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

(19) World Intellectual Property Organization
International Bureau

(10) International Publication Number
WO 2017/182796 A1

(43) International Publication Date
26 October 2017 (26.10.2017)

(51) International Patent Classification:
A61K 35/742 (2015.01) *A61P 31/04* (2006.01)

(21) International Application Number:
PCT/GB2017/051083

(22) International Filing Date:
19 April 2017 (19.04.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1606801.7 19 April 2016 (19.04.2016) GB

(71) Applicant: GENOME RESEARCH LIMITED [GB/GB]; Wellcome Trust Genome Campus, Hinxton, Cambridge Cambridgeshire, CB10 1SA (GB).

(72) Inventors: **NEVILLE, Anne**; Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge Cambridgeshire CB10 1SA (GB). **BROWNE, Hilary**; Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge Cambridgeshire CB10 1SA (GB). **FORSTER, Sam**; Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge Cambridgeshire CB10 1SA (GB). **LAWLEY, Trevor**; Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge Cambridgeshire CB10 1SA (GB).

(74) Agent: **KEIRSTEAD, Tanis** et al.; Mewburn Ellis LLP, City Tower, 40 Basinghall Street, London Greater London EC2V 5DE (GB).

(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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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).

Declarations under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *with international search report (Art. 21(3))*

(54) Title: BACTERIOTHERAPY

(57) Abstract: The invention relates to therapeutic compositions comprising at least one isolated bacterium and a pharmaceutically acceptable excipient, as well as methods of preparing such therapeutic compositions. The therapeutic compositions find application in the treatment of dysbiosis, in particular dysbiosis of the gastrointestinal tract.

Bacteriotherapy

Field of the Invention

5 The present invention relates to therapeutic compositions comprising at least one isolated bacterium as defined herein and a pharmaceutically acceptable excipient, as well as methods of preparing such therapeutic compositions. The therapeutic compositions find application in the treatment of dysbiosis, in particular dysbiosis of the gastrointestinal tract. The dysbiosis may be a dysbiosis associated with an enteric bacterial infection, inflammatory 10 bowel disease, pouchitis, irritable bowel syndrome, a metabolic disease, a neuropsychiatric disorder, an autoimmune disease, an allergic disorder, a cancer, or hepatic encephalopathy.

Background to the invention

15 A typical human intestinal microbiota contains 100-1000 bacterial species. There is extensive compositional diversity between individuals, such that each individual's microbiota is as unique as a fingerprint (Qin, Li et al. 2010; Nielsen, Almeida et al. 2014). The majority of the bacterial species within the adult human microbiota are derived from four high level taxonomic classifications or phyla, the Firmicutes, Bacteroidetes, Actinobacteria and 20 Proteobacteria. These groups change in abundance from birth to adulthood to old age, reflecting changing environmental influences such as initial mode of delivery, diet, insults such as pathogen infection and in many cases antibiotic usage (Dominguez-Bello, Costello et al. 2010, Koenig, Spor et al. 2011, Ottman, Smidt et al. 2012). In adulthood, the intestinal microbiota is dominated by the Firmicutes and the Bacteroidetes, both of which are strict 25 anaerobes.

The intestinal microbiota plays a key role in digesting food inaccessible to the human gastrointestinal tract, such as metabolizing carbohydrates into short chain fatty acids (Sekirov, Russell et al. 2010), interacting with the immune system to maintain homeostasis 30 (Hooper, Littman et al. 2012), promoting maturation of the gut (Hooper, Wong et al. 2001) and development of the immune system. The intestinal microbiota also plays an important role in resisting pathogen invasion, termed 'colonisation resistance'. This functions through the diversity and abundance of commensal species present and through the occupation of key niches and utilization of nutrients (Lawley and Walker 2013; Britton and Young 2014). If 35 microbial homeostasis is disturbed, for example through use of antibiotics, a shift towards dysbiosis can occur.

Dysbiosis provides the opportunity for pathogens to establish themselves and cause disease to the individual in question. This has been best studied in terms of a single implicated pathogen such as *Clostridium difficile* (Lawley et al. 2012; Britton and Young 2014; Buffie et al. 2015), but dysbiosis has also been linked with other more complex, multi-factorial diseases such as Inflammatory Bowel Disease (IBD), pouchitis (Angeriman et al. 2014), Irritable Bowel Syndrome (IBS), hepatic encephalopathy (Bajaj 2014; Bajaj et al. 2012) metabolic diseases (including metabolic syndrome, malnutrition, and obesity), neuropsychiatric disorders such as Parkinson's and Alzheimer's disease, autoimmune diseases, allergic disorders, and cancer (Jostins, Ripke et al. 2012, Collins 2014, Hold, Smith et al. 2014, Perez Martinez, Bauerl et al. 2014, Scheperjans, Aho et al. 2015; Blanton et al. 2016, Xu et al. 2015).

Faecal microbiota transplantation (FMT) has proved successful in resolving *C. difficile* associated dysbiosis (Petrof et al. 2013, van Nood et al. 2013), and the administration of specific bacteria has also proved effective for this purpose (Lawley et al. 2012, Buffie et al. 2015). FMT has also showed promising results in the treatment of other intestinal diseases, as well as the management of extra-intestinal disorders associated with gut microbiota, including metabolic diseases, neuropsychiatric disorders, autoimmune diseases, allergic disorders, and tumours (Xu et al. 2015).

Recent years have seen great advances in understanding the role the intestinal microbiota plays in health and disease and how it can be manipulated for the benefit of the host. The majority of our understanding has to date been derived by culture-independent studies, i.e. by studying the compositional components of the microbiota and how they change during disease using molecular and genomic techniques. This process allows identification of potential therapeutic candidates that can resolve disease. However, the isolation, purification and acquisition of such candidate therapeutic bacteria has proven difficult.

There is therefore a need in the art to identify and isolate specific bacteria, as well as combinations of bacteria, which can be used to treat dysbiosis. Therapeutic compositions based on known, defined, bacteria or bacterial mixtures are advantageous as they improve patient safety because they comprise only defined and well characterised bacteria that are known to promote, and not harm, human health, and eliminate the possibility of inadvertently transferring pathogenic material to a recipient by FMT. In addition, such therapeutic compositions can be prepared *in vitro* in a large-scale manner using standardised,

reproducible procedures, thereby providing batch consistency, and do not rely on regular donations from healthy human donors. Therapeutic compositions comprising known, defined, bacteria or bacterial mixtures can also be therapeutically delivered e.g. in a capsule, as a tablet, or as an enema, which is more acceptable to patients and health care professionals than suspensions of faecal material used in the case of FMT. The bacteria included in such therapeutic compositions can further be tailored to the treatment of specific dysbiotic states and diseases associated therewith by specifically altering the bacterial composition to optimally resolve the dysbiotic state in question and thus improve efficacy.

However, in order to isolate such candidate therapeutic bacteria for the treatment of dysbiosis, a thorough understanding of the biology of the candidates in question is required, as well as a large initial panel of candidates to select from. This poses a problem as the majority of the bacteria in the intestinal microbiota are considered to be unculturable and have never been isolated in the laboratory (Eckburg, Bik et al. 2005, Hattori and Taylor 2009, Stewart 2012). Thus, gaining a basic understanding of the functional attributes of the microbiota and developing a multi-species bacteria-based therapeutic with fastidious, anaerobic commensal isolates presents a formidable challenge. While recent efforts have made progress in resolving this issue (Goodman, Kallstrom et al. 2011, Lagier, Hugon et al. 2015), there remains a need in the art to identify and isolate bacteria capable of treating dysbiosis.

Statements of invention

The present invention relates to therapeutic compositions, in particular therapeutic compositions for use in the treatment of dysbiosis in an individual. Dysbiosis can occur in any part of the human or animal body which is normally colonized by bacteria and other microbes. The present invention particularly concerns dysbiosis of the gastrointestinal tract in humans.

The present inventors have surprisingly found that the majority of bacteria present in the human intestinal microbiota can be cultured, contrary to the prevailing view in the art which was that the majority of the human intestinal microbiota is unculturable. This major breakthrough now allows the majority of bacteria present in the human microbiota to be isolated and characterised, and evaluated for their activity in treating dysbiosis. This is possible not only for individual bacterial isolates but also for combinations of bacteria isolated from the intestinal microbiota. In addition, isolation of these bacteria allows the

bacteria to be screened, for example, for the absence of virulence factors and antibiotic resistance prior to their inclusion in a therapeutic composition, thereby improving safety. In addition, the bacteria included in a therapeutic composition can be tailored to the treatment of a specific dysbiotic state and/or disease associated therewith by optimising the bacterial composition to resolve the dysbiosis in question, thereby improving efficacy. None of this is possible in FMT where undefined mixes of bacteria are used, usually obtained from a faecal sample of a healthy human donor. The use of isolated bacteria for the treatment of dysbiosis has the further advantage that it allows the bacteriotherapy treatment to be standardised, making patient outcomes more predictable, as well as facilitating evaluation of the therapeutic potential of bacteriotherapy in the context of particular diseases by removing the variability in bacterial composition associated with the use of FMT.

Through surprisingly being able to culture the majority of bacteria present in the human intestinal microbiota, the present inventors were able to prepare libraries of intestinal bacteria which were then subjected to whole-genome sequencing and screened using both *in silico* analysis and *in vitro* experiments to identify bacteria which are expected to be useful in treating dysbiosis, in particular dysbiosis of the gastrointestinal tract. Using this approach, the present inventors identified 51 bacteria which are expected to be useful for this purpose, including several families, genera, and species of bacteria which have not previously been described, let alone isolated or employed in the treatment of dysbiosis. As already explained above, the majority of the human microbiota was thought in the art to be unculturable, so the mere disclosure of a 16S ribosomal RNA sequence of one of these bacteria does not in itself enable the isolation of such a bacterium from its natural environment. Nor does the disclosure of such a 16S ribosomal RNA sequence suggest that a bacterium with such a sequence has previously been isolated, as 16S ribosomal RNA sequence information can be obtained from bacterial populations, including faecal samples, without the need to isolate individual bacteria. However, the ability to isolate bacteria in pure form from their natural environment is a prerequisite for their inclusion in therapeutic compositions according to the present invention.

Thus, in a first aspect, the present invention provides a therapeutic composition comprising at least one isolated bacterium and a pharmaceutically acceptable excipient. The bacterium preferably comprises a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51.

The therapeutic composition may comprise more than one isolated bacterium, in addition to the pharmaceutically acceptable excipient. Where more than one bacterium is included in the therapeutic composition, the bacteria are preferably distinct, wherein each bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51. Thus, for example, the therapeutic composition may comprise two distinct isolated bacteria, wherein the first bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1 and the second bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 2.

As mentioned above, the therapeutic compositions of the present invention find application in the treatment of dysbiosis, in particular the treatment of a dysbiosis of the intestinal tract. Thus, in a second aspect, the present invention provides a therapeutic composition according to the invention for use in a method of treating a dysbiosis, preferably a dysbiosis of the gastrointestinal tract, in an individual. Also provided is a method of treating a dysbiosis in an individual, the method comprising administering a therapeutically effective amount of a therapeutic composition according to the invention to an individual in need thereof, as well as the use of a therapeutic composition according to the present invention for the manufacture of a medicament for the treatment of a dysbiosis in an individual. Also provided is the use of at least one isolated bacterium, as described herein, and optionally a pharmaceutically acceptable excipient, for the manufacture of a medicament for the treatment of a dysbiosis in an individual, the bacterium preferably comprising a gene encoding a 16S rRNA and said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51.

Methods of preparing or manufacturing a therapeutic composition according to the present invention also form part of the invention. Thus, in a third aspect, the present invention provides a method of preparing or manufacturing a therapeutic composition according to the present invention, wherein the method preferably comprises the steps of:

(i) culturing an isolated bacterium as described herein; and

(ii) mixing the bacteria obtained in (i) with a pharmaceutically acceptable excipient.

As mentioned above, the therapeutic compositions of the present invention may comprise at least two distinct isolated bacteria as described herein. Where the therapeutic composition

comprise more than one distinct isolated bacteria, the method of preparing or manufacturing a therapeutic composition preferably comprises steps of:

(i) culturing a first isolated bacterium as described herein;

(ii) culturing a second isolated bacterium as described herein; and

5 (ii) mixing the bacteria obtained in (i) and (ii) with a pharmaceutically acceptable excipient. The bacteria cultured in steps (i) and (ii) preferably have distinct 16S rRNA sequences. Steps (i) and (ii) are preferably performed independently. The above method can be adapted to include further steps to allow the culturing of more than two distinct isolated bacteria, preferably bacteria with distinct 16S rRNA sequences, by including an additional
10 step or steps for the culturing of a third or further isolated bacterium as disclosed herein. In this case, all bacteria cultured in the method are mixed with a pharmaceutically acceptable excipient.

A therapeutic composition obtainable by a method of preparing or manufacturing a
15 therapeutic composition, as disclosed herein also forms part of the present invention.

Brief Description of the Figures

Figure 1 shows a schematic diagram of the workflow used to culture, archive and
20 characterise the intestinal microbiota. The process incorporates several steps which are: culture, re-streak, archive and phenotype. (**A**) Fresh faecal samples were left untreated or were treated to select for bacteria with a desired phenotype (such as sporulation). The stool was homogenised and then serially diluted before aliquots of the homogenate were inoculated on YCFA agar to culture the bacteria present in the faecal samples. (**B**) Bacterial
25 isolates were identified by selecting single colonies that were then streaked to purity before full-length 16S rRNA gene amplification and sequencing was performed. (**C**) Each unique, novel and desired bacterial isolate was archived frozen in a culture collection and a whole genome sequence was generated for each. (**D**) Phenotypic characterisation and functional validation of metagenomics studies was then performed using *in vitro* and *in vivo* methods.

30

Figure 2: Targeted phenotypic culturing facilitates bacterial discovery from healthy human faecal microbiota. Figure 2 shows the relative abundance of bacteria in faecal samples (x axis) compared to the relative abundance of bacteria growing on YCFA agar plates (y axis) as determined by metagenomic sequencing. The results demonstrate that the bacteria
35 grown on YCFA agar are representative of the bacteria present in complete faecal samples as indicated by Spearman Rho =0.75.

Figure 3 shows a principal component analysis (PCoA) plot of 16S rRNA gene sequences detected from 6 donor faecal samples representing bacteria in the complete faecal samples (unfilled circles), faecal bacterial colonies recovered from YCFA agar plates without ethanol pre-treatment (filled black squares) or with ethanol pre-treatment to select for ethanol-resistant spore-forming bacteria (filled black circles). These results demonstrate that culturing without ethanol selection is representative of the complete faecal sample, while ethanol treatment shifts the profile, enriching for ethanol-resistant spore-forming bacteria and allowing their subsequent isolation.

Figure 4: Archiving of bacterial diversity and novelty through anaerobic culturing. **Figure 4A** shows that the culture conditions employed by the present inventors enabled isolation of representatives from 21 of the 25 most abundant faecal bacterial genera as determined by metagenomic sequencing. A black dot indicates the number of species cultured and archived from each genus. Lachnospiraceae incertae sedis, unclassified Lachnospiraceae, Clostridium IV and Clostridium XI are not strict genera and represent currently unclassified species. **Figure 4B** shows the 24 most abundant bacterial species (comprising 90% of the total bacterial abundance at the species level) as determined by metagenomic sequencing. All were cultured and archived except for *Odoribacter splanchnicus*. **Figure 4C** shows that intestinal microbiota members present at low abundance were also cultured. At least one representative species from each of the genera presented was cultured. Genera are listed in order of decreasing abundance.

Figure 5: Phylogeny of cultured and archived bacteria from healthy human faecal microbiota. Figure 5 shows a phylogenetic tree of bacteria cultured from the 6 donors constructed from full length 16S rRNA gene sequences. Novel candidate species (filled black circles), genera (grey filled circles) and families (filled stars) are shown by dot colours. Major phyla and family names are indicated. Proteobacteria were not cultured, but are included for context.

Figure 6: Faecal Microbiota Transplant (FMT) restores the intestinal microbiota of patients with recurrent *C. difficile* infection to a healthy state. **Figure 6** shows the results of a principal component analysis of the donors, recipients and controls at 2-3 months after FMT. The clustering of the faecal samples indicates a similar microbial community structure. Antibiotic use and exposure to *C. difficile* likely leads to a shift from the healthy state, as seen in the metronidazole-treated control samples. Treatment of *C. difficile* infection (CDI) with

vancomycin, an antibiotic that targets Gram-positive organisms would presumably lead to further disruption of the intestinal microbiota. FMT led to a shift from a diseased microbiota to a healthy one with most of the post-FMT samples clustering with the donors and healthy control samples.

5

Figure 7: Taxonomic summary of the bacteria isolated from faecal samples of the donors, recipients and controls from a study to treat *C. difficile* using FMT. These bacterial isolates represent a large cross-section of intestinal microbiota diversity.

10 **Figure 8:** Average relative abundance of bacteriotherapy candidates in healthy individuals. The bacteriotherapy candidates occur at an average abundance of greater than 0.001% within the gastrointestinal microbiota of 1883 healthy individuals (3218 samples).

Figure 9: The bacteriotherapy candidates are depleted in dysbiotic and disease states.
15 The average fold change in the relative abundance of each bacteriotherapy candidate in diseased and dysbiotic states in comparison to its relative abundance in a healthy microbiota is plotted. *Escherichia coli*, a marker of dysbiosis is also included for comparison. The relative abundance of a bacterium refers to the proportion of the total microbiota represented by the bacterium in question.

20

Figure 10 demonstrates how the zone of clearing around bacteriotherapy candidates was measured in the overlay assays. Figure 10 shows part of a YCFA agar plate on which a bacteriotherapy candidate was streaked in an X-shape and allowed to grow. Following growth of the bacteriotherapy candidate, the plate was covered with overlay agar comprising
25 *C. difficile* or *E.coli*. Inhibition of *C. difficile* or *E.coli* growth by a bacteriotherapy candidate was measured by determining the width of the zone of clearing around the bacteriotherapy candidate strain grown on the plate. The black diagonal line in Figure 10 indicates the distance measured and recorded as the width of the zone of clearing for an exemplary bacteriotherapy candidate. Four such measurements were taken per plate.

30

Figure 11 shows the results of *C. difficile* and *E. coli* growth overlay assays to determine the anti-pathogen activity of bacteriotherapy candidates. The zones of clearing were measured with a ruler as described in Figure 10. Millimetre (mm) measurements were taken. Figure 11 shows the mean measurement ± standard deviation from a representative experiment.

35

Figure 12 shows the results of *C. difficile* and *E. coli* growth inhibition assays to determine the anti-pathogen activity of bacteriotherapy candidates. **Figure 12A** shows the reduced relative *C. difficile* growth in Cell Free Supernatant (CFS) of bacteriotherapy candidate cultures at the 18.17h time-point, while **Figure 12B** shows reduced relative *E. coli* growth in
5 CFS from the candidate bacteriotherapy isolates. The relative growth of either pathogen in the control YCFA medium was high (*C. difficile* = 8.96±0.39 rel. growth units; *E. coli* 11.61±2.55 rel. growth units, at the 18.17h time-point). When the mean and two standard deviations of relative growth of either pathogen in CFS derived from a bacteriotherapy candidate culture was more than two standard deviations below its mean growth in YCFA at
10 the 18.17 h time-point, growth of the pathogen was considered to be inhibited. Where only one relative growth value was available for a particular CFS (vs *C. difficile*, Figure 12A: HMI 15, HMI 26, HMI 27, HMI 28), the bacteriotherapy candidate was considered to be inhibitory if the relative growth of *C. difficile* was two standard deviations below the mean growth of *C. difficile* in YCFA broth. Only the results from inhibitory CFS are shown.
15

Figure 13 shows a summary of the results obtained in the growth overlay and growth inhibition assays. Bacterial isolates shown to have inhibitory activity in the *E. coli* (AIEC) overlay assay, *C. difficile* overlay assay and *C. difficile* and *E. coli* growth assays are shown. Bacterial isolates showing inhibitory activity in two or more assays are shown in the
20 overlapping regions. Bacterial isolates are referred to by their isolate number. See Table 1 for details of the bacterial isolates listed.

Figure 14: Dendrogram and bar plots showing the relative abundance of each genus indicated in **Figure 14** at 3 months post-FMT in both donors and recipients. The dendrogram
25 clusters samples based on the phylogenetic relationship of the microbial community present within the samples. The composition of donor and recipient profiles were similar when assessed at the genus level post-FMT.

Detailed Description
30

The bacteria disclosed herein have been obtained from human stool samples, and thus are naturally present in the gastrointestinal tract of at least some healthy human individuals. However, these bacteria have been cultured *in vitro* for the first time by the present inventors, thereby isolating them from their environment in pure form, and making it possible
35 to include them as defined active ingredients in therapeutic compositions. The bacterium present in the therapeutic composition of the present invention is thus isolated. In other

words, the bacterium present in the therapeutic composition is provided in isolated and/or purified form, e.g. isolated and/or purified from the environment in which it is normally present, such as the gastrointestinal tract and/or stool samples. The isolated bacterium present in the therapeutic composition may be in substantially pure, or in homogeneous form. For example, the bacterium may be free, or substantially free, of material with which it is found in the environment in which it is normally present (e.g. the gastrointestinal tract and/or stool samples).

The bacterium present in the therapeutic composition of the present invention is preferably a human intestinal bacterium, i.e. a bacterium found in the human intestine. The bacteria whose 16S rRNA gene sequences are set out in SEQ ID NOs 1 to 51 are all intestinal bacteria.

The bacterium is preferably a non-pathogenic bacterium. In other words, the bacterium preferably does not cause disease in a healthy human individual when administered to said individual, in particular the gastrointestinal tract of said individual. The therapeutic composition can be administered to an individual in a variety of ways as described in more detailed elsewhere herein, including in the form of a tablet or enema.

The bacterium present in the therapeutic composition of the present invention is preferably susceptible to treatment with one or more antibiotics. In other words, the bacterium is preferably not resistant to treatment with at least one antibiotic. This allows antibiotic treatment of an individual in the event that one or more of the bacteria included in a therapeutic composition administered to the individual causes disease in the individual, contrary to expectations. All of the 51 bacteria disclosed herein where found to carry no known genes conferring resistance to the following antibiotics: beta-lactams, fusidic acid, elfamycin, aminoglycoside, fosfomycin, and tunicamycin. Thus, in a preferred embodiment, the bacterium is susceptible to treatment with one or more antibiotics selected from the group consisting of: a beta-lactam, fusidic acid, elfamycin, aminoglycoside, fosfomycin, and tunicamycin. *In vitro* and *in silico* methods for screening bacteria for antibiotic resistance are known in the art. Exemplary *in silico* methods are also described in Example 1.

The bacterium included in the therapeutic composition of the present invention preferably does not comprise one or more genes encoding one or more virulence factors and/or preferably does not produce one or more virulence factors. Virulence factors in this context are properties which enhance the potential of a bacterium to cause disease in an individual.

Virulence factors include the production of bacterial toxins, such as endotoxins and exotoxins by a bacterium, as well as the production of hydrolytic enzymes that may contribute to the pathogenicity of the bacterium. Methods for screening bacteria for genes encoding virulence factors are known in the art and include the *in silico* methods described in Example 1. The 51 bacteria disclosed herein were found not to carry any known virulence factors using *in silico* analysis. Methods for screening bacteria for the production of virulence factors are similarly known in the art.

Bacteria can be taxonomically classified based on the sequence of the gene encoding the 16S ribosomal RNA (rRNA) in the bacterium. This gene sequence is also referred to as the ribosomal DNA sequence (rDNA). A bacterium comprising a gene which encodes a 16S rRNA which has 90% or more sequence identity with the 16S rRNA encoded by a second bacterium belongs to the same family as said second bacterium. A bacterium comprising a gene which encodes a 16S rRNA which has 95% or more sequence identity with the 16S rRNA encoded by a second bacterium belongs to the same genus as said second bacterium. A bacterium comprising a gene which encodes a 16S rRNA which has 97% or more, or 98.7% or more sequence identity with the 16S rRNA encoded by a second bacterium belongs to the same species as said second bacterium. A bacterium included in the therapeutic composition of the present invention may be a bacterium which belongs to the same family, genus, and/or species as a bacterium disclosed herein.

A bacterium which belongs to the same family, genus, and/or species as a bacterium disclosed herein is expected to retain one or more properties of the disclosed bacterium. Thus, in a preferred embodiment, a bacterium present in the therapeutic composition of the present invention belongs to the same family, genus, and/or species as a bacterium disclosed herein and retains at least one property of the bacterium disclosed herein. Various properties of the bacteria disclosed herein are described and include, for example, a lack of production of one or more virulence factors, susceptibility to treatment with one or more antibiotics, and a lack of pathogenicity.

The therapeutic composition of the present invention may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51.

For example, the therapeutic composition of the present invention may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51. In a preferred embodiment, the therapeutic composition comprises an isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, and wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1. In addition, or alternatively, the therapeutic composition may comprise an isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, and wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 21.

In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51. In a preferred embodiment, the therapeutic composition comprises at least one isolated bacterium, wherein said bacterium comprises a gene encoding a 16S rRNA, and wherein said gene comprises a sequence with at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 2 to 20, or 22 to 51, more preferably any one of SEQ ID NOs 5, 6, 11, 13, 14, 15, 17, 18, 19, 20, 22, 23, 24, 26, 29, 33, 35, 41, 43, 45, 46, 47, 49, or 50, yet more preferably any one of SEQ ID NOs 5, 6, 11, 13, 15, 19, 22, 23, 29, 33, 35, 41, 43, 45, 46, or 50.

In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 97%, or at least 98.7%, sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51. In a preferred embodiment, the therapeutic composition comprises at least one isolated bacterium, wherein said bacterium comprises a gene encoding a 16S rRNA, and wherein said gene comprises a sequence with at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 2 to 20, or 22 to 51, more preferably any one of SEQ ID NOs 2 to 3, 5 to 8, 10 to 20, 22 to 26, 29 to 37, or 39 to 50, yet more preferably any one of SEQ ID NOs 3, 5 to 8, 10 to 13, 15, 16, 19, 22, 23, 29, 32 to 37, 39 to 46, or 48 to 50. In an alternative preferred embodiment, the therapeutic composition may comprise at least one isolated bacterium, wherein said bacterium comprises a gene encoding a 16S rRNA, and wherein said gene comprises a sequence with at least 98.7% sequence identity with the sequence set forth in any one of

SEQ ID NOs 2 to 20, or 22 to 51, more preferably any one of SEQ ID NOs 2 to 4, 5 to 20, 22 to 26, 29 to 37 to 51, yet more preferably any one of SEQ ID NOs 2 to 8, 10 to 13, 15, 16, 17, 19, 20, 22, 23, 29, 31, 32 to 37 to 46, or 48 to 51.

5 As mentioned above, in a preferred embodiment, the therapeutic composition of the present invention may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 21. In addition, or alternatively, the therapeutic composition may comprise at least one isolated 10 bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 91% sequence identity with the sequence set forth in SEQ ID NO: 29. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 92% sequence identity with the 15 sequence set forth in any one of SEQ ID NOs 6, 11, 19 or 24. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 93% sequence identity with the sequence set forth in any one of SEQ ID NOs 13, 22, 26 or 35. In addition, or alternatively, the therapeutic composition may 20 comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 5, 14, 15, 17, 18, 23, or 50. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said 25 gene comprises a sequence with at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 20, 33, 41, 43, 45, 46, 47, or 49. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 96% sequence identity with the sequence set forth in any one of SEQ 30 ID NOs 2, 7, 8, 10, 12, 30, 32, 39, 42, 44, or 48. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 3, 16, 25, 31, 34, 36, 37, or 40. In addition, or alternatively, the therapeutic composition may comprise at least 35 one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 98% sequence identity with the

sequence set forth in any one of SEQ ID NOs 4 or 9. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 27, 28, 38, or 51.

More preferably, the therapeutic composition of the present invention may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 21. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 92% sequence identity with the sequence set forth in any one of SEQ ID NOs 6, or 11. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 93% sequence identity with the sequence set forth in SEQ ID NO: 35. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 5, 19, 22, 23, or 50. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 13, 15, 29, 33, 41, 43, 45, or 46. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 7, 12, 32, 39, 42, or 44, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 3, 8, 10, 16, 34, 36, 37, 40, 48, or 49. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 98% sequence identity with the sequence set forth in any one of SEQ ID NOs 4, 9, 17 or 31. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 20, 38, or 51. In addition, or alternatively,

the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, 18, 24, 25, 26, 27, 28, 30, or 47.

Sequence identity is commonly defined with reference to the algorithm GAP (Wisconsin GCG package, Accelerys Inc, San Diego USA). GAP uses the Needleman and Wunsch algorithm to align two complete sequences that maximizes the number of matches and minimizes the number of gaps. Generally, default parameters are used, with a gap creation penalty = 12 and gap extension penalty = 4. Other algorithms suitable for the alignment of nucleotide sequences may be used instead of GAP, e.g. BLAST (Basic Local Alignment Search Tool) (which uses the method of Altschul et al. (1990) J. Mol. Biol. 215: 405-410), FASTA (which uses the method of Pearson and Lipman (1988) PNAS USA 85: 2444-2448), the Smith-Waterman algorithm (Smith and Waterman (1981) J. Mol Biol. 147: 195-197), the TBLASTN program, of Altschul et al. (1990) supra, or the psi-Blast algorithm (Nucl. Acids Res. (1997) 25 3389-3402), generally employing default parameters. In particular, BLAST may be used, preferably employing default parameters.

Sequence alignment algorithms, such as BLAST, calculate the similarity score between a query sequence and a subject sequence. The sequence identity of the query sequence to the subject sequence may be dependent on the percentage of the query sequence that is required to overlap with the subject sequence. This is also referred to as query coverage. In a preferred embodiment, the isolated bacterium present in the therapeutic composition of the present invention comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence, which (in addition to the specified sequence identity) has a query coverage of at least 98%, at least 99%, or 100%, preferably at least 98%. The query coverage refers to the percentage of said sequence which overlaps with the sequence with which it has the specified sequence identity, e.g. SEQ ID NO: 1. For example, the bacterium present in the therapeutic composition may comprise a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51 and at least 98% query coverage.

Yet more preferably, the therapeutic composition of the present invention may comprise at least one isolated bacterium, wherein the bacterium is a bacterium as deposited under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedures at the Leibniz-Institut DSMZ - Deutsche Sammlung von

Mikroorganismen und Zellkulturen GmbH (DSMZ), Inhoffenstr. 7B, 38124 Braunschweig by Genome Research Limited under an accession number as listed in Table 1 below. Specifically, the therapeutic composition of the present invention may comprise at least one isolated bacterium, wherein the bacterium is a bacterium as deposited under the Budapest Treaty at DSMZ under one of the following accession numbers (the date of deposit with DSMZ for each bacterium deposited is indicated in brackets after the accession number): DSM32191 (27 October 2015), DSM32147 (23 September 2015), DSM32149 (23 September 2015), DSM32175 (6 October 2015), DSM32153 (27 October 2015), DSM32152 (23 September 2015), DSM32158 (23 September 2015), DSM32192 (27 October 2015), DSM32148 (23 September 2015), DSM32166 (6 October 2015), DSM32151 (23 September 2015), DSM32150 (23 September 2015), DSM32193 (27 October 2015), DSM32162 (6 October 2015), DSM32194 (27 October 2015), DSM32163 (6 October 2015), DSM32205 (1 March 2016), DSM32195 (27 October 2015), DSM32164 (6 October 2015), DSM32177 (13 October 2015), DSM32167 (6 October 2015), DSM32165 (6 October 2015), DSM32169 (6 October 2015), DSM32168 (6 October 2015), DSM32178 (13 October 2015), DSM32182 (13 October 2015), DSM32179 (13 October 2015), DSM32180 (13 October 2015), DSM32184 (13 October 2015), DSM32181 (13 October 2015), DSM32183 (13 October 2015), DSM 32262 (2 February 2016), DSM32211 (2 December 2015), DSM 32219 (8 December 2015), DSM 32222 (8 December 2015), DSM 32261 (2 February 2016), DSM32212 (2 December 2015), DSM32220 (8 December 2015), DSM32213 (2 December 2015), DSM 32226 (8 December 2015), DSM32215 (2 December 2015), DSM32216 (2 December 2015), DSM 32217 (2 February 2016), DSM32221 (8 December 2015), DSM32218 (2 December 2015), DSM 32224 (8 December 2015), DSM 32214 (2 December 2015), DSM 32263 (2 February 2016), DSM 32223 (8 December 2015), DSM 32225 (8 December 2015), and DSM 32265 (10 February 2016). The putative genus and species names of the deposited bacteria, as well as their known characteristics, are listed in Table 1 below.

Yet more preferably, the therapeutic composition of the present invention comprises at least one isolated bacterium, wherein the bacterium is a bacterium as deposited under the Budapest Treaty at DSMZ under one of the following accession numbers: DSM32191 and DSM32177. In addition, or alternatively, the therapeutic composition may comprise at least one isolated bacterium, wherein the bacterium is a bacterium as deposited under the Budapest Treaty at DSMZ under one of the following accession numbers: DSM32153, DSM32152, DSM32151, DSM32193, DSM32162, DSM32194, DSM32205, DSM32195, DSM32164, DSM32177, DSM32165, DSM32169, DSM32168, DSM32182, DSM32184, DSM32211, DSM32222, DSM32215, DSM32217, DSM32218, DSM32224, DSM32214,

DSM32223, and DSM32225; more preferably a bacterium as deposited under one of the following accession numbers: DSM32153, DSM32152, DSM32151, DSM32193, DSM32194, DSM32164, DSM32165, DSM32169, DSM32184, DSM32211, DSM32222, DSM32215, DSM32217, DSM32218, DSM32224, and DSM32225.

5

Alternatively, the therapeutic composition of the present invention may comprise at least one isolated bacterium, said bacterium comprising a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least

10 99%, or 100% sequence identity with the sequence of the gene encoding the 16S rRNA in a deposited bacterium as described above.

The therapeutic composition of the invention may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at

15 least ten, at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, at least seventeen, at least eighteen, at least nineteen, at least nineteen, at least twenty, at least twenty-one, at least twenty-two, at least twenty-three, at least twenty-four, at least twenty-five, at least twenty-six, at least twenty-seven, at least twenty-eight, at least twenty-nine, at least thirty, at least thirty-one, at least thirty-two, at least thirty-three, at

20 least thirty-four, at least thirty-five, at least thirty-six, at least thirty-seven, at least thirty-eight, at least thirty-nine, at least forty, at least forty-one, at least forty-two, at least forty-three, at least forty-four, at least forty-five, at least forty-six, at least forty-seven, at least forty-eight, at least forty-nine, at least fifty, or at least fifty-one bacteria as disclosed herein.

25 The therapeutic composition of the invention may comprise one, up to two, up to three, up to four, up to five, up to six, up to seven, up to eight, up to nine, up to ten, up to eleven, up to twelve, up to thirteen, up to fourteen, up to fifteen, up to sixteen, up to seventeen, up to eighteen, up to nineteen, up to nineteen, up to twenty, up to twenty-one, up to twenty-two, up to twenty-three, up to twenty-four, up to twenty-five, up to twenty-six, up to twenty-seven,

30 up to twenty-eight, up to twenty-nine, up to thirty, up to thirty-one, up to thirty-two, up to thirty-three, up to thirty-four, up to thirty-five, up to thirty-six, up to thirty-seven, up to thirty-eight, up to thirty-nine, up to forty, up to forty-one, up to forty-two, up to forty-three, up to forty-four, up to forty-five, up to forty-six, up to forty-seven, up to forty-eight, up to forty-nine, up to fifty, or up to fifty-one bacteria as disclosed herein. Preferably, the therapeutic

35 composition of the invention comprise up to twenty, preferably up to ten, bacteria as disclosed herein.

Where a therapeutic composition comprises more than one isolated bacterium, the isolated bacteria are preferably distinct. “Distinct” may refer to the isolated bacteria encoding distinct 16S rRNA sequences.

The therapeutic composition of the invention may comprise at least one isolated bacterium which forms spores. Such a bacterium is also referred to as a spore-forming bacterium. Spores are metabolically dormant structures that are resilient to environmental insults and are used by certain bacteria as a survival strategy upon encountering adverse conditions. Bacteriotherapy candidates HMI_1, HMI_2, HMI_4, HMI_6, HMI_10, HMI_15, HMI_17, HMI_21, HMI_22, HMI_33, HMI_36, HMI_37, HMI_38, HMI_44, HMI_47, HMI_48, HMI_50, HMI_51, and HMI_52 were isolated from ethanol-treated samples and are thus expected to be capable of forming spores. In addition, HMI_3, HMI_7, HMI_8, HMI_16, HMI_18, HMI_19, HMI_24, HMI_25, HMI_26, HMI_27, HMI_28, HMI_29, HMI_30, HMI_34, HMI_41, and HMI_46 are expected to be spore formers based on phylogenetic analysis.

Thus, the therapeutic composition of the present invention may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, at least seventeen, at least eighteen, at least nineteen, at least twenty, at least twenty-one, at least twenty-two, at least twenty-three, at least twenty-four, at least twenty-five, at least twenty-six, at least twenty-seven, at least twenty-eight, at least twenty-nine, at least thirty, at least thirty-one, at least thirty-two, at least thirty-three, at least thirty-four, or at least thirty-five isolated spore-forming bacteria. In one embodiment, the bacteria in the therapeutic composition may consist of spore forming bacteria.

The spore-forming bacterium may thus be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA) , wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1, 2, 3, 4, 6, 7, 8, 10, 14, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 29, 32, 33, 35, 36, 37, 40, 43, 45, 46, 47, 49, 50, or 51. Alternatively, the spore forming bacterium may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1, 2, 4, 6, 10, 14, 16, 20, 21, 32, 35, 36, 37, 43, 46, 47, 49, 50, or 51.

The spore-forming bacterium may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 1, 2, 3, 4, 6, 7, 8, 10, 14, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 29, 32, 33, 35, 36, 37, 40, 43, 45, 46, 47, 49, 50, or 51. Alternatively, the spore forming bacterium may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 1, 2, 4, 6, 10, 14, 16, 20, 21, 32, 35, 36, 37, 43, 46, 47, 49, 50, or 51.

Preferably, the spore forming bacterium is a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NOs 1 or SEQ ID NO: 21, and/or at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 3, 4, 6, 7, 8, 10, 14, 15, 16, 17, 18, 20, 23, 24, 25, 26, 27, 28, 29, 32, 33, 35, 36, 37, 40, 43, 45, 46, 47, 49, 50, or 51. More preferably, the spore forming bacterium may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NOs 1 or SEQ ID NO: 21, and/or at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID Nos 2, 4, 6, 10, 14, 16, 20, 32, 35, 36, 37, 43, 46, 47, 49, 50, or 51.

More preferably, the spore forming bacterium may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth SEQ ID NO: 1 or SEQ ID NO: 21, at least 92% sequence identity with the sequence set forth in SEQ ID NO: 29, at least 92% sequence identity with the sequence set forth in SEQ ID NOs 6, or 24 at least 93% sequence identity with the sequence set forth in SEQ ID NOs 35 or 26, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, 15, 17, 18, 23, or 50, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 20, 33, 43, 45, 46, 47, or 49, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 7, 8, 10, or 32, at least 97% sequence identity with the sequence

set forth in any one of SEQ ID NOs 3,16, 25, 36, 37 or 40, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 4, or at least 99% sequence identity with the sequence set forth in SEQ ID NOs 27, 28, or 51. Yet more preferably, the spore forming bacterium may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth SEQ ID NO: 1 or SEQ ID NO: 21, at least 92% sequence identity with the sequence set forth in SEQ ID NO: 6, at least 93% sequence identity with the sequence set forth in SEQ ID NO: 35, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, or 50, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 20, 43, 46, 47, or 49, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 10, or 32, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 16, 36, or 37, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 4, or at least 99% sequence identity with the sequence set forth in SEQ ID NO: 51.

More preferably, the spore forming bacterium may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 21, at least 92% sequence identity with the sequence set forth SEQ ID NO: 6, at least 93% sequence identity with the sequence set forth in SEQ ID NO: 35, at least 94% sequence identity with the sequence set forth in SEQ ID NOs 23 or 50, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 15, 29, 33, 43, 45 or 46, at least 96% sequence identity with the sequence set forth in SEQ ID NOs 7 or 32, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 3, 10, 16, 36, 37, 40 or 49, at least 98% sequence identity with the sequence set forth in SEQ ID NOs 4, 8 or 17, at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 20, or 51, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, 18, 25, 26, 27, 28, or 47. Even more preferably, the spore forming bacterium may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 21, at least 92% sequence identity with the sequence set forth SEQ ID NO: 6, at least 93% sequence identity with the sequence set forth in SEQ ID NO: 35, at least 94% sequence identity with the sequence set forth in SEQ ID NO: 50, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 43, or 46, at least 96% sequence identity with the sequence set forth in SEQ ID NO: 32, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 10, 16, 36, 37, or

49, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 4, at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 20, or 51, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, or 47.

5 Yet more preferably, the spore forming bacterium may be a bacterium as deposited at DSMZ under accession number DSM32191, DSM32147, DSM32175, DSM32152, DSM32166, DSM32162, DSM32163, DSM32177, DSM32167, DSM 32262, DSM 32222, DSM 32261, DSM32212, DSM32217, DSM32224, DSM32214, DSM32223, DSM32225, DSM32265, DSM32149, DSM32158, DSM32192, DSM32194, DSM32205, DSM32195, DSM32169,
10 DSM32168, DSM32178, DSM32182, DSM32179, DSM32180, DSM32211, DSM32226, or DSM32218. Most preferably, the spore forming bacterium is a bacterium as deposited at DSMZ under accession number DSM32191, DSM32147, DSM32175, DSM32152, DSM32166, DSM32162, DSM32163, DSM32177, DSM32167, DSM 32262, DSM 32222, DSM 32261, DSM32212, DSM32217, DSM32224, DSM32214, DSM32223, DSM32225, or
15 DSM32265. Alternatively, the therapeutic composition of the present invention may comprise at least one isolated spore-forming bacterium, said bacterium comprising a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene
20 encoding the 16S rRNA in a deposited bacterium as described above.

The isolated bacterium present in the therapeutic composition may be antagonistic towards an intestinal bacterium, inhibit or prevent the growth or sporulation of an intestinal bacterium, and/or neutralize or protect against a toxin produced by an intestinal bacterium. Preferably
25 the bacterium inhibits or prevents the growth of an intestinal bacterium. The intestinal bacterium may be a pathogenic or non-pathogenic intestinal bacterium. Preferably, the intestinal bacterium is a pathogenic bacterium. This is particularly preferred in the context of a therapeutic composition for use in the treatment of a dysbiosis associated with an enteric bacterial infection. However, other diseases are also known to be characterised by an
30 increase in certain types of bacteria in the gastrointestinal tract. For example, inflammatory bowel disease is known to be characterised by an increase in bacteria from the Proteobacteria phylum, such as *Escherichia coli,* in the intestinal microbiota. Similarly, irritable bowel syndrome, obesity and malnutrition are known to be characterised by an increase in certain types of bacteria in the gastrointestinal tract. A bacterial composition
35 comprising at least one bacterium which is antagonistic towards an intestinal bacterium, inhibits or prevents the growth or sporulation of an intestinal bacterium, and/or neutralizes or

protects against a toxin produced by an intestinal bacterium thus also finds application in the treatment of dysbiosis associated with inflammatory bowel disease, irritable bowel syndrome, obesity, or malnutrition.

5 The pathogenic bacterium may be a Gram positive bacterium, or a Gram negative bacterium. Exemplary pathogenic bacteria include pathogenic bacteria of the genera *Clostridium, Escherichia, Enterococcus, Klebsiella*, *Enterobacter, Proteus*, *Salmonella*, *Shigella*, *Staphylococcus*, *Vibrio, Aeromonas, Campylobacter, Bacillus, Helicobacter, Listeria, Plesiomonas, or Yersinia*. In a preferred embodiment, the pathogenic bacterium is a
10 pathogenic bacterium of the genera *Clostridium* or *Escherichia*, such as *Clostridium difficile* or *Escherichia coli.*

Examples of pathogenic *Escherichia coli* include adherent-invasive *Escherichia coli* (AIEC), enteroaggregative *Escherichia coli,* enterohemorrhagic *Escherichia coli,* enteroinvasive
15 *Escherichia coli*, enterotoxigenic *Escherichia coli,* and *Escherichia coli 0157:H7.* An enterotoxigenic *Escherichia coli* may produce a heat-labile enterotoxin, or heat-stable enterotoxin.

For example, the pathogenic bacterium may be *Clostridium difficile* or adherent-invasive *E.*
20 *coli* (AIEC).

Bacteriotherapy candidates HMI_14, HMI_25, HMI_42, HMI_26, HMI_28, HMI_35 and HMI_46 have been shown to inhibit growth of *Clostridium difficile* in an overlay assay. In addition, HMI_2, HMI_4, HMI_5, HMI_6, HMI_15, HMI_26, HMI_27, HMI_28, HMI_34,
25 HMI_35, HMI_39, HMI_40, HMI_43, HMI_44, HMI_46 and HMI_47 have been shown to inhibit growth of *Clostridium difficile* in a CFS-relative growth inhibition assay (see Example 2, Figure 13, and Table 1).

Bacteriotherapy candidates HMI_4, HMI_10, HMI_11, HMI_14, HMI_26, HMI_28, HMI_33,
30 HMI_35, HMI_42 and HMI_46 have been shown to inhibit growth of *Escherichia coli* in an overlay assay. In addition, HMI_46 and HMI_28, have been shown to inhibit growth of *Escherichia coli* in a CFS-relative growth inhibition assay (see Example 2, Figure 13, and Table 1).

35 It is expected that a bacterium which inhibits the growth of *Escherichia coli* also inhibits the growth of other Proteobacteria. Thus, the pathogenic bacterium may be a proteobacterium.

Proteobacteria include (apart from *Escherichia* species), *Salmonella* species, Campylobacter species, *Vibrio* species, *Helicobacter* species, and *Yersinia* species.

It is expected that a bacterium which inhibits the growth of *Clostridium difficile* also inhibits
5 the growth of other bacteria of the genus *Clostridium*. Thus, the pathogenic bacterium may be a bacterium of the genus *Clostridium*. Pathogenic bacteria of the genus *Clostridium* (apart from *Clostridium difficile*), include *Clostridium perfringens*, *Clostridium botulinum*, and *Clostridium tetani*.

10 The therapeutic composition may thus comprise at least one isolated bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli*. For example, the therapeutic composition may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, at
15 least seventeen, at least eighteen, at least nineteen, at least twenty, at least twenty-one, or at least twenty-two isolated bacteria which inhibit the growth of *Clostridium difficile* and/or *Escherichia coli*. In one embodiment, the bacteria in the therapeutic composition may consist of one or more isolated bacteria which have been shown to inhibit the growth of *Clostridium difficile* and/or *Escherichia coli.*

20

In a preferred embodiment, the therapeutic composition may comprise at least one isolated bacterium which has been shown to inhibit the growth of *Clostridium difficile*. This is preferred in the context of a therapeutic composition for use in the treatment of a dysbiosis associated with an enteric infection, in particular a dysbiosis associated with an infection with
25 a pathogenic *Clostridium*-related species, *such as Clostridium difficile*, *Clostridium perfringens*, *Clostridium botulinum*, or *Clostridium tetani, most preferably Clostridium difficile*.

For example, the therapeutic composition may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at
30 least ten, at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, at least seventeen, at least eighteen, or at least nineteen isolated bacteria which inhibit the growth of *Clostridium difficile*. In one embodiment, the bacteria in the therapeutic composition may consist of bacteria which inhibit the growth of *Clostridium difficile.*

35

A bacterium inhibits the growth of *Clostridium difficile* may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 4, 5, 6, 13, 14, 24, 25, 26, 27, 33, 34, 38, 39, 41, 42, 43, 45 and 46.

Alternatively, a bacterium which inhibits the growth of *Clostridium difficile* may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 91%, at least 92%, at least 93%,at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 4, 5, 6, 13, 14, 24, 25, 26, 27, 33, 34, 38, 39, 41, 42, 43, 45 and 46.

More preferably, a bacterium which inhibits the growth of *Clostridium difficile* may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 92% sequence identity with the sequence set forth in any one of SEQ ID NOs 6, or 24, at least 93% sequence identity with the sequence set forth in any one of SEQ ID NOs 13, or 26, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 5, or 14, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 33, 41, 43, 45, or 46, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 39, or 42, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 25, or 34, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 4, or at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 27, or 38.

Even more preferably, a bacterium which inhibits the growth of *Clostridium difficile* may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 92% sequence identity with the sequence set forth SEQ ID NO: 6, at least 94% sequence identity with the sequence set forth in SEQ ID NO: 5, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 13, 33, 41, 43, 45, or 46, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 39, or 42, at least 97% sequence identity with the sequence set forth in SEQ ID NO: 34, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 4, at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, or 38, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, 24, 25, 26, or 27.

Most preferably, the bacterium which inhibits the growth of *Clostridium difficile* may be a bacterium as deposited at DSMZ under accession number DSM32147, DSM32175, DSM32153, DSM32152, DSM32193, DSM32162, DSM32168, DSM32178, DSM32182, DSM32179, DSM32211, DSM 32219, DSM32220, DSM32213, DSM32215, DSM32216, DSM 32217, DSM32218, DSM 32224. Alternatively, the therapeutic composition of the present invention may comprise at least one isolated bacterium which inhibits the growth of *Clostridium difficile*, wherein said bacterium comprising a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene encoding the 16S rRNA in a deposited bacterium as described above.

In an alternative preferred embodiment, the therapeutic composition may comprise at least one isolated bacterium which inhibits the growth of *Escherichia coli*. This is preferred in the context of a therapeutic composition for use in the treatment of a dysbiosis associated with an enteric infection, in particular a dysbiosis associated with an infection with a Proteobacterium, such as *Escherichia* species, *Salmonella* species, Campylobacter species, *Vibrio* species, *Helicobacter* species, and *Yersinia* species, most preferably a dysbiosis associated with an infection with *Escherichia coli*.

For example, the therapeutic composition may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or at least ten, isolated bacteria which inhibit the growth of *Escherichia coli*. In one embodiment, the bacteria in the therapeutic composition may consist of bacteria which inhibit the growth of *Escherichia coli.*

A bacterium which inhibits the growth of *Escherichia coli* may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 4, 10, 11, 13, 25, 27, 32, 34, 41, and 45.

Alternatively, a bacterium which inhibits the growth of *Escherichia coli* may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 91%, at least 92%, at least 93%,at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 4, 10, 11, 13, 25, 27, 32, 34, 41, and 45.

More preferably, a bacterium which inhibits the growth of *Escherichia coli* may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 92% sequence identity with the sequence set forth in SEQ ID NO: 11, at least 93% sequence identity with the sequence set forth in SEQ ID NO: 13, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 41, or 45, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 10, or 32, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 25, or 34, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 4, or at least 99% sequence identity with the sequence set forth in SEQ ID NO: 27.

Even more preferably, a bacterium which inhibits the growth of *Escherichia coli* may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 92% sequence identity with the sequence set forth in SEQ ID NO: 11, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 13, 41, or 45, at least 96% sequence identity with the sequence set forth in SEQ ID NO: 32, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 10, or 34, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 4, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 25, or 27.

Most preferably, the bacterium which inhibits the growth of *Escherichia coli* may be a bacterium as deposited at DSMZ under accession number DSM32175, DSM32166, DSM32151, DSM32193, DSM32178, DSM32179, DSM 32262, DSM 32219, DSM32215, DSM32218. Alternatively, the therapeutic composition of the present invention may comprise at least one isolated bacterium which inhibits the growth of *Escherichia coli*, wherein said bacterium comprising a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene encoding the 16S rRNA in a deposited bacterium as described above.

In addition to an isolated bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli*, the therapeutic composition may comprise at least one isolated bacterium which co-occurs with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein. Bacteriotherapy candidates which have been shown to

co-occur with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein are HMI_2, HMI_5, HMI_6, HMI_7, HMI_8, HMI_9, HMI_10, HMI_11, HMI_12, HMI_14, HMI_15, HMI_16, HMI_17, HMI_18, HMI_19, HMI_20, HMI_26, HMI_27, HMI_31, HMI_33, HMI_34, HMI_35, HMI_37, HMI_38, HMI_39, HMI_41, HMI_42, HMI_43, HMI_44, HMI_46, HMI_47, HMI_48, HMI_50, HMI_51, and HMI_52 (see Table 1 for details).

Thus, the therapeutic composition may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, at least seventeen, at least eighteen, at least nineteen, at least nineteen, at least twenty, at least twenty-one, at least twenty-two, at least twenty-three, at least twenty-four, at least twenty-five, at least twenty-six, at least twenty-seven, at least twenty-eight, at least twenty-nine, at least thirty, at least thirty-one, at least thirty-two, at least thirty-three, at least thirty-four, or thirty-five isolated bacteria which co-occur with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein.

The bacterium which co-occurs with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein, may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 25, 26, 30, 32, 33, 34, 36, 37, 38, 40, 41, 42, 43, 45, 46, 47, 49, 50, or 51.

Alternatively, the bacterium which co-occurs with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein, may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 25, 26, 30, 32, 33, 34, 36, 37, 38, 40, 41, 42, 43, 45, 46, 47, 49, 50, or 51.

Preferably, the bacterium which co-occurs with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein, is a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with

at least 92% sequence identity with the sequence set forth in any one of SEQ ID NOs 6, 11, or 19, at least 93% sequence identity with the sequence set forth in any one of SEQ ID NOs 13 or 26, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 5, 14, 15, 17, 18, or 50, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 33, 41, 43, 45, 46, 47, or 49, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 7, 8, 10, 12, 30, 32, or 42, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 16, 25, 34, 36, 37, or 40, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 9, or at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 38, or 51.

More preferably, the bacterium which co-occurs with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein, is a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 92% sequence identity with the sequence set forth in any one of SEQ ID NOs 6, or 11, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 5, 19, or 50, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 13, 15, 33, 41, 43, 45, or 46, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 7, 12, 32, or 42, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 8, 10, 16, 34, 36, 37, 40, or 49, at least 98% sequence identity with the sequence set forth in any one of SEQ ID NOs 9 or 17, at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 38, 51, or at least 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, 18, 25, 26, 30, or 47.

Most preferably, the bacterium which co-occurs with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein, is a bacterium as deposited at DSMZ under accession number DSM32147, DSM32153, DSM32152, DSM32158, DSM32192, DSM32148, DSM32166, DSM32151, DSM32150, DSM32193, DSM32162, DSM32194, DSM32163, DSM32205, DSM32195, DSM32164, DSM32178, DSM32182, DSM32181, DSM32262, DSM32211, DSM32219, DSM32261, DSM32212, DSM32220, DSM32226, DSM32215, DSM32216, DSM32217, DSM32218, DSM32224, DSM32214, DSM32223, DSM32225, or DSM32265. Alternatively, the therapeutic composition of the present invention may comprise at least one isolated bacterium which co-occurs with a bacterium which inhibits the growth of *Clostridium difficile* and/or *Escherichia coli* as disclosed herein, wherein said bacterium comprises a gene encoding a 16S rRNA, wherein

said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene encoding the 16S rRNA in a deposited bacterium as described above.

5

Bacteria related to the genus *Clostridium* have been shown to be beneficial in reducing inflammation through interactions with the host immune system (Atarashi, Tanoue *et al.* 2013). The isolated bacterium present in the therapeutic composition may thus be a bacterium which has immunomodulatory activity. For example, the bacterium may reduce
10 inflammation in the individual, e.g. in the gastrointestinal tract of the individual. Bacteriotherapy candidates which are in the same genus as bacteria which have been shown to be beneficial in reducing inflammation through interactions with the host immune system are HMI_4, HMI_9, HMI_10, HMI_15, HMI_27, HMI_28 and HMI_38. The bacteria ar therefore expected to have immunomodulatory activity, such as reducing inflammation in the
15 individual, e.g. in the gastrointestinal tract of the individual.

Thus, the therapeutic composition may comprise at least one, at least two, at least three, at least four, at least five, at least six, or at least seven isolated bacteria which have immunomodulatory activity. In one embodiment, the bacteria in the therapeutic composition
20 may consist of bacteria which reduce inflammation in the individual.

The bacterium which has immunomodulatory activity may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 4, 9, 10,
25 14, 26, 27, or 37.

Alternatively, a bacterium which has immunomodulatory activity may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 91%, at least 92%, at least 93%,at least 94%, at least 95%, at least
30 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 4, 9, 10, 14, 26, 27, or 37.

Preferably, a bacterium which has immunomodulatory activity may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a
35 sequence with at least 93% sequence identity with the sequence set forth in SEQ ID NO: 26, at least 94% sequence identity with the sequence set forth in SEQ ID NO: 14, at least 96%

sequence identity with the sequence set forth in SEQ ID NO:10, at least 97% sequence identity with the sequence set forth in SEQ ID NO: 37, at least 98% sequence identity with the sequence set forth in any one of SEQ ID NOs 4 or 9, or at least 99% sequence identity with the sequence set forth in SEQ ID NO: 27.More preferably, a bacterium which has
5 immunomodulatory activity may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 97% sequence identity with the sequence set forth in in any of SEQ ID NOs 10 or 37, at least 98% sequence identity with the sequence set forth in any one of SEQ ID NOs 4 or 9, or at least 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, 26,
10 or 27.

Most preferably, the bacterium which has immunomodulatory activity may be a bacterium as deposited at DSMZ under accession number DSM32175, DSM32148, DSM32166, DSM32162, DSM32182, DSM32179, or DSM32212. Alternatively, the therapeutic
15 composition of the present invention may comprise at least one isolated bacterium which has immunomodulatory activity, wherein said bacterium comprising a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene
20 encoding the 16S rRNA in a deposited bacterium as described above.

The therapeutic composition of the invention may comprise at least one isolated bacterium which is on the Human Microbiome Project's (HMP) "most wanted" list. Bacteriotherapy candidates HMI_1, HMI_2, HMI_4, HMI_5, HMI_7, HMI_11, HMI_12, HMI_15, HMI_16,
25 HMI_17, HMI_18, HMI_19, HMI_35, HMI_37, HMI_38, HMI_39, HMI_45, HMI_50, and HMI_51 are on HMP's "most wanted" list.

Thus, the therapeutic composition of the present invention may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight,
30 at least nine, at least ten, at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, at least seventeen, at least eighteen, or nineteen bacteria which are on HMP's "most wanted" list. In one embodiment, the bacteria in the therapeutic composition may consist of bacteria which are on HMP's "most wanted" list.

The bacterium which is on HMP's "most wanted" list may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1, 2, 4, 5, 7, 11, 12, 14, 15, 16, 17, 18, 34, 36, 37, 38, 44, 49, or 50.

Alternatively, the bacterium which is on HMP's "most wanted" list may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 1, 2, 4, 5, 7, 11, 12, 14, 15, 16, 17, 18, 34, 36, 37, 38, 44, 49, or 50.

Preferably, the bacterium which is on HMP's "most wanted" list is a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth SEQ ID NO: 1, at least 92% sequence identity with the sequence set forth SEQ ID NO: 11, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 5, 14, 15, 17, 18, or 50, at least 95% sequence identity with the sequence set forth SEQ ID NO: 49, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 7, 12, or 44, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 16, 34, 36, or 37, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 4, or at least 99% sequence identity with the sequence set forth in SEQ ID NO: 38.

More preferably, the bacterium which is on HMP's "most wanted" list is a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth SEQ ID NO: 1, at least 92% sequence identity with the sequence set forth SEQ ID NO: 11, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 5 or 50, at least 95% sequence identity with the sequence set forth SEQ ID NO: 15, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 7, 12, or 44, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 16, 34, 36, 37, 49, at least 98% sequence identity with the sequence set forth in any one of SEQ ID NOs 4, or 17, at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, or 38, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, or 18.

Most preferably, the bacterium which is on HMP's "most wanted" list is a bacterium as deposited at DSMZ under accession number DSM32191, DSM32147, DSM32175, DSM32153, DSM32158, DSM32151, DSM32150, DSM32162, DSM32194, DSM32163, DSM32205, DSM32195, DSM32219, DSM32261, DSM32212, DSM32220, DSM32221, DSM32223, or DSM32225. Alternatively, the therapeutic composition of the present invention may comprise at least one bacterium which is on HMP's "most wanted" list, wherein said bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene encoding the 16S rRNA in a deposited bacterium as described above.

The therapeutic composition of the invention may comprise at least one isolated bacterium which is a keystone species. Bacteriotherapy candidates HMI_17, HMI_23, HMI_24, HMI_25, HMI_26, HMI_27, HMI_28, HMI_29, HMI_30, HMI_31, HMI_32, HMI_45, HMI_49, HMI_51, and HMI_52 are expected to be keystone species.

Thus, the therapeutic composition of the present invention may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, at least eleven, at least twelve, at least thirteen, at least fourteen, or fifteen bacteria which are keystone species. In one embodiment, the bacteria in the therapeutic composition may consist of bacteria which are keystone species.

The bacterium which is a keystone species may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA) , wherein said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 16, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 44, 48, 50, or 51.

Alternatively, the bacterium which is a keystone species may be a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA) wherein said gene comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 16, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 44, 48, 50, or 51.

Preferably, the bacterium which is a keystone species is a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA) wherein said gene comprises a sequence with at least 91% sequence identity with the sequence set forth in SEQ ID NO: 29, at least 92% sequence identity with the sequence set forth in SEQ ID NO: 24, at least 93% sequence identity with the sequence set forth in any one of SEQ ID NOs 22, or 26, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 23, or 50, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 30, 44, or 48, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 16, 25, or 31, or at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 27, 28, or 51.

More preferably, the bacterium which is a keystone species is a bacterium comprising a gene encoding a 16S ribosomal RNA (rRNA) wherein said gene comprises a sequence with at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 22, 23, or 50, at least 95% sequence identity with the sequence set forth in SEQ ID NO: 29, at least 96% sequence identity with the sequence set forth in SEQ ID NO: 44, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 16, or 48, at least 98% sequence identity with the sequence set forth in SEQ ID NO: 31, at least 99% sequence identity with the sequence set forth in SEQ ID NO: 51, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 24, 25, 26, 27, 28, or 30.

Most preferably, the bacterium which is a keystone species is a bacterium as deposited at DSMZ under accession number DSM32163, DSM32165, DSM32169, DSM32168, DSM32178, DSM32182, DSM32179, DSM32180, DSM32184, DSM32181, DSM32183, DSM32221, DSM32263, DSM32225, or DSM32265. Alternatively, the therapeutic composition of the present invention may comprise at least one bacterium which is a keystone species, wherein said bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene encoding the 16S rRNA in a deposited bacterium as described above.

In addition, or alternatively, the therapeutic composition of the invention may comprise at least one isolated bacterium which has been shown to be present post-FMT. The bacteriotherapy candidates to which this applies are set out in Table 1. For example, the therapeutic composition may comprise at least one, at least two, at least three, at least four,

at least five, at least six, at least seven, at least eight, at least nine, at least ten, at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, at least seventeen, at least eighteen, at least nineteen, at least twenty, at least twenty-one or twenty two bacteria which has been shown to be present post-FMT. For example the bacterium which has been shown to be present post-FMT may comprise a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 36, 38, 40, 43, 44, 45, 46, 48, or 51. Preferably, the bacterium which has been shown to be present post-FMT is a bacterium as deposited at DSMZ under accession number DSM32165, DSM32169, DSM32168, DSM32178, DSM32182, DSM32179, DSM32180, DSM32184, DSM32181, DSM32183, DSM32262, DSM32211, DSM32219, DSM32261, DSM32220, DSM32226, DSM32217, DSM32221, DSM32218, DSM32224, DSM32263, or DSM32265. Alternatively, the therapeutic composition of the present invention may comprise at least one bacterium which has been shown to be present post-FMT, wherein said bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene encoding the 16S rRNA in a deposited bacterium as described above.

In addition, or alternatively, the therapeutic composition of the invention may comprise at least one isolated bacterium which is expected to produce one or more beneficial metabolites, such as short chain fatty acids (SCFA). The bacteriotherapy candidates to which this applies are set out in Table 1. For example, the therapeutic composition may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, at least eleven, at least twelve, at least thirteen, or fourteen bacteria which produce one or more beneficial metabolites. For example the bacterium which produces one or more beneficial metabolites may comprise a gene encoding a 16S ribosomal RNA (rRNA), wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 9, 12, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, or 31. Preferably, the bacterium which is expected to produce one or more beneficial metabolites is a bacterium as deposited at DSMZ under accession number DSM32148, DSM32150, DSM32164, DSM32177, DSM32165, DSM32169, DSM32168, DSM32178,

DSM32182, DSM32179, DSM32180, DSM32184, DSM32181, or DSM32183. Alternatively, the therapeutic composition of the present invention may comprise at least one bacterium which produces one or more beneficial metabolites, wherein said bacterium comprises a gene encoding a 16S rRNA, wherein said gene comprises a sequence with at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 98.7%, at least 99%, or 100% sequence identity with the sequence of the gene encoding the 16S rRNA in a deposited bacterium as described above.

The isolated bacterium or isolated bacteria present in a therapeutic composition may make up at least 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 7%, 75%, 80%, 85%, or 90% of the therapeutic composition by volume or weight.

The therapeutic composition may comprise no other active ingredient other than the isolated bacterium or isolated bacteria in question, and optionally a prebiotic. Thus, the active ingredient of the therapeutic composition may consist of one or more isolated bacteria disclosed herein, and optionally a prebiotic. This may also be referred to as a defined active ingredient.

The therapeutic composition of the present invention is not a faecal microbiota transplant (FMT). FMTs usually consist of a stool sample from a healthy human donor which is administered directly to the recipient, e.g. in the form of an enema, without bacteria present in the stool sample being isolated prior to the administration of the FMT to the recipient. An advantage of the therapeutic composition of the invention is that it may comprise no undefined components, which are present in FMTs, thereby allowing the therapeutic composition to be standardised and increasing safety.

The therapeutic composition of the present invention may be prepared by a method comprising culturing the one or more isolated bacteria present in the therapeutic composition in a suitable medium or media. Media and conditions suitable for culturing the bacteria to be included in the therapeutic composition of the present invention are described in detail elsewhere herein. For example, a method of preparing a therapeutic composition according to the present invention may comprise the steps of:

(i) culturing a first isolated bacterium;

(ii) optionally culturing a second isolated bacterium; and

(iii) mixing the bacteria obtained in (i) and optionally (ii) to prepare the therapeutic composition. The isolated bacteria to be included in the therapeutic composition are

preferably cultured in separate steps. In other words, a separate culture of each bacterium to be included in the therapeutic composition is preferably prepared. This allows the growth of each bacterium to be evaluated and the amount of each bacterium to be included in the pharmaceutical composition to be controlled as desired. The bacteria cultured in steps (i) and (ii) preferably have distinct 16S rRNA sequences.

The above method may include steps of culturing each isolated bacterium which is to be included in the therapeutic composition. Thus, the method may e.g. further include steps of culturing a third, fourth, fifth, sixth, seventh, eighth, ninth, and/or tenth distinct isolated bacterium, as required. In this way, the method comprise steps of culturing up to 51 distinct isolated bacteria. The bacterium or bacteria cultured by said method may be any bacterium as disclosed herein.

The method may optionally comprise one or more further steps in which the bacteria are mixed with one or more additional ingredients, such as a pharmaceutically acceptable excipient, prebiotic, carrier, insoluble fibre, buffer, osmotic agent, antifoaming agent, and/or preservative. In addition, or alternatively, the method may comprise suspending the bacteria obtained in (i) and optionally (ii) in a chemostat medium, or saline, e.g. 0.9% saline. The bacteria obtained in (i) and optionally (ii) may be provided under a reduced atmosphere, such as N_2, CO_2, H_2, or a mixture thereof, e.g. N_2:CO_2:H_2. The gases may be present in appropriate ratios for the preservation of the bacteria present in the therapeutic composition. For Example, the reduced atmosphere may comprise 80% N_2, 10% CO_2 and 10% H_2. In addition, or alternatively, the method may comprise a step of lyophilising the bacteria obtained in (i) and optionally (ii), optionally in the presence of a stabiliser and/or cryprotectant. The method may also comprise a step of preparing a capsule, tablet, or enema comprising the bacteria obtained in (i) and optionally (ii). The capsule or tablet may be enteric-coated, pH dependant, slow-release, and/or gastro-resistant.

The present invention also encompasses a therapeutic composition obtainable by, or obtained by, a method as disclosed herein. Such a therapeutic composition may further be used for a therapeutic purpose, in a therapeutic method, or for the manufacture of a medicament, as described herein, such as treatment of a dysbiosis, in particular a dysbiosis of the gastrointestinal tract.

It is expected that the bacteria disclosed herein will be suitable for the treatment of a dysbiosis, in particular a dysbiosis of the gastrointestinal tract. Without wishing to be limited

by theory, it is expected that administration of one or more of the bacteria disclosed herein to an in individual will resolve a gastrointestinal dysbiosis, where present, and/or prevent the occurrence of gastrointestinal dysbiosis, in the individual. “Individual”, as used herein, refers to a human individual or human patient.

5

Treatment of a dysbiosis may refer to the cure, prevention, or amelioration of a dysbiosis or the amelioration of at least one symptom associated with dysbiosis. Where the dysbiosis is associated with a disease, such as inflammatory bowel disease, treatment of the dysbiosis may refer to the cure, prevention, or amelioration of said disease, or the amelioration of at
10 least one symptom associated with said disease.

The therapeutic compositions of the present invention thus find application in the treatment of dysbiosis, in particular dysbiosis of the gastrointestinal tract. Accordingly, the invention provides a method of treating a dysbiosis comprising administering a therapeutically
15 effective amount of a therapeutic composition of the invention to an individual in need thereof, a therapeutic composition according to the invention for use in a method of treating a dysbiosis in an individual, and the use of a therapeutic composition of the invention for the manufacture of a medicament for the treatment of a dysbiosis in an individual.

20 “Dysbiosis” in the context of the present invention refers to a state in which the normal diversity and/or function of the microbiota or microbiome, in particular the human gastrointestinal microbiota, is disrupted. Any disruption from the normal state of the microbiota in a healthy individual can be considered a dysbiosis, even if the dysbiosis does not result in a detectable decrease in health in the individual. In a preferred embodiment, the
25 dysbiosis may be associated with one or more pathological symptoms. For example, “dysbiosis” may refer to a decrease in the microbial diversity of the microbiota. In addition, or alternatively, “dysbiosis” may refer to an increase in the abundance of one or more bacteria, e.g. one or more pathogenic bacteria, in the microbiota of an individual relative to the abundance of said bacterium or bacteria in the microbiota of a healthy individual, i.e. an
30 individual without a dysbiosis. The pathogenic bacteria present during dysbiosis are often Proteobacteria and resistant to one or more antibiotics. Examples of Proteobacteria include *Escherichia, Salmonella, Campylobacter, Vibrio, Helicobacter*, and *Yersinia* species.

The dysbiosis may be a dysbiosis associated with an enteric bacterial infection, such as an
35 infection of the gastrointestinal tract with a pathogenic bacterium. Many bacteria capable of causing infections of the gastrointestinal tract in humans are known and include: gram

positive bacteria, and gram negative bacteria. The pathogenic bacterium is preferably a pathogenic species of the genus *Clostridium, Escherichia, Enterococcus, Klebsiella*, *Enterobacter, Proteus*, *Salmonella*, *Shigella*, *Staphylococcus*, *Vibrio, Aeromonas, Campylobacter, Plesiomonas, Bacillus, Helicobacter, Listeria, or Yersinia*. Preferred examples of such pathogenic bacteria include *Clostridium difficile, Clostridium perfringens, Clostridium botulinum, Escherichia coli, Salmonella typhi, Staphylococcus aureus, Vibrio cholerae, Vibrio parahaemolyticus, Vibrio vulnificus, Campylobacter fetus, Campylobacter jejuni, Aeromonas hydrophila, Plesiomonas shigelloides, Bacillus cereus, Helicobacter pylori, Listeria monocytogenes,* and *Yersinia enterocolitica.* More preferably, the pathogenic bacterium is a pathogenic species of the genus *Clostridium or Escherichia.* Most preferably, the pathogenic bacterium is *Clostridium difficile* or *Escherichia coli.*

The pathogenic bacterium may resistant to one or more antibiotics. For example, the pathogenic bacterium, e.g. *Clostridium difficile*, may be resistant to fluoroquinolones. In addition, or alternatively, the pathogenic bacterium may be resistant to one or more carbapenems. Carbapenems are antibiotics used for the treatment of infections by multidrug-resistant (MDR) bacteria, and examples include imipenem, meropenem, ertapenem, doripenem, panipenem, and biapenem.

Treatment of a dysbiosis associated with an infection with a pathogenic bacterium may comprise reducing the abundance of the pathogenic bacterium, e.g. in the gastrointestinal tract of the individual, relative to the abundance of the pathogenic bacterium prior to treatment.

The dysbiosis may be a recurrent or chronic dysbiosis. For example, *Clostridium difficile* is known to result in recurrent infections in some individuals, with the infection reoccurring once antibiotic treatment is stopped. This may be referred to as a recurrent or chronic dysbiosis.

Dysbiosis of the gastrointestinal tract is known to be associated with, and is thought to play a causal role in, a number of diverse diseases, including inflammatory bowel disease, irritable bowel syndrome, metabolic disease, a neuropsychiatric disorder, an autoimmune disease, an allergic disorder, or a cancer. Thus the dysbiosis may be a dysbiosis associated with inflammatory bowel disease, irritable bowel syndrome, a metabolic disease, a neuropsychiatric disorder, an autoimmune disease, an allergic disorder, a cancer, or hepatic encephalopathy. Examples of inflammatory bowel disease include ulcerative colitis and Crohn’s disease.

Metabolic disease in which dysbiosis of the gastrointestinal tract has been shown to play a role include metabolic syndrome, obesity, type 2 diabetes mellitus, a cardiovascular disease, and non-alcoholic fatty liver.

Neuropsychiatric disorder in which dysbiosis of the gastrointestinal tract has been shown to play a role include Parkinson's disease, Alzheimer's disease, multiple sclerosis, myoclonus dystonia, autism and chronic fatigue syndrome.

Autoimmune diseases in which dysbiosis of the gastrointestinal tract has been shown to play a role include idiopathic thrombocytopenic purpura, arthritis, Sjögren's syndrome, systemic lupus erythematosus, and Hashimoto's thyroiditis.

Allergic disorder in which dysbiosis of the gastrointestinal tract has been shown to play a role include atopy, and asthma.

Cancers in which dysbiosis of the gastrointestinal tract has been shown to play a role include colorectal cancer, extra-intestinal tumours, mammary tumours, hepatocellular carcinoma, lymphoma, melanoma, and lung cancer.

The therapeutic composition of the invention may comprise a pharmaceutically acceptable excipient, carrier, buffer, stabilizer or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the isolated bacteria present in the therapeutic composition. The precise nature of the pharmaceutically acceptable excipient or other material will depend on the route of administration, which may be oral or rectal. Many methods for the preparation of therapeutic compositions are known to those skilled in the art. See e.g. Robinson ed., Sustained and Controlled Release Drug Delivery Systems, Marcel Dekker, Inc., New York, 1978.

The therapeutic composition of the invention may comprise a prebiotic, a carrier, insoluble fibre, a buffer, an osmotic agent, an anti-foaming agent and/or a preservative.

Prebiotics may provide nutrients for the isolated bacteria present in the therapeutic composition to assist their early growth and colonisation after administration to the individual. Any prebiotic known in the art may be used. Non-limiting examples of prebiotics include oligosaccharides, e.g., fructooligosaccharides such as oligofructose and inulin, mannan

oligosaccharides and galactooligosaccharides, soluble, oligofructose-enriched inulin and soluble fiber. Insoluble fiber may be included in the therapeutic composition as a carrier, e.g., to provide protection during transit or storage. A buffer may be included in the therapeutic composition to promote the viability of the isolated bacteria present. An anti-
5 fungal agent may included in the therapeutic composition as a preservative.

The therapeutic composition may be made or provided in chemostat medium. Alternatively, the therapeutic composition may be made or provided in saline, e.g., 0.9% saline. It will be understood that any carrier or solution which does not impair viability of the bacteria present
10 in the therapeutic composition and is compatible with administration to an individual may be used.

The therapeutic composition may be made or provided under reduced atmosphere, i.e., in the absence of oxygen. The synthetic stool preparation may be made or provided under N_2,
15 CO_2, H_2, or a mixture thereof, optionally with controlled levels of partial pressure of N_2:CO_2:H_2.

The therapeutic composition may be for oral or rectal administration to the individual. Where the therapeutic composition is for oral administration, the therapeutic composition may be in
20 the form of a capsule, or a tablet. Where the therapeutic composition is for rectal administration, the therapeutic composition may be in the form of an enema. The preparation of suitable capsules, tablets and enema is well-known in the art. The capsule or tablet may comprise a coating to protect the capsule or tablet from stomach acid. For example, the capsule or tablet may be enteric-coated, pH dependant, slow-release, and/or gastro-
25 resistant. Such capsules and tablets are used, for example, to minimize dissolution of the capsule or tablet in the stomach but allow dissolution in the small intestine.

The therapeutic composition may be lyophilized. The lyophilized therapeutic composition may comprise one or more stabilisers and/or cryoprotectants. The lyophilized therapeutic
30 composition may be reconstituted using a suitable diluent prior to administration to the individual.

A therapeutic composition according to the present invention may be administered alone or in combination with other treatments, concurrently or sequentially or as a combined
35 preparation with another therapeutic agent or agents, for the treatment of dysbiosis, or a disease associated with dysbiosis as described herein. For example, a conjugate of the

invention may be used in combination with an existing therapeutic agent for inflammatory bowel disease, irritable bowel syndrome, a metabolic disease, a neuropsychiatric disorder, an autoimmune disease, an allergic disorder, a cancer, or hepatic encephalopathy.

5 For example, where the therapeutic composition is for the treatment of a dysbiosis associated with cancer, the therapeutic composition may optionally be administered in combination a cancer immunotherapy, such as an immune check-point inhibitor, to the individual. Examples of check-point inhibitors which may be employed in this context include Programmed cell death protein 1 (PD-1) inhibitors, Programmed death-ligand 1 (PD-L1)
10 inhibitors, cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) inhibitors. Manipulation of the gut microbiota in combination with immune check-point inhibitor treatment has been shown to improve efficacy of immune check-point inhibitors in treating cancer (Snyder et al. 2015). In a preferred embodiment, the cancer in this context is lung cancer or melanoma. Immune check-point inhibitors have been approved for the treatment of these cancers and
15 bacteriotherapy has been shown to improve efficacy of check-point inhibitors in the treatment of melanoma (Snyder et al. 2015).

The therapeutic compositions of the invention may be administered to an individual, preferably a human individual. Administration may be in a "therapeutically effective amount",
20 this being sufficient to show benefit to the individual. Such benefit may be at least amelioration of at least one symptom. Thus “treatment” of a specified disease refers to amelioration of at least one symptom. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated, the particular patient being treated, the clinical condition of the individual patient, the cause of
25 the dysbiosis, the site of delivery of the composition, the type of therapeutic composition, the method of administration, the scheduling of administration and other factors known to medical practitioners. Prescription of treatment, e.g. decisions on dosage etc., is within the responsibility of general practitioners and other medical doctors, and may depend on the severity of the symptoms and/or progression of a disease being treated. A therapeutically
30 effective amount or suitable dose of a therapeutic composition of the invention can be determined by comparing its *in vitro* activity and *in vivo* activity in an animal model. Methods for extrapolation of effective dosages in mice and other test animals to humans are known. The precise dose will depend upon a number of factors, including whether the therapeutic composition is for prevention or for treatment.

35

Further aspects and embodiments of the invention will be apparent to those skilled in the art given the present disclosure including the following experimental exemplification.

All documents mentioned in this specification are incorporated herein by reference in their
5 entirety.

Unless the context dictates otherwise, the singular includes the plural.

"and/or" where used herein is to be taken as specific disclosure of each of the two specified
10 features or components with or without the other. For example "A and/or B" is to be taken as specific disclosure of each of (i) A, (ii) B and (iii) A and B, just as if each is set out individually herein.

Unless context dictates otherwise, the descriptions and definitions of the features set out
15 above are not limited to any particular aspect or embodiment of the invention and apply equally to all aspects and embodiments which are described.

Certain aspects and embodiments of the invention will now be illustrated by way of example and with reference to the figures described above.
20

Examples

Example 1 – Identification and isolation of bacteriotherapy candidates

25 *Materials and Methods*

Two different approaches were used to isolate bacterial species for inclusion in a therapeutic composition for treating dysbiosis. The first relied on a broad culturing approach from healthy adult donors to establish a culture collection which is as representative as possible of the
30 bacterial component of the intestinal microbiota of healthy individuals. This process also incorporated a targeted culturing approach to preferentially select bacteria displaying a particular phenotype or function e.g. spore formation. The second approach was more targeted in nature and aimed to isolate bacterial species specifically associated with resolving gastrointestinal dysbiosis by comparing the microbiota of individuals before and
35 after Faecal Microbiota Transplantation (FMT) administered to resolve dysbiosis associated

with *C. difficile* infection. These two approaches are respectively referred to as Candidate Isolation Process 1 (CIP1) and Candidate Isolation Process 2 (CIP2) below.

Sample collection and culturing

5

For CIP1, fresh faecal samples were obtained from six consenting healthy adult human donors (one faecal sample per donor - minimum 0.5g). The samples were placed in anaerobic conditions within one hour of passing to preserve the viability of anaerobic bacteria. All sample processing and culturing took place under anaerobic conditions in a
10 Whitley DG250 workstation (Don Whitley, West Yorkshire, UK) at 37°C. Culture media, phosphate-buffered saline (PBS) and all other materials that were used for culturing were placed in the anaerobic cabinet 24 hours before use. The faecal samples were divided into two portions. One portion was homogenised in reduced PBS (0.1g stool/ml PBS) and was serially diluted and plated directly onto YCFA (Duncan, Hold et al. 2002) agar supplemented
15 with 0.002g/ml each of glucose, maltose and cellobiose in large (13.5cm diameter) petri dishes. This sample was also subjected to metagenomic sequencing to profile the entire community. The other portion was treated with an equal volume of 70% (v/v) ethanol for 4 hours at room temperature under ambient aerobic conditions to kill vegetative cells. Then, the solid material was washed 3 times with PBS and it was eventually resuspended in PBS.
20 Plating was performed in same manner as described for the non-ethanol treated samples above.

For the ethanol-treated samples of CIP1, the medium was supplemented with 0.1% sodium taurocholate to stimulate spore germination. Colonies were picked 72 hours after plating
25 from petri dishes of both ethanol-treated and non-ethanol-treated conditions harbouring non-confluent growth, (i.e. plates on which the colonies were distinct and not touching). The colonies that were picked were re-streaked to confirm purity.

For CIP2, twelve individuals who had each experienced more than three recurrences of
30 *Clostridium difficile* infection (CDI), having failed treatment with metronidazole and vancomycin were selected for Faecal Microbiota Transplantation (FMT). The donors were screened for pathogens and other viral infections as previously described (Landy, Al-Hassi et al. 2011). The patients discontinued oral vancomycin 1-2 days before FMT. FMT was administered to recipients by enema (n = 3), pills (n = 6), a combination of both (n = 2, R8
35 and R10), or by nasogastric infusion (n = 1, R7). Faecal samples were collected from the patients 1-2 days after stopping vancomycin treatment (pre-FMT) and at different times post-

FMT. Stool samples used for FMT were also collected from the donors. Samples from healthy individuals and individuals infected with *C. difficile* treated with antibiotics were also included as controls. Faecal samples were collected in sterile containers and were frozen at -80^{0}C. DNA was extracted from all samples for 454 sequencing and subsequent analysis as
5 described below.

For the culturing of samples from the faecal microbiota transplant (FMT) recipients (CIP2), 50 mg of each faecal sample was mixed thoroughly in 0.5 ml sterile, reduced phosphate buffered saline (PBS). The homogenate was serially diluted to 10^{-6} and aliquots of this
10 dilution were plated on a panel of media under anaerobic conditions. The following media were used: fastidious anaerobic agar (FAA, Lab M Ltd, Lancashire, UK) containing 2% defibrinated horse blood, Brain Heart Infusion (BHI, Oxoid UK), de Man Rogosa Sharpe and CCEY (Bioconnections, UK) agars with and without the addition of 10 µg/ml vancomycin (AppliChem, Germany). All plated media were incubated anaerobically at 37°C for 48-72
15 hours except for the BHI agar, which was incubated aerobically at 37°C for 24-48 hours.

Microbiota profiling and sequencing

Identification of each cultured isolate was performed by PCR amplification of the full length
20 16S rRNA gene (using 7F (5'–AGAGTTTGATYMTGGCTCAG-3') forward primer and 1510R (5'-ACGGYTACCTTGTTACGACTT-3') reverse primer followed by capillary sequencing. For both CIP1 and CIP2, 16S rRNA gene sequence reads were aligned in the Ribosomal Database Project (RDP) and manually curated in ARB (Ludwig, Strunk et al. 2004). For CIP1, the R package seqinr version 3.1 was used to determine sequence similarity between
25 16S rRNA gene sequences and, as full-length 16S rRNA gene sequence reads were generated, 98.7% was used as the species-level cut-off to classify reads to Operational Taxonomic Units (OTUs) (Bosshard, Abels et al. 2003, Clarridge 2004). As only partial length 16S rRNA gene sequence reads were generated for candidate bacteria from CIP2, 97% was used as the species-level cut-off (Bosshard, Abels et al. 2003, Clarridge 2004) and
30 the OTUs at this cut-off were determined using mothur (Schloss, Westcott et al. 2009). For both CIP1 and CIP2, the 16S rRNA gene sequence of each species-level OTU was then compared to the Ribosomal Database Project (RDP) reference database to assign taxonomic designations down to the genus level (Wang, Garrity et al. 2007). A BLASTn search was then performed with the 16S rRNA gene sequences to determine whether the
35 OTU represented either a previously characterised or a novel species (Altschul, Gish et al.1990).

Comparisons of the OTUs with the Human Microbiome Project (HMP) “Most Wanted” list and reference genomes database were carried out using 97% sequence identity of the 16S rRNA gene sequences to define a bacterial species because only partial 16S rRNA gene sequences were available for the bacteria on the HMP “Most Wanted” list and reference genomes database. HMP data regarding the most wanted taxa and the completed sequencing projects were downloaded from the NIH Human Microbiome Project’s "Most Wanted" Taxa from the Human Microbiome for Whole Genome Sequencing (Web 8 March 2016 <http://hmpdacc.org/most_wanted/#data>) and the NIH Human Microbiome Project’s Reference Genomes Data (Web 8 March 2016 <http://hmpdacc.org/HMRGD/>), respectively. Genomic DNA was extracted from at least one representative of each unique OTU using a phenol-chloroform based DNA isolation procedure. DNA was sequenced on the Illumina HiSeq platform generating read lengths of 100bp and these were assembled and annotated for further analysis.

DNA was also extracted directly from each faecal sample for whole community metagenomic and 16S rRNA gene amplicon sequencing using the MP Biomedical FastDNA SPIN Kit for soil. To enable comparisons with the complete community samples, non-confluent cultures were scraped from agar plates 72 hours after inoculation with the initial faecal sample and DNA was extracted from this community using the same DNA isolation process. 16S rRNA gene amplicon libraries were made by PCR amplification of variable regions 1 and 2 of the 16S rRNA gene using the Q5 High-Fidelity Polymerase Kit supplied by New England Biolabs. Primers 27F AATGATACGGCGACCACCGAGATCTACAC TATGGTAATT CC AGMGTTYGATYMTGGCTCAG (1st part = Illumina adapter, 2nd = forward primer pad, 3rd= Forward primer linker and 4th = Forward primer) and 338R CAAGCAGAAGACGGCATACGAGAT ACGAGACTGATT AGTCAGTCAG AA GCTGCCTCCCGTAGGAGT (1st part= reverse complement of 3′ Illumina adapter, 2nd = golay barcode, 3rd =reverse primer pad, 4th =reverse primer linker and 5th =reverse primer) were used. Four PCR amplification reactions per sample were carried out; products were pooled and combined in equimolar amounts for sequencing using the Illumina MiSeq platform, generating 150 bp reads.

For 454 amplicon sequencing of the CIP2 derived faecal samples, DNA was extracted directly from the faecal samples (70 mg) using the FastDNA Spin Kit for Soil on a Fastprep instrument (MP Biomedicals, USA) following the manufacturer’s instructions. The V3-V5 regions of the 16S rRNA gene were amplified using barcoded primers 338F (5′-

ACTCCTACGGGAGGCAGCAG-3′) and 926R (5′-CCG TCA ATT CMT TTR AGT-3′) adapted with linkers. Thermocycling involved an initial 2-min denaturation step at 94^{0}C followed by 20 cycles of denaturation (94^{0}C for 30 s), annealing (53^{0}C for 30 s) and elongation (68^{0}C for 2 min). The PCR products were purified using the Wizard SV Gel and PCR Clean-Up System
5 (Promega, UK) following the manufacturer's protocol and quantified using the Qubit® dsDNA HS Assay Kit (Life_Technologies, UK). Equimolar volumes of each cleaned-up products of each PCR reaction were sequenced on the Roche 454 FLX-Titanium platform.

Microbiota analysis

10

A maximum likelihood phylogeny of the culture derived bacteria from CIP1 was generated from the aligned RDP sequence using FastTree version 2.1.3 (Price, Dehal et al. 2010) with the following settings: a Generalised Time-Reversible (GTR) model of nucleotide substitution and CAT approximation of the variation in rates across sites with 20 rate categories. The
15 ethanol resistant phylogeny was derived directly from the entire culture phylogeny. All phylogenetic trees were edited in ITOL (Letunic and Bork 2011).

Analysis of the partial 16S rRNA gene sequence generated from the 16S rRNA gene amplicon libraries from the CIP1 derived faecal samples was carried out using the mothur
20 MiSeq SOP (Kozich, Westcott et al. 2013) on August 29th 2014, identifying 7549 OTUs across all samples. A sequence identity threshold of ≥ 97% was again used to define an OTU.

For the 454 sequence analysis from CIP2 derived faecal samples the sequence reads were
25 trimmed, filtered and pre-processed using the mothur software 454 SOP accessed in November 2012 (Schloss, Westcott et al. 2009, Schloss, Gevers et al. 2011). To ensure high quality sequence data for analysis, the sequences were trimmed using a window size of 50 bp (average quality score of 35 bp), homopolymers > 8 bp were removed and no ambiguous bases or mismatches in the primer sequence were allowed. Redundant sequence reads
30 were removed to generate unique sequences, which were aligned to the SILVA alignment database (Pruesse, Quast et al. 2007). These aligned sequences were screened to ensure that sequences overlapped in the same alignment space using the *screen.seqs* command in mothur. Unique sequences were again generated and the sequences were preclustered to remove sequences that were likely due to pyrosequencing errors (Huse, Dethlefsen et al.
35 2008). Chimeric sequences were removed using Perseus (Quince, Lanzen et al. 2011) and other contaminants such as chloroplast and mitochrondria were also removed. Sequences

with ≥ 97% sequence identity and their assigned taxonomy from phylum to genus level according to the Ribosomal Database Project (RDP) (Cole, Wang et al. 2014) and SILVA (Pruesse, Quast et al. 2007) databases were regarded as belonging to the same operational taxonomic units (OTUs). The species diversity in each sample was measured by calculating the Shannon diversity Index (SDI), which takes into account both species richness and relative proportional abundance (Schloss, Westcott et al. 2009). The OTUs were then used to cluster dendrograms, using the Bray Curtis calculator in the mothur package. Other analyses, such as Invsimpson index, principal component analysis (PCA) and the UniFrac method of comparing microbial communities were performed as described previously using the mothur software (Lozupone and Knight 2005, Lawley, Clare et al. 2012).

Metagenomic Analysis

Microbial abundance was calculated using the Human Pan-Microbe Community Database (Forster, Browne et al. 2015) against 1883 healthy individuals (3218 samples) and 458 diseased individuals (628 samples). Occurrence was calculated as greater than 1000, independent, normalised reads with abundance calculated relative to total high quality reads within the sample. Antimicrobial resistance and virulence factor identification were performed using automated sequence homology search against protein sequences annotated in the complete genome sequence. The antimicrobial resistance reference list was defined based on the comprehensive antimicrobial CARD database (McArthur, Waglechner et al. 2013) while toxins were identified by occurrence in the Database of Bacterial Exotoxins for Humans (DBETH) (Chakraborty, Ghosh et al. 2012).

Experimental set-up and results

The inventors established methods to isolate and identify bacteria for incorporation into a therapeutic composition tailored to the treatment of dysbiosis of the gastrointestinal tract, as well as e.g. enteric infections, such as, but not limited to, those caused by *Clostridium difficile*. As mentioned above, two different approaches for acquiring bacterial candidates for inclusion in a therapeutic were employed. The first (CIP1) relied on a broad culturing approach from healthy adult donors to establish a culture collection which is as representative as possible of the bacterial component of the healthy human intestinal microbiota. This process also incorporated a targeted culturing approach to preferentially select bacteria displaying a particular phenotype or function e.g. spore formation. The second process (CIP2) was more targeted in nature and aimed to acquire bacterial species specifically associated with resolving gastrointestinal dysbiosis by comparing individuals before and FMT to resolve *C. difficile* associated dysbiosis. These two approaches are described in more detail below.

CIP1- *broad culturing approach to identify therapeutic candidates:*

The inventors first sought to establish a genomic-based workflow that could be used as a platform for targeted culturing of specific bacterial phenotypes (**Figure 1**). Fresh faecal samples were collected from 6 healthy humans and defined the resident bacterial communities with a combined metagenomic sequencing and bacterial culturing approach. Applying shotgun metagenomic sequencing, the inventors profiled and compared the bacterial species present in the original faecal samples to those that grew as distinct colonies on agar plates containing the complex, broad range bacteriological medium, YCFA (Duncan, Hold et al. 2002) supplemented with glucose, maltose and cellobiose. Importantly, a strong correlation was observed between the two samples at the species level (Spearman Rho = 0.75, $p < 0.01$) (**Figure 2**). When sequenced, the original faecal sample and the cultured bacterial community shared an average of 93% of raw reads across the 6 donors.

These results demonstrate that surprisingly, and contrary to the established view in the art, a significant proportion of the bacteria within the faecal microbiota can be cultured with a single growth medium. Thus, a broad range culturing method was established that, when combined with high throughput archiving or specific phenotypic selection, can be utilised to isolate and identify novel bacteria from the human gastrointestinal tract.

The human intestinal microbiota is dominated by strict anaerobic bacteria that are extremely sensitive to ambient oxygen. Certain members of the Firmicutes, including *Clostridium difficile*, produce metabolically dormant and highly resistant spores during colonisation that facilitate both persistence within the host and environmental transmission (Lawley, Clare et al. 2009, Francis, Allen et al. 2013, Janoir, Deneve et al. 2013). Relatively few intestinal spore-forming bacteria have been cultured to date and while metagenomic studies suggest that other unexpected members of the intestinal microbiota possess potential sporulation genes, they remain poorly characterised (Galperin, Mekhedov et al. 2012, Abecasis, Serrano et al. 2013, Meehan and Beiko 2014, Rajilic-Stojanovic and de Vos 2014).

The inventors hypothesized that sporulation might be an unappreciated basic phenotype of the human intestinal microbiota that may have a profound impact on microbiota persistence and spread between humans. Spore-formation is also viewed as desirable for bacteriotherapy formulations since the resistant nature of the spore structure would promote survival of the medicine during production and subsequent storage. Spores from *C. difficile* are resistant to ethanol and this phenotype can be used to select for spores from a mixed population of spores and ethanol-sensitive vegetative cells (Riley, Brazier et al. 1987). Faecal samples with or without ethanol treatment were processed using our combined culture and metagenomics workflow (**Figure 1**). Principle component analysis demonstrated that ethanol treatment profoundly altered the culturable bacterial composition and when compared to the original profile, efficiently enriched for ethanol-resistant bacteria, facilitating their isolation (**Figure 3**). ~2,000 individual bacterial colonies were picked from both ethanol-treated and non-ethanol-treated conditions, re-streaked them to purity and performed full-length 16S rRNA gene sequencing to enable taxonomic characterisation. Unique taxa were then archived as frozen stocks for future phenotypic analysis.

In total, bacteria representing 96% of the bacterial abundance at the genus level and 90% of the bacterial abundance at the species level based on average relative abundance across the 6 donors (**Figure 4A and 4B**) were archived. Even genera that were present at low average relative abundance (<0.2%) were isolated and purified (**Figure 4C**). Ethanol-resistant species were isolated from 5 known families (*Clostridiaceae, Peptostreptococcaceae, Lachnospiraceae, Ruminococcaceae* and *Erysipelotrichaceae*) and 2 newly identified candidate families (bacterial isolates HMI_1 and HMI_22) (see Table 1 for details). The identification of these new and unexpected spore-formers highlights the broad taxonomic distribution of this phenotype among the enteric species of the Firmicutes. Overall, 137 distinct bacterial species including 45 candidate novel species and isolates

representing 20 candidate novel genera and 2 candidate novel families were archived (**Figure 5**). Our collection contains 90 species from the Human Microbiome Project's (HMP) 'most wanted' list of previously uncultured and unsequenced microbes (Fodor, DeSantis et al. 2012). 19 of the deposited bacterial isolates listed in Table 1 are included in the HMP's
5 "Most Wanted" list, namely: HMI_1, HMI_2, HMI_4, HMI_5, HMI_7, HMI_11, HMI_12, HMI_15, HMI_16, HMI_17, HMI_18, HMI_19, HMI_35, HMI_37, HMI_38, HMI_39, HMI_45, HMI_50 and HMI_51 (see Table 1 for details). Thus, our broad-range YCFA based culturing approach led to the discovery of large numbers of novel bacteria (including novel families, genera, species, and isolates) and challenges the prevailing perception in the art that the
10 majority of the intestinal microbiota is "unculturable".

CIP2*: Targeted identification of candidates to resolve gastrointestinal dysbiosis*

As described above, FMT has proven effective in resolving CDI. The inventors therefore
15 sought to culture from faecal samples from FMT donors and recipients to isolate candidate bacteria that could be used in therapy. A panel of different microbiological media were tested to recover the broadest range of bacterial species from the faecal samples (see Methods). This approach allowed culturing and archival of bacterial candidates. Over 2600 bacterial isolates were cultured and using 16S rRNA gene sequencing these were taxonomically
20 classified (**Figure 7**). These bacterial isolates were members of the 4 major phyla (Actinobacteria, Bacteroidetes, Proteobacteria and Firmicutes) in the intestinal microbiota. These bacterial isolates represented more than 350 different OTUs based on alignments of the partial length 16S rRNA gene.

25 *In silico* analysis of candidate bacteria:

Having established a culture collection through the two approaches described above (CIP1 and CIP2), the inventors next sought to screen these bacteria to identify bacterial candidates for bacteriotherapy.

30

The inventors first sought to analyse the isolates cultured from FMT donors and recipients. At one to three months post-FMT, the faecal microbiota profiles of the recipients were similar to those of the donors and the healthy controls. In particular, the relative abundances of the four major bacterial phyla present in the gut microbiota were also similar across these
35 groups. The microbiota community structure in the donors and recipients (before and after FMT) was visualised for evaluation using principal component analysis (PCA) (**Figure 6**).

The PCA plot demonstrates the presence of two distinct compositional profiles representing a “health-associated” microbiota, and a “vancomycin” microbiota. The health-associated profile contained samples from FMT donors, healthy controls and FMT recipients at 2-3 months post treatment. The “vancomycin” microbiota profile was separated from the health-associated microbiota along Principal Component 1 and only contained individuals treated with vancomycin. These vancomycin control individuals became infected with *C. difficile* while taking antibiotics to treat other disease conditions. Additionally, the “metronidazole-associated” profile was separated from the “health-associated” profile along Principal Component 2 and contained samples from *C. difficile* infected patients treated with metronidazole.

The donor-recipient profiles of each pair were compared before and after FMT to identify taxa that were present in the donor sample, and which increased in relative abundance in the recipient’s profile after FMT. A total of 786 OTUs from all recipient samples were detected after FMT but singleton OTUs present at the different time points were removed. This resulted in 375 OTUs for further analysis. Given that recurrence of CDI typically occurs 3-4 weeks after the withdrawal of antibiotic treatment, (Cornely, Miller et al. 2012, Abujamel, Cadnum et al. 2013) OTUs that were increased in relative proportional abundances at 2-3 months post-FMT were analysed further.

Next, the inventors undertook *in silico* analysis to further screen the bacteriotherapy candidates from both of our culturing approaches (CIP1 and CIP2). As described above a healthy intestinal microbiota is based on a diverse and abundant microbial community. Using the whole genome sequences that the inventors generated from bacterial isolates from CIP1 and CIP2, the inventors computationally assessed their prevalence in healthy and diseased individuals in public metagenomic data-sets using the HPMC database tool (Forster, Browne et al. 2015). Candidate bacteria were first filtered to include only those isolates with greater than 0.001% average abundance within the bacterial community across all healthy individuals in which they were detected (**Figure 8**). All of the bacteria deposited with DSMZ thus had greater than 0.001% average abundance within the bacterial community across all healthy individuals in which they were detected (see Table 1). In addition to being health-associated, preferred candidates for bacteriotherapy applications are expected to ameliorate gastrointestinal dysbiosis. To identify such candidates, the distribution of each of our isolates in publicly available metagenomics datasets was examined. Bacterial species whose total average abundance was substantially decreased (greater than four-fold decrease) in individuals with gastrointestinal dysbiosis relative to healthy individuals were selected and

subjected to further analysis as described below (**Figure 9**). All of the bacteria deposited with DSMZ thus showed a decreased total average abundance (greater than four-fold decrease) in individuals with gastrointestinal dysbiosis relative to healthy individuals (see Table 1).

The list of bacteriotherapy candidates was further analysed on the basis of computationally predicted antimicrobial resistance (AMR) and virulence factors. Bacteriotherapy candidates with overall predicted resistance scores below 20% of the overall predicted resistance scores of the known pathogens *C. difficile*, *Enterococcus faecalis* and *Escherichia coli* were included. Candidates were also selected for the absence of *in-silico* predicted resistance to beta-lactams, fusidic acid, elfamycin, aminoglycoside, fosfomycin and tunicamycin and by the absence of known toxins as listed in Chakrabory A. et al, 2012, A Database of Bacterial Exotoxins for Humans (DBETH). Based on this analysis the inventors identified 51 candidates for use in bacteriotherapy from CIP1 and CIP2 (see **Table 1**). 10 of these bacteriotherapy candidates were identified using CIP2, namely: HMI_23, HMI_24, HMI_25, HMI_26, HMI_27, HMI_28, HMI_29, HMI_30, HMI_31 and HMI_32 (see Table 1 for details). All of these 10 isolates were cultured from healthy donors. The remaining bacteriotherapy candidates were identified using CIP1.

The bacteriotherapy candidates identified using CIP1 and CIP2, with the exception of HMI_17, were then subjected to *in vitro* analysis to establish their therapeutic efficacy in treating *C. difficile* and *E. coli* infection as described in Example 2 below.

Example 2 – *In vitro* analysis of bacteriotherapy candidates

Detection of anti-pathogen activities of the bacteriotherapy candidates by an overlay assay

The bacterial isolates of interest identified in Example 1 were streaked in an “X” shape over the surface of a standard Petri dish containing warmed and reduced YCFA agar. These inoculated plates were then incubated anaerobically at 37°C for 3 to 6 days, until bacterial growth was clearly visible. Overlay agar was prepared by adding 0.8% agar to an appropriate broth. For *C. difficile*, BHI broth + 0.8% agar was used. For *E. coli*, LB + 0.8% agar was prepared. The overlay agar was held molten at 50°C before use. The overlay agar was inoculated (1% inoculum) with an aliquot of a turbid culture of the pathogen of interest, in this case either *C. difficile* M7404 or *E. coli* (AIEC). A 10ml aliquot of the inoculated overlay agar was added to the surface of the agar plates bearing each

commensal strain of interest. The overlay agar was allowed to set and the plates were incubated anaerobically at 37°C for one to two days. Following incubation, zones of clearing could be observed if the commensal strain of interest was capable of inhibiting the growth of the pathogen in the overlay layer. The width of each zone of clearing was measured with a
5 ruler as shown in **Figure 10**. The results are shown in **Figure 11**.

Detection of anti-pathogen activity by CFS-relative growth inhibition assay.

Bacteriotherapy candidates were grown in 1ml aliquots of reduced YCFA broth at 37°C
10 under anaerobic conditions for two days. Cell free supernatant (CFS) was prepared by centrifuging each culture to remove the bacteria and by passing the resulting supernatant through a 0.22µm filter to sterilise it. Uninoculated YCFA broth was also filter sterilised. The CFSs and filtered YCFA broth aliquots were frozen at -20°C until they were required. These filtrates were thawed under anaerobic conditions at 37°C and a 100µl aliquot of each CFS
15 was added to one well of a flat-bottomed 96-well plate. Several wells were filled with filter-sterilised YCFA broth to serve as positive controls for pathogen growth. Each well was inoculated (2-5% inoculum) with a turbid, early-mid exponential phase *C. difficile* M7404 culture. Alternatively, a 5% inoculum of a stationary phase *E. coli* culture adjusted to OD_{600} ≈1 was used. The 96-well plate was sealed with an optically clear film and it was transferred
20 to a FLUOstar Omega microplate-reader (BMG Labtech). The plate was incubated static at 37°C in the plate-reader and OD_{600} readings were taken every 10min for 18.17 hours. The plate was shaken for 10 seconds before each OD reading was taken. All isolates, except HMI17 were tested.

25 The relative growth of the pathogen of interest in each of the CFSs tested was then calculated as follows: For each CFS tested, every raw-data value was expressed relative to its OD_{600} reading that was taken at the ten minute time-point. Such data normalisation permitted direct comparison of *C. difficile* or *E. coli* growth in the various CFSs by eliminating the initial inherent variation in the optical density of the CFSs (due to the pre-fermentation of
30 the media) from consideration. The relative growth achieved at the 18.17h time-point by *C. difficile* or *E. coli* in each of the CFSs was compared to the relative growth of the pathogen of interest that was achieved in YCFA broth. A commensal strain was considered a potential inhibitor of *C. difficile* or *E. coli* if the relative growth plus two standard deviations of the pathogen of interest in the CFS derived from that same commensal isolate, was less than
35 the mean minus two standard deviations of the relative growth of the pathogen in YCFA broth. Where only one relative growth value was available, a CFS was considered as

potentially inhibitory if the relative pathogen growth was more than two standard deviations below the mean relative growth in YCFA broth. The results for bacteriotherapy candidates found to have inhibitory activity are shown in **Figure 12**.

5 *Results*

A summary of the results obtained in the growth overlay and growth inhibition assays is shown in **Figure 13** and **Table 1**. The bacteriotherapy candidates which showed activity in each of the *in vitro* assays are indicated in this figure.

10

Of the 50 bacteriotherapy candidates tested, 22 demonstrate growth inhibition of at least one of *C. difficile* M7404 or *E. coli* (AIEC) in one of the assays performed. 11 of the bacteriotherapy candidates inhibited the growth of at least one of either *C. difficile* or *E. coli* in overlay assays, suggesting that the inhibition conferred by these bacteriotherapy
15 candidates is direct. According to the overlay assay data, 5 of the bacteriotherapy candidates inhibit only the growth of either *C. difficile* or *E. coli*, suggesting that the inhibitory activity of these bacteriotherapy candidates is not generic, i.e. that the inhibitory activity is specific for one or more pathogenic bacteria.

20 Of the 50 bacteriotherapy candidates tested, 6 inhibited the growth of both *C. difficile* and *E. coli* in overlay assays, suggesting that they have a broad-spectrum of inhibitory activity and are likely to also have inhibitory activity against other pathogenic bacteria.

The results from the CFS-relative growth inhibition assay demonstrate that the CFS from 16
25 of the 50 bacteriotherapy candidates tested, only supported the relative growth of *C. difficile* to levels more than two standard deviations below the mean relative growth in YCFA broth at the 18.17h time-point. These bacteriotherapy candidates are thus considered to inhibit *C. difficile* growth. 5 of these bacteriotherapy candidates were also shown to directly inhibit *C. difficile* and/or *E. coli* growth in the overlay assays. This suggests that these 5
30 bacteriotherapy candidates secrete one or more substances which inhibit the growth of these pathogenic bacteria. The remaining 11 bacteriotherapy candidates which showed inhibitory activity in the CFS-relative growth inhibition assay are likely to compete with *C. difficile* for nutrients. CFS from two of the bacteriotherapy candidates did not support growth of *E. coli* to within two standard deviations of the mean growth observed for *E. coli* in YCFA
35 broth. These isolates are therefore considered to inhibit the growth of *E. coli*.

Example 3 - *In silico* co-abundance network analysis

To identify bacteria that, while not capable of directly inhibiting pathogen growth as tested in Example 2, may support the growth or survival of those bacteria that exhibited direct inhibition of pathogen growth in Example 2, co-abundance network analysis was performed. This analysis was performed as described previously using the complete list of healthy datasets in the HPMC database tool (Forster, Browne et al. 2015). For each candidate bacterium that demonstrated inhibition of pathogen growth in Example 2, a complete list of first degree neighbour species that exhibited co-occurrence with the candidate bacterium across at least 95% of faecal samples with an average abundance greater than 0.001% and a minimum of 100 reads was generated. Bacteria that exhibited extensive co-occurrence with candidate bacteria exhibiting direct inhibitory activity of pathogen growth are predicted to provide a metabolic, environmental and/or immunomodulatory support function required for colonization of the gastrointestinal tract by the candidate bacteria. The deposited bacteria demonstrating such co-occurrence is indicated in Table 1.

Discussion

Bacterial isolates that inhibited the growth of one or more pathogenic bacteria as shown in Example 2 are expected to be suitable for treating gastrointestinal dysbiosis in humans.

However, bacterial isolates that did not show evidence of pathogen inhibition in Example 2 are still expected to be useful for the treatment of gastrointestinal dysbiosis.

Firstly, based on the co-occurrence data obtained in Example 3, a large number of the deposited bacteria are expected to support the colonization of the gastrointestinal tract by the inhibitory bacteria identified in Example 2 through direct or indirect interaction. Metabolic networks in which consortia of bacteria thrive by cross-feeding, structural networks, such as biofilms, or the interactions of 'keystone species', allow the microbiota to establish and stabilise (Ze and Mougen et al. 2013). Co-occurrence analysis identified 35 candidates that formed first degree co-occurrence neighbours with direct inhibitors at a rate above 95% (HMI_2, HMI_5, HMI_6, HMI_7, HMI_8, HMI_9, HMI_10, HMI_11, HMI_12, HMI_14, HMI_15, HMI_16, HMI_17, HMI_18, HMI_19, HMI_20, HMI_26, HMI_27, HMI_31, HMI_33, HMI_34, HMI_35, HMI_37, HMI_38, HMI_39, HMI_41, HMI_42, HMI_43, HMI_44, HMI_46, HMI_47, HMI_48, HMI_50, HMI_51, HMI_52; see Table 1 for details). In addition, several of the bacterial isolates listed in Table 1 reside within the same genera as known keystone

species (HMI_17, HMI_23 to HMI_32, HMI_45, HMI_49, HMI_51 and HMI_52; see Table 1 for details) and thus are expected to represent keystone species themselves.

Secondly, the bacterial isolates listed in Table 1 are shown in Example 1 to contribute to the overall diversity of the gastrointestinal microbiota, which is low during dysbiosis. Specifically, a number of these bacteria (HMI_23 to HMI_32 inclusive) were recovered from the intestinal microbiota of FMT donors as part of CIP2. When the microbiota of a healthy donor was transferred to an individual with dysbiosis due to antibiotic treatment for recurrent *C. difficile* infection, all were restored to health (**Figure 6**), which was determined as the absence of *C. difficile* at 2-3 months post-FMT. The criteria for identifying bacteriotherapy candidates by the CIP2 process required that a certain candidate bacterial species was present in the microbiota of half the recipients at more than >0.6% average relative abundance at 2-3 months post-FMT. Moreover, genera representing several of the 51 bacteria listed in Table 1 were also identified in healthy donors and in cured recipients post-FMT (**Figure 14**). Furthermore, the *in silico* analysis presented herein (**Figures 8** and **9**) revealed that the 51 candidate bacteriotherapy isolates are prevalent in healthy individuals, in whom they occur at an average relative abundance >0.001% and these bacteria tend to be depleted under conditions of dysbiosis (**Figure 9**). Together, these data strongly suggest that the 51 bacterial isolates listed in Table 1 are suitable in the treatment of gastrointestinal dysbiosis.

Thirdly, the bacterial isolates listed in Table 1 are expected to compete with enteric pathogens in the gastrointestinal tract, and thus find application in the treatment of gastrointestinal dysbiosis. Specifically, the widespread occurrence of these bacteria in healthy individuals implies that they efficiently colonise the gastrointestinal tract. When the microbiota is populated by these health-associated bacteria, the likelihood of enteric infection with any pathogenic bacterium is known to be low, as such infections usually do not occur in individuals with a healthy gastrointestinal microbiota. Indeed, following FMT, during which genera representing many of the 51 bacteria listed in Table 1 were identified in individuals treated for a dysbiosis of the gastrointestinal tract following antibiotic treatment for *C. difficile* infection (**Figure 14**), a healthy microbiota profile was restored (**Figure 6** and **Figure 14**) and *C. difficile* infection did not recur within 3 months. This indicates that these bacteria promote health according to the principles of colonisation resistance, in which pathogens are excluded or suppressed by competition with the resident health-associated bacteria for nutrients and attachment sites (Britton & Young 2014; Lawley & Walker 2013).

Fourthly, several of the bacterial isolates listed in Table 1 are expected, on the basis of extrapolation from investigations of other species in the same genera or clades (Louis & Flint, 2009), to produce metabolites, such as short chain fatty acids, which have known benefits for gastrointestinal health (e.g. HMI_9, HMI_12, HMI_20, HMI_21 and HMI_23-
5 HMI_32; see table 1 for details).

Finally some *Clostridium* related species have been shown to be immunomodulatory and can be beneficial in reducing inflammation (Atarashi, Tanoue et al. 2013). Based on a comparison of 16S rRNA gene sequences using 95% sequence identity as a cut-off to define
10 a genus (Bosshard, Abels et al. 2003), examples in this context that are in the same genus as these bacteria are HMI_4, HMI_9, HMI_10, HMI_15, HMI_27, HMI_28 and HMI_38.

Table 1 – Characteristics of deposited bacteriotherapy candidates

16S rRNA gene sequence	Reference number of deposited bacterial isolate	Closest characterised bacterial species as determined by BLAST analysis	Leibniz-Institut DSMZ (DSMZ) accession number for the deposited bacterial isolate	On Human Microbiome Project's "Most Wanted List"	Average abundance > 0.001%	Reduced average abundance in gastrointestinal dysbiosis	Spore-former (isolated from an ethanol-treated sample)	Present post FMT (CIP2 criteria)	Adheres to CIP2 selection criteria	Pathogen (*C.difficile* and/or *E.coli*) inhibition	Co-occurrence with inhibitor	Production of beneficial metabolites (SCFA)	Keystone species	Immunomodulation activity
1	HMI_1	*Clostridium thermocellum*	DSM32191											
2	HMI_2	*Flavonifractor plautii*	DSM32147											
3	HMI_3	*Flavonifractor plautii*	DSM32149											
4	HMI_4	*Clostridium orbiscindens*	DSM32175											
5	HMI_5	*Ruminococcus flavefaciens*	DSM32153											
6	HMI_6	*Anaerotruncus colihominis*	DSM32152											
7	HMI_7	*Clostridium xylanolyticum*	DSM32158											
8	HMI_8	*Clostridium oroticum*	DSM32192											
9	HMI_9	*Eubacterium contortum*	DSM32148											
10	HMI_10	*Clostridium oroticum*	DSM32166											
11	HMI_11	*Lachnospira pectinoschiza*	DSM32151											
12	HMI_12	*Roseburia faecis*	DSM32150											
13	HMI_14	*Clostridium hathewayi*	DSM32193											
14	HMI_15	*Fusicatenibacter saccharivorans*	DSM32162											
15	HMI_16	*Clostridium clostridioforme*	DSM32194											
16	HMI_17	*Ruminococcus torques*	DSM32163											
17	HMI_18	*Clostridium celerecrescens*	DSM32205											
18	HMI_19	*Clostridium celerescens*	DSM32195											
19	HMI_20	*Eubacterium infirmum*	DSM32164											
20	HMI_21	*Eubacterium infirmum*	DSM32177											
21	HMI_22	*Clostridium thermocellum*	DSM32167											
22	HMI_23	*Anaerovorax odorimutans*	DSM32165											
23	HMI_24	*Clostridium saccharogumia*	DSM32169											
24	HMI_25	*Clostridium saccharogumia*	DSM32168											
25	HMI_26	*Blautia luti*	DSM32178											
26	HMI_27	*Clostridium clostridioforme*	DSM32182											
27	HMI_28	*Blautia producta*	DSM32179											

16S rRNA gene sequence	Reference number of deposited bacterial isolate	Closest characterised bacterial species as determined by BLAST analysis	Leibniz-Institut DSMZ (DSMZ) accession number for the deposited bacterial isolate	On Human Microbiome Project's "Most Wanted List"	Average abundance > 0.001%	Reduced average abundance in gastrointestinal dysbiosis	Spore-former (isolated from an ethanol-treated sample)	Present post FMT (CIP2 criteria)	Adheres to CIP2 selection criteria	Pathogen (*C.difficile* and/or *E.coli*) inhibition	Co-occurrence with inhibitor	Production of beneficial metabolites (SCFA)	Keystone species	Immunomodulation activity
28	HMI_29	*Blautia glucerasea*	DSM32180		Grey	Grey		Grey	Grey			Grey	Grey	/
29	HMI_30	*Clostridium straminisolvens*	DSM32184		Grey	Grey		Grey	Grey			Grey	Grey	/
30	HMI_31	*Butyricicoccus pullicaecorum*	DSM32181		Grey	Grey		Grey	Grey		Grey	Grey	Grey	/
31	HMI_32	*Clostridium maritimum*	DSM32183		Grey	Grey		Grey	Grey			Grey	Grey	/
32	HMI_33	*Eubacterium fissicatens*	DSM32262		Grey	Grey	Grey	Grey	/	Grey	Grey	/	/	/
33	HMI_34	*Clostridium saccharolyticum*	DSM32211		Grey	Grey		Grey	/	Grey	Grey	/	/	/
34	HMI_35	*Ruminococcus obeum*	DSM32219	Grey	Grey	Grey		Grey	/	Grey	Grey	/	/	/
35	HMI_36	*Clostridium methylpentosum*	DSM32222		Grey	Grey	Grey		/			/	/	/
36	HMI_37	*Clostridium xylanolyticum*	DSM32261	Grey	Grey	Grey	Grey	Grey	/		Grey	/	/	/
37	HMI_38	*Oscillibacter valericigenes*	DSM32212	Grey	Grey	Grey	Grey		/		Grey	/	/	Grey
38	HMI_39	*Ruminococcus obeum*	DSM32220	Grey	Grey	Grey		Grey	/	Grey	Grey	/	/	/
39	HMI_40	*Megasphaera elsdenii*	DSM32213		Grey	Grey			/	Grey		/	/	/
40	HMI_41	*Blautia luti*	DSM32226		Grey	Grey		Grey	/		Grey	/	/	/
41	HMI_42	*Bacteroides coprocola*	DSM32215		Grey	Grey			/	Grey	Grey	/	/	/
42	HMI_43	*Bacteroides plebius*	DSM32216		Grey	Grey			/	Grey	Grey	/	/	/
43	HMI_44	*Roseburia inulinivorans*	DSM32217		Grey	Grey	Grey	Grey	/	Grey	Grey	/	/	/
44	HMI_45	*Ruminococcus albus*	DSM32221	Grey	Grey	Grey		Grey	/			/	Grey	/
45	HMI_46	*Blautia producta*	DSM32218		Grey	Grey		Grey	/	Grey	Grey	/	/	/
46	HMI_47	*Clostridium nexile*	DSM32224		Grey	Grey	Grey	Grey	/	Grey	Grey	/	/	/
47	HMI_48	*Butyricicoccus pullicaecorum*	DSM32214		Grey	Grey	Grey		/		Grey	/	/	/
48	HMI_49	*Ruminococcus flavefaciens*	DSM32263		Grey	Grey		Grey	/			/	Grey	/
49	HMI_50	*Flavonifractor plautii*	DSM32223	Grey	Grey	Grey	Grey		/		Grey	/	/	/
50	HMI_51	*Ruminococcus bromii*	DSM32225	Grey	Grey	Grey	Grey		/		Grey	/	Grey	/
51	HMI_52	*Ruminococcus albus*	DSM32265		Grey	Grey	Grey	Grey	/		Grey	/	Grey	/

Grey fill = Positive/Yes; no fill = Negative/No; slash = No data available/Not tested

5

Sequence listing

16S rRNA gene sequences of the 51 deposited bacterial isolates listed in Table 1 are set out below below. For each bacteriotherapy candidate a putative genus and species name is give. The genus was and species names were assigned based on the Ribosomal Database Project (RDP) reference database and BLASTn analysis as explained in Example 1. The genus and species names assigned to each of the bacteriotherapy candidates are thus that of the most closely related known bacterium and hence subject to change.

HMI_1 Clostridium thermocellum 16S rDNA sequence (SEQ ID NO: 1)

```
CAGGACGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGAGAATCTTTGAACAGA
TCTTTTCGGAGTGACGTTCAAAGAGGAAAGTGGCGGACGGGCGAGTAACGCGTGAGTAAC
CTGCCCATAAGAGGGGGATAATCCATGGAAACGTGGACTAATACCGCATATTGTAGTCAA
GTCGCATGACTAGATTATGAAAGATTTATCGCTTATGGATGGACTCGCGTCAGATTAGAT
AGTTGGTGAGGTAACGGCTCACCAAGTCAACGATCTGTAGCCGAACTGAGAGGTTGATCG
GCCGCATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTG
CGCAATGGGGGCAACCCTGACGCAGCAACGCCGCGTGCAGGAAGAAGGTCTTCGGATTGT
AAACTGTTGTCGCAAGGGAAGAAGACAGTGACGGTACCTTGTGAGAAAGTCACGGCTAAC
TACGTGCCAGCAGCCGCGGTAATACGTAGGTGACAAGCGTTGTCCGGATTTACTGGGTGT
AAAGGGCGCGTAGGCGGACTGTCAAGTCAGTCGTGAAATACCGGGGCTTAACCCCGGGGC
TGCGATTGAAACTGACAGCCTTGAGTATCGGAGAGGAAAGCGGAATTCCTAGTGTAGCGG
TGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGACGACAACT
GACGCTGAGGCGCGAAAGTGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACC
GTAAACGATGGATACTAGGTGTAGGAGGTATCGACCCCTTCTGTGCCGCAGTTAACACAA
TAAGTATCCCACCTGGGGAGTACGACCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGC
CCGCACAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCTGGGC
TTGACATCCCTGGAATCGAGTAGAGATACTTGAGTGCCTTCGGGAATCAGGTGACAGGTG
GTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCA
ACCCCTATTGTCAGTTGCCATCATTAAGTTGGGCACTCTGGCGAGACTGCCGGTGACAAA
TCGGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCCTTATGCCCAGGGCTACACACG
TACTACAATGGCCGATAACAAAGTGCAGCGAAACCGTGAGGTGGAGCGAATCACAAAACT
CGGTCTCAGTTCAGATTGCAGGCTGCAACTCGCCTGCATGAAGTTGGAATTGCTAGTAAT
CGCGGATCAGAATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACAC
CATGAGAGTCGATAACACCCGAAGCCTGTGAGCTAACCTATTAGGAGGCAGCAGTCGAAG
GTGGGGTTGATGATTGGGGTGAAGTCG
```

HMI_2 Flavonifractor plautii 16S rDNA sequence (SEQ ID NO: 2)

GAGTGCTCATGACAGAGGATTCGTCCAATGGAGTGAGTTACTTAGTGGCGGACGGGTGAGTAACGCGTGAGTAAC
CTGCCTTGGAGTGGGGAATAACAGGTGGAAACATCTGCTAATACCGCATGATGCAGTTGGGTCGCATGGCTCTGA
CTGCCAAAGATTTATCGCTCTGAGATGGACTCGCGTCTGATTAGCTGGTTGGCGGGGTAACGGCCCACCAAGGCG
ACGATCAGTAGCCGGACTGAGAGGTTGGCCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCA
GCAGTGGGGAATATTGGGCAATGGGCGCAAGCCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGCTTTCGGGTTG
TAAACTTCTTTTCTCAGGGACGAAGCAAGTGACGGTACCTGAGGAATAAGCCACGGCTAACTACGTGCCAGCAGC
CGCGGTAATACGTAGGTGGCGAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGGACTGCAAGTCA
GATGTGAAAACCATGGGCTCAACCTGTGGCCTGCATTTGAAACTGTAGTTCTTGAGTACTGGAGAGGCAGACGGA
ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGTCTGCTGGACAGCAAC
TGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCTGTAAACGATGGATA
CTAGGTGTGGGGGGTCTGACCCCCTCCGTGCCGCAGTTAACACAATAAGTATCCCACCTGGGGAGTACGATCGCA
AGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAA
GAACCTTACCAGGGCTTGACATCCCGGTGACCGGTGTAGAGATACACCTTCTTCTTCGGAAGCGCCGGTGACAGG
TGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATTGTTA
GTTGCTACGCAAGAGCACTCTAGCGAGACTGCCGTTGACAAAACGGAGGAAGGTGGGGACGACGTCAAATCATCA
TGCCCCTTATGTCCTGGGCCACACACGTACTACAATGGTGGTCAACAGAGGGAAGCAAGACCGCGAGGTGGAGCA
AACCCCTAAAAGCCATCCCAGTTCGGATTGCAGGCTGCAACTCGCCTGTATGAAGTTGGAATCGCTAGTAATCGC
GGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGAGAGTCGGGAACAC
CCGAAGTCCGTAGCCTAACCGCAAGGGGGGCGCGGCCGAAGGTGGGTTCGATAATTGGGGTGAAGTCGT

HMI_3 Flavonifractor plautii 16S rDNA sequence (SEQ ID NO: 3)

TGGCTGTTTAGTGGCGGACGGGTGAGTAACGCGTGAGTAACCTGCCTTGGAGTGGGGAATAACACAGTGAAAACT
GTGCTAATACCGCATGACATATTGGTGTCGCATGGCACTGATATCAAAGATTTATCGCTCTGAGATGGACTCGCG
TCTGATTAGATAGTTGGCGGGGTAACGGCCCACCAAGTCGACGATCAGTAGCCGGACTGAGAGGTTGGCCGGCCA
CATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGGCAATGGGCGCAAGCCTG
ACCCAGCAACGCCGCGTGAAGGAAGAAGGCTTTCGGGTTGTAAACTTCTTTTAACAGGGACGAAGTAAGTGACGG
TACCTGTTGAATAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAT
TTACTGGGTGTAAAGGGCGTGTAGGCGGGACTGCAAGTCAGATGTGAAAACTATGGGCTCAACCCATAGCCTGCA
TTTGAAACTGTAGTTCTTGAGTGTCGGAGAGGCAATCGGAATTCCGTGTGTAGCGGTGAAATGCGTAGATATACG
GAGGAACACCAGTGGCGAAGGCGGATTGCTGGACGATAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAG
GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTAGGTGTGGGGGGTCTGACCCCCTCCGTGCCGCA
GCTAACGCAATAAGTATCCCACCTGGGGAGTACGATCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCA
CAAGCGGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGGCTTGACATCCTACTAACGAAC
CAGAGATGGATTAGGTGCCCTTCGGGGAAAGTAGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGA
TGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATTGTTAGTTGCTACGCAAGAGCACTCTAGCGAGACTGCCG
TTGACAAAACGGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCCTTATGTCCTGGGCCACACACGTACTACA

ATGGCGGTTAACAGAGGGAGGCAAAGCCGCGAGGCAGAGCAAACCCCTAAAAGCCGTCCCAGTTCGGATTGCAGG
CTGAAACCCGCCTGTATGAAGTCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGC
CTTGTACACACCGCCCGTCACACCATGAGAGTCGGGAACACCCGAAGTCCGTAGCCTAACTGCAAAGGGGGCGCG
GCCGAAGGTGGGTTCGATAATTGGGGTGAAGTCGTAACAGGGTAACCG

5

HMI_4 Clostridium orbiscindens 16S rDNA sequence (SEQ ID NO: 4)

TGGCGGACGGGTGAGTAACGCGTGAGGAACCTGCCTCGGAGTGGGGAATAACAGACCGAAAGGCCTGCTAATACC
GCATGATGCAGTTGGACCGCATGGTCCTGACTGCCAAAGATTTATCGCTCTGAGATGGCCTCGCGTCTGATTAGC
10 TTGTTGGCGGGGTAATGGCCCACCAAGGCGACGATCAGTAGCCGGACTGAGAGGTTGGCCGGCCACATTGGGACT
GAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGGCAATGGGCGCAAGCCTGACCCAGCAAC
GCCGCGTGAAGGAAGAAGGCTTTCGGGTTGTAAACTTCTTTTCTCAGGGACGAACAAATGACGGTACCTGAGGAA
TAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGATTTACTGGGTGT
AAAGGGCGTGTAGGCGGGAAGGCAAGTCAGATGTGAAAACTATGGGCTCAACCCATAGCCTGCATTTGAAACTGT
15 TTTTCTTGAGTGCTGGAGAGGCAATCGGAATTCCGTGTGTAGCGGTGAAATGCGTAGATATACGGAGGAACACCA
GTGGCGAAGGCGGATTGCTGGACAGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACC
CTGGTAGTCCACGCTGTAAACGATGGATACTAGGTGTGGGGGGTCTGACCCCCTCCGTGCCGCAGTTAACACAAT
AAGTATCCCACCTGGGGAGTACGATCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGA
GTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGGCTTGACATCCTACTAACGAAGCAGAGATGCAT
20 TAGGTGCCCTTCGGGGAAAGTAGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAA
GTCCCGCAACGAGCGCAACCCTTATTGTTAGTTGCTACGCAAGAGCACTCTAGCGAGACTGCCGTTGACAAAACG
GAGGAAGGCGGGGACGACGTCAAATCATCATGCCCCTTATGTCCTGGGCTACACACGTACTACAATGGTGGTAAA
CAGAGGGAAGCAAGACCGCGAGGTGGAGCAAATCCCTAAAAGCCATCCCAGTTCGGATTGCAGGCTGAAACCCGC
CTGTATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACAC
25 CGCCCGTCACACCATGAGAGTCGGGAACACCCGAAGTCCGTAGTCTAACCGCAAGGGGGACGCGGCCGAAGGTGG
GTTCGATAATTGGGGTGAAGTCGTAACAGGGTAACC

HMI_5 Ruminococcus flavefaciens 16S rDNA sequence (SEQ ID NO: 5)

30 CGGATCAGTGGCGGACGGGTGAGTAACACGTGAGCAACCTGCCTTTAAGAGGGGGATAACGTTTGGAAACGAACG
CTAATACCGCATAACATAGAAGATTCACATGTTTCTTCTATCAAAGATTTATCGCTTAAAGATGGGCTCGCGTCT
GATTAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCGACGATCAGTAGCCGTACTGAGAGGTAGAACGGCCACAT
TGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGAGGAAACTCTGATG
CAGCGATGCCGCGTGAGGGAAGAAGGTTTTCGGATTGTAAACCTCTGTCTTCAGGGACGATAATGACGGTACCTG
35 AGGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCGAGCGTTGTCCGGAATTACTG
GGTGTAAAGGGAGCGTAGGCGGGATCTTAAGTCAGGTGTGAAAACTATGGGCTCAACCCATAGACTGCACTTGAA
ACTGAGGTTCTTGAGTGAAGTAGAGGCAGGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAA
CATCAGTGGCGAAGGCGGCCTGCTGGGCTTTTACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAG
ATACCCTGGTAGTCCACGCTGTAAACGATGATTACTAGGTGTGGGGGGACTGACCCCTTCCGTGCCGCAGTTAAC

ACAATAAGTAATCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCA
GTGGAGTATGTGGTTTAATTCGAAGCACGCGAAGAACCTTACCGGGTCTTGACATCTACAGAATCCTTTAGAGAT
AAGGGAGTGCCCTTCGGGGAACTGTAAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGT
TAAGTCCCGCAACGAGCGCAACCCCTATCATTAGTTGCTACGCAAGAGCACTCTAATGAGACTGCCGTTGACAAA
ACGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCCGGGCTACACACGTACTACAATGGCGTA
ACAGAGGGAAGCAACATCGCGAGGTGAAGCAAATCTCTAAAAAACGTCCCAGTTCAGATTGCAGGCTGCAACTCG
CCTGCATGAAGACGGAATTGCTAGTAATCGCAGATCAGCATGCTGCGGTGAATACGTTCCCGGGCCTTGTACACA
CCGCCCGTCACACCATGGGAGTCGGTAACACCCGAAGTCGCTTGTCTAA

HMI_6 Anaerotruncus colihominis 16S rDNA sequence (SEQ ID NO: 6)

AGTCGACGGACACATCCGACGGAATAGCTTGCTAGGAAGATGGATGTTGTTAGTGGCGGACGGGTGAGTAACACG
TGAGCAACCTACCTCAGAGTGGGGGACAACAGTTGGAAACGACTGCTAATACCGCATAAGATGGCAGGGTCGCAT
GGCCTGGTCATAAAAGGAGCAATTCGCTCTGAGATGGGCTCGCGTCTGATTAGCTAGTTGGTGAGGTAACGGCTC
ACCAAGGCAACGATCAGTAGCCGGACTGAGAGGTTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTA
CGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAGGGAAGACGGTT
TTCGGATTGTAAACCTCTGTCTTGTGGGACGATAGTGACGGTACCACAGGAGGAAGCCATGGCTAACTACGTGCC
AGCAGCCGCGGTAATACGTAGATGGCGAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGTGTAGGCGGGCTGGT
AAGTTGAATGTGAAACCTTCGGGCTCAACCCGGAGCGTGCGTTCAAAACTGCTGGTCTTGAGTGAAGTAGAGGCA
GGCGGAATTCCCGGTGTAGCGGTGGAATGCGTAGATATCGGGAGGAACACCAGTGGCGAAGGCGGCCTGCTGGGC
TTTTACTGACGCTGAGGCTCGAAAGCATGGGTAGCAAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGA
TGATTACTAGGTGTGGGGGGATTGACCCCCTCCGTGCCGGAGTTAACACAATAAGTAATCCACCTGGGGAGTACG
ACCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAA
CGCGAAAAACCTTACCAGGTCTTGACATCCATCGCCAGGCTAAGAGATTAGCTGTTCCCTCCGGGGACGATGAGA
CAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTACT
ATTAGTTGCTACGCAAGAGCACTCTAATGGGACTGCCGTTGACAAAACGGAGGAAGGTGGGGATGACGTCAAATC
ATCATGCCCCTTATGACCTGGGCTACACACGTACTACAATGGCCGTTAACAGAGAGCAGCGATACCGCGAGGTGG
AGCGAATCTAGAAAAACGGTCTCAGTTCGGATTGCAGGCTGAAACTCGCCTGCATGAAGTCGGAATTGCTAGTAA
TCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGAGAGCCGGTA
ACACCCGAAGTCAGTAGCCTAACCGCAAGGAGGGCGCTGCCGAAGGTGGGGCTGGTAATTGGGGTGAAGTCGTAA
C

HMI_7 Clostridium xylanolyticum 16S rDNA sequence (SEQ ID NO: 7)

GTAACGCGTGGGTAACCTGCCTCATACAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAG
GGTCGCATGACCTAGTGTGAAAAACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGTGAGGTAA
CGGCTCACCAAGGCGACGATCAGTAGCCGACCTGAGAGGGTGATCGGCCACATTGGGACTGAGACACGGCCCAAA
CTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAAGGAAG
AAGTATTTCGGTATGTAAACTTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTA

CGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGGCGG
TCTGACAAGTCAGAAGTGAAAGCCCGGGGCTCAACTCCGGGACTGCTTTTGAAACTGCCGGACTAGATTGCAGGA
GAGGTAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTA
CTGGACTGTAAATGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGT
5 AAACGATGAATACTAGGTGTTGGGGAGCACAGCTCTTCGGTGCCGCAGCAAACGCAATAAGTATTCCACCTGGGG
AGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCG
AAGCAACGCGAAGAACCTTACCAAGTCTTGACATCCCGATGACCGTCCCGTAACGGGGGCTTCTCTTCGGAGCAT
CGGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACC
CTTATCTTTAGTAGCCAGCGGTACGGCCGGGCACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGA
10 TGACGTCAAATCATCATGCCCCTTATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAAA
CTGTGAAGTCTAGCAAATCTCAAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGG
AATCGCTAGTAATCGCGAATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCA
TGGGAGTTGGAAACGCCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGCGGGTCTGATAACTGGG
GTGAAGTCGTAACAAGGTAACCG

15

HMI_8 *Clostridium oroticum* 16S rDNA sequence (SEQ ID NO: 8)

TTTTGATTGATTTCTTCGGAAAGAGAGAGACTGTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGC
CTCATACAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCACACAGCTTCGCATGAAGCAGTGTGA
20 AAAACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGGTAGTTGGTGGGGTAACGGCCCACCAAGCCGACGAT
CAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGT
GGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCGGTATGTAAAC
TTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAA
TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTAATGCAAGTCTGGAGTGAA
25 AACCCGGGGCTCAACCCCGGGACTGCTTTGGAAACTGTGTAACTAGAGTGTCGGAGAGGCAAGTGGAATTCCTAG
TGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGCTGGACGATGACTGACGTTG
AGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGT
CGGTAAGCAAAGCTTATCGGTGCCGCAGCAAACGCAATAAGTAGTCCACCTGGGGAGTACGTTCGCAAGAATGAA
ACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTA
30 CCTGGTCTTGACATCCCTCTGACAGCTGAGTAATGTCGGTTTTCTTTCGGGACAGAGGAGACAGGTGGTGCATGG
TTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCA
TATGAGATGGGCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCC
CCTTATGATCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGCCTGCGAGGGGGAGCAAATCC
CAAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAAT
35 CAGAATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGA
AGTCAGTGACCCAACCTTTCAGGAGGGAGCTGCCGAAGGCGGGACCGATAACTGGGGTGAAGTCGT

HMI_9 Eubacterium contortum 16S rDNA sequence (SEQ ID NO: 9)

CTTAAGTTTGATTCTTCGGATGAAGACTTTTGTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCC
TCATACAGGTGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCACAGCACCGCATGGTGCAGGGGTAA
AAACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGTGGGGTAACGGCCTACCAAGGCGACGATC
AGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTG
5 GGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAAGGATGAAGTATTTCGGTATGTAAACT
TCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAAT
ACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTATGGCAAGTCTGATGTGAAA
GGCCAGGGCTCAACCCTGGGACTGCATTGGAAACTGTCGAACTAGAGTGTCGGAGAGGCAAGTGGAATTCCTAGT
GTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTGCTGGACGATGACTGACGTTGA
10 GGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGACTACTAGGTGTC
GGGTAGCAGAGCTATTCGGTGCCGCAGCCAACGCAATAAGTAGTCCACCTGGGGAGTACGTTCGCAAGAATGAAA
CTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTAC
CTGCTCTTGACATCTCCCTGACCGGCAAGTAATGTTGCCTTTCCTTCGGGACAGGGATGACAGGTGGTGCATGGT
TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCTTTAGTAGCCAGCGG
15 TTTGGCCGGGCACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCC
CTTATGAGCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGCCTGCGAGGGTAAGCAAATCTC
AAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAATC
AGAATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTTGGTAACGCCCGAA
GTCAGTGACCCAACCGCAAGGAGGGAGCTGCCGAAGGTGGGACCGATAACTGGGGTGAAGTCGTAACAAGGTAAC
20 CG

HMI_10 Clostridium oroticum 16S rDNA sequence (SEQ ID NO: 10)

ACATGCAAGTCGAGCGAGCGCTTTAGTGGAATTCTACGGAAGGAAAGTGAAGTGACTGAGCGGCGGACGGGTGAG
25 TAACGCGTGGGTAACCTGCCTCATACAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGACCCCAGT
ACCGCATGGTACAGAGGTAAAAACTGAGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGTGAGGTAGA
GGCTCACCAAGGCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAAC
TCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAGCGAAGA
AGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCACCGGCTAAATAC
30 GTGCCAGCAGCCGCGGTAATACGTATGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGA
GCAGCAAGTCTGATGTGAAAACCCGGGGCTCAACCCCGGGAGTGCATTGGAAACTGTTGATCTAGAGTGCTGGAG
AGGTAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTAC
TGGACAGTGACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA
AACGATGACTACTAGGTGTCGGGTAGCAAAGCTATTCGGTGCCGCAGCCAACGCAATAAGTAGTCCACCTGGGGA
35 GTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGA
AGCAACGCGAAGAACCTTACCTGCCCTTGACATCCGGGTGACCGGCGAGTAATGTCGCCTTCTCTTCGGAGCAGC
CGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCC
TTATCTTTAGTAGCCAGCGGATAAGCCGGGGACTCTAGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGAT
GACGTCAAATCATCATGCCCCTTATGGGCAGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAAGC

TGTGAAGCGGAGCGAATCTCAAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGA
ATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCAT
GGGAGTCAGTAACGCCCGAAGTCAGTGACCCAACCGTAAGGAGGGAGCTGCCGAAGGCGGGACGGATAACTGGGG
TGAAGTCGTAACAAGGTAACCG

5

HMI_11 Lachnospira pectinoschiza 16S rDNA sequence (SEQ ID NO: 11)

AGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCCTGTACAGGGGGACAACAGCTGGAAACGGCTGCTAATA
CCGCATAAGCCCTTAGCACTGCATGGTGCATAGGGAAAAGGAGCAATCCGGTACAGGATGGACCCGCGTCTGATT
10 AGCCAGTTGGCAGGGTAACGGCCTACCAAAGCGACGATCAGTAGCCGATCTGAGAGGATGTACGGCCACATTGGG
ACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGAGGAAACTCTGATGCAGC
GACGCCGCGTGAGTGAAGAAGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTA
AGAAGCACCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTATGGTGCAAGCGTTATCCGGATTTACTGGGTG
TAAAGGGAGCGTAGGTGGCAAGGCAAGCCAGAAGTGAAAACCCGGGGCTCAACCGCGGGATTGCTTTTGGAACTG
15 TCATGCTAGAGTGCAGGAGGGGTGAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACC
GGAGGCGAAGGCGGCTCACTGGACTGTAACTGACACTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATAC
CCTGGTAGTCCACGCGGTAAACGATGAATACTAGATGTCGGGTAGCAAAGCTACTCGGTGTCGTCGCAAACGCAA
TAAGTATTCCACCTGGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGG
AGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCTGCTCTTGACATCCCATTCGATAGAGGGTAATGCT
20 TCTAGCCCTTCGGGGGAATGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGT
CCCGCAACGAGCGCAACCCTTATTGTCAGTAGCCAGCAGGTGAAGCTGGGCACTCTGATGAGACTGCCGGGGATA
ACCCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGAGCAGGGCTACACACGTGCTACAATGGCG
TAAACAGAGGGAAGCGAAGGAGTGATCTGGAGCAAATCTCAAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAA
CTCGACTACATGAAGCTGGAATCGCTAGTAATCGCAGATCAAAATGCTGCGGTGAATACGTTCCCGGGTCTTGTA
25 CACACCGCCCGTCACACCATGGGAGTCGGTAATGCCCGAAGTCAGTGACTCAACCGAAAGGAAAAAGCTGCCGAA
GGCAGGACTGGTAACTGGGGTGAAGTCGT

HMI_12 *Roseburia faecis* 16S rDNA sequence (SEQ ID NO: 12)

30 AGTCGAACGAAGCACTTTATTACGATTTCTTCGGAATGACGATTTAGTGACTGAGTGGCGGACGGGTGAGTAACG
CGTGGGTAACCTGCCTTATACAGGGGGATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGGATTGC
ATGATCTGGTGTGAAAAACTCCGGTGGTATAAGATGGACCCGCGTCTGATTAGCTGGTTGGTGAGGTAACGGCCC
ACCAAGGCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTA
CGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAGCGAAGAAGTAT
35 TTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGT
GCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGTGC
GGCAAGTCTGATGTGAAAGCCCGGGGCTCAACCCCGGGACTGCATTGGAAACTGTCGTACTTGAGTATCGGAGAG
GTAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTG
GACGATAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAA

CGATGAATACTAGGTGTCGGGGGACATAGTCCTTCGGTGCCGCAGCAAACGCAATAAGTATTCCACCTGGGGAGT
ACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAG
CAACGCGAAGAACCTTACCAAGTCTTGACATCCCGGTGACAAAGTATGTAATGTACTCTTTCTTCGGAACACCGG
TGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCT
5 GTTCTTAGTAGCCAGCGGTTCGGCCGGGCACTCTAGGGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGGATGA
CGTCAAATCATCATGCCCCTTATGACTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAAAAGG
TGACTTCTAGCAAATCCCAAAAATAACGTCCCAGTTCGGACTGTAGTCTGCAACTCGACTACACGAAGCTGGAAT
CGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGG
GAGTCGGGAATGCCCGAAGCCGGTGACTCAACCGAAAGGAGAGAGCCGTCGAAGGCAGGTCTGATAACTGGGGTG
10 AAGTCGTAACAAGGTAACC

HMI_14 Clostridium hathewayi 16S rDNA sequence (SEQ ID NO: 13)

AGTCGACGGAGATGCGATGTGAGCGAGAGGTGCTTGCACTGATCAATCTTTTCGTATCTTAGTGGCGGACGGGTG
15 AGTAACGCGTGGGTAACCTGCCTTATACCGGGGGATAACACTTAGAAATAGGTGCTAATACCGCATAAGCGCACG
GTGTCGCATGACACAGTGTGAAAAACTCCGGTGGTATAAGATGGACCCGCGTCTGATTAGCCAGTTGGCAGGGTA
ACGGCCTACCAAAGCGACGATCAGTAGCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAA
ACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAGTGAA
GAAGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAGAAATGACGGTACCTGACTAAGAAGCCCCGGCTAA
20 CTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGA
CGGTGAAGCAAGTCTGAAGTGAAAGGTTGGGGCTCAACCCCGAAACTGCTTTGGAAACTGTTTAACTGGAGTACA
GGAGAGGTAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGC
TTACTGGACTGTAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGC
CGTAAACGATGATTACTAGGTGTTGGTGGATATGGATCCATCGGTGCCGCAGCAAACGCAATAAGTAATCCACCT
25 GGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAA
TTCGAAGCAACGCGAAGAACCTTACCTGATCTTGACATCCCTATGAATACAGGGTAATGCCTGTAGTACTTCGGT
ACATAGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGC
AACCCCTATCTTTAGTAGCCAGCAGTAAGATGGGCACTCTAGAGAGACTGCCGGGGATAACCCGGAGGAAGGTGG
GGATGACGTCAAATCATCATGCCCCTTATGACCAGGGCTACACACGTGCTACAATGGCGTAAACAGAGGGAAGCG
30 AAGTGGTGACATGGAGCAAATCCCAAAAATAACGTCCCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGC
TGGAATCGCTAGTAATCGCAGATCAGAATGCTGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACA
CCATGGGAGTAGGTAATGCCCGAAGTCGGTGACCTAACCGCAAGGAAGGAGCCGCCGAAGGCAGGACTTATAACT
GGGGTGAAGTCGTAACAAGGTAACCGT

35 HMI_15 Fusicatenibacter_saccharivorans 16S rDNA sequence (SEQ ID NO: 14)

CGGATGGAATCGGTATAACTTAGTGGCGGACGGGTGAGTAACGCGTGGGAAACCTGCCCTGTACCGGGGGATAAC
ACTTAGAAATAGGTGCTAATACCGCATAAGCGCACGGAACTGCATGGTTCTGTGTGAAAAACTCCGGTGGTACAG
GATGGTCCCGCGTCTGATTAGCCAGTTGGCAGGGTAACGGCCTACCAAAGCGACGATCAGTAGCCGGCCTGAGAG

GGTGAACGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATG
GGGGAAACCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAGA
AAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGT
TATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCATGGCAAGCCAGATGTGAAAACCCAGGGCTCAACCTT
5 GGGATTGCATTTGGAACTGCCAGGCTGGAGTGCAGGAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGT
AGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACTGTAACTGACGTTGAGGCTCGAAAGCGTGGGG
AGCAAACAGGATTAGATACCCTGGTAGTCCACGCGGTAAACGATGATTGCTAGGTGTAGGTGGGTATGGACCCAT
CGGTGCCGCAGCTAACGCAATAAGCAATCCACCTGGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACG
GGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAAGTCTTGACATCCC
10 AATGACGCACCTGTAAAGAGGTGTTCCCTTCGGGGCATTGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGT
CGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATTCTTAGTAGCCAGCAGGTGAAGCTGGGCACTC
TAAGGAGACTGCCGGGGATAACCCGGAGGAAGGCGGGGATGACGTCAAATCATCATGCCCCTTATGATTTGGGCT
ACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGACAGTGATGTGGAGCAAATCCCAGAAATAACGTCTCA
GTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAATCAGCATGTCGCGGTGA
15 ATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTTGGAAATGCCCGAAGTCTGTGACCTAACC
GAAAGGGAGGAGCAGCCGAAGGCAGGTCTGATAACTGGGGTGAAGTCGTA

HMI_16 Clostridium_clostridioforme 16S rDNA sequence (SEQ ID NO: 15)

20 CTGCTTTGATGAAGTTTTCGGATGGATTTAAAACAGCTTAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGC
CTCACACTGGGGGATAACAGTTAGAAATAGCTGCTAATACCGCATAAGCGCACGGTTCCGCATGGAACAGTGTGA
AAAACTCCGGTGGTGTGAGATGGACCCGCGTCTGATTAGCCAGTTGGCGGGGTAACGGCCCACCAAAGCGACGAT
CAGTAGCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGT
GGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCGGTATGTAAAG
25 CTCTATCAGCAGGGAAGAAAGTGACGGTACCTGAATAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAA
TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCAAGGCAAGTCTGAAGTGAA
AGCCCGGTGCTTAACGCCGGGACTGCTTTGGAAACTGTTTAGCTGGAGTGCCGGAGAGGTAAGCGGAATTCCTAG
TGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACTGACGTTG
AGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATTGCTAGGTGT
30 AGGTGGGTATGGACCCATCGGTGCCGCAGCTAACGCAATAAGCAATCCACCTGGGGAGTACGTTCGCAAGAATGA
AACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTT
ACCAGGTCTTGACATCCCGATGAAAAACCCGTAACGGGGTTCCCTCTTCGGAGCATCGGAGACAGGTGGTGCATG
GTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATTCTTAGTAGCCAGC
AGGTAAGGCTGGGCACTCTAAGGAGACTGCCGGGGATAACCCGGAGGAAGGTGGGGATGACGTCAAATCATCATG
35 CCCCTTATGATCTGGGCTACACACGTGCTACAATGGCGTAACAAAGGGAAGCGAGCCTGCGAGGGTGAGCAAATC
CCAAAAATAACGTCCCAGTTCGGACTGTAGTCTGCAACCCGACTACACGAAGCTGGAATCGCTAGTAATCGCGAA
TCAGAATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCGGAAATGCCCG
AAGTCTGTGACTCAACCGCAAGGAGAGAGCAGCCGAAGGCAGGTCTGATAACTGGGGTGAAGTCGT

HMI_17 Ruminococcus_torques 16S rDNA sequence (SEQ ID NO: 16)

CGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGTGGGGTAACGGCCTACCAAGGCG
ACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAAA
CTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGAC
GCCGCGTGAGCGAAGAAGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAATGAC
GGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGG
GCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGAGGCAAGTCTGAT
GTGAAAACCCGGGGCTCAACCCCGTGACTGCATTGGAAACTGTTTTGCTTGAGTGCCGGA
GAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGT
GGCGAAGGCGGCTTACTGGACGGCAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAAC
AGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCGGGGAGCAA
AGCTCTTCGGTGCCGCCGCAAACGCAATAAGTATTCCACCTGGGGAGTACGTTCGCAAGA
ATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCG
AAGCAACGCGAAGAACCTTACCAAGCCTTGACATCCCATTGACAGAGCATGTAATGTGCT
TTCCCTTCGGGGCAGTGGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGAT
GTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCTTCAGTAGCCAGCGGTTTGGCCGG
GCACTCTGGAGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATC
ATGCCCCTTATGGCTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGC
CTGCGAGGGGGAGCAAATCCCAAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCG
ACTACATGAAGCTGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCC
CGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCGGCAACGCCCGAAGCCAGTGAC
CCAACCGAAAG

HMI_18 Clostridium_celerecrescens 16S rDNA sequence (SEQ ID NO: 17)

AGTCGACGAGGTAATGAGATGAAGTTTTCGGATGGATTCTTATTTCCGAGTGGCGGACGGGTGAGTAACGCGTGG
GTAACCTGCCTCATACAGGGGGATAACGATTGGAAACGATTGCTAATACCGCATAAGCGCACAGTACCACATGGT
ACAGTGTGAAAAACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGTGAGGTAACGGCCCACCAA
GGCAACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGA
GGCAGCAGTGGGGGATATTGCACAATGGAGGAAACTCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCGG
TATGTAAAGCTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAG
CCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGATGCAAGTC
TGAAGTGAAATACCCGGGCTCAACCTGGGAACTGCTTTGGAAACTGTATGGCTAGAGTGCTGGAGAGGTAAGCGG
AATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGGCTTACTGGACAGTAA
CTGACGTTCAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAAT
ACTAGGTGTCGGGGGACAAAGTCTTTCGGTGCCGCCGCAAACGCAATAAGTATTCCACCTGGGGAGTACGTTCGC
AAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGA
AGAACCTTACCAAATCTTGACATCCCTCTGAAAAGCCTTTAATCGAGCTCCTCCTTCGGGACAGAGGTGACAGGT

GGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATTGTCAG
TAGCCAGCAGGTAAAGCTGGGCACTCTGATGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAA
TCATCATGCCCCTTATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAGGCGAAGCTGTGAGGCA
GAGCAAATCTCAAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGT
5 AATCGCGGATCAGAATGCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCGG
AAATGCCCGAAGCCAGTGACCCAAGCGAAAGCAGGGAGCTGTCGAAGGCAGGTCTGATAACTGGGGTGAAGTCGT

HMI_19 Clostridium_celerescens 16S rDNA sequence (SEQ ID NO: 18)

10 TCGACGAGGTATTTTGATTGAAGTTTTCGGATGGATTTCAGATACCGAGTGGCGGACGGGTGAGTAACGCGTGGG
TAACCTGCCTCATACAGGGGGATAACGGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTACCGCATGGTA
CGGTGTGAAAAACTCCGGTGGTATGAGATGGACCCGCGTCTGATTAGCTAGTTGGTGGGGTAACGGCCCACCAAG
GCGACGATCAGTAGCCGACCTGAGAGGGTGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAG
GCAGCAGTGGGGGATATTGCACAATGGAGGAAACTCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCGGT
15 ATGTAAAGCTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGC
CGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGACGCAAGTCT
GAAGTGAAATACCCGGGCTCAACCTGGGAACTGCTTTGGAAACTGTGTTGCTAGAGTGCTGGAGAGGTAAGCGGA
ATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGGCTTACTGGACAGTAAC
TGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATA
20 CTAGGTGTTGGTGAGCAAAGCTCATCGGTGCCGCCGCAAACGCAATAAGTATTCCACCTGGGGAGTACGTTCGCA
AGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAA
GAACCTTACCAAATCTTGACATCCCTCTGAAACGCCCTTAATCGGGCTCCTCCTTCGGGACAGAGGTGACAGGTG
GTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATTGTCAGT
AGCCAGCAGGTAAAGCTGGGCACTCTGATGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAAT
25 CATCATGCCCCTTATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGAGAAGCGAGCCTGCGAGGGGG
AGCAAATCTCAAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTA
ATCGCAGATCAGAATGCTGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCGGA
AATGCCCGAAGCCAGTGACCCAAGCGAAAGCAGGGAGCTGTCGAAGGCAGGTCTGATAACTGGGGTGAAGTCGTA
ACAGGGTAACCG

30

HMI_20 Eubacterium_infirmum 16S rDNA sequence (SEQ ID NO: 19)

GAGCTCATCACAGATGCTTCGGTTGAAGTGATGAGTGGAAAGCGGCGGACGGGTGAGTAA
CGCGTAGGCAACCTGCCCTTTGCAGAGGGATAGCCTCGGGAAACCGGGATTAAAACCTCA
35 TGACACCTCTTAAAGACATCTTTGAGAGGTCAAAGATTTATCGGCAGAGGATGGGCCTGC
GTCTGATTAGCTAGTTGGTGGGGTAACGGCCTACCAAGGCGACGATCAGTAGCCGACCTG
AGAGGGTGATCGGCCACATTGGAACTGAGACACGGTCCAAACTCCTACGGGAGGCAGCAG
TGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCAACGCCGCGTGAAGGAAGAAGG
CCTTTGGGTCGTAAACTTCTGTTCTAAGGGAAGATAATGACGGTACCTTAGGAGCAAGTC

CCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGAATT
ATTGGGCGTAAAGAGTACGTAGGTGGTTACCTAAGCACGAGGTATAAGGCAATGGCTTAA
CCATTGTTCGCCTTGTGAACTGGGCTACTTGAGTGCAGGAGAGGAAAGCGGAATTCCTAG
TGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGA
CTGTAACTGACACTGAGGTACGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAG
TCCACGCCGTAAACGATGAGCACTAGGTGTCGGGGTCGCAAGACTTCGGTGCCGCAGTTA
ACGCAATAAGTGCTCCGCCTGGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGAC
GGGGACCCGCACAAGCAGCGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTAC
CAGGACTTGACATCCCTCTGACAGCCTTTTAATCGAGGTTTTCTACGGACAGAGGAGACA
GGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAG
CGCAACCCTTGTCATTAGTTGCCAGCAGTAAGATGGGCACTCTAGTGAGACTGCCGGGGA
TAACTCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGTTCTGGGCTACA
CACGTGCTACAATGGCCGGTACAGAGAGAAAGCGAGACTGCGAAGTGGAGCGAAACTCAA
AAGCCGGTCCCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAGTTGCTAG
TAATCGCAGATCAGAATGCTGCGGTGAATGCGTTCCCGGGTCTTGTACACACCGCCCGTC
ACACCATGGAAGTTGGGGGCGCCCGAAGTTGGCAGATAAATATGTTACCTAAGGCGAAAT
CAATGACTGGGGTGAAGTCGT

HMI_21 Eubacterium_infirmum 16S rDNA sequence (SEQ ID NO: 20)

TCGGTAAAGGGATATGGCGGAAAGCGGCGGACGGGTGAGTAACGCGTAGGCAACCTGCCC
CTTACAGAGGGATAGCCATTGGAAACGATGATTAAGACCTCATAACGCCTCCCTCCCACA
TGAGGGGGAGGCCAAAGATTCATCGGTAAGGGATGGGCCTGCGTCTGATTAGCTTGTTGG
CGGGGTAACGGCCCACCAAGGCGACGATCAGTAGCCGACCTGAGAGGGTGATCGGCCACA
TTGGAACTGAGACACGGTCCAAACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAAT
GGGCGGAAGCCTGATGCAGCAACGCCGCGTGAGGGATGAAGGCCTTCGGGTCGTAAACCT
CTGTCCTTGGGGAAGAAACAAATGACGGTACCCATGGAGGAAGCCCCGGCTAACTACGTG
CCAGCAGCCGCGGTAATACGTAGGGGGCGAGCGTTATCCGGAATTATTGGGCGTAAAGAG
TGCGTAGGTGGTTACCTAAGCGCAGGGTCTAAGGCAATGGCTCAACCATTGTTCGCCCTG
CGAACTGGGCTACTTGAGTGCAGGAGAGGAAAGCGGAATTCCTAGTGTAGCGGTGAAATG
CGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGACTGTTACTGACACTG
AGGCACGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG
ATGAGCACTAGGTGTCGGGGCCGCAAGGCTTCGGTGCCGCAGTTAACGCATTAAGTGCTC
CGCCTGGGGAGTACGCACGCAAGTGTGAAACTCAAAGGAATTGACGGGGACCCGCACAAG
CAGCGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGACTTGACATCC
CCCTGACAGATCCTTAACCGGATCCTTCTTCGGACAGGGGAGACAGGTGGTGCATGGTTG
TCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGCCAT
TAGTTGCCATCATTCAGTTGGGCACTCTAATGGGACTGCCGGGGACAACTCGGAGGAAGG
TGGGGATGACGTCAAATCATCATGCCCCTTATGTTCTGGGCTACACACGTGCTACAATGG
CCGGTACAGCAGGAAGCGATCCCGCGAGGGGGAGCAAATCCCAAAAACCGGTCCCAGTTC

GGACTGCAGGCTGCAACCCGCCTGCACGAAGCCGGAGTTGCTAGTAATCGTGGATCAGAA
TGCCACGGTGAATGCGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGAAGTTGG
GGGTGCCCGAAGCCGGCAGGGAGATATGCTGTCTAAGGCAAAACCAAT

5 HMI_22 Clostridium_thermocellum 16S rDNA sequence (SEQ ID NO: 21)

GGATGAGGAAATGCTTCGGCATGGAGACATCCGATCTAGTGGCGGACGGGTGAGTAACGC
GTGAGCAACCTGTCCTGCACAGGGGGATAACACTGAGAAATCAGTGCTAATACCGCATGA
GACCACAGTATCACATGGTACAGGGGTCAAAGGAGAAATCCGGTGCAGGGTGGGCTCGCG
10 TCCCATTAGCTAGTTGGTAGGGTAAAGGCCTACCAAGGCGACGATGGGTAGCCGGACTGA
GAGGTTGGCCGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGT
GGGGAATATTGGGCAATGGGCGAAAGCCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGT
CTTTGGATTGTAAACTTTTGTCCTATGGGAAGAAGGAAGTGACGGTACCATGGGAGGAAG
CCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCGAGCGTTGTCCGGAA
15 TTACTGGGCGTAAAGGGCGCGCAGGCGGCCGATCAAGTTAGATGTGAAATACCCGGGCTT
AACCTGGGAACTGCATTTAAAACTGGTTGGCTAGGAGTGCAGGAGAGGGAAGCGGAATTC
CTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTTC
TGGACTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTG
GTAGTCCACGCTGTAAACGATGAATACTAGGTGTAGGGGGTATCGACCCCCCCTGTGCCG
20 GAGCAAACGCAATAAGTATTCCGCCTGGGGAGTACGGCCGCAAGGTTGAAACTCAAAGGA
ATTGACGGGGGCCCGCACAAGCAGCGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAA
CCTTACCAGGTCTTGACATCCCTCGAAGTGCATAGAGATATGTACGTCCTTCGGGACGAG
GAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGC
AACGAGCGCAACCCCTACAGTTAGTTACCAGCGGGTAAAGCCGGGGACTCTAACAGGACT
25 GCCGTGGATAACACGGAGGAAGGTGGGGACGACGTCAAATCATCATGCTCCTTATGACCT
GGGCTACACACGTGCTACAATGGCCGGTACAAAGAGAAGCGAGACCGTAAGGTGGAGCGG
ATCTCAAAAAACCGGTCCCAGTTCGGATTGTGGGCTGCAACCCGCCCACATGAAGTTGGA
GTTGCTAGTAATCGCGAATCAGCATGTCGCGGTGAATGCGTTCCCGGGCCTTGTACACAC
CGCCCGTCACACCATGGGAGTTGGGAGCGCCCGAAGTCGTTGAGGTAACCCGCAAGGGAG
30 CCAGGCGCCGAAGGTGAGACCGATAACTGGGGTGAAGTCGT

HMI_23 Anaerovorax_odorimutans 16S rDNA sequence (SEQ ID NO: 22)

AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAGC
35 GGGAAATCTTGGAACGATACTTCGGTAAAGGGAAGAGATGGATAGCGGCGGACGGGTGAG
TAACGCGTAGGTAACCTGCCTCATGCAGAGGGATAGCCTCGGGAAACTGGGATTAATACC
TCATAATGCGGAGGAGTCACATGGCTCCATCGCCAAAGATTTATCGGCATGAGATGGACC
TGCGTCTGATTAGTTAGTTGGTGAGGTAACGGCTCACCAAGGCAGCGATCAGTAGCCGAC
CTGAGAGGGTAATCGGCCACATTGGAACTGAGACACGGTCCAAACTCCTACGGGAGGCAG

CAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCAACGCCGCGTGAGCGATGA
AGGTCTTCGGATCGTAAAGCTCTGTCCTAGGGGAAGAATATATGACGGTACCCTTGGAGG
AAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCGAGCGTTATCCG
GAATTATTGGGCGTAAAGAGTTCGTAGGTGGTTTTGTAAGCGCGGGGTTTAAGGCAACGG
CTCAACCGTTGTTCGCCTTGCGAACTGCAAGACTTGAGTGCGGGAGAGGAAAGTGGAATT
CCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGACTTT
CTGGACCGTAACTGACACTGAGGAACGAAAGCGTGGGGAGCAAACAGGATTAGATACCCT
GGTAGTCCACGCCGTAAACGATGAGCACTAGGTGTCGGGGCCGCAAGGTTTCGGTGCCGC
AGTTAACGCATTAAGTGCTCCGCCTGGGGAGTACGCACGCAAGTGTGAAACTCAAAGGAA
TTGACGGGGACCCGCACAAGCAGCGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAAC
CTTACCAGGGCTTGACATCCCGATGACCGGCGGGTAACGCCGCCTTCTCTTCGGAGCATC
GGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCG
CAACGAGCGCAACCCTTGTCATTAGTTGCCAGCAGTTCGGCTGGGCACTCTAGTGAGACT
GCCGGGGACAACTCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGTTCT
GGGCTACACACGTGCTACAATGGCCGGTACAGAGAGACGCAAGACTGTGAAGTGGAGCAA
AACTCTAAAACCGGTCCCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGTTGGAG
TTGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATGCGTTCCCGGGTCTTGTACACACC
GCCCGTCACACCATGGAAGTTGGGGGCGCCCGAAGTTGGTCAACAAATCGATTACCTAAG
GCGAAACCAATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGG
ATCACCT

HMI_24 Clostridium_saccharogumia 16S rDNA sequence (SEQ ID NO: 23)

AGCCACCGGCTTCGGGTGTTATCAACTCTCATGGTGTGACGGGCGGTGTGTACAAGGCCC
GAGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCATCTTCATGCAG
GCGAGTTGCAGCCTGCAATCCGAACTGAGAACGGGTTTTTGAGTTTCGCTCCAAGTCGCC
TCTTCGCTTCCCTTTGATCCGTCCATTGTAGCACGTGTGTAGCCCAGGTCATAAGGGGCA
TGATGATTTGACGTCATCCCCGCCTTCCTCCGGCTTGTCACCGGCTGTCTCGTTAGAGTC
CCCATCTTACTGCTGGTAACTAACGACAAGGGTTGCGCTCGTTGCGGGACTTAACCCAAC
ATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCTTGAGTATATCTATCCCTC
TATCTCTAGAGTCTTTACTCTGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAA
TTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCATTCTT
GCGAACGTACTACTCAGGCGGAGTACTTATTGCGTTAACTGCAGCACTGAGGCTTGTCCC
CCCAACACTTAGTACTCATCGTTTACGGCGTGGACTACTAGGGTATCTAATCCTATTTGC
TCCCCACGCTTTCGGGACTGAGCGTCAGTTACAGACCAGATCGTCGCCTTCGCCACTGGT
GTTCCTCCATATATCTACGCATTTCACCGCTACACATGGAATTCCACGATCCTCTTCTGC
ACTCTAGCTATTTGGTTTCCATGGCTTACTGAAGTTAAGCTTCAGCCTTTTACCACAGAC
CTCCATTGCCGCCTGCTCCCTCTTTACGCCCAATAATTCCGGATAACGCTTGCCACCTAC
GTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTCCTCACAAAGTACCGTCACTC
TAATACCATTCCCTGTATTAGTCGTTCTTCCTTTATAACAGAAGTTTACAACCCGAAGGC

CTTCTTCCTTCACGCGGCGTTGCTCGGTCAGGGTTCCCCCATTGCCGAAAATTCCCTAC
TGCTGCCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCCGTTCACCCTCTC
AGGCCGGCTATGCATCGTCGCCTTGGTAGGCCGTTACCCCTCCAACTAGCTAATGCACCA
TAAGCCCATCTGTTCCCTATCCCTTAGGATATTTAACTTAGAGAAAATGCTTCCTCTAAG
5 CCTATGCGGTGTTAGCGCATGTTTCCACGCGTTATCCCCCTGGTACAGCCAGGTTGCTTA
TGTCTTACTCACCCGTTCGCCACTCATCACCGAAGTGATGCGTTCGACTTGCATGTAT

HMI_25 Clostridium_saccharogumia 16S rDNA sequence (SEQ ID NO: 24)

10 GGCATCTACAGGGGGATAACTGATGGAAACGTCAGCTAAGACCGCATAGGTGTAGAGATC
GCATGAACTCTATATGAAAAGTGCTACGGGACTGGTAGATGATGGACTTATGGCGCATTA
GCTTGTTGGTAGGGTAACGGCCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGA
CCGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATT
TTCGGCAATGGGGGAAACCCTGACCGAGCAACGCCGCGTGAAGGAAGAAGTAATTCGTTA
15 TGTAAACTTCTGTCATAGAGGAAGAACGGTGGATATAGGGAATGATATCCAAGTGACGGT
ACTCTATAAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCG
AGCGTTATCCGGAATTATTGGGCGTAAAGAGGGAGCAGGCGGCACTAAGGGTCTGTGGTG
AAAGATCGAAGCTTAACTTCGGTAAGCCATGGAAACCGTAGAGCTAGAGTGTGTGAGAGG
ATCGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGATATATGGAGGAACACCAGTGGCG
20 AAGGCGACGATCTGGCGCATAACTGACGCTCAGTCCCGAAAGCGTGGGGAGCAAATAGGA
TTAGATACCCTAGTAGTCCACGCCGTAAACGATGAGTACTAAGTGTTGGGTGTCAAAGCT
CAGTGCTGCAGTTAACGCAATAAGTACTCCGCCTGAGTAGTACGTTCGCAAGAATGAAAC
TCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAAC
GCGAAGAACCTTACCAGGTCTTGACATCGATCTAAAGGCTCCAGAGATGGAGAGATAGCT
25 ATAGAGAAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAG
TCCCGCAACGAGCGCAACCCCTGTTGCCAGTTGCCAGCATTAAGTTGGGGACTCTGGCGA
GACTGCCGGTGACAAGCCGGAGGAAGGCGGGGATGACGTCAAATCATCATGCCCCTTATG
ACCTGGGCTACACACGTGCTACAATGGACAGAGCAGAGGGAAGCGAAGCCGCGAGGTGGA
GCGAAACCCATAAAACTGTTCTCAGTTCGGACTGCAGTCTGCAACTCGACTGCACGAAGA
30 TGGAATCGCTAGTAATCGCGAATCAGCATGTCGCGGTGAATACGTTCTCGGGCCTTGTAC
ACACCGCCCGTCACACCATGAGAGTCGGTAACACCCGAAGCCGGTGGCCTAACCGCAAGG
AAGGAGCTGTCTAAGGTGGGACTGATGATTGGGGTGAAGTCGTAACAAGGGTAACC

HMI_26 Blautia_luti 16S rDNA sequence (SEQ ID NO: 25)

35

CGGGAATACTTTATTGAAACTTCGGTGGATTTAATTTATTTCTAGTGGCGGACGGGTGAGTAACGCGTGGGTAAC
CTGCCTTATACTGGGGGATAACAGCCAGAAATGACTGCTAATACCGCATAAGCGCACAGAACCGCATGGTTCCGT
GTGAAAAACTCCGGTGGTATAAGATGGACCCGCGTTGGATTAGCTAGTTGGCAGGGCAGCGGCCTACCAAGGCGA
CGATCCATAGCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAG

CAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAAGGAAGAAGTATCTCGGTATGT
AAACTTCTATCAGCAGGGAAGATAATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCG
GTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGCAGCAAGTCTGATG
TGAAAGGCAGGGGCTTAACCCCTGGACTGCATTGGAAACTGCTGTGCTTGAGTGCCGGAGGGGTAAGCGGAATTC
5 CTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGTAACTGAC
GTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAG
GTGTCAGGGAGCACAGCTCTTTGGTGCCGCCGCAAACGCATTAAGTATTCCACCTGGGGAGTACGTTCGCAAGAA
TGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAAC
CTTACCAAATCTTGACATCCCTCTGACCGGGACTTAACCGTCCCTTTCCTTCGGGACAGGGGAGACAGGTGGTGC
10 ATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTAGCC
AGCACGTAATGGTGGGCACTCTGAGGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGGATGACGTCAAATCATC
ATGCCCCTTATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAACCCGCGAGGGTGGGCA
AATCTCAAAAATAACGTCCCAGTTCGGACTGCAGTCTGCAACTCGACTGCACGAAGCTGGAATCGCTAGTAATCG
CGGATCAGAATGCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACG
15 CCCGAAGTCAG

HMI_27 Clostridium_clostridioforme 16S rDNA sequence (SEQ ID NO: 26)

TTGCGGTAGGTCACAGGCTTCGGGCATTTCCAACTCCCATGGTGTGACGGGCGGTGTGTA
20 CAAGACCCGGGAACGTATTCACCGCGACATGCTGATTCGCGATTACTAGCGATTCCAGCT
TCATGTAGTCGAGTTGCAGACTACAATCCGAACTGAGACGTTATTTCTGGGATTTGCTCA
ACATCACTGTCTCGCTTCCCTTTGTTTACGCCATTGTAGCACGTGTGTAGCCCAAATCAT
AAGGGGCATGATGATTTGACGTCATCCCCGCCTTCCTCCGGGTTATCCCCGGCAGTCTCC
CTAGAGTGCCCAGCTCTACCTGCTGGCTACTAAGGATAAGGGTTGCGCTCGTTGCGGGAC
25 TTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCTCCAATGCT
CCGAAGAGAATGCCCCGTTACGGACACGTCATTGGGATGTCAAGACTTGGTAAGGTTCTT
CGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGTCCCCGTCAATTCCTT
TGAGTTTCATTCTTGCGAACGTACTCCCCAGGTGGATTGCTTATTGCGTTAGCTGCGGCA
CCGATGGGTCCATACCCACCTACACCTAGCAATCATCGTTTACCGCGTGGACTACCAGGG
30 TATCTAATCCTGTTTGCTCCCCACGCTTTCGAGCCTCAACGTCAGTTACAGTCCAGTAAG
CCGCCTTCGCCACTGGTGTTCCTCCTAATATCTACGCATTTCACCGCTACACTAGGAATT
CCGCTTACCTCTCCTGCACTCCAGCCTGGCAGTTCCAAATGCAGTCCCAGGGTTGAGCCC
TGGGTTTTCACATCTGGCTTGTCATGCCGTCTACGCTCCCTTTACACCCAGTAAATCCGG
ATAACGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGGGCTTCTT
35 AGTCAGGTACCGTCATTTTCTTCCCTGCTGATAGAGCTTTACATACCGAAATACTTCTTC
ACTCACGCGGCGTCGCTGCATCAGGGTTTCCCCCATTGTGCAATATTCCCCACTGCTGCC
TCCCGTAGGAGTTTGGGCCGTGTCTCAGTCCCAATGTGGCCGTTCACCCTCTCAGGCCGG
CTACTGATCGTCGCTTTGGTAGGCCGTTACCCTGCCAACTGGCTAATCAGACGCGGGACC
ATCCTGTACCACCGGAGTTTTTCACACTGCCTCATGTGAAGCTGTGCGCTTATGCGGTAT
40 TAGCACCTATTTCTAAGTGTTATCCCCCGGTACAGGGCAGGTTTCCCACGCGTTACTCAC

CCGTCCGCCACTAAGTTACGCCGATTCCATCCGAAAACTTCCTCTGCATAACTCCGTCGA
CTG

HMI_28 Blautia_producta 16S rDNA sequence (SEQ ID NO: 27)

5

CCGGTGGTCG CATCGGCGCT CCTCCTGTAG GTTGGGTCAC TGACTTCGGG
CGTTACTGAC TCCCATGGTG TGACGGGCGG TGTGTACAAG ACCCGGGAAC
GTATTCACCG CGACATTCTG ATTCGCGATT ACTAGCGATT CCAGCTTCGT
GCAGTCGAGT TGCAGACTGC AGTCCGAACT GGGACGTTAT TTTTGGGATT
10 TGCTCAACAT CGCTGTCTCG CTTCCCTTTG TTTACGCCAT TGTAGCACGT
GTGTAGCCCA AATCATAAGG GGCATGATGA TTTGACGTCG TCCCCGCCTT
CCTCCGGGTT ATCCCCGGCA GTCTCCCTAG AGTGCCCAGC TTCACCTGCT
GGCTACTAAG GATAGGGGTT GCGCTCGTTG CGGGACTTAA CCCAACATCT
CACGACACGA GCTGACGACA ACCATGCACC ACCTGTCTCC TCTGCCCCGA
15 AGGGAAGGCC CCGTTACGGG CCGGTCAGAG GGATGTCAAG ACTTGGTAAG
GTTCTTCGCG TTGCTTCGAA TTAAACCACA TGCTCCACCG CTTGTGCGGG
TCCCCGTCAA TTCCTTTGAG TTTCATTCTT GCGAACGTAC TCCCCAGGTG
GAATACTTAT TGCGTTTGCT GCGGCACCGA ATGGGCTTTG CCACCCGACA
CCTAGTATTC ATCGTTTACG GCGTGGACTA CCAGGGTATC TAATCCTGTT
20 TGCTCCCCAC GCTTTCGAGC CTCAACGTCA GTTACCGTCC AGAAAGCCGC
CTTCGCCACT GGTGTTCCTC CTAATATCTA CGCATTTCAC CGCTACACTA
GGAATTCCGC TTACCTCTCC GGCACTCTAG AAAAACAGTT TCCAATGCAG
TCCTGGGGTT AAGCCCCAGC CTTTCACATC AGACTTGCTC TTCCGTCTAC
GCTCCCTTTA CACCCAGTAA ATCCGGATAA CGCTTGCCCC CTACGTATTA
25 CCGCGGCTGA TGGCACGTAG TTAGCCGGGG CTTCTTAGTC AGGTACCGTC
ATTTTCTTCC CTGCTGATAG AAGTTTACAT ACCGAGATAC TTCTTCCTTC
ACGCGGCGTC GCTGCATCAG GGTTTCCCCC ATTGTGCAAT ATTCCCCACT
GCTGCCTCCC GTAGGAGTCT GGGCCGTGTC TCAGTCCCAA TGTGGCCGTT
CACCCTCTCA GGCCGGCTAC TGATCGTCGC CTTGGTGGGC CGTTACCCCT
30 CCAACTAGCT AATCAGACGC GGGTCCATCT CATACCACCG GAGTTTTTCA
CACCAGACCA TGCGGTCCTG TGCGCTTATG CGGTATTAGC AGCCATTTCT
AACTGTTATC CCCCTGTATG AGGCAGGTTA CCCACGCGTT ACTCAGCCCG
TCCGCCGCTC AGTCAAATAA GTTTCAATCC GAAGAGATCC ACTTAAGTGC
TTCGCTCGAC TTGCATGTGT TAAGCACGCC GCCAGCGTTC ATCCT

35

HMI_29 Blautia glucerasea 16S rDNA sequence (SEQ ID NO: 28)

GCCTTCGGCAGCTCCGTCCTTTCGGTTCGGTCACTGACTTCGGGCGTTACTGACTCCCAT
GGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCATTCTGATCCGC

GATTACTAGCGATTCCAGCTTCGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGGGACG
TTATTTTTGGGATTTGCTTAAGCTCACACTCTCGCTTCCCTTTGTTTACGCCATTGTAGC
ACGTGTGTAGCCCAAATCATAAGGGGCATGATGATTTGACGTCATCCCCGCCTTCCTCCA
GGTTATCCCTGGCAGTCTCCTCAGAGTGCCCGGCCAAACCGCTGGCTACTAAGGATAGGG
5 GTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGC
ACCACCTGTCTCCGATGCTCCGAAGAAAAGGCGACGTTACTCGCCGGTCATAGGGATGTC
AAGACTTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGC
GGGTCCCCGTCAATTCCTTTGAGTTTCATTCTTGCGAACGTACTCCCCAGGTGGAATACT
TACTGCGTTTGCTGCGGCACCGAATGGCTCTGCCACCCGACACCTAGTATTCATCGTTTA
10 CGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGAGCCTCAACGT
CAGTTACCGTCCAGTAAGCCGCCTTCGCCACTGGTGTTCCTCCTAATATCTACGCATTTC
ACCGCTACACTAGGAATTCCGCTTACCTCTCCGGTACTCAAGATCAACAGTTTCCAATGC
AGTCCGGGGGTTGAGCCCCCGCCTTTCACATCAGACTTGCTGCTCCGTCTACGCTCCCTT
TACACCCAGTAAATCCGGATAACGCTTGCCCCCTACGTATTACCGCGGCTGCTGGCACGT
15 AGTTAGCCGGGGCTTCTTAGTCAGGTACCGTCATTTTCTTCCCTGCTGATAGAAGTTTAC
ATACCGAGATACTTCTTCCTTCACGCGGCGTCGCTGCATCAGGGTTTCCCCCATTGTGCA
ATATTCCCCACTGCTGCCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAATGTGGCCG
TCCACCCTCTCAGGCCGGCTATGGATCGTCGCTTTGGTAGGCCGTTACCCTGCCAACTGG
CTAATCCAACGCGGGTCCATCTCACACCACCGGAGTTTTTCACACTGGATCATGCAATCC
20 CGTGCGCTTATGCGGTATTAGCAGTCATTTCTGACTGTTATCCCCCAGTGTGAGGCAGGT
TACCCACGCGTTACTCACCCGTCCGCCACTAGGATTATAACGACTTCAACCGAAGTCTCT
GTCAAAATAATCCCCGTTCGACTTGCATGTGT

HMI_30 Clostridium_straminisolvens 16S rDNA sequence (SEQ ID NO: 29)

25

AGCGGCGGACGGGTGAGTAACGCGTGAGTAACCTGCCTTTAGGAGGGGGACAACATTCCGAAAGGGATGCTAATA
CCGCATAAAATTATTGTATCGCATGGTATAATAATCAAAGATTTATCGCCTAAAGATGGACTCGCGTCCGATTAG
CTAGTTGGTGGGGTAAAAGCCTACCAAGGCGACGATCGGTAGCCGAACTGAGAGGTTGATCGGCCACATTGGGAC
TGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGGATATTGCGCAATGGGGGAAACCCTGACGCAGCAA
30 CGCCGCGTGAAGGAAGAAGGCCTTCGGGTTGTAAACTTCTTTAAGTGTGGAAGATAATGACGGTACACACAGAAT
AAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGTGTA
AAGGGCGTGTAGGCGGGTAGACAAGTCAGATGTGAAATACCGGGGCTCAACTCCGGGGCTGCATTTGAAACTGTA
TATCTTGAGTGTCGGAGAGGAAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAG
TGGCGAAGGCGGCTTTCTGGACGATAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCC
35 TGGTAGTCCACGCCGTAAACGATGGATACTAGGTGTAGGAGGTATCGACCCCTTCTGTGCCGCAGTTAACACAAT
AAGTATCCCACCTGGGGAGTACGGTCGCAAGATTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCAGTGGA
GTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGACTTGACATCCCACGCATAGCCTAGAGATAGGT
GAAGTCCTACGGGACGTGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCC
CGCAACGAGCGCAACCCTTACTGTCAGTTACCATCATTAAGTTGGGGACTCTGGCAGGACTGCCGGTGACAAATC
40 GGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCCTTATGTCCTGGGCTACACACGTACTACAATGGCTGTTA

ACAAAGTGAAGCAAAGCAGTGATGTGGAGCAAAACACAAAAAGCAGTCTCAGTTCAGATTGTAGGCTGAAACTCG
CCTATATGAAGTCGGAATTGCTAGTAATCGCAGATCAGCATGCTGCGGTGAATACGTTCCCGGGCCTTGTACACA
CCGCCCGTCACACCATGAGAGTCGATAACACCCGAAGCCTGT

5 HMI_31 Butyricoccus_pullicaecorum 16S rDNA sequence (SEQ ID NO: 30)

AGTGGCGGACGGGTGAGTAACGCGTGAGCAATCTGCCTTTAAGAGGGGGATAACAGTCGGAAACGGCTGCTAATA
CCGCATAAAGCATCGAAACCGCATGATTTTGATGCCAAAGGAGCAATCCGCTTTTAGATGAGCTCGCGTCTGATT
AGCTGGTTGGCGGGGTAACGGCCCACCAAGGCGACGATCAGTAGCCGGACTGAGAGGTTGAACGGCCACATTGGG
10 ACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCGCAATGGGGGAAACCCTGACGCAGC
AACGCCGCGTGATTGAAGAAGGCCTTCGGGTTGTAAAGATCTTTAATCAGGGACGAAACAAATGACGGTACCTGA
AGAATAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCAAGCGTTATCCGGATTTACTGG
GTGTAAAGGGCGCGCAGGCGGGCCGGTAAGTTGGAAGTGAAATCTATGGGCTTAACCCATAAACTGCTTTTCAAA
CTGCTGGTCTTGAGTGATGGAGAGGCAGGCGGAATTCCGTGTGTAGCGGTGAAATGCGTAGATATACGGAGGAAC
15 ACCAGTGGCGAAGGCGGCCTGCTGGACATTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGA
TACCCTGGTAGTCCACGCCGTAAACGATGGATACTAGGTGTGGGAGGTATTGACCCCTTCCGTGCCGCAGTTAAC
ACAATAAGTATCCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCA
GTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCCGATGACCGCCTCAGAG
ATGAGCCTTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTT
20 AAGTCCCGCAACGAGCGCAACCCTTACGGTTAGTTGATACGCAAGATCACTCTAGCCGGACTGCCGTTGACAAAA
CGGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTACTACAATGGCAGTC
ATACAGAGGGAAGCAAAACCGCGAGGTGGAGCAAATCCCTAAAAGCTGTCCCAGTTCAGATTGCAGGCTGCAACC
CGCCTGCATGAAGTCGGAATTGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACA
CACCGCCCGTCACACCATGAGAGCCGTCAATACCCGAAGTCCGT

25

HMI_32 Clostridium_maritium 16S rDNA sequence (SEQ ID NO: 31)

CTACACTGCA GTCGAGCGAT TTCTTCGGTA AGAGCGGCGG ACGGGTGAGT
AACGCGTGGG TAACCTGCCC TATACACACG GATAACATAC CGAAAGGTAT
30 GCTAATACGA GATAACATAA GAGATTCGCA TGGATTTCTT ATCAAAGCTT
TTGCGGTATA GGATGGACCC GCGTCTGATT AGCTAGTTGG TAAGGTAACG
GCTTACCAAG GCGACGATCA GTAGCCGACC TGAGAGGGTG ATCGGCCACA
TTGGAACTGA GACACGGTCC AAACTCCTAC GGGAGGCAGC AGTGGGGAAT
ATTGCACAAT GGGCGAAAGC CTGATGCAGC AACGCCGCGT GAGCGATGAA
35 GGCCTTCGGG TCGTAAAGCT CTGTCCTCAA GGAAGATAAT GACGGTACTT
GAGGAGGAAG CCCCGGCTAA CTACGTGCCA GCAGCCGCGG TAATACGTAG
GGGGCTAGCG TTATCCGGAA TTACTGGGCG TAAAGGGTGC GTAGGTGGTT
TCTTAAGTCA GAGGTGAAAG GCTACGGCTC AACCGTAGTA AGCCTTTGAA
ACTGAGAAAC TTGAGTGCAG GAGAGGAGAG TAGAATTCCT AGTGTAGCGG

TGAAATGCGT AGATATTAGG AGGAATACCA GTTGCGAAGG CGGCTCTCTG
GACTGTAACT GACACTGAGG CACGAAAGCG TGGGGAGCAA ACAGGATTAG
ATACCCTGGT AGTCCACGCC GTAAACGATG AGTACTAGCT GTCGGAGGTT
ACCCCCTTCG GTGGCGCAGC TAACGCATTA AGTACTCCGC C

5

HMI_33_Eubacterium fissicatens 16S rDNA sequence (SEQ ID NO: 32)

AGTGGCGGAC GGGTGAGTAA CGCGTGGGTA ACCTGCCTTG TACAGGGGGA
TAACAGTTAG AAATGACTGC TAATACCGCA TAAGCGCACA GTATCGCATG
10 GTACAGTGTG AAAAACTCCG GTGGTACAAG ATGGACCCGC GTCTGATTAG
CTAGTTGGTA AGGTAACGGC TTACCAAGGC AACGATCAGT AGCCGACTTG
AGAGAGTGAT CGGCCACATT GGGACTGAGA CACGGCCCAA ACTCCTACGG
GAGGCAGCAG TGGGGAATAT TGCACAATGG GGGAAACCCT GATGCAGCGA
CGCCGCGTGA GTGAAGAAGT ATTTCGGTAT GTAAAACTCT ATCAGCAAGG
15 AAGATAATGA CGGTACTTGA CTAAGAAGCC CCGGCTAACT ACGTGCCAGC
AGCCGCGGTA ATACGTAGGG GGCAAGCGTT ATCCGGATTT ACTGGGTGTA
AAGGGAGCGT AGACGGTATG GTAAGTCAGA TGTGAAAGCC CGGGGCTTAA
CCCCGGAACT GCATTTGAAA CTATCAAACT AGAGTGTCGG AGAGGTAAGT
GGAATTCCTA GTGTAGCGGT GAAATGCGTA GATATTAGGA GGAACACCAG
20 TGGCGAAGGC GGCTTACTGG ACGATAACTG ACGTTGAGGC TCGAAAGCGT
GGGGAGCAAA CAGGATTAGA TACCCTGGTA GTCCACGCCG TAAACGATGA
ATACTAGGTG TCAGGGAACA ATAGTTCTTT GGTGCCGCAG CAAACGCATT
AAGTATTCCA CCTGGGGAGT ACGTTCGCAA GAATGAAACT CAAAGGAATT
GACGGGGACC CGCACAAGCG GTGGAGCATG TGGTTTAATT CGAAGCAACG
25 CGAAGAACCT TACCTGGTCT TGACATCCCA ATGACGCCTC TTTAATCGGA
GGTTTCCTTC GGGACATTGG AGACAGGTGG TGCATGGTTG TCGTCAGCTC
GTGTCGTGAG ATGTTGGGTT AAGTCCCGCA ACGAGCGCAA CCCTTATCTT
TAGTAGCCAG CAGTTCGGCT GGGCACTCTA GAGAGACTGC CAGGGATAAC
CTGGAGGAAG GTGGGGATGA CGTCAAATCA TCATGCCCCT TATGACCAGG
30 GCTACACACG TGCTACAATG GCGTAAACAA AGGGAAGCAA AACTGTGAGG
TTGAGCAAAT CCCAAAAATA ACGTCTCAGT TCGGATTGTA GTCTGCAACT
CGACTACATG AAGCTGGAAT CGCTAGTAAT CGCAGATCAG AATGCTGCGG
TGAATACGTT CCCGGGTCTT GTACACACCG CCCGTCACAC CATGGGAGTC
GGATATGCCC GAAGTCAGTG ACCCAACCGT AAGGAGGGAG CTGCCGAAGG
35 TGGAGCCGAT AACTGGGGTG AAGTCGT

HMI_34_Clostridium_saccharolyticum 16S rDNA sequence (SEQ ID NO: 33)

AGCGGCGGAC GGGTGAGTAA CGCGTGGGTA ACCTGCCTCA TACAGGGGGA

TAACAGTTAG AAATGACTGC TAATACCGCA TAAGCGCACA GTGCTGCATG
GCACAGTGTG AAAAACTCCG GTGGTATGAG ATGGACCCGC GTTGGATTAG
GCAGTTGGCG GGGTAACGGC CCACCAAACC GACGATCCAT AGCCGGCCTG
AGAGGGTGAA CGGCCACATT GGGACTGAGA CACGGCCCAA ACTCCTACGG
5 GAGGCAGCAG TGGGGAATAT TGCACAATGG GGGAAACCCT GATGCAGCGA
CGCCGCGTGA GTGAAGAAGT AATTCGTTAT GTAAAGCTCT ATCAGCAGGG
AAGAAAATGA CGGTACCTGA CTAAGAAGCC CCGGCTAACT ACGTGCCAGC
AGCCGCGGTA ATACGTAGGG GGCAAGCGTT ATCCGGATTT ACTGGGTGTA
AAGGGAGCGT AGACGGCCGT GCAAGTCTGA TGTGAAAGGC TGGGGCTCAA
10 CCCCGGGACT GCATTGGAAA CTGTATGGCT GGAGTGCCGG AGAGGTAAGC
GGAATTCCTA GTGTAGCGGT GAAATGCGTA GATATTAGGA GGAACACCAG
TGGCGAAGGC GGCTTACTGG ACGGTAACTG ACGTTGAGGC TCGAAAGCGT
GGGGAGCAAA CAGGATTAGA TACCCTGGTA GTCCACGCCG TAAACGATGA
TTACTAGGTG TTGGGGGACA TGGTCCTTCG GTGCCGCCGC AAACGCAGTA
15 AGTAATCCAC CTGGGGAGTA CGTTCGCAAG AATGAAACTC AAAGGAATTG
ACGGGGACCC GCACAAGCGG TGGAGCATGT GGTTTAATTC GAAGCAACGC
GAAGAACCTT ACCAAGTCTT GACATCGAGA GGACAGAGTA TGTAATGTAC
TTTCCCTTCG GGGCCTCGAA GACAGGTGGT GCATGGTTGT CGTCAGCTCG
TGTCGTGAGA TGTTGGGTTA AGTCCCGCAA CGAGCGCAAC CCCTATCTTC
20 AGTAGCCAGC AATTCGGATG GGCACTCTGG AGAGACTGCC GGGGATAACC
CGGAGGAAGG CGGGGATGAC GTCAAATCAT CATGCCCCTT ATGACTTGGG
CTACACACGT GCTACAATGG CGTAAACAAA GGGAAGCGAG GGAGTGATCC
GGAGCAAATC CCAAAAATAA CGTCTCAGTT CGGATTGTAG TCTGCAACTC
GACTACATGA AGCTGGAATC GCTAGTAATC GCGAATCAGC ATGTCGCGGT
25 GAATACGTTC CCGGGTCTTG TACACACCGC CCGTCACACC ATGGGAGTCG
ATAACGCCCG AAGTCAGTGA CCCAACCGAA AGGAGGGAGC TGCCGAAGGC
GGGATTGGTA ACTGGGGTGA AGTCGT

HMI_35 *Blautia_luti* 16S rDNA sequence (SEQ ID NO: 34)

30

AGTGGCGGAC GGGTGAGTAA CGCGTGGGTA ACCTGCCTTA TACTGGGGGA
TAACAGCCAG AAATGGCTGC TAATACCGCA TAAGCGCACG GGGCCGCATG
GTCCTGTGTG AAAAACTCCG GTGGTATAAG ATGGACCCGC GTTGGATTAG
CTAGTTGGCA GGGCAGCGGC CTACCAAGGC GACGATCCAT AGCCGGCCTG
35 AGAGGGTGAA CGGCCACATT GGGACTGAGA CACGGCCCAG ACTCCTACGG
GAGGCAGCAG TGGGGAATAT TGCACAATGG GGGAAACCCT GATGCAGCGA
CGCCGCGTGA AGGAAGAAGT ATCTCGGTAT GTAAACTTCT ATCAGCAGGG
AAGATAATGA CGGTACCTGA CTAAGAAGCC CCGGCTAACT ACGTGCCAGC
AGCCGCGGTA ATACGTAGGG GGCGAGCGTT ATCCGGATTT ACTGGGTGTA
40 AAGGGAGCGT AGACGGCGTA TCAAGTCTGA TGTGAAAGGC AGGGGCTTAA

CCCCTGGACT GCATTGGAAA CTGGTATGCT TGAGTGCCGG AGGGGTAAGC
GGAATTCCTA GTGTAGCGGT GAAATGCGTA GATATTAGGA GGAACACCAG
TGGCGAAGGC GGCTTACTGG ACGGTAACTG ACGTTGAGGC TCGAAAGCGT
GGGGAGCAAA CAGGATTAGA TACCCTGGTA GTCCACGCCG TAAACGATGA
5 ATACTAGGTG TCTGGGAGCA CAGCTCTTAG GTGCCGCCGC AAACGCATTA
AGTATTCCAC CTGGGGAGTA CGTTCGCAAG AATGAAACTC AAAGGAATTG
ACGGGGACCC GCACAAGCGG TGGAGCATGT GGTTTAATTC GAAGCAACGC
GAAGAACCTT ACCAAATCTT GACATCCCTC TGACAGAGTA TGTAATGTAC
TTTTCCTTCG GGACAGGGGA GACAGGTGGT GCATGGTTGT CGTCAGCTCG
10 TGTCGTGAGA TGTTGGGTTA AGTCCCGCAA CGAGCGCAAC CCCTATCCTT
AGTAGCCAGC AAGTAATGTT GGGCACTCTG AGGAGACTGC CAGGGATAAC
CTGGAGGAAG GCGGGGATGA CGTCAAATCA TCATGCCCCT TATGATTTGG
GCTACACACG TGCTACAATG GCGTAAACAA AGGGAAGCGA ACCTGTGAGG
GTGGGCAAAT CTCAAAAATA ACGTCCCAGT TCGGACTGCA GTCTGCAACT
15 CGACTGCACG AAGCTGGAAT CGCTAGTAAT CGCGGATCAG AATGCCGCGG
TGAATACGTT CCCGGGTCTT GTACACACCG CCCGTCACAC CATGGGAGTC
AGTAACGCCC GAAGTCAGTG ACCTAACCGT AAGGAAGGAG CTGCCGAAGG
CGGGACGGAT GACTGGGGTG AAGTCGT

20 HMI_36 Clostridium_methylpentosum 16S rDNA sequence (SEQ ID NO: 35)

GGTTACCTTGTTACGACTTCACCCCAATCATCAACCCCACCTTCGACGACGTCCCCCTTG
CGGTTAGACTATCGGCTTCGGGTGTTGCCAACTCTCATGGTGTGACGGGCGGTGTGTACA
AGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCAATTCCGGCTTC
25 ATGCAGGCGGGTTGCAGCCTGCAATCCGAACTGAGACTATTTTTAGGGGTTTGCTCCATG
TCACCATCTTGCTTCCCTCTGTTAATAGCCATTGTAGTACGTGTGTAGCCCAGGTCATAA
GGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGTTTTGTCAACGGCAGTCCGTCT
AGAGTGCTCTTGCGTAGCAACTAAACGTAAGGGTTGCGCTCGTTGCGGGACTTAACCCAA
CATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCTCGGTGCCCCGAAGGGCT
30 TCACCTATCTCTAGGCTATGCACCGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTT
CGAATTAAACCACATACTCCACTGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAA
CCTTGCGGTCGTACTCCCCAGGTGGATTACTTATTGTGTTAACTCCGGCACGGAAGGGGT
CAGTCCCCCCACACCTAGTAATCATCGTTTACAGCGTGGACTACCAGGGTATCTAATCCT
GTTTGCTACCCACGCTTTCGAGCCTCAGCGTCAGTTAAAGCCCAGCAGGCCGCCTTCGCC
35 ACTGGTGTTCCTCCTAATATTTACGCATTTCACCGCTACACTAGGAATTCCGCCTGCCTC
TACTTCACTCAAGAACTGCAGTTTTGAACGCGGCTATGGGTTGAGCCCATAGATTTAACA
TTCAACTTGCAATCCCGCCTACGCTCCCTTTACACCCAGTAATTCCGGACAACGCTCGCT
ACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTAGCTTCCTCCTTGGTTACCG
TCATTATCTTCACCAAGGACAGAGGTTTACAATCCGAAAACCTTCTTCCCTCACTCGGCG
40 TCGCTGCATCAGGGTTTCCCCCATTGTGCAATATTCCCCACTGCTGCCTCCCGTAGGAGT

CTGGGCCGTGTCTCAGTCCCAATGTGGCCGTTCAACCTCTCAGTCCGGCTACCAATCGTC
GCCTTGGTGGGCCGTTACCTCACCAACTAGCTAATTGGACGCGAGTCCATCTTTCAGCGG
ATTGCTCCTTTGATATCAGCTCCATGCGAAACCAATATGTTATGCGGTATTAGCGTCCGT
TTCCAGACGTTATCCCCCTCTGAAAGGCAGGTTACTCACGCGTTACTCACCCGTCCGCCA
5 CTAAGTTGAATCAAATTCCTTCCGAAGAATTCATTCAAAGCAACTTCGTCGACTTGCATG
TGTAAGGCGCGCCGACAGCGTTCGT

HMI_37 Clostridium_xylanolyticum 16S rDNA sequence (SEQ ID NO: 36)

10 AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAAC
GGAATTTACATGAAGCCTAGCGATTGTAAATTTAGTGGCGGACGGGTGAGTAACGCGTGG
GTAACCTGCCTTGTACTGGGGGACAACAGTTGGAAACGACTGCTAATACCGCATAAGCGC
ACAGCTTCGCATGAAGCAGTGTGAAAAACTCCGGTGGTACAAGATGGACCCGCGTCTGAT
TAGCTGGTTGGTGAGGTAACGGCCCACCAAGGCGACGATCAGTAGCCGGCCTGAGAGGGT
15 GAACGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTGGGGAA
TATTGCACAATGGGGGAAACCCTGATGCAGCAACGCCGCGTGAGTGAAGAAGTATTTCGG
TATGTAAAGCTCTATCAGCAGGAAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGCTA
ACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGT
GTAAAGGGAGCGTAGACGGTTTTGCAAGTCTGAAGTGAAAGCCCGGGGCTTAACCCCGGG
20 ACTGCTTTGGAAACTGTAGGACTAGAGTGCAGGAGAGGTAAGTGGAATTCCTAGTGTAGC
GGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACTGTAA
CTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACG
CCGTAAACGATGATTACTAGGTGTTGGTGGGTACGACCCATCGGTGCCGCAGCAAACGCA
ATAAGTAATCCACCTGGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGA
25 CCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCTGGT
CTTGACATCCCTATGAATAACGGGCAATGCCGTTAGTACTTCGGTACATAGGAGACAGGT
GGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGC
AACCCTTATCTTTAGTAGCCAGCAGTAAGATGGGCACTCTAGAGAGACTGCCGGGGATAA
CCCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCAGGGCTACACAC
30 GTGCTACAATGGCGTAAACAAAGAGAAGCGAAGTCGTGAGGCAGAGCGAATCTCAAAAAT
AACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAA
TCGCAGATCAGAATGCTGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACA
CCATGGGAGTCGGAAATGCCCGAAGTCGGTGACCTAACCGAA

35 HMI_38 Oscillibacter_valericigenes 16S rDNA sequence (SEQ ID NO: 37)

CTCAGGACGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGGAGCACCCTTGACT
GAGGTTTCGGCCAAATGATAGGAATGCTTAGTGGCGGACTGGTGAGTAACGCGTGAGGAA
CCTACCTTCCAGAGGGGGACAACAGTTGGAAACGACTGCTAATACCGCATGACGCATGAC

CGGGGCATCCCGGGCATGTCAAAGATTTTATCGCTGGAAGATGGCCTCGCGTCTGATTAG
CTAGATGGTGGGGTAACGGCCCACCATGGCGACGATCAGTAGCCGGACTGAGAGGTTGAC
CGGCCACATTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATAT
TGGGCAATGGACGCAAGTCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGCTTTCGGGTT
5 GTAAACTTCTTTTGTCAGGGAAGAGTAGAAGACGGTACCTGACGAATAAGCCACGGCTAA
CTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGTG
TAAAGGGCGTGCAGCCGGGCCGGCAAGTCAGATGTGAAATCTGGAGGCTTAACCTCCAAA
CTGCATTTGAAACTGTAGGTCTTGAGTACCGGAGAGGTTATCGGAATTCCTTGTGTAGCG
GTGAAATGCGTAGATATAAGGAAGAACACCAGTGGCGAAGGCGGATAACTGGACGGCAAC
10 TGACGGTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGC
TGTAAACGATGGATACTAGGTGTGCGGGGACTGACCCCCTGCGTGCCGCAGTTAACACAA
TAAGTATCCCACCTGGGGAGTACGATCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGC
CCGCACAAGCGGTGGATTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGGC
TTGACATCCTACTAACGAAGTAGAGATACATCAGGTGCCCTTCGGGGAAAGTAGAGACAG
15 GTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGC
GCAACCCCTATTGTTAGTTGCTACGCAAGAGCACTCTAGCGAGACTGCCGTTGACAAAAC
GGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCCTTATGTCCTGGGCTACACACGTA
ATACAATGGCGGTCAACAGAGGGAGGCAAAGCCGCGAGGCAGAGCAAACCCCCAAAAGCC
GTCCCAGTTCGGATCGCAGGCTGCAACCCGCCTGCGTGAAGTCGGAATCGCTAGTAATCG
20 CGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCA
TGAGAGTCGGGAACACCCGAAGTCCGTAGCCTAACCGCAAGGAGGGCGCGGCCGAAGGTG
GGTTCGATAATTGGGGTGAAGTCGTAACAAGGTAACCG

HMI_39 *Ruminococcus_obeum* 16S rDNA sequence (SEQ ID NO: 38)

25

AGTCGAACGGGAACCTTTTATTGAAGCTTCGGCAGATTTAGCTGGTTTCTAGTGGCGGAC
GGGTGAGTAACGCGTGGGTAACCTGCCCTATACAGGGGGATAACAACCAGAAATGGTTGC
TAATACCGCATAAGCGCACAGGACCGCATGGTCCGGTGTGAAAAACTCCGGTGGTATAGG
ATGGACCCGCGTTGGATTAGCCAGTTGGCAGGGTAACGGCCTACCAAAGCGACGATCCAT
30 AGCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGG
GAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGA
AGGAAGAAGTATCTCGGTATGTAAACTTCTATCAGCAGGGAAGATAGTGACGGTACCTGA
CTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTT
ATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGATTAGCAAGTCTGATGTGAAAGGC
35 AGGGGCTCAACCCCTGGACTGCATTGGAAACTGCCAGTCTTGAGTGCCGGAGAGGTAAGC
GGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGC
GGCTTACTGGACGGCAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGA
TACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTTGGGGAGCAAAGCTCTTCG
GTGCCGCCGCAAACGCATTAAGTATTCCACCTGGGGAGTACGTTCGCAAGAATGAAACTC
40 AAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGC

GAAGAACCTTACCAAGTCTTGACATCCCTCTGACGGACTCTTAACCGAGTCTTTCCTTCG
GGACAGAGGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTA
AGTCCCGCAACGAGCGCAACCCCTATCCCCAGTAGCCAGCATTTCGGATGGGCACTCTGA
GGAGACTGCCAGGGATAACCTGGAGGAAGGCGGGGATGACGTCAAATCATCATGCCCCTT
5 ATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAGCCTGCGAGGG
TAAGCAAATCCCAAAAATAACGTCCCAGTTCGGACTGCAGTCTGCAACTCGACTGCACGA
AGCTGGAATCGCTAGTAATCGCGGATCAGAATGCCGCGGTGAATACGTTCCCGGGTCTTG
TACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGTGACCTAACCGCA
AGGGAGGAGCTGCCGAAGGCGGGACCGATGACTGGGGTGAAGTCGTAACAAGGTAACCGT
10 GACTACACGAAGCTGGAATCGCTAGTAATCGCGGATCAGAATGCCGCGGTGAATACGTTCC
CGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGTGACC
TAACCGCAAGGAAGGAGCTGCCGAAGGCGGGACCGATGACTGGGGTGAAGTCGTAACA

HMI_40 *Megasphaera elsdenii* 16S rDNA sequence (SEQ ID NO: 39)

15

ACGCGTAAGCAACCTGCCCTCCGGATGGGGACAACAGCTGGAAACGGCTGCTAATACCGA
ATACGTTTCCATTGCCGCATGGCAGTGGGAAGAAAGGTGGCCTCTGAATATGCTACCGCC
GGGGGAGGGGCTTGCGTCTGATTAGCTAGTTGGAGGGGTAACGGCCCACCAAGGCGACGA
TCAGTAGCCGGTCTGAGAGGATGAACGGCCACATTGGAACTGAGACACGGTCCAGACTCC
20 TACGGGAGGCAGCAGTGGGGAATCTTCCGCAATGGGCGAAAGCCTGACGGAGCAACGCCG
CGTGAGCGAAGACGGCCTTCGGGTTGTAAAGCTCTGTTATACGGGACGAACGGCTAGTGT
GCCAATACCACATTAGAATGACGGTACCGTAAGAGAAAGCCACGGCTAACTACGTGCCAG
CAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGGGCGCG
CAGGCGGTTTCATAAGTCTGTCTTAAAAGTGCGGGGCTTAACCCCGTGAGGGGACGGAAA
25 CTGTGAGACTGGAGTGTCGGAGAGGAAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTA
GATATTAGGAGGAACACCAGTGGCGAAAGCGGCTTTCTGGACGACAACTGACGCTGAGGC
GCGAAAGCCAGGGGAGCGAACGGGATTAGATACCCCGGTAGTCCTGGCCGTAAACGATGG
ATACTAGGTGTAGGGGGTATCGACCCCTCCTGTGCCGGAGTTAACGCAATAAGTATCCCG
CCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCG
30 GTGGAGTATGTGGTTTAATTCGACGCAACGCGAAGAACCTTACCAAGCCTTGACATTGAG
TGCTATCCTCAGAGATGAGGAGTTCTTCTTCGGAAGACGCGAAAACAGGTGGTGCACGGC
TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATC
TTCTGTTGCCAGCGCGTCATGGCGGGGACTCAGGAGAGACTGCCGCAGACAATGCGGAGG
AAGGCGGGGATGACGTCAAGTCATCATGCCCCTTATGGCTTGGGCTACACACGTACTACA
35 ATGGCTCTTAATAGAGGGAAGCGAAGGAGCGATCCGGAGCAAACCCCAAAAACAGAGTCC
CAGTTCGGATTGCAGGCTGCAACCCGCCTGCATGAAGCAGGAATCGCTAGTAATCGCAGG
TCAGCATACTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAA
AGTCATTCACACCCGAAGCCGGTGAGGTAACCGTAAGGAGCCAGCCGTCGAAGGTGGGGG
CGATGATTGGGGTGAAGTCGTAA

40

HMI_41 *Blautia_luti* 16S rDNA sequence (SEQ ID NO: 40)

GGTGAGTAACGCGTGGGTAACCTGCCTTATACAGGGGGATAACAGTCAGAAATGGCTGCT
AATACCGCATAAGCGCACAGGGCCGCATGGCCCGGTGTGAAAAACTGAGGTGGTATAAGA
5 TGGACCCGCGTTGGATTAGCCAGTTGGCAGGGTAACGGCCTACCAAAGCGACGATCCATA
GCCGGCCTGAGAGGGTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGG
AGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAA
GGAAGAAGTATCTCGGTATGTAAACTTCTATCAGCAGGGAAGAAAATGACGGTACCTGAC
TAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTA
10 TCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCATAACAAGTCTGATGTGAAAGGCT
GGGGCTTAACCCCGGGACTGCATTGGAAACTGTTAAGCTTGAGTGCCGGAGGGGTAAGCG
GAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCG
GCTTACTGGACGGTAACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGAT
ACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTCGGGGAGCACAGCTCTTCGG
15 TGCCGCCGCAAACGCATTAAGTATTCCACCTGGGGAGTACGTTCGCAAGAATGAAACTCA
AAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCG
AAGAACCTTACCAAGTCTTGACATCTGCCTGACCGGTGAGTAACGTCACCTTTCCTTCGG
GACAGGCAAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAA
GTCCCGCAACGAGCGCAACCCCTATCCCCAGTAGCCAGCATGTAAAGGTGGGCACTCTGA
20 GGAGACTGCCAGGGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTT
ATGATTTGGGCTACACACGTGCTACAATGGCGTAAACAGAGGGAAGCGAAAGGGTGACCT
GGAGCAAATCCCAAAAATAACGTCCCAGTTCGGACTGTAGTCTGCAACCCGACTACACGA
AGCTGGAATCGCTAGTAATCGCGGATCAGAATGCCGCGGTGAATACGTTCCCGGGTCTTG
TACACACCGCCCGTCACACCATGGGAGTCAGTAACGCCCGAAGTCAGTGACCTAACCGAA
25 AGGGAGGAGCTGCCGAAGGCGGGACGGATGACTGGGGTGAAGTCGTAAC

HMI_42 *Bacteroides coprocola* 16S rDNA sequence (SEQ ID NO: 41)

GTATCCAACCTTCCGTTTACTCAGGGATAGCCTTTCGAAAGAAAGATTAATACCTGATAG
30 TATGGTAAGATTGCATGATAATACCATTAAAGATTCATCGGTAAACGATGGGGATGCGTT
CCATTAGGTAGTAGGCGGGGTAACGGCCCACCTAGCCGACGATGGATAGGGGTTCTGAGA
GGAAGGTCCCCCACATTGGAACTGAGACACGGTCCAAACTCCTACGGGAGGCAGCAGTGA
GGAATATTGGTCAATGGGCGAGAGCCTGAACCAGCCAAGTAGCGTGAAGGATGAAGGTTC
TATGGATTGTAAACTTCTTTTATAAGGGAATAAAGTGCTTTACGTGTAGAGTTTTGTATG
35 TACCTTATGAATAAGCATCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGATGC
GAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGACGGGATGTTAAGTCAGCTGT
GAAAGTTTGGGGCTCAACCTTAAAATTGCAGTTGAAACTGGCGTTCTTGAGTGCGGTAGA
GGCAGGCGGAATTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACCCCGATTG
CGAAGGCAGCTTGCTGGAGCGTAACTGACGTTGATGCTCGAAAGTGTGGGTATCAAACAG

GATTAGATACCCTGGTAGTCCACACGGTAAACGATGGATACTCGCTGTTGGCGATATACG
GTCAGCGGCCAAGCGAAAGCATTAAGTATCCCACCTGGGGAGTACGCCGGCAACGGTGAA
ACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGAGGAACATGTGGTTTAATTCGATGAT
ACGCGAGGAACCTTACCCGGGCTTAAATTATGCATGAATGATCTGGAGACAGATCAGCCG
5 CAAGGCATGTATGAAGGTGCTGCATGGTTGTCGTCAGCTCGTGCCGTGAGGTGTCGGCTT
AAGTGCCATAACGAGCGCAACCCTTTCTGCCAGTTACTAACAGGCAATGCTGAGGACTCT
GGCGGTACTGCCATCGTAAGATGTGAGGAAGGTGGGGATGACGTCAAATCAGCACGGCCC
TTACGTCCGGGGCTACACACGTGTTACAATGGGGGGTACAGAAGGCAGCTTACCGGCGAC
GGTTGGCCAATCCCTAAAGCCCCTCTCAGTTCGGACTGGAGTCTGCAACCCGACTCCACG
10 AAGCTGGATTCGCTAGTAATCGCGCATCAGCCACGGCGCGGTGAATACGTTCCCGGGCCT
TGTACACACCGCCCGTCAAGCCATGAAAGCCGGGAGTACCTGAAGTGCGTAACCGCGAGG
AGCGCCCTAGGGTAACACTGGTAATTGGGGCTAAGTCGT

HMI_43 *Bacteroides plebius* 16S rDNA sequence (SEQ ID NO: 42)

15

GGGGCAGCATGAACTTAGCTTGCTAAGTTCGATGGCGACCGGCGCACCGTTGAGTAACGC
GTATCCAACCTTCCGTACACTCAGGAATAGCCTTTCGAAAGAAAGATTAATACCTGATGG
TATGATGGGATTGCATGAAATCATCATTAAAGATTCATCGGTGTACGATGGGGATGCGTT
CCATTAGATAGTAGGCGGGGTAACGGCCCACCTAGTCGACGATGGATAGGGGTTCTGAGA
20 GGAAGGTCCCCCACATTGGAACTGAGACACGGTCCAAACTCCTACGGGAGGCAGCAGTGA
GGAATATTGGTCAATGGGCGCGAGCCTGAACCAGCCAAGTAGCGTGAAGGATGAAGGTCC
TACGGATTGTAAACTTCTTTTATAAGGGAATAAAGTCACCCACGTGTGGGTGTTTGTATG
TACCTTATGAATAAGCATCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGATGC
GAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGACGGGTCGTTAAGTCAGCTGT
25 GAAAGTTCGGGGCTCAACCTTGAAATTGCAGTTGATACTGGCGTCCTTGAGTACGGTTGA
GGCAGGCGGAATTCGTGGTGTAGCGGTGAAATGCTTAGATATCACGAAGAACCCCGATTG
CGAAGGCAGCCTGCTAAACCGCCACTGACGTTGAGGCTCGAAAGTGTGGGTATCAAACAG
GATTAGATACCCTGGTAGTCCACACGGTAAACGATGGATACTCGCTGTTGGCGATAGACT
GTCAGCGGCTTAGCGAAAGCGTTAAGTATCCCACCTGGGGAGTACGCCGGCAACGGTGAA
30 ACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGAGGAACATGTGGTTTAATTCGATGAT
ACGCGAGGAACCTTACCCGGGCTTGAATTGCAGACGAATTGCTTGGAAACAGGCAAGCCG
CAAGGCGTCTGTGAAGGTGCTGCATGGTTGTCGTCAGCTCGTGCCGTGAGGTGTCGGCTT
AAGTGCCATAACGAGCGCAACCCTCGTGTCCAGTTGCTAGCAGGTAGTGCTGAGGACTCT
GGACAGACTGCCATCGTAAGATGTGAGGAAGGTGGGGATGACGTCAAATCAGCACGGCCC
35 TTACGTCCGGGGCTACACACGTGTTACAATGGGGGGTACAGCAGGCAGCTACCGGGCGAC
CGGATGCCAATCCCGAAAGCCTCTCTCAGTTCGGACTGGAGTCTGCAACCCGACTCCACG
AAGCTGGATTCGCTAGTAATCGCGCATCAGCCACGGCGCGGTGAATACGTTCCCGGGCCT
TGTACACACCGCCCGTCAAGCCATGAAAGCCGGGGGTACCTGAAGTGCGTAACCGCAAGG
AGCGCCCTAGGGTAAAACTGGTAATT

40

HMI_44 Roseburia inulinivorans 16S rDNA sequence (SEQ ID NO: 43)

GCACTTTTGCCGATTTTCTTCGGAACTGAAGTAATAGTGACTGAGTGGCGGACGGGTGAG
TAACGCGTGGATAACCTGCCTCACACAGGGGGATAACAGTTAGAAATGACTGCTAATACC
5 GCATAAGCGCACAGTACCGCATGGTACAGTGTGAAAAACTCCGGTGGTGTGAGATGGATC
CGCGTCTGATTAGCCAGTTGGCGGGGTAACGGCCCACCAAAGCGACGATCAGTAGCCGGC
CTGAGAGGGCGACCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAG
CAGTGGGGAATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAGCGAAGA
AGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAA
10 GCTCCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTATGGAGCAAGCGTTATCCGGA
TTTACTGGGTGTAAAGGGAGCGCAGGCGGTATGACAAGTCTGATGTGAAAGGCTGGGGCT
CAACCCCAGGACTGCATTGGAAACTGTCAGACTAGAGTGTCGGAGAGGTAAGTGGAATTC
CTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTAC
TGGACGACAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTG
15 GTAGTCCACGCCGTAAACGATGAATACTAGGTGTCGGGAGGCAGAGCCTTTCGGTGCCGC
AGCAAACGCAGTAAGTATTCCACCTGGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAA
TTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAAC
CTTACCAGGCCTTGACATCCCCCTGACGGGACAGTAATGTGTCCGTTCCTTCGGGACAGA
GGAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCG
20 CAACGAGCGCAACCCTTATCCTCAGTAGCCAGCGGATAAAGCCGGGCACTCTGTGGAGAC
TGCCAGGGACAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGGCC
TGGGCTACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGAAGCTGTGAAGTGAAGCA
AATCCCAAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGG
AATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGTCTTGTACACA
25 CCGCCCGTCACACCATGGGAGTCGGGAATGCCCGAAGCCGGTGACCCAACCTTAAGGAGG
GAGCCGTCGAAGGCAGGCCTGATAACTGGGGTGAAGTCGT

HMI_45 *Ruminococcus albus* 16S rDNA sequence (SEQ ID NO: 44)

30 CTGATCTAGTGGCGGACGGGTGAGTAACACGTGAGCAATCTGCCTTTCAGAGGGGGATAC
CGATTGGAAACGATCGTTAATACCGCATAACATAATTGAACCGCATGATTTGATTATCAA
AGATTTATCGCTGAAAGATGAGCTCGCGTCTGATTAGCTAGTTGGTAAGGTAACGGCTTA
CCAAGGCGACGATCAGTAGCCGGACTGAGAGGTTGATCGGCCACATTGGGACTGAGACAC
GGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGAGGAAACTCTGAT
35 GCAGCGATGCCGCGTGAGGGAAGAAGGTTTTAGGATTGTAAACCTCTGTCTTCAGGGACG
AAAAAAGACGGTACCTGAGGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAAT
ACGTAGGGAGCGAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATCGC
AAGTCAGATGTGAAAACTATGGGCTTAACCCATAAACTGCATTTGAAACTGTGGTTCTTG
AGTGAAGTAGAGGTAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAGG

AACATCAGTGGCGAAGGCGGCTTACTGGGCTTTAACTGACGCTGAGGCTCGAAAGCGTGG
GGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATTACTAGGTGTG
GGGGGACTGACCCCTTCCGTGCCGCAGCAAACGCAATAAGTAATCCACCTGGGGAGTACG
ACCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCAGTGGAGTATGTGG
5 ATTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCGTATGCATAGCTCAGA
GATGAGTGAAATCTCTTCGGAGACATATAGACAGGTGGTGCATGGTTGTCGTCAGCTCGT
GTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTACTGTTAGTTGCTACGC
AAGAGCACTCTAGCAGGACTGCCGTTGACAAAACGGAGGAAGGTGGGGATGACGTCAAAT
CATCATGCCCCTTATGACCTGGGCCTCACACGTACTACAATGGCTGTTAACAGAGGGATG
10 CAAAGCCGCGAGGTAGAGCGAACCCCTAAAAGCAGTCTTAGTTCGGATTGTAGGCTGCAA
CCCGCCTACATGAAGTCGGAATTGCTAGTAATCGCAGATCAGCATGCTGCGGTGAATACG
TTCCCGGGCCTTGTACACACCGCCCGTCACGCCATGGGAGTCGGTAACACCCGAAGCCTG
TAGTCTAACCGCAAGGAGGACGCAGTCGAAGGTGGGATTGATGACTGGGGTGAAGTCGTA
ACAGGGTAACCG

15

HMI_46 *Blautia producta* 16S rDNA sequence (SEQ ID NO: 45)

TGGACAGATTCTTCGGATGAAGTCCTTAGTGACTGAGTGGCGGACGGGTGAGTAACGCGT
GGGTAACCTGCCTCATACAGGGGGATAACAGTTAGAAATGGCTGCTAATACCGCATAAGC
20 GCACGGTACTGCATGGTACAGTGTGAAAAACTCCGGTGGTATGAGATGGACCCGCGTTGG
ATTAGCTAGTTGGCAGGGTAACGGCCTACCAAGGCGACGATCCATAGCCGGCCTGAGAGG
GTGGACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGG
AATATTGCACAATGGGGGAAACCCTGATGCAGCGACGCCGCGTGAGCGAAGAAGTATTTC
GGTATGTAAAGCTCTATCAGCAGGGAAGAAAATGACGGTACCTGACTAAGAAGCCCCGGC
25 TAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGG
GTGTAAAGGGAGCGTAGACGGAATGGCAAGTCTGATGTGAAAGGCCGGGGCTCAACCCCG
GGACTGCATTGGAAACTGTCAATCTAGAGTACCGGAGGGGTAAGTGGAATTCCTAGTGTA
GCGGTGAAATGCGTAGATATTAGGAGGAACACCAGTGGCGAAGGCGGCTTACTGGACGGT
AACTGACGTTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA
30 CGCCGTAAACGATGAATACTAGGTGTTGGGGAGCAAAGCTCTTCGGTGCCGCAGCAAACG
CAATAAGTATTCCACCTGGGGAGTACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGG
GACCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAA
ATCTTGACATCGATCTGACCGGACTGTAATGAGTCCTTTCCCTTCGGGGACAGAGAAGAC
AGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGA
35 GCGCAACCCTTATCCTCAGTAGCCAGCAAGTGAAGTTGGGCACTCTGTGGAGACTGCCAG
GGATAACCTGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGATTTGGGCT
ACACACGTGCTACAATGGCGTAAACAAAGGGAAGCGATCACGCGAGTGTGAGCAAATCTC
AAAAATAACGTCTCAGTTCGGATTGTAGTCTGCAACTCGACTACATGAAGCTGGAATCGC
TAGTAATCGCAGGTCAGCATACTGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCC
40 GTCACACCATGGGAGTCAGTAACACCCGAAGCCGGTGACCTAACCGAAAGGAAGGAGCCG

TCGAAGGTGGGACCGATAACTGGGGTGAAGTCGT

HMI_47 *Clostridium nexile* 16S rDNA sequence (SEQ ID NO: 46)

5 GTTTGTGACTTAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTTATACAGGGGG
ATAACAGTTAGAAATGACTGCTAATACCGCATAAGCGCACAGTCTCGCATGGGACAGTGT
GAAAAACTAAGGTGGTATAAGATGGACCCGCGTCTGATTAGCTAGTTGGTGGGGTAAAGG
CCTACCAAGGCGACGATCAGTAGCCGACCTGAGAGGGTGATCGGCCACATTGGGACTGAG
ACACGGCCCAAACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACCC
10 TGATGCAGCAACGCCGCGTGAAGGAAGAAGTATCTCGGTATGTAAACTTCTATCAGCAGG
GAAGAAAATGACGGTACCTGACTAAGAAGCTCCGGCTAAATACGTGCCAGCAGCCGCGGT
AATACGTATGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGGCGGTTA
TGCAAGTCAGATGTGAAAGCCCGGGGCTTAACCCCGGGACTGCATTTGAAACTGTGTAAC
TAGAGTGTCGGAGAGGTAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGG
15 AGGAACACCAGTGGCGAAGGCGGCTTACTGGACGATAACTGACGCTGAGGCTCGAAAGCG
TGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGT
GTTGGGGAGCAAAGCTCTTCGGTGCCGCAGCAAACGCAATAAGTATTCCACCTGGGGAGT
ACGTTCGCAAGAATGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATG
TGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAAGTCTTGACATCTGGATGACCGGAC
20 CGTAATGGGTCCTTTCCTTCGGGACATCCAAGACAGGTGGTGCATGGTTGTCGTCAGCTC
GTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATCCTTAGTAGCCAG
CAGTAAGATGGGCACTCTAGGGAGACTGCCGGAGACAATCCGGAGGAAGGTGGGGATGAC
GTCAAATCATCATGCCCCTTATGACTTGGGCTACACACGTGCTACAATGGCGTAAACAAA
GGGAAGCGAGACCGCGAGGTTAAGCAAATCTCAAAAATAACGTCTCAGTTCGGATTGTAG
25 TCTGCAACTCGACTACATGAAGCTGGAATCGCTAGTAATCGCGAATCAGCATGTCGCGGT
GAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTCGATAACGCCCG
AAGCCGGTGACTCAACCGAAAGGAGAGAGCCGTCGAAGGCGGGATGGATAACTGGGGTGA
AGTCGTAAC

30 HMI_48 Butyricicoccus pullicaecorum 16S rDNA sequence (SEQ ID NO: 47)

ATCTCTTCGGAGATGGAATTCTTAACCTAGTGGCGGACGGGTGAGTAACGCGTGAGCAAT
CTGCCTTTAGGAGGGGGATAACAGTCGGAAACGGCTGCTAATACCGCATAATACGTTTGG
GAGGCATCTCTTGAACGTCAAAGATTTTATCGCCTTTAGATGAGCTCGCGTCTGATTAGC
35 TGGTTGGCGGGGTAACGGCCCACCAAGGCGACGATCAGTAGCCGGACTGAGAGGTTGAAC
GGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATT
GCGCAATGGGGGAAACCCTGACGCAGCAACGCCGCGTGATTGAAGAAGGCCTTCGGGTTG
TAAAGATCTTTAATCAGGGACGAAAAATGACGGTACCTGAAGAATAAGCTCCGGCTAACT
ACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCAAGCGTTATCCGGATTTACTGGGTGTA

AAGGGCGCGCAGGCGGGCCGGCAAGTTGGGAGTGAAATCCCGGGGCTTAACCCCGGAACT
GCTTTCAAAACTGCTGGTCTTGAGTGATGGAGAGGCAGGCGGAATTCCGTGTGTAGCGGT
GAAATGCGTAGATATACGGAGGAACACCAGTGGCGAAGGCGGCCTGCTGGACATTAACTG
ACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCG
5 TAAACGATGGATACTAGGTGTGGGAGGTATTGACCCCTTCCGTGCCGCAGTTAACACAAT
AAGTATCCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCC
CGCACAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCT
TGACATCCCGATGACCGGCGTAGAGATACGCCCTCTCTTCGGAGCATCGGTGACAGGTGG
TGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAA
10 CCCTTACGGTTAGTTGATACGCAAGATCACTCTAGCCGGACTGCCGTTGACAAAACGGAG
GAAGGTGGGGACGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTACTAC
AATGGCAGTCATACAGAGGGAAGCAATACCGCGAGGTGGAGCAAATCCCTAAAAGCTGTC
CCAGTTCAGATTGCAGGCTGCAACCCGCCTGCATGAAGTCGGAATTGCTAGTAATCGCGG
ATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGA
15 GAGCCGTCAATACCCGAAGTCCGTAGCCTAACCGCAAGGGGGGCGCGGCCGAAGGTAGGG
GTGGTAATTAGGGTGAAGTCGTAC

HMI_49 Ruminococcus flavefaciens 16S rDNA sequence (SEQ ID NO: 48)

20 AGTCGACGGACGAGGAGGAGCTTGCTTCTCCGAGTTAGTGGCGGACGGGTGAGTAACACG
TGAGCAACCTACCCTTGAGAGGGGGATAGCTTCTGGAAACGGATGGTAATACCCCATAAC
ATATATTTTAGGCATCTAAGATATATCAAAGAAATTCGCTCAAGGATGGGCTCGCGTCTG
ATTAGATAGTTGGTGAGGTAACGGCCCACCAAGTCGACGATCAGTAGCCGGACTGAGAGG
TTGAACGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGG
25 AATATTGCACAATGGGGGGAACCCTGATGCAGCGATGCCGCGTGGAGGAAGAAGGTTTTC
GGATTGTAAACTCCTTTTAACAGGGACGATAATGACGGTACCTGAAGAAAAAGCTCCGGC
TAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCGAGCGTTGTCCGGAATTACTGG
GTGTAAAGGGAGCGTAGGCGGGACGGTAAGTCAGGTGTGAAATATACGTGCTCAACATGT
AGACTGCACTTGAAACTGCTGTTCTTGAGTGAAGTAGAGGTAAGCGGAATTCCTAGTGTA
30 GCGGTGAAATGCGTAGATATTAGGAGGAACATCGGTGGCGAAGGCGGCTTACTGGGCTTT
TACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA
CGCTGTAAACGATGATTACTAGGTGTGGGGGGACTGACCCCTTCCGTGCCGCAGTTAACA
CAATAAGTAATCCACCTGGGGAGTACGACCGCAAGGTTGAAACTCAAAGGAATTGACGGG
GGCCCGCACAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAG
35 GTCTTGACATCGTATGCATAGTCTAGAGATAGATGAAATCCCTTCGGGGACATATAGACA
GGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAG
CGCAACCCTTACCTTTAGTTGCTACGCAAGAGCACTCTAGAGGGACTGCCGTTGACAAAA
CGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGT
ACTACAATGGCAATTAACAGAGGGAAGCAAAACAGCGATGTGGAGCAAATCCCGAAAAAT
40 TGTCCCAGTTCAGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAATTGCTAGTAATC

GCAGATCAGAATGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC
ATGGGAGTCGGTAACACCCGAAGCCTGTAGTCTAACCTTATAGGAGGACGCAGTCGAAGG
TGGGATTGATGACTGGGGTGAAGTCGT

5 HMI_50 Clostridium orbiscindens 16S rDNA sequence (SEQ ID NO: 49)

AAAGGGAATGCTTAGTGGCGGACGGGTGAGTAACGCGTGAGTAACCTGCCTTGGAGTGGG
GAATAACAGCCGGAAACGGCTGCTAATACCGCATGATGTATCTGGATCGCATGGTTCTGG
ATACCAAAGATTTATCGCTCTGAGATGGACTCGCGTCTGATTAGCTAGTTGGTGAGGTAA
10 CGGCTCACCAAGGCGACGATCAGTAGCCGGACTGAGAGGTTGGCCGGCCACATTGGGACT
GAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGGCAATGGGCGAAA
GCCTGACCCAGCAACGCCGCGTGAAGGAAGAAGGCCCTCGGGTTGTAAACTTCTTTTGTC
AGGGACGAAGCAAGTGACGGTACCTGACGAATAAGCCACGGCTAACTACGTGCCAGCAGC
CGCGGTAATACGTAGGTGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGTGTAGG
15 CGGGAGTGCAAGTCAGATGTGAAAACTATGGGCTCAACCCATAGCCTGCATTTGAAACTG
TACTTCTTGAGTGATGGAGAGGCAGGCGGAATTCCCTGTGTAGCGGTGAAATGCGTAGAT
ATAGGGAGGAACACCAGTGGCGAAGGCGGCCTGCTGGACATTAACTGACGCTGAGGCGCG
AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATA
CTAGGTGTGGGGGGTCTGACCCCCTCCGTGCCGCAGTTAACACAATAAGTATCCCACCTG
20 GGGAGTACGATCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGG
AGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGACTTGACATCCTACTAA
CGAAGCAGAGATGCATAAGGTGCCCTTCGGGGAAAGTAGAGACAGGTGGTGCATGGTTGT
CGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATTGTT
AGTTGCTACGCAAGAGCACTCTAGCGAGACTGCCGTTGACAAAACGGAGGAAGGTGGGGA
25 CGACGTCAAATCATCATGCCCCTTATGTCCTGGGCCACACACGTACTACAATGGCGGTCA
ACAGAGGGAAGCAAAGCCGCGAGGTGGAGCAAATCCCTAAAAGCCGTCCCAGTTCGGATT
GCAGGCTGAAACTCGCCTGTATGAAGTCGGAATCGCTAGTAATCGCGGATCAGCATGCCG
CGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGAGAGTCGGGAACA
CCCGAAGTCCGTAGCCTAACAGCAATGGG

30

HMI_51 *Ruminococcus bromii* 16S rDNA sequence (SEQ ID NO: 50)

ACGAAGCTTTGAGGAGCTTGCTTTTTAAGCTTAGTGGCGGACGGGTGAGTAACGCGTGAG
CAACCTGCCTCTCAGAGGGGAATAACGTTTTGAAAAGAACGCTAATACCGCATAACATAT
35 CGGAACCGCATGATTCTGATATCAAAGGAGCAATCCGCTGAGAGATGGGCTCGCGTCCGA
TTAGTTAGTTGGTGAGGTAACGGCTCACCAAGACTACGATCGGTAGCCGGACTGAGAGGT
TGATCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGG
ATATTGCGCAATGGGGGAAACCCTGACGCAGCAACGCCGCGTGAAGGAAGAAGGTCTTCG
GATTGTAAACTTCTTTTGTCAGGGACGAAGAAAGTGACGGTACCTGACGAATAAGCTCCG

GCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCGAGCGTTGTCCGGATTTACT
GGGTGTAAAGGGTGCGTAGGCGGCCGAGCAAGTCAGTTGTGAAAACTATGGGCTTAACCC
ATAACGTGCAATTGAAACTGTCCGGCTTGAGTGAAGTAGAGGTAGGCGGAATTCCCGGTG
TAGCGGTGAAATGCGTAGAGATCGGGAGGAACACCAGTGGCGAAGGCGGCCTACTGGGCT
5 TTAACTGACGCTGAGGCACGAAAGCATGGGTAGCAAACAGGATTAGATACCCTGGTAGTC
CATGCCGTAAACGATGATTACTAGGTGTGGGGGGACTGACCCCTTCCGTGCCGCAGTTAA
CACAATAAGTAATCCACCTGGGGAGTACGGCCGCAAGGTTGAAACTCAAAGGAATTGACG
GGGGCCCGCACAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACC
AGGTCTTGACATCCTGAGAATCCTTAAGAGATTAGGGAGTGCCTTCGGGAACTCAGAGAC
10 AGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGA
GCGCAACCCTTGCTATTAGTTGCTACGCAAGAGCACTCTAATAGGACTGCCGTTGACAAA
ACGGAGGAAGGTGGGGACGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACG
TACTACAATGGCCATTAACAGAGGGAAGCAAAACCGCGAGGCAGAGCAAACCCCTAAAAA
TGGTCCCAGTTCGGATTGTAGGCTGCAACCCGCCTACATGAAGTTGGAATTGCTAGTAAT
15 CGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACAC
CATGGGAGCCGGTAATACCCGAAGTCAGTAGTCTAACAGCAATGAGGACGCTGCCGAAGG
TAGGATTGGCGACTGGGGTGAAGTCGTAACAAGGTAACCG

HMI_52 *Ruminococcus albus* 926R 16S rDNA sequence (SEQ ID NO: 51)

20

AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCACGCTTAACACATGCAAGTCGAAC
GAGAGAAGAGAAGCTTGCTTTTCTGATCTAGTGGCGGACGGGTGAGTAACACGTGAGCAA
TCTGCCTTTCAGAGGGGGATACCGATTGGAAACGATCGTTAATACCGCATAACATAATTG
AACCGCATGATTTGATTATCAAAGATTTATCGCTGAAAGATGAGCTCGCGTCTGATTAGC
25 TAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATCAGTAGCCGGACTGAGAGGTTGATC
GGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATT
GCACAATGGAGGAAACTCTGATGCAGCGATGCCGCGTGAGGGAAGAAGGTTTTAGGATTG
TAAACCTCTGTCTTCAGGGACGAAAAAAAAGACGGTACCTGAGGAGGAAGCTCCGGCTAA
CTACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCGAGCGTTGTCCGGAATTACTGGGTG
30 TAAAGGGAGCGTAGGCGGGATCGCAAGTCAGATGTGAAAACTATGGGCTTAACCCATAAA
CTGCATTTGAAACTGTGGTTCTTGAGTGAAGTAGAGGTAAGCGGAATTCCTAGTGTAGCG
GTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCGGCTTACTGGGCTTTAAC
TGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGC
CGTAAACGATGATTACTAGGTGTGGGGGGACTGACCCCTTCCGTGCCGCAGCAAACGCAA
35 TAAGTAATCCACCTGGGGAGTACGACCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGC
CCGCACAAGCAGTGGAGTATGTGGATTAATTCGAAGCAACGCGAAGAACCTTACCAGGTC
TTGACATCGTATGCATAGCTCAGAGATGAGTGAAATCTCTTCGGAGACATATAGACAGGT
GGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGC
AACCCTTACTGTTAGTTGCTACGCAAGAGCACTCTAGCAGGACTGCCGTTGACAAAACGG
40 AGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCCTCACACGTACT

ACAATGGCTGTCAACAGAGGGATGCAAAGCCGCGAGGTGGAGCGAACCCCTAAAAGCAGT
CTTAGTTCGGATTGTAGGCTGCAACCCGCCTACATGAAGTCGGAATTGCTAGTAATCGCA
GATCAGCATGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGCCATG
GGAGTCGGTAACACCCGAAGCCTGTAGTCTAACCGCAAGGAGGACGCAGTCGAAGGTGGG
5 ATTGATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCG

References

All documents mentioned in this specification are incorporated herein by reference in their entirety.

5

Abecasis, A. B., M. Serrano, R. Alves, L. Quintais, J. B. Pereira-Leal and A. O. Henriques (2013). "A genomic signature and the identification of new sporulation genes." J Bacteriol 195(9): 2101-2115.

10 Abujamel, T et al. (2013). "Defining the Vulnerable Period for Re-Establishment of Clostridium difficile Colonization after Treatment of C. difficile Infection with Oral Vancomycin or Metronidazole." PLoS ONE 8(10): e76269.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local
15 alignment search tool." Journal of Molecular Biology 215(3): 403-410.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol 215(3): 403-410.

20 Angriman, I., M. Scarpa and I. Castagliuolo (2014). “Relationship between pouch microbiota and pouchitis following restorative proctocolectomy for ulcerative colitis.” World J Gastroenterol 20(29): 9665-9674.

Atarashi, K., T. Tanoue, K. Oshima, W. Suda, Y. Nagano, H. Nishikawa, S. Fukuda, T. Saito,
25 S. Narushima, K. Hase, S. Kim, J. V. Fritz, P. Wilmes, S. Ueha, K. Matsushima, H. Ohno, B. Olle, S. Sakaguchi, T. Taniguchi, H. Morita, M. Hattori and K. Honda (2013). “T_{reg} induction by a rationally selected mixture of Clostridia strains from the human microbiota”. Nature 500(7461): 232-236.

30 Blanton, L. V., M. R. Charbonneau, T. Salih, M. J. Barratt, S. Venkatesh, O. Ilkaveya, S. Subramanian, M. J. Manary, I. Trehan, J. M. Jorgensen, Y. M. Fan, B. Henrissat, S. A. Leyn, D. A. Rodionov, A. L. Osterman, K. M. Maleta, C. B. Newgard, P. Ashorn, K. G. Dewey and J. I. Gordon (2016). "Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children." Science **351**(6275).
35

Bajaj, J. S. (2014). “The role of microbiota in hepatic encephalopathy.” Gut Microbes 5(3): 397-403.

Bajaj, J. S., P. B. Hylemon and Z. Younossi (2012). “The intestinal microbiota and liver
40 disease.” Am J Gastroenterol. Suppl. 1:9-14

Bosshard, P. P., S. Abels, R. Zbinden, E. C. Bottger and M. Altwegg (2003). "Ribosomal DNA sequencing for identification of aerobic gram-positive rods in the clinical laboratory (an 18-month evaluation)." J Clin Microbiol 41(9): 4134-4140.
45

Britton, R. A. and V. B. Young (2014). "Role of the intestinal microbiota in resistance to colonization by Clostridium difficile." Gastroenterology 146(6): 1547-1553.

50 Buffie, C. G., et al. (2015). "Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile." Nature 517(7533): 205-208.

Chakraborty, A., S. Ghosh, G. Chowdhary, U. Maulik and S. Chakrabarti (2012). "DBETH: a Database of Bacterial Exotoxins for Human." Nucleic Acids Res 40(Database issue): D615-620.

Clarridge, J. E., 3rd (2004). "Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases." Clin Microbiol Rev 17(4): 840-862, table of contents.

Cole, J. R., et al. (2014). "Ribosomal Database Project: data and tools for high throughput rRNA analysis." Nucleic Acids Research 42(D1): D633-D642.

Collins, S. M. (2014). "A role for the gut microbiota in IBS." Nat Rev Gastroenterol Hepatol 11(8): 497-505.

Cornely, O. A., M. A. Miller, T. J. Louie, D. W. Crook and S. L. Gorbach (2012). "Treatment of First Recurrence of Clostridium difficile Infection: Fidaxomicin Versus Vancomycin." Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America 55(Suppl 2): S154-S161.

Dominguez-Bello, M. G., E. K. Costello, M. Contreras, M. Magris, G. Hidalgo, N. Fierer and R. Knight (2010). "Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns." Proc Natl Acad Sci U S A 107(26): 11971-11975.

Duncan, S. H., G. L. Hold, H. J. Harmsen, C. S. Stewart and H. J. Flint (2002). "Growth requirements and fermentation products of Fusobacterium prausnitzii, and a proposal to reclassify it as Faecalibacterium prausnitzii gen. nov., comb. nov." Int J Syst Evol Microbiol 52(Pt 6): 2141-2146.

Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson and D. A. Relman (2005). "Diversity of the human intestinal microbial flora." Science 308(5728): 1635-1638.

Fodor, A. A., T. Z. DeSantis, K. M. Wylie, J. H. Badger, Y. Ye, T. Hepburn, P. Hu, E. Sodergren, K. Liolios, H. Huot-Creasy, B. W. Birren and A. M. Earl (2012). "The "most wanted" taxa from the human microbiome for whole genome sequencing." PLoS One 7(7): e41294.

Forster, S. C., H. P. Browne, N. Kumar, M. Hunt, H. Denise, A. Mitchell, R. D. Finn and T. D. Lawley (2015). "HPMCD: the database of human microbial communities from metagenomic datasets and microbial reference genomes." Nucleic Acids Res.

Francis, M. B., C. A. Allen, R. Shrestha and J. A. Sorg (2013). "Bile acid recognition by the Clostridium difficile germinant receptor, CspC, is important for establishing infection." PLoS Pathog 9(5): e1003356.

Galperin, M. Y., S. L. Mekhedov, P. Puigbo, S. Smirnov, Y. I. Wolf and D. J. Rigden (2012). "Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes." Environ Microbiol 14(11): 2870-2890.

Goodman, A. L., G. Kallstrom, J. J. Faith, A. Reyes, A. Moore, G. Dantas and J. I. Gordon (2011). "Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice." Proc Natl Acad Sci U S A 108(15): 6252-6257.

5 Hattori, M. and T. D. Taylor (2009). "The human intestinal microbiome: a new frontier of human biology." DNA Res 16(1): 1-12.

Hold, G. L., M. Smith, C. Grange, E. R. Watt, E. M. El-Omar and I. Mukhopadhya (2014). "Role of the gut microbiota in inflammatory bowel disease pathogenesis: what have we 10 learnt in the past 10 years?" World J Gastroenterol 20(5): 1192-1210.

Hooper, L. V., D. R. Littman and A. J. Macpherson (2012). "Interactions between the microbiota and the immune system." Science 336(6086): 1268-1273.

15 Hooper, L. V., M. H. Wong, A. Thelin, L. Hansson, P. G. Falk and J. I. Gordon (2001). "Molecular analysis of commensal host-microbial relationships in the intestine." Science 291(5505): 881-884.

Huse, S. M., L. Dethlefsen, J. A. Huber, D. M. Welch, D. A. Relman and M. L. Sogin (2008). 20 "Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing." PLoS Genet 4(11): e1000255.

Janoir, C., C. Deneve, S. Bouttier, F. Barbut, S. Hoys, L. Caleechum, D. Chapeton-Montes, F. C. Pereira, A. O. Henriques, A. Collignon, M. Monot and B. Dupuy (2013). "Adaptive 25 strategies and pathogenesis of Clostridium difficile from in vivo transcriptomics." Infect Immun 81(10): 3757-3769.

Johnson, M., I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis and T. L. Madden (2008). "NCBI BLAST: a better web interface." Nucleic Acids Research 36(Web Server 30 issue): W5-W9.

Jostins, L., et al. (2012). "Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease." Nature 491(7422): 119-124.

35 Koenig, J. E., et al. (2011). "Succession of microbial consortia in the developing infant gut microbiome." Proc Natl Acad Sci U S A 108 Suppl 1: 4578-4585.

Kozich, J. J., S. L. Westcott, N. T. Baxter, S. K. Highlander and P. D. Schloss (2013). "Development of a dual-index sequencing strategy and curation pipeline for analyzing 40 amplicon sequence data on the MiSeq Illumina sequencing platform." Appl Environ Microbiol 79(17): 5112-5120.

Lagier, J. C., P. Hugon, S. Khelaifia, P. E. Fournier, B. La Scola and D. Raoult (2015). "The Rebirth of Culture in Microbiology through the Example of Culturomics To Study Human Gut 45 Microbiota." Clin Microbiol Rev 28(1): 237-264.

Landy, J., H. O. Al-Hassi, S. D. McLaughlin, A. W. Walker, P. J. Ciclitira, R. J. Nicholls, S. K. Clark and A. L. Hart (2011). "Review article: faecal transplantation therapy for gastrointestinal disease." Alimentary Pharmacology & Therapeutics 34(4): 409-415.

50

Lawley, T. D., et al. (2009). "Antibiotic treatment of clostridium difficile carrier mice triggers a supershedder state, spore-mediated transmission, and severe disease in immunocompromised hosts." Infect Immun 77(9): 3661-3669.

Lawley, T. D., et al. (2012). "Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy resolves relapsing Clostridium difficile disease in mice." PLoS Pathog 8(10): e1002995.

5

Lawley, T. D. and A. W. Walker (2013). "Intestinal colonization resistance." Immunology 138(1): 1-11.

Letunic, I. and P. Bork (2011). "Interactive Tree Of Life v2: online annotation and display of
10 phylogenetic trees made easy." Nucleic Acids Res 39(Web Server issue): W475-478.

Louis, P and H. J. Flint (2009). “Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine.” FEMS Microbiol. Lett. 294(1): 1-8.

15 Lozupone, C. and R. Knight (2005). "UniFrac: a New Phylogenetic Method for Comparing Microbial Communities." Applied and Environmental Microbiology 71(12): 8228-8235.

Ludwig, W., et al. (2004). "ARB: a software environment for sequence data." Nucleic Acids Res 32(4): 1363-1371.

20

McArthur, A. G. et al. (2013). "The comprehensive antibiotic resistance database." Antimicrob Agents Chemother 57(7): 3348-3357.

25 Meehan, C. J. and R. G. Beiko (2014). "A phylogenomic view of ecological specialization in the Lachnospiraceae, a family of digestive tract-associated bacteria." Genome Biol Evol 6(3): 703-713.

Nielsen, H. B., et al. (2014). "Identification and assembly of genomes and genetic elements
30 in complex metagenomic samples without using reference genomes." Nat Biotechnol 32(8): 822-828.

Ottman, N., H. Smidt, W. M. de Vos and C. Belzer (2012). "The function of our microbiota: who is out there and what do they do?" Front Cell Infect Microbiol 2: 104.

35

Perez Martinez, G., C. Bauerl and M. C. Collado (2014). "Understanding gut microbiota in elderly's health will enable intervention through probiotics." Benef Microbes 5(3): 235-246.
Petrof, E. O., G. B. Gloor, S. J. Vanner, S. J. Weese, D. Carter, M. C. Daigneault, E. M.
40 Brown, K. Schroeter and E. Allen-Vercoe (2013). "Stool substitute transplant therapy for the eradication of Clostridium difficile infection: 'RePOOPulating' the gut." Microbiome 1(1): 3.

Price, M. N., P. S. Dehal and A. P. Arkin (2010). "FastTree 2--approximately maximum-likelihood trees for large alignments." PLoS One 5(3): e9490.

45

Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies and F. O. Glöckner (2007). "SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB." Nucleic acids research 35(21): 7188-7196.

50 Qin, J., et al. (2010). "A human gut microbial gene catalogue established by metagenomic sequencing." Nature 464(7285): 59-65.

Quince, C., A. Lanzen, R. Davenport and P. Turnbaugh (2011). "Removing Noise From Pyrosequenced Amplicons." BMC Bioinformatics 12(1): 38.

Rajilic-Stojanovic, M. and W. M. de Vos (2014). "The first 1000 cultured species of the
5 human gastrointestinal microbiota." FEMS Microbiol Rev 38(5): 996-1047.

Riley, T. V., J. S. Brazier, H. Hassan, K. Williams and K. D. Phillips (1987). "Comparison of alcohol shock enrichment and selective enrichment for the isolation of Clostridium difficile." Epidemiol Infect 99(2): 355-359.
10

Scheperjans, F., et al. (2015). "Gut microbiota are related to Parkinson's disease and clinical phenotype." Mov Disord 30(3): 350-358.

Schloss, P. D., D. Gevers and S. L. Westcott (2011). "Reducing the Effects of PCR
15 Amplification and Sequencing Artifacts on 16S rRNA-Based Studies." PLoS ONE 6(12): e27310.

Schloss, P. D., et al. (2009). "Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities."
20 Applied and Environmental Microbiology 75(23): 7537-7541.

Schloss, P. D., et al. (2009). "Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities." Appl Environ Microbiol 75(23): 7537-7541.
25

Sekirov, I., S. L. Russell, L. C. Antunes and B. B. Finlay (2010). "Gut microbiota in health and disease." Physiol Rev 90(3): 859-904.

Snyder, A (2014). “Genetic basis for clinical response to CTLA-4 blockade in melanoma”. N
30 Engl J Med. 371(23):2189-99.

Stewart, E. J. (2012). "Growing unculturable bacteria." J Bacteriol 194(16): 4151-4160.
van Nood, E., A. Vrieze, M. Nieuwdorp, S. Fuentes, E. G. Zoetendal, W. M. de Vos, C. E. Visser, E. J. Kuijper, J. F. Bartelsman, J. G. Tijssen, P. Speelman, M. G. Dijkgraaf and J. J.
35 Keller (2013). "Duodenal infusion of donor feces for recurrent Clostridium difficile." N Engl J Med 368(5): 407-415.

Wang, Q., G. M. Garrity, J. M. Tiedje and J. R. Cole (2007). "Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy." Appl Environ
40 Microbiol 73(16): 5261-5267.

Xu et al. (2015). “Fecal microbiota transplantation broadening its application beyond intestinal disorders”. World J Gastroenterol. 21(1): 102-111.

45 Ze, X., F. Le Mougen, S. H. Duncan, P. Louis and H. J. Flint (2013). "Some are more equal than others: the role of "keystone" species in the degradation of recalcitrant substrates." Gut Microbes 4(3): 236-240

Claims

1. A therapeutic composition comprising at least one isolated bacterium and a pharmaceutically acceptable excipient, wherein the bacterium comprises a gene encoding a 16S ribosomal RNA (rRNA) and said gene comprises a sequence with at least 90% sequence identity with the sequence set forth in any one of SEQ ID NOs 1 to 51.

2. The therapeutic composition according to claim 1, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1.

3. The therapeutic composition according to claim 1 or 2, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 2.

4. The therapeutic composition according to any one of claims 1 to 3, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 3.

5. The therapeutic composition according to any one of claims 1 to 4, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 4.

6. The therapeutic composition according to any one of claims 1 to 5, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 5.

7. The therapeutic composition according to any one of claims 1 to 6, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 6.

8. The therapeutic composition according to any one of claims 1 to 7, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 7.

9. The therapeutic composition according to any one of claims 1 to 8, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 8.

5 10. The therapeutic composition according to any one of claims 1 to 9, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 9.

11. The therapeutic composition according to any one of claims 1 to 10, wherein the 10 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 10.

12. The therapeutic composition according to any one of claims 1 to 11, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity 15 with the sequence set forth in SEQ ID NO: 11.

13. The therapeutic composition according to any one of claims 1 to 12, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 12.

20

14. The therapeutic composition according to any one of claims 1 to 13, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 13.

25 15. The therapeutic composition according to any one of claims 1 to 14, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 14.

16. The therapeutic composition according to any one of claims 1 to 15, wherein the 30 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 15.

17. The therapeutic composition according to any one of claims 1 to 16, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity 35 with the sequence set forth in SEQ ID NO: 16.

18. The therapeutic composition according to any one of claims 1 to 17, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 17.

5 19. The therapeutic composition according to any one of claims 1 to 18, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 18.

20. The therapeutic composition according to any one of claims 1 to 19, wherein the 10 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 19.

21. The therapeutic composition according to any one of claims 1 to 20, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity 15 with the sequence set forth in SEQ ID NO: 20.

22. The therapeutic composition according to any one of claims 1 to 21, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 21.

20

23. The therapeutic composition according to any one of claims 1 to 22, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 22.

25 24. The therapeutic composition according to any one of claims 1 to 23, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 23.

25. The therapeutic composition according to any one of claims 1 to 24, wherein the 30 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 24.

26. The therapeutic composition according to any one of claims 1 to 25, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity 35 with the sequence set forth in SEQ ID NO: 25.

27. The therapeutic composition according to any one of claims 1 to 26, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 26.

5 28. The therapeutic composition according to any one of claims 1 to 27, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 27.

29. The therapeutic composition according to any one of claims 1 to 28, wherein the
10 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 28.

30. The therapeutic composition according to any one of claims 1 to 29, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity
15 with the sequence set forth in SEQ ID NO: 29.

31. The therapeutic composition according to any one of claims 1 to 30, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 30.
20

32. The therapeutic composition according to any one of claims 1 to 31, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 31.

25 33. The therapeutic composition according to any one of claims 1 to 32, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 32.

34. The therapeutic composition according to any one of claims 1 to 33, wherein the
30 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 33.

35. The therapeutic composition according to any one of claims 1 to 34, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity
35 with the sequence set forth in SEQ ID NO: 34.

36. The therapeutic composition according to any one of claims 1 to 35, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 35.

5 37. The therapeutic composition according to any one of claims 1 to 36, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 36.

38. The therapeutic composition according to any one of claims 1 to 37, wherein the
10 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 37.

39. The therapeutic composition according to any one of claims 1 to 38, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity
15 with the sequence set forth in SEQ ID NO: 38.

40. The therapeutic composition according to any one of claims 1 to 39, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 39.

20

41. The therapeutic composition according to any one of claims 1 to 40, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 40.

25 42. The therapeutic composition according to any one of claims 1 to 41, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 41.

43. The therapeutic composition according to any one of claims 1 to 42, wherein the
30 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 42.

44. The therapeutic composition according to any one of claims 1 to 43, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity
35 with the sequence set forth in SEQ ID NO: 43.

45. The therapeutic composition according to any one of claims 1 to 44, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 44.

5 46. The therapeutic composition according to any one of claims 1 to 45, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 45.

47. The therapeutic composition according to any one of claims 1 to 46, wherein the 10 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 46.

48. The therapeutic composition according to any one of claims 1 to 47, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity 15 with the sequence set forth in SEQ ID NO: 47.

49. The therapeutic composition according to any one of claims 1 to 48, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 48.

20

50. The therapeutic composition according to any one of claims 1 to 49, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 49.

25 51. The therapeutic composition according to any one of claims 1 to 50, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 50.

52. The therapeutic composition according to any one of claims 1 to 51, wherein the 30 gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 51.

53. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is at least 91%.

35

54. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is at least 92%.

55. The therapeutic composition according to any one of claims 1 to 52, wherein the
5 sequence identity is at least 93%.

56. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is at least 94%.

10 57. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is at least 95%.

58. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is at least 96%.

15

59. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is at least 97%.

60. The therapeutic composition according to any one of claims 1 to 52, wherein the
20 sequence identity is at least 98%.

61. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is at least 98.7%.

25 62. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is at least 99%.

63. The therapeutic composition according to any one of claims 1 to 52, wherein the sequence identity is 100%.

30

64. The therapeutic composition according to any one of claims 1 to 52, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 21, at least 91% sequence identity with the sequence set forth in SEQ ID NO: 29, at least 92% sequence identity with
35 the sequence set forth in any one of SEQ ID NOs 6, 11, 19 or 24, at least 93% sequence identity with the sequence set forth in any one of SEQ ID NOs 13, 22, 26 or 35, at least 94%

sequence identity with the sequence set forth in any one of SEQ ID NOs 5, 14, 15, 17, 18, 23, or 50, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 20, 33, 41, 43, 45, 46, 47, or 49, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 7, 8, 10, 12, 30, 32, 39, 42, 44, or 48, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 3, 16, 25, 31, 34, 36, 37, or 40, at least 98% sequence identity with the sequence set forth in any one of SEQ ID NOs 4 or 9, or at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 27, 28, 38, or 51.

65. The therapeutic composition according to any one of claims 1 to 52, wherein the gene encoding the 16S rRNA comprises a sequence with at least 90% sequence identity with the sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 21, at least 92% sequence identity with the sequence set forth in any one of SEQ ID NOs 6, or 11, at least 93% sequence identity with the sequence set forth in SEQ ID NO: 35, at least 94% sequence identity with the sequence set forth in any one of SEQ ID NOs 5, 19, 22, 23, or 50, at least 95% sequence identity with the sequence set forth in any one of SEQ ID NOs 13, 15, 29, 33, 41, 43, 45, or 46, at least 96% sequence identity with the sequence set forth in any one of SEQ ID NOs 7, 12, 32, 39, 42, or 44, at least 97% sequence identity with the sequence set forth in any one of SEQ ID NOs 3, 8, 10, 16, 34, 36, 37, 40, 48, or 49, at least 98% sequence identity with the sequence set forth in any one of SEQ ID NOs 4, 9, 17 or 31, at least 99% sequence identity with the sequence set forth in any one of SEQ ID NOs 2, 20, 38, or 51, or 100% sequence identity with the sequence set forth in any one of SEQ ID NOs 14, 18, 24, 25, 26, 27, 28, 30, or 47.

66. The therapeutic composition according to any one of claims 1 to 65, wherein the isolated bacterium is a bacterium as deposited at the Leibniz-Institut DSMZ under accession number DSM32191, DSM32147, DSM32149, DSM32175, DSM32153, DSM32152, DSM32158, DSM32192, DSM32148, DSM32166, DSM32151, DSM32150, DSM32193, DSM32162, DSM32194, DSM32163, DSM32205, DSM32195, DSM32164, DSM32177, DSM32167, DSM32165, DSM32169, DSM32168, DSM32178, DSM32182, DSM32179, DSM32180, DSM32184, DSM32181, DSM32183, DSM32262, DSM32211, DSM32219, DSM32222, DSM32261, DSM32212, DSM32220, DSM32213, DSM32226, DSM32215, DSM32216, DSM32217, DSM32221, DSM32218, DSM32224, DSM32214, DSM32263, DSM32223, DSM32225, or DSM32265.

67. The therapeutic composition according to any one of claims 1 to 66, wherein the composition comprises at least two distinct isolated bacteria, and wherein the bacteria are as defined in any one of claims 1 to 66.

5 68. The therapeutic composition according to any one of claims 1 to 67, wherein the composition is for use in a method of treating a dysbiosis of the gastrointestinal tract in an individual.

69. A method of treating a dysbiosis of the gastrointestinal tract in an individual, the
10 method comprising administering a therapeutically effective amount of a therapeutic composition according to any one of claims 1 to 67 to an individual in need thereof.

70. The therapeutic composition for use according to claim 68, or method according to claim 69, wherein the dysbiosis is a dysbiosis associated with an enteric bacterial infection.
15

71. The therapeutic composition for use, or method, according to claim 70, wherein the enteric bacterial infection is an infection with a pathogenic bacterium.

72. The therapeutic composition for use, or method, according to claim 71, wherein the
20 pathogenic bacterium is resistant to treatment with one or more antibiotics.

73. The therapeutic composition for use, or method, according to any one of claims 70 to 72, wherein the enteric bacterial infection is an infection with a pathogenic bacterium of the genus *Clostridium, Escherichia, Enterococcus, Klebsiella*, *Enterobacter, Proteus*,
25 *Salmonella*, *Shigella*, *Staphylococcus*, *Vibrio, Aeromonas, Campylobacter, Bacillus, Helicobacter, Listeria, Plesiomonas, or Yersinia*.

74. The therapeutic composition for use, or method, according to any one of claims 70 to 73, wherein the enteric bacterial infection is an infection with a pathogenic bacterium, and
30 wherein the pathogenic bacterium is *Clostridium difficile,* or *Escherichia coli*.

75. The therapeutic composition for use, or method, according to claim 74, wherein the enteric bacterial infection is an infection with pathogenic *Clostridium difficile*.

35 76. The therapeutic composition for use, or method, according to any one of claims 70 to 75, wherein the enteric bacterial infection is a recurrent or chronic infection.

77. The therapeutic composition for use, or method, according to any one of claims 68 or 69, wherein the dysbiosis is a dysbiosis associated with inflammatory bowel disease (IBD) or pouchitis.

78. The therapeutic composition for use, or method, according to claim 77, wherein the IBD is ulcerative colitis (UC) or Crohn's disease.

79. The therapeutic composition for use, or method, according to any one of claims 68 or 69, wherein the dysbiosis is a dysbiosis associated with irritable bowel syndrome (IBS).

80. The therapeutic composition for use, or method, according to any one of claims 68 or 69, wherein the dysbiosis is a dysbiosis associated with a metabolic disease, a neuropsychiatric disorder, an autoimmune disease, an allergic disorder, or a cancer.

81. The therapeutic composition for use, or method, according to claim 80, wherein the metabolic disease is selected from the group consisting of: metabolic syndrome, obesity, type 2 diabetes mellitus, a cardiovascular disease, and non-alcoholic fatty liver.

82. The therapeutic composition for use, or method, according to claim 80, wherein the neuropsychiatric disorder is selected from the group consisting of: Parkinson's disease, Alzheimer's disease, multiple sclerosis, myoclonus dystonia, autism and chronic fatigue syndrome.

83. The therapeutic composition for use, or method, according to claim 80, wherein the autoimmune disease is selected from the group consisting of: idiopathic thrombocytopenic purpura, arthritis, Sjögren's syndrome, systemic lupus erythematosus, and Hashimoto's thyroiditis.

84. The therapeutic composition for use, or method, according to claim 80, wherein the allergic disorder is selected from the group consisting of: atopy, and asthma.

85. The therapeutic composition for use, or method, according to claim 80, wherein the cancer is selected from the group consisting of: colorectal cancer, extra-intestinal tumours, mammary tumours, hepatocellular carcinoma, lymphoma, melanoma, and lung cancer.

86. The therapeutic composition for use, or method, according to any one of claims 68 or 69, wherein the dysbiosis is a dysbiosis associated with hepatic encephalopathy.

87. The therapeutic composition, therapeutic composition for use, or method, according
5 to any one of claims 1 to 86, wherein the isolated bacterium is a spore-forming bacterium.

88. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 87, wherein the isolated bacterium is susceptible to treatment with one or more antibiotics.

10

89. The therapeutic composition, therapeutic composition for use, or method, according to claim 88, wherein the antibiotic is a beta-lactam, fusidic acid, elfamycin, aminoglycoside, fosfomycin, and/or tunicamycin.

15 90. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 89, wherein at least one of bacteria in the composition is antagonistic towards a pathogenic intestinal bacterium, inhibits or prevents the growth of a pathogenic intestinal bacterium, or neutralizes or protects against a toxin produced by an intestinal bacterium.

20

91. The therapeutic composition, therapeutic composition for use, or method, according to claim 90, wherein the pathogenic bacterium is a pathogenic bacterium of the genus *Clostridium, Escherichia, Enterococcus, Klebsiella*, *Enterobacter, Proteus*, *Salmonella*, *Shigella*, *Staphylococcus*, *Vibrio, Aeromonas, Campylobacter, Bacillus, Helicobacter,*
25 *Listeria, Plesiomonas, or Yersinia*.

92. The therapeutic composition, therapeutic composition for use, or method, according to claim 9, wherein the pathogenic bacterium is *Clostridium difficile* or *Escherichia coli*.

30 93. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 92, wherein at least one of bacteria in the composition has immunomodulatory activity.

94 The therapeutic composition, therapeutic composition for use, or method, according
35 to any one of claims 1 to 93, wherein the composition further comprises a prebiotic.

95. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 94, wherein the composition further comprises a carrier.

96. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 95, wherein the composition further comprises insoluble fiber, a buffer, an osmotic agent, an antifoaming agent, and/or a preservative.

97. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 96, wherein the composition is provided in chemostat medium.

98. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 97, wherein the composition is provided in saline, optionally 0.9% saline.

99. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 98, wherein the composition is provided under reduced atmosphere.

100. The therapeutic composition, therapeutic composition for use, or method, according to claim 99, wherein the composition is under N_2, CO_2, H_2, or a mixture thereof.

101. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 100, wherein the composition is for oral administration or rectal administration.

102. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 101, wherein the composition is for oral administration and is lyophilised.

103. The therapeutic composition, therapeutic composition for use, or method, according to claim 102, wherein the composition further comprises a stabiliser and/or a cryoprotectant.

104. The therapeutic composition, therapeutic composition for use, or method, according to any one of claims 1 to 103, wherein the composition is the form of a capsule, a tablet, or an enema.

105. The therapeutic composition, therapeutic composition for use, or method, according to claim 104, wherein the capsule or tablet is enteric-coated, pH dependant, slow-release, and/or gastro-resistant.

5 106. A method of preparing a therapeutic composition according to any one of claims 1 to 67, wherein the method comprises the steps of:

(i) culturing an isolated bacterium as set out in any one of claims 1 to 67; and

(ii) mixing the bacteria obtained in (i) with a pharmaceutically acceptable excipient.

10 107. A method of preparing a therapeutic composition according to 106, wherein the method comprises the steps of:

(i) culturing a first isolated bacterium as set out in any one of claims 1 to 67;

(ii) culturing a second isolated bacterium as set out in any one of claims 1 to 67; and

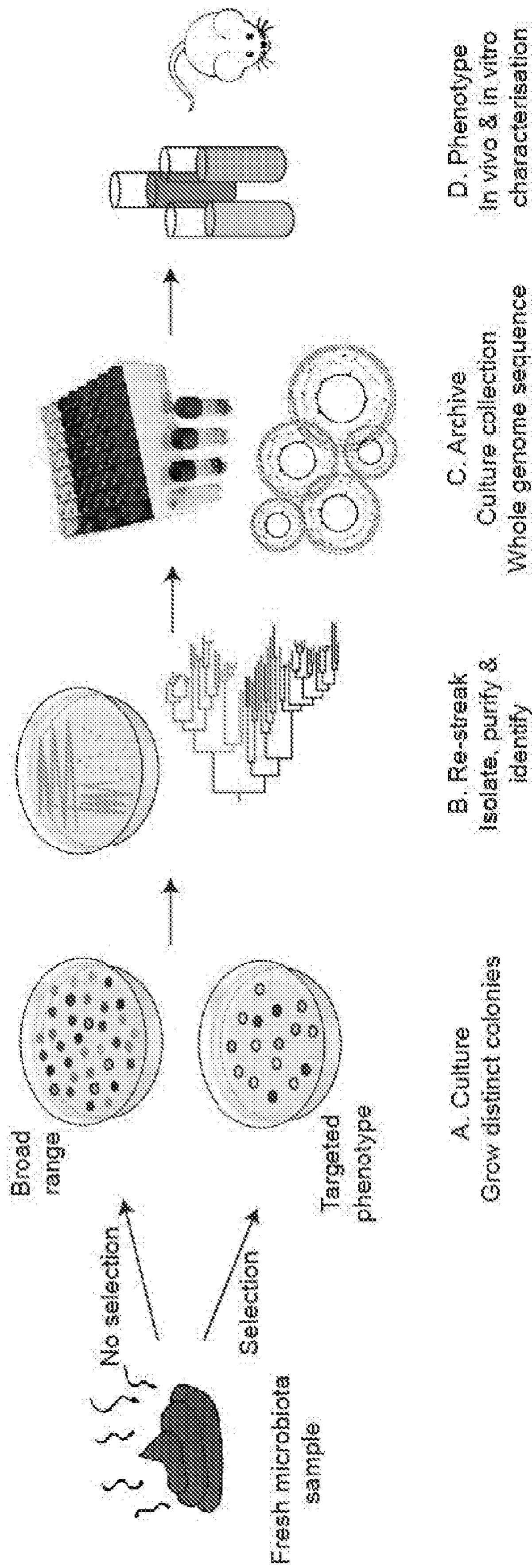
(ii) mixing the bacteria obtained in (i) and (ii) with a pharmaceutically acceptable

15 excipient,

wherein the bacteria cultured in steps (i) and (ii) have distinct 16S rRNA sequences, and wherein steps (i) and (ii) are performed independently.

108. A therapeutic composition obtainable by the method set out in claims 106 or 107.

20



5 Figure 1

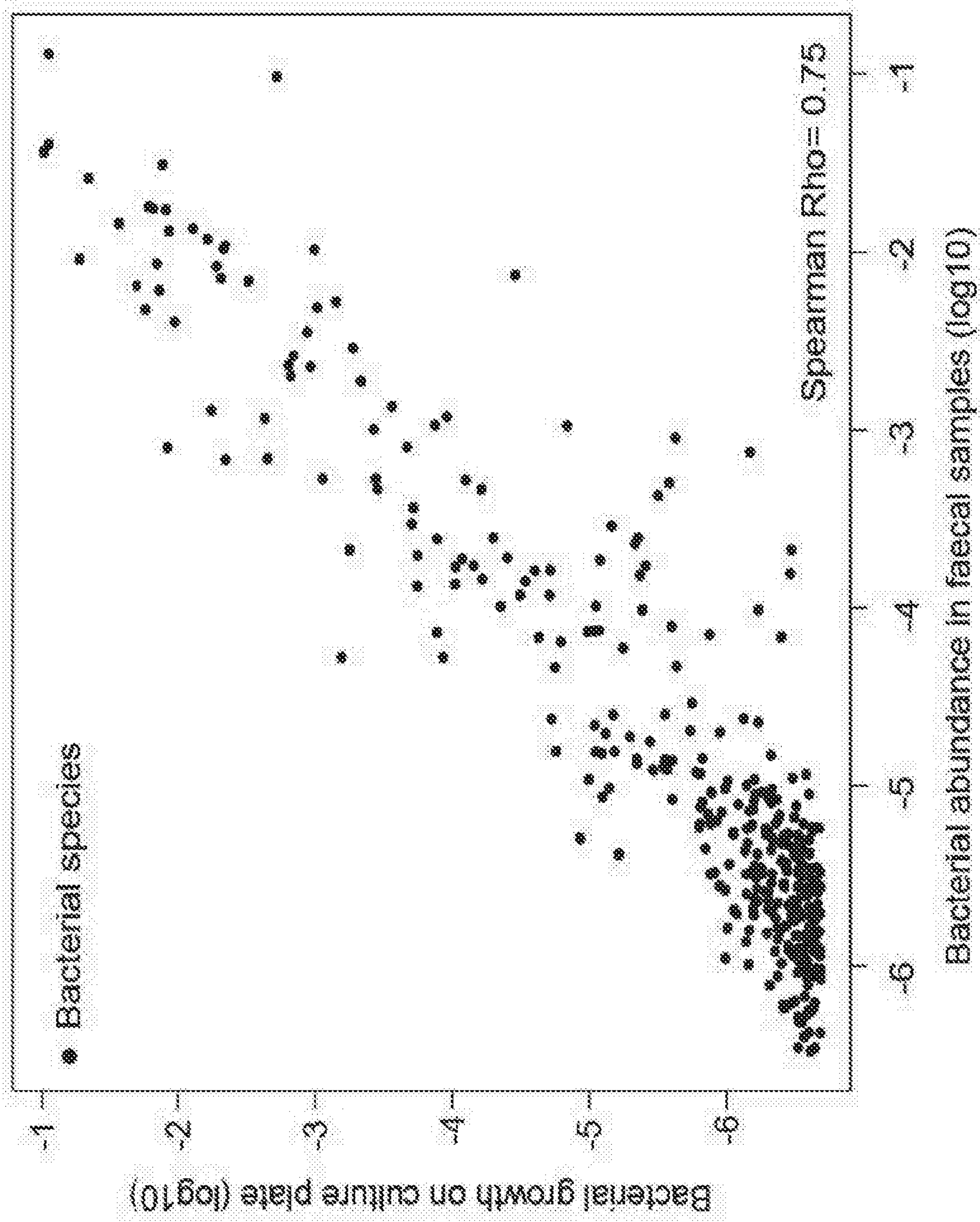


Figure 2

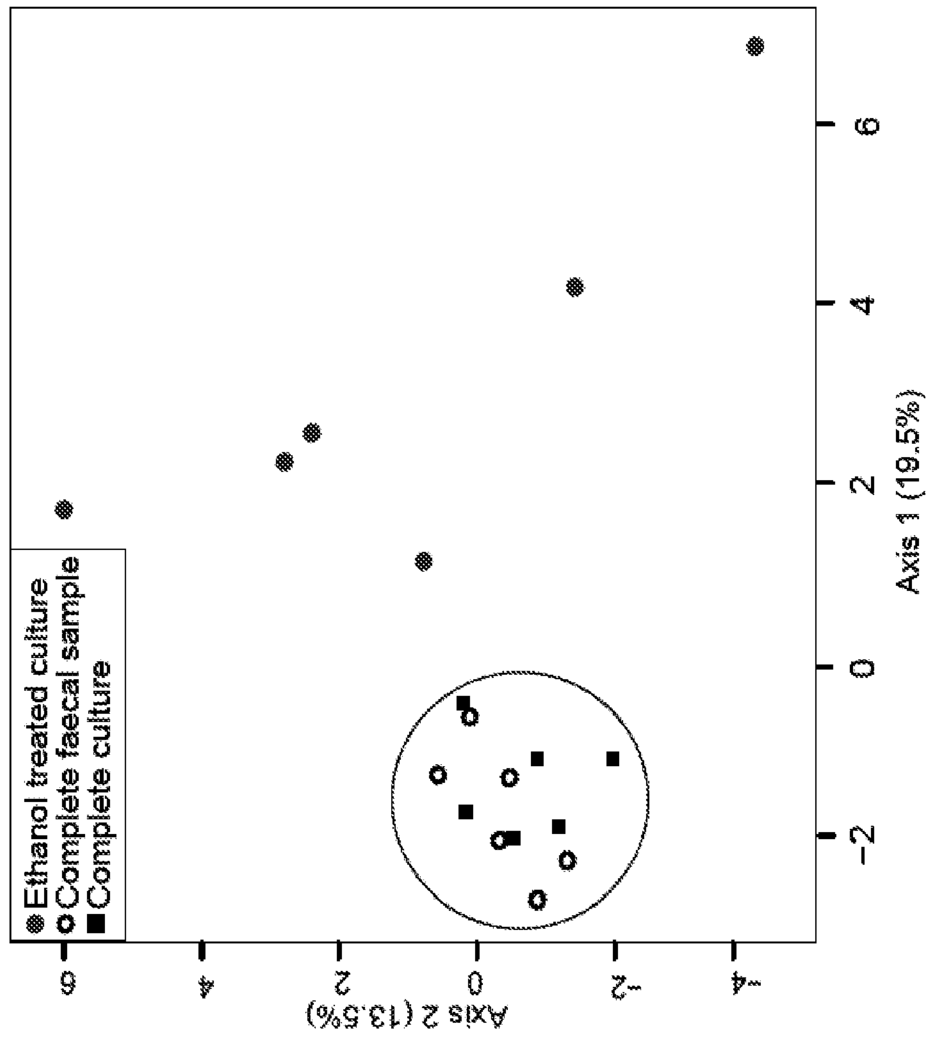


Figure 3

A

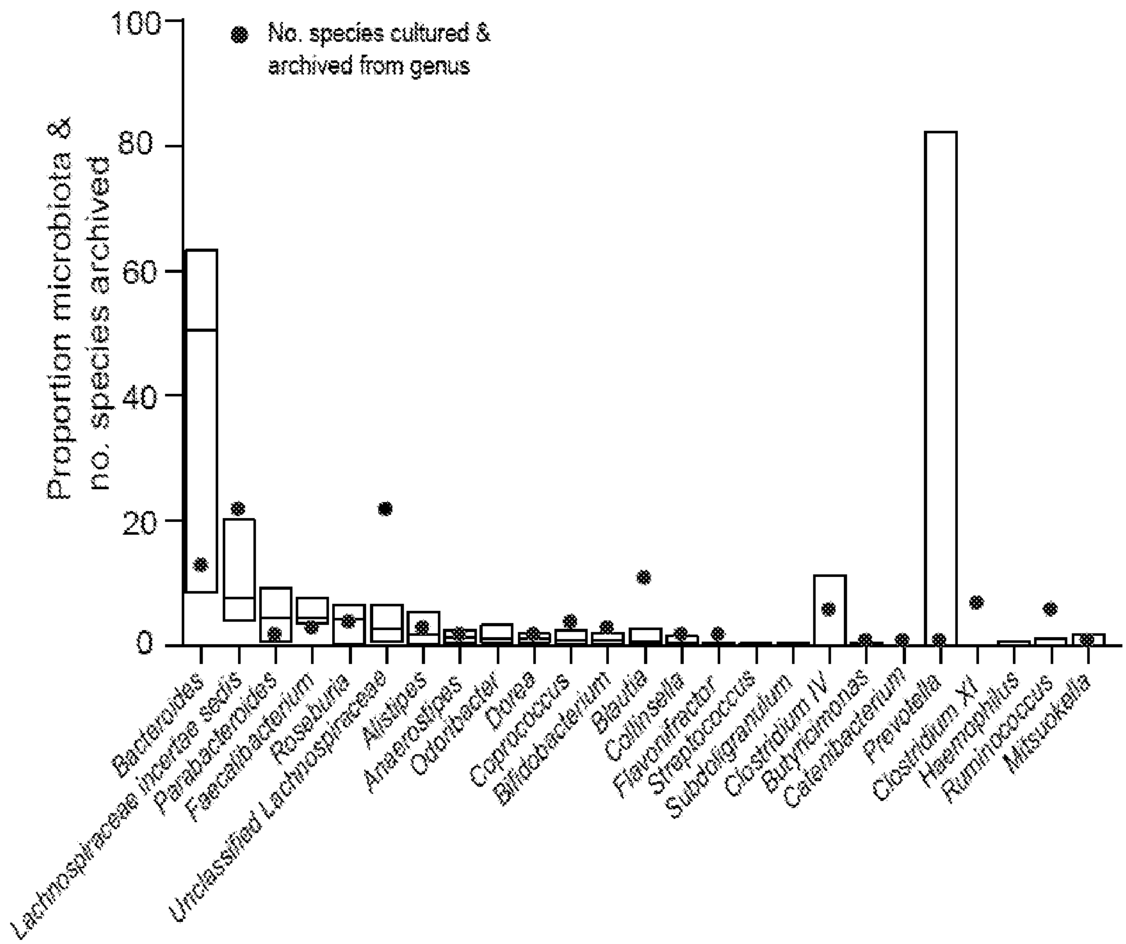


Figure 4

B

Proportion microbiota

100
80
60
40
20
0

Bacteroides vulgatus
Bacteroides uniformis
Faecalibacterium prausnitzii
Fusicatenibacter saccharivorans
Bacteroides ovatus
Bacteroides xylanisolvens
Eubacterium rectale
Parabacteroides distasonis
Bacteroides thetaiotaomicron
Alistipes finegoldii
Bacteroides caccae
Lachnospira pectinoschiza
Ruminococcus obeum
Anaerostipes hadrus
Eubacterium eligens
Odoribacter splanchnicus
Ruminococcus torques
Parabacteroides merdae
Roseburia faecis
Dorea longicatena
Roseburia intestinalis
Eubacterium siraeum
Prevotella copri
Bacteroides salyersiae

Figure 4 continued

C

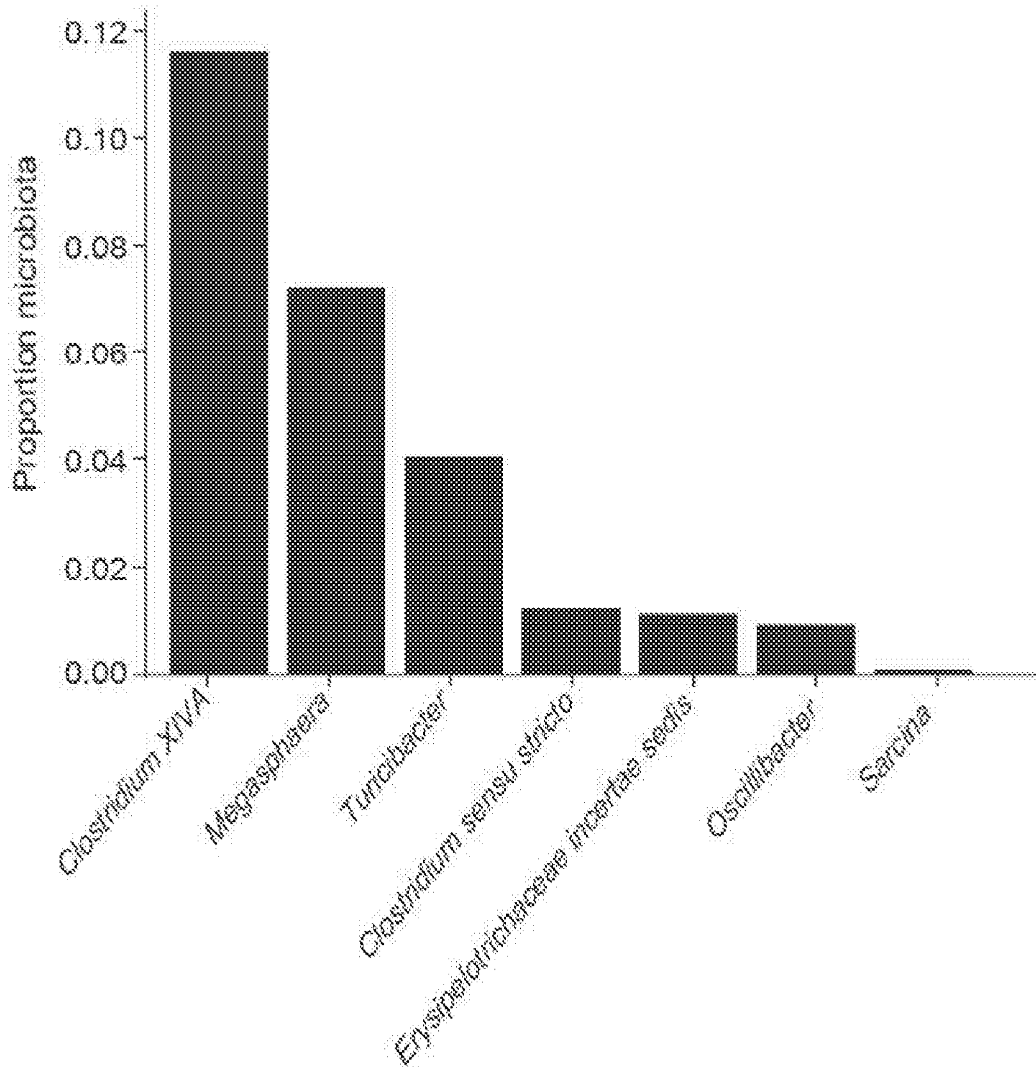


Figure 4 continued

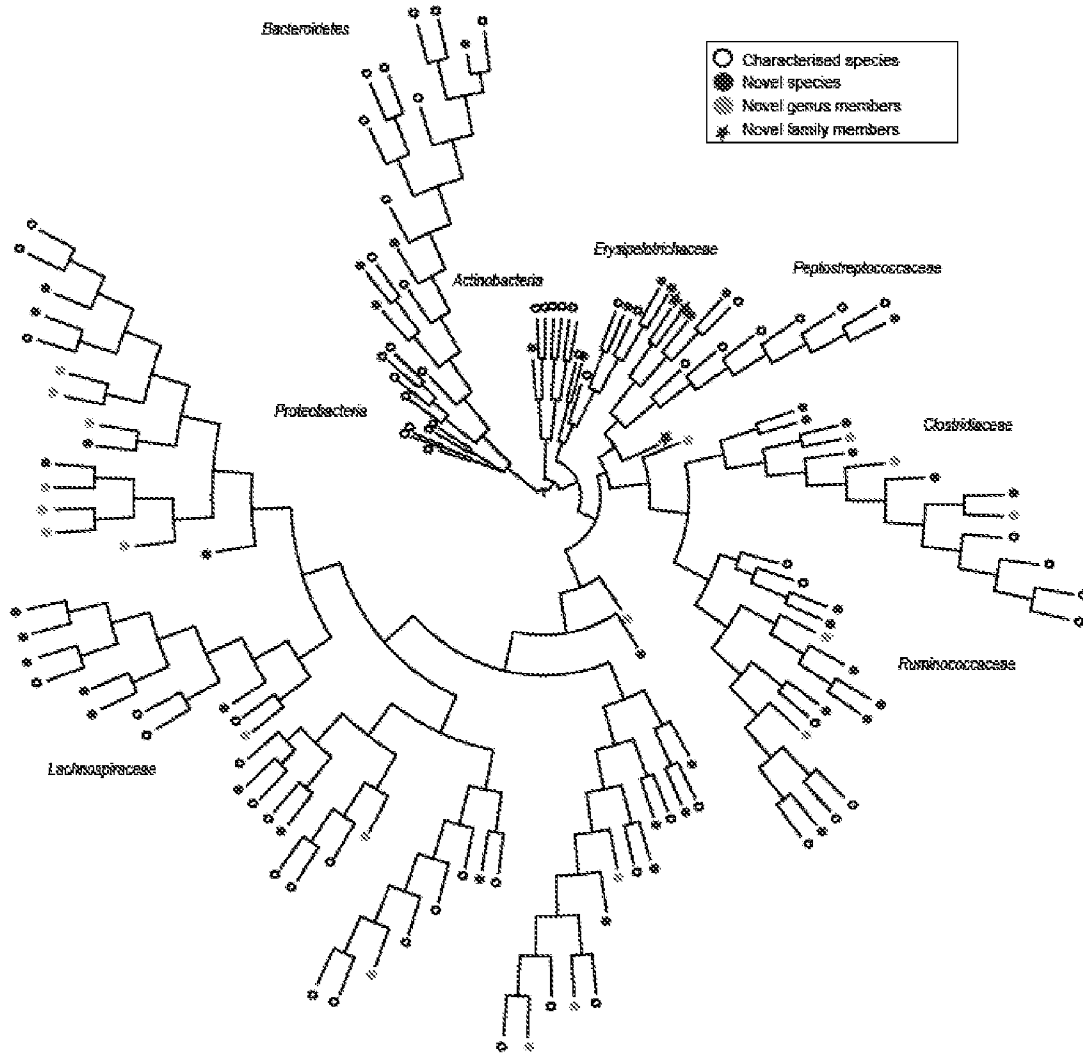


Figure 5

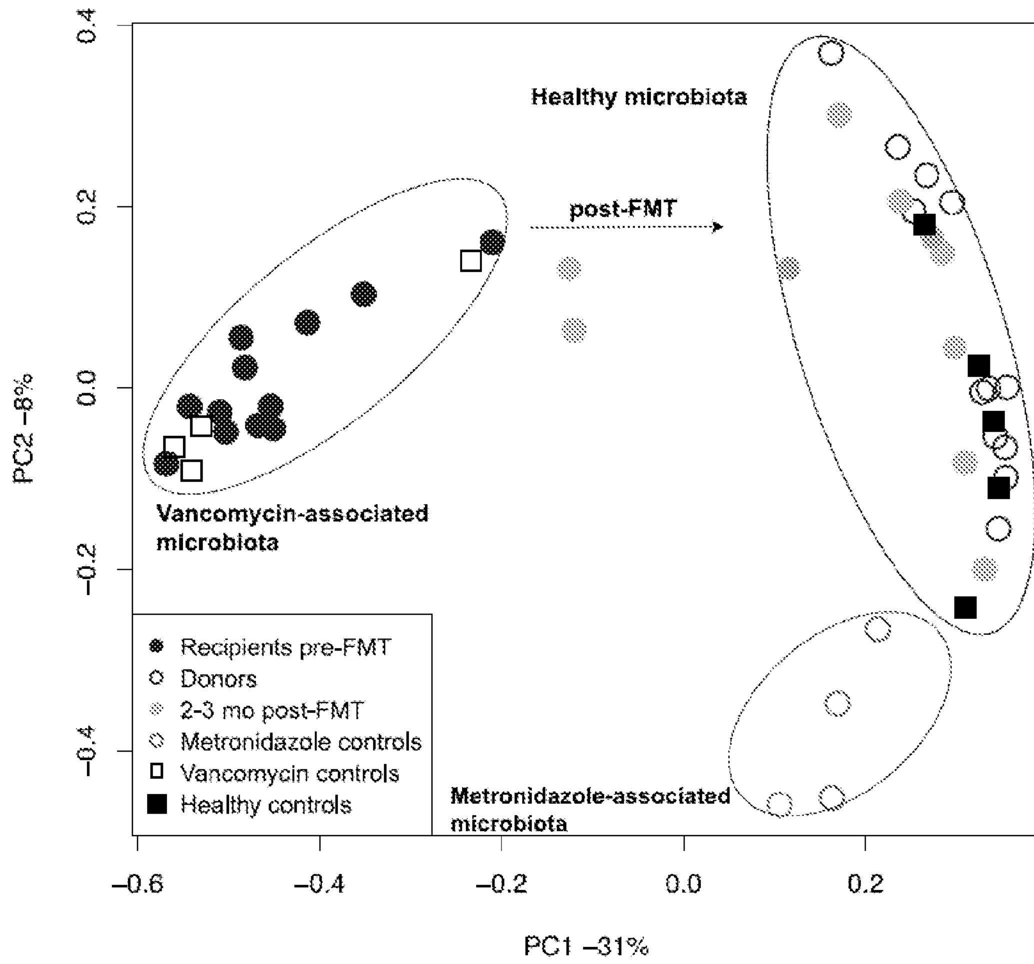


Figure 6

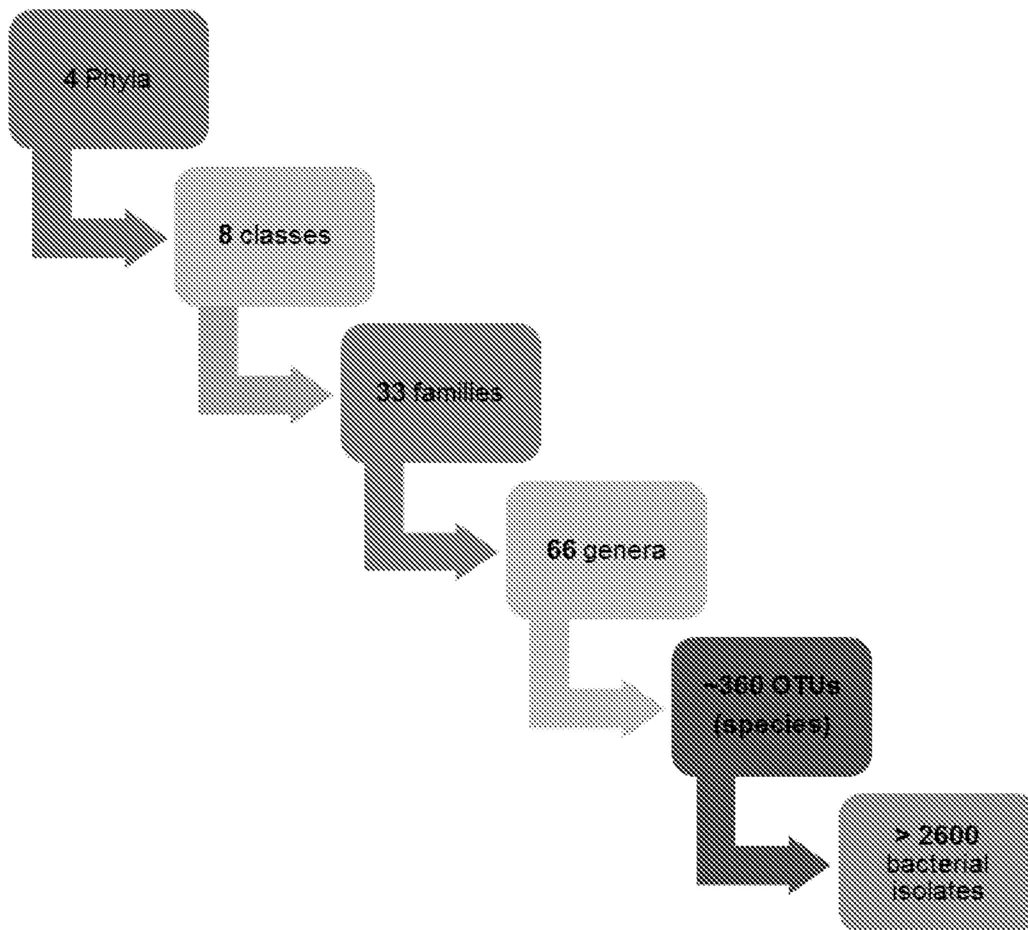


Figure 7

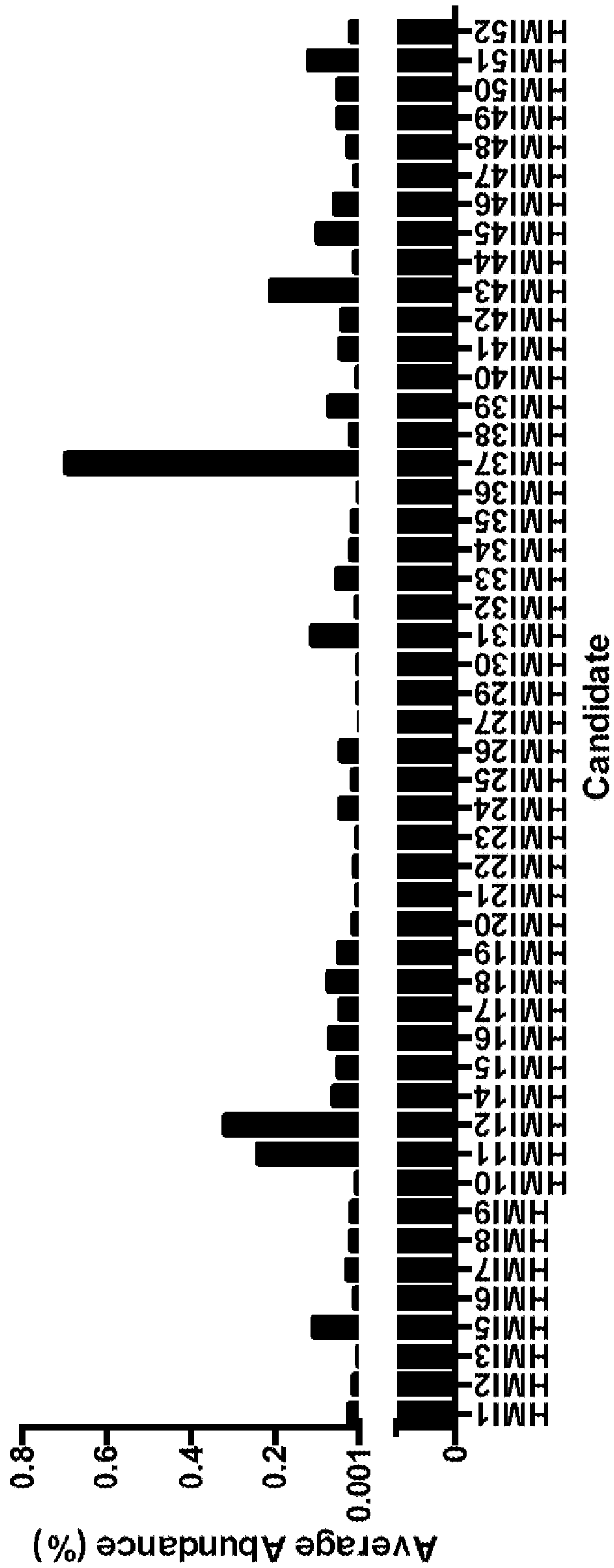


Figure 8

Figure 9

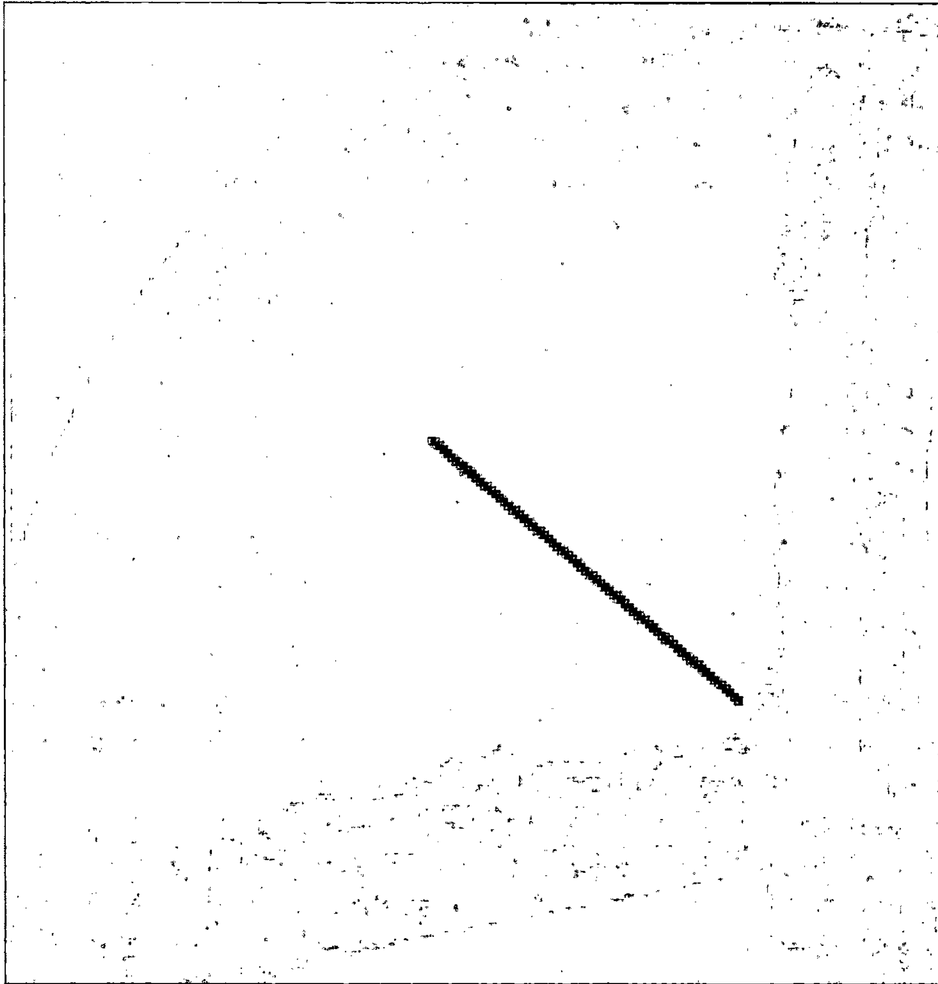
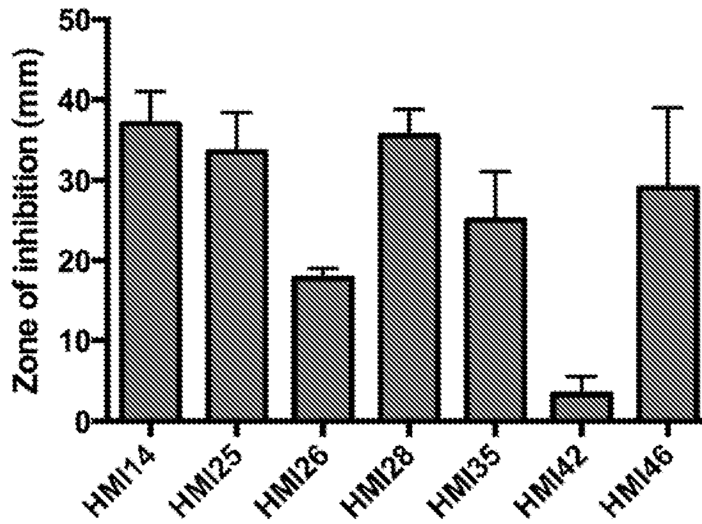


Figure 10

A



B

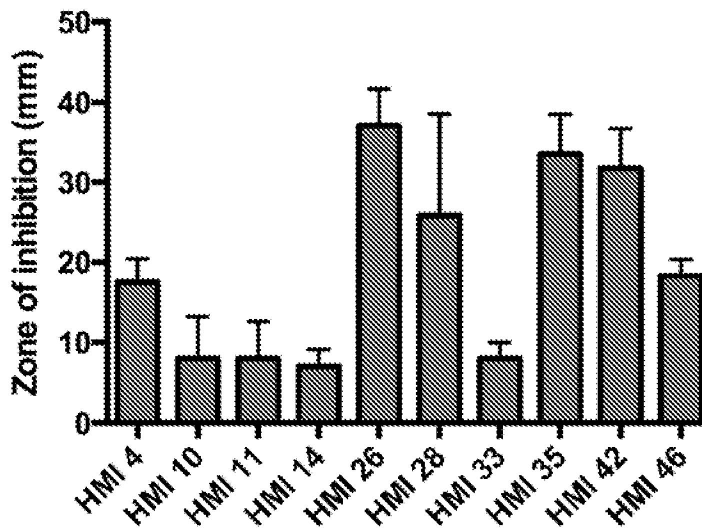


Figure 11

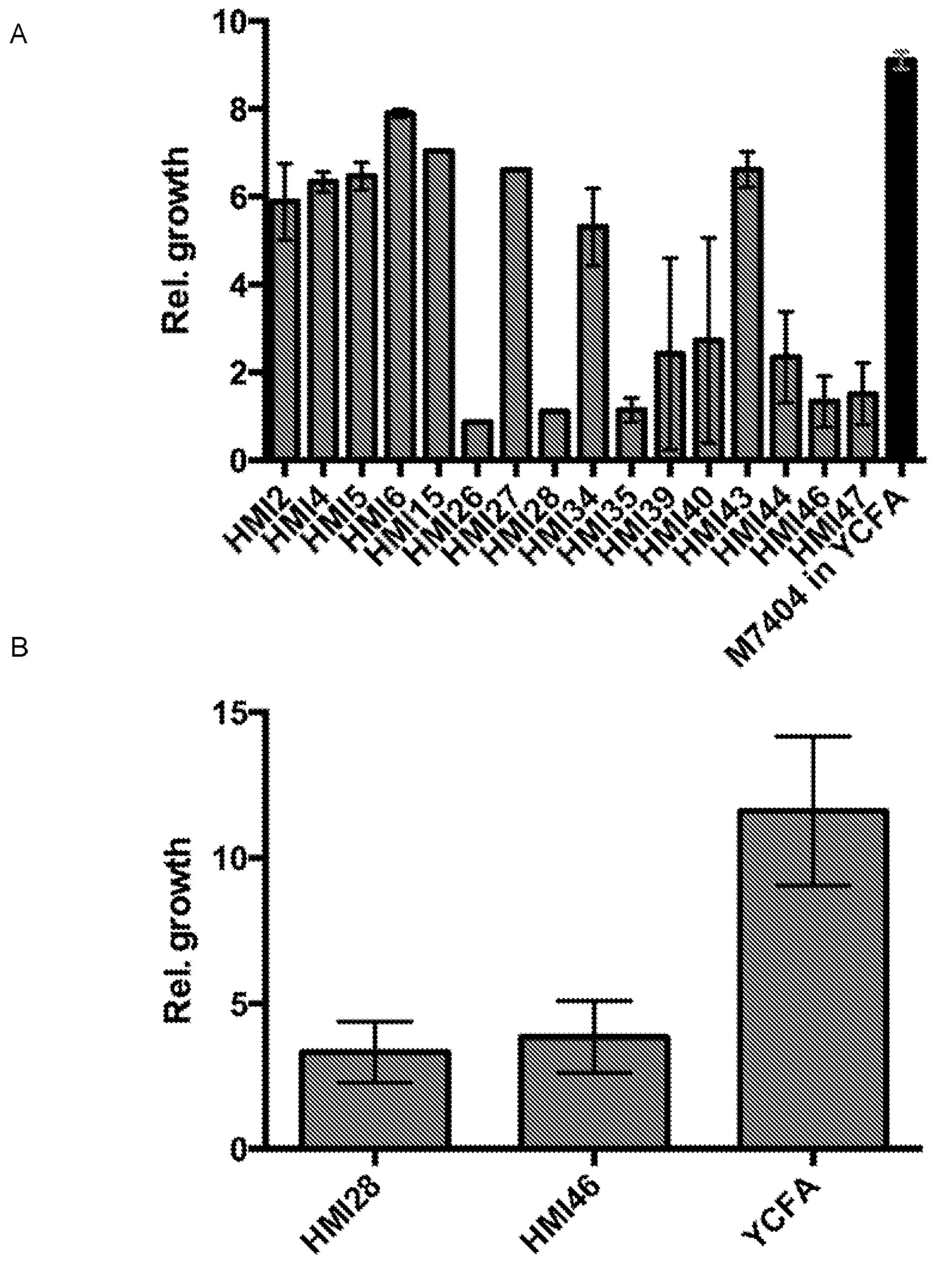


Figure 12

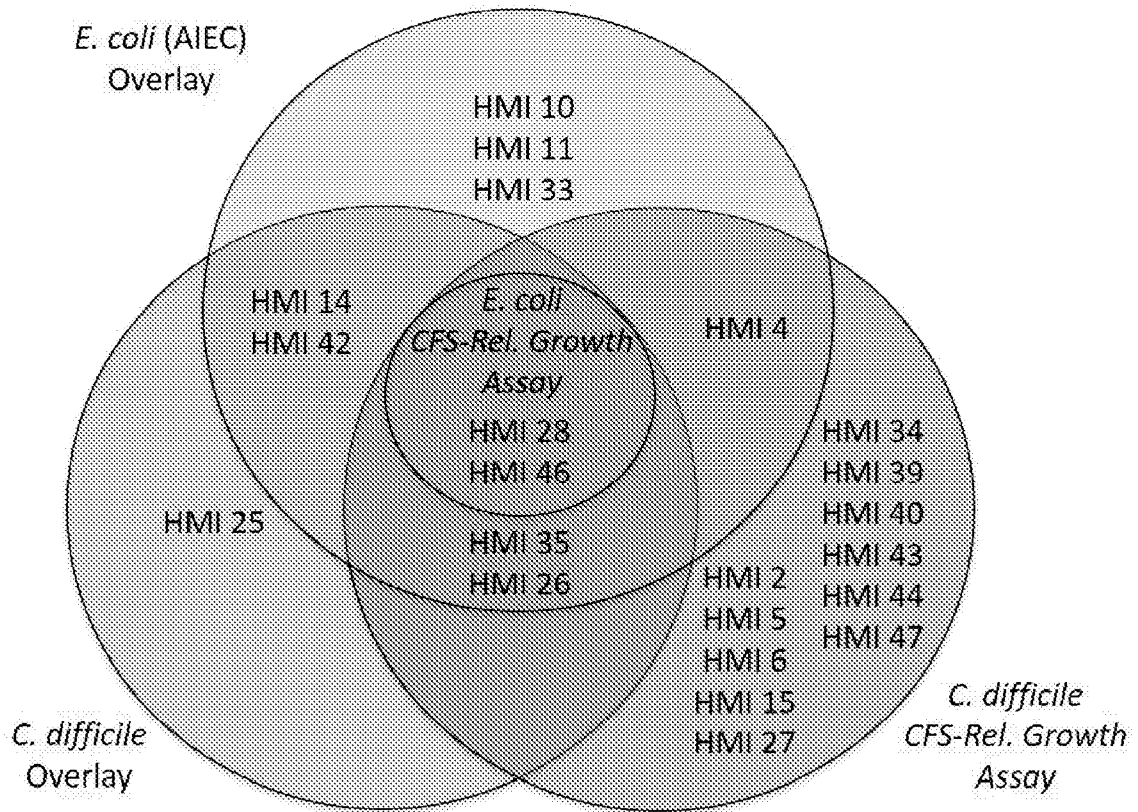


Figure 13

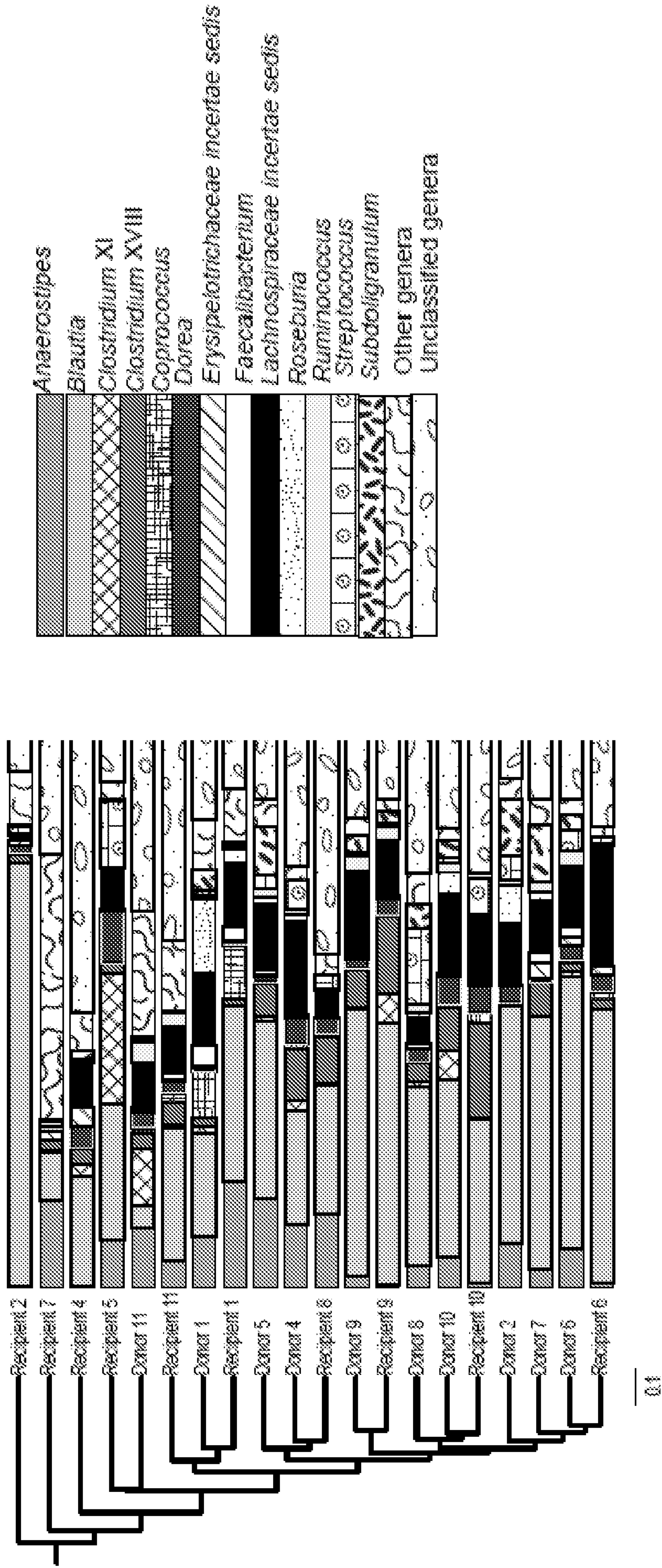


Figure 14

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2017/051083

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61K35/742 A61P31/04
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)
A61K

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, WPI Data, BIOSIS, COMPENDEX, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2014/147425 A1 (HENN MATTHEW R [US] ET AL) 29 May 2014 (2014-05-29)	1,3-21, 23-31, 33-36, 38-50, 52-65, 67-76, 87-96, 101-108
Y	paragraphs [0008], [0012], [0025]; examples P7,P8; tables 4,9; sequences 4, 548, 576, 776, 978, 1474, 1532	2,22,37, 66, 77-86, 97-100

	-/--	

[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
21 June 2017	05/07/2017

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	Bochelen, Damien

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2017/051083

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. ☐ forming part of the international application as filed:

      ☐ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☒ furnished subsequent to the international filing date for the purposes of international search only:

      ☒ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2017/051083

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2015/077794 A1 (SERES HEALTH INC [US]) 28 May 2015 (2015-05-28)	1,3-21, 23-31, 33-36, 38-50, 52-65, 67-76, 87-96, 101-108
Y	paragraphs [0009] - [0010], [0015]; sequences 552,886,1050,1057,1659	2,22,37, 66, 77-89, 97-100
X	WO 2012/039615 A2 (WINCLOVE BIO IND B V [NL]; TIMMERMAN HERMAN MARTIN [NL]; RIJKERS GERRI) 29 March 2012 (2012-03-29)	1,31,32, 53-65, 68-81, 95,96, 101-108
Y	page 14, lines 3-8; sequence 21 page 21, line 30	82-89, 97-100
X	WO 2015/156419 A1 (RIKEN [JP]; UNIV TOKYO [JP]; SCHOOL CORP AZABU VETERINARY MEDICINE EDU) 15 October 2015 (2015-10-15) paragraphs [0044], [0051]; claim 3; sequence 8	1,51-64, 68-75, 90-92

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2017/051083

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2014147425 A1	29-05-2014	US 2014147425 A1	29-05-2014
		US 2014199281 A1	17-07-2014
		US 2015190435 A1	09-07-2015
		US 2017165302 A1	15-06-2017
WO 2015077794 A1	28-05-2015	AU 2014352643 A1	30-06-2016
		CA 2931317 A1	28-05-2015
		CN 105979952 A	28-09-2016
		EP 3074027 A1	05-10-2016
		JP 2016537434 A	01-12-2016
		KR 20160095006 A	10-08-2016
		US 2017151291 A1	01-06-2017
		WO 2015077794 A1	28-05-2015
WO 2012039615 A2	29-03-2012	NONE	
WO 2015156419 A1	15-10-2015	AU 2015244700 A1	24-11-2016
		CA 2944846 A1	15-10-2015
		CN 106659746 A	10-05-2017
		EP 3129035 A1	15-02-2017
		KR 20170015286 A	08-02-2017
		US 2017028061 A1	02-02-2017
		WO 2015156419 A1	15-10-2015