(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 June 2025 (19.06.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/125682 A1

(51) International Patent Classification:
C12Q 1/689 (2018.01)

(21) International Application Number:
PCT/EP2024/086588

(22) International Filing Date:
16 December 2024 (16.12.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
23217079.5 15 December 2023 (15.12.2023) EP

(71) Applicants: **VIB VZW** [BE/BE]; Suzanne Tassierstraat 1, 9052 Gent (BE). **KATHOLIEKE UNIVERSITEIT LEUVEN** [BE/BE]; KU LEUVEN R&D, Waaistraat 6 - bus 5105, 3000 Leuven (BE).

(72) Inventors: **RAES, Jeroen**; c/o VIB-KU Leuven Center for Microbiology, Rega Institute, Herestraat 49, 3000 Leuven (BE). **PROOST, Sebastian**; c/o VIB-KU Leuven Center for Microbiology, Rega Institute, Herestraat 49, 3000 Leuven (BE). **FALONY, Gwen**; c/o VIB-KU Leuven Center for Microbiology, Rega Institute, Herestraat 49, 3000 Leuven (BE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*
— *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

(54) **Title:** MEANS AND METHODS TO CLASSIFY THE GUT MICROBIOME

(57) **Abstract:** The present invention relates to the field of the human gut microbiome, more particularly to its effect on health and disease. Provided herein are means and methods of classifying the gut microbiome based on qPCR data. The invention also provides methods to diagnose and treat or reduce the severity of gut flora dysbiosis as well as of gastro-intestinal inflammation and inflammation-associated disorders or conditions in a subject in need thereof.

WO 2025/125682 A1

MEANS AND METHODS TO CLASSIFY THE GUT MICROBIOME

Field of the invention

The present invention relates to the field of the human gut microbiome and its effect on health and disease. Provided herein are means and methods of classifying the gut microbiome, more particularly of determining the dysbiotic Bacteroides_2 enterotype from a biological sample of a subject. The invention also provides methods to diagnose and treat or reduce the severity of gut flora dysbiosis as well as of gastro-intestinal inflammation and inflammation-associated disorders or conditions in a subject in need thereof.

Background

The last decade, the collection of bacteria living on and in our bodies, i.e. the human microbiome, has been the subject of various studies. With the number of bacteria roughly equalling the number of cells in our bodies, the number of microbial genes outnumbers ours by 100-fold. It is now getting widely accepted that these micro-organisms can play a significant role in the host's health condition. Butyrate producers for example are known to reduce intestinal inflammation (Hamer et al 2007 Aliment Pharmacol Ther 27), *Akkermansia muciniphila* has positive effects on the intestine lining while reducing risks for heart disease and insulin resistance (Hasani et al 2021 J Med Microbiol 70) and *Bifidobacteria* help digest fiber and alleviate diarrhea and constipation (Zhao et al 2017 Medicine 96). On the other hand, various diseases of the gastro-intestinal tract such as Irritable Bowel Syndrome, Ulcerative Colitis, or Crohn's disease have been associated with changes in the gut microbiome composition. Interestingly, the impact of gut microbiota ranges far beyond the gastro-intestinal tract, as links with Type 2 Diabetes and even mental well-being have been found as well (Gurung et al 2020 EBioMedicine 51; Doumatey et al 2020 Front Cell Infect Microbiol 25; Fan and Pedersen 2020 Nat Rev Microbiol 19; Valles-Colomer et al 2019 Nat Microbiol 4; Rogers et al 2016 Molecular Psychiatry 21).

While the gut microbiome is a complex ecosystem with many genera interacting, the flora seems to be classifiable into four non-discreet microbial constellations, called the enterotypes (Arumugam et al 2011 Nature 473; Falony et al 2016 Science 352). Named after the species that set them apart (Bacteroides_1, Bacteroides_2, Prevotella and Ruminococcus) these subjects with these different gut compositions were found to have different outlooks on health. Specifically, the Bacteroides_2 (herein further referred to as B2) enterotype was found to be associated with several intestinal disorders, but also subjects with B2 enterotype were found to have lower self-reported mental Quality-of-Life and are more prone to have diabetes.

While a combination of external factors, such as host genetics, lifestyle, diet, health, and environment have an impact on the gut flora, it is not yet determined which factors matter most. Current methods to determine the enterotype require samples to be processed as described in Falony et al (2016 Science 352). In a nutshell, a well-managed cold chain is required, followed by DNA extraction, 16S amplicon sequencing followed by complex bioinformatical processing where new samples need to be processed along with a sufficiently large set of samples from a wide segment of the population (e.g. samples from the Flemish Gut Flora Project (FGFP), Falony et al 2016 Science 352) to determine the enterotype using a Dirichlet Multinomial Mixture model (DMM) (Holmes et al 2012 PLoS One). These requirements make it impractical, if not impossible for a standard laboratory, to efficiently determine the enterotype of novel samples. The inclusion of a sequencing step, which typically is carried out through an external company, often induces a substantial delay as samples need to be shipped. Furthermore, to make the sequencing cost-efficient there might be additional requirements depending on the service used (such as a full 96-well plate to be submitted) which pose an additional hurdle for this method to be widely adopted.

There is a need for a provision of a simplified, yet accurate and robust means and methods to classify a gut microbiome, and allow for diagnose or reduce the severity of gut flora dysbiosis as well as of gastro-intestinal inflammation and inflammation-associated disorders or conditions in a subject in need. For example, an easy-to-use enterotype test could open doors for novel applications, e.g. regular checks of the enterotype could alert a subject when its diet, lifestyle or environment starts to have a negative effect on its microbiome. Furthermore, the impact of dietary interventions (or use of pre- and pro-biotics) as well as novel drugs on the microbiome could be assessed to objectively determine if there is a negative, neutral, or positive effect on the gut flora.

Summary

The present invention and its aspects and embodiments mentioned below aims to resolve at least a problem present in the current art. The present invention, discloses a novel, preferably qPCR-based approach that makes gut microbiome enterotyping significantly easier.

In a first aspect, the application provides a method of detecting gut flora dysbiosis in a subject, said method comprising measuring abundances of at least two bacterial genera, preferably measuring said abundances of three or four bacterial genera, in a biological sample obtained from said subject, then comparing said measured abundances with reference abundances and determining that the subject has gut flora dysbiosis if the measured abundances are statistically significantly different compared to the reference abundances. In one embodiment, said at least two bacterial genera comprise at least two genera selected from the list consisting of *Sporobacter*, *Coprococcus*, *Fusicatenibacter, Ruminococcus*,

Oscillibacter and *Faecalibacterium* and/or from the list consisting of *Clostridium_IV*, *Butyricimonas*, *Butyricicoccus*, *Prevotella*, *Roseburia, Flavonifractor* and *Bacteroides*. In a preferred embodiment, said method comprises at least two genera selected from the list consisting of *Bacteroides*, *Fecalibacterium, Ruminococcus*, *Oscillibacter*, and *Prevotella*. In a particularly preferred embodiment, said at least two bacterial genera comprise *Oscillibacter and Faecalibacterium*. In a particular embodiment, said at least two bacterial genera do not consist of *Faecalibacterium* and *Bacteroides*. The reference abundances, referred to in the method of the invention, are obtained from a plurality of control samples. Said control samples, can be, for example samples obtainable from the healthy subjects not having a gut dysbiosis, for example having an enterotype other than B2 enterotype. In a particular embodiment, said control samples are samples from non-disease or healthy subjects or from subject not suffering from gut flora dysbiosis. In a preferred embodiment, a subject will be diagnosed with gut flora dysbiosis if each of the measured abundances of each of the selected bacterial genera are decreased compared to the reference abundances, except for *Bacteroides* and *Flavonifractor*, where an increased abundance compared to the reference abundances is predictive for gut flora dysbiosis. It was surprisingly found that by selection, detection and quantification of selected bacterial genera in a biological sample obtained from a subject, the gut dysbiosis can be detected without an extensive sequencing and genome analyses of the complex gut microbiome data. By relying on equipment and expertise which is readily available in many labs and, eliminating the requirement for 16S Amplicon Sequencing and downstream bioinformatics analyses, results can be obtained in a fraction of the time needed using the known methods. The method according to the first aspect of the present invention can be suitable for wide applicability in diagnostic labs, can be designed to minimise the cost of such testing and dysbiosis detection. The method according to the first aspect is robust and reliable, meaning that it shows consistency and minimal influence of the sample preparation.

In a preferred embodiment, the method of the invention comprises the step of measuring of abundances of at least genera *Oscillibacter and Faecalibacterium*. In a particularly preferred embodiment, said step of measuring the abundances comprises measuring of at least one further genus selected from a list consisting of *Ruminococcus*, *Prevotella*, and *Bacteroides*. In another particularly preferred embodiment, said step of measuring the abundances comprises measuring of at least two, or at least three further genera selected from a list consisting of *Ruminococcus*, *Prevotella*, and *Bacteroides*. By measuring abundances of said selected genera, dysbiosis or B2 enterotype could be detected by analysis of complex biological samples without investigation of full microbiome composition, which can be costly, lengthy and/or cumbersome.

In another embodiment, the method of the invention allows for association of the gut flora dysbiosis with an inflammatory disorder, obesity, diabetes type 2, depression or anxiety. In a particular

embodiment, the inflammatory disorder is selected from the list consisting of spondyloarthritis, ankylosing spondylitis, reactive arthritis, psoriatic arthritis, enteropathic arthritis, undifferentiated spondyloarthritis, juvenile idiopathic arthritis, primary sclerosing cholangitis, multiple sclerosis and any gut inflammation associated therewith. Alternatively, the inflammatory disorder is a gut inflammatory disorder selected from the list consisting of Crohn's disease, irritable bowel syndrome, inflammatory bowel disease, ulcerative colitis and celiac disease. The method of the invention can be used to detect the gut dysbiosis and systemic disorders associated with the dysbiosis, for example, associated with B2 enterotype, by a reliable and lab-applicable method which is of a reduced cost compared to the whole genome sequencing, and takes less time.

In a particular embodiment, said biological sample is selected from the list consisting of a stool sample, a mucosal biopsy sample and a sample of the lumen content. The method of the invention allows for detecting gut flora dysbiosis in a subject using a non-invasive sampling method, for example, allowing to detect said dysbiosis associated with an inflammatory disorder, obesity, diabetes type 2, depression, and/or anxiety by using samples such as stool or lumen content.

In another embodiment, the quantification of the at least two, at least three, at least four or at least five bacterial genera is described above is performed by quantifying DNA sequences specific for said at least two, at least three, at least four or at least five bacterial genera. In a particular embodiment, said DNA sequences are 16S rDNA sequences. In another embodiment, the quantification of the DNA sequences is performed by quantitative Polymerase chain reaction (qPCR), wherein in a more particular embodiment the abundances of the at least two bacterial genera are expressed as log cells/g. Alternatively, the abundances may be expressed as copies/µL or the Cycle threshold (Ct) value. In even more particular embodiments, said 16S rDNA sequences are selected from SEQ ID No. 1-68. In another particular embodiment, the methods include a step of DNA extraction from the biological sample, prior to the quantification of said DNA sequences. In a preferred embodiment, a quantification of said genera according to the first aspect of the invention is done by the qPCR, as said qPCR quantification of the bacterial genera in the method of the invention allows for higher speed and better cost-effectiveness of gut flora dysbiosis detection.

In a second aspect, a bacterial genera panel is provided comprising at least two bacterial genera selected from the list consisting of *Oscillibacter*, *Faecalibacterium*, *Ruminococcus*, *Prevotella* and *Bacteroides*, wherein the panel does not consist of *Faecalibacterium* and *Bacteroides*. Preferably, said bacterial genera panel comprises at least genera *Oscillibacter* and *Faecalibacterium*. More preferably, said bacterial genera panel comprises at least a further genus selected from the group consisting of *Ruminococcus*, *Prevotella* and *Bacteroides*. In one embodiment, the bacterial genera panel is provided

for use to diagnose a subject with gut flora dysbiosis, more particularly for determining gut flora dysbiosis in a subject according to any of the methods herein disclosed. The bacterial genera panel of the invention, herein further referred to as bacterial panel, allows for simplified and low-cost detection and/or diagnosis of the gut flora dysbiosis in a biological sample obtained from a subject. The bacterial panel could be used, for example for simplified and low-cost detection and/or diagnosis of the gut flora dysbiosis is associated with an inflammatory disorder, obesity, diabetes type 2, depression, and/or anxiety.

In a third aspect, a kit is provided comprising

i) at least two pairs of nucleic acid primers designed for specifically amplifying DNA sequences of at least two different bacterial genera, wherein the bacterial genera are selected from a list consisting of *Ruminococcus*, *Oscillibacter*, *Faecalibacterium, Prevotella* and *Bacteroides*; and

ii) preferably instructions for quantifying the levels of two or more DNA sequences from a stool sample.

In one embodiment, the kit comprises at least five pairs of nucleic acid primers capable of specifically amplifying DNA sequences of at least five different bacterial genera selected from the list consisting of *Ruminococcus*, *Oscillibacter*, *Faecalibacterium*, *Prevotella*, and *Bacteroides*. In another embodiment, said kit is provided wherein the DNA sequences to be amplifies are 16S rDNA sequences or fragments thereof, more particularly the 16S rDNA sequences are selected from SEQ ID No. 1-68. Also provided is said kit for diagnosing or early detecting gut flora dysbiosis.

Brief description of the Figures

Figure 1 illustrates the performance obtained using an optimized random forest classifier on all 13 selected genera. (a) The area under the curve here is 0.97 which is an outstanding performance for a classifier. The confusion matrix (b) shows where the mistakes the model makes are located.

Figure 2 shows the SHAP output to assess feature importance within the full model. Light gray dots are samples with a high content of the specific genus, dark gray indicates a low abundance. If dots are on the left of the vertical gray line, they are negatively associated with B2, on the right they indicate a positive correlation. The SHAP value indicates the impact on the model, and genera are ranked from top to bottom for high importance to low.

Figure 3 shows the B2 scores for all FGFP samples per enterotype. As intended the score is high for B2 samples while low for samples from other enterotypes. A Kruskal-Wallis test was used to determine there were differences (p-value = 2.27e-169, followed by a Posthoc Dunn test (with FDR correction) to determine which differences were significant. As shown here, sample assigned to the B2 enterotype get

significantly higher scores than those from other enterotypes. Therefore, the goal of the analysis has been obtained. (ns: not significant, *: 1.00e-2 < p <= 5.00e-2, **: 1.00e-3 < p <= 1.00e-2, ***: 1.00e-4 < p <= 1.00e-3, ****: p <= 1.00e-4)

Figure 4 shows relative abundances of selected genera used to detect the gut dysbiosis according the method of the invention. The abundances shown are obtained using qPCR indicating the differences between the enterotypes. These closely match results observed in 16S studies with e.g. high *Bacteroides*/*Faecalibacterium* ratio in B2, and lower abundances of genera such as *Oscillibacter, Faecalibacterium, Ruminococcus*, and *Prevotella* with respect to the corresponding abundances of said genera detected in biological samples related to other enterotypes.

Figure 5 shows Receiver Operator Curve (ROC) showing the Area Under the Curve (ROC AUC) as a metric for classification performance of B2 versus non-B2 enterotypes with the qPCR model results.

Figure 6 shows confusion matrix for binary predictions. Normalized for predictions the model outputs, here if the model predicts B2 that is the correct label in 93% of the predictions.

Detailed description

Definitions

The present invention will be described with respect to particular embodiments and with reference to certain drawings but the invention is not limited thereto but only by the claims. Any reference signs in the claims shall not be construed as limiting the scope. The drawings described are only schematic and are non-limiting. In the drawings, the size of some of the elements may be exaggerated and not drawn on scale for illustrative purposes. Where the term "comprising" is used in the present description and claims, it does not exclude other elements or steps. Where an indefinite or definite article is used when referring to a singular noun e.g. "a" or "an", "the", this includes a plural of that noun unless something else is specifically stated.

Furthermore, the terms first, second, third and the like in the description and in the claims, are used for distinguishing between similar elements and not necessarily for describing a sequential or chronological order. It is to be understood that the terms so used are interchangeable under appropriate circumstances and that the embodiments of the invention described herein are capable of operation in other sequences than described or illustrated herein.

The following terms or definitions are provided solely to aid in the understanding of the invention. Unless specifically defined herein, all terms used herein have the same meaning as they would to one skilled in

the art of the present invention. Practitioners are particularly directed to Michael R. Green and Joseph Sambrook, Molecular Cloning: A Laboratory Manual, 4th ed., Cold Spring Harbor Laboratory Press, Plainsview, New York (2012); and Ausubel et al., Current Protocols in Molecular Biology (Supplement 47), John Wiley & Sons, New York (1999), for definitions and terms of the art. The definitions provided herein should not be construed to have a scope less than understood by a person of ordinary skill in the art.

The term "gut microbiome composition" is equivalent in wording as "gut microbiome configuration" or "gut microbiome profile" and these wordings are used interchangeably herein. A gut microbiome profile represents the presence, absence or the abundance of one or more of bacterial genera identified in a stool sample. The gut microbiome profile can be determined based on an analysis of amplification products of DNA and/or RNA of the gut microbiota, e.g. based on an analysis of amplification products of genes coding for one or more of small subunit rRNA, etc. and/or based on an analysis of proteins and/or metabolic products present in the biological sample. Gut microbiome profiles may be "compared" by any of a variety of statistical analytic procedures.

In microbiology, "16S sequencing" or "16S" refers to a sequence derived by characterizing the nucleotides that comprise the 16S ribosomal RNA gene(s). The bacterial 16S rRNA is approximately 1500 nucleotides in length and is used in reconstructing the evolutionary relationships and sequence similarity of one bacterial isolate to another using phylogenetic approaches.

As used herein, "abundance" and "level" of a genus X are equivalents. The term "abundance" as used herein refers to a measure of the quantity of a target microorganism within a biological sample. It can also be referred to as "load". In one embodiment, the "abundance of genus X" is a relative abundance of genus X with respect to or compared to a plurality of other genera present in the same sample. Hence, a "higher abundance of genus X" or "increased abundance of genus X" then means that the relative abundance of genus X within a test sample is higher or increased compared to the relative abundance of genus X in a control sample or compared to relative reference abundance or abundances of genus X obtained from one or more control samples. Similarly, a "lower abundance of genus X" or "decreased abundance of genus X" then means that the relative abundance of genus X within a test sample is lower or decreased compared to the relative abundance of genus X in a control sample or compared to relative reference abundance or abundances of genus X obtained from one or more control samples.

In another embodiment, the "abundance of genus X" is an absolute abundance of genus X within a sample. The absolute abundance of a genus can be quantitatively measured in a sample for example by adding or spiking a precisely know amount of control DNA or by measuring the amount of bacterial cells (i.e. cell count) present in the sample. By combining relative abundances with cell counts, one or more genera can be absolutely quantified.

Hence, a "higher abundance of genus X" or "increased abundance of genus X" then means that the absolute abundance of genus X within a test sample is higher or increased compared to the absolute abundance of genus X in a control sample or compared to absolute reference abundance or abundances of genus X obtained from one or more control samples. Similarly, a "lower abundance of genus X" or "decreased abundance of genus X" then means that the absolute abundance of genus X within a test sample is lower or decreased compared to the absolute abundance of genus X in a control sample or compared to absolute reference abundance or abundances of genus X obtained from one or more control samples.

"Higher" or "increased" as used herein is defined herein as a statistically significantly increased, more particularly an at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 1.5 fold, at least 2 fold, at least 3 fold, at least 5 fold or at least 10 fold higher or increased abundance compared to the abundance of the genus X in the stool sample of a healthy subject or compared to one or a plurality of references abundances of genus X.

"Lower" or "decreased" as used herein is defined herein as a statistically significantly decreased, more particularly an at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 1.5 fold, at least 2 fold, at least 3 fold, at least 5 fold or at least 10 fold lower or decreased abundance compared to the abundance of the genus X in the stool sample of a healthy subject or compared to one or a plurality of references abundances of genus X.

The term "statistically significant" or "statistically significantly" different is well known by the person skilled in the art. Statistical significance plays a pivotal role in statistical hypothesis testing. It is used to determine whether the null hypothesis should be rejected or retained. It states that the results are obtained because of chance and are not supporting a real change or difference between two data sets. The null hypothesis is the default assumption that what one is trying to prove did not happen. In contrast the alternative hypotheses states that the obtained results support the theory being investigated. For the null hypothesis to be rejected (and thus the alternative hypothesis to be accepted), an observed result has to be statistically significant, i.e. the observed *p*-value is less than the pre-specified significance level α. The *p* stands for probability and measures how likely it is that the null hypothesis is incorrectly rejected and thus that any observed difference between data sets is purely due to chance. In most cases the significance level α is set at 0.05.

The term "Highest Density Interval" or "HDI" is well known by the person skilled in the art. When analysing distributions, the HDI can be used to determine the smallest interval which contains a desired

fraction of all values. Paired with probabilistic programming, where variables in a model are described by probability density functions rather than point estimates, the HDI can be used to summarize those densities into credible intervals which can easily be interpreted as the interval which contains the true value with a certain probability. The HDI is typically determined for 94% or 95% of the values.

A "biological sample" as used herein refers to a stool sample, a mucosal biopsy sample or a sample of the lumen content. In a most particular embodiment, said biological sample is a stool or faecal sample. "Stool sample" and "faecal sample" are used interchangeably and refer to as a sample or aliquot of the stool or faeces of a subject, more particularly a mammal, even more particularly a human being, most particularly a patient. The stool sample as used herein comprises the gut microbiome from a human patient to be diagnosed. As used herein, the term "microflora" refers to the collective bacteria in an ecosystem of a host (e.g. an animal, such as a human) or in a single part of the host's body, e.g. the gut. An equivalent term is "microbiota". As used herein, the term "microbiome" refers to the totality of bacteria, their genetic elements (genomes) in a defined environment, e.g. within the gut of a host, the latter then being referred to as the "gut microbiome".

As used herein, the term "patient" or "individual" or "subject" typically denotes humans, but may also encompass reference to non-human animals, preferably warm-blooded animals, more preferably mammals, such as, e.g. non-human primates, rodents, canines, felines, equines, ovines, porcines, and the like.

As used herein, the term "gut" generally comprises the stomach, the colon, the small intestine, the large intestine, cecum and the rectum. In addition, regions of the gut may be subdivided, e.g. the right versus the left side of the colon may have different microflora populations due to the time required for digesting material to move through the colon, and changes in its composition in time. Synonyms of gut include the "gastrointestinal tract", "intestinal tract" or possibly the "digestive system", although the latter is generally also understood to comprise the mouth, oesophagus, etc.

Regarding the current application, "method to detect gut flora dysbiosis" is equivalent to a "method of detecting the presence or of assessing the risk of developing gut flora dysbiosis".

In one embodiment, gut flora dysbiosis is associated with an inflammatory disorder or gut flora dysbiosis is equivalent to a gut inflammatory disorder. Hence, the application also provides a "method of detecting an inflammatory disorder" which is equivalent to a "method to detect the presence or to assess the risk of developing an inflammatory disease". In another embodiment of the invention, a "method to detect gut flora dysbiosis" is equivalent to a "method to detect and alert for a suboptimal lifestyle leading to gut flora dysbiosis and/or for suboptimal dietary choices leading to gut flora dysbiosis".

The term "control sample" or "plurality of control samples", as used herein, means a group of biological samples of healthy, non-b2 or non-dysbiotic subjects, and should be understood as a healthy control

group which consists of participants who do not have the gut dysbiosis. The detected levels of selected bacterial genera and their mean values detected in said "plurality of control samples group" serve as a the reference abundance, for example a baseline to compare the bacterial abundances with respect to each other with the aim to detect dysbiotic, i.e., B2 or non-dysbiotic enterotype.

A gut dysbiotic group of samples refers to the samples of participants who have an imbalance in their gut microbiome, known as dysbiosis. The term "dysbiosis", as used herein, occurs when there is a disruption in the normal balance of microorganisms in the gut, which can lead to various health issues.

The term "inflammation", "inflammatory disorder" or "inflammatory disease" refers to complex – but to the skilled person well known – biological response of body tissues to harmful stimuli, such as pathogens, damaged cells, or irritants. Inflammation is not a synonym for infection though. Infection describes the interaction between the action of microbial invasion and the reaction of the body's inflammatory response — the two components are considered together when discussing an infection, and the word is used to imply a microbial invasive cause for the observed inflammatory reaction. Inflammation on the other hand describes purely the body's immunovascular response, whatever the cause may be. Inflammation is a protective response involving immune cells, blood vessels and molecular mediators. The function of inflammation is to eliminate the initial cause of cell injury, clear out necrotic cells and tissues damaged from the original insult and the inflammatory process, and to initiate tissue repair. The classical signs of inflammation are heat, pain, redness, swelling, and loss of function. Inflammation is a generic response, and therefore it is considered as a mechanism of innate immunity, as compared to adaptive immunity, which is specific for each pathogen. Inflammation can be classified as either acute or chronic. Acute inflammation is the initial response of the body to harmful stimuli and is achieved by the increased movement of plasma and leukocytes (especially granulocytes) from the blood into the injured tissues. A series of biochemical events propagates and matures the inflammatory response, involving the local vascular system, the immune system, and various cells within the injured tissue. Prolonged inflammation, known as chronic inflammation, leads to a progressive shift in the type of cells present at the site of inflammation, such as mononuclear cells, and is characterized by simultaneous destruction and healing of the tissue from the inflammatory process.

In a particular embodiment of the whole application, "inflammatory disorder" as used herein is selected from the list consisting of spondyloarthritis, ankylosing spondylitis, reactive arthritis, psoriatic arthritis, enteropathic arthritis, undifferentiated spondyloarthritis, juvenile idiopathic arthritis, primary sclerosing cholangitis, multiple sclerosis, a gut inflammatory disorder, inflammatory bowel disease (IBD), Crohn's disease (CD), ulcerative colitis (UC), irritable bowel syndrome (IBS), celiac disease and any combination thereof and any gut inflammation associated with one of the above listed inflammatory disorders. In

another particular embodiment, said inflammatory disorder is characterized by a TH1, TH17, TH2 and/or TH9 response.

The term "ROC" or Receiver Operating Characteristic curve refers to a graphical plot that illustrates the diagnostic ability of a binary classifier system or alternatively phrased a probability curve and includes all the possible decision thresholds from a diagnostic test result. Alternatively defined, an ROC curve is a plot of the sensitivity (or True Positive Rate) on the y-axis versus 1 - specificity (or False Positive Rate) on the x-axis of a diagnostic test. The different points on the curve correspond to the different cut points used to determine whether the test results are positive. A ROC curve can be considered as the average value of the sensitivity for a test over all possible values of specificity or vice versa (Mandrekar 2010 J Thorac Oncol 5: 1315-1316). The ROC curve is a well-known tool to the skilled person in the assessment of the performance of a diagnostic test.

The "area under the ROC curve" (often referred to as simply the "AUC") is an effective way to summarize the overall diagnostic accuracy of the test. It refers then to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. It thus tells how much the model is capable of distinguishing between classes. The AUC takes values from 0 to 1, where a value of 0 indicates a perfectly inaccurate test and a value of 1 reflects a perfectly accurate test (Mandrekar 2010 J Thorac Oncol 5: 1315-1316).

The term "biomarker" as used herein refers to an indicator of a phenotype of a patient, e.g., a pathological state or likely responsiveness to a therapeutic agent, which can be detected in a biological sample obtained from the patient. Biomarkers include, but are not limited to microorganisms, bacteria, DNA, RNA, protein, carbohydrate, or glycolipid-based molecular markers.

The term "diagnosis" is used herein to refer to the identification or classification of a molecular or pathological state, disease or condition. For example, "diagnosis" may refer to identification of a particular type of IBD, e.g., UC or Crohn's disease. "Diagnosis" may also refer to the classification of a particular subtype of IBD, e.g., by histopathological criteria or by molecular features (e.g., a subtype characterized by expression of one or a combination of particular genes or proteins encoded by said genes).

The term "kit", as used herein, relates to a set articles or equipment needed for a purpose of detecting a gut dysbiosis in a subject. A non-limiting example is a qPCR (quantitative Polymerase Chain Reaction) kit. Said qPCR kit is arranged for amplifying and simultaneously quantifying a targeted DNA molecule. The qPCR as a technique, is also known as real-time PCR, and allows to measure the amount of DNA in a sample in real-time as the PCR process progresses. Said qPCR kit can comprise at least one or more of the following components: i) *DNA Polymerase*: An enzyme that synthesizes new DNA strands from the template DNA; *ii) primers*: Short DNA sequences that initiate the DNA synthesis; *iii) probes or dyes*:

Fluorescent molecules that bind to the DNA and emit a signal that can be measured; *iv) buffer solution*: Maintains the optimal conditions for the PCR reaction; v) *template DNA*: The DNA sample that contains the target sequence to be amplified.

The bacterial genera as used in embodiments of all aspects of the present invention, are defined herein based on the LPSN (List of Prokaryotic names with Standing in Nomenclature) website https://lpsn.dsmz.de/, as available and/or in force on 9 December 2024. All the species validly published under the ICNP (International Code of Nomenclature of Prokaryotes) with full names in each genus are included in the lists provided below, with an example 16S rRNA DNA sequence, via the reference number of GenBank NCBI for their DNA sequence, to provide an unambiguous definition. Each one of the following lists are extensive, i.e. the species listed herein are limiting.

The *Ruminococcus* genus is defined with 10 species: *Ruminococcus albus* (AB538438), *R. bovis* (NR_180520), *R. bromii* (NR_025930), *R. callidus* (NR_029160), *R. champanellensis* (AB910742), *R. flavefaciens* (KP689131), *R. gauvreauii* (EF529620), *R. lactaris* (NR_027579), *R. torques* (AB910746) and *R. turbiniformis* (OK510360).

The *Bacteroides* genus is defined with 52 species: *Bacteroides acidifaciens (EU136694), B. caccae (EU136686), B. caecicola (AB910337), B. caecigallinarum (AB861981), B. caecimuris (KR364741), B. cellulolyticus (NZ_JAOQJJ010000028 region 30677-32204), B. cellulosilyticus (AB510698), B. clarus (AB547638), B. coprosuis (AF319778), B. eggerthii (EU136695), B. facilis (MT905157), B. faecalis (MK207058), B. faecichinchillae (AB574480), B. faecis (AB547640), B. faecium (MN853409), B. finegoldii (AB222699), B. fluxus (AB547642), B. fragilis (KP326374), B. galacturonicus (DQ497994), B. gallinaceus (AB910339), B. gallinarum (AB253732), B. graminisolvens (AB547643), B. helcogenes (EU136692), B. heparinolyticus (L16487), B. hominis (OK236009), B. humanifaecis (MW228414), B. intestinalis (AB214328), B. koreensis (KX025133), B. kribbi (KX025134), B. luhongzhouii (MK584158), B. luti (AB787271), B. muris (ON325392), B. nordii (EU136693), B. oleiciplenus (AB547644), B. ovatus (AB510705), B. parvus (JACOOG010000001 region 258551-260084), B. pectinophilus (ABVQ01000036), B. polypragmatus (no sequence found), B. propionicifaciens (AB510706), B. propionicigenes (MZ127654), B. pyogenes (EU136683), B. reticulotermitis (AB692943), B. rhinocerotis (OP931997), B. rodentium (AB531489), B. salyersiae (EU136690), B. stercorirosoris (AB574481), B. stercoris (EU136684), B. thetaiotaomicron (M58763), B. uniformis (EU136680), B. xylanisolvens (AB510713), B. zhangwenhongii (MK583535)* and *B. zoogleoformans (L16488).*

Although the *Prevotella* genus was reorganised (split up) recently and partially renamed, these names such as *Segatella* are currently synonyms. As such, the *Prevotella* genus is defined with 60 species: *Prevotella albensis (AJ011683), P. amnii (AB547670), P. aurantiaca (AB480534), P. baroniae (AB547671), P. bergensis (AB547672), P. bivia (AB547673), P. brevis (AJ011682), P. brunnea (MK424351), P. bryantii*

*(AJ006457), P. buccae (AB547675), P. buccalis (AB547676), P. cerevisiae (LC331291), P. colorans (KT886066), P. communis (CP091792), P. copri (AB649279), P. corporis (AB547677), P. dentalis (AB547678), P. dentasini (AB477014), P. denticola (AB547679), P. disiens (AB547682), P. enoeca (AB547684), P. falsenii (AB429504), P. fusca (AB683252), P. herbatica (AB298732), P. histicola (AB547685), P. hominis (MK567960), P. illustrans MT675969), P. intermedia (AB547686), P. jejuni (JQ778983), P. lacticifex (LC639954), P. lascolaii (LN998055), P. loescheii (AB547688), P. maculosa (EF534314), P. marshii (AB547691), P. melaninogenica (AB547693), P. merdae (LT984637), P. micans (AB547692), P. mizrahii (*MN537546*), P. multiformis (AB182483), P. multisaccharivorax (AB200414), P. nanceiensis (AB547695), P. nigrescens (AB547697), P. oralis (AB547698), P. oris (AB547700), P. oryzae (AB588018), P. oulorum (AB547702), P. pallens (AB547703), P. paludivivens (AB547704), P. phocaeensis (LN998069), P. pleuritidis (AB278593), P. rara (KX198132), P. ruminicola (L16482), P. saccharolytica (AB616135), P. salivae (AB108826), P. scopos (AB683253), P. shahii (AB108825), P. stercorea (AB244774), P. timonensis (AB547706), P. veroralis (AB547707)* and *P. vespertine (MK424350).*

The genus Faecalibacterium is defined with 7 species: *Faecalibacterium butyricigenerans (KX146426), F. duncaniae (AJ270469), F. gallinarum (DQ057476), F. hattorii (LC670256), F. longum (KX150462), F. prausnitzii (AJ413954),* and *F. taiwanense (OQ749744).*

The *Oscillibacter* genus is defined with 4 species: *Oscillibacter acetigenes (*NZ_JAOQKG010000009.1 region 66310-67840), *O. hominis (MT905210), O. ruminantium (JF750939)* and *O. valericigenes (AB238598).*

The starting point of the design of the method has been substantiated by the RDP 16 database. Although DNA sequences remain the same, the bacterial taxonomy and nomenclature should be understood as a dynamic field with regular changes. However, above -mentioned lists of each and every genera used in the method, panel and kit of the invention are extensive as on date 9 December 2024, thus the scope of each genus is unambiguously defined.

A bacterial profile predicting the B2 gut enterotype

The human gut is the natural habitat for a large and dynamic bacterial community. These human digestive-tract associated microbes are referred to as the gut microbiome. The human gut microbiome and its role in both health and disease has been the subject of extensive research (Fan and Pedersen 2020 Nat Rev Microbiol 19). The human gut microbiome can be classified in genera-driven clusters or enterotypes (Arumugam et al 2011 Nature 473:174-180; Falony et al 2016 Science 352:560-564), i.e. the Prevotella, Ruminococcus, Bacteroides_1 and Bacteroides_2 (herein referred to as Bact2 or B2) enterotype. The latter is a recently described intestinal microbiota configuration embodying gut flora dysbiosis. It has also been demonstrated that Bacteroides_2 is associated with systemic inflammation, inflammatory bowel disease, primary sclerosing cholangitis, obesity, depression, multiple sclerosis and

has a high prevalence in loose stools in humans (Vandeputte et al 2017 Nature 551: 507-511; Valles-Colomer et al 2019 Nat Microbiol 4: 623-632; Veira-Silva et al 2019 Nat Microbiol 4: 1826-1831; Veira-Silva et al 2020 Nature 581: 310-315; Reynders et al 2020 Ann Clin Transl Neur 7: 406-419). B2 is characterized by a high proportion of *Bacteroides*, a low proportion of *Faecalibacterium* and low microbial cell densities (WO2019/115755A1). Its prevalence varies from 13% in a general population cohort to as high as 78% in patients with inflammatory bowel disease.

As discussed earlier, the prior-art enterotyping techniques, including diagnosing the dysbiotic B2 enterotype solely relies on 16S rRNA sequencing of bacteria obtained from a stool sample. In order to develop a simple enterotyping assay, the present invention utilizes, among others, machine learning models which could predict the dysbiotic B2 enterotype based on the presence or abundance of specific bacterial genera obtained from the biological sample, preferably of a stool sample. The present invention also aims to provide an reliable, fast and cost-saving enterotyping method which method would be applicable in a diagnostic lab, and preferably would include qPCR diagnostic method.

From the moment samples and a labelled dataset, of which the enterotype is known, can be accessed, a classifier can be trained on (a subset of) the measured bacterial genera. Any classifier that can predict a class label from one or more continuous features can be used, this includes, but isn't limited to Decision Trees, Random Forest Classifiers, Support Vector Classifiers, Stochastic Gradient Decent Classifier and the ADABoostClassifier (Smola and Schölkopf 2004 Statistics and Computing 14; Schölkopf et al 2000 Neural Computation 12; Crammer and Singer 2001 Journal of Machine Learning Research 2; Bottou et al 2018 SIAM Review 60; Freund and Schapire 1999 Journal of Japanese Society for Artificial Intelligence 14). Various implementations for these classifiers are available in Scikit-learn (for the Python language) (Pedregosa et al 2011 J Machine Learning Research 12), Machine Learning for R (mlr library for R) (Bischl et al 2016 J Machine Learning Research 17). Once trained on a labelled set, abundances of bacterial genera from patients' samples with an unknown enterotype can be provided to the trained classifier to obtain a predicted class (in the context of current application this is B2 or non-B2).

In current application a set of 13 bacterial genera is described, selected using state-of-the art machine learning, which when measured – alone or in combination – in human faecal material, can be used to predict the dysbiotic B2 enterotype. The abundances of one or more genera can for example be measured using standard PCR techniques.

Surprisingly, some of the selected thirteen bacterial genera, when measuring their abundances alone or particular combinations, as were shown as reliable predictors of B2 enterotype and/or dysbiosis, when using qPCR technique. Preferably, if said qPCR technique is used to detect the abundances of bacterial genera, at least two of the genera selected from the group consisting of *Oscillibacter*, *Faecalibacterium*,

Ruminococcus, Prevotella and *Bacteroides* can be considered as a reliable indicative of the dysbiosis in a subject.

In the methods described herein, the determination step is based on an increased or decreased abundance of each of at least 2 bacterial genera selected from the list of 13 bacterial genera consisting of *Sporobacter*, *Coprococcus*, *Fusicatenibacter*, *Ruminococcus, Oscillibacter, Faecalibacterium, Clostridium_IV, Butyricimonas, Butyricicoccus, Prevotella, Roseburia, Flavonifractor* and *Bacteroides*, in the subject sample compared to those in the control sample or to reference abundances. If *Faecalibacterium, Ruminococcus, Fusicatenibacter, Sporobacter, Oscillibacter, Coprococus, Clostridium_IV, Butyricimonas, Prevotella, Butyricicoccus* or *Roseburia* is measured as one of the at least 2 bacterial genera then a decreased level is predictive for the dysbiotic B2 enterotype (and thus, for example, for gut flora dysbiosis and/or B2-associated disease or disorder such as inflammatory disorders, obesity, diabetes type 2 and depression) or for a gut microbiome associated with or predictive for B2-associated disease or disorder.

If *Flavonifractor* and/or *Bacteroides* is measured as one of the at least two bacterial genera then an increased level is predictive for the dysbiotic B2 enterotype or for a gut microbiome associated with or predictive for B2-associated disease or disorder.

In a particular embodiment of the whole application, differences in the measured abundance or level of each of the at least two bacterial genera between the subject sample and each of those of the control sample or each of the reference abundances are statistically significant.

By way of example, if *Faeacalibacterium* and *Ruminococcus* would be selected as the bacterial biomarkers, then a test subject has a high probability of having a B2 gut enterotype, if the abundance of *Faecalibacterium* in a stool sample from the test subject is decreased compared to the abundance of *Faecalibacterium* in a stool sample from a healthy control and the abundance of *Ruminococcus* in the stool sample from the test subject is decreased compared to the abundance of *Ruminococcus* in the stool sample from the healthy control.

In a preferred embodiment, the method of the invention comprises step of measuring the abundances of at least genera *Oscillibacter* and *Faecalibacterium*. The relative abundances of both genera *Oscillibacter* and *Faecalibacterium* are decreased in the stool sample of the subject having dysbiosis compared to a reference value detected in stool samples from the healthy control group.

Preferably said step of measuring the abundances includes measuring of at least one further genus, preferably at least two further genera selected from a list consisting of *Ruminococcus, Prevotella*, and *Bacteroides*. The relative abundances of genera *Ruminococcus* and *Prevotella* are decreased while the

relative abundance of genus *Bacteroides* is increased compared to the reference value obtained by measuring the abundance of said genera in a plurality of control samples.

In another embodiment, the method of the invention comprises step of measuring the abundances of five bacterial genera selected from the list consisting of *Oscillibacter*, *Faecalibacterium, Ruminococcus, Prevotella* and *Bacteroides*.

In an alternative embodiment, the method of the invention comprises step of measuring the abundances of at least two bacterial genera comprising *Oscillibacter* and *Bacteroides*.

The bacterial genera disclosed herein are define according to the taxonomy based on the 11.5 release of the RDP database.

In some embodiments, the bacterial genera used in the method of the invention could be defined by some non-limiting, exemplary 16S rDNA sequences of selected genera, said some non-limiting examples being shown in Table 1.

"Bacteroides" as used herein refers to a genus of Gram-negative, obligate anaerobic bacteria. Bacteroides species are normally mutualistic, making up the most substantial portion of the mammalian gastrointestinal flora. The *Bacteroides* genus belongs to the family of Bacteroidaceae and a non-limiting example of a *Bacteroides* species is *B. fragilis*. Non-limiting examples of 16S rDNA sequences of the *Bacteriodes* genus are SEQ ID No. 1-14.

"Butyricicoccus" as used herein refers to a genus of spherical, gram-positive, non-motile bacteria, which includes but isn't limited to *B. desmolans*, *B. faecihominis*, *B. porcorum* and, *B. pullicaecorum* (Eeckhaut et al 2016 Front Microbiol 7). Isolates from different animals are known to produce butyrates. Non-limiting examples of 16S rDNA sequences of the *Butyricicoccus* genus are SEQ ID No. 15-17.

"Butyricimonas" as used herein refers to a genus of gram-negative, anaerobic bacteria found in the gastrointestinal track of humans (and other mammals) (Sakamoto et al 2009 Int J Syst Evol Microbiol 59). Which includes, but isn't limited to, species *B. faecalis*, *B. faecihominis*, *B. paravirosa*, *B. synergistica* and, *B. virosa*. Various isolates are known to produce butyrates. Non-limiting examples of 16S rDNA sequences of the *Butyricimonas* genus are SEQ ID No. 18-20.

"Clostridium_IV" also known as "Clostridium cluster IV" as used herein refers to a genus of rod-shaped, gram-positive, anaerobic, spore-forming bacteria that are known to inhabit the gut (Guo et al 2020 J An Sc Biotechnol 11; Zweilehner et al 2009 Experimental Gerontology 44). As many Clostridia can process a wide variety of nutrients, while producing metabolites with potential health benefits these species are considered beneficial for the gut flora of mammals (incl. humans). Non-limiting examples of 16S rDNA sequences of the Clostridium_IV genus are SEQ ID No. 21-27.

"Coprococcus" as used herein refers to a genus of spherical (cocci), gram-positive, anaerobic bacteria which contains several characterized species (e.g. *Coprococcus catus*, *Coprococcus comes*, *Coprococcus*

eutactus) as well as various unclassified species. The genus comprises known residents of the human gut which have the capacity to ferment carbohydrates and produce butyric acid, an anti-inflammatory compound. Non-limiting examples of 16S rDNA sequences of the *Coprococcus* genus are SEQ ID No. 28-32.

"Faecalibacterium" as used herein refers to a genus of bacteria of which its sole known species, *Faecalibacterium prausnitzii* is gram-positive, mesophilic, rod-shaped, anaerobic and is one of the most abundant and important commensal bacteria of the human gut microbiota. It is non-spore forming and non-motile. These *Faecalibacterium* bacteria produce butyrate and other short-chain fatty acids through the fermentation of dietary fiber. Non-limiting examples of 16S rDNA sequences of the *Faecalibacterium* genus are SEQ ID No. 33-38.

"Flavonifractor" as used herein refers to a genus of anaerobic, gram-positive, rod-shaped bacteria (Berger et al 2018 IDCases 14). *Flavonifractor* species, including but not limited to *Flavonifractor plautii*, are common inhabitants of the human gut (Rodriguez-Castaño et al 2020 PLoS One 15) and produce butyrate although *F. plautii* can also break down flavonoids e.g. quercetin. Non-limiting examples of 16S rDNA sequences of the *Flavonifractor* genus are SEQ ID No. 39-40.

"Fusicatenibacter" as used herein refers to a genus of gram-positive, obligately anaerobic, non-motile, non-spore-forming, spindle-shaped bacteria originally isolated from human faeces. From this genus belonging to the family Lachnospiraceae a single species has been described, *Fusicatenibacter saccharivorans*, though there are several additional species proposed (*Fusicatenibacter intestinigallinarum*, *Fusicatenibacter intestinipullorum* and, *Fusicatenibacter merdavium*) which currently have the candidatus status indicating they are well characterized, but to date not cultured. Non-limiting examples of 16S rDNA sequences of the *Fusicatenibacter* genus are SEQ ID No. 41-42.

"Oscillibacter" as used herein refers to a genus of motile (flagellated), gram-negative, anaerobic bacteria with a slightly curved rod-shape. The type-strain for the genus *Oscillibacter valericigenes* was originally isolated from the digestive track of Japanese Corbicula clams, though has been detected as an inhabitant in the human gut since. Furthermore, "Oscillibacter" has been found in the human gut and has been linked with diet (Rosés et al 2021 Nutrients 13) and disease status (Metwaly et al 2020 Nat Comm 11). Other species within the genus include, but aren't limited to, *Oscillibacter massiliensis* and, *Oscillibacter ruminantium*, though additional species have been characterized but were not cultured and therefore still have the status candidatus. Non-limiting examples of 16S rDNA sequences of the *Oscillibacter* genus are SEQ ID No. 43-49.

"Prevotella" as used herein refers to a genus of gram-negative, anaerobic bacteria, shaped like small, short rods. *Prevotella* can be found in the human oral cavity, vaginal mucosa and gastro-intestinal track. Though species were also isolated from the rumen of cattle. Species include (but aren't limited to): *P.*

copri, P. melaninogenica and, P. intermedia. P. corpi has been found to be particularly common in non-Western populations and has been linked with gut inflammation. Though depending on the context, *P. corpi* might also have beneficial effects. (Ley 2016 Nat Rev Gastrol Hepatol 13). Non-limiting examples of 16S rDNA sequences of the *Prevotella* genus are SEQ ID No. 50-53.

"Roseburia" as used herein refers to a genus of anaerobic, gram-positive, motile (flagellated) slightly curved rod-shaped bacteria (Nie et al 2021 Front Cell Infect Microbiol 11). The genus consists of, but isn't limited to, five species (*R. intestinalis*, *R. hominis*, *R. inulinivorans*, *R. faecis*, and *R. cecicola*). These are all known to produce butyrate along with proprionate, acetate and other short-chain fatty acids. Non-limiting examples of 16S rDNA sequences of the Roseburia genus are SEQ ID No. 54-60.

"Ruminococcus" as used herein refers to a genus of gram-positive bacteria found in the human gut. This genus is known for its ability to break down cellulose and form methane in the process. While recently the taxonomy of "Ruminococcus" has been subject of scrutiny and might require reclassification of several genera, within this work it comprises, but isn't limited to; *Ruminococcus albus*, Ruminococcus bicirculans, *Ruminococcus bromii*, *Ruminococcus callidus*, *Ruminococcus flavefaciens*, *Ruminococcus gauvreauii*, *Ruminococcus gnavus*, *Ruminococcus lactaris*, *Ruminococcus obeum* and, *Ruminococcus torques*. Non-limiting examples of 16S rDNA sequences of the *Ruminococcus* genus are SEQ ID No. 61-64.

"Sporobacter" as used herein refers to a genus of gram-positive, spore-forming, rod-shaped (albeit with a slight curve) bacteria which are obligately anaerobic (Grech-Mora et al 1996 Intern J Sys Evol Microbiol 46). From this genus currently only a single identified species "*Sporobacter termitidis*", isolated from a wood-feeding termite *Nasutitermes lujae*, has been described. Though additional unclassified and uncultured species are known to exist, and the genus was found in various 16S based studies on the human gut microbiome (Rinninella et al 2019 Microorganisms 7; Senghor et al 2018 Human Microbiome Journal 7-8; Yao et al 2022 Psychol Med 25). Non-limiting examples of 16S rDNA sequences of the *Sporobacter* genus are SEQ ID No. 65-68.

Table 1: Non-limiting examples of 16S rDNA sequences, shown as SEQ ID No. 1-68.

SEQ ID No.	**sequence**
1	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGATGGATGTTTAAGTCAGTTGTGAAAGTTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGATATCTTGAGTGCAGTTGAGGCAGGCGGAATTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCCTGCTAAGCTGCAACTGACATTGAGGCTCGAAAGTGTGGGT

SEQ ID No.	sequence
2	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGACGCTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGGTGTCTTGAGTACAGTAGAGGCAGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTTGCTGG ACTGTAACTGACGCTGATGCTCGAAAGTGTGGGT
3	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGACAGTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGCTGTCTTGAGTACAGTAGAGGTGGGCGGAAT TCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTGGA CTGCAACTGACACTGATGCTCGAAAGTGTGGGT
4	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGATGGATGTTTAAGTCAGTTGTGAAAGT TTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGATGTCTTGAGTGCAGTTGAGGCAGGCGGAATT CGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCCTGCTAAGC TGCAACTGACATTGAGGCTCGAAAGTGTGGGT
5	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGACGCTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGGTGTCTTGAGTACAGTAGAGGCAGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCCTGCTGG ACTGTAACTGACGCTGATGCTCGAAAGTGTGGGT
6	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGATTGTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGCAGTCTTGAGTGCAGTAGAGGTGGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTGG AGTGTAACTGACGCTGATGCTCGAAAGTGTGGGT
7	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGATTGTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGAAACTGGCAGTCTTGAGTACAGTAGAGGTGGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTAG ACTGCAACTGACACTGATGCTCGAAAGTGTGGGT
8	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGATGGGTTGTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAATTGATACTGGCAGTCTTGAGTACAGTTGAGGTAGGCGGAAT TCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTTACTAACC TGTAACTGACATTGATGCTCGAAAGTGTGGGT
9	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGGTTGTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGCGACCTTGAGTGCAACAGAGGTAGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTTACTGG ATTGTAACTGACGCTGATGCTCGAAAGTGTGGGT

SEQ ID No.	sequence
10	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGATTGTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGAAACTGGCAGTCTTGAGTACAGTAGAGGTGGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTAG ACTGTCACTGACACTGATGCTCGAAAGTGTGGGT
11	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGATTGTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGAAACTGGCAGTCTTGAGTACAGTAGAGGTGGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTAG ACTGTTACTGACACTGATGCTCGAAAGTGTGGGT
12	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGACTATTAAGTCAGCTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGTCGTCTTGAGTGCAGTAGAGGTAGGCGGAAT TCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTTACTGGA CTGTAACTGACGCTGATGCTCGAAAGTGTGGGT
13	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGACTGGTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGTCAGTCTTGAGTACAGTAGAGGTGGGCGGAAT TCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTGGA CTGCAACTGACACTGATGCTCGAAAGTGTGGGT
14	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGGTGCTTAAGTCAGTTGTGAAAG TTTGCGGCTCAACCGTAAAATTGCAGTTGATACTGGGCGCCTTGAGTGCAGCATAGGTAGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTTACTGG ACTGTAACTGACGCTGATGCTCGAAAGTGTGGGT
15	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGCGCAGGCGGGCCGGTAAGTTGGAAGTGAAA TCTATGGGCTTAACCCATAAACTGCTTTCAAAACTGCTGGTCTTGAGTGATGGAGAGGCAGGCGGAA TTCCGTGTGTAGCGGTGAAATGCGTAGATATACGGAGGAACACCAGTGGCGAAGGCGGCCTGCTG GACATTAACTGACGCTGAGGCGCGAAAGCGTGGGG
16	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGCGCAGGCGGGCCGGCAAGTTGGAAGTGAAA TCTATGGGCTTAACCCATAAACTGCTTTCAAAACTGCTGGTCTTGAGTGATGGAGAGGCAGGCGGAA TTCCGTGTGTAGCGGTGAAATGCGTAGATATACGGAGGAACACCAGTGGCGAAGGCGGCCTGCTG GACATTAACTGACGCTGAGGCGCGAAAGCGTGGGG
17	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGCGCAGGCGGGCCGGCAAGTTGGAAGTGAAA TCCGGGGGCTTAACCCCCGAACTGCTTTCAAAACTGCTGGTCTTGAGTGATGGAGAGGCAGGCGGA ATTCCGTGTGTAGCGGTGAAATGCGTAGATATACGGAGGAACACCAGTGGCGAAGGCGGCCTGCTG GACATTAACTGACGCTGAGGCGCGAAAGCGTGGGG

SEQ ID No.	sequence
18	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGCGCGTAGGCGGGACGCCAAGTCAGCGGTAAAAG ACTGCAGCTAAACTGTAGCACGCCGTTGAAACTGGCGACCTGGAGACGAGACGAGGGAGGCGGAA CAAGTGAAGTAGCGGTGAAATGCTTAGATATCACTTGGAACCCCGATAGCGAAGGCAGCTTCCCAG GCTCGATCTGACGCTGATGCGCGAGAGCGTGGGT
19	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGCGCGTAGGCGGGACGCCAAGTCAGCGGTAAAAG ACTGCAGCTAAACTGTAGCACGCCGTTGAAACTGGCGCCCTCGAGACGAGACGAGGGAGGCGGAA CAAGTGAAGTAGCGGTGAAATGCTTAGATATCACTTGGAACCCCGATAGCGAAGGCAGCTTCCCAG GCTCGATCTGACGCTGATGCGCGAGAGCGTGGGT
20	CGAGCGTTATCCGGATTTATTGGGTTTAAAGGGCGCGTAGGCGGGACGTCAAGTCAGCGGTAAAAG ACTGCAGCTAAACTGTAGCACGCCGTTGAAACTGGCGCCCTGGAGACGAGACGAGGGAGGCGGAA CAAGTGAAGTAGCGGTGAAATGCATAGATATCACTTGGAACCCCGATAGCGAAGGCAGCTTCCCAG GCTCGTTCTGACGCTGATGCGCGAGAGCGTGGGT
21	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGTGCGTAGGCGGATTGGCAAGTCAGAAGTGAAAT CCATGGGCTTAACCCATGAACTGCTTTTGAAACTGTTAGTCTTGAGTGAAGTAGAGGTAGGCGGAAT TCCCGGTGTAGCGGTGAAATGCGTAGAGATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTGG GCTTTAACTGACGCTGAGGCACGAAAGTGTGGGT
22	CAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGTGTAGGCGGGATATCAAGTCAGAAGTGAAA ATTACGGGCTCAACTCGTAACCTGCTTTTGAAACTGACATTCTTGAGTGAAGTAGAGGCAAGCGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGCTGG GCTTTTACTGACGCTGAGGCTCGAAAGCGTGGGG
23	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGTGCGTAGGCGGCTAGGCAAGTCAGACGTGAAAA CCATGGGCTCAACCTGTGGACTGCGTTTGAAACTGTTTAGCTTGAGTGAAGTAGAGGTAGGCGGAA TTCCCGGTGTAGCGGTGAAATGCGTAGAGATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTG GGCTTTAACTGACGCTGAGGCACGAAAGCATGGGT
24	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGGAGAGCAAGTCAGAAGTGAAA TCTATGGGCTTAACCCATAAACTGCTTTTGAAACTGTTCTTCTTGAGTATCGGAGAGGCAGGCGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCCTGCTGG ACGACAACTGACGCTGAGGCGCGAAAGCGTGGGG
25	CGAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGTGTAGGCGGGAAGGCAAGTCAGAAGTGAAA ATTATGGGCTTAACCCATAACCTGCTTTTGAAACTGTTTTTCTTGAGTGAGGCAGAGGCAAGCGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGCTGG GCCTTTACTGACGCTGAGGCTCGAAAGCGTGGGG

SEQ ID No.	sequence
26	CGAGCGTTGTCCGGATTTACTGGGTGTAAAGGGTGCGTAGGCGGCGAGGCAAGTCAGGCGTGAAA TCTATGGGCTTAACCCATAAACTGCGCTTGAAACTGTCTTGCTTGAGTGAAGTAGAGGTAGGCGGAA TTCCCGGTGTAGCGGTGAAATGCGTAGAGATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTG GGCTTTAACTGACGCTGAAGCACGAAAGCATGGGT
27	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGGACTGCAAGTTGGATGTGAAAT ACCGTGGCTTAACCACGGAACTGCATCCAAAACTGTAGTTCTTGAGTGAAGTAGAGGCAAGCGGAA TTCCGAGTGTAGCGGTGAAATGCGTAGATATTCGGAGGAACACCAGTGGCGAAGGCGGCTTGCTGG GCTTTAACTGACGCTGAGGCTCGAAAGTGTGGGG
28	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCTGTGTAAGTCTGAAGTGAAAG CCCGGGGCTCAACCCCGGGACTGCTTTGGAAACTATGCAGCTAGAGTGTCGGAGAGGTAAGTGGAA TTCCCAGTGTAGCGGTGAAATGCGTAGATATTGGGAGGAACACCAGTGGCGAAGGCGGCTTACTGG ACGATGACTGACGTTGAGGCTCGAAAGCGTGGGG
29	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGGCGGCGGAGCAAGTCAGAAGTGAAA GCCCGGGGCTCAACCCCGGGACGGCTTTTGAAACTGCCCTGCTTGATTTCAGGAGAGGTAAGCGGA ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTG GACTGACAATGACGCTGAGGCTCGAAAGCGTGGGG
30	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGTGCGTAGGTGGCAAGGCAAGTCTGAAGTGAAAA TCCGGGGCTCAACCCCGGAACTGCTTTGGAAACTGTTTAGCTGGAGTACAGGAGAGGTAAGTGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGACTTACTGG ACTGCTACTGACACTGAGGCACGAAAGCGTGGGG
31	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGTGCGTAGGTGGTGAGACAAGTCTGAAGTGAAAA TCCGGGGCTTAACCCCGGAACTGCTTTGGAAACTGCCTGACTAGAGTACAGGAGAGGTAAGTGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGACTTACTGG ACTGCTACTGACACTGAGGCACGAAAGCGTGGGG
32	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGTGCGTAGGTGGCAAGGCAAGTCTGAAGTGAAAA TCCGGGGCTCAACCCCGGAACTGCTTTGGAAACTGTTTAGCTAGAGTACAGGAGAGGTAAGTGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGACTTACTGG ACTGCTACTGACACTGAGGCACGAAAGCGTGGGG
33	CAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAAGACAAGTTGGAAGTGAAA TCTATGGGCTCAACCCATAAACTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGGCGGAAT TCCCGGTGTAGCGGTGGAATGCGTAGATATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTGG GCACCAACTGACGCTGAGGCTCGAAAGTGTGGGT

SEQ ID No.	sequence
34	CAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAAGACAAGTTGGAAGTGAAA TCCATGGGCTCAACCCATGAACTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGGCGGAA TTCCCGGTGTAGCGGTGGAATGCGTAGATATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTG GGCACCAACTGACGCTGAGGCTCGAAAGTGTGGGT
35	CAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAGAACAAGTTGGAAGTGAAA TCCATGGGCTCAACCCATGAACTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGGCGGAA TTCCCGGTGTAGCGGTGGAATGCGTAGATATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTG GGCACCAACTGACGCTGAGGCTCGAAAGTGTGGGT
36	CAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAAGACAAGTTGGAAGTGAAA ACCATGGGCTCAACCCATGAATTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGATGGAA TTCCCGGTGTAGCGGTGGAATGCGTAGATATCGGGAGGAACACCAGTGGCGAAGGCGGTCTACTG GGCACCAACTGACGCTGAGGCTCGAAAGCATGGGT
37	CAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGCGATCAAGTTGGAAGTGAAA TCCATGGGCTCAACCCATGAACTGCTTTCAAAACTGGTCGTCTTGAGTAGTGCAGAGGTAGGCGGAA TTCCCGGTGTAGCGGTGGAATGCGTAGATATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTG GGCACCAACTGACGCTGAGGCTCGAAAGTGTGGGT
38	CAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAAGGCAAGTTGGAAGTGAAA TCCATGGGCTCAACCCATGAACTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGGCGGAA TTCCCGGTGTAGCGGTGGAATGCGTAGATATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTG GGCACCAACTGACGCTGAGGCTCGAAAGTGTGGGT
39	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGGAAGGCAAGTCAGATGTGAAA ACTATGGGCTCAACCCATAGCCTGCATTTGAAACTGTTTTTCTTGAGTGCTGGAGAGGCAATCGGAA TTCCGTGTGTAGCGGTGAAATGCGTAGATATACGGAGGAACACCAGTGGCGAAGGCGGATTGCTG GACAGTAACTGACGCTGAGGCGCGAAAGCGTGGGG
40	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGGATTGCAAGTCAGATGTGAAAA CTGGGGGCTCAACCTCCAGCCTGCATTTGAAACTGTAGTTCTTGAGTGCTGGAGAGGCAATCGGAAT TCCGTGTGTAGCGGTGAAATGCGTAGATATACGGAGGAACACCAGTGGCGAAGGCGGATTGCTGG ACAGTAACTGACGCTGAGGCGCGAAAGCGTGGGG
41	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCAAGGCAAGTCTGATGTGAAAA CCCAGGGCTTAACCCTGGGACTGCATTGGAAACTGTCTGGCTCGAGTGCCGGAGAGGTAAGCGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGGCTTACTGG ACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGG

SEQ ID No.	sequence
42	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCAAGGCAAGTCTGAAGTGAAA GCCCGGTGCTTAACGCCGGGACTGCTTTGGAAACTGTTTGGCTGGAGTGCCGGAGAGGTAAGCGGA ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGGCTTACTG GACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGG
43	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGCCGGGAGGGCAAGTCAGATGTGAAAT CCACGGGCTCAACTCGTGAACTGCATTTGAAACTACTCTTCTTGAGTATCGGAGAGGCAATCGGAAT TCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGATTGCTGG ACGACAACTGACGGTGAGGCGCGAAAGCGTGGGG
44	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGCAGCCGGGCATGCAAGTCAGATGTGAAAT CTCAGGGCTTAACCCTGAAACTGCATTTGAAACTGTATGTCTTGAGTGCCGGAGAGGTAATCGGAAT TCCTTGTGTAGCGGTGAAATGCGTAGATATAAGGAAGAACACCAGTGGCGAAGGCGGATTACTGGA CGGTAACTGACGGTGAGGCGCGAAAGCGTGGGG
45	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGCAGCCGGGCCGGCAAGTCAGATGTGAAAT CTGGAGGCTTAACCTCCAAACTGCATTTGAAACTGTAGGTCTTGAGTACCGGAGAGGTTATCGGAAT TCCTTGTGTAGCGGTGAAATGCGTAGATATAAGGAAGAACACCAGTGGCGAAGGCGGATAACTGG ACGGCAACTGACGGTGAGGCGCGAAAGCGTGGGG
46	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGCCGGGAAGGCAAGTCAGATGTGAAAT CCACGGGCTTAACTCGTGAACTGCATTTGAAACTACTTTTCTTGAGTATCGGAGAGGCAATCGGAAT TCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGATTGCTGG ACGACAACTGACGGTGAGGCGCGAAAGCGTGGGG
47	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGCCGGGGAGACAAGTCAGATGTGAAAT CTGGAGGCTCAACCTCCAAACTGCATTTGAAACTGTTTCTCTTGAGTATCGGAGAGGTAATCGGAAT TCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGGATTACTGGA CGACAACTGACGGTGAGGCGCGAAAGCGTGGGG
48	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGCCGGGAAGGCAAGTCAGATGTGAAAT CCACGGGCTCAACTCGTGAACTGCATTTGAAACTGTTTTTCTTGAGTATCGGAGAGGCAATCGGAAT TCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGATTGCTGG ACGACAACTGACGGTGAGGCGCGAAAGCGTGGGG
49	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGCAGCCGGGTGCGCAAGTCAGATGTGAAAT CTCAGGGCTCAACCCTGAAACTGCATTTGAAACTGTGCATCTTGAGTGCCGGAGAGGTAATCGGAAT TCCTTGTGTAGCGGTGAAATGCGTAGATATAAGGAAGAACACCAGTGGCGAAGGCGGATTACTGGA CGGTAACTGACGGTGAGGCGCGAAAGCGTGGGG

SEQ ID No.	sequence
50	CGGGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCCGGAGATTAAGCGTGTTGTGAAAT GTAGACGCTCAACGTCTGCACTGCAGCGCGAACTGGTTTCCTTGAGTACGCACAAAGTGGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTGG AGCGCAACTGACGCTGAAGCTCGAAAGTGCGGGT
51	CGGGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCCGGAGATTAAGCGTGTTGTGAAAT GTAGATGCTCAACATCTGAACTGCAGCGCGAACTGGTTTCCTTGAGTACGCACAAAGTGGGCGGAA TTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTGG AGCGCAACTGACGCTGAAGCTCGAAAGTGCGGGT
52	CGGGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCCGGAGATTAAGCGTGTTGTGAAAT GTAGATGCTCAACATCTGCACTGCAGCGCGAACTGGTTTCCTTGAGTACGCACAAAGTGGGCGGAAT TCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTCACTGGA GCGCAACTGACGCTGAAGCTCGAAAGTGCGGGT
53	CGGGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGCAGGCCGTCCTTTAAGCGTGCTGTGAAAT GCCGCGGCTCAACCGTGGCACTGCAGCGCGAACTGGAGGACTTGAGTACGCACGAGGTAGGCGGA ATTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACTCCGATTGCGAAGGCAGCTTACCG GAGCGCAACTGACGCTGAGGCTCGAAAGCGCGGGT
54	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGTGCGGCAAGTCTGATGTGAAA GCCCGGGGCTCAACCCCGGTACTGCATTGGAAACTGTCGTACTAGAGTGTCGGAGGGGTAAGCGGA ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTG GACGATAACTGACGCTGAGGCTCGAAAGCGTGGGG
55	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGTGCGGCAAGTCTGATGTGAAA GCCCGGGGCTCAACCCCGGTACTGCATTGGAAACTGTCGTACTAGAGTGTCGGAGGGGTAAGTGGA ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTG GACGATAACTGACGCTGAGGCTCGAAAGCGTGGGG
56	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGAAGGCTAAGTCTGATGTGAAA GCCCGGGGCTCAACCCCGGTACTGCATTGGAAACTGGTCATCTAGAGTGTCGGAGGGGTAAGTGGA ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTG GACGATAACTGACGCTGAGGCTCGAAAGCGTGGGG
57	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGTACGGCAAGTCTGATGTGAAAG CCCGGGGCTCAACCCCGGTACTGCATTGGAAACTGTCGGACTAGAGTGTCGGAGGGGTAAGTGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGG ACGATTACTGACGCTGAGGCTCGAAAGCGTGGGG

SEQ ID No.	sequence
58	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGTACGGCAAGTCTGATGTGAAAT CCCGGGGCTCAACCCCGGTACTGCATTGGAAACTGTCGGACTAGAGTGTCGGAGGGGTAAGTGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGG ACGATTACTGACGCTGAGGCTCGAAAGCGTGGGG
59	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGAAGGCTAAGTCTGATGTGAAA GCCCGGGGCTCAACCCCGGTACTGCATTGGAAACTGTCGTACTAGAGTGTCGGAGGGGTAAGTGGA ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTG GACGATAACTGACGCTGAGGCTCGAAAGCGTGGGG
60	CAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGTGCGGCAAGTCTGATGTGAAA GCCCGGGGCTCAACCCCGGGACTGCATTGGAAACTGTCGTACTTGAGTATCGGAGAGGTAAGTGGA ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTG GACGATAACTGACGCTGAGGCTCGAAAGCGTGGGG
61	CGAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATCGCAAGTCAGATGTGAAA ACTATGGGCTTAACCCATAAACTGCATTTGAAACTGTGGTTCTTGAGTGAAGTAGAGGTAAGCGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCGGCTTACTGG GCTTTAACTGACGCTGAGGCTCGAAAGCGTGGGG
62	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGTGCGTAGGCGGCTTTGCAAGTCAGATGTGAAAT CTATGGGCTCAACCCATAAACTGCATTTGAAACTGTAGAGCTTGAGTGAAGTAGAGGCAGGCGGAA TTCCCCGTGTAGCGGTGAAATGCGTAGAGATGGGGAGGAACACCAGTGGCGAAGGCGGCCTGCTG GGCTTTAACTGACGCTGAGGCACGAAAGCGTGGGT
63	CGAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATGGCAAGTCAGATGTGAAA ACTATGGGCTCAACCCATAGACTGCATTTGAAACTGTTGTTCTTGAGTGAGGTAGAGGTAAGCGGAA TTCCTGGTGTAGCGGTGAAATGCGTAGAGATCAGGAGGAACATCGGTGGCGAAGGCGGCTTACTG GGCCTTTACTGACGCTGAGGCTCGAAAGCGTGGGG
64	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGTGCGTAGGCGGCTTTGCAAGTCAGATGTGAAAT CTATGGGCTCAACCCATAGCCTGCATTTGAAACTGCAGAGCTTGAGTGAAGTAGAGGCAGGCGGAA TTCCCCGTGTAGCGGTGAAATGCGTAGAGATGGGGAGGAACACCAGTGGCGAAGGCGGCCTGCTG GGCTTTAACTGACGCTGAGGCACGAAAGCGTGGGT
65	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGAGATGCAAGTCAGATGTGAAAT CCTCGGGCTTAACCCGGGAACTGCATTTGAAACTGTATCCCTTGAGTATCGGAGAGGCAGGCGGAA TTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCCTGCTGG ACGACAACTGACGCTGAGGCGCGAAAGCGTGGGG

SEQ ID No.	sequence
66	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGCGTAGGCGGGATGGCAAGTCAGATGTGAAAT CCATGGGCTCAACCCATGAACTGCATTTGAAACTGTCGTTCTTGAGTATCGGAGAGGCAAGCGGAAT TCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGCTGGA CGACAACTGACGCTGAGGCGCGAAAGCGTGGGG
67	CAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGGAAGACAGGTCAGATGTGAAA TGTCGGGGCTCAACCCCGGAACTGCATTTGAAACCGTTTTTCTTGAGTATCGGAGAGGCAGGCGGA ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCCTGCTG GACGACAACTGACGCTGAGGCGCGAAAGCGTGGGG
68	CGAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGCGTAGGCGGGAATGCAAGTCAGATGTGAAAT CCAGGGGCTTAACCCTTGAACTGCATTTGAAACTGTATTTCTTGAGTGTCGGAGAGGTTGACGGAAT TCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGTCAACTGG ACGATAACTGACGCTGAGGCGCGAAAGCGTGGGG

Preferably, the method according to any one of the above-mentioned embodiments is the method to detect the gut flora dysbiosis associated with an inflammatory disorder, obesity, diabetes type 2, depression, and/or anxiety.

Preferably, the method according to any one of the above-mentioned embodiments is the method to detect the inflammatory disorder selected from the list consisting of spondyloarthritis, ankylosing spondylitis, reactive arthritis, psoriatic arthritis, enteropathic arthritis, undifferentiated spondyloarthritis, juvenile idiopathic arthritis, primary sclerosing cholangitis, multiple sclerosis and any gut inflammation associated therewith.

Preferably, the method according to any one of the above-mentioned embodiments is the method to detect the inflammatory disorder is a gut inflammatory disorder selected from the list consisting of Crohn's disease, irritable bowel syndrome, inflammatory bowel disease, ulcerative colitis and celiac disease.

Preferably, in the method according to any one of the above-mentioned embodiments, said biological sample is a stool sample, a mucosal biopsy sample or a sample of the lumen content.

Preferably, in the method according to any one of the above-mentioned embodiments, the quantification of the at least two bacterial genera is performed by quantifying DNA sequences specific for said at least two bacterial genera, preferably at least three or at least four bacterial genera. Preferably, said DNA sequences are 16S rDNA sequences. Preferably, said quantification of the DNA sequences is performed by quantitative PCR.

In a preferred embodiment, said abundances of the at least two bacterial genera are expressed as the log cells/g. Alternatively, the abundances may be expressed as copies/µL or Cycle threshold (Ct) value.

Preferably, prior to quantification of said DNA sequences, DNA is extracted from the biological sample.

Biomarker panels

As will be appreciated by the skilled person, ROC curves are standard tool in the assessment of the performance of a diagnostic test with the AUC being a measure of discrimination and allowing investigators to compare the performance of two or more diagnostic tests. In general, an AUC of 0.5 suggests no discrimination (i.e. no ability to diagnose patients with and without the disease or condition based on the test). An AUC between 0.7 and 0.8 is considered acceptable, while an AUC of 0.8 to 0.9 is considered excellent, and an AUC of more than 0.9 is considered outstanding (Hosmer and Lemeshow 2000 Applied Logistic Regression, 2nd Ed. Chapter 5, p 160-164).

The inventors of current application analysed all possible combinations within the set of 13 bacterial genera that yielded an AUC of 0.7 or more in predicting the B2 enterotype or gut flora dysbiosis. These combinations of genera are provided herein as biomarker panels.

In a first aspect, a biomarker, bacterial genera panel is provided comprising at least two bacterial genera selected from the list consisting of *Oscillibacter*, *Faecalibacterium*, *Ruminococcus*, *Prevotella* and *Bacteroides* wherein the panel does not consist of *Faecalibacterium* and *Bacteroides*.

In a preferred embodiment, biomarker, bacterial genera panel is provided comprising at least genera *Oscillibacter* and *Faecalibacterium*.

In a preferred embodiment, said panel comprises at least one further genus, preferably at least two further genera selected from a list consisting of *Ruminococcus*, *Prevotella*, and *Bacteroides*.

In another embodiment said panel comprises or consists of five bacterial genera selected from the list consisting of *Oscillibacter*, *Faecalibacterium*, *Ruminococcus*, *Prevotella* and *Bacteroides*.

In an alternative embodiment, said bacterial genera panel comprises at least two bacterial genera comprising *Oscillibacter* and *Bacteroides*.

Preferably, said bacterial genera panel is for use to diagnose a subject with gut flora dysbiosis.

Preferably, said bacterial genera panel is suitable for determining gut flora dysbiosis in a subject according to the method of the invention.

In one embodiment, the biomarker panel or the bacterial genera panel consists of 2 bacterial genera wherein one bacterial genus is selected from Table 2 and one bacterial genus is selected from Table 3, wherein the panel does not consist of *Faecalibacterium* and *Bacteroides*.

In another embodiment, the biomarker panel or the bacterial genera panel comprises at least 3 bacterial genera and comprises at least one bacterial genus selected from Table 2 and one bacterial genus selected from Table 3, or comprises at least 2 bacterial genera from Table 2 and at least one bacterial genus from

Table 3, or comprises at least 1 bacterial genus from Table 2 and at least 2 bacterial genera from Table 3. In a particular embodiment, the panel does not consist of *Faecalibacterium* and *Bacteroides*.

In another embodiment, a biomarker panel is provided comprising at least one genus from Table 2 and at least one genus from Table 4 and/or 5. In another embodiment, a biomarker panel is provided comprising at least two bacterial genera from Table 2 and at least one from Table 4 and/or *Roseburia* or *Flavonifractor* or comprising at least one bacterial genus from Table 2 and at least two from Table 4 and/or *Roseburia* or *Flavonifractor*. In another embodiment, a biomarker panel is provided comprising at least two bacterial genera from Table 2 and at least two from Table 4 and/or genera *Roseburia* or *Flavonifractor* or comprising at least 3 bacterial genera from Table 2 and at least one from Table 4 and/or *Roseburia* or *Flavonifractor* comprising at least one bacterial genus from Table 2 and at least 3 from Table 3 and/or *Roseburia* or *Flavonifractor* .

Table 2. Selected bacterial genera.

Sporobacter	Fusicatenibacter	Oscillibacter
Coprococcus	Ruminococcus	Faecalibacterium

Table 3. Selected bacterial genera.

Clostridium_IV	Roseburia	Bacteroides
Butyricimonas	Prevotella	
Butyricicoccus	Flavonifractor	

Table 4. Selected bacterial genera.

Clostridium_IV	Butyricicoccus
Butyricimonas	Prevotella

In a particular embodiment, the biomarker panel comprises or consists of *Coprococcus* and at least one or at least two genera selected from *Faecalibacterium*, *Oscillibacter*, *Ruminococcus*, *Sporobacter* and *Fusicatenibacter*. In another particular embodiment, the biomarker panel comprises or consists of *Faecalibacterium* and at least one or at least two genera selected from *Oscillibacter*, *Ruminococcus*, *Sporobacter* and *Fusicatenibacter*. In another particular embodiment, the biomarker panel comprises or consists of *Fusicatenibacter* and at least one or at least two genera selected from *Coprococcus*, *Faecalibacterium, Oscillibacter, Ruminococcus* and *Sporobacter*. In another particular embodiment, the

biomarker panel comprises or consists of at least *Sporobacter* and/or *Ruminococcus* and at least *Faecalibacterium* and/or *Oscillibacter*.

In yet a particular embodiment, the biomarker panel comprises or consists of at least one genus selected from *Ruminococcus*, *Sporobacter*, *Coprococcus*, *Faecalibacterium* and *Fusicatenibacter* and at least two genera selected from *Butyricicoccus, Bacteroides, Prevotella* and *Clostridium_IV* or at least one genus selected from *Butyricicoccus, Bacteroides, Prevotella* and *Clostridium_IV* and at least *Flavonifractor* and/or *Roseburia*.

In another particular embodiment, the biomarker panel comprises *Fusicatenibacter* and *Faecalibacterium* and *Coprococcus*. In another particular embodiment, the biomarker panel comprises *Fusicatenibacter* and *Ruminococcus* and at least 1 from the group consisting of *Coprococcus*, *Faecalibacterium* and *Sporobacter*. In another particular embodiment, the biomarker panel comprises *Ruminococcus* and *Coprococcus*. In another particular embodiment, the biomarker panel comprises *Ruminococcus* and *Faecalibacterium* and at least 1 from the group consisting of *Coproccus* and *Sporobacter*.

It was also found that an AUC of 0.7 or more is obtained by any combination of at least 5 bacterial genera selected from Table 5. Therefore, in another embodiment, a biomarker panel is provided comprising at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12 or 13 bacterial biomarkers selected from Table 5.

Table 5. Selected bacterial genera.

Sporobacter	Oscillibacter	Butyricicoccus
Coprococcus	Faecalibacterium	Prevotella
Fusicatenibacter	Clostridium_IV	Roseburia
Ruminococcus	Butyricimonas	Flavonifractor
Bacteroides		

<u>Uses of the biomarker panels</u>

The biomarker panels herein disclosed are designed to be predictive for the dysbiotic B2 enterotype, more particularly to be predictive for the dysbiotic B2 enterotype in the methods herein disclosed. The B2 enterotype represents a dysbiotic gut flora and is associated with health problems and several inflammatory disorders. People who have this dysbiotic enterotype have a higher blood concentration of C-reactive protein – a hallmark of inflammation – than do individuals who have other enterotypes (Costea et al 2018 Nat Microbiol 3: 8-16). More than 75% of individuals who have IBD have the B2

enterotype in contrast to fewer than 15% of people who do not have the disease (Veira-Silva et al 2019 Nat Microbiol 4: 1826-1831). The B2 enterotype is also correlated to primary sclerosing cholangitis (Veira-Silva et al 2019 Nat Microbiol 4: 1826-1831), multiple sclerosis (Reynders et al 2020 Ann Clin Transl Neur 7: 406-419), depression (Valles-Colomer et al 2019 Nat Microbiol 4: 623-632) and obesity (Veira-Silva et al 2020 Nature 581: 310-315).

Hence, the biomarkers herein disclosed have multiple applications, for example and without the purpose of limiting, the biomarker panels can be used in (methods of) diagnosing or detecting a dysbiotic gut flora or a disease or disorder associated with gut flora dysbiosis.

More particularly, for the purpose of diagnosing or detecting a dysbiotic gut flora in a subject, a biological sample from the subject is needed. Said biological sample is preferably a stool sample, a mucosal biopsy sample of a sample from the lumen content. Such samples comprise a huge quantity and diversity of microorganisms such as bacterial and fungal strains. For the purpose of current invention, only the bacterial strains are taken into account. From the sample, the abundances or levels of separate bacterial genera are to be measured or quantified. This can for example be done by first extracting DNA from the sample according to established methods of which the skilled person is familiar with, followed by a quantification of genus-specific DNA fragments. In microbiology, the bacterial 16s rRNA or regions thereof (e.g. the V4 region) is standardly used for reconstructing evolutionary relationships and identification of bacteria isolated on genus and species level. Hence, based on the amount of genus-specific DNA fragments such as quantification of different genus-specific 16S rRNA sequences, the relative but also absolute abundances of bacterial genera in a sample can be determined. This can for example be done by 16S rRNA sequencing but can also be done more easily by a quantitative Polymerase Chain Reaction (or qPCR). These abundances can then be compared with those of a control sample from a different subject, for example a healthy person. In practice, comparisons are done with reference abundances obtained from control samples from a plurality of healthy persons. Based on the invention, only the abundance of a limited number of bacterial genera should be compared between a test sample and a reference in order to diagnose the test sample.

Therefore any of the above-mentioned biomarker panels according to any embodiment of the second aspect of the invention is provided for use in detecting a dysbiotic gut flora in a biological sample obtained from a subject, more particularly a stool sample. In some embodiments, said dysbiotic gut flora is associated with an inflammatory disorder in the subject. In other embodiments, said dysbiotic gut flora is associated with obesity, diabetes type 2, depression, a suboptimal lifestyle or suboptimal dietary choices.

In another embodiment, the use of any of the biomarker panels described in the first aspect of the application is provided to classify, categorize or distinguish different gut flora microbiomes based on isolated biological samples. The use of any of said biomarker panels of the application is also provided to distinguish a B2 enterotype or a dysbiotic gut microbiome or a gut microbiome associated with gut flora dysbiosis from a healthy gut microbiome. Hence the biomarker panels of the invention can be used to diagnose gut flora dysbiosis and/or an inflammatory disorder in a subject or patient.

Diagnostic methods

The biomarkers panels of the invention herein described are of use in methods of diagnosing or detecting gut flora dysbiosis in a subject.

In an embodiment, a method of detecting gut flora dysbiosis in a subject is provided. Said method comprises the following steps:

- measuring or quantifying in a biological sample obtained from said subject the abundances or levels of at least one bacterial genus selected from the genera listed in Table 2;
- comparing the measured abundances or levels of said at least one genus measured in said subject's sample to those of a healthy control sample or to reference abundances or reference levels of said at least one genus obtained from one or more healthy control samples; and
- determining that the subject suffers from gut flora dysbiosis if the measured abundances or levels of the at least one genus in the subject sample are decreased compared to those of the healthy control sample

In the method of the invention, the abundances or the levels of the bacterial genera are measured or quantified for each of the selected bacterial genera separately. The level of each selected bacterial genus is then compared to the level of the same bacterial genus in the control sample. The decision to diagnose a subject with gut flora dysbiosis using the methods of the invention is taken only if all the abundances of all selected bacterial genera are decreased compared to the abundances of the same bacterial genera in the control sample. In one embodiment, the control sample is a single biological sample from a control subject. In another embodiment, the control sample is a plurality of biological samples obtained from a plurality of control subjects. Hence, in one embodiment, the measured abundances of said at least one genus can also be compared to reference abundances or references levels of said at least one genus obtained from control samples of a group or population of healthy subjects. In that case and concerning the third aspect of the application, decreased abundances or levels of the least one genus compared to those reference abundances or levels is indicative of gut flora dysbiosis.

The steps of the method of the third aspect can be alternatively phrased as:

- measuring or quantifying in a biological sample obtained from said subject the abundances or levels of at least one bacterial genus selected from the list consisting of *Sporobacter*, *Coprococcus*, *Fusicatenibacter, Ruminococcus, Oscillibacter* and *Faecalibacterium*; and
- diagnosing gut flora dysbiosis in said subject if the abundances or levels of said at least one bacterial genus are decreased compared to those of a control sample obtained from a healthy subject or plurality of healthy subjects; or
- determining whether or not a stool sample from said subject comprises a decreased level of said at least one bacterial genus as compared to a healthy control or plurality of healthy controls, wherein a decreased level of said at least one bacterial genus indicates that said subject has gut flora dysbiosis.

The methods of the invention can alternatively be phrased as methods of measuring the probability of a subject developing or having gut flora dysbiosis. Said methods comprise the following steps:

a) measuring or quantifying the abundances of at least one bacterial genus selected from the genera listed in Table 2 in a biological sample from the subject;

b) determining the probability of the subject developing or having gut flora dysbiosis based on the abundances measured in the previous step, wherein decreased abundances of said at least one genus in the subject's sample compared to a control sample from a healthy subject or plurality of healthy subjects indicates a high probability of the subject developing or having gut flora dysbiosis.

In alternative embodiment, a method of detecting gut flora dysbiosis in a subject is provided. Said method comprises the following steps:

a) measuring or quantifying in a biological sample obtained from said subject the abundance or level of each of at least two bacterial genera;

b) comparing the measured abundances or levels of said genera measured in said subject's sample separately to those of a healthy control sample or to reference abundances or reference levels of said genera obtained from a plurality of healthy control samples; and

c) determining that the subject suffers from gut flora dysbiosis if the measured abundances or levels of each of the at least two bacterial genera in the subject sample are statistically significantly different compared to those of the healthy control sample or to the reference abundances or reference levels of said genera obtained from a plurality of healthy control samples.

It should be clear to the skilled person that the abundances or the levels of the bacterial genera are measured or quantified for each of the selected bacterial genera separately. The level of each selected bacterial genus is then compared to the level of the same bacterial genus in the control sample. It is thus not the general or cumulative abundance of the selected bacterial genera that is compared between a test and control sample, but the abundances of each of the selected bacterial genera separately.

In the methods of the invention, the decision to diagnose a subject with gut flora dysbiosis is taken only if all of the separate abundances of each of the selected bacterial genera are statistically significantly different compared to the abundances of the same bacterial genera in the control sample.

The method steps of the methods of the invention can alternatively be phrased as:

- measuring or quantifying in a biological sample obtained from said subject the abundances or levels of each of at least two bacterial genera; and
- diagnosing gut flora dysbiosis in said subject if the abundances or levels of each of the at least two bacterial genera are statistically significantly different compared to those of a control sample obtained from a healthy subject; or
- determining whether or not a stool sample from said subject comprises a statistically significantly different level of each of said at least 2 bacterial genera as compared to a healthy control, wherein a statistically significantly different level of each of said at least 2 bacterial genera indicates that said subject has gut flora dysbiosis.

The methods of the invention can alternatively be phrased as methods of measuring the probability of a subject developing or having gut flora dysbiosis. Said methods comprise the following steps:

a) measuring or quantifying the abundances or levels of each of at least 2 bacterial genera in a biological sample from the subject;
b) determining the probability of the subject developing or having gut flora dysbiosis based on the abundances or levels measured in the previous step, wherein statistically significantly different abundances or levels of each of said bacterial genera in the subject's sample compared to a control sample from a healthy subject indicates a high probability of the subject developing or having gut flora dysbiosis.

In one embodiment of, said at least two bacterial genera comprise at least one bacterial genus or at least two bacterial genera selected from Table 2. Said bacterial genera selected from Table 2 are predictive for diagnosing or detecting gut flora dysbiosis when each of the abundances of all selected bacterial genera separately is decreased, more particularly statistically significantly decreased. Hence, in one

embodiment, the statistically significant different requirement in the methods herein described means decreased or more particularly statistically significantly decreased. And thus a method of detecting gut flora dysbiosis in a subject is provided comprising the following steps:

a) measuring or quantifying in a biological sample obtained from said subject the abundance or level of each of at least two bacterial genera, said two bacterial genera comprising at least one genus or at least two bacterial genera selected from the list consisting of *Sporobacter, Coprococcus, Fusicatenibacter, Ruminococcus, Oscillibacter* and *Faecalibacterium*;

b) comparing the measured abundances or levels of each of said at least one genus or two genera measured in said subject's sample separately to those of a healthy control sample or to reference abundances or reference levels of each of said at least one genus or two genera obtained from a plurality of healthy control samples; and

c) determining that the subject suffers from gut flora dysbiosis if the measured abundances or levels of each of the at least one genus or two genera in the subject sample are decreased compared to those of the healthy control sample or compared to said reference abundances or reference levels.

In a particular embodiment, said at least two bacterial genera selected from Table 2 comprise or consist of the genera *Coprococcus* and *Ruminococcus, Faecalibacterium* and *Ruminococcus, Faecalibacterium* and *Sporobacter, Fusicatenibacter* and *Ruminoccus, Fusicatenibacter* and *Sporobacter* or *Sporobacter* and *Ruminococcus*.

In another embodiment of all the methods of the invention, said at least two bacterial genera comprise at least one genus from Table 2 and at least one genus from Table 4 and/or *Roseburia* or *Flavonifractor*, wherein the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased, except for genus *Flavonifractor* that is predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when its abundance is increased, more particularly statistically significantly increased compared to the abundance of *Flavonifractor* in a healthy control or compared to the reference abundances obtained from a plurality of healthy control samples.

In a particular embodiment of all the methods of the fourth aspect, said at least two bacterial genera comprise or consist of *Faecalibacterium* and *Clostridium_IV* or comprise or consist of *Sporobacter* and *Butyricimonas*.

In another embodiment of all the methods of the invention, said at least two bacterial genera comprise at least one genus from Table 6 and at least one genus from Table 3, 4 and/or *Roseburia* or *Flavonifractor,* wherein the bacterial genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when their abundance is decreased, more particularly statistically significantly decreased, except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when the abundance of *Bacteroides* and the abundance of *Flavonifractor* is increased, more particularly statistically significantly increased compared to the abundance of *Bacteroides* and the abundance of *Flavonifractor* in a healthy control.

Table 6.

Sporobacter	Fusicatenibacter	Oscillibacter
Coprococcus	Ruminococcus	

In another embodiment of all the methods of the invention, said at least two bacterial genera comprise at least one genus from Table 2 and at least *Flavonifractor* or *Roseburia*, wherein the bacterial genera from Table 2 as well as the genus *Roseburia* are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when their abundance is decreased, more particularly statistically significantly decreased compared to the abundance of the selected genera in a healthy control and wherein the bacterial genus *Flavonifractor* is predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when its abundance is increased, more particularly statistically significantly increased compared to the abundance of *Flavonifractor* in a healthy control. In a particular embodiment, the at least one genus from Table 2 and at least *Flavonifractor* or *Roseburia* comprise or consist of *Roseburia* and *Sporobacter* or comprise or consist of *Flavonifractor* and *Faecalibacterium*.

In another embodiment of all the methods of the fourth aspect, said at least two bacterial genera are selected from the bacterial genera listed in Table 3, more particularly listed in Table 4 and *Bacteroides*, wherein the bacterial genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when their abundance is decreased, more particularly statistically significantly decreased, except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when their abundance is increased, more particularly statistically significantly increased.

In an embodiment, a method of detecting gut flora dysbiosis in a subject is provided. Said method comprises the following steps:

a) measuring or quantifying in a biological sample obtained from said subject the abundances or levels of each of at least 3 bacterial genera;

b) comparing the measured abundances or levels of each of said genera measured in said subject's sample to each of those of a healthy control sample or to reference abundances or reference levels of each of said genera obtained from a plurality of healthy control sample; and

c) determining that the subject suffers from gut flora dysbiosis if the measured abundances or levels of each of the at least 3 bacterial genera in the subject sample are statistically significantly different compared to each of those of the healthy control sample or to each of said reference abundances or reference levels.

It should be clear to the skilled person that the abundances or the levels of the bacterial genera are measured or quantified for each of the selected bacterial genera separately. The level of each selected bacterial genus is then compared to the level of the same bacterial genus in the control sample. In the methods of the invention, the decision to diagnose a subject with gut flora dysbiosis is taken only if all of the abundances of each of the selected bacterial genera are statistically significantly different compared to the abundances of the same bacterial genera in the control sample.

The steps of the methods can be alternatively phrased as:

- measuring or quantifying in a biological sample obtained from said subject the abundances or levels of each of at least 3 bacterial genera; and
- diagnosing gut flora dysbiosis in said subject if the abundances or levels of each of the at least 3 bacterial genera are statistically significantly different compared to each of those of a control sample obtained from a healthy subject; or
- determining whether or not a stool sample from said subject comprises a statistically significantly different level of each of said at least 3 bacterial genera as compared to each of those of a healthy control, wherein a statistically significantly different level of each and all of said at least 3 bacterial genera indicates that said subject has gut flora dysbiosis.

The methods of the invention are equivalent to methods of measuring the probability of a subject developing or having gut flora dysbiosis. Said methods comprise the following steps:

a) measuring or quantifying the abundances or levels of each of at least 3 bacterial genera in a biological sample from the subject;

b) determining the probability of the subject developing or having gut flora dysbiosis based on each of the abundances or levels measured in the previous step, wherein statistically significantly different abundances or levels of each and all of said bacterial genera in the subject's sample compared to each of those of a control sample from a healthy subject indicates a high probability of the subject developing or having gut flora dysbiosis.

In one embodiment of all the methods of the invention, said at least 3 bacterial genera comprise at least one bacterial genus or at least two bacterial genera or at least three bacterial genera selected from Table 2. Said bacterial genera selected from Table 2 are predictive for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased. Hence, in one embodiment, the statistically significant different requirement in the methods means decreased or more particularly statistically significantly decreased.

In a particular embodiment, said at least 3 bacterial genera comprise or consist of the genus *Coprococcus* and at least two bacterial genera selected from the list consisting of *Faecalibacterium, Oscillibacter, Ruminococcus, Sporobacter* and *Fusicatenibacter*.

In another particular embodiment, said at least 3 bacterial genera comprise or consist of the genus *Fusicatenibacter* and at least two bacterial genera selected from the list consisting of *Coprococcus, Faecalibacterium, Oscillibacter, Ruminococcus* and *Sporobacter*.

In another particular embodiment, said at least 3 bacterial genera comprise or consist of *Sporobacter, Ruminococcus, Oscillibacter* and/or *Faecalibacterium*.

In one embodiment of all the methods herein disclosed, said at least 3 bacterial genera comprise at least two bacterial genera selected from Table 2 and at least one bacterial genus selected from Table 3, 4, *Roseburia* or *Flavonifractor* and/or *Bacteroides*, wherein each of the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased, except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is increased, more particularly statistically significantly increased compared to the abundance of Bacteroides and/or *Flavonifractor* respectively in the control sample of a healthy person.

In another embodiment of all the methods herein disclosed, said at least 3 bacterial genera comprise at least one bacterial genus selected from Table 1 and at least two bacterial genera selected from Table 3, 4, *Roseburia* or *Flavonifractor* and/or *Bacteroides*, wherein the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased, compared to each of the abundances in a control sample, except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is increased, more particularly statistically significantly increased compared to the abundance of *Bacteroides* and/or *Flavonifractor* respectively in the control sample of a healthy person.

In another embodiment of all the methods herein disclosed, said at least 3 bacterial genera comprise at least 3 bacterial genera selected from Table 3 or from Table 4 and *Bacteroides*, wherein the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased, compared to each of the abundances in a control sample, except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is increased, more particularly statistically significantly increased compared to the abundance of *Bacteroides* and/or *Flavonifractor* respectively in the control sample of a healthy person.

In a particular embodiment, said at least 3 bacterial genera comprise at least one bacterial genus selected from the list consisting of *Ruminococcus, Sporobacter, Coprococcus, Faecalibacterium* and *Fusicatenibacter* and at least 2 bacterial genera selected from the list consisting of *Butyricicoccus*, *Bacteroides, Prevotella* and *Clostridium_IV*, wherein the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased, compared to each of the abundances in a control sample, except for *Bacteroides* that is predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when its abundance is increased, more particularly statistically significantly increased compared to the abundance of Bacteroides in the control sample of a healthy person.

In another particular embodiment, said at least 3 bacterial genera comprise at least one bacterial genus selected from the list consisting of *Ruminococcus, Sporobacter, Coprococcus, Faecalibacterium* and *Fusicatenibacter*, at least one bacterial genus selected from the list consisting of *Butyricicoccus, Bacteroides, Prevotella* and *Clostridium_IV* and at least one genus selected from *Roseburia* or

Flavonifractor, wherein the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when their abundance is decreased, more particularly statistically significantly decreased, compared to each of the abundances in a control sample, except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when their abundance is increased, more particularly statistically significantly increased compared to the abundance of *Bacteroides* and/or *Flavonifractor* respectively in the control sample of a healthy person.

In an embodiment, a method of detecting gut flora dysbiosis in a subject is provided. Said method comprises the following steps:

a) measuring or quantifying in a biological sample obtained from said subject the abundances or levels of each of at least 4 bacterial genera;

b) comparing the measured abundances or levels of each of said genera measured in said subject's sample to each of those of a healthy control sample or to reference abundances or reference levels of each of said genera obtained from a plurality of healthy control samples; and

c) determining that the subject suffers from gut flora dysbiosis if each of the measured abundances or levels of each of the at least 4 bacterial genera in the subject sample are statistically significantly different compared to each of those of the healthy control sample

It should be clear to the skilled person that the abundances or the levels of the bacterial genera are measured or quantified for each of the selected bacterial genera separately. The level of each selected bacterial genus is then compared to the level of the same bacterial genus in the control sample. In the methods of the invention, the decision to diagnose a subject with gut flora dysbiosis is taken only if all of the abundances of each of the selected bacterial genera are statistically significantly different compared to each of the abundances of the same bacterial genera in the control sample.

The method steps of the methods of the invention can be alternatively phrased as:

- measuring or quantifying in a biological sample obtained from said subject the abundances or levels of each of at least 4 bacterial genera; and
- diagnosing gut flora dysbiosis in said subject if the abundances or levels of each of the at least 4 bacterial genera are statistically significantly different compared to each of those of a control sample obtained from a healthy subject; or

- determining whether or not a stool sample from said subject comprises a statistically significantly different level of each of said at least 4 bacterial genera as compared to each of those of a healthy control, wherein a statistically significantly different level of each of said at least 4 bacterial genera indicates that said subject has gut flora dysbiosis.

The methods of the sixth aspect can be equivalently phrased as methods of measuring the probability of a subject developing or having gut flora dysbiosis. Said methods comprise the following steps:

a) measuring or quantifying the abundances or levels of each of at least 4 bacterial genera in a biological sample from the subject;

b) determining the probability of the subject developing or having gut flora dysbiosis based on the abundances or levels measured in the previous step, wherein statistically significantly different abundances or levels of each of said bacterial genera in the subject's sample compared to a control sample from a healthy subject indicates a high probability of the subject developing or having gut flora dysbiosis.

In one embodiment of all the methods of the sixth aspect, said at least 4 bacterial genera comprise at least one bacterial genus or at least two, three or four bacterial genera selected from Table 2. Said bacterial genera selected from Table 2 are predictive for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased compared to each of the abundances in a control sample. Hence, in one embodiment, the statistically significant different requirement in the methods herein described means decreased or more particularly statistically significantly decreased.

In one embodiment of all the methods of the sixth aspect, said at least 4 bacterial genera comprise at least two bacterial genera selected from Table 2 and at least two bacterial genera selected from Table 3 or from Table 4 and *Bacteroides* or from *Roseburia* or *Flavonifractor*, wherein the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased, compared to each of the abundances in a control sample; except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of its abundance is increased, more particularly statistically significantly increased compared to the abundance of *Bacteroides* and/or *Flavonifractor* respectively in the control sample of a healthy person.

In another embodiment of all the methods of the invention, said at least 4 bacterial genera comprise at least one bacterial genus selected from Table 2 and at least three bacterial genera selected from Table

3 or from Table 4 and *Bacteroides* or said at least 4 bacterial genera comprise at least three bacterial genera selected from Table 2 and at least one genus selected from Table 3, 4, *Roseburia* or *Flavonifractor* or *Bacteroides*, wherein the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased, compared to each of the abundances in a control sample; except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of its abundance is increased, more particularly statistically significantly increased compared to the abundance of *Bacteroides* and/or *Flavonifractor* respectively in the control sample of a healthy person.

In another embodiment of all the methods of the invention, said at least 4 bacterial genera comprise at least 3 or at least 4 bacterial genera selected from Table 3 or from Table 4 and *Bacteroides*, wherein the genera are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when each of their abundance is decreased, more particularly statistically significantly decreased, compared to each of the abundances in a control sample; except for *Bacteroides* and *Flavonifractor* that are predictive in the diagnostic methods herein provided or more particularly for diagnosing or detecting gut flora dysbiosis when its abundance is increased, more particularly statistically significantly increased compared to the abundance of *Bacteroides* and/or *Flavonifractor* respectively in the control sample of a healthy person.

It is herein demonstrated that any combination of at least 5 bacterial genera selected from Table 5 is predictive for gut flora dysbiosis and the diseases associated therewith with an AUC of at least 0.7. Therefore, in another embodiment, a method of detecting gut flora dysbiosis in a subject is provided. Said method comprises the following steps:

a) measuring or quantifying in a biological sample obtained from said subject the abundances or levels of each of at least 5 bacterial genera selected from Table 5;

b) comparing the measured abundances or levels of each of said genera measured in said subject's sample to each of those of a healthy control sample or to reference abundances or reference levels of each of said genera obtained from a plurality of healthy control samples; and

c) determining that the subject suffers from gut flora dysbiosis if the measured abundances or levels of each of the at least 5 bacterial genera in the subject sample are decreased compared to each of those of the healthy control sample or to said reference abundances or reference levels, except for Bacteroides and *Flavonifractor* that are predictive for glut flora dysbiosis if

each of their abundance is increased compared to those of a healthy control sample or to said references abundances or reference levels.

It should be clear to the skilled person that the abundances or the levels of the bacterial genera are measured or quantified for each of the selected bacterial genera separately. The level of each selected bacterial genus is then compared to the level of the same bacterial genus in the control sample. In the methods of the invention, the decision to diagnose a subject with gut flora dysbiosis is taken only if all of the abundances of each of the selected bacterial genera are decreased compared to each of the abundances of the same bacterial genera in the control sample, or in case Bacteroides and/or *Flavonifractor* are selected, if the abundances of *Bacteroides* and/or *Flavonifractor* from the subject's sample are increased compared to the abundances of *Bacteroides* and/or *Flavonifractor* in the control sample.

The method steps of the methods of the invention can be alternatively phrased as:

- measuring or quantifying in a biological sample obtained from said subject the abundances or levels of each of at least 5 bacterial genera selected from Table 5; and
- diagnosing gut flora dysbiosis in said subject if the abundances or levels of each of the at least 5 bacterial genera are decreased compared to each of the abundances of the same bacterial genera in the control sample, or in case *Bacteroides* and/or *Flavonifractor* are selected, if the abundances of *Bacteroides* and/or *Flavonifractor* from the subject's sample are increased compared to the abundances of *Bacteroides* and/or *Flavonifractor* in the control sample; or
- determining whether or not a stool sample from said subject comprises a statistically significantly different level of each of said at least 5 bacterial genera as compared to each of those of a healthy control, wherein a statistically significantly increased level of *Bacteroides* and *Flavonifractor* and a statistically significantly decreased level of each of the other selected genera indicate that said subject has gut flora dysbiosis.

The methods of the invention can be equivalently phrased as methods of measuring the probability of a subject developing or having gut flora dysbiosis. Said methods comprise the following steps:

a) measuring or quantifying the abundances or levels of each of at least 5 bacterial genera from Table 5 in a biological sample from the subject;

b) determining the probability of the subject developing or having gut flora dysbiosis based on the abundances or levels measured in the previous step, wherein increased levels of each of *Flavonifractor* and *Bacteroides* and decreased levels of each of the other genera from Table 5

in the subject's sample compared to a control sample from a healthy subject indicates a high probability of the subject developing or having gut flora dysbiosis.

In one embodiment of the methods of the seventh aspect, said at least 5 bacterial genera are at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12 or 13 bacterial genera selected from Table 5.

In another embodiment, the application also provides methods to detect the presence or to assess the risk of developing a disease or disorder, or a gut microbiome associated with or predictive of a disease or disorder in a patient, comprising the steps of:

- determining a gut microbiome profile from a biological sample obtained from said patient and comparing said profile to one or more gut microbiome reference profiles, wherein said one or more gut microbiome reference profiles comprise at least one of a positive gut microbiome reference profile based on results from control subjects with said disease or disorder or with a gut microbiome associated with or predictive of said disease or disorder, and a negative gut microbiome reference profile based on results from control subjects without said disease or disorder or without a gut microbiome associated with or predictive of said disease or disorder,
- if said gut microbiome profile for said patient statistically significantly matches said positive gut microbiome reference profile, then concluding that said patient has or is at risk of developing said disease or disorder or of a gut microbiome associated with or predictive of said disease or disorder in a patient; and/or
- if said gut microbiome profile for said patient statistically significantly matches said negative gut microbiome reference profile, then concluding that said patient does not have or is not at risk of developing said disease or disorder or does not have a gut microbiome associated with or predictive of said disease or disorder in a patient;

wherein the gut microbiome profile is determined by at least one of the biomarker panels described in the first aspect of the application.

In one embodiment, a positive gut microbiome reference profile is a gut microbiome reference profile from a subject with a Bacteroides_2 (B2) enterotype and a negative gut microbiome reference profile is a gut microbiome reference profile from a subject not having a Bacteroides_2 enterotype or alternatively phrased having a Bacteroides_1, Ruminococcus or Prevotella enterotype.

In one embodiment, said disease or disorder is gut flora dysbiosis and/or an inflammatory disorder. In another embodiment, said disease or disorder is obesity, diabetes type 2 or depression. In a particular embodiment, said inflammatory disorder is selected from the list consisting of spondyloarthritis, ankylosing spondylitis, reactive arthritis, psoriatic arthritis, enteropathic arthritis, undifferentiated spondyloarthritis, juvenile idiopathic arthritis, primary sclerosing cholangitis, multiple sclerosis, a gut inflammatory disorder, inflammatory bowel disease (IBD), Crohn's disease (CD), ulcerative colitis (UC), irritable bowel syndrome (IBS), celiac disease and any combination thereof and any gut inflammation associated with one of the above listed inflammatory disorders. In another particular embodiment, said inflammatory disorder is characterized by a TH1, TH17, TH2 and/or TH9 response.

In another embodiment, said gut microbiome profile is determined from a biological sample which can be a stool sample, a mucosal biopsy sample and/or a sample of the lumen content.
In another embodiment, said gut microbiome profile comprises an indication of the presence and/or abundance of at least 5 bacterial biomarkers selected from Table 5.

Other embodiments

The application also provides the methods herein described comprising the methods steps herein described, wherein the methods are methods of determining or detecting the Bacteroides_2 enterotype in a stool sample from a subject or methods of detecting or diagnosing in a subject a gut microbiome associated with or predictive for gut flora dysbiosis and/or an inflammatory disorder. Even more particular, said methods steps are also provided for methods of distinguishing or predicting or diagnosing different gut flora microbiomes, more particularly a gut flora microbiome associated with gut flora dysbiosis or inflammation, most particularly a Bacteroides_2 enterotype.

In an independent aspect, the present invention relates to a method of analysing abundances of at least two bacterial genera in a biological sample of a subject, the method comprising:

a) measuring in a biological, preferably a stool sample, obtained from the subject the abundances of at least two genera selected from a list consisting of *Oscillibacter, Faecalibacterium, Ruminococcus, Prevotella* and *Bacteroides* wherein said at least two genera do not consist of *Faecalibacterium* and *Bacteroides*;
b) diagnosing that the subject suffers from gut flora dysbiosis, if:
in case the at least two bacterial genera do not comprise Bacteroides, the abundances of each of the at least two bacterial genera as measured in step a) are decreased in the subject sample compared to the reference abundances; or

in case the at least two bacterial genera comprise Bacteroides, the abundances of Bacteroides as measured in step a) are increased in the subject sample compared to the reference abundances and the abundances of each of the other selected bacterial genera as measured in step a) are decreased in the subject sample compared to the reference abundances,

wherein said reference abundances of each of the at least two bacterial genera are obtained from a plurality of control samples.

In a preferred embodiment said step of measuring the abundances includes measuring of abundances of at least genera *Oscillibacter* and *Faecalibacterium*.

In another preferred embodiment, said step of measuring the abundances includes measuring of at least one further genus selected from a list consisting of *Ruminococcus, Prevotella*, and *Bacteroides*.

In another preferred embodiment said step of measuring the abundances includes measuring of at least two further genera selected from the list consisting of *Ruminococcus, Prevotella*, and *Bacteroides*.

The methods herein disclosed can thus also be of use of detecting or diagnosing a disease or disorder. In one embodiment, said disease or disorder is gut flora dysbiosis and/or an inflammatory disorder. In another embodiment, said disease or disorder is obesity, diabetes type 2 or depression. In a particular embodiment, said inflammatory disorder is selected from the list consisting of spondyloarthritis, ankylosing spondylitis, reactive arthritis, psoriatic arthritis, enteropathic arthritis, undifferentiated spondyloarthritis, juvenile idiopathic arthritis, primary sclerosing cholangitis, multiple sclerosis, a gut inflammatory disorder, inflammatory bowel disease (IBD), Crohn's disease (CD), ulcerative colitis (UC), irritable bowel syndrome (IBS), celiac disease and any combination thereof and any gut inflammation associated with one of the above listed inflammatory disorders. In another particular embodiment, said inflammatory disorder is characterized by a TH1, TH17, TH2 and/or TH9 response.

In one embodiment of all aspects and embodiments thereof, the control sample is representative of matched human subjects. More particularly, said control sample is a sample from a subject with a non-Bacteroides_2 enterotype or alternatively phrased a subject with a Bacteriodes_1, Prevotella or Ruminococcaceae enterotype. In a further embodiment, said control sample is a sample from a subject with a gut microbiome that is not associated with or predictive for gut flora dysbiosis and/or inflammatory disorder or obesity or diabetes type 2 or depression. In other further embodiments, said control sample is a negative control sample from a healthy individual, i.e. comparable individual not suffering from or diagnosed with gut flora dysbiosis and/or inflammatory disorders or obesity or diabetes type 2 or depression or a comparable individual not having an enterotype or a gut microbiome associated with or predictive for gut flora dysbiosis and/or inflammatory disorders.

In another embodiment of all aspects and embodiments thereof, the abundance of a bacterial genus in a control sample is equivalent to a reference abundance of said bacterial genus obtained from a plurality of control samples. In a particular embodiment, the reference abundance is determined by the average of the abundances of said bacterial genus in the plurality of samples obtained from a population of control subjects. In another particular embodiment, the reference abundance is a range determined by the abundances of said bacterial genus in the plurality of samples obtained from a population of control subjects with a lower border and a higher border.

In another aspect, methods of diagnosing and treating an inflammatory disorder in a subject are provided. Said methods comprise the steps from the method according t any embodiment of the first aspect of the invention and further comprising a step of administering an effective amount of anti-inflammatory drugs to the subject. This is equivalent as saying that methods are provided of diagnosing and treating an inflammatory disorder in a patient, comprising administering anti-inflammatory therapy to said patient if the gut microbiome profile for said patient statistically significantly matches that of a Bacteroides_2 enterotype. In a particular embodiment, said match is performed by using one of the biomarker panels from the application, more particularly described in the second aspect and embodiment thereof.

Diseases to be diagnosed

In further embodiments, said inflammatory disorder is selected from the list consisting of spondyloarthritis, ankylosing spondylitis, reactive arthritis, psoriatic arthritis, enteropathic arthritis, undifferentiated spondyloarthritis, juvenile idiopathic arthritis, primary sclerosing cholangitis, multiple sclerosis, a gut inflammatory disorder, inflammatory bowel disease (IBD), Crohn's disease (CD), ulcerative colitis (UC), irritable bowel syndrome (IBS), celiac disease and any combination thereof and any gut inflammation associated with one of the above listed inflammatory disorders.

As used herein, the term "spondyloarthritis" or abbreviated "SpA" refers to a group of closely related, but clinically heterogeneous, inflammatory arthritis diseases with common features, including inflammation of the spine, eyes, skin, joints and gastrointestinal tract. This SpA group is also sometimes referred to as spondylitis and spondyloarthropathies. As used herein, SpA includes ankylosing spondylitis (including non-radiographic axial SpA, i.e. ankylosing spondylitis diagnosed using MRI), reactive arthritis, psoriatic arthritis, enteropathic arthritis (arthritis associated with inflammatory bowel disease or IBD related arthritis), undifferentiated spondyloarthritis, juvenile idiopathic arthritis and juvenile-onset SpA. Characteristics of these SpA diseases include inflammatory arthritis of the spine, peripheral arthritis that

differs from rheumatoid arthritis, extra articular manifestations of inflammatory bowel disease, arthritis and uveitis, seronegativity for rheumatoid factor and some degree of heritability, including the presence of the gene HLA-B27. It is thus clear that in current application SpA is not rheumatoid arthritis.

"Primary sclerosing cholangitis" or "PSC" as used herein refers to a severe chronic liver disease characterized by progressive biliary inflammation and fibrosis. The development of multifocal bile duct structures can lead to liver fibrosis and subsequent cirrhosis. Patients with PSC are usually asymptomatic and the diagnostic work up is triggered by incidental findings of altered liver enzymes. In symptomatic patients, fatigue, pruritus, abdominal pain and jaundice are the most reported symptoms (Lazaridis et al 2016 N Engl J Med 375:1161-1170). Following clinical suspicion and a suggestive biochemistry, magnetic resonance cholangiography or endoscopic retrograde cholangiopancreatography are used to establish the diagnosis. Presently, liver biopsy is reserved to diagnose suspected small duct PSC or to exclude other diagnosis (Lindor et al 2015 Am J Gastroenterol 110:646-659). It would thus be highly advantageous to develop presymptomatic diagnostic methods or non-invasive diagnostic methods. The diagnostic methods disclosed above solve this technical problem. Therefore, in a particular embodiment, the herein disclosed methods are provided of diagnosing primary sclerosing cholangitis, more particularly gut inflammation associated with primary sclerosing cholangitis.

A systematic review of the epidemiologic studies in PSC reported an incidence varying between 0 and 1.3 cases per 100 000 individuals and a prevalence of 0-16.2 cases per 100 000 individuals (Boonstra et al 2012 J Hepatol 56:1181-1188). Most commonly, PSC affects men at the age of 40 and the concomitant diagnose of IBD is very common. Between 60 to 80% of the patients with PSC have concomitantly IBD, most frequently UC, pointing towards the possible role of the colon in the pathogenesis of PSC (Boonstra et al 2013 Hepatology 58:2045-2055). This role is further evidenced by transplantation data showing that colectomy before liver transplantation is a protective factor for recurrence of PSC after liver transplantation (Alabraba et al 2009 Liver Transpl 15:330-340). Interestingly, the absence of intestinal microbiota is associated with increased severity of the disease in mice models (Tabibian et al 2016 Hepatology 63:185-196). Therefore, intestinal microbiota may play an important role in the pathogenesis of PSC by modulating the gut-associated immune system to a more immunogenic or tolerogenic phenotype. In patients with IBD, the prevalence of PSC varies from 0.4 to 6.4%. However, in a recent study using magnetic resonance to diagnose PSC in patients with IBD the prevalence of PSC was 3-fold higher than previously reported, mainly due to subclinical PSC without symptoms or altered liver enzymes (Lunder et al 2016 Gastroenterology 151:660-669).

Genome-wide association studies suggested a role for immune-related pathways in the pathogenesis of PSC. Patients with PSC have a higher activity of TH17 cells. These lymphocytes help in the defence against

bacteria and fungi by promoting inflammation and are involved in autoimmune diseases (Katt et al 2013 Hepatology 58:1084-1093. Moreover, Treg cells (CD4+CD25+FOXP3+CD127-), which suppress inflammation, are reduced in PSC (Sebode et al 2014 J Hepatol 60:1010-1016). Therefore, in a very particular embodiment, the inflammatory disorder as mentioned in the application refers to inflammatory disorders characterized by a TH17 response.

"Multiple sclerosis" or "MS" as used herein refers to a chronic inflammatory and neurodegenerative disease characterized by substantial clinical heterogeneity. Both genetic and immunologic factors, as well as environmental elements contribute to its aetiology. Most MS patients present with recurrent periods of relapses and remissions, with relapses thought to be provoked by the infiltration of adaptive immune cells into the central nervous system (CNS), hereby resulting in focal inflammation and myelin loss (Franciotta et al 2008 Lancet neurology 7:852-588). In a minority of patients, slow progression is observed from onset. Therefore, three clinical phenotypes can be distinguished: relapsing-remitting (RR), secondary progressive (SP) or primary progressive (PP) MS. Lublin et al (2014 Neurology 83:278-286) further described these phenotypes as active, not active, and with or without progression. While not recognized as a separate phenotype, a subset of RRMS patients appears to have a mild course, often referred to as benign MS (BMS) (Amato et al 2006 J Neurol 253:1054-1059; Calabrese et al 2013 Mult Scler 19:904-911). Patients experience a wide variety of symptoms, ranging from physical and cognitive symptoms to even bowel dysfunction, with the latter being reported in more than 70% of cases (Wiesel et al 2001 Eur J Gastroenterol Hepatol 13:441-448). Studies in experimental allergic encephalomyelitis (EAE), a widely used mouse model for MS, have provided evidence for a substantial effect of gut microbiota on central nervous system (CNS)-specific autoimmune disease (Berer et al 2014 FEBS letters 588:4207-4013). The absence of gut microbes (germ-free conditions) or the alteration of the gut microbial flora composition with antibiotics resulted in a shift in T cell responses (decreased concentration of IL-17, increased number of regulatory T and B cells) and affected disease severity (Ochoa-Reparaz et al 2009 J Immunol 183:6041-6050). Additionally, mice raised in a germ-free environment were highly resistant to developing spontaneous EAE, unless exposed to specific pathogen-free condition-derived fecal material or a fecal transplant from MS twin-derived microbiota (Berer K et al 2011 Nature 479:538-541; Berer et al 2017 Proc Nat Ac Sc USA). Immune cells from mouse recipients of MS-twin samples produced less IL-10 than immune cells from mice colonized with healthy-twin samples. IL-10 may have a regulatory role in spontaneous CNS autoimmunity, as neutralization of the cytokine in mice colonized with healthy-twin fecal samples increased disease incidence. This evidence suggests that the microbiota may be capable of altering the individual at a phenotypic level and influence the onset, severity and progression of MS. Therefore, in a particular embodiment, the methods disclosed

herein are provided for detecting multiple sclerosis or gut inflammation associated with multiple sclerosis.

The wording “gut inflammation” is equivalent to the wording “microscopic gut inflammation” as used herein and refers to an inflammatory response in the gut as defined above. The inflammation can affect the entire gastrointestinal tract, can be more limited to for example the small intestine or large intestine but can also be limited to specific components or structures such as the bowel walls.

As used herein, the term “inflammatory bowel disease” or abbreviated “IBD” refers to an umbrella term for inflammatory conditions of the gut under which both Crohn's disease and ulcerative colitis fall. In people with IBD, the immune system mistakes food, bacteria, or other materials in the gut for foreign substances and responds by sending white blood cells into the lining of the bowels. The result of the immune system's attack is chronic inflammation. Crohn’s disease and ulcerative colitis are the most common forms of IBD. Less common IBDs include microscopic colitis, diverticulosis-associated colitis, collagenous colitis, lymphocytic colitis and Behçet's disease. In the case of CD, transmural inflammation commonly affects the terminal ileum, although any part of the gastrointestinal system can be affected. Discontinuous inflammation and the presence of non-caveating granulomas are also characteristic of the inflammation in patients with CD. In contrast, UC is characterized by continuous mucosal inflammation starting in the rectum and extending proximally until the caecum (Harries et al 1982 Br Med J Clin Res Ed, 284:706). These are chronic relapsing diseases originating mostly during adolescence and young adulthood and are characterized by chronic inflammation of the gastrointestinal tract leading to invalidating symptoms of bloody diarrhea, weight loss and fatigue (Wilks 1859 Med Times Gazette 2:264-265). Recent epidemiologic data from France reported a mean incidence of 4.4 cases per 100 000 individuals (Ghione et al 2017 Am J Gastroenterol). Worldwide, the incidence and prevalence of CD range from 0.0-29.3 per 100 000 person-years and 0.6-318.5 per 100 000 persons, respectively. The incidence and prevalence of UC varies from 0.0-19.2 per 100 000 person-years and 2.42-298.5 per 100 000 persons, respectively (Molodecky et al 2012 Gastroenterology 142:46-54).

Several defects in innate and adaptive immunity have been described both in UC and CD (de Souza et al 2016 Nat Rev Gastroenterol Hepatol 13:13-27). In normal conditions, intestinal macrophages exhibit inflammatory anergy which allows the interaction with commensal flora without inducing strong inflammatory responses (Smythies et al 2005 J Clin Invest 115:66-75). However, CD14+ intestinal macrophages are more abundant in patients with CD than in healthy individuals. These CD14+ intestinal macrophages produce more proinflammatory cytokines, such as interleukin(IL)-6, IL-23 and tumor necrosis factor (TNF)-a, than the common CD14-intestinal macrophages (Kamada et al 2008 J Clin Invest

118:2269-2280). Adaptive immunity also plays a role in the pathogenesis of IBD. T helper (TH) lymphocytes are cytokine producing lymphocytes that potentiate or regulate immune responses by interacting with other immune cells such as macrophages, CD8+ T cells, eosinophils and basophils. Following an initial trigger (e.g. impaired barrier function by injury or exposure to xenobiotics) the microbe-associated molecular patterns will induce the secretion of cytokines by dendritic cells, epithelial cells and macrophages, among others. Different cytokine milieus will induce TH1, TH2, TH17 or regulatory T-cell (Treg) subsets (de Souza et al 2016 Nat Rev Gastroenterol Hepatol 13:13-27). In susceptible individuals, an interplay between TH1 and TH17 immune responses seem to be linked with inflammation associated with CD. On the other hand, UC has been described as a TH2-like condition with possible implication of a newly discovered TH9 lymphocytes (de Souza et al 2016 Nat Rev Gastroenterol Hepatol 13:13-27; Gerlach et al 2014 Nat Immunol 15:676-686). In both diseases, an insufficient Treg response seems to be involved in the impaired regulation of inflammatory responses (Maul et al 2005 Gastroenterology 128:1868-1878). In active IBD, the immune system shows an increased response to bacterial stimulation, thereby contributing even further to the chronic inflammatory state. This inflammatory state also produces an increase in the intestinal permeability, allowing bacterial antigens to contact with the immune system, hereby perpetuating the inflammatory state.

In particular embodiments, said inflammation or inflammatory disorder as used in the methods of the fifth aspect is inflammation or an inflammatory disorder characterized by a TH1, TH17, TH2 and/or TH9 response. In even more particular embodiments, said inflammation or inflammatory disorder is characterized by a TH1 and/or TH17 response.

Methods of treatment

The therapeutic options of the inflammatory disorder diagnosed using the methods herein provided comprise the commonly used anti-inflammatory drugs such as inhibitors of cyclooxygenase activity (aspirin, celecoxib, diclofenac, diflunisal, etodolac, ibuprofen, indomethacin, ketoprofen, ketorolac, meloxicam, nabumetone, naproxen, oxaprozin, piroxicam, salsalate, sulindac, tolmetin, among others) or corticosteroids (prednisone, dexamethasone, hydrocortisone, methylprednisolone, among others) or in combination with commonly used analgesics (acetaminophen, duloxetine, paracetamol, among others) or in any combination thereof. In particular embodiments, said anti-inflammatory therapy includes a biological therapy, such as TNF-alpha blockers, anti-IL17A monoclonal antibodies, anti-CD20 antibodies.

The therapeutic options for CD or UC include corticosteroids, aminosalicylates, immunosuppressive agents and biological therapies. Due to the chronic relapsing and remitting disease-course of IBD, the goal of medical therapy is to induce (induction phase) and maintain remission (maintenance phase). The

choice between the different medical therapies depends on several factors such as disease location and severity, medical and surgical history, age, co-morbidities, extra-intestinal manifestations and treatment availability (Gomollon et al 2017 J Crohns Colitis 11:3-25; Harbord et al 2017 J Crohns Colitis 2017).

An "effective amount" of a composition is equivalent to the dosage of the composition that leads to treatment, prevention or a reduction of the severity of inflammation status in a patient. Said inflammation can be gut inflammation for which several methods are known to the person skilled in the art to evaluate or thus to diagnose the severity of the inflammation.

Recently, Vieira-Silva et al. (2020 Nature 581: 310-315) reported that a higher prevalence of the Bacteroides_2 enterotype correlates with a higher body-mass index and obesity. Interestingly, the pattern of enterotypes found in the population of obese individuals differed significantly depending on whether people were taking cholesterol-lowering drugs called statins. Obese participants taking statins had a significantly lower prevalence of the B2 enterotype than did their obese counterparts not taking statins.

Therefore, methods of diagnosing and treating gut flora dysbiosis are provided. Methods of diagnosing a gut microbiome associated with or predictive for gut flora dysbiosis and/or inflammatory disorder and changing said gut microbiome to a healthy or non-disease associated gut flora are also provided. Said methods comprise the steps from the methods according to any embodiment of the first aspect of the invention further comprising a step of administering an effective amount of a statin to the subject.

In one embodiment, said methods comprise the following steps:

- Measuring in a biological sample obtained from a subject the level of each of the bacterial genera from one of the biomarker panels described in the second aspect of the invention and embodiments thereof;
- Comparing the measured level of each of the bacterial genera of said subject sample to that of a control sample or to reference levels obtained from a plurality of control samples;
- Treating the subject with an effective amount of a statin when the measured level of each of the bacterial genera is statistically significantly decreased compared to the level of each of the bacterial genera in a control sample or compared to a reference level of each of the bacterial genera obtained from a plurality of control samples, except for *Flavonifractor* and *Bacteroides* that are predictive for glut flora dysbiosis when each of their measured level is statistically significantly increased compared to those of a control sample or to said reference levels.

In one embodiment, the biological sample is selected from the list consisting of a stool sample, a mucosal biopsy sample or a sample of the lumen content.

Statins, also known as HMG-CoA reductase inhibitors, are a class of lipid-lowering medications that are often prescribed to reduce illness and mortality in those who are at high risk of cardiovascular disease. Statins are the most common cholesterol-lowering drugs. Non-limiting examples of statins are lovastatin, fluvastatin, pravastin, rosuvastatin, pitavastatin, atorvastatin, simvastatin, cerivastatin, mevastatin.

An "effective amount" of a statin is equivalent to the dosage of the statin that leads to change in gut microbiome in a subject. Said change is a change from a B2 enterotype to a non-B2 enterotype or from a gut microbiome associated with gut flora dysbiosis and/or an inflammatory disorder to a healthy gut microbiome.

Bacterial quantification

Central in the methods of current application is the quantification of one or more bacterial genera. One of ordinary skill in the art knows several methods and devices for the quantification and analysis of the bacterial genera of the disclosure. The term "quantifying" refers to the ability to quantify the amount in a sample of one or more bacterial genera or of a nucleic acid sequence specific to a bacterial genus. Indeed, in most cases bacterial quantification in a sample is reduced to quantifying the amount of target nucleic acid sequences or target DNA that are specific for the bacteria that need to be quantified. Such molecular biology methods of detecting and quantifying target DNA are well known in the art.

In some embodiments, the step of quantifying one or more bacterial genera listed in Table 5 equals quantifying one or more target DNA sequences that are specific for said one or more bacterial genera. The target DNA can be the genes encoding the different 16S ribosomal RNA (the 16S rRNA gene) of the different bacterial genera listed in Table 5, or other genes or genomic sequences of interest possessed by said bacterial genera. In a particular embodiment, the bacterial quantification is carried out by determining the number of 16S rRNA gene copies of the one or more bacterial genera disclosed in current application. In another particular embodiment, said bacterial quantification includes a polynucleotide amplification step. A variety of polynucleotide amplification methods are well established and frequently used in research. In some embodiments, the amplification assay is a polymerase chain reaction (PCR) assay. The general methods of polymerase chain reaction (PCR) for polynucleotide sequence amplification are well known in the art and are thus not described in detail herein. For a review of PCR methods, protocols, and principles in designing primers, see, e.g., Innis, et al., PCR Protocols: A Guide to Methods and Applications, Academic Press, Inc. N.Y., 1990. PCR reagents and protocols are also available from commercial vendors, such as Roche Molecular Systems.

Non-limiting examples of more specific PCR techniques are end point-PCR, competitive PCR, reverse transcriptase-PCR (RT-PCR), quantitative PCR (qPCR) and reverse transcriptase qPCR (RT-qPCR). In particular embodiments, the quantification of the one or more bacterial genera according to the

methods herein described is performed by quantitative PCR assay (qPCR), also known as real-time PCR. The output of the quantification step using qPCR is the Ct value. The Ct (cycle threshold) value is defined as the number of qPCR cycles required for the fluorescent signal to cross the threshold. Ct levels are inversely proportional to the amount of target nucleic acid in the sample (i.e. the lower the Ct level the greater the amount of target nucleic acid in the sample). The abundances of the at least two bacterial genera can be expressed as the log cells/g. Alternatively, the abundances may be expressed as copies/µL or Cycle threshold (Ct) value. q-PCR technology overview and protocols are available from vendors such as from Sigma-Aldrich on SYBR Green qPCR applications, see for instance http://www.sigmaaldrich.com/technical-documents/protocols/biology/sybr-green-qpcr.html or http://www.sigmaaldrich.com/life-science/molecular-biology/pcr/quantitative-pcr/qpcr-technical-guide.html. For a review of qPCR methods to quantify the abundance and expression of bacterial gene markers see, e.g., Smith CJ and Osborn AM., FEMS Microbiol Ecol., 2009;67(1):6-20.

In other particular embodiments, the PCR assays are performed in multiplex. The term "multiplex-PCR" refers to a single PCR reaction carried out on nucleic acid obtained from a single source (e.g. a stool sample from a patient) using more than one primer set for the purpose of amplifying two or more DNA sequences in a single reaction.

Although PCR amplification is typically used in practicing the present invention, one of skill in the art will recognize that amplification of the relevant genomic sequence may be accomplished by any known method, such as the ligase chain reaction (LCR), transcription-mediated amplification, and self-sustained sequence replication or nucleic acid sequence-based amplification (NASBA), each of which provides sufficient amplification. More recently developed branched-DNA technology may also be used to quantitatively determining the amount of specific bacterial mRNA markers. For a detailed description of branched-DNA signal amplification for direct quantitation of nucleic acid sequences in clinical samples, see, for example, Nolte, Adv. Clin. Chem. 33: 201-235, 1998.

In some embodiments, the step of detecting and quantifying target DNA includes sequence-specific probe/primer hybridization, which can occur in the absence of polynucleotide amplification. Non-limiting examples of techniques to detect and quantify target DNA is PCR-ELISA, DNA microarrays, in situ hybridization assays such as dot-blot or Fluorescence In Situ Hybridization assay (FISH), 16S sequencing and to multiplex versions of said methods (see for instance, Andoh et al., Current Pharmaceutical Design, 2009; 15,2066-2073).

In one particular embodiment, the quantification step in the methods herein disclosed is performed by FISH, which combines probe hybridization with fluorescent light microscopy, confocal laser microscopy or flow cytometry for direct quantification of individual bacterial sequences. For reviews of FISH

methodology, see, e.g., Harmsen et al.,Appl Environ Microbiol, 2002;68 2982-2990,Kalliomaki et al., J AllergClinImmunol, 2001;107 129-134; Tkachuk et al., Genet. Anal. Tech. Appl., 1991;8:67-74; Trask et al., Trends Genet., 1991;7 (5): 149-154; and Weier et al., Expert Rev. Mol. Diagn., 2002,2(2):109-119; and U.S. Pat. No. 6,174,681.

Quantification of the abundance of a target nucleic acid sequence within a stool sample might be absolute or relative. “Relative quantification” is generally based on one or more internal reference genes, i.e., 16S rRNA genes from reference strains, such as determination of total bacteria using universal primers and expressing the abundance of the target nucleic acid sequence as a percentage of total bacterial 16S rRNA gene copies or normalized by E.coli 16S rRNA gene copies. "Absolute quantification" gives the exact number of target molecules by comparison with DNA standards or normalizing by DNA concentration.

Further preferred is a method in which the one or more bacterial genera or the total number of target gene copies are quantified in an absolute manner by using an external quantification standard and universal primer pairs that are added to a quantitative real-time PCR. The total number of marker gene copies in the mixture of sample and standards is then determined for example by quantitative real-time PCRs.

In particular embodiments of the invention, the methods according to first aspect and any embodiment thereof as herein disclosed, are provided, wherein the quantification of the at least two bacterial genera is performed by quantifying DNA sequences specific for said at least two bacterial genera. In more particular embodiment, said DNA sequences are 16S rDNA sequences or a fragment thereof. Non-limiting examples of 16S rDNA sequences specific for the bacterial genera disclosed herein are SEQ ID No. 1-68 in Table 1. In other more particular embodiments, the quantification of the DNA sequences is performed by quantitative PCR. In the latter case, the abundances of the at least two bacterial genera are expressed as the log cells/g. Alternatively, the abundances may be expressed as copies/µL or Cycle threshold (Ct) value.

In another particular embodiment of the methods of the invention, prior to the quantification of the one or more bacterial genera and thus of the one or more bacterial sequences which are specific for said one or more bacterial genera, DNA is extracted from the stool sample. There are numerous methods for extracting bacterial DNA from a stool sample and are routinely practiced in the art of molecular biology. These methods rely on chemical or mechanical disruption, lysis using detergents, or a combination of these approaches (Kennedy A. et al., PLoS One, 2014;9{2):e88982). Methods for extraction of bacterial DNA in fecal samples are known for example from Corist et al. (2002 Journal of Microbiological Methods 50; Whitney et al. 2004 Journal of Molecular Diagnostics, American Society for Investigative Pathology,

6; and WO2003/068788). In some embodiments, the extraction methods use a combination of mechanical disruption, such as high speed bead beating extraction, chemical lysis and a final purification step, for example using silica membrane columns such as those included in the commercially available DNA extraction kits "MobioPowerSoil DNA extraction procedure" (Mo-Bio Laboratories Inc.), FastDNA SPIN Kit for soil procedure (MP biomedicals), RNeasy PowerMicrobiome Kit – Stool and Gut RNA Extraction" (Qiagen), and NucteoSpin Soil (Macherey-Nagel Gmbh& Co. KG). As the presence of PGR inhibitors in the DNA extracts from faecal samples such as bilirubins, bile salts and complex carbohydrates is one of the difficulties faced for the determination of DNA biomarkers in DNA extracts from feces (Fleckna et al. 2007 Mol Cell Probes 21), preferably DNA extraction methods are used that provide fecal extracts with a low amount of PCR inhibitors, such as less than 5%, preferably less than 2%, more preferably less than 1 %, even more preferably less than 0.5%, such as less than 0.25%, 0.1 %, 0.05% or 0.01 %.

Kits for diagnosing gut flora dysbiosis

As described in the previous section, the quantification of the bacterial genera as one of the steps in the diagnostic methods of the application can be performed by quantifying DNA sequences which are specific for said bacterial genera. Therefore, current application also provides a kit comprising:

- at least two pairs of nucleic acid primers, each pair of primers being designed to specifically amplify a DNA sequence of a different bacterial genus selected from the list consisting of *Sporobacter*, *Coprococcus*, *Fusicatenibacter*, *Ruminococcus*, *Oscillibacter* and *Faecalibacterium*; or selected from the list consisting of *Sporobacter*, *Coprococcus*, *Fusicatenibacter*, *Ruminococcus*, *Oscillibacter* and *Faecalibacterium* and/or from the list consisting of *Clostridium_IV*, *Butyricimonas*, *Butyricicoccus*, *Prevotella*, *Roseburia* and *Flavonifractor*, ; or selected from the list consisting of *Sporobacter*, *Coprococcus*, *Fusicatenibacter*, *Ruminococcus* and *Oscillibacter* and/or from the list consisting of *Clostridium_IV*, *Butyricimonas*, *Butyricicoccus*, *Prevotella*, *Roseburia*, *Flavonifractor* and *Bacteroides*; and
- preferably instructions for quantifying the levels of two or more DNA sequences from a stool sample.

In a preferred embodiment, said at least two pairs of primers are designed to specifically amplify a DNA sequence of bacterial genera *Oscillibacter and Faecalibacterium.* In one embodiment, said at kit comprises at least 3 or at least 4 pairs of nucleic acid primers, wherein at least two primers thereof are designed to specifically amplify a DNA sequence of bacterial genera *Oscillibacter* and *Faecalibacterium*. In a particular embodiment, the kit of the invention comprises at least five pairs of primers designed to specifically amplify a DNA sequence of bacterial genera *Ruminococcus*, *Oscillibacter*, *Faecalibacterium Prevotella* and *Bacteroides*.

In another embodiment, said at least two pairs of nucleic acid primers are at least 5 pairs of nucleic acid primers capable of specifically amplifying DNA sequences of at least 5 different bacterial genera selected from the list consisting of *Sporobacter*, *Coprococcus*, *Fusicatenibacter*, *Ruminococcus*, *Oscillibacter*, *Faecalibacterium*, *Clostridium_IV*, *Butyricimonas*, *Butyricicoccus*, *Prevotella*, *Roseburia*, *Bacteroides* and *Flavonifractor*. In a particular embodiment, said DNA sequences are 16S rDNA sequences or fragments thereof. In an even more particular embodiment, said DNA sequences are selected from SEQ ID No. 1-68 or a fragment thereof. Also provided herein, are the above described kits for diagnosing or early detecting gut flora dysbiosis.

The following examples are intended to promote a further understanding of the invention. While the invention is described herein with reference to illustrated embodiments, it should be understood that the invention is not limited hereto. Those having ordinary skill in the art and access to the teachings herein will recognize additional modifications and embodiments within the scope thereof. Therefore, the invention is limited only by the claims attached herein.

EXAMPLES

Example 1: Data preparation

Quantitative microbial profiles (QMP) were obtained from a full set of 2919 samples from the FGFP, processed as described in Vandeputte et al. (2017 Nature) and were agglomerated at the genus level using DADA2 after identifying each read's taxonomy using RDP. From the 2919 samples where 296 microbial features were recorded, features which were definable on the genus level were retained which were present in at least half of the samples. This resulted in an abundance matrix containing 44 genera. Enterotypes (Vandeputte et al 2017 Nature) were assigned to each of the samples using a DMM based approach, resulting in 518 Bacteroides_2 samples, with other samples either assigned to the Bacteroides_1 (1054 samples), Prevotella (515 samples) or Ruminococcus (911 samples) enterotype.

Example 2: Predicting the Bacteroides 2 enterotype from QMP data

Relevant genera were selected using a custom-pipeline implemented in Python (Python version 3.7.4) (as described below). First, the dataset was randomly under sampled (package imblearn 0.8.1) retaining an equal set of samples from subjects with the B2 enterotype as samples from subjects with another enterotype. The remaining set of 994 samples was divided into training and testing data at an 80-20 ratio as is customary in machine learning.

Microbial features were scaled so each feature would have a mean of zero and a standard deviation of zero (scikit-learn 1.0.1, StandardScaler with default parameters). After scaling the training data, features relevant for the prediction of the B2 enterotype were picked up using feature selection (SelectFdr with

default parameters) resulting in a final selection (using an ANOVA based selection combined with FDR correction, corrected p-value < 0.05) of 13 genera: *Roseburia*, *Fusicatenibacter*, *Butyricicoccus*, *Flavonifractor*, *Coprococcus*, *Faecalibacterium*, *Clostridium_IV*, *Oscillibacter*, *Sporobacter*, *Ruminococcus*, *Bacteroides*, *Butyricimonas* and *Prevotella*.

A random forest classifier with optimized settings (grid search, 5x cross validation using GridSearchCV) was trained on the training dataset to predict the enterotype from the quantitative abundances from selected genera. The optimized model was found to require 100 estimators using the entropy criterion. By applying this model onto the test set, which was not included during training, the performance of the model was measured using Receiver Operator Characteristics (ROC) as well as a confusion matrix. Here an area under the curve (AUC) of 0.97 was obtained, which is considered outstanding (Figure 1a). The confusion matrix shows additional details on the prediction errors (Figure 1b). The average precision of the model is 0.90 with 0.89 recall. Considering the goal of the classifier is to detect samples with B2, the sensitivity (recall of B2) is 0.83 with a specificity (recall of non-B2) of 0.95.

Example 3: Importance of individual genera for B2 predictions

Next, an initial impression of the genera's importance within the prediction was obtained using SHAP, which will assess each genus' impact on the model and if there is a positive correlation with B2 or not. From the SHAP output, shown in Figure 2, it can be concluded that high values of *Ruminococcus*, *Faecalibacterium*, *Sporobacter*, *Coprococcus*, *Clostridium_IV*, *Oscillibacter* and *Prevotella*, are indicative for a non-B2, while high *Flavonifractor* and *Bacteroides* levels are driving predictions towards B2.

Example 4: Design and evaluation of the B2 Score

A Bayesian linear regression model, which assigns a weight to each of the 13 selected genera was implemented in PYMC3 (version 3.11.2) and trained to produce a low score for non-B2 samples, and a high score for B2 samples. After sampling (default settings, 4 chains, 1000 tuning and 1000 sampling steps) the average weight for each genus, the standard deviation and the 94% highest density interval (HDI, generated using ArViz 0.11.2) were extracted from the model. The model works as intended (shown in Figure 3), generating a higher score for B2 samples compared to non-B2 samples. In this model a weight of zero would indicate a genus had no effect on the B2 predictions, a positive weight would indicate the genus is more abundant in B2 while a negative weight indicates a reduction of a genus' abundance in B2. By evaluating the weights assigned to the genera in the model, and their HDI, it can be concluded with 94% certainty that genera where 0 is outside this interval have a meaning full effect on the model. As shown in Table 8 there are 11 genera that have an impact on the final score.

Table 8 Weights in the Bayesian linear regression model along with their 94% Highest Density Interval (HDI). For 13 of the genera zero is outside of the interval indicating with confidence these have an impact on the score.

Genus	Weight	sd	hdi_3%	hdi_97%
Faecalibacterium	-0.0745	0.0082	-0.0903	-0.0595
Ruminococcus	-0.038	0.0074	-0.0519	-0.0238
Sporobacter	-0.0377	0.0081	-0.0541	-0.0234
Fusicatenibacter	-0.0301	0.0073	-0.0441	-0.0167
Oscillibacter	-0.0283	0.0077	-0.0424	-0.013
Clostridium_IV	-0.0231	0.0065	-0.0353	-0.0111
Butyricimonas	-0.0222	0.0065	-0.0343	-0.0102
Coprococcus	-0.0203	0.0068	-0.0338	-0.0078
Prevotella	-0.0141	0.0063	-0.0269	-0.0032
Roseburia	-0.0102	0.0071	-0.0239	0.0026
Butyricicoccus	-0.0102	0.007	-0.0228	0.0032
Flavonifractor	0.0665	0.0067	0.0542	0.0792
Bacteroides	0.089	0.0075	0.0743	0.1026

Example 5. B2 predictions with one genus

As 13 genera is substantial to measure within a qPCR experiment, the performance of models with fewer genera was assessed. First, it was assessed whether the abundance of single genera would be sufficient to diagnose the dysbiotic B2 enterotype. Surprisingly, it was found that 6 genera on their own were predictive for B2 with an AUC of 0.7 or more (Table 9).

Table 9. List of genera that on their own are predictive for B2 with an AUC of at least 0.7.

Genus	AUC_mean	AUC_std
Sporobacter	0.774	0.002
Coprococcus	0.728	0.005
Fusicatenibacter	0.725	0.014
Ruminococcus	0.722	0.009
Faecalibacterium	0.714	0.011
Oscillibacter	0.705	0.015

Example 6. Exhaustive search of useable combinations

From Table 9 it is clear that six genera (*Sporobacter, Coprococcus, Ruminococcus, Faecalibacterium, Oscillibacter*, and *Fusicatenibacter*) are sufficient on their own to get a Bacteroides_2 prediction with a ROC AUC larger than 0.7. As any combination with these single predictors will yield a result as good or better than the single genus, the overwhelming majority (1011 out of 1023) of the possible combinations considered here also had a ROC AUC of 0.7 or more. Additionally, there are 15 combinations without the

single predictors, of which 6 combinations are sufficient to get a ROC AUC 0.7 or higher. These are *Clostridium_IV* and *Prevotella*, *Clostridium_IV* and *Bacteroides*, *Clostridium_IV* and *Flavonifractor* and *Bacteroides*, *Clostridium_IV* and *Bacteroides* and *Prevotella*, *Clostridium_IV* and *Flavonifractor* and *Prevotella*, and *Clostridium_IV* and *Flavonifractor* and *Bacteroides* and *Prevotella*.

Table 10 below lists all combinations of two of the 13 bacterial genera disclosed herein that predict the B2 enterotype with and ROC AUC of 0.7 or more.

Table 10. Combinations of 2 bacterial genera with an AUC of at least 0.7. *Coprococcus*, Cop; *Faecalibacterium*, Fae; *Oscillibacter*, Osc; *Ruminococcus*, Rum; *Sporobacter*, Spo; *Fusicatenibacter*, Fus; *Clostridium_IV*, Clo; *Butyricimonas*, Bmo; *Butyricicoccus*, Bco; *Prevotella*, Pre; *Flavonifractor*, Fla; *Bacteroides*, Bac; *Roseburia*, Ros.

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Cop,Fae	0.780	0.009	Cop,Bac	0.757	0.006
Cop,Osc	0.752	0.010	Cop,Bmo	0.725	0.008
Cop,Rum	0.804	0.010	Cop,Clo	0.776	0.007
Cop,Spo	0.788	0.005	Clo,Rum	0.771	0.009
Fae,Osc	0.763	0.012	Clo,Spo	0.733	0.006
Fae,Rum	0.807	0.007	Bco,Osc	0.706	0.009
Fae,Spo	0.800	0.007	Bco,Rum	0.759	0.012
Fus,Cop	0.751	0.007	Bco,Spo	0.753	0.007
Fus,Fae	0.767	0.008	Cop,Pre	0.732	0.009
Fus,Osc	0.759	0.009	Fae,Bac	0.792	0.008
Fus,Rum	0.790	0.011	Fae,Bmo	0.724	0.009
Fus,Spo	0.809	0.006	Fae,Clo	0.799	0.008
Osc,Rum	0.760	0.008	Fae,Pre	0.752	0.009
Osc,Spo	0.748	0.009	Fus,Bac	0.748	0.009
Spo,Rum	0.804	0.013	Fus,Bmo	0.710	0.008
			Fus,Clo	0.771	0.011
Fla,Cop	0.708	0.007	Fus,Pre	0.764	0.007
Fla,Fae	0.786	0.009	Osc,Pre	0.741	0.008
Fla,Osc	0.748	0.008	Rum,Bac	0.788	0.008
Fla,Rum	0.764	0.009	Rum,Bmo	0.746	0.011
Fla,Spo	0.755	0.010	Rum,Pre	0.787	0.005
Ros,Cop	0.722	0.008	Spo,Bac	0.732	0.007
Ros,Fae	0.760	0.007	Spo,Bmo	0.793	0.007
Ros,Osc	0.702	0.010	Spo,Pre	0.756	0.008
Ros,Rum	0.736	0.004			
Ros,Spo	0.801	0.008	Bco,Clo	0.709	0.007
			Clo,Bac	0.737	0.013
			Clo,Bmo	0.715	0.010
			Clo,Pre	0.719	0.007

Table 11 lists all combinations of three of the 13 bacterial genera herein disclosed that predict the B2 enterotype with a ROC AUC of 0.7 or more.

Table 11. Combinations of 3 bacterial genera with an AUC of at least 0.7. *Coprococcus*, Cop; *Faecalibacterium*, Fae; *Oscillibacter*, Osc; *Ruminococcus*, Rum; *Sporobacter*, Spo; *Fusicatenibacter*, Fus; *Clostridium_IV*, Clo; *Butyricimonas*, Bmo; *Butyricicoccus*, Bco; *Prevotella*, Pre; *Flavonifractor*, Fla; *Bacteroides*, Bac; *Roseburia*, Ros.

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Bco,Clo,Osc	0.752	0.007	Bco,Spo,Rum	0.799	0.008
Bco,Clo,Spo	0.754	0.009	Bco,Cop,Fae	0.763	0.006
Bco,Cop,Bac	0.755	0.009	Bco,Cop,Osc	0.738	0.009
Bco,Cop,Clo	0.757	0.007	Bco,Cop,Rum	0.802	0.007
Bco,Fae,Bac	0.789	0.007	Bco,Cop,Spo	0.780	0.006
Bco,Fae,Bmo	0.701	0.009	Bco,Fae,Osc	0.778	0.013
Bco,Fae,Clo	0.781	0.007	Bco,Fae,Rum	0.799	0.009
Bco,Fae,Pre	0.764	0.005	Bco,Fae,Spo	0.804	0.005
Bco,Fla,Cop	0.737	0.007	Bco,Osc,Rum	0.796	0.009
Bco,Fla,Fae	0.779	0.005	Bco,Osc,Spo	0.775	0.008
Bco,Fla,Osc	0.758	0.008	Clo,Osc,Rum	0.797	0.008
Bco,Fla,Rum	0.786	0.007	Clo,Osc,Spo	0.756	0.007
Bco,Fla,Spo	0.793	0.007	Clo,Spo,Rum	0.792	0.006
Bco,Osc,Bac	0.752	0.006	Cop,Clo,Osc	0.793	0.006
Bco,Osc,Pre	0.794	0.008	Cop,Clo,Rum	0.819	0.004
Bco,Rum,Bac	0.809	0.007	Cop,Clo,Spo	0.791	0.006
Bco,Rum,Bmo	0.746	0.008	Cop,Fae,Bac	0.828	0.005
Bco,Rum,Pre	0.807	0.011	Cop,Fae,Bmo	0.782	0.006
Bco,Spo,Bac	0.791	0.004	Cop,Fae,Clo	0.822	0.008
Bco,Spo,Bmo	0.753	0.012	Cop,Fae,Pre	0.801	0.005
Bco,Spo,Pre	0.764	0.006	Cop,Osc,Bac	0.787	0.009
Clo,Osc,Bac	0.766	0.007	Cop,Osc,Bmo	0.763	0.007
Clo,Osc,Bmo	0.715	0.011	Cop,Osc,Pre	0.747	0.009
Clo,Osc,Pre	0.749	0.005	Cop,Rum,Bac	0.855	0.005
Bco,Clo,Rum	0.810	0.008	Cop,Rum,Bmo	0.811	0.008
Clo,Rum,Bac	0.844	0.006	Cop,Rum,Pre	0.830	0.006
Clo,Rum,Bmo	0.795	0.005	Cop,Spo,Bac	0.804	0.005
Clo,Rum,Pre	0.812	0.009	Cop,Spo,Bmo	0.799	0.007
Clo,Spo,Bac	0.774	0.008	Cop,Spo,Pre	0.784	0.006
Clo,Spo,Bmo	0.773	0.006	Fae,Clo,Osc	0.792	0.010
Clo,Spo,Pre	0.763	0.010	Fae,Clo,Rum	0.841	0.007
Cop,Bac,Bmo	0.760	0.006	Fae,Clo,Spo	0.824	0.003
Cop,Bac,Pre	0.759	0.005	Fae,Osc,Bac	0.853	0.007
Cop,Bmo,Pre	0.740	0.006	Fae,Osc,Bmo	0.766	0.006
Cop,Clo,Bac	0.821	0.006	Fae,Osc,Pre	0.827	0.010
Cop,Clo,Bmo	0.799	0.004	Fae,Rum,Bac	0.877	0.008

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Cop,Clo,Pre	0.777	0.006	Fae,Rum,Bmo	0.795	0.011
Fae,Bac,Bmo	0.785	0.006	Fae,Rum,Pre	0.856	0.004
Fae,Bac,Pre	0.831	0.007	Fae,Spo,Bac	0.873	0.005
Fae,Bmo,Pre	0.762	0.008	Fae,Spo,Bmo	0.798	0.007
Fae,Clo,Bac	0.854	0.006	Fae,Spo,Pre	0.830	0.009
Fae,Clo,Bmo	0.792	0.008	Fla,Cop,Fae	0.814	0.007
Fae,Clo,Pre	0.827	0.010	Fla,Cop,Osc	0.806	0.007
Fla,Clo,Osc	0.761	0.004	Fla,Cop,Rum	0.808	0.005
Fla,Clo,Rum	0.798	0.005	Fla,Cop,Spo	0.796	0.007
Fla,Clo,Spo	0.769	0.005	Fla,Fae,Osc	0.851	0.007
Fla,Cop,Bac	0.742	0.007	Fla,Fae,Rum	0.856	0.006
Fla,Cop,Bmo	0.742	0.007	Fla,Fae,Spo	0.857	0.006
Fla,Cop,Clo	0.774	0.005	Fla,Osc,Rum	0.799	0.005
Fla,Cop,Pre	0.731	0.007	Fla,Osc,Spo	0.790	0.007
Fla,Fae,Bac	0.811	0.009	Fla,Spo,Rum	0.826	0.007
Fla,Fae,Bmo	0.780	0.007	Fus,Bco,Cop	0.753	0.004
Fla,Fae,Clo	0.835	0.008	Fus,Bco,Fae	0.775	0.004
Fla,Fae,Pre	0.823	0.004	Fus,Bco,Osc	0.759	0.006
Fla,Osc,Bac	0.751	0.007	Fus,Bco,Rum	0.799	0.008
Fla,Osc,Bmo	0.752	0.008	Fus,Bco,Spo	0.823	0.005
Fla,Osc,Pre	0.789	0.006	Fus,Clo,Osc	0.816	0.003
Fla,Rum,Bac	0.800	0.006	Fus,Clo,Rum	0.842	0.005
Fla,Rum,Bmo	0.776	0.010	Fus,Clo,Spo	0.826	0.006
Fla,Rum,Pre	0.825	0.006	Fus,Cop,Bac	0.832	0.004
Fla,Spo,Bac	0.781	0.008	Fus,Cop,Bmo	0.779	0.007
Fla,Spo,Bmo	0.802	0.005	Fus,Cop,Clo	0.811	0.004
Fla,Spo,Pre	0.778	0.010	Fus,Cop,Pre	0.787	0.009
Fus,Bac,Bmo	0.755	0.009	Fus,Fae,Bac	0.853	0.006
Fus,Bac,Pre	0.801	0.008	Fus,Fae,Bmo	0.749	0.005
Fus,Bco,Bac	0.759	0.009	Fus,Fae,Clo	0.838	0.006
Fus,Bco,Bmo	0.720	0.004	Fus,Fae,Pre	0.815	0.007
Fus,Bco,Clo	0.768	0.004	Fus,Fla,Cop	0.766	0.008
Fus,Bco,Fla	0.721	0.006	Fus,Fla,Fae	0.806	0.004
Fus,Bco,Pre	0.769	0.010	Fus,Fla,Osc	0.807	0.007
Fus,Bmo,Pre	0.752	0.010	Fus,Fla,Rum	0.817	0.007
Fus,Clo,Bac	0.856	0.011	Fus,Fla,Spo	0.822	0.005
Fus,Clo,Bmo	0.797	0.006	Fus,Osc,Bac	0.814	0.007
Fus,Clo,Pre	0.827	0.005	Fus,Osc,Bmo	0.756	0.005
Fus,Fla,Bac	0.749	0.010	Fus,Osc,Pre	0.817	0.009
Fus,Fla,Bmo	0.726	0.009	Fus,Rum,Bac	0.868	0.008
Fus,Fla,Clo	0.797	0.007	Fus,Rum,Bmo	0.819	0.005
Fus,Fla,Pre	0.745	0.009	Fus,Rum,Pre	0.831	0.008
Osc,Bac,Bmo	0.713	0.008	Fus,Spo,Bac	0.846	0.007
Osc,Bac,Pre	0.732	0.008	Fus,Spo,Bmo	0.826	0.007
Ros,Bco,Cop	0.707	0.007	Fus,Spo,Pre	0.836	0.006

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Ros,Bco,Fae	0.774	0.006	Osc,Rum,Bac	0.833	0.005
Ros,Bco,Osc	0.729	0.006	Osc,Rum,Bmo	0.765	0.008
Ros,Bco,Rum	0.751	0.011	Osc,Rum,Pre	0.796	0.006
Ros,Bco,Spo	0.789	0.005	Osc,Spo,Bac	0.768	0.008
Ros,Clo,Osc	0.761	0.008	Osc,Spo,Bmo	0.764	0.009
Ros,Clo,Rum	0.791	0.006	Osc,Spo,Pre	0.749	0.015
Ros,Clo,Spo	0.774	0.007	Ros,Cop,Fae	0.801	0.006
Ros,Cop,Bac	0.789	0.005	Ros,Cop,Osc	0.768	0.007
Ros,Cop,Bmo	0.765	0.007	Ros,Cop,Rum	0.812	0.009
Ros,Cop,Clo	0.778	0.005	Ros,Cop,Spo	0.823	0.006
Ros,Cop,Pre	0.753	0.007	Ros,Fae,Osc	0.814	0.009
Ros,Fae,Bac	0.824	0.010	Ros,Fae,Rum	0.846	0.007
Ros,Fae,Bmo	0.784	0.010	Ros,Fae,Spo	0.852	0.005
Ros,Fae,Clo	0.813	0.005	Ros,Fus,Cop	0.761	0.008
Ros,Fae,Pre	0.806	0.007	Ros,Fus,Fae	0.807	0.010
Ros,Fla,Clo	0.770	0.008	Ros,Fus,Osc	0.755	0.009
Ros,Fla,Cop	0.735	0.009	Ros,Fus,Rum	0.794	0.011
Ros,Fla,Fae	0.796	0.007	Ros,Fus,Spo	0.834	0.006
Ros,Fla,Osc	0.769	0.006	Ros,Osc,Rum	0.788	0.007
Ros,Fla,Rum	0.769	0.008	Ros,Osc,Spo	0.787	0.003
Ros,Fla,Spo	0.799	0.006	Ros,Spo,Rum	0.818	0.007
Ros,Fus,Bac	0.771	0.008	Spo,Rum,Bac	0.835	0.007
Ros,Fus,Bco	0.709	0.008	Spo,Rum,Bmo	0.783	0.006
Ros,Fus,Bmo	0.754	0.012	Spo,Rum,Pre	0.816	0.007
Ros,Fus,Clo	0.789	0.007			
Ros,Fus,Pre	0.742	0.008	Cop,Fae,Osc	0.822	0.006
Ros,Osc,Bac	0.769	0.007	Cop,Fae,Rum	0.857	0.004
Ros,Osc,Bmo	0.714	0.008	Cop,Fae,Spo	0.833	0.007
Ros,Osc,Pre	0.754	0.009	Cop,Osc,Rum	0.832	0.006
Ros,Rum,Bac	0.822	0.005	Cop,Osc,Spo	0.788	0.006
Ros,Rum,Bmo	0.792	0.008	Cop,Spo,Rum	0.829	0.006
Ros,Rum,Pre	0.802	0.005	Fae,Osc,Rum	0.836	0.009
Ros,Spo,Bac	0.828	0.006	Fae,Osc,Spo	0.824	0.007
Ros,Spo,Bmo	0.819	0.009	Fae,Spo,Rum	0.838	0.005
Ros,Spo,Pre	0.801	0.011	Fus,Cop,Fae	0.812	0.005
Rum,Bac,Bmo	0.826	0.006	Fus,Cop,Osc	0.799	0.006
Rum,Bac,Pre	0.846	0.006	Fus,Cop,Rum	0.837	0.006
Rum,Bmo,Pre	0.764	0.004	Fus,Cop,Spo	0.824	0.006
Spo,Bac,Bmo	0.746	0.010	Fus,Fae,Osc	0.812	0.006
Spo,Bac,Pre	0.756	0.009	Fus,Fae,Rum	0.833	0.005
Spo,Bmo,Pre	0.778	0.009	Fus,Fae,Spo	0.845	0.006
Bco,Cop,Pre	0.732	0.010	Fus,Osc,Rum	0.834	0.008
			Fus,Osc,Spo	0.820	0.006
Bco,Fla,Clo	0.764	0.006	Fus,Spo,Rum	0.847	0.005
Bco,Fla,Pre	0.708	0.011	Osc,Spo,Rum	0.804	0.007

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Fla,Clo,Bac	0.759	0.010			
Fla,Clo,Bmo	0.743	0.011	Bco,Clo,Bac	0.776	0.007
Fla,Clo,Pre	0.731	0.011	Bco,Clo,Bmo	0.725	0.008
Ros,Bac,Bmo	0.729	0.007	Bco,Clo,Pre	0.756	0.010
Ros,Bco,Clo	0.737	0.005	Clo,Bac,Bmo	0.752	0.011
Ros,Bco,Pre	0.704	0.012	Clo,Bac,Pre	0.778	0.008
Ros,Bmo,Pre	0.711	0.009	Clo,Bmo,Pre	0.729	0.008
Ros,Clo,Bac	0.789	0.006			
Ros,Clo,Bmo	0.751	0.011			
Ros,Clo,Pre	0.761	0.011			

Table 12 lists all combinations of four of the 13 bacterial genera herein disclosed that predict the B2 enterotype with a ROC AUC of 0.7 or more.

Table 12. Combinations of 4 bacterial genera with an AUC of at least 0.7. *Coprococcus*, Cop; *Faecalibacterium*, Fae; *Oscillibacter*, Osc; *Ruminococcus*, Rum; *Sporobacter*, Spo; *Fusicatenibacter*, Fus; *Clostridium_IV*, Clo; *Butyricimonas*, Bmo; *Butyricicoccus*, Bco; *Prevotella*, Pre; *Flavonifractor*, Fla; *Bacteroides*, Bac; *Roseburia*, Ros.

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Bco,Bac,Bmo,Pre	0.712	0.013	Fus,Bco,Rum,Bmo	0.808	0.005
Bco,Clo,Bac,Bmo	0.781	0.008	Fus,Bco,Rum,Pre	0.843	0.009
Bco,Clo,Bac,Pre	0.813	0.008	Fus,Bco,Spo,Bac	0.846	0.005
Bco,Clo,Bmo,Pre	0.789	0.009	Fus,Bco,Spo,Bmo	0.826	0.007
Clo,Bac,Bmo,Pre	0.768	0.008	Fus,Bco,Spo,Pre	0.841	0.007
Bco,Clo,Osc,Bac	0.809	0.005	Fus,Bco,Spo,Rum	0.852	0.005
Bco,Clo,Osc,Bmo	0.757	0.007	Fus,Clo,Bac,Bmo	0.843	0.004
Bco,Clo,Osc,Pre	0.802	0.008	Fus,Clo,Bac,Pre	0.868	0.007
Bco,Clo,Rum,Bac	0.849	0.005	Fus,Clo,Bmo,Pre	0.838	0.007
Bco,Clo,Rum,Bmo	0.806	0.004	Fus,Clo,Osc,Bac	0.867	0.005
Bco,Clo,Rum,Pre	0.833	0.006	Fus,Clo,Osc,Bmo	0.820	0.006
Bco,Clo,Spo,Bac	0.799	0.005	Fus,Clo,Osc,Pre	0.847	0.004
Bco,Clo,Spo,Bmo	0.780	0.005	Fus,Clo,Osc,Rum	0.858	0.005
Bco,Clo,Spo,Pre	0.779	0.006	Fus,Clo,Osc,Spo	0.847	0.004
Bco,Cop,Bac,Bmo	0.775	0.010	Fus,Clo,Rum,Bac	0.892	0.004
Bco,Cop,Bac,Pre	0.785	0.005	Fus,Clo,Rum,Bmo	0.854	0.003
Bco,Cop,Bmo,Pre	0.754	0.006	Fus,Clo,Rum,Pre	0.868	0.007
Bco,Cop,Clo,Bac	0.814	0.006	Fus,Clo,Spo,Bac	0.856	0.004
Bco,Cop,Clo,Bmo	0.780	0.004	Fus,Clo,Spo,Bmo	0.845	0.006
Bco,Cop,Clo,Pre	0.797	0.008	Fus,Clo,Spo,Pre	0.855	0.007
Bco,Fae,Bac,Bmo	0.788	0.006	Fus,Clo,Spo,Rum	0.857	0.007
Bco,Fae,Bac,Pre	0.829	0.006	Fus,Cop,Bac,Bmo	0.818	0.006
Bco,Fae,Bmo,Pre	0.778	0.007	Fus,Cop,Bac,Pre	0.829	0.008

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Bco,Fae,Clo,Bac	0.850	0.005	Fus,Cop,Bmo,Pre	0.799	0.007
Bco,Fae,Clo,Bmo	0.790	0.005	Fus,Cop,Clo,Bac	0.860	0.006
Bco,Fae,Clo,Pre	0.840	0.009	Fus,Cop,Clo,Bmo	0.827	0.006
Bco,Osc,Bac,Bmo	0.745	0.008	Fus,Cop,Clo,Osc	0.824	0.005
Bco,Osc,Bac,Pre	0.798	0.004	Fus,Cop,Clo,Pre	0.832	0.009
Bco,Osc,Bmo,Pre	0.768	0.005	Fus,Cop,Clo,Rum	0.848	0.006
Bco,Rum,Bac,Bmo	0.822	0.009	Fus,Cop,Clo,Spo	0.843	0.006
Bco,Rum,Bac,Pre	0.845	0.008	Fus,Cop,Fae,Bac	0.853	0.005
Bco,Rum,Bmo,Pre	0.789	0.008	Fus,Cop,Fae,Bmo	0.813	0.006
Bco,Spo,Bac,Bmo	0.797	0.008	Fus,Cop,Fae,Clo	0.836	0.006
Bco,Spo,Bac,Pre	0.802	0.005	Fus,Cop,Fae,Osc	0.837	0.006
Bco,Spo,Bmo,Pre	0.788	0.006	Fus,Cop,Fae,Pre	0.827	0.006
Clo,Osc,Bac,Bmo	0.784	0.006	Fus,Cop,Fae,Rum	0.869	0.007
Clo,Osc,Bac,Pre	0.791	0.004	Fus,Cop,Fae,Spo	0.854	0.005
Clo,Osc,Bmo,Pre	0.768	0.008	Fus,Cop,Osc,Bac	0.836	0.005
Bco,Clo,Osc,Rum	0.805	0.005	Fus,Cop,Osc,Bmo	0.814	0.005
Bco,Clo,Osc,Spo	0.778	0.009	Fus,Cop,Osc,Pre	0.821	0.007
Bco,Clo,Spo,Rum	0.802	0.007	Fus,Cop,Osc,Rum	0.854	0.006
Bco,Cop,Clo,Osc	0.782	0.004	Fus,Cop,Osc,Spo	0.831	0.005
Bco,Cop,Clo,Rum	0.806	0.006	Fus,Cop,Rum,Bac	0.892	0.006
Bco,Cop,Clo,Spo	0.784	0.010	Fus,Cop,Rum,Bmo	0.848	0.007
Bco,Cop,Fae,Bac	0.818	0.005	Fus,Cop,Rum,Pre	0.852	0.007
Bco,Cop,Fae,Bmo	0.779	0.007	Fus,Cop,Spo,Bac	0.859	0.004
Bco,Cop,Fae,Clo	0.813	0.007	Fus,Cop,Spo,Bmo	0.836	0.007
Bco,Cop,Fae,Pre	0.802	0.006	Fus,Cop,Spo,Pre	0.844	0.008
Bco,Cop,Osc,Bac	0.792	0.006	Fus,Cop,Spo,Rum	0.861	0.006
Bco,Cop,Osc,Bmo	0.758	0.008	Fus,Fae,Bac,Bmo	0.815	0.006
Bco,Cop,Osc,Pre	0.766	0.005	Fus,Fae,Bac,Pre	0.859	0.006
Bco,Cop,Rum,Bac	0.856	0.004	Fus,Fae,Bmo,Pre	0.803	0.008
Bco,Cop,Rum,Bmo	0.795	0.006	Fus,Fae,Clo,Bac	0.889	0.005
Bco,Cop,Rum,Pre	0.828	0.006	Fus,Fae,Clo,Bmo	0.841	0.005
Bco,Cop,Spo,Bac	0.810	0.004	Fus,Fae,Clo,Osc	0.845	0.004
Bco,Cop,Spo,Bmo	0.786	0.010	Fus,Fae,Clo,Pre	0.861	0.005
Bco,Cop,Spo,Pre	0.800	0.008	Fus,Fae,Clo,Rum	0.864	0.004
Bco,Fae,Clo,Osc	0.803	0.007	Fus,Fae,Clo,Spo	0.858	0.005
Bco,Fae,Clo,Rum	0.824	0.005	Fus,Fae,Osc,Bac	0.869	0.005
Bco,Fae,Clo,Spo	0.821	0.005	Fus,Fae,Osc,Bmo	0.802	0.008
Bco,Fae,Osc,Bac	0.848	0.005	Fus,Fae,Osc,Pre	0.853	0.008
Bco,Fae,Osc,Bmo	0.777	0.007	Fus,Fae,Osc,Rum	0.856	0.007
Bco,Fae,Osc,Pre	0.838	0.010	Fus,Fae,Osc,Spo	0.850	0.006
Bco,Fae,Rum,Bac	0.861	0.005	Fus,Fae,Rum,Bac	0.895	0.009
Bco,Fae,Rum,Bmo	0.790	0.008	Fus,Fae,Rum,Bmo	0.827	0.006
Bco,Fae,Rum,Pre	0.850	0.005	Fus,Fae,Rum,Pre	0.855	0.006
Bco,Fae,Spo,Bac	0.872	0.003	Fus,Fae,Spo,Bac	0.892	0.004
Bco,Fae,Spo,Bmo	0.813	0.005	Fus,Fae,Spo,Bmo	0.849	0.006

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Bco,Fae,Spo,Pre	0.839	0.004	Fus,Fae,Spo,Pre	0.864	0.007
Bco,Osc,Rum,Bac	0.835	0.007	Fus,Fae,Spo,Rum	0.871	0.005
Bco,Osc,Rum,Bmo	0.789	0.005	Fus,Fla,Bac,Bmo	0.764	0.009
Bco,Osc,Rum,Pre	0.835	0.006	Fus,Fla,Bac,Pre	0.785	0.008
Bco,Osc,Spo,Bac	0.797	0.006	Fus,Fla,Bmo,Pre	0.780	0.005
Bco,Osc,Spo,Bmo	0.784	0.009	Fus,Fla,Clo,Bac	0.856	0.006
Bco,Osc,Spo,Pre	0.788	0.007	Fus,Fla,Clo,Bmo	0.819	0.007
Bco,Spo,Rum,Bac	0.842	0.005	Fus,Fla,Clo,Osc	0.839	0.008
Bco,Spo,Rum,Bmo	0.800	0.006	Fus,Fla,Clo,Pre	0.838	0.007
Bco,Spo,Rum,Pre	0.831	0.006	Fus,Fla,Clo,Rum	0.844	0.004
Clo,Osc,Rum,Bac	0.840	0.004	Fus,Fla,Clo,Spo	0.845	0.006
Clo,Osc,Rum,Bmo	0.813	0.007	Fus,Fla,Cop,Bac	0.818	0.008
Clo,Osc,Rum,Pre	0.822	0.008	Fus,Fla,Cop,Bmo	0.791	0.006
Clo,Osc,Spo,Bac	0.800	0.005	Fus,Fla,Cop,Clo	0.817	0.006
Clo,Osc,Spo,Bmo	0.786	0.007	Fus,Fla,Cop,Fae	0.832	0.005
Clo,Osc,Spo,Pre	0.771	0.006	Fus,Fla,Cop,Osc	0.821	0.008
Bco,Cop,Fae,Rum	0.834	0.006	Fus,Fla,Cop,Pre	0.793	0.009
Bco,Cop,Fae,Spo	0.836	0.006	Fus,Fla,Cop,Rum	0.847	0.006
Bco,Cop,Fae,Osc	0.808	0.006	Fus,Fla,Cop,Spo	0.846	0.005
Bco,Cop,Osc,Rum	0.820	0.005	Fus,Fla,Fae,Bac	0.850	0.006
Bco,Cop,Osc,Spo	0.794	0.006	Fus,Fla,Fae,Bmo	0.796	0.005
Bco,Cop,Spo,Rum	0.817	0.007	Fus,Fla,Fae,Clo	0.861	0.003
Bco,Fae,Osc,Rum	0.830	0.006	Fus,Fla,Fae,Osc	0.860	0.008
Bco,Fae,Osc,Spo	0.830	0.006	Fus,Fla,Fae,Pre	0.834	0.005
Bco,Fae,Spo,Rum	0.843	0.005	Fus,Fla,Fae,Rum	0.852	0.006
Bco,Osc,Spo,Rum	0.814	0.007	Fus,Fla,Fae,Spo	0.869	0.004
Bco,Fla,Bac,Bmo	0.708	0.007	Fus,Fla,Osc,Bac	0.833	0.004
Bco,Fla,Bac,Pre	0.741	0.007	Fus,Fla,Osc,Bmo	0.824	0.005
Bco,Fla,Bmo,Pre	0.740	0.007	Fus,Fla,Osc,Pre	0.839	0.005
Bco,Fla,Clo,Bac	0.809	0.006	Fus,Fla,Osc,Rum	0.846	0.004
Bco,Fla,Clo,Bmo	0.783	0.006	Fus,Fla,Osc,Spo	0.839	0.007
Bco,Fla,Clo,Pre	0.813	0.009	Fus,Fla,Rum,Bac	0.865	0.004
Bco,Fla,Clo,Osc	0.799	0.006	Fus,Fla,Rum,Bmo	0.834	0.005
Bco,Fla,Clo,Rum	0.818	0.005	Fus,Fla,Rum,Pre	0.849	0.006
Bco,Fla,Clo,Spo	0.800	0.006	Fus,Fla,Spo,Bac	0.853	0.005
Bco,Fla,Cop,Bac	0.775	0.007	Fus,Fla,Spo,Bmo	0.835	0.004
Bco,Fla,Cop,Bmo	0.757	0.005	Fus,Fla,Spo,Pre	0.844	0.007
Bco,Fla,Cop,Clo	0.785	0.007	Fus,Fla,Spo,Rum	0.863	0.005
Bco,Fla,Cop,Fae	0.810	0.007	Fus,Osc,Bac,Bmo	0.806	0.004
Bco,Fla,Cop,Osc	0.797	0.008	Fus,Osc,Bac,Pre	0.837	0.005
Bco,Fla,Cop,Pre	0.767	0.007	Fus,Osc,Bmo,Pre	0.802	0.006
Bco,Fla,Cop,Rum	0.811	0.005	Fus,Osc,Rum,Bac	0.881	0.004
Bco,Fla,Cop,Spo	0.803	0.005	Fus,Osc,Rum,Bmo	0.839	0.005
Bco,Fla,Fae,Bac	0.822	0.007	Fus,Osc,Rum,Pre	0.862	0.006
Bco,Fla,Fae,Bmo	0.785	0.005	Fus,Osc,Spo,Bac	0.844	0.005

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Bco,Fla,Fae,Clo	0.842	0.008	Fus,Osc,Spo,Bmo	0.838	0.007
Bco,Fla,Fae,Osc	0.853	0.004	Fus,Osc,Spo,Pre	0.841	0.007
Bco,Fla,Fae,Pre	0.816	0.005	Fus,Osc,Spo,Rum	0.859	0.004
Bco,Fla,Fae,Rum	0.836	0.004	Fus,Rum,Bac,Bmo	0.874	0.004
Bco,Fla,Fae,Spo	0.859	0.004	Fus,Rum,Bac,Pre	0.881	0.006
Bco,Fla,Osc,Bac	0.785	0.007	Fus,Rum,Bmo,Pre	0.852	0.008
Bco,Fla,Osc,Bmo	0.769	0.007	Fus,Spo,Bac,Bmo	0.839	0.007
Bco,Fla,Osc,Pre	0.830	0.007	Fus,Spo,Bac,Pre	0.857	0.006
Bco,Fla,Osc,Rum	0.816	0.007	Fus,Spo,Bmo,Pre	0.837	0.006
Bco,Fla,Osc,Spo	0.817	0.006	Fus,Spo,Rum,Bac	0.887	0.009
Bco,Fla,Rum,Bac	0.818	0.005	Fus,Spo,Rum,Bmo	0.852	0.008
Bco,Fla,Rum,Bmo	0.787	0.006	Fus,Spo,Rum,Pre	0.875	0.007
Bco,Fla,Rum,Pre	0.827	0.004	Osc,Bac,Bmo,Pre	0.732	0.005
Bco,Fla,Spo,Bac	0.813	0.004	Osc,Rum,Bac,Bmo	0.847	0.006
Bco,Fla,Spo,Bmo	0.819	0.005	Osc,Rum,Bac,Pre	0.848	0.005
Bco,Fla,Spo,Pre	0.818	0.007	Osc,Rum,Bmo,Pre	0.796	0.006
Bco,Fla,Spo,Rum	0.833	0.008	Osc,Spo,Bac,Bmo	0.772	0.007
Clo,Osc,Spo,Rum	0.810	0.009	Osc,Spo,Bac,Pre	0.773	0.008
Clo,Rum,Bac,Bmo	0.854	0.005	Osc,Spo,Bmo,Pre	0.770	0.009
Clo,Rum,Bac,Pre	0.856	0.003	Osc,Spo,Rum,Bac	0.849	0.004
Clo,Rum,Bmo,Pre	0.818	0.006	Osc,Spo,Rum,Bmo	0.814	0.007
Clo,Spo,Bac,Bmo	0.785	0.002	Osc,Spo,Rum,Pre	0.829	0.006
Clo,Spo,Bac,Pre	0.781	0.005	Ros,Bac,Bmo,Pre	0.745	0.009
Clo,Spo,Bmo,Pre	0.793	0.007	Ros,Bco,Bac,Bmo	0.737	0.009
Clo,Spo,Rum,Bac	0.834	0.007	Ros,Bco,Bac,Pre	0.741	0.007
Clo,Spo,Rum,Bmo	0.812	0.005	Ros,Bco,Bmo,Pre	0.752	0.009
Clo,Spo,Rum,Pre	0.815	0.005	Ros,Bco,Clo,Bac	0.798	0.006
Cop,Bac,Bmo,Pre	0.763	0.005	Ros,Bco,Clo,Bmo	0.763	0.010
Cop,Clo,Bac,Bmo	0.822	0.005	Ros,Bco,Clo,Osc	0.765	0.006
Cop,Clo,Bac,Pre	0.823	0.005	Ros,Bco,Clo,Pre	0.787	0.009
Cop,Clo,Bmo,Pre	0.804	0.007	Ros,Bco,Clo,Rum	0.797	0.009
Cop,Clo,Osc,Bac	0.828	0.005	Ros,Bco,Clo,Spo	0.783	0.008
Cop,Clo,Osc,Bmo	0.802	0.006	Ros,Bco,Cop,Bac	0.766	0.006
Cop,Clo,Osc,Pre	0.793	0.007	Ros,Bco,Cop,Bmo	0.750	0.009
Cop,Clo,Osc,Rum	0.821	0.004	Ros,Bco,Cop,Clo	0.765	0.008
Cop,Clo,Osc,Spo	0.797	0.005	Ros,Bco,Cop,Fae	0.784	0.006
Cop,Clo,Rum,Bac	0.867	0.007	Ros,Bco,Cop,Osc	0.754	0.007
Cop,Clo,Rum,Bmo	0.833	0.007	Ros,Bco,Cop,Pre	0.747	0.008
Cop,Clo,Rum,Pre	0.835	0.006	Ros,Bco,Cop,Rum	0.800	0.007
Cop,Clo,Spo,Bac	0.814	0.005	Ros,Bco,Cop,Spo	0.808	0.007
Cop,Clo,Spo,Bmo	0.810	0.007	Ros,Bco,Fae,Bac	0.833	0.007
Cop,Clo,Spo,Pre	0.802	0.006	Ros,Bco,Fae,Bmo	0.781	0.006
Cop,Clo,Spo,Rum	0.817	0.006	Ros,Bco,Fae,Clo	0.812	0.005
Cop,Fae,Bac,Bmo	0.827	0.006	Ros,Bco,Fae,Osc	0.814	0.007
Cop,Fae,Bac,Pre	0.837	0.005	Ros,Bco,Fae,Pre	0.818	0.006

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Cop,Fae,Bmo,Pre	0.805	0.005	Ros,Bco,Fae,Rum	0.829	0.007
Cop,Fae,Clo,Bac	0.867	0.006	Ros,Bco,Fae,Spo	0.852	0.005
Cop,Fae,Clo,Bmo	0.830	0.007	Ros,Bco,Fla,Bmo	0.729	0.007
Cop,Fae,Clo,Osc	0.827	0.007	Ros,Bco,Fla,Clo	0.792	0.006
Cop,Fae,Clo,Pre	0.839	0.004	Ros,Bco,Fla,Cop	0.753	0.008
Cop,Fae,Clo,Rum	0.851	0.005	Ros,Bco,Fla,Fae	0.813	0.006
Cop,Fae,Clo,Spo	0.842	0.005	Ros,Bco,Fla,Osc	0.784	0.005
Cop,Fae,Osc,Bac	0.859	0.006	Ros,Bco,Fla,Pre	0.752	0.009
Cop,Fae,Osc,Bmo	0.816	0.006	Ros,Bco,Fla,Rum	0.787	0.007
Cop,Fae,Osc,Pre	0.831	0.006	Ros,Bco,Fla,Spo	0.826	0.005
Cop,Fae,Osc,Rum	0.865	0.005	Ros,Bco,Osc,Bac	0.783	0.007
Cop,Fae,Osc,Spo	0.850	0.006	Ros,Bco,Osc,Bmo	0.742	0.008
Cop,Fae,Rum,Bac	0.893	0.006	Ros,Bco,Osc,Pre	0.791	0.008
Cop,Fae,Rum,Bmo	0.841	0.007	Ros,Bco,Osc,Rum	0.795	0.006
Cop,Fae,Rum,Pre	0.866	0.005	Ros,Bco,Osc,Spo	0.797	0.007
Cop,Fae,Spo,Bac	0.877	0.006	Ros,Bco,Rum,Bac	0.823	0.006
Cop,Fae,Spo,Bmo	0.844	0.006	Ros,Bco,Rum,Bmo	0.783	0.007
Cop,Fae,Spo,Pre	0.851	0.003	Ros,Bco,Rum,Pre	0.815	0.006
Cop,Fae,Spo,Rum	0.863	0.006	Ros,Bco,Spo,Bac	0.828	0.004
Cop,Osc,Bac,Bmo	0.793	0.009	Ros,Bco,Spo,Bmo	0.803	0.008
Cop,Osc,Bac,Pre	0.795	0.005	Ros,Bco,Spo,Pre	0.811	0.005
Cop,Osc,Bmo,Pre	0.776	0.005	Ros,Bco,Spo,Rum	0.807	0.005
Cop,Osc,Rum,Bac	0.868	0.007	Ros,Clo,Bac,Bmo	0.820	0.007
Cop,Osc,Rum,Bmo	0.835	0.005	Ros,Clo,Bac,Pre	0.802	0.010
Cop,Osc,Rum,Pre	0.836	0.007	Ros,Clo,Bmo,Pre	0.782	0.010
Cop,Osc,Spo,Bac	0.813	0.004	Ros,Clo,Osc,Bac	0.821	0.006
Cop,Osc,Spo,Bmo	0.803	0.007	Ros,Clo,Osc,Bmo	0.785	0.008
Cop,Osc,Spo,Pre	0.794	0.008	Ros,Clo,Osc,Pre	0.789	0.006
Cop,Osc,Spo,Rum	0.835	0.007	Ros,Clo,Osc,Rum	0.804	0.005
Cop,Rum,Bac,Bmo	0.867	0.004	Ros,Clo,Osc,Spo	0.801	0.004
Cop,Rum,Bac,Pre	0.866	0.007	Ros,Clo,Rum,Bac	0.842	0.003
Cop,Rum,Bmo,Pre	0.836	0.006	Ros,Clo,Rum,Bmo	0.818	0.007
Cop,Spo,Bac,Bmo	0.811	0.010	Ros,Clo,Rum,Pre	0.822	0.005
Cop,Spo,Bac,Pre	0.811	0.005	Ros,Clo,Spo,Bac	0.823	0.008
Cop,Spo,Bmo,Pre	0.805	0.008	Ros,Clo,Spo,Bmo	0.801	0.005
Cop,Spo,Rum,Bac	0.856	0.006	Ros,Clo,Spo,Pre	0.797	0.007
Cop,Spo,Rum,Bmo	0.834	0.005	Ros,Clo,Spo,Rum	0.803	0.006
Cop,Spo,Rum,Pre	0.848	0.004	Ros,Cop,Bac,Bmo	0.801	0.006
Fae,Bac,Bmo,Pre	0.808	0.008	Ros,Cop,Bac,Pre	0.789	0.008
Fae,Clo,Bac,Bmo	0.834	0.003	Ros,Cop,Bmo,Pre	0.775	0.005
Fae,Clo,Bac,Pre	0.865	0.008	Ros,Cop,Clo,Bac	0.831	0.005
Fae,Clo,Bmo,Pre	0.827	0.006	Ros,Cop,Clo,Bmo	0.803	0.006
Fae,Clo,Osc,Bac	0.868	0.005	Ros,Cop,Clo,Osc	0.797	0.006
Fae,Clo,Osc,Bmo	0.795	0.009	Ros,Cop,Clo,Pre	0.792	0.003
Fae,Clo,Osc,Pre	0.843	0.008	Ros,Cop,Clo,Rum	0.817	0.007

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Fae,Clo,Osc,Rum	0.840	0.006	Ros,Cop,Clo,Spo	0.804	0.008
Fae,Clo,Osc,Spo	0.829	0.005	Ros,Cop,Fae,Bac	0.834	0.007
Fae,Clo,Rum,Bac	0.890	0.005	Ros,Cop,Fae,Bmo	0.810	0.003
Fae,Clo,Rum,Bmo	0.833	0.005	Ros,Cop,Fae,Clo	0.817	0.006
Fae,Clo,Rum,Pre	0.865	0.009	Ros,Cop,Fae,Osc	0.818	0.003
Fae,Clo,Spo,Bac	0.874	0.006	Ros,Cop,Fae,Pre	0.811	0.003
Fae,Clo,Spo,Bmo	0.837	0.006	Ros,Cop,Fae,Rum	0.859	0.005
Fae,Clo,Spo,Pre	0.852	0.004	Ros,Cop,Fae,Spo	0.863	0.006
Fae,Clo,Spo,Rum	0.842	0.002	Ros,Cop,Osc,Bac	0.806	0.005
Fae,Osc,Bac,Bmo	0.834	0.005	Ros,Cop,Osc,Bmo	0.781	0.008
Fae,Osc,Bac,Pre	0.860	0.006	Ros,Cop,Osc,Pre	0.776	0.005
Fae,Osc,Bmo,Pre	0.798	0.005	Ros,Cop,Osc,Rum	0.819	0.007
Fae,Osc,Rum,Bac	0.898	0.004	Ros,Cop,Osc,Spo	0.816	0.006
Fae,Osc,Rum,Bmo	0.816	0.007	Ros,Cop,Rum,Bac	0.852	0.004
Fae,Osc,Rum,Pre	0.865	0.007	Ros,Cop,Rum,Bmo	0.831	0.008
Fae,Osc,Spo,Bac	0.890	0.004	Ros,Cop,Rum,Pre	0.834	0.006
Fae,Osc,Spo,Bmo	0.828	0.007	Ros,Cop,Spo,Bac	0.835	0.004
Fae,Osc,Spo,Pre	0.844	0.007	Ros,Cop,Spo,Bmo	0.829	0.005
Fae,Osc,Spo,Rum	0.849	0.005	Ros,Cop,Spo,Pre	0.814	0.004
Fae,Rum,Bac,Bmo	0.858	0.008	Ros,Cop,Spo,Rum	0.834	0.004
Fae,Rum,Bac,Pre	0.890	0.004	Ros,Fae,Bac,Bmo	0.820	0.005
Fae,Rum,Bmo,Pre	0.832	0.007	Ros,Fae,Bac,Pre	0.847	0.006
Fae,Spo,Bac,Bmo	0.863	0.006	Ros,Fae,Bmo,Pre	0.813	0.009
Fae,Spo,Bac,Pre	0.870	0.006	Ros,Fae,Clo,Bac	0.861	0.006
Fae,Spo,Bmo,Pre	0.825	0.004	Ros,Fae,Clo,Bmo	0.811	0.006
Fae,Spo,Rum,Bac	0.888	0.006	Ros,Fae,Clo,Osc	0.820	0.008
Fae,Spo,Rum,Bmo	0.837	0.005	Ros,Fae,Clo,Pre	0.834	0.006
Fae,Spo,Rum,Pre	0.868	0.005	Ros,Fae,Clo,Rum	0.840	0.005
Fla,Clo,Bac,Bmo	0.785	0.009	Ros,Fae,Clo,Spo	0.846	0.005
Fla,Clo,Bac,Pre	0.798	0.011	Ros,Fae,Osc,Bac	0.867	0.003
Fla,Clo,Bmo,Pre	0.768	0.013	Ros,Fae,Osc,Bmo	0.804	0.005
Fla,Clo,Osc,Bac	0.795	0.005	Ros,Fae,Osc,Pre	0.843	0.008
Fla,Clo,Osc,Bmo	0.795	0.006	Ros,Fae,Osc,Rum	0.853	0.006
Fla,Clo,Osc,Pre	0.802	0.007	Ros,Fae,Osc,Spo	0.857	0.005
Fla,Clo,Osc,Rum	0.814	0.006	Ros,Fae,Rum,Bac	0.892	0.004
Fla,Clo,Osc,Spo	0.799	0.007	Ros,Fae,Rum,Bmo	0.832	0.006
Fla,Clo,Rum,Bac	0.839	0.006	Ros,Fae,Rum,Pre	0.869	0.004
Fla,Clo,Rum,Bmo	0.826	0.006	Ros,Fae,Spo,Bac	0.891	0.003
Fla,Clo,Rum,Pre	0.832	0.006	Ros,Fae,Spo,Bmo	0.863	0.006
Fla,Clo,Spo,Bac	0.792	0.007	Ros,Fae,Spo,Pre	0.862	0.005
Fla,Clo,Spo,Bmo	0.809	0.007	Ros,Fae,Spo,Rum	0.866	0.006
Fla,Clo,Spo,Pre	0.788	0.006	Ros,Fla,Bac,Bmo	0.719	0.005
Fla,Clo,Spo,Rum	0.822	0.005	Ros,Fla,Bac,Pre	0.711	0.006
Fla,Cop,Bac,Bmo	0.761	0.007	Ros,Fla,Bmo,Pre	0.743	0.008
Fla,Cop,Bac,Pre	0.766	0.006	Ros,Fla,Clo,Bac	0.806	0.009

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Fla,Cop,Bmo,Pre	0.770	0.006	Ros,Fla,Clo,Bmo	0.796	0.008
Fla,Cop,Clo,Bac	0.817	0.007	Ros,Fla,Clo,Osc	0.819	0.008
Fla,Cop,Clo,Bmo	0.806	0.004	Ros,Fla,Clo,Pre	0.806	0.006
Fla,Cop,Clo,Osc	0.802	0.004	Ros,Fla,Clo,Rum	0.814	0.006
Fla,Cop,Clo,Pre	0.793	0.008	Ros,Fla,Clo,Spo	0.811	0.006
Fla,Cop,Clo,Rum	0.826	0.008	Ros,Fla,Cop,Bac	0.776	0.010
Fla,Cop,Clo,Spo	0.800	0.004	Ros,Fla,Cop,Bmo	0.785	0.005
Fla,Cop,Fae,Bac	0.836	0.004	Ros,Fla,Cop,Clo	0.796	0.006
Fla,Cop,Fae,Bmo	0.822	0.007	Ros,Fla,Cop,Fae	0.817	0.007
Fla,Cop,Fae,Clo	0.846	0.006	Ros,Fla,Cop,Osc	0.799	0.007
Fla,Cop,Fae,Osc	0.855	0.004	Ros,Fla,Cop,Pre	0.773	0.006
Fla,Cop,Fae,Pre	0.823	0.006	Ros,Fla,Cop,Rum	0.809	0.006
Fla,Cop,Fae,Rum	0.865	0.005	Ros,Fla,Cop,Spo	0.828	0.003
Fla,Cop,Fae,Spo	0.862	0.005	Ros,Fla,Fae,Bac	0.829	0.007
Fla,Cop,Osc,Bac	0.806	0.007	Ros,Fla,Fae,Bmo	0.797	0.006
Fla,Cop,Osc,Bmo	0.816	0.007	Ros,Fla,Fae,Clo	0.851	0.006
Fla,Cop,Osc,Pre	0.806	0.005	Ros,Fla,Fae,Osc	0.847	0.005
Fla,Cop,Osc,Rum	0.840	0.004	Ros,Fla,Fae,Pre	0.838	0.008
Fla,Cop,Osc,Spo	0.814	0.008	Ros,Fla,Fae,Rum	0.856	0.005
Fla,Cop,Rum,Bac	0.845	0.004	Ros,Fla,Fae,Spo	0.876	0.003
Fla,Cop,Rum,Bmo	0.824	0.005	Ros,Fla,Osc,Bac	0.792	0.006
Fla,Cop,Rum,Pre	0.831	0.002	Ros,Fla,Osc,Bmo	0.783	0.008
Fla,Cop,Spo,Bac	0.812	0.005	Ros,Fla,Osc,Pre	0.803	0.007
Fla,Cop,Spo,Bmo	0.832	0.006	Ros,Fla,Osc,Rum	0.814	0.006
Fla,Cop,Spo,Pre	0.809	0.004	Ros,Fla,Osc,Spo	0.820	0.008
Fla,Cop,Spo,Rum	0.843	0.005	Ros,Fla,Rum,Bac	0.823	0.006
Fla,Fae,Bac,Bmo	0.801	0.007	Ros,Fla,Rum,Bmo	0.815	0.005
Fla,Fae,Bac,Pre	0.850	0.007	Ros,Fla,Rum,Pre	0.837	0.004
Fla,Fae,Bmo,Pre	0.813	0.010	Ros,Fla,Spo,Bac	0.825	0.006
Fla,Fae,Clo,Bac	0.866	0.009	Ros,Fla,Spo,Bmo	0.835	0.005
Fla,Fae,Clo,Bmo	0.832	0.004	Ros,Fla,Spo,Pre	0.827	0.005
Fla,Fae,Clo,Osc	0.854	0.005	Ros,Fla,Spo,Rum	0.840	0.005
Fla,Fae,Clo,Pre	0.866	0.007	Ros,Fus,Bac,Bmo	0.781	0.009
Fla,Fae,Clo,Rum	0.871	0.004	Ros,Fus,Bac,Pre	0.801	0.004
Fla,Fae,Clo,Spo	0.858	0.004	Ros,Fus,Bco,Bac	0.766	0.013
Fla,Fae,Osc,Bac	0.878	0.003	Ros,Fus,Bco,Bmo	0.744	0.005
Fla,Fae,Osc,Bmo	0.846	0.007	Ros,Fus,Bco,Clo	0.783	0.006
Fla,Fae,Osc,Pre	0.884	0.007	Ros,Fus,Bco,Cop	0.751	0.008
Fla,Fae,Osc,Rum	0.868	0.005	Ros,Fus,Bco,Fae	0.802	0.005
Fla,Fae,Osc,Spo	0.873	0.005	Ros,Fus,Bco,Fla	0.736	0.004
Fla,Fae,Rum,Bac	0.886	0.007	Ros,Fus,Bco,Osc	0.759	0.005
Fla,Fae,Rum,Bmo	0.835	0.008	Ros,Fus,Bco,Pre	0.757	0.009
Fla,Fae,Rum,Pre	0.877	0.005	Ros,Fus,Bco,Rum	0.799	0.005
Fla,Fae,Spo,Bac	0.885	0.007	Ros,Fus,Bco,Spo	0.834	0.006
Fla,Fae,Spo,Bmo	0.866	0.004	Ros,Fus,Bmo,Pre	0.787	0.005

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Fla,Fae,Spo,Pre	0.856	0.005	Ros,Fus,Clo,Bac	0.852	0.006
Fla,Fae,Spo,Rum	0.878	0.007	Ros,Fus,Clo,Bmo	0.818	0.008
Fla,Osc,Bac,Bmo	0.768	0.005	Ros,Fus,Clo,Osc	0.817	0.006
Fla,Osc,Bac,Pre	0.780	0.006	Ros,Fus,Clo,Pre	0.822	0.006
Fla,Osc,Bmo,Pre	0.781	0.006	Ros,Fus,Clo,Rum	0.829	0.004
Fla,Osc,Rum,Bac	0.831	0.004	Ros,Fus,Clo,Spo	0.837	0.005
Fla,Osc,Rum,Bmo	0.809	0.004	Ros,Fus,Cop,Bac	0.829	0.008
Fla,Osc,Rum,Pre	0.832	0.005	Ros,Fus,Cop,Bmo	0.796	0.007
Fla,Osc,Spo,Bac	0.795	0.005	Ros,Fus,Cop,Clo	0.806	0.007
Fla,Osc,Spo,Bmo	0.825	0.006	Ros,Fus,Cop,Fae	0.814	0.003
Fla,Osc,Spo,Pre	0.805	0.009	Ros,Fus,Cop,Osc	0.793	0.004
Fla,Osc,Spo,Rum	0.823	0.008	Ros,Fus,Cop,Pre	0.783	0.004
Fla,Rum,Bac,Bmo	0.823	0.007	Ros,Fus,Cop,Rum	0.835	0.005
Fla,Rum,Bac,Pre	0.839	0.005	Ros,Fus,Cop,Spo	0.844	0.004
Fla,Rum,Bmo,Pre	0.816	0.007	Ros,Fus,Fae,Bac	0.856	0.004
Fla,Spo,Bac,Bmo	0.805	0.009	Ros,Fus,Fae,Bmo	0.801	0.006
Fla,Spo,Bac,Pre	0.793	0.007	Ros,Fus,Fae,Clo	0.842	0.002
Fla,Spo,Bmo,Pre	0.810	0.006	Ros,Fus,Fae,Osc	0.828	0.007
Fla,Spo,Rum,Bac	0.846	0.003	Ros,Fus,Fae,Pre	0.829	0.006
Fla,Spo,Rum,Bmo	0.835	0.004	Ros,Fus,Fae,Rum	0.846	0.006
Fla,Spo,Rum,Pre	0.850	0.004	Ros,Fus,Fae,Spo	0.870	0.004
Fus,Bac,Bmo,Pre	0.788	0.007	Ros,Fus,Fla,Bac	0.759	0.009
Fus,Bco,Bac,Bmo	0.761	0.006	Ros,Fus,Fla,Bmo	0.766	0.005
Fus,Bco,Bac,Pre	0.804	0.009	Ros,Fus,Fla,Clo	0.812	0.004
Fus,Bco,Bmo,Pre	0.778	0.008	Ros,Fus,Fla,Cop	0.780	0.007
Fus,Bco,Clo,Bac	0.843	0.007	Ros,Fus,Fla,Fae	0.816	0.004
Fus,Bco,Clo,Bmo	0.800	0.008	Ros,Fus,Fla,Osc	0.801	0.007
Fus,Bco,Clo,Osc	0.805	0.005	Ros,Fus,Fla,Pre	0.754	0.007
Fus,Bco,Clo,Pre	0.829	0.006	Ros,Fus,Fla,Rum	0.809	0.008
Fus,Bco,Clo,Rum	0.842	0.006	Ros,Fus,Fla,Spo	0.842	0.007
Fus,Bco,Clo,Spo	0.820	0.005	Ros,Fus,Osc,Bac	0.813	0.007
Fus,Bco,Cop,Bac	0.811	0.006	Ros,Fus,Osc,Bmo	0.772	0.006
Fus,Bco,Cop,Bmo	0.765	0.005	Ros,Fus,Osc,Pre	0.797	0.008
Fus,Bco,Cop,Clo	0.797	0.006	Ros,Fus,Osc,Rum	0.823	0.009
Fus,Bco,Cop,Fae	0.803	0.006	Ros,Fus,Osc,Spo	0.829	0.005
Fus,Bco,Cop,Osc	0.782	0.006	Ros,Fus,Rum,Bac	0.861	0.007
Fus,Bco,Cop,Pre	0.789	0.008	Ros,Fus,Rum,Bmo	0.830	0.005
Fus,Bco,Cop,Rum	0.831	0.006	Ros,Fus,Rum,Pre	0.822	0.005
Fus,Bco,Cop,Spo	0.825	0.009	Ros,Fus,Spo,Bac	0.860	0.005
Fus,Bco,Fae,Bac	0.837	0.005	Ros,Fus,Spo,Bmo	0.851	0.006
Fus,Bco,Fae,Bmo	0.770	0.006	Ros,Fus,Spo,Pre	0.852	0.005
Fus,Bco,Fae,Clo	0.836	0.006	Ros,Fus,Spo,Rum	0.852	0.005
Fus,Bco,Fae,Osc	0.812	0.005	Ros,Osc,Bac,Bmo	0.786	0.006
Fus,Bco,Fae,Pre	0.815	0.008	Ros,Osc,Bac,Pre	0.788	0.008
Fus,Bco,Fae,Rum	0.832	0.009	Ros,Osc,Bmo,Pre	0.765	0.006

Genera	AUC_mean	AUC_std	Genera	AUC_mean	AUC_std
Fus,Bco,Fae,Spo	0.837	0.004	Ros,Osc,Rum,Bac	0.840	0.006
Fus,Bco,Fla,Bac	0.775	0.009	Ros,Osc,Rum,Bmo	0.817	0.007
Fus,Bco,Fla,Bmo	0.745	0.006	Ros,Osc,Rum,Pre	0.811	0.007
Fus,Bco,Fla,Clo	0.802	0.005	Ros,Osc,Spo,Bac	0.825	0.005
Fus,Bco,Fla,Cop	0.786	0.007	Ros,Osc,Spo,Bmo	0.812	0.009
Fus,Bco,Fla,Fae	0.794	0.006	Ros,Osc,Spo,Pre	0.803	0.007
Fus,Bco,Fla,Osc	0.809	0.007	Ros,Osc,Spo,Rum	0.811	0.004
Fus,Bco,Fla,Pre	0.781	0.009	Ros,Rum,Bac,Bmo	0.849	0.006
Fus,Bco,Fla,Rum	0.813	0.008	Ros,Rum,Bac,Pre	0.853	0.007
Fus,Bco,Fla,Spo	0.840	0.005	Ros,Rum,Bmo,Pre	0.826	0.007
Fus,Bco,Osc,Bac	0.814	0.008	Ros,Spo,Bac,Bmo	0.831	0.008
Fus,Bco,Osc,Bmo	0.756	0.005	Ros,Spo,Bac,Pre	0.826	0.006
Fus,Bco,Osc,Pre	0.815	0.007	Ros,Spo,Bmo,Pre	0.824	0.009
Fus,Bco,Osc,Rum	0.824	0.006	Ros,Spo,Rum,Bac	0.853	0.006
Fus,Bco,Osc,Spo	0.823	0.006	Ros,Spo,Rum,Bmo	0.832	0.006
Fus,Bco,Rum,Bac	0.860	0.006	Ros,Spo,Rum,Pre	0.839	0.004
Spo,Rum,Bac,Bmo	0.847	0.007	Rum,Bac,Bmo,Pre	0.837	0.006
Spo,Rum,Bac,Pre	0.859	0.003	Spo,Bac,Bmo,Pre	0.768	0.008
Spo,Rum,Bmo,Pre	0.821	0.004			

As selection criteria for combinations of bacterial genera that are predictive for the dysbiotic B2 enterotype in the methods described herein, an AUC of 0.7 or more was used (see Table 10-12). Given the full set of 13 genera described herein, there are 8191 combinations of one or more genera. All combinations of 2, 3 or 4 bacterial genera are listed in Table 10, 11 or 12 respectively. However as reporting even more combinations is impractical, we considered to list the non-redundant combinations with a ROC AUC of 0.7 or above. A combination is considered redundant if a subset of its genera already yields predictions with a ROC AUC of 0.7. Indeed, when a combination of two bacterial genera is already sufficient to predict the B2 enterotype, any further combination with an additional genus or genera will logically also be selected. Therefore, in the full list of all possible combinations of the 13 bacterial genera disclosed herein we have removed the redundant combinations.

When singles are ignored, there is a list of 65 non-redundant combinations. These non-redundant combinations are listed in Table 13 and should be read as any combination of the 13 bacterial genera from Table 5 (i.e. *Coprococcus, Faecalibacterium, Oscillibacter, Ruminococcus, Sporobacter, Fusicatenibacter, Clostridium_IV, Butyricimonas, Butyricicoccus, Prevotella, Flavonifractor, Roseburia* and *Bacteroides*) comprising at least the mentioned bacterial genera. For example, the biomarker panel consisting of *Clostridium_IV* and *Sporobacter* already has an AUC of 0.733, any biomarker panel consisting of *Clostridium_IV*, *Sporobacter* and any additional genus from Table 5 also has an AUC of 0.7 or more and is thus useful in the methods described herein.

Table 13. Non-redundant combinations of the bacterial genera from Table 5 with an AUC of at least 0.7. *Coprococcus*, Cop; *Faecalibacterium*, Fae; *Oscillibacter*, Osc; *Ruminococcus*, Rum; *Sporobacter*, Spo; *Fusicatenibacter*, Fus; *Clostridium_IV*, Clo; *Butyricimonas*, Bmo; *Butyricicoccus*, Bco; *Prevotella*, Pre; *Flavonifractor*, Fla; *Bacteroides*, Bac; *Roseburia*, Ros.

Genera	roc_auc_mean	roc_auc_std	Genera	roc_auc_mean	roc_auc_std
Clo,Spo	0.733	0.006	Fae,Rum	0.807	0.007
Spo,Bac	0.732	0.007	Fus,Spo	0.809	0.006
Cop,Pre	0.732	0.009	Fus,Rum	0.790	0.011
Cop,Bmo	0.725	0.008	Fus,Fae	0.767	0.008
Rum,Bmo	0.746	0.011	Fus,Cop	0.751	0.007
Clo,Bac	0.737	0.013	Fae,Pre	0.752	0.009
Osc,Pre	0.741	0.008	Cop,Osc	0.752	0.010
Ros,Cop	0.722	0.008	Bco,Spo	0.753	0.007
Ros,Rum	0.736	0.004	Fla,Spo	0.755	0.010
Bco,Clo	0.709	0.007	Spo,Pre	0.756	0.008
Clo,Pre	0.719	0.007	Cop,Bac	0.757	0.006
Ros,Osc	0.702	0.010	Bco,Rum	0.759	0.012
Bco,Osc	0.706	0.009	Fus,Osc	0.759	0.009
Fla,Cop	0.708	0.007	Osc,Rum	0.760	0.008
Fus,Bmo	0.710	0.008	Ros,Fae	0.760	0.007
Clo,Bmo	0.715	0.010	Fae,Osc	0.763	0.012
Osc,Spo	0.748	0.009	Fla,Rum	0.764	0.009
Fla,Osc	0.748	0.008	Fus,Pre	0.764	0.007
Clo,Rum	0.771	0.009	Fus,Bac	0.748	0.009
Fus,Clo	0.771	0.011	Fae,Bmo	0.724	0.009
Cop,Clo	0.776	0.007	Ros,Fla,Clo	0.770	0.008
Cop,Fae	0.780	0.009	Ros,Bac,Bmo	0.729	0.007
Fla,Fae	0.786	0.009	Ros,Fus,Bco	0.709	0.008
Rum,Pre	0.787	0.005	Bco,Fla,Pre	0.708	0.011
Cop,Spo	0.788	0.005	Ros,Bco,Pre	0.704	0.012
Rum,Bac	0.788	0.008	Fus,Bco,Fla	0.721	0.006
Fae,Bac	0.792	0.008	Osc,Bac,Bmo	0.713	0.008
Spo,Bmo	0.793	0.007	Ros,Bmo,Pre	0.711	0.009
Fae,Clo	0.799	0.008	Ros,Bco,Fla,Bmo	0.729	0.007
Fae,Spo	0.800	0.007	Bco,Fla,Bac,Bmo	0.708	0.007
Ros,Spo	0.801	0.008	Ros,Fla,Bac,Pre	0.711	0.006
Spo,Rum	0.804	0.013	Bco,Bac,Bmo,Pre	0.712	0.013
Cop,Rum	0.804	0.010			

Example 7. Any combination of at least 5 selected genera can predict gut flora dysbiosis

By summarizing the full set of all possible combinations between the 13 bacterial genera, we observed that all combinations of 5 or more genera out of the 13 genera listed in Table 5 allow predictors with a ROC AUC of 0.7 or more to be created (Table 14).

Table 14. Summary of all possible combinations of the 13 selected genera and the minimum and maximum performance for each possible combination with n genera.

number of genera	number of combinations	min_AUC	max_AUC
1	13	0.535	0.774
2	78	0.543	0.809
3	286	0.633	0.877
4	715	0.681	0.898
5	1287	0.738	0.910
6	1716	0.777	0.914
7	1716	0.816	0.915
8	1287	0.832	0.918
9	715	0.853	0.918
10	286	0.871	0.917
11	78	0.879	0.916
12	13	0.896	0.913
13	1	0.908	0.908

Example 8. Relative improvements of predictions when including additional genera

Combinations of *n* genera could be evaluated by their relative improvement in ROC AUC over the best performing subset of *n-1* genera. When only considering a relative improvement of 10% or higher, there are 7 combinations of 2 genera which outperform predictions bases on the best single genus in the pair by that amount. When considering combinations of 3 genera, 2 combinations exist which improve the ROC AUC by 10% over the best subset of 2 genera. The full list is provided in Table 15.

Table 15. Combinations of 2 or 3 genera (n) that show a relative improvement in ROC AUC of 10% or higher compared to the ROC AUC of the single genus or the highest ROC AUC of 2 of the genera respectively. *Roseburia*, Ros; *Bacteroides*, Bac; *Faecalibacterium*, Fae; *Clostridium_IV*, Clo; *Ruminococcus*, Rum; *Butyricicoccus*, Bco; *Prevotella*, Pre; *Coprococcus*, Cop; *Flavonifractor*, Fla; *Fusicatenibacter*, Fus.

genera	n	AUC_mean n genera	AUC_mean n-1 genera	relative improvement (%)
Ros, Bac	2	0.623	0.548	13.586
Fae, Clo	2	0.799	0.714	11.940
Fae, Rum	2	0.807	0.722	11.722

genera	n	AUC_mean n genera	AUC_mean n-1 genera	relative improvement (%)
Bco, Pre	2	0.672	0.603	11.342
Fae, Bac	2	0.792	0.714	10.874
Cop, Rum	2	0.804	0.728	10.559
Fla, Fae	2	0.786	0.714	10.046
Fus, Clo, Bac	3	0.856	0.771	10.970
Ros, Flav, Clo	3	0.770	0.699	10.134

A final metric to evaluate the predictive power of a single genus in combination with other genera is to evaluate the relative improvement after adding a genus to another genus (or combination of genera) and tracking how much the original ROC AUC improves by including the extra genus. This approach shows for example that *Faecalibacterium* and *Ruminococcus*, which on their own are sufficient to obtain a ROC AUC above 0.7, when combined yield a ROC AUC of 0.81 an improvement of 11.7% and 13.0% over the performance of the respective individual genera, showing these genera are complementary rather than correlated (Table 16). Six genera, *Sporobacter, Coprococcus, Ruminococcus, Faecalibacterium, Oscillibacter,* and *Fusicatenibacter*, can be added to various combinations, with ROC AUC > 0.7, of one or two other genera to improve predictions by 10% or more. Highlighting that these genera are not only powerful indicators of B2 by themselves but provide significant improvements in combination with other genera. The full set of single and paired genera with a ROC AUC of 0.7 or more which can be improved with 10% or more by adding a single organism is provided in Table 16.

Table 16. Improvements of the ROC AUC of the original selection by adding an extra genus. The 1st column (Genera) shows the original selection of bacterial genera and the 2nd column the corresponding ROC AUC (AUC_mean). The 4th column shows the ROC AUC of the extra genus (which is listed in the 3rd column) to the original selection and the 5th column the ROC AUC after adding the extra genus. Finally, the columns entitled “improvement (genera)” and “improvement (extra genus)” show the percentage increase in ROC AUC compared to the original selection and extra genus respectively.

Genera	AUC_mean	extra genus	AUC_mean extra genus	AUC with extra genus	Improvement % (genera)	improvement % (extra genus)
Bco,Clo	0.7088	Rum	0.7220	0.8095	14.2068	10.8087
Bco,Osc	0.7056	Pre	0.5990	0.7938	12.5009	24.5371
Clo,Bac	0.7371	Cop	0.7275	0.8211	11.3850	11.3913
Clo,Bac	0.7371	Rum	0.7220	0.8438	14.4743	14.4355
Clo,Bac	0.7371	Fus	0.7249	0.8556	16.0728	15.2738
Clo,Bac	0.7371	Fae	0.7140	0.8543	15.8912	16.4228
Clo,Pre	0.7190	Rum	0.7220	0.8120	12.9386	11.0850
Clo,Pre	0.7190	Fus	0.7249	0.8274	15.0718	12.3817
Clo,Pre	0.7190	Fae	0.7140	0.8269	15.0058	13.6556
Clo,Spo	0.7327	Fus	0.7249	0.8257	12.6908	12.2060
Clo,Spo	0.7327	Fae	0.7140	0.8239	12.4364	13.3364

Genera	AUC_mean	extra genus	AUC_mean extra genus	AUC with extra genus	Improvement % (genera)	improvement % (extra genus)
Cop	0.7275	Rum	0.7220	0.8044	10.5586	10.2353
Cop,Bac	0.7569	Rum	0.7220	0.8549	12.9424	15.5415
Cop,Bmo	0.7249	Rum	0.7220	0.8105	11.8093	10.9166
Cop,Bmo	0.7249	Clo	0.6980	0.7986	10.1713	12.6010
Cop,Osc	0.7519	Rum	0.7220	0.8324	10.7144	13.2644
Cop,Pre	0.7318	Rum	0.7220	0.8296	13.3591	12.9629
Fae	0.7140	Rum	0.7220	0.8067	12.9801	10.4921
Fae	0.7140	Clo	0.6980	0.7992	11.9396	12.6667
Fae	0.7140	Fla	0.6022	0.7857	10.0460	23.3583
Fae	0.7140	Bac	0.5348	0.7916	10.8743	32.4436
Fae,Bac	0.7916	Spo	0.7739	0.8725	10.2199	11.3075
Fae,Bac	0.7916	Rum	0.7220	0.8765	10.7246	17.6263
Fae,Osc	0.7628	Fla	0.6022	0.8511	11.5848	29.2474
Fae,Osc	0.7628	Bac	0.5348	0.8533	11.8696	37.3250
Fae,Pre	0.7516	Osc	0.7048	0.8272	10.0500	14.7958
Fae,Pre	0.7516	Clo	0.6980	0.8269	10.0130	15.5893
Fae,Pre	0.7516	Rum	0.7220	0.8556	13.8356	15.6153
Fae,Pre	0.7516	Bac	0.5348	0.8308	10.5304	35.6284
Fla,Cop	0.7084	Rum	0.7220	0.8082	14.1023	10.6672
Fla,Cop	0.7084	Fae	0.7140	0.8144	14.9750	12.3325
Fla,Cop	0.7084	Osc	0.7048	0.8064	13.8463	12.6032
Fla,Osc	0.7485	Fae	0.7140	0.8511	13.7142	16.1113
Fla,Rum	0.7636	Fae	0.7140	0.8560	12.0979	16.5875
Fla,Spo	0.7550	Fae	0.7140	0.8565	13.4382	16.6396
Fus,Bac	0.7483	Cop	0.7275	0.8317	11.1420	12.5244
Fus,Bac	0.7483	Fae	0.7140	0.8529	13.9766	16.2887
Fus,Bac	0.7483	Rum	0.7220	0.8678	15.9703	16.8013
Fus,Bac	0.7483	Clo	0.6980	0.8556	14.3384	18.4223
Fus,Bmo	0.7096	Rum	0.7220	0.8191	15.4295	11.8552
Fus,Bmo	0.7096	Clo	0.6980	0.7974	12.3714	12.4698
Fus,Clo	0.7710	Bac	0.5348	0.8556	10.9696	37.4963
Fus,Cop	0.7508	Rum	0.7220	0.8366	11.4335	13.6948
Fus,Cop	0.7508	Bac	0.5348	0.8317	10.7816	35.6988
Fus,Fae	0.7670	Bac	0.5348	0.8529	11.2071	37.2979
Osc,Pre	0.7410	Fus	0.7249	0.8174	10.3096	11.3164
Osc,Pre	0.7410	Fae	0.7140	0.8272	11.6243	13.6846
Osc,Spo	0.7482	Fae	0.7140	0.8236	10.0728	13.3034
Ros,Cop	0.7215	Fae	0.7140	0.8009	11.0040	10.8565
Ros,Cop	0.7215	Rum	0.7220	0.8119	12.5292	11.0745
Ros,Fae	0.7605	Rum	0.7220	0.8459	11.2419	14.6480
Ros,Osc	0.7018	Fae	0.7140	0.8139	15.9716	12.2759
Ros,Rum	0.7364	Cop	0.7275	0.8119	10.2557	10.3959
Ros,Rum	0.7364	Fae	0.7140	0.8459	14.8719	15.5985
Ros,Rum	0.7364	Bac	0.5348	0.8219	11.6108	34.9338
Rum	0.7220	Fae	0.7140	0.8067	11.7220	11.4889
Rum,Bac	0.7882	Fus	0.7249	0.8678	10.1021	16.4661

Genera	AUC_mean	extra genus	AUC_mean extra genus	AUC with extra genus	Improvement % (genera)	improvement % (extra genus)
Rum,Bac	0.7882	Fae	0.7140	0.8765	11.2049	18.5436
Rum,Bmo	0.7460	Bac	0.5348	0.8257	10.6855	35.2347
Spo,Bac	0.7322	Rum	0.7220	0.8347	14.0043	13.4990
Spo,Bac	0.7322	Fus	0.7249	0.8459	15.5363	14.3021
Spo,Bac	0.7322	Fae	0.7140	0.8725	19.1709	18.1707
Spo,Bac	0.7322	Ros	0.5482	0.8280	13.0910	33.7930
Spo,Pre	0.7555	Fus	0.7249	0.8362	10.6873	13.3099

Previously, it was disclosed that a high abundance of Bacteroides and a low abundance of *Faecalibacterium* in a stool sample of a subject compared to that of a healthy control is predictive for the Bacteroides_2 enterotype (WO2019115755A1). Interestingly, additionally determining the level of *Ruminicoccus*, *Sporobacter, Clostridium_IV, Oscillibacter, Fusicatenibacter, Prevotella, Coprococcus, Roseburia* or *Flavonifractor* significantly increases the AUC (Table 17).

Table 17. The ROC AUC for the combination of *Faecalibacterium* and *Bacteroides* and at least one additional bacterial genus selected from Table 5. Included is the relative improvement of the additional genus over the *Faecalibacterium* and *Bacteroides* combination.

Genera	ROC AUC (mean)	ROC AUC (std)	Relative improvement
Faecalibacterium,*Bacteroides* (reference)	0.7916	0.0082	-
+ *Ruminococcus*	0.8765	0.0077	10.72%
+ *Sporobacter*	0.8725	0.0049	10.22%
+ *Clostridium_IV*	0.8543	0.0062	7.91%
+ *Oscillibacter*	0.8533	0.0068	7.79%
+ *Fusicatenibacter*	0.8529	0.0064	7.74%
+ *Prevotella*	0.8308	0.0073	4.95%
+ *Coprococcus*	0.8278	0.0052	4.56%
+ *Roseburia*	0.8241	0.0097	4.10%
+ *Flavonifractor*	0.8115	0.0090	2.51%

Example 9. qPCR analysis and design of kits

During development of the method, panel and kit according to the invention, it was surprisingly found that some genera could be better detectable and more reliable indicators or biomarkers of dysbiosis, i.e., B2 enterotype than other bacterial genera present in a biological, preferably stool sample. In the present invention, it was demonstrated that quantitative measuring a limited set of key genera in a biological sample, preferably a stool sample, by qPCR in combination with machine learning based predictions can be simplified and reliable method to detect dysbiosis or B2 enterotype. In previous examples, this was done on a combination of 16S amplicon sequencing data combined with flow cytometry (Quantitative Microbiome Profiling, QMP), which served as a proxy for other detection

methods. The way the patent was written allowed for the inclusion of other quantitative methods, such as qPCR, to be considered as input features for predicting the enterotype.

To this end, the current qPCR prototype for B2 detection is based on 389 samples, with known enterotypes assigned by the reference method to assess method agreement, i.e. concordance, and thus the feasibility of changing from 16S amplicon sequencing (16S NGS) to qPCR for enterotype assignment. Furthermore, performance of the qPCR based method was thoroughly evaluated and compared to the QMP based predictions, to pinpoint potential future improvements for the method. Prior to training these samples were balanced, using random under-sampling, to create dataset with an equal amount of B2 and non B2 samples and split into a training and test set using 80-20 ratio. This resulted into a training set of 148 and test set with 38 samples. The latter was used to evaluate and report performance (for example confusion matrices, ROC AUC, etc.).

The present invention surprisingly shows that detection of dysbiosis and/or enterotype can be done by qPCR using (a sub-set of) the genera identified in studies based on quantitative microbiome profiles (QMP). A random forest classifier tailored to qPCR data showed similar performance in classification of stool samples in B2 versus non-B2 enterotype. The Receiver Operator Curve Area Under the Curve (ROC AUC) was 0.94 for the model on QMP (1590 training samples) and 0.89 for the model on qPCR data (148 training samples). Moreover, classification for B2 enterotype was also possible at similar good performance with ROC AUC of 0.87 for B2. The method concordance, i.e. agreement with the same enterotype assignment with qPCR and QMP, was 71% for B2.

Inspecting the qPCR model and evaluating performance against the QMP pilot shows both models perform similarly in similar circumstances. Key genera and their impact on the models' output are comparable. This suggest insights obtained with the QMP work translate well to qPCR.

When inspecting relative abundances per genus, obtained using qPCR clear differences between the enterotypes can be observed. These closely match results observed in 16S studies with e.g. high *Bacteroides/Faecalibacterium* ratio in B2, high relative abundance of slow growing genera such as *Oscillibacter* in *Ruminococcus* and *Prevotella* reaching high relative abundance samples assigned to its namesake enterotype (**Figure 4**).

A new random forest classifier tailored to qPCR data showed similar performance in classification of stool samples in B2 versus non-B2 enterotype. The Receiver Operator Curve (ROC) is an indicator for classification performance by plotting the sensitivity (true positive rate = true positives / true+false positives) in function of the 1-specificity (false positives) for different thresholds. The higher the Area Under the Curve (ROC AUC), the lower the false positive and false negative classification and thus the better the classification performance. A perfect test without any misclassifications would have a ROC

AUC of 1.0. A useless test has a ROC AUC of 0.5. ROC curves are frequently used to compare tests and to select a specific threshold with emphasis on sensitivity or specificity if one is more important than the other in the specific test. In this use case, there is an equal importance.

The ROC AUC reported previously was 0.94 for the model on QMP data (which was trained on 1590 samples including 13 genera) and 0.89 for the model on qPCR data (148 samples 6 genera within the relative qPCR results), (Figure 5 and Figure 6). Though when limiting the QMP model to six genera the performance drops to 0.91, highlighting that the current qPCR panel performs on par with the QMP based methods, despite the qPCR-based training data being an order of magnitude smaller.

While all features included in the qPCR model are relevant for predicting the enterotype, all 63 possible subsets, ranging from a single genus to the full set, were evaluated. This shows that even with a single genus, *Oscillibacter*, a ROC AUC of 0.7 can be reached, at least two are needed to go over 0.8 (*Bacteroides* + *Oscillibacter* or *Faecalibacterium* + *Oscillibacter*) and with three genera (*Bacteroides* + *Faecalibacterium* + *Oscillibacter*) a ROC AUC comparable to the full complement of genera can be reached. Though other combinations with adequate performance can be found below in Table 18.

Table 18. ROC AUC performance by qPCR analysis for combinations of selected genera.

n_genera	genera	ROC AUC (mean)
1	*Oscillibacter*	0.72175
1	*Bacteroides*	0.69187
1	*Faecalibacterium*	0.67741
1	*Ruminococcus*	0.66088
1	*Prevotella*	0.54370
2	*Faecalibacterium, Oscillibacter*	0.83824
2	*Bacteroides, Oscillibacter*	0.81633
2	*Bacteroides, Faecalibacterium*	0.78619
2	*Faecalibacterium, Ruminococcus*	0.77484
2	*Oscillibacter, Ruminococcus*	0.75272
2	*Bacteroides, Ruminococcus*	0.73460
2	*Oscillibacter, Prevotella*	0.71954
2	*Faecalibacterium, Prevotella*	0.70664
2	*Bacteroides, Prevotella*	0.69939
2	*Prevotella, Ruminococcus*	0.64206
3	*Bacteroides, Faecalibacterium, Oscillibacter*	0.87076
3	*Faecalibacterium, Oscillibacter, Prevotella*	0.86492
3	*Faecalibacterium, Oscillibacter, Ruminococcus*	0.84593
3	*Bacteroides, Faecalibacterium, Ruminococcus*	0.82274
3	*Bacteroides, Oscillibacter, Prevotella*	0.80688

n_genera	genera	ROC AUC (mean)
3	*Bacteroides, Oscillibacter, Ruminococcus*	0.80533
3	*Faecalibacterium, Prevotella, Ruminococcus*	0.78402
3	*Bacteroides, Faecalibacterium, Prevotella*	0.77397
3	*Bacteroides, Prevotella, Ruminococcus*	0.74841
3	*Oscillibacter, Prevotella, Ruminococcus*	0.74495
4	*Bacteroides, Faecalibacterium, Oscillibacter, Prevotella*	0.85978
4	*Bacteroides, Faecalibacterium, Oscillibacter, Ruminococcus*	0.85941
4	*Faecalibacterium, Oscillibacter, Prevotella, Ruminococcus*	0.84646
4	*Bacteroides, Faecalibacterium, Prevotella, Ruminococcus*	0.81910
4	*Bacteroides, Oscillibacter, Prevotella, Ruminococcus*	0.80781
5	*Bacteroides, Faecalibacterium, Oscillibacter, Prevotella, Rumino-coccus*	0.85886

qPCR reactions and methodology.

The qPCR assay used in this example consists of 6 singleplex reactions, all targeting regions in the 16S rRNA gene and shown in detail in **Table 19**.

Table 19: qPCR reactions setup.

qPCR reaction	Tem-plate di-lution	Tm (°C)	Product (bp)	Fluoro-phore	16S rRNA gene	Reference
1. Ruminococcus	100fold	64	246	FAM	V5 region	Own design
2. Bacteroides	100fold	60	215	FAM	V2 region	Gómez-Doñate et al. (2016). Microbiology Open, 14; 5(1):83–94. doi: 10.1002/mbo3.313
3. Prevotella	100fold	63	139	FAM	V6 region	Own design
4. Faecalibacterium	100fold	64	122	FAM	V6 region	Own design
5. Oscillibacter	10fold	66	118	FAM	V4 region	Own design
6. Total bacte-ria	100fold	64	223	FAM	V3 region	Brukner et al. (2015) Diagn Microbiol Infect Dis 83 (1), 1-6. doi: 10.1016/j.diagmicro-bio.2015.04.005

"Tm" for the melting temperature here defined as the melting temperature of the oligo.

For specific embodiments, example primers and probes for each reaction are provided. The qPCR reaction for total bacteria was taken from literature: Brukner et al. (2015) Diagn Microbiol Infect Dis 83 (1), 1-6. doi: 10.1016/j.diagmicrobio.2015.04.005. The qPCR reaction for the genus Bacteroides was also taken from literature: Gómez-Doñate et al. (2016). Development of new host-specific Bacteroides qPCRs for the identification of fecal contamination sources in water. Microbiology Open, 14; 5(1):83–94. doi:

10.1002/mbo3.313. The other qPCR reactions were designed for the purpose of kit for the detection of dysbiosis and/or B2 enterotype, and are shown in Table 20.

Table 20: The primers and probes (oligos) used for the qPCR reactions.

qPCR reaction	Oligo	Sequence (5' to 3')	Length (bp)	GC%	Melting temperature (°C)	Concentration in qPCR (µM)
1. Ruminococcus	Forward primer	TGTAGCGGTGAAATGCGTAGA	21	47,6	63,8	1.0
	Reverse primer	CCCCGTCAATTCCTTTGAGT	20	50.0	62,2	1.0
	Probe	FAM-GGATTAGATACCCTGGTAG-TCCACGC-QSY	26	53.8	66,9	0.4
2. Bacteroides	Forward primer	GGCGCACGGGTGAGTAAC	18	66.7	64.9	0.9
	Reverse primer	TGTGGGGGACCTTCCTCTC	19	63.2	64.7	0.9
	Probe	FAM-CGGGGTAACGGCCCA-QSY	15	73.3	63.3	0.25
3. Prevotella	Forward primer	CACAAGCGGAGGAACATGTG	20	55.0	63.2	0.5
	Reverse primer	ACGAGCTGACGACAACCATG	20	55.0	64.4	0.5
	Probe	FAM-GCGAGGAACCTTACCCGGGC-QSY	20	70.0	68.6	0.2
4. Faecalibacterium	Forward primer	GCAACGCGAAGAACCTTACC	20	55.0	63,4	0.5
	Reverse primer	CCCAACATCTCACGACACGA	20	55.0	63,9	0.5
	Probe	FAM-ACAGGTGGTGCATGGTTGTCGT-QSY	22	54.5	68,5	0.8
5. Oscillibacter	Forward primer	AGCGGTGAAATGCGTAGATAT	21	42,9	61.9	0.5
	Reverse primer	ACTACCAGGGTATCTAATCCTGTT	24	41.7	62.6	0.5
	Probe	FAM-ACTGACGGTGAGGCGCGAAA-QSY	20	60.0	68,4	0.2
6. Total bacteria	Forward primer	AATAAATCATAAACTCCTACGGGAGGCAG-CAGT	33	42,4	70.1	0.5
	Reverse primer	AATAAATCATAACCTAGCTATTAC-CGCGGCTGCT	34	41,2	70.2	0.5
	Probe	FAM-CGGCTAACTACGTGCCAG-QSY	18	61.1	61,9	0.2

"Tm" for the melting temperature here defined as the melting temperature of the oligo.

Materials and Methods

Reagents used: Ultrapure™ DNase/RNase-Free Distilled Water (10977035, Fisher Scientific), TE Buffer : Tris-EDTA pH 8.0 (BP2473-100, Fisher Scientific), TaqPath ProAmp mastermix (A30867, Life Technologies), MicroAmp™ qPCR plates (4346907, Thermo Fisher).

Instrument used: Quantstudio5, fast 0.1 mL 96-well system

The standards for the qPCR reactions consist of relevant 16S rRNA genes cloned into a plasmid. Synthetic 16S rRNA genes of a representative bacterial strain for each target genus (see Table 3 below) were cloned into a pCR(R)4Blunt-TOPO vector in TOP10 electrocompetent E. coli using Zero Blunt TOPO PCR Cloning

Kit (MAN0000110, ThermoFisher Scientific). Plasmid DNA was purified with a QIAprep spin Miniprep Kit (27106, Qiagen) and sequenced to verify the correct fragment insertion, identity and copy number.

Standard series were prepared as 10fold serial dilution for each stock of plasmid DNA in TE buffer to obtain a range of 10^8, 10^7, 10^6, 10^5, 10^4, 10^3, and 10^2 copies/µL in 1.5 mL DNA LoBind Eppendorf tubes (10051232, Fisher Scientific). This allows absolute quantification in this concentration range, which was previously determined to be the appropriate dynamic range for the qPCR reactions.
A random selection was made of 100 samples per enterotype (400 total) from VDP.3000 cohort (Flemish Gut Flora Project, FGFP) with known enterotypes assigned by the reference method next-generation sequencing of 16S amplicons (16S-NGS) clustered into four groups using Dirichlet Multinomial Mixtures (DirichletMultinomial package version 1.39.0) after agglomerating ASVs at genus level. The same DNA extracts from stools which were used for 16S-NGS were also used in this experiment to exclude variation from sample preparation (sampling and DNA extraction from stool), which is not in scope for concordance testing as here the aim is to directly compare methodologies on a given DNA sample without additional preprocessing or biological variation. The DNA extracts were diluted 10fold and 100fold in TE buffer for qPCR analysis.

Quality Control (QC) sample. To assess to performance of the qPCR reactions, one reference sample was repeated in 7/10 plates per reaction. This QC reference sample consists of a DNA extract from a random healthy donor (D23002) with a manual DNA extraction kit (MDKT00230096, Magtivio).
This sample contained approx. 11 log cells/g of which approx. 9 log cells/g *Bacteroides*, and was negative for *Ruminococcus* and *Prevotella*. The rest of the compositions was unknown (not tested), but *Faecalibacterium* is always expected. As such, the absent genera except *Faecalibacterium* were added to this sample by spiking the corresponding plasmids, resulting in the new sample S23087 containing 8 log copies/µL of *Prevotella copri*, 7 log copies/µL of *Ruminococcus*, and 6 log copies/µL of *Oscillibacter rumanitium*.

qPCR assay protocols. Master mix is prepared in the pre-PCR UV cabinet after cleaning the cabinet surface and materials with RNase AWAY Surface Decontaminant (10666421, Fisher Scientific) and UV lamp decontamination for 15 min using separated materials and gowning. The master mix is a commercial product (TaqPath ProAmp, A30867, Life Technologies). The product is double concentrated so that it can be diluted by adding the appropriate amounts of primers and probes (see Table 2 above), nuclease-free water and samples. Reactions are run in a total of 10 µL with 9 µL mastermix and 1 µL sample. For example, to the composition of 1 well for the Faecilbacterium reaction is provided below. In reality, the master mix is made in bulk, in the concordance study for 5 plates, and 9 µL is added to each

well with a Eppendorf multi-dispenser pipette (the Eppendorf™ Combitips 0.5 mL, 9 µL dispense, 48x). The plates with master mix are sealed and stored cold and dark until sample addition.

Mastermix (MM) composition for 1 well (10 µL) for the Faecalibacterium reaction

Contents	Stock concentration	Final concentration	Volume (µL)
TaqPath ProAmp master mix	2X	1X	5,0 µL
Forward primer	10 µM	0,5 µM	0,5 µL
Reverse primer	10 µM	0,5 µM	0,5 µL
Probe	10 µM	0,8 µM	0,5 µL
Nuclease-free water	-	-	2,2 µL
Template (sample, standard or water)	-	-	1,0 µL

Sample addition is performed in the bio-lab on the bench. Before and during testing the surfaces are thoroughly cleaned with RNase AWAY™ Surface Decontaminant. Samples of 1 µL are added to each well, either 1) DNA extracted from a stool sample, 2) Standards for quantification and positive control, or 3) nuclease-free water as negative control (No Template Control, NTC). All 3 samples were present in each plate. In the concordance study all samples were run in duplicate. The samples were added per 8 according to the plate layout, using the 12.5 µL electronic multichannel pipette (Integra) with the repeat dispense program (dispense = 1 µL, pre-dispense = 1 µL, post-dispense = 0.5 µL, asp. speed 6, re-use post dispense = no, count = 2).

Data analysis. The fluorescent signals from qPCR were converted to Ct values (Threshold Cycle) by the Quantstudio Design and Analysis software 1.5.2 using the automatic settings for baseline threshold determination for all reactions except *Ruminococcus*. A fixed threshold of 0.2 was used for the *Ruminococcus* reaction.

Standard curves were run in each plate in duplicate. Linear regression was performed on the means between the concentration in log copies/µL on the x-axis and the Ct value (raw data from qPCR) on the Y-axis. The acceptance criteria for a good fit was $R^2 \geq 0.99$.

The mean Ct values of unknown samples were converted using the equation of linear regression. When the standard deviation (SD) of duplicates was higher than 1 Ct value, the sample was flagged for elevated variation. When no Ct value was obtained, the sample was indicated as negative. If the mean Ct value was below the lowest standard, the sample is flagged as outside the quantification range. Since these values are not reliable, they are assigned the mean concentration between the lowest standard and zero, i.e. 1 log copies/µL. No samples were above the highest standard.

The results in copies/µL were multiplied by the dilution factor and converted to cells/µL using the generalized copy numbers per genome (for the total bacteria, *Ruminococcus*, *Bacteroides*, *Prevotella*,

Faecalibacterium, and *Oscillibacter* 4.8, 4.0, 4.9, 4.3, 6.0 and 3.0 16S copies/cell, were used respectively). These copy numbers were obtained by consulting and averaging the copy numbers per genome reported on NCBI and RDP for the species targeted for each qPCR reaction. For total bacterial, the average of the top 11 most abundant bacteria in human stool was used. The data in cells/µL was converted to a percentage by dividing all genus/species specific reactions by the total bacteria, resulting in the relative concentrations. The results in cells/µL were also converted to absolute concentrations in cells/g stool using the exact stool mass input for DNA extraction.

Claims

1. A method of detecting gut flora dysbiosis in a subject, said method comprising:
    a. Measuring in a biological sample obtained from said subject abundances of each of at least two bacterial genera, said at least two genera selected from the list consisting of *Sporobacter*, *Coprococcus*, *Fusicatenibacter*, *Ruminococcus*, *Oscillibacter* and *Faecalibacterium* and/or from the list consisting of *Clostridium_IV*, *Butyricimonas*, *Butyricicoccus*, *Prevotella*, *Roseburia*, *Flavonifractor* and *Bacteroides*, wherein said at least two bacterial genera do not consist of *Faecalibacterium* and *Bacteroides*;
    b. Comparing the measured abundances of each of the at least two bacterial genera of said biological sample to reference abundances of each of the at least two bacterial genera obtained from a plurality of control samples; and
    c. Determining that the subject suffers from gut flora dysbiosis, if:
        i. in case the at least two bacterial genera do not comprise *Bacteroides* and/or *Flavonifractor*, the abundances of each of the at least two bacterial genera as measured in step a) are decreased in the subject sample compared to the reference abundances; or
        ii. in case the at least two bacterial genera comprise *Bacteroides* and/or *Flavonifractor*, the abundances of Bacteroides and/or Flavonifractor as measured in step a) are increased in the subject sample compared to the reference abundances and the abundances of each of the other selected bacterial genera as measured in step a) are decreased in the subject sample compared to the reference abundances.
2. The method according to claim 1, wherein said step of measuring the abundances includes measuring of abundances of at least genera *Oscillibacter and Faecalibacterium*.
3. The method according to claim 2, wherein said step of measuring the abundances includes measuring of at least one further genus selected from a list consisting of *Ruminococcus, Prevotella,* and *Bacteroides*.
4. The method according to claim 2, wherein said step of measuring the abundances includes measuring of at least two further genera selected from the list consisting of *Ruminococcus, Prevotella,* and *Bacteroides*.
5. The method according to any one of claims 1 to 4, wherein the gut flora dysbiosis is associated with an inflammatory disorder, obesity, diabetes type 2, depression, and/or anxiety.
6. The method according to claim 5, wherein the gut flora dysbiosis is associated with the inflammatory disorder, wherein said inflammatory disorder is selected from the list consisting of spondyloarthritis, ankylosing spondylitis, reactive arthritis, psoriatic arthritis, enteropathic arthritis, undifferentiated

spondyloarthritis, juvenile idiopathic arthritis, primary sclerosing cholangitis, multiple sclerosis and any gut inflammation associated therewith.

7. The method according to claim 5, wherein the gut flora dysbiosis is associated with the inflammatory disorder, wherein the inflammatory disorder is a gut inflammatory disorder selected from the list consisting of Crohn's disease, irritable bowel syndrome, inflammatory bowel disease, ulcerative colitis and celiac disease.
8. The method according to any one of the preceding claims, wherein the biological sample is a stool sample, a mucosal biopsy sample or a sample of the lumen content.
9. The method according to any one of the preceding claims, wherein the quantification of the at least two bacterial genera preferably at least three or at least four bacterial genera is performed by quantifying DNA sequences specific for said at least two bacterial genera.
10. The method according to claim 9, wherein said DNA sequences are 16S rDNA sequences.
11. The method according to claim 9 or 10, wherein the quantification of the DNA sequences is performed by quantitative PCR.
12. The method according to claim 11, wherein the abundances of the at least two bacterial genera are expressed as the log cells/g.
13. The method according to any one of claims 9 to 12, wherein prior to quantification of said DNA sequences, DNA is extracted from the biological sample.
14. A bacterial genera panel comprising at least at least two bacterial genera selected from the list consisting of *Oscillibacter*, *Faecalibacterium*, *Ruminococcus*, *Prevotella* and *Bacteroides* wherein the panel does not consist of *Faecalibacterium* and *Bacteroides*.
15. The bacterial genera panel according to claim 14 for use to diagnose a subject with gut flora dysbiosis.
16. The bacterial genera panel according to claim 15 for determining gut flora dysbiosis in a subject the by the method according to any of claims 1 to 13.
17. A kit comprising:
    i) at least two pairs of nucleic acid primers designed for specifically amplifying DNA sequences of at least two different bacterial genera, wherein the bacterial genera are selected from a list consisting of *Ruminococcus*, *Oscillibacter*, *Faecalibacterium* *Prevotella* and *Bacteroides*; and
    ii) preferably instructions for quantifying the levels of two or more DNA sequences from a stool sample.
18. The kit according to claim 17 wherein said kit comprises at least two pairs of nucleic acid primers designed for specifically amplifying DNA sequences of at least two different bacterial genera,

wherein said bacterial genera are selected from a list consisting of *Faecalibacterium* and *Oscillibacter*.

19. The kit according to claim 18, wherein said kit comprises at least a further pair of nucleic acid primers designed for specifically amplifying DNA sequences of a further bacterial genus, said genus being selected from a list comprising *Prevotella*, *Ruminococcus* and *Bacteroides*.
20. The kit according to any one of claims 17 to 19, wherein the DNA sequences are 16S rDNA sequences or fragments thereof.
21. The kit according to any one of claims 17 to 20 for use in detecting gut flora dysbiosis.
22. The kit according to any one of claims 17 to 20 for use according to claim 21, wherein the gut flora dysbiosis is associated with an inflammatory disorder, obesity, diabetes type 2 or depression.
23. The kit according to any one of claims 17 to 20 for use according to claim 22, wherein said gut flora dysbiosis is associated with the inflammatory disorder, said inflammatory disorder being selected from the list consisting of spondyloarthritis, ankylosing spondylitis, reactive arthritis, psoriatic arthritis, enteropathic arthritis, undifferentiated spondyloarthritis, juvenile idiopathic arthritis, primary sclerosing cholangitis, multiple sclerosis and any gut inflammation associated therewith.
24. The kit according to any one of claims 17 to 20 for use according to claim 22, wherein the inflammatory disorder is a gut inflammatory disorder selected from the list consisting of Crohn's disease, irritable bowel syndrome, inflammatory bowel disease, ulcerative colitis and celiac disease.

Figure 1

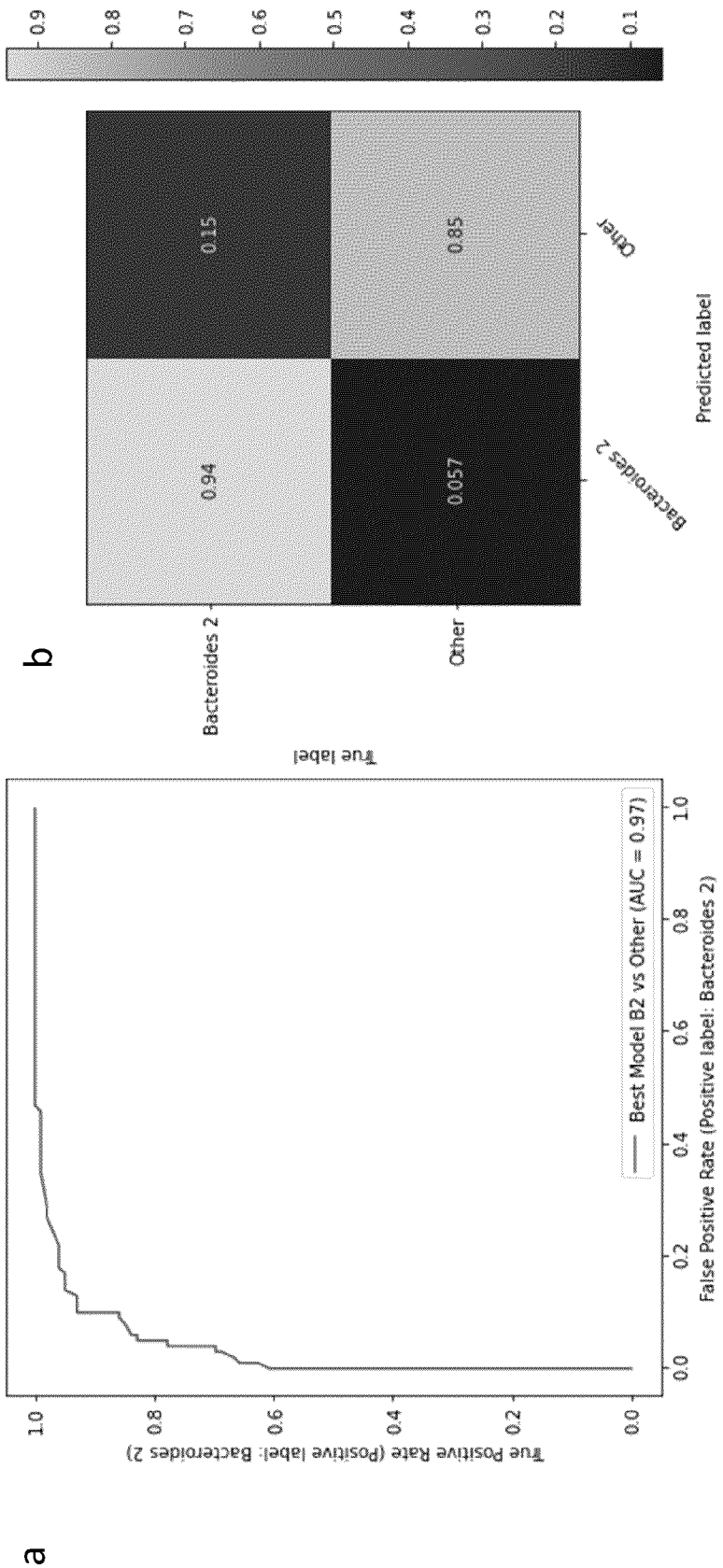

a
True Positive Rate (Positive label: Bacteroides 2)
False Positive Rate (Positive label: Bacteroides 2)
Best Model B2 vs Other (AUC = 0.97)
b
True label
Bacteroides 2
Other
Predicted label
0.94
0.057
0.15
0.85

Figure 2

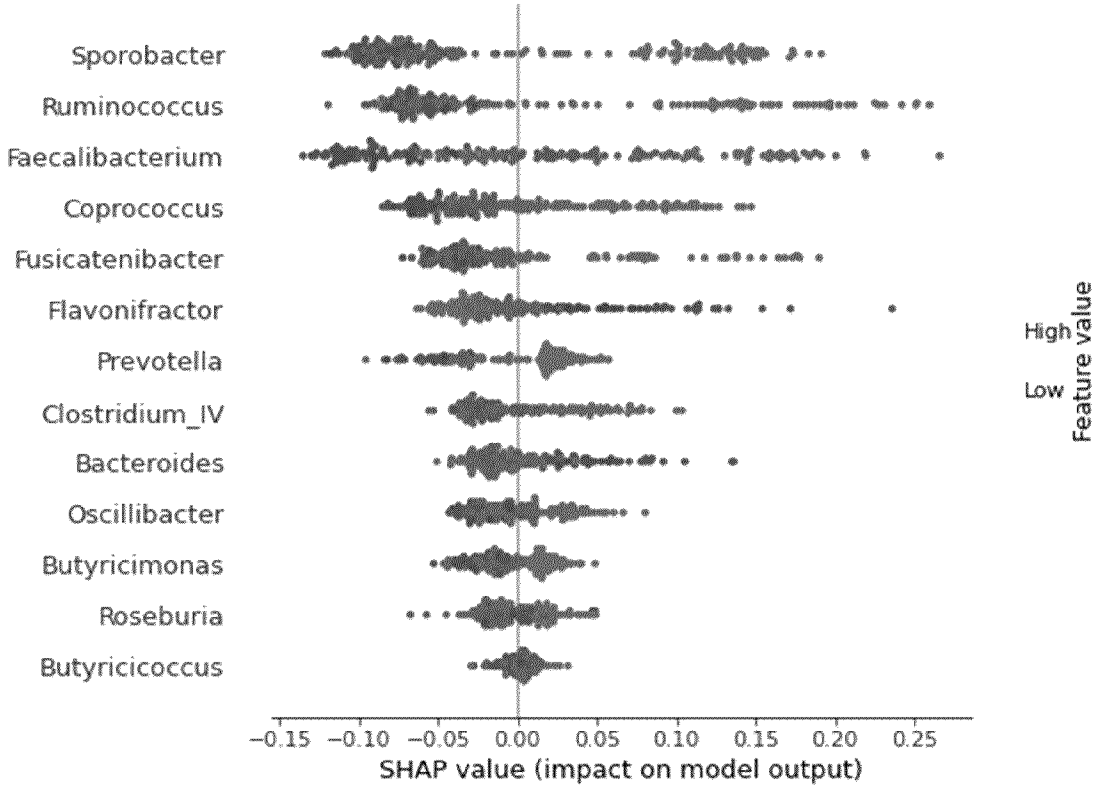


Figure 3

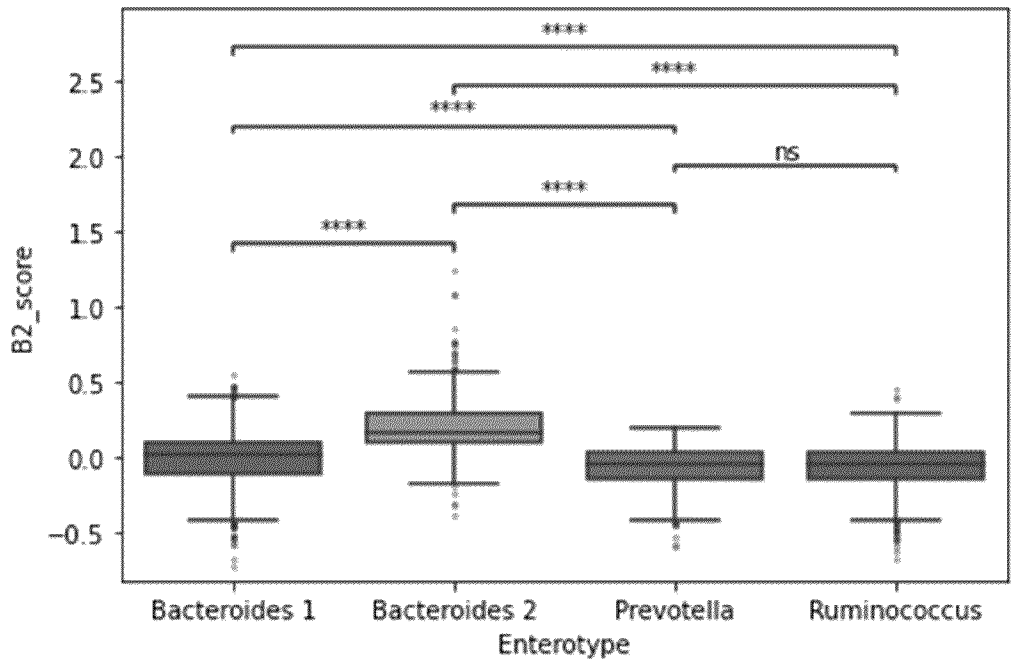


Figure 4

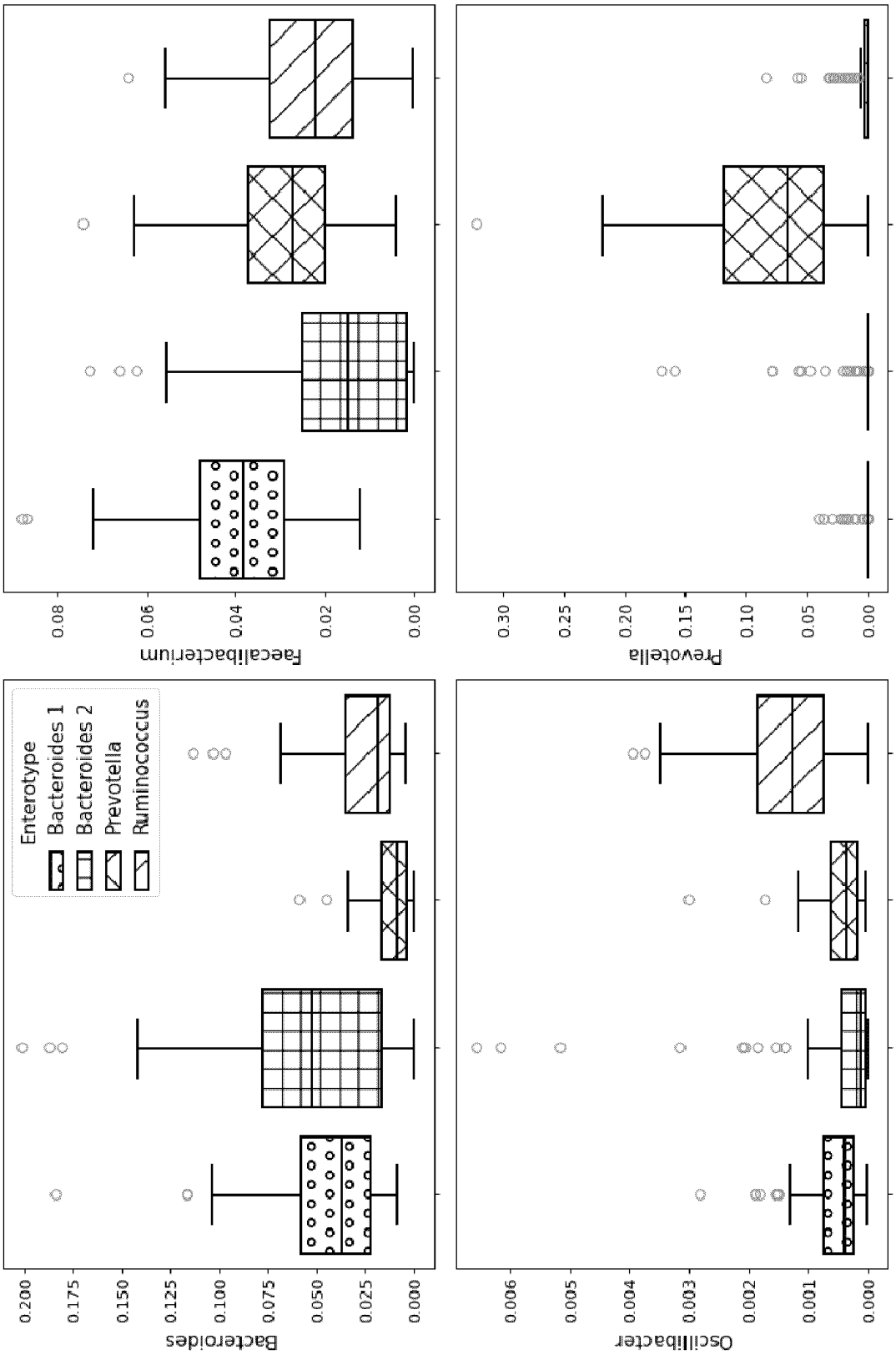

Enterotype
Bacteroides 1
Bacteroides 2
Prevotella
Ruminococcus
Bacteroides
Oscillibacter
Faecalibacterium
Prevotella

Figure 4 continued

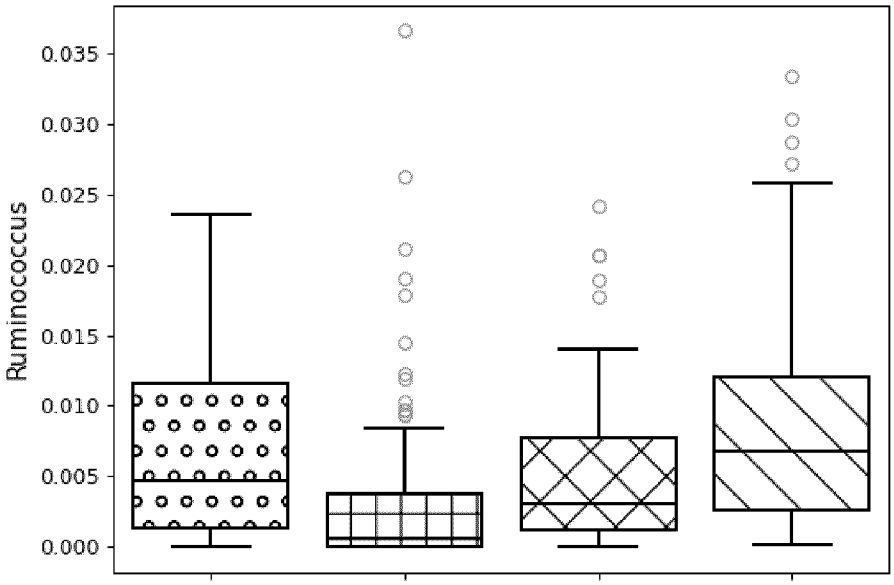


Figure 5

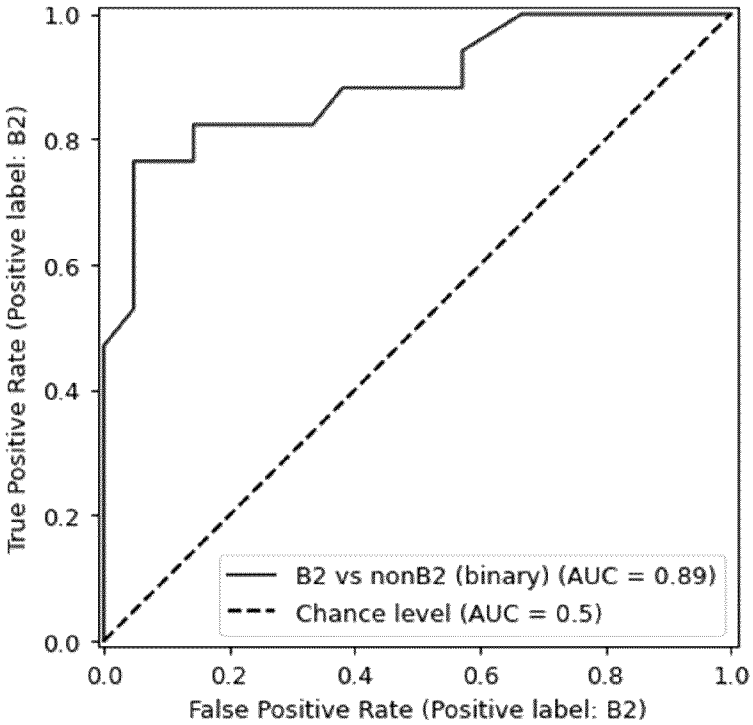
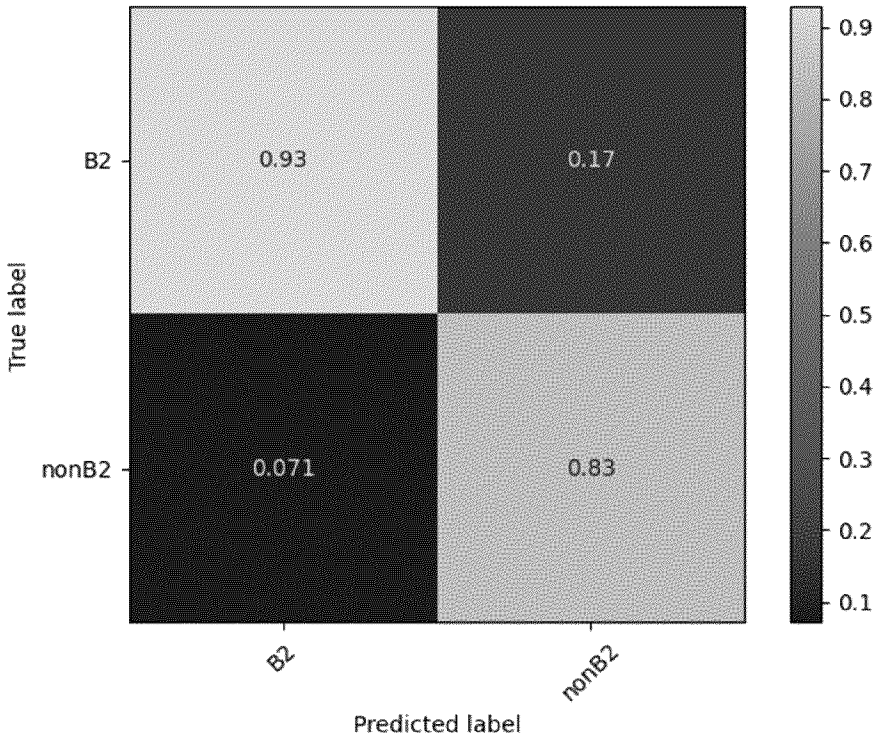


Figure 6


B2
nonB2
True label
0.93
0.17
0.071
0.83
B2
nonB2
Predicted label
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2024/086588

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/086588

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/689
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2023/192815 A2 (BAYLOR COLLEGE MEDICINE [US]) 5 October 2023 (2023-10-05) claims 60-65, par 25-28	14-17, 21-24
X	US 2019/388482 A1 (URBAN RANDALL J [US] ET AL) 26 December 2019 (2019-12-26) table 2	14-17, 20-24
X	S. MONDOT ET AL: "Highlighting new phylogenetic specificities of Crohn?s disease microbiota", HHS PUBLIC ACCESS AUTHOR MANUSCRIPT, vol. 17, no. 1, 18 August 2010 (2010-08-18), pages 185-192, XP055234413, US ISSN: 1078-0998, DOI: 10.1002/ibd.21436 the whole document	14-17, 20-24

-/--

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 April 2025	17/04/2025

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Cornelis, Karen

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/086588

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOROBEINIKOVA ANNA V. ET AL: "Gut Microbiota Patterns in Patients with Non-Alcoholic Fatty Liver Disease: A Comprehensive Assessment Using Three Analysis Methods", INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES, vol. 24, no. 20, 17 October 2023 (2023-10-17), page 15272, XP093266864, Basel, CH ISSN: 1422-0067, DOI: 10.3390/ijms242015272 Retrieved from the Internet: URL:https://www.mdpi.com/1422-0067/24/20/15272/pdf> the whole document	14-24
A	WO 2019/115755 A1 (VIB VZW [BE]; UNIV LEUVEN KATH [BE]; UNIV BRUSSEL VRIJE [BE]) 20 June 2019 (2019-06-20)	1-24
A	WO 2019/014714 A1 (SMARTDNA PTY LTD [AU]) 24 January 2019 (2019-01-24)	1-24
A	OH HYUN-SEOK ET AL: "Proposal of a health gut microbiome index based on a meta-analysis of Korean and global population datasets", JOURNAL OF MICROBIOLOGY, THE MICROBIOLOGICAL SOCIETY OF KOREA, SEOUL, vol. 60, no. 5, 31 March 2022 (2022-03-31), pages 533-549, XP037813812, DOI: 10.1007/S12275-022-1526-0 [retrieved on 2022-03-31]	1-24

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2024/086588

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2023192815 A2	05-10-2023	EP 4499878 A2	05-02-2025
		WO 2023192815 A2	05-10-2023
US 2019388482 A1	26-12-2019	US 2019388482 A1	26-12-2019
		US 2021275608 A1	09-09-2021
WO 2019115755 A1	20-06-2019	EP 3724358 A1	21-10-2020
		US 2021172006 A1	10-06-2021
		WO 2019115755 A1	20-06-2019
WO 2019014714 A1	24-01-2019	AU 2018253466 A1	31-01-2019
		AU 2019203763 A1	20-06-2019
		AU 2020202974 A1	21-05-2020
		AU 2023202551 A1	18-05-2023
		CN 110892081 A	17-03-2020
		EP 3655544 A1	27-05-2020
		JP 7317821 B2	31-07-2023
		JP 2020527055 A	03-09-2020
		JP 2023089141 A	27-06-2023
		SG 11201913701V A	27-02-2020
		US 2020181674 A1	11-06-2020
		US 2023130204 A1	27-04-2023
		WO 2019014714 A1	24-01-2019