions, techniques based on droplet PCR, or any other suitable method wherein the abundance of taxa is expressed as read count, percentage, cell count, or value expressing intensity of any other suitable signal.

It is also disclosed a kit for use in a method of the invention, the comprising means for determining alterations in the gut microbiota. The kit comprises reagents and means for determining the abundance of one or more of the taxa mentioned in claims of the present invention in a fecal sample.

It is also disclosed a composition comprising a gut microbiome altering agent that, when administered to an individual, increases the abundance of *Prevotellacae* in the intestine. The composition may comprise live or killed microbes and optionally a pharmaceutically acceptable carrier. Preferably the microbes are from the taxon *Prevotellacae*. The composition may also have a decreasing effect on the abundance of one or more of the taxa *Enterobacteriaceae, Ruminococcaceae, Bacteroidaceae,* Lachnospiraceae, Verrucomicrobiaceae, Bradyrhizobiaceae, and Clostridiales Incertae Sedis IV in the intestines.

The disclosed composition may be in the form of a food composition, pharmaceutical composition, nutraceutical, or supplement. The composition may be administered mixed in food or drink, for example, or separately in the form of a tablets, capsules, microcapsules, powders, solutions, pastes, etc. Food composition may be any kind of food (functional, conventional and novel), food supplement, formula for nutritional purposes, or nutraceutical and it may contain any suitable additives and excipients. The effect of the bacterial

20145492 PRH 07-11-2018

supplementation may be enhanced by adding prebiotics such as fibre to the composition. The composition may also be a prebiotic and optionally contain a pharmaceutically acceptable carrier. The composition may contain live microbes of the taxon *Prevotellacae* and a prebiotic supporting growth of these taxa and optionally a pharmaceutically acceptable carrier. The composition may be used in treatment or prevention of PD.

It is also disclosed a method for treatment or prevention of PD in an individual by administering said individual a composition comprising a gut microbiome altering agent that, when administered to an individual, increases the abundance of *Prevotellacae* in the intestine. The composition to be administered may be the composition as disclosed above. Preferably the treatment is initiated before the diagnosis of PD based on results of the fecal analysis of the present invention. The goal of the treatment is to slow the disease process or to even stop it, or even prevent PD. The term “treatment” may refer to both therapeutic treatment and prophylactic treatment or preventative measures, wherein the goal of the treatment is to slow the disease process or even stop it, or prevent PD.

The administration of a gut microbiome altering agent can be done via a bacterial composition given orally, via a nasogastric tube, as a suppository, or by fecal microbiota transplantation. Preferably the feces donor is previously selected based on a high abundance of Prevotellacae and/or low abundance of one or more of the taxa Enterobacteriaceae, Ruminococcaceae, Bacteroidaceae, Lachnospiraceae, Verrucomicrobiaceae, Bradyrhizobiaceae, and Clostridiales Incertae Sedis IV in feces. It is also preferred that the donor feces is previously selected based on method of the present invention showing that the donor does not have PD.

Fecal microbiota transplantation (FMT) or a stool transplant is the process of transplantation of fecal bacteria from a healthy individual into a recipient. It involves restoration of the colonic flora by introducing healthy bacterial flora through infusion of stool, e.g. by enema.

It will be obvious to a person skilled in the art that, as the technology advances, the inventive concept can be implemented in various ways. The invention and its embodiments are not limited to the examples described below, but may vary within the scope of the claims.

20145492 PRH 07-11-2018

Examples

The study was approved by the ethics committee of the Hospital District of Helsinki and Uusimaa and all participants gave informed consent. The study was registered at clinicaltrials.gov (NCT01536769).

Study subjects

This case-control study compared patients with a diagnosis of PD according to the Queen Square Brain Bank criteria {{231 Berardelli,A. 2013}} with gender- and age-matched (± five years) control subjects without any signs of parkinsonism or potential premotor symptoms. The chance of recruiting patients with monogenic parkinsonism was minimized by restricting the age of motor-symptom onset to above 50 years and by excluding subjects with more than one relative or one first-degree relative with PD {{9 Alcalay,R.N. 2010}}. From the control group subjects with symptoms associated with premotor PD such as hyposmia, REM sleep behaviour disorder (RBD), and accumulation of other NMS were excluded. Further exclusion criteria covered a broad range of conditions and medications that could independently affect the fecal microbiome.

Subjects were recruited from in- and outpatient departments of the participating hospitals as well as via referrals from cooperating neurologists. We also accepted referrals via participants and invited subjects via announcements in patient journals and meetings. Between November 2011 and December 2012, 277 subjects were screened for the study. Of these, 152 (76 patients, 76 controls) were included. All patients were under regular follow-up by a neurologist who was confident in the diagnosis. Patient records were reviewed if available. Due to sample processing 72 patients and 72 controls were included in the final analysis.

Clinical data

Parkinsonian symptoms were measured using the Unified Parkinson's Disease Rating Scale (UPDRS) and the modified Hoehn & Yahr scale (H&Y).{{237 Jankovic,J. 1990}} {{238 Goetz,C.G. 2004; 237 Jankovic,J. 1990}} The PD patients were classified into postural instability and gait difficulty (PIGD), tremor dominant (TD), and mixed phenotypes (MX) as described by Jankovic et al.{{237 Jankovic,J. 1990}} In addition, an akinetic-rigid symptom score was calculated as described by Poletti et al., but items 29 (gait) and 30 (postural stability) were excluded since those are included in the PIGD score.

20145492 PRH 07-11-2018

Overall NMS severity was assessed using the Non-Motor Symptoms Scale (NMSS). The degree of constipation was quantified in more detail using the Wexner constipation score. Diagnosis of active irritable bowel syndrome (IBS) was an exclusion criterion in this study since it is associated with alterations of gut microbiota.

Analysis of fecal microbiota

Sample collection

The subjects collected the fecal samples at home into collection tubes pre-filled with Stool DNA Stabilizer (PSP® Spin Stool DNA Plus Kit, STRATEC Molecular). Within three days, the tubes were transferred to a storing temperature of -80°C until further processing.

DNA extraction and PCR

Total DNA was extracted using the PSP® Spin Stool DNA Plus Kit (STRATEC Molecular). PCR amplification was carried out in an ARKTIK Thermal Cycler (Finnzymes Diagnostics / Thermo Scientific). In the first stage, the V1–V3 regions of the bacterial 16S rRNA gene were amplified in three replicate reactions with universal bacterial primers pA (AGAGTTTGATCMTGGCTCAG) and pD′ (GTATTACCGCGGCTGCTG) with 18-mer overhangs added to the 5' ends of the primers {{232 Edwards,U. 1989; 233 Lane, David J. 1991}}. The replicate PCR products were pooled and purified with Agencourt® AMPure® XP magnetic beads (Agencourt Bioscience) and subjected to a second PCR round with barcoded forward primers and a reverse primer, both of which attached to the respective 18-mer overhang sequences from the primers of the first PCR amplification. Phusion polymerase (Thermo Fisher Scientific / Finnzymes) with HF buffer and 2.5% DMSO were used. Cycling conditions for both PCR reactions consisted of an initial denaturation at 98°C for 30 s, followed by 15 cycles at 98°C for 10 s, 65°C for 30 s, and 72°C for 10 s, and then a final extension for 5 min. Between 3.6 and 60 ng of template DNA were used in the initial reaction. DNA concentration and quality were measured with Qubit (Invitrogen) and Bioanalyzer 2100 (Agilent).

20145492 PRH 07-11-2018

Sequencing and sequence quality control

PCR products were sequenced using 454-GS FLX Titanium chemistry, with an average read length of ~400 bp (Roche Diagnostics). The dataset consisted of 2549217 raw reads (mean read count 17224 per subject). Basic sequence quality control and taxonomical assignment were performed with mothur, following the standard operating procedure for 454 sequenced 16S data.[12,]{{236 Schloss,P.D. 2011}} The phylotype based approach was selected for further analysis. The final dataset included 1131504 reads, with a mean read count of 7645 per subject, and a mean read length of 218 bases. The sequences have been deposited in the Sequence Read Archive at the European Bioinformatics Institute (accession no. PRJEB4927).

Statistical analysis

From each subject, random subsamples of 4500 sequences were used for statistical analysis of family level data. Alpha diversity indices were calculated and beta diversity dendrograms describing the between subject differences were generated and compared with mothur. Differences in bacterial communities between patients and controls and between PIGD and TD patients (a priori defined subgroup analysis) were analyzed using Metastats.[14] The following analyses were performed using IBM® SPSS® Statistics Version 19.0.0.1 (IBM Corp.): Group differences of clinical parameters were analyzed using T-tests for normally distributed variables, otherwise non-parametric tests were used. Differences regarding categorical variables were tested using Fisher's exact test. Associations of microbiome features with clinical parameters were explored using generalized linear models (GLM) employing linear, negative binomial, or logistic distributions depending on the distribution of the target variable.

In the Metastats results, differences with Q-values below 0.05 were considered significant and are the basis of the main conclusions of this paper. We also performed exploratory analyses on bacterial families with higher Q-values if P-values were below 0.05. In the remaining analyses (SPSS), P-values below 0.05 (two-sided if not indicated otherwise) were considered significant.

20145492 PRH 07-11-2018

Results

Demographics and clinical data

The patient and control groups were similar with respect to most of the studied variables (Tables 1).

Table 1

Selected demographic and clinical parameters of the cohort including all those parameters that showed significantly different distributions between groups.

	Patients	Controls	P-Value
Demographics			
n	72	72	
female subjects	48.6%	50.0%	1.000
Age (years, mean ± SD)	65.3 ± 5.5	64.5 ± 6.9	0.448
Medical History			
Atrial Fibrillation	4.2%	18.1%	**0.015**
TIA or Ischaemic Stroke	7.0%	37.5%	**< 0.001**
Non-motor Symptoms			
Overall NMS severity (NMSS	40 [25.25-55.00]	8 [4.00-11.75]	**< 0.001**
Constipation (Wexner score)	5 [3-9]	2 [1-4]	**< 0.001**
Medication			
Levodopa	54.2%	0%	**< 0.001**
COMT Inhibitor	15.3%	0%	**0.001**
Dopamine Agonist	77.8%	0%	**< 0.001**
MAO Inhibitor	70.8%	0%	**< 0.001**
Anticholinergic	8.3%	0%	**0.028**
Warfarin	1.4%	15.3%	**0.004**
Statin	20.8%	54.2%	**< 0.001**

20145492 PRH 07-11-2018

A dopamine transporter SPECT study was documented for 26 (36.1%) patients (TD: 56.5%, MX: 22.2%, PIGD: 27.5%). For 14 (19.4%) patients it was not known whether a SPECT study had been performed at time of diagnosis. The median of the time from motor symptom onset to study visit was 5 [3-8] years (TD: 5 [3-6], MX: 6 [2-12.5], PIGD: 6 [3-10]). The median H&Y stage was 2.5 [2.0-2.5] (Table 2).

Table 2

Distribution of PD patients into the modified Hoehn & Yahr (H&Y) stages

H&Y Stage	1	1·5	2	2·5	3
n (%)	4 (5.6%)	2 (2.8%)	23 (31.9%)	27 (37.5%)	16 (22.2%)

The mean (± SD) scores for total UPDRS and UPDRS-III were 45.5 ± 13.6 and 31.6 ± 9.0 points, respectively. 97.3% of PD patients were using antiparkinsonian medication (Table 1). Two patients (2.8%) were treated by deep brain stimulation (DBS). 55.6% of our patients had a PIGD phenotype, 31.9% were TD, and 12.5% were classified as MX. PD patients had significantly more severe NMS overall (median NMSS score 40 [25.25-55.00] vs. 8 [4.00-11.75] points; $P<0.001$) and constipation in particular than control subjects (median Wexner constipation score 5 [3-9] vs. 2 [1-4] points; $P<0.001$).

Microbiome

The full dataset included bacteria from 360 genera, 125 families, 60 orders, 29 classes and 18 phyla. The majority of the reads (94%) represented the phyla Firmicutes and Bacteroidetes, which are typically the dominant phyla in the gut microbiome.{{15 Arumugam,M. 2011}} In the subsampled dataset the 10 and 5 most common families accounted for 91.6% and 81.3% of all reads, respectively (Figure 1). While no statistically significant differences were found with respect to commonly used alpha diversity indices (Chao1, ACE, Shannon, Inverse Simpson, data not shown), comparisons of the clustering of patient and control samples in dendrograms based on beta diversity metrics (Yue & Clayton theta, Morisita-Horn index and Bray-Curtis index, calculated with family-level data) showed a significant difference between groups (unweighted UniFrac $P<0.02$ and weighted UniFrac $P<0.001$ for all three indices).

20145492 PRH 07-11-2018

The mean abundance of Prevotellaceae in the feces of PD subjects was reduced by 77.6% in comparison to control subjects (Q=0.031; Table 3). High levels of Prevotellaceae were rare in the PD group whereas low levels were found in both groups (Figure 2). Samples devoid of Prevotellaceae were equally frequent in both groups (PD: 29.2%; control: 27.8%; P=1.000; Fisher's exact test). The explorative analysis suggested that five families were more abundant in patients than in controls, but the absolute differences between groups were smaller than for Prevotellaceae (Table 3).

Table 3

Bacterial families showing different abundances between PD and control groups with a P-value less than 0.05.

Taxonomy	Patients*	Controls*	P-Value	Q-Value
Prevotellaceae	2.70 ± 0.32	12.06 ± 3.73	0.001	0.031
Lactobacillaceae	0.44 ± 0.04	0.02 ± 0.00	0.004	0.063
Verrucomicrobiaceae	0.06 ± 0.00	0.02 ± 0.00	0.014	0.146
Bradyrhizobiaceae	0.16 ± 0.00	0.03 ± 0.00	0.021	0.151
Clostridiales Incertae Sedis IV	2.49 ± 0.30	1.01 ± 0.03	0.025	0.151
Ruminococcaceae	33.63 ± 1.99	28.54 ± 1.81	0.029	0.151

* mean % ± variance

Mean abundance and variation of taxa at the genus level (Table 4a) and family level (Table 4b) that showed different abundances in PD and control groups (Metastats analysis) are shown in Table 4a and b. Data is based on same sample as family data.

20145492 PRH 07-11-2018

Table 4a. Mean abundance and variation of taxa at the genus level that showed different abundances in PD and control groups

	Parkinson			Control				
Genus	mean (group1)	variance (group1)	stderr (group1)	mean (group2)	variance (group2)	stderr (group2)	p-value	q-value
Prevotella	0.013423	0.001995	0.005264	0.104250	0.035548	0.022220	0.000999	0.021104
Lactobacillus	0.004358	0.000390	0.002329	0.000216	0.000001	0.000087	0.000999	0.021104
Saccharofermentans	0.047272	0.002385	0.005755	0.029738	0.000912	0.003560	0.016983	0.198570
Anaerotruncus	0.024364	0.000787	0.003307	0.015238	0.000310	0.002074	0.018981	0.198570
Mahella	0.025262	0.003024	0.006480	0.010022	0.000322	0.002114	0.020979	0.198570
Sutterella	0.005485	0.000075	0.001019	0.010182	0.000217	0.001735	0.020979	0.198570
Agromonas	0.001469	0.000021	0.000545	0.000306	0.000001	0.000144	0.022977	0.198570
Phaeovibrio	0.003355	0.000175	0.001561	0.000438	0.000003	0.000203	0.041958	0.332389

20145492 PRH 07-11-2018

Table 4b. Mean abundance and variation of taxa at the family level that showed different abundances in PD and control groups

	Parkinson			Control				
Taxonomy	mean (group1)	variance (group1)	stderr (group1)	mean (group2)	variance (group2)	stderr (group2)	p-value	q-value
Prevotellaceae	0,0270	0,0032	0,0066	0,1206	0,0373	0,0228	0,0010	0,0313
Lactobacillaceae	0,0044	0,0004	0,0023	0,0002	0,0000	0,0001	0,0040	0,0627
Verrucomicrobiaceae	0,0006	0,0000	0,0002	0,0002	0,0000	0,0000	0,0140	0,1462
Bradyrhizobiaceae	0,0016	0,0000	0,0006	0,0003	0,0000	0,0001	0,0210	0,1514
Clostridiales_Incertae_Sedis_IV	0,0249	0,0030	0,0065	0,0101	0,0003	0,0022	0,0250	0,1514
Ruminococcaceae	0,3363	0,0199	0,0166	0,2854	0,0181	0,0158	0,0290	0,1514

20145492 PRH 07-11-2018

It was investigated whether antiparkinsonian medication was a possible confounder by testing each of the abovementioned families for significantly different abundances between users and non-users of each antiparkinsonian drug class in the PD group. Patients using COMT-inhibitors had significantly higher levels (median % [IQR]) of Lactobacillaceae (0.22 [0-2.07] vs. 0 [0-0.02]; P=0.001) and lower levels of Clostridiales Incertae Sedis IV (0.00 [0-0.20] vs. 0.84 [0.02-2.73]; P=0.005) than those not using this class of drugs. No significant differences were found for any other antiparkinsonian drug class.

To estimate effects of possible confounders (Table 1) on the observed group differences GLMs was applied to model the distribution of bacterial abundances (Table 5). For those bacterial families that were more abundant in PD subjects, Prevotellaceae abundance was also included in the model to disentangle a group related effect from an unspecific effect compensating low Prevotellaceae levels. Study group was the only factor that was significantly associated with Prevotellaceae abundance (Table 5). Thus, the decreased abundance of Prevotellaceae was not explained by e.g. more severe constipation in the PD group or differences in medications or medical history. The abundances of all the other families except for Ruminococcaceae were independently related to PD diagnosis. Ruminococcaceae abundance was significantly associated only with levels of Prevotellaceae suggesting that the higher levels of Ruminococcaceae in the PD group were not related to PD itself, but rather compensating lower levels of Prevotellaceae (Table 5). Abundances of Verrucomicrobiaceae and Bradyrhizobiaceae were associated with the severity of NMS and in particular constipation. Abundances of Lactobacillaceae, Bradyrhizobiaceae, and Clostridiales Incertae Sedis IV were associated with statin medication and abundance of Lactobacillaceae also with previous TIA/ischaemic stroke. The independent associations of bacterial abundances with PD diagnosis remained significant when subjects fulfilling Rome-III criteria for IBS were excluded from the analysis.).

20145492 PRH 07-11-2018

20145492 PRH 07-11-2018

Table 5

GLM results

Results of the GLMs for bacterial abundances (sequence counts) based on the group factor and possible confounders. Normal distribution for Ruminococcaceae. Negative binomial distribution with log link for all others. COMT-inhibitor medication effect was nested within the study group effect for Lactobacillaceae and Clostridiales Incertae Sedis IV. Results shown as: B [95% CI], Wald Chi-Square, P-Value

Family	Control vs. PD	Atrial fibrillation No vs. Yes	TIA or ischaemic stroke No vs. Yes	Warfarin No vs. Yes	Statin No vs. Yes	Total NMSS Score (Z-transformed)	Total Wexner Score (Z-transformed)	Prevotellaceae abundance (Z-transformed)	COMT-inhibitor No vs. Yes
Prevotellaceae	**1.490 [1.005-1.976], 36.179, <0.001**	0.432 [-0.489-1.353], 0.845, 0.358	0.025 [-0.517-0.567], 0.008, 0.927	-0.145 [-1.218-0.928], 0.070, 0.791	-0.096 [-0.542-0.349], 0.180, 0.671	0.038 [-0.154-0.230], 0.152, 0.696	-0.040 [-0.270-0.191], 0.114, 0.736		
Lactobacillaceae	**-5.095 [-6.190 - -4.000], 83.151, <0.001**	0.017 [-1.502-1.535], 0.000, 0.983	**1.754 [0.958-2.549], 18.667, <0.001**	-0.258 [-1.943-1.427], 0.090, 0.764	**-1.859 [-2.373 - -1.344], 50.134, <0.001**	-0.199 [-0.590-0.193], 0.988, 0.320	-0.015 [-0.280-0.250], 0.012, 0.911	**0.377 [0.142-0.611], 9.925, 0.002**	**-3.998 [-4.756 - -3.241], 107.025, <0.001**
Verrucomicrobiaceae	**-1.126 [-1.784 - -0.468], 11.254, 0.001**	1.031 [-0.714-2.775], 1.341, 0.247	0.228 [-0.475-0.931], 0.405, 0.525	-1.768 [-3.654-0.117], 3.378, 0.066	0.081 [-0.496-0.658], 0.076, 0.783	**-0.340 [-0.626 - -0.054], 5.444, 0.020**	**0.406 [0.141-0.672], 9.001, 0.003**	**-0.954 [-1.523 - -0.385], 10.801, 0.001**	
Bradyrhizobiaceae	**-2.368 [-3.069 - -1.666], 43.748, <0.001**	-0.109 [-1.307-1.088], 0.032, 0.858	-0.440 [-1.237-0.356], 1.174, 0.279	-0.213 [-1.631-1.204], 0.087, 0.768	**0.687 [0.004-1.370], 3.891, 0.049**	**-0.376 [-0.703 - -0.049], 5.076, 0.024**	**-0.595 [-0.888 - -0.301], 15.769, <0.001**	-0.285 [-0.618-0.048], 2.813, 0.094	

Clostridiales Incertae Sedis IV	**1.441 [0.613-2.269], 11.628, 0.001**	0.173 [-0.418-0.764], 0.330, 0.565	-0.080 [-0.549-0.388], 0.113, 0.737	0.481 [-0.208-1.170], 1.871, 0.171	**-0.856 [-1.303 - -0.409], 14.116, <0.001**	-0.144 [-0.397-0.109], 1.239, 0.266	0.153 [-0.074-0.380], 1.742, 0.187	**-0.712 [-0.938 - -0.485], 37.966, <0.001**	**2.237 [1.557-2.918], 41.535, <0.001**
Ruminococcaceae	-52.526 [-317.204-212.151], 0.151, 0.697	-243.780 [-719.394-231.834], 1.009, 0.315	96.856 [-175.166-368.877], 0.487, 0.485	190.885 [-347.816-729.585], 0.482, 0.487	-112.421 [-339.258-114.416], 0.944, 0.331	-27.605 [-157.321-102.112], 0.174, 0.677	95.836 [-16.063-207.736], 2.818, 0.093	**-212.358 [-311.798 - -112.917], 17.519, <0.001**	

20145492 PRH 07-11-2018

A receiver operating characteristic (ROC) curve analysis of Prevotellaceae abundance was performed with respect to discrimination between PD patients and controls (Figure 3). Using the optimal inclusive cut-off of 6.5% for classifying a subject with low Prevotellaceae abundance into the PD group, PD patients were identified with 86.1% sensitivity but only 38.9% specificity (LR+ = 1.41, LR- = 0.36). It was evaluated whether inclusion of other bacterial families improved the discriminative power by performing a logistic regression analysis with study group as the dependent variable and all six bacterial families as covariates (likelihood ratio based backward elimination method). This model retained Prevotellaceae, Lactobacillaceae, Bradyrhizobiaceae and Clostridiales Incertae Sedis IV as predictors of study group (Table 6).

Table 6

Logistic Regression

Results of logistic regression analysis with study group as the dependent variable. Initially all six bacterial families were included as covariates. We employed a backward elimination method based on the likelihood ratio.

Covariate	Wald	P-Value	OR per 1% increase in abundance for belonging to PD group [95% CI]
Prevotellaceae	7.478	0.006	0.945 [0.907-0.984]
Lactobacillaceae	1.805	0.179	23.441 [0.235-2337.377]
Bradyrhizobiaceae	3.965	0.046	15.788 [1.044-238.716]
Clostridiales Incertae Sedis IV	2.166	0.141	1.112 [0.966-1.280]

This classifier (Figure 3) achieved a significantly higher AUC than Prevotellaceae alone (P=0.020 one-sided) and provided 90.3% specificity and 47.2% sensitivity (optimal probability cut-off 0.54; LR+ = 4.86, LR- = 0.58). When Wexner total score as a clinical measure of constipation was included (Table 7) the discriminative power was further increased significantly (P=0.022 one-sided{{261 Hanley,J.A. 1982}}) due to better sensitivity (66.7%) and preserved 90.3% specificity (optimal probability cut-off 0.55; LR+ = 6.86, LR- =

20145492 PRH 07-11-2018

0.37; Figure 3). The AUC of this model was also larger than that of Wexner score alone, but this difference just missed the level of significance (P=0.055 one-sided).

Table 7

Logistic Regression

Results of logistic regression analysis with study group as the dependent variable after adding Wexner constipation score to the model. Based on Z-transformed values.

Covariate	Wald	P-Value	OR per 1 unit Z-score increase for belonging to PD group [95% CI]
Prevotellaceae	7·053	0·008	0·429 [0·230-0·801]
Lactobacillaceae	1·752	0·186	156·372 [0·088-277179·530]
Bradyrhizobiaceae	3·562	0·059	2·590 [0·964-6·957]
Clostridiales Incertae Sedis IV	1·084	0·298	1·349 [0·768-2·371]
Wexner total score	17·742	<0·001	2·969 [1·789-4·927]

Enterobacteriaceae were signficantly more abundant in patients with a PIGD phenotype than in TD patients (Q=0.018; Figure 4, Table 8).

Table 8

Bacterial families showing differences in abundances between TD and PIGD phenotypes with a p-value less than 0.05.

Taxonomy	TD*	PIGD*	P-Value	Q-Value
Enterobacteriaceae	0.28 ± 0.00	2.31 ± 0.26	0.004	0.018
Erysipelotrichaceae	1.17 ± 0.02	2.19 ± 0.04	0.024	0.103

* mean % ± variance

20145492 PRH 07-11-2018

Potential confounders such as gender ratio, medical history, medication, Wexner score, or time from motor symptom onset did not differ significantly between TD and PIGD subgroups. However, PIGD subjects tended to be older (66.3 ± 5.7 vs. 63.4 ± 5.2 years; mean ± SD; P=0.051) and to have higher NMSS total scores (42.0 [32.0-69.5] vs. 33.0 [19.0-51.0]; P=0.053) than TD subjects. The only antiparkinsonian drug class significantly associated with Enterobacteriaceae abundance was COMT-inhibitors (users: 0.56 [0.24-7.40]; non-users: 0.09 [0.02-0.62]; median % [IQR]; P=0.043). A GLM was used to quantify effects of these possible confounders versus the effect of PD phenotype on Enterobacteriaceae abundance. In this model, PIGD and akinetic-rigid subscores, were positively associated with Enterobacteriaceae abundance whereas the negative association with tremor subscore slightly missed the level of significance (Table 9). When patients with motor fluctuations or DBS treatment were excluded from the analysis (n=43; Table 9), the association with PIGD subscore remained highly significant, but not the association with akinetic-rigid subscore. However, in this model the COMT-inhibitor medication showed a positive association with Enterobacteriaceae abundance.

20145492 PRH 07-11-2018

Table 9

GLM results

Results of the GLMs for Enterobacteriaceae abundance (sequence counts) based on phenotype subscores and possible confounders. Negative binomial distribution with log link. Results shown as: B [95% CI], Wald Chi-Square, P-Value

	Tremor subscore (Z-transformed)	PIGD subscore (Z-transformed)	Akinetic-rigid subscore (Z-transformed)	NMSS total score (Z-transformed)	Age (Z-transformed)	COMT-inhibitor No vs. Yes
Enterobacteriaceae (all patients included; n=72)	-0.254 [-0.517- -0.009], 3.572, 0.059	**0.931 [0.500-1.362], 17.955, <0.001**	**0.445 [0.134-0.757], 7.857, 0.005**	-0.237 [-0.638-0.164], 1.340, 0.247	0.156 [-0.150-0.462], 0.996, 0.318	-0.712 [-1.495 - 0.071], 3.173, 0.075
Enterobacteriaceae (motor fluctuations and DBS excluded; n=43)	-0.143 [-0.619-0.332], 0.349, 0.555	**2.643 [1.698-3.588], 30.075, <0.001**	-0.210 [-0.695-0.274], 0.724, 0.395	-0.344 [-1.065-0.377], 0.876, 0.349	-0.366 [-0.920-0.188], 1.676, 0.196	**-3.502 [-5.836 - -1.167], 8.643, 0.003**

20145492 PRH 07-11-2018

It was found that, although no subject had a known diagnosis of active irritable bowel syndrome (IBS), 18 of 72 (25.0%) of PD patients but only 4 of 72 (5.6%) of control subjects fulfilled Rome-III criteria for IBS (Fisher's exact test p=0.002). While this prevalence in the control group matched that reported in previous studies, this is the first documentation of a clearly increasd prevalence of IBS symptoms in PD patients indicating that IBS may be a non-motor symptom of PD.

References

Cersosimo,M.G. 2012;
Shannon,K.M. 2012; 44
Braak,H. 2006; 287
Bloch,A. 2006}}
Poewe,W. 2008;
Muller,B. 2013}}
Korczyn,A.D. 1990;
Savica,R. 2009; 265
Jost,W.H. 1997;
Poewe,W. 2008;
Noyce,A.J. 2012}}
Kieburtz,K. 2013;
Wirdefeldt,K. 2011}}
Kang,D.W. 2013}}
Wu,G.D. 2011}}
Arumugam,M. 2011;
Scher,J.U. 2013}}
Berardelli,A. 2013}}
Alcalay,R.N. 2010}}.
Hoehn & Yahr scale (H&Y).
Jankovic,J. 1990}}
Goetz,C.G. 2004;
Hanley,J.A. 1982

A. Wilke, T. Harrison, J. Wilkening, D. Field, E.M. Glass, N. Kyrpides, K. Mavrommatis, and F. Meyer. The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. BMC Bioinformatics, 13:141, 2012.). Access date May 2014.

20145492 PRH 07-11-2018

Segata N[1], Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic biomarker discovery and explanation. Genome Biol. 2011 Jun 24;12(6):R60. doi: 10.1186/gb-2011-12-6-r60.

Ding T, Schloss PD. Dynamics and associations of microbial community types across the human body. Nature. 2014 May 15;509(7500):357-60. doi: 10.1038/nature13178. Epub 2014 Apr 16.

20145492 PRH 07-11-2018

Claims

1. A method for measuring the probability of a subject developing or having Parkinson's disease (PD), wherein the definition of PD includes the premotor period of the disorder in which the subject has no motor symptoms, the method comprising

a) obtaining a sample from a subject;

b) measuring the relative abundances of at least *Prevotellaceae* taxa in the sample; and

c) determining the probability of the subject developing or having PD based on the abundances measured in step b),

wherein a high relative abundance of said taxa indicates a low probability of the subject developing or having PD.

2. The method according to claim 1, wherein a sample from a subject is a faecal sample.

3. The method according to claim 1 or 2, wherein subject is human.

4. The method according to any of the previous claims where the taxon is *Prevotella*.

5. A method according to any of the previous claims where the taxon is one or more of the following: *Prevotella amnii, Prevotella bergensis, Prevotella bivia, Prevotella bryantii, Prevotella buccae, Prevotella buccalis, Prevotella copri, Prevotella disiens, Prevotella marshii, Prevotella melaninogenica, Prevotella oralis, Prevotella oris, Prevotella ruminicola, Prevotella salivae, Prevotella sp oral taxon 299, Prevotella sp oral taxon 317, Prevotella sp oral taxon 472, Prevotella tannerae, Prevotella timonensis and Prevotella veroralis*.

6. The method according to claim 1, wherein low relative abundance of *Prevotellaceae* and high relative abundances of one or more of taxa selected from a group consisting of *Lactobacillaceae, Verrucomicrobiaceae, Bradyrhizobiaceae, Clostridiales Incertae Sedis IV, Bacteroidacae* and/or *Ruminococcaceae* indicates a high probability of the subject developing or having PD.

7. The method of claim 1, wherein relative abundances of further bacteria are measured and low relative abundance of *Prevotella* and/or *Prevotellaceae* and of one or more of the following further taxa selected from a group consisting of *Bacteroidia, Clostridiaceae, Clostridium sensu stricto, Faecalibacterium, Mogibacterium, Oscillibacter, Sutterella and Roseburia* and the high relative abundance of one or more of the taxa selected from the following

20145492 PRH 23-07-2018

group: *Acetivibrio, Actinobacteria, Allobaculum, Agromonas, Anaerotruncus, Bacteroides, Blautia, Butyricicoccus, Clostridium IV, Collinsella, Coriobacteriaceae, Coriobacteriales, Eggerthella, Firmicutes, Lactobacillus, Lachnospiraceae, Mahella, Parabacteroides, Phaeovibrio, Porphyromonadaceae, Ruminococcaceae, Oscillibacter, Saccharofermentans* and *Sporobacterium* , indicates a high probability of the subject developing or having PD.

20145492 PRH 23 -07- 2018

Patenttivaatimukset

1. Menetelmä, jolla voidaan mitata todennäköisyyttä, jolla tutkittava sairastaa tai tulee tulevaisuudessa sairastamaan Parkinsonin tautia (PD), jolloin PD:n määritelmä sisältää taudin premotorisen vaiheen, jossa tutkittavalla ei ole motorisia oireita, jolloin menetelmä käsittää:

a. otetaan näyte tutkittavalta;

b. mitataan vähintään *Prevotellaceae* taksonien suhteelliset runsaudet näytteessä; ja

c. määritetään todennäköisyys, jolla tutkittava sairastaa tai tulee tulevaisuudessa sairastamaan PD:a perustuen vaiheessa b) mitattuihin runsauksiin, jolloin mainittujen taksonien suuri suhteellinen runsaus viittaa pieneen todennäköisyyteen, että tutkittava sairastaa tai tulee tulevaisuudessa sairastamaan PD:a.

2. Vaatimuksen 1 mukainen menetelmä, jossa tutkittavalta otettu näyte on ulostenäyte.

3. Vaatimusten 1 tai 2 mukainen menetelmä, jossa tutkittava on ihminen.

4. Minkä tahansa edeltävän patenttivaatimuksen mukainen menetelmä, jossa taksoni on *Prevotella*.

5. Minkä tahansa edeltävän patenttivaatimuksen mukainen menetelmä, jossa taksoni on yksi tai useampi seuraavista: *Prevotella amnii, Prevotella bergensis, Prevotella bivia, Prevotella bryantii, Prevotella buccae, Prevotella buccalis, Prevotella copri, Prevotella disiens, Prevotella marshii, Prevotella melaninogenica, Prevotella oralis, Prevotella oris, Prevotella ruminicola, Prevotella salivae, Prevotella sp oral taxon 299, Prevotella sp oral taxon 317, Prevotella sp oral taxon 472, Prevotella tannerae, Prevotella timonensis* ja *Prevotella veroralis*.

6. Vaatimuksen 1 mukainen menetelmä, jossa *Prevotellaceae*n pieni suhteellinen runsaus ja suuret suhteelliset runsaudet yhdestä tai useammasta taksonista, jotka valitaan seuraavista koostuvasta ryhmästä: *Lactobacillaceae, Verrucomicrobiaceae, Bradyrhizobiaceae, Clostridiales Incertae Sedis IV, Bacteroidacae* ja/tai *Ruminococcaceae,* viittaa suureen todennäköisyyteen, että tutkittava sairastaa tai tulee tulevaisuudessa sairastamaan PD:a.

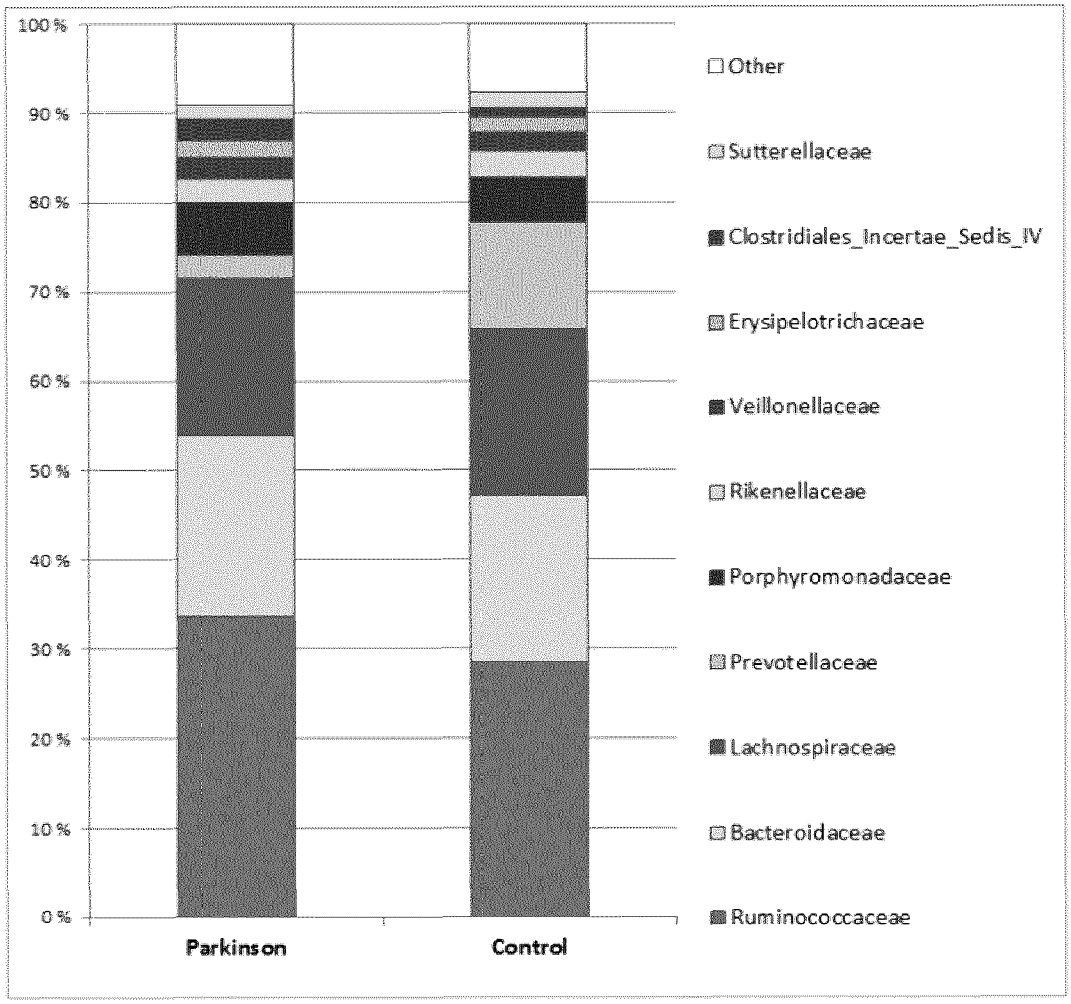
7. Vaatimuksen 1 mukainen menetelmä, jossa mitataan useamman bakteerin runsaudet, ja *Prevotella* ja/tai *Prevotellaceaen* sekä yhden tai use-

20145492 PRH 23 -07- 2018

amman seuraavista lisätaksoneista, jotka valitaan seuraavista koostuvasta ryhmästä: *Bacteroidia, Clostridiaceae, Clostridium sensu stricto, Faecalibacterium, Mogibacterium, Oscillibacter, Sutterella* ja *Roseburia,* pieni suhteellinen runsaus sekä yhden tai useamman taksoneista, jotka valitaan seuraavasta ryhmästä: *Acetivibrio, Actinobacteria, Allobaculum, Agromonas, Anaerotruncus, Bacteroides, Blautia, Butyricicoccus, Clostridium IV, Collinsella, Coriobacteriaceae, Coriobacteriales, Eggerthella, Firmicutes, Lactobacillus, Lachnospiraceae, Mahella, Parabacteroides, Phaeovibrio, Porphyromonadaceae, Ruminococcaceae, Oscillibacter, Saccharofermentans* ja *Sporobacterium,* suuri suhteellinen runsaus viittaa suureen todennäköisyyteen, että tutkittava sairastaa tai tulee tulevaisuudessa sairastamaan PD:a.

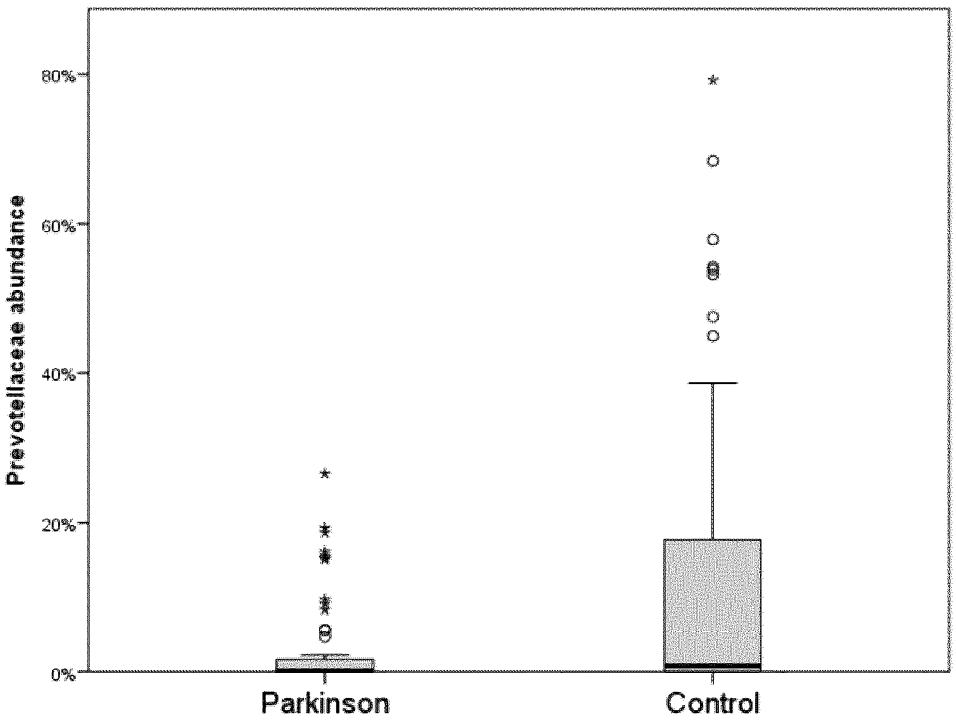
20145492 PRH 23 -07- 2018

Figure 1



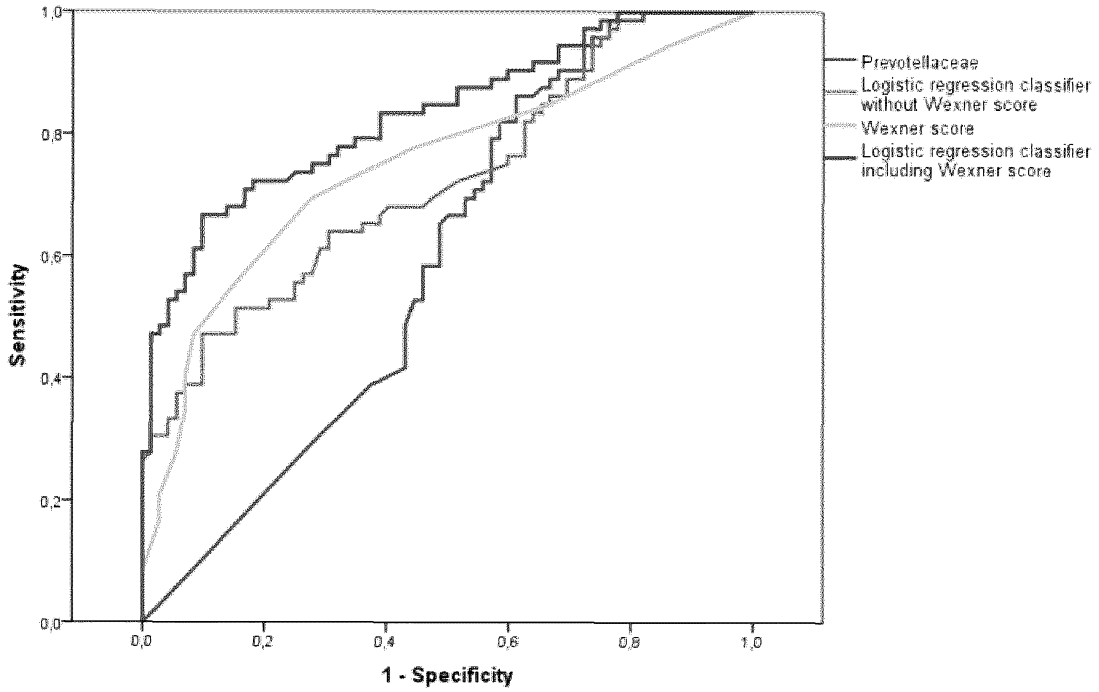
20145492 PRH 07 -11- 2018

Figure 2



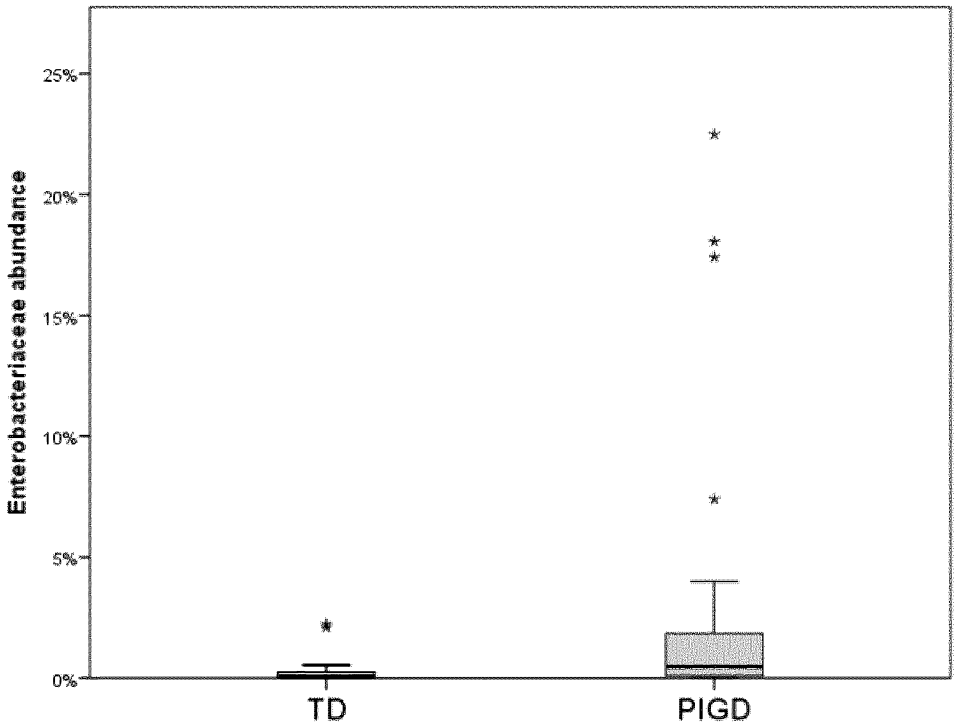
20145492 PRH 07-11-2018

Figure 3



20145492 PRH 07-11-2018

Figure 4



20145492 PRH 07-11-2018

Figure 5

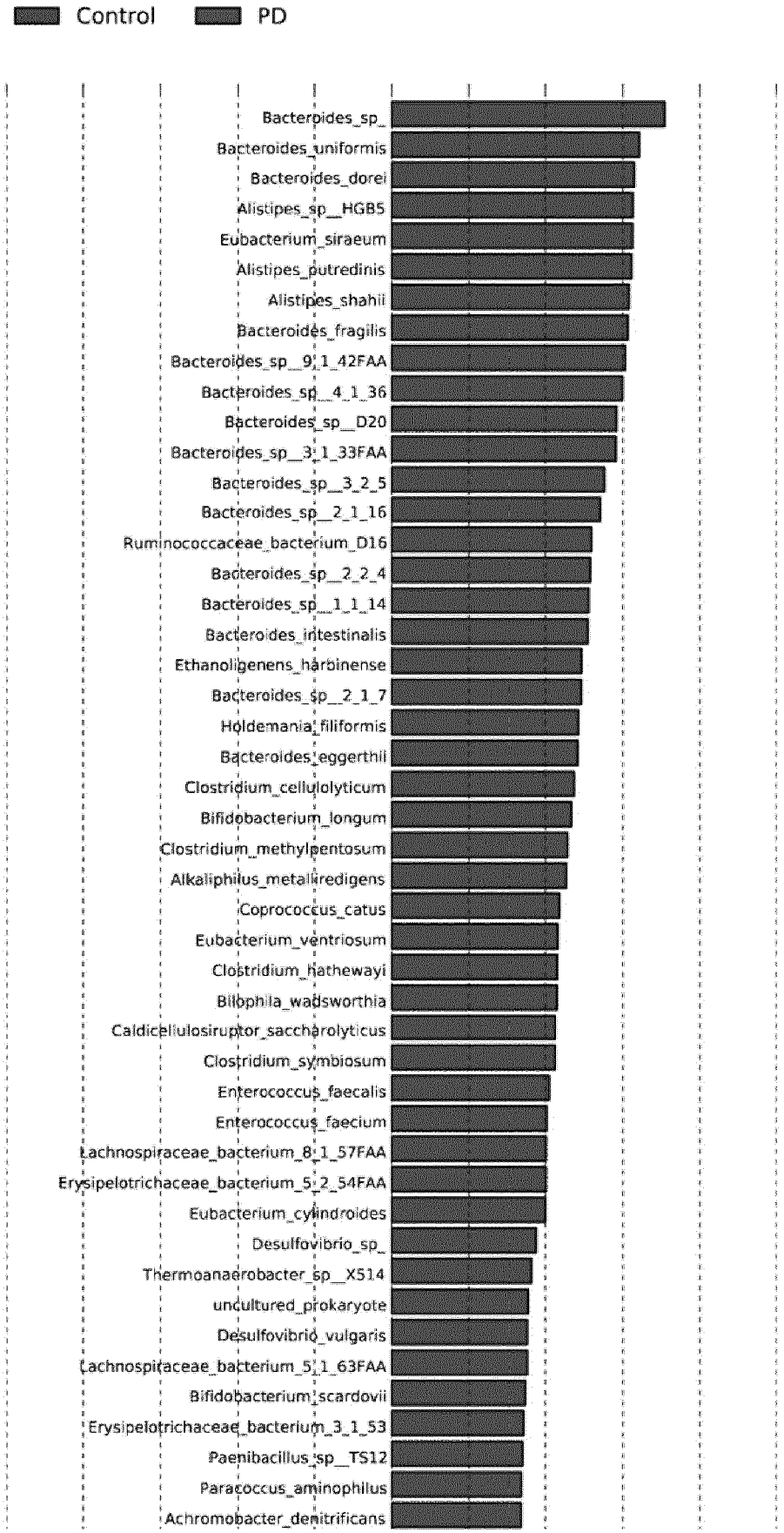
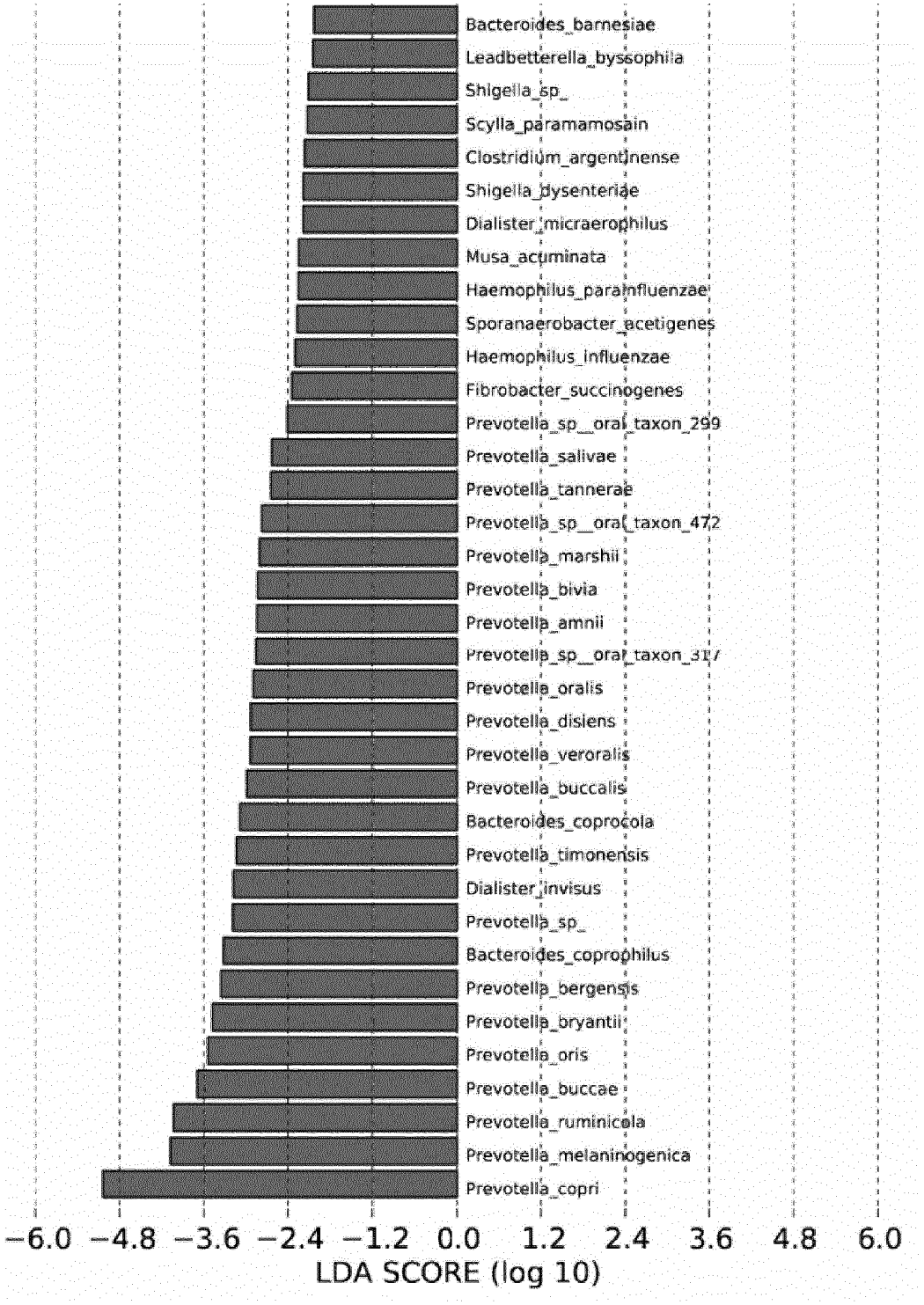


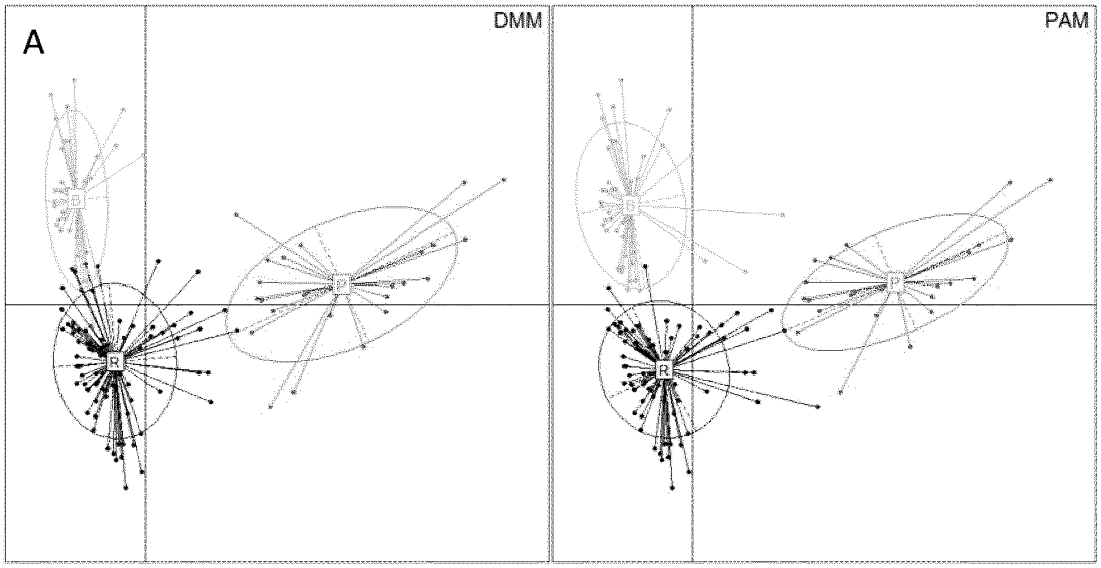
Figure 5 (continued)

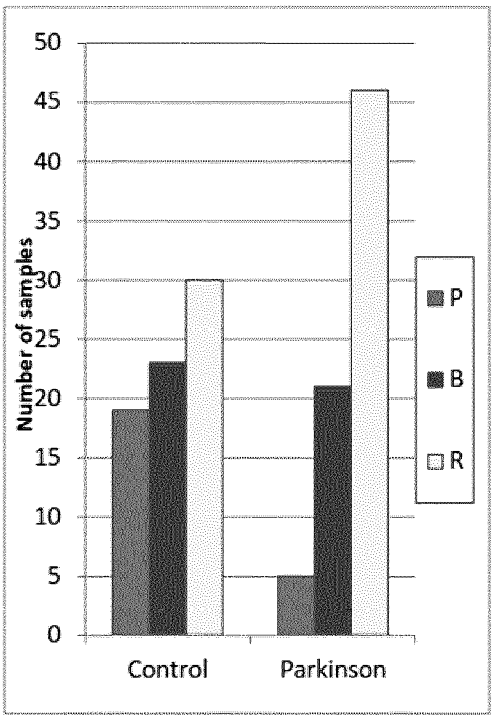
20145492 PRH 07-11-2018



20145492 PRH 07-11-2018

Figure 6A-6B

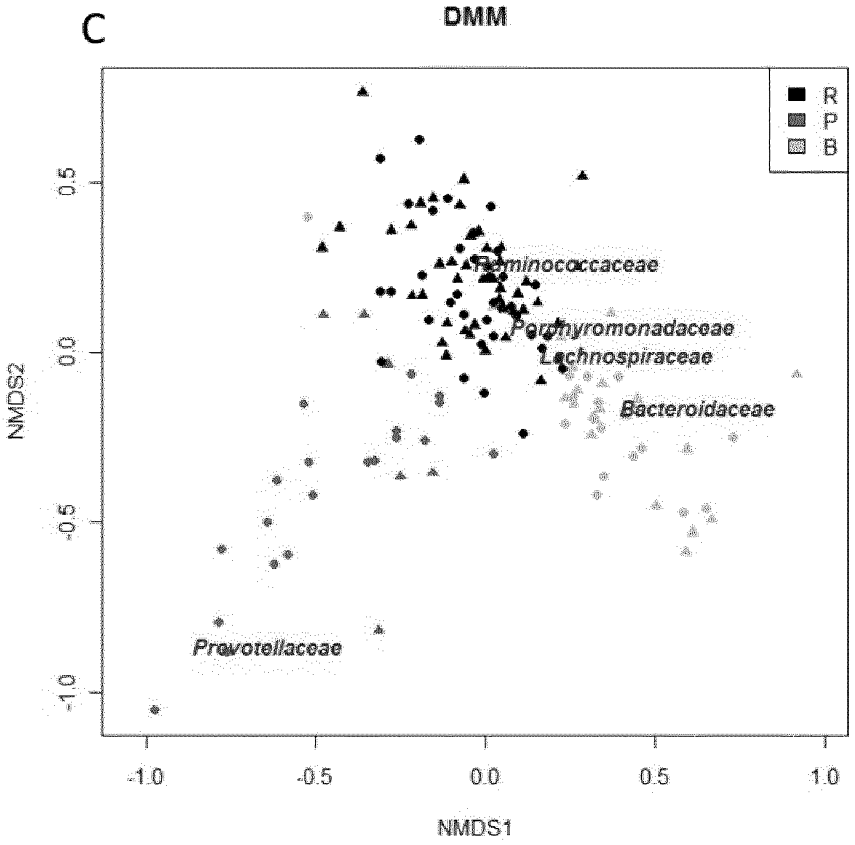




20145492 PRH 07-11-2018

Figure 6C-6D

C



D

Group	Control	Parkinson	Total
P	19	5	24
B	23	21	44
R	30	46	76
Total	72	72	144

Chi square

p=0,003

20145492 PRH 07 -11- 2018

Figure 6E

E
Ruminococcaceae
Bacteroidaceae
Lachnospiraceae
Prevotellaceae
Porphyromonadaceae
Relative abundance
R
P
B

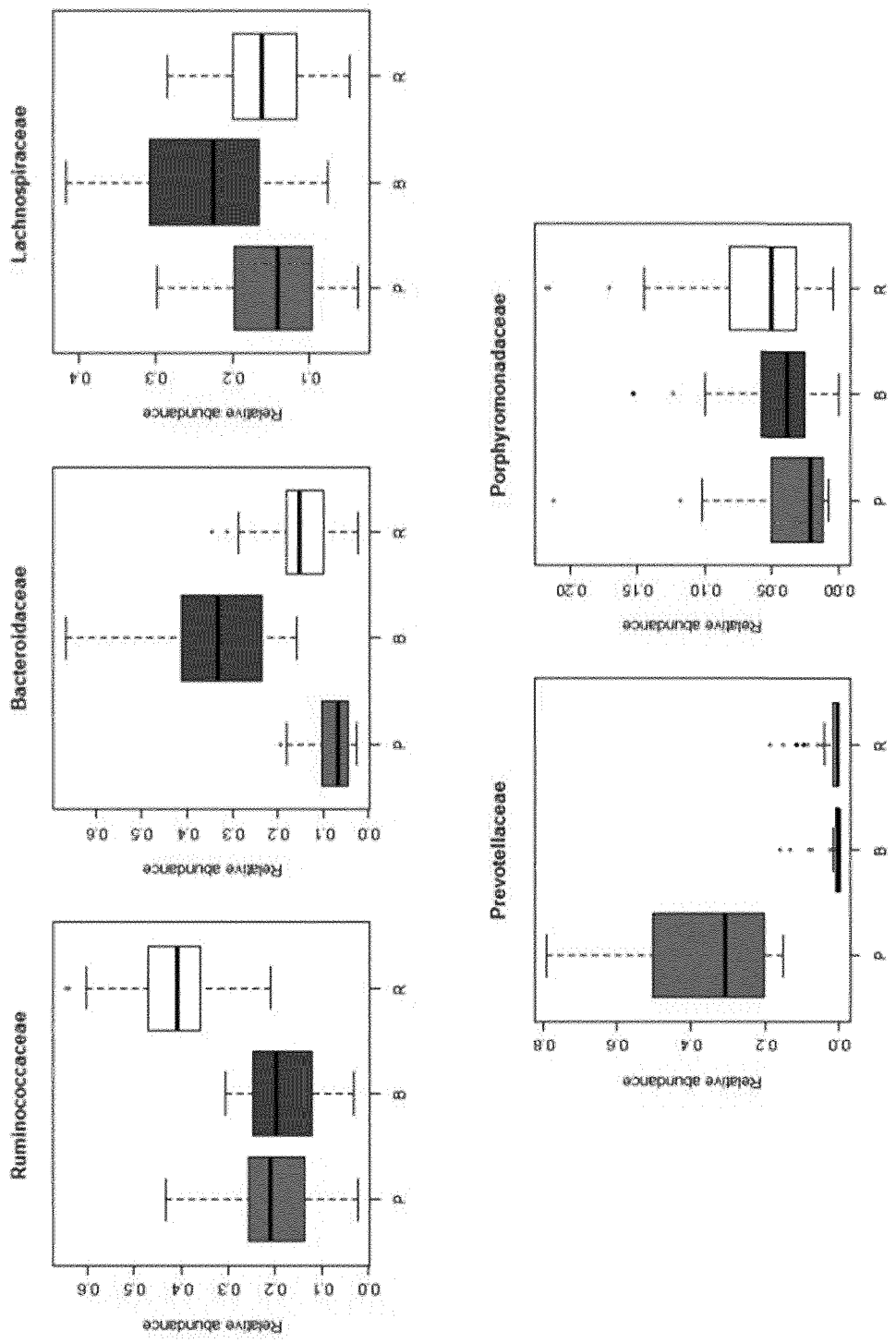
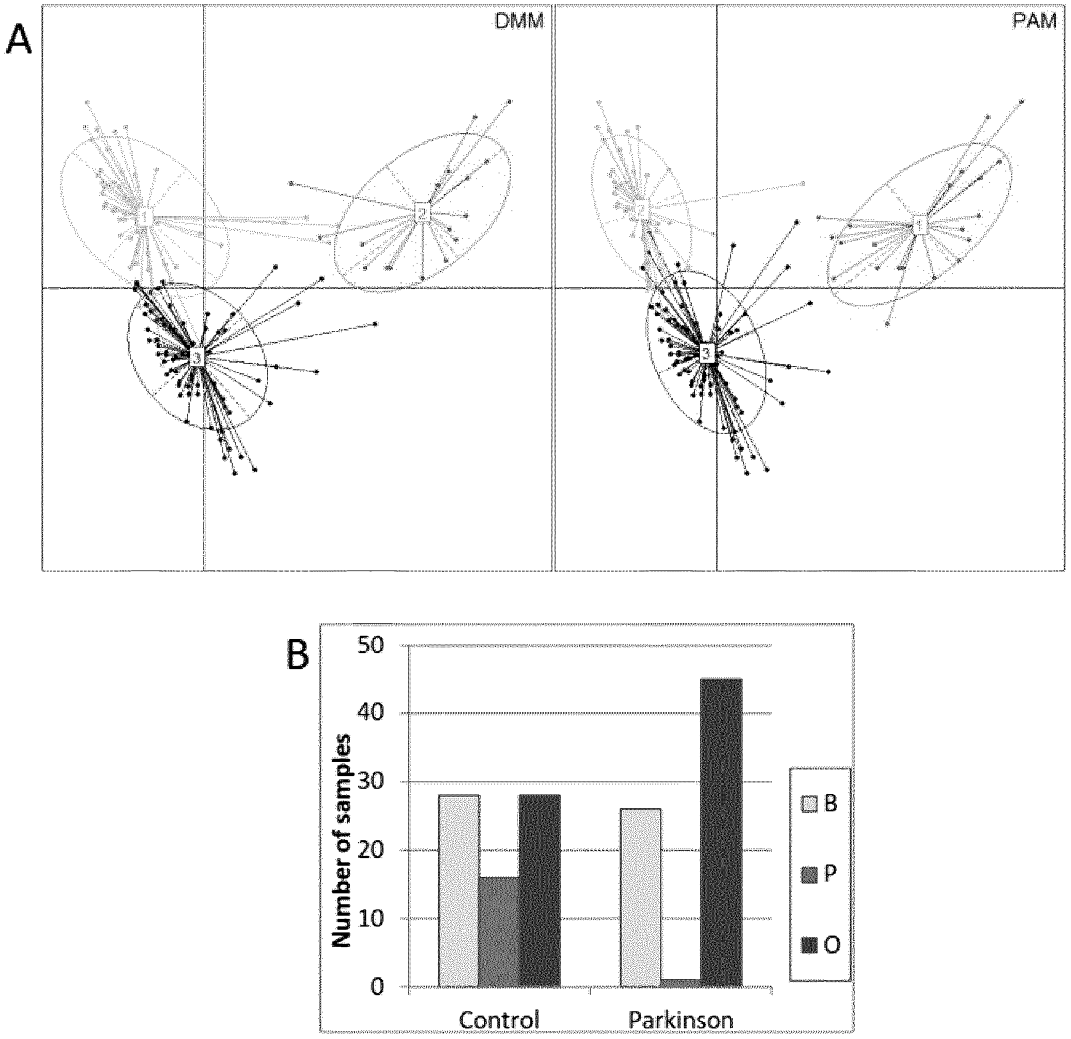

F
Ruminococcaceae
Bacteroidaceae
Lachnospiraceae
Prevotellaceae
Porphyromonadaceae
Relative abundance
P
B
R

Figure 6F

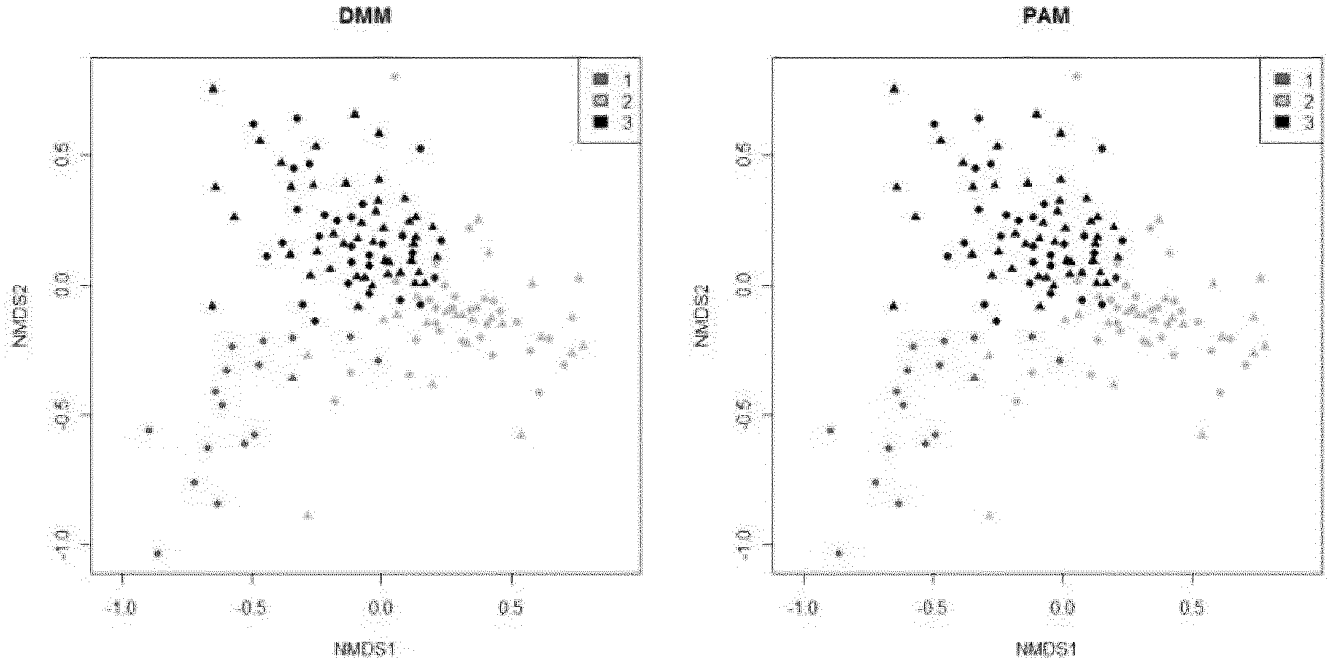
Figure 7A-7B



20145492 PRH 07 -11- 2018

Figure 7C-7D

C



D

Group	Control	Parkinson	Total
B	28	26	54
P	16	1	17
O	28	45	73
Total	72	72	144

Chi square

p=0,0002

20145492 PRH 07 -11- 2018

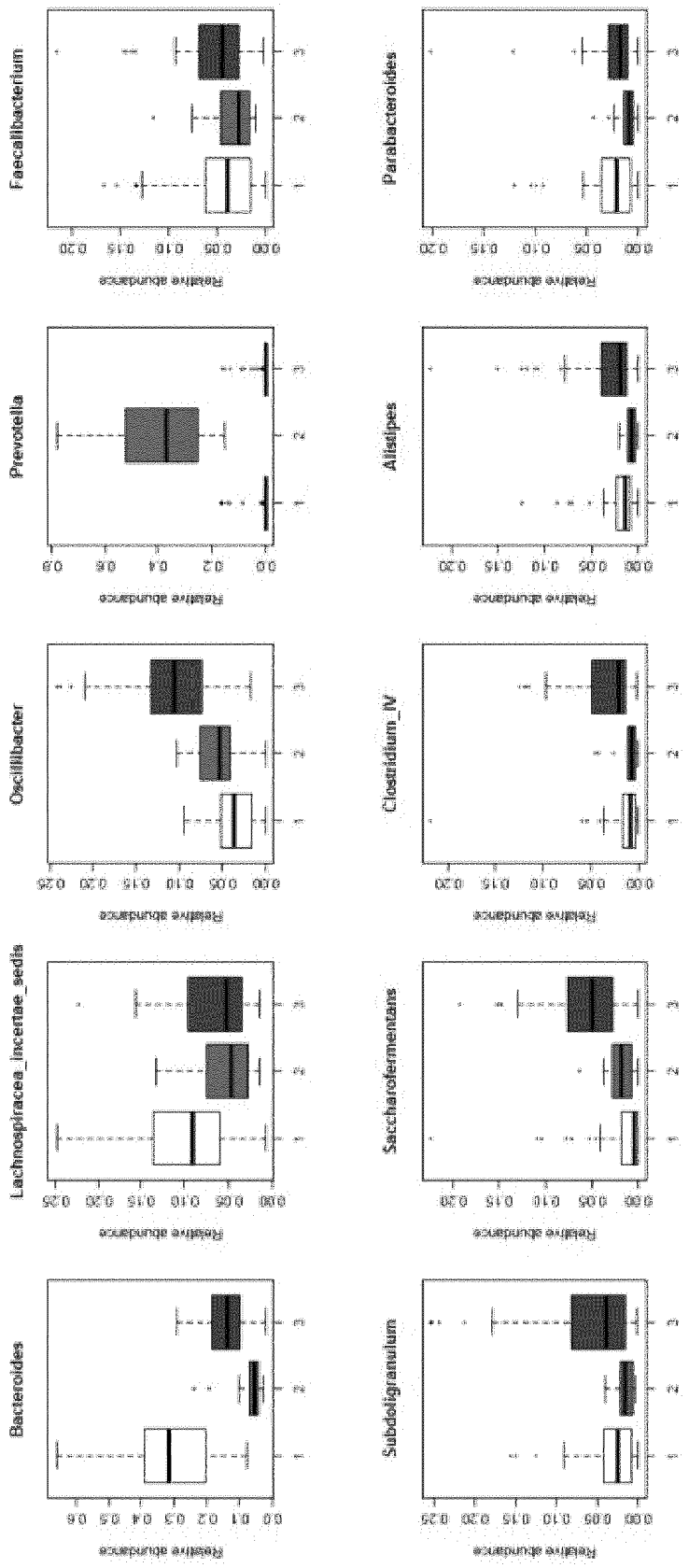

E
Bacteroides
Lachnospiracea_incertae_sedis
Oscillibacter
Prevotella
Faecalibacterium
Subdoligranulum
Saccharofermentans
Clostridium_IV
Alistipes
Parabacteroides
Relative abundance

Figure 7E

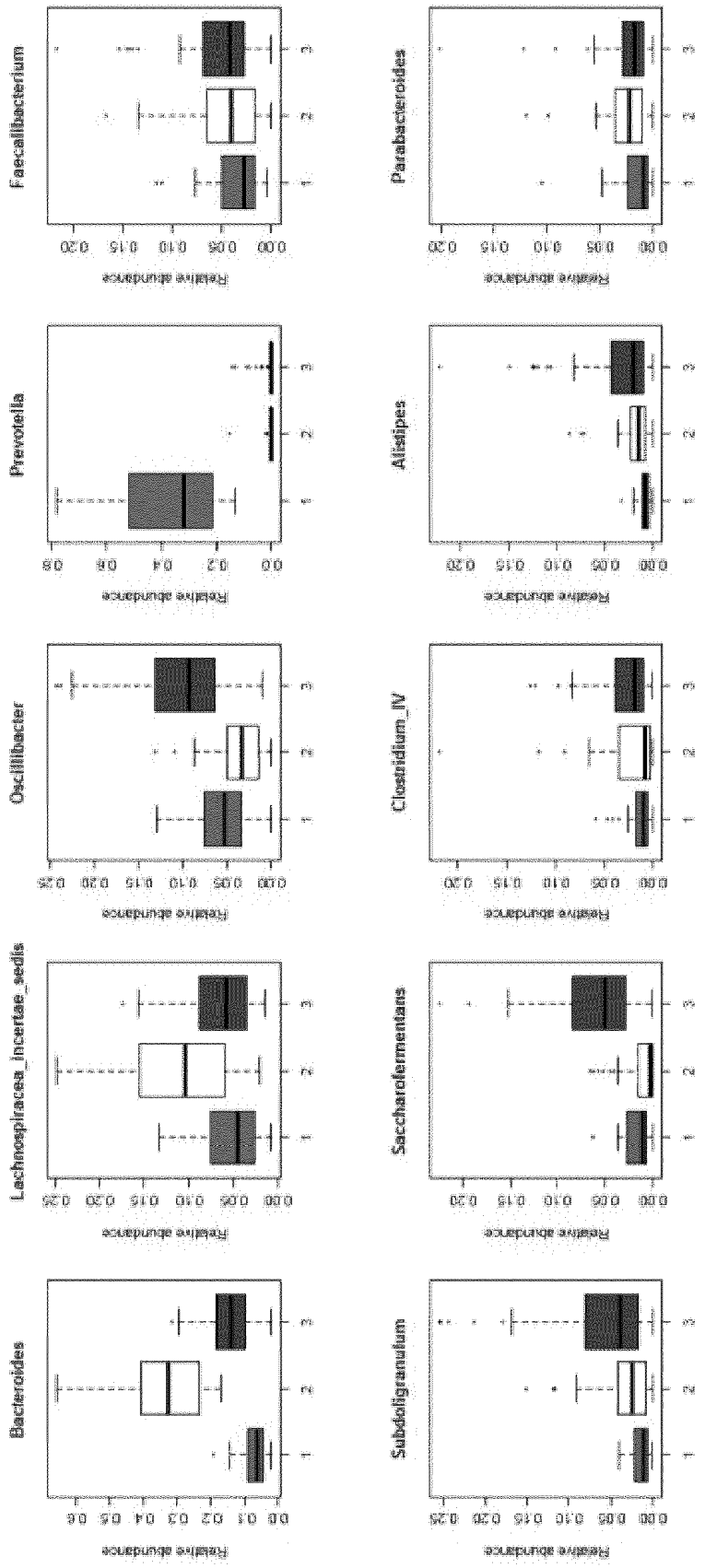
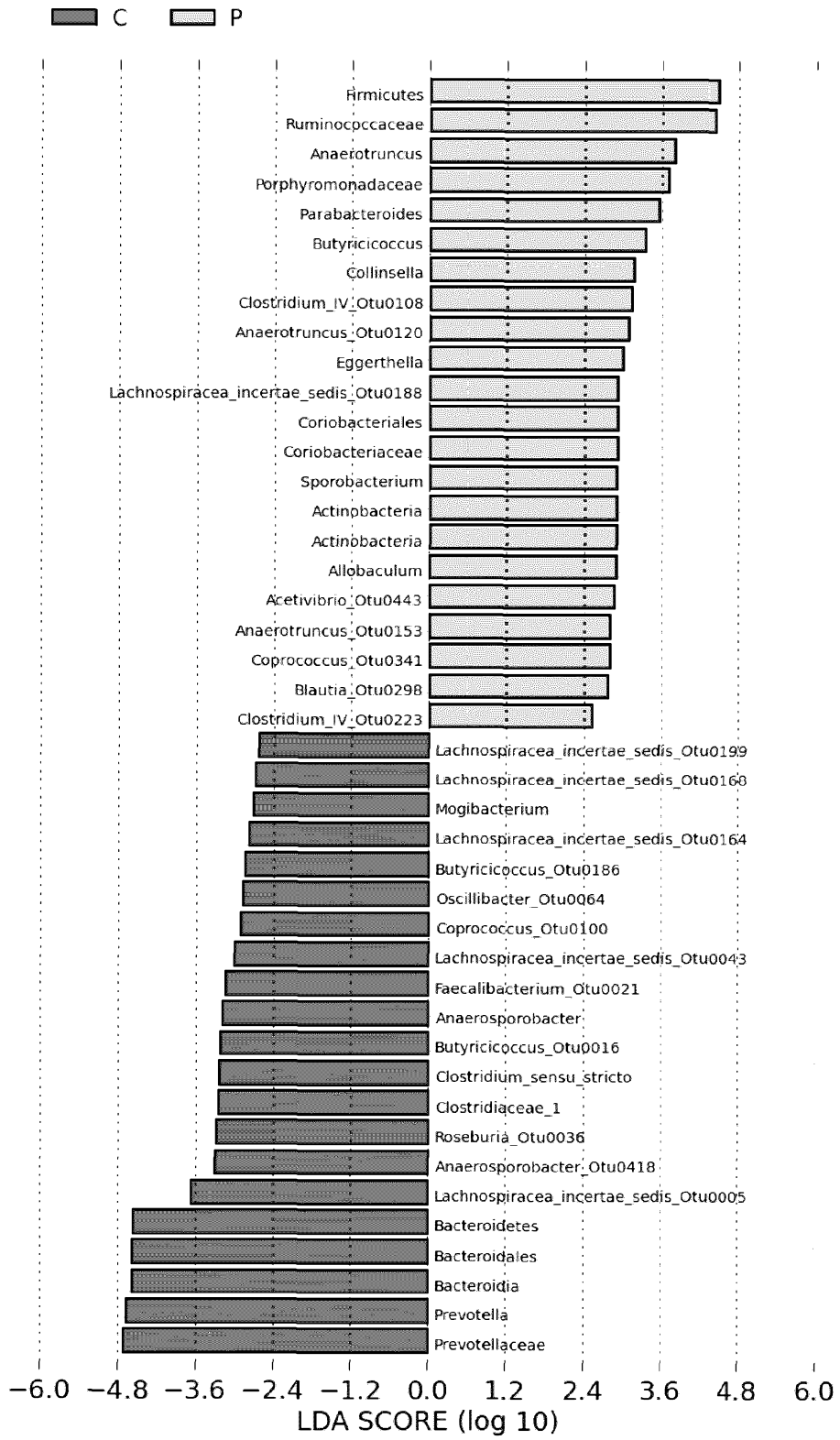

F
Bacteroides
Lachnospiraceae_incertae_sedis
Oscillibacter
Prevotella
Faecalibacterium
Subdoligranulum
Saccharofermentans
Clostridium_IV
Alistipes
Parabacteroides
Relative abundance

Figure 7F

Figure 8



20145492 PRH 07-11-2018